Combinatorial and Probabilistic Approaches to Planar Tanglegrams, Colorful Permutations, and Cosine Functions

Kevin Liu

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2024

Reading Committee:
Sara Billey, Chair
Ricky Liu
Farbod Shokrieh

Program Authorized to Offer Degree:
Mathematics

University of Washington

Abstract

Combinatorial and Probabilistic Approaches to
Planar Tanglegrams, Colorful Permutations, and Cosine Functions

Kevin Liu

Chair of the Supervisory Committee:
Sara Billey
Department of Mathematics

Through combinatorial and probabilistic approaches, we study the structure of three objects: tanglegrams, colored permutations, and cosine functions. Our work on tanglegrams includes a characterization of planar tanglegram layouts and a method for sampling planar tanglegrams. The former generalizes a result of Lozano, Pinter, Rokhlenko, Valiente, and Ziv-Ukelson for finding a planar layout of a planar tanglegram, and the latter is a planar analog of an algorithm of Billey, Konvalinka, and Matsen for generating tanglegrams uniformly at random. Our work on colored permutations involves analyzing the moments of statistics on conjugacy classes without "short" cycles, with particular emphasis on the descent, major index, and flag-major index statistics. These generalize results of Fulman involving the descent and major index statistics on the symmetric group, as well as certain cases of recent work by Hamaker and Rhoades applying representation theory to the study of moments on conjugacy classes. Our work on cosine functions involves studying minimal cases for the correlation of their signs when the input is scaled by various integers. This is motivated by the study of Schrödinger operators due to Gonçalves, Oliveira e Silva, and Steinerberger, who characterized the minimal cases for $n = 2$ cosine functions. We provide the corresponding characterization for $n = 3$ cosine functions.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I am grateful to numerous individuals, without whom this dissertation would not have been possible. For some, the path to obtaining a Ph.D. or becoming a mathematician is direct, but my own journey was more akin to a biased random walk with several individuals playing instrumental roles.

I am grateful to my Ph.D. advisor at the University of Washington (UW), Sara Billey. Her guidance, discussions, and feedback have pushed my mathematical growth and writing in immeasurable ways. I am particularly grateful for her feedback on drafts of this dissertation and the publications that led to it. The conventional wisdom is that your choice of advisor is the most important decision in graduate school, and I could not agree with this more.

I am also grateful to several other professors from my time at UW. In particular, I would like to thank Ricky Liu, Shayan Oveis Gharan, and Farbod Shokrieh for serving on my thesis committee. I would also like to thank Stefan Steinerberger for being the faculty supervisor on two Washington Experimental Mathematics Lab (WXML) projects, which gave me opportunities to mentor undergraduate research projects. Finally, I would like to thank Marina Meila for a research assistantship that supported my work for a quarter.

I would also like to thank my fellow graduate students at UW, who have contributed significantly to making my time at UW a positive experience despite the difficult times that arose from the COVID-19 pandemic. In particular, I would like to thank Paige Helms and Albert Artiles for being constant sources of laughter and feedback, and I would like to thank my academic sibling and collaborator Herman Chau for many fruitful mathematical discussions.

I would also like to acknowledge several professors from my undergraduate years pursuing mathematics education at Vanderbilt University. My undergraduate advisors, Peg Cagle and Emily Shahan, taught me about mathematics as a fun endeavor and showed me the potential for what mathematics education can be. Their influence continues to drive my current practices as a teacher and mentor. Additionally, I am grateful to my discrete mathematics professor, Paul Edelman. My initial struggles with mathematical proofs as an undergraduate convinced me that I should not pursue mathematics, and I doubt that I would have changed my mind without Paul's positive encouragement and influence.

I am also grateful to my former colleagues at the University School of Nashville (USN), where I worked as a high school teacher for three years. I am particularly grateful to Cynthia Crenshaw for supporting my hiring, mentoring me during my early years as a teacher, and encouraging my departure to pursue a Ph.D. despite the challenges this created for USN. I would also like to thank Justin Fitzpatrick for entertaining my independent study of abstract algebra. When I finished my undergraduate degree, I felt certain that I no longer wished to be a student, but this experience taught me about my desire to become a lifelong learner.

My transition from teaching to mathematics also would not have been possible without Iowa State University's post-baccalaureate program in mathematics. When I initially explored the possibilities for returning to school, it appeared that entering the mathematical community would be out of reach for me due to a shortage of financial resources and mentorship. The program and faculty at Iowa State addressed these needs and gave me opportunities that I would not have otherwise had, and I hope to one day pay this forward to others.

I would also like to acknowledge my many mathematical collaborators, and parts of our work are included in this thesis. My collaboration with Alex Black, Alex McDonough, Garrett Nelson, Michael Wigal, Mei Yin, and Youngho Yoo resulted in a publication [BLM+23], portions of which are included in Section 3.2 and Section 3.3. My collaboration with Jesse Campion Loth, Michael Levet, Sheila Sundaram, and Mei Yin resulted in a preprint

[CLL+23]. These results and ongoing research with Michael Levet, Sheila Sundaram, and Mei Yin are included in Chapter 4. I would like to thank the Graduate Research Workshop in Combinatorics and its organizers, where the aforementioned collaborations began. Finally, my collaboration with UW undergraduates Shilin Dou, Ansel Goh, Madeline Legate, and Gavin Pettigrew resulted in a publication [DGL+24] that is included in Chapter 5. I would also like to thank the WXML for organizing and sponsoring our project.

Finally, I am also grateful for my mathematics education cohort at Vanderbilt, my family, and all of the friends that I made in the past several years. Each of these individuals has contributed to my growth from a naive college freshman to the mathematician and educator that I am today.

DEDICATION

To my parents, Jing Fang Liu and Feng Zhen Liu.

Chapter 1

INTRODUCTION

The intersection of combinatorics with other fields provides profound insights into both combinatorics and other areas in mathematics, and exploring these connections continues to fuel modern research in mathematics. We will focus on combinatorial and probabilistic methods that further our understanding of three objects: tanglegrams, colored permutations, and cosine functions. In the process, connections with other fields arise.

1.1 Tanglegrams

A *tanglegram* $\mathcal{T} = (L, R, \sigma)$ is a special type of graph constructed from two rooted binary trees L, R and a perfect matching between their leaves σ. They initially arose in mathematical biology [Pag93], and they also have applications in computer science [BBB+12]. Tanglegrams are drawn in the plane using *layouts* such as the ones shown in Figure 1.1. Similar to drawings of a graph, a tanglegram can have many layouts. In biology, layouts with the fewest number of crossings possible are of interest in applications such as determining how two species may have co-evolved [MBKK15] or estimating the number of horizontal gene transfers between species [VASJG09]. Applications in computer science include clustering, decomposition of programs into layers, or analyzing the difference in hierarchy between similar programs [BBB+12].

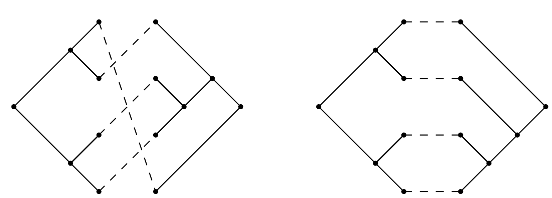


Figure 1.1: Two layouts for the same tanglegram.

Billey, Konvalinka, and Matsen [BKM17] formalized tanglegrams as mathematical objects and gave the first enumeration formulas in the literature. Note that their mathematical formalization preceded [MBKK15], despite the difference in publication date. Since then, many mathematical properties of tanglegrams have been studied. These include enumerating various families of tanglegrams [Ges21, RRW18], sampling tanglegrams uniformly at random [BKM17, Fus16], and studying properties of random tanglegrams [KW16]. One particularly interesting subfamily consists of the *planar* tanglegrams, which have a layout with no crossings. The tanglegram in Figure 1.1 is an example of one, as demonstrated by the second layout. An efficient algorithm for finding a planar layout was established in [LPR+07], and a characterization in terms of forbidden subtanglegrams was established in [CSW19].

1.1.1 Main Results

Our work builds on the existing literature for planar tanglegrams. The first result is a characterization of the planar layouts of a planar tanglegram. Our characterization is based on *leaf-matched pairs* of a planar tanglegram $\mathcal{T} = (L, R, \sigma)$, which are pairs of internal vertices $u \in L, v \in R$ where the subtrees rooted at u and v have their leaves matched by σ. On these pairs, we define a *paired flip* operation that reflects the subtrees rooted at u and v in a layout. Using this operation, we establish the following result.

Theorem 1.1.1 (Liu '23)**.** *Let $\mathcal{T}$ be a planar tanglegram, and let $\mathscr{P}(\mathcal{T})$ be its set of planar layouts. Each planar layout in $\mathscr{P}(\mathcal{T})$ can be obtained by starting with any planar layout in $\mathscr{P}(\mathcal{T})$ and performing some sequence of paired flips at leaf-matched pairs of vertices.*

One way to visualize this result on a planar tanglegram $\mathcal{T}$ is using a graph $G(\mathcal{T})$. The vertices of $G(\mathcal{T})$ are the planar layouts of $\mathcal{T}$, and edges correspond to pairs of planar layouts that can be obtained from one another using paired flips at leaf-matched pairs. An example is shown in Figure 1.2. Theorem 1.1.1 is equivalent to the connectedness of this flip graph.

We also consider enumerative results involving leaf-matched pairs, which directly generalize certain enumerative results in [RRW18]. For any tanglegram $\mathcal{T} = (L, R, \sigma)$, its *size*

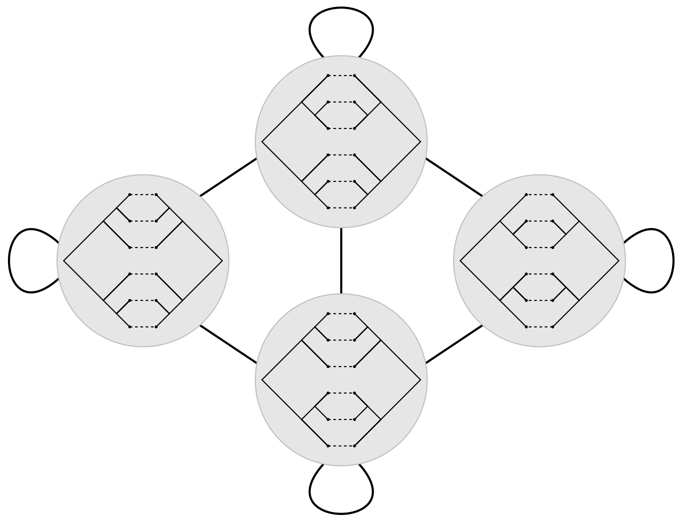


Figure 1.2: An example of a graph $G(\mathcal{T})$ on the planar layouts of a planar tanglegram.

is the common number of leaves in L and R. Additionally, a tanglegram is *irreducible* if its only leaf-matched pairs are the roots of L and R, and for any planar tanglegram, one can construct its *irreducible component* by contracting the subtrees at non-root leaf-matched pairs into a pair of matched leaves. Our generating function results in Theorem 3.2.1 and Theorem 3.2.2 establish relations that can be used to count the number of tanglegrams based on size and number of leaf-matched pairs, or size and irreducible component size. Some initial values are shown in Table 1.1. Additional terms can be found at [OEI24, A349409] and [OEI24, A371659].

n,k	1	2	3	4	5
2	1				
3	1	1			
4	5	4	2		
5	34	28	11	3	
6	273	239	102	29	6

n,k	2	3	4	5	6
2	1				
3	1	1			
4	3	3	5		
5	13	9	20	34	
6	90	46	70	170	273

Table 1.1: The number of planar tanglegrams of size n with k leaf-matched pairs (left), and the number of planar tanglegrams with size n and irreducible component of size k (right).

We next consider the problem of efficiently sampling planar tanglegrams uniformly at random. Note that one can use rejection sampling with the algorithms in [BKM17] and [Fus16] for sampling general tanglegrams to sample planar tanglegrams uniformly at random. However, if we let T_n and t_n respectively denote the number of tanglegrams and planar tanglegrams of size n, then [BKM17, Corollary 8] and [RRW18, Theorem 2] respectively imply

$$T_n \sim \frac{2^{2n-3/2} n^{n-5/2}}{\sqrt{\pi} e^{n-1/8}} \quad \text{and} \quad t_n \sim \frac{(0.00788\ldots)(15.77553\ldots)^n}{n^3}.$$

Consequently, the proportion of tanglegrams that are planar

$$\frac{t_n}{T_n} \sim \frac{\sqrt{\pi}(0.00788\ldots)}{\sqrt{8} e^{1/8} n^{1/2}} \left(\frac{e(15.77553\ldots)}{4n} \right)^n$$

approaches 0 rapidly as n increases.

Our approach for planar tanglegrams uses Markov chains to approximate sampling uniformly at random. We first use our enumerative results to reduce the problem of generating planar tanglegrams to the corresponding problem for irreducible planar tanglegrams. In their work enumerating planar tanglegrams, Ralaivaosaona, Ravelomanana, and Wagner [RRW18] established a bijection between irreducible planar tanglegrams of size $n+1$ and pairs of triangulations of a convex n-gon that do not share a diagonal. We call these *pairs of disjoint triangulations* of an n-gon, and starting with $n = 3$, there are

$$1, 5, 34, 273, 2436, 23391, 237090, 2505228, \ldots$$

such pairs. See [OEI24, A257887] for more terms. We construct a flip graph $\mathcal{D}_n$ whose vertices are pairs of disjoint triangulations and whose edges are given by a flip operation that generalizes the typical flip on triangulations. The flip graph $\mathcal{D}_5$ is shown in Figure 1.3. Our flip operation implies many surprising properties for $\mathcal{D}_n$.

Theorem 1.1.2 (Black, Liu, McDonough, Nelson, Wigal, Yin and Yoo '23)**.** *For any positive integer $n \geq 5$, the graph $\mathcal{D}_n$ is simple, connected, and $2(n-3)$-regular. Furthermore,* *a random walk on $\mathcal{D}_n$ that starts at an arbitrary vertex and chooses neighboring vertices uniformly at random will converge to the uniform distribution on the vertices in $\mathcal{D}_n$.*

One natural follow-up question is the mixing time of the random walk on $\mathcal{D}_n$, which is the number of time steps needed to be within any $\epsilon > 0$ of the uniform distribution with respect to the total variation metric on probability distributions. Explicit computations for $n \leq 9$ show that the mixing times for these values of n are relatively small, despite the rapid growth in the number of vertices in $\mathcal{D}_n$. A bound on mixing time for general n remains an open problem for future work.

1.1.2 Related work

Our work on pairs of disjoint triangulations connects to various problems involving triangulations of a convex n-gon. These are one of many objects enumerated by the Catalan numbers; see [Sta15] and [OEI24, A000108] for more details on Catalan objects. Prior work suggests that pairs of disjoint triangulations are significantly more complicated than individual triangulations. In particular, there is no known simple formula in n for the number of pairs of disjoint triangulations of an n-gon, even after fixing one of the triangulations [AR16, RRW18].

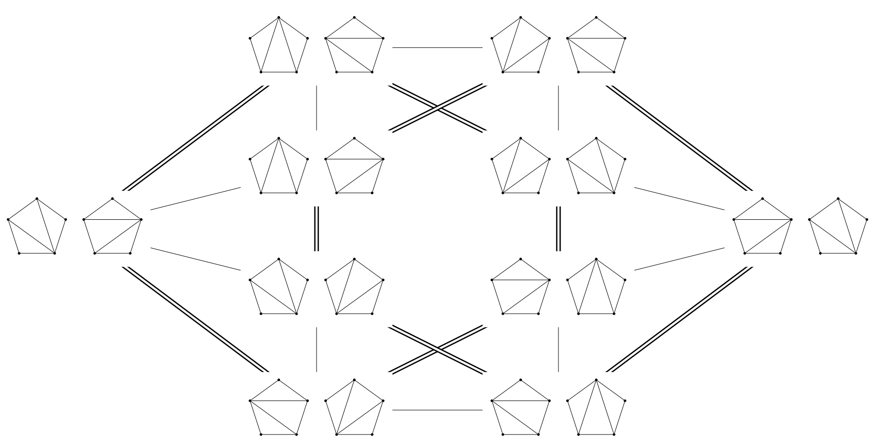


Figure 1.3: The flip graph $\mathcal{D}_5$. Single lines indicate when only one triangulation is changed, and double lines indicate when both are changed.

Sampling Catalan objects uniformly or approximately uniformly at random has been an active area of research over the last several decades. Approaches include direct methods using properties of Catalan numbers [AS92, BBJ17, Rém85], Boltzmann sampling [DFLS03], and random walks on flip graphs of Catalan objects [EF23, MT97, MRS01], and our approach for pairs of disjoint triangulations is based on the latter. The mixing time of a random walk on the classical flip graph on triangulations of an n-gon is bounded by $O(n^3 \log^3 n)$ [EF23], and it may be possible to adapt these techniques to bound the mixing time of our $\mathcal{D}_n$ graph. Note that the proof of Theorem 1.1.2 also provides novel insight into the classical flip graph on triangulations of an n-gon. Namely, we show in Theorem 3.3.6 that the induced subgraph on the triangulations disjoint from any fixed triangulation is always connected.

Our general construction of $\mathcal{D}_n$ is also not the first flip graph on pairs of Catalan objects satisfying some restriction. Previously, Heitsch and Tetali [HT11] studied *meanders*, which are pairs of noncrossing matchings of $2n$ points that form a cycle. An example is shown in Figure 1.4. Heitsch and Tetali also constructed a flip graph and random walk that converges to the uniform distribution, and the mixing time of their random walk also remains open.

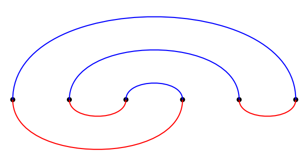


Figure 1.4: An example of a meander, with each noncrossing matching color-coded.

1.2 Colored permutation statistics

The symmetric group consists of bijections on $[n] = \{1, 2, \ldots, n\}$ with group operation given by function composition. A permutation $\omega \in \mathfrak{S}_n$ can be expressed in multiple ways. We will primarily use the two-line, one-line, and cycle notations. The two-line notation is a $2 \times n$ array where $\omega(1), \ldots, \omega(n)$ appears under $1, 2, \ldots, n$ surrounded by a pair of square brackets. The one-line notation is formed by removing the first line. The cycle notation decomposes $[n]$ into disjoint cycles denoted by pairs of parentheses. Within each cycle, $\omega(i)$ appears after $i \in [n]$, where the first element in each cycle is considered to be after the last one. For example, the permutation $\omega \in \mathfrak{S}_n$ defined by

$$\omega(1) = 4,\ \omega(2) = 5,\ \omega(3) = 1,\ \omega(4) = 3,\ \omega(5) = 2$$

has two-line, one-line, and cycle notations respectively given by

$$\omega = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 3 & 2 \end{bmatrix} = [45132] = (143)(25). \tag{1.2.1}$$

One can also view elements of $\mathfrak{S}_n$ as $n \times n$ permutation matrices, which have a single 1 in each row or column and have 0 in all remaining entries. The group operation is then given by the usual matrix multiplication.

The cycle notation is related to the conjugacy classes of $\mathfrak{S}_n$. The *cycle type* of a permutation is the multiset of cycle lengths that appear in the cycle notation. This can be expressed using a partition $\lambda = (1^{m_1(\lambda)}, 2^{m_2(\lambda)}, \ldots, n^{m_n(\lambda)})$ of n, where $m_i(\lambda)$ indicates the multiplicity of i. For brevity, entries of the form i^0 are omitted, and i^1 is abbreviated as i. For example, the permutation in (1.2.1) has cycle type $(1^0, 2^1, 3^1, 4^0, 5^0) = (2, 3)$. Two elements in $\mathfrak{S}_n$ are in the same conjugacy class if and only if they have the same cycle type, and the conjugacy class corresponding to cycle type λ will be denoted C_λ.

A permutation statistic X is any function from the symmetric group $\mathfrak{S}_n$ to $\mathbb{R}$. By equipping $\mathfrak{S}_n$ with the uniform distribution, X can be interpreted as a random variable with a discrete probability distribution given by

$$\Pr_{\mathfrak{S}_n}[X = i] = \frac{|X^{-1}(i)|}{|\mathfrak{S}_n|}. \tag{1.2.2}$$

The random variable X then has k-th moments for each integer $k \geq 1$ given by

$$\mathbb{E}_{\mathfrak{S}_n}[X^k] = \sum_{i \in \mathbb{R}} i^k \cdot \Pr_{\mathfrak{S}_n}[X = i] = \frac{1}{|\mathfrak{S}_n|} \sum_{\omega \in \mathfrak{S}_n} X(\omega)^k. \tag{1.2.3}$$

The distributions of many statistics are well-understood, including joint distributions and asymptotic behavior, e.g., see [BZ11, BS21, DH23, Fel45, MR08]. We will focus our attention on three statistics in particular: descents, inversions, and major index.

Definition 1.2.1. The *descent set* of $\omega \in \mathfrak{S}_n$ is

$$\mathrm{Des}(\omega) = \{i \in [n-1] \mid \omega(i) > \omega(i+1)\}.$$

The *descent statistic* and *major index* of $\omega \in \mathfrak{S}_n$ are respectively defined as

$$\mathrm{des}(\omega) = |\,\mathrm{Des}(\omega)| \quad \text{and} \quad \mathrm{maj}(\omega) = \sum_{i \in \mathrm{Des}(\omega)} i.$$

Definition 1.2.2. The *inversion set* of $\omega \in \mathfrak{S}_n$ is

$$\mathrm{Inv}(\omega) = \{(i,j) \in [n] \times [n] \mid i < j \text{ and } \omega(i) > \omega(j)\}.$$

The *inversion statistic* is defined as $\mathrm{inv}(\omega) = |\,\mathrm{Inv}(\omega)|$.

Example 1.2.3. Consider the permutation $\omega = [45132] \in \mathfrak{S}_5$. From the definitions above,

$$\mathrm{Des}(\omega) = \{2,4\} \quad \text{and} \quad \mathrm{Inv}(\omega) = \{(1,3),(1,4),(1,5),(2,3),(2,4),(2,5),(4,5)\}.$$

Hence, one calculates $\mathrm{des}(\omega) = 2$, $\mathrm{maj}(\omega) = 6$, and $\mathrm{inv}(\omega) = 7$.

The descent, inversion, and major index statistics arise in many contexts. For example, inversions appear in the study of sorting objects [Knu98] and testing randomness [Ros42]. Descents appear in the study of card shuffling [BD92] and carrying when adding numbers

[DF09, Hol97]. Descents and inversions also play a fundamental role in the weak Bruhat order on any Coxeter group, which contains $\mathfrak{S}_n$ as a special case [BB05]. Versions of the major index have been used in the study of the representation theory of $\mathfrak{S}_n$ [BKS20, Sta79].

The generating functions for the descent, inversion, and major index statistics are well-understood. For the descent statistic, the Carlitz identity for the Eulerian polynomials states

$$\sum_{\omega\in\mathfrak{S}_n} q^{\mathrm{des}(\omega)} = (1-q)^{n+1}\sum_{k=0}^{\infty}(k+1)^n q^k.$$

See [Pet15, Corollary 1.1] for details. For the inversion statistic, first define the *q-analog* of n to be $[n]_q = 1+q+q^2+\ldots+q^{n-1}$ and the *q-factorial* of n to be $[n]_q! = [n]_q[n-1]_q\ldots[2]_q[1]_q$. Then

$$\sum_{\omega\in\mathfrak{S}_n} q^{\mathrm{inv}(\omega)} = [n]_q!.$$

See [Pet15, Theorem 5.1] for details. Major Percy MacMahon [Mac13] originally introduced the major index statistic and established the following surprising result.

Theorem 1.2.4. *[Mac13, p. 285] The following holds:*

$$\sum_{\omega\in\mathfrak{S}_n} q^{\mathrm{maj}(\omega)} = [n]_q!.$$

Consequently, inv *and* maj *have the same distribution on* $\mathfrak{S}_n$*.*

When studying permutation statistics, one area of interest is the general or asymptotic behavior of their distributions. For the statistics above, we will see that when n is large, their distributions are approximated well by normal distributions. See Figure 1.5 for an illustration.

Definition 1.2.5. Let $\{X_n\}_{n\geq 1}$ and Y be real-valued random variables with cumulative distribution functions $F_n(x) = \Pr[X_n \leq x]$ and $F(x) = \Pr[Y \leq x]$. X_n *converges in distribution* to Y if $\lim_{n\to\infty} F_n(x) = F(x)$ for all $x \in \mathbb{R}$ where $F(x)$ is continuous.

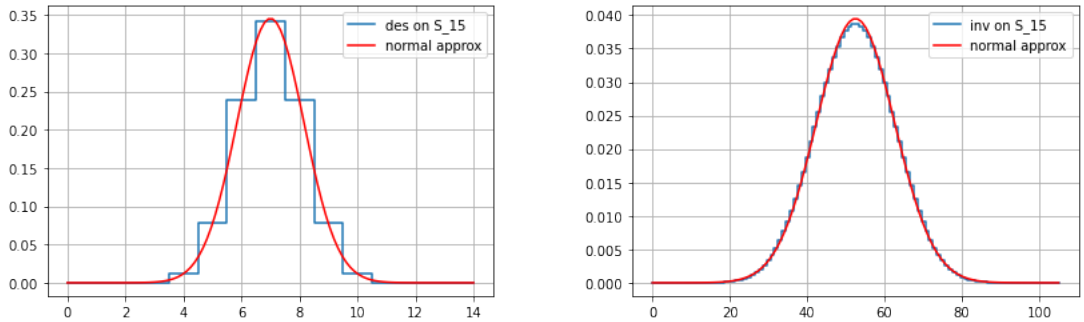


Figure 1.5: The distribution of the descent and inversion statistics on $\mathfrak{S}_{15}$, where the corresponding normal curves given in Theorems 1.2.6 and 1.2.7 are shown in red.

Theorem 1.2.6. *[DB62, pp. 150-154] The descent statistic on $\mathfrak{S}_n$ has mean $\mu_n = (n-1)/2$ and variance $\sigma_n^2 = (n+1)/12$. Furthermore, as $n \to \infty$, the random variable $\frac{\text{des} - \mu_n}{\sigma_n}$ converges in distribution to the standard normal distribution.*

Theorem 1.2.7. *[Fel45, pp. 814-815] The inversion statistic on $\mathfrak{S}_n$ has mean $\mu_n = n(n-1)/4$ and variance $\sigma_n^2 = n(2n^2+3n-5)/72$. Furthermore, as $n \to \infty$, the random variable $\frac{\text{inv} - \mu_n}{\sigma_n}$ converges in distribution to the standard normal distribution.*

One can also consider statistics and random variables on conjugacy classes of the symmetric group. Restricting a random variable $X : \mathfrak{S}_n \to \mathbb{R}$ to C_λ results in the conditional distribution

$$\Pr_{\mathfrak{S}_n}[X = i \mid C_\lambda] = \frac{|X^{-1}(i) \cap C_\lambda|}{|C_\lambda|}. \tag{1.2.4}$$

For brevity, we will use $\Pr_\lambda[X = i]$ to denote this. The distributions of various statistics refined to conjugacy classes are known, e.g., see [Bre93, CJZ20, DMP95, GR93]. Our work is inspired by that of Fulman, who considered moments of the descent and major index statistics on conjugacy classes without "short" cycles.

Theorem 1.2.8. *[Ful98, Corollary 5] Let C_λ be a conjugacy class of $\mathfrak{S}_n$. If C_λ contains no cycles of length $1, 2, \ldots, 2k$, then the k-th moments of* des *and* maj *on C_λ match their respective k-th moments on $\mathfrak{S}_n$.*

Using this result, Theorems 1.2.6 and 1.2.7, and a tool called the Method of Moments, Fulman obtained the following corollary. This shows that when n is large and C_λ has no "short" cycles, the distributions of des and maj on C_λ can be approximated well by their corresponding distributions on $\mathfrak{S}_n$. See Figure 1.6 for an illustration.

Corollary 1.2.9. *[Ful98, Corollary 6] For every $n \geq 1$, let C_{λ_n} be a conjugacy class of $\mathfrak{S}_n$ such that for all i, we have $\lim_{n\to\infty} m_i(\lambda_n) = 0$. Let X_n be the descent or major index statistic on C_{λ_n} with mean μ_n and variance σ_n^2. Then as $n \to \infty$, the random variable $(X_n - \mu_n)/\sigma_n$ converges in distribution to the standard normal distribution.*

Since Fulman's work, various analogs for other statistics have been established. Hamaker and Rhoades [HR22, Theorem 3.16] used representation-theoretic methods to show that for an arbitrary statistic, any fixed moment depends only on n and the number of "short" cycles in λ. In particular, this fixed moment will coincide on all conjugacy classes C_λ without "short" cycles.

Our work focuses on the more general setting of colored permutation statistics. The colored permutation group $\mathfrak{S}_{n,r}$ is the wreath product $\mathbb{Z}_r \wr \mathfrak{S}_n$, and these groups play an essential role in the classification of irreducible complex reflection groups [ST54]. The case of $r = 1$ corresponds to the usual symmetric group $\mathfrak{S}_n$. Similar to how one can view the elements in $\mathfrak{S}_n$ as $n \times n$ permutation matrices, one can view $\mathfrak{S}_{n,r}$ as a more general form of

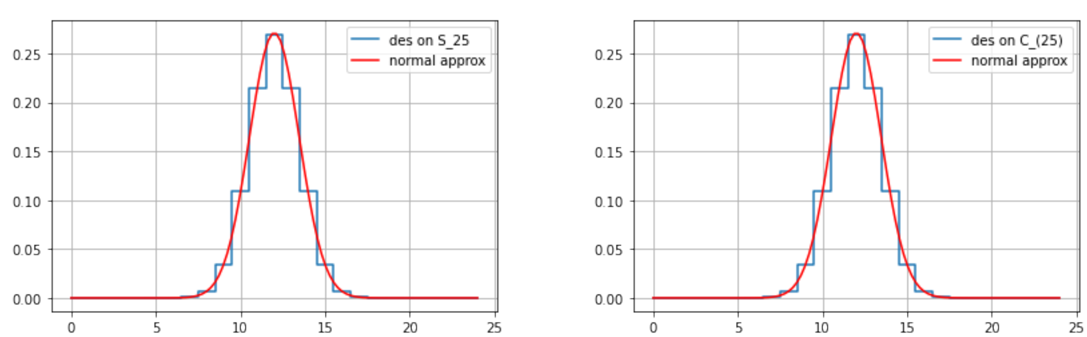


Figure 1.6: The distributions of the descent statistic on $\mathfrak{S}_{25}$ and $C_{(25)}$, the elements in $\mathfrak{S}_{25}$ whose cycle type is a single cycle of length 25.

these matrices where any r-th root of unity can appear in the nonzero entry in each row and column. The group operation is then given by the usual matrix multiplication.

In contrast to the vast literature on permutation statistics, there has been considerably less work on statistics for general colored permutation groups $\mathfrak{S}_{n,r}$. The distributions of specific colored permutation statistics over the entire group $\mathfrak{S}_{n,r}$ have been studied, e.g., see [AR01a, CM12, Fir05, Mou18, Mou21, Ste94]. However, we are not aware of any prior work on individual conjugacy classes except when $r = 2$. In this case, the colored permutation groups $\mathfrak{S}_{n,2}$ are isomorphic to the hyperoctahedral groups B_n, which are the type B Coxeter groups. Statistics on the entire group B_n and its conjugacy classes have been studied extensively, e.g., see [ABR01, AGR05, AR01a, AR01b, Rei93a, Rei93b, Rei95].

1.2.1 Main Results

Since [HR22, Theorem 3.16] implies that any fixed moment of a permutation statistic coincides on all conjugacy classes of $\mathfrak{S}_n$ with no "short" cycles, one might expect similar properties to hold for colored permutation statistics. We show that this is indeed the case. Similar to $\mathfrak{S}_n$, conjugacy classes of $\mathfrak{S}_{n,r}$ are determined by cycle type. Similar to the usage of C_λ for conjugacy classes of $\mathfrak{S}_n$, we will use $C_{\boldsymbol{\lambda}}$ to denote a conjugacy class of $\mathfrak{S}_{n,r}$, where $\boldsymbol{\lambda}$ records cycle type. Note that the definition of cycle type in $\mathfrak{S}_{n,r}$ is somewhat technical, so we will not give a precise definition here. A careful treatment will be given in Section 2.2.

For a colored permutation statistic $X : \mathfrak{S}_{n,r} \to \mathbb{R}$ and a conjugacy class $C_{\boldsymbol{\lambda}}$, we consider the probability distribution

$$\Pr_{\boldsymbol{\lambda}}[X = i] = \frac{|X^{-1}(i) \cap C_{\boldsymbol{\lambda}}|}{|C_{\boldsymbol{\lambda}}|}, \tag{1.2.5}$$

which has corresponding moments $\mathbb{E}_{\boldsymbol{\lambda}}[X^k]$ for all $k \geq 1$. We will define a notion of a colored permutation statistic being *realizable over constraints of size m* in Definition 4.1.5. Using this, we establish the following result, which formalizes the statement that the k-th moment of a colored permutation statistic stabilizes on conjugacy classes without "short" cycles, where “short” is in terms of m and k.

Theorem 1.2.10 (Campion Loth, Levet, Liu, Sundaram, and Yin ’24+)**.** *Suppose* $X : \mathfrak{S}_{n,r} \to \mathbb{R}$ *is realizable over constraints of size* m*, where* m *is some positive integer. For any* $k \geq 1$*, the* k*-th moment* $\mathbb{E}_{\boldsymbol{\lambda}}[X^k]$ *coincides on all conjugacy classes* $C_{\boldsymbol{\lambda}}$ *of* $\mathfrak{S}_{n,r}$ *with no cycles of length* $1, 2, \ldots, mk$*.*

Many statistics have natural decompositions that show they are realizable over constraints of some size. We will show for example that the descent, major index, and flag-major index statistics on $\mathfrak{S}_{n,r}$ are realizable over constraints of size 2. The definitions for these statistics are also somewhat technical, so we will give them in Section 2.2. Theorem 1.2.10 then implies that the k-th moments for the descent, major index, or flag-major index statistic will coincide on all $C_{\boldsymbol{\lambda}}$ without cycles of length $1, 2, \ldots, 2m$. For these statistics, we will also show a stronger statement.

Theorem 1.2.11 (Levet, Liu, Sundaram, and Yin ’24+)**.** *Let* X *be the descent, major index, or flag-major index statistic on* $\mathfrak{S}_{n,r}$*, and let* $C_{\boldsymbol{\lambda}}$ *be a conjugacy class of* $\mathfrak{S}_{n,r}$*. If* $C_{\boldsymbol{\lambda}}$ *has no cycles of length* $1, 2, \ldots, 2k$*, then the* k*-th moment of* X *on* $C_{\boldsymbol{\lambda}}$ *matches the* k*-th moment on* $\mathfrak{S}_{n,r}$*.*

Analogs of Theorem 1.2.6 and Theorem 1.2.7 are known for the statistics in Theorem 1.2.11 [CM12]. Consequently, the preceding theorem can be used to establish the following corollary. Since the descent, major index, and flag-major index statistics on $\mathfrak{S}_{n,r}$ generalize the descent and major index statistics on the symmetric group $\mathfrak{S}_n$, these results generalize Theorem 1.2.8 and Corollary 1.2.9.

Corollary 1.2.12 (Levet, Liu, Sundaram, and Yin ’24+)**.** *For every* $n \geq 1$*, let* $C_{\boldsymbol{\lambda}_n}$ *be a conjugacy class of* $\mathfrak{S}_{n,r}$ *such that for all* i*, the number of cycles of length* i *in* $\boldsymbol{\lambda}_n$ *approaches 0 as* $n \to \infty$*. Let* X_n *be the descent, major index, or flag-major index statistic on* $C_{\boldsymbol{\lambda}_n}$ *with mean* μ_n *and variance* σ_n^2*. Then as* $n \to \infty$*, the random variable* $(X_n - \mu_n)/\sigma_n$ *converges in distribution to the standard normal distribution.*

It is interesting to note that the hyperoctahedral group $B_n \cong \mathfrak{S}_{n,2}$ also has a descent statistic from its Coxeter group presentation, but this does not align with the general descent statistic on $\mathfrak{S}_{n,r}$. However, in Theorem 4.2.8 and Corollary 4.2.9, we will show similar results for the Coxeter-descent statistic on B_n.

1.2.2 Related Work

In the far-reaching paper [HR22] of Hamaker and Rhoades, the authors introduced *local* statistics and used character theory to study their moments on conjugacy classes. Their notion of "k-local" coincides with our notion of "realizable over constraints of size k." In particular, our work applied to $\mathfrak{S}_n$ gives combinatorial interpretations for some of the results established in [HR22]. Prior to the work of Hamaker and Rhoades, others also used the irreducible characters of $\mathfrak{S}_n$ to study permutation statistics, e.g., see [GR20, Gil13, Hul14].

In addition to Corollary 1.2.9, asymptotic results for some other statistics on conjugacy classes of $\mathfrak{S}_n$ are known. For descents, Kim and Lee [KL20] extended Corollary 1.2.9 to arbitrary conjugacy classes of $\mathfrak{S}_n$, where the asymptotic distribution depends only on the limiting proportion of fixed points. Our Theorem 1.2.11 is an analog of Fulman's original result for $\mathfrak{S}_{n,r}$, and exploring the asymptotic distribution of descents on arbitrary conjugacy classes of $\mathfrak{S}_{n,r}$ is one potential direction for future work.

1.3 Cosine functions

Finally, we consider a seemingly elementary problem at the intersection of analysis and combinatorics introduced by Stefan Steinerberger. For any finite set of positive integers $\{a_1, \ldots, a_n\} \subseteq \mathbb{Z}_+$, consider the associated functions $\cos(a_1 x)$, $\cos(a_2 x)$,$\ldots$, $\cos(a_n x)$ and ask the following question: if x is chosen uniformly at random from $[0, 2\pi]$, what is the probability that all of these n numbers have the same sign? Letting $|S|$ denote the total length of any subset $S \subseteq \mathbb{R}$ that is a union of bounded intervals, we are interested in

$$\mathbb{P}(a_1, \ldots, a_n) = \frac{1}{2\pi} \left| \left\{ x \in [0, 2\pi] : \min_{1 \leq i \leq n} \cos(a_i x) > 0 \quad \text{or} \quad \max_{1 \leq i \leq n} \cos(a_i x) < 0 \right\} \right|. \quad (1.3.1)$$

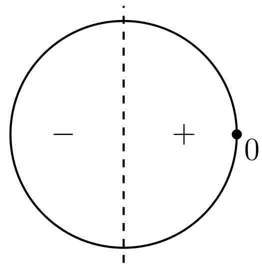


Figure 1.7: For n runners on a circular track of length 1 all starting at 0 and running at a constant integer speed, what is the proportion of time that they all spend in the right half (+) or the left half (−)?

One way to interpret this probability is illustrated in Figure 1.7. Imagine n runners starting on the right end of a circular track. They then run around this track at varying speeds, where $a_i \in \mathbb{Z}_+$ is the number of loops completed by runner i per unit of time. The probability in (1.3.1) then calculates the proportion of time that the runners are either all together on the right-hand side of the track or all together on the left-hand side of the track. This situation also relates to the Lonely Runner Conjecture of Cusick [Cus73] and Wills [Wil67], which states that regardless of the speed of the n runners, each runner gets *lonely* at some time, meaning that they are distance at least $1/n$ from all other runners. This conjecture is known to be difficult and only understood for small n and special settings. We refer the reader to [BS08, BHK01, Cze12, GW06, Kra21, PS14, Ren04, Tao17] for an incomplete list of results.

Returning to (1.3.1), it is clear that for any choice of $\{a_1, \ldots, a_n\} \subseteq \mathbb{Z}_+$, the probability in (1.3.1) has to be positive because for values of x near 0 or 2π, all of the cosine functions are close to 1. Using the convention $a_1 < a_2 < \ldots < a_n$, one can use this idea to show $\mathbb{P}(a_1, \ldots, a_n) \geq 1/(2a_n)$. A natural question is to determine what values of $a_1, \ldots, a_n$ minimize $\mathbb{P}(a_1, \ldots, a_n)$. Hence, we define

$$p_n = \inf_{\{a_1,\ldots,a_n\} \subseteq \mathbb{Z}_+} \mathbb{P}(a_1, \ldots, a_n).$$

It is less clear whether p_n is strictly positive, what general upper bounds exist as n grows large, or if there exists $\{a_1, \ldots, a_n\} \subseteq \mathbb{Z}_+$ achieving $\mathbb{P}(a_1, \ldots, a_n) = p_n$. When such a subset $\{a_1, \ldots, a_n\}$ exists, we call it a *subset achieving* p_n. Since $\mathbb{P}(a_1, \ldots, a_n) = \mathbb{P}(ca_1, \ldots, ca_n)$ for any $c \in \mathbb{Z}_+$, it suffices to consider only the sets $\{a_1, \ldots, a_n\} \subseteq \mathbb{Z}_+$ satisfying $\gcd(a_1, \ldots, a_n) = 1$, where gcd denotes greatest common divisor.

1.3.1 Main Results

A natural intuition is that if we fix $\{a_1, \ldots, a_{n-1}\} \subseteq \mathbb{Z}_+$ and let a_n grow very large, then $\mathbb{P}(a_1, \ldots, a_{n-1}, a_n)$ will approach half of $\mathbb{P}(a_1, \ldots, a_{n-1})$, as large values of a_n result in $\cos(a_n x)$ oscillating significantly on any interval where $\cos(a_1 x), \ldots, \cos(a_{n-1} x)$ all share the same sign. Hence, if we choose $a_1, \ldots, a_n$ of different magnitudes, we expect (1.3.1) to be close to $1/2^{n-1}$. The following result shows that there are subsets that are better than this.

Theorem 1.3.1 (Dou, Goh, Liu, Legate, and Pettigrew '24)**.** *For any $n \geq 2$, we have that*

$$\mathbb{P}\left(1, 3, 9, \ldots, 3^{n-1}\right) = \frac{1}{3^{n-1}}.$$

Consequently, we have that $p_n \leq 1/3^{n-1}$.

There is a precise result for p_2 due to Gonçalves, Oliveira e Silva, and Steinerberger [GOS21], which we now state. For any $S \subseteq \mathbb{Z}_+$ and $c \in \mathbb{Z}_+$, we use $c \cdot S$ for the set $\{c \cdot s \mid s \in S\}$.

Theorem 1.3.2. *[GOS21, Proof of Theorem 1] We have*

$$\mathbb{P}(a_1, a_2) \geq \frac{1}{3}$$

with equality if and only if $\{a_1, a_2\} = \gcd(a_1, a_2) \cdot \{1, 3\}$. Hence, we have that $p_2 = 1/3$.

This shows that in the case $n = 2$, multiples of the subset in Theorem 1.3.1 are precisely the ones achieving p_2. Our next result provides the corresponding characterization for p_3.

Theorem 1.3.3 (Dou, Goh, Liu, Legate, and Pettigrew '24)**.** *We have*

$$\mathbb{P}(a_1, a_2, a_3) \geq \frac{1}{9}$$

with equality if and only if $\{a_1, a_2, a_3\} = \gcd(a_1, a_2, a_3) \cdot \{1, 3, 9\}$. *Hence, we have that* $p_3 = 1/9$.

Our proof uses Fourier analysis to establish that if $\{a_1, a_2, a_3\}$ is a subset achieving p_3, then most of the elements in $\{a_1, a_2, a_3\}$ must be relatively small. Using the conventions $a_1 < a_2 < a_3$ and $\gcd(a_1, a_2, a_3) = 1$, we specifically show that $a_1 = 1$ and $a_2 \leq 7$. We then establish that the remaining element a_3 cannot be too much larger than a_2. Our specific result is that $a_3 \leq 84$, which reduces the problem to a finite search space.

Naturally, one may be tempted to conjecture a general pattern and expect that the powers of 3 given in Theorem 1.3.1 characterize subsets achieving p_n. This is not the case, as an explicit computation shows

$$\mathbb{P}\left(1, 3, 11, 33\right) = \frac{1}{33} < \frac{1}{27} = \mathbb{P}\left(1, 3, 9, 27\right).$$

Using Monte-Carlo sampling to narrow down a list of candidates $1 \leq a_1 < a_2 < a_3 < a_4 \leq 105$ and then performing an explicit calculation using Lemma 5.1.4, we believe that multiples of $\{1, 3, 11, 33\}$ are the subsets achieving p_4. As for $n = 5$, a similar process has identified $\{1, 3, 11, 35, 105\}$ as a potential subset achieving $p_5 \leq \mathbb{P}(1, 3, 11, 35, 105) = 1/105$. These examples show that finding p_n for general n is not an elementary problem.

1.3.2 Related Work

Theorem 1.3.2 is established in [GOS21], but phrased in a different setting. In that paper, Gonçalves, Oliveira e Silva, and Steinerberger explored the eigenvalues and eigenfunctions of differential operators. For certain operators, each eigenfunction can be expressed as a trigonometric function up to a small error, and this expression is called a WKB expansion, named after Wentzel, Kramers, and Brillouin. See [Hal13] for the general theory of WKB expansions, including their applications to the Schrödinger operators used in mathematical physics. The problem of finding p_n arises from analyzing sign patterns of these eigenfunctions at different points using the WKB expansion. We will briefly outline the main ideas below, focusing our attention on an example based on what is described in [GOS21] and [GOS17]. We refer the reader to those papers and the references therein for a more thorough treatment, including mathematically rigorous statements.

Consider the Schrödinger operator $H = -\frac{d^2}{dt^2} + t^2$ on functions defined over $\mathbb{R}$. This is a special case of the quantum harmonic oscillator, which models oscillation around an equilibrium point with some potential energy function around this point. The corresponding eigenfunctions are the Hermite polynomials

$$H_m(t) = (-1)^n e^{t^2} \frac{d^m}{dt^m}(e^{-t^2}),$$

where m ranges over the nonnegative integers. Ordinarily, one might expect the sign of H_m at different points $a \neq b$ to be unrelated as m varies. However, computer calculations for $a = 1/2$, $b = 5/2$, and various values of m suggest that

$$\lim_{m\to\infty} \frac{1}{m} |\{1 \leq i \leq m \mid \operatorname{sgn}(H_i(1/2)) = \operatorname{sgn}(H_i(5/2))\}| \approx \frac{3}{5}, \tag{1.3.2}$$

where sgn is the sign function. For any fixed real numbers $a \neq b$, Gonçalves, Oliveira e Silva, and Steinerberger called the above quantity the *sign correlation limit* of a and b. Their result [GOS21, Theorem 1] establishes bounds on the sign correlation limit of two points, and Theorem 1.3.2 arises when establishing the lower bound.

The connection to our problem involving cosine functions can be seen using the WKB expansion of the Hermite functions

$$\frac{\Gamma(m/2+1)}{\Gamma(m+1)} e^{-t^2/2} H_m(t) = \cos\left(t\sqrt{2m+1} - \frac{m\pi}{2}\right) + O\left(\frac{1}{\sqrt{m}}\right),$$

where Γ is the gamma function. One can consider various cases for m mod 4, each of which can be expressed in terms of a cosine function using trigonometric properties. For example,

$$\frac{\Gamma(2m+1)}{\Gamma(4m+1)} e^{t^2/2} H_{4m}(t) = \cos(t\sqrt{8m+1}) + O\left(\frac{1}{\sqrt{m}}\right),$$

so up to an error that disappears as $m \to \infty$, we see that $H_{4m}(t)$ and $\cos(t\sqrt{8m+1})$ have the same sign. Additionally, one can show that the sequence $\sqrt{8m+1}$ mod 2π is equidistributed on $[0, 2\pi]$. A similar analysis can be performed for other cases of m mod 4, and consequently, one can analyze the proportion of time that $\cos(ax)$ and $\cos(bx)$ share the same sign to study the sign correlation limit of a and b.

Applying this at the values $a = 1/2$ and $b = 5/2$, we consider the proportion of time $\cos\left(\frac{1}{2}x\right)$ and $\cos(\frac{5}{2}x)$ share the same sign. Scaling both parameters, one can instead consider $\cos(x)$ and $\cos(5x)$, and we can restrict our attention to the interval $[0, 2\pi]$ since these functions complete an integer number of cycles on $[0, 2\pi]$. This general approach leads to the definition given in (1.3.1), and one can show $\mathbb{P}(1, 5) = 3/5$, which is consistent with (1.3.2). Our problem of calculating p_n for general n corresponds to finding lower bounds on the sign correlation limit at n different points that can be scaled simultaneously into integers.

Outline

We start in Chapter 2 by outlining general notation and background for our work. Chapter 3 will focus on planar tanglegrams. We prove Theorem 1.1.1 in Section 3.1, establish our enumerative results in Section 3.2, and use these to study sampling in Section 3.3, which includes establishing Theorem 1.1.2. Chapter 4 focuses on colored permutation statistics. We establish Theorem 1.2.10 in Section 4.1, and establish individual parts of Theorem 1.2.11 and Corollary 1.2.12 throughout Section 4.2. Finally, Chapter 5 considers cosine functions. We begin with results for general n in Section 5.1, including establishing Theorem 1.3.1. We then study $n = 3$ in Section 5.2 and prove Theorem 1.3.3. At the end of various sections, we will also discuss open problems that arise from our results to promote future work.

Chapter 2
BACKGROUND AND NOTATION

We start by giving a general background and defining our notation for several areas relevant to our work. These will be organized into two sections involving material related to graph theory and material related to colored permutation groups. Throughout, we will assume some background in algebra, analysis, probability theory, and combinatorics at a level that appears in undergraduate courses. Specifically, we assume familiarity with the symmetric group as given in [Hun12], integration as given in [Rud76], univariate probability distributions as given in [ASV17], and graphs and enumeration as given in [Bón17].

2.1 Graphs and tanglegrams

In this section, we outline our general terminology and notation for graphs. We then give background on trees and tanglegrams.

2.1.1 Graph theory

A graph will be denoted $G = (V, E)$ where V is a set of vertices and E is a set of edges, which we express as pairs uv with $u, v \in V$. The *size* of G, denoted $|G|$, is its number of vertices. Unless otherwise stated, a graph is undirected, so the pairs in E are not ordered. In general, a graph $G = (V, E)$ can have repeated edges or loops, and when neither of these exists, G is called *simple*. For any $v \in V$, we use $\deg(v)$ to denote its degree in G, and a graph is d-regular if all vertices have degree d. Additionally, recall that a graph is *bipartite* if V can be expressed as a disjoint union $X \sqcup Y$ such that all edges in E consist of a vertex from X and a vertex from Y. In this case, we also express G as (X, E, Y).

A graph $G = (V, E)$ is typically illustrated in the plane with vertices as distinct points and edges as curves that can potentially cross. We call this a *drawing* of G. In the case of a directed graph, edges $uv \in E$ are drawn with directed arrows from u to v. We give examples of these definitions below.

Example 2.1.1. The d-dimensional hypercube $Q_d = (V_d, E_d)$ is the graph with vertex set V_d given by d-dimensional vectors with entries in $\{0, 1\}$ and edges between any two vertices that differ in exactly one coordinate. In general, Q_d is d-regular and bipartite. The latter can be shown by letting X_d be the vertices whose entries sum to be an even number and letting Y_d be the vertices whose entries sum to an odd number. Two drawings for Q_3 are shown in Figure 2.1, one based on coordinates in $\mathbb{R}^3$ and the other based on its bipartite structure.

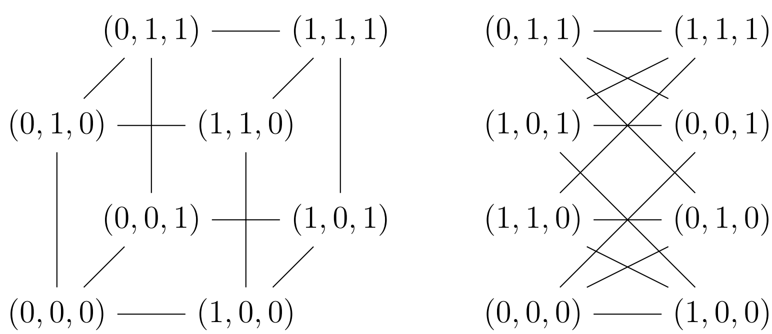


Figure 2.1: Two drawings for the 3-dimensional hypercube Q_3, the first based on coordinates in $\mathbb{R}^3$ and the second based on its bipartite structure.

Example 2.1.2. Let V_n be the triangulations of a convex n-gon. These are one of many objects enumerated by the Catalan numbers. In any triangulation, replacing a diagonal with the other possible diagonal in the unique quadrilateral containing it is called a *flip*. The triangulation flip graph $G_n = (V_n, E_n)$ is the graph formed by the vertex set V_n with edges E_n between any triangulations that can be obtained from one another using a flip. One can show that in general, G_n is connected and $(n-3)$-regular. The graph G_5 is shown in Figure 2.2.

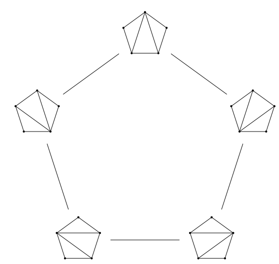


Figure 2.2: The flip graph G_5 on triangulations of a convex pentagon.

We now consider a specific random walk on a graph $G = (V, E)$. While the terminology below can be stated in the language of *Markov chains*, we will not need this level of generality.

Definition 2.1.3. Let $G = (V, E)$ be a graph. The *simple random walk* on G is defined as follows: choose X_0 to be an arbitrary vertex $u \in V$, and for $t \geq 1$, choose X_t from the neighbors of X_{t-1} uniformly at random.

Once a starting vertex $u \in V$ is chosen, each X_t in the random walk follows a probability distribution on the vertices in V. We use $\Pr_u[X_t = v]$ for the value of this distribution at $v \in V$. We will need a few more technical definitions and results involving these distributions.

Definition 2.1.4. Let $G = (V, E)$ be a graph. The *period* of $u \in V$ is

$$\gcd(t \in \mathbb{Z}_+ \mid \Pr_u[X_t = u] > 0). \tag{2.1.1}$$

Lemma 2.1.5. *[LP17, Lemma 1.6 and Exercise 1.2] If G is connected, then the period of every vertex $u \in V$ is the same.*

Definition 2.1.6. The *period* of the simple random walk $(X_t)_{t\geq 0}$ on a connected graph $G = (V, E)$ is the common period of all vertices $u \in V$, and $(X_t)_{t\geq 0}$ is *aperiodic* if its period is 1.

The following theorem shows that when $(X_t)_{t\geq 0}$ is aperiodic, the asymptotic distribution as $t \to \infty$ is well-understood. We will be particularly interested in the case of a regular graph, where this asymptotic distribution is the uniform distribution on the vertices in G.

Theorem 2.1.7. *[LP17, Section 1.5 and Theorem 4.9] Let $(X_t)_{t\geq 0}$ be a simple random walk on a connected graph $G = (V, E)$ that starts at $u \in V$. If $(X_t)_{t\geq 0}$ is aperiodic, then for any $v \in V$,*

$$\lim_{t\to\infty} \Pr_u[X_t = v] = \frac{\deg(v)}{\sum_{w\in V} \deg(w)}.$$

In particular, if G is d-regular, then

$$\lim_{t\to\infty} \Pr_u[X_t = v] = \frac{1}{|V|}.$$

One consequence of this theorem is that sampling from some set of objects V uniformly at random can be approximated by using a random walk on an appropriate d-regular graph on V, where edges correspond to some well-understood "local" operation. The triangulation flip graphs described in Example 2.1.2 are examples of this approach.

Note that the aperiodicity assumption is essential in Theorem 2.1.7 for the convergence to occur. For example, Theorem 2.1.7 does not hold on the hypercube Q_d since its bipartite structure causes $\Pr_u[X_t = v]$ to behave differently depending on whether t is odd or even. Aperiodicity is a relatively mild assumption though. One can always introduce *lazyness*, which is a nonzero probability of remaining at the same vertex. Adding loops at each vertex would be one method of achieving this.

2.1.2 Trees and tanglegrams

Recall that a *tree* is a connected graph that has no cycles. A *rooted binary tree* T is a tree in which every vertex has either zero or two children. In this case, a unique vertex has degree 2. This is called the root of T and will be denoted $\mathrm{root}(T)$. A vertex that has children is called an *internal vertex*, and a vertex with no children is called a *leaf*. Unless otherwise stated, rooted binary trees will be unlabeled and considered up to isomorphism. By ordering the children of each vertex, one obtains a *plane tree*, which has essentially a unique drawing

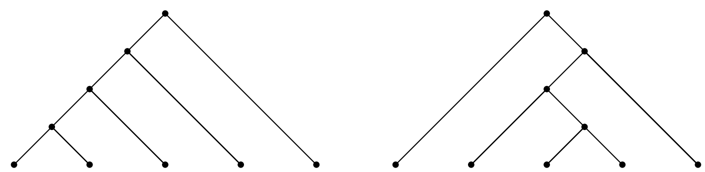


Figure 2.3: Two plane trees corresponding to the same rooted binary tree.

in the sense that this drawing is unique up to homeomorphism of the plane. Examples are shown in Figure 2.3.

If v has children v_1 and v_2, we call v the *parent* of v_1 and v_2. We say vertex v_1 is a *descendant* of v_k or v_k is an *ancestor* of v_1 if there is a sequence of vertices $v_1, v_2, \ldots, v_k$ such that v_{i+1} is the parent of v_i for $i = 1, 2, \ldots, k-1$, and we use the notation $v_1 < v_k$ or $v_k > v_1$ to denote this. For an internal vertex $v \in T$, the *subtree rooted at* v is the tree formed by all vertices u with $u \leq v$, and this subtree then has v as its root. For any vertex $v \in T$, we use $\mathrm{Lf}(v)$ to denote the leaves in the subtree rooted at v.

A tanglegram $\mathcal{T} = (L, R, \sigma)$ is formed from a pair of rooted binary trees L, R and a bijection σ matching their leaves. The *size* of $\mathcal{T}$ is the common number of leaves in L or R, and this is denoted $|\mathcal{T}|$. We will call the edges in L and R *tree edges* and call the edges induced by ϕ *between-tree edges*. Two tanglegrams $\mathcal{T} = (L, R, \sigma)$ and $\mathcal{T}' = (L', R', \sigma')$ are *isomorphic* if there is an isomorphism of graphs that maps L to L' and R to R'. As with trees, we will consider tanglegrams up to isomorphism. Additionally, in general, tanglegrams will be unlabeled, though at times, it will be convenient to fix labelings of the leaves in L and R using $[n]$.

A layout of a tanglegram $\mathcal{T} = (L, R, \sigma)$ embeds L in the plane left of $x = 0$ with leaves on $x = 0$, embeds R in the plane right of $x = 1$ with leaves on $x = 1$, and draws σ using straight lines. Examples are shown in Figure 2.4. A *crossing* is any pair of crossing edges induced by σ, and a tanglegram is *planar* if it has a layout with no crossings, which is called a *planar layout*. Of the thirteen tanglegrams shown in Figure 2.4, eleven are planar. The remaining two are not planar, and it is shown in [CSW19] that any non-planar tanglegram of size 4

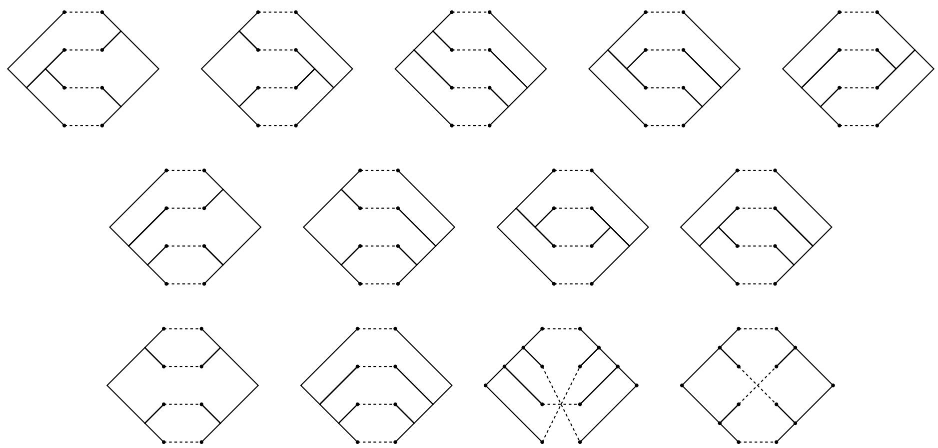


Figure 2.4: A layout for each of the 13 tanglegrams of size four.

or more "contains" one of these two. Note that crossings only depend on the relative order of the matched leaves in $\mathcal{T}$ rather than specific coordinates. Consequently, we will consider layouts up to redrawing the plane trees corresponding to L and R.

For a tanglegram $\mathcal{T} = (L, R, \sigma)$, if $\mathrm{Lf}(u)$ and $\mathrm{Lf}(v)$ are matched by σ for some internal non-root vertices $u \in L$ and $v \in R$, then the subtrees rooted at v_1 and v_2 with the matching induced by σ is called a *proper subtanglegram* of $\mathcal{T}$. A tanglegram $\mathcal{T}$ is *irreducible* if it contains no proper subtanglegrams. Of the eleven planar tanglegrams in Figure 2.4, only five are irreducible. A classical theorem of Whitney states that a 3-planar graph has only one planar drawing up to automorphism of the plane, and the following result of Ralaivaosaona, Ravelomanana, and Wagner is the corresponding result for planar layouts of an irreducible planar tanglegram.

Proposition 2.1.8. *[RRW18, Proposition 5] Every irreducible planar tanglegram $\mathcal{T}$ with $|\mathcal{T}| \geq 3$ has exactly two planar layouts. Moreover, the two planar layouts are mirror images of one another.*

The planar layouts for irreducible planar tanglegrams can be viewed as certain pairs of plane trees where the "horizontal" matching does not produce proper subtanglegrams. Since plane trees are Catalan objects, one natural question is to find bijections into other pairs

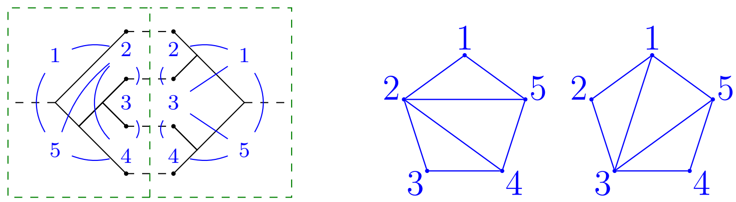


Figure 2.5: The plane dual bijection from Theorem 2.1.9.

of Catalan objects. For triangulations of a convex polygon, the following characterization is known. Note that two triangulations are *disjoint* if they do not share any diagonals.

Theorem 2.1.9. *[RRW18, Theorem 4] For any $n \geq 2$, there is a bijection between the following sets:*

- *the set of planar layouts of irreducible planar tanglegrams of size n, and*
- *the set of ordered pairs of disjoint triangulations of an $(n+1)$-gon.*

We describe this bijection. Starting with the two plane binary trees in a planar layout of an irreducible planar tanglegram $\mathcal{T} = (L, R, \sigma)$, draw lines from the root and all leaves to infinity. Then take the plane dual. Label the region above the root in each tree as 1. In L, proceed clockwise, and in R, proceed counterclockwise. An example is shown in Figure 2.5.

We conclude this section with some known enumerative results. Let $\mathcal{P}$ and $\mathcal{I}$ respectively denote the set of planar and irreducible planar tanglegrams. Define the generating functions

$$T(x) = \sum_{\mathcal{T} \in \mathcal{P}} x^{|\mathcal{T}|} \quad \text{and} \quad H(x) = \frac{1}{2}x^2 + \sum_{\mathcal{T} \in \mathcal{I}:\ |\mathcal{T}|>2} x^{|\mathcal{T}|}. \tag{2.1.2}$$

Ralaivaosaona, Ravelomanana, and Wagner showed the following results.

Theorem 2.1.10. *[RRW18, Theorem 1] The generating functions $H(x)$ and $T(x)$ are de-*

termined by the following relations involving a generating function $A(x)$:

$$\begin{aligned} A(x) &= \frac{x}{2}\sum_{r=1}^{\infty}\frac{1}{(r+1)^2}\binom{2r}{r}^2 x^r(1-A(x))^{r+1}, \\ H(x) &= \frac{x}{2}A(x), \\ T(x) &= H(T(x)) + x + \frac{1}{2}T(x^2). \end{aligned} \tag{2.1.3}$$

Theorem 2.1.10 can be used to compute coefficients of $A(x), H(x)$, and $T(x)$. For $A(x)$, one can express $A(x) = \sum_{n=1}^{\infty} a_n x^n$ and use this to expand $\frac{x}{2}\sum_{r=1}^{\infty}\frac{1}{(r+1)^2}\binom{2r}{r}^2 x^r(1-A(x))^{r+1}$ in terms of $\{a_n\}_{n=1}^{\infty}$. Using the relation for $A(x)$ from Theorem 2.1.10, one can set up and solve a system of equations to obtain $a_1, \ldots, a_n$ for arbitrary n. The coefficients of $H(x)$ can then be computed directly from $A(x)$, and the coefficients for $T(x)$ can then be computed using a similar technique as the one for $A(x)$.

2.2 Colored permutation groups

In this section, we give background on the colored permutation groups $\mathfrak{S}_{n,r}$, their conjugacy classes, and some commonly studied statistics. Our definitions are based on what is given in [Ste94], and we refer to [Mac98, Chapter 1, Appendix B] for results involving conjugacy classes of $\mathfrak{S}_{n,r}$. We also give additional background specific to $r = 2$, as these correspond to the type B Coxeter groups.

Before beginning, we briefly outline our notation for partitions, which index conjugacy classes of $\mathfrak{S}_n$. We will express each partition $\lambda \vdash n$ in multiplicative notation as $\lambda = (1^{m_1(\lambda)}, 2^{m_2(\lambda)}, \ldots, n^{m_n(\lambda)})$, where $i^{m_i(\lambda)}$ indicates that there are $m_i(\lambda)$ cycles of length i in λ. Terms of the form i^0 will be omitted, and terms of the form i^1 will just be denoted as i. The number of parts in λ will be denoted $\ell(\lambda)$, so that $\ell(\lambda) = \sum_i m_i(\lambda)$ and $n = \sum_i i \cdot m_i(\lambda)$. The centralizer of an element $\omega \in \mathfrak{S}_n$ with cycle type λ has order

$$z_\lambda = \prod_i i^{m_i(\lambda)} m_i(\lambda)!, \tag{2.2.1}$$

and the size of the conjugacy class C_λ in $\mathfrak{S}_n$ is given by $|\mathfrak{S}_n|/z_\lambda$.

2.2.1 Colored permutation groups

We start by describing colored permutation groups. Since $\mathfrak{S}_n \cong \mathfrak{S}_{n,1}$, these definitions specialize to the typical ones in $\mathfrak{S}_n$ when $r = 1$.

Definition 2.2.1. For positive integers n and r, the *colored permutation group* $\mathfrak{S}_{n,r}$ is the wreath product $\mathbb{Z}_r \wr \mathfrak{S}_n$, where $\mathfrak{S}_n$ is the symmetric group on n elements and $\mathbb{Z}_r$ is the cyclic group on r elements. A *colored permutation* is an element in $\mathfrak{S}_{n,r}$, and it can be expressed as (ω, τ) where $\omega \in \mathfrak{S}_n$ and $\tau : [n] \to \mathbb{Z}_r$ is a function called a *coloring*. From its construction as a wreath product, the group operation is defined as follows: for all $(\omega_1, \tau_1), (\omega_2, \tau_2) \in \mathfrak{S}_{n,r}$,

$$(\omega_1, \tau_1)(\omega_2, \tau_2) = (\omega_1\omega_2, (\tau_1 \circ \omega_2) + \tau_2),$$

where $\circ$ denotes function composition.

The colored permutation group $\mathfrak{S}_{n,r}$ can be embedded as a subgroup of the symmetric group $\mathfrak{S}_{rn}$, which we describe explicitly as follows. Let $[n]^r$ denote the set of rn elements

$$\{i^c \mid i \in [n],\, c \in \mathbb{Z}_r\},$$

where the exponent indicates the *color* of an element in $[n]$. We can view the colored permutation (ω, τ) as a bijection on $[n]^r$. We abuse notation and also denote this bijection (ω, τ), and it is defined by $(\omega, \tau)(i^c) = \omega(i)^{\tau(i)+c}$ for all $i \in [n]$ and $c \in \mathbb{Z}_r$. Observe that for all $i \in [n]$,

$$(\omega_1, \tau_1)(\omega_2, \tau_2)(i^0) = (\omega_1, \tau_1)(\omega_2(i)^{\tau_2(i)}) = \omega_1\omega_2(i)^{\tau_1(\omega_2(i))+\tau_2(i)} = (\omega_1\omega_2, (\tau_1 \circ \omega_2) + \tau_2)(i^0),$$

so this identification makes $\mathfrak{S}_{n,r}$ into a subgroup of $\mathfrak{S}_{nr}$.

Using the aforementioned identification, we can express $\mathfrak{S}_{n,r}$ in two-line and one-line notations as in $\mathfrak{S}_n$. However, the values of $(\omega,\tau)(i^0)$ for $i \in [n]$ determine $(\omega,\tau)(i^c)$ for any $c \in \mathbb{Z}_r$, so we do not need to specify the images of all elements in $[n]^r$. Instead, the two-line notation is a $2 \times n$ array that expresses $(\omega,\tau)(1^0), \ldots, (\omega,\tau)(n^0)$ under $1^0, 2^0, \ldots, n^0$ surrounded by a pair of square brackets, and the one-line notation is formed by removing the first line. An example is shown below.

Example 2.2.2. Let $\omega = [45132] = (143)(25) \in \mathfrak{S}_5$. Define $\tau : [5] \to \mathbb{Z}_4$ by

$$\tau(1) = 3,\ \tau(2) = 0,\ \tau(3) = 1,\ \tau(4) = 1,\ \tau(5) = 3.$$

Combined, this defines an element $(\omega,\tau) \in \mathfrak{S}_{5,4}$. Interpreting (ω,τ) as a bijection from $[5]^4$ to itself, the images of i^0 for $i \in [5]$ are given by

$$(\omega,\tau)(1^0) = 4^3, (\omega,\tau)(2^0) = 5^0, (\omega,\tau)(3^0) = 1^1, (\omega,\tau)(4^0) = 3^1, (\omega,\tau)(5^0) = 2^3.$$

One can express this colored permutation in two-line and one-line notations, as shown below:

$$(\omega,\tau) = \begin{bmatrix} 1^0 & 2^0 & 3^0 & 4^0 & 5^0 \\ 4^3 & 5^0 & 1^1 & 3^1 & 2^3 \end{bmatrix} = [4^3 5^0 1^1 3^1 2^3].$$

Colored permutations also have a cycle notation. Starting with (ω,τ), one can express ω in the usual cycle notation with color 0 on all elements and then write $\omega(i)^{\tau(i)}$ under i^0 for $i \in [n]$. We will refer to this as the *two-line cycle notation.* Removing the first row in every cycle then results in the cycle notation for (ω,τ).

Example 2.2.3. Let $(\omega,\tau) = [4^3 5^0 1^1 3^1 2^3] \in \mathfrak{S}_{5,4}$ be the permutation from Example 2.2.2. The two-line cycle notation is given by

$$(\omega,\tau) = \begin{pmatrix} 1^0 & 4^0 & 3^0 \\ 4^3 & 3^1 & 1^1 \end{pmatrix} \begin{pmatrix} 2^0 & 5^0 \\ 5^0 & 2^3 \end{pmatrix}.$$

The cycle notation is then given by $(4^3 3^1 1^1)(5^0 2^3)$. Note that as in $\mathfrak{S}_n$, the cycle notation is not unique, as one can change the element that appears first in each cycle and modify the remaining ones appropriately to express the same colored permutation. For example, $(1^1 4^3 3^1)(2^3 5^0)$ defines the same colored permutation.

We next describe the conjugacy classes of $\mathfrak{S}_{n,r}$. Similar to permutations in $\mathfrak{S}_n$, colored permutations have a notion of cycle type derived from the cycle notation. We will use a generalization of partitions to record cycle type, and we will see that the conjugacy classes of $\mathfrak{S}_{n,r}$ are well understood in terms of cycle type.

Definition 2.2.4. An *r-partition* of $n \in \mathbb{Z}_+$ is an r-tuple of partitions $\boldsymbol{\lambda} = (\lambda^j)_{j=0}^{r-1}$ where each λ^j is a partition of some nonnegative integer n_j such that $\sum_{j=0}^{r-1} n_j = n$. When $r = 2$, we also call this a *bi-partition*.

Definition 2.2.5. For any cycle in the cycle notation of $(\omega, \tau) \in \mathfrak{S}_{n,r}$, its *length* is the number of elements in it, and its *color* is the sum of the colors that appear (as an element in $\mathbb{Z}_r$). The *cycle type* of $(\omega, \tau) \in \mathfrak{S}_{n,r}$ is an r-partition $\boldsymbol{\lambda}$ where λ^j records the cycle lengths for the cycles with color j. For any such r-partition $\boldsymbol{\lambda}$, define $C_{\boldsymbol{\lambda}}$ to denote the subset of elements in $\mathfrak{S}_{n,r}$ with cycle type $\boldsymbol{\lambda}$.

Example 2.2.6. Consider the colored permutation in $\mathfrak{S}_{5,4}$ from Example 2.2.3 expressed in cycle notation as $(1^1 4^3 3^1)(2^3 5^0)$. The first cycle has length 3 and color 1, and the second cycle has length 2 and color 3. Since $r = 4$, the cycle type of this colored permutation is

$$\boldsymbol{\lambda} = (\lambda^0, \lambda^1, \lambda^2, \lambda^3) = (\emptyset, (3), \emptyset, (2)).$$

Example 2.2.7. For a larger example, consider the following colored permutation in $\mathfrak{S}_{9,3}$:

$$(\omega, \tau) = (1^0 3^2 7^1)(2^1)(4^2 5^0)(8^0)(9^1).$$

Since $r = 3$, the cycle type of this colored permutation is

$$\boldsymbol{\lambda} = (\lambda^0, \lambda^1, \lambda^2) = ((1,3), (1^2), (2)).$$

Proposition 2.2.8. *[Mac98, Chapter I, Appendix B.3] Two colored permutations in $\mathfrak{S}_{n,r}$ are conjugate if and only if they share the same cycle type. Hence, the conjugacy classes of $\mathfrak{S}_{n,r}$ are given by $C_{\boldsymbol{\lambda}}$, where $\boldsymbol{\lambda}$ is an r-partition of n.*

Throughout, we use $\Pr_{\mathfrak{S}_{n,r}}[X = i]$ and $\Pr_{\boldsymbol{\lambda}}[X = i]$ for the distribution of $X : \mathfrak{S}_{n,r} \to \mathbb{R}$ when considered as a random variable on the respective sets $\mathfrak{S}_{n,r}$ and $C_{\boldsymbol{\lambda}}$ equipped with the uniform distribution. We similarly use $\mathbb{E}_{\mathfrak{S}_{n,r}}[X]$ and $\mathbb{E}_{\boldsymbol{\lambda}}[X]$ for the corresponding expected values. It will sometimes be convenient to consider more general $\Omega \subseteq \mathfrak{S}_{n,r}$, and we similarly use $\Pr_{\Omega}[X = i]$ for the distribution X as a random variable on Ω.

We will primarily focus on three statistics on $\mathfrak{S}_{n,r}$: descents, major index, and flag-major index. In the case of $r = 1$, these will reduce to the usual descent and major index statistics on $\mathfrak{S}_n$.

For any $(\omega, \tau) \in \mathfrak{S}_{n,r}$, an index $i \in [n]$ is a *descent* of (ω, τ) if $\tau(i) > \tau(i+1)$, or $\tau(i) = \tau(i+1)$ and $\omega(i) > \omega(i+1)$, where we use the convention $\tau(n+1) = 0$ and $\omega(n+1) = n+1$. One can alternatively fix the following total order on r colored copies of $\mathbb{Z}$:

$$1^0 < 2^0 < \ldots < j^0 < \ldots < 1^1 < 2^1 < \ldots < j^1 < \ldots < 1^{r-1} < 2^{r-1} < \ldots < j^{r-1} < \ldots. \tag{2.2.2}$$

Viewing (ω, τ) as a bijection on $[n]^r$, a descent is then any $i \in [n]$ such that $(\omega, \tau)(i^0) > (\omega, \tau)((i+1)^0)$ with respect to this ordering, where we use the convention that N^0 for $N > n$ are fixed points. Note that a descent at position n occurs precisely when $\tau(n) \neq 0$, or equivalently when $(\omega, \tau)(n^0)$ is any element with nonzero color.

Letting $\operatorname{Des}(\omega, \tau)$ denote the set of descents of $(\omega, \tau) \in \mathfrak{S}_{n,r}$, the *descent* and *major index* statistics on $\mathfrak{S}_{n,r}$ are respectively defined as

$$\operatorname{des}_{n,r}(\omega,\tau) = |\operatorname{Des}(\omega,\tau)| \quad \text{and} \quad \operatorname{maj}_{n,r}(\omega,\tau) = \sum_{i\in \operatorname{Des}(\omega,\tau)\cap[n-1]} i. \tag{2.2.3}$$

The *color* and *flag-major index* statistics on $\mathfrak{S}_{n,r}$ are defined by

$$\operatorname{col}_{n,r}(\omega,\tau) = \sum_{i=1}^{n} \tau(i) \quad \text{and} \quad \operatorname{fmaj}_{n,r}(\omega,\tau) = r\cdot \operatorname{maj}_{n,r}(\omega,\tau) + \operatorname{col}_{n,r}(\omega,\tau). \tag{2.2.4}$$

Note that the col statistic uses $\{0,1,\ldots,r-1\}$ as representative elements in $\mathbb{Z}_r$ and adds them as elements in $\mathbb{Z}$. We give an example below.

Example 2.2.9. Consider $(\omega,\tau)\in\mathfrak{S}_{8,3}$ expressed in one line-notation as

$$(\omega,\tau) = [3^1 8^0 5^0 6^1 2^2 1^2 4^0 7^1]$$

The descent set of (ω,τ) is $\{1,2,5,6,8\}$, and the sum of the colors that appear is 7. Using this, we calculate

$$\operatorname{des}_{8,3}(\omega,\tau) = 5, \quad \operatorname{maj}_{8,3}(\omega,\tau) = 14, \quad \text{and} \quad \operatorname{fmaj}_{8,3}(\omega,\tau) = 3\cdot 14 + 7 = 49.$$

The asymptotic distributions of $\operatorname{des}_{n,r}$ and $\operatorname{fmaj}_{n,r}$ are due to Chow and Mansour. Since these will be relevant to our results, we state them below.

Theorem 2.2.10. *[CM12, Theorems 3.1 and 3.4] For any positive integers n and r, $\operatorname{des}_{n,r}$ has mean $\mu_{n,r} = \frac{rn+r-2}{2r}$ and variance $\sigma^2_{n,r} = \frac{n+1}{12}$, and as $n\to\infty$, the random variable $\frac{\operatorname{des}_{n,r}-\mu_{n,r}}{\sigma_{n,r}}$ converges in distribution to the standard normal distribution.*

Theorem 2.2.11. *[CM12, Theorems 4.1 and 4.3] For any positive integers n and r, $\operatorname{fmaj}_{n,r}$ has mean $\mu_{n,r} = \frac{n(rn+r-2)}{4}$ and variance $\sigma^2_{n,r} = \frac{2r^2n^3+3r^2n^2+(r^2-6)n}{72}$, and as $n\to\infty$, the random variable $\frac{\operatorname{fmaj}_{n,r}-\mu_{n,r}}{\sigma_{n,r}}$ converges in distribution to the standard normal distribution.*

For our results on the asymptotic distributions of certain statistics, we will need additional tools from probability theory. In general, two different probability distributions can share the same moments. We will be primarily interested in normal distributions, which are uniquely determined by their moments, so once we have that the moments of a random variable X align with those of a normal distribution, we can conclude that the distribution of X coincides with a normal distribution. See [ASV17, Section 5.1] for details. We will use this property for normal distributions in conjunction with the following tool.

Theorem 2.2.12 (Method of Moments)**.** *Suppose $\{X_n\}_{n\geq 1}$ and Y are real-valued random variables with finite k-th moments for all k. If Y is uniquely determined by its moments and*

$$\lim_{n\to\infty} \mathbb{E}[X_n^k] = \mathbb{E}[Y^k],$$

for all k, then X_n converges in distribution to Y.

2.2.2 Hyperoctahedral groups

We now consider the hyperoctahedral group $B_n \cong \mathfrak{S}_{n,2}$. While one can use the previous conventions involving $\mathfrak{S}_{n,2}$, the alternatives introduced here are intended to align with the literature on B_n. We will adopt the conventions given in this section whenever specifically considering B_n.

Let $[\pm n]$ denote the set $\{\pm 1, \pm 2, \ldots, \pm n\}$. The hyperoctahedral group B_n is the group of permutations ω on $[\pm n]$ satisfying $\omega(i) = -\omega(-i)$ with group operation given by function composition. A comparison with the definitions in the preceding section shows that this is isomorphic to the 2-colored permutation group $\mathfrak{S}_{n,2}$, where the set $\{1, 2, \ldots, n\}$ corresponds to elements colored 0 and the set $\{-1, -2, \ldots, -n\}$ corresponds to elements colored 1. Note that we will simply use ω to denote elements of B_n rather than the (ω, τ) notation used for general colored permutations.

Similar to general colored permutation groups, it suffices to specify the images of the set $[n]$ to determine an element in B_n. Using these values, elements in B_n can again be represented using the two-line, one-line, and cycle notations. In the cycle notation, a cycle is *even* (resp. *odd*) if there is an even (resp. odd) number of negatives in the cycle. These respectively correspond to cycles with color 0 and 1 in Definition 2.2.5. From Proposition 2.2.8, the conjugacy classes in B_n are uniquely determined by a bi-partition (λ, μ) of n, where λ records cycle lengths for even cycles and μ records cycle lengths for odd cycles. We give an example below.

Example 2.2.13. Consider

$$\begin{aligned}\omega &= \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 7 & 3 & -4 & 8 & -1 & -6 & -5 \end{bmatrix} \\ &= [2, 7, 3, -4, 8, -1, -6, -5] \\ &= (2, 7, -6, -1)(3)(-4)(8, -5).\end{aligned}$$

Its even cycles are $(2, 7, -6, -1)$ and (3), respectively of lengths 4 and 1, and its odd cycles are (-4) and $(8, -5)$, respectively of lengths 1 and 2. Hence the cycle type is $(\lambda, \mu) = ((1, 4), (1, 2))$.

Throughout, $C_{\lambda,\mu}$ denotes the conjugacy class of B_n indexed by the ordered pair (λ, μ). Additionally, we respectively use $\Pr_{B_n}[X = i]$ and $\Pr_{\lambda,\mu}[X = i]$ to denote the distribution of X as a random variable on B_n and $C_{\lambda,\mu}$ equipped with the uniform distribution. We use $\mathbb{E}_{B_n}[X]$ and $\mathbb{E}_{\lambda,\mu}[X]$ to denote the corresponding expected values.

The size of any conjugacy class in B_n is well-known. Recall the z_λ constants from (2.2.1), which gives the order of the centralizer for any $\omega \in \mathfrak{S}_n$ with cycle type λ.

Definition 2.2.14. For any bi-partition (λ, μ) of n, define $z_{\lambda,\mu}$ to be $2^{\ell(\lambda)} z_\lambda 2^{\ell(\mu)} z_\mu$.

Lemma 2.2.15. *[Mac98, Chapter I, Appendix B.3] Let $C_{\lambda,\mu}$ be the conjugacy class of B_n indexed by the bi-partition (λ, μ) of n. The order of the centralizer of an element of cycle type (λ, μ) is $z_{\lambda,\mu}$, and the size of the conjugacy class is*

$$|C_{\lambda,\mu}| = \frac{n! \cdot 2^n}{2^{\ell(\lambda)} z_\lambda 2^{\ell(\mu)} z_\mu} = \frac{|B_n|}{z_{\lambda,\mu}}.$$

We will be particularly interested in the descent statistic on B_n related to its presentation as a Coxeter group; see [BB05] for a detailed account. The *descent* statistic on B_n is defined as

$$\mathrm{des}_{B_n}(\omega) = |\{i \in \{0\} \cup [n-1] \mid \omega(i) > \omega(i+1)\}|,$$

with the convention that $\omega(0) = 0$. Note that the condition $\omega(i) > \omega(i+1)$ is with respect to the usual ordering on $[\pm n] \cup \{0\}$. The distribution of des_{B_n} is well-understood.

Theorem 2.2.16. *[Bre94, Theorem 3.4] For any positive integer n,*

$$\sum_{\omega \in B_n} t^{\mathrm{des}_{B_n}(\omega)} = (1-t)^{n+1} \sum_{k \geq 0} (2k+1)^n t^k.$$

Remark 2.2.17. The definition of des_{B_n} is not equivalent to the definition of $\mathrm{des}_{n,2}$ on $\mathfrak{S}_{n,2}$ under the usual isomorphism between these groups. Steingrímmson originally defined $\mathrm{des}_{n,r}$ for arbitrary positive integers n and r, and he showed in [Ste94, Theorem 17] that

$$\sum_{(\omega,\tau) \in \mathfrak{S}_{n,r}} t^{\mathrm{des}_{n,r}(\omega,\tau)} = (1-t)^{n+1} \sum_{k \geq 0} (rk+1)^n t^k.$$

Consequently, the distributions of des_{B_n} and $\mathrm{des}_{n,2}$ coincide on $B_n \cong \mathfrak{S}_{n,2}$, and Theorem 2.2.10 also applies to des_{B_n}. While the distributions for des_{B_n} and $\mathrm{des}_{n,2}$ coincide on $B_n \cong \mathfrak{S}_{n,2}$, they do not coincide on all conjugacy classes. For example, we can consider cycle type $(\emptyset, (1^n))$ in $B_n \cong \mathfrak{S}_{n,2}$. The single permutation with this cycle type satisfies

$$\mathrm{des}_{B_n}([-1,-2,\ldots,-n]) = n \neq 1 = \mathrm{des}_{n,2}([1^1, 2^1, \ldots, n^1]).$$

Chapter 3

PLANAR TANGLEGRAMS

In this chapter, we analyze planar tanglegrams. We prove Theorem 1.1.1 in Section 3.1 and consider the problem of counting the number of planar layouts of a planar tanglegram. We then establish some enumerative results in Section 3.2 and apply these in Section 3.3 to construct an algorithm for sampling planar tanglegrams. Many of the results up to Theorem 3.2.1 appear in [Liu23]. The remaining results are from a collaboration with Alex Black, Alex McDonough, Garrett Nelson, Michael Wigal, Mei Yin, and Youngho Yoo, and these results appear in [BLM+23].

3.1 Planar tanglegram layouts

In this section, we will give our characterization of the planar layouts of a planar tanglegram. This characterization leads to a connected graph on planar tanglegram layouts, and we will describe how to determine its number of vertices, as well as characterize extremal cases. Our results can be viewed as the analog of Proposition 2.1.8 for all planar tanglegrams.

3.1.1 A characterization of planar layouts

We start by characterizing the planar layouts of a planar tanglegram. Throughout this subsection, we will consider a planar tanglegram $\mathcal{T} = (L, R, \sigma)$ and fix some labeling of the vertices in L and R. In any layout, the number of crossings is completely determined by the order of the leaves in the two trees and the bijection σ matching these leaves, as the between-tree edges u_1v_1 and u_2v_2 intersect when u_1 is embedded above u_2 and v_1 is embedded below v_2. This leads to the following definition based on the notation used in [LPR+07].

Definition 3.1.1. Let $\mathcal{T} = (L, R, \sigma)$ be a tanglegram with some arbitrary labeling of the vertices in L and R. The *leaf order* of a layout is a pair of ordered lists (X, Y), where X and Y respectively list the leaves of T and S in order of appearance from top to bottom in the layout.

Note that one can recover a layout from the ordered lists (X, Y) by drawing the leaves listed in X and Y from top to bottom respectively on $x = 0$ and $x = 1$, and then using the information from L, R, and σ to draw the trees and between-tree edges. Throughout this section, we abuse terminology and refer to this pair of lists (X, Y) also as a layout.

We now give our `ModifiedUntangle` algorithm, based on the `Untangle` algorithm by Lozano et al. [LPR+07] for finding a planar layout of a planar tanglegram. This algorithm starts with the roots of L and R and iteratively replaces vertices with their children until only leaves remain. Lozano et al. called the ordered lists of vertices (X, Y) from each step *partial layouts*, and we retain this terminology. Our modifications from the original `Untangle` algorithm involve the set $\mathcal{L}$. We will explain the significance of this set later.

Input: ordered lists of vertices (A, B), $u \in A$, edges E on $A \cup B$, Boolean table P
Output: A, E after u has been replaced with its children

$u_1, u_2 :=$ children of u in $T \cup S$
for $j \in [m]$ *such that* $(u, b_j) \in E$ *where* $B = (b_1, \ldots, b_m)$ **do**
  update $E := E \setminus \{(u, b_j)\}$ `// delete edges involving` u
  for $i \in \{1, 2\}$ **do**
    if $P[u_i, b_j] = True$ **then**
      update $E := E \cup \{(u_i, b_j)\}$ `// insert edges involving` u_1 `or` u_2
    end
  end
end
$k := \max\{j \in [m] : (u_1, b_j) \in E\}$ `// last vertex in` B `adjacent to` u_1
if $j > k$ *for all* $(u_2, b_j) \in E$ **then**
  replace u with $u_1 u_2$ in A
end
else
  replace u with $u_2 u_1$ in A
end
return A, E

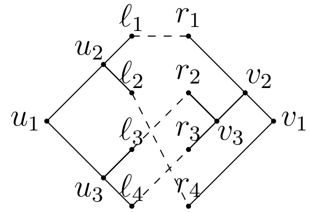
Algorithm 3.1.1: `Refine` (based on [LPR+07, Algorithm 3])

Input: planar tanglegram (L, R, σ) with some fixed labeling
Output: a planar layout (X, Y) of (L, R, σ) and set $\mathcal{L} \subseteq T \times S$ consisting of pairs of leaves

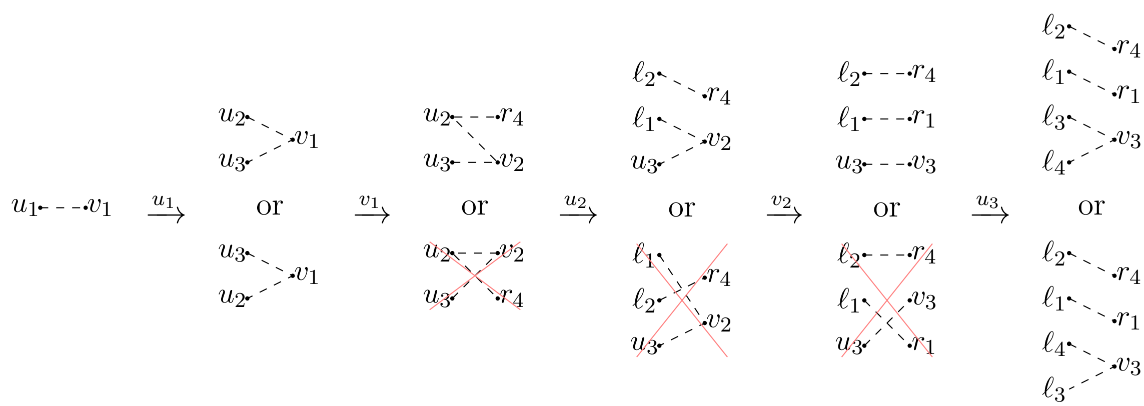
$P \coloneqq$ Boolean table with $P[u, v] = False \ \forall$ vertices $u \in L$, $v \in R$
set $P[u, v] = True \ \forall$ leaves $u \in L$ and $v \in R$, and then recursively set $P[u, v] = True$ for internal vertices $u \in L, v \in R$ if there exists $u' \leq_T u$, $v' \leq_S v$ with $P[u', v'] = True$
$X \coloneqq (\text{root}(L))$, $Y \coloneqq (\text{root}(R))$ as ordered lists
$E \coloneqq \{(\text{root}(L), \text{root}(R))\}$ as a set of edges
$\mathcal{L} \coloneqq \emptyset$
while $X \cup Y$ *contains an internal vertex of* L *or* R **do**
  $u \coloneqq$ internal vertex of $L \cup R$ with highest degree in the bipartite graph $G = (X, E, Y)$
  if $u \in X$ **then**
    if u *has degree 1 in* G **then**
      update $\mathcal{L} \coloneqq \mathcal{L} \cup (u, v)$, where v is the unique neighbor of u in G
    end
    update $X, E \coloneqq$ `Refine`(X, Y, u, E, P)
  end
  else if $u \in Y$ **then**
    if u *has degree 1 in* G **then**
      update $\mathcal{L} \coloneqq \mathcal{L} \cup (v, u)$, where v is the unique neighbor of u in G
    end
    update $Y, E \coloneqq$ `Refine`(Y, X, u, E, P)
  end
end
return $(X, Y), L$

Algorithm 3.1.2: `ModifiedUntangle` (based on [LPR+07, Algorithm 2])

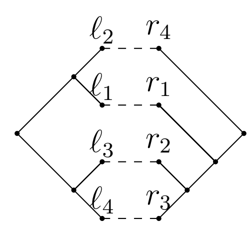
Example 3.1.2. Consider the tanglegram with labeling shown below.



`Untangle` starts with the roots of L and R. At each step, it draws edges according to the Boolean table P. It then uses `Refine` on a vertex u of highest degree in (X, E, Y), which replaces u with its children in some order. One way to visualize the chosen order is to draw the two bipartite graphs resulting from the two possible orders for the children of u and choose the one that does not have crossings. The following is a potential sequence of bipartite graphs considered by the algorithm, and the corresponding partial layout (X, Y) at each step is obtained by listing the vertices on each side from top to bottom. Note that at some steps, both choices lead to no crossings, and in this example, the graphs given in the top row are the ones chosen for future steps in the algorithm.



After one more application of `Refine` on the top right partial layout, the algorithm outputs the layout $(\ell_2\ell_1\ell_3\ell_4,\ r_4r_1r_2r_3)$ visualized below.



Lozano et al. showed that their `Untangle` algorithm results in a planar layout when the inputted tanglegram is planar. Since results from their proof will be relevant to our work, we give these below for the convenience of the reader. Note that Lemma 3.1.4 is stated more generally here than in [LPR$^+$07].

Definition 3.1.3. A partial layout (X, Y) for a tanglegram is called *promising* if it can be extended to a planar layout by successively replacing vertices with their children in some order.

Lemma 3.1.4. *[LPR+07, Lemma 3] Let (X, Y) be a promising partial layout of a planar tanglegram $\mathcal{T} = (L, R, \sigma)$, and let E be the set of edges on $X \cup Y$ generated using the Boolean table P, that is, for all $u \in X$ and $v \in Y$, $(u, v) \in E$ if and only if $P[u, v] = True$. Let u be a vertex of highest degree in the bipartite graph (X, E, Y).*

(a) If $\deg(u) = 1$, then replacing u with u_1u_2 or u_2u_1 results in a promising partial layout.

(b) If $\deg(u) > 1$, then either replacing u with u_1u_2 or replacing u with u_2u_1 results in a promising partial layout, but not both.

In particular, if (X, Y) is promising at the beginning of an iteration of the ***while*** *loop in* ***Untangle****, then it is promising at the end.*

Proof. Without loss of generality, we will assume $u \in L$, as the result when $u \in R$ is done similarly. We let (X_1, Y) and (X_2, Y) be the partial layouts obtained by replacing u with u_1u_2 or u_2u_1, respectively. Since (X, Y) is promising, at least one of these partial layouts must be promising, so assume that (X_1, Y) is promising.

First, suppose $\deg(u) = 1$ in (X, E, Y). Since u is a vertex of maximum degree, the unique neighbor of u, denoted v, also has degree 1. Since u and v have degree 1, $\mathrm{Lf}(u)$ and $\mathrm{Lf}(v)$ must be matched by ϕ. Extend (X_1, Y) to a planar layout (X', Y') of (T, S, ϕ), where $\mathrm{Lf}(u_1)$ appears before $\mathrm{Lf}(u_2)$. If we replace the drawings of the subtrees rooted at u and v with their mirror images, then we obtain a layout where $\mathrm{Lf}(u_2)$ appears before $\mathrm{Lf}(u_1)$, as shown in Figure 3.1. Notice that this is a planar layout that can be obtained by replacing u with u_2u_1 instead, implying (X_2, Y) is also promising. Then regardless of how we replace u with u_1 and u_2, the result is promising.

Next, suppose that $\deg(u) > 1$. As before, we suppose (X_1, Y) is promising. Let E' be the edges on $X_1 \cup Y$ constructed using the Boolean table P, and for $i = 1, 2$, we let $N(u_i)$

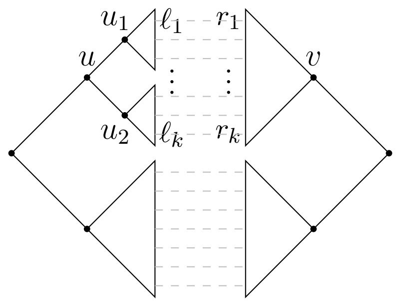


Figure 3.1: Starting with the layout (X', Y') on the left with $\{\ell_i\}_{i=1}^k$ and $\{r_i\}_{i=1}^k$ denoting leaves, performing flips at u and v produces another planar layout.

denote the set of neighbors of u_i in (X_1, E', Y). We claim that $N(u_1)\Delta N(u_2) \neq \emptyset$, where Δ denotes the symmetric difference of two sets. To see this, suppose that $N(u_1) = N(u_2)$. If $|N(u_1)| = |N(u_2)| = 1$, then this would imply $\deg(u) = 1$ in (X, E, Y), which by assumption cannot be the case. Hence, $|N(u_1)| = |N(u_2)| \geq 2$. Then there exists some pair of vertices $v_1, v_2 \in N(u_1) = N(u_2)$ that are each adjacent to both u_1 and u_2. However, this implies a crossing occurs in both (X_1, E', Y) and (X_2, E', Y), as shown in Figure 3.2. Then by construction of the Boolean table P, there exist some $\ell_{11}, \ell_{12} \in \text{Lf}(u_1)$ and $\ell_{21}, \ell_{22} \in \text{Lf}(u_2)$ where each $\ell_{ij} \in \text{Lf}(u_i)$ is matched to some $r_{ij} \in \text{Lf}(v_j)$. Regardless of any refinements of (X, Y), the resulting layout will have either a crossing formed from $\ell_{12}r_{12}$ and $\ell_{21}r_{21}$, or a crossing formed from $\ell_{11}r_{11}$ and $\ell_{22}r_{22}$. Since (X_1, Y) is assumed to be promising, it must be that $N(u_1) \neq N(u_2)$.

With $N(u_1)\Delta N(u_2) \neq \emptyset$ established, we let $v \in N(u_1)\Delta N(u_2)$. First, we assume that $v = v_1 \in N(u_1) \backslash N(u_2)$. Since (X_1, Y) is promising, it must be that drawing (X_1, E', Y) with

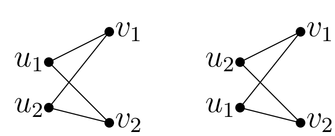


Figure 3.2: Arrangements of the vertices u_1, u_2, v_1 and v_2 when $N(u_1) = N(u_2)$ and $\deg(u) \geq 2$.

vertices in the orders indicated by X_1 and Y produces no crossings. Then v_1 must appear before any $v_2 \in N(u_2)$ in the list Y. Furthermore, for any $v_2 \in N(u_2)$, if we interchange u_1 and u_2, then the edges u_1v_1 and u_2v_2 will intersect. Regardless of our future refinements, there will exist some leaves $\ell_1 \in \mathrm{Lf}(u_1), \ell_2 \in \mathrm{Lf}(u_2), r_1 \in \mathrm{Lf}(v_1)$, and $r_2 \in \mathrm{Lf}(v_2)$ such that $\ell_1 r_1$ and $\ell_2 r_2$ intersect, which implies (X_2, Y) cannot be promising. A similar argument applies when $v = v_2 \in N(u_2) \setminus N(u_1)$.

Finally, consider the **while** loop in `Untangle`. The algorithm uses `Refine` on a vertex u of highest degree in (X, E, Y), which replaces u with u_1 and u_2. If $\deg(u) = 1$, then (a) shows that (X, Y) is promising regardless of the choice at u. If $\deg(u) > 1$, `Refine` replaces (X, Y) with (X_1, Y) or (X_2, Y) based on whichever bipartite graph (X_1, E', Y) or (X_2, E', Y) does not have crossings, and our proof of (b) shows that this results in a promising partial layout. □

Theorem 3.1.5. *[LPR+07, Theorem 1] For any planar tanglegram* $\mathcal{T} = (L, R, \sigma)$*, the* ***`Untangle`*** *algorithm terminates in a planar layout* (X, Y)*.*

Proof. If $\mathcal{T} = (L, R, \sigma)$ is planar, then $(X, Y) = (\mathrm{root}(L), \mathrm{root}(R))$ in line 3 is promising. By Lemma 3.1.4, (X, Y) is promising after each iteration of the **while** loop. This loop terminates when (X, Y) contains only leaves of L and R, which must then be a planar layout. □

Remark 3.1.6. In the proofs of Lemma 3.1.4 and Theorem 3.1.5, the arguments hold regardless of which vertex of highest degree u is selected. In fact, we do not even need to select the vertex of highest degree at every step in `Untangle`. We specified a vertex of highest degree for simplicity. As long as `Untangle` does not use `Refine` on a vertex in (X, E, Y) with degree one while its neighbor has degree more than one, the algorithm will still output a planar layout for a planar tanglegram.

Remark 3.1.7. Lozano et al. also showed that for a planar tanglegram of size n, their `Untangle` algorithm runs in $O(n^2)$ time and space, and the bottleneck occurs from computing the Boolean table P. The additional steps in `ModifiedUntangle` involve the set $\mathcal{L}$, which has size at most $n-1$. Hence, for a planar tanglegram of size n, `ModifiedUntangle` also runs in $O(n^2)$ time and space.

We now consider the additional steps involving the set $\mathcal{L}$ in the `ModifiedUntangle` algorithm. While the `Untangle` algorithm produces a single planar layout for any given planar tanglegram, one might ask how to generate all possible planar layouts. As observed in the proof of Lemma 3.1.4 and implicitly in Proposition 2.1.8, if (X, Y) is a planar layout of (L, R, σ), then one method to generate additional planar layouts is using mirror images at $u \in T$ and $v \in S$ where $\mathrm{Lf}(u)$ and $\mathrm{Lf}(v)$ are matched by ϕ. We give a name for these pairs of vertices and the operation involving mirror images at both vertices, followed by an example in Figure 3.3.

Definition 3.1.8. Let $\mathcal{T} = (L, R, \sigma)$ be a tanglegram. A pair of internal vertices (u, v) with $u \in L$ and $v \in R$ is a *leaf-matched pair* of $\mathcal{T}$ if $\mathrm{Lf}(u)$ and $\mathrm{Lf}(v)$ are matched by ϕ. Given a fixed planar layout, a *paired flip* at (u, v) in (X, Y) is a reflection at each of the subtrees rooted at u and v.

Note that in the leaf order of a layout (X, Y), this operation reverses the orders of $\mathrm{Lf}(u)$ and $\mathrm{Lf}(v)$ in X and Y, respectively. Additionally, the roots of L and R always form a leaf-matched pair, and when a leaf-matched pair does not consist of the roots, then they are the roots of a proper subtanglegram of $\mathcal{T}$. We now show that the set $\mathcal{L}$ from `ModifiedUntangle` is precisely the set of leaf-matched pairs.

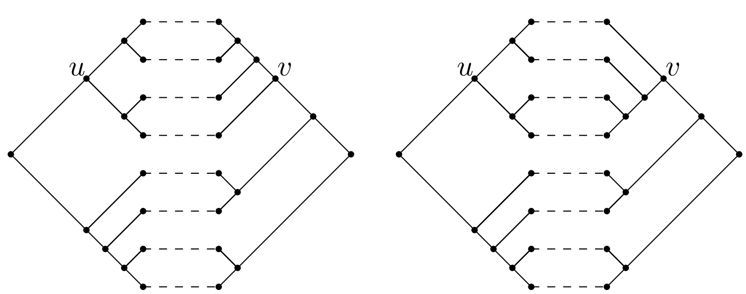


Figure 3.3: A paired flip at (u, v) maps each layout to the other one.

Lemma 3.1.9. *Let $\mathcal{T} = (L, R, \sigma)$ be a planar tanglegram. A pair of internal vertices (u, v) is a leaf-matched pair of $\mathcal{T}$ if and only if at some step of the* `ModifiedUntangle` *algorithm, the internal vertices $u \in L$ and $v \in R$ appear as adjacent degree one vertices in (X, E, Y). Hence, the set $\mathcal{L}$ returned by* `ModifiedUntangle` *is the set of leaf-matched pairs of $\mathcal{T}$.*

Proof. Suppose the internal vertices $u \in L$ and $v \in R$ appear as adjacent degree one vertices in (X, E, Y) during some step of the `ModifiedUntangle` algorithm. Since both vertices have degree one, the construction of the edges in (X, E, Y) using the Boolean table P in line 4 of `Refine` implies that $\mathrm{Lf}(u)$ and $\mathrm{Lf}(v)$ are matched under σ, and thus (u, v) is a leaf-matched pair.

Conversely, suppose (u, v) is a leaf-matched pair. If u and v are the root vertices of T and S, then these appear as degree one vertices at the first step of `ModifiedUntangle`. Otherwise, at some step of the algorithm, either u or v will appear for the first time in a partial layout (X, Y), and without loss of generality, we assume it is $u \in X$. Since v has not appeared in a partial layout yet, there is some vertex $v' \in Y$ that is an ancestor of v. Then $\mathrm{Lf}(u)$ is matched with a proper subset of $\mathrm{Lf}(v')$ in the tanglegram (L, R, σ), so $\deg(v') > 1$ in (X, E, Y). From line 7 of `ModifiedUntangle`, observe that v' will be replaced with its children before u is. Repeating this argument, v must appear before `Refine` is used on u, and at this time, u and v will be adjacent degree one vertices. □

We know that given a planar layout (X, Y) of (T, S, ϕ), paired flips will generate additional planar layouts, but we do not yet know that they generate all possible planar layouts. It may be possible that operations not equivalent to a sequence of paired flips also result in a planar layout. We will show that this is not the case. Our proof for this uses Lemma 3.1.4, which holds for arbitrary promising partial layouts, not just ones considered in `ModifiedUntangle`.

First, notice that for a tanglegram $\mathcal{T} = (L, R, \sigma)$ of size n, `ModifiedUntangle` starts with a promising partial layout $(X_1, Y_1) = (\mathrm{root}(L), \mathrm{root}(R))$. At each step, it replaces an internal vertex of L or R using `Refine`. Since a tree on n leaves has $n - 1$ internal vertices, the algorithm uses `Refine` a total of $2n - 2$ times. Thus, it produces a sequence of promising partial layouts $\{(X_k, Y_k)\}_{k=1}^{2n-1}$ that terminates at $(X, Y) = (X_{2n-1}, Y_{2n-1})$. We give a name for such a sequence.

Definition 3.1.10. Let (X, Y) be a planar layout of a tanglegram $\mathcal{T} = (L, R, \sigma)$. We call $\{(X_k, Y_k)\}_{k=1}^{2n-1}$ a *partial sequence for* (X, Y) if

- for $k = 1, 2, \ldots, 2n-1$, (X_k, Y_k) is a partial layout,
- $(X_1, Y_1) = (\mathrm{root}(L), \mathrm{root}(R))$,
- $(X_{2n-1}, Y_{2n-1}) = (X, Y)$, and
- for $k = 1, 2, \ldots, 2n-2$, the partial layout (X_{k+1}, Y_{k+1}) is obtained from (X_k, Y_k) by refining some vertex $u \in X_k \cup Y_k$, that is, replacing u with its children in an appropriate order.

A partial sequence is *promising* if all (X_k, Y_k) are promising partial layouts, or equivalently, if $(X, Y) = (X_{2n-1}, Y_{2n-1})$ is a planar layout.

If $\{(X_k, Y_k)\}_{k=1}^{2n-1}$ is a promising partial sequence for (X, Y), then for any other planar layout (X', Y') of $\mathcal{T}$, we can use this sequence to find a promising partial sequence $\{(X'_k, Y'_k)\}_{k=1}^{2n-1}$ for (X', Y'). We do this by constructing each (X'_k, Y'_k) as follows.

- Draw the trees L and R with leaves from top to bottom in the order described by (X', Y').
- For each $u \in X_k$, contract the subtree of L rooted at u to the vertex u itself. Do the same for each $v \in Y_k$, and call the resulting trees L_k and R_k.
- Let X'_k be the leaves of L_k listed from top to bottom, and similarly for Y'_k and R_k.

An example of these steps is shown in Figure 3.4. Note that by construction, (X_k, Y_k) and (X'_k, Y'_k) contain the same vertices, but possibly in different orders. We now show that

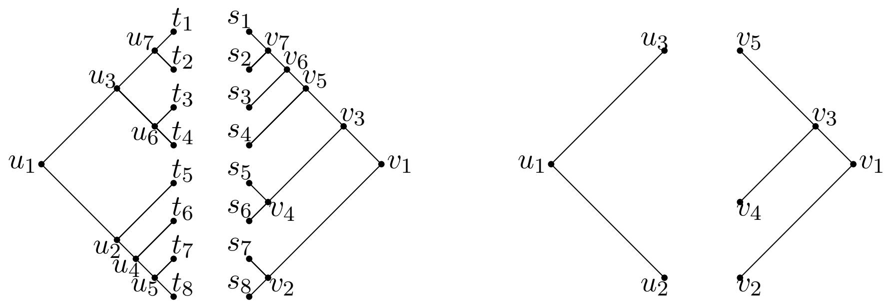


Figure 3.4: Starting with drawings of L and R corresponding to (X', Y'), we use $(X_k, Y_k) = (u_2u_3, v_2v_4v_5)$ to form the contracted trees L_k and R_k shown on the right. Listing leaves from top to bottom gives us the partial layout $(X'_k, Y'_k) = (u_3u_2, v_5v_4v_2)$.

$\{(X'_k, Y'_k)\}_{k=1}^{2n-1}$ constructed in this manner is a promising partial sequence for (X', Y') and then use this to establish our characterization of planar tanglegram layouts.

Lemma 3.1.11. *The sequence* $\{(X'_k, Y'_k)\}_{k=1}^{2n-1}$ *is a promising partial sequence for* (X', Y')*.*

Proof. By construction, each (X'_k, Y'_k) is a partial layout, $(X'_1, Y'_1) = (\mathrm{root}(L), \mathrm{root}(R))$, and $(X'_{2n-1}, Y'_{2n-1}) = (X', Y')$. It remains to show that refining a vertex $u \in X'_k \cup Y'_k$ produces (X'_{k+1}, Y'_{k+1}). Denote the vertex refined in (X_k, Y_k) to obtain (X_{k+1}, Y_{k+1}) as u_k, and without loss of generality, we assume $u_k \in X_k$. This implies that (X'_k, Y'_k) and (X'_{k+1}, Y'_{k+1}) have almost the same vertices, except (X'_k, Y'_k) contains u_k, while (X'_{k+1}, Y'_{k+1}) contains its two children $u_{k,1}$ and $u_{k,2}$. By our construction using contractions, the tree L_k can be obtained from L_{k+1} by contracting the two children of u_k onto the vertex itself. Thus, we can obtain (X'_k, Y'_k) from (X'_{k+1}, Y'_{k+1}) by replacing the adjacent children of u_k with u_k itself. Then we can also obtain (X'_{k+1}, Y'_{k+1}) from (X'_k, Y'_k) by refining u_k. Thus, all conditions in Definition 3.1.10 are satisfied, so $\{(X'_k, Y'_k)\}_{k=1}^{2n-1}$ is a promising partial sequence for (X', Y'). □

Theorem 3.1.12 (Theorem 1.1.1 revisited)**.** *Let $\mathcal{T}$ be a planar tanglegram, and let $\mathscr{P}(\mathcal{T})$ be its set of planar layouts. Each planar layout in $\mathscr{P}(\mathcal{T})$ can be obtained by starting with any planar layout in $\mathscr{P}(\mathcal{T})$ and performing some sequence of paired flips at leaf-matched pairs of vertices.*

Proof. Let the outputs of `ModifiedUntangle`($\mathcal{T}$) be (X, Y) and $\mathcal{L}$. By Theorem 3.1.5, we have that $(X, Y) \in \mathscr{P}(\mathcal{T})$, and by Lemma 3.1.9, we have that $\mathcal{L}$ is the set of leaf-matched pairs of $\mathcal{T}$. It is clear that if $(u, v) \in \mathcal{L}$, then starting with (X, Y) and performing a paired flip at (u, v) produces another layout in $\mathscr{P}(\mathcal{T})$, so the same is true if we perform any sequence of paired flips starting with (X, Y). We show that all planar layouts can be obtained this way.

Let $(X', Y') \in \mathscr{P}(\mathcal{T})$ be distinct from (X, Y). Let $\{(X_k, Y_k)\}_{k=1}^{2n-1}$ be the promising partial sequence for (X, Y) produced in `ModifiedUntangle`, with corresponding bipartite graphs $\{(X_k, E_k, Y_k)\}_{k=1}^{2n-1}$. By Lemma 3.1.11, we can use this sequence for (X, Y) to construct a corresponding promising partial sequence $\{(X'_k, Y'_k)\}_{k=1}^{2n-1}$ for (X', Y') where (X_k, Y_k) and (X'_k, Y'_k) contain the same vertices, though possibly in different orders. Thus, if we use the Boolean table P to construct edges E'_k on the vertices in $X'_k \cup Y'_k$, then $E'_k = E_k$. Since $(X', Y') \neq (X, Y)$, there must be some minimal m such that $(X_{m+1}, Y_{m+1}) \neq (X'_{m+1}, Y'_{m+1})$. Without loss of generality, we assume the refined vertex at this step was $u \in X_m$ and that `Refine` replaced u with u_1u_2 to obtain (X_{m+1}, Y_{m+1}), while (X'_{m+1}, Y'_{m+1}) requires replacing u with u_2u_1.

Consider the degree of u in the bipartite graph $(X_m, E_m, Y_m) = (X'_m, E_m, Y'_m)$. If $\deg(u) \geq 2$, then Lemma 3.1.4 implies that (X'_{m+1}, Y'_{m+1}) is not promising, which is not the case since $\{(X'_k, Y'_k)\}_{k=1}^{2n-1}$ is a promising partial sequence for (X', Y'). Thus, it must be that $\deg(u) = 1$. Since `ModifiedUntangle` always refines a vertex of highest degree, this implies that all vertices in $(X_m, E_m, Y_m) = (X'_m, E_m, Y'_m)$ must have degree 1. Then let $v \in Y_m$ be the unique neighbor of u in (X_m, E_m, Y_m). Once u is replaced with its children, notice that v will be the unique vertex in $(X_{m+1}, E_{m+1}, Y_{m+1})$ with $\deg(v) \geq 2$. Thus, after `ModifiedUntangle` replaces u with u_1u_2 to obtain (X_{m+1}, Y_{m+1}), it will replace v with its children in some order to obtain (X_{m+2}, Y_{m+2}).

Lemma 3.1.9 implies that $(u, v) \in \mathcal{L}$, so we let $(\widetilde{X}, \widetilde{Y}) \in \mathscr{P}(T, S, \phi)$ be (X, Y) after a paired flip at (u, v). Using $\{(X_k, Y_k)\}_{k=1}^{2n-1}$, we again use Lemma 3.1.11 to construct a promising partial sequence $\{(\widetilde{X}_k, \widetilde{Y}_k)\}_{k=1}^{2n-1}$ for $(\widetilde{X}, \widetilde{Y})$. By construction, $(\widetilde{X}_k, \widetilde{Y}_k) = (X_k, Y_k)$

for all $k \leq m$, and because of the paired flip at (u, v), we have $(\widetilde{X}_k, \widetilde{Y}_k) = (X'_k, Y'_k)$ for all $k \leq m+1$. Furthermore, the preceding paragraph implies that $(\widetilde{X}_{m+2}, \widetilde{Y}_{m+2})$ is obtained from $(\widetilde{X}_{m+1}, \widetilde{Y}_{m+1})$ by refining the vertex v. Since $\deg(v) = 2$ in $(\widetilde{X}_{m+1}, E_{m+1}, \widetilde{Y}_{m+1})$, Lemma 3.1.4 implies that a unique choice for the children of v results in a promising partial layout, and thus it must be that $(\widetilde{X}_{m+2}, \widetilde{Y}_{m+2}) = (X'_{m+2}, Y'_{m+2})$.

If $(\widetilde{X}, \widetilde{Y}) \neq (X', Y')$, then we can repeat the above argument. Eventually, this process terminates in a planar layout $(\widetilde{X}, \widetilde{Y})$ obtained from a sequence of paired flips starting at (X, Y), where $(\widetilde{X}_k, \widetilde{Y}_k) = (X'_k, Y'_k)$ for all k. Hence, we conclude $(X', Y') = (\widetilde{X}, \widetilde{Y})$, and any $(X', Y') \in \mathscr{P}(\mathcal{T})$ can be obtained using a sequence of paired flips starting with (X, Y). $\square$

We now define a graph on the planar layouts of a planar tanglegram. Theorem 3.1.12 can be restated in terms of connectedness for this graph.

Definition 3.1.13. Let $\mathcal{T}$ be a planar tanglegram. Define the *flip graph of* $\mathcal{T}$ as $G(\mathcal{T}) = (V, E)$ with vertices in V corresponding to planar layouts, and edges in E between vertices whenever the corresponding planar layouts that can be obtained from one another using paired flips at leaf-matched pairs of vertices.

Corollary 3.1.14. *The flip graph of a planar tanglegram is connected.*

3.1.2 The flip graph of a planar tanglegram

Since we consider planar layouts of a tanglegram $\mathcal{T}$ up to drawings of plane trees, the structure of $G(\mathcal{T})$ is not immediately clear. In the example shown in Figure 1.2, deleting certain edges results in a hypercube graph. We will show that this is always the case, and using this, we can determine the number of vertices in $G(\mathcal{T})$.

We start by defining *degenerate* and *non-degenerate* leaf-matched pairs. See Figure 3.5 for examples.

Definition 3.1.15. Let (u, v) be a leaf-matched pair of a planar tanglegram $\mathcal{T} = (L, R, \sigma)$. We call (u, v) a *degenerate* leaf-matched pair if $\mathcal{T}|_{(u,v)}$, the subtanglegram formed by the subtrees rooted at u and v, can be constructed by starting with the unique tanglegram of size 2 and replacing each pair of matched leaves with a copy of the same tanglegram $\mathcal{T}'$. Otherwise, (u, v) is a *non-degenerate* leaf-matched pair. We call a paired flip at (u, v) *degenerate* (resp. *non-degenerate*) if (u, v) is degenerate (resp. non-degenerate).

Definition 3.1.16. Let $\mathcal{T}$ be a planar tanglegram. The *reduced flip graph* of $\mathcal{T}$, denoted $G'(\mathcal{T})$, is the subgraph of $G(\mathcal{T})$ formed by including only the edges corresponding to non-degenerate paired flips.

Corollary 3.1.14 implies that paired flips allow us to generate all planar layouts of a planar tanglegram. We now show that this still holds when restricting to non-degenerate paired flips, so $G'(\mathcal{T})$ is also connected.

Lemma 3.1.17. *Let $\mathcal{T}$ be a planar tanglegram. Every planar layout of $\mathcal{T}$ can be obtained from any other planar layout of $\mathcal{T}$ using a combination of non-degenerate paired flips.*

Proof. It suffices to show each degenerate paired flip can be replaced by a sequence of non-degenerate paired flips, which we do using induction on the number m of degenerate leaf-matched pairs. The case of $m = 0$ is trivial. Now assume the result holds when there are less than m degenerate pairs, and consider a tanglegram $\mathcal{T}$ with m degenerate leaf-matched pairs.

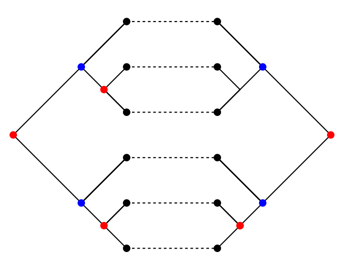


Figure 3.5: Vertices that form degenerate leaf-matched pairs are shown in red, while vertices that form non-degenerate leaf-matched pairs are shown in blue.

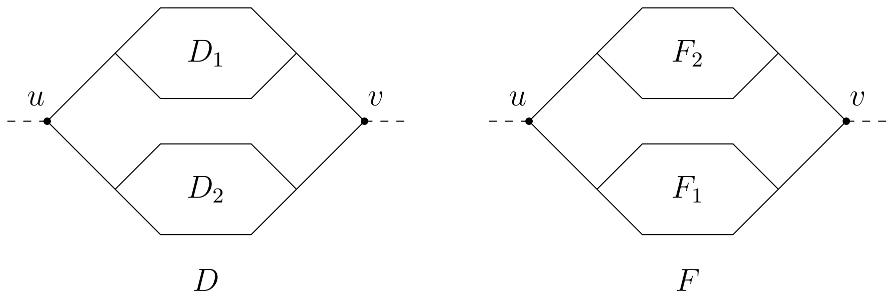


Figure 3.6: An illustration of the argument in Lemma 3.1.17.

We consider two planar layouts D and F for $\mathcal{T}$ such that F can be obtained from D by a paired flip at some degenerate leaf-matched pair (u, v). Let (u_1, v_1) and (u_2, v_2) be the leaf-matched pairs formed by the children of u and v with corresponding proper subtanglegrams $\mathcal{T}_1 \cong \mathcal{T}_2$. The layouts D and F restrict to planar layouts on these tanglegrams, which we can denote D_i and F_i for $i \in \{1, 2\}$, as in Figure 3.6.

Since $\mathcal{T}_1 \cong \mathcal{T}_2$, D_i and F_i for $i \in \{1, 2\}$ are all layouts for the same tanglegram. There are fewer than m degenerate leaf-matched pairs in $\mathcal{T}_1 \cong \mathcal{T}_2$, so the inductive hypothesis on these subtanglegrams implies all of the D_i's and F_i's can be obtained from one another using non-degenerate paired flips, allowing us to obtain F from D using non-degenerate paired flips. By induction on m, the result follows. □

We now show the uniqueness of the non-degenerate paired flips chosen. We will need the following definition.

Definition 3.1.18. The *irreducible component* of a tanglegram $\mathcal{T}$, denoted $\mathrm{Irr}(\mathcal{T})$, is the irreducible tanglegram formed by contracting each non-root leaf-matched pair of $\mathcal{T}$ to a single pair of matched leaves. When $\mathcal{T}' = \mathrm{Irr}(\mathcal{T})$, we say $\mathcal{T}'$ *extends to* $\mathcal{T}$.

Lemma 3.1.19. *The combination of paired flips in Lemma 3.1.17 is unique.*

Proof. The result is clear for the unique size 1 and unique size 2 tanglegrams, which have only one layout. For tanglegrams of size 3 or more, we use induction on the total number of leaf-matched pairs ℓ. The base case of $\ell = 1$ leaf-matched pair corresponds to irreducible tanglegrams, and the result, in this case, follows from Proposition 2.1.8.

Now assume the result holds when there are less than ℓ leaf-matched pairs in a planar tanglegram, and suppose $\mathcal{T} = (L, R, \sigma)$ has ℓ leaf-matched pairs. Let $\mathcal{N}$ denote the set of non-degenerate leaf-matched pairs of $\mathcal{T}$, and suppose that starting with a layout D and performing flips at $S_1, S_2 \subseteq \mathcal{N}$ produces the same layout D'. In particular, these layouts D and D' restrict to the same layout of $\text{Irr}(\mathcal{T})$. If $\text{Irr}(\mathcal{T})$ has size at least 3, then Proposition 2.1.8 implies S_1 and S_2 either both contain or don't contain a paired flip at the roots of L and R. Applying the inductive hypothesis on each of the maximal proper subtanglegrams of $\mathcal{T}$ then implies the remaining elements of S_1 and S_2 must coincide.

Alternatively, if $\text{Irr}(\mathcal{T})$ has size 2, then we have two cases involving the two maximal proper subtanglegrams $\mathcal{T}_1$ and $\mathcal{T}_2$. If $\mathcal{T}_1 \ncong \mathcal{T}_2$, then S_1 and S_2 either both contain or don't contain a paired flip at the roots of L and R. We then again apply the induction hypothesis on $\mathcal{T}_1$ and $\mathcal{T}_2$ to conclude the remaining elements of S_1 and S_2 are the same. Otherwise, $\mathcal{T}_1 \cong \mathcal{T}_2$, so the roots of L and R form a degenerate leaf-matched pair, and hence are not in $\mathcal{N}$ nor in its subsets S_1 and S_2. Again, we apply the induction hypothesis to $\mathcal{T}_1$ and $\mathcal{T}_2$ to conclude $S_1 = S_2$. The result now follows by induction. □

We now establish a graph isomorphism between $G'(T)$ and the hypercube graph of the appropriate size. Recall that the hypercube graphs are described in Example 2.1.1.

Theorem 3.1.20. *Let $\mathcal{T}$ be a planar tanglegram with ℓ' non-degenerate leaf-matched pairs. Then $G'(\mathcal{T})$ is isomorphic to the ℓ'-hypercube graph.*

Proof. We produce a graph isomorphism between $G'(\mathcal{T})$ and the ℓ'-hypercube $Q = Q_{\ell'}$, where ℓ' is the number of non-degenerate leaf-matched pairs in $\mathcal{T}$. We denote the vertices in Q as ℓ'-tuples with elements in $\{0, 1\}$, where edges occur between vertices that differ in exactly one coordinate. Arbitrarily order the non-degenerate leaf-matched pairs of $\mathcal{T}$ as $\mathcal{N} = \{(u_i, v_i)\}_{i=1}^{\ell'}$. By the preceding lemma, after fixing an arbitrary initial layout D, each

$S \subseteq \mathcal{N}$ corresponds to a unique planar layout D_S of $\mathcal{T}$. Hence, we obtain a bijection

$$D_S \quad \longleftrightarrow \quad \left(1_{\{(u_1,v_1)\in S\}}, \ldots, 1_{\{(u_{\ell'},v_{\ell'})\in S\}}\right),$$

where $1_{\{(u_i,v_i)\in S\}}$ is the indicator function for the pair (u_i, v_i) being in S. This defines a bijection between the vertices of $G'(\mathcal{T})$ and Q. Two layouts D_{S_1}, D_{S_2} are adjacent in $G'(\mathcal{T})$ if and only if the corresponding sets S_1, S_2 differ by a single element, and hence if and only if the corresponding vertices in Q differ in a single coordinate. Consequently, this map between the vertices of $G'(\mathcal{T})$ and Q preserves all edges, implying that it is a graph isomorphism. □

Corollary 3.1.21. *Let $\mathcal{T}$ be a planar tanglegram with ℓ' non-degenerate leaf-matched pairs. Then $\mathcal{T}$ has exactly $2^{\ell'}$ planar layouts which can all be obtained from one another using paired flips at non-degenerate leaf-matched pairs.*

Remark 3.1.22. For a planar tanglegram of size n, a planar layout and the set of leaf-matched pairs can be identified in $O(n^2)$ time using `ModifiedUntangle`, and there are at most $n-1$ leaf-matched pairs. To identify which pairs are non-degenerate, we can consider the subtanglegram $\mathcal{T}|_{(u,v)}$ for each leaf-matched pair (u, v), check if its irreducible component has size 2, and test isomorphism of the two maximal proper subtanglegrams when this is the case.

Since an isomorphism between tanglegrams $\mathcal{T} = (L, R, \sigma)$ and $\mathcal{T}' = (L', R', \sigma')$ can be viewed as an isomorphism of graphs that maps the root of L to the root of L' and the root of R to the root of R', we can adapt isomorphism testing for planar graphs to planar tanglegrams. For example, one can test the isomorphism of the planar graphs obtained by attaching a single leaf to $\text{root}(L)$ and $\text{root}(L')$. Note that isomorphism testing of planar graphs can be done asymptotically in linear time with respect to the number of vertices [HW74], though faster algorithms exist for graphs that are not very large [KHC04].

Combined, we conclude that identifying the non-degenerate leaf-matched pairs of a tanglegram can be completed in $O(n^2)$ time. Consequently, one can efficiently compute the number of planar layouts of a planar tanglegram $\mathcal{T}$.

We conclude this section by analyzing the extremal cases in Corollary 3.1.21 for the number of planar layouts of a planar tanglegram. We start by defining several families of tanglegrams that will appear. Examples are shown in Figure 3.7.

Definition 3.1.23. For $h \geq 0$, let B_h denote the *complete binary tree* of height h, which is the tree with 2^h leaves and the maximal number of leaves at each level. Define $\mathcal{B}_h$ to be the tanglegram formed from two copies of B_h with matching induced by an automorphism of B_h.

Definition 3.1.24. For $n \geq 1$, let C_n denote the caterpillar with n leaves, that is, the tree with n leaves whose internal vertices form a path on $n - 1$ vertices. Define $\mathcal{C}_n$ to be the tanglegram formed from two copies of C_n with matching induced by an automorphism of C_n.

Definition 3.1.25. Define $\mathcal{C}'_3$ to be the planar tanglegram on 3 leaves that is not $\mathcal{C}_3$. For $n \geq 4$, define $\mathcal{C}'_n$ to be the tanglegram formed by starting with $\mathcal{C}_{n-2}$ and replacing a pair of matched leaves of maximal distance from the roots with $\mathcal{C}'_3$.

We now consider the minimum possible number of planar layouts. Proposition 2.1.8 and Theorem 2.1.10 show that for $n \geq 3$, there exist tanglegrams with exactly 2 planar layouts. The $\mathcal{B}_h$ tanglegrams defined above are precisely the ones that achieve 1 planar layout.

Lemma 3.1.26. *Let $\mathcal{T}$ be a planar tanglegram. Then $|G(\mathcal{T})| = 1$ if and only if $\mathcal{T} = \mathcal{B}_h$ for some h.*

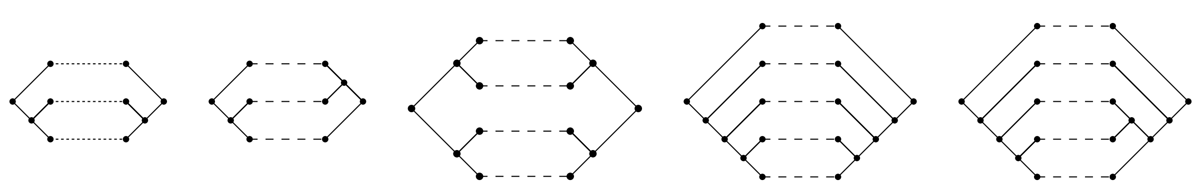


Figure 3.7: The tanglegrams $\mathcal{C}_3$, $\mathcal{C}'_3$, $\mathcal{B}_2$, $\mathcal{C}_5$, and $\mathcal{C}'_5$.

Proof. Corollary 3.1.21 implies that $\mathcal{B}_h$ has 1 planar layout, as all leaf-matched pairs are degenerate. We use induction on size to show this is the only possibility. Observe that if $\mathcal{T}$ has size 1 or 2, then the result is clear since $\mathcal{B}_0$ and $\mathcal{B}_1$ are the unique tanglegrams of their respective sizes.

Now assume the result holds for all tanglegrams of size $k < n$, and assume $\mathcal{T} = (L, R, \sigma)$ has size n and exactly one planar layout. Then the roots of L and R must form a degenerate leaf-matched pair, implying $\mathcal{T}$ is constructed by starting with $\mathcal{B}_1$ and replacing the matched leaves with some tanglegram $\mathcal{T}'$. Since $\mathcal{T}$ has a unique layout, the same holds for the subtanglegram $\mathcal{T}'$. Using the induction hypothesis, we conclude $\mathcal{T}' \cong \mathcal{B}_h$ for some h. Then $\mathcal{T} \cong \mathcal{B}_{h+1}$, and the result follows by induction. □

We now consider tanglegrams with many planar layouts. The following result establishes an upper bound for the number of planar layouts and gives some necessary conditions for achieving this bound.

Lemma 3.1.27. *Let $\mathcal{T} = (L, R, \sigma)$ be a planar tanglegram of size $n \geq 2$. Then $|G(\mathcal{T})| \leq 2^{n-2}$. When $n \geq 4$, each of the following conditions implies that $|G(\mathcal{T})| \leq 2^{n-3}$:*

(a) $\mathrm{Irr}(\mathcal{T})$ *has size at least* 3, *or*

(b) $\mathrm{Irr}(\mathcal{T}) = \mathcal{B}_1$ *has size* 2 *and $\mathcal{T}$ is formed by replacing the matched leaves in $\mathcal{B}_1$ with two tanglegrams of size larger than* 1.

Proof. Observe that the bound $|G(\mathcal{T})| \leq 2^{n-1}$ follows from Corollary 3.1.21 and the fact that L and R each have $n-1$ internal vertices. To improve this upper bound, consider a tanglegram $\mathcal{T} = (L, R, \sigma)$ of size $n \geq 2$ with $n-1$ leaf-matched pairs. Notice that L must have an internal vertex u that has only two children, and this vertex forms a leaf-matched pair (u, v), which must be degenerate. Hence, a planar tanglegram has at most $n-2$ non-degenerate leaf-matched pairs, so the improved upper bound of 2^{n-2} follows.

We now consider the cases above for the 2^{n-3} bound. As before, we rule out vertices that cannot form non-degenerate leaf-matched pairs. For (a), first note that if $\mathrm{Irr}(\mathcal{T})$ has size at least 4, then at least two of the $n-1$ internal vertices in L and R cannot be leaf-matched pairs, so $|G(\mathcal{T})| \leq 2^{n-3}$ follows. If $\mathrm{Irr}(\mathcal{T})$ has size 3, then at least one internal vertex in each of the component trees of $\mathrm{Irr}(\mathcal{T})$ cannot form a leaf-matched pair. Since $\mathcal{T}$ has size at least 4, one of its proper subtanglegrams of size $k \geq 2$ has at most $k-2$ non-degenerate pairs, and hence $\mathcal{T}$ has at most $n-3$ non-degenerate pairs, again implying $|G(\mathcal{T})| \leq 2^{n-3}$. For (b), the two maximal subtanglegrams of size $k \geq 2$ and $n-k \geq 2$ respectively have at most $k-2$ and $n-k-2$ non-degenerate pairs, so $\mathcal{T}$ has at most $1+(k-2)+(n-k-2) = n-3$ non-degenerate pairs, and we again conclude $|G(\mathcal{T})| \leq 2^{n-3}$. □

We conclude with our characterization for the extremal cases of $|G(\mathcal{T})|$.

Theorem 3.1.28. *Let $\mathcal{T}$ be a planar tanglegram of size $n \geq 2$. Then $1 \leq |G(\mathcal{T})| \leq 2^{n-2}$. Furthermore,*

(a) $|G(\mathcal{T})| = 1$ if and only if $\mathcal{T} = \mathcal{B}_{\log_2(n)}$ with $\log_2(n) \in \mathbb{Z}_+ \cup \{0\}$, and

(b) $|G(\mathcal{T})| = 2^{n-2}$ if and only if $\mathcal{T} = \mathcal{C}_n$ or $\mathcal{T} = \mathcal{C}'_n$.

Proof. The lower bound of 1 follows from the definition of a planar tanglegram and the upper bound of 2^{n-2} follows by Lemma 3.1.27. Equality in the lower bound follows from Lemma 3.1.26. For equality in the upper bound, Corollary 3.1.21 shows that $\mathcal{C}_n$ and $\mathcal{C}'_n$ achieve this bound. To show that they are the only tanglegrams, we first note that the results for tanglegrams of size 2 or 3 are clear, as the only tanglegrams of these sizes are $\mathcal{B}_1 = \mathcal{C}_2$, $\mathcal{C}_3$, and $\mathcal{C}'_3$. For tanglegrams of larger size, we use induction with base case $n = 3$.

Assume the result holds for tanglegrams of size $k < n$, and suppose $\mathcal{T}$ is a tanglegram of size $n \geq 4$ with 2^{n-2} distinct planar layouts. Then Lemma 3.1.27 implies that $\mathcal{T}$ is formed by starting with $\mathcal{B}_1$ and replacing a single pair of matched leaves with some tanglegram $\mathcal{T}'$ of size $n-1$. A planar layout of $\mathcal{T}$ is determined by whether $\mathcal{T}'$ is drawn on top or bottom, and then a layout of $\mathcal{T}'$ itself. Hence, $\mathcal{T}'$ must have 2^{n-3} layouts, and by the induction hypothesis, is either $\mathcal{C}_{n-1}$ or $\mathcal{C}'_{n-1}$. Then $\mathcal{T}$ must be $\mathcal{C}_n$ or $\mathcal{C}'_n$, and the result follows. □

3.2 Enumerative results

In this section, we give two enumerative results involving planar tanglegrams, both of which will generalize the relation between $H(x)$ and $T(x)$ given in Theorem 2.1.10. Let $\mathcal{P}$ denote the set of planar tanglegrams. Our first result will involve the following generating function that generalizes $T(x)$ by also considering the number of leaf-matched pairs:

$$F(x,q) = \sum_{\mathcal{T}\in\mathcal{P}} x^{|\mathcal{T}|} q^{|\{\text{leaf-matched pairs of } \mathcal{T}\}|}. \tag{3.2.1}$$

Theorem 3.2.1. *The generating function $F(x,q)$ satisfies the relation*

$$F(x,q) = x + q\cdot H(F(x,q)) + \frac{q\cdot F(x^2,q^2)}{2}. \tag{3.2.2}$$

Proof. Equation (3.2.2) is equivalent to

$$F(x,q) = x + q\cdot\left(H(F(x,q)) - \frac{F(x,q)^2}{2}\right) + q\cdot\frac{F(x,q)^2 + F(x^2,q^2)}{2}, \tag{3.2.3}$$

so we establish this relation instead. The term x accounts for the unique tanglegram of size 1, which has no leaf-matched pairs. For the remaining tanglegrams, we can form each tanglegram $\mathcal{T}$ by starting with its irreducible component $\mathrm{Irr}(\mathcal{T})$ and replacing matched leaves with planar tanglegrams (possibly of size 1). The remaining summands correspond to tanglegrams where $\mathrm{Irr}(\mathcal{T})$ has size larger than two or equal to two.

To interpret the first summand, observe that $q\cdot[H(x) - x^2/2]$ counts irreducible planar tanglegrams of size $n \geq 3$ by size and number of leaf-matched pairs. Each term qx^k corresponds to an irreducible planar tanglegram $\mathcal{T}$ of size k. Proposition 2.1.8 implies that irreducible tanglegrams of size at least 3 have no symmetry. Consequently, fixing a layout of $\mathcal{T}$ and replacing matched leaves from top to bottom with planar tanglegrams $(\mathcal{T}_i)_{i=1}^k$ produces a distinct tanglegram for each selection of $(\mathcal{T}_i)_{i=1}^k$. Replacing the k matched leaves of $\mathcal{T}$ with k planar tanglegrams corresponds to replacing x in qx^k with $F(x,q)$. Hence,

$q \cdot [H(F(x,q)) - F(x,q)^2/2]$ enumerates tanglegrams with irreducible component of size $n \geq 3$.

For the second summand, observe that tanglegrams with irreducible components of size 2 are formed by starting with the unique planar tanglegram of size two corresponding to the term qx^2 and replacing the two pairs of leaves with two planar tanglegrams $\mathcal{T}_1$ and $\mathcal{T}_2$, where the order is not relevant. The generating function $q \cdot F(x,q)^2$ would count ordered pairs of planar tanglegrams. This correctly counts the case when $\mathcal{T}_1 \cong \mathcal{T}_2$ but double-counts the remaining cases. To remedy this over-counting, we can add $q \cdot F(x^2, q^2)$, which counts the pairs where $\mathcal{T}_1 \cong \mathcal{T}_2$, and then divide the result by two to account for the order not being relevant. Hence, $q \cdot \frac{F(x,q)^2 + F(x^2,q^2)}{2}$ enumerates tanglegrams with irreducible component of size two. Combined, we obtain (3.2.3). $\square$

Note that substituting $q = 1$ results in the original relation given in Theorem 2.1.10. Using Theorem 3.2.1 and the coefficients of $H(x)$ from [OEI24, A257887], one can generate the coefficients of $F(x,q)$. We collect some of these coefficients in Table 3.1. See [OEI24, A349409] for more terms.

We can similarly generalize $T(x)$ by also considering size of irreducible component. Let

n, k	1	2	3	4	5	6	7	8	total
2	1								1
3	1	1							2
4	5	4	2						11
5	34	28	11	3					76
6	273	239	102	29	6				649
7	2436	2283	1045	325	73	11			6173
8	23391	23475	11539	3852	968	181	23		63429
9	237090	254309	133690	47640	12923	2756	444	46	688898

Table 3.1: The number of tanglegrams of size n with k leaf-matched pairs.

$\mathcal{P}$ and $\mathcal{I}$ respectively denote the set of planar and irreducible planar tanglegrams, and define

$$\begin{aligned} T(x,y) &= \sum_{\mathcal{T}\in\mathcal{P}} x^{|\mathcal{T}|} y^{|\operatorname{Irr}(\mathcal{T})|}, \\ H(x,y) = H(xy) &= \frac{1}{2}x^2y^2 + \sum_{\mathcal{T}\in\mathcal{I}:\ |\mathcal{T}|>2} x^{|\mathcal{T}|}y^{|\mathcal{T}|}. \end{aligned} \tag{3.2.4}$$

The following result also generalizes Theorem 2.1.10, and substituting $y = 1$ similarly recovers the original result.

Theorem 3.2.2. *The following holds:*

$$T(x,y) = H(T(x),y) + \frac{T(x^2)y^2}{2} + xy. \tag{3.2.5}$$

Proof. As in Theorem 3.2.1, we first rewrite the right-hand side of Equation (3.2.5) as

$$\left(H(T(x),y) - \frac{T(x)^2y^2}{2}\right) + \left(\frac{T(x)^2y^2 + T(x^2)y^2}{2}\right) + xy. \tag{3.2.6}$$

The third summand xy accounts for the unique tanglegram where both trees consist of a single vertex. For the remaining summands, we consider those with an irreducible component of size two and greater than two separately.

To interpret the first summand, observe that $H(x,y) - \frac{x^2y^2}{2}$ counts irreducible tanglegrams of size at least 3, where each term x^ky^k corresponds to an irreducible planar tanglegram of size k. As in Theorem 3.2.1, for each irreducible tanglegram $\mathcal{T}$ of size $k > 2$, we can fix a layout of $\mathcal{T}$ and replace matched leaves from top-to-bottom with planar tanglegrams $(\mathcal{T}_i)_{i=1}^k$, which produces a distinct tanglegram for each selection of $(\mathcal{T}_i)_{i=1}^k$. Replacing pairs of matched leaves with planar tanglegrams corresponds to replacing x with $T(x)$. Hence, $H(T(x),y) - \frac{T(x)^2y^2}{2}$ is the generating function for planar tanglegrams with $|\operatorname{Irr}(\mathcal{T})| \geq 3$.

To interpret the second summand, tanglegrams with $|\operatorname{Irr}(\mathcal{T})| = 2$, we must start with the unique layout for the unique tanglegram of size two and replace matched leaves with $\mathcal{T}_1$ and $\mathcal{T}_2$. While $T(x)^2y^2$ would correctly count the cases when $\mathcal{T}_1 = \mathcal{T}_2$, it double-counts the other cases. As in Theorem 3.2.1, we add $T(x^2)y^2$ so that all cases are double-counted, and dividing by 2 then results in the second summand enumerating planar tanglegrams with irreducible components of size 2. □

Note that once we can efficiently compute the coefficients of $T(x)$ and $H(x)$, we can efficiently perform the composition to generate the coefficients of $T(x, y)$. Using known values of $T(x)$ and $H(x)$ from [OEI24, A257887 and A349408], we give some coefficients of $T(x, y)$ in Table 3.2. Additional terms can be found at [OEI24, A371659].

We respectively use the notation t_n and $t_{n,k}$ for the coefficient of x^n in $T(x)$ and $x^n y^k$ in $T(x, y)$. We also use the notation h_n for $t_{n,n}$, which is the number of irreducible planar tanglegrams of size n. Recall that a *composition* of n is a decomposition of n into an ordered sum of positive integers, and we use the notation $(a_i)_{i=1}^k \models n$ to denote this. The preceding theorem and proof imply the following corollary involving the coefficients $t_{n,k}$.

Corollary 3.2.3. *Let $2 \leq k \leq n$. The number of ways to extend a size k irreducible planar tanglegram into a size n planar tanglegram, denoted $c_{n,k}$, is independent of the irreducible planar tanglegram. Moreover,*

(a) if n is even, then $t_{n,2} = \frac{1}{2}\left(\sum_{i=1}^{n-1} t_i t_{n-i} + t_{n/2}\right)$,

(b) if n is odd, then $t_{n,2} = \frac{1}{2}\sum_{i=1}^{n-1} t_i t_{n-i}$, and

n, k	2	3	4	5	6	7	8	9	total
2	1								1
3	1	1							2
4	3	3	5						11
5	13	9	20	34					76
6	90	46	70	170	273				649
7	747	312	360	680	1638	2436			6173
8	7040	2580	2435	3570	7371	17052	23391		63429
9	71736	24056	19800	23970	39858	85260	187128	237090	688898

Table 3.2: The number of tanglegrams of size n with irreducible component size k.

(c) if $k \neq 2$*, then* $t_{n,k} = h_k \cdot \sum_{(a_i)_{i=1}^k \models n} t_{a_1} t_{a_2} \dots t_{a_k}$.

Additionally, note that $c_{n,k} = t_{n,k}/h_k$.

3.3 Sampling planar tanglegrams

In this section, we give an algorithm for sampling planar tanglegrams. We start by applying the $c_{n,k}$ constants from the previous section to reduce the problem of generating planar tanglegrams to generating irreducible planar tanglegrams or their layouts. Note that in the following theorem, we assume an algorithm for uniformly sampling irreducible planar tanglegram layouts.

Theorem 3.3.1. *The following procedure generates a planar tanglegram of size* $n \geq 3$ *uniformly at random.*

1. *Choose an integer* $2 \leq k \leq n$ *with probability* $\frac{h_k c_{n,k}}{t_n}$ *and generate a layout* D *for an irreducible planar tanglegram of size* k *uniformly at random.*

2. (a) *If* $k \neq 2$*, select* $(a_i)_{i=1}^k \models n$ *with probability* $\frac{t_{a_1} t_{a_2} \dots t_{a_k}}{c_{n,k}}$ *and independently generate planar tanglegrams* $(\mathcal{T}_i)_{i=1}^k$ *of sizes* $(a_i)_{i=1}^k$ *uniformly at random.*

   (b) *If* n *is odd and* $k = 2$*, select* $(a_1, a_2) \models n$ *with probability* $\frac{t_{a_1} t_{a_2}}{2c_{n,2}}$ *and independently generate planar tanglegrams* $(\mathcal{T}_1, \mathcal{T}_2)$ *of sizes* (a_1, a_2) *uniformly at random.*

   (c) *If* n *is even and* $k = 2$,

   - *with probability* $\frac{t_{n/2}}{2c_{n,2}}$*, generate a single tanglegram* $\mathcal{T}_1 = \mathcal{T}_2$ *of size* $n/2$ *uniformly at random, and*

   - *otherwise, select* $(a_1, a_2) \models n$ *with probability* $\frac{t_{a_1} t_{a_2}}{\sum_{i=1}^{n-1} t_i t_{n-i}}$ *and independently generate planar tanglegrams* $(\mathcal{T}_1, \mathcal{T}_2)$ *of sizes* (a_1, a_2) *uniformly at random.*

3. *In all cases, output the tanglegram corresponding to* D *with matched leaves replaced from top to bottom by* $\{\mathcal{T}_i\}_{i=1}^k$.

Proof. The definition of $c_{n,k}$ and the results of Corollary 3.2.3 imply that all of the necessary quantities in steps (1) and (2) sum to 1. We show that each tanglegram of size n can be generated in two ways, and each of these possibilities has probability $\frac{1}{2t_n}$.

Consider a planar tanglegram $\mathcal{T}$ with $|\operatorname{Irr}(\mathcal{T})| \geq 3$. To generate $\mathcal{T}$ in the algorithm, we must first generate one of the layouts of $\operatorname{Irr}(\mathcal{T})$ in step (1). Proposition 2.1.8 implies that there are two possibilities D_1 and D_2, and observe that each of them has probability $\frac{h_k c_{n,k}}{t_n} \cdot \frac{1}{2h_k} = \frac{c_{n,k}}{2t_n}$ of being generated. For each D_i, a unique list of tanglegrams $(\mathcal{T}_{i,j})_{j=1}^k$ must replace the matched leaves in D_i from top-to-bottom to construct $\mathcal{T}$. Letting $a_{i,j} = |\mathcal{T}_{i,j}|$, the probability $(\mathcal{T}_{i,j})_{j=1}^k$ is generated in step (2) is given by $\frac{t_{a_1} t_{a_2} \cdots t_{a_k}}{c_{n,k}} \cdot \frac{1}{t_{a_1} t_{a_2} \cdots, t_{a_k}} = \frac{1}{c_{n,k}}$. Hence, each of the two ways of generating $\mathcal{T}$ has probability $\frac{c_{n,k}}{2t_n} \cdot \frac{1}{c_{n,k}} = \frac{1}{2t_n}$ of being generated, so the probability $\mathcal{T}$ is generated is $\frac{1}{t_n}$.

Next, consider a planar tanglegram $\mathcal{T}$ with $|\operatorname{Irr}(\mathcal{T})| = 2$. This requires first generating the unique layout D for the unique planar tanglegram of size 2. If $n = |\mathcal{T}|$ is odd, then two possibilities (T_1, T_2) and (T_2, T_1) extend D to $\mathcal{T}$. Letting $a_1 = |\mathcal{T}_1|$ and $a_2 = |\mathcal{T}_2|$, the probability of obtaining $\mathcal{T}$ is given by

$$\frac{h_2 c_{n,2}}{t_n} \cdot \left(\frac{t_{a_1} t_{a_2}}{2c_{n,2}} \cdot \frac{1}{t_{a_1} t_{a_2}} + \frac{t_{a_2} t_{a_1}}{2c_{n,2}} \cdot \frac{1}{t_{a_2} t_{a_1}} \right) = \frac{1}{t_n}.$$

Note that $h_2 = 1$, so this term disappears in the product above.

Now consider when $n = |\mathcal{T}|$ is even. Suppose $\mathcal{T}$ requires replacing the matched leaves in $\mathcal{L}$ with the same tanglegram $\mathcal{T}'$. With probability $\frac{t_{n/2}}{2c_{n,2}} \cdot \frac{1}{t_{n/2}}$, we generate $\mathcal{T}'$ twice in the first case of (2c), so $\mathcal{T}$ has a

$$\frac{h_2 c_{n,2}}{t_n} \cdot \frac{t_{n/2}}{2c_{n,2}} \cdot \frac{1}{t_{n/2}} = \frac{1}{2t_n}$$

probability of being generated this way. The probability of generating $\mathcal{T}$ by generating $\mathcal{T}'$ twice in the second case of (2c) is

$$\frac{h_2 c_{n,2}}{t_n} \cdot \left(1 - \frac{t_{n/2}}{2c_{n,2}} \cdot \right) \cdot \frac{t_{n/2} t_{n/2}}{\sum_{i=1}^{n-1} t_i t_{n-i}} \cdot \frac{1}{t_{n/2} t_{n/2}}.$$

Using Corollary 3.2.3, $h_2 = 1$, and $c_{n,2} = t_{n,2}$, this simplifies to

$$\frac{c_{n,2}}{t_n} \cdot \frac{\sum_{i=1}^{n-1} t_i t_{n-i}}{2c_{n,2}} \cdot \frac{1}{\sum_{i=1}^{n-1} t_i t_{n-i}} = \frac{1}{2t_n}.$$

The case when $\mathcal{T}$ requires replacing matched leaves in $\mathcal{L}$ with two distinct tanglegrams is done using the same properties, where we note that there is no way to generate $\mathcal{T}$ in the first case of (2c) but two ways to generate it in the second case. □

The procedure in Theorem 3.3.1 can be applied recursively when generating $\{\mathcal{T}_i\}_{i=1}^k$, except that the tanglegrams of size 1 and 2 should be generated directly since they are unique. For efficiency reasons, one can also directly generate all planar tanglegrams below a certain size. From this, we conclude that uniform sampling of planar tanglegrams can be reduced to computation of the coefficients in $T(x, y)$ and uniform sampling of irreducible planar tanglegram layouts.

We now consider the problem of uniformly sampling pairs of disjoint triangulations, which is equivalent to uniformly sampling irreducible planar tanglegram layouts by Theorem 2.1.9. Throughout, fix a labeling of the convex n-gon using $[n] = \{1, 2, \ldots, n\}$, and use pairs (a, b) with $a, b \in [n]$ to denote diagonals. For a triangulation T, we use the notation $(a, b) \in T$ to denote that the diagonal (a, b) is in the triangulation T.

Recall from Example 2.1.2 that given a triangulation T and a diagonal $(a, b) \in T$, a *flip* replaces (a, b) with the other diagonal in the unique quadrilateral containing (a, b). We extend the flip operation to pairs of disjoint triangulations. An example is shown in Figure 3.8.

Definition 3.3.2. Let (T_1, T_2) be an ordered pair of disjoint triangulations of an n-gon, and suppose $(a, b) \in T_i$ for some $i \in [2]$. A *flip* at $(a, b)_i \in (T_1, T_2)$ is defined as

(a) flip $(a, b) \in T_i$, and

(b) if the resulting diagonal (a', b') is in T_j for $j \neq i$, then flip $(a', b') \in T_j$.

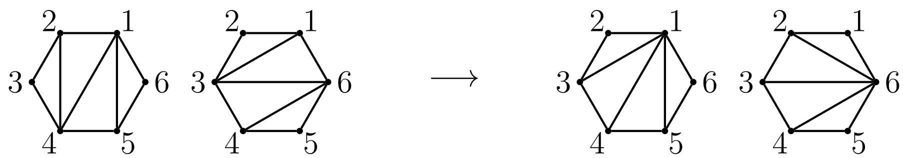


Figure 3.8: A (double) flip at $(2,4)_1$.

When only (a) is performed, we refer to this as a *single flip*, and when both (a) and (b) are performed, we refer to this as a *double flip*.

Lemma 3.3.3. *Let (T_1, T_2) be an ordered pair of disjoint triangulations of an n-gon. If (T_1', T_2') is obtained from (T_1, T_2) by a flip at $(a,b)_i$, then (T_1', T_2') is also a pair of disjoint triangulations. Furthermore, (T_1, T_2) can also be obtained from (T_1', T_2') by some flip.*

Proof. Without loss of generality, assume we flip the edge $(a,b) \in T_1$ to obtain T_1'. If after the flip, the new diagonal (a', b') does not coincide with any diagonals of T_2, then we are done. Note that in this case, we have that $T_2 = T_2'$, and the flip $(a', b')_1$ allows us to obtain (T_1, T_2) from (T_1', T_2').

Otherwise, we flip (a', b') in the second polygon to obtain T_2'. The resulting diagonal (a'', b'') crosses (a', b'), and hence cannot appear in T_1'. Hence, (T_1', T_2') is a pair of disjoint triangulations. In this case, observe that the flip $(a'', b'')_2$ allows us to obtain (T_1, T_2) from (T_1', T_2'). □

The mutual reachability between two pairs of disjoint triangulations allows us to now formally define the *flip graph on pairs of disjoint triangulations*. An example was previously shown in Figure 1.3.

Definition 3.3.4. Let $\mathcal{D}_n$ denote the (undirected) graph with vertices corresponding to ordered pairs of disjoint triangulations of an n-gon and edges corresponding to flips as defined in Definition 3.3.2.

Observe that $\mathcal{D}_3$ is a single vertex graph and $\mathcal{D}_4$ is a path graph on two vertices. Hence, we are primarily interested in the cases $n \geq 5$. For any pair $(T_1, T_2) \in \mathcal{D}_n$, there are $2(n-3)$ diagonals on which a flip can be performed. It is not difficult to see that if $n \geq 5$, then flipping at different diagonals results in different pairs of disjoint triangulations.

Lemma 3.3.5. *For any integer $n \geq 5$, the flip graph $\mathcal{D}_n$ is simple and $2(n-3)$-regular.*

Proof. It suffices to show that for any pair (T_1, T_2) of disjoint triangulations, a flip at $(a, b) \in T_1$ and a flip at $(c, d) \in T_2$ cannot result in the same pair (T_1', T_2'). Let $(a', b') \in T_1'$ and $(c', d') \in T_2'$ be the diagonals obtained by flipping $(a, b) \in T_1$ and $(c, d) \in T_2$ respectively. If (T_1', T_2') is obtained from (T_1, T_2) by flipping $(a, b) \in T_1$, then $(a', b') = (c, d)$, and if (T_1', T_2') is obtained from (T_1, T_2) by flipping $(c, d) \in T_2$, then $(c', d') = (a, b)$. This implies that (a, c, b, d) is a quadrilateral in both T_1 and T_2. Since $n \geq 5$, this implies that T_1 and T_2 share a diagonal, which is a contradiction. □

Let G_n denote the flip graph for triangulations of an n-gon, as described in Example 2.1.2. For any triangulation S of an n-gon, the subgraph of G_n induced by the set of triangulations disjoint from S is denoted $G_n(S)$. Pournin showed in [Pou14] that the diameter of G_n is $2n - 10$ for $n > 12$. We show the connectedness of $\mathcal{D}_n$ and a linear diameter bound for $\mathcal{D}_n$ by first showing corresponding statements for $G_n(S)$.

Theorem 3.3.6. *Let $n \geq 5$, and let S be a triangulation of the n-gon. Then $G_n(S)$ is connected, and its diameter is at most $2n - 8$.*

Proof. Every triangulation contains a diagonal of the form $(i, i{+}2)$, so we assume without loss of generality that S contains the diagonal $(2, n)$. Let Δ denote the triangulation consisting of $\{(1, i) : 3 \leq i \leq n{-}1\}$, which is called a *standard triangulation* in Chapter 1 of [DLRS10]. We show that every triangulation T disjoint from S is connected to Δ by a path in $G_n(S)$ of length at most $n{-}4$. In fact, we claim that if T contains d edges of the form $\{(1, i) : 3 \leq i \leq n{-}1\}$, then it is connected to Δ by a path in $G_n(S)$ of length at most $n - 3 - d$. Note that $d \geq 1$, as $(2, n) \notin T$ implies the existence of some diagonal of the form $(1, i)$.

We prove the claim by induction on $n-3-d$. If $d=n-3$, then $T=\Delta$, and these triangulations are connected by a path of length $n-3-(n-3)=0$ as needed. Now suppose $d<n-3$, and let $3 \leq i_1 < i_2 < \ldots < i_d \leq n-1$ denote the indices such that $(1, i_j) \in T$ for all i_j. Since $d<n-3$, it must be that $i_{j+1} - i_j > 1$ for some $j \in \{1, \ldots, d\}$. Consider the polygon with vertices $\{1, i_j, i_j+1, i_j+2, \ldots, i_{j+1}\}$ with triangulation T_1 induced by T. Observe that T_1 cannot contain any edges of the form $(1,i)$, and since $i_j, 1, i_{j+1}$ are consecutive vertices, this can only occur if $(i_j, i_{j+1}) \in T_1$. Flipping this diagonal results in $(1, i')$ for some $i_j < i' < i_{j+1}$. Note that this diagonal $(1, i')$ cannot appear in S since S contains $(2, n)$. Hence, flipping (i_j, i_{j+1}) in T results in some triangulation T' disjoint from S that contains $d+1$ diagonals of the form $\{(1,i) : 3 \leq i \leq n-1\}$. By the inductive hypothesis, T' and Δ are connected by a path in $G_n(S)$ of length at most $n-3-d-1$, and hence T is connected to Δ by a path in $G_n(S)$ of length at most $n-3-d$.

Now let T_1 and T_2 be any two triangulations disjoint from S. Choosing Δ as above, each T_i is connected to Δ by a path of length at most $n-4$ in $G_n(S)$, implying T_1 and T_2 are connected by a path of length at most $2n-8$. □

Pournin's result for $\text{diam}(G_n)$ when $n>12$ combined with known values for $5 \leq n \leq 12$ imply that for all $n \geq 5$, we have that $\text{diam}(G_n) \leq 2n-8$. Hence, the above result implies the following statements for $\mathcal{D}_n$.

Corollary 3.3.7. *For $n \geq 5$, the flip graph $\mathcal{D}_n$ is connected and*

$$\textit{diam}(\mathcal{D}_n) \leq \textit{diam}(G_n) + 2n - 8 \leq 4n - 16.$$

Proof. Let $(T_1, T_2), (T_3, T_4) \in \mathcal{D}_n$. Then there is a path of length at most $\text{diam}(G_n)$ in $\mathcal{D}_n$ from (T_1, T_2) to (T_3, T) for some T disjoint from T_3. By Theorem 3.3.6, there is a path from (T_3, T) to (T_3, T_4) of length at most $2n-8$ in $\mathcal{D}_n$. □

Theorem 3.3.8 (Theorem 1.1.2 revisited)**.** *For any positive integer $n \geq 5$, the graph $\mathcal{D}_n$ is simple, connected, and $2(n-3)$-regular. Furthermore, a random walk on $\mathcal{D}_n$ that starts at*

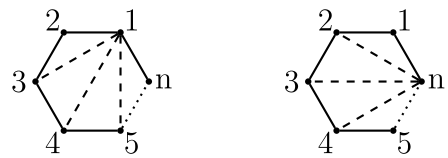


Figure 3.9: A pair of disjoint triangulations where two sequences of flips $(n,2)_2, (2,4)_1$ and $(n,2)_2$, $(1,3)_2$, $(2,4)_1$ both result in the original pair again.

an arbitrary vertex and chooses neighboring vertices uniformly at random will converge to the uniform distribution on the vertices in $\mathcal{D}_n$.

Proof. Lemma 3.3.5 and Corollary 3.3.7 imply that $\mathcal{D}_n$ is simple, connected, and $2(n-3)$-regular. For aperiodicity, Lemma 2.1.5 implies that it suffices to consider the random walk $(X_t)_{t\geq 0}$ starting at the pair (T_1, T_2) shown in Figure 3.9. The flips described in Figure 3.9 imply that

$$\Pr_{(T_1,T_2)}[X_2 = (T_1,T_2)] > 0 \quad \text{and} \quad \Pr_{(T_1,T_2)}[X_3 = (T_1,T_2)] > 0.$$

From this, we find

$$1 = \gcd(2,3) \mid \gcd(t \in \mathbb{Z}_+ \mid \Pr_{(T_1,T_2)}[X_t = (T_1,T_2)] > 0),$$

so the greatest common divisor on the right must be 1. We conclude that this random walk is aperiodic. Convergence to the uniform distribution then follows from Theorem 2.1.7. □

Running the random walk on $\mathcal{D}_n$ from Theorem 3.3.8 for sufficiently many iterations allows for approximately uniform sampling of pairs of disjoint triangulations. Determining the number of iterations needed remains a direction open for future work.

Open Problem. *Determine the mixing time of the random walk from Theorem 3.3.8.*

Remark 3.3.9. We note that a trivial upper bound for the mixing time is $O(|V(\mathcal{D}_n)|^2)$ [LP17, Proposition 10.28]. We suspect that this bound can be improved significantly because

n	5	6	7	8	9
$\lvert V(\mathcal{D}_n)\rvert$	10	68	546	4872	46782
iterations	3	7	14	25	39
σ_2	$0.5590\ldots$	$0.7287\ldots$	$0.8478\ldots$	$0.9512\ldots$	$0.9677\ldots$

Table 3.3: For each n, the number of iterations needed for the total variation distance from the uniform distribution to be smaller than $1/4$ regardless of the initial vertex chosen, and the second largest eigenvalue of the transition matrix.

the mixing time for a simple random walk on the triangulation flip graph G_n is polynomial in n [MT97, MRS01], while $|V(\mathcal{D}_n)|$ grows rapidly with respect to n [RRW18]. The computer data given in Table 3.3 supports our suspicion that $O(|V(\mathcal{D}_n)|^2)$ is not a useful upper bound for the mixing time of $\mathcal{D}_n$.

Chapter 4

COLORED PERMUTATION STATISTICS

In this chapter, we study statistics on the colored permutation groups and their conjugacy classes. We establish Theorem 1.2.10 in Section 4.1, which shows that any fixed moment of a statistic will coincide on all conjugacy classes when there are no cycles of “short” length. In Section 4.2, we show that for certain statistics, this moment aligns with the corresponding ones on $\mathfrak{S}_{n,r}$, allowing us to translate results on the asymptotic behavior of these statistics on $\mathfrak{S}_{n,r}$ to conjugacy classes without “short” cycles. The results in Section 4.1 and Section 4.2.1 are from a collaboration with Jesse Campion Loth, Michael Levet, Sheila Sundaram, and Mei Yin, and these appear in [CLL$^+$23]. The remaining results are from ongoing work with Michael Levet, Sheila Sundaram, and Mei Yin [LLSY24].

4.1 Moments of statistics on conjugacy classes without short cycles

In this section, we will introduce colored permutation constraints and realizability over constraints of a given size. Using constraints, we prove that any fixed moment of a colored permutation statistic coincides on all conjugacy classes without cycles of sufficiently short length. Before giving formal definitions, we start with a motivating example.

Example 4.1.1. Consider the statistic des on $\mathfrak{S}_n$. One method for calculating $\operatorname{des}(\omega)$ is to count the number of $i \in [n-1]$ and $j_1, j_2 \in [n]$ such that $j_1 < j_2$, $\omega(i) = j_2$, and $\omega(i+1) = j_1$. We can formalize this by defining an indicator function

$$I_{\{(i,j_2),(i+1,j_1)\}}(w) = \begin{cases} 1 & \text{if } \omega(i) = j_2 \text{ and } \omega(i+1) = j_2 \\ 0 & \text{otherwise} \end{cases} \tag{4.1.1}$$

and expressing des as

$$\mathrm{des} = \sum_{i=1}^{n-1} \sum_{j_1<j_2} I_{\{(i,j_2),(i+1,j_1)\}}.$$

By replacing each summand with $i \cdot I_{(i,j_2),(i+1,j_1)}$, we also obtain a decomposition of this form for maj on $\mathfrak{S}_n$. Hence, one can view these indicator functions $I_{(i,j_2),(i+1,j_1)}$ as building blocks for the descent and major index statistics. Recall from Theorem 1.2.8 that the k-th moments of des and maj will coincide on all conjugacy classes $C_\lambda \subseteq \mathfrak{S}_n$ that have no cycles of length at most $2k$. One might expect that this factor of 2 is related to the fact that the indicator functions given in (4.1.1) check two conditions, and we will eventually see that this is indeed the case.

We now give our definition of colored permutation constraints. The reader should keep the preceding example in mind as motivation, though our definition applies to $\mathfrak{S}_{n,r}$ for any r. Similar to how colored permutations in $\mathfrak{S}_{n,r}$ consist of two components, a colored permutation constraint will also consist of two components.

Definition 4.1.2. A *colored permutation constraint* on $\mathfrak{S}_{n,r}$ is a pair (K, κ), where

- $K = \{(i_h, j_h)\}_{h=1}^m$ is a set of ordered pairs of elements in $[n]$ where the sets $\{i_1\}_{h=1}^m$ and $\{j_h\}_{h=1}^m$ have m elements, and

- $\kappa : \{i_1, \ldots, i_m\} \to \mathbb{Z}_r$ is any function, which we can also represent as ordered pairs $\{(i_h, \kappa(i_h))\}_{h=1}^m$.

We call m the *size* of (K, κ), and we denote it as $|(K, \kappa)|$. For brevity, we will sometimes denote a constraint using a single set of ordered pairs

$$(K, \kappa) = \left\{ \left(i_h^0, j_h^{\kappa(i_h)} \right) \right\}_{h=1}^m \tag{4.1.2}$$

of elements in $[n]^r$. When specifically considering $\mathfrak{S}_n$, the function κ and all colors will be omitted. When specifically considering B_n, the codomain of κ will be $\{+, -\}$, and ordered pairs in (4.1.2) will be denoted (i_h, j_h) with $i_h \in [n]$ and $j_h \in [\pm n]$.

Definition 4.1.3. Let (K, κ) be a colored permutation constraint on $\mathfrak{S}_{n,r}$ of size m. A permutation $\omega \in \mathfrak{S}_n$ *satisfies* K if $\omega(i_h) = j_h$ for all $h \in [m]$. A coloring $\tau : [n] \to \mathbb{Z}_r$ *satisfies* κ if $\tau(i_h) = \kappa(i_h)$ for all $h \in [m]$. A colored permutation $(\omega, \tau) \in \mathfrak{S}_{n,r}$ *satisfies* (K, κ) if ω satisfies K and τ satisfies κ. In the bijective notation, this is equivalent to $(\omega, \tau)(i_h^0) = j_h^{\kappa(i_h)}$ for all $h \in [m]$.

Example 4.1.4. The colored permutation $(\omega, \tau) = (1^1 4^3 3^1)(2^3 5^0)$ from Example 2.2.2 satisfies the size 3 constraint $(K, \kappa) = \{(1^0, 4^3), (4^0, 3^1), (5^0, 2^3)\}$, but does not satisfy the size 2 constraint $\{(1^0, 4^2), (4^0, 3^1)\}$.

Recall that a colored permutation statistic is a function $X : \mathfrak{S}_{n,r} \to \mathbb{R}$, and equipping $\mathfrak{S}_{n,r}$ with the uniform distribution allows us to consider X as a random variable. For any colored permutation constraint (K, κ), one useful statistic will be the indicator function $I_{(K,\kappa)} : \mathfrak{S}_{n,r} \to \mathbb{R}$ that takes value 1 on colored permutations satisfying (K, κ) and 0 otherwise. We will view these indicator functions as building blocks for many permutation statistics through the following definition.

Definition 4.1.5. A colored permutation statistic $X : \mathfrak{S}_{n,r} \to \mathbb{R}$ is *realizable over constraints of size* m if X is in the $\mathbb{R}$-vector space spanned by $\{I_{(K,\kappa)} \mid |(K, \kappa)| \leq m\}$. The *size* of a statistic is the minimum possible m such that X is realizable over constraints of size m.

Many statistics have natural decompositions in terms of constraints. We give some examples using the ordered pair notation from Definition 4.1.2.

Example 4.1.6. From Example 4.1.1, the descent and major index statistics on $\mathfrak{S}_n$ are realizable over constraints of size 2. More generally, the descent, major index, and flag-major index statistics on $\mathfrak{S}_{n,r}$ for any r are realizable over constraints of size 2:

$$\mathrm{des}_{n,r} = \sum_{i=1}^{n-1} \sum_{j_1^{c_1} < j_2^{c_2}} I_{\{(i^0, j_2^{c_2}), ((i+1)^0, j_1^{c_1})\}} + \sum_{j=1}^{n} \sum_{c=0}^{r-1} I_{\{(n^0, j^c)\}},$$

$$\mathrm{maj}_{n,r} = \sum_{i=1}^{n-1} \sum_{j_1^{c_1} < j_2^{c_2}} i \cdot I_{\{(i^0, j_2^{c_2}),((i+1)^0, j_1^{c_1})\}},$$

$$\mathrm{fmaj}_{n,r} = r \cdot \sum_{i=1}^{n-1} \sum_{j_1^{c_1} < j_2^{c_2}} i \cdot I_{\{(i^0, j_2^{c_2}),((i+1)^0, j_1^{c_1})\}} + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{c=0}^{r-1} c \cdot I_{\{(i^0, j^c)\}},$$

where the conditions $j_1^{c_1} < j_2^{c_2}$ are with respect to the total order on $[n]^r$ given in (2.2.2).

Remark 4.1.7. A colored permutation $(\omega, \tau) \in \mathfrak{S}_{n,r}$ is itself a colored permutation constraint of size n. Hence, any statistic X is realizable over constraints of size n as

$$X = \sum_{(\omega,\tau) \in \mathfrak{S}_{n,r}} X(\omega, \tau) \cdot I_{(\omega,\tau)}.$$

For the full strength of our results, we wish to realize statistics over constraints of small size.

For any colored permutation constraint (K, κ) on $\mathfrak{S}_{n,r}$ of size m, we now consider the mean of $I_{(K,\kappa)}$ on $\mathfrak{S}_{n,r}$ and its conjugacy classes. As $I_{(K,\kappa)}$ is an indicator function, its mean on $\Omega \subseteq \mathfrak{S}_{n,r}$ is equivalent to the probability that $I_{(K,\kappa)}$ takes value 1 on Ω. Since generating $(\omega, \tau) \in \mathfrak{S}_{n,r}$ uniformly at random can be decomposed into generating $\omega \in \mathfrak{S}_n$ and $\tau : [n] \to \mathbb{Z}_r$ uniformly at random, it is not difficult to show that

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[I_{(K,\kappa)}] = \mathrm{Pr}_{\mathfrak{S}_{n,r}}[(\omega, \tau) \text{ satisfies } (K, \kappa)] = \frac{1}{n(n-1)\dots(n-m+1)} \cdot \frac{1}{r^m}.$$

In general, the corresponding expectation on arbitrary $C_{\boldsymbol{\lambda}} \subseteq \mathfrak{S}_{n,r}$ is complex. However, when all cycles in $C_{\boldsymbol{\lambda}}$ have sufficiently long lengths, we will see that $\mathbb{E}_{\boldsymbol{\lambda}}[I_{(K,\kappa)}]$ has a reasonably nice formulation. We will need the following notions of constraint graphs and acyclicity.

Definition 4.1.8. Let (K, κ) be a colored permutation constraint on $\mathfrak{S}_{n,r}$. The *constraint graph* of (K, κ), denoted $G(K, \kappa)$, is the directed graph with vertex set $[n]$, directed edge set K, and coloring of each edge $(i, j) \in K$ given by $\kappa(i)$.

Definition 4.1.9. A colored permutation constraint (K, κ) on $\mathfrak{S}_{n,r}$ is *acyclic* if its constraint graph $G(K, \kappa)$ does not contain any cycles. Observe that in this case, $G(K, \kappa)$ consists of a set of directed paths, each of which has a source vertex. Additionally, acyclicity implies that the total number of edges in $G(K, \kappa)$ is $|(K, \kappa)| < n$.

We now show that on any conjugacy class without short cycles,

$$\mathbb{E}_{\boldsymbol{\lambda}}[I_{(K,\kappa)}] = \Pr_{\lambda}[(\omega, \tau) \text{ satisfies } (K, \kappa)]$$

takes one of two values determined entirely by acyclicity of (K, κ). We do this by explicitly calculating the probability above.

Lemma 4.1.10. *Let (K, κ) be a colored permutation constraint on $\mathfrak{S}_{n,r}$ of size m, and let $C_{\boldsymbol{\lambda}}$ be a conjugacy class of $\mathfrak{S}_{n,r}$ with no cycles of length $1, 2, \ldots, m$. If (K, κ) is not acyclic, then*

$$\Pr_{\boldsymbol{\lambda}}[\omega \text{ satisfies } K] = 0.$$

If (K, κ) is acyclic, then

$$\Pr_{\boldsymbol{\lambda}}[\omega \text{ satisfies } K] = \frac{1}{(n-1)(n-2)\cdots(n-m)}.$$

Proof. We first consider the case when (K, κ) is not acyclic. In this case, in order for an element $(\omega, \tau) \in C_{\boldsymbol{\lambda}}$ to have the property that ω satisfies K, it must be that ω contains a cycle induced by the conditions in K. Since K has size m, this cycle has length at most m. However, we assumed $C_{\boldsymbol{\lambda}}$ has no cycles of length $1, 2, \ldots, m$, so this is not possible. Hence, no such elements exist, and we conclude that $\Pr_{\boldsymbol{\lambda}}[\omega \text{ satisfies } K] = 0$.

For the case when (K, κ) is acyclic, we use induction on n and m to compute the probability that a colored permutation of cycle type $\boldsymbol{\lambda}$ chosen uniformly at random satisfies K. As base cases, consider when n is an arbitrary positive integer and $m = 1$. We express $K = \{(i, j)\}$ and analyze $\Pr_{\boldsymbol{\lambda}}[\omega(i) = j]$. Consider any $k \in [n] \setminus \{i, j\}$. Letting $\mathbf{0} : [n] \to \mathbb{Z}_r$ denote the zero coloring, conjugating by the colored permutation $((j, k), \mathbf{0})$ induces bijections between

$$\{(\omega, \tau) \in C_{\boldsymbol{\lambda}} \mid \omega(i) = j\} \quad \text{and} \quad \{(\omega, \tau) \in C_{\boldsymbol{\lambda}} \mid \omega(i) = k\}.$$

Therefore, $\Pr_{\boldsymbol{\lambda}}[\omega(i) = k]$ is invariant under our choice of $k \in [n] \setminus \{i\}$. Since $C_{\boldsymbol{\lambda}}$ does not contain cycles of length $1, 2, \ldots, m$, we have that $\Pr_{\boldsymbol{\lambda}}[\omega(i) = i] = 0$. Combined, we conclude

$$\Pr_{\boldsymbol{\lambda}}[\omega(i) = j] = \frac{1}{n-1}.$$

This establishes our base case of $m = 1$ and positive integers $n \geq 2$.

Now fix $n > m > 1$, and suppose that the result holds for $n-1$ and $m-1$. Consider an acyclic constraint (K, κ) on $\mathfrak{S}_{n,r}$ of size m and a conjugacy class $C_{\boldsymbol{\lambda}}$ of $\mathfrak{S}_{n,r}$ without cycles of length $1, 2, \ldots m$. Express $K = \{(i_1, j_1), \ldots, (i_m, j_m)\}$, and partition $C_{\boldsymbol{\lambda}}$ according to the cycle containing i_m. Namely, let $C_{\boldsymbol{\lambda}} = \bigsqcup_{k=m+1}^{n} \bigsqcup_{c \in \mathbb{Z}_r} \Omega_{k,c}$, where

$$\Omega_{k,c} = \{(\omega, \tau) \in C_{\boldsymbol{\lambda}} \mid i_m \text{ appears in a cycle of length } k \text{ and color } c\}.$$

Using conditional expectations, the law of total probability, and our result in the preceding paragraph, we express $\Pr_{\boldsymbol{\lambda}}[\omega \text{ satisfies } K]$ as

$$\begin{aligned}
&\Pr_{\boldsymbol{\lambda}}\left[\bigcap_{h=1}^{m}\{\omega(i_h) = j_h\}\right] \\
&= \Pr_{\boldsymbol{\lambda}}\left[\bigcap_{h=1}^{m-1}\{\omega(i_h) = j_h\} \,\middle|\, \omega(i_m) = j_m\right] \cdot \Pr_{\boldsymbol{\lambda}}[\omega(i_m) = j_m] \\
&= \frac{1}{n-1} \cdot \sum_{k=m+1}^{n} \sum_{c \in \mathbb{Z}_r} \Pr_{\boldsymbol{\lambda}}\left[\bigcap_{h=1}^{m-1}\{\omega(i_h) = j_h\} \cap \Omega_{k,c} \,\middle|\, \omega(i_m) = j_m\right] \\
&= \frac{1}{n-1} \cdot \sum_{k=m+1}^{n} \sum_{c \in \mathbb{Z}_r} \Pr_{\boldsymbol{\lambda}}\left[\bigcap_{h=1}^{m-1}\{\omega(i_h) = j_h\} \,\middle|\, \Omega_{k,c} \cap \{\omega(i_m) = j_m\}\right] \cdot \Pr_{\boldsymbol{\lambda}}[\Omega_{k,c} \mid \omega(i_m) = j_m].
\end{aligned} \tag{4.1.3}$$

We now consider the remaining pairs $\{(i_1, j_1), \ldots, (i_{m-1}, j_{m-1})\}$. By reordering the pairs in K if necessary, we can assume without loss of generality that i_m is a source vertex in $G(K, \kappa)$, so i_m is not an element in (i_h, j_h) for $1 \leq h \leq m-1$. For any fixed $k \in [n]$ and $c \in \mathbb{Z}_r$ where $\Pr_{\boldsymbol{\lambda}}[\Omega_{k,c}] \neq 0$, define $\boldsymbol{\lambda}'$ to be the r-partition of $[n-1]$ obtained by starting with $\boldsymbol{\lambda}$ and replacing a part in λ^c of size k with a part of size $k-1$. Since $\boldsymbol{\lambda}$ contains only partitions whose parts have lengths larger than m, $\boldsymbol{\lambda}'$ contains only partitions whose parts have lengths larger than $m-1$. Additionally, let $f : [n] \setminus \{i_m\} \to [n-1]$ be the unique order-preserving function. Using this, define a function

$$\pi_{k,c} : \{(\omega, \tau) \in \Omega_{k,c} \mid \omega(i_m) = j_m\} \to C_{\boldsymbol{\lambda}'}$$

that in the cycle notation of (ω, τ) replaces $i_m^a j_m^b$ with $f(j_m)^{a+b}$ and replaces all other elements $x \in [n] \setminus \{i_m, j_m\}$ with $f(x)$. Observe that π is an r-to-1 map. Additionally, for a colored permutation (ω, τ) in the domain with image $(\omega', \tau') \in C_{\lambda'}$, we have that $\omega(i_h) = j_h$ if and only if $\omega'(f(i_h)) = f(j_h)$, where $1 \le h \le m-1$. In particular, if we let $(\omega', \tau') \in C_{\boldsymbol{\lambda}'}$ be generated uniformly at random, then this observation combined with our induction hypothesis implies that for each fixed k and c,

$$\begin{aligned}\Pr_{\boldsymbol{\lambda}}\left[\bigcap_{h=1}^{m-1}\{\omega(i_h) = j_h\} \,\middle|\, \Omega_{k,c} \cap \{\omega(i_m) = j_m)\}\right] &= \Pr_{\boldsymbol{\lambda}'}\left[\bigcap_{h=1}^{m-1}\{\omega'(f(i_h)) = f(j_h)\}\right] \\ &= \frac{1}{(n-2)(n-3)\cdots(n-m)}.\end{aligned}$$

Note that the first term is $(n-1)-1$ and the last term is $(n-1)-(m-1)$ since this probability involves $\mathfrak{S}_{n-1,r}$ and $\{(f(i_h), f(j_h)\}_{h=1}^{m-1}$ has size $m-1$. Returning to (4.1.3), we conclude

$$\begin{aligned}&\Pr_{\boldsymbol{\lambda}}\left[\bigcap_{h=1}^{m}\{\omega(i_h) = j_h\}\right] \\ &= \frac{1}{n-1}\sum_{k=m+1}^{n}\sum_{c \in \mathbb{Z}_r}\frac{1}{(n-2)(n-3)\cdots(n-m)} \cdot \Pr_{\boldsymbol{\lambda}}[\Omega_{k,c} \mid \omega(i_m) = j_m] \\ &= \frac{1}{(n-1)(n-2)\cdots(n-m)} \cdot \sum_{k=m+1}^{n}\sum_{c \in \mathbb{Z}_r}\Pr_{\boldsymbol{\lambda}}[\Omega_{k,c} \mid \omega(i_m) = j_m] \\ &= \frac{1}{(n-1)(n-2)\cdots(n-m)}.\end{aligned}$$

□

Lemma 4.1.11. *Let* (K,κ) *be a colored permutation constraint on* $\mathfrak{S}_{n,r}$ *of size* m*, and let* $C_{\boldsymbol{\lambda}}$ *be a conjugacy class of* $\mathfrak{S}_{n,r}$ *with no cycles of length* $1,2,\ldots,m$*. If* (K,κ) *is not acyclic, then*

$$\Pr_{\boldsymbol{\lambda}}[(\omega,\tau) \text{ satisfies } (K,\kappa)] = 0.$$

If (K,κ) *is acyclic, then*

$$\Pr_{\boldsymbol{\lambda}}[(\omega,\tau) \text{ satisfies } (K,\kappa)] = \frac{1}{(n-1)(n-2)\cdots(n-m)} \cdot \frac{1}{r^m}.$$

Proof. First, we express

$$\Pr_{\boldsymbol{\lambda}}[(\omega,\tau) \text{ satisfies } (K,\kappa)] = \Pr_{\boldsymbol{\lambda}}[\omega \text{ satisfies } K] \cdot \Pr_{\boldsymbol{\lambda}}[\tau \text{ satisfies } \kappa \mid \omega \text{ satisfies } K]. \quad (4.1.4)$$

After applying Lemma 4.1.10 to the first term on the right side of (4.1.4), it suffices to show that when (K,κ) is acyclic, the second term is $1/r^m$.

Express $K = \{(i_h, j_h)\}_{h=1}^m$. For each $h \in [m]$, define e_h to be the m-tuple with 0 everywhere and 1 in position h. Using this, define an action of $\mathbb{Z}_r^m$ on $C_{\boldsymbol{\lambda}}$ as follows: e_h acts on (ω,τ) by

- adding 1 to the color $\tau(i_h)$, and
- subtracting 1 from the color $\tau(x)$, where $x \in [n] \setminus \{i_1,\ldots,i_m\}$ is the smallest element that appears in the cycle containing i_h.

Since $C_{\boldsymbol{\lambda}}$ contains no cycles of length $1,2,\ldots,m$, the element x always exists. Observe that this action on (ω,τ) does not affect ω, and using this fact, it is straightforward to see that extending this action linearly to all elements in $\mathbb{Z}_r^m$ results in a well-defined group action on $C_{\boldsymbol{\lambda}}$, and each orbit has size r^m. It is clear that each orbit contains an element $(\omega^\star,\tau^\star)$ such that $\tau^\star$ satisfies κ, and since the action of any nonzero element in $\mathbb{Z}_r^m$ on $(\omega^\star,\tau^\star)$ results in a colored permutation $(\omega^\star,\tau')$ where τ' does not satisfy κ, we conclude that exactly one element in each orbit has this property. Furthermore, the subset $\{(\omega,\tau)\in C_{\boldsymbol{\lambda}} \mid \omega \text{ satisfies } K\}$ is invariant under this action, and this allows us to conclude

$$\Pr_{\boldsymbol{\lambda}}[\tau \text{ satisfies } \kappa \mid \omega \text{ satisfies } K] = \Pr_{\boldsymbol{\lambda}}[\tau \text{ satisfies } \kappa] = \frac{1}{r^m}. \tag{4.1.5}$$

Combined, we conclude the result when (K,κ) is acyclic. □

Our proof in the preceding lemma also implies the following corollary. One can view this corollary as stating that satisfying K and satisfying κ are independent when all cycles in $C_{\boldsymbol{\lambda}}$ have sufficiently long lengths.

Corollary 4.1.12. *Let (K,κ) be a colored permutation constraint on $\mathfrak{S}_{n,r}$ of size m, and let $C_{\boldsymbol{\lambda}}$ be a conjugacy class of $\mathfrak{S}_{n,r}$ with no cycles of length $1,2,\dots,m$. Then*

$$\Pr_{\boldsymbol{\lambda}}[(\omega,\tau) \textit{ satisfies } (K,\kappa)] = \Pr_{\boldsymbol{\lambda}}[\omega \textit{ satisfies } K]\cdot \Pr_{\boldsymbol{\lambda}}[\tau \textit{ satisfies } \kappa].$$

Remark 4.1.13. Our preceding results also imply that when (K,κ) is an acyclic constraint of size m and $C_{\boldsymbol{\lambda}}$ has no cycles of length $1,2,\dots,m$,

$$\Pr_{\boldsymbol{\lambda}}[\omega \text{ satisfies } K] = \Pr_{\boldsymbol{\lambda}}[\omega \text{ satisfies } K \mid \tau \text{ satisfies } \kappa] = \frac{1}{(n-1)(n-2)\dots(n-m)}.$$

While one can attempt to show these statements directly, we found this to be much more technical than the proof of Lemma 4.1.10.

Lemma 4.1.11 can be used to conclude results involving the means of statistics on conjugacy classes with sufficiently long cycle lengths, but our main result stated in Theorem 1.2.10 involves arbitrary moments. To connect these, we will need to analyze products of indicator functions $I_{(K,\kappa)}$. The terminology below is adapted from [HR22].

Definition 4.1.14. Two colored permutation constraints (K_1,κ_1) and (K_2,κ_2) on $\mathfrak{S}_{n,r}$ are *compatible* if there exists a colored permutation $(\omega,\kappa)\in\mathfrak{S}_{n,r}$ satisfying both constraints.

Note that when two constraints are compatible, we can use the unions $K_1 \cup K_2$ and $\kappa_1 \cup \kappa_2$ to define a new constraint, and the size of $(K_1 \cup K_2, \tau_1 \cup \tau_2)$ is bounded by the sum of the sizes of (K_1, τ_1) and (K_1, τ_2). Using this observation, the following two results are straightforward exercises.

Lemma 4.1.15. *Let (K_1, κ_1) and (K_2, κ_2) be two colored permutation constraints on $\mathfrak{S}_{n,r}$. If (K_1, κ_1) and (K_2, κ_2) are not compatible, then $I_{(K_1,\kappa_1)} \cdot I_{(K_2,\kappa_2)}$ is identically zero. If (K_1, κ_1) and (K_2, κ_2) are compatible, then $I_{(K_1,\kappa_1)} \cdot I_{(K_2,\kappa_2)} = I_{(K_1 \cup K_2, \kappa_1 \cup \kappa_2)}$.*

Corollary 4.1.16. *Suppose X_1 and X_2 are statistics on $\mathfrak{S}_{n,r}$ that are realizable over constraints of size m_1 and m_2, respectively. Then $X_1 \cdot X_2$ is realizable over constraints of size $m_1 + m_2$. In particular, for any integer $k \geq 1$ such that $m_1 k \leq n$, we have that X_1^k is realizable over constraints of size km_1.*

Combining all of our results, we now establish Theorem 1.2.10. This formalizes the statement that the k-th moment of a statistic X will coincide on all $C_{\boldsymbol{\lambda}}$ without "short" cycles.

Theorem 4.1.17 (Theorem 1.2.10 revisited)**.** *Suppose $X : \mathfrak{S}_{n,r} \to \mathbb{R}$ is realizable over constraints of size m, where m is some positive integer. For any $k \geq 1$, the k-th moment $\mathbb{E}_{\boldsymbol{\lambda}}[X^k]$ coincides on all conjugacy classes $C_{\boldsymbol{\lambda}}$ with no cycles of length $1, 2, \ldots, mk$.*

Proof. We first consider the case when $k = 1$ and $C_{\boldsymbol{\lambda}}$ has no cycles of length $1, 2, \ldots, m$. Express $X = \sum_{(K,k)} c_{(K,\kappa)} I_{(K,\kappa)}$ where each (K, κ) has size at most m. Linearity of expectation implies

$$\mathbb{E}_{\boldsymbol{\lambda}}[X] = \sum_{(K,\kappa)} c_{(K,\kappa)} \cdot \mathbb{E}_{\boldsymbol{\lambda}}[I_{(K,\kappa)}]. \tag{4.1.6}$$

Applying Lemma 4.1.11 to each summand, either $\mathbb{E}_{\boldsymbol{\lambda}}[I_{(K,\kappa)}] = 0$ or

$$\mathbb{E}_{\boldsymbol{\lambda}}[I_{(K,\kappa)}] = \frac{1}{(n-1)(n-2)\ldots(n-|(K,\kappa)|)} \cdot \frac{1}{r^{|(K,\kappa)|}}, \tag{4.1.7}$$

and the latter case corresponds to acyclicity of (K,κ). Acyclicity of (K,κ) is independent of $C_{\boldsymbol{\lambda}}$, so (4.1.6) is equivalent to the following expression that is independent of $C_{\boldsymbol{\lambda}}$:

$$\sum_{\text{acyclic }(K,\kappa)} \frac{c_{(K,\kappa)}}{(n-1)(n-2)\dots(n-|(K,\kappa)|)} \cdot \frac{1}{r^{|(K,\kappa)|}}.$$

For $k \geq 2$, it suffices to consider when $mk < n$, as there are no conjugacy classes $C_{\boldsymbol{\lambda}}$ without cycles of length $1, 2, \dots, mk$ when $mk \geq n$. Corollary 4.1.16 implies that if X is realizable over constraints of size m, then X^k is realizable over constraints of size mk. The general result for $\mathbb{E}_{\boldsymbol{\lambda}}[X^k]$ now follows by combining this with the above case of $k = 1$. □

In the case of $\mathfrak{S}_n$, there are known representation-theoretic explanations for Theorem 4.1.17. Let Class$(\mathfrak{S}_n, \mathbb{C})$ be the set of class functions from S_n to $\mathbb{R}$. The irreducible characters of $\mathfrak{S}_n$ form a basis for these class functions, and these characters are indexed by partitions. See [Mac98] for details. We use χ^λ for the irreducible character indexed by a partition $\lambda \vdash n$. The following is a result of Hamaker and Rhoades.

Theorem 4.1.18. *[HR22, Theorem 3.16] For any $1 \leq k \leq n$ and positive integer m,*

$$\mathit{Class}(S_n, \mathbb{C}) \cap \mathit{span}_{\mathbb{C}} \{I_K : |K| \leq m\}$$

is the subspace of class functions spanned by the irreducible characters χ^λ for partitions $\lambda \vdash n$ whose largest part has size at least $n - m$.

Letting $X : \mathfrak{S}_n \to \mathbb{R}$ be a permutation statistic, the class function

$$\overline{X}(\omega) = \frac{1}{|S_n|} \sum_{\pi \in S_n} X(\pi^{-1}\omega\pi).$$

maps each $\omega \in \mathfrak{S}_n$ to the mean of X on the conjugacy class containing ω. Hamaker and Rhoades show that if X is realizable over constraints of size m, then so is $\overline{X}$. Theorem 4.1.18 implies that $\overline{X}$ can be expressed as a linear combination of characters χ^λ where the largest part of λ is at least $n-m$. The theory of character polynomials [GG09] implies that χ^λ_μ depends only on the number of cycles of length $1, 2, \ldots, m$ in μ. Combined, we can conclude that the value of $\overline{X}$ on a conjugacy class C_μ depends only on the number of cycles with lengths $1, 2, \ldots, m$ in μ. Theorem 4.1.17 on the symmetric group $\mathfrak{S}_n$ can be viewed as the case where there are 0 cycles of length $1, 2, \ldots, m$. Given that Theorem 4.1.17 holds on any general colored permutation group $\mathfrak{S}_{n,r}$, it is likely that a generalization of Theorem 4.1.18 exists for $\mathfrak{S}_{n,r}$. This leads to the following problem, which is a direction for future work.

Open Problem. *Use the representation theory of $\mathfrak{S}_{n,r}$ to obtain an analog of Theorem 4.1.18 for $\mathfrak{S}_{n,r}$. For various statistics X, study the resulting expansion for $\overline{X}$ in terms of the irreducible characters of $\mathfrak{S}_{n,r}$.*

Interested readers can consult [Mac98] for general background on the representation theory of $\mathfrak{S}_{n,r}$. Since much of this theory builds on the representation theory of $\mathfrak{S}_n$, certain techniques from [HR22] should generalize from $\mathfrak{S}_n$ to $\mathfrak{S}_{n,r}$. However, one may need to also generalize additional representation-theoretic results from $\mathfrak{S}_n$ to $\mathfrak{S}_{n,r}$. For example, we have not seen an analog of character polynomials in $\mathfrak{S}_{n,r}$ in the literature, though a Murnaghan-Nakayama Rule is known [Ste89].

4.2 Asymptotics on conjugacy classes without short cycles

Our results in the preceding section show that any fixed moment of a statistic will coincide on all conjugacy classes of $\mathfrak{S}_{n,r}$ without short cycles. As stated in Theorem 1.2.8, Fulman showed that for the descent and major index statistics on $\mathfrak{S}_n$, these moments align with the corresponding ones on all of $\mathfrak{S}_n$. While this does not hold for all statistics, we show in this section that this holds for des_{B_n} on $B_n \cong \mathfrak{S}_{n,2}$ and the $\mathrm{des}_{n,r}, \mathrm{maj}_{n,r}$, and $\mathrm{fmaj}_{n,r}$ statistics on $\mathfrak{S}_{n,r}$. Note that Remark 2.2.17 shows that the result for des_{B_n} does not follow immediately from the one for $\mathrm{des}_{n,r}$, as these statistics do not necessarily share the same distribution on conjugacy classes.

4.2.1 The Coxeter group descent statistic on B_n

In this subsection, we consider des_{B_n} on conjugacy classes of B_n, as defined in Section 2.2.2. We will derive properties involving the generating function for $\mathrm{des}_{B_n}+1$ on a conjugacy class $C_{\lambda,\mu}$, as well as a formula for the number of elements in $C_{\lambda,\mu}$ with a fixed number of descents. We then consider the moments of des_{B_n} on conjugacy classes without short cycles.

We first recall the definition of the descent statistic

$$\mathrm{des}_{B_n}(\omega) = |\{i \in \{0\} \cup [n-1] \mid \omega(i) > \omega(i+1)\}|$$

with the convention that $\omega(0) = 0$. Note that the condition $\omega(i) > \omega(i+1)$ is based on the usual order on $[\pm n]$. Using the above definition, one can realize des_{B_n} over constraints of size 2 using

$$\mathrm{des}_{B_n} = \sum_{j\in[n]} I_{(1,-j)} + \sum_{i\in[n-1]} \sum_{\substack{k,\ell\in[\pm n]\\ k<\ell}} I_{\{(i,\ell),(i+1,k)\}}.$$

Theorem 4.1.17 implies that the k-th moment of des_{B_n} coincides on all $C_{\lambda,\mu}$ with no cycles of length $1, 2, \ldots, 2k$. We will see later that this moment aligns with the k-th moment of des_{B_n} on all of B_n.

We note that Reiner [Rei93b] uses a different notion of descents, which we describe now. Under the ordering

$$1 < 2 < \cdots < n < -n < \cdots < -2 < -1, \tag{4.2.1}$$

ω has a descent at position $i \in [n-1]$ if $\omega(i) > \omega(i+1)$, and ω has a descent at position n if $\omega(n) < 0$. While the two definitions are different, [FKLP21, Remark 5.1] shows that the generating function

$$\sum_{\omega\in B_n} t^{\mathrm{des}_{B_n}(\omega)+1} \prod_i x_i^{m_i(\lambda(\omega))} y_i^{m_i(\mu(\omega))} \tag{4.2.2}$$

is unaffected, where $(\lambda(\omega), \mu(\omega))$ denotes the cycle type of $\omega \in B_n$.

Following Fulman's analysis for the symmetric group in [Ful98], we analyze the generating function given in (4.2.2) for des_{B_n} on a conjugacy class and derive an alternative expression for it. We will need several known results, which we state now. We begin with a definition for an expression that will come up frequently in our analysis.

Definition 4.2.1. Let $\mu(d)$ be the number-theoretic Möbius function on the positive integers,

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^k & \text{if } n = p_1 \dots p_k \text{ for distinct primes } p_1, \dots, p_k \\ 0 & \text{otherwise.} \end{cases}$$

For nonnegative integers r and m, define

$$N(r, 2m) = \frac{1}{2m} \sum_{\substack{d|m \\ d \text{ odd}}} \mu(d) \left(r^{m/d} - 1\right). \tag{4.2.3}$$

Remark 4.2.2. Reiner [Rei93b, Theorem 4.1 and Theorem 4.2] describes two sets of objects, *primitive blinking necklaces* and *primitive twisted necklaces*. Reiner then shows that if $D_{m,i}^{(k)}$ is the number of primitive blinking necklaces D of size m with $|D| = i$ and $\max(D) \leq k-1$, and $P_{m,i}^{(k)}$ is the number of primitive twisted necklaces P of size m with $|P| = i$ and $\max(P) \leq k-1$, then

$$\sum_{i=0}^{(k-1)m} D_{m,i}^{(k)} = \begin{cases} N(2k-1, 2m) & \text{if } k > 0 \text{ and } m > 1 \\ k & \text{if } k > 0 \text{ and } m = 1 \\ 0 & \text{if } k = 0, \end{cases} \tag{4.2.4}$$

$$\sum_{i=0}^{(k-1)m} P_{m,i}^{(k)} = \begin{cases} N(2k-1, 2m) & \text{if } k > 0 \text{ and } m > 1 \\ k-1 & \text{if } k > 0 \text{ and } m = 1 \\ 0 & \text{if } k = 0. \end{cases} \tag{4.2.5}$$

Consequently, $N(2k-1, 2m)$ enumerates primitive blinking necklaces D of size m where $0 \leq |D| \leq (k-1)m$ and $\max(D) \leq k-1$, or primitive twisted necklaces P of size m where $0 \leq |P| \leq (k-1)m$ and $\max(D) \leq k-1$. In particular, $N(2k-1, 2m)$ must be a nonnegative integer for all $k, m \geq 1$.

We use the notation $(\lambda(\omega), \mu(\omega))$ for the cycle type of $\omega \in B_n$. The following result, which is stated without proof in [FKLP21, Theorem 5.3], is a special case of [Rei93b, Theorem 4.1]. In view of some possibly confusing typos in these papers, and the fact that this result plays a key role in our analysis, we include a proof. Additionally, [Rei93b, Theorem 4.1] uses the convention that B_0 is the trivial group containing a single permutation with 0 descents, and we maintain this convention throughout our work.

Theorem 4.2.3. *[FKLP21, Theorem 5.3] The following holds:*

$$\begin{aligned}\sum_{n\geq 0}\frac{u^n}{(1-t)^{n+1}}\left(\sum_{\omega\in B_n} t^{\mathrm{des}_{B_n}(\omega)+1}\prod_i x_i^{m_i(\lambda(\omega))}y_i^{m_i(\mu(\omega))}\right)\\ = 1+\sum_{k\geq 1} t^k\frac{1}{1-x_1u}\prod_{m\geq 1}\left(\frac{1+y_m u^m}{1-x_m u^m}\right)^{N(2k-1,2m)}.\end{aligned} \tag{4.2.6}$$

Proof. In [Rei93b, Theorem 4.1], it was established that

$$\sum_{n\geq 0}\frac{u^n}{(t;q)_{n+1}}\left(\sum_{\omega\in B_n} t^{\mathrm{des}(\omega)+1}q^{\mathrm{maj}(\omega)}\prod_i x_i^{m_i(\lambda(\omega))}y_i^{m_i(\mu(\omega))}\right) \tag{4.2.7}$$

is equal to

$$\sum_{k\geq 0} t^k \prod_{m\geq 1}\prod_{i=0}^{(k-1)m}(1-x_m u^m q^i)^{-D_{m,i}^{(k)}}(1+y_m u^m q^i)^{P_{m,i}^{(k)}}, \tag{4.2.8}$$

where

- $(t;q)_{n+1} = (1-t)(1-tq)\cdots(1-tq^n)$,
- $\mathrm{des}(\omega)$ and $\mathrm{maj}(\omega)$ are the descent and major index statistics with respect to the ordering in (4.2.1), and
- $D_{m,i}^{(k)}$, $P_{m,i}^{(k)}$ are the nonnegative integers defined by Reiner as described in Remark 4.2.2.

Setting $q = 1$, the equality of (4.2.7) and (4.2.8) implies

$$\begin{aligned}
&\sum_{n\geq 0} \frac{u^n}{(1-t)^{n+1}} \left(\sum_{\omega \in B_n} t^{\mathrm{des}(\omega)+1} \prod_i x_i^{m_i(\lambda(\omega))} y_i^{m_i(\mu(\omega))} \right) \\
&= \sum_{k\geq 0} t^k \prod_{m\geq 1} \left[\left((1-x_m u^m)^{-\sum_{i=0}^{(k-1)m} D_{m,i}^{(k)}} \right) \cdot \left((1+y_m u^m)^{\sum_{i=0}^{(k-1)m} P_{m,i}^{(k)}} \right) \right] \qquad (4.2.9) \\
&= 1 + \sum_{k\geq 1} t^k \prod_{m\geq 1} \left[\left((1-x_m u^m)^{-\sum_{i=0}^{(k-1)m} D_{m,i}^{(k)}} \right) \cdot \left((1+y_m u^m)^{\sum_{i=0}^{(k-1)m} P_{m,i}^{(k)}} \right) \right].
\end{aligned}$$

As noted in the discussion surrounding (4.2.1), we can replace des in the first expression of (4.2.9) with des_{B_n}. The result then follows by applying Remark 4.2.2 on the last expression in (4.2.9), noting that in the boundary case $m = k = 1$, the exponent of $(1 - x_1 u)$ is $N(2k-1, 2) + 1$ from (4.2.4), while the exponent of $(1 - y_1 u)$ is precisely $N(2k-1, 2)$ from (4.2.5). □

For a fixed bi-partition (λ, μ) of n, we can now derive the following expression for the generating function

$$B_{\lambda,\mu}(t) = \sum_{\omega \in C_{\lambda,\mu}} t^{\mathrm{des}_{B_n}(\omega)+1} \qquad (4.2.10)$$

of descents over the conjugacy class $C_{\lambda,\mu}$.

Lemma 4.2.4. *Let $C_{\lambda,\mu}$ be a conjugacy class of B_n where $(\lambda, \mu) \neq ((1^n), \emptyset)$. Then*

$$\frac{B_{\lambda,\mu}}{(1-t)^{n+1}} = \sum_{k\geq 2} t^k \binom{N(2k-1,2)+m_1(\lambda)}{m_1(\lambda)} \prod_{i\geq 1} \binom{N(2k-1,2i)}{m_i(\mu)} \prod_{i\geq 2} \binom{N(2k-1,2i)+m_i(\lambda)-1}{m_i(\lambda)}. \qquad (4.2.11)$$

Proof. We set $u = 1$ in (4.2.6), and the first line can be rewritten as

$$\sum_{n\geq 0} \sum_{C_{\lambda,\mu}\subseteq B_n} \frac{B_{\lambda,\mu}(t)}{(1-t)^{n+1}} \prod_i x_i^{m_i(\lambda)} y_i^{m_i(\mu)}. \qquad (4.2.12)$$

We now examine the coefficient of t^k in the second line of (4.2.6), which can then be written as

$$
\begin{aligned}
&\frac{1}{1-x_1}\prod_{i\geq 1}\left(\frac{1+y_i}{1-x_i}\right)^{N(2k-1,2i)} \\
&=\left(\prod_{i\geq 1}(1+y_i)^{N(2k-1,2i)}\prod_{i\geq 2}(1-x_i)^{-N(2k-1,2i)}\right)(1-x_1)^{-N(2k-1,2)-1}.
\end{aligned} \tag{4.2.13}
$$

Note that $k = 1$ corresponds to permutations with no descents, which are the identity elements in $\{B_n\}_{n\geq 1}$. For our result, we will not be interested in this term.

For $k \geq 2$, we use Newton's generalized binomial theorem

$$(1-x)^{-m} = \sum_{j\geq 0} x^j \binom{m+j-1}{j}$$

for $m > 0$ to find that the coefficient of the product $\prod_i x_i^{m_i(\lambda)} y_i^{m_i(\mu)}$ in (4.2.13) is

$$\left(\prod_{i\geq 1}\binom{N(2k-1,2i)}{m_i(\mu)}\prod_{i\geq 2}\binom{N(2k-1,2i)+m_i(\lambda)-1}{m_i(\lambda)}\right)\cdot\binom{N(2k-1,2)+1+m_1(\lambda)-1}{m_1(\lambda)}.$$

This must equal the coefficient of $t^k \prod_i x_i^{m_i(\lambda)} y_i^{m_i(\mu)}$ in (4.2.12), so (4.2.11) follows. □

Corollary 4.2.5. *The number of permutations $\omega \in B_n$ that are of cycle type (λ, μ) and have $d-1$ descents is*

$$\sum_{k=1}^{d}(-1)^{d-k}\binom{n+1}{d-k}\binom{N(2k-1,2)+m_1(\lambda)}{m_1(\lambda)}\prod_{i\geq 2}\binom{N(2k-1,2i)+m_i(\lambda)-1}{m_i(\lambda)}\prod_{i\geq 1}\binom{N(2k-1,2i)}{m_i(\mu)}.$$

Proof. Starting with Lemma 4.2.4, multiply by $(1-t)^{n+1}$ to obtain the generating function $B_{\lambda,\mu}(t)$. The result follows by extracting the coefficient of t^d. □

We next derive an elegant analog of a result of Fulman [Ful98, Proof of Theorem 2], which will relate $B_{\lambda,\mu}(t)$ and

$$B_n(t) = \sum_{\omega\in B_n} t^{\mathrm{des}_{B_n}(\omega)+1}. \tag{4.2.14}$$

We start by describing the tools that we need. First, Theorem 2.2.11 can be restated as

$$\frac{B_n(t)}{(1-t)^{n+1}} = \sum_{k\geq 1}(2k-1)^n t^k. \tag{4.2.15}$$

Additionally, denote $\Delta^2(\lambda,\mu) = m_1(\lambda)^2 - m_1(\mu)^2$. Finally, let $s_n^{(i)}$ be the Stirling number of the first kind, whose absolute value is the number of permutations in $\mathfrak{S}_n$ with i disjoint cycles. The following generating functions are well-known, e.g., see [Sta99, Chapter 1]:

$$\begin{aligned} \sum_{i=1}^{n} s_n^{(i)} y^i &= y(y-1)\cdots(y-n+1) = n!\binom{y}{n}, \\ \sum_{i=1}^{n} |s_n^{(i)}| y^i &= y(y+1)\cdots(y+n-1) = n!\binom{y+n-1}{n}. \end{aligned} \tag{4.2.16}$$

Lemma 4.2.6. *Let $C_{\lambda,\mu}$ be a conjugacy class of B_n where $(\lambda,\mu)\neq((1^n),\emptyset)$. Then*

$$\frac{B_{\lambda,\mu}(t)}{|C_{\lambda,\mu}|} = \frac{B_n(t)}{2^n n!} + \frac{1-t}{2n}\frac{B_{n-1}(t)}{2^{n-1}(n-1)!}\Delta^2(\lambda,\mu) + (1-t)^2 g(t), \tag{4.2.17}$$

where $g(t)$ is some polynomial in t.

Proof. From Theorem 2.2.11, Lemma 2.2.15, and Lemma 4.2.4,

$$\begin{aligned} \frac{B_{\lambda,\mu}(t)}{|C_{\lambda,\mu}|} = (1-t)^{n+1}\frac{z_{\lambda,\mu}}{2^n n!}\Bigg[\sum_{k\geq 2} t^k \qquad\qquad\qquad\qquad\qquad\qquad (4.2.18)\\ \cdot\left(\prod_{i\geq 1}\binom{N(2k-1,2i)}{m_i(\mu)}\prod_{i\geq 2}\binom{N(2k-1,2i)+m_i(\lambda)-1}{m_i(\lambda)}\right)\cdot\binom{N(2k-1,2)+m_1(\lambda)}{m_1(\lambda)}\Bigg]. \end{aligned}$$

Using (4.2.16), the coefficient of t^k for $k\geq 2$ can be rewritten as

$$\begin{aligned} &\left(\prod_{i\geq 1}\frac{\sum_{b=1}^{m_i(\mu)} s_{m_i(\mu)}^{(b)} N(2k-1,2i)^b}{m_i(\mu)!}\prod_{i\geq 2}\frac{\sum_{a=1}^{m_i(\lambda)} |s_{m_i(\lambda)}^{(a)}| N(2k-1,2i)^a}{m_i(\lambda)!}\right) \\ &\cdot\left[\frac{1}{m_1(\lambda)!}\sum_{a=1}^{m_1(\lambda)} |s_{m_1(\lambda)}^{(a)}|\left(\frac{(2k-1)+1}{2}\right)^a\right], \end{aligned} \tag{4.2.19}$$

the last sum arising from the fact that $N(2k-1,2)+1=k$. Combined with the definition of $N(2k-1,2i)$ given in (4.2.3), we may view the coefficient of t^k for any $k \geq 2$ as a polynomial in $(2k-1)$.

For each $k \geq 2$, the largest power of $(2k-1)$ in $B_{\lambda,\mu}(t)$ is obtained by taking the largest power of $(2k-1)$ in each factor of (4.2.19). From (4.2.3), $N(2k-1,2i)$ is a polynomial in $(2k-1)$ with leading term $\frac{(2k-1)^i}{2i}$. It follows that the largest power of $2k-1$ occurs when we respectively take the summands corresponding to $b = m_i(\mu)$, $a = m_i(\lambda)$, and $a = m_1(\lambda)$ in the three parts of (4.2.19), yielding

$$\prod_{i\geq 1} \frac{(N(2k-1,2i))^{m_i(\lambda)+m_i(\mu)}}{m_i(\lambda)!m_i(\mu)!} \sim \prod_{i\geq 1} \frac{\left(\frac{(2k-1)^i}{2i}\right)^{m_i(\lambda)+m_i(\mu)}}{m_i(\lambda)!m_i(\mu)!}. \tag{4.2.20}$$

Collecting terms with the highest power of $(2k-1)$ in $B_{\lambda,\mu}(t)$ for each $k \geq 2$ then results in

$$\begin{aligned} &\frac{z_{\lambda,\mu}}{2^n n!}(1-t)^{n+1} \sum_{k\geq 2} t^k \prod_{i\geq 1} \frac{\left(\frac{(2k-1)^i}{2i}\right)^{m_i(\lambda)+m_i(\mu)}}{m_i(\lambda)!m_i(\mu)!} \\ &= \frac{1}{2^n n!}(1-t)^{n+1} \sum_{k\geq 2} t^k (2k-1)^n \\ &= \frac{B_n(t) - t(1-t)^{n+1}}{2^n n!}. \end{aligned} \tag{4.2.21}$$

Observe that in (4.2.21), the highest power of $(2k-1)$ for each k is given by $(2k-1)^n$, so we next consider the coefficient of $(2k-1)^{n-1}$ in (4.2.20). We note that for $i > 1$, setting $b \neq m_i(\mu)$ in the first summation of (4.2.19), or $a \neq m_i(\lambda)$ in the second summation of (4.2.19), or $d \neq 1$ in (4.2.3) will result in a power of $(2k-1)$ strictly less than $n-1$ in (4.2.20). So for $i > 1$, we still respectively set $b = m_i(\mu), a = m_i(\lambda)$ in each of the summations in (4.2.19) and still take $d = 1$ in (4.2.3) to identify any terms involving $(2k-1)^{n-1}$. This term must come from the $i = 1$ term in (4.2.19), so we turn our attention to

$$\frac{\sum_{a=1}^{m_1(\lambda)} |s^{(a)}_{m_1(\lambda)}| \frac{1}{2^a}((2k-1)+1)^a}{m_1(\lambda)!} \cdot \frac{\sum_{b=1}^{m_1(\mu)} s^{(b)}_{m_1(\mu)} \frac{1}{2^b}((2k-1)-1)^b}{m_1(\mu)!},$$

and find the coefficient of $(2k-1)^{n-1}$ is

$$\left(|s_{m_1(\lambda)}^{(m_1(\lambda))}|\cdot s_{m_1(\mu)}^{(m_1(\mu))}\cdot\left(\frac{m_1(\lambda)-m_1(\mu)}{2}\right)+|s_{m_1(\lambda)}^{(m_1(\lambda)-1)}|\cdot s_{m_1(\mu)}^{(m_1(\mu))}+|s_{m_1(\lambda)}^{(m_1(\lambda))}|\cdot s_{m_1(\mu)}^{(m_1(\mu)-1)}\right)$$
$$\cdot\frac{1}{m_1(\lambda)!m_1(\mu)!}\left(\frac{2k-1}{2}\right)^{m_1(\lambda)+m_1(\mu)-1}$$
$$=\left(\frac{m_1(\lambda)-m_1(\mu)}{2}+\binom{m_1(\lambda)}{2}-\binom{m_1(\mu)}{2}\right)\frac{1}{m_1(\lambda)!m_1(\mu)!}\left(\frac{2k-1}{2}\right)^{m_1(\lambda)+m_1(\mu)-1}$$
$$=\frac{1}{2m_1(\lambda)!m_1(\mu)!}\left(m_1(\lambda)^2-m_1(\mu)^2\right)\left(\frac{2k-1}{2}\right)^{m_1(\lambda)+m_1(\mu)-1}.$$

Hence the terms involving $(2k-1)^{n-1}$ can be expressed as

$$\begin{aligned}&\left(m_1(\lambda)^2-m_1(\mu)^2\right)\frac{1}{2^n n!}(1-t)^{n+1}\sum_{k\geq 2}t^k(2k-1)^{n-1}\\ &=\frac{1-t}{2n}\frac{B_{n-1}(t)-t(1-t)^n}{2^{n-1}(n-1)!}\Delta^2(\lambda,\mu)\\ &=\frac{1-t}{2n}\frac{B_{n-1}(t)}{2^{n-1}(n-1)!}\Delta^2(\lambda,\mu)-\frac{t(1-t)^{n+1}}{2^n n!}\Delta^2(\lambda,\mu).\end{aligned}\tag{4.2.22}$$

To conclude, we express $B_{\lambda,\mu}(t)/|C_{\lambda,\mu}|$ as a sum of (4.2.21), (4.2.22), and the remaining terms in (4.2.18) resulting from lower order terms of the form $t^k(2k-1)^{k-j}$ for $2\leq j\leq k$. Equations (4.2.21) and (4.2.22) contain the first two terms in our claimed result (4.2.17), and the other terms in (4.2.21) and (4.2.22) can be expressed as $(1-t)^2 f_1(t)$ for some polynomial $f_1(t)$. For the lower order terms of the form $t^k(2k-1)^{k-j}$ with $2\leq j\leq k$, we combine these into

$$(1-t)^{n+1}\frac{z_{\lambda,\mu}}{2^n n!}\sum_{k\geq 2}t^k h(2k-1),\tag{4.2.23}$$

where $h(2k-1)$ is a polynomial in $2k-1$ of degree at most $n-2$ obtained from removing the $(2k-1)^n$ and $(2k-1)^{n-1}$ terms from (4.2.19). For each $2\leq j\leq k$, Theorem 2.2.11 implies

$$\sum_{k\geq 2}t^k(2k-1)^{k-j}=-t+\sum_{k\geq 1}t^k(2k-1)^{k-j}=-t+\frac{B_{n-j}(t)}{(1-t)^{n+1-j}}.$$

Since $j \geq 2$ and $B_{n-j}(t)$ is a polynomial in t, multiplying the above expression by $(1-t)^{n+1}$ results in a polynomial where $(1-t)^2$ can be factored out. Hence, each term in (4.2.23) can be expressed as $(1-t)^2$ multiplied by some polynomial in t, so the entire expression (4.2.23) can be expressed as $(1-t)^2 f_2(t)$ for some polynomial $f_2(t)$. The result now follows by letting $g(t) = f_1(t) + f_2(t)$. □

In the special case of no short cycles, we obtain the following variation of Lemma 4.2.6. This variation will be used to establish an analog of Theorem 1.2.11 and Corollary 1.2.12 for des_{B_n}.

Corollary 4.2.7. *Let $C_{\lambda,\mu}$ be a conjugacy class of B_n that contains no cycles of length $1, 2, \ldots, 2\ell$. Then*

$$\frac{B_{\lambda,\mu}(t)}{|C_{\lambda,\mu}|} = \frac{B_n(t)}{2^n n!} + (1-t)^{\ell+1} h(t),$$

where $h(t)$ is some polynomial in t.

Proof. As in the proof of Lemma 4.2.6,

$$\frac{B_{\lambda,\mu}(t)}{|C_{\lambda,\mu}|} = \frac{1}{2^n n!}(1-t)^{n+1} \prod_{i>2\ell} (2i)^{m_i(\lambda)+m_i(\mu)} m_i(\lambda)! m_i(\mu)! \cdot \sum_{k\geq 1} t^k \prod_{i>2\ell} \binom{m_i(\lambda) + N(2k-1, 2i) - 1}{m_i(\lambda)} \binom{N(2k-1, 2i)}{m_i(\mu)}.$$

We note that, as a polynomial in $(2k-1)$, the leading term of

$$\prod_{i>2\ell} \binom{m_i(\lambda) + N(2k-1, 2i) - 1}{m_i(\lambda)} \binom{N(2k-1, 2i)}{m_i(\mu)}$$

is $(2k-1)^n$ as in the general case, but the second highest-degree term is at most $(2k-1)^{n-\ell-1}$ under the long cycle assumption. This is because a lower order term must have either some $a \neq m_i(\lambda)$ or some $b \neq m_i(\mu)$ in (4.2.19), or $d \neq 1$ in (4.2.3). If some $a \neq m_i(\lambda)$ or some $b \neq m_i(\mu)$, then the power of $2k-1$ from such a term is at most $n - i < n - 2\ell$. If some

$d \neq 1$, then the power of $2k-1$ from such a term is at most $n-(i-i/3) < n-\ell$. We then proceed similarly as in the proof of Lemma 4.2.6. $\square$

Theorem 4.2.8. *Let $C_{\lambda,\mu}$ be a conjugacy class of B_n. If $C_{\lambda,\mu}$ contains no cycles of length $1, 2, \ldots, 2k$, then*

$$\mathbb{E}_{\lambda,\mu}[\mathrm{des}_{B_n}^k] = \mathbb{E}_{B_n}[\mathrm{des}_{B_n}^k].$$

Proof. We start with the result from Corollary 4.2.7 given by

$$\frac{B_{\lambda,\mu}(t)}{|C_{\lambda,\mu}|} = \frac{B_n(t)}{2^n n!} + (1-t)^{k+1} g(t). \tag{4.2.24}$$

Observe that $B_{\lambda,\mu}(t)/|C_{\lambda,\mu}|$ and $B_n(t)/(2^n n!)$ are the probability generating functions of $\mathrm{des}_{B_n}+1$ on $C_{\lambda,\mu}$ and B_n, respectively, so the coefficient of x^d in each generating function is the probability that $\mathrm{des}_{B_n}+1$ takes value d. Letting $t \cdot \frac{d}{dt}$ be the operator that differentiates with respect to t and then multiplies the result by t, the j-th moments of $\mathrm{des}_{B_n}+1$ on $C_{\lambda,\mu}$ and B_n can respectively be obtained by applying $t \cdot \frac{d}{dt}$ to $B_{\lambda,\mu}(t)/|C_{\lambda,\mu}|$ and $B_n(t)/(2^n n!)$ a total of j times and then evaluating the result at $t=1$. If we apply these operations to both sides of (4.2.24), then whenever $j \leq k$, the second term on the right side becomes 0. Hence, $\mathbb{E}_{\lambda,\mu}[(\mathrm{des}_{B_n}+1)^j] = \mathbb{E}_{B_n}[(\mathrm{des}_{B_n}+1)^j]$ for all $j \leq k$. It is then straightforward to use this to show that $\mathbb{E}_{\lambda,\mu}[\mathrm{des}_{B_n}^j] = \mathbb{E}_{B_n}[\mathrm{des}_{B_n}^j]$ for all $j \leq k$. $\square$

Corollary 4.2.9. *For every $n \geq 1$, let C_{λ_n,μ_n} be a conjugacy class of B_n. Suppose that for all i, the number of cycles of length i in λ_n and μ_n approaches 0 as $n \to \infty$. Then for sufficiently large n, des_{B_n} has mean $\frac{n}{2}$ and variance $\frac{n+1}{12}$ on C_{λ_n,μ_n}, and as $n \to \infty$, the random variable $(\mathrm{des}_{B_n}-n/2)/\sqrt{(n+1)/12}$ converges in distribution to the standard normal distribution.*

Proof. The mean and variance follow from applying Theorem 4.2.8 and Theorem 2.2.10 on the first two moments of des_{B_n} with the hypothesis that there are no cycles of length $1, 2, 3$, or 4 when n is sufficiently large. For the asymptotic behavior, fix k and expand

$$\left(\frac{\mathrm{des}_{B_n}-n/2}{\sqrt{(n+1)/12}}\right)^k = \frac{1}{((n+1)/12)^{k/2}} \sum_{i=0}^{k} \binom{k}{i} \left(\frac{n}{2}\right)^{k-i} \mathrm{des}_{B_n}^i .$$

Linearity of expectation then implies that the k-th moment of $(\text{des}_{B_n} - n/2)/\sqrt{(n+1)/12}$ on $\mathfrak{S}_n$ or C_{λ_n,μ_n} can be expressed as a linear combination of the first k moments of des_{B_n} on those respective sets. Applying Theorem 4.2.8 with the hypothesis that there are no cycles of length $1, 2, \ldots, 2k$ when n is sufficiently large, we conclude that

$$\lim_{n\to\infty} \mathbb{E}_{B_n}\left[\left(\frac{\text{des}_{B_n} - n/2}{\sqrt{(n+1)/12}}\right)^k\right] = \lim_{n\to\infty} \mathbb{E}_{\lambda_n,\mu_n}\left[\left(\frac{\text{des}_{B_n} - n/2}{\sqrt{(n+1)/12}}\right)^k\right].$$

The result now follows from the Method of Moments and Theorem 2.2.10. □

4.2.2 Descents and major index in $\mathfrak{S}_{n,r}$

We now consider $\text{des}_{n,r}$ and $\text{maj}_{n,r}$ on $\mathfrak{S}_{n,r}$ as defined in (2.2.3). Throughout this section, we define X_i to be the indicator function for a descent at position i,

$$X_i(\omega, \tau) = \begin{cases} 1 & \text{if } i \in \text{Des}(\omega, \tau) \\ 0 & \text{otherwise.} \end{cases}$$

The descent and major index statistics can then be expressed as

$$\text{des}_{n,r} = \sum_{i=1}^{n} X_i \quad \text{and} \quad \text{maj}_{n,r} = \sum_{i=1}^{n-1} i \cdot X_i.$$

Observe that the above decompositions also allow us to decompose the k-th powers of the descent and major index statistics in terms of $X_1, \ldots, X_n$ as

$$\text{des}_{n,r}^k = \sum_{a_1,\ldots,a_k \in [n]} X_{a_1} \ldots X_{a_k} \quad \text{and} \quad \text{maj}_{n,r}^k = \sum_{a_1,\ldots,a_k \in [n-1]} a_1 \ldots a_k X_{a_1} \ldots X_{a_k}.$$

Note that the $a_1, \ldots, a_k$ need not be distinct. Since expectation is linear, an understanding of the mean of $X_{a_1} \ldots X_{a_k}$ on $\mathfrak{S}_{n,r}$ or $C_{\boldsymbol{\lambda}}$ informs us of the k-th moments of $\text{des}_{n,r}$ and $\text{maj}_{n,r}$ on these sets. We start with the following definitions based on [Ful98].

Definition 4.2.10. The *Young subgroup* generated by $a_1, \ldots, a_k \in [n]$ is the subgroup of $\mathfrak{S}_n$ generated by the adjacent transpositions $\{(a_1, a_1+1), \ldots, (a_k, a_k+1)\} \setminus \{(n, n+1)\}$.

Definition 4.2.11. Let J be the Young subgroup of $\mathfrak{S}_n$ generated by $a_1, \ldots, a_k \in [n]$. The *blocks* induced by $a_1, \ldots, a_k \in [n]$ are the equivalence classes $\mathscr{B}_1, \ldots, \mathscr{B}_t \subseteq [n]$ generated by the following property: $i, j \in [n]$ are in the same equivalence class if some $\omega \in J$ maps i to j. Observe that one can equivalently express

$$J = \mathfrak{S}_{\mathscr{B}_1} \times \ldots \times \mathfrak{S}_{\mathscr{B}_t},$$

where $\mathfrak{S}_{\mathscr{B}_i}$ is the group of permutations on the elements in $\mathscr{B}_i$.

Example 4.2.12. The blocks induced by $1, 2, 4, 7 \in [8]$ are $\mathscr{B}_1 = \{1, 2, 3\}$, $\mathscr{B}_2 = \{4, 5\}$, $\mathscr{B}_3 = \{6\}$, and $\mathscr{B}_4 = \{7, 8\}$. Note that the blocks induced by $1, 2, 4, 7, 8 \in [8]$ will be the same. Observe that the number of blocks t depends on the specific choices of $a_1, \ldots, a_k$ rather than only k.

Fulman shows in [Ful98, Proof of Theorem 3] that when the blocks induced by $a_1, \ldots, a_k \in [n-1]$ are $\mathscr{B}_1, \ldots, \mathscr{B}_t$,

$$\mathbb{E}_{\mathfrak{S}_n}[X_{a_1} X_{a_2} \ldots X_{a_k}] = \prod_{i=1}^{t} \frac{1}{|\mathscr{B}_i|!}. \tag{4.2.25}$$

In $\mathfrak{S}_{n,r}$, we will derive the corresponding formulas for $\mathbb{E}_{\mathfrak{S}_n}[X_{a_1} X_{a_2} \ldots X_{a_k}]$, and there will be two cases depending on whether or not $a_1, \ldots, a_k$ contains n. When this is not the case, we show that Equation (4.2.25) translates directly to $\mathfrak{S}_{n,r}$.

Lemma 4.2.13. *Let $a_1, \ldots, a_k \in [n-1]$ with induced blocks $\mathscr{B}_1, \ldots, \mathscr{B}_t$. Then*

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1} X_{a_2} \ldots X_{a_k}] = \mathbb{E}_{\mathfrak{S}_n}[X_{a_1} X_{a_2} \ldots X_{a_k}] = \prod_{i=1}^{t} \frac{1}{|\mathscr{B}_i|!}. \tag{4.2.26}$$

Proof. Let $\mathfrak{S}_n$ act on $\mathfrak{S}_{n,r}$ by permuting entries in the one-line notation. This partitions $\mathfrak{S}_{n,r}$ into orbits based on the elements that appear in the one-line notation. Each orbit Ω_c

can be indexed by $c = (c_1, \dots, c_n)$, where $c_i \in \mathbb{Z}_r$ is the color of element i in the one-line notation. Let $f_c : \{i^{c_i}\}_{i=1}^n \to [n]$ be the unique order-preserving bijection from $\{i^{c_i}\}_{i=1}^n$ with the ordering in (2.2.2) to $[n]$ with the usual ordering. This induces a bijection $F_c : \Omega_c \to \mathfrak{S}_n$ that on the one-line notation is given by

$$F_c([(\omega,\tau)(1^0),(\omega,\tau)(2^0),\dots,(\omega,\tau)(n^0)]) = [f_c((\omega,\tau)(1^0)), f_c((\omega,\tau)(2^0)),\dots,f_c((\omega,\tau)(n^0))].$$

Furthermore, since f_c is order-preserving, F_c preserves descents. Therefore, for all $(\omega,\tau) \in \Omega_c$, we have that $X_{a_1}X_{a_2}\dots X_{a_k}(\omega,\tau) = X_{a_1}X_{a_2}\dots X_{a_k}(F_c(\omega,\tau))$. As F_c is a bijection, this implies that

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1}X_{a_2}\dots X_{a_k} \mid \Omega_c] = \mathbb{E}_{\mathfrak{S}_n}[X_{a_1}X_{a_2}\dots X_{a_k}]. \tag{4.2.27}$$

Equation (4.2.27) holds for every Ω_c, so we use the Law of Total Expectation to conclude

$$\begin{aligned}\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1}X_{a_2}\dots X_{a_k}] &= \sum_c \Pr_{\mathfrak{S}_{n,r}}[\Omega_c]\cdot\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1}X_{a_2}\dots X_{a_k} \mid \Omega_c]\\ &= \sum_c \Pr_{\mathfrak{S}_{n,r}}[\Omega_c]\cdot\mathbb{E}_{\mathfrak{S}_n}[X_{a_1}X_{a_2}\dots X_{a_k}]\\ &= \mathbb{E}_{\mathfrak{S}_n}[X_{a_1}X_{a_2}\dots X_{a_k}].\end{aligned}$$

The result now follows from (4.2.25). □

It remains now to consider products involving X_n. We start with the case of consecutive indices $X_{m+1}\dots X_n$ containing n. Since $X_n(\omega,\tau) = 1$ occurs precisely when $\tau(n) \neq 0$, observe that $X_{m+1}\dots X_n(\omega,\tau) = 1$ implies $\tau(i) \neq 0$ for all $i \geq m$. Conditioning on this property, we compute the expectation of $X_{m+1}\dots X_n$ on $\mathfrak{S}_{n,r}$.

Lemma 4.2.14. *For any $1 \leq m < n$, the following holds:*

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{m+1}\dots X_n] = \left(\frac{r-1}{r}\right)^{n-m}\cdot\frac{1}{(n-m)!}. \tag{4.2.28}$$

Proof. We first express

$$\begin{aligned}\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{m+1}\dots X_n] &= \Pr_{\mathfrak{S}_{n,r}}[X_{m+1}\dots X_n = 1]\\ &= \Pr_{\mathfrak{S}_{n,r}}[\{\tau(i)\neq 0\,\forall\, i>m\}\cap\{X_{m+1}\dots X_{n-1}=1\}]\\ &= \Pr_{\mathfrak{S}_{n,r}}[\tau(i)\neq 0\,\forall\, i>m]\cdot \Pr_{\mathfrak{S}_{n,r}}[X_{m+1}\dots X_{n-1}=1 \mid \tau(i)\neq 0\,\forall\, i>m].\end{aligned}$$

The first term is equal to $((r-1)/r)^{n-m}$, so it suffices to show the second term is $1/(n-m)!$. For this, we let $\mathfrak{S}_{n-m}$ act on the set $\{(\omega,\tau)\in\mathfrak{S}_{n,r}\mid \tau(i)\neq 0\,\forall\, i>m\}$ by permuting the last $n-m$ entries in the one-line notation. Under this action, each orbit has size $(n-m)!$, and exactly one element in each orbit has these last $n-m$ elements in descending order. The same argument as in Lemma 4.2.13 shows then that

$$\Pr_{\mathfrak{S}_{n,r}}[X_{m+1}\dots X_{n-1}=1 \mid \tau(i)\neq 0\,\forall\, i>m] = \frac{1}{(n-m)!}. \qquad \square$$

Corollary 4.2.15. *For any positive integers m and n,*

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[X_1X_2\dots X_n] = \mathbb{E}_{\mathfrak{S}_{m+n,r}}[X_{m+1}X_{m+2}\dots X_{m+n}].$$

Finally, we consider the expectation of arbitrary products of the X_i statistics that contain X_n. Our approach is to again use an action by a symmetric group of appropriate size.

Lemma 4.2.16. *For any $a_1,\dots,a_j\in[m-1]\subseteq[n]$,*

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1}\dots X_{a_j}X_{m+1}X_{m+2}\dots X_n] = \mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1}\dots X_{a_j}]\cdot\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{m+1}X_{m+2}\dots X_n].$$

Proof. Express

$$\begin{aligned}&\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1}\dots X_{a_j}X_{m+1}X_{m+2}\dots X_n]\\ &= \Pr_{\mathfrak{S}_{n,r}}[X_{a_1}\dots X_{a_j}X_{m+1}X_{m+2}\dots X_n=1] \qquad (4.2.29)\\ &= \Pr_{\mathfrak{S}_{n,r}}[X_{a_1}\dots X_{a_j}=1]\cdot\Pr_{\mathfrak{S}_{n,r}}[X_{m+1}\dots X_n=1 \mid X_{a_1}\dots X_{a_j}=1]\end{aligned}$$

The first term is $\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1} \dots X_{a_j}]$, and the group action and argument from Lemma 4.2.14 shows that

$$\Pr_{\mathfrak{S}_{n,r}}[X_{m+1} \dots X_n = 1 \mid X_{a_1} \dots X_{a_k} = 1] = \left(\frac{r-1}{r}\right)^{n-m} \cdot \frac{1}{(n-m)!}. \qquad \square$$

Corollary 4.2.17. *Consider any $a_1, \dots, a_k \in [n]$ with induced blocks $\mathscr{B}_1, \dots, \mathscr{B}_t$, where $\mathscr{B}_t$ contains n. If $n \in \{a_1, \dots, a_k\}$, then*

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1} \dots X_{a_k}] = \left(\frac{r-1}{r}\right)^{|\mathscr{B}_t|} \cdot \prod_{i=1}^{t} \frac{1}{|\mathscr{B}_i|!}.$$

Proof. Since $n \in \{a_1, \dots, a_k\}$, we can express $X_{a_1} \dots X_{a_k}$ as $X_{a_1} \dots X_{a_j} X_{m+1} X_{m+2} \dots X_n$, where $a_1, \dots, a_j \in [m-1]$. By Lemma 4.2.16,

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1} \dots X_{a_k}] = \mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1} \dots X_{a_j}] \cdot \mathbb{E}_{\mathfrak{S}_{n,r}}[X_{m+1} \dots X_n].$$

The result follows by applying Lemma 4.2.13 and Lemma 4.2.14. $\square$

We now consider $\mathbb{E}_{\boldsymbol{\lambda}}[X_{a_1} \dots X_{a_k}]$ on $C_{\boldsymbol{\lambda}}$ without cycles of length $1, 2, \dots, 2k$ and establish analogs of Lemma 4.2.13 and Corollary 4.2.17. Many of our techniques for proving Lemma 4.2.13 and Corollary 4.2.17 involved group actions where orbits have exactly one element with $X_{a_1} \dots X_{a_k}(\omega, \tau) = 1$, and we will define an appropriate action on $C_{\boldsymbol{\lambda}}$ with the same property.

Fix $a_1, \dots, a_k \in [n]$, let $\mathscr{B}_1, \dots, \mathscr{B}_t \subseteq [n]$ be blocks induced by $a_1, \dots, a_k$, and let $J = \mathfrak{S}_{\mathscr{B}_1} \times \dots \times \mathfrak{S}_{\mathscr{B}_t}$ be the Young subgroup of $\mathfrak{S}_n$ generated by $a_1, \dots, a_k$. Define an action of J on $\mathfrak{S}_{n,r}$ as follows: for all $\pi \in J$ and $(\omega, \tau) \in \mathfrak{S}_{n,r}$,

$$\pi \cdot (\omega, \tau) = (\pi, \mathbf{0})(\omega, \tau)(\pi, \mathbf{0})^{-1}, \tag{4.2.30}$$

where $\mathbf{0}$ is the zero coloring. Alternatively, this is the conjugation action of J on $\mathfrak{S}_{n,r}$ after identifying J with the subgroup $J \times \mathbf{0}$. The following result describes orbits under the action given in (4.2.30).

Lemma 4.2.18. *Let $(\omega, \tau) \in \mathfrak{S}_{n,r}$. Let $\pi \in \mathfrak{S}_n$ and $\mathbf{0}$ be the zero coloring. If $(i_1^{c_1}, i_2^{c_2}, \ldots, i_\ell^{c_\ell})$ is a cycle in (ω, τ), then*

$$(\pi, \mathbf{0})(i_1^{c_1}, i_2^{c_2}, \ldots, i_\ell^{c_\ell})(\pi, \mathbf{0})^{-1} = (\pi(i_1)^{c_1}, \pi(i_2)^{c_2}, \ldots, \pi(i_j)^{c_\ell}).$$

Consequently, $(\pi(i_1)^{c_1}, \pi(i_2)^{c_2}, \ldots, \pi(i_\ell)^{c_\ell})$ is a cycle in $(\pi, \mathbf{0})(\omega, \tau)(\pi, \mathbf{0})^{-1}$.

Proof. For any i_j, we consider the image of $\pi(i_j)^0$ under $(\pi, \mathbf{0})^{-1}(\omega, \tau)(\pi, \mathbf{0})$:

$$\begin{aligned}(\pi, \mathbf{0})(\omega, \tau)(\pi, \mathbf{0})^{-1}(\pi(i_j)^0) &= (\pi, \mathbf{0})(\omega, \tau)(i_j^0) \\ &= (\pi, \mathbf{0})(i_{j+1}^{c_{j+1}}) \\ &= \pi(i_{j+1})^{c_{j+1}},\end{aligned}$$

where in the case of $j = \ell$, we replace $j + 1$ with 1. Hence, $\pi(i_{j+1})^{c_{j+1}}$ follows $\pi(i_j)^{c_j}$ in the cycle notation as claimed. □

Lemma 4.2.18 implies that the orbit of any $(\omega, \tau) \in \mathfrak{S}_{n,r}$ under the action in (4.2.30) consists of colored permutations that can be obtained by starting with the cycle notation of (ω, τ) and permuting elements within each block $\mathscr{B}_1, \ldots, \mathscr{B}_t$ without changing the location of colors. On conjugacy classes $C_{\boldsymbol{\lambda}}$ without cycles of length $1, 2, \ldots, 2k$, we will show that these orbits are particularly well-behaved.

Lemma 4.2.19. *Let $a_1, \ldots, a_k \in [n-1]$ with induced blocks $\mathscr{B}_1, \ldots, \mathscr{B}_t$, and let $J = \mathfrak{S}_{\mathscr{B}_1} \times \ldots \times \mathfrak{S}_{\mathscr{B}_t}$ act on a conjugacy class $C_{\boldsymbol{\lambda}}$ of $\mathfrak{S}_{n,r}$ by (4.2.30). If $C_{\boldsymbol{\lambda}}$ contains no cycles of length $1, 2, \ldots, 2k$, then each orbit under this action has size $|J| = \prod_{i=1}^{t} |\mathscr{B}_t|!$. Furthermore, there is a unique element in each orbit that has descents at $a_1, \ldots, a_k$.*

To prove Lemma 4.2.19, we will define an algorithm that identifies necessary conditions for descents at $a_1, \ldots, a_k$ to appear and replaces elements in each block $\mathscr{B}_1, \ldots, \mathscr{B}_t$ appropriately.

This algorithm will generalize one used by Fulman in [Ful98, Proof of Theorem 3]. Since our algorithm is very technical, we will start with an example.

Example 4.2.20. Consider indices $1, 2, 4, 5 \in [9]$ and $(\omega, \tau) = (1^0 3^1 8^2 5^2 2^0 7^0 4^1 9^0 6^2) \in \mathfrak{S}_{9,3}$. The blocks induced by $1, 2, 4, 5$ are $\mathscr{B}_1 = \{1, 2, 3\}$, $\mathscr{B}_2 = \{4, 5, 6\}$, $\mathscr{B}_3 = \{7\}$, $\mathscr{B}_4 = \{8\}$, and $\mathscr{B}_5 = \{9\}$. We wish to find an element in the orbit of (ω, τ) under the action in (4.2.30) that has descents at positions $1, 2, 4$, and 5, so we start by replacing elements in the cycle notation with the smallest number in its corresponding block, resulting in

$$(1^0 1^1 8^2 4^2 1^0 7^0 4^1 9^0 4^2). \tag{4.2.31}$$

We now must find an appropriate way to replace the instances of 1 and 4 with elements in the same block. Ignoring colors for the moment, we observe that the elements $7, 8$, and 9 appear exactly once, and they are respectively preceded by $1, 1$, and 4. Both 1 and 4 appear multiple times in (4.2.31), so we can try to use the information involving $7, 8$, or 9 to change this. For simplicity, we choose the largest element 9, which is preceded by a 4 in (4.2.31). The elements directly after appearances of 4's are 1^0, 9^0, and 1^0. Regardless of how these two appearances of 1 are replaced with other elements in $\mathcal{B}_1 = \{1, 2, 3\}$, the element 9^0 will still be the largest. Then for descents at positions 4 and 5 to occur, the element 4^0 must map to 9^0. Using this information, we next consider

$$(1^0 1^1 8^2 5^2 1^0 7^0 4^1 9^0 5^2), \tag{4.2.32}$$

as we have determined the image of 4^0, but we have not determined the images of 5^0 or 6^0. Observe that since 9 appeared exactly once in (4.2.31), the element preceding it in (4.2.32) now appears exactly once.

Continuing in this manner, 8 is now the largest element that appears only once but whose preceding element 1 in (4.2.32) appears multiple times. The elements that follow these appearances of 1 are 1^1, 8^2, and 7^0. We wish for descents at positions 1 and 2, and the unique option for this is $(2^0 1^1 8^2 5^2 3^0 7^0 4^1 9^0 5^2)$. Finally, we consider repeated instances of 5 to obtain $(2^0 1^1 8^2 5^2 3^0 7^0 4^1 9^0 6^2)$. Observe that this is in the orbit of (ω, τ) under the action (4.2.30), and it has descents at positions $1, 2, 4$, and 5.

We now give an algorithm that formalizes the example above. We then use this to establish Lemma 4.2.19.

Input: $(\omega, \tau) \in \mathfrak{S}_{n,r}$ with no cycles of length $1, 2, \ldots, 2k$; indices $a_1, \ldots, a_k \in [n]$
Output: a colored permutation $(\omega', \tau') \in \mathfrak{S}_{n,r}$ in the orbit of (ω, τ) under (4.2.30)

$\mathscr{B}_1, \ldots, \mathscr{B}_t :=$ blocks induced by $a_1, \ldots, a_k$
$\sigma_1, \ldots, \sigma_m :=$ cycles of (ω, τ)
$\sigma'_1, \ldots \sigma'_m :=$ cycles obtained by starting with $\sigma_1, \ldots, \sigma_m$ and replacing each $i \in [n]$ with the smallest number from the block that contains it
while $\sigma'_1, \ldots, \sigma'_m$ *contains repeated integers from* $[n]$ **do**
  $S :=$ subset of $[n]$ consisting of elements that appear exactly once in $\sigma'_1, \ldots, \sigma'_m$
  $j :=$ largest element in S whose preceding element i in $\sigma'_1, \ldots, \sigma'_m$ appears multiple times
  $\mathscr{B} :=$ block containing i
  $i_1, \ldots, i_\ell :=$ elements in $\sigma'_1, \ldots, \sigma'_m$ that are in the block $\mathscr{B}$
  $j_1^{c_1}, \ldots, j_\ell^{c_\ell} :=$ elements respectively following $i_1, \ldots, i_\ell$ in $\sigma'_1, \ldots, \sigma'_m$
  $\leq :=$ partial order on $j_1^{c_1}, \ldots, j_\ell^{c_\ell}$ given by (2.2.2) with repeated elements treated as distinct, incomparable elements
  $\preceq :=$ the partial order on $i_1, \ldots, i_\ell$ formed by starting with $\leq$, replacing each $j_h^{c_h}$ with i_h, and reversing the relations in $\leq$
  $\sigma'_1, \ldots, \sigma'_m := \sigma'_1, \ldots, \sigma'_m$ after replacing instances of $i_1, \ldots, i_\ell$ with minimal elements in $\mathscr{B}$ in a manner that respects $\preceq$
end
return $\sigma'_1, \ldots, \sigma'_m$

Algorithm 4.2.1: `ColoredDescents`

Proof of Lemma 4.2.19. It was shown in [Ful98, Proof of Theorem 3] that the conjugation action of J on any conjugacy class C_λ of $\mathfrak{S}_n$ with no cycles of length $1, 2, \ldots, 2k$ results in orbits of size $|J| = \prod_{i=1}^{t} |\mathscr{B}_i|!$. Define $f : \mathfrak{S}_{n,r} \to \mathfrak{S}_n$ to be the projection $f(\omega, \tau) = \omega$.

Combining all of this with Lemma 4.2.18, we conclude that for any $(\omega, \tau) \in C_{\boldsymbol{\lambda}}$,

$$\begin{aligned} |J \cdot (\omega, \tau)| &= |\{(\pi, \mathbf{0})(\omega, \tau)(\pi, \mathbf{0})^{-1} \mid \pi \in J\}| \\ &\geq |\{f((\pi, \mathbf{0})(\omega, \tau)(\pi, \mathbf{0})^{-1}) \mid \pi \in J\}| \\ &= |\{\pi\omega\pi^{-1} \mid \pi \in J\}| \\ &= |J|. \end{aligned} \tag{4.2.33}$$

Since $|J \cdot (\omega, \tau)| \leq |J|$ always holds, we conclude $|J \cdot (\omega, \tau)| = |J|$. It now suffices to show that there is a unique element in each orbit with descents at $a_1, \ldots, a_k$, which we do using `ColoredDescents`. We start by showing that this algorithm is well-defined.

First, observe that the k elements in $a_1, \ldots, a_k$ can induce at most k blocks of size larger than 1, which accounts for at most $2k$ of the elements in $[n]$. Hence, some blocks in $\mathscr{B}_1, \ldots, \mathscr{B}_t$ must initially consist of only one element. If $(\omega, \tau) \in \mathfrak{S}_{n,r}$ has no cycles of lengths $1, 2, \ldots, 2k$, then each cycle $\sigma'_1, \ldots, \sigma'_m$ in line 3 must contain some element from a block of size 1. Consequently, choosing j in the `while` loop is well-defined in the first iteration. After each iteration of the `while` loop, the number of elements that appear exactly once increases in at least one cycle in $\sigma'_1, \ldots, \sigma'_m$, as the element that precedes j appears multiple times at the start of the loop but appears exactly once at the end of the loop. Consequently, future iterations of the `while` loop are well-defined, and the algorithm will continue until it terminates at a colored permutation. Furthermore, `ColoredDescents` only replaces elements in the cycle notation with others in the same block while leaving colors unchanged, so by Lemma 4.2.18, the output of this algorithm is in the same J-orbit as the original colored permutation.

Now observe that at the start of the algorithm, the cycles $\sigma'_1, \ldots, \sigma'_m$ in `ColoredDescents` satisfy the property that whenever $i_1 \leq i_2$ appear in $\sigma_1, \ldots, \sigma_m$ and belong to the same block, the elements $j_1^{c_1}$ and $j_2^{c_2}$ that follow them in $\sigma'_1, \ldots, \sigma'_m$ satisfy $j_1^{c_1} \geq j_2^{c_2}$ with respect to the ordering for descents given in (2.2.2). This property is preserved after each iteration of the `while` loop, so the colored permutation resulting from `ColoredDescents` has the property that when $i_1 < i_2$ are in the same block, the elements following them in the cycle notation satisfy $j_1^{c_1} > j_2^{c_2}$. Consequently, the colored permutation resulting from the algorithm has descents at $a_1, \ldots, a_k$. Additionally, it is clear that at each iteration of line 12, the algorithm identifies necessary conditions for descents to eventually occur at $a_1, \ldots, a_k$, and the replacement used at this line is unique. Consequently, the outputted colored permutation must be the unique permutation in the orbit of (ω, τ) that has a descent at all $a_1, \ldots, a_k$. □

Lemma 4.2.21. *Let $a_1, \ldots, a_k \in [n]$ with induced blocks $\mathscr{B}_1, \ldots, \mathscr{B}_t$, where $\mathscr{B}_t$ contains n. Let $C_{\boldsymbol{\lambda}}$ be any conjugacy class of $\mathfrak{S}_{n,r}$ that contains no cycles of length $1, 2, \ldots, 2k$. If $a_1, \ldots, a_k \in [n-1]$, then*

$$\mathbb{E}_{\boldsymbol{\lambda}}[X_{a_1} X_{a_2} \ldots X_{a_k}] = \prod_{i=1}^{t} \frac{1}{|\mathscr{B}_i|!}. \tag{4.2.34}$$

Otherwise,

$$\mathbb{E}_{\boldsymbol{\lambda}}[X_{a_1} X_{a_2} \ldots X_{a_k}] = \left(\frac{r-1}{r}\right)^{|\mathscr{B}_t|} \cdot \prod_{i=1}^{t} \frac{1}{|\mathscr{B}_i|!}. \tag{4.2.35}$$

Proof. First consider when $a_1, \ldots, a_k \in [n-1]$. Define $J = \mathfrak{S}_{\mathscr{B}_1} \times \ldots \times \mathfrak{S}_{\mathscr{B}_t}$, and let $\omega \in J$ act on $C_{\boldsymbol{\lambda}}$ as given in (4.2.30). Lemma 4.2.19 shows that this action decomposes $C_{\boldsymbol{\lambda}}$ into orbits of size $|J|$ where exactly one element in each orbit has descents at $a_1, \ldots, a_k$. This implies (4.2.34).

For (4.2.35), we assume without loss of generality that $a_k = n$ and $a_1, \ldots, a_{k-1} \in [n-1]$. Expressing $B_t = \{m+1, \ldots, n\}$, we have that

$$\mathbb{E}_{\boldsymbol{\lambda}}[X_{a_1} X_{a_2} \ldots X_{a_k}] = \Pr_{\boldsymbol{\lambda}}[\tau(i) \neq 0 \,\forall\, i > m] \cdot \Pr_{\boldsymbol{\lambda}}[X_{a_1}, \ldots X_{a_{k-1}} = 1 \mid \tau(i) \neq 0 \,\forall\, i > m]. \tag{4.2.36}$$

By summing over all choices of nonzero colors and applying (4.1.5), the first term is $((r-1)/r)^{n-m}$. For the second term, Lemma 4.2.18 shows that the action of J preserves the property that $\tau(i) \neq 0$ for all $i > m$, as the colors on the elements following $m+1, \ldots, n$ in the cycle notation are unaffected by the conjugation action of J. Hence, this action stabilizes the subset in $C_{\boldsymbol{\lambda}}$ where $\tau(i) \neq 0$ for all $i > m$. Lemma 4.2.19 then implies that the second term in (4.2.36) is $1/|J|$ as needed. ☐

Combining our results, we can now establish Theorem 1.2.11 for $\mathrm{des}_{n,r}$ and $\mathrm{maj}_{n,r}$. In fact, this result holds for any statistic that is a linear combination of the X_i statistics.

Theorem 4.2.22. *Let $X = \sum_{i=1}^{n} c_i X_i$ with $c_i \in \mathbb{R}$, and let $C_{\boldsymbol{\lambda}}$ be the conjugacy class of $\mathfrak{S}_{n,r}$ indexed by $\boldsymbol{\lambda}$. If $C_{\boldsymbol{\lambda}}$ contains no cycles of length $1, 2, \ldots, 2k$, then $\mathbb{E}_{\boldsymbol{\lambda}}[X^k] = \mathbb{E}_{\mathfrak{S}_{n,r}}[X^k]$. Furthermore, if $c_n = 0$, then this is also equal to $\mathbb{E}_{\mathfrak{S}_n}[X^k]$.*

Proof. Using the decomposition $X = \sum_{i=1}^{n} c_i X_i$ with $c_i \in \mathbb{R}$ and expanding, we obtain

$$\mathbb{E}_{\boldsymbol{\lambda}}[X^k] = \sum_{a_1,\ldots,a_k \in [n]} \left(\prod_{i=1}^{k} c_i\right) \cdot \mathbb{E}_{\boldsymbol{\lambda}}[X_{a_1} \ldots X_{a_k}]. \tag{4.2.37}$$

Note that the summation ranges over all possible $a_1, \ldots, a_k$, so it is possible that some of the X_i's in the product $X_{a_1} \ldots X_{a_k}$ are redundant. Regardless, using the fact that $C_{\boldsymbol{\lambda}}$ has no cycles of length $1, 2, \ldots, 2k$ with Lemma 4.2.13, Corollary 4.2.17, and Lemma 4.2.21, each of the summands in (4.2.37) is equal to the corresponding summand in

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[X^k] = \sum_{a_1,\ldots,a_k \in [n]} \left(\prod_{i=1}^{k} c_i\right) \mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1} \ldots X_{a_k}], \tag{4.2.38}$$

so the k-th moments of X on $C_{\boldsymbol{\lambda}}$ and $\mathfrak{S}_{n,r}$ coincide.

In the case where $c_n = 0$, we can restrict the summation in (4.2.38) to $a_1, \ldots, a_k \in [n-1]$. Lemma 4.2.21 then implies that each term in the summation for $\mathbb{E}_{\mathfrak{S}_{n,r}}[X^k]$ equals the corresponding one in

$$\mathbb{E}_{\mathfrak{S}_n}[X^k] = \sum_{a_1,\ldots,a_k \in [n-1]} \left(\prod_{i=1}^{k} c_i\right) \mathbb{E}_{\mathfrak{S}_n}[X_{a_1} \ldots X_{a_k}],$$

so the k-th moments of X on $\mathfrak{S}_{n,r}$ and $\mathfrak{S}_n$ coincide. ☐

Corollary 4.2.23. *Let $C_{\boldsymbol{\lambda}}$ be a conjugacy class of $\mathfrak{S}_{n,r}$. If $C_{\boldsymbol{\lambda}}$ contains no cycles of length*

$1, 2, \ldots, 2k$, *then*

$$\mathbb{E}_{\boldsymbol{\lambda}}[\mathrm{des}_{n,r}^k] = \mathbb{E}_{\mathfrak{S}_{n,r}}[\mathrm{des}_{n,r}^k] \quad \textit{and} \quad \mathbb{E}_{\boldsymbol{\lambda}}[\mathrm{maj}_{n,r}^k] = \mathbb{E}_{\mathfrak{S}_{n,r}}[\mathrm{maj}_{n,r}^k] = \mathbb{E}_{\mathfrak{S}_n}[\mathrm{maj}^k].$$

We conclude this section with Corollary 1.2.12 for $\mathrm{des}_{n,r}$ and $\mathrm{maj}_{n,r}$. These follow from a similar argument as in Corollary 4.2.9, where we apply the preceding result with the Method of Moments and either Theorem 2.2.10 or a combination of Theorem 1.2.4 and Theorem 1.2.7.

Corollary 4.2.24. *For every $n \geq 1$, let $C_{\boldsymbol{\lambda}_n}$ be a conjugacy class of $\mathfrak{S}_{n,r}$. Suppose that for all i, the number of cycles of length i in $\boldsymbol{\lambda}_n$ approaches 0 as $n \to \infty$. Then for sufficiently large n, $\mathrm{des}_{n,r}$ has mean $\mu_{n,r} = \frac{rn+r-2}{2r}$ and variance $\sigma_{n,r}^2 = \frac{n+1}{12}$ on $C_{\boldsymbol{\lambda}_n}$, and as $n \to \infty$, the random variable $(\mathrm{des}_{n,r} - \mu_{n,r})/\sigma_{n,r}$ converges in distribution to the standard normal distribution.*

Corollary 4.2.25. *For every $n \geq 1$, let $C_{\boldsymbol{\lambda}_n}$ be a conjugacy class of $\mathfrak{S}_{n,r}$. Suppose that for all i, the number of cycles of length i in $\boldsymbol{\lambda}_n$ approaches 0 as $n \to \infty$. Then for sufficiently large n, $\mathrm{maj}_{n,r}$ has mean $\mu_{n,r} = \frac{n(n-1)}{4}$ and variance $\sigma_{n,r}^2 = \frac{n(2n^2+3n-5)}{72}$ on $C_{\boldsymbol{\lambda}_n}$, and as $n \to \infty$, the random variable $(\mathrm{maj}_{n,r} - \mu_{n,r})/\sigma_{n,r}$ converges in distribution to the standard normal distribution.*

4.2.3 Flag-major index

In this section, we consider the flag-major index statistic $\mathrm{fmaj}_{n,r}$ and show that its k-th moment coincides on $\mathfrak{S}_{n,r}$ and any conjugacy class with no cycles of length $1, 2, \ldots, 2k$. Our general approach follows the approach we used for $\mathrm{des}_{n,r}$ and $\mathrm{maj}_{n,r}$, but with several technical modifications to account for the $\mathrm{col}_{n,r}$ statistic.

Throughout this section, we define $Y_{i,c}$ to be the indicator function for the color of $i \in [n]$ being $c \in \mathbb{Z}_r$,

$$Y_{i,c}(\omega, \tau) = \begin{cases} 1 & \text{if } \tau(i) = c \\ 0 & \text{otherwise.} \end{cases}$$

Using the same X_i indicator functions for descents, this allows us to express $\mathrm{fmaj}_{n,r}$ from

(2.2.4) as

$$\mathrm{fmaj}_{n,r} = r \cdot \sum_{i=1}^{n-1} iX_i + \sum_{i=1}^{n}\sum_{c=0}^{r-1} cY_{i,c},$$

In particular, $\mathrm{fmaj}_{n,r}^k$ can be expressed as linear combinations of the random variables

$$X_{a_1} \dots X_{a_j} Y_{a_{j+1},c_{j+1}} \dots Y_{a_k,c_k} \tag{4.2.39}$$

where $a_1, \dots, a_j \in [n-1]$, $a_{j+1}, \dots, a_k \in [n]$, and $c_{j+1}, \dots, c_k \in \mathbb{Z}_r$. We will consider products of this form, and show that their expectations align on $\mathfrak{S}_{n,r}$ and all $C_{\boldsymbol{\lambda}}$ with no cycles of length $1, 2, \dots, 2k$. We start with a definition and then a result on realizability.

Definition 4.2.26. Let $a_1, \dots, a_j \in [n-1]$, $a_{j+1}, \dots, a_k \in [n]$, and $c_{j+1}, \dots, c_k \in \mathbb{Z}_r$. The *support* of the statistic $X_{a_1} \dots X_{a_j} Y_{a_{j+1},c_{j+1}} \dots Y_{a_k,c_k}$ is

$$\mathrm{supp}(X_{a_1} \dots X_{a_j} Y_{a_{j+1},c_{j+1}} \dots Y_{a_k,c_k}) = \left(\bigcup_{i=1}^{j} \{a_i, a_i+1\}\right) \bigcup \left(\bigcup_{i=j+1}^{k} \{a_i\}\right).$$

Lemma 4.2.27. *Let $a_1, \dots, a_j \in [n-1]$, $a_{j+1}, \dots, a_k \in [n]$, and $c_{j+1}, \dots, c_k \in \mathbb{Z}_r$. Then $Z = X_{a_1} \dots X_{a_j} Y_{a_{j+1},c_{j+1}} \dots Y_{a_k,c_k}$ is realizable over constraints of size $j+k$. Consequently, its mean coincides on all conjugacy classes $C_{\boldsymbol{\lambda}}$ of $\mathfrak{S}_{n,r}$ without cycles of length $1, 2, \dots, j+k$. The same holds for $ZY_{i,c}$ when $i \in \mathrm{supp}(Z)$ and $c \in \mathbb{Z}_r$ is arbitrary.*

Proof. By Theorem 4.1.17, it suffices to show that Z and $ZY_{i,c}$ are realizable over constraints of size $j+k$. Observe that summands for $\mathrm{fmaj}_{n,r}$ in Example 4.1.6 can be used to realize each X_{a_i} using constraints of size 2 and each $Y_{a_i,c}$ using constraints of size 1. Using Lemma 4.1.15, Z is then realizable over constraints of size $2j + (k-j) = j+k$.

For $ZY_{i,c}$, first observe the resulting expansion described above for Z consists of products of linear combinations of statistics of the form

$$\prod_{i=1}^{j} I_{\{(a_i^0, x_i^{c_i}),((a_i+1)^0, y_i^{d_i})\}} \cdot \prod_{i=j+1}^{n} I_{\{(a_i^0, z_i^{c_i})\}}, \tag{4.2.40}$$

where $x_i^{c_i} < y_i^{d_i}$ are elements in $[n]^r$ and $z_i^{c_i}$ are elements in $[n]^r$ with color c_i. Additionally, we can express

$$Y_{i,c} = \sum_{x=1}^{n} I_{\{(i^0,\, x^c)\}}. \tag{4.2.41}$$

It now suffices to show that the product of (4.2.40) with any summand $I_{\{(i^0,\, x^c)\}}$ of (4.2.41) is realizable over constraints of size $j + k$.

Since $i \in \operatorname{supp}(Z)$, there is some $I_{(K,\kappa)}$ in the product (4.2.40) where (K, κ) contains an ordered pair of the form (i^0, z^d). If $x^c = z^d$, then multiplying (4.2.40) by $I_{(i^0,x^c)}$ has no effect, and hence, this additional indicator function can be omitted. Otherwise, $x^c \neq z^d$ implies that $I_{(K,\kappa)}$ is not compatible with $I_{(i^0,x^c)}$, so the product of (4.2.40) and $I_{\{(i^0,\, x^c)\}}$ is identically 0 by Lemma 4.1.15. Combined, we conclude that $ZY_{i,c}$ is also realizable over constraints of size $j + k$. □

In statistics of the form $X_{a_1} \dots X_{a_j} Y_{a_{j+1},c_{j+1}} \dots Y_{a_k,c_k}$, some of the elements in $a_{j+1}, \dots, a_k$ may be involved with descents at positions $a_1, \dots, a_j$, while others are not. Our next result allows us to reduce to when all elements in $a_{j+1}, \dots, a_k$ are involved in descents at $a_1, \dots, a_j$.

Lemma 4.2.28. *Let* $a_1, \dots, a_j \in [n-1]$*,* $a_{j+1}, \dots, a_k \in [n]$*, and* $c_{j+1}, \dots, c_k \in \mathbb{Z}_r$*. If* $a_k \notin \operatorname{supp}(X_{a_1} \dots X_{a_j} Y_{a_{j+1},c_{j+1}} \dots Y_{a_{k-1},c_{k-1}})$*, then*

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1} \dots X_{a_j} Y_{a_{j+1},c_{j+1}} \dots Y_{a_{k-1},c_{k-1}} Y_{a_k,c_k}] = \frac{1}{r} \cdot \mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1} \dots X_{a_j} Y_{a_{j+1},c_{j+1}} \dots Y_{a_{k-1},c_{k-1}}].$$

The same holds on any $C_{\boldsymbol{\lambda}}$ *with no cycles of length* $1, 2, \dots, j+k$*.*

Proof. For simplicity, we let $Z = X_{a_1} \dots X_{a_j} Y_{a_{j+1},c_{j+1}} \dots Y_{a_{k-1},c_{k-1}}$ and express

$$\begin{aligned}
\mathbb{E}_{\mathfrak{S}_{n,r}}[ZY_{a_k,c_k}] &= \Pr_{\mathfrak{S}_{n,r}}[ZY_{a_k,c_k} = 1] \\
&= \Pr_{\mathfrak{S}_{n,r}}[Z = 1] \cdot \Pr_{\mathfrak{S}_{n,r}}[Y_{a_k,c_k} = 1 \mid Z = 1] \\
&= \mathbb{E}_{\mathfrak{S}_{n,r}}[Z] \cdot \Pr_{\mathfrak{S}_{n,r}}[Y_{a_k,c_k} = 1 \mid Z = 1].
\end{aligned}$$

It now suffices to show the second term is $1/r$. Define an action of $\mathbb{Z}_r$ on $\mathfrak{S}_{n,r}$ as follows: m

acts on (ω, τ) by adding m to $\tau(a_k)$. Since a_k is not in the support of Z, this fixes the set of elements where $Z = 1$. Within each orbit of size r, exactly one satisfies $\tau(a_k) = c_k$. Hence, $\Pr_{\mathfrak{S}_{n,r}}[Y_{a_k,c_k} = 1 \mid Z = 1] = 1/r$ as desired.

For conjugacy classes with no cycles of length $1, 2, \ldots, j+k$, we let Ω_c be the conjugacy class of $\mathfrak{S}_{n,r}$ consisting of permutations with a single cycle of length n and color c, and we consider $\Omega = \bigcup_{c \in \mathbb{Z}} \Omega_c$. The same action of $\mathbb{Z}_r$ on $\mathfrak{S}_{n,r}$ given above is stable on Ω, implying

$$\mathbb{E}_{\Omega}[ZY_{i_k,c_k}] = \frac{1}{r}\mathbb{E}_{\Omega}[Z].$$

Lemma 4.2.27 with the Law of Total Expectation implies that

$$\mathbb{E}_{\Omega}[Z] = \sum_{c \in \mathbb{Z}_r} \Pr_{\Omega}[\Omega_c] \cdot \mathbb{E}_{\Omega_c}[Z] = \sum_{c \in \mathbb{Z}_r} \frac{1}{r} \cdot \mathbb{E}_{\Omega_0}[Z] = \mathbb{E}_{\Omega_0}[Z],$$

and the same holds when Z is replaced with ZY_{i_k,c_k}. Applying Lemma 4.2.27 again allows us to conclude that on any $C_{\boldsymbol{\lambda}}$ with no cycles of length $1, 2, \ldots, j+k$

$$\mathbb{E}_{\boldsymbol{\lambda}}[ZY_{i_k,c_k}] = \mathbb{E}_{\Omega_0}[ZY_{i_k,c_k}] = \mathbb{E}_{\Omega}[ZY_{i_k,c_k}] = \frac{1}{r}\mathbb{E}_{\Omega}[Z] = \frac{1}{r}\mathbb{E}_{\Omega_0}[Z] = \frac{1}{r}\mathbb{E}_{\boldsymbol{\lambda}}[Z]. \qquad \square$$

We now consider when $a_{j+1}, \ldots, a_k$ is the support of $X_{a_1} \ldots X_{a_j}$. In this case, our preceding work with `ColoredDescents` shows that the mean coincides on $\mathfrak{S}_{n,r}$ and any conjugacy classes without cycles of length $j+k$.

Lemma 4.2.29. *Let $a_1, \ldots, a_j \in [n-1]$, $a_{j+1}, \ldots, a_k \in [n]$, and $c_{j+1}, \ldots, c_k \in \mathbb{Z}_r$, and define $Z = X_{a_1} \ldots X_{a_j} Y_{a_{j+1},c_{j+1}} \ldots Y_{a_k,c_k}$. If*

$$\operatorname{supp}(X_{a_1} \ldots X_{a_j}) = \operatorname{supp}(Y_{a_{j+1},c_{j+1}} \ldots Y_{a_k,c_k}),$$

then on any conjugacy class $C_{\boldsymbol{\lambda}}$ of $\mathfrak{S}_{n,r}$ with no cycles of length $1, 2, \ldots, 2j$,

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[Z] = \mathbb{E}_{\boldsymbol{\lambda}}[Z].$$

Proof. We can assume without loss of generality that all of the elements $a_1, \ldots, a_j$ are distinct, and all of the elements $a_{j+1}, \ldots, a_k$ are distinct. Let $\mathscr{B}_1, \ldots, \mathscr{B}_t$ be the blocks induced by $a_1, \ldots, a_j$. We first consider when there exists some i and $i+1$ in the same block, and both $Y_{i,c}$ and $Y_{i+1,c'}$ appear in Z for some c, c'. If $c < c'$, then a descent at position i is impossible. This implies $Z = 0$ on both the entire group and on conjugacy classes, so the result is clear. If $c > c'$, then removing X_i from the product defining Z results in the same statistic. Iterating this argument, we can assume without loss of generality that for any i and $i+1$ in the same block where some $Y_{i,c}$ and $Y_{i+1,c'}$ appear in Z, we have $c = c'$. Combined with the fact that

$$\operatorname{supp}(X_{a_1} \ldots X_{a_j}) = \operatorname{supp}(Y_{a_{j+1},c_{j+1}} \ldots Y_{a_k,c_k}),$$

this implies that the property $Y_{a_{j+1},c_{j+1}} \ldots Y_{a_k,c_k}(\omega, \tau) = 1$ is equivalent to τ satisfying some fixed $\kappa : \{a_{j+1}, \ldots, a_k\} \to \mathbb{Z}_r$ that is constant on any block $\mathscr{B}_1, \ldots, \mathscr{B}_t$ of size larger than 1.

We now show the claimed equality, first by considering $\mathfrak{S}_{n,r}$. We first express

$$\begin{aligned}\mathbb{E}_{\mathfrak{S}_{n,r}}[Z] &= \Pr_{\mathfrak{S}_{n,r}}[X_{a_1} \ldots X_{a_j} = 1 \mid Y_{a_{j+1},c_{j+1}} \ldots Y_{a_k,c_k} = 1] \cdot \Pr_{\mathfrak{S}_{n,r}}[Y_{a_{j+1},c_{j+1}} \ldots Y_{a_k,c_k} = 1] \\ &= \Pr_{\mathfrak{S}_{n,r}}[X_{a_1} \ldots X_{a_j} = 1 \mid \tau \text{ satisfies } \kappa] \cdot \Pr_{\mathfrak{S}_{n,r}}[\tau \text{ satisfies } \kappa]\end{aligned}$$

There are $k - j$ elements in the domain of κ, so the second term is $1/r^{k-j}$. For the first term, we use a similar approach as the one for descents. Let $J = \mathfrak{S}_{\mathscr{B}_1} \times \ldots \times \mathfrak{S}_{\mathscr{B}_t}$ act by permuting the one-line notation within each block so that $\sigma \in \mathfrak{S}_{\mathscr{B}_j}$ permutes the images of i^0 for $i \in \mathscr{B}_j$. Since κ is constant on each block of size larger than 1, this action stabilizes the subset of colored permutations satisfying κ. Each orbit has size $|J|$ and contains exactly one element where the one-line notation within each block is in descending order. Hence, exactly one element in each orbit has the appropriate descents at $a_1, \ldots, a_j$, and we conclude

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[Z] = \frac{1}{r^{k-j}} \prod_{i=1}^{t} \frac{1}{|\mathscr{B}_i|}.$$

For a conjugacy class $C_{\boldsymbol{\lambda}}$ with no cycles of length $1, 2, \ldots 2j$, we similarly express

$$\mathbb{E}_{\boldsymbol{\lambda}}[Z] = \Pr_{\boldsymbol{\lambda}}[X_{a_1} \ldots X_{a_j} = 1 \mid \tau \text{ satisfies } \kappa] \cdot \Pr_{\boldsymbol{\lambda}}[\tau \text{ satisfies } \kappa].$$

Since we assumed that

$$\operatorname{supp}(X_{a_1} \ldots X_{a_j}) = \operatorname{supp}(Y_{a_{j+1},c_{j+1}} \ldots Y_{a_k,c_k})$$

and the left side has size at most $2j$, we conclude that the domain of κ has size $k - j \leq 2j$. Since $C_{\boldsymbol{\lambda}}$ has no cycles of length $1, 2, \ldots, 2j$, the second term in the product above is $1/r^{k-j}$ by (4.1.5). For the first term, let $\pi \in J$ act on (ω, τ) as conjugation by $(\pi, \mathbf{0})$. We know κ is constant on any block $\mathscr{B}_j$ of size larger than 1, so for any $(\omega, \tau) \in C_{\boldsymbol{\lambda}}$ where τ satisfies κ, all elements following $i \in \mathscr{B}_j$ in the cycle notation have the same color. Then Lemma 4.2.18 implies that this property is preserved under the action of J, and hence J stabilizes the elements in $C_{\boldsymbol{\lambda}}$ satisfying κ. This allows us to apply Lemma 4.2.19 to conclude that exactly one element in each orbit of size $|J|$ has descents at $a_1, \ldots, a_j$. Thus,

$$\mathbb{E}_{\boldsymbol{\lambda}}[Z] = \frac{1}{r^{k-j}} \prod_{i=1}^{t} \frac{1}{|\mathscr{B}_i|} = \mathbb{E}_{\mathfrak{S}_{n,r}}[Z]. \qquad \square$$

Finally, we show that the mean of $X_{a_1} \ldots X_{a_j} Y_{a_{j+1},c_{j+1}} \ldots Y_{a_k,c_k}$ coincides on $\mathfrak{S}_{n,r}$ and appropriate $C_{\boldsymbol{\lambda}}$. We then conclude with Theorem 1.2.11 and Corollary 1.2.12 for $\operatorname{fmaj}_{n,r}$.

Lemma 4.2.30. *Let $Z = X_{a_1} \ldots X_{a_j} Y_{a_{j+1},c_{j+1}} \ldots Y_{a_k,c_k}$ where $a_1, \ldots, a_j \in [n-1]$, $a_{j+1}, \ldots, a_k \in [n]$, and $c_{j+1}, \ldots, c_k \in \mathbb{Z}_r$. Then on any conjugacy class $C_{\boldsymbol{\lambda}}$ of $\mathfrak{S}_{n,r}$ with no cycles of length $1, 2, \ldots, j + k$,*

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[Z] = \mathbb{E}_{\boldsymbol{\lambda}}[Z].$$

Proof. We can assume without loss of generality that all of the elements $a_1, \ldots, a_j$ are distinct, and all of the elements $a_{j+1}, \ldots, a_k$ are distinct. Starting with Y, observe that if some $a_i \in \{a_{j+1}, \ldots, a_k\}$ is not in the support of $X_{a_1} \ldots X_{a_j}$, then Lemma 4.2.28 implies that it suffices to remove Y_{a_i,c_i} and prove the result for the resulting statistic. Applying this repeatedly, we see that we can assume $\operatorname{supp}(Y_{a_{j+1},c_{j+1}}\dots Y_{a_k,c_k}) \subseteq \operatorname{supp}(X_{a_1}\dots X_{a_j})$.

Now suppose that this is a proper subset, so there exists some $i \in \operatorname{supp}(X_{a_1}\dots X_{a_j}) \setminus \{a_{j+1},\dots,a_k\}$. In this case, we can express

$$Z = \sum_{c=0}^{r-1} Z \cdot Y_{i,c},$$

where each statistic in the sum is still realizable over constraints of size $j+k$ by Lemma 4.2.27. Hence, it suffices to show the statements for each $Z \cdot Y_{i,c}$. Iterating this process, we see that it suffices to consider when $\operatorname{supp}(Y_{a_{j+1},c_{j+1}}\dots Y_{a_k,c_k}) = \operatorname{supp}(X_{a_1}\dots X_{a_j})$, and this case follows from Lemma 4.2.29. □

Theorem 4.2.31. *Let $C_{\boldsymbol{\lambda}}$ be a conjugacy class of $\mathfrak{S}_{n,r}$ with no cycles of length $1, 2, \dots, 2k$. Then*

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[\mathrm{fmaj}_{n,r}^k] = \mathbb{E}_{\boldsymbol{\lambda}}[\mathrm{fmaj}_{n,r}^k].$$

Proof. As noted in (4.2.39), fmaj^k can be expressed as linear combinations of

$$X_{a_1}\dots X_{a_j} Y_{a_{j+1},c_{j+1}} \dots Y_{a_k,c_k}$$

where $a_1,\dots,a_j \in [n-1]$, $a_{j+1},\dots,a_k \in [n]$, and $c_{j+1},\dots,c_k \in \mathbb{Z}_r$. The previous result implies that on any $C_{\boldsymbol{\lambda}} \subseteq \mathfrak{S}_{n,r}$ with no cycles of length $1, 2, \dots, 2k$,

$$\mathbb{E}_{\mathfrak{S}_{n,r}}[X_{a_1}\dots X_{a_j} Y_{a_{j+1},c_{j+1}} \dots Y_{a_k,c_k}] = \mathbb{E}_{\boldsymbol{\lambda}}[X_{a_1}\dots X_{a_j} Y_{a_{j+1},c_{j+1}} \dots Y_{a_k,c_k}].$$

Linearity of expectation allows us to conclude $\mathbb{E}_{\boldsymbol{\lambda}}[\mathrm{fmaj}_{n,r}^k] = \mathbb{E}_{\mathfrak{S}_{n,r}}[\mathrm{fmaj}_{n,r}^k]$. □

Corollary 4.2.32. *For every $n \geq 1$, let $C_{\boldsymbol{\lambda}_n}$ be a conjugacy class of $\mathfrak{S}_{n,r}$. Suppose that for all i, the number of cycles of length i in $\boldsymbol{\lambda}_n$ approaches 0 as $n \to \infty$. Then for sufficiently large n, $\mathrm{fmaj}_{n,r}$ has mean $\mu_{n,r} = \frac{n(rn+r-2)}{4}$ and variance $\sigma_{n,r}^2 = \frac{2r^2n^3+3r^2n^2+(r^2-6)n}{72}$ on $C_{\boldsymbol{\lambda}_n}$.*

Furthermore, as $n \to \infty$, the statistic $(\mathrm{fmaj}_{n,r} - \mu_{n,r})/\sigma_{n,r}$ *converges in distribution to the standard normal distribution.*

Our methods for analyzing the asymptotic distributions of $\mathrm{des}_{n,r}$, $\mathrm{maj}_{n,r}$, and $\mathrm{fmaj}_{n,r}$ on conjugacy classes without short cycles involved moments. However, one might be interested in the actual distributions for these statistics on conjugacy classes, and this is one potential direction for future work.

Open Problem. *Study the generating functions for* $\mathrm{des}_{n,r}$, $\mathrm{maj}_{n,r}$, *and* $\mathrm{fmaj}_{n,r}$ *on conjugacy classes of* $\mathfrak{S}_{n,r}$. *Establish analogs of the results in Section 4.2.1 for these statistics, and explore asymptotic distributions on arbitrary conjugacy classes of* $\mathfrak{S}_{n,r}$.

Chapter 5

COSINE SIGN CORRELATION

We now consider sign patterns of cosine functions as described in Section 1.3. Recall from (1.3.1) that for positive integers $\{a_1, a_1, \ldots, a_n\} \subseteq \mathbb{Z}_+$,

$$\mathbb{P}(a_1, \ldots, a_n) = \frac{1}{2\pi} \left| \left\{ x \in [0, 2\pi] : \min_{1 \le i \le n} \cos(a_i x) > 0 \quad \text{or} \quad \max_{1 \le i \le n} \cos(a_i x) < 0 \right\} \right|$$

is the probability that $\cos(a_1 x)$, $\cos(a_2 x)$,..., $\cos(a_n x)$ all share the same sign when x is chosen uniformly at random from $[0, 2\pi]$. In particular, we minimize this probability by finding

$$p_n = \inf_{\{a_1, \ldots, a_n\} \subseteq \mathbb{Z}_+} \mathbb{P}(a_1, \ldots, a_n).$$

We begin with some general results in Section 5.1, including a more general case of Theorem 1.3.1. We then focus on $n = 3$ in Section 5.2, including the characterization of p_3 given in Theorem 1.3.3. All results in this section are from a collaboration with Shilin Dou, Ansel Goh, Madeline Legate, and Gavin Pettigrew, and these appear in [DGL+24].

5.1 Properties for general n

Throughout this section, we assume without loss of generality that $a_1 < a_2 < \ldots < a_n$. While we are primarily interested in cosine sign correlation on the interval $[0, 2\pi]$, it is useful at times to consider other intervals.

Definition 5.1.1. The *cosine sign correlation* of $\{a_1, \ldots, a_n\} \subseteq \mathbb{Z}_+$ on a bounded interval $I \subseteq \mathbb{R}$ is defined as

$$\mathbb{P}_I(a_1, \ldots, a_n) = \frac{1}{|I|} \cdot \left| \left\{ x \in I : \min_{1 \le i \le n} \cos(a_i x) > 0 \text{ or } \max_{1 \le i \le n} \cos(a_i x) < 0 \right\} \right|.$$

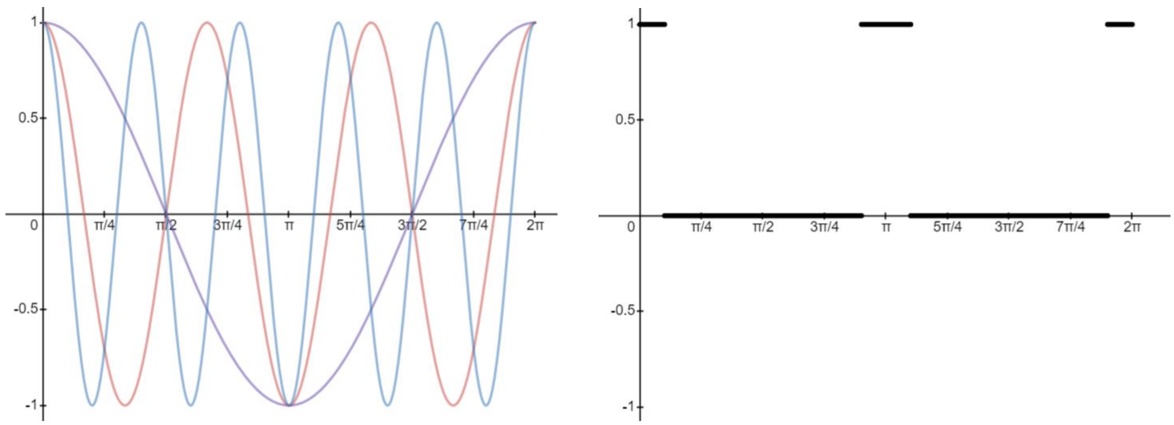


Figure 5.1: The functions $\cos(x), \cos(3x), \cos(5x)$ on $[0, 2\pi]$ are shown on the left, and $\chi_{(1,3,5)}(x)$ is shown on the right. The intervals where $\chi_{(1,3,5)}(x) = 1$ are $[0, \pi/10)$, $(9\pi/10, 11\pi/10)$, and $(19\pi/10, 2\pi]$, implying $\mathbb{P}(1, 3, 5) = 1/5$.

When $I = [0, 2\pi]$, we omit the subscript.

Given $\{a_1, \ldots, a_n\} \subseteq \mathbb{Z}_+$, we consider the indicator function

$$\chi_{a_1,\ldots,a_n}(x) = \begin{cases} 1 & \text{if } \min_{1\leq i\leq n} \cos(a_i x) > 0 \text{ or } \max_{1\leq i\leq n} \cos(a_i x) < 0, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that the cosine sign correlation of $\{a_1, \ldots, a_n\}$ can be equivalently expressed as

$$\mathbb{P}_I(a_1, \ldots, a_n) = \frac{1}{|I|} \int_I \chi_{a_1,\ldots,a_n}(x)\, dx. \tag{5.1.1}$$

Note that continuity of $\cos(a_1 x), \ldots, \cos(a_n x)$ implies that $\chi_{a_1,\ldots,a_n}$ takes value 1 on a union of open intervals in $\mathbb{R}$, so $\chi_{a_1,\ldots,a_n}$ is Riemann-integrable on any bounded interval $I \subseteq \mathbb{R}$. An example is shown in Figure 5.1. Applying these definitions, we establish Theorem 1.3.1.

Lemma 5.1.2. *Suppose that $\mathbb{P}(a_1, a_2, \ldots, a_n) = 1/a_n$ and a_i is odd for all i. Then for any positive integer m,*

$$\mathbb{P}(a_1, a_2, \ldots, a_n, 3a_n, \ldots, 3^m a_n) = \frac{1}{3^m a_n}.$$

Proof. We first show that when $a_1, \ldots, a_n$ are all odd and $\mathbb{P}(a_1, a_2, \ldots, a_n) = 1/a_n$, we have that

$$\chi^{-1}_{a_1,\ldots,a_n}(1) = \left[0, \frac{\pi}{2a_n}\right) \cup \left(\pi - \frac{\pi}{2a_n}, \pi + \frac{\pi}{2a_n}\right) \cup \left(2\pi - \frac{\pi}{2a_n}, 2\pi\right]. \tag{5.1.2}$$

Since all a_i are odd, we know that all $\cos(a_i x)$ are positive in a neighborhood of 0, negative in a neighborhood of π, and positive in a neighborhood of 2π. Since $\cos(a_i x)$ has period $2\pi/a_i$ and a_n is the largest of the a_i's, we have

$$\left[0, \frac{\pi}{2a_n}\right) \cup \left(\pi - \frac{\pi}{2a_n}, \pi + \frac{\pi}{2a_n}\right) \cup \left(2\pi - \frac{\pi}{2a_n}, 2\pi\right] \subseteq \chi^{-1}_{a_1,\ldots,a_n}(1). \tag{5.1.3}$$

Note that the total length of the intervals given in (5.1.3) is $2\pi/a_n$. By assumption, we know

$$\mathbb{P}(a_1, a_2, \ldots, a_n) = \frac{1}{2\pi} \cdot |\chi^{-1}_{a_1,\ldots,a_n}(1)| = \frac{1}{a_n},$$

so $|\chi^{-1}_{a_1,\ldots,a_n}(1)| = 2\pi/a_n$. If (5.1.2) does not hold, then there is some x^* not in the left-hand side of (5.1.3) satisfying $\chi_{a_1,\ldots,a_n}(x^*) = 1$. However, since each $\cos(a_i x)$ is continuous, this would imply $\chi_{a_1,\ldots,a_n} \equiv 1$ in some open interval containing x that is not contained in (5.1.2), which would contradict the fact that $|\chi^{-1}_{a_1,\ldots,a_n}(1)| = 2\pi/a_n$. We conclude that (5.1.2) holds.

Second, consider $\{a_1, \ldots, a_n, 3a_n\} \subseteq \mathbb{Z}_+$. The period of $\cos(3a_n x)$ is $2\pi/(3a_n)$, so on the interval $I = [0, \pi/(2a_n))$, we have that $\cos(3a_n x) > 0$ only for $x \in [0, \pi/(6a_n))$. On the remaining intervals in $\chi^{-1}_{a_1,\ldots,a_n}(1)$, we have that $\cos(3a_n x) < 0$ precisely when $x \in (\pi - \pi/(6a_n), \pi + \pi/(6a_n))$ and $\cos(3a_n x) > 0$ precisely when $x \in (2\pi - \pi/(6a_n), 2\pi]$. Combined, we find

$$\chi^{-1}_{a_1,\ldots,a_n,3a_n}(1) = \left[0, \frac{\pi}{6a_n}\right) \cup \left(\pi - \frac{\pi}{6a_n}, \pi + \frac{\pi}{6a_n}\right) \cup \left(2\pi - \frac{\pi}{6a_n}, 2\pi\right]. \tag{5.1.4}$$

From this, we conclude

$$\mathbb{P}(a_1, \ldots, a_n, 3a_n) = \frac{1}{2\pi} \int_0^{2\pi} \chi_{a_1,\ldots,a_n,3a_n}(x)\, dx = \frac{1}{2\pi} \cdot \frac{2\pi}{3a_n} = \frac{1}{3a_n}.$$

Finally, observe that if a_n is odd, then $3a_n$ is also odd. Hence, the general result follows from induction on m. □

Theorem 5.1.3 (Theorem 1.3.1 revisited)**.** *For any $n \geq 2$, we have that*

$$p_n \leq \mathbb{P}\left(1, 3, 9, \ldots, 3^{n-1}\right) = \frac{1}{3^{n-1}}.$$

Proof. A direct calculation shows $\mathbb{P}(1, 3) = 1/3$. The result now follows from the preceding lemma. □

Our general approach in the results above is to consider where $\chi_{a_1,\ldots,a_n}$ has value 1. Using this idea, we derive a general method for calculating $\mathbb{P}(a_1, \ldots, a_n)$.

Lemma 5.1.4. *Let $\ell = \operatorname{lcm}(a_1, \ldots, a_n)$. For each $m \in \{0, 1, \ldots, 4\ell - 1\}$, choose a sample point $x_m^* \in (\pi m/(2\ell), \pi(m+1)/(2\ell))$. Then*

$$\mathbb{P}(a_1, \ldots, a_n) = \frac{\#\{x_m^* : \chi_{a_1,\ldots,a_n}(x_m^*) = 1\}}{4\ell}.$$

Proof. The function $\cos(a_i x)$ is 0 when $a_i x = \pi/2 + \pi k$ for some $k \in \mathbb{Z}$. Then zeros can only occur when

$$x = \frac{\pi}{2a_i} + \frac{\pi k}{a_i} = \frac{\pi(1 + 2k)}{2a_i} = \pi \cdot \frac{(1 + 2k) \cdot \ell/a_i}{2\ell}.$$

Hence, $\chi_{a_1,\ldots,a_n}$ is constant on intervals of the form $(\pi m/(2\ell), \pi(m+1)/(2\ell))$. Using this, we find that

$$\begin{aligned}
\mathbb{P}(a_1, \ldots, a_n) &= \frac{1}{2\pi} \int_0^{2\pi} \chi_{a_1,\ldots,a_n}(x)\, dx \\
&= \frac{1}{2\pi} \sum_{m=0}^{4\ell-1} \int_{\pi m/(2\ell)}^{\pi(m+1)/(2\ell)} \chi_{a_1,\ldots,a_n}(x)\, dx \\
&= \frac{1}{2\pi} \cdot \frac{\pi}{2\ell} \cdot \#\{x_m^* : \chi_{a_1,\ldots,a_n}(x_m^*) = 1\} \\
&= \frac{\#\{x_m^* : \chi_{a_1,\ldots,a_n}(x_m^*) = 1\}}{4\ell}.
\end{aligned}$$

Note that the finitely many points of the form $\pi m/(2\ell)$ on $[0, 2\pi]$ do not affect the integral above. □

Theorem 1.3.2 implies that $p_2 = 1/3$, and Lemma 5.1.2 implies $p_3 \leq 1/9$, so $p_3 \leq p_2/3$. Focusing on this factor of 1/3, we show that $\mathbb{P}(a_1, \ldots, a_n) \leq \mathbb{P}(a_1, \ldots, a_{n-1})/3$ can only hold when a_n is sufficiently small with respect to the remaining integers $\{a_1, \ldots, a_{n-1}\}$.

Lemma 5.1.5. *If* $a_n > 12 \cdot \operatorname{lcm}(a_1, \ldots, a_{n-1})$, *then*

$$\mathbb{P}(a_1, \ldots, a_n) > \frac{1}{3} \cdot \mathbb{P}(a_1, \ldots, a_{n-1}).$$

Proof. Let $\ell = \operatorname{lcm}(a_1, \ldots, a_{n-1})$. As observed in Lemma 5.1.4, $\chi_{a_1,\ldots,a_{n-1}}$ is constant on any interval of the form $I = (\pi m/(2\ell), \pi(m+1)/(2\ell)) \subseteq [0, 2\pi]$. Suppose $\chi_{a_1,\ldots,a_{n-1}}|_I = 1$. The function $\cos(a_n t)$ completes r full cycles on I for some $r \in \mathbb{Z}_+$. We denote the intervals for these cycles $I_1, \ldots, I_r$, and let I_{r+1} be the remaining portion of I. Decompose

$$\begin{aligned} \mathbb{P}_I(a_1, \ldots, a_n) &= \frac{\sum_{j=1}^r |I_j| \cdot \mathbb{P}_{I_j}(a_1, \ldots, a_n) + |I_{r+1}| \cdot \mathbb{P}_{I_{r+1}}(a_1, \ldots, a_n)}{|I|} \\ &\geq \frac{\sum_{j=1}^r |I_j| \cdot \mathbb{P}_{I_j}(a_1, \ldots, a_n)}{|I|}. \end{aligned} \tag{5.1.5}$$

Since $\cos(a_n t)$ completes one full cycle in each I_j and all remaining components have the same sign, we have that

$$\mathbb{P}_{I_j}(a_1, \ldots, a_n) = \frac{1}{2} \cdot \mathbb{P}_{I_j}(a_1, \ldots, a_{n-1}) = \frac{1}{2}.$$

All intervals $I_1, \ldots, I_r$ have the same length, so this implies

$$\mathbb{P}_I(a_1, \ldots, a_n) \geq \frac{r|I_1|}{2|I|}.$$

Since $|I| = \pi/(2\ell)$, $|I_1| = 2\pi/a_n$, and $r = \lfloor (\pi/(2\ell))/(2\pi/a_n) \rfloor = \lfloor a_n/(4\ell) \rfloor$, we find

$$\mathbb{P}_I(a_1, \ldots, a_n) \geq \frac{4\ell}{2a_n} \cdot \left\lfloor \frac{a_n}{4\ell} \right\rfloor \geq \frac{2\ell}{a_n} \left(\frac{a_n}{4\ell} - 1 \right) = \frac{1}{2} - \frac{2\ell}{a_n}.$$

Since the assumption $a_n > 12\ell$ implies $2\ell/a_n < 1/6$, we have

$$\mathbb{P}_I(a_1, \ldots, a_n) > \frac{1}{2} - \frac{1}{6} = \frac{1}{3} = \frac{1}{3}\mathbb{P}_I(a_1, \ldots, a_{n-1})$$

on any I where $\chi_{a_1,\ldots,a_{n-1}} = 1$. Hence, $\mathbb{P}(a_1, \ldots, a_n) > \mathbb{P}(a_1, \ldots, a_{n-1})/3$. □

Remark 5.1.6. By also considering an upper bound in the proof of Lemma 5.1.5, one can obtain the bounds

$$\frac{1}{2} - \frac{2\ell}{a_n} \leq \mathbb{P}_I(a_1, \ldots, a_n) \leq \frac{1}{2} + \frac{4\ell}{a_n} \tag{5.1.6}$$

on any interval I where $\cos(a_1 x), \ldots, \cos(a_{n-1} x)$ share the same sign. Hence, as $a_n \to \infty$, we have $\mathbb{P}_I(a_1, \ldots, a_n) \to 1/2$. This allows us to conclude that

$$\lim_{a_n \to \infty} \mathbb{P}(a_1, \ldots, a_n) = \frac{1}{2} \cdot \mathbb{P}(a_1, \ldots, a_{n-1}).$$

This formalizes the idea that large values of a_n will multiply cosine sign correlation by a factor of approximately 1/2. The bounds in (5.1.6) also allow us to find a_n so that the factor is arbitrarily close to 1/2, and Lemma 5.1.5 is a special case of this.

Remark 5.1.7. When $a_n \leq 12 \cdot \mathrm{lcm}(a_1, \ldots, a_{n-1})$, the conclusion of Lemma 5.1.5 need not hold. For example, Lemma 5.1.4 allows us to calculate $\mathbb{P}(1, 3, 11) = 5/33$ and $\mathbb{P}(1, 3, 11, 33) = 1/33$.

5.2 *The subsets achieving* p_3

We now focus on the three-dimensional case. In the case of three distinct positive integers $\{a, b, c\} \subseteq \mathbb{Z}_+$ with $\gcd(a, b, c) = 1$, we will show that subsets achieving p_3 have at least two of the elements in $\{a,b,c\}$ relatively small and then apply Lemma 5.1.5 to bound the remaining element.

We start by revisiting a key lemma used in establishing Theorem 1.3.2. Gonçalves, Oliveira e Silva, and Steinerberger considered

$$\Phi(x,y) = \operatorname{sgn}(\cos(2\pi x)\cos(2\pi y)). \tag{5.2.1}$$

Using Fourier Analysis, they established the following result for lines on the two-dimensional torus $\mathbb{T}^2$, which we view as $\mathbb{R}^2$ with points separated by integer coordinates identified.

Lemma 5.2.1. *[GOS21, Lemma 3] Let $a,b \in \mathbb{R}$ be nonzero such that $a/b = p/q$ for some coprime $p,q \in \mathbb{Z}$. Let $\alpha,\beta \in \mathbb{R}$ and let $\gamma(t) = (at-\alpha, bt-\beta)$ be the corresponding ray on $\mathbb{T}^2$. If p or q is even, then*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \Phi(\gamma(t))\,dt = 0. \tag{5.2.2}$$

If both p and q are odd, then

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \Phi(\gamma(t))\,dt = (-1)^{\frac{p+q}{2}} \frac{8}{\pi^2 pq} \sum_{\ell=0}^{\infty} \frac{\cos(2\pi(2\ell+1)(p\beta - q\alpha))}{(2\ell+1)^2}. \tag{5.2.3}$$

There is one important consequence of this result, which we use multiple times. We state and prove this below.

Corollary 5.2.2. *Let $a,b \in \mathbb{R}$ be nonzero such that $a/b = p/q$ for some coprime $p,q \in \mathbb{Z}$, and define $\gamma(t) = (at, bt)$ to be a ray on $\mathbb{T}^2$. Then*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \Phi(\gamma(t)) = \int_0^1 \Phi(\gamma(t))\,dt. \tag{5.2.4}$$

Furthermore, if p or q is even, then

$$\int_0^1 \Phi(\gamma(t)) = 0, \tag{5.2.5}$$

and if both p and q are odd, then

$$\left|\int_0^1 \Phi(\gamma(t))\right| = \frac{1}{|pq|}. \tag{5.2.6}$$

Proof. For $\alpha = \beta = 0$, the function $\Phi(\gamma(t))$ is 1-periodic, so for any positive integer k, the integral of $\Phi(\gamma(t))$ on $[k, k+1]$ is the same. Then for any positive integer T,

$$\frac{1}{T}\int_0^T \Phi(\gamma(t)) = \frac{1}{T}\cdot\sum_{k=0}^{T-1}\int_k^{k+1}\Phi(\gamma(t))\,dt = \int_0^1 \Phi(\gamma(t))\,dt.$$

Taking the limit as $T \to \infty$, we conclude that (5.2.4) holds. If p or q is even, then (5.2.2) implies the left side of (5.2.4) is 0, and we conclude (5.2.5) holds. For the case when p and q are odd, first note that $\alpha = \beta = 0$ implies that $p\beta - q\alpha = 0$, so $\cos(2\pi(2\ell+1)(p\beta - q\alpha)) = 1$ for all ℓ in the summation on the right-hand side of (5.2.3). Since $\sum_{\ell=0}^{\infty} 1/(2\ell+1)^2 = \pi^2/8$, we conclude that

$$\left|\int_0^1 \Phi(\gamma(t))\,dt\right| = \left|\lim_{T\to\infty}\frac{1}{T}\int_0^T \Phi(\gamma(t))\right| = \left|\frac{8}{\pi^2 pq}\sum_{\ell=0}^{\infty}\frac{1}{(2\ell+1)^2}\right| = \frac{1}{|pq|}. \qquad \square$$

We consider lines of the form $\gamma(t) = (at, bt, ct)$ on the three-dimensional torus $\mathbb{T}^3$, which is $\mathbb{R}^3$ with points separated by integer coordinates identified. Define the function

$$\Psi(\gamma(t)) = \frac{\Phi(at, bt) + \Phi(at, ct) + \Phi(bt, ct) - 1}{2}, \tag{5.2.7}$$

which takes value 1 when $\cos(2\pi at), \cos(2\pi bt), \cos(2\pi ct)$ have the same sign and -1 otherwise. Letting I denote the set of all $x \in [0, 2\pi]$ such that $\Psi(\gamma(x/2\pi)) = 1$, a change of variables shows

$$\begin{aligned}\int_0^1 \Psi(\gamma(t))\,dt &= \frac{1}{2\pi}\int_0^{2\pi}\Psi\left(\gamma\left(\frac{x}{2\pi}\right)\right)\,dx = \frac{1}{2\pi}\left(|I| - (2\pi - |I|)\right)\\ &= \frac{1}{2\pi}\left(2\cdot|I| - 2\pi\right) = 2\cdot\mathbb{P}(a,b,c) - 1.\end{aligned} \tag{5.2.8}$$

For the remainder of this section, we fix distinct $a, b, c \in \mathbb{Z}_+$ and select $p, q, r, s, u, v \in \mathbb{Z}_+$ such that $a/b = p/q$, $a/c = r/s$, and $b/c = u/v$ with $\gcd(p, q) = \gcd(r, s) = \gcd(u, v) = 1$. We now give a reduction to the case when a, b, c are all odd.

Lemma 5.2.3. *Suppose* $\gcd(a, b, c) = 1$ *and* $\mathbb{P}(a, b, c) \leq 1/9$*. Then* a, b, c *are odd and*

$$\frac{1}{|pq|} + \frac{1}{|rs|} + \frac{1}{|uv|} \geq \frac{5}{9}.$$

Proof. Since $\mathbb{P}(a, b, c) \leq 1/9$, (5.2.8) implies

$$\left|\int_0^1 \Psi(at, bt, ct)dt\right| = |1 - 2 \cdot \mathbb{P}(a, b, c)| \geq \left|1 - 2 \cdot \frac{1}{9}\right| = \frac{7}{9}. \tag{5.2.9}$$

We show that if a, b, or c are even, then

$$\left|\int_0^1 \Psi(at, bt, ct)dt\right| \leq \frac{2}{3},$$

so (5.2.9) is not satisfied.

First, assume that we have exactly one even integer in $\{a, b, c\}$. Without loss of generality, suppose it is c. Then v and s are even integers and p and q are distinct odd integers, so by the triangle inequality and Corollary 5.2.2,

$$\begin{aligned}\left|\int_0^1 \Psi(at, bt, ct)dt\right| &\leq \frac{1}{2}\left(\left|\int_0^1 \Phi(at, bt)dt\right| + \left|\int_0^1 \Phi(at, ct)dt\right| + \left|\int_0^1 \Phi(bt, ct)dt\right| + 1\right) \\ &\leq \frac{1}{2|pq|} + \frac{1}{2} \\ &\leq \frac{2}{3}.\end{aligned} \tag{5.2.10}$$

Next, assume that we have exactly two even integers in $\{a, b, c\}$. Without loss of generality, suppose these are a and b. Then r and u are both even. If p and q are both odd, then

(5.2.10) again holds. If p or q is even, then Corollary 5.2.2 implies a smaller bound of

$$\left|\int_0^1 \Psi(at,bt,ct)dt\right| \leq \frac{1}{2}\left(\left|\int_0^1 \Phi(at,bt)dt\right| + \left|\int_0^1 \Phi(at,ct)dt\right| + \left|\int_0^1 \Phi(bt,ct)dt\right| + 1\right) = \frac{1}{2}.$$

Combined, we conclude that all three of a, b, and c must be odd for (5.2.9) to hold.

To conclude, we use the triangle inequality on the left-hand side of (5.2.9) and apply Corollary 5.2.2 to obtain

$$\frac{1}{2|pq|} + \frac{1}{2|rs|} + \frac{1}{2|uv|} + \frac{1}{2} \geq \frac{7}{9}.$$

Rewriting this, we conclude that

$$\frac{1}{|pq|} + \frac{1}{|rs|} + \frac{1}{|uv|} \geq \frac{5}{9}.$$ □

Finally, we rule out the case $a \neq 1$. We then conclude with our result characterizing subsets achieving p_3.

Lemma 5.2.4. *Suppose* $gcd(a,b,c) = 1$ *and* $a < b < c$. *If* $a \neq 1$, *then* $\mathbb{P}(a,b,c) > \frac{1}{9}$.

Proof. If a, b, or c is even, then the result follows from Lemma 5.2.3. Assume then that a, b, and c are all odd, so that p, q, r, s, u, and v are all odd as well. We do not consider the case when both $a \mid b$ and $a \mid c$ since this violates $\gcd(a,b,c) = 1$. We also do not consider the case when both $a \mid b$ and $b \mid c$ since this also violates $\gcd(a,b,c) = 1$. The remaining cases can be grouped into the following situations:

1. $a \nmid b, a \nmid c$,
2. $a \nmid c, b \nmid c$,
3. $a \nmid b, b \nmid c$, and
4. $a \nmid b, a \mid c, b \mid c$.

By Lemma 5.2.3, it suffices to show that in these cases,

$$\frac{1}{|pq|}+\frac{1}{|rs|}+\frac{1}{|uv|}<\frac{5}{9}.$$

We will consider cases (1), (2), and (3) simultaneously. Note that $1<a<b<c$ implies $q>p\geq 1$, $s>r\geq 1$, and $v>u\geq 1$. If $a\nmid b$, it follows that $p\geq 3$ and $q\geq 5$. Likewise, $a\nmid c$ implies $r\geq 3$ and $s\geq 5$, and $b\nmid c$ implies $u\geq 3$ and $v\geq 5$. In (1), (2), and (3), two out of the following three hold: $a\nmid b$, $a\nmid c$, or $b\nmid c$. Then

$$\frac{1}{|pq|}+\frac{1}{|rs|}+\frac{1}{|uv|}\leq\frac{1}{15}+\frac{1}{15}+\frac{1}{3}=\frac{7}{15}<\frac{5}{9}.$$

Thus, in these cases, we have $\mathbb{P}(a,b,c)>1/9$.

Now consider case (4). Note that $a\nmid b$, so $p\geq 3$ and $q\geq 5$. We consider r, s, u, and v. We know that $a\mid c$ and $b\mid c$ with $a<b$, so $s>v$. If $s\geq 7$, then

$$\frac{1}{|pq|}+\frac{1}{|rs|}+\frac{1}{|uv|}\leq\frac{1}{15}+\frac{1}{7}+\frac{1}{3}=\frac{19}{35}<\frac{5}{9}.$$

Note that $v>3$ would imply $s\geq 7$ since $s>v$. Hence, the only remaining possibility to rule out is $s=5$ and $v=3$. From the definition of r, s, u, and v, this implies that $a/c=1/5$ and $b/c=1/3$. We conclude that $c=5a$ and $c=3b$, which implies $b=5a/3$. Thus, we consider triples of the form $\{a,5a/3,5a\}$. Since we assumed $\gcd(a,b,c)=1$, $\{3,5,15\}$ is the only possibility. A direct check with Lemma 5.1.4 shows that $\mathbb{P}(3,5,15)>1/9$. □

Theorem 5.2.5 (Theorem 1.3.3 revisited)**.** *We have*

$$\mathbb{P}(a_1,a_2,a_3)\geq\frac{1}{9}$$

with equality if and only if $\{a_1,a_2,a_3\}=\gcd(a_1,a_2,a_3)\cdot\{1,3,9\}$. *Hence, we have that* $p_3=1/9$.

Proof. By Theorem 1.3.1, $1/9$ is achieved by $k\cdot\{1,3,9\}$ for any $k\in\mathbb{Z}_+$. We show that no other choices of a, b, c can attain $\mathbb{P}(a, b, c) \leq 1/9$. It suffices to consider $a < b < c$ with $\gcd(a, b, c) = 1$. Recall from Lemma 5.2.3 that if $\mathbb{P}(a, b, c) \leq 1/9$, then a, b, and c are odd and

$$\frac{1}{|pq|} + \frac{1}{|rs|} + \frac{1}{|uv|} \geq \frac{5}{9}.$$

In addition, Lemma 5.2.4 shows that $a = 1$, which forces $p = r = 1$, $q = b$, and $s = c$. Since $b < c$, we also have $u \geq 1$ and $v \geq 3$. Additionally, c is odd, so $c \geq b + 2$. Combined, we conclude

$$\frac{1}{|pq|} + \frac{1}{|rs|} + \frac{1}{|uv|} \leq \frac{1}{b} + \frac{1}{b+2} + \frac{1}{3}.$$

Note that for $b \geq 9$, we have

$$\frac{1}{b} + \frac{1}{b+2} + \frac{1}{3} \leq \frac{1}{9} + \frac{1}{11} + \frac{1}{3} = \frac{53}{99} < \frac{5}{9}.$$

Hence, we must have $b < 9$, and since b cannot be even, we conclude that $b \leq 7$. Theorem 1.3.2 implies $\mathbb{P}(a, b) \geq 1/3$ for any a and b, so it follows from Lemma 5.1.5 that $\mathbb{P}(a, b, c) \leq 1/9$ can only occur if $c \leq 12b \leq 84$. Therefore, if $\mathbb{P}(a, b, c) \leq 1/9$, we must have $a = 1$, $b \leq 7$, and $c \leq 84$. Computer verification using Lemma 5.1.4 then establishes the result. □

The cases of $n = 4$ and $n = 5$ are natural follow-up problems. The subsets achieving p_2 and p_3 suggest that powers of 3 may be involved in subsets achieving general p_n. However, as stated in Section 1.3, the subsets $\{1, 3, 11, 33\}$ and $\{1, 3, 11, 35, 105\}$ achieve smaller cosine sign correlations for $n = 4$ and $n = 5$, respectively. Consequently, using only powers of 3 is not sufficient. These observations suggest several directions for future work.

Open Problem. *Determine if p_4 is achieved precisely by multiples of $\{1, 3, 11, 33\}$.*

Open Problem. *Determine if either of the following conditions are necessary for $a_1 < a_2 < \ldots < a_n$ to be a subset achieving p_n: $a_2 = 3a_1$ or $a_n = 3a_{n-1}$.*

In general, we know that $p_n \geq 0$, and Theorem 1.3.1 gives us an upper bound $p_n \leq 1/3^{n-1}$. This upper bound can also be improved slightly by using $p_4 \leq \mathbb{P}(1,3,11,33) = 1/33$ or $p_5 \leq \mathbb{P}(1,3,11,35,105) = 1/105$ with Lemma 5.1.2, though the resulting bound still involves powers of 1/3. Another natural future direction is to improve these bounds for general n.

Open Problem. *Find $f(n)$ and $g(n)$ such that the following hold for all $n \in \mathbb{Z}_+$:*

(a) $f(n) \leq p_n \leq g(n)$,

(b) $f(n) > 0$, *and*

(c) $g(n+1) < \frac{1}{3}g(n)$ *when n is sufficiently large.*

BIBLIOGRAPHY

[ABR01] R. M. Adin, F. Brenti, and Y. Roichman. Descent numbers and major indices for the hyperoctahedral group. *Advances in Applied Mathematics*, 27(2):210–224, 2001.

[AGR05] R. Adin, I. Gessel, and Y. Roichman. Signed mahonians. *Journal of Combinatorial Theory, Series A*, 109:25–43, 01 2005.

[AR01a] R. Adin and Y. Roichman. The flag major index and group actions on polynomial rings. *European Journal of Combinatorics*, 22:431–446, 05 2001.

[AR01b] R. M. Adin and Y. Roichman. Descent functions and random Young tableaux. *Combin. Probab. Comput.*, 10(3):187–201, 2001.

[AR16] A. Asinowski and A. Regev. Triangulations with few ears: Symmetry classes and disjointness. *Integers*, 16:A5, 2016.

[AS92] M. D. Atkinson and J.-R. Sack. Generating binary trees at random. *Inf. Process. Lett.*, 41:21–23, 1992.

[ASV17] D. F. Anderson, T. Seppäläinen, and B. Valkó. *Introduction to Probability*. Cambridge Mathematical Textbooks. Cambridge University Press, 2017.

[BB05] A. Björner and F. Brenti. *Combinatorics of Coxeter groups*, volume 231 of *Graduate Texts in Mathematics*. Springer, New York, 2005.

[BBB$^+$12] K. Buchin, M. Buchin, J. Byrka, M. Nöllenburg, Y. Okamoto, R. I. Silveira, and A. Wolff. Drawing (complete) binary tanglegrams: hardness, approximation, fixed-parameter tractability. *Algorithmica*, 62(1-2):309–332, 2012.

[BBJ17] A. Bacher, O. Bodini, and A. Jacquot. Efficient random sampling of binary and unary-binary trees via holonomic equations. *Theoretical Computer Science*, 695:42–53, 2017.

[BD92] D. Bayer and P. Diaconis. Trailing the dovetail shuffle to its lair. *The Annals of Applied Probability*, 2, 05 1992.

[BHK01] T. Bohman, R. Holzman, and D. Kleitman. Six lonely runners. *Electr. J. Comb.*, 8, 02 2001.

[BKM17] S. C. Billey, M. Konvalinka, and F. A. Matsen. On the enumeration of tanglegrams and tangled chains. *J. Combin. Theory Ser. A*, 146:239–263, 2017.

[BKS20] S. C. Billey, M. Konvalinka, and J. P. Swanson. Tableau posets and the fake degrees of coinvariant algebras. *Advances in Mathematics*, 371:107252, 2020.

[BLM+23] A. E. Black, K. Liu, A. McDonough, G. Nelson, M. C. Wigal, M. Yin, and Y. Yoo. Sampling planar tanglegrams and pairs of disjoint triangulations. *Advances in Applied Mathematics*, 149:Paper No. 102550, 2023.

[Bón17] M. Bóna. *A walk through combinatorics*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2017.

[Bre93] F. Brenti. Permutation enumeration, symmetric functions, and unimodality. *Pacific J. Math.*, 157(1):1–28, 1993.

[Bre94] F. Brenti. q-eulerian polynomials arising from coxeter groups. *European Journal of Combinatorics*, 15(5):417–441, 1994.

[BS08] J. Barajas and O. Serra. The lonely runner with seven runners. *Electr. J. Comb.*, 15, 03 2008.

[BS21] N. Blitvic and E. Steingrímsson. Permutations, moments, measures. *Transactions of the American Mathematical Society*, 374(08):5473–5508, 2021.

[BZ11] A. Baxter and D. Zeilberger. The number of inversions and the major index of permutations are asymptotically joint-independently normal, 2011. arXiv:1004.1160.

[CJZ20] M. C. Cooper, W. S. Jones, and Y. Zhuang. On the joint distribution of cyclic valleys and excedances over conjugacy classes of S_n. *Adv. in Appl. Math.*, 115:Paper No. 101999, 2020.

[CLL+23] J. Campion Loth, M. Levet, K. Liu, S. Sundaram, and M. Yin. Moments of colored permutation statistics by conjugacy class, 2023. arXiv:2305.11800.

[CM12] C.-O. Chow and T. Mansour. Asymptotic probability distributions of some permutation statistics for the wreath product $C_r \wr S_n$. *Online Analytic Journal of Combinatorics*, 7:Article #2, 12 2012.

[CSW19] É. Czabarka, L. A. Székely, and S. Wagner. A tanglegram Kuratowski theorem. *J. Graph Theory*, 90(2):111–122, 2019.

[Cus73] T. Cusick. View-obstruction problems. *Aequationes Mathematicae*, 9, 01 1973.

[Cze12] S. Czerwiński. Random runners are very lonely. *Journal of Combinatorial Theory, Series A*, 119(6):1194–1199, 2012.

[DB62] F. N. David and D. E. Barton. *Combinatorial Chance*. Hafner Publishing Company, 1962.

[DF09] P. Diaconis and J. Fulman. Carries, shuffling, and an amazing matrix. *The American Mathematical Monthly*, 116(9):788–803, 2009.

[DFLS03] P. Duchon, P. Flajolet, G. Louchard, and G. Schaeffer. Boltzmann samplers for the random generation of combinatorial structures. *Combinatorics Probability and Computing*, 13, 02 2003.

[DGL$^+$24] S. Dou, A. Goh, K. Liu, M. Legate, and G. Pettigrew. Cosine sign correlation. *J. Fourier Anal. Appl.*, 30(13), 2024.

[DH23] P. D orr and J. Heiny. Extremes of joint inversions and descents on finite coxeter groups, 2023. arXiv:2309.17314.

[DLRS10] J. De Loera, J. Rambau, and F. Santos. *Triangulations*, volume 25 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, 2010. Structures for algorithms and applications.

[DMP95] P. Diaconis, M. McGrath, and J. Pitman. Riffle shuffles, cycles, and descents. *Combinatorica*, 15(1):11–29, mar 1995.

[EF23] D. Eppstein and D. Frishberg. Improved Mixing for the Convex Polygon Triangulation Flip Walk. In Kousha Etessami, Uriel Feige, and Gabriele Puppis, editors, *50th International Colloquium on Automata, Languages, and Programming (ICALP 2023)*, volume 261 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 56:1–56:17, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[Fel45] W. Feller. The fundamental limit theorems in probability. *Bulletin of the American Mathematical Society*, 51:800–832, 1945.

[Fir05] M. Fire. Statistics on wreath products, 2005. arXiv:math/0409421.

[FKLP21] J. Fulman, G. B. Kim, S. Lee, and T. K. Petersen. On the joint distribution of descents and signs of permutations. *Electron. J. Combin.*, 28(3):Paper No. 3.37, 30, 2021.

[Ful98] J. Fulman. The distribution of descents in fixed conjugacy classes of the symmetric groups. *J. Combin. Theory Ser. A*, 84(2):171–180, 1998.

[Fus16] É. Fusy. On symmetries in phylogenetic trees. *The Electronic Journal of Combinatorics*, 23, 02 2016.

[Ges21] I. M. Gessel. Counting tanglegrams with species. *J. Comb. Theory, Ser. A*, 184:105498, 2021.

[GG09] A. Garsia and Alain Goupil. Character polynomials, their q-analogs and the kronecker product. *Electr. J. Comb.*, 16, 07 2009.

[Gil13] J. Gill. *The k-assignment polytope, phylogenetic trees, and permutation patterns.* PhD thesis, Linköping University, Department of Mathematics, Mathematics and Applied Mathematics. Linköping University, The Institute of Technology., 2013.

[GOS17] F. Gonçalves, D. Oliveira e Silva, and S. Steinerberger. Hermite polynomials, linear flows on the torus, and an uncertainty principle for roots. *Journal of Mathematical Analysis and Applications*, 451(2):678–711, 2017.

[GOS21] F. Gonçalves, D. Oliveira e Silva, and S. Steinerberger. A universality law for sign correlations of eigenfunctions of differential operators. *Journal of Spectral Theory*, 11, 03 2021.

[GR93] I. M. Gessel and C. Reutenauer. Counting permutations with given cycle structure and descent set. *Journal of Combinatorial Theory, Series A*, 64(2):189–215, 1993.

[GR20] C. Gaetz and C. Ryba. Stable characters from permutation patterns. *Selecta Mathematica*, 27:1–13, 2020.

[GW06] L. A. Goddyn and E. B. Wong. Tight instances of the lonely runner problem. *Integers*, 6, 2006.

[Hal13] B. C. Hall. *Quantum Theory for Mathematicians.* Graduate Texts in Mathematics. Springer, New York, 2013.

[Hol97] J. M. Holte. Carries, combinatorics, and an amazing matrix. *The American Mathematical Monthly*, 104(2):138–149, 1997.

[HR22] Z. Hamaker and B. Rhoades. Characters of local and regular permutation statistics, 2022. arXiv:2206.06567.

[HT11] C. E. Heitsch and P. Tetali. Meander Graphs. pages 469–480. Discrete Mathematics and Theoretical Computer Science, 2011.

[Hul14] A. Hultman. Permutation statistics of products of random permutations. *Advances in Applied Mathematics*, 54:1–10, 2014.

[Hun12] T. W. Hungerford. *Abstract Algebra: An Introduction*. Cengage Learning, 2012.

[HW74] J. E. Hopcroft and J. K. Wong. Linear time algorithm for isomorphism of planar graphs (preliminary report). In *Symposium on the Theory of Computing*, 1974.

[KHC04] J. Kukluk, L. Holder, and D. Cook. Algorithm and experiments in testing planar graphs for isomorphism. *J. Graph Algorithms Appl.*, 8:313–356, 01 2004.

[KL20] G. B. Kim and S. Lee. Central limit theorem for descents in conjugacy classes of S_n. *Journal of Combinatorial Theory, Series A*, 169:105123, 2020.

[Knu98] D. E. Knuth. *The Art of Computer Programming, Volume 3: (2nd Ed.) Sorting and Searching*. Addison Wesley Longman Publishing Co., Inc., USA, 1998.

[Kra21] N. Kravitz. Barely lonely runners and very lonely runners: a refined approach to the lonely runner problem. *Combinatorial Theory*, 2021.

[KW16] M. Konvalinka and S. Wagner. The shape of random tanglegrams. *Adv. Appl. Math.*, 78:76–93, 2016.

[Liu23] K. Liu. Characterizing planar tanglegram layouts and applications to edge insertion problems. *Electronic Journal of Combinatorics*, 30(2):Paper No. 2.32, 2023.

[LLSY24] M. Levet, K. Liu, S. Sundaram, and M. Yin. Descents and flag-major index on conjugacy classes of $\mathfrak{S}_{n,r}$ without short cycles, In preparation 06/06/2024.

[LP17] D. A. Levin and Y. Peres. *Markov Chains and Mixing Times*. American Mathematical Society, 2017.

[LPR+07] A. Lozano, R. Y. Pinter, O. Rokhlenko, G. Valiente, and M. Ziv-Ukelson. Seeded tree alignment and planar tanglegram layout. In *WABI*, pages 98–110, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.

[Mac13] P. A. MacMahon. The indices of permutations and the derivation therefrom of functions of a single variable associated with the permutations of any assemblage of objects. *American Journal of Mathematics*, 35(3):281–322, 1913.

[Mac98] I. G. Macdonald. *Symmetric Functions and Hall Polynomials*. Oxford classic texts in the physical sciences. Clarendon Press, 1998.

[MBKK15] F. A. Matsen, S. C. Billey, A. Kas, and M. Konvalinka. Tanglegrams: a reduction tool for mathematical phylogenetics. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 15(1):343–349, 2015.

[Mou18] V.-D. P. Moustakas. The Eulerian distribution on the involutions of the hyperoctahedral group is unimodal. *Graphs and Combinatorics*, 35:1077 – 1090, 2018.

[Mou21] V.-D. Moustakas. Specializations of colored quasisymmetric functions and euler–mahonian identities. *Advances in Applied Mathematics*, 127:102177, 2021.

[MR08] A. Mendes and J. Remmel. Descents, inversions, and major indices in permutation groups. *Discrete Mathematics*, 308(12):2509–2524, 2008.

[MRS01] M. Molloy, B. Reed, and W. Steiger. On the mixing rate of the triangulation walk. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 43, 09 2001.

[MT97] L. McShine and P. Tetali. On the mixing time of the triangulation walk and other catalan structures. In *Randomization Methods in Algorithm Design*, 1997.

[OEI24] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences. `http://oeis.org`, 2024.

[Pag93] R. D. M. Page. Parasites, phylogeny and cospeciation. *International Journal for Parasitology*, 23(4):499–506, 1993.

[Pet15] T. K. Petersen. *Eulerian numbers*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser/Springer, New York, 2015.

[Pou14] L. Pournin. The diameter of associahedra. *Advances in Mathematics*, 259:13–42, 2014.

[PS14] G. Perarnau and O. Serra. Correlation among runners and some results on the lonely runner conjecture. *Electron. J. Comb.*, 23:1, 2014.

[Rei93a] V. Reiner. Signed permutation statistics. *European J. Combin.*, 14(6):553–567, 1993.

[Rei93b] V. Reiner. Signed permutation statistics and cycle type. *European J. Combin.*, 14(6):569–579, 1993.

[Rei95] V. Reiner. Descents and one-dimensional characters for classical Weyl groups. *Discrete Math.*, 140(1-3):129–140, 1995.

[Rém85] J.-L. Rémy. Un procédé itératif de dénombrement d'arbres binaires et son application á leur génération aléatoire. *RAIRO Theor. Informatics Appl.*, 19(2):179–195, 1985.

[Ren04] J. Renault. View-obstruction: a shorter proof for 6 lonely runners. *Discrete Mathematics*, 287(1):93–101, 2004.

[Ros42] A. C. Rosander. The use of inversions as a test of random order. *Journal of the American Statistical Association*, 37(219):352–358, 1942.

[RRW18] D. Ralaivaosaona, J. B. Ravelomanana, and S. Wagner. Counting Planar Tanglegrams. In *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018)*, volume 110 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 32:1–32:18, Dagstuhl, Germany, 2018. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

[Rud76] W. Rudin. *Principles of mathematical analysis*. McGraw-Hill, 1976.

[ST54] G. C. Shephard and J. A. Todd. Finite unitary reflection groups. *Canadian Journal of Mathematics*, 6:274–304, 1954.

[Sta79] R. P. Stanley. Invariants of finite groups and their applications to combinatorics. *Bulletin of the American Mathematical Society*, 1:475–511, 1979.

[Sta99] R. P. Stanley. *Enumerative Combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[Sta15] R. P. Stanley. *Catalan Numbers*. Cambridge University Press, 2015.

[Ste89] J. R. Stembridge. On the eigenvalues of representations of reflection groups and wreath products. *Pacific Journal of Mathematics*, 140(2):353 – 396, 1989.

[Ste94] E. Steingrímsson. Permutation statistics of indexed permutations. *European Journal of Combinatorics*, 15(2):187–205, 1994.

[Tao17] T. Tao. Some remarks on the lonely runner conjecture. *Contributions to Discrete Mathematics*, 13, 01 2017.

[VASJG09] B. Venkatachalam, J. Apple, K. St. John, and D. Gusfield. Untangling tanglegrams: Comparing trees by their drawings. In *Bioinformatics Research and Applications*, pages 88–99, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[Wil67] J. M. Wills. Zwei sätze über inhomogene diophantische approximation von irrationalzehlen. *Monatshefte für Mathematik*, 71:263–269, 1967.