THE REVERSAL POSET OF SIGNED PERMUTATIONS

By Fadi Awik

A Thesis

Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Science

in Mathematics

Northern Arizona University

April 2021

Approved:

Dana C. Ernst, Ph.D., Chair

Bahattin Yildiz, Ph.D.

Nándor Sieben, Ph.D.

ABSTRACT

THE REVERSAL POSET OF SIGNED PERMUTATIONS

FADI AWIK

The set of signed permutations $S^{\pm}(n)$ has a fascinating structure. A reversal acting on a permutation $\pi \in S^{\pm}(n)$ reverses the order of elements in consecutive positions and changes their signs. As a group, $S^{\pm}(n)$ is generated by the collection of reversals. The reversal distance of a signed permutation $\pi \in S^{\pm}(n)$ is equal to the minimal number of reversals needed to transform π into the identity permutation. The reversal poset on $S^{\pm}(n)$ is a poset whose elements are signed permutations with covering relations determined by: $u \lessdot v$ if and only if there exists a reversal that transforms v into u and the reversal distance of u is one less than the reversal distance of v. The reversal poset is ranked by reversal distance. We refer to a signed permutation that attains the maximal reversal distance in $S^{\pm}(n)$ as a maximal permutation. These are the permutations of maximal rank in the reversal poset on $S^{\pm}(n)$. It turns out that maximal permutations in $S^{\pm}(n)$ have reversal distance $n+1$ when $n \neq 1, 3$. In this thesis, we derive several results pertaining to the structure of the reversal poset and enumerate permutations of rank $0, 1, 2$, and $n+1$, and obtain partial results for reversal distance n. Our main result is an enumeration of the maximal permutations.

Contents

List of Figures

Chapter 1

Introduction

A *permutation* of the numbers 1 through n is a list of these number with a specific ordering. For example, $[3, 2, 4, 5, 1]$ is a permutation of 1 through 5. A *signed permutation* is a permutation of the numbers 1 through n in which each number is signed. For example, $[4, -5, -2, -1, -3]$ is a signed permutation of the numbers 1 through 5. Signed permutations arise in many areas, both in and outside of mathematics. In particular, signed permutations are useful for modeling large-scale mutations of genomes [2, 7].

A *reversal* of a signed permutation is the act of swapping the order of a consecutive subsequence of numbers and changing the sign of each number in the subsequence. For example, if we perform a reversal involving the first, second third, and fourth, entries in $[4, -5, -2, -1, -3]$ we obtain $[1, 2, 5, -4, -3]$. Given a signed permutation π of the numbers 1 through n, it is always possible to transform π into the identity permutation $[1, 2, \ldots, n]$. This can always be accomplished in a variety of ways and not necessarily using the same number of reversals. In the worst case scenario, we need two reversals to put the value 1 in the first position, two reversals to put the value 2 in the second position, etc. At each step, we need one reversal to put the number in the proper position and potentially a second reversal to change its sign. Therefore, we need at most $2n$ reversals to transform one permutation into the identity.

The *reversal distance* of signed permutation π is the minimum number of reversals required to transform π into the identity permutation. For example, consider the permutation $[4, -5, -2, -1, -3]$. We can transform this permutation into the identity using 3 reversals:

$$[\underline{4, -5, -2, -1}, -3] \to [1, 2, \underline{5}, -4, -3] \to [1, 2, \underline{-5, -4, -3}] \to [1, 2, 3, 4, 5].$$

The first reversal involved the first, second, third, and fourth entries, the second reversal only involved the third entry, and the final reversal involved the last three entries. There are other sequences of reversals that would also work for this example, but what is not immediately obvious is that we could not have obtained the identity permutation in fewer than three reversals. Thus, $[4, -5, -2, -1, -3]$ has reversal distance 3. Since the order in which we apply two reversals often matters, it can be challenging to find an optimal sequence of reversals that transforms a signed permutation into the identity.

The genome of a species can be viewed of as a collection of chromosomes, where each chromosome is an ordered sequences of genes. Each gene has an orientation given by its location on the DNA double strand. Genomes of different species differ from one another. These differences are the result of *point mutations*, in which a single nucleotide is modified, and *genome rearrangements*, where clusters of nucleotides are modified. Genome rearrangements manifest themselves as a shuffling of the genes (possibly inserting new genes or deleting existing genes). Despite the fact that genome rearrangements occur infrequently, over time, continued genome rearrangements cause the order of the genes on a chromosome to become more and more scrambled with respect to the original ordering. Two closely-related species will typically have similar gene orders while the gene orders of two more distant species often differ substantially. Comparing two similar sequences of genes yields two signed permutations, one for each species. Each number in the signed permutation represents either a single gene or a conserved block of genes, where the sign of the number indicates the orientation. Note that we are only considering linear chromosomes (eukaryotes), as opposed to circular ones.

There are several types of genome rearrangements that act on a single chromosome: deletion, insertion, duplication, transposition, and reversal (also called inversion). Each of these rearrangements can be interpreted in terms of signed permutations. The *genetic distance* between two closely-related species is the minimum number of rearrangements necessary to transform one genome into the other. For example, the genomes for cabbage and turnip differ by three reversals while the genomes for a human and a mouse differ by 251 rearrangements, 149 of which are reversals [7]. Since reversals typically make up the bulk of the rearrangements, genetic distance can be approximated by computing the reversal distance between the corresponding signed permutations. For more background on the connection between genome rearrangements and the mathematics of signed permutations being acted on by reversals, see [1, 2, 3, 4, 5, 8, 9].

Except when n is 1 or 3, it turns out that the maximal reversal distance of a signed permutation of the numbers 1 through n is $n + 1$ (see Chapter 3). That is, every signed permutation on the numbers 1 through n can be sorted in at most $n + 1$ reversals. This is substantially better than the worst case scenario of $2n$ that we mentioned above. For each n (not equal to 1 or 3), there are always permutations that attain the upper bound of $n + 1$, which establishes a cap on the genetic distance between two genomes that are related by reversal mutations. We refer to a signed permutation that attains the maximal reversal distance as a *maximal permutation*. It turns out that $[5, 1, 3, 2, 4]$ is an example of a maximal permutation having reversal distance 6.

We define the *reversal poset* on $S^{\pm}(n)$ as a poset with covering relation $\pi \lessdot \gamma$ if there exists a reversal that transforms γ into π such that the reversal lowers the reversal distance by one. Consider our example from earlier. We have the chain

$$[1,2,3,4,5] < [1,2,-5,-4,-3] < [1,2,5,-4,-3] < [4,-5,-2,-1,-3]$$

since we can transform $[4,-5,-2,-1,-3]$ to $[1,2,3,4,5]$ while lowering the reversal distance at each step. It is clear that the reversal poset is ranked by reversal distance. If a signed permutation is maximal in the poset, we say the permutation is *terminal* to avoid confusion with permutations that have maximal reversal distance. Notice that all maximal permutations are terminal. Perhaps surprisingly, there are terminal permutations that are not maximal. For example, the permutation $[2,5,1,3,4]$ is terminal but not maximal having reversal distance 5. Very little is known about terminal yet non-maximal permutations. Genetically speaking, the existence of terminal yet non-maximal permutations implies that certain sequences of reversal mutations can lead to genetic distance topping out prior to attaining the maximum possible genetic distance. In other words, certain mutations can limit the future differences between related species.

In this thesis, our goal is to make progress on enumerating signed permutations of the same rank. Specifically, we tackle the cases of rank $0,1,2$, and $n+1$ for $n \neq 1,3$, and obtain partial results for rank n. In order to explore the reversal poset, we also discuss the *breakpoint diagram* first introduced by Hanenhalli and Pevzner [7]. Hanenhalli and Pevzner give us a formula to calculate reversal distance using the breakpoint diagram, which forms a foundation for understanding how to enumerate maximal permutations. Several of the results in this thesis were inspired by the work that Tanner Rosenberg did during an undergraduate research project. However, most of our proofs are original work. We have indicated when this is not the case throughout the thesis.

Chapter 2

Signed Permutations and Breakpoint Diagrams

We begin by introducing some terminology. Define $S(n)$ to be the group of permutations on $\{1, 2, \ldots, n\}$ and $S^{\pm}(n)$ to be the group of signed permutations on $\{1, 2, \ldots, n\}$. Recall that $|S(n)| = n!$ and $|S^{\pm}(n)| = 2^n n!$. We will represent a signed permutation $\pi \in S^{\pm}(n)$ using one-line notation:

$$\pi = [\pi_1, \pi_2, \ldots, \pi_n],$$

where $\pi_i := \pi(i)$. For $\pi \in S^{\pm}(n)$ and $i \leq j$, we define the *reversal* $\delta_{i,j}$ acting on π via

$$\delta_{i,j}(\pi) = [\pi_1, \ldots, -\pi_j, -\pi_{j-1}, \ldots, -\pi_{i+1}, -\pi_i, \pi_{j+1}, \ldots, \pi_n].$$

That is, $\delta_{i,j}$ takes the elements in positions $i, i+1, \ldots, j-1, j$, reverses their order and changes all the signs. We can consider each reversal $\delta_{i,j}$ as a signed permutation itself. In particular, we have

$$\delta_{i,j} = [1, \ldots, i-1, -j, -(j-1), \ldots, -(i+1), -i, j+1, \ldots, n].$$

Note that $\delta_{i,i}$ is the reversal that changes the sign in the ith position. It is well known that $S^{\pm}(n)$ is generated by the set of reversals $R := \{\delta_{i,j} \mid 1 \leq i \leq j \leq n\}$ (for a fun reference, see [14]), and so every group element can be written as a word in terms of reversals. Every reversal $\delta_{i,j}$ has order 2. When viewing a signed permutation as a product of reversals, we will compose the reversals from right to left. Note that $|R| = \binom{n+1}{2} = T_n$, where T_n is the nth triangular number.

As expected, the order in which we apply two reversals often matters. Two reversals commute if and only if the sets of positions being acted on are disjoint from each other or the two reversals share a common center [14]. Since the set of reversals generate $S^{\pm}(n)$, we can always perform a sequence of reversals on a permutation to obtain the identity permutation $I_n := [1, 2, ..., n-1, n]$. Given a signed permutation π, the *reversal distance*, denoted $d(\pi)$, is defined as the minimum number of reversals needed to transform π to I_n. This is also the minimum number of reversals in a product that yields π.

Example 2.1. Consider the permutation $\pi = [-5, 1, 2, -4, -3, 6, 7] \in S^{\pm}(7)$. Figure 2.1 depicts a sequence of three reversals that transforms π to the identity permutation. Thus, $d(\pi) \leq 3$. It turns out that we cannot use fewer reversals, and so $d(\pi) = 3$. Reversing the sequence of reversals indicated in the figure determines a product that yields π. That is, $\pi = \delta_{4,5}\delta_{2,5}\delta_{1,5}$.

$$\begin{array}{c}
\pi = [-5,\ \ 1,\ \ 2,\ \boxed{-4, -3,}\ \ 6,\ \ 7] \\
\big\downarrow \delta_{4,5} \\
[-5,\ \boxed{1,\ \ 2,\ \ 3,\ \ 4,}\ \ 6,\ \ 7] \\
\big\downarrow \delta_{2,5} \\
[\boxed{-5, -4, -3, -2, -1,}\ \ 6,\ \ 7] \\
\big\downarrow \delta_{1,5} \\
I_7 = [\ 1,\ \ 2,\ \ 3,\ \ 4,\ \ 5,\ \ 6,\ \ 7]
\end{array}$$

Figure 2.1: Sequence of reversals.

Next, we introduce the notion of a breakpoint diagram in order to visualize the structure of a signed permutation. This diagram proves to be useful when computing reversal distance. We mimic the development described in [7] by Hannenhalli and Pevzner, although we have adapted some notation and terminology for our purposes.

Define $S_0(2n)$ to be the set of unsigned permutations on $\{0, 1, 2, \ldots, 2n+1\}$ such that 0 and $2n+1$ are fixed points. We define the *expansion transformation* from a signed permutation $\pi \in S^{\pm}(n)$ to an unsigned permutation $\pi' \in S_0(2n)$ as follows:

$$\pi'_0 = 0, \pi'_{2n+1} = 2n+1,$$

and for all other values, if $\pi_i > 0$, then

$$\pi'_{2i-1} = 2\pi_i - 1, \pi'_{2i} = 2\pi_i,$$

while if $\pi_i < 0$, then

$$\pi'_{2i-1} = 2|\pi_i|, \pi'_{2i} = 2|\pi_i| - 1.$$

Note that the expansion transformation is injective, which implies that the process is uniquely reversible for an unsigned permutation in the image. The *breakpoint diagram* of π, denoted $B(\pi)$, is a graph with colored edges constructed as follows.

1. vertex set: $\{\pi'_0, \pi'_1, \ldots, \pi'_{2n+1}\}$;
2. black edge set: $\{\{\pi'_{2i}, \pi'_{2i+1}\} \mid 0 \leq i \leq n\}$;
3. gray edge set: $\{\{2i, 2i+1\} \mid 0 \leq i \leq n\}$.

By convention, we lay out the vertices $\pi'_0, \ldots, \pi'_{2n+1}$ from left to right, draw the black edges horizontally, and draw the gray edges as arcs above the black edges. Using this convention, every other consecutive pair of vertices is connected by a black edge.

Example 2.2. Figure 2.2 depicts the breakpoint diagram for $\pi = [-5, 1, 3, 2, 4, 6, -7]$. To make $B(\pi)$, we first obtain

$$\pi' = [0, 10, 9, 1, 2, 5, 6, 3, 4, 7, 8, 11, 12, 14, 13, 15],$$

and then organize the vertices determined by π' from left to right and connect with black and gray edges accordingly.

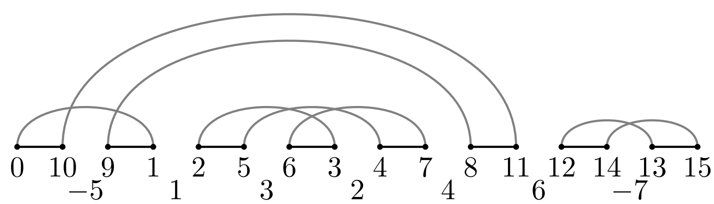


Figure 2.2: Breakpoint diagram for $[-5, 1, 3, 2, 4, 6, -7]$.

Let $\pi \in S^{\pm}(n)$ and let $\pi' \in S_0(2n)$ be its corresponding expansion transformation. If $\{\pi'_i, \pi'_l\}$ is a gray edge such that $|i - l| \geq 1$, we define the *support* of $\{\pi'_i, \pi'_l\}$ to be the interval of elements of π' between, and including, its endpoints. We say that the edge $\{\pi'_i, \pi'_l\}$ is *oriented* if its support contains an odd number of elements, and it is *unoriented* otherwise.

Example 2.3. Refer to π in Example 2.2. The gray edge $\{12, 13\}$ has support $\{12, 14, 13\}$, which has odd cardinality. Thus, $\{12, 13\}$ is oriented. The gray edge $\{2, 3\}$ has support $\{2, 5, 6, 3\}$, which has even cardinality, and hence $\{2, 3\}$ is an unoriented edge.

Visually, an oriented edge occurs if it connects two left vertices of a black edge or two right vertices of a black edge. An edge is unoriented if it connects a left vertex of a black edge and a right vertex of a black edge. Figures 2.3(a) and 2.3(b) depict unoriented edges in blue, and Figures 2.3(c) and 2.3(d) depict oriented edges in green. A cycle in $B(\pi)$ containing at least two gray edges and two black edges is *oriented* if at least one edge in the cycle is oriented, and *unoriented* otherwise. A *trivial cycle* is a cycle that only contains one black edge and one gray edge. By definition, a trivial cycle is neither oriented nor unoriented. A *component* of the breakpoint diagram is a collection of cycles such that each cycle of the component has at least one gray edge that intersects with a gray edge of another cycle in the component (drawn using our convention). A *trivial component* is a component that contains a single trivial cycle. We note that a non-trivial component contains at least two black edges. When a non-trivial component has at least one oriented cycle, it is an *oriented component*. Otherwise, a non-trivial component is *unoriented*. A trivial component is neither oriented nor unoriented.

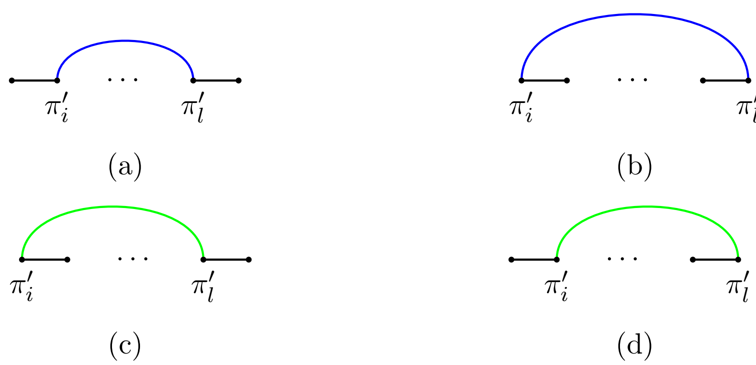


Figure 2.3: Visual representations of unoriented edges (blue) and oriented edges (green).

Example 2.4. Let $\lambda = [-5, 1, 3, 2, 4, 6, -7, 8, 11, 10, 9] \in S^{\pm}(11)$. The breakpoint diagram for λ is given in Figure 2.4, where unoriented edges are colored in blue and oriented edges colored in green. There are two oriented cycles and three unoriented cycles. The two components with no green edges are unoriented, whereas the two components containing green edges are oriented. Looking from left to right, the first three components each contain a single cycle, whereas the last component contains two unoriented cycles that intersect. In summary, there are two oriented components and two unoriented components.

It is important to note that not every seemingly potential breakpoint diagram corresponds to the breakpoint diagram for a signed permutation.

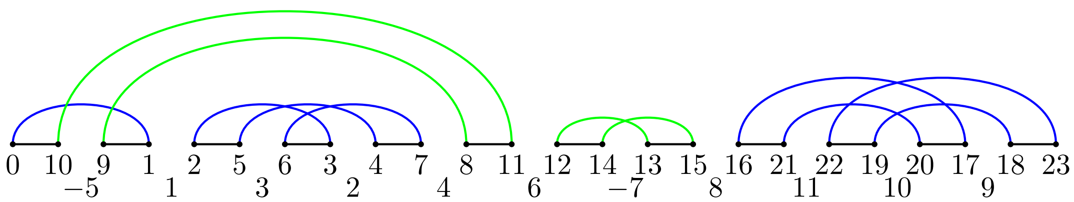


Figure 2.4: Breakpoint diagram of $\lambda = [-5, 1, 3, 2, 4, 6, -7, 8, 11, 10, 9]$ with unoriented edges colored in blue and oriented edges colored in green.

Example 2.5. Consider the diagram in Figure 2.5. This diagram does not represent the breakpoint diagram for a signed permutation since the last vertex in the diagram should be labeled 16 by convention, but the configuration of the gray edges dictates that the last vertex would be labeled by 3.

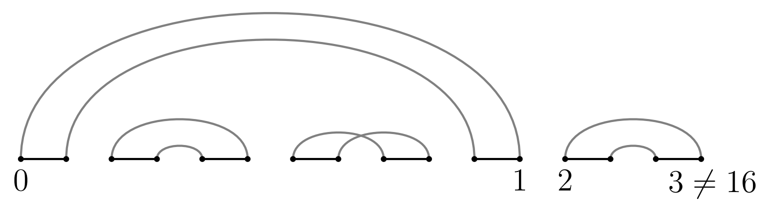


Figure 2.5: Diagram that does not represent a breakpoint diagram.

Let $\pi \in S^{\pm}(n)$. We say that a component U_1 of $B(\pi)$ *covers* component U_2 if all black edges of U_2 appear between two black edges of U_1 in the breakpoint diagram. Covers is a transitive relation.

Example 2.6. In Figure 2.6, the component U_1 covers U_2. In Figure 2.7, the component U_2 covers both U_1 and U_3.

Let $\pi \in S^{\pm}(n)$. An unoriented component U_2 *separates* two other unoriented components U_1 and U_3 when there are black edges of U_2 occurring between U_1, and U_3 and U_2 covers U_1 or U_3 (possibly both). Note that an unoriented component does not separate two other unoriented components if, ignoring black edges for oriented components, all of its black edges appear consecutively in the breakpoint diagram or if all of its black edges occur only on the far left and far right of the breakpoint diagram.

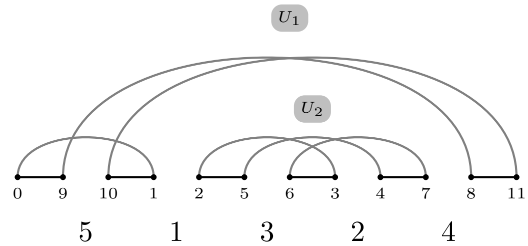


Figure 2.6: Breakpoint diagram where component U_1 covers component U_2.

Example 2.7. Consider the breakpoint diagram in Figure 2.7 with unoriented components U_1, U_2, and U_3. Component U_2 separates components U_1 and U_3, since the black edges of U_2 occur between U_1 and U_3 and U_2 covers both U_1 and U_3. In the breakpoint diagram in Figure 2.9, the unoriented components U_2, U_4, and U_6 do not separate two other unoriented components, since all of their black edges appear consecutively in the breakpoint diagram. However, the component U_1 separates U_2 and U_3, for example.

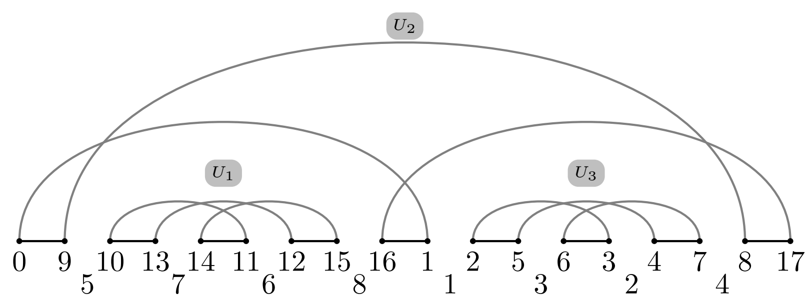


Figure 2.7: Breakpoint diagram that has a component that separates.

An unoriented component U of $B(\pi)$ is a *hurdle* if it does not separate two other unoriented components and it either covers all other unoriented components or does not cover any unoriented components. If U is a hurdle that covers all unoriented components, we call U a *maximal hurdle*. If U is a hurdle that does not cover any unoriented components, U is a *minimal hurdle*. Note that there is at most one maximal hurdle, but there could be several minimal hurdles.

Example 2.8. In Figure 2.8, we have two unoriented components (U_1 and U_2) and one oriented component (U_3). We see that U_1 covers all of the other unoriented components

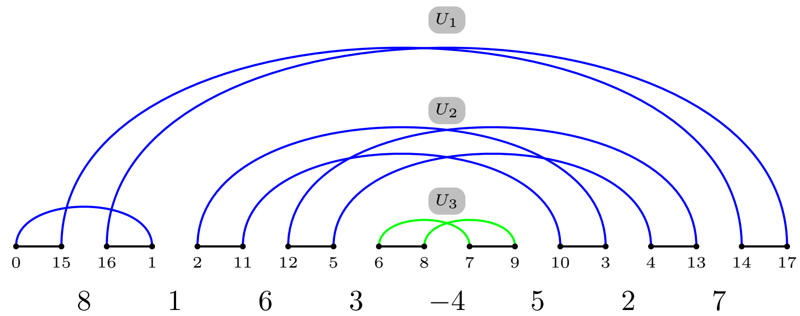


Figure 2.8: Breakpoint diagram with two hurdles.

and U_2 covers no other unoriented component, and neither component separates two other unoriented components. Thus, U_1 and U_2 are both hurdles. Specifically, U_1 is a maximal hurdle and U_2 is a minimal hurdle. In Figure 2.9, the components U_2, U_4, and U_6 are each minimal hurdles and there is no maximal hurdle.

For $\pi \in S^{\pm}(n)$, we define $c(\pi)$ and $h(\pi)$ to be the number of cycles and hurdles in $B(\pi)$, respectively.

Example 2.9. Let $\pi = [-5, 1, 3, 2, 4, 6, -7]$. From its breakpoint diagram depicted in Figure 2.2, we can infer that $c(\pi) = 3$ and $h(\pi) = 1$. Consider $\lambda = [-5, 1, 3, 2, 4, 6, -7, 8, 11, 10, 9]$, whose breakpoint diagram is given in Figure 2.4. Then $c(\lambda) = 5$ and $h(\lambda) = 2$.

We say that a hurdle U of $B(\pi)$ is a *superhurdle* if there exists an unoriented non-hurdle T such that the removal of U from $B(\pi)$ results in T becoming a hurdle.

Example 2.10. Consider $\pi = [5, 1, 3, 2, 4, 6, 8, 10, 9, 11, 7, 12, 17, 13, 15, 14, 16] \in S^{\pm}(16)$, whose breakpoint diagram is displayed in Figure 2.9. Observe that $B(\pi)$ contains three superhurdles (components U_2, U_4, and U_6) and three unoriented non-hurdles (components U_1, U_3, and U_5). For instance, component U_2 is a superhurdle because removing U_2 results in U_1 becoming a hurdle, since it would not cover any other unoriented component if U_2 was removed.

We say $\pi \in S^{\pm}(n)$ is a *fortress* if it contains an odd number of superhurdles and each hurdle in the breakpoint diagram is a superhurdle.

Example 2.11. Since the permutation given in Example 2.10 (see Figure 2.9) has three hurdles, all of which are superhurdles, π is a fortress.

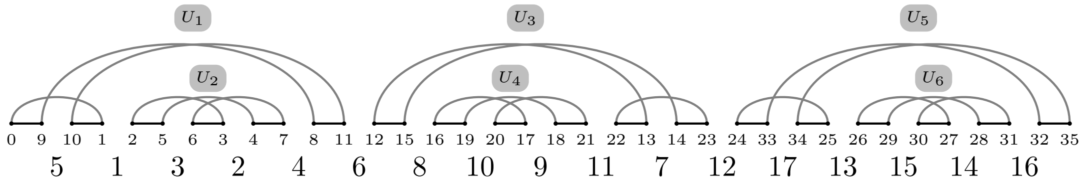


Figure 2.9: Breakpoint diagram for a fortress.

Notice that hurdles and superhurdles are properties of components in a breakpoint diagram, while being a fortress is a property of a permutation.

We can use the following result of Hannenhalli and Pevzner [7] to compute the reversal distance of a signed permutation using its breakpoint diagram.

Theorem 2.12. The reversal distance of any signed permutation $\pi \in S^{\pm}(n)$ is given by

$$d(\pi) = n + 1 - c(\pi) + h(\pi) + f(\pi)$$

where $f(\pi)$ is 1 if π is a fortress and 0 otherwise.

Bergeron, Mixtacki, and Stoye [2] present a different approach to reversal distance that does not involve hurdles and fortresses. However, the original description introduced by Hannenhalli and Pevzner [7] is useful for our purposes.

Example 2.13. Let π be the permutation from Examples 2.9 and 2.10. We see that $n = 17, c(\pi) = 6, h(\pi) = 3$, and $f(\pi) = 1$, which means

$$d(\pi) = 17 + 1 - 6 + 3 + 1 = 16.$$

Let λ be the permutation from Example 2.8. We conclude that $n = 8, c(\lambda) = 3, h(\lambda) = 2$, and $f(\lambda) = 0$, meaning

$$d(\lambda) = 8 + 1 - 3 + 2 + 0 = 8.$$

It is also useful to see how the structure of breakpoint diagrams changes as we sort a permutation to the identity.

Example 2.14. Let $\pi = [-5, 1, 2, -4, -3, 6, 7]$ be the permutation from Example 2.1. Figure 2.10 visualizes what happens as we transform π to the identity using an optimal sequence of reversals, which happens to be the same sequence of reversals given in Example 2.1. Since the sequence of reversals is optimal, we must lower the reversal distance by 1 at each step.

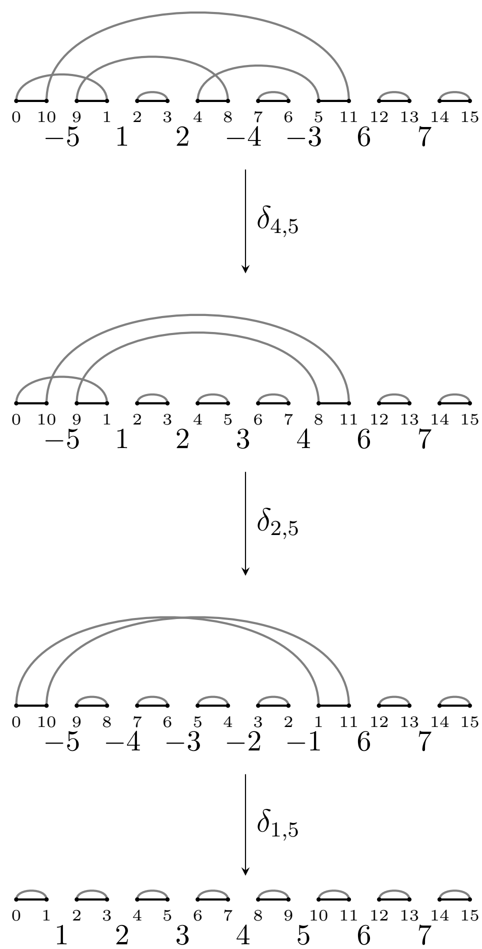


Figure 2.10: Sequence of reversals in terms of the corresponding breakpoint diagrams.

Chapter 3

Maximal Permutations

In this chapter, we discuss the breakpoint diagram in further detail and then use it to prove several results concerning fortresses and permutations that attain the maximal reversal distance in $S^{\pm}(n)$. To start, we introduce the *circular breakpoint diagram* of a permutation, which uses the same convention as the breakpoint diagram defined previously, but wrapped in a circle. We place 0 and $2n+1$ on the top of the circle (with $2n+1$ appearing to the left of 0), and all gray edges are drawn inside the circle.

Example 3.1. Consider the signed permutation $\pi = [8, 1, 6, 3, -4, 5, 2, 7]$. The breakpoint diagram for this permutation is displayed in Figure 2.8, while Figure 3.1 displays the circular breakpoint diagram of π.

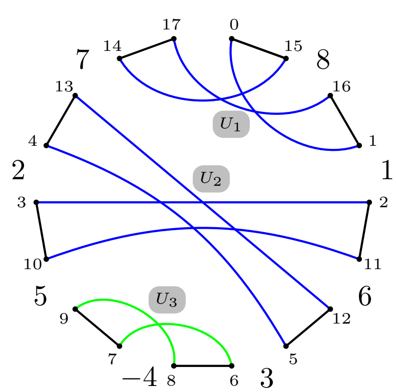


Figure 3.1: Circular breakpoint diagram for the permutation $[8, 1, 6, 3, -4, 5, 2, 7]$.

Notice that the notion of a component covering all other components is no longer sensible. In the circular breakpoint diagram for a permutation, an unoriented component U is a hurdle if all black edges of U appear consecutively along the circle while ignoring any oriented components that may appear between black edges of U. In Example 3.1, components U_1 and U_2 are hurdles, since all black edges appear consecutively along the circle while ignoring the oriented component U_3.

A *cyclic shift* of the circular breakpoint diagram for a permutation is done by moving all black edges counterclockwise while preserving the connections of the gray and black edges between vertices and then relabeling the vertices in the following way:

1. Label the vertices at the top of the circle 0 and $2n+1$ with 0 on the right and $2n+1$ on the left. The vertex connected to 0 by a gray edge is labeled 1. If 1 is the most clockwise vertex of a black edge, the next vertex clockwise is labeled 2. Otherwise the next vertex counterclockwise is labeled 2.

2. The vertex connected to 2 by a gray edge is labeled 3. If 3 is the most clockwise vertex of a black edge, the next vertex clockwise is labeled 4. Otherwise the next vertex counterclockwise is labeled 4.

3. Continue this process until all vertices are labeled.

Since we started with the breakpoint diagram for a permutation, it is clear that this labeling is possible. In this context, a cyclic shift cyclically permutes the breakpoint diagram counterclockwise. The cyclic shift of a breakpoint diagram for $\pi \in S^{\pm}(n)$ is denoted by $\text{shift}(B(\pi))$.

Example 3.2. Consider the permutation $\pi = [8, 1, 6, 3, -4, 5, 2, 7] \in S^{\pm}(8)$ whose breakpoint diagram is given on the left in Figure 3.2. If we perform a cyclic shift, we obtain the diagram on the right in Figure 3.2, which corresponds to the permutation $[2, 7, 4, -5, 6, 3, 8, 1]$. Observe that both permutations have reversal distance 7. Note that the resulting permutation is not simply a cyclic shift of π.

Since we are always able to label the vertices of $\text{shift}(B(\pi))$ for $\pi \in S^{\pm}(n)$ as described above, we can reverse the expansion transformation to obtain a permutation whose breakpoint diagram is $\text{shift}(B(\pi))$. In light of this, we define $\text{shift}(\pi)$ to be the resulting permutation represented by $\text{shift}(B(\pi))$. Moreover, the cyclic shift of $B(\pi)$ preserves orientation and maintains the number of hurdles, superhurdles, components, and cycles. This implies that $c(\pi) = c(\text{shift}(\pi))$, $h(\pi) = h(\text{shift}(\pi))$, and π is a fortress if and only if $\text{shift}(\pi)$ is a fortress. This implies that $d(\pi) = d(\text{shift}(\pi))$. We summarize these results in the following theorem.

Theorem 3.3. If $\pi \in S^{\pm}(n)$, then $\text{shift}(B(\pi))$ is the breakpoint diagram for a signed permutation in $S^{\pm}(n)$, denoted $\text{shift}(\pi)$. Moreover, $d(\pi) = d(\text{shift}(\pi))$.

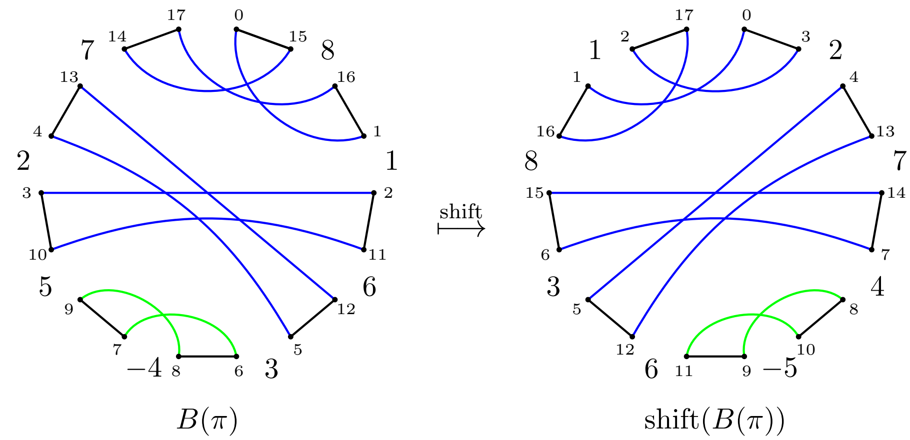


Figure 3.2: Cyclic shift of a circular breakpoint diagram.

We now discuss how to perform a cyclic shift of our typical breakpoint diagrams. Let $\pi \in S^{\pm}(n)$ and consider $B(\pi)$. Let $b_1, \ldots, b_{n+1}$ denote the black edges of $B(\pi)$ when looking at the breakpoint diagram from left to right. In this context, the cyclic shift of $B(\pi)$ is the diagram obtained by shifting b_i to b_{i-1} (mod $n+1$) while preserving the connections of the gray and black edges between vertices and then relabeling the vertices accordingly.

Example 3.4. Consider $\pi = [2, 1, 3, 5, 4, 6, 8, 7] \in S^{\pm}(8)$, whose breakpoint diagram is given in Figure 3.3. When applying a cyclic shift to $B(\pi)$, the result is $\mathrm{shift}(B(\pi))$, where each black edge in $B(\pi)$ is cyclically permuted to the left. The connections of black and gray edges in $B(\pi)$ have been preserved in $\mathrm{shift}(B(\pi))$. It turns out that $\mathrm{shift}(\pi) = [8, 1, 3, 2, 4, 6, 5, 7]$.

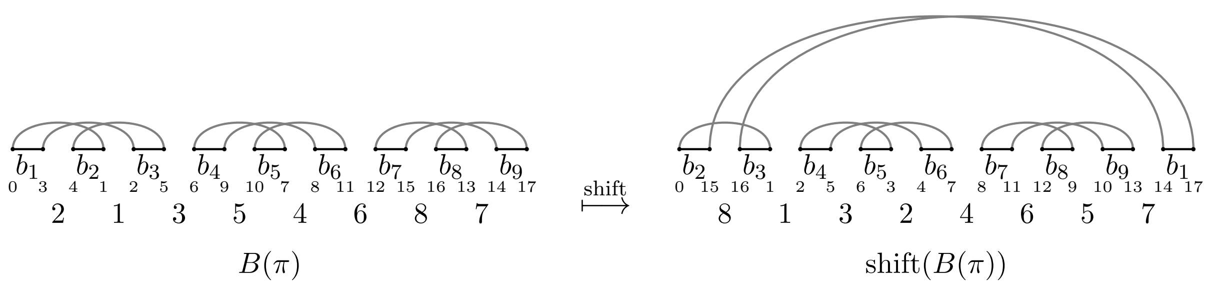


Figure 3.3: Cyclic shift of a breakpoint diagram.

For $\pi, \gamma \in S^{\pm}(n)$, define $\pi \sim \gamma$ if we can obtain $B(\gamma)$ from $B(\pi)$ by a sequence of cyclic shifts. It is readily seen that $\sim$ is an equivalence relation. If $\pi \sim \gamma$, we say that π and γ are *shift equivalent*. Define the *shift equivalence class* of $\pi \in S^{\pm}(n)$ via

$$[\pi] = \{\gamma \in S^{\pm}(n) \mid \gamma \sim \pi\}.$$

Example 3.5. Consider the permutation $\pi = [2,1,3,5,4,6,9,7,10,8] \in S^{\pm}(10)$. There are 11 black edges in $B(\pi)$, which implies the shift equivalence class of π contains at most ten additional permutations. It turns out that there are exactly 10 additional permutations. The breakpoint diagrams for these cyclic shifts are displayed in Figure 3.4, and each permutation corresponding to these breakpoint diagrams are in the shift equivalence class of π.

Here are some interesting results about fortresses and reversal distance that become relevant later. A few of these results are likely known, but we could not find references in the literature. The next two lemmas follow immediately from the definitions.

Lemma 3.6. If S is a superhurdle of $B(\pi)$ for $\pi \in S^{\pm}(n)$, then there exists a unique unoriented component P such that P is not a hurdle, and either

1. S covers P if S is a maximal superhurdle, or

2. P covers S if S is a minimal superhurdle,

and there does not exist an unoriented component between P and S.

In Lemma 3.6, P is the component that becomes a hurdle when S is removed.

Lemma 3.7. If U is an unoriented component of $B(\pi)$ for $\pi \in S^{\pm}(n)$ that covers at least one other unoriented component, then there exists an unoriented component P covered by U, where P covers no other unoriented component. In this case, P is a minimal hurdle.

Loosely speaking, Lemma 3.7 tells us that for every unoriented component that covers at least one other unoriented component, there exists at least one minimal hurdle at the "bottom of the pile." The rest of this chapter up to Theorem 3.16 is inspired by results or conjectures obtained by Tanner Rosenberg during an undergraduate research project. However, we have resequenced the results and retooled most of the proofs.

Theorem 3.8. If $\pi \in S^{\pm}(n)$ is a fortress, then π has at least three superhurdles.

Proof. Let $\pi \in S^{\pm}(n)$ be a fortress such that $B(\pi)$ has k superhurdles. By definition, k is odd and every hurdle is a superhurdle. For sake of a contradiction, assume $k = 1$. Suppose the unique superhurdle (and hence unique hurdle) of $B(\pi)$ is S. Note that S is the only hurdle in $B(\pi)$. We consider two cases.

First, assume S is a minimal superhurdle of $B(\pi)$. By Lemma 3.6, there exists a unique unoriented non-hurdle M_1 such that M_1 covers S and there does not exist an unoriented

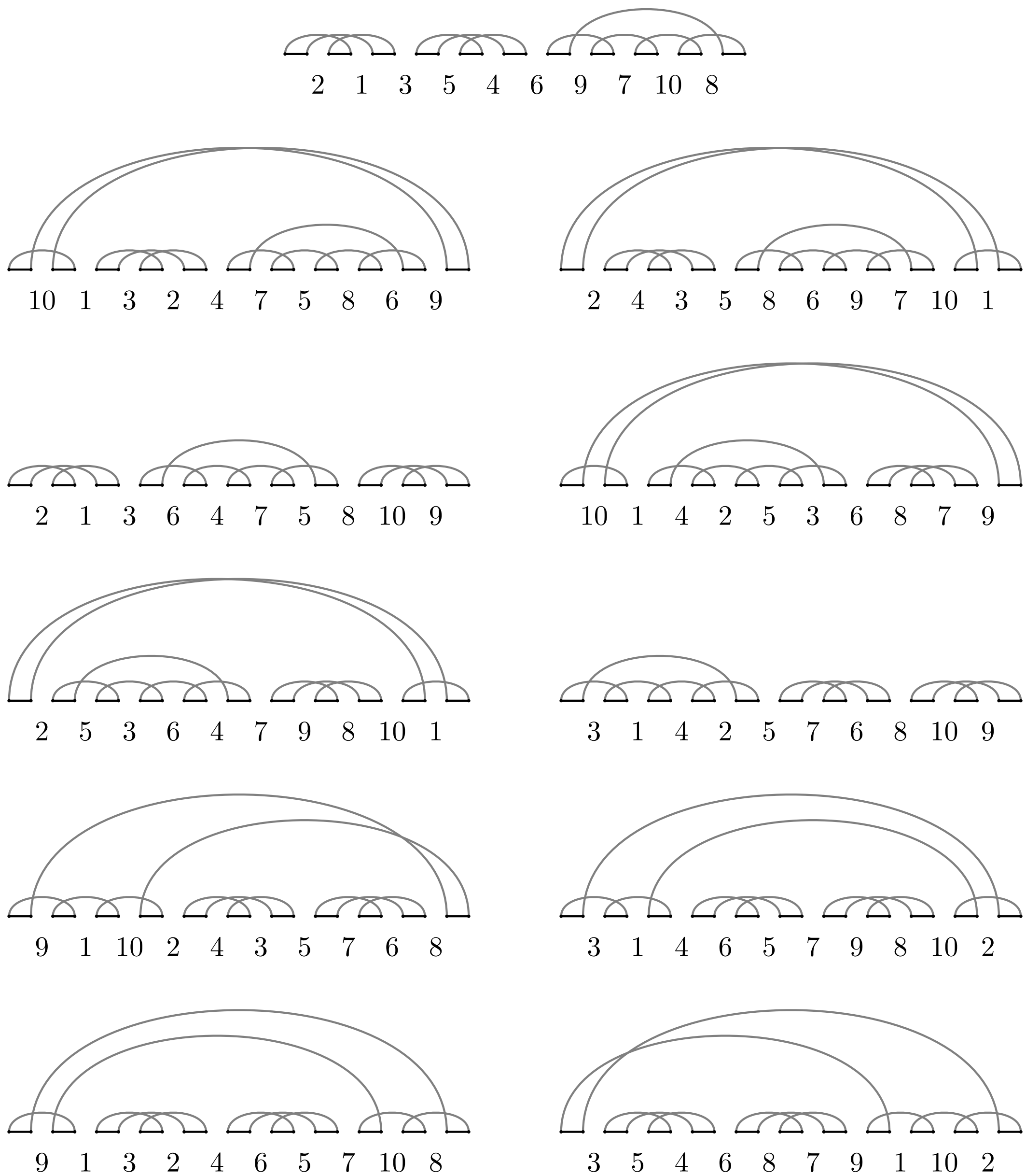


Figure 3.4: The breakpoint diagrams for the shift equivalence class of the permutation $[2, 1, 3, 4, 6, 9, 7, 10, 8]$.

component between M_1 and S. Since M_1 is not a hurdle, there exists an unoriented component $M_2 \neq M_1$ that either appears entirely on the right or left of M_1, or covers M_1. Assume M_2 appears to the right of M_1. The case involving left is similar. Since π is a fortress with S being the unique superhurdle, M_2 cannot be a hurdle. Thus, it must cover another unoriented component. By Lemma 3.7, there exists a minimal hurdle P that is covered by M_2, a contradiction. Now, assume M_2 covers M_1. Since M_2 cannot be a hurdle, and we have expelled the possibility of an unoriented component appearing to the right or left of M_1, there exists an unoriented component Q that covers all other unoriented components. This implies Q is a hurdle, again a contradiction.

For the second case, assume S is a maximal superhurdle of $B(\pi)$. By Lemma 3.7, there exists a minimal hurdle P that is covered by S. This implies P is a hurdle, which once again is a contradiction.

Therefore, $k \neq 1$, so $k \geq 3$ and odd. □

Theorem 3.9. If $\pi \in S^{\pm}(n)$ is a fortress, then the number of superhurdles is less than or equal to the number of unoriented non-hurdles of $B(\pi)$.

Missing case: P_i might cover S_i.

Proof. Let $\pi \in S^{\pm}(n)$ be a fortress, where $B(\pi)$ has $k \geq 3$ superhurdles, say $S_1, \ldots, S_k$. By Lemma 3.6, for each superhurdle S_i of $B(\pi)$, there exists a unique unoriented component P_i covered by S_i such that there is no other unoriented component between S_i and P_i. We know each P_i is not a hurdle. Therefore, there are at least k unoriented non-hurdles of $B(\pi)$, completing the proof. □

Some consequences of the previous result are given below.

Corollary 3.10. If $\pi \in S^{\pm}(n)$ is a fortress, then $c(\pi) \geq 2h(\pi)$.

Proof. Consider a fortress $\pi \in S^{\pm}(n)$ with $k \geq 3$ superhurdles, say $S_1, \ldots, S_k$. By Theorem 3.9, there are at least k unoriented non-hurdle components. Thus there are at least $2k$ cycles, one for each superhurdle and one for each unoriented component. Therefore, $c(\pi) \geq 2h(\pi)$. □

Corollary 3.11. If $\pi \in S^{\pm}(n)$ is a fortress, then $d(\pi) \leq n - 1$.

Proof. From Theorem 2.12, the reversal distance for $\pi \in S^{\pm}(n)$ is

$$d(\pi) = n + 1 - c(\pi) + h(\pi) + f(\pi).$$

Thus for a fortress $\pi \in S^{\pm}(n)$, we see that

$$\begin{aligned}
d(\pi) &= n + 1 - c(\pi) + h(\pi) + 1 \\
&= n + 2 - c(\pi) + h(\pi) \\
&\leq n + 2 - 2h(\pi) + h(\pi) && \text{(by Corollary 3.10)} \\
&= n + 2 - h(\pi) \\
&\leq n - 1 && \text{(since } h(\pi) \geq 3 \text{ by Theorem 3.8)}
\end{aligned}$$

as desired. □

By examining the definition of fortress together with the structure of the breakpoint diagram, it follows from Theorem 3.8 that the breakpoint diagram for a fortress requires at least 18 black edges. Hence there are no fortresses in $S^{\pm}(n)$ for $n \leq 16$, so we know $n-1$ is not a sharp bound on reversal distance for small n. We are not sure if this bound on reversal distance for fortresses is a sharp bound for all $n \geq 17$, but evidence suggests it might be.

Theorem 3.12. Suppose $\pi \in S^{\pm}(n)$. Then π includes at least one negative number if and only if $B(\pi)$ contains an oriented cycle.

Proof. Assume $\pi \in S^{\pm}(n)$ includes at least one negative number. Choose the smallest $|\pi_i|$ such that $\pi_i < 0$. Suppose $\pi_i = -k$ for $k \in [n]$. We consider two cases.

Assume $k = 1$. A gray edge connected to 0 must connect to $\pi'_{2i} = 1$ in $B(\pi)$, which is the left endpoint of a black edge since $|\pi_i| < 0$. Therefore, the gray edge is oriented, implying $B(\pi)$ has an oriented cycle.

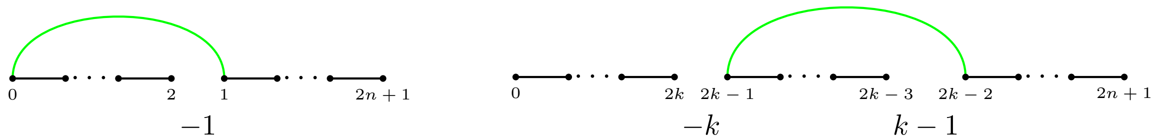
For the second case, assume $k > 1$. Since $|\pi_i|$ is the smallest value for which π is negative, there exists π_j such that $\pi_j = k - 1$. So, $\pi'_{2i} = 2k - 1$ and $\pi'_{2j} = 2k - 2$ are left endpoints of black edges in $B(\pi)$. Thus, the gray edge connecting $2k-1$ and $2k-2$ is oriented, implying $B(\pi)$ contains an oriented cycle.

Figure 3.5 displays both cases, where we have assumed $j > i$ in Figure 3.5(b). In either case, $B(\pi)$ contains an oriented cycle.

It is well known (for example, see [6]) that the breakpoint diagrams for unsigned permutations consist only of unoriented edges, which handles the reverse implication. □

We call a signed permutation $\pi \in S^{\pm}(n)$ *maximal* if it has maximal reversal distance. Before we can prove a result about the reversal distance for maximal permutations, it will be important to show that a maximal permutation is never a fortress.

Theorem 3.13. If $\pi \in S^{\pm}(n)$ is a maximal permutation, then π is not a fortress.



(a) (b)

Figure 3.5: Breakpoint diagrams to display the two cases of the proof of Theorem 3.12.

Proof. For each n, we exhibit a permutation with reversal distance greater than $n-1$. We will handle the cases involving $n = 1$ and $n = 3$ separately.

If $n = 1$, then $[-1]$ is a permutation with reversal distance $1 = n > n-1$. If $n = 3$, then $[2, -3, 1]$ is a permutation with reversal distance $3 = n > n-1$. In either case, there exists a permutation with reversal distance n when $n = 1, 3$.

Now, assume $n \neq 1, 3$. If n is even, let $m = \frac{n}{2}$, and consider the permutation

$$[n, 1, m+1, 2, m+2, 3, m+3, \cdots, m-1, 2m-1, m].$$

If n is odd, let $m = \frac{n-1}{2}$, and consider the permutation

I don't think this is true for n odd?

$$[n, 1, m+1, 2, m+2, 3, \cdots, m, 2m].$$

Both permutations consist of a single unoriented cycle (according to the OEIS entry A131209 [10], see Figure 3.6 for an example with $n = 6$), and so its reversal distance is greater than or equal to n by Theorem 2.12. But then by Corollary 3.11, neither is a fortress, and hence each has reversal distance equal to $n+1$. □

It follows from the proof of Theorem 3.13 that every maximal permutation in $S^{\pm}(n)$ must have distance at least n when $n = 1, 3$ and at least $n+1$ when $n \neq 1, 3$.

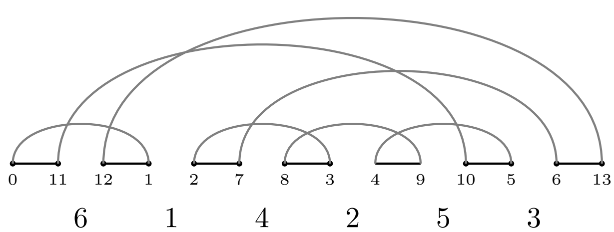


Figure 3.6: Breakpoint Diagram for the permutation $[6, 1, 4, 2, 5, 3]$.

Lemma 3.14. If $\pi \in S^{\pm}(n)$ is maximal with $n \neq 1, 3$, then $c(\pi) = h(\pi)$.

Proof. Assume $\pi \in S^{\pm}(n)$ with $n \neq 1, 3$ is maximal. We know that π is not a fortress by Theorem 3.13, so $d(\pi) = n + 1 - c(\pi) + h(\pi)$. Since each hurdle is a component, and components can be comprised of multiple cycles, $h(\pi) \leq c(\pi)$. Therefore, we need $c(\pi) = h(\pi)$ in order to maximize $d(\pi)$. Indeed, we exhibited permutations in the proof of Theorem 3.13 that achieve this. $\square$

One consequence of Lemma 3.14 is that if $\pi \in S^{\pm}(n)$ is maximal with $n \neq 1, 3$, then each component in $B(\pi)$ is a hurdle and consists of a single cycle.

Theorem 3.15. If $\pi \in S^{\pm}(n)$ is a maximal permutation with $n \neq 1, 3$, then it cannot include a negative number.

remove "unoriented" here.

Proof. Assume π is a maximal permutation in $S^{\pm}(n)$ for $n \neq 1, 3$. By Lemma 3.14, $c(\pi) = h(\pi)$. Therefore, all unoriented cycles are hurdles, and so $B(\pi)$ has no oriented cycles. Thus, by Theorem 3.12, π does not contain a negative entry. $\square$

We note that the converse of Theorem 3.15 is false. For example, consider π from Examples 2.10 and 2.11. Note that π does not include a negative number. However, π is a fortress and hence not maximal by Theorem 3.13. The next theorem is likely well known, but we were unable to find a reference other than a mention in the OEIS entry A131209 [10].

Theorem 3.16. If $\pi \in S^{\pm}(n)$ is maximal, then

$$d(\pi) = \begin{cases} n, & n = 1, 3 \\ n + 1, & \text{otherwise.} \end{cases}$$

Proof. A brute-force check verifies the result for $n = 1$ and $n = 3$. Assume $n \neq 1, 3$ and $\pi \in S^{\pm}(n)$ is maximal. We know by Theorem 3.13 that π is not a fortress. Therefore, by Theorem 2.12, we have

$$d(\pi) = n + 1 - c(\pi) + h(\pi).$$

By Lemma 3.14, $c(\pi) = h(\pi)$. Therefore, $d(\pi) = n + 1$. Thus,

$$d(\pi) = \begin{cases} n, & n \neq 1, 3 \\ n + 1, & \text{otherwise} \end{cases}$$

as desired. $\square$

Indeed, the permutations we exhibited in the proof of Theorem 3.13 are maximal if $n \neq 1, 3$. However, it is important to point out that there exist maximal permutations consisting of more than one cycle. The next result follows from Theorem 3.3 and Theorem 3.16.

Corollary 3.17. If $\pi \in S^{\pm}(n)$ is maximal, then every $\gamma \in [\pi]$ is maximal.

In other words, given a maximal permutation, all permutations arising from cyclic shifts of the breakpoint diagram are also maximal.

Example 3.18. Consider π from Example 3.5. Using Theorem 2.12 and $B(\pi)$ in Figure 3.4, we find $d(\pi) = 11$. So, π is indeed maximal. Each permutation corresponding to the various cyclic shifts of $B(\pi)$ are also maximal by Corollary 3.17.

The next result characterizes the breakpoint diagrams for cyclic shifts of a maximal permutation and follows from the definition of cyclic shift and Corollary 3.17.

Corollary 3.19. For $n \neq 1, 3$, if $\pi \in S^{\pm}(n)$ is a maximal permutation, then either every cycle of $B(\pi)$ is a minimal hurdle or there exists a unique cycle that is a maximal hurdle and all other cycles are minimal hurdles. Moreover, if $B(\pi)$ has a maximal hurdle, then there exists $\gamma \in [\pi]$ where $B(\gamma)$ does not have a maximal hurdle.

Corollary 3.20. If $\pi \in S^{\pm}(n)$ is a maximal permutation with $n \neq 1, 3$, then for all $i \leq j$, $d(\delta_{i,j}(\pi)) = n$.

Proof. Let π be a maximal permutation. Since π is maximal, all elements of π must be positive by Theorem 3.15. Thus, when we apply a reversal, our resulting permutation has a negative element and hence cannot be maximal. Thus, $d(\delta_{i,j}(\pi)) = n$ for all $i \leq j$ by Theorem 3.16. $\square$

From the proof of Corollary 3.20, one might be tempted to conclude that for $n \neq 1, 3$, if $d(\pi) = n$, then π contains at least one negative entry. However, this is not true (see Example 4.3). Something wrong here.

The next result summarizes our results involving maximal permutations.

Remark 3.21. Let $\pi \in S^{\pm}(n)$ be a maximal permutation. Then

1. π is not a fortress;
2. π only contains positive entries;
3. $d(\pi) = n + 1$, if $n \neq 1, 3$ (and n otherwise);
4. All components of $B(\pi)$ are unoriented cycles (and hence hurdles); and
5. Every sequence of cyclic shifts of $B(\pi)$ represents a breakpoint diagram for a maximal permutation.

Chapter 4

The Reversal Poset

In this chapter, we discuss the reversal poset. Define the *reversal poset* as $R_n := (S^{\pm}(n), \leq)$, determined by the following covering relations: $\pi \lessdot \gamma$ if there exists $i < j$ such that $\delta_{ij}(\gamma) = \pi$ and $d(\pi) + 1 = d(\gamma)$. In general, $\pi \leq \gamma$ if and only if there exists a sequence of reversals that transforms γ into π such that reversal distance goes down at each step. Note that R_n is analagous to the absolute order of $S(n)$ (see [13, Chapter 3]).

Example 4.1. Figure 4.1 depicts the reversal poset R_2. We have labeled each edge in the Hasse diagram with the corresponding reversal.

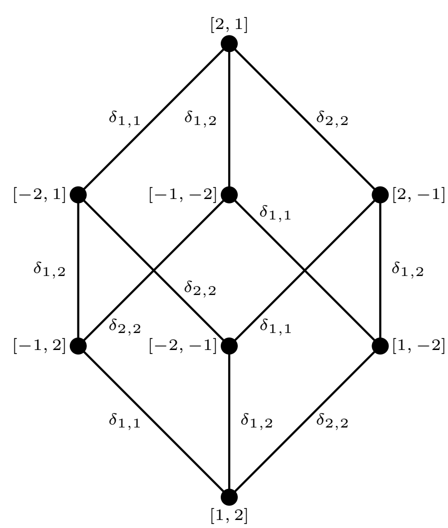


Figure 4.1: The Hasse diagram for the reversal poset R_2.

Given a product $\delta_{x_1}\delta_{x_2}\cdots\delta_{x_n}$ in terms of a minimal number of reversals that yields $\pi \in S^{\pm}(n)$,

$$I_n < \delta_{x_n} < \delta_{x_{n-1}}\delta_{x_n} < \cdots < \delta_{x_1}\cdots\delta_{x_n}$$

is a chain in R_n such that the reversal distance increases by 1 at each step. Therefore, $I_n \leq \pi$ for all $\pi \in S^{\pm}(n)$. In other words, I_n is the unique minimal element in R_n.

It follows immediately from the definition that R_n is a ranked poset, where rank corresponds precisely to reversal distance. While R_1 and R_2 are each lattices, for $n \geq 3$, R_n is not a lattice since there is not a unique maximal element.

By definition, each vertex in the Hasse diagram for R_n can have degree at most T_n since there are T_n distinct reversals we can apply to a permutation. From the way the covering relations are defined, one may wonder if a reversal acting on a permutation yields a permutation with reversal distance that is either one more or one less. However, there exists incomparable permutations π and γ such that $\delta_{i,j}(\pi) = \gamma$ while $d(\pi) = d(\gamma)$. In this case, we say that π and γ are *laterally related.* The presence of laterally-related permutations implies that a vertex of the Hasse diagram for R_n can have degree less than T_n. This is dissimilar from what happens in the absolute order on $S(n)$.

Example 4.2. Let $\pi = [2, -3, 1, -4] \in S^{\pm}(4)$. The breakpoint diagram for π is depicted in Figure 4.2(a). From the breakpoint diagram, we find $c(\pi) = 1$ and $h(\pi) = 0$. Since π is not a fortress (since there are no superhurdles), we can conclude by Theorem 2.12 that $d(\pi) = 4$. Applying the reversal $\delta_{3,3}$ to π results in $\gamma = [2, -3, -1, -4]$. The breakpoint diagram for this permutation is depicted in Figure 4.2(b). We find that $h(\gamma) = 0$ and $c(\gamma) = 1$. Since γ is not a fortress, $d(\gamma) = 4$ by Theorem 2.12. Since both of these permutations have the same reversal distance, π and γ are laterally related by $\delta_{3,3}$.

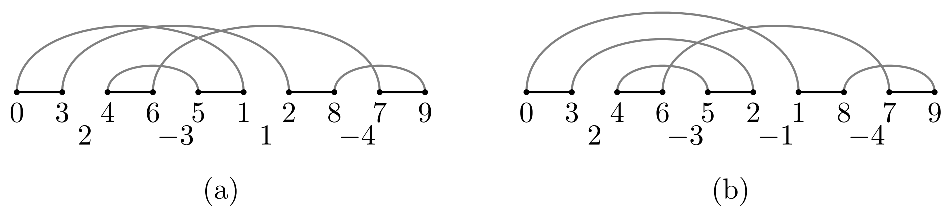


Figure 4.2: Example of two laterally related permutations.

The reversal poset is defined in such a way that laterally-related permutations are hidden. We do not have a characterization for when laterally-related pairs occur and it is unknown how often they occur.

If a signed permutation is maximal in the poset, we say the permutation is *terminal* to avoid confusion with permutations that have maximal reversal distance. In other words, π

is terminal if $d(\delta_{ij}(\pi)) \leq d(\pi)$ for all δ_{ij}. Note that every maximal permutation in $S^{\pm}(n)$ is terminal in R_n. However, there exists terminal permutations in R_n that are not maximal.

Example 4.3. Consider $\pi = [2, -3, 1, -4] \in S^{\pm}(4)$. We calculated $d(\pi) = 4$ in Example 4.2. We know from Theorem 3.16 that the maximal reversal distance for a permutation in $S^{\pm}(4)$ is 5. It turns out that $d(\delta_{i,j}(\pi)) \leq 4$ for all reversals $\delta_{i,j}$, which implies that π is terminal but not maximal.

It is unknown when terminal non-maximal permutations occur, but they appear to be somewhat common among permutations of reversal distance n for $n \geq 4$. This next result is a corollary of Theorem 3.16.

Corollary 4.4. Let E_n be the undirected Cayley graph of $S^{\pm}(n)$ generated by reversals. If R_n is interpreted as a graph, R_n is a subgraph of E_n built by removing all edges connecting laterally-related permutations. Moreover, the diameter of both graphs is $n+1$ for $n \neq 1, 3$ and n otherwise.

Define $R_{n,i} := \{\pi \in S^{\pm}(n) \mid d(\pi) = i\}$ and $r_{n,i} := |R_{n,i}|$. Note that $R_{n,i}$ is the collection of all permutations in $S^{\pm}(n)$ with the same reversal distance i, which we can think of as a row in the poset since the poset is ranked. Also, notice that by Theorem 3.16, $R_{n,i} \neq \emptyset$ if and only if $0 \leq i \leq n+1$ for $n \neq 1, 3$ and $0 \leq i \leq n$ for $n = 1, 3$. In addition, by Theorem 3.3, it follows that each $R_{n,i}$ is partitioned into various shift classes. We end this chapter by discussing $r_{n,0}, r_{n,1}$, and $r_{n,2}$.

The reversal poset has a fascinating structure. We know that $R_{n,0}$ only contains the identity permutation on n elements since this is the only permutation with reversal distance zero. Hence $r_{n,0} = 1$ for all $n \in \mathbb{N}$. Below we discuss how to count $R_{n,1}$ and $R_{n,2}$.

Theorem 4.5. We have $r_{n,1} = T_n = \binom{n+1}{2}$ for all $n \in \mathbb{N}$.

Proof. We know that for any $\pi \in R_{n,1}$, $\pi = \delta_{i,j}(I_n)$ for some $1 \leq i \leq j \leq n$. Since there are $T_n = \binom{n+1}{2}$ distinct reversals, it follows that $r_{n,1} = T_n = \binom{n+1}{2}$. $\square$

Theorem 4.6. We have $r_{n,2} = \frac{(n-1)n(n+1)^2}{6}$ for all $n \in \mathbb{N}$.

Proof. Let $\pi \in S^{\pm}(n)$ with reversal distance 2. Therefore, $\pi = \delta_{i,j}\delta_{k,l}$, where $\delta_{i,j} \neq \delta_{k,l}$. We will refer to i, j, k, l as the ends of our reversals. There are three cases we consider:

- Case 1: There are two distinct values amongst i, j, k, l;
- Case 2: There are three distinct values amongst i, j, k, l;
- Case 3: There are four distinct values amongst i, j, k, l.

	Subcases	Count
Case 1	$i = j, k = l$	$\binom{n}{2}$
	$i = j = k, k < l$	$\binom{n}{2}$
	$i = j = l, k < l$	$\binom{n}{2}$
Case 2	$i < k < l, j = l$	$\binom{n}{3}$
	$i < l < j, k = i$	$\binom{n}{3}$
	$i < j < l, k = j$	$\binom{n}{3}$
	$k < l < j, l = j$	$\binom{n}{3}$
	$i < j < k, k = l$	$2\binom{n}{3}$
	$i < k < j, k = l$	$\binom{n}{3}$
Case 3	$i < j < k < l$	$\binom{n}{4}$
	$i < k < j < l$	$\binom{n}{4}$
	$k < i < l < j$	$\binom{n}{4}$
	$i < k < l < j$	$\binom{n}{4}$

Table 4.1: Summary of cases for the proof of Theorem 4.6.

Each of these cases and their respective subcases are outlined in Table 4.1.

In Case 1, the first subcase implies both subintervals have size one. In this case, there are $\binom{n}{2}$ ways to choose the two distinct values amongst i, j, k, l. In the second and third subcases, there is one subinterval of size one and the second subinterval is larger, where the subinterval of size 1 is "nested" in the other. In the second subcase, note that if $i = j = k < l$, then $\delta_{i,j}\delta_{k,l} = \delta_{k,l}\delta_{l,l}$, and so we can restrict ourselves to this situation. Just as in the first case, the count is $\binom{n}{2}$. The third subcase is similar.

In Case 2, subcases 1 and 2 are similar to subcases 2 and 3 in Case 1, but with a smaller subinterval of a different size. Subcases 3 and 4 involve our two reversals overlapping in a single common position. The last subcase happens when one reversal is a subinterval of size 1. This subcase covers the situation where $k < i < j$ and $i = j$. In each of these subcases, there are $\binom{n}{3}$ ways to choose three distinct values for i, j, k, l. The fifth subcase counts the number of ways in which one of our reversals is a subinterval of size one and the other is a subinterval of size greater than 1, where the reversals do not intersect in positions. There are $\binom{n}{3}$ ways to pick three distinct endpoints and two ways we can choose which reversal is a subinterval of size greater than 1. There are three other situations in which this happens: $k < i < j$ and $k = l$, $k < l < j$ and $i = j$, and $i < k < l$ and $i = j$. These subcases are counted in subcase 5.

In Case 3, subcase 1 occurs when the reversals have no overlap in position. Subcases

2 and 3 occur when the two reversals overlap in position. In subcase 2, we note that $\delta_{i,j}\delta_{k,l} = \delta_{i+j-l,i+j-k}\delta_{i,j}$. Subcase 3 has a similar analogue. Subcase 4 happens when one reversal is "nested" in the other. In each subcase, there are $\binom{n}{4}$ ways to choose four distinct values for i, j, k, l.

Thus, there are $3\binom{n}{2} + 7\binom{n}{3} + 4\binom{n}{4}$ elements in $R_{n,2}$. One can verify that

$$3\binom{n}{2} + 7\binom{n}{3} + 4\binom{n}{4} = \frac{(n-1)n(n+1)^2}{6}.$$

Thus, $r_{n,2} = \frac{(n-1)n(n+1)^2}{6}$ for all $n \in \mathbb{N}$. □

The first few terms in the sequence for $r_{n,2}$ are $0, 3, 16, 50, 120, 245, 448, 756$. Note that we verified these values using the software package `baobabLUNA` [4]. It turns out that $r_{n,2}$ is the number of lattice rectangles (squares included) inside half of an Aztec diamond of order $n+1$, according to the OEIS entry A004320 [11]. This shape is obtained by stacking $n+1$ rows of consecutive unit lattice squares, with the centers of rows vertically aligned and consisting successively of $2(n+1), 2(n+1)-2, ..., 4, 2$ squares. In Chapter 6, we enumerate $r_{n,n+1}$ and discuss partial results for $r_{n,n}$ when $n \neq 1, 3$.

Chapter 5

Important Sets Of Compositions

Compositions are going to play an important role in proving results about the reversal poset for larger ranks. A *composition* of n is an ordered list of positive integers whose sum is n, denoted

$$\alpha = (\alpha_1, ..., \alpha_k).$$

We refer to each α_i as a *part* of the composition. Let $C(n)$ denote the set of all compositions on n.

Example 5.1. Consider $C(4)$. We see that

$$C(4) = \{(1, 1, 1, 1), (1, 2, 1), (1, 1, 2), (2, 1, 1), (3, 1), (1, 3), (2, 2), (4)\}.$$

One of these compositions has 4 parts, three have 3 parts, three have 2 parts, and one has 1 part.

As we shall see, each part of a composition will correspond to the number of black edges in a cycle in the breakpoint diagram for certain permutations. Here are four subsets of compositions in which we are interested.

1. $C_1(n) := \{(\alpha_1, \ldots, \alpha_k) \in C(n) \mid \text{each } \alpha_i \text{ is odd and greater than } 1\}$,

2. $C_2(n) := \{(\alpha_1, \ldots, \alpha_k) \mid \alpha_i > 1$ for all i, there exists j such that $\alpha_j \geq 4$ and for all $i \neq j$, α_i is odd$\}$,

3. $C_3(n) := \{(\alpha_1, \ldots, \alpha_k) \in C_1(n) \mid k \geq 3\}$.

For each i, define $c_i(n) := |C_i(n)|$. It is quite obvious that $C_3(n) \subseteq C_1(n)$, which means $c_3(n) \leq c_1(n)$.

Example 5.2. Let's look at compositions of 9. We see that $(3, 3, 3)$ is an element of $C_3(9)$, and hence an element of $C_1(9)$ since each part is odd and greater than 1. An element of $C_2(9)$ is $(3, 6)$ since all parts are odd except one part that is greater than or equal to 4.

Below, we enumerate the set $C_1(n)$ using a recursive formula.

Theorem 5.3. We have $c_1(1) = c_1(2) = 0, c_1(3) = 1$ and for $n \geq 4$

$$c_1(n) = c_1(n-2) + c_1(n-3).$$

Proof. Compositions of 1 and 2 do not have any compositions consisting of only odd parts greater than 1, so $c_1(1) = c_1(2) = 0$. The only composition of 3 that meets the criteria is (3), so $c_1(3) = 1$. Suppose $n \geq 4$. A composition $\alpha \in C_1(n)$ either ends in 3 or it does not. If α ends in 3, then delete the last part to obtain a composition of $n - 3$. If α does not end in 3, subtract 2 from the last part to obtain a composition of $n - 2$. The aforementioned process is reversible. Therefore, $c_1(n) = c_1(n-2) + c_1(n-3)$. □

The first few terms of the sequence determined by $c_1(n)$ are $0, 0, 1, 0, 1, 1, 1, 2, 2, 3$. It turns out that $c_1(n)$ is the Padovan sequence (OEIS entry A000931 [12]), which counts the number of compositions of n into parts congruent to 2 (mod 3). We do not have results on how to count $C_2(n)$ and $C_3(n)$. The first non zero term of $c_2(n)$ is 1 when $n = 4$, corresponding to the composition (4). The first nonzero term of $c_3(n)$ is 1 when $n = 9$ corresponding to the composition (3,3,3). Neither of these sequences appear in the OEIS.

Chapter 6

Enumeration of Maximal Permutations

The goal of this chapter is to enumerate $R_{n,n+1}$ and provide partial results for $R_{n,n}$. First, we introduce the Hultman numbers.

The *Hultman numbers* $H(n, k)$ are defined to be the number of unsigned permutations in $S(n)$ whose breakpoint diagram consists of k cycles [6]. We are specifically interested in $H(n, 1)$, which counts the number of unsigned permutations having a breakpoint diagram consisting of a single cycle. According to [6], we have

$$H(n,1) = \begin{cases} \frac{2n!}{n+2}, & n \text{ is even} \\ 0, & n \text{ is odd.} \end{cases}$$

Recall that the breakpoint diagram for permutations in $S^{\pm}(n)$ consists of $n + 1$ black edges. The values for $H(n, 1)$ and Remark 3.21 lead to the following lemma.

Lemma 6.1. If $\pi \in S^{\pm}(n)$ is a maximal permutation such that every hurdle of $B(\pi)$ is minimal, then each cycle consists of an odd number of black edges.

The next theorem is the main result of this thesis.

Theorem 6.2. For $n \neq 1, 3$, we have

$$r_{n,n+1} = \sum_{(\alpha_1,\ldots,\alpha_k)\in C_1(n+1)} \left(\prod_{i=1}^{k} \frac{2(\alpha_i - 1)!}{\alpha_i + 1} \right) \cdot \begin{cases} \alpha_1, & \text{if } k \neq 1 \\ 1, & \text{if } k = 1 \end{cases}.$$

Proof. Assume π is maximal. Then by Remark 3.21, π consists of all positive entries and each component of $B(\pi)$ is a single unoriented cycle that is a hurdle. There are three cases in which π is maximal.

1. $c(\pi) = h(\pi) = 1$;

2. $c(\pi) = h(\pi) > 1$ and all hurdles are minimal;

3. $c(\pi) = h(\pi) > 1$ and there exists a maximal hurdle and the rest are minimal.

Notice that Case 1 only occurs when n is even by Lemma 6.1, in which case we have $H(n, 1) = \frac{2n!}{n+1}$ many permutations. This case corresponds to $k = 1$ in our formula.

For Case 2, we number the minimal hurdles from left to right: $N_1, \ldots, N_k$, where α_i is the number of black edges in N_i. Then each permutation in Case 2 is associated with a composition $(\alpha_1, \alpha_2, \ldots, \alpha_k) \in C_1(n + 1)$ (since $k > 1$ and each $\alpha_i > 1$), which tracks the number of black edges in each cycle. Since the number of choices for each N_i is $H(\alpha_i - 1, 1)$, the number of permutations associated with a fixed composition $(\alpha_1, \alpha_2, \ldots, \alpha_k)$ is $\prod_{i=1}^{k} H(\alpha_i - 1, 1) = \prod_{i=1}^{k} \frac{2(\alpha_i-1)!}{\alpha_i+1}$.

In Case 3, each breakpoint diagram for a maximal permutation π with a maximal hurdle can be obtained from a unique breakpoint diagram for a maximal permutation γ with no maximal hurdles by doing i cyclic shifts, where $1 \leq i \leq \alpha_1 - 1$, where α_1 is the number of black edges in the leftmost cycle in the breakpoint diagram for γ by Corollary 3.19. We note that the resulting breakpoint diagram is indeed maximal by Corollary 3.17. So, each permutation in Case 3 results from a sequence of cyclic shifts of a permutation in Case 2. Since $k > 1$, here is how we construct the breakpoint diagram in this case. First, choose a composition $(\alpha_1, \ldots, \alpha_k) \in C_1(n + 1)$. For each α_i, choose a breakpoint diagram with α_i many black edges that consists of a single cycle. There are $H(\alpha_i - 1, 1)$ many choices for each i. Lay out the cycles side by side to obtain a breakpoint diagram for a permutation in Case 2. Now, to obtain a permutation in Case 3, do a sequence of i cyclic shifts where $1 \leq i \leq \alpha_1 - 1$.

That is, given a choice of a sequence of hurdles, we construct α_1 many breakpoint diagrams for a maximal permutation, the breakpoint diagram we started with (counted in Case 2) plus the $\alpha_1 - 1$ cyclic shifts. We obtain the desired formula from these three cases. □

The first few terms of $r_{n,n+1}$ are $0, 1, 0, 8, 3, 180, 64, 8067$. We verified these values using `baobabLUNA` [4]. This sequence does not appear in the OEIS.

We have a classification of the types of permutations of size n that have reversal distance n, which can help us estimate the size of $R_{n,n}$. We know by Corollary 3.11 that a permutation with reversal distance n cannot be a fortress, so this reduces the cases that we have to examine. It is clear there are two cases when examining the formula for reversal distance from Theorem 2.12.

- Case 1: Either all the cycles in the breakpoint diagram are not oriented, or

- Case 2: There is one oriented cycle in the breakpoint diagram.

We focus on Case 1 in our analysis. We do not attempt to classify permutations in Case 2, since it appears to be quite complex due to the inclusion of negative values in the permutation (by Theorem 3.12). For Case 1, there are three possibilities:

(a) All cycles are distinct components and all are unoriented except for one trivial cycle.

(b) All cycles are unoriented and exactly one pair of cycles overlap to form a single component. The remaining cycles are distinct components.

(c) Every cycle is a distinct unoriented component and we have one (unoriented) non-hurdle while every other cycle is a hurdle.

Let X_n, Y_n, and Z_n be the set of permutations in $S^{\pm}(n)$ following Cases 1(a), 1(b), and 1(c) respectively, and let $x_n := |X_n|, y_n := |Y_n|$, and $z_n := |Z_n|$. By Theorem 3.15, we know permutations counted by x_n, y_n, and z_n are unsigned permutations. Moreover, analysing the formula for reversal distance makes it clear that $X_n \cup Y_n \cup Z_n$ is the set of permutations in $R_{n,n}$ that are unsigned. In particular, $X_n \cup Y_n \cup Z_n$ is exactly the collection of terminal non-maximal permutations with reversal distance equal to n. Since the three sets are pairwise disjoint,

$$|X_n \cup Y_n \cup Z_n| = x_n + y_n + z_n.$$

Next we enumerate the permutations in Case 1(a) above. For convenience, we define $r_{0,1} := 0$.

Theorem 6.3. We have $x_n = (n+1) \cdot r_{n-1,n}$.

Proof. First, note that $r_{n-1,n} = 0$ when $n = 2$ and $n = 4$ by Theorem 3.16. The cases involving $n = 1, 2, 3, 4$ are easily verified by hand. Assume $n \geq 5$ and let $\pi \in X_n$. Then $B(\pi)$ consists of a single trivial cycle whose removal results in a breakpoint diagram for a maximal permutation in $S^{\pm}(n-1)$. Let $\hat{\pi} \in S^{\pm}(n-1)$ be the maximal permutation that results from removal of this single trivial cycle. Thus, $\hat{\pi} \in R_{n-1,n}$. There are $r_{n-1,n}$ many choices for $\hat{\pi}$ and $n+1$ many ways to reinsert the trivial cycle. This results in the desired formula. $\square$

The first few terms of x_n are $0, 0, 4, 0, 48, 21, 1440, 576$. We verified these values using `baobabLUNA` [4]. This sequence does not appear in the OEIS. We conjecture a formula for y_n in Chapter 7. Next we provide an upper bound for z_n.

Theorem 6.4. We have

$$z_n \leq \sum_{(\alpha_1,\ldots,\alpha_k)\in C_3(n+1)} \left(\left(\prod_{i=1}^{k-1} \frac{2(\alpha_i-1)!}{\alpha_i+1} \right) \left(\left(\binom{k+\alpha_k-1}{\alpha_k} - (\alpha_k+k-1) \right. \right. \right.$$
$$\left. \left. \left. + (\alpha_1-1)\left(\binom{k+\alpha_k-2}{\alpha_k} - (k-1) \right) \right) \right) \right).$$

Proof. Let $\pi \in Z_n$. Note that every cycle of $B(\pi)$ is a distinct unoriented component and there is exactly one non-hurdle component (while every other component is a hurdle). Assume the breakpoint diagram for π consists of k cycles, say $N_1, N_2, \cdots, N_k$, where α_i corresponds to the number of black edges in N_i and N_k is the non-hurdle. In order for the breakpoint diagram to have an unoriented non-hurdle, $k \geq 3$. Also, since there are no trivial cycles, each $\alpha_i > 1$. Therefore, the composition $(\alpha_1, \alpha_2, \ldots, \alpha_k) \in C_3(n+1)$. We consider two cases.

First, assume each N_i is a minimal hurdle for $1 \leq i \leq k-1$. Then place $N_1, \cdots, N_{k-1}$ in the breakpoint diagram from left to right. There are $\prod_{i=1}^{k-1} \frac{2(\alpha_i-1)!}{2\alpha_i}$ ways to do this. We now place N_k in the breakpoint diagram such that N_k is an unoriented non-hurdle. We must place the α_k black edges in k gaps in the breakpoint diagram (there are $k-2$ gaps in between our $k-1$ cycles and two gaps on either side of the breakpoint diagram). In total, we can do this in $\binom{\alpha_k+k-1}{\alpha_k}$ ways (which is the formula for counting the number of arrangements of α_k unlabeled balls in k labeled boxes). However, some of these configurations are not allowable. Note that there are $\alpha_k + k - 1$ ways to place the black edges of N_k so that N_k covers none of the other cycles or covers all other cycles and does not separate, which occurs when we place all α_k edges on either the far left or far right of the breakpoint diagram, or if all α_k black edges are placed in a gap between two cycles. Thus, there are

$$\binom{\alpha_k+k-1}{\alpha_k} - (\alpha_k+k-1)$$

allowable ways to place N_k such that N_k is a non-hurdle in this case.

Next, assume N_1 is the maximal hurdle in the breakpoint diagram. By Corollary 3.19, each breakpoint diagram for a permutation with a maximal hurdle can be obtained from a unique breakpoint diagram for a permutation γ with no maximal hurdles by doing i cyclic shifts, where $1 \leq i \leq \alpha_1 - 1$, where α_1 is the number of black edges in the leftmost cycle in the breakpoint diagram for γ. Thus, there are $\alpha_1 - 1$ distinct cyclic shifts of a breakpoint diagram resulting in N_1 being a maximal hurdle. So, there are

$$(\alpha_1-1)\prod_{i=1}^{k-1} \frac{2(\alpha_i-1)!}{2\alpha_i}$$

ways to rearrange the each of the first $k-1$ cycles corresponding to the various cyclic shifts. However, N_k must be covered by N_1 in order to be a non-hurdle. Thus, there are α_k black edges that need to be placed in the $k-1$ gaps that occur under N_1. There are $\binom{k+\alpha_k-2}{\alpha_k}$ ways to do this in general, but some configurations are not allowed. There are $k-1$ ways in which all α_k black edges fall between two cycles in the breakpoint diagram, meaning N_k would not cover any cycles. So, there are

$$\binom{k+\alpha_k-2}{\alpha_k}-(k-1)$$

allowable ways to insert the black edges of N_k. □

Using `baobabLUNA` [4], we have verified that the formula given in Theorem 6.4 agrees with the actual count of permutations in Z_n for $1 \leq n \leq 8$. The first few terms of z_n are $0, 0, 0, 0, 0, 0, 0, 9$. The first few terms are 0 since $C_3(n+1) = \emptyset$ for $n \leq 7$. This gives us some intuition that our formula does count the number of unsigned permutations with one unoriented non-hurdle. However, it is feasible that our formula for z_n is an upper bound since it is possible that our construction in the proof of Theorem 6.4 yields breakpoint diagrams that do not correspond to permutations, in which case we have overcounted. However, we conjecture that the formula in Theorem 6.4 is exact.

Chapter 7

Open Problems

We conclude with a list of open problems:

- Enumerate the signed permutations on 1 through n that have fixed reversal distance. In this thesis, we solved the problem for reversal distance 0, 1, 2, and $n+1$ and obtained partial results for reversal distance n. In particular, fully characterize and enumerate signed permutations with reversal distance n. To do this, we need to determine y_n and z_n and handle the cases involving the existence of a single oriented cycle.
- Characterize the terminal non-maximal permutations according to reversal length. Such a characterization will likely involve necessary and/or sufficient conditions on the structure of the permutation or the corresponding breakpoint diagram.
- Characterize the pairs of signed permutations that are laterally related.
- Determine whether $n-1$ is a sharp bound on reversal distance for a fortress for all $n \geq 17$.
- We conjecture that

$$y_n = \sum_{(\alpha_1,\ldots,\alpha_k)\in C_2(n+1)} \left(\sum_{j=1}^{k} \left(\left(\prod_{i\neq j} \frac{2(\alpha_i-1)!}{\alpha_i+1} \right) \left(H(\alpha_j-1,2) - \left\lceil \frac{\alpha_j-3}{2} \right\rceil \right) \right) \cdot \begin{cases} \alpha_1, & \text{if } k \neq 1 \\ 1, & \text{if } k = 1 \end{cases} \right),$$

$$z_n = \sum_{(\alpha_1,\ldots,\alpha_k)\in C_3(n+1)} \left(\left(\prod_{i=1}^{k-1} \frac{2(\alpha_i-1)!}{\alpha_i+1} \right) \left(\left(\binom{k+\alpha_k-1}{\alpha_k} - (\alpha_k+k-1) \right.\right.\right.$$
$$\left.\left.\left. + (\alpha_1-1)\left(\binom{k+\alpha_k-2}{\alpha_k} - (k-1) \right) \right) \right) \right).$$

- Determine the distribution of maximal permutations among all signed permutations of length n. We conjecture that

$$\lim_{n\to\infty} \frac{r_{n,n+1}}{2(n-1)!} = 1 \quad \text{if } n \text{ is odd,}$$
$$\lim_{n\to\infty} \frac{r_{n,n+1}}{2(n-3)!} = 1 \quad \text{if } n \text{ is even.}$$

  This would imply that if we choose a signed permutation uniformly at random, the probability of selecting a maximal permutation is about $n/2^n$ for n odd and $n(n-1)(n-2)/2^n$ for n even. That is, as n grows, it is exponentially unlikely to happen.

Bibliography

[1] A. Auyeung and A. Abraham. Estimating genome reversal distance by genetic algorithm. In *2003 Congress on Evolutionary Computation, CEC 2003 - Proceedings*, volume 2, pages 1157–1161, 2003.

[2] A. Bergeron. A very elementary presentation of the Hannenhalli–Pevzner theory. *Discrete Applied Mathematics*, 146:134–145, 2005.

[3] A. Bergeron, J. Mixtacki, and J. Stoye. Reversal distance without hurdles and fortresses. In S. Sahinalp, S. Muthukrishnan, and U. Dogrusoz, editors, *Combinatorial Pattern Matching*, pages 388–399. Springer Berlin Heidelberg, 2004.

[4] M.D.V. Braga. baobabLUNA: The solution space of sorting by reversals. *Bioinformatics*, 25(14):1833–1835, 2009.

[5] R.X. Chen and C. Reidys. A simple framework on sorting permutations. `arXiv:1502.07971`, 2015.

[6] J.P. Doignon and A. Labarre. On Hultman numbers. *Journal of Integer Sequences*, 10(6):1–13, 2007.

[7] S. Hannenhalli and P.A. Pevzner. Transforming Cabbage into Turnip: Polynomial Algorithm for Sorting Signed Permutations by Reversals. *Journal of the ACM*, 46(1):1–22, 1999.

[8] B. Hayes. Sorting Out the Genome. *American Scientist*, 95, 2007.

[9] J. Kececioglu and D. Sankoff. Exact and approximation algorithms for sorting by reversals, with application to genome rearrangement. *Algorithmica*, 12(1-2):180–210, 1995.

[10] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences. *Published electronically at* `https://oeis.org/A131209`*, Sequence A131209*, 2021.

[11] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences. *Published electronically at* `https://oeis.org/A004320`, *Sequence A004320*, 2021.

[12] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences. *Published electronically at* `https://oeis.org/A000931`, *Sequence A000931*, 2021.

[13] T.K. Petersen. *Eulerian Numbers.* Birkhäuser Basel, 2015.

[14] A. Sutherland and A. Sutherland. The mathematics of Spinpossible. `arXiv:1110.6645`, 2011.