University of Oklahoma

Honors Math Thesis

MATH 3980

Polynomial Formulae for the k-Slice of the Symmetric Group under Various Genome Rearrangement Models

Author:
T. Rhyker Benavidez

Supervisor:
Dr. Jonathan Kujawa

January 10, 2014

Abstract

A central goal in computational molecular biology is to be able to determine the evolutionary distance between two individuals based on their genomes. As Dobzhansky and Sturtevant [6] proposed, the fewer mutations needed to transform the genome of one individual to the genome of another, the more closely the two individuals are related. Thus, this paper seeks to provide insight into the genome rearrangement problem: given two genomes and a set of allowed mutations, what is the fewest number of mutations transforming one into the other. Building off the work of Galvão and Dias [4] and Konstantinova and Medvedev [10], we seek to determine formulae, as functions of n, for the k-slices of the symmetric group S_n under various genome rearrangement models. The models considered include reversals, prefix reversals, transpositions, and prefix transpositions for unsigned linear permutations. We conjecture that the size of each k-slice is given by a polynomial function of the size of the symmetric group, for n sufficiently large. We prove that this conjecture holds for the 1-slice and give computer calculations supporting the conjecture for all k-slices.

1 Permutations

Definition A *permutation* π is a bijection of the set $\{1, 2, \ldots, n\}$. The image of $i \in \{1, 2, \ldots, n\}$ by π is denoted by π_i. Permutations are often represented using a two-line notation, in which the first row contains the inputs and the second row contains the outputs:

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi_1 & \pi_2 & \cdots & \pi_n \end{pmatrix}$$

In this paper, however, we usually adopt a one-line notation: $\pi = [\pi_1, \pi_2, \ldots, \pi_n]$. In this way, a permutation can be thought of as an arrangement of n genes in a genome.

Definition The *product* of two permutations π and σ over the set $\{1, 2, \ldots, n\}$ is defined as their composition:

$$\pi\sigma = \pi \circ \sigma = \pi \circ [\sigma_1, \sigma_2, \ldots, \sigma_n] = [\pi_{\sigma_1}, \pi_{\sigma_2}, \ldots, \pi_{\sigma_n}]$$

Remark Multiplication is performed from right to left.

A permutation can also be thought of as a mutation, or rearrangement, of the n genes in the genome. In the above example, σ transforms the genome π into $\pi\sigma$.

Remark Mutations must be multiplied on the right of the permutation they are mutating, since they involve switching the positions of the permutation.

Definition The *symmetric group* on n, S_n, is defined to be the set of all permutations on $\{1, 2, \ldots, n\}$ together with the multiplication operation. Thus, the symmetric group on n is the set of all the possible arrangements of n genes.

Theorem 1.1. *S_n is a group.*

Proof. Since composition of functions is associative, the multiplication operation is associative. Define $e_n \in S_n$ to be the permutation [1,2,...,n]. Let $\pi \in S_n$. Then $\pi \circ e_n = \pi \circ [1, 2, \ldots, n] = [\pi_1, \pi_2, \ldots, \pi_n] = \pi$ and $e_n \circ \pi = e_n \circ [\pi_1, \pi_2, \ldots, \pi_n] = [\pi_1, \pi_2, \ldots, \pi_n] = \pi$. Thus, e_n is the *multiplicative identity* of S_n. Define π^{-1} by $\pi^{-1}_{\pi_i} = i$ for $1 \leq i \leq n$. Applying π to both sides yields $\pi \circ \pi^{-1}_{\pi_i} = \pi_i$ for $1 \leq i \leq n$. Since π is a bijection, this is equivalent to $\pi \circ \pi_i^{-1} = i$ for $1 \leq i \leq n$. So we have $\pi \circ \pi^{-1} = \pi \circ [\pi_1^{-1}, \pi_2^{-1}, \ldots, \pi_n^{-1}] = [1, 2, \ldots, n] = e_n$ and $\pi^{-1} \circ \pi = \pi^{-1} \circ [\pi_1, \pi_2, \ldots, \pi_n] = [1, 2, \ldots, n] = e_n$. Thus, π^{-1} is the *multiplicative inverse* of S_n. □

2 Genome Rearrangement Problem

Definition The *genome rearrangement problem* [7]: Given any two permutations π and σ in S_n and a set G of allowed mutations, find a minimum length sequence $\rho_1, \rho_2, \ldots, \rho_k$ of elements of G such that $\pi \circ \rho_1 \circ \rho_2 \circ \ldots \circ \rho_k = \sigma$. In other words, find a minimum length sequence of elements of G transforming π into σ.

Definition A related problem is the *distance problem*: Given any permutation π in S_n and a set G of allowed mutations, find a mimimum length sequence $\rho_1, \rho_2, \ldots, \rho_k$ of elements of G such that $\rho_1 \circ \rho_2 \circ \ldots \circ \rho_k = \pi$. In other words, find a minimum length factorization of π that consists only of elements of G.

Note that in general the genome rearrangement problem (GRP) and the distance problem (DP) are not solvable in all cases. However, by Lemma 2.1, requiring G to be a generating set ensures that both the GRP and the DP are solvable.

Definition A *generating set* G of S_n is a subset of S_n such that any element of S_n can be written as a product of the elements of G. The elements of G are said to be *generators* of S_n. Note that the empty product in S_n is equal to the identity e_n by definition.

In this paper, the only sets of allowed mutations (rearrangement models) we consider are generating sets. Thus, by Lemma 2.1, both the GRP and the DP are solvable in the cases we consider.

Lemma 2.1. *If G is a generating set, then both the genome rearrangement problem and the distance problem are solvable.*

Proof. Let G be a generating set of S_n. Let $\pi, \sigma \in S_n$. Then $\pi^{-1} \circ \sigma \in S_n$. Since G is a generating set, there exists a sequence of elements of G such that $\pi^{-1} \circ \sigma$ is a product of these elements. Let $\rho_1, \rho_2, \ldots, \rho_k$ be a minimum length sequence meeting this requirement. Thus, $\rho_1 \circ \rho_2 \circ \ldots \circ \rho_k = \pi^{-1} \circ \sigma$. This implies $\pi \circ \rho_1 \circ \rho_2 \circ \ldots \circ \rho_k = \sigma$. We need to show that $\rho_1, \rho_2, \ldots, \rho_k$ is a minimum length sequence satisfying this equation. Assume to the contrary, then there exists $\rho'_1, \rho'_2, \ldots, \rho'_l$ with $l < k$ such that $\pi \circ \rho'_1 \circ \rho'_2 \circ \ldots \circ \rho'_l = \sigma$. This would imply $\rho'_1 \circ \rho'_2 \circ \ldots \circ \rho'_l = \pi^{-1} \circ \sigma$ with $l < k$. This contradicts the fact that $\rho_1, \rho_2, \ldots, \rho_k$ was minimal. Thus, $\rho_1, \rho_2, \ldots, \rho_k$ is a solution to the GRP.

Likewise, there exists a sequence of elements of G such that π is a product of these elements. Let $\tau_1, \tau_2, \ldots, \tau_m$ be a minimum length sequence meeting

this requirement. This implies $\tau_1 \circ \tau_2 \circ \ldots \circ \tau_m = \pi$. Thus, $\tau_1, \tau_2, \ldots, \tau_m$ is a solution to the DP. □

Theorem 2.2. *For a given generating set G, the genome rearrangement problem and the distance problem are equivalent.*

Proof. Let G be a generating set. We first need to show that if we have solved the genome rearrangement problem for all $(\pi, \sigma) \in S_n \times S_n$, then we have solved the distance problem for all $\pi' \in S_n$.

Let $\tau \in S_n$. Let $\rho_1, \rho_2, \ldots, \rho_k$ be a solution to the genome rearrangement problem with input (e_n, τ). Then, $e_n \circ \rho_1 \circ \rho_2 \circ \ldots \circ \rho_k = \tau$. This implies $\rho_1 \circ \rho_2 \circ \ldots \circ \rho_k = \tau$. We need to show that $\rho_1, \rho_2, \ldots, \rho_k$ is a minimum length sequence satisfying this equation. Assume to the contrary, then there exists $\rho'_1, \rho'_2, \ldots, \rho'_l$ with $l < k$ such that $\rho'_1 \circ \rho'_2 \circ \ldots \circ \rho'_l = \tau$. This would imply $e_n \circ \rho'_1 \circ \rho'_2 \circ \ldots \circ \rho'_l = \tau$ with $l < k$. This contradicts the fact that $\rho_1, \rho_2, \ldots, \rho_k$ is a solution to the genome rearrangement problem with input (e_n, τ). Thus, $\rho_1, \rho_2, \ldots, \rho_k$ is a solution to the distance problem with input τ. Since τ was arbitrary, we have solved the distance problem.

We now need to show that if we have solved the distance problem for all $\pi \in S_n$, then we have solved the genome rearrangement problem for all $(\pi', \sigma') \in S_n \times S_n$.

Let $\tau, \upsilon \in S_n$. Then, $\tau^{-1} \circ \upsilon \in S_n$. Let $\rho_1, \rho_2, \ldots, \rho_k$ be a solution to the distance problem with input $\tau^{-1} \circ \upsilon$. Then, $\rho_1 \circ \rho_2 \circ \ldots \circ \rho_k = \tau^{-1} \circ \upsilon$. This implies $\tau \circ \rho_1 \circ \rho_2 \circ \ldots \circ \rho_k = \upsilon$. We need to show that $\rho_1, \rho_2, \ldots, \rho_k$ is a minimum length sequence satisfying this equation. Assume to the contrary, then there exists $\rho'_1, \rho'_2, \ldots, \rho'_l$ with $l < k$ such that $\tau \circ \rho'_1 \circ \rho'_2 \circ \ldots \circ \rho'_l = \upsilon$. This would imply $\rho'_1 \circ \rho'_2 \circ \ldots \circ \rho'_l = \tau^{-1} \circ \upsilon$ with $l < k$. This contradicts the fact that $\rho_1, \rho_2, \ldots, \rho_k$ is a solution to the distance problem with input $\tau^{-1} \circ \upsilon$. Thus, $\rho_1, \rho_2, \ldots, \rho_k$ is a solution to the genome rearrangement problem with input (τ, υ). Since τ and υ were arbitrary, we have solved the genome rearrangement problem. □

In this paper, we specifically investigate the distance problem. However, by Theorem 2.2, our results are also pertinent to the genome rearrangement problem.

3 Distance, Diameter, and *k*-Slices

Definition For any permutation $\pi \in S_n$ and generating set G of S_n, the *distance* of π under G, denoted by $d_G(\pi)$, is defined to be the number of elements in a minimum length factorization of π that consists only of elements of G.

The distance of a permutation τ can be thought of as the fewest number of mutations necessary to transform the identity permutation into τ. Likewise, by Theorem 2.2, if $\tau = \pi^{-1} \circ \sigma$ for some $\pi, \sigma \in S_n$, then $d_G(\tau)$ is equivalent to the fewest number of mutations necessary to transform π into σ.

Remark Since the minimum length factorization of e_n is the empty product, $d_G(e_n) = 0$.

Definition Let G be a generating set of S_n. The *diameter* of S_n under G, denoted by D_G, is given by:

$$D_G = \max_{\pi \in S_n} d_G(\pi)$$

The diameter gives an indicator for how closely related the genomes in the symmetric group are. The greater the diameter, the greater the distance of the genome or genomes that are furthest from the identity.

Definition Let G be a generating set of S_n. The *k-slice* of S_n under G, denoted by S_n^k, is given by:

$$S_n^k = \{\pi \in S_n \mid d_G(\pi) = k\}$$

Remark Since e_n is the only permutation with distance 0, $S_n^0 = \{e_n\}$. Furthermore, it is easily seen that $S_n^1 = G$ for any G that does not include the identity.

Remarkably, breaking the symmetric group into k-slices led to the observation of patterns concerning the size of the k-slice in relation to the size of the symmetric group. For now, however, we go into more detail concerning the rearrangement models under consideration in this paper.

4 Rearrangement Models

Definition A *rearrangement* is a permutation in S_n that acts on another permutation according to a specific rule. Common genome rearrangements in nature include reversals, prefix reversals, transpositions, and prefix transpositions. In this paper, we consider genome rearrangements only on unsigned linear permutations.

Definition Let $1 \leq a < b \leq n$. A *reversal* on S_n, denoted by $\rho(a, b)$, is a rearrangement given by:

$$\rho = \begin{pmatrix} 1 & \cdots & a-1 & \boxed{a} & \boxed{a+1} & \cdots & \boxed{b-1} & \boxed{b} & b+1 & \cdots & n \\ 1 & \cdots & a-1 & b & b-1 & \cdots & a+1 & a & b+1 & \cdots & n \end{pmatrix}$$

Thus, a reversal reverses the closed interval determined by a and b. Equivalently,

$$\rho_i = \begin{cases} i & \text{if } 1 \leq i \leq a-1 \text{ or } b+1 \leq i \leq n \\ (b+a) - i & \text{if } a \leq i \leq b \end{cases}$$

Definition Let $1 < b \leq n$. A *prefix reversal* on S_n, denoted by $\rho(1, b)$, is a reversal with $a = 1$. Equivalently,

$$\rho_i = \begin{cases} i & \text{if } b+1 \leq i \leq n \\ (b+1) - i & \text{if } 1 \leq i \leq b \end{cases}$$

Definition Let $1 \leq a < b < c \leq n+1$. A *transposition* on S_n, denoted by $\tau(a, b, c)$, is a rearrangement given by:

$$\tau = \begin{pmatrix} 1 & \cdots & a-1 & \boxed{a} & \boxed{a+1} & \cdots & \boxed{b-2} & \boxed{b-1} & \boxed{b} & \boxed{b+1} & \cdots & \boxed{c-1} & c & \cdots & n \\ 1 & \cdots & a-1 & \boxed{b} & \boxed{b+1} & \cdots & \boxed{c-1} & \boxed{a} & \boxed{a+1} & \cdots & \boxed{b-2} & \boxed{b-1} & c & \cdots & n \end{pmatrix}$$

Thus, a transposition swaps two adjacent blocks. Equivalently,

$$\tau_i = \begin{cases} i & \text{if } 1 \leq i \leq a-1 \text{ or } c \leq i \leq n \\ (b-a)+i & \text{if } a \leq i \leq a-b+c-1 \\ (b-c)+i & \text{if } a-b+c \leq i \leq c-1 \end{cases}$$

Definition Let $1 < b < c \leq n+1$. A *prefix transposition* on S_n, denoted by $\tau(1,b,c)$, is a transposition with $a = 1$. Equivalently,

$$\tau_i = \begin{cases} i & \text{if } c \leq i \leq n \\ (b-1)+i & \text{if } 1 \leq i \leq c-b \\ (b-c)+i & \text{if } c-b+1 \leq i \leq c-1 \end{cases}$$

Lemma 4.1. *Let $\rho(a,b), \rho(c,d) \in S_n$. Then $\rho(a,b) = \rho(c,d) \iff a = b$ and $c = d$.*

Proof. This is straightforward from the definition of a reversal. □

Lemma 4.2. *Let $\tau(a,b,c),\ \tau(d,e,f) \in S_n$. Then $\tau(a,b,c) = \tau(d,e,f) \iff a = d$ and $b = e$ and $c = f$.*

Proof. This is straightforward from the definition of a transposition. □

Theorem 4.3. *Let $\rho^{-1}(a,b)$ and $\tau^{-1}(a,b,c)$ denote the inverse of $\rho(a,b)$ and $\tau(a,b,c)$ respectively. Then $\rho^{-1}(a,b) = \rho(a,b)$ and $\tau^{-1}(a,b,c) = \tau(a,a-b+c,c)$.*

Proof. We need to calculate $(\rho(a,b) \circ \rho(a,b))_i$ for $1 \leq i \leq n$. Case 1: $1 \leq i \leq a-1$ or $b+1 \leq i \leq n$. Then $\rho(a,b)_{\rho(a,b)_i} = \rho(a,b)_i = i$. Case 2: $a \leq i \leq b$. This implies $a \leq (b+a)-i \leq b$. Thus, $\rho(a,b)_{\rho(a,b)_i} = \rho(a,b)_{(b+a)-i} = (b+a)-((b+a)-i) = i$. Thus, $\rho(a,b) \circ \rho(a,b) = e_n$ which implies $\rho^{-1}(a,b) = \rho(a,b)$.

We also need to calculate $(\tau(a,b,c) \circ \tau(a,a-b+c,c))_i$ and $(\tau(a,a-b+c,c) \circ \tau(a,b,c))_i$ for $1 \leq i \leq n$. We first calculate $(\tau(a,b,c) \circ \tau(a,a-b+c,c))_i$ for $1 \leq i \leq n$. Case 1: $1 \leq i \leq a-1$ or $c \leq i \leq n$. Then $\tau(a,b,c)_{\tau(a,a-b+c,c)_i} = \tau(a,b,c)_i = i$. Case 2: $a \leq i \leq a-(a-b+c)+c-1$. This implies $a-b+c \leq c-b+i \leq c-1$. Thus, $\tau(a,b,c)_{\tau(a,a-b+c,c)_i} = \tau(a,b,c)_{((a-b+c)-a)+i} = \tau(a,b,c)_{c-b+i} = (b-c)+(c-b+i) = i$. Case 3: $a-(a-b+c)+c \leq i \leq c-1$. This implies $a \leq a-b+i \leq a-b+c-1$. Thus, $\tau(a,b,c)_{\tau(a,a-b+c,c)_i} = \tau(a,b,c)_{((a-b+c)-c)+i} = \tau(a,b,c)_{a-b+i} = (b-a)+(a-b+i) = i$. Thus, $\tau(a,b,c) \circ \tau(a,a-b+c,c) = e_n$.

We now calculate $(\tau(a,a-b+c,c) \circ \tau(a,b,c))_i$ for $1 \leq i \leq n$. Case 1: $1 \leq i \leq a-1$ or $c \leq i \leq n$. Then $\tau(a,a-b+c,c)_{\tau(a,b,c)_i} = \tau(a,a-b+c,c)_i = i$. Case 2: $a \leq i \leq a-b+c-1$. This implies $a-(a-b+c)+c \leq (b-a)+i \leq c-1$. Thus, $\tau(a,a-b+c,c)_{\tau(a,b,c)_i} = \tau(a,a-b+c,c)_{(b-a)+i} = ((a-b+c)-c)+((b-a)+i) = i$. Case 3: $a-b+c \leq i \leq c-1$. This implies $a \leq (b-c)+i \leq a-(a-b+c)+c-1$. Thus, $\tau(a,a-b+c,c)_{\tau(a,b,c)_i} = \tau(a,a-b+c,c)_{(b-c)+i} = ((a-b+c)-a)+((b-c)+i) = i$. Thus, $\tau(a,a-b+c,c) \circ \tau(a,b,c) = e_n$. This gives $\tau^{-1}(a,b,c) = \tau(a,a-b+c,c)$ as required. □

Definition A *rearrangement model* is another name for the set of allowed mutations. Define R_n, PR_n, T_n, and PT_n to be the set of all reversals, the set of all prefix reversals, the set of all transpositions, and the set of all prefix transpositions on S_n.

Definition Let $1 \leq a \leq n-1$. A *simple transposition* on S_n, denoted by $\mu(a)$, is a rearrangement given by:

$$\mu = \begin{pmatrix} 1 & \cdots & a-1 & \boxed{a} & \boxed{a+1} & a+2 & \cdots & n \\ 1 & \cdots & a-1 & a+1 & a & a+2 & \cdots & n \end{pmatrix}$$

Thus, a simple transposition switches the elements a and $a+1$. Equivalently,

$$\mu_i = \begin{cases} i & \text{if } 1 \leq i \leq a-1 \text{ or } a+2 \leq i \leq n \\ (2a+1) - i & \text{if } a \leq i \leq a+1 \end{cases}$$

Lemma 4.4. *The simple transposition $\mu(a)$ can be factored into reversals: $\mu(a) = \rho(a, a+1)$ for $1 \leq a \leq n-1$; into prefix reversals: $\mu(a) = \rho(1, a+1) \circ \rho(1,2) \circ \rho(1, a+1)$ for $1 \leq a \leq n-1$; into transpositions: $\mu(a) = \tau(a, a+1, a+2)$ for $1 \leq a \leq n-1$; and into prefix transpositions: $\mu(1) = \tau(1,2,3)$ and $\mu(a) = \tau(1, a, a+1) \circ \tau(1, 2, a+2)$ for $2 \leq a \leq n-1$.*

Proof. We first show that $\mu(a)$ is equivalent to $\rho(a, a+1)$. Case 1: $1 \leq i \leq a-1$ or $a+2 \leq i \leq n$. Then, $\rho(a, a+1)_i = i$. Case 2: $a \leq i \leq a+1$. Then $\rho(a, a+1)_i = ((a+1) + a) - i = (2a+1) - i$. So, $\mu(a) = \rho(a, a+1)$.

Second, we show that $\mu(a)$ is equivalent to $\rho(1, a+1) \circ \rho(1,2) \circ \rho(1, a+1)$. Case 1: $1 \leq i \leq a-1$. This implies $1 \leq i \leq a+1$, $3 \leq a+2-i \leq n$, and $1 \leq a+2-i \leq a+1$. Thus, $\rho(1,a+1)_{\rho(1,2)_{\rho(1,a+1)_i}} = \rho(1,a+1)_{\rho(1,2)_{((a+1)+1)-i}} = \rho(1,a+1)_{\rho(1,2)_{a+2-i}} = \rho(1,a+1)_{a+2-i} = ((a+1)+1) - (a+2-i) = i$. Case 2: $a+2 \leq i \leq n$. This implies $3 \leq i \leq n$. Thus, $\rho(1,a+1)_{\rho(1,2)_{\rho(1,a+1)_i}} = \rho(1,a+1)_{\rho(1,2)_i} = \rho(1,a+1)_i = i$. Case 3: $a \leq i \leq a+1$. This implies $1 \leq i \leq a+1$, $1 \leq a+2-i \leq 2$, and $1 \leq 1-a+i \leq a+1$. Thus, $\rho(1,a+1)_{\rho(1,2)_{\rho(1,a+1)_i}} = \rho(1,a+1)_{\rho(1,2)_{((a+1)+1)-i}} = \rho(1,a+1)_{\rho(1,2)_{a+2-i}} = \rho(1,a+1)_{(2+1)-(a+2-i)} = \rho(1,a+1)_{1-a+i} = ((a+1)+1)-(1-a+i) = (2a+1)-i$. So, $\mu(a) = \rho(1,a+1) \circ \rho(1,2) \circ \rho(1,a+1)$.

Third, we show that $\mu(a)$ is equivalent to $\tau(a, a+1, a+2)$. Case 1: $1 \leq i \leq a-1$ or $a+2 \leq i \leq n$. Then $\tau(a, a+1, a+2)_i = i$. Case 2: $i = a$. This implies $a \leq i \leq a-(a+1)+(a+2)-1$. Thus, $\tau(a, a+1, a+2)_i = ((a+1)-a)+i = i+1 = a+1 = (2a+1)-a = (2a+1)-i$. Case 3: $i = a+1$. This implies $a-(a+1)+(a+2) \leq i \leq (a+2)-1$. Thus, $\tau(a, a+1, a+2)_i = ((a+1)-(a+2))+i = -1+i = -1+(a+1) = a = (2a+1)-(a+1) = (2a+1)-i$. So, $\mu(a) = \tau(a, a+1, a+2)$.

Fourth, $\mu(1) = \tau(1,2,3)$ follows directly from the previous section of this proof with $a = 1$.

Lastly, we show that $\mu(a)$ is equivalent to $\tau(1, a, a+1) \circ \tau(1, 2, a+2)$ for $2 \leq a \leq n-1$. Case 1: $1 \leq i \leq a-1$. This implies $1 \leq i \leq 1-2+(a+2)-1$, and $1-a+(a+1) \leq 1+i \leq (a+1)-1$. Then, $\tau(1, a, a+1)_{\tau(1,2,a+2)_i} =$

$\tau(1,a,a+1)_{(2-1)+i} = \tau(1,a,a+1)_{1+i} = (a-(a+1))+(1+i) = i$. Case 2: $a+2 \leq i \leq n$. This implies $a+1 \leq i \leq n$. Thus, $\tau(1,a,a+1)_{\tau(1,2,a+2)_i} = \tau(1,a,a+1)_i = i$. Case 3: $i = a$. This implies $1 \leq i \leq 1-2+(a+2)-1$, and $a+1 \leq 1+i \leq n$. Thus, $\tau(1,a,a+1)_{\tau(1,2,a+2)_i} = \tau(1,a,a+1)_{(2-1)+i} = \tau(1,a,a+1)_{1+i} = 1+i = 1+a = (2a+1)-a = (2a+1)-i$. Case 4: $i = a+1$. This implies $1-2+(a+2) \leq i \leq (a+2)-1$, and $1 \leq -a+i \leq 1-a+(a+1)-1$. Thus, $\tau(1,a,a+1)_{\tau(1,2,a+2)_i} = \tau(1,a,a+1)_{(2-(a+2))+i} = \tau(1,a,a+1)_{-a+i} = (a-1)+(-a+i) = -1+i = -1+(a+1) = a = (2a+1)-(a+1) = (2a+1)-i$. So, $\mu(a) = \tau(1,a,a+1) \circ \tau(1,2,a+2)$ for $2 \leq a \leq n-1$. □

Theorem 4.5. *The rearrangement models R_n, PR_n, T_n, and PT_n are each generating sets of S_n.*

Proof. It is a well known fact that the symmetric group S_n is generated by simple transpositions [8]. Thus, for each $\pi \in S_n$, there exists a sequence of simple transpositions $\mu_1, \mu_2, \ldots, \mu_k$ such that $\mu_1 \circ \mu_2 \circ \ldots \circ \mu_k = \pi$. By Lemma 4.4, each of these simple transpositions can factor into just reversals, into just prefix reversals, into just transpositions, or into just prefix transpositions. Thus, π can be factored into just reversals, into just prefix reversals, into just transpositions, or into just prefix transpositions. □

Remark By Theorem 4.3, all of the rearrangement models we consider are closed under inverses.

Definition Let $\pi \in S_n$. We call $d_{R_n}(\pi)$, $d_{PR_n}(\pi)$, $d_{T_n}(\pi)$, and $d_{PT_n}(\pi)$ the *reversal distance*, *prefix reversal distance*, *transposition distance*, and *prefix transposition distance* of π respectively. We call D_{R_n}, D_{PR_n}, D_{T_n}, and D_{PT_n} the *reversal diameter*, *prefix reversal diameter*, *transposition diameter*, and *prefix transposition diameter* of S_n respectively.

5 Recursive Formulae for 1-Slice

Definition Let $m, n \in \mathbb{N}$ with $m < n$. The *symmetric group on n restricted to m*, denoted by $S_{n|m}$, is given by:

$$S_{n|m} = \{\pi \in S_n \mid \pi_i = i \text{ for } m+1 \leq i \leq n\}$$

Theorem 5.1. *$S_{n|m}$ is a group.*

Proof. $S_{n|m}$ is a subset of S_n by definition. Thus, we need to show that $S_{n|m}$ is not empty, that it is closed under multiplication, and that it is closed under inverses. Since $e_n \in S_{n|m}$, it is nonempty. Let $\pi, \sigma \in S_{n|m}$. Then $\pi \circ \sigma = \pi \circ [\sigma_1, \sigma_2, \ldots, \sigma_m, \sigma_{m+1}, \ldots, \sigma_n] = \pi \circ [\sigma_1, \sigma_2, \ldots, \sigma_m, m+1, \ldots, n] = [\pi_{\sigma_1}, \pi_{\sigma_2}, \ldots, \pi_{\sigma_m}, \pi_{m+1}, \ldots, \pi_n] = [\pi_{\sigma_1}, \pi_{\sigma_2}, \ldots, \pi_{\sigma_m}, m+1, \ldots, n] \in S_{n|m}$. Let $m+1 \leq i \leq n$. Then $\pi_i^{-1} = \pi_{\pi_i}^{-1} = i$. Thus, $\pi^{-1} \in S_{n|m}$. □

Definition The *restriction function*, $\phi : S_{n|m} \to S_m$, is given by:

$$\phi([\pi_1, \pi_2, \ldots, \pi_m, m+1, \ldots, n]) = [\pi_1, \pi_2, \ldots, \pi_m]$$

Theorem 5.2. *The restriction function, $\phi : S_{n|m} \to S_m$, is a group isomorphism.*

Proof. Let $\pi, \sigma \in S_{n|m}$. Then $\phi(\pi \circ \sigma) = \phi([\pi_{\sigma_1}, \pi_{\sigma_2}, \ldots, \pi_{\sigma_m}, m+1, \ldots, n]) = [\pi_{\sigma_1}, \pi_{\sigma_2}, \ldots, \pi_{\sigma_m}] = [\pi_1, \pi_2, \ldots, \pi_m] \circ [\sigma_1, \sigma_2, \ldots, \sigma_m] = \phi([\pi_1, \pi_2, \ldots, \pi_m, m+1, \ldots, n]) \circ \phi([\sigma_1, \sigma_2, \ldots, \sigma_m, m+1, \ldots, n]) = \phi(\pi) \circ \phi(\sigma)$. We must also show ϕ is a bijection. Let $\rho, \tau \in S_{n|m}$ such that $\phi(\rho) = \phi(\tau)$. This implies $\phi([\rho_1, \rho_2, \ldots, \rho_m, m+1, \ldots, n]) = \phi([\tau_1, \tau_2, \ldots, \tau_m, m+1, \ldots, n])$ which implies $[\rho_1, \rho_2, \ldots, \rho_m] = [\tau_1, \tau_2, \ldots, \tau_m]$. This implies $\rho_i = \tau_i$ for $1 \leq i \leq m$ which implies $\rho = \tau$. Thus, ϕ is one-to-one. Now let $\upsilon \in S_m$. Then $[\upsilon_1, \upsilon_2, \ldots, \upsilon_m, m+1, \ldots, n] \in S_{n|m}$ with $\phi([\upsilon_1, \upsilon_2, \ldots, \upsilon_m, m+1, \ldots, n]) = [\upsilon_1, \upsilon_2, \ldots, \upsilon_m] = \upsilon$. Thus, ϕ is surjective. □

Corollary 5.3. $|S_{n|m}| = |S_m|$.

Proof. This follows directly from Theorem 5.2. □

Definition Let $m, n \in \mathbb{N}$ with $m < n$ and G_n be a generating set of S_n. The *generating set on n restricted to m*, denoted by $G_{n|m}$, is given by:

$$G_{n|m} = \{\pi \in G_n \mid \pi_i = i \text{ for } m+1 \leq i \leq n\}$$

Theorem 5.4. *Let G_n be any generating set mentioned in Theorem 4.5 and let $\pi \in G_{n|m}$. Then $\phi(\pi) \in G_m$.*

Proof. This is straightforward from the definitions of reversals and transpositions. □

Theorem 5.5. *Let G_n be any generating set mentioned in Theorem 4.5. Then $\phi : G_{n|m} \to G_m$ is a bijection.*

Proof. This is straightforward from Theorem 5.2 and Theorem 5.4. □

Corollary 5.6. $|G_{n|m}| = |G_m|$.

Proof. This follows directly from Theorem 5.5. □

Theorem 5.7. *Let G_n be any generating set mentioned in Theorem 4.5. Then $S_n^1 = G_n$.*

Proof. The identity e_n is not in G_n for any of the generating sets in Theorem 4.5. Thus, by the remark at the end of Section 3, $S_n^1 = G_n$. □

Theorem 5.8. *Let $G = R_n$. Then $|S_n^1| = |S_{n-1}^1| + (n-1)$.*

Proof. Since $S_n^1 = G_n$ and $S_{n-1}^1 = G_{n-1}$ by Theorem 5.7, we must show that $|G_n| = |G_{n-1}| + (n-1)$. Let $\sigma \in G_n$. Then $\sigma = \rho(a, b)$ for $1 \leq a < b \leq n$.

Case 1: $\sigma \in G_{n|n-1}$. By Corollary 5.6, this case has cardinality $|G_{n-1}|$.

Case 2: $\sigma \notin G_{n|n-1}$. This implies $\sigma_n \neq n \Rightarrow b = n \Rightarrow 1 \leq a \leq n-1$. By Lemma 4.1, there are exactly $n-1$ reversals that satisfy this criteria. Thus, this case has cardinality $(n-1)$.

Since G_n is a disjoint union of these two cases, $|G_n| = |G_{n-1}| + (n-1)$. □

Theorem 5.9. *Let $G = PR_n$. Then $|S_n^1| = |S_{n-1}^1| + 1$.*

Proof. Since $S_n^1 = G_n$ and $S_{n-1}^1 = G_{n-1}$ by Theorem 5.7, we must show that $|G_n| = |G_{n-1}| + 1$. Let $\sigma \in G_n$. Then $\sigma = \rho(1, b)$ for $1 < b \leq n$.

Case 1: $\sigma \in G_{n|n-1}$. By Corollary 5.6, this case has cardinality $|G_{n-1}|$.

Case 2: $\sigma \notin G_{n|n-1}$. This implies $\sigma_n \neq n \Rightarrow b = n$. By Lemma 4.1, the only prefix reversal that satisfies this criteria is $\rho(1, n)$. Thus, this case has cardinality 1.

Since G_n is a disjoint union of these two cases, $|G_n| = |G_{n-1}| + 1$. □

Theorem 5.10. *Let $G = T_n$. Then $|S_n^1| = |S_{n-1}^1| + \binom{n}{2}$.*

Proof. Since $S_n^1 = G_n$ and $S_{n-1}^1 = G_{n-1}$ by Theorem 5.7, we must show that $|G_n| = |G_{n-1}| + \binom{n}{2}$. Let $\sigma \in G_n$. Then $\sigma = \tau(a, b, c)$ for $1 \leq a < b < c \leq n+1$.

Case 1: $\sigma \in G_{n|n-1}$. By Corollary 5.6, this case has cardinality $|G_{n-1}|$.

Case 2: $\sigma \notin G_{n|n-1}$. This implies $\sigma_n \neq n \Rightarrow c = n + 1 \Rightarrow 1 \leq a < b \leq n$. By Lemma 4.2, there are exactly $\binom{n}{2}$ transpositions that satisfy this criteria. Thus, this case has cardinality $\binom{n}{2}$.

Since G_n is a disjoint union of these two cases, $|G_n| = |G_{n-1}| + \binom{n}{2}$. □

Theorem 5.11. *Let $G = PT_n$. Then $|S_n^1| = |S_{n-1}^1| + (n - 1)$.*

Proof. Since $S_n^1 = G_n$ and $S_{n-1}^1 = G_{n-1}$ by Theorem 5.7, we must show that $|G_n| = |G_{n-1}| + (n - 1)$. Let $\sigma \in G_n$. Then $\sigma = \tau(1, b, c)$ for $1 < b < c \leq n+1$.

Case 1: $\sigma \in G_{n|n-1}$. By Corollary 5.6, this case has cardinality $|G_{n-1}|$.

Case 2: $\sigma \notin G_{n|n-1}$. This implies $\sigma_n \neq n \Rightarrow c = n + 1 \Rightarrow 2 \leq b \leq n$. By Lemma 4.2, there are exactly $n - 1$ prefix transpositions that satisfy this criteria. Thus, this case has cardinality $(n - 1)$.

Since G_n is a disjoint union of these two cases, $|G_n| = |G_{n-1}| + (n - 1)$. □

Theorem 5.8, Theorem 5.9, Theorem 5.10, and Theorem 5.11 prove recursive formulae for the 1-slice for the four rearrangement models under consideration. These formulae give polynomial equations as shown below.

Theorem 5.12 (Polynomial Formulae for 1-Slice)**.** *If $G = R_n$ then $|S_n^1| = \frac{n^2}{2} - \frac{n}{2}$. If $G = PR_n$ then $|S_n^1| = n - 1$. If $G = T_n$ then $|S_n^1| = \frac{n^3}{6} - \frac{n}{6}$. If $G = PT_n$ then $|S_n^1| = \frac{n^2}{2} - \frac{n}{2}$.*

Proof. We prove this theorem using induction. Note that for each genome rearrangement, $|S_1^1| = 0$.

Case 1: $G = R_n$. We first show that $|S_k^1| = \frac{k^2}{2} - \frac{k}{2}$ holds for $k = 1$. $|S_1^1| = 0 = \frac{1^2}{2} - \frac{1}{2}$ as needed. Now, for $k \geq 2$, we assume that $|S_{k-1}^1| = \frac{(k-1)^2}{2} - \frac{(k-1)}{2}$ holds. Then by Theorem 5.8, $|S_k^1| = |S_{k-1}^1| + (k-1) = \frac{(k-1)^2}{2} - \frac{(k-1)}{2} + (k-1) =$

$\frac{(k^2-2k+1)}{2} - \frac{(k-1)}{2} + (k-1) = \frac{k^2}{2} - \frac{k}{2}$. By induction, $|S_n^1| = \frac{n^2}{2} - \frac{n}{2}$ holds for $n \geq 1$.

Case 2: $G = PR_n$. We first show that $|S_k^1| = k - 1$ holds for $k = 1$. $|S_1^1| = 0 = 1 - 1$ as needed. Now, for $k \geq 2$, we assume that $|S_{k-1}^1| = (k-1) - 1$ holds. Then by Theorem 5.9, $|S_k^1| = |S_{k-1}^1| + 1 = (k-1) - 1 + 1 = k - 1$. By induction, $|S_n^1| = n - 1$ holds for $n \geq 1$.

Case 3: $G = T_n$. We first show that $|S_k^1| = \frac{k^3}{6} - \frac{k}{6}$ holds for $k = 1$. $|S_1^1| = 0 = \frac{1^3}{6} - \frac{1}{6}$ as needed. Now, for $k \geq 2$, we assume that $|S_{k-1}^1| = \frac{(k-1)^3}{6} - \frac{(k-1)}{6}$ holds. Then by Theorem 5.10, $|S_k^1| = |S_{k-1}^1| + \binom{k}{2} = \frac{(k-1)^3}{6} - \frac{(k-1)}{6} + \frac{k^2}{2} - \frac{k}{2} = \frac{(k^3-3k^2+3k-1)}{6} - \frac{(k-1)}{6} + \frac{k^2}{2} - \frac{k}{2} = \frac{k^3}{6} - \frac{k}{6}$. By induction, $|S_n^1| = \frac{n^3}{6} - \frac{n}{6}$ holds for $n \geq 1$.

Case 4: $G = PT_n$. See Case 1 above. □

Lemma 5.13. *If $\sigma \in S_n^k$ and $g \in G_n$, then $k - 1 \leq d_G(\sigma g) \leq k + 1$.*

Proof. $\sigma \in S_n^k$ implies that $\sigma = g_1 g_2 \cdots g_k$ is minimal for some $g_1, g_2, \ldots, g_k \in G_n$. Thus, $\sigma g = g_1 g_2 \cdots g_k g$ which implies that $d_G(\sigma g) \leq k + 1$. For the other half, assume to the contrary that $d_G(\sigma g) < k - 1$. This implies that $\sigma g = h_1 h_2 \cdots h_l$ for some $h_1, h_2, \ldots, h_l \in G_n$ and $l < k - 1$. Thus, $\sigma = h_1 h_2 \cdots h_l g^{-1}$ which implies that $d_G(\sigma) \leq l + 1 < k - 1 + 1 = k$. This contradicts the fact that $\sigma \in S_n^k$. □

Definition Let $f(n) \leq O(g(n))$ denote that there is a positive constant C and an integer n_0 such that for all $n \geq n_0$, $f(n) \leq Cg(n)$.

Theorem 5.14. *Let G be any generating set mentioned in Theorem 4.5. Then, for all $k \geq 0$, $|S_n^k|$ is bounded above by a polynomial in n. Specifically, if $G = R_n$ then $|S_n^k| \leq O(n^{2k})$, if $G = PR_n$ then $|S_n^k| \leq O(n^k)$, if $G = T_n$ then $|S_n^k| \leq O(n^{3k})$, and if $G = PT_n$ then $|S_n^k| \leq O(n^{2k})$.*

Proof. Let $SG = \{\sigma g \mid \sigma \in S_n^k, g \in G_n\}$. If $\pi \in S_n^{k+1}$ then $\pi = g_1 g_2 \cdots g_k g_{k+1}$ is minimal for some $g_1, g_2, \ldots, g_k, g_{k+1} \in G_n$. This implies that $\pi \in SG$ since $g_1 g_2 \cdots g_k \in S_n^k$ and $g_{k+1} \in G_n$. Thus, $S_n^{k+1} \subseteq SG$.

Case 1: $G = R_n$. $|S_n^1| = \frac{n^2}{2} - \frac{n}{2}$ by Theorem 5.12. Using induction, we first show that $|S_n^k| \leq O(n^{2k})$ holds for $k = 1$. $|S_n^1| = \frac{n^2}{2} - \frac{n}{2} \leq O(n^2) = O(n^{2(1)})$ as needed. Now, for $k \geq 2$, we assume that $|S_n^k| \leq O(n^{2k})$. Then $|S_n^{k+1}| \leq |SG| \leq |S_n^k| \cdot |G_n| = |S_n^k| \cdot |S_n^1| \leq O(n^{2k}) O(n^2) = O(n^{2k+2}) = O(n^{2(k+1)})$.

Case 2: $G = PR_n$. $|S_n^1| = n - 1$ by Theorem 5.12. Using induction, we first show that $|S_n^k| \leq O(n^k)$ holds for $k = 1$. $|S_n^1| = n - 1 \leq O(n) = O(n^1)$ as needed. Now, for $k \geq 2$, we assume that $|S_n^k| \leq O(n^k)$. Then $|S_n^{k+1}| \leq |SG| \leq |S_n^k| \cdot |G_n| = |S_n^k| \cdot |S_n^1| \leq O(n^k) O(n^1) = O(n^{k+1})$.

Case 3: $G = T_n$. $|S_n^1| = \frac{n^3}{6} - \frac{n}{6}$ by Theorem 5.12. Using induction, we first show that $|S_n^k| \leq O(n^{3k})$ holds for $k = 1$. $|S_n^1| = \frac{n^3}{6} - \frac{n}{6} \leq O(n^3) = O(n^{3(1)})$ as needed. Now, for $k \geq 2$, we assume that $|S_n^k| \leq O(n^{3k})$. Then $|S_n^{k+1}| \leq |SG| \leq |S_n^k| \cdot |G_n| = |S_n^k| \cdot |S_n^1| \leq O(n^{3k}) O(n^3) = O(n^{3k+3}) = O(n^{3(k+1)})$.

Case 4: $G = PT_n$. $|S_n^1| = \frac{n^2}{2} - \frac{n}{2}$ by Theorem 5.12. Using induction, we first show that $|S_n^k| \leq O(n^{2k})$ holds for $k = 1$. $|S_n^1| = \frac{n^2}{2} - \frac{n}{2} \leq O(n^2) = O(n^{2(1)})$ as needed. Now, for $k \geq 2$, we assume that $|S_n^k| \leq O(n^{2k})$. Then $|S_n^{k+1}| \leq |SG| \leq |S_n^k| \cdot |G_n| = |S_n^k| \cdot |S_n^1| \leq O(n^{2k})O(n^2) = O(n^{2k+2}) = O(n^{2(k+1)})$. □

Conjecture 5.15. *Let G be any generating set mentioned in Theorem 4.5. Then, for all $k \geq 0$, $|S_n^k|$ is given by a polynomial in n for n sufficiently large. Specifically, if $G = R_n$ then $|S_n^k|$ is given by polynomial in n of degree $2k$ for n sufficiently large, if $G = PR_n$ then $|S_n^k|$ is given by polynomial in n of degree k for n sufficiently large, if $G = T_n$ then $|S_n^k|$ is given by polynomial in n of degree $3k$ for n sufficiently large, and if $G = PT_n$ then $|S_n^k|$ is given by polynomial in n of degree $2k$ for n sufficiently large.*

Since $|S_n^0| = 1$, the conjecture holds for $k = 0$. By Theorem 5.12, the conjecture holds for $k = 1$. According to Konstantinova [9], the conjecture holds for $k = 2$ and $k = 3$ for prefix reversals. Further work needs to be done to prove that Conjecture 5.15 holds for all k-slices.

6 Numerical Computations

In this section, we supply computer calculations supporting Conjecture 5.15. We have calculated polynomials for the k-slices under the rearrangement models based on the data provided by Galvão and Dias [4]. Note that the degrees of the polynomials support Conjecture 5.15.

6.1 Prefix Reversal Computations

0-slice: $1\binom{n}{0}$
1-slice: $-1\binom{n}{0} + 1\binom{n}{1}$
2-slice: $2\binom{n}{0} - 2\binom{n}{1} + 2\binom{n}{2}$
3-slice (for $n \geq 3$): $-5\binom{n}{0} + 4\binom{n}{1} - 4\binom{n}{2} + 6\binom{n}{3}$
4-slice (for $n \geq 4$): $-17\binom{n}{0} + 11\binom{n}{1} - 2\binom{n}{2} - 9\binom{n}{3} + 24\binom{n}{4}$
5-slice (for $n \geq 5$): $265\binom{n}{0} - 219\binom{n}{1} + 150\binom{n}{2} - 67\binom{n}{3} - 20\binom{n}{4} + 120\binom{n}{5}$
6-slice (for $n \geq 6$): $-967\binom{n}{0} + 1546\binom{n}{1} - 1616\binom{n}{2} + 1294\binom{n}{3} - 716\binom{n}{4} + 34\binom{n}{5} + 720\binom{n}{6}$
7-slice (for $n \geq 6$): $5037\binom{n}{0} - 9854\binom{n}{1} + 13080\binom{n}{2} - 13701\binom{n}{3} + 11509\binom{n}{4} - 7002\binom{n}{5} + 1286\binom{n}{6} + 5037\binom{n}{7}$

6.2 Reversal Computations

0-slice: $1\binom{n}{0}$
1-slice: $1\binom{n}{2}$
2-slice (for $n \geq 3$): $7\binom{n}{0} - 3\binom{n}{1} + 4\binom{n}{3} + 4\binom{n}{4}$
3-slice (for $n \geq 4$): $310\binom{n}{0} - 211\binom{n}{1} + 130\binom{n}{2} - 65\binom{n}{3} + 16\binom{n}{4} + 70\binom{n}{5} + 35\binom{n}{6}$
4-slice (for $n \geq 5$): $-40924\binom{n}{0} + 26450\binom{n}{1} - 15956\binom{n}{2} + 8700\binom{n}{3} - 4022\binom{n}{4} + 1346\binom{n}{5} + 2\binom{n}{6} + 1379\binom{n}{7} + 413\binom{n}{8}$

6.3 Prefix Transposition Computations

0-slice: $1\binom{n}{0}$
1-slice: $1\binom{n}{2}$
2-slice: $2\binom{n}{3} + 6\binom{n}{4}$
3-slice: $3\binom{n}{4} + 40\binom{n}{5} + 90\binom{n}{6}$
4-slice: $4\binom{n}{5} + 170\binom{n}{6} + 1324\binom{n}{7} + 2520\binom{n}{8}$
5-slice: $5\binom{n}{6} + 527\binom{n}{7} + 11176\binom{n}{8} + 68410\binom{n}{9} + 113400\binom{n}{10}$
6-slice: $3\binom{n}{7} + 1137\binom{n}{8} + 63842\binom{n}{9} + 989244\binom{n}{10} + 513953\binom{n}{11} + 7484400\binom{n}{12}$

6.4 Transposition Computations

0-slice: $1\binom{n}{0}$
1-slice: $1\binom{n}{2} + 1\binom{n}{3}$
2-slice: $1\binom{n}{3} + 8\binom{n}{4} + 18\binom{n}{5} + 11\binom{n}{6}$
3-slice: $1\binom{n}{4} + 26\binom{n}{5} + 209\binom{n}{6} + 656\binom{n}{7} + 841\binom{n}{8} + 369\binom{n}{9}$
4-slice: $45\binom{n}{6}+1198\binom{n}{7}+11156\binom{n}{8}+44324\binom{n}{9}+84987\binom{n}{10}+76917\binom{n}{11}+26251\binom{n}{12}$

7 Coxeter Matrix for Prefix Reversals

Definition Let S be a set and $i, j \in S$. A matrix $m : S \times S \to \{1, 2, \ldots, \infty\}$ is called a *Coxeter matrix* if it satisfies $m(i, j) = m(j, i)$ and $m(i, j) = 1 \Leftrightarrow i = j$.

For our purposes, let $a, b \in \mathbb{N}$ and assume without loss of generality that $a \leq b$. Let $n \geq b$ and $S = PR_n$. Let $\rho_a = \rho(1, a)$ and $\rho_b = \rho(1, b)$ be prefix reversals in PR_n.

Theorem 7.1. *Let $m(\rho_a, \rho_b)$ denote the order of $\rho_a\rho_b$. Then $m : PR_n \times PR_n \to \{1, 2, \ldots, \infty\}$ is a Coxeter matrix.*

Proof. For simplicity, let $m(\rho_a, \rho_b)$ be denoted by $m(a, b)$. We first need to show that $m(a, b) = m(b, a)$. First note that $(\rho_a\rho_b)^{-1} = \rho_b\rho_a$ since $(\rho_a\rho_b)(\rho_b\rho_a) = \rho_a\rho_b^2\rho_a = \rho_a\rho_a = \rho_a^2 = e_n$ and $(\rho_b\rho_a)(\rho_a\rho_b) = \rho_b\rho_a^2\rho_b = \rho_b\rho_b = \rho_b^2 = e_n$. Now let $k = m(a, b)$. Then $(\rho_b\rho_a)^k = ((\rho_a\rho_b)^{-1})^k = ((\rho_a\rho_b)^k)^{-1} = e_n^{-1} = e_n$. Now we have to show that there does not exist an $l < k$ such that $(\rho_b\rho_a)^l = e_n$. Assume to the contrary. Then $((\rho_a\rho_b)^{-1})^l = e_n$ which implies $((\rho_a\rho_b)^l)^{-1} = e_n$. This implies $(\rho_a\rho_b)^l = e_n$ which implies that k is not the order of $\rho_a\rho_b$, a contradiction. Thus, k is the order of $\rho_b\rho_a$ which implies that $m(a, b) = m(b, a)$. We now need to show that $m(a, b) = 1 \Leftrightarrow \rho_a = \rho_b$. $m(a, b) = 1 \Leftrightarrow (\rho_a\rho_b)^1 = e_n \Leftrightarrow \rho_a = \rho_b^{-1} \Leftrightarrow \rho_a = \rho_b$. □

Note that for all $(\rho_a, \rho_b) \in PR_n^2$, $m(a, b) \neq \infty$. Thus, according to [1], the Coxter matrix m determines a group W with the presentation:

$$\begin{cases} \text{Generators: } PR_n \\ \text{Relations: } (\rho_a\rho_b)^{m(a,b)} = e_n, \text{ for all } (\rho_a, \rho_b) \in PR_n^2 \end{cases}$$

In particular, this implies there is an onto homomorphism from W to S_n which sends the generators of W to the prefix reversals in S_n. Since Coxeter groups are well understood, the group W and this homomorphism may be helpful in

studing prefix reversals in S_n. In the rest of this section, we will prove a formula for the entries of the Coxeter matrix. Let $d = b - a$. If $d = 0$, then $\rho_a \rho_b = \rho_a^2 = e$ by Theorem 4.3. Thus, $m(a, b) = 1$. If $d \geq 1$, divide b by d and write $b = qd + r$ according to the Euclidean division algorithm with $0 \leq r < d$. Let $\bar{l}_d$ denote the residue of the integer l modulo d. For $1 \leq l \leq b$, the equivalence class of l, denoted by E_l, is given by:

$$E_l = \{i \in \{1, 2, \ldots, b\} \mid i \in \bar{l}_d\}$$

Let $\alpha = \left\lfloor \frac{r+1}{2} \right\rfloor$ and $\beta = \left\lfloor \frac{d+(r+1)}{2} \right\rfloor$. Let L_α be the set $\{1, 2, \ldots, \alpha\}$, L_β be the set $\{r+1, r+2, \ldots, \beta\}$, and $L = L_\alpha \cup L_\beta$. For $l \in L$, the cycle class of l, denoted by Z_l, is given by:

$$Z_l = \begin{cases} E_l \cup E_{(r+1)-l} & \text{if } l \in L_\alpha \\ E_l \cup E_{(d+1)-l+r} & \text{if } l \in L_\beta \end{cases}$$

Lemma 7.2. $r - \alpha \leq \alpha$.

Proof. Note that $\frac{r}{2} \leq \lfloor \frac{r+1}{2} \rfloor$. This implies $-\lfloor \frac{r+1}{2} \rfloor \leq -\frac{r}{2}$. Thus, $r - \alpha = r - \lfloor \frac{r+1}{2} \rfloor \leq r - \frac{r}{2} = \frac{r}{2} \leq \lfloor \frac{r+1}{2} \rfloor = \alpha$. □

Lemma 7.3. $d + r - \beta \leq \beta$.

Proof. Note that $\frac{d+r}{2} \leq \lfloor \frac{d+(r+1)}{2} \rfloor$. This implies $-\lfloor \frac{d+(r+1)}{2} \rfloor \leq -\frac{d+r}{2}$. Thus, $d + r - \beta = d + r - \lfloor \frac{d+(r+1)}{2} \rfloor \leq d + r - \frac{d+r}{2} = \frac{d+r}{2} \leq \lfloor \frac{d+(r+1)}{2} \rfloor = \beta$. □

Lemma 7.4. *If $l = \alpha$ and r is odd, then $(r + 1) - l = \alpha$.*

Proof. If r is odd then $\alpha = \frac{r+1}{2}$ which implies $l = \frac{r+1}{2}$. Thus, $(r + 1) - l = (r + 1) - \frac{r+1}{2} = \frac{r+1}{2} = \alpha$. □

Lemma 7.5. *If $l = \alpha$ and r is even, then $(r + 1) - l = \alpha + 1$.*

Proof. If r is even then $\alpha = \frac{r}{2}$ which implies $l = \frac{r}{2}$. Thus, $(r + 1) - l = (r + 1) - \frac{r}{2} = \frac{r}{2} + 1 = \alpha + 1$. □

Lemma 7.6. *If $l = \beta$ and $d - r$ is odd, then $(d + 1) - l + r = \beta$.*

Proof. If $d - r$ is odd then $d - r + (2r + 1) = d + (r + 1)$ is even. This implies $\beta = \frac{d+(r+1)}{2}$ which implies $l = \frac{d+(r+1)}{2}$. Thus, $(d + 1) - l + r = (d + 1) - \frac{d+(r+1)}{2} + r = d + (r + 1) - \frac{d+(r+1)}{2} = \frac{d+(r+1)}{2} = \beta$. □

Lemma 7.7. *If $l = \beta$ and $d - r$ is even, then $(d + 1) - l + r = \beta + 1$.*

Proof. If $d - r$ is even then $d - r + (2r + 1) = d + (r + 1)$ is odd. This implies $\beta = \frac{d+r}{2}$ which implies $l = \frac{d+r}{2}$. Thus, $(d + 1) - l + r = (d + 1) - \frac{d+r}{2} + r = (d + r) - \frac{d+r}{2} + 1 = \frac{d+r}{2} + 1 = \beta + 1$. □

Theorem 7.8. *If $i \in \{1, 2, \ldots, b\}$, then $i \in Z_l$ for some $l \in L$. Furthermore, this l is unique.*

Proof. Let j be the unique integer in E_i with $1 \leq j \leq d$. Then $i \in E_j$.

Case 1: $1 \leq j \leq \alpha$. Let $l = j$. This implies $1 \leq l \leq \alpha$ which implies $l \in L_\alpha$. Thus, $Z_l = E_j \cup E_{(r+1)-j} \supseteq E_j$ which implies $i \in Z_l$.

Case 2: $\alpha+1 \leq j \leq r$. Let $l = (r+1)-j$. This implies $1 \leq l \leq r-\alpha$ which implies $1 \leq l \leq \alpha$ by Lemma 7.2. Thus, $l \in L_\alpha$ and $Z_l = E_{(r+1)-j} \cup E_{(r+1)-((r+1)-j)} = E_{(r+1)-j} \cup E_j \supseteq E_j$ which implies $i \in Z_l$.

Case 3: $r+1 \leq j \leq \beta$. Let $l = j$. This implies $r+1 \leq l \leq \beta$ which implies $l \in L_\beta$. Thus, $Z_l = E_j \cup E_{(d+1)-j+r} \supseteq E_j$ which implies $i \in Z_l$.

Case 4: $\beta+1 \leq j \leq d$. Let $l = (d+1)-j+r$. This implies $r+1 \leq l \leq d+r-\beta$ which implies $r+1 \leq l \leq \beta$ by Lemma 7.3. Thus, $l \in L_\beta$ and $Z_l = E_{(d+1)-j+r} \cup E_{(d+1)-((d+1)-j+r)+r} = E_{(d+1)-j+r} \cup E_j \supseteq E_j$ which implies $i \in Z_l$.

Now, we must show that this l is unique.

Case 1: $l \in L_\alpha$. If $l = 1$, then $Z_l = E_1 \cup E_r$. If $l = 2$, then $Z_l = E_2 \cup E_{r-1}$. This pattern continues until $l = \alpha$ at which point $Z_l = E_\alpha \cup E_\alpha = E_\alpha$ if r is odd, or $Z_l = E_\alpha \cup E_{\alpha+1}$ if r is even by Lemma 7.4 and Lemma 7.5. Each Z_l is a union of equivalence classes. Since no equivalence class is a part of more than one Z_l, each Z_l is mutually distinct from the others. Also note that this case accounts for the equivalence classes from E_1 to E_r.

Case 2: $l \in L_\beta$. If $l = r+1$, then $Z_l = E_{r+1} \cup E_d$. If $l = r+2$, then $Z_l = E_{r+2} \cup E_{d-1}$. This pattern continues until $l = \beta$ at which point $Z_l = E_\beta \cup E_\beta = E_\beta$ if $d-r$ is odd, or $Z_l = E_\beta \cup E_{\beta+1}$ if $d-r$ is even by Lemma 7.6 and Lemma 7.7. Each Z_l is a union of equivalence classes. Since no equivalence class is a part of more than one Z_l, each Z_l is mutually distinct from the others. Also note that this case accounts for the equivalence classes from E_{r+1} to E_d.

Since the two cases do not deal with any overlapping equivalence classes, no Z_l from one case overlaps with any Z_l from the other. □

7.1 Boundaries For Alpha

Lemma 7.9. *If $1 \leq l \leq \alpha$ and $q \geq k \geq 1$, then $kd + (r+1) - l \leq b$ and $a + 1 - ((k-1)d + (r+1) - l) \leq a$.*

Proof. Note that $d > 0$, $k \leq q$, $b = qd+r$, and $-l \leq -1$. Thus, $kd+(r+1)-l \leq qd + (r+1) - l = b + 1 - l \leq b + 1 - 1 = b$. Furthermore note that $d > 0$, $-k \leq -1$, and $l \leq r$ since $l \leq \alpha$ and $\alpha \leq r$. Thus, $a+1-((k-1)d+(r+1)-l) = a+1-kd+d-r-1+l \leq a+1-d+d-r-1+l = a-r+l \leq a-r+r = a$. □

Lemma 7.10. *If $1 \leq l \leq \alpha$, then $(r+1) - l \leq b$ and $qd + l > a$.*

Proof. Note that $-l \leq -1$ and $0 \leq qd$ since $0 \leq d$ and $0 \leq q$. Thus, $(r+1)-l \leq (r+1)-1 = r = 0+r \leq qd+r = b$. Furthermore note that $l > 0$ and $-r > -d$. Thus, $qd + l > qd + 0 = qd + r - r > qd + r - d = b - d = a$. □

Lemma 7.11. *If $1 \leq l \leq \alpha$ and $q \geq k \geq 1$, then $kd + l \leq b$ and $a + 1 - ((k - 1)d + l) \leq a$.*

Proof. Note that $d > 0$, $k \leq q$, and $l \leq r$. Thus, $kd + l \leq qd + l \leq qd + r = b$. Furthermore note that $d > 0$, $-k \leq -1$, and $-l \leq -1$. Thus, $a+1-((k-1)d+l) = a+1-kd+d-l \leq a+1-d+d-l = a+1-l \leq a+1-1 = a$. □

Lemma 7.12. *If $1 \leq l \leq \alpha$, then $l \leq b$ and $qd + (r + 1) - l > a$.*

Proof. Note that $l \leq r$ and $r \leq b$ since $r \leq d$ and $d \leq b$. Thus, $l \leq r \leq b$. Furthermore note that $-l > -r - 1$ since $l \leq r$ and $-r > -d$. Thus, $qd+(r+1)-l > qd+(r+1)-r-1 = qd = qd+r-r > qd+r-d = b-d = a$. □

7.2 Boundaries for Beta

Lemma 7.13. *If $r + 1 \leq l \leq \beta$ and $q - 1 \geq k \geq 1$, then $kd + (d + 1) - l + r \leq b$ and $a + 1 - ((k - 1)d + (d + 1) - l + r) \leq a$.*

Proof. Note that $d > 0$, $k \leq q - 1$, and $-l \leq -1$. Thus, $kd + (d + 1) - l + r \leq (q-1)d+(d+1)-l+r = qd-d+d+1-l+r = qd+r+1-l = b+1-l \leq b+1-1 = b$. Furthermore note that $d > 0$, $-k \leq -1$, $l \leq d$ since $l \leq \beta$ and $\beta \leq d$, and $-r \leq 0$. Thus, $a + 1 - ((k - 1)d + (d + 1) - l + r) = a + 1 - kd + d - d - 1 + l - r \leq a+1-d+d-d-1+l-r = a-d+l-r \leq a-d+d-r = a-r \leq a+0 = a$. □

Lemma 7.14. *If $r + 1 \leq l \leq \beta$, then $(d + 1) - l + r \leq b$ and $(q - 1)d + l > a$.*

Proof. Note that $-l \leq -r - 1$ and $0 \leq a$. Thus, $(d + 1) - l + r \leq d + 1 - r - 1 + r = d = d + 0 \leq d + a = b$. Furthermore note that $l > r$. Thus, $(q - 1)d + l > (q - 1)d + r = qd + r - d = b - d = a$. □

Lemma 7.15. *If $r + 1 \leq l \leq \beta$ and $q - 1 \geq k \geq 1$, then $kd + l \leq b$ and $a + 1 - ((k - 1)d + l) \leq a$.*

Proof. Note that $d > 0$, $k \leq (q - 1)$, $l \leq d$, and $0 \leq r$. Thus, $kd + l \leq (q - 1)d + l \leq (q - 1)d + d = qd = qd + 0 \leq qd + r = b$. Furthermore note that $d > 0$, $-k \leq -1$, $-l \leq -1$. Thus, $a + 1 - ((k - 1)d + l) = a + 1 - kd + d - l \leq a + 1 - d + d - l = a + 1 - l \leq a + 1 - 1 = a$. □

Lemma 7.16. *If $r + 1 \leq l \leq \beta$, then $l \leq b$ and $(q - 1)d + (d + 1) - l + r > a$.*

Proof. Note that $l \leq \beta$ and $\beta \leq b$ since $\beta \leq d$ and $d \leq b$. Thus, $l \leq \beta \leq b$. Furthermore note that $-l > -d-1$ since $l \leq d$. Thus, $(q-1)d+(d+1)-l+r = qd - d + d + 1 - l + r > qd - d + d + 1 - d - 1 + r = qd + r - d = b - d = a$. □

7.3 Cycles for Alpha

Lemma 7.17. *If $1 \leq l \leq \alpha$ and $q \geq k \geq 1$, then $(\rho_a\rho_b)_{kd+(r+1)-l} = (k - 1)d + (r + 1) - l$.*

Proof. Keeping in mind Lemma 7.9, $(\rho_a\rho_b)_{kd+(r+1)-l} = (\rho_a)_{b+1-(kd+(r+1)-l)} = (\rho_a)_{a+d+1-(kd+(r+1)-l)} = (\rho_a)_{a+1-((k-1)d+(r+1)-l)} = (k - 1)d + (r + 1) - l$. □

Lemma 7.18. *If $1 \leq l \leq \alpha$, then $(\rho_a\rho_b)_{(r+1)-l} = qd + l$.*

Proof. Keeping in mind Lemma 7.10, $(\rho_a\rho_b)_{(r+1)-l} = (\rho_a)_{b+1-((r+1)-l)} = (\rho_a)_{qd+r+1-r-1+l} = (\rho_a)_{qd+l} = qd+l$. □

Lemma 7.19. *If $1 \leq l \leq \alpha$ and $q \geq k \geq 1$, then $(\rho_a\rho_b)_{kd+l} = (k-1)d+l$.*

Proof. Keeping in mind Lemma 7.11, $(\rho_a\rho_b)_{kd+l} = (\rho_a)_{b+1-(kd+l)} = (\rho_a)_{a+d+1-(kd+l)} = (\rho_a)_{a+1-((k-1)d+l))} = (k-1)d+l$. □

Lemma 7.20. *If $1 \leq l \leq \alpha$, then $(\rho_a\rho_b)_l = qd+(r+1)-l$.*

Proof. Keeping in mind Lemma 7.12, $(\rho_a\rho_b)_l = (\rho_a)_{b+1-l} = (\rho_a)_{qd+(r+1)-l} = qd+(r+1)-l$. □

7.4 Cycles for Beta

Lemma 7.21. *If $r+1 \leq l \leq \beta$ and $q-1 \geq k \geq 1$, then $(\rho_a\rho_b)_{kd+(d+1)-l+r} = (k-1)d+(d+1)-l+r$.*

Proof. Keeping in mind Lemma 7.13,
$(\rho_a\rho_b)_{kd+(d+1)-l+r} = (\rho_a)_{b+1-(kd+(d+1)-l+r)} = (\rho_a)_{a+d+1-(kd+(d+1)-l+r)} = (\rho_a)_{a+1-((k-1)d+(d+1)-l+r)} = (k-1)d+(d+1)-l+r$. □

Lemma 7.22. *If $r+1 \leq l \leq \beta$, then $(\rho_a\rho_b)_{(d+1)-l+r} = (q-1)d+l$.*

Proof. Keeping in mind Lemma 7.14, $(\rho_a\rho_b)_{(d+1)-l+r} = (\rho_a)_{b+1-((d+1)-l+r)} = (\rho_a)_{qd+r+1-d-1+l-r} = (\rho_a)_{(q-1)d+l} = (q-1)d+l$. □

Lemma 7.23. *If $r+1 \leq l \leq \beta$ and $q-1 \geq k \geq 1$, then $(\rho_a\rho_b)_{kd+l} = (k-1)d+l$.*

Proof. Keeping in mind Lemma 7.15, $(\rho_a\rho_b)_{kd+l} = (\rho_a)_{b+1-(kd+l)} = (\rho_a)_{a+d+1-(kd+l)} = (\rho_a)_{a+1-((k-1)d+l)} = (k-1)d+l$. □

Lemma 7.24. *If $r+1 \leq l \leq \beta$, then $(\rho_a\rho_b)_l = (q-1)d+(d+1)-l+r$.*

Proof. Keeping in mind Lemma 7.16, $(\rho_a\rho_b)_l = (\rho_a)_{b+1-l} = (\rho_a)_{qd+r+1-l} = (\rho_a)_{(q-1)d+(d+1)-l+r} = (q-1)d+(d+1)-l+r$. □

7.5 Prefix Reversal Cycles

Theorem 7.25. *Let $i \in \{1,2,\ldots,b\}$. Let j be the unique integer in E_i with $1 \leq j \leq d$. If $1 \leq j \leq r$, then $|E_i| = q+1$. If $r+1 \leq j \leq d$, then $|E_i| = q$.*

Proof. Writing $\{1,2,\ldots,b\}$ as $\{1,\ldots,r,r+1,\ldots,d,d+1,\ldots,d+r,d+(r+1),\ldots,2d,\ldots,(q-1)d+1,\ldots,(q-1)d+r,(q-1)d+(r+1),\ldots,qd,qd+1,\ldots,qd+r\} = \left(\bigcup_{k=0}^{q-1}\{kd+1,\ldots,kd+r,kd+(r+1),\ldots,kd+d\}\right) \cup \{qd+1,\ldots,qd+r\}$ makes the theorem clear. □

Theorem 7.26. *$\rho_a\rho_b : Z_l \to Z_l$ is a bijection for all $l \in L$.*

Proof. It is enough to show that if $i \in Z_l$, $(\rho_a\rho_b)_i \in Z_l$.

Case 1: $l \in L_\alpha$. Then $1 \leq l \leq \alpha$. If $i \in Z_l$, then $i = kd + l$ for some $0 \leq k \leq q$, or $i = kd + (r+1) - l$ for some $0 \leq k \leq q$ by Theorem 7.25. By Lemmas 7.9, 7.10, 7.11, and 7.12, $(\rho_a\rho_b)_i = kd + l$ for some $0 \leq k \leq q$, or $(\rho_a\rho_b)_i = kd + (r+1) - l$ for some $0 \leq k \leq q$. Thus, $(\rho_a\rho_b)_i \in Z_l$.

Case 2: $l \in L_\beta$. Then $r + 1 \leq l \leq \beta$. If $i \in Z_l$, then $i = kd + l$ for some $0 \leq k \leq q - 1$, or $i = kd + (d+1) - l + r$ for some $0 \leq k \leq q - 1$ by Theorem 7.25. By Lemmas 7.13, 7.14, 7.15, and 7.16, $(\rho_a\rho_b)_i = kd + l$ for some $0 \leq k \leq q - 1$, or $(\rho_a\rho_b)_i = kd + (d+1) - l + r$ for some $0 \leq k \leq q - 1$. Thus, $(\rho_a\rho_b)_i \in Z_l$. □

Theorem 7.27. $m(a, b) = \underset{l \in L}{\mathrm{LCM}} |Z_l|$.

Proof. It is well known that the order of any permutation of a finite set written as the product of disjoint cycles is the lowest common multiple of the lengths of the cycles. Note that for $b < i \leq n$, the cycle of $\rho_a\rho_b$ containing i is simply the identity cycle (i). Since these have length 1, they do not contribute to the lowest common multiple. By Theorem 7.8 and Theorem 7.26, the elements of each Z_l correspond to the elements in each (non-identity) cycle of $\rho_a\rho_b$. Since these cycles are disjoint by Theorem 7.8, $m(a, b) = \underset{l \in L}{\mathrm{LCM}} |Z_l|$. □

Theorem 7.28. *If $l \in \{1, 2, \ldots, \alpha - 1\}$ or $l = \alpha$ and r is even, then the cycle for $\rho_a\rho_b$ in Z_l is given by: $(qd + (r+1) - l, (q-1)d + (r+1) - l, \ldots, (r+1) - l, qd + l, (q-1)d + l, \ldots, l)$ and $|Z_l| = 2(q+1)$. If $l = \alpha$ and r is odd, then the cycle for $\rho_a\rho_b$ in Z_l is given by: $(qd + l, (q-1)d + l, \ldots, l)$ and $|Z_l| = q + 1$.*

Proof. Case 1: $l \in \{1, 2, \ldots, \alpha-1\}$ or $l = \alpha$ and r is even. If $l \in \{1, 2, \ldots, \alpha-1\}$, then $l < \alpha$. Thus, $r + 1 - l > r + 1 - \alpha = r + 1 - \lfloor \frac{r+1}{2} \rfloor \geq r + 1 - \frac{r+1}{2} = \frac{r+1}{2} \geq \lfloor \frac{r+1}{2} \rfloor = \alpha > l$. This implies $r + 1 - l \neq l$. If $l = \alpha$ and r is even, then $(r+1) - l = \alpha + 1 = l + 1 \neq l$ by Lemma 7.5. Either way, $E_l \cap E_{r+1-l} = \emptyset$. Thus, $|Z_l| = |E_l \cup E_{r+1-l}| = |E_l| + |E_{r+1-l}| = (q+1) + (q+1) = 2(q+1)$ by Theorem 7.25. By Lemmas 7.17, 7.18, 7.19, and 7.20, we verify that the cycle above is correct.

Case 2: $l = \alpha$ and r is odd. Then $r + 1 - l = \alpha = l$ by Lemma 7.4. Thus, $|Z_l| = |E_l \cup E_{r+1-l}| = |E_l \cup E_l| = |E_l| = q + 1$ by Theorem 7.25. By Lemmas 7.19 and 7.20, we verify that the cycle above is correct. □

Theorem 7.29. *If $l \in \{r+1, r+2, \ldots, \beta-1\}$ or $l = \beta$ and $d - r$ is even, then the cycle for $\rho_a\rho_b$ in Z_l is given by: $((q-1)d+(d+1)-l+r, (q-2)d+(d+1)-l+r, \ldots, (d+1)-l+r, (q-1)d+l, (q-2)d+l, \ldots, l)$ and $|Z_l| = 2q$. If $l = \beta$ and $d-r$ is odd, then the cycle for $\rho_a\rho_b$ in Z_l is given by: $((q-1)d+l, (q-2)d+l, \ldots, l)$ and $|Z_l| = q$.*

Proof. Case 1: $l \in \{r + 1, r + 2, \ldots, \beta - 1\}$ or $l = \beta$ and $d - r$ is even. If $l \in \{r+1, r+2, \ldots, \beta-1\}$, then $l < \beta$. Thus, $(d+1) - l + r > d + 1 - \beta + r = d + 1 - \lfloor \frac{d+(r+1)}{2} \rfloor + r \geq d + 1 - \frac{d+(r+1)}{2} + r = \frac{d+(r+1)}{2} \geq \lfloor \frac{d+(r+1)}{2} \rfloor = \beta > l$. This implies $(d+1) - l + r \neq l$. If $l = \beta$ and $d - r$ is even, then $(d+1) - l + r = \beta + 1 = l + 1 \neq l$ by Lemma 7.7. Either way, $E_l \cap E_{(d+1)-l+r} = \emptyset$. Thus,

$|Z_l| = |E_l \cup E_{(d+1)-l+r}| = |E_l| + |E_{(d+1)-l+r}| = q + q = 2q$ by Theorem 7.25. By Lemmas 7.21, 7.22, 7.23, and 7.24, we verify that the cycle above is correct.

Case 2: $l = \beta$ and $d - r$ is odd. Then $(d + 1) - l + r = \beta = l$ by Lemma 7.6. Thus, $|Z_l| = |E_l \cup E_{(d+1)-l+r}| = |E_l \cup E_l| = |E_l| = q$ by Theorem 7.25. By Lemmas 7.23 and 7.24, we verify that the cycle above is correct. □

Remark Note that Theorem 7.28 and Theorem 7.29 account for all possible lengths of cycles for $\rho_a\rho_b$. Thus, if no l satisfies the conditions of the theorems, then there is no Z_l with the length given by the theorems.

Lemma 7.30. *If $r = 0$, then there are no cycles of length $2(q+1)$ and no cycles of length $q + 1$. If $r = 1$, then there are no cycles of length $2(q + 1)$ and exactly one cycle of length $q + 1$. If $r = 2$, then there is exactly one cycle of length $2(q + 1)$ and no cycles of length $q + 1$. If $r \geq 3$ and r is even, then are are at least two cycles of length $2(q + 1)$ and no cycles of length $q + 1$. If $r \geq 3$ and r is odd, then there is at least one cycle of length $2(q + 1)$ and exactly one cycle of length $q + 1$.*

Proof. Case 1: $r = 0$. Then $\alpha = 0 < 1 \leq l$. So $\{1, 2, \ldots, \alpha - 1\}$ is empty and $l \neq \alpha$. By Theorem 7.28 and the above remark, there are no cycles of length $2(q + 1)$ and no cycles of length $q + 1$.

Case 2: $r = 1$. Then $\alpha = 1$. So $\{1, 2, \ldots, \alpha - 1\}$ is empty and $l = \alpha$ when $l = 1$. Since r is odd, there are no cycles of length $2(q+1)$ and exactly one cycle of length $q + 1$ by Theorem 7.28 and the above remark.

Case 3: $r = 2$. Then $\alpha = 1$. So $\{1, 2, \ldots, \alpha - 1\}$ is empty and $l = \alpha$ when $l = 1$. Since r is even, there is exactly one cycle of length $2(q+1)$ and no cycles of length $q + 1$ by Theorem 7.28 and the above remark.

Case 4: $r \geq 3$ and r is even. Then $\alpha \geq 2$ which implies that $\alpha - 1 \geq 1$. So $\{1, 2, \ldots, \alpha - 1\}$ is nonempty and $l = \alpha$ when $l = \lfloor \frac{r+1}{2} \rfloor$. Since r is even, there are at least two cycles of length $2(q + 1)$ and no cycles of length $q + 1$ by Theorem 7.28 and the above remark.

Case 5: $r \geq 3$ and r is odd. Then $\alpha \geq 2$ which implies that $\alpha - 1 \geq 1$. So $\{1, 2, \ldots, \alpha - 1\}$ is nonempty and $l = \alpha$ when $l = \lfloor \frac{r+1}{2} \rfloor$. Since r is odd, there is at least one cycle of length $2(q+1)$ and exactly one cycle of length $q+1$ by Theorem 7.28 and the above remark. □

Lemma 7.31. *If $d - r = 1$, then there are no cycles of length $2q$ and exactly one cycle of length q. If $d - r = 2$, then there is exactly one cycle of length $2q$ and no cycles of length q. If $d - r \geq 3$ and $d - r$ is even, then are are at least two cycles of length $2q$ and no cycles of length q. If $d - r \geq 3$ and $d - r$ is odd, then there is at least one cycle of length $2q$ and exactly one cycle of length q.*

Proof. Case 1: $d - r = 1$. Then $\beta = r + 1$. So $\{r + 1, r + 2, \ldots, \beta - 1\}$ is empty and $l = \beta$ when $l = r + 1$. Since $d - r$ is odd, there are no cycles of length $2q$ and exactly one cycle of length q by Theorem 7.29 and the above remark.

Case 2: $d - r = 2$. Then $\beta = r + 1$. So $\{r + 1, r + 2, \ldots, \beta - 1\}$ is empty and $l = \beta$ when $l = \lfloor \frac{d+(r+1)}{2} \rfloor$. Since $d - r$ is even, there is exactly one cycle of length $2q$ and no cycles of length q by Theorem 7.29 and the above remark.

Case 3: $d - r \geq 3$ and $d - r$ is even. Then $\beta \geq r + 2$ which implies that $\beta - 1 \geq r + 1$. So $\{r + 1, r + 2, \ldots, \beta - 1\}$ is nonempty and $l = \beta$ when $l = \lfloor \frac{d+(r+1)}{2} \rfloor$. Since $d - r$ is even, there are at least two cycles of length $2q$ and no cycles of length q by Theorem 7.29 and the above remark.

Case 4: $d - r \geq 3$ and $d - r$ is odd. Then $\beta \geq r + 2$ which implies that $\beta - 1 \geq r + 1$. So $\{r + 1, r + 2, \ldots, \beta - 1\}$ is nonempty and $l = \beta$ when $l = \lfloor \frac{d+(r+1)}{2} \rfloor$. Since $d - r$ is odd, there is at least one cycle of length $2q$ and exactly one cycle of length q by Theorem 7.29 and the above remark. □

Theorem 7.32 (Coxeter Matrix for Prefix Reversals)**.** *If $d = 0$, then $m(a, b) = 1$. If $r = 0$ and $d - r = 1$, then $m(a, b) = q$. If $r = 0$ and $d - r \geq 2$, then $m(a, b) = 2q$. If $r = 1$ and $d - r = 1$, then $m(a, b) = q(q + 1)$. If $r = 1$ and $d - r \geq 2$ and q is odd, then $m(a, b) = q(q + 1)$. If $r = 1$ and $d - r \geq 2$ and q is even, then $m(a, b) = 2q(q + 1)$. If $r \geq 2$ and $d - r = 1$ and q is odd, then $m(a, b) = 2q(q+1)$. If $r \geq 2$ and $d-r = 1$ and q is even, then $m(a, b) = q(q+1)$. If $r \geq 2$ and $d - r \geq 2$, then $m(a, b) = 2q(q + 1)$.*

Proof. First note that the cases in this theorem are exhaustive. If $d = 0$, then $m(a, b) = 1$ as shown at the beginning of this section. If $d \geq 1$, then $0 \leq r < d$ by the Euclidean division algorithm. This implies $0 \leq r$ and $1 \leq d - r$. The cases fully account for these ranges.

Case 1: $r = 0$ and $d - r = 1$. By Lemma 7.30 and Lemma 7.31, there are no cycles of length $2(q+1)$, no cycles of length $q+1$, no cycles of length $2q$ and exactly one cycle of length q. Thus, $m(a, b) = \underset{l \in L}{\text{LCM}} |Z_l| = q$.

Case 2: $r = 0$ and $d - r \geq 2$. By Lemma 7.30, there are no cycles of length $2(q+1)$ and no cycles of length $q+1$. We are also in one of the last three cases of Lemma 7.31. In any case, the lowest common multiple for these cycle lengths is $2q$. Thus, $m(a, b) = \underset{l \in L}{\text{LCM}} |Z_l| = 2q$.

Case 3: $r = 1$ and $d - r = 1$. By Lemma 7.30 and Lemma 7.31, there are no cycles of length $2(q+1)$, exactly one cycle of length $q+1$, no cycles of length $2q$ and exactly one cycle of length q. Thus, $m(a, b) = \underset{l \in L}{\text{LCM}} |Z_l| = q(q + 1)$.

Case 4: $r = 1$ and $d - r \geq 2$. By Lemma 7.30, there are no cycles of length $2(q + 1)$ and exactly one cycle of length $q + 1$. We are also in one of the last three cases of Lemma 7.31. In any case, the lowest common multiple for these cycle lengths is $2q$. Thus, $m(a, b) = \underset{l \in L}{\text{LCM}} |Z_l| = \text{LCM}\{(q + 1), 2q\}$. If q is odd, then $m(a, b) = \text{LCM}\{(q + 1), 2q\} = q(q + 1)$. If q is even, then $m(a, b) = \text{LCM}\{(q + 1), 2q\} = 2q(q + 1)$.

Case 5: $r \geq 2$ and $d - r = 1$. We are in one of the last three cases of Lemma 7.30. In any case, the lowest common multiple for these cycle lengths is $2(q + 1)$. By Lemma 7.31, there are no cycles of length $2q$ and exactly

one cycle of length q. Thus, $m(a,b) = \underset{l \in L}{\text{LCM}} |Z_l| = \text{LCM}\{2(q+1), q\}$. If q is odd, then $m(a,b) = \text{LCM}\{2(q+1), q\} = 2q(q+1)$. If q is even, then $m(a,b) = \text{LCM}\{2(q+1), q\} = q(q+1)$.

Case 6: $r \geq 2$ and $d - r \geq 2$. We are in one of the last three cases of Lemma 7.30. In any case, the lowest common multiple for these cycle lengths is $2(q+1)$. We are also in one of the last three cases of Lemma 7.31. In any case, the lowest common multiple for these cycle lengths is $2q$. Thus, $m(a,b) = \underset{l \in L}{\text{LCM}} |Z_l| = \text{LCM}\{2(q+1), 2q\} = 2q(q+1)$. □

8 Acknowledgements

I would like to thank the Joe C. and Carole Kerr McClendon Honors College for reviewing this thesis and granting an honors distinction for my Bachelor of Science in Mathematics. I would also like to give thanks to the Department of Mathematics at the University of Oklahoma for the use of their facilities and computer workstation. I would like to express my gratitude to the University of Oklahoma McNair Scholars Program for allowing me to present this research for McNair conferences at the University of New Mexico, at the University of North Texas, and in Kansas City. My sincere appreciation goes to Sophia Morren for her continuous support throughout my scholastic endeavors. I am deeply indebted to Dr. Kujawa for his invaluable insights, comments, and suggestions that he has given me for the past three years. I am especially grateful for the generous support I received from the National Science Foundation (NSF Grant: DMS-1160763) during the summer of 2013 when much of the work for this thesis was completed.

References

[1] A. Björner and F. Brenti, *Combinatorics of Coxeter Groups*, Graduate Texts in Mathematics, Springer, Berlin, 2005.

[2] J. Cibulka, *On Average and Highest Number of Flips in Pancake Sorting*, Theoretical Computer Science, 412(8-10), 2010, 822-834.

[3] J. Cibulka, *bfs-unb.c* [Computer Program], 2011. Available at `kam.mff.cuni.cz/~cibulka/pancakes/`

[4] Z. Dias and G. R. Galvão, *Computing Rearrangement Distance of Every Permutation in the Symmetric Group*, Proceedings of the ACM Symposium on Applied Computing, 2011, 106-107.

[5] Z. Dias and G. R. Galvão, *Rearrangement Distance Database* [Database], 2011. Available at `mirza.ic.unicamp.br:8080/bioinfo/index.jsf`

[6] T. Dobzhansy and A. H. Sturtevant, *Inversions in the Third Chromosome of Wild Races of Drosophila Pseudoöbscura, and Their Use in the Study of the History of the Species*, Proceedings of the National Academy of Sciences (USA), 22(7), 1936, 448-450.

[7] G. Fertin, A. Labarre, I. Rusu, E. Tannier, and S. Vialette, *Combinatorics of Genome Rearrangements*, The MIT Press, Cambridge, MA, 2009.

[8] D. E. Knuth, *The Art of Computer Programming*, Volume 3: *Sorting and Searching*, 2nd ed., Addison-Wesley, 1995.

[9] E. Konstantinova, *On Some Structural Properties of Star and Pancake Graphs*, Lecture Notes in Computer Science, 7777, 2013, 472-487.

[10] E. Konstantinova and A. Medvedev, *Small Cycles in the Pancake Graph*, Ars Mathematica Contemporanea, 7(1), 2014, 237-246.