Major Index Statistics: Cyclic Sieving, Branching Rules, and Asymptotics

Joshua P. Swanson

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

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Sara C. Billey, Chair

Max Lieblich

Soumik Pal

Program Authorized to Offer Degree:
Mathematics

University of Washington

Abstract

Major Index Statistics:
Cyclic Sieving, Branching Rules, and Asymptotics

Joshua P. Swanson

Chair of the Supervisory Committee:
Professor Sara C. Billey
Department of Mathematics

Major index statistics have been studied for over a century in many guises and appear throughout algebraic combinatorics. We pursue major index statistics from two complementary perspectives: algebraic and asymptotic. We first prove an instance of refined cyclic sieving for the major index statistic on words with a fixed cyclic descent type. We next connect this cyclic sieving result to Schur expansions due to Kraskiewicz–Weyman, Stembridge, and Schocker related to certain reflection group branching rules and higher Lie modules. This leads to a conjectured approach to a generalization of Thrall's problem. Afterwards, we transition between the algebraic and the probabilistic by classifying the irreducible components appearing in some of these induced representations. The argument uses the underlying representation theory to prove a uniform local limit theorem, answering a conjecture of Sundaram. We then study the distribution of the major index on standard tableaux of straight shape and certain skew shapes. In particular, we classify all possible limit laws, most of them normal, providing a common generalization of results due to Canfield–Janson–Zeilberger, Chen–Wang–Wang, Diaconis, Feller, Mann–Whitney, and others. We also provide a combinatorial and constructive characterization of the irreducible representations appearing in each degree of the type A coinvariant algebra. Finally we describe a new approach to a result of Baxter–Zeilberger on the limiting joint distribution of the inversion number and major index on permutations using a generating function of Roselle, answering a \$300 question of Romik and Zeilberger.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to heartily thank my advisor, Sara Billey, for so many things: careful readings, discussions, introductions, travel support, literature references, collaboration, mentorship, a lively seminar, friendship—the list goes on. Looking back, I am exceptionally happy with my time in grad school in no small part because of her efforts. I would also like to thank my partner, Andrew Ohana, for his enormous support over the years, his work on some of my projects, and his tremendous patience as a sounding board. Further thanks go to my remaining collaborators on the work that follows, Connor Ahlbach and Matjaž Konvalinka; all of their contributions were instrumental. Still more thanks go to Vic Reiner, Brendon Rhoades, Yuval Roichman, Dan Romik, Sheila Sundaram, and Doron Zeilberger for providing their conjectures, preprints, questions, expositions, and insights for these and related projects. The University of Washington Mathematics Department also provided a wonderfully friendly, collaborative, and supportive environment. I'd like to thank many faculty members and fellow graduate students for their dedication to teaching, willingness to attend my seminars and talks, and the thousand day-to-day things that add up beyond the sum of their parts to make a happy work environment. In particular, I would like to thank the remaining members of my committee, Max Lieblich, Soumik Pal, and Boris Blinov.

I do not, however, thank Seattle's interminable, dreary weather, in spite of the fact that it likely increased my productivity.

DEDICATION

For Stu, Donna, Michael, and Andrew.

Chapter 1

INTRODUCTION

The *major index* statistic was introduced a century ago by MacMahon [62]. For a permutation $w = w_1 \cdots w_n$ in the symmetric group S_n, the major index of w is the sum of all i such that $w_i > w_{i+1}$. For example, the identity permutation $\mathrm{id} = 1\ 2\ \cdots\ n$ has major index 0 and the "longest element" $w_0 = n\ n-1\ \cdots\ 1$ has major index $1 + 2 + \cdots + (n-1) = \binom{n}{2}$. Almost magically, variations on the major index appear again and again throughout algebraic combinatorics. Baxter and Zeilberger [7] describe the major index as the "second most important permutation statistic" (after inversion number). The following are some examples of the many uses of major index statistics, most of which will be discussed in more detail shortly.

1. There is a homogeneous polynomial basis for the cohomology ring of the complete flag manifold indexed by permutations whose major index is the degree of the corresponding polynomial [33].

2. The type A coinvariant algebra is a graded S_n-module over the complex numbers. The multiplicity of each irreducible component in each homogeneous component is given by counting standard tableaux by their major index [87, Prop. 4.11]

3. The stable principal specializations of several classic symmetric function bases including the complete homogeneous and Schur bases have beautiful, compact expressions in terms of major index generating functions [91, Prop. 7.19.11].

4. Families of irreducible complex $\mathrm{GL}_n(\mathbb{F}_q)$-representations have polynomial degree as

a function of q. These polynomials are essentially given by major index generating functions [39, Lemma 7.4].

5. The branching rules from *arbitrary* cyclic subgroups of S_n are given by counting tableaux with a given generalized major index up to a modular congruence [94, Thm. 3.3].

1.1 Generating Functions and Representation Theory

The canonical launching point for explorations of the major index is the following beautiful generating function identity due to MacMahon.

Theorem 1.1.1 (MacMahon, [60, Art. 6])**.** *The ordinary generating function for the major index statistic on the symmetric group S_n is*

$$\sum_{w \in S_n} q^{\mathrm{maj}(w)} = [n]_q! := [n]_q [n-1]_q \cdots [1]_q$$

where $[n]_q := 1 + q + \cdots + q^{n-1} = \frac{1-q^n}{1-q}$ is a q-integer.

The ordinary generating function of the inversion statistic

$$\mathrm{inv}(w_1 \cdots w_n) := \#\{i < j : w_i > w_j\}$$

was also known by MacMahon to be $[n]_q!$. Consequently, the two most important permutation statistics, inv and maj on S_n, are equidistributed. Many bijective proofs of this surprising fact are now known, the first being due to Foata [26].

MacMahon considered major index statistics essentially from the perspective of what would now be termed enumerative combinatorics. The representation theoretic importance of major index statistics was realized piece by piece in the decades following MacMahon's explorations. To describe these connections, we require some of the basics of the representation theory of symmetric groups. All representations will be over $\mathbb{C}$.

The conjugacy classes of the symmetric group S_n are given by *cycle types*. The cycle type of a permutation is encoded in the list of the lengths of the cycles in its disjoint cycle decomposition, written in weakly decreasing order to avoid redundancy. We call such a list a *partition* of n, denoted $\lambda \vdash n$. The complex irreducible inequivalent representations of a finite group are equinumerous with its conjugacy classes, but for symmetric groups something extraordinary occurs: the irreducible representations of S_n are *canonically* indexed by partitions of n. The corresponding irreducible S^λ was constructed by Alfred Young [102, QSA IV] and are usually referred to as Specht modules[1]. The naturality of the construction is perhaps best illustrated by the notion of *Schur functors*. These are categorifications of the S^λ which "interpolate" between the symmetric power and exterior power endofunctors on vector spaces. See for instance [31, §8.1] for more details.

The representation theory of symmetric groups is intimately tied to the algebra of symmetric functions. A formal power series $f(x_1, x_2, \ldots)$ in infinitely many commuting indeterminates is called *symmetric* if it is unchanged upon swapping any two inputs. The most well-studied basis of symmetric functions is undoubtedly the *Schur functions* $s_\lambda(x_1, x_2, \ldots)$, which are again indexed by partitions. The *Frobenius characteristic* of an S_n-representation is given by sending the irreducible S^λ to the Schur function s_λ and extending additively. Many problems in algebraic combinatorics originate in representation theory and are transferred to algebra via the characteristic map. The process is often reversed as well: given a Schur-positive symmetric function, one may ask for an intrinsic construction of a corresponding S_n-module.

One particularly straightforward class of S_n-modules is given as follows. The *content* of a word $w \in \mathbb{Z}_{\geq 1}^n$ is the sequence $(\alpha_1, \alpha_2, \ldots)$ where α_i is the number of times the letter i appears in w. We write such a sequence as $\alpha \vDash n$, were n is the sum of the entries in α. Let W_α denote the set of words of content α. Permuting the letters of a word preserves its content, so S_n acts on the set W_α, and hence also on formal $\mathbb{C}$-linear combinations of elements of W_α. This S_n-module is almost always reducible. Under the Frobenius characteristic map,

[1] Perhaps unfairly. This situation is deplored by Adriano Garsia [34, Rem. 1.1], who suggests Young's peculiar writing style doomed his earlier work to relative obscurity.

this S_n-module corresponds with the well-known *complete homogeneous* symmetric function $h_\alpha := h_{\alpha_1} h_{\alpha_2} \cdots$. The stable principal specialization of the complete homogeneous symmetric functions have the following elegant expression in terms of the major index on W_α, which is defined verbatim as for S_n.

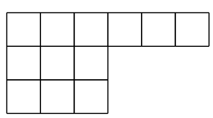
Theorem 1.1.2 ([91, Prop. 7.8.3] and [60, Art. 6])**.** *Let* $\alpha \vDash n$*. Then*

$$h_\alpha(1, q, q^2, \ldots) = \frac{\mathrm{W}_\alpha^{\mathrm{maj}}(q)}{(1-q)\cdots(1-q^n)}$$

where

$$\mathrm{W}_\alpha^{\mathrm{maj}}(q) := \sum_{w \in \mathrm{W}_\alpha} q^{\mathrm{maj}(w)} = \binom{n}{\alpha}_q := \frac{[n]_q!}{[\alpha_1]_q! [\alpha_2]_q! \cdots}.$$

The S_n-module $\mathbb{C}\{\mathrm{W}_\alpha\}$ associated with Theorem 1.1.2 by construction has basis given by words in W_α. Analogously, the irreducible modules S^λ have bases indexed by combinatorial objects called *standard Young tableaux*, which we next describe.



(a) The Young diagram of λ.

8	7	6	3	2	1
4	3	2			
3	2	1			

(b) Hook lengths of cells of λ.

1	2	4	7	9	12
3	6	10			
5	8	11			

(c) A standard tableau of shape λ.

Figure 1.1: Constructions related to the partition $\lambda = (6, 3, 3)$. The standard tableau has descents at $2, 4, 7, 9, 10$ and major index 32.

One frequently visualizes a partition $\lambda \vdash n$ using its *Young diagram*, which is the upper-left justified arrangement of unit squares called *cells* where the ith row from the top has λ_i cells[2]; see Figure 1.1a. Partitions and their Young diagrams are frequently considered synonymous. The *transpose* λ' of a partition λ is obtained by reflecting its Young diagram through the line $y = -x$. The *hook length* of a cell $c \in \lambda$ is the number h_c of cells in λ in the same row as c to the right of c or in the same column as c and below c, including c itself; see Figure 1.1b. A *standard tableau* of shape $\lambda \vdash n$ is a filling of the cells of the Young diagram of λ with the numbers $1, 2, \ldots, n$, each used once, which increases along rows and columns; see Figure 1.1c. The set of standard tableaux of shape λ is denoted $\mathrm{SYT}(\lambda)$. The count $f^\lambda := \#\,\mathrm{SYT}(\lambda)$ is given by the famous Frame–Robinson–Thrall hook length formula [29], which we describe shortly. The *descent set* of $T \in \mathrm{SYT}(\lambda)$ is the set of all labels i such that $i+1$ appears in a strictly lower row of T than i; see Figure 1.1. The *major index* of $T \in \mathrm{SYT}(\lambda)$ is

$$\mathrm{maj}(T) := \sum_{i \in \mathrm{Des}(T)} i.$$

We have the following beautiful analogue of Theorem 1.1.2 for Schur functions and the irreducibles S^λ due to Stanley. The first equality was first explicitly stated by Stanley and also occurred in unpublished work of Lusztig. For more historical details and references, see [91, p. 401-403].

Theorem 1.1.3 ([91, Prop. 7.19.11, Cor. 7.21.3])**.** *Let $\lambda \vdash n$. Then*

$$s_\lambda(1, q, q^2, \ldots) = \frac{\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)}{(1-q)\cdots(1-q^n)}$$

where

$$\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) := \sum_{T \in \mathrm{SYT}(\lambda)} q^{\mathrm{maj}(T)} = q^{b(\lambda)} \frac{[n]_q!}{\prod_{c \in \lambda} [h_c]_q}$$

and $b(\lambda) := \sum_i (i-1)\lambda_i$.

---

[2]Following English notation.

Corollary 1.1.4 (Hook Length Formula, Frame–Robinson–Thrall [29])**.** *Let $\lambda \vdash n$. Then*

$$f^\lambda = \frac{n!}{\prod_{c\in\lambda} h_c}.$$

Proof. Let $q \to 1$ in Theorem 1.1.3. $\square$

1.2 Cyclic Sieving and Evaluations at Roots of Unity

A common theme when encountering generating functions with a q-parameter is to find interpretations of related expressions when q is a prime-power or a complex root of unity. For example, Green [39], building on work of Steinberg [93], gave an interpretation for $(1-q)\cdots(1-q^n)s_\lambda(1,q,q^2,\ldots)$ when q is a prime-power. Using Theorem 1.1.3, Green's interpretation says $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$ is the dimension of a certain irreducible $\mathrm{GL}_n(\mathbb{F}_q)$-representation.

We pause here to fill a minor gap in the literature. Steinberg's formula [93, (2.9)] together with Green's work and an identity of Littlewood [39, Lemma 7.4, (41)] and Stanley's formula, Theorem 1.1.3, gives the following alternate product formula for $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$. We give a direct proof of the equivalence of the two formulas.

Corollary 1.2.1. *Let $\lambda = (\lambda_1, \ldots, \lambda_k) \vdash n$. Here we allow trailing 0's, i.e. the length λ is at most k. Set $\ell_i \coloneqq \lambda_i + k - i$. Then*

$$q^{b(\lambda)}\frac{[n]_q!}{\prod_{c\in\lambda}[h_c]_q} = q^{-\binom{k}{3}}\frac{[n]_q!}{[\ell_1]_q!\cdots[\ell_k]_q!}\Delta([\ell_1]_q,\ldots,[\ell_k]_q)$$

where

$$\Delta(x_1,\ldots,x_k) \coloneqq \prod_{i<j}(x_i - x_j)$$

is the Vandermonde determinant.

Proof. By definition, we have $\Delta([\ell_1]_q,\ldots,[\ell_k]_q) = \prod_{1\le i<j\le k}([\ell_i]_q - [\ell_j]_q)$. Note that $\ell_1 > \ell_2 >$

$\cdots > \ell_k$ and that $[\ell_i]_q - [\ell_j]_q = q^{\ell_j}[\ell_i - \ell_j]_q$. The equality is thus equivalent to

$$q^{b(\lambda)+\binom{k}{3}}[\ell_1]_q! \cdots [\ell_k]_q! = \prod_{c \in \lambda}[h_c]_q \prod_{1 \le i < j \le k} q^{\ell_j}[\ell_i - \ell_j]_q. \tag{1.1}$$

We have

$$\begin{aligned}
\sum_{1 \le i < j \le k} \ell_j &= \sum_{1 \le j \le k} (j-1)(\lambda_j + k - j) \\
&= \sum_{1 \le j \le k} (j-1)\lambda_j + \sum_{1 \le j \le k} (j-1)(k-j) \\
&= b(\lambda) + \binom{k}{3},
\end{aligned}$$

so the q-shifts on either side of (1.1) cancel. Now (1.1) reduces to the first half of [91, Lemma 7.21.1], which gives a simple bijective proof of the equality of multisets

$$\bigcup_{i=1}^{k} \{1, \ldots, \ell_i\} = \{h_c : c \in \lambda\} \cup \left(\bigcup_{1 \le i < j \le k} \{1, \ldots, \ell_i - \ell_j\} \right).$$

□

We further note that the $q = 1$ specialization of the right-hand side of Corollary 1.2.1 arises from the Frobenius character formula for $f^\lambda = K_{\lambda,(1^n)}$ [31, Exercise 7.3.6]. The $q = 1$ specialization of Corollary 1.2.1 is [31, Exercise 4.3.9].

We now return to our theme and describe further interpretations for $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$ and $\mathrm{W}_\alpha^{\mathrm{maj}}(q)$ when evaluated at complex roots of unity. The "long cycle" $\sigma_n := (1\ 2\ \cdots\ n) \in S_n$ is the unique (up to conjugacy) *Coxeter element* in S_n and hence frequently plays a special role in the representation theory of symmetric groups. One may then be led to ask the following.

Question 1.2.2. *Given an S_n-module M, how does the restricted module $M\downarrow^{S_n}_{\langle \sigma_n \rangle}$ over the cyclic group $\langle \sigma_n \rangle$ of order n decompose into irreducibles?*

One may reduce Question 1.2.2 to the case when $M = S^\lambda$ is irreducible. The n irreducible representations χ^r of the cyclic group $\langle \sigma_n \rangle$, all linear, may be identified[3] with $\mathbb{Z}/n$. Consequently, Question 1.2.2 may be rephrased as follows.

Question 1.2.3. *Give a rule for determining the counts*

$$a_{\lambda,r} := \textit{multiplicity of } \chi^r \textit{ in } S^\lambda\downarrow^{S_n}_{\langle \sigma_n \rangle}$$

where $\lambda \vdash n$ *and* $r \in \mathbb{Z}/n$.

Stembridge [94] studied the so-called *cyclic exponents* of complex reflection groups, which are essentially a generalization of the counts $a_{\lambda,r}$. He proved that the cyclic exponents for S_n are given by counting certain tableaux by major index as follows.

Theorem 1.2.4 (Stembridge, [94, Thm. 3.3])**.** *The cyclic exponents of* S_n *are given by*

$$a_{\lambda,r} = \#\{T \in \mathrm{SYT}(\lambda) : \mathrm{maj}(T) \equiv_n r\}$$

where $\lambda \vdash n$ *and* $r \in \mathbb{Z}/n$.

A straightforward argument (for instance, using the standard scalar product for C_n-characters) can be used to show that Stembridge's result is equivalent to the claim

$$\chi^\lambda(\sigma_n^r) = \mathrm{SYT}(\lambda)^{\mathrm{maj}}(\omega_n^r) \tag{1.2}$$

where χ^λ is the trace character of S^λ, ω_n is any fixed primitive nth root of unity, and $\lambda \vdash n$. Later, Reiner–Stanton–White introduced the following notion of *cyclic sieving*, which bears a strong resemblance to (1.2).

Definition 1.2.5 (Reiner–Stanton–White, [73])**.** Let W be a finite set on which a cyclic group C generated by an element σ_n of order n acts. Let $f(q) \in \mathbb{Z}_{\geq 0}[q]$ be a polynomial. We

[3] So long as we choose $\chi^r \leftrightarrow r$ in such a way that tensor products correspond to multiplication.

say the triple $(W, C, f(q))$ *exhibits the cyclic sieving phenomenon* (CSP) if

$$\#W^{\sigma_n^r} := \#\{w \in W : \sigma_n^r \cdot w = w\} = f(\omega_n^r) \tag{1.3}$$

for all $r \in \mathbb{Z}$, where ω_n is any primitive nth root of unity.

Intuitively, evaluations of the polynomial f at nth roots of unity give the characters of the C-action on W, leading to efficient techniques for counting fixed point sets. A key example arises from words of a fixed content using the major index generating function.

Theorem 1.2.6 (Reiner–Stanton–White, [73, Prop. 4.4])**.** *Suppose $\alpha \vDash n$ and let $\langle \sigma_n \rangle$ act on W_α by rotation. Then, the triple*

$$(\mathrm{W}_\alpha, \langle \sigma_n \rangle, \mathrm{W}_\alpha^{\mathrm{maj}}(q))$$

exhibits the CSP.

Equivalently, Theorem 1.2.6 gives the following analogue of (1.2):

$$\tau^\alpha(\sigma_n^r) = \mathrm{W}_\alpha^{\mathrm{maj}}(\omega_n^r) \tag{1.4}$$

where τ^α is the trace character of the $\langle \sigma_n \rangle$-action on W_α, or equivalently on the S_n-module $\mathbb{C}\{\mathrm{W}_\alpha\}$. The proof of Theorem 1.2.6 involves quite a bit of algebra and representation theory. It uses Springer's regular elements [86] and the invariant theory of complex reflection groups [87].

It is natural to ask for a more direct, combinatorial proof of Theorem 1.2.6. With Connor Ahlbach, we provide such a proof in Chapter 3 which has been published as [5]. An "extended abstract" of this work appeared in [3]. We in fact prove a stronger result, which we next describe.

Definition 1.2.7. The *cyclic descent number* of a word $w = w_1 \cdots w_n \in \mathbb{Z}_{\geq 1}^n$ is

$$\operatorname{cdes}(w) := \#\{i \in \mathbb{Z}/n : w_i > w_{i+1}\}.$$

Here we allow a descent between w_n and w_1. Let $w^{(i)}$ denote the word w with all letters strictly larger than i removed. The *cyclic descent type* of w is the sequence

$$\operatorname{CDT}(w) := (\operatorname{cdes}(w^{(1)}), \operatorname{cdes}(w^{(2)}) - \operatorname{cdes}(w^{(1)}), \operatorname{cdes}(w^{(3)}) - \operatorname{cdes}(w^{(2)}), \ldots).$$

Example 1.2.8. Suppose $w = 143124114223$, so

$$\begin{aligned}
w^{(1)} &= 1111 & \operatorname{cdes}(w^{(1)}) &= 0, \\
w^{(2)} &= 112.1122. & \operatorname{cdes}(w^{(2)}) &= 2, \\
w^{(3)} &= 13.12.11223. & \operatorname{cdes}(w^{(3)}) &= 3, \\
w^{(4)} &= 14.3.124.114.223. & \operatorname{cdes}(w^{(4)}) &= 5,
\end{aligned}$$

where periods have been inserted to indicate cyclic descents. Hence, $\operatorname{CDT}(143124114223) = (0, 2-0, 3-2, 5-3) = (0, 2, 1, 2)$.

With these definitions in place, our refinement of Theorem 1.2.6 is as follows.

Theorem 1.2.9. *Let $\mathrm{W}_{\alpha,\delta}$ denote the set of words of content $\alpha \vDash n$ and cyclic descent type δ. Then*

$$(\mathrm{W}_{\alpha,\delta}, \langle \sigma_n \rangle, \mathrm{W}_{\alpha,\delta}^{\mathrm{maj}}(q))$$

exhibits the CSP.

Attempting to apply the Reiner–Stanton–White representation-theoretic argument in [73] to this refinement encounters immediate difficulties since some of the S_n-modules it uses are irreducible. It would be very interesting to find a representation-theoretic interpretation of Theorem 1.2.9. The argument in Chapter 3 is completely different from that given by Reiner–Stanton–White. It involves first recursively constructing the sets $\mathrm{W}_{\alpha,\delta}$ using a Carlitz-style insertion procedure and tracking changes in the major index. This results in a product formula for $\mathrm{W}_{\alpha,\delta}^{\mathrm{maj}}(q)$ modulo $q^n - 1$, Theorem 3.5.19. The rest of the argument is fundamentally inductive and relies on a certain "extension lemma," Lemma 3.3.3, which is used to enlarge the cyclic group for which a set exhibits the CSP. The algebraic notion of *modular periodicity* is introduced in Section 3.3 which avoids explicit evaluations at roots of unity. Another interesting coarsening of Theorem 1.2.9 fixes just the number of circular descents instead of the full cyclic descent type.

Corollary 1.2.10. *Let $\mathrm{W}_{\alpha,k}$ denote the set of words of content $\alpha \vDash n$ with k cyclic descents. Then*

$$(\mathrm{W}_{\alpha,k}, \langle \sigma_n \rangle, \mathrm{W}_{\alpha,k}^{\mathrm{maj}}(q))$$

exhibits the CSP.

1.3 Branching Rules

Kraśkiewicz–Weyman [54] gave an earlier, different, and long-unpublished proof of (1.2) involving results of Stanley [87, Prop. 4.11] and Lusztig (also unpublished) on the type A coinvariant algebra as well as an intricate though beautiful argument involving ℓ-decomposable partitions. One of their main results is the following, which is equivalent to Theorem 1.2.4 using Frobenius reciprocity.

Theorem 1.3.1 (Kraśkiewicz–Weyman, [54])**.** *The multiplicity of S^λ in the induced representation $\chi^r\uparrow_{\langle \sigma_n \rangle}^{S_n}$ is*

$$a_{\lambda,r} = \#\{T \in \mathrm{SYT}(\lambda) : \mathrm{maj}(T) \equiv_n r\}.$$

One may ask about the relationship between the key equation (1.2) and the cyclic sieving result equation (1.4). They are equivalent by translating between the Schur and complete homogeneous bases in the following sense:

$$\begin{aligned}
\chi^\lambda(\sigma_n^r) &= \mathrm{SYT}(\lambda)^{\mathrm{maj}}(\omega_n^r) = (1-q)\cdots(1-q^n)s_\lambda(1,q,q^2,\ldots)|_{q=\omega_n^r} \\
\tau^\lambda(\sigma_n^r) &= \mathrm{W}_\lambda^{\mathrm{maj}}(\omega_n^r) \qquad\quad = (1-q)\cdots(1-q^n)h_\lambda(1,q,q^2,\ldots)|_{q=\omega_n^r}
\end{aligned}$$

where the first equality is (1.2), the second is Theorem 1.1.3, the third is (1.4), and the fourth is Theorem 1.1.2. Effectively, the top line and both Kraśkiewicz–Weyman and Stembridge's approaches to Theorem 1.3.1 and Theorem 1.2.4 are "in the Schur basis" while the bottom line is "in the h-basis."

One may instead ask for an approach to Theorem 1.3.1 and Theorem 1.2.4 "in the h-basis." In Chapter 4, with Ahlbach, we give such an approach to Theorem 1.3.1 hinging on the cyclic sieving result above, Theorem 1.2.6, or equivalently on (1.4). The argument is remarkably straightforward and applies more generally to a further result of Stembridge [94, Thm. 3.3] and a result of Schocker [81, Thm. 3.1]. It begins by generalizing an argument due to Klyachko which quickly expresses the Schur character of the $\mathrm{GL}(\mathbb{C}^m)$-module corresponding to the S_n-module $\chi^r\uparrow_{\langle\sigma_n\rangle}^{S_n}$ under Schur–Weyl duality as a generating function on necklaces, Proposition 4.2.5. This yields a monomial expansion for the Frobenius characteristic of $\chi^r\uparrow_{\langle\sigma_n\rangle}^{S_n}$. Turning this into a graded Frobenius series tracking branching rules for the inclusion $\langle\sigma_n\rangle \hookrightarrow S_n$, we are then able to directly apply cyclic sieving and the RSK algorithm to convert from the monomial to the Schur basis. The same basic outline holds for the two further results of Stembridge and Schocker mentioned above. The approach has several additional benefits: it is "nearly bijective" as discussed in Section 4.3, and certain generalized major index statistics maj_μ described in Definition 4.4.7 fall out very naturally from the combinatorics of orbits and necklaces.

The representations $\chi^r\uparrow_{\langle\sigma_n\rangle}^{S_n}$ are intimately related to the so-called *Lie modules*, which are summarized in Section 4.2.4. The decomposition of the Lie modules into irreducibles was considered in the 1940's by Thrall [99] and is referred to as *Thrall's problem*. Theorem 1.3.1 can be interpreted as solving this problem in an important special case. The general problem reduces to a difficult *plethysm* problem involving an induced module involving a certain inclusion $C_a \wr S_b \hookrightarrow S_{ab}$. In Section 4.5, we discuss applying our general approach to the determination of these higher Lie multiplicities. We give a monomial expansion for the graded Frobenius series tracking all branching rules involving the inclusion $C_a \wr S_b \hookrightarrow S_{ab}$, Theorem 4.5.4. In Open Problem 4.5.5, we conjecture the existence of a statistic interpolating in a certain sense between the major index modulo n and the shape under RSK. This statistic would allow us to turn our monomial expansion into a Schur expansion, thereby giving a combinatorial interpretation for the higher Lie multiplicities. Unfortunately, finding such a statistic has proven elusive outside of easy special cases.

Future work in this direction has also been begun with Ahlbach and Rhoades. We plan to study "Euler–Mahonian" refined cyclic sieving. We for instance have a conjectured refinement of Rhoades' CSP on rectangular tableaux [75], with a proof in special cases. We also have a refined CSP arising from work of Elizalde–Roichman [22]. Recent work of Adin–Reiner–Roichman [1] offers tantalizing clues to more results in this direction, though there are significant difficulties which have yet to be overcome.

1.4 Generating Functions and Asymptotics

The above generating function identities for $h_\alpha(1, q, q^2, \ldots)$ and $s_\lambda(1, q, q^2, \ldots)$, Theorem 1.1.2 and Theorem 1.1.3, are beautiful and algebraically useful. They may also be exploited from a probabilistic perspective. In Chapter 5, we use a blend of representation theory, generating function identities, and explicit estimates to answer a conjecture of Sundaram, which we next describe.

Question 1.4.1 (Sundaram [95])**.** *Fix a cycle type $\lambda \vdash n$. Let M^λ consist of formal $\mathbb{C}$-linear combinations of permutations of cycle type λ. Let S_n act by conjugation on M^λ. For which λ does every possible S_n-irreducible appear in M^λ?*

Sundaram answered the question subject to a conjectured description of which irreducibles appear when $\lambda = (n)$, corresponding to permutations with a single cycle. This in turn is equivalent to classifying which irreducibles appear in $1\uparrow_{\langle\sigma_n\rangle}^{S_n}$. By Kraśkiewicz–Weyman's theorem, Theorem 1.3.1, one may try to answer the question by constructing tableaux with major index divisible by n. This approach has several issues. First, earlier work by Johnson [46] suggested it would be combinatorially intricate. Second, it would not give an idea of the actual magnitude of $a_{\lambda,r}$. The approach we take in Chapter 5 is instead fundamentally probabilistic and gives a conceptual reason behind the classification. We prove the following *local limit theorem* which has been published in [97].

Theorem 1.4.2. *Let $\lambda \vdash n$ be a partition where $f^\lambda := \#\,\mathrm{SYT}(\lambda) \geq n^5 \geq 1$. Then for all r,*

$$\left|\frac{a_{\lambda,r}}{f^\lambda} - \frac{1}{n}\right| < \frac{1}{n^2}.$$

In particular, if $n \geq 81$, $\lambda_1 < n-7$, and $\lambda'_1 < n-7$, then $f^\lambda \geq n^5$ and the inequality holds.

Intuitively, the result says that the statistic "major index modulo n" on $\mathrm{SYT}(\lambda)$ is asymptotically uniformly distributed outside of a few degenerate cases. Consequently, this value may be zero only for a small class of shapes for which f^λ is particularly tiny. The argument in Chapter 5 proceeds by classifying such shapes using a curious inequality involving opposite hook lengths, Proposition 5.4.5. This inequality was discovered independently around the same time by Morales–Pak–Panova, [65, Proposition 12.1]. The resulting classification for when $a_{\lambda,r} = 0$, Theorem 5.1.3, answers Sundaram's conjecture, completing her approach to Question 1.4.1. See [95] for more details.

The proof of the local limit theorem, Theorem 1.4.2, involves exploiting Kraśkiewicz–Weyman's result, Foulkes' formula for the p-expansion of $\chi^r\uparrow_{\langle\sigma_n\rangle}^{S_n}$, some simple manipulations with Ramanujan sums, the generating function identity Theorem 1.1.3, a bound due to Fomin–Lulov [27], and Stirling's approximation. Chapter 5 also gives a new proof of a generalization of the hook length formula; see Section 5.5. The argument gives a particularly explicit description of the motion of hook lengths through a partition as ribbons are added. It hinges on Lemma 5.5.2, which appears to be new and which describes changes in the number of hook lengths modulo ℓ as length ℓ ribbons are added. A related set of criteria for ℓ-decomposability of a partition is summarized in Corollary 5.5.3.

Chapter 6 and Chapter 7 continue the theme of interplay between algebraic combinatorics and discrete probability. This interplay is neatly exemplified by a simple observation which we next describe.

Definition 1.4.3. Let X be a real-valued random variable. The *characteristic function* of X is the expectation of the corresponding random variable e^{isX} for $s \in \mathbb{R}$,

$$\phi_X(s) := \mathbb{E}[e^{isX}].$$

Characteristic functions are a very useful tool in asymptotic analysis. For instance, there is a standard and surprisingly easy proof of the central limit theorem using characteristic functions and Taylor approximations. On the other hand, ordinary generating functions arise in discrete probability as follows.

Definition 1.4.4. Let W be a finite set, and let $\mathrm{stat}\colon W \to \mathbb{Z}_{\geq 0}$ be any function. The *probability generating function* of stat on W taken uniformly at random is

$$P(q) := \frac{1}{\#W} \sum_{w \in W} q^{\mathrm{stat}(w)}.$$

A simple but important observation is the following: if X is the random variable corresponding to stat taken uniformly at random, then

$$\phi_X(s) = P(e^{is}).$$

Thus, compact expressions for generating functions of combinatorial statistics correspond exactly to compact expressions for the characteristic function of the underlying random variable. Chapter 6 and Chapter 7 largely hinge upon this interplay.

With Sara Billey and Matjaž Konvalinka, in Chapter 6 we are motivated by the following questions raised by Chapter 5.

Question 1.4.5. *Suppose $\lambda \vdash n$ and $r \in \mathbb{Z}$. Let*

$$b_{\lambda,r} := \#\{T \in \mathrm{SYT}(\lambda) : \mathrm{maj}(T) = r\}.$$

What is the approximate distribution of $b_{\lambda,r}$ for a fixed λ?

Question 1.4.6. *When is $b_{\lambda,r}$ zero?*

Proceeding experimentally, if one were to pick a random partition λ and plot a histogram of the values of maj on $\mathrm{SYT}(\lambda)$, the result would very likely be a bell curve as in Figure 6.1. The notion of asymptotic normality allows us to make this precise. Given a real-valued random variable X with mean μ and variance $\sigma^2 > 0$, let $X^* := (X - \mu)/\sigma$ be the corresponding normalized random variable with mean 0 and variance 1.

Definition 1.4.7. A sequence of real-valued random variables $X_1, X_2, \ldots$ is *asymptotically normal* if for all $t \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}[X_n^* \leq t] = \mathbb{P}[Z \leq t]$$

where Z has the standard normal distribution.

We show in Chapter 6 that, under very mild conditions, given a sequence of partitions $\lambda^{(1)}, \lambda^{(2)}, \ldots$, the sequence of major index statistics on $\mathrm{SYT}(\lambda^{(N)})$ is asymptotically normal, see Theorem 6.1.3. We more generally are able to determine the limit law for a completely arbitrary sequence of partitions, see Theorem 6.1.7. It turns out that certain degenerate families result in uniform-sum distributions rather than normal distributions. The argument involves first translating Theorem 1.1.3 into explicit expressions for the cumulants of maj on $\mathrm{SYT}(\lambda)$, see Theorem 6.1.1. We then use some constructions involving reverse standard tableaux to prove explicit asymptotic estimates for these cumulant expressions, Corollary 6.3.6 and Lemma 6.4.1. The argument is presented more generally for certain skew partitions. This allows us to give a uniform generalization of a series of asymptotic normality results due to Canfield–Janson–Zeilberger [13], Chen–Wang–Wang [16], Diaconis [21, p. 128-129], Mann–Whitney [63], and others. Future work not included in this chapter will connect its results to the distribution of irreducible constituents in the homogeneous components of the coinvariant algebras of very general complex reflection groups.

One may ask for a local limit theorem for the $b_{\lambda,r}$'s analogous to Theorem 1.4.2, though this appears quite difficult. Sufficiently powerful bounds on the characteristic function have been found in neighborhoods around -1 and 1 using an interesting generalization of Fomin–Lulov's bound to non-decomposable partitions. We provide a conjectured local limit theorem, Conjecture 6.1.10 supported by strong computational evidence.

The asymptotic description in Chapter 6 provides a satisfactory answer to Question 1.4.5 by giving us strong qualitative intuition about the distribution of maj on $\mathrm{SYT}(\lambda)$. It does not directly answer Question 1.4.6, determining when $b_{\lambda,r} \neq 0$. In Section 6.6, we answer this latter question by combinatorial manipulations on tableaux, resulting in Theorem 6.1.9. Roughly speaking, we show that there are "no internal zeroes" of the generating function $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$, except for at most two coefficients when λ is a rectangle. The proof constructs a map $\phi\colon\ \mathrm{SYT}(\lambda) - \mathcal{E}(\lambda) \to \mathrm{SYT}(\lambda)$, where $\mathcal{E}(\lambda)$ is a small set of "exceptional" tableaux, with the property that $\mathrm{maj}(\phi(T)) = \mathrm{maj}(T) + 1$. The map ϕ is explicitly computable. The key construction involves cyclically rotating a set of adjacent values of T in such a way that the result is still standard and a single descent has been incremented. Sometimes this is not possible, so we show that several additional rules may be used to handle these cases. The argument gives $\mathrm{SYT}(\lambda)$ the structure of a poset ranked, up to a shift, by maj, see Corollary 6.6.17. The approach also gives a combinatorial solution to the classification for when $a_{\lambda,r} = 0$, thereby giving an alternate proof of Theorem 5.1.3.

Returning to inv and maj in permutations, Chapter 7 discusses the following result of Baxter and Zeilberger.

Theorem 1.4.8 (Baxter–Zeilberger, [7]). *Let X_n and Y_n be the normalized random variables associated with* inv *and* maj *on S_n, respectively, taken uniformly at random. Then, for all*

$s, t \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}[X_n \le s, Y_n \le t] = \mathbb{P}[Z_1 \le s, Z_2 \le t]$$

where Z_1 and Z_2 are independent random variables with the standard normal distribution.

Informally speaking, inv and maj on S_n are jointly independently asymptotically normally distributed.

The argument in [7] involves identifying leading terms in combinatorial recursions for mixed factorial moments. Romik asked whether a generating function approach could instead be used involving a result of Roselle [77]. Zeilberger subsequently offered a \$300 reward for such an argument, which is provided by Chapter 7. Our approach translates Roselle's formula into the claim that the characteristic function of $(\mathrm{inv}, \mathrm{maj})$ on S_n is the product of the characteristic functions of inv and maj and a certain "correction factor." The argument involves identifying a sort of leading term of the correction factor and showing the remaining contributions are negligible in an appropriate sense. It uses generating function manipulatorics, Möbius inversion on the set partition lattice, and some elementary inequalities related to Stirling numbers of the first kind. One may again ask for a local limit theorem in this context, which is the subject of promising and ongoing research [96].

The rest of this thesis is organized as follows. Chapter 2 provides further background for use in later sections. Chapter 3 proves our refined cyclic sieving result, quoted as Theorem 1.2.9 above. Chapter 4 discusses our approach using cyclic sieving to Kraśkiewicz–Weyman's theorem, Theorem 1.3.1 above, and generalizations. Chapter 5 gives our argument completing Sundaram's approach to Question 1.4.1. Chapter 6 analyzes the distribution and support of maj on $\mathrm{SYT}(\lambda)$ and generalizations. Chapter 7 gives our new approach to Baxter and Zeilberger's result, Theorem 1.4.8.

Chapter 2

BACKGROUND

Here we include common background for use in later chapters supplementing that given in Chapter 1. See [91] for more.

Given a set S, we write

$$\binom{S}{k} := \{\text{all } k\text{-element subsets of } S\}, \tag{2.1}$$

$$\left(\!\!\binom{S}{k}\!\!\right) := \{\text{all } k\text{-element multisubsets of } S\}. \tag{2.2}$$

We typically use $\#S$ for the cardinality of S. Set $[n] \coloneqq \{1, 2, \ldots, n\}$. Representatives for $\mathbb{Z}/n$ will typically be taken in $[n]$.

2.1 *Words and Necklaces*

A *word* $w = w_1 w_2 \cdots w_n \in \mathbb{Z}_{\geq 1}^n$ has *letters* $w_1, \ldots, w_n$ and *length* $|w| \coloneqq n$. A sequence $\alpha = (\alpha_1, \alpha_2, \ldots)$ where $\alpha_i \in \mathbb{Z}_{\geq 0}$ sums to n is a (weak) *composition* of n, denoted $\alpha \vDash n$. The *content* of w is the composition where α_i is the number of times i appears as a letter in w. If the parts of a composition weakly decrease, it is called a *partition*, denoted $\lambda \vdash n$. We frequently omit trailing zeros when writing compositions and partitions. From Chapter 1, partitions index conjugacy classes in the symmetric group. Similarly, compositions index orbits of words under the natural symmetric group action

$$\sigma \cdot w_1 \cdots w_n \coloneqq w_{\sigma^{-1}(1)} \cdots w_{\sigma^{-1}(n)}$$

where $\sigma \in S_n$ is a permutation in the symmetric group S_n. The long cycle $\sigma_n \coloneqq (1\ 2\ \cdots\ n) \in S_n$ acts by rotation as

$$\sigma_n \cdot w_1 \cdots w_n \coloneqq w_n w_1 w_2 \cdots w_{n-1}.$$

The *descent set* of a word $w = w_1 w_2 \cdots w_n \in \mathbb{Z}_{\geq 1}^n$ is

$$\mathrm{Des}(w) \coloneqq \{i \in [n-1] : w_i > w_{i+1}\}$$

and the *major index* of w is

$$\mathrm{maj}(w) := \sum_{i \in \mathrm{Des}(w)} i.$$

The *inversion* number of w is

$$\mathrm{inv}(w) := \#\{(i,j) : i < j, w_i > w_j\}.$$

The set of all words with letters from $\mathbb{Z}_{\geq 1}$ is a monoid under concatenation. A word is *primitive* if it is not a power of a smaller word. Any non-empty word w may be written uniquely as $w = v^f$ for $f \geq 1$ with v primitive. We call $|v|$ the *period* of w, written period(w), and f the *frequency* of w, written freq(w). An orbit of a word under rotation is a *necklace*, usually denoted $[w]$. We have period$(w) = \#[w]$ and freq$(w) \cdot$ period$(w) = |w|$. Content, primitivity, period, and frequency are all constant on necklaces.

Example 2.1.1. The necklace of $w = 15531553 = (1553)^2$ is

$$[w] := \{15531553, 55315531, 53155315, 31553155\}$$

which has period 4 and frequency 2.

Given a composition $\alpha = (\alpha_1, \ldots, \alpha_m) \vDash n$ and $k \in \mathbb{Z}_{\geq 0}$ we use the following standard

q-analogues:

$$[n]_q := 1 + q + \cdots + q^{n-1} = \frac{q^n - 1}{q - 1},$$
$$[n]_q! := [n]_q[n-1]_q \cdots [1]_q,$$
$$\binom{n}{\alpha}_q := \frac{[n]_q!}{[\alpha_1]_q! \cdots [\alpha_m]_q!} \in \mathbb{Z}_{\geq 0}[q],$$
$$\left(\!\!\binom{n}{k}\!\!\right)_q := \binom{n+k-1}{k}_q := \binom{n+k-1}{k, n-1}_q.$$

2.2 Tableaux

The *size* of a partition $\lambda \vdash n$ is written $|\lambda| := n$. The *length* of λ is the number of non-zero entries and is written $\ell(\lambda)$. See Figure 1.1 for the Young diagram and the hook length of a partition. We sometimes write a partition in exponential form as $\lambda = 1^{m_1}2^{m_2}\cdots \vdash n$ where m_i is the number of rows of λ of length i. In this case, the number of elements of S_n with cycle type λ is $\frac{n!}{z_\lambda}$ where $z_\lambda := 1^{m_1}2^{m_2}\cdots m_1!m_2!\cdots$.

A *semistandard Young tableau* (briefly, a *tableau*) of shape λ is a filling of the Young diagram of λ with entries from $\mathbb{Z}_{\geq 1}$ which weakly increases along rows and strictly increases along columns. The set of semistandard Young tableaux of shape λ is denoted $\mathrm{SSYT}(\lambda)$. The *content* of $P \in \mathrm{SSYT}(\lambda)$, denoted $\mathrm{cont}(P)$, is the composition whose j-th entry is the number of j's in P. The set of *standard Young tableaux* of shape λ, denoted $\mathrm{SYT}(\lambda)$, is the subset of $\mathrm{SSYT}(\lambda)$ consisting of tableaux of content $(1,\ldots,1) \vDash n$. The *descent set* of a tableau $Q \in \mathrm{SYT}(\lambda)$, denoted $\mathrm{Des}(Q)$, is the set of all $i \in [n-1]$ such that $i+1$ lies in a lower row of Q than i. The *major index* of Q is $\mathrm{maj}(Q) := \sum_{i \in \mathrm{Des}(Q)} i$.

Example 2.2.1. The semistandard tableau

$$P = \begin{array}{|c|c|c|c|c|c|} \hline 1 & 1 & 2 & 3 & 3 & 4 \\ \hline 2 & 3 & 4 & 4 & 6 \\ \cline{1-5} 3 \\ \cline{1-1} \end{array} \in \mathrm{SSYT}(6,5,1)$$

has $\mathrm{cont}(P) = (2,2,4,3,0,1)$. The standard tableau

$$Q = \begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 \\ \cline{1-2} 6 \\ \cline{1-1}\end{array} \in \mathrm{SYT}(3,2,1)$$

has $\mathrm{Des}(Q) = \{2,5\}$, and $\mathrm{maj}(Q) = 7$.

2.3 Generating Functions

Given $\mathrm{stat}\colon W \to \mathbb{Z}_{\geq 0}$, we write the corresponding generating function as

$$W^{\mathrm{stat}}(q) \coloneqq \sum_{w \in W} q^{\mathrm{stat}(w)}.$$

We say two statistics $\mathrm{stat}, \mathrm{stat}'\colon W \to \mathbb{Z}_{\geq 0}$ are *equidistributed* on W if $W^{\mathrm{stat}}(q) = W^{\mathrm{stat}'}(q)$, and we say they are *equidistributed modulo n* on W if $W^{\mathrm{stat}}(q) \equiv W^{\mathrm{stat}'}(q) \pmod{q^n - 1}$. For example, MacMahon's result in Section 1.1 shows that inv and maj are equidistributed on S_n.

We use natural multivariable analogues of this notation as well. For example, if $\mathrm{W}_n \coloneqq \mathbb{Z}_{\geq 1}^n$ is the set of length n words, then the joint content-major index generating function is

$$\mathrm{W}_n^{\mathrm{cont},\mathrm{maj}}(\mathbf{x}; q) \coloneqq \sum_{w \in \mathrm{W}_n} x^{\mathrm{cont}(w)} q^{\mathrm{maj}(w)} \in \mathbb{Z}_{\geq 0}[[x_1, x_2, \ldots]][q]$$

where $x^\alpha \coloneqq x_1^{\alpha_1} \cdots x_m^{\alpha_m}$. To give another example, for a partition λ, the *Schur function* s_λ is the generating function

$$s_\lambda(\mathbf{x}) \coloneqq \mathrm{SSYT}(\lambda)^{\mathrm{cont}}(\mathbf{x}) \coloneqq \sum_{T \in \mathrm{SSYT}(\lambda)} x^{\mathrm{cont}(T)}.$$

This thesis is fundamentally about content and major index generating functions and their generalizations and relations. In many instances of cyclic sieving, and all of those considered in this thesis, $f(q)$ is the generating function for some statistic on W.

Chapter 3

REFINED CYCLIC SIEVING ON WORDS FOR THE MAJOR INDEX STATISTIC

This chapter is joint work with Connor Ahlbach. A version of it has been published as [5]. An earlier "extended abstract" for this work appeared in [3].

3.1 Main Results

Since Reiner, Stanton, and White introduced the *cyclic sieving phenomenon* (CSP) in 2004 [73], it has become an important companion to any cyclic action on a finite set. See Definition 1.2.5. Some remarkable examples of the CSP involve the action of a Springer regular element on Coxeter groups [73, Theorem 1.6], the action of Schutzenberger's promotion on Young tableaux of fixed rectangular shape [75], and the creation of new CSPs from old using multisets and plethysms with homogeneous symmetric functions [9, Proposition 8]. See [79] for Sagan's thorough introduction to the cyclic sieving phenomenon. More recent work on the CSP includes, for instance, [6, 70, 72].

An interesting example of the CSP is Theorem 1.2.6 above. Reiner, Stanton, and White deduced Theorem 1.2.6 from the following more general result about Coxeter systems.

Theorem 3.1.1. *[73, Theorem 1.6]. Let (W, S) be a finite Coxeter system and $J \subseteq S$. Let W_J be the corresponding parabolic subgroup, W^J the set of minimal length representatives for left cosets $X := W/W_J$, and $X^\ell(q) := \sum_{w \in W^J} q^{\ell(w)}$. Let C be a cyclic subgroup of W generated by a Springer regular element. Then $(X, C, X^\ell(q))$ exhibits the cyclic sieving phenomenon.*

Theorem 1.2.6 follows from Theorem 3.1.1 when $W = S_n$ by identifying W/W_J with words of fixed content α, where α is the composition recording the lengths of consecutive subsequences of J, and C is generated by an n-cycle. One must also use MacMahon's result, Theorem 1.1.1, to translate from inv to maj.

Our main result in this chapter, Theorem 1.2.9, is an example of the following notion.

Definition 3.1.2. A *refinement* of a CSP triple

$$(W, C_n, W^{\text{stat}}(q))$$

is a CSP triple

$$(V, C_n, V^{\text{stat}}(q))$$

where $V \subset W$ has the restricted C_n-action.

If $(V, C_n, V^{\text{stat}}(q))$ refines $(W, C_n, W^{\text{stat}}(q))$, then so does $(U, C_n, U^{\text{stat}}(q))$ where $U := W - V$. Thus, a CSP refinement partitions W into smaller CSPs with the same statistic. If W is an orbit, its only refinements are W and $\varnothing$. In Section 3.8, we define a statistic on words, *flex*, which is universal in the sense that it refines to all C_n-orbits on words of length n. We observe in Section 3.8 that such universal statistics are essentially equivalent to the choice of a total ordering for each orbit $\mathcal{O}$ of W.

As in Section 1.2, we let

$$\mathrm{W}_{\alpha,\delta} := \{\text{words } w : \operatorname{cont}(w) = \alpha, \operatorname{CDT}(w) = \delta\} \tag{3.1}$$

where $\operatorname{CDT}(w)$ is defined in Definition 1.2.7 and Example 1.2.8. Intuitively, one computes $\operatorname{CDT}(w)$ by building up w by adding all 1's, 2's, …, and counting the number of cyclic descents introduced at each step. We restate the main result of this chapter, Theorem 1.2.9 in terms of Definition 3.1.2.

Theorem 3.1.3. *Let $\alpha \vDash n$ and δ be any composition. The triple*

$$(\mathrm{W}_{\alpha,\delta}, C_n, \mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q))$$

refines the CSP triple $(\mathrm{W}_\alpha, C_n, \mathrm{W}_\alpha^{\mathrm{maj}}(q))$.

Indeed, we derive an explicit product formula for $\mathrm{W}_{\alpha,\delta}^{\mathrm{maj}}(q) \bmod (q^n - 1)$ involving q-binomial coefficients, see Theorem 3.5.19. The formula results in a q-identity similar to the Vandermonde convolution identity, see Corollary 3.5.20. The argument involves constructing $\mathrm{W}_{\alpha,\delta}$ algorithmically by recursively building a certain tree.

The two-letter case of Theorem 3.1.3 can be rephrased as follows. Fix $n \in \mathbb{Z}_{\geq 1}$ and $k, b \in \mathbb{Z}_{\geq 0}$. Let $S_{k,b}$ denote the set of subsets Δ of $\mathbb{Z}/n$ of size k where $\#\{i \in \Delta : i+1 \notin \Delta\} = b$. Define the statistic $\mathrm{mbs}\colon S_{k,b} \to \mathbb{Z}_{\geq 0}$ by identifying $\mathbb{Z}/n$ with $\{1, \ldots, n\}$ and setting $\mathrm{mbs}(\Delta) := \sum_{i \in \Delta : i+1 \notin \Delta} i$, which sums the maximum of the cyclic blocks of Δ.

Corollary 3.1.4. *The triple*

$$(S_{k,b}, C_n, S_{k,b}^{\mathrm{mbs}}(q))$$

exhibits the CSP.

Example 3.1.5. When $n = 5, k = 3, b = 2$,

$$S_{k,b} = \{\{1,2,4\}, \{2,3,5\}, \{3,4,1\}, \{4,5,2\}, \{5,1,3\}\},$$

which have mbs statistic $6, 8, 5, 7, 4$, respectively, so $S_{k,b}^{\mathrm{mbs}}(q) = q^4 + q^5 + q^6 + q^7 + q^8$. We then have $S_{k,b}^{\mathrm{mbs}}(\omega_5) = 0, S_{k,b}^{\mathrm{mbs}}(1) = 5$, in agreement with (1.3).

Theorem 8.3 in [73] and hence Theorem 1.2.6 builds on a representation-theoretic result due to Springer [86, Proposition 4.5]. Our argument is highly combinatorial, but it is not entirely bijective. Finding an explicit bijection would be quite interesting. See Section 3.8 for more details.

A key building block of our proof of Theorem 3.1.3 involves cyclic sieving on multisubsets and subsets, which was also first stated in [73]. We describe refinements of these results as well, Theorem 3.7.4 and Theorem 3.7.11, restricting to certain gcd requirements in the subset case. We present a completely different inductive proof of our subset refinement in the spirit of our proof of Theorem 3.1.3. Both our proofs of Theorem 3.1.3 and Theorem 3.7.11 use an extension lemma, Lemma 3.3.3, which allows us to extend CSPs from smaller cyclic groups to larger ones.

The rest of the chapter is organized as follows. In Section 3.2, we recall relevant background. In Section 3.3, we introduce the concept of *modular periodicity* and prove our extension lemma, Lemma 3.3.3. In Section 3.4, we discuss cyclic descent type. In Section 3.5, we decompose words with fixed content and cyclic descent type and prove a product formula for $\mathrm{W}_{\alpha,\delta}^{\mathrm{maj}}(q)$ modulo $q^n - 1$, Theorem 3.5.19. Section 3.6 uses the results of Section 3.5 to prove our main result, Theorem 3.1.3. Section 3.7 refines cyclic sieving on multisubsets and subsets with respect to shifted sum statistics. In Section 3.8, we introduce the *flex* statistic and use it to reinterpret Theorem 3.1.3.

3.2 Cyclic Descents and Cyclic Sieving

Recall the notions introduced in Section 2.1. We continue to use the alphabet of positive integers $\mathbb{Z}_{\geq 1}$ throughout unless otherwise noted. Let W_n denote the set of words of length n. Given $\alpha \vDash n$, the set W_α is the set of words of content α, which is a single orbit under the S_n-action. The *cyclic descent set* of w is $\mathrm{CDes}(w) := \{1 \leq i \leq n : w_i > w_{i+1}\}$, where the subscripts are taken modulo n, and we write $\mathrm{cdes}(w) := \#\,\mathrm{CDes}(w)$ for the number of cyclic descents. Any position $1 \leq i \leq n$ that is not a cyclic descent is a *cyclic weak ascent*.

Cyclic descents were introduced by Cellini in an algebraic context; see [15]. Since then, cyclic descents have been used by Lam and Postnikov in studying alcoved polytopes [56] and by Petersen in studying P-partitions [71]. We use lower dots between letters to indicate cyclic descents and upper dots to indicate cyclic weak ascents throughout the chapter as in the following example.

Example 3.2.1. If $w = 155.3.155.3. = 1^{\cdot}5^{\cdot}531^{\cdot}5^{\cdot}53$, then $|w| = 8$, $\mathrm{Des}(w) = \{3, 4, 7\}$, $\mathrm{des}(w) = 3$, $\mathrm{CDes}(w) = \{3, 4, 7, 8\}$, $\mathrm{cdes}(w) = 4$, $\mathrm{maj}(w) = 14$, and $\mathrm{inv}(w) = 9$.

A composition $\alpha = (\alpha_1, \ldots, \alpha_m)$ is *strong* if $\alpha_i > 0$ for all i. The cyclic group $C_n := \langle \sigma_n \rangle$

of order n generated by the long cycle $\sigma_n \coloneqq (1\ 2\ \cdots\ n)$ which acts on W_n by rotation. The necklace of a word $w \in \mathrm{W}_n$ is the C_n-orbit of w, denoted $[w]$. The notions of content, primitivity, period, frequency, and cdes are all constant on necklaces. For instance, the necklace in Example 2.1.1 has cdes 4.

Reiner-Stanton-White gave several equivalent conditions for a triple $(W, C_n, f(q))$ to exhibit the CSP. In place of (1.3) in Definition 1.2.5, we may instead require

$$f(q) \equiv \sum_{\text{orbits } \mathcal{O} \subset W} \frac{q^n - 1}{q^{n/|\mathcal{O}|} - 1} \pmod{q^n - 1}, \tag{3.2}$$

where the sum is over all orbits $\mathcal{O}$ under the action of C_n on W. Note that for $d \mid n$,

$$\frac{q^n - 1}{q^d - 1} = \sum_{i=0}^{n/d-1} q^{di} \not\equiv 0 \pmod{q^n - 1}.$$

This means every C_n-action on a finite set W gives rise to a CSP $(W, C_n, f(q))$, where $f(q)$ is the right hand side of (3.2). We refer the interested reader to [73, Proposition 2.1] for the proof of the equivalence of (1.3) and (3.2).

Remark 3.2.2. If $(V, C_n, f(q))$ exhibits the CSP, then so do both the triples $(V, C_g, f(q))$ and $(V, C_n, f(q^{-1}))$ when $g \mid n$ by (1.3). In the latter case we have relaxed the constraint $f(q) \in \mathbb{Z}_{\geq 0}[q]$ to $f(q) \in \mathbb{Z}_{\geq 0}[q, q^{-1}]$, which does no harm since (1.3) involves evaluations at roots of unity. Further, if $(V, C_n, f(q))$ and $(W, C_n, h(q))$ exhibit the CSP, then $(V \coprod W, C_n, f(q)+h(q))$ and $(V \times W, C_n, f(q)h(q))$ exhibit the CSP, where C_n acts on $V \times W$ by $\tau \cdot (v, w) := (\tau \cdot v, \tau \cdot w)$ [9, Prop. 2.2].

We write $[a, b] := \{i \in \mathbb{Z} : a \leq i \leq b\}$. Observe that the cyclic group $C_n = \langle \sigma_n \rangle$ of order n acts on $[0, n-1]$ by $\sigma_n(i) := i + 1 \pmod{n}$. This induces actions of C_n on $\binom{[0,n-1]}{k}$ and $\left(\!\binom{[0,n-1]}{k}\!\right)$ by acting on values in each subset or multisubset. For example, $\sigma_4 \cdot \{0, 0, 0, 2, 2, 3\} = \{0, 1, 1, 1, 3, 3\}$. These actions, in slightly more generality, appear in one of the original, foundational CSP results as follows.

Theorem 3.2.3. *[73, Thm. 1.1]. In the notation above, the triples*

$$\left(\binom{[0,n-1]}{k}, C_n, \binom{n}{k}_q\right) \qquad \textit{and} \qquad \left(\left(\!\!\binom{[0,n-1]}{k}\!\!\right), C_n, \left(\!\!\binom{n}{k}\!\!\right)_q\right)$$

exhibit the CSP.

We will also have use of the following principal specializations (see [59, Example I.2.2] or [91, Proposition 7.8.3]):

$$\binom{[0,n-1]}{k}^{\mathrm{sum}}(q) = e_k(1,q,q^2,\ldots,q^{n-1}) = q^{\binom{k}{2}}\binom{n}{k}_q, \tag{3.3}$$

$$\left(\!\!\binom{[0,n-1]}{k}\!\!\right)^{\mathrm{sum}}(q) = h_k(1,q,q^2,\ldots,q^{n-1}) = \left(\!\!\binom{n}{k}\!\!\right)_q. \tag{3.4}$$

Here the sum statistic denotes the sum of the elements of a subset or submultiset of $\mathbb{Z}$.

Recall from Section 1.1 MacMahon's result, that inv and maj are equidistributed on W_α. Despite this, maj and inv are not equidistributed even modulo n on $\mathrm{W}_{\alpha,\delta}$ in general, so $(\mathrm{W}_{\alpha,\delta}, C_n, \mathrm{W}^{\mathrm{inv}}_{\alpha,\delta}(q))$ does not generally exhibit the CSP. As an explicit example, an easy computation shows that $\mathrm{W}_{(2,2),(0,2)} = \{1212, 2121\}$. The corresponding major index generating function is q^2+q^4, while the inversion generating function is q^1+q^3, which are not even congruent modulo q^4-1.

3.3 Modular Periodicity and an Extension Lemma

We now introduce the concept of *modular periodicity* and use it to give an extension lemma, Lemma 3.3.3, which allows us to extend CSP's from certain subgroups to larger groups. We will verify the hypotheses of Lemma 3.3.3 in the subsequent sections to deduce Theorem 3.1.3.

Definition 3.3.1. We say a statistic $\mathrm{stat}\colon W \to \mathbb{Z}$ has *period a modulo b on W* if for all $i \in \mathbb{Z}$,

$$\#\{w \in W : \mathrm{stat}(w) \equiv_b i\} = \#\{w \in W : \mathrm{stat}(w) \equiv_b i + a\}.$$

Similarly, we say a Laurent polynomial $f(q) \in \mathbb{C}[q, q^{-1}]$ has *period a modulo b* if

$$q^a f(q) \equiv f(q) \pmod{q^b - 1},$$

or equivalently if $(q^b - 1) \mid (q^a - 1)f(q)$.

For example, $1 + 5q + q^2 + 5q^3 + q^4 + 5q^5$ has period 2 modulo 6. Note that stat has period a modulo b on W if and only if $W^{\text{stat}}(q)$ has period a modulo b. The following basic properties of modular periodicity will be useful throughout the chapter.

Lemma 3.3.2. *Let $f(q) \in \mathbb{C}[q, q^{-1}]$ and $a, b, c \in \mathbb{Z}$.*

(i) If $f(q)$ has period a modulo c and period b modulo c, then $f(q)$ has period $ua + vb$ modulo c for any $u, v \in \mathbb{Z}$. In particular, $f(q)$ has period $\gcd(a, b)$ modulo c.

(ii) If $f(q)$ has period a modulo b and period b modulo c, then $f(q)$ has period a modulo c.

(iii) If $f(q)$ has period a modulo c and $b \mid c$, then $f(q)$ has period a modulo b.

(iv) If $f(q)$ has period a modulo b, then so does $f(q)h(q)$ for any Laurent polynomial $h(q)$.

(v) If $f(q)$ has period a modulo b and $a \mid b$, then

$$f(q) \equiv \frac{a}{b}\left(\frac{q^b - 1}{q^a - 1}\right) f(q) \pmod{q^b - 1}.$$

Proof. (i), (iii), (iv), and (v) are straightforward. For (ii), suppose

$$(q^b - 1) \mid (q^a - 1)f(q), \qquad (q^c - 1) \mid (q^b - 1)f(q).$$

Write $q^c - 1 = \prod_{k=1}^{c}(q - \omega_c^k)$. If $q - \omega_c^k$ does not divide $f(q)$, then it must divide $q^b - 1$ and hence $q^a - 1$. It follows that

$$(q^c - 1) \mid (q^a - 1)f(q).$$

□

Lemma 3.3.3. *Suppose $C_n = \langle \sigma_n \rangle$ acts on W. Let $g \mid n$ and $C_g := \langle \sigma_n^{n/g} \rangle \subset C_n$. If*

(i) $(W, C_g, f(q))$ exhibits the CSP,

(ii) $f(q)$ has period g modulo n, and

(iii) for all C_n-orbits $\mathcal{O} \subset W$, we have $\frac{n}{|\mathcal{O}|} \mid g$,

then $(W, C_n, f(q))$ exhibits the CSP.

Proof. Let

$$F(q) := \sum_{C_n\text{-orbits } \mathcal{O} \subset W} \frac{q^n - 1}{q^{n/|\mathcal{O}|} - 1}.$$

By (3.2), $(W, C_n, F(q))$ exhibits the CSP, so $(W, C_g, F(q))$ also exhibits the CSP by Remark 3.2.2. Thus, by (3.2) and condition (i),

$$f(q) = F(q) + p(q)(q^g - 1) \tag{3.5}$$

for some $p(q) \in \mathbb{C}[q]$. Each summand of $F(q)$ has period g modulo n since

$$(q^n - 1) \mid (q^g - 1) \frac{q^n - 1}{q^{n/|\mathcal{O}|} - 1},$$

by condition (iii). Putting this together with condition (ii), $f(q)$ and $F(q)$ have period g modulo n. Using Lemma 3.3.2(v) twice along with (3.5) now gives

$$\begin{aligned} f(q) &\equiv \frac{g}{n} \frac{q^n - 1}{q^g - 1} f(q) \\ &= \frac{g}{n} \frac{q^n - 1}{q^g - 1} (F(q) + p(q)(q^g - 1)) \\ &\equiv \frac{g}{n} \frac{q^n - 1}{q^g - 1} F(q) \\ &\equiv F(q) \qquad \bmod (q^n - 1). \end{aligned}$$

□

3.4 Cyclic Descent Type

In this section, we more thoroughly introduce the *cyclic descent type* of a word. We also verify hypothesis (iii) of Lemma 3.3.3 for $\mathrm{W}_{\alpha,\delta}$ for a particular g; see Lemma 3.4.2.

Let $w^{(i)}$ denote the subsequence of w with all letters larger than i removed. We have a "filtration"

$$\varnothing \preceq w^{(1)} \preceq w^{(2)} \preceq \cdots \preceq w^{(m-1)} \preceq w^{(m)} = w,$$

where $u \preceq v$ means that u is a subsequence of v. We think of this filtration as building up w by recursively adding all of the copies of the next largest letter "where they fit." The cyclic descent type of a word w, denoted $\mathrm{CDT}(w)$, is the sequence which tracks the number of new cyclic descents at each stage of the filtration. Precisely, we have the following.

Definition 3.4.1. The *cyclic descent type* (CDT) of a word w is the weak composition of $\mathrm{cdes}(w)$ given by

$$\mathrm{CDT}(w) := (\mathrm{cdes}(w^{(1)}), \mathrm{cdes}(w^{(2)}) - \mathrm{cdes}(w^{(1)}), \ldots, \mathrm{cdes}(w^{(m)}) - \mathrm{cdes}(w^{(m-1)})). \tag{3.6}$$

Note that CDT is constant on necklaces since rotating w rotates each $w^{(i)}$ and cdes is constant under rotations. Furthermore, $\mathrm{cdes}(w^{(1)}) = 0$ always, so $\mathrm{CDT}(w)$ always begins with 0. Recall from (3.1) that

$$\mathrm{W}_{\alpha,\delta} := \{w \in \mathrm{W}_n : \mathrm{cont}(w) = \alpha,\ \mathrm{CDT}(w) = \delta\}.$$

We could define $\mathrm{W}_{\alpha,\delta}$ more "symmetrically" by replacing cont with "cyclic weak ascent type," which would be the point-wise difference of cont and CDT. However, content is ubiquitous in the literature, so we use it.

Lemma 3.4.2. *If $\alpha = (\alpha_1, \ldots, \alpha_m)$, $\delta = (\delta_1, \ldots, \delta_m)$, $N \subset \mathrm{W}_{\alpha,\delta}$ is a necklace, and $g :=$ gcd$(\alpha_1, \ldots, \alpha_m, \delta_1, \ldots \delta_m)$, then $\frac{n}{|N|} \mid g$.*

Proof. We can write $N = [w]$ with $w = u^{\frac{n}{|N|}}$ since $\mathrm{freq}(w) = \frac{n}{|N|}$. Hence, using pointwise multiplication,

$$\mathrm{cont}(w) = \frac{n}{|N|} \cdot \mathrm{cont}(u), \qquad \mathrm{CDT}(w) = \frac{n}{|N|} \cdot \mathrm{CDT}(u).$$

In particular, $\frac{n}{|N|}$ divides $\alpha_1, \ldots, \alpha_m, \delta_1, \ldots \delta_m$, so $\frac{n}{|N|} \mid g$.

$\square$

3.5 Runs and Falls

In this section, we give a method to algorithmically construct $\mathrm{W}_{\alpha,\delta}$ and use it to prove a product formula for $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$, Theorem 3.5.19. We conclude the section by using this formula to verify hypothesis (ii) of Lemma 3.3.3 for $\mathrm{W}_{\alpha,\delta}$; see Lemma 3.5.21.

3.5.1 A Tree Decomposition for $\mathrm{W}_{\alpha,\delta}$

We now describe a way to create words with a fixed content and CDT in terms of insertions into *runs* and *falls*. This procedure is organized into a tree, Definition 3.5.11, whose edges are labeled with sets and multisets. Lemma 3.5.8 describes changes in the major index upon traversing an edge of this tree.

Definition 3.5.1. Write $w = w_1 \cdots w_n \in \mathrm{W}_n$. A *fall* in w is a maximal set of distinct consecutive indices $i, i+1, \ldots, j-1, j$ such that $w_i > w_{i+1} > \cdots > w_j$, where we take indices modulo n. A *run* in a non-constant word w is a maximal set of distinct consecutive indices $i, i+1, \ldots, j$ such that $w_i \leq w_{i+1} \leq \cdots \leq w_j$, where we take indices modulo n. The constant word $w = \ell^n$ by convention has no runs, and it has n falls.

Note that each letter in w is part of a unique fall and a unique run, except when $w = \ell^n$ is constant. Index falls from 0 from left to right starting at the fall containing the first letter of w, and do the same with runs. It is easy to see that w has $n - \mathrm{cdes}(w)$ falls and $\mathrm{cdes}(w)$ runs, since they are separated by cyclic weak ascents and cyclic descents, respectively.

Definition 3.5.2. We write

$$F(w) := [0, |w| - \mathrm{cdes}(w) - 1] \qquad \text{and} \qquad R(w) := [0, \mathrm{cdes}(w) - 1]$$

for the indices of the falls and runs of w, respectively.

Example 3.5.3. Let $w = 26534611 = 2^{\cdot}653^{\cdot}4^{\cdot}61^{\cdot}1^{\cdot} = 26.5.346.11 \in W_8$, where upper dots indicate cyclic weak ascents and lower dots indicate cyclic descents. Since $\mathrm{cdes}(w) = 3$, we have $F(w) = [0, 4]$ and $R(w) = [0, 2]$. The 5 falls of w are $2, 653, 4, 61, 1$, with respective indices $0, 1, 2, 3, 4$. The 3 runs of w are $1126, 5, 346$, with respective indices $0, 1, 2$.

Definition 3.5.4. Fix a letter ℓ and pick a subset F of the falls $F(w)$. Assume ℓ does not appear in any of the falls in F. We *insert ℓ into falls F* by successively inserting ℓ into each fall $w_i > w_{i+1} > \cdots > w_j$ in F so that $w_i \cdots \ell \cdots w_j$ is still decreasing.

Similarly, we may fix a letter ℓ and pick a *multisubset* R of $R(w)$ (this time ℓ may already appear in a run in R). We *insert ℓ into runs R* by successively inserting ℓ into each run $w_i \leq w_{i+1} \leq \cdots \leq w_j$ in R so that $w_i \cdots \ell \cdots w_j$ is still weakly increasing.

When inserting ℓ into a run already containing ℓ, the resulting word is independent of precisely which of the possible positions is used. This is the reason we insert into runs and falls instead of positions.

Note that there is a slight ambiguity in our description of insertion into falls and runs, since it may be possible to insert either at the beginning or at the end of w while still satisfying the relevant inequalities. Given the choice, we always insert at the beginning of w.

Example 3.5.5. Let $w = 2^{\cdot}653^{\cdot}4^{\cdot}61^{\cdot}1^{\cdot}$. Insert 7 into falls of w with indices 0 and 3 to successively obtain $\underline{7}2^{\cdot}653^{\cdot}4^{\cdot}61^{\cdot}1^{\cdot}$ and then $w' := 72^{\cdot}653^{\cdot}4^{\cdot}\underline{7}61^{\cdot}1^{\cdot}$. Note that $w' = 7.26.5.347.6.11$ has two more runs (or cyclic descents) than w. Now insert 7 into the runs of w' with multiset of indices $\{0, 2, 3, 3\}$ to successively obtain $\underline{7}7.26.5.347.6.11$, $77.26.5\underline{7}.347.6.11$, $77.26.57.34\underline{7}7.6.11$, and $w'' := 77.26.57.34\underline{7}77.6.11$.

Let

$$\widetilde{W}_n = \{w \in \mathrm{W}_n : w \text{ ends in a } 1\}, \tag{3.7}$$

$$\widetilde{W}_{\alpha,\delta} = \{w \in \mathrm{W}_{\alpha,\delta} : w \text{ ends in a } 1\}. \tag{3.8}$$

We restrict to $\widetilde{W}_n$ and $\widetilde{W}_{\alpha,\delta}$ since the major index generating function is easier to find and extends to $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q) \pmod{q^n - 1}$.

Definition 3.5.6. Fix $w \in \widetilde{W}_n$, a letter ℓ not in w, and

$$F \subset F(w) = [0, |w| - \mathrm{cdes}(w) - 1] \qquad \text{and} \qquad R \underset{\text{mult.}}{\subset} [0, \mathrm{cdes}(w) + |F| - 1]$$

where $\underset{\text{mult.}}{\subset}$ denotes a multisubset. Let w' be obtained by inserting ℓ into falls F of w. Note that $[0, \mathrm{cdes}(w)+|F|-1] = R(w')$ indexes the runs of w'. Now let w'' be obtained by inserting ℓ into runs R of w'. We say w'' is obtained by *inserting the triple* (ℓ, F, R) into w. Observe that $\mathrm{cdes}(w'') = \mathrm{cdes}(w') = \mathrm{cdes}(w) + |F|$ and $w'' \in \widetilde{W}_{n+|F|+|R|}$.

First we define the cyclic descent type $\delta = \mathrm{CDT}(w)$ of any $w \in \mathrm{W}_\alpha$. Then we give a product formula for $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ modulo $q^n - 1$, Theorem 3.5.19. The $q = 1$ specialization gives a formula for $\# \mathrm{W}_{\alpha,\delta}$, Proposition 3.5.16. Along the way, we describe how to build words in $\mathrm{W}_{\alpha,\delta}$ by walking along a tree whose edges are labeled by sets and multisets. We describe a fixed point lemma arising from the tree which will play a key role in Section 3.6. We also introduce the notion of modular periodicity in order to transfer certain results between different cyclic groups or moduli.

We next describe the effect of inserting a single letter on maj. We restrict to $\widetilde{W}_n$ so we preserve a cyclic weak ascent at the end and never add a letter to the end. The fact that the increments in major index from inserting a new letter into all possible positions form a permutation was first observed by Gupta [40]. Lemma 3.5.7 tells us exactly the increment in major index based on which run or fall the newly inserted letter fits into.

Lemma 3.5.7. *Suppose $w' \in \widetilde{W}_{n+1}$ is obtained by adding a letter ℓ to $w \in \widetilde{W}_n$ in any position. Then w' is obtained by inserting ℓ into some run or fall of w, and*

$$\mathrm{maj}(w') - \mathrm{maj}(w) = \begin{cases} \mathrm{cdes}(w) - r & \text{if } \ell \text{ is inserted into run } r \text{ of } w \\ \mathrm{cdes}(w) + 1 + f & \text{if } \ell \text{ is inserted into fall } f \text{ of } w. \end{cases} \tag{3.9}$$

Proof. If $\mathrm{cdes}(w') = \mathrm{cdes}(w)$, then ℓ is inserted into some run of w, and otherwise $\mathrm{cdes}(w') = \mathrm{cdes}(w) + 1$ and ℓ is inserted into some fall of w. Inserting ℓ into run r of w will increment the position of $\mathrm{cdes}(w) - r$ descents by 1 each, so

$$\mathrm{maj}(w') - \mathrm{maj}(w) = \mathrm{cdes}(w) - r.$$

Let $\mathrm{comaj}(w) := 1 + 2 + \cdots + (|w| - 1) - \mathrm{maj}(w)$, which is the sum of $i \in [|w| - 1]$ where $w_i \leq w_{i+1}$. Inserting ℓ into fall f of w will increment the position of $(|w| - 1) - \mathrm{cdes}(w) - f$ weak ascents by 1 each, so

$$\mathrm{comaj}(w') - \mathrm{comaj}(w) = (|w| - 1) - \mathrm{cdes}(w) - f,$$

from which it follows that

$$\mathrm{maj}(w') - \mathrm{maj}(w) = \mathrm{cdes}(w) + 1 + f.$$

□

Lemma 3.5.8. *Suppose w'' is obtained by inserting the triple (ℓ, F, R) into $w \in \widetilde{W}_n$. Then*

$$\mathrm{maj}(w'') - \mathrm{maj}(w) = \binom{|F| + 1}{2} + (\mathrm{cdes}(w))(|F| + |R|) + |F||R| + \sum_{f \in F} f - \sum_{r \in R} r. \tag{3.10}$$

Proof. Let w' be obtained by inserting ℓ into falls F of w. It suffices to show

$$\operatorname{maj}(w') - \operatorname{maj}(w) = \binom{|F|+1}{2} + (\operatorname{cdes}(w))|F| + \sum_{f \in F} f \tag{3.11}$$

and

$$\operatorname{maj}(w'') - \operatorname{maj}(w') = (\operatorname{cdes}(w'))|R| - \sum_{r \in R} r \tag{3.12}$$

since $\operatorname{cdes}(w') = \operatorname{cdes}(w'') = \operatorname{cdes}(w) + |F|$. Both (3.11) and (3.12) follow from iterating Lemma 3.5.7 and recalling cdes is incremented by 1 each time we insert into a fall. □

Notation 3.5.9. For the rest of this section, fix a strong composition $\alpha = (\alpha_1, \ldots, \alpha_m)$ of $n \geq 1$ and $\delta = (\delta_1, \ldots, \delta_m) \vDash k$ with $\delta_1 = 0$. We emphasize that α and δ have the same number, m, of parts. For $\ell = 1, \ldots, m$, let

$$n_\ell := \alpha_1 + \cdots + \alpha_\ell, \tag{3.13}$$

$$k_\ell := \delta_1 + \cdots + \delta_\ell. \tag{3.14}$$

For $w \in \mathrm{W}_{\alpha,\delta}$, we have the defining conditions $|w^{(\ell)}| = n_\ell$ and $\operatorname{cdes}(w^{(\ell)}) = k_\ell$. Furthermore, let

$$S_\ell := \binom{[0, n_{\ell-1} - k_{\ell-1} - 1]}{\delta_\ell}, \qquad M_\ell := \left(\!\!\binom{[0, k_\ell - 1]}{\alpha_\ell - \delta_\ell}\!\!\right)$$

and

$$g := \gcd(\alpha_1, \alpha_2, \ldots, \alpha_m, \delta_1, \delta_2, \ldots, \delta_m).$$

If $w \in \widetilde{W}_{\alpha,\delta}$, then the set S_ℓ consists of all subsets of the falls $F(w^{(\ell-1)})$ which, when ℓ is inserted into those falls of $w^{(\ell-1)}$, result in a word w' with k_ℓ cyclic descents. The multiset M_ℓ similarly consists of all choices of runs $R(w')$ which, when ℓ is inserted into those runs, result in a word with length n_ℓ.

Remark 3.5.10. We restrict to strong compositions α for notational simplicity, though the results in this section may easily be generalized to arbitrary weak compositions by “flattening” weak compositions to strong ones by removing zeros.

Definition 3.5.11. Construct a rooted, vertex-labeled and edge-labeled tree $T_{\alpha,\delta}$ recursively as follows. Begin with a tree $T^{(1)}$ containing only a root labeled by the word 1^{α_1}. For $\ell = 2, \ldots, m$, to obtain $T^{(\ell)}$, do the following. For each leaf w of $T^{(\ell-1)}$ and for each triple (ℓ, F, R) with

$$F \in S_\ell \qquad \text{and} \qquad R \in M_\ell,$$

add an edge labeled by (F, R) to $T^{(\ell-1)}$ from w to w'' where w'' is obtained by inserting (ℓ, F, R) into w. Define $T_{\alpha,\delta} := T^{(m)}$.

Example 3.5.12. Let $\alpha = (4, 2, 3)$ and $\delta = (0, 2, 1)$. Figure 3.1 is the subgraph of $T_{\alpha,\delta}$ consisting of paths from the root to leaves that are rotations of 112113323.

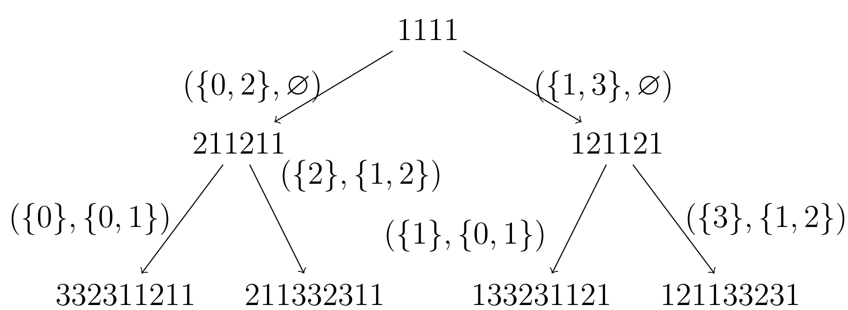


Figure 3.1: Subgraph of tree $T_{\alpha,\delta}$ with $\alpha = (4, 2, 3)$, $\delta = (0, 2, 1)$.

For this full $T_{\alpha,\delta}$, the root has $\binom{4}{2} = 6$ children since 1111 has 4 falls. Each child of the root itself has $\binom{4}{1}\left(\!\binom{3}{2}\!\right) = 24$ children. Hence, $T_{\alpha,\delta}$ has 144 leaves. Notice that the cyclic rotations of 311211332 appearing as leaves in Figure 3.1 are precisely those ending in 1. It will shortly become apparent that in this example, $\#\,\mathrm{W}_{\alpha,\delta} = \frac{9}{4} \cdot 144 = 324$.

Lemma 3.5.13. *The vertices of $T_{\alpha,\delta}$ which are $\ell < m$ edges away from the root are precisely the elements of $\{w^{(\ell+1)} : w \in \widetilde{W}_{\alpha,\delta}\}$, each occurring once. In particular, the leaves of $T_{\alpha,\delta}$ are precisely the elements of $\widetilde{W}_{\alpha,\delta}$, each occurring once.*

Proof. By definition of S_ℓ and M_ℓ, any leaf of $T^{(\ell)}$ has content $(\alpha_1, \ldots, \alpha_\ell)$, cyclic descent type $(\delta_1, \ldots, \delta_\ell)$, and ends in a 1, so is in $\{w^{(\ell)} : w \in \widetilde{W}_{\alpha,\delta}\}$. Conversely, given any $w \in \widetilde{W}_{\alpha,\delta}$, the word $w^{(\ell)}$ is obtained by inserting a unique triple (ℓ, F, R) into $w^{(\ell-1)}$ by repeated applications of Lemma 3.5.7. □

Definition 3.5.14. By Lemma 3.5.13, the tree $T_{\alpha,\delta}$ encodes a bijection

$$\Phi \colon \widetilde{W}_{\alpha,\delta} \xrightarrow{\sim} \prod_{\ell=2}^{m} S_\ell \times M_\ell$$

given by reading the edge labels from the root to w. We suppress the dependence of Φ on α and δ from the notation since they can be computed from the input w.

Lemma 3.5.15. *For any $w \in \mathrm{W}_{\alpha,\delta}$,*

$$\#[w] = \frac{n}{\alpha_1} \cdot \# \left([w] \cap \widetilde{W}_{\alpha,\delta} \right). \tag{3.15}$$

Consequently,

$$\# \, \mathrm{W}_{\alpha,\delta} = \frac{n}{\alpha_1} \# \widetilde{W}_{\alpha,\delta}. \tag{3.16}$$

Proof. Each $w \in \mathrm{W}_{\alpha,\delta}$ has $\mathrm{period}(w) = n / \mathrm{freq}(w)$ distinct cyclic rotations, of which $\alpha_1 / \mathrm{freq}(w)$ end in 1. □

Proposition 3.5.16. *Using Notation 3.5.9, we have*

$$\# \, \mathrm{W}_{\alpha,\delta} = \frac{n}{\alpha_1} \prod_{\ell=2}^{m} \binom{n_{\ell-1} - k_{\ell-1}}{\delta_\ell} \left(\!\!\binom{k_\ell}{\alpha_\ell - \delta_\ell}\!\!\right). \tag{3.17}$$

In particular, $\mathrm{W}_{\alpha,\delta} \neq \varnothing$ if and only if

$$\begin{aligned} 0 \le \delta_\ell &\le \alpha_\ell & &\text{for all } 1 \le \ell \le m, \text{ and} \\ \delta_1 + \cdots + \delta_{\ell+1} &\le \alpha_1 + \cdots + \alpha_\ell & &\text{for all } 1 \le \ell < m. \end{aligned} \tag{3.18}$$

Proof. The product in (3.17) is $\#\prod_{\ell=2}^{m} S_\ell \times M_\ell$, which is $\#\widetilde{W}_{\alpha,\delta}$ by the bijection Φ. Now (3.17) follows from (3.16), and (3.18) follows from (3.17). □

3.5.2 Major Index Generating Functions

We next use the bijection Φ and Lemma 3.5.8 to give a product formula for $\widetilde{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$, Theorem 3.5.17. We then use modular periodicity to obtain an analogous expression for $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ modulo $q^n - 1$, Theorem 3.5.19.

Theorem 3.5.17. *Using Notation 3.5.9, we have*

$$\widetilde{W}^{\mathrm{maj}}_{\alpha,\delta}(q) = \prod_{\ell=2}^{m} q^{k_\ell \alpha_\ell} \binom{n_{\ell-1} - k_{\ell-1}}{\delta_\ell}_q \left(\!\!\binom{k_\ell}{\alpha_\ell - \delta_\ell}\!\!\right)_{q^{-1}} \tag{3.19}$$

$$= q^{\eta(\alpha,\delta)} \prod_{\ell=2}^{m} \binom{n_{\ell-1} - k_{\ell-1}}{\delta_\ell}_q \left(\!\!\binom{k_\ell}{\alpha_\ell - \delta_\ell}\!\!\right)_{q} \tag{3.20}$$

where

$$\eta(\alpha,\delta) := n - \alpha_1 + \binom{k}{2} + \sum_{\ell=2}^{m} \binom{\delta_\ell}{2}.$$

Proof. Combining Φ with Lemma 3.5.8 shows that

$$\widetilde{W}^{\mathrm{maj}}_{\alpha,\delta}(q) = \prod_{\ell=2}^{m} \sum_{\substack{F \in S_\ell \\ R \in M_\ell}} q^{\epsilon(\ell,F,R)}, \tag{3.21}$$

where

$$\epsilon(\ell,F,R) := \binom{\delta_\ell + 1}{2} + k_{\ell-1}\alpha_\ell + \delta_\ell(\alpha_\ell - \delta_\ell) + \mathrm{sum}(F) - \mathrm{sum}(R).$$

Noting that

$$\binom{\delta_\ell + 1}{2} + k_{\ell-1}\alpha_\ell + \delta_\ell(\alpha_\ell - \delta_\ell) = k_\ell \alpha_\ell - \binom{\delta_\ell}{2},$$

simplifying (3.21) gives

$$\widetilde{W}^{\mathrm{maj}}_{\alpha,\delta}(q) = \prod_{\ell=2}^{m} q^{k_\ell \alpha_\ell - \binom{\delta_\ell}{2}} S_\ell^{\mathrm{sum}}(q) M_\ell^{\mathrm{sum}}(q^{-1}). \tag{3.22}$$

Equation (3.19) now follows from (3.3), (3.4), and the definition of S_ℓ and M_ℓ. As for (3.20), consider the reversal bijection $r \colon M_\ell \to M_\ell$ induced by

$$x \mapsto k_\ell - 1 - x$$

on $[0, k_\ell - 1]$. This bijection satisfies $\mathrm{sum}(r(A)) = (k_\ell - 1)(\alpha_\ell - \delta_\ell) - \mathrm{sum}(A)$, so

$$M_\ell^{\mathrm{sum}}(q^{-1}) = q^{-(k_\ell - 1)(\alpha_\ell - \delta_\ell)} M_\ell^{\mathrm{sum}}(q). \tag{3.23}$$

Plugging (3.23) into (3.22) and noting that

$$\begin{aligned}
\sum_{\ell=2}^{m} \left(k_\ell \alpha_\ell - \binom{\delta_\ell}{2} - (k_\ell - 1)(\alpha_\ell - \delta_\ell) \right) &= \sum_{\ell=2}^{m} \left(\alpha_\ell - \frac{\delta_\ell}{2} - \frac{\delta_\ell^2}{2} + k_\ell \delta_\ell \right) \\
&= n - \alpha_1 - \frac{k}{2} + \sum_{\ell=2}^{m} \left(-\frac{\delta_\ell^2}{2} + \sum_{j=2}^{\ell} \delta_j \delta_\ell \right) \\
&= n - \alpha_1 - \frac{k}{2} + \frac{1}{2} \sum_{\ell=2}^{m} \sum_{j=2}^{m} \delta_j \delta_\ell \\
&= n - \alpha_1 - \frac{k}{2} + \frac{k^2}{2}
\end{aligned}$$

gives

$$\widetilde{W}^{\mathrm{maj}}_{\alpha,\delta}(q) = q^{n - \alpha_1 + \binom{k}{2}} \prod_{\ell=2}^{m} S_\ell^{\mathrm{sum}}(q) M_\ell^{\mathrm{sum}}(q).$$

Using (3.3) and (3.4) now yields (3.20). □

Lemma 3.5.18. *Let $\alpha \vDash n$, $\delta \vDash k$. The statistic* maj *has period k modulo n on* $\mathrm{W}_{\alpha,\delta}$.

Moreover, maj *is constant modulo* $d := \gcd(n,k)$ *on necklaces in* $\mathrm{W}_{\alpha,\delta}$*, and*

$$W^{\mathrm{maj}}_{\alpha,\delta}(q) \equiv \frac{n}{\alpha_1}\widetilde{W}^{\mathrm{maj}}_{\alpha,\delta}(q) \qquad (\textit{mod } q^d - 1). \tag{3.24}$$

Proof. Since cyclically rotating $w \in \mathrm{W}_{\alpha,\delta}$ increments each cyclic descent by 1 modulo n, we have

$$\mathrm{maj}(\sigma_n \cdot w) \equiv_n \mathrm{maj}(w) + k. \tag{3.25}$$

In particular, maj has period k modulo n on necklaces in $\mathrm{W}_{\alpha,\delta}$. Furthermore, maj is constant on necklaces in $\mathrm{W}_{\alpha,\delta}$ modulo d. Now (3.24) follows from (3.15). ☐

Theorem 3.5.19. *Using Notation 3.5.9, let* $d := \gcd(n,k)$*. Then, modulo* $q^n - 1$*,*

$$\begin{aligned} \mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q) &\equiv \frac{d}{\alpha_1}\left(\frac{q^n-1}{q^d-1}\right)\prod_{\ell=2}^{m} q^{k_\ell \alpha_\ell}\binom{n_{\ell-1}-k_{\ell-1}}{\delta_\ell}_q \left(\!\!\binom{k_\ell}{\alpha_\ell-\delta_\ell}\!\!\right)_{q^{-1}} \\ &\equiv \frac{d}{\alpha_1}\left(\frac{q^n-1}{q^d-1}\right) q^{\binom{k}{2}+\sum_{\ell=2}^{m}\binom{\delta_\ell}{2}-\alpha_1}\prod_{\ell=2}^{m}\binom{n_{\ell-1}-k_{\ell-1}}{\delta_\ell}_q \left(\!\!\binom{k_\ell}{\alpha_\ell-\delta_\ell}\!\!\right)_{q}. \end{aligned} \tag{3.26}$$

Proof. By Lemma 3.5.18, maj has period k modulo n on $\mathrm{W}_{\alpha,\delta}$. Hence by Lemma 3.3.2(i), maj has period d modulo n on $\mathrm{W}_{\alpha,\delta}$. Using Lemma 3.3.2(v) and (3.24) gives

$$\begin{aligned} \mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q) &\equiv \frac{d}{n}\left(\frac{q^n-1}{q^d-1}\right)\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q) \\ &\equiv \frac{d}{n}\left(\frac{q^n-1}{q^d-1}\right)\left(\frac{n}{\alpha_1}\widetilde{W}^{\mathrm{maj}}_{\alpha,\delta}(q) + p(q)(q^d-1)\right) \\ &\equiv \frac{d}{\alpha_1}\left(\frac{q^n-1}{q^d-1}\right)\widetilde{W}^{\mathrm{maj}}_{\alpha,\delta} \qquad (\mathrm{mod}\, q^n - 1), \end{aligned}$$

where $p(q) \in \mathbb{C}[q]$. Theorem 3.5.19 now follows from Theorem 3.5.17. ☐

Corollary 3.5.20. *Using Notation 3.5.9, let* $d := \gcd(n,k)$*. Then, modulo* $q^n - 1$*,*

$$\mathrm{W}_{\alpha}^{\mathrm{maj}}(q) = \binom{n}{\alpha}_q \equiv \sum_{\delta} \frac{d}{\alpha_1} \left(\frac{q^n-1}{q^d-1}\right) \prod_{\ell=2}^{m} q^{k_\ell \alpha_\ell} \binom{n_{\ell-1}-k_{\ell-1}}{\delta_\ell}_q \left(\!\!\binom{k_\ell}{\alpha_\ell - \delta_\ell}\!\!\right)_{q^{-1}} \tag{3.27}$$

where the sum is over weak compositions δ *of* k *satisfying* (3.18)*. In particular,*

$$\#\,\mathrm{W}_\alpha = \binom{n}{\alpha} = \sum_{\delta} \frac{n}{\alpha_1} \prod_{\ell=2}^{m} \binom{n_{\ell-1}-k_{\ell-1}}{\delta_\ell} \left(\!\!\binom{k_\ell}{\alpha_\ell - \delta_\ell}\!\!\right). \tag{3.28}$$

Note that the two-letter case of (3.28) is a special case of the classical Vandermonde convolution identity [92, Ex. 1.1.17].

3.5.3 Verifying Hypothesis (ii) of Lemma 3.3.3 for $\mathrm{W}_{\alpha,\delta}$

Lemma 3.5.21. *Using Notation 3.5.9,* $\mathrm{W}_{\alpha,\delta}^{\mathrm{maj}}(q)$ *has period* g *modulo* n*.*

Proof. Let $d = \gcd(n,k)$. By Theorem 3.5.19,

$$\mathrm{W}_{\alpha,\delta}^{\mathrm{maj}}(q) \equiv \frac{d}{\alpha_1} \left(\frac{q^n-1}{q^d-1}\right) \prod_{\ell=2}^{m} q^{k_\ell \alpha_\ell} \binom{n_{\ell-1}-k_{\ell-1}}{\delta_\ell}_q \left(\!\!\binom{k_\ell}{\alpha_\ell - \delta_\ell}\!\!\right)_{q^{-1}}$$

modulo $q^n - 1$. The action of rotation on elements of $S_\ell = \binom{[0, n_{\ell-1}-k_{\ell-1}]}{\delta_\ell}$ increases their sum by δ_ℓ modulo $n_{\ell-1} - k_{\ell-1}$. Thus by (3.3), $\binom{n_{\ell-1}-k_{\ell-1}}{\delta_\ell}_q$ has period δ_ℓ modulo $n_{\ell-1} - k_{\ell-1}$. Similarly by (3.4), $\left(\!\binom{k_\ell}{\alpha_\ell-\delta_\ell}\!\right)_{q^{-1}}$ has period $\alpha_\ell - \delta_\ell$ modulo k_ℓ. For $\ell = 2, \ldots, m$, by Lemma 3.3.2(iv) we then have

$$\mathrm{W}_{\alpha,\delta}^{\mathrm{maj}}(q) \text{ has period } \delta_\ell \text{ modulo } n_{\ell-1} - k_{\ell-1}, \text{ and} \tag{3.29}$$

$$\mathrm{W}_{\alpha,\delta}^{\mathrm{maj}}(q) \text{ has period } \alpha_\ell - \delta_\ell \text{ modulo } k_\ell. \tag{3.30}$$

We show $\mathrm{W}_{\alpha,\delta}^{\mathrm{maj}}(q)$ has period α_ℓ and δ_ℓ modulo n by downward induction on ℓ, for $m \geq \ell \geq 2$. Note that the base case $\ell = m$ is accounted for by our argument as well.

Suppose $\mathrm{W}_{\alpha,\delta}^{\mathrm{maj}}(q)$ has period α_j and δ_j modulo n for all $j > \ell$. By Lemma 3.5.18, $\mathrm{W}_{\alpha,\delta}^{\mathrm{maj}}(q)$

has period k modulo n. By Lemma 3.3.2(i), $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ thus has period

$$k_\ell = k - (\delta_m + \cdots + \delta_{\ell+1})$$

modulo n. Since $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ has period $\alpha_\ell - \delta_\ell$ modulo k_ℓ, $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ has period $\alpha_\ell - \delta_\ell$ modulo n by Lemma 3.3.2(ii).

As noted, $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ has period δ_ℓ modulo $n_{\ell-1} - k_{\ell-1}$. By Lemma 3.3.2(i), $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ also has period

$$n_{\ell-1} - k_{\ell-1} = n - (\alpha_m + \cdots + \alpha_{\ell+1}) - k + (\delta_m + \cdots + \delta_{\ell+1}) - (\alpha_\ell - \delta_\ell)$$

modulo n. Hence, as $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ has period δ_ℓ modulo $n_{\ell-1} - k_{\ell-1}$, $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ has period δ_ℓ modulo n by Lemma 3.3.2(ii). By another application of Lemma 3.3.2(i), $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ has period α_ℓ modulo n as well, completing the induction.

Indeed, $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ has period $\delta_1 = 0$ modulo n trivially, and $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ has period $\alpha_1 = n - (\alpha_m + \cdots + \alpha_2)$ modulo n by Lemma 3.3.2(i). Putting everything together, $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ has periods $\alpha_1, \ldots, \alpha_m, \delta_1, \ldots, \delta_m$ modulo n, so by one more application of Lemma 3.3.2(i), $\mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ has period g modulo n. □

3.6 Refining the CSP to Fixed Content and Cyclic Descent Type

In this section, we verify the final hypothesis (i) of Lemma 3.3.3 for $\mathrm{W}_{\alpha,\delta}$ and deduce Theorem 3.1.3. Throughout this section we continue to follow Notation 3.5.9. We recall in particular that

$$S_\ell := \binom{[0, n_{\ell-1} - k_{\ell-1} - 1]}{\delta_\ell}, \qquad M_\ell := \left(\!\!\binom{[0, k_\ell - 1]}{\alpha_\ell - \delta_\ell}\!\!\right)$$

and

$$g := \gcd(\alpha_1, \ldots, \alpha_m, \delta_1, \ldots, \delta_m).$$

3.6.1 A Fixed Point Lemma

To prove our main result, Theorem 3.1.3, one approach would be to find a C_n-equivariant isomorphism between a known CSP triple and $(\mathrm{W}_{\alpha,\delta}, C_n, \mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q))$. Such a triple is hinted at by (3.19) and the bijection Φ using products of CSP's coming from Theorem 3.2.3, though the approach encounters immediate difficulties. For instance, $\widetilde{W}_{\alpha,\delta}$ is not generally closed under the C_n-action. In this section, we instead give a fixed point lemma, Lemma 3.6.5, which is intuitively a weakened version of the equivariant isomorphism approach.

Definition 3.6.1. Since $g \mid n_{\ell-1} - k_{\ell-1}$, $g \mid k_\ell$, and $g \mid n$, C_g acts on each of S_ℓ, M_ℓ, and $\mathrm{W}_{\alpha,\delta}$ by restricting the actions of $C_{n_{\ell-1}-k_{\ell-1}}$, C_{k_ℓ}, and C_n to their unique subgroups of size g. For instance, the action of C_g on $\mathrm{W}_{\alpha,\delta}$ is generated by rotation by n/g.

We further let C_g act diagonally on the products $S_\ell \times M_\ell$ and $\prod_{\ell=2}^m S_\ell \times M_\ell$. We emphasize that despite having C_g-actions on $\mathrm{W}_{\alpha,\delta}$ and $\prod_{\ell=2}^m S_\ell \times M_\ell$, the bijection $\Phi \colon \widetilde{W}_{\alpha,\delta} \xrightarrow{\sim} \prod_{\ell=2}^m S_\ell \times M_\ell$ is not in general equivariant since $\widetilde{W}_{\alpha,\delta}$ is not closed under the C_g action on $\mathrm{W}_{\alpha,\delta}$.

Definition 3.6.2. Given a multisubset of some set $[0, a]$, we may encode it as a multiplicity word $w_0 w_1 \ldots w_a$ where w_i is the multiplicity of i. In particular, we may consider the bijection $\Phi \colon \widetilde{W}_{\alpha,\delta} \xrightarrow{\sim} \prod_{\ell=2}^m S_\ell \times M_\ell$ as mapping words to sequences of pairs of certain words.

Example 3.6.3. Consider the leaf $w = 211332311$ in Figure 3.1 from Example 3.5.12. Reading edge labels gives $\Phi(w) = ((\{0,2\}, \varnothing), (\{2\}, \{1,2\}))$. Recalling that S_2 consists of subsets of $[0, 4-1]$, M_2 consists of multisubsets of $\varnothing$, S_3 consists of subsets of $[0, 4-1]$, and M_3 consists of multisubsets of $[0, 3-1]$, the corresponding sequence of words is $((1010, \epsilon), (0010, 011))$, where ϵ denotes the empty word. Table 3.1 summarizes several similar translations.

Lemma 3.6.4. *Suppose $w = u^k$ for some word u. If $\Phi(u) = ((x_2, y_2), \ldots, (x_m, y_m))$ encoded as multiplicity words as in Definition 3.6.2, then $\Phi(w) = ((x_2^k, y_2^k), \ldots, (x_m^k, y_m^k))$.*

Proof. The insertion triples needed to build w are the sequences of k shifted copies of the insertion triples needed to build u. $\square$

w	$\Phi(w)$	sequence of pairs of words
211332311	$((\{0,2\},\varnothing),(\{2\},\{1,2\}))$	$((1010,\epsilon),(0010,011))$
121133231	$((\{1,3\},\varnothing),(\{3\},\{1,2\}))$	$((0101,\epsilon),(0001,011))$
$(211332311)^2$	$((\{0,2,4,6\},\varnothing),(\{2,6\},\{1,2,4,5\}))$	$((1010^2,\epsilon),(0010^2,011^2))$
2221123311	$((\{0,2\},\{0,0\}),(\varnothing,\{1,1\}))$	$((1010,20),(000000,02))$

Table 3.1: Examples of words, corresponding sequences of edge labels in $T_{\alpha,\delta}$, and corresponding sequences of words. Note that the second word is a cyclic rotation of the first.

Lemma 3.6.5. *An element $\tau \in C_g$ fixes $w \in \widetilde{W}_{\alpha,\delta}$ if and only if τ fixes $\Phi(w)$.*

Proof. For $\tau \in C_n$, let $o(\tau)$ denote the order of τ. It is easy to see that $\tau \in C_n$ fixes $w \in \mathrm{W}_n$ if and only if there is some word u such that $w = u^{o(\tau)}$.

Suppose $\tau \in C_g$ fixes w, so that $w = u^{o(\tau)}$. By Lemma 3.6.4,

$$\Phi(w) = ((x_2^{o(\tau)}, y_2^{o(\tau)}), \dots, (x_m^{o(\tau)}, y_m^{o(\tau)})).$$

Each of the words $x_i^{o(\tau)}$ and $y_i^{o(\tau)}$ is fixed by τ, so $\Phi(w)$ is fixed by τ. The reverse implication follows analogously using the fact that Φ is a bijection. $\square$

3.6.2 Verifying Hypothesis (i) of Lemma 3.3.3 for $\mathrm{W}_{\alpha,\delta}$

Lemma 3.6.6. *Using Notation 3.5.9,* $(\mathrm{W}_{\alpha,\delta}, C_g, \mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q))$ *exhibits the CSP.*

Proof. We use the notation and actions in Definition 3.6.1. Recall that

$$S_\ell := \binom{[0, n_{\ell-1} - k_{\ell-1} - 1]}{\delta_\ell}, \qquad M_\ell := \left(\!\!\binom{[0, k_\ell - 1]}{\alpha_\ell - \delta_\ell}\!\!\right).$$

From Theorem 3.2.3, for each $2 \le \ell \le m$, $\left(S_\ell, C_g, \binom{n_\ell - k_\ell}{\delta_\ell}_q\right)$ and $\left(M_\ell, C_g, \left(\!\binom{k_\ell}{\alpha_\ell - \delta_\ell}\!\right)_{q^{-1}}\right)$ exhibit the CSP. Taking products,

$$\left(\prod_{\ell=2}^{m} S_\ell \times M_\ell, C_g, \prod_{\ell=2}^{m} \binom{n_\ell - k_\ell}{\delta_\ell}_q \left(\!\!\binom{k_\ell}{\alpha_\ell - \delta_\ell}\!\!\right)_{q^{-1}}\right)$$

exhibits the CSP. Comparing this to Theorem 3.5.17, we have

$$\widetilde{W}^{\mathrm{maj}}_{\alpha,\delta} \equiv \prod_{\ell=2}^{m} \binom{n_\ell - k_\ell}{\delta_\ell}_q \left(\!\!\binom{k_\ell}{\alpha_\ell - \delta_\ell}\!\!\right)_{q^{-1}}$$

modulo $q^g - 1$, as $\sum_{\ell=2}^{m} k_\ell \alpha_\ell \equiv_g 0$ because $g \mid \alpha_\ell$ for all ℓ. Thus,

$$\left(\prod_{\ell=2}^{m} S_\ell \times M_\ell, C_g, \widetilde{W}^{\mathrm{maj}}_{\alpha,\delta}(q)\right) \tag{3.31}$$

exhibits the CSP.

By Lemma 3.5.15, for any $w \in \mathrm{W}_{\alpha,\delta}$,

$$\#[w] = \frac{n}{\alpha_1} \cdot \# \left([w] \cap \widetilde{W}_{\alpha,\delta}\right).$$

Since $[w]$ is an orbit under C_n, an element $\tau \in C_n$ fixes w if and only if τ fixes $[w]$ pointwise. Thus, for any $\tau \in C_n$,

$$\#\, \mathrm{W}^{\tau}_{\alpha,\delta} = \frac{n}{\alpha_1} \cdot \# \widetilde{W}^{\tau}_{\alpha,\delta}. \tag{3.32}$$

Combining (3.32) and Lemma 3.6.5 now shows that for any $\tau \in C_g$,

$$\#\, \mathrm{W}^{\tau}_{\alpha,\delta} = \frac{n}{\alpha_1} \cdot \# \left(\prod_{\ell=2}^{m} S_\ell \times M_\ell\right)^{\tau}. \tag{3.33}$$

Hence, by (3.33), the CSP in (3.31), and (1.3),

$$\left(\mathrm{W}_{\alpha,\delta}, C_g, \frac{n}{\alpha_1} \widetilde{W}^{\mathrm{maj}}_{\alpha,\delta}(q)\right)$$

exhibits the CSP. By (3.24), $\frac{n}{\alpha_1}\widetilde{W}_{\alpha,\delta}(q) \equiv \mathrm{W}^{\mathrm{maj}}_{\alpha,\delta}(q)$ modulo $q^d - 1$, hence also modulo $q^g - 1$ since $g \mid d$, completing the proof. $\square$

We have now finished the verification of the conditions in Lemma 3.3.3 for $\mathrm{W}_{\alpha,\delta}$. Condition (i) is Lemma 3.6.6, Condition (ii) is Lemma 3.5.21, and Condition (iii) is Lemma 3.4.2. This completes the proof of Theorem 3.1.3.

3.7 Refinements of Binomial CSP's

A key step in the proof of Lemma 3.6.6 was Theorem 3.2.3 due to Reiner-Stanton-White, which says that the triples

$$\left(\binom{[0,n-1]}{k}, C_n, \binom{n}{k}_q\right) \qquad \text{and} \qquad \left(\left(\!\!\binom{[0,n-1]}{k}\!\!\right), C_n, \left(\!\!\binom{n}{k}\!\!\right)_q\right)$$

exhibit the CSP. Indeed, [73] contains two proofs, one via representation theory [73, §3] and another by direct calculation [73, §4]. In this section, we give two refinements of related CSP's involving an action of C_d on sets of subsets (Theorem 3.7.11) and multisubsets (Theorem 3.7.4) for all $d \mid n$, using shifted sum statistics. Our proof of the subset refinement, Theorem 3.7.11, does not use Theorem 3.2.3, so it can be used as an alternative proof of the subset case of Theorem 3.2.3. Our method is inspired by the rotation of subintervals used by Wagon and Wilf in [100, §3].

3.7.1 Cyclic Actions and Notation

We define two different cyclic actions of the cyclic group of order d on $[0, n-1]$ and induce these actions to $\binom{[0,n-1]}{k}$ and $\left(\!\binom{[0,n-1]}{k}\!\right)$. We also fix notation for the rest of the section.

Notation 3.7.1. Fix $n \in \mathbb{Z}_{\geq 1}, k \in \mathbb{Z}_{\geq 0}$, and $d \mid n$. Let

$$S = \binom{[0,n-1]}{k}, \qquad M = \left(\!\!\binom{[0,n-1]}{k}\!\!\right)$$

For all $j \in [1, \frac{n}{d}]$, let

$$I_d^j := [(j-1)d, jd-1],$$

which we call a *d-interval*. For any composition $\alpha = (\alpha_1, \ldots, \alpha_{n/d}) \vDash k$ with n/d parts, let

$$S_\alpha := \{A \in S : \#(A \cap I_d^j) = \alpha_j \text{ for all } j\}, \tag{3.34}$$

$$M_\alpha := \{A \in M : \#(A \cap I_d^j) = \alpha_j \text{ for all } j\}, \tag{3.35}$$

where the intersection in (3.35) preserves the multiplicity of A. We also fix cyclic groups C_d, C'_d of order d whose actions are described below.

Let C_d act on $[0, n-1]$ by simultaneous rotation of d-intervals, which is generated by the permutation

$$\sigma_d := (0\ 1\ \ldots\ (d-1)) \ldots ((n-d)\ (n-d+1)\ \ldots\ (n-1)) \tag{3.36}$$

in cycle notation. On the other hand, C_n has a unique subgroup C'_d of order d which also acts on $[0, n-1]$ and is generated by the permutation

$$\sigma_n^{n/d} = \left(0\ \left(\frac{n}{d}\right)\ \ldots \left(n - \frac{n}{d}\right)\right) \ldots \left(\left(\frac{n}{d} - 1\right)\ \left(\frac{2n}{d} - 1\right)\ \ldots (n-1)\right). \tag{3.37}$$

Induce these actions of C_d and C'_d up to S and M by

$$g \cdot \{a_1, \ldots, a_k\} := \{g \cdot a_1, \ldots, g \cdot a_k\}.$$

Notice that the action of C_d restricts to S_α and M_α for any $\alpha = (\alpha_1, \ldots, \alpha_{n/d}) \vDash k$.

Let (G, X) be a pair where G is a group acting on a set X. A morphism of group actions $(G, X) \to (G', X')$ is a pair (ϕ, ψ) where $\phi\colon G \to G'$ is a group homorphism and $\psi\colon X \to X'$

is a map of sets which satisfy

$$\psi(g \cdot x) = \phi(g) \cdot \psi(x) \text{ for all } g \in G, x \in X.$$

Remark 3.7.2. The actions of C_d and C'_d on $[0, n-1]$ are isomorphic since σ_d and $\sigma_n^{n/d}$ have the same cycle type. This isomorphism explicitly arises from $\phi\colon \sigma_d \mapsto \sigma_n^{n/d}$ with $\psi\colon 0 \mapsto 0$, $1 \mapsto \frac{n}{d}$, etc. Thus the actions of C_d and C'_d on S and M are isomorphic as well.

Recall the sum statistic sums the elements of a set of multiset. We also use the following shifted sum statistic. For $A \in S$, let

$$\operatorname{sum}'(A) := \sum_{a \in A} a - \sum_{i=0}^{k-1} i = \operatorname{sum}(A) - \binom{k}{2}. \tag{3.38}$$

Recall from (3.3) and (3.4) that

$$S^{\mathrm{sum}'}(q) = \binom{n}{k}_q, \qquad M^{\mathrm{sum}}(q) = \left(\!\!\binom{n}{k}\!\!\right)_q. \tag{3.39}$$

Using (3.39), we may restate Theorem 3.2.3 as saying that both the triples $(S, C_n, S^{\mathrm{sum}'}(q))$ and $(M, C_n, M^{\mathrm{sum}}(q))$ exhibit the CSP. Moreover, under the restricted action of $C'_d \subset C_n$ on M and S, $(S, C'_d, S^{\mathrm{sum}'}(q))$ and $(M, C'_d, M^{\mathrm{sum}}(q))$ exhibit the CSP by Remark 3.2.2. By Remark 3.7.2,

$$(S, C_d, S^{\mathrm{sum}'}(q)) \qquad \text{and} \qquad (M, C_d, M^{\mathrm{sum}}(q)) \tag{3.40}$$

also exhibit the CSP.

Example 3.7.3. Let $n = 8$, $k = 4$, and $d = 4$. Abbreviating $\{0, 4, 5, 6\}$ as 0456, etc., gives

$$\begin{aligned} S_{(1,3)} = \{&0456, 0457, 0467, 0567, 1456, 1457, 1467, 1567,\\ &2456, 2457, 2467, 2567, 3456, 3457, 3467, 3567\}. \end{aligned}$$

Here, C_4 acts on $[0, 8-1]$ by the permutation $(0123)(4567)$, and C'_4 acts by $(0246)(1357)$.

$M_{(1,3)}$ contains $S_{(1,3)}$ in addition to, for instance, 0444.

3.7.2 A Multisubset Refinement

We next prove a refinement of the CSP triple $(M, C_d, M^{\text{sum}}(q))$ in (3.40) by fixing sizes of intersections with the d-intervals.

Theorem 3.7.4. *Recall Notation 3.7.1, and fix a composition $\alpha = (\alpha_1, \ldots, \alpha_{n/d}) \vDash k$. Then, $(M_\alpha, C_d, M_\alpha^{\text{sum}}(q))$ refines the CSP triple $(M, C_d, M^{\text{sum}}(q))$.*

Proof. Separating the d-intervals into different multisubsets gives

$$M_\alpha \cong \left(\!\!\binom{[0, d-1]}{\alpha_1}\!\!\right) \times \cdots \times \left(\!\!\binom{[0, d-1]}{\alpha_{n/d}}\!\!\right), \tag{3.41}$$

which preserves the natural C_d-action and sum statistic modulo d. Since

$$\left(\left(\!\!\binom{[0, d-1]}{\alpha_j}\!\!\right), C_d, \left(\!\!\binom{[0, d-1]}{\alpha_j}\!\!\right)^{\text{sum}}(q) \right)$$

exhibits the CSP for all j, the result follows from Remark 3.2.2. $\square$

The following analogous result holds for subsets.

Proposition 3.7.5. *Recall Notation 3.7.1, and fix a composition $\alpha = (\alpha_1, \ldots, \alpha_{n/d}) \vDash k$. Then $(S_\alpha, C_d, S_\alpha^{\text{sum}^*}(q))$ exhibits the CSP, where*

$$\text{sum}^*(A) := \text{sum}(A) - \sum_{j=1}^{n/d} \binom{\alpha_j}{2}. \tag{3.42}$$

Proof. Separating the d-intervals into different subsets gives

$$S_\alpha \cong \binom{[0, d-1]}{\alpha_1} \times \cdots \times \binom{[0, d-1]}{\alpha_{n/d}}, \tag{3.43}$$

which preserves the C_d-action and sum statistic modulo d. Since

$$\left(\binom{[0,d-1]}{\alpha_j}, C_d, \binom{[0,d-1]}{\alpha_j}^{\mathrm{sum}-\binom{\alpha_j}{2}}(q)\right)$$

exhibits the CSP for all j, $(S_\alpha, C_d, S_\alpha^{\mathrm{sum}^*}(q))$ exhibits the CSP by Remark 3.2.2.

□

Remark 3.7.6. Since we must shift the sum statistic by different amounts depending on α, Proposition 3.7.5 is not a CSP refinement, in contrast to Theorem 3.7.4.

3.7.3 A Subset Refinement

We next prove an honest refinement of the CSP triple $(S, C_d, S^{\mathrm{sum}'}(q))$ in (3.40). To do so, we restrict to certain subsets of S for each divisibility chain ending in n. Our proof again inductively extends CSP's up from cyclic subgroups of C_d using Lemma 3.3.3. In this subsection we first define our restricted subsets and give some examples. We then present a series of lemmata verifying the conditions of Lemma 3.3.3 before proving our refinement, Theorem 3.7.11.

Definition 3.7.7. Suppose $e \mid d \mid n$. Let

$$G_{d,e} := \{A \in S : \gcd(d, \#(A \cap I_d^1), \#(A \cap I_d^2), \ldots, \#(A \cap I_d^{n/d})) = e\}. \tag{3.44}$$

We have $G_{n,\gcd(n,k)} = S$ and $G_{n,e} = \varnothing$ for all other e.

Remark 3.7.8. Note that $G_{d,e}$ decomposes as the disjoint union

$$G_{d,e} = \coprod S_\alpha, \tag{3.45}$$

ranging over all $\alpha = (\alpha_1, \ldots, \alpha_{n/d}) \vDash k$ satisfying

$$\gcd(d, \alpha_1, \ldots, \alpha_{n/d}) = e.$$

Example 3.7.9. If $n = 4, k = 2$, then abbreviating $\{0, 2\}$ as 02, etc., gives

$$G_{1,1} = \{01, 02, 03, 12, 13, 23\} = S,$$
$$G_{2,1} = \{02, 03, 12, 13\}, \qquad G_{2,2} = \{01, 23\},$$
$$G_{4,1} = \varnothing, \qquad G_{4,2} = \{01, 02, 03, 12, 13, 23\} = S, \qquad G_{4,4} = \varnothing.$$

Consequently, $G_{4,2} \cap G_{2,1} = \{02, 03, 12, 13\}$ and $G_{4,2} \cap G_{2,2} = \{01, 23\}$.

Definition 3.7.10. Suppose D is a totally ordered chain in the divisibility lattice ending with $\gcd(n, k) \mid n$, i.e. $D = d_p \mid d_{p-1} \mid \cdots \mid d_0 \mid n$ where $d_0 := \gcd(n, k)$. Write

$$G_D := G_{n,d_0} \cap G_{d_0,d_1} \cap \cdots \cap G_{d_{p-1},d_p} \subset S.$$

We may now state our subset refinement. The proof is postponed to the end of this subsection.

Theorem 3.7.11. *Using Notation 3.7.1, let D be a totally ordered chain in the divisibility lattice ending with $\gcd(n, k) \mid n$ and starting with $e \mid d$. Then, $(G_D, C_d, G_D^{\mathrm{sum}'}(q))$ refines the CSP triple $(S, C_d, S^{\mathrm{sum}'}(q))$.*

Example 3.7.12. If $n = 4$, $k = 2$, and $D = 1 \mid 2 \mid 4$, then $G = G_{4,2} \cap G_{2,1}$ has C_2 orbits $\{02, 13\}$ and $\{03, 12\}$. Moreover,

$$G^{\mathrm{sum}'}(q) = q^1 + 2q^2 + q^3 \equiv 2(q^0 + q^1) \pmod{q^2 - 1},$$

so $(G, C_2, G^{\mathrm{sum}'}(q))$ exhibits the CSP by (3.2).

In fact, the subset case of Theorem 3.2.3 is the special case $D = \gcd(n, k) \mid n$ of Theorem 3.7.11, so the proof below of Theorem 3.7.11 yields an alternative proof of the subset case of Theorem 3.2.3.

Corollary 3.7.13. $(S, C_n, S^{\mathrm{sum}'}(q))$ *exhibits the CSP.*

Lemma 3.7.14. *Let D be a totally ordered chain in the divisibility lattice ending with* $\gcd(n,k) \mid n$ *and beginning with $e \mid d$. Suppose C'_e is the unique subgroup of C_d of order e.*

(i) $G_D = \coprod S_\alpha$*, where the disjoint union is over some set of α satisfying* $\alpha = (\alpha_1, \ldots, \alpha_{n/d}) \vDash k$ *and* $\gcd(d, \alpha_1, \ldots, \alpha_{n/d}) = e$.

(ii) G_D *is closed under the C_d and C_e-actions on S.*

(iii) *The C'_e and C_e-actions on G_D are isomorphic.*

(iv) *For any C_d-orbit $\mathcal{O}$ of G_D, we have* $\frac{d}{|\mathcal{O}|} \mid e$.

(v) *The* sum' *statistic has period e modulo d on G_D.*

Proof. For (i), by (3.45) we have $G_{d,e} = \coprod S_\alpha$ where α ranges over all compositions satisfying the constraints in (i). Similarly, if $c \mid b$, then $G_{b,c} = \coprod S_\beta$ where in particular $\beta = (\beta_1, \ldots, \beta_{n/b}) \vDash k$. Now if $d \mid c$, we may break up each b-interval into b/d d-intervals. It is then easy to see that

$$S_\alpha \cap S_\beta = \varnothing \text{ or } S_\alpha. \tag{3.46}$$

Now (i) follows inductively.

For (ii), by (i) it suffices to show that each S_α is closed under the C_d and C_e-actions. Since σ_d rotates d-intervals, it preserves the size of each d-interval, so σ_d indeed maps S_α to itself. The same argument applies with σ_e in place of σ_d.

For (iii), by (i), it suffices to show the C_e and C'_e-actions on S_α are isomorphic. Recalling (3.43), we have

$$S_\alpha \cong \binom{[0, d-1]}{\alpha_1} \times \cdots \times \binom{[0, d-1]}{\alpha_{n/d}}.$$

By Remark 3.7.2, the actions of C_e and C'_e on $\binom{[0,d-1]}{\alpha_j}$ are isomorphic for each j, so their actions on S_α are isomorphic as well.

For (iv), pick $A \in \mathcal{O}$ with $A \in S_\alpha$ for α as in (i). Let $A_j := A \cap I_d^j$, which has α_j elements. Viewing A_j as a multiplicity word w_j as in Definition 3.6.2, we see that A_j has $d - \alpha_j$ zeros and α_j ones. For all j, w_j is some word repeated $\frac{d}{|\mathcal{O}|}$ times. Using the two-letter case of Lemma 3.4.2, we have $\frac{d}{|\mathcal{O}|} \mid \mathrm{freq}(w_j) \mid \alpha_j$. Thus $\frac{d}{|\mathcal{O}|} \mid \gcd(d, \alpha_1, \ldots, \alpha_{n/d}) = e$.

For (v), it suffices to show that sum' has period e modulo d on S_α for α as in (i). By the gcd condition, there exist $c_1, \ldots, c_{n/d} \in \mathbb{Z}$ such that

$$c_1\alpha_1 + \cdots + c_{n/d}\alpha_{n/d} \equiv e \pmod{d}.$$

For some particular $A \in S_\alpha$, consider cyclically rotating the elements of $A \cap I_d^j$ forward by c_j in I_d^j for all j. The result is a bijection $\phi\colon S_\alpha \to S_\alpha$ such that for all $A \in S_\alpha$,

$$\mathrm{sum}'(\phi(A)) \equiv \mathrm{sum}'(A) + e \pmod{d}.$$

Hence sum' indeed has period e modulo d on S_α. $\square$

Example 3.7.15. Let $n = 12, k = 8$, and $D = 1 \mid 2 \mid 4 \mid 12$. Then

$$G_D = G_{12,4} \cap G_{4,2} \cap G_{2,1} = G_{4,2} \cap G_{2,1}.$$

We have $G_{2,1} = \coprod S_\alpha$ where $\alpha = (\alpha_1, \ldots, \alpha_6) \vDash 8$ and $\gcd(2, \alpha_1, \ldots, \alpha_6) = 1$. Similarly $G_{4,2} = \coprod S_\beta$ where $\beta = (\beta_1, \beta_2, \beta_3) \vDash 8$ and $\gcd(4, \beta_1, \beta_2, \beta_3) = 2$. In fact,

$$\varnothing \subsetneq G_D \subsetneq G_{2,1}$$

since, for instance, $S_\alpha \subset G_D$ when $\alpha = (4, 0, 1, 1, 1, 1)$ while $S_\alpha \subset G_{2,1} - G_D$ when $\alpha = (2, 0, 2, 1, 2, 1)$.

Lemma 3.7.16. *Let $d \mid n$. The C_d action on $G_{d,d}$ is trivial and $(G_{d,d}, C_d, G_{d,d}^{\mathrm{sum}'}(q))$ exhibits the CSP.*

Proof. All subsets in $G_{d,d}$ have each d-interval either full or empty, so C_d fixes every $A \in G_{d,d}$. By (3.2), $(G_{d,d}, C_d, G_{d,d}^{\mathrm{sum}'}(q))$ thus exhibits the CSP if and only if $G_{d,d}^{\mathrm{sum}'}(q) \equiv \#G_{d,d} \bmod (q^d - 1)$. If $G_{d,d} = \varnothing$ the result is trivial, so take $G_{d,d} \neq \varnothing$. For any $A \in G_{d,d}$, since each d-interval is full or empty, we have $d \mid k$ and

$$\mathrm{sum}'(A) \equiv \frac{k}{d}\binom{d}{2} - \binom{k}{2} \equiv \frac{k(d-k)}{2} \equiv 0 \pmod{d}. \tag{3.47}$$

□

We may now prove Theorem 3.7.11.

Proof of Theorem 3.7.11. We induct on d. If $d = 1$, then $(G_D, C_1, G_D^{\mathrm{sum}'}(q))$ exhibits the CSP trivially. For the induction step, we first claim that $(G_D, C_e, G_D^{\mathrm{sum}'}(q))$ exhibits the CSP. If $e = d$, then $G_D \subset G_{d,d}$, so by Lemma 3.7.16 the C_e action is trivial. It is easy to see that CSP's with trivial actions refine to arbitrary subsets, so $(G_D, C_e, G_D^{\mathrm{sum}'}(q))$ exhibits the CSP in this case. If $e < d$, by conditioning on the sizes of the intersections of the e-intervals, we can write

$$G_D = \coprod_{f \mid e} G_{f|D} \tag{3.48}$$

where $f \mid D$ denotes the chain with f prepended to D. Hence $(G_{f|D}, C_e, G_{f|D}^{\mathrm{sum}'}(q))$ exhibits the CSP by induction for each $f \mid e$, since $f \mid D$ begins with $f \mid e$. Thus $(G_D, C_e, G_D^{\mathrm{sum}'}(q))$ exhibits the CSP by (3.48), proving the claim.

In order to realize the $(G_D, C_d, G_D^{\mathrm{sum}'}(q))$ CSP triple from the $(G_D, C_e, G_D^{\mathrm{sum}'}(q))$ CSP triple, we verify the conditions of Lemma 3.3.3. From Lemma 3.7.14(ii), the restriction of the C_d-action on G_D to the subgroup $C'_e \subset C_d$ of size e is isomorphic to the C_e-action on G_D, giving Condition (i). Condition (ii) is Lemma 3.7.14(v), and Condition (iii) is Lemma 3.7.14(iv). Thus $(G_D, C_d, G_D^{\mathrm{sum}'}(q))$ exhibits the CSP by Lemma 3.3.3. □

3.8 The Flex Statistic

We conclude by formalizing the notion of *universal* sieving statistics and giving an example, *flex*, in the context of words. We end with an open problem.

Definition 3.8.1. Given a set W with a C_n-action, we say $\mathrm{stat}\colon W \to \mathbb{Z}_{\geq 0}$ is a *universal CSP statistic* for (W, C_n) if $(\mathcal{O}, C_n, \mathcal{O}^{\mathrm{stat}}(q))$ exhibits the CSP for all C_n-orbits $\mathcal{O}$ of W.

Definition 3.8.2. Let $\mathrm{lex}(w)$ denote the index at which w appears when lexicographically ordering the necklace $[w]$, starting from 0. Let *flex* be the product

$$\mathrm{flex}(w) := \mathrm{freq}(w)\,\mathrm{lex}(w).$$

For example, the necklace in Example 2.1.1 has lex statistics $0, 3, 2, 1$, respectively, so that we have $\mathrm{lex}(55315531) = 3$ and $\mathrm{flex}(55315531) = 2 \cdot 3 = 6$.

Lemma 3.8.3. *The function* flex *is a universal CSP statistic for* (W_n, C_n).

Proof. Let N be any necklace of length n words. Since $\mathrm{freq}(N) = \frac{n}{|N|}$, and $\mathrm{lex}(N) = \{0, 1, \ldots, |N| - 1\}$, we have

$$N^{\mathrm{flex}}(q) = \sum_{j=0}^{|N|-1} q^{j \cdot \frac{n}{|N|}} = \frac{q^n - 1}{q^{n/|N|} - 1}, \tag{3.49}$$

so $(N, C_n, N^{\mathrm{flex}}(q))$ exhibits the CSP by (3.2). $\square$

Given a universal sieving statistic stat on some set W, stat takes on each of the values $\{0, n/d, \ldots, n - n/d\}$ modulo n on any orbit of size d. The converse holds as well. In this sense, up to shifting values by n, universal sieving statistics are equivalent to total orderings on each orbit $\mathcal{O}$ of W.

Standing in contrast to Lemma 3.8.3, $(N, C_n, N^{\mathrm{maj}}(q))$ does not exhibit the CSP when $N = [123123]$, so maj is not a universal CSP statistic on (W_n, C_n). However, maj trivially refines to the orbit $N = \{1^n\}$ for any n. Since refinement is not generally closed under intersections, it is not clear if there is any useful sense in which maj on words can be "maximally refined."

It follows from Lemma 3.8.3 and (3.2) that Theorem 3.1.3 is equivalent to the following.

Theorem 3.8.4. *The statistics* flex *and* maj *are equidistributed modulo* n *on* $\mathrm{W}_{\alpha,\delta}$.

Corollary 3.8.5. *For all* $\alpha \vDash n$,

$$\mathrm{W}_{\alpha}^{\mathrm{maj}_n}(q) = \mathrm{W}_{\alpha}^{\mathrm{flex}}(q)$$

where maj_n *is the major index modulo* n *taking values in* $[0, n-1]$.

Indeed, we were originally led to Theorem 3.1.3 through an exploration of the irreducible multiplicities of the so-called higher Lie modules, which are described in detail in Chapter 4, which uncovered the fact that flex and maj are equidistributed modulo n on W_{α}. Data exploration led us originally to conjecture this equidistribution refined to fixed cyclic descent type as in Theorem 3.8.4. These observations naturally suggest the problem of finding explicit bijections proving Theorem 3.8.4, which we leave as an open problem.

Open Problem 3.8.6. *For* $\alpha \vDash n$ *and* δ *any weak composition, find a bijection* $\varphi\colon \mathrm{W}_{\alpha,\delta} \to \mathrm{W}_{\alpha,\delta}$ *satisfying*

$$\mathrm{maj}(\varphi(w)) \equiv \mathrm{flex}(w) \pmod{n}. \tag{3.50}$$

Chapter 4

CYCLIC SIEVING, NECKLACES, AND BRANCHING RULES RELATED TO THRALL'S PROBLEM

This chapter is based on joint work with Connor Ahlbach. A version of it will be submitted for publication shortly [4].

4.1 *Main Results*

Thrall [99] famously introduced a certain $\mathrm{GL}(V)$-module decomposition $\oplus_\lambda \mathcal{L}_\lambda$ of the tensor algebra of a vector space V, where the sum is over all partitions of all $n \in \mathbb{Z}_{\geq 0}$. The decomposition arises from the Poincaré–Birkhoff–Witt theorem applied to the free Lie algebra generated by V. The $\mathcal{L}_\lambda$ are called the *Lie modules* and the determination of the multiplicity of the irreducible V^μ in $\mathcal{L}_\lambda$ is called *Thrall's problem*. See Section 4.2 for further background.

Kraśkiewicz–Weyman [54] solved Thrall's problem in terms of major indices of standard Young tableaux in the important special case $\lambda = (n)$; see Theorem 1.3.1. Klyachko [50] had earlier and independently computed the Schur–Weyl dual of $\mathcal{L}_{(n)}$ as an induced representation $\chi^1\uparrow_{C_n}^{S_n}$. His argument identified the corresponding character as a generating function on primitive necklaces; see Proposition 4.2.5 below. More recently, Schocker [81] gave a general formula for the multiplicity of V^μ in $\mathcal{L}_\lambda$ using different techniques which reduces to Kraśkiewicz–Weyman's result when $\lambda = (n)$, though it involves many subtractions and divisions in general. Stembridge [94] separately, using yet another set of techniques, proved a conjecture of Stanley describing the irreducible multiplicities of general $\chi^r\uparrow_{\langle\sigma\rangle}^{S_n}$.

In this chapter, we present a unified approach to these results and certain generalizations using the cyclic sieving phenomenon, Definition 1.2.5. The CSP $(\mathrm{W}_\alpha, C_n, \mathrm{W}_\alpha^{\mathrm{maj}}(q))$ from Theorem 1.2.6 above is fundamental to our arguments in this chapter. In Section 4.3, we interpret Theorem 1.2.6 as a "universality" result for sieving on words under a symmetric group action using maj modulo n. We contrast this with the universality result for flex from Section 3.8. We show that Kraśkiewicz–Weyman's result can be interpreted as essentially equivalent to these two universality statements. The resulting proof of Kraśkiewicz–Weyman's result is quite conceptual and "nearly bijective." For instance, a bijective proof of the maj_n, flex equidistribution result, Corollary 3.8.5, together with our other arguments would provide a bijective proof of the well-known symmetry result, Corollary 4.2.8, following from Kraśkiewicz–Weyman's theorem. Such a proof would be quite interesting and is not currently known. Our argument also provides a thus far relatively rare example of an instance of cyclic sieving being used to prove other results rather than vice versa. See Section 4.3 for more details.

In Section 4.4, we apply Theorem 1.2.6 and direct manipulations with necklace generating functions to provide a new proof of Stembridge's generalization of Kraśkiewicz–Weyman's result to all branching rules for inclusions $\langle\sigma\rangle \hookrightarrow S_n$. The corresponding generalized major index statistics arise very naturally from the combinatorics of orbits and cyclic sieving. We also have analogous arguments proving Schocker's formula [81, Theorem 3.1] for the so-called higher Lie multiplicities and generalize the result to all branching rules for all one-dimensional representations for the natural inclusion $C_a \wr S_b \hookrightarrow S_{ab}$. This latter argument will appear in a subsequent publication.

The basic outline of each argument is the same: we obtain an orbit generating function from an explicit basis of a $\mathrm{GL}(V)$-module, we construct an appropriate necklace generating function, we use cyclic sieving to rewrite this generating function using words and descent statistics like the major index, and we finally apply RSK to get a Schur expansion. In Section 4.5, we discuss applying aspects of this approach to Thrall's problem in general. It suggests attacking Thrall's problem by considering all branching rules $C_a \wr S_b \hookrightarrow S_{ab}$ rather than considering only one such rule. We identify the analogue of the flex statistic in this context and isolate key properties of an analogue of maj. Assuming that such an analogue of maj can be found, we give a complete solution to the higher Lie multiplicities problem and more generally determine all branching rules for the inclusion $C_a \wr S_b \hookrightarrow S_{ab}$. While finding this generalized maj statistic has proven elusive outside of certain special cases, we hope that our approach will inspire further progress on this difficult problem.

The rest of this chapter is organized as follows. In Section 4.2, we give background. In Section 4.3, we present our proof of Kraśkiewicz–Weyman's result using cyclic sieving. In Section 4.4, we give an analogous proof of Stembridge's result. In Section 4.5 we give a monomial expansion of a certain graded Frobenius series, Theorem 4.5.4, and we discuss how the approach might be used to find the branching rules for the inclusion $C_a \wr S_b \hookrightarrow S_{ab}$.

4.2 Necklaces, Thrall's Problem, and Wreath Products

Here we provide background on necklaces, RSK, Schur–Weyl duality, Thrall's problem, and certain wreath products for use in later sections. See also the background material from Chapter 2. All representations will be over $\mathbb{C}$.

4.2.1 Necklaces

Definition 4.2.1. In analogy with the sets W_n and W_α for the set of words with letters from $\mathbb{Z}_{\geq 1}$ either of length n or with content α, we use the following notions on necklaces. For $n \geq 1$ and $\alpha \vDash n$, we write

$$\begin{aligned} \mathrm{N}_n &= \{\text{necklaces of length } n \text{ words}\}, \\ \mathrm{N}_\alpha &= \{\text{necklaces with content } \alpha\}, \\ \mathrm{PN}_n &= \{\text{primitive necklaces of length } n \text{ words}\}, \\ \mathrm{NFD}_{n,r} &= \{\text{necklaces of length } n \text{ with frequency dividing } r\}. \end{aligned}$$

In particular, $\mathrm{NFD}_{n,n} = \mathrm{N}_n$, and $\mathrm{NFD}_{n,1} = \mathrm{PN}_n$.

Example 4.2.2. Consider $w = 15531553 \in \mathrm{W}_8$. Then, $|w| = 8$, $\mathrm{Des}(w) = \{3, 4, 7\}$, $\mathrm{maj}(w) = 14$, and $\mathrm{cont}(w) = (2, 0, 2, 0, 4)$, so $w \in W_{(2,0,2,0,4)}$. Since $w = 15531553 = (1553)^2$ and 1553 is primitive, w is not primitive, $\mathrm{period}(w) = 4$, and $\mathrm{freq}(w) = 2$. The necklace of w

is

$$[w] := \{15531553, 55315531, 53155315, 31553155\} \in \mathrm{N}_{(2,0,2,0,4)} \subset \mathrm{N}_8\,.$$

4.2.2 RSK

We write Par for the set of all partitions over all n. The Robinson–Shensted–Knuth (RSK) correspondence is a bijection

$$\begin{aligned}\mathrm{RSK}\colon\ \mathrm{W}_n &\to \bigsqcup_{\lambda\vdash n} \mathrm{SSYT}(\lambda)\times\mathrm{SYT}(\lambda),\\ w &\mapsto (P(w), Q(w)).\end{aligned}$$

The *shape* of w under RSK, denoted $\mathrm{sh}(w)$, is the common shape of $P(w)$ and $Q(w)$. Two well-known properties of the RSK correspondence are

$$\mathrm{cont}(w) = \mathrm{cont}(P(w)), \qquad \mathrm{Des}(w) = \mathrm{Des}(Q(w)). \tag{4.1}$$

See [78, Chapter 3] for more details on RSK.

4.2.3 Schur–Weyl Duality

We next summarize a few key points from the representation theory of S_n and $\mathrm{GL}(\mathbb{C}^m)$. See [31] for more. Recall from Section 1.1 that the complex irreducible inequivalent representations S^λ of S_n are canonically indexed by partitions $\lambda \vdash n$. The Frobenius characteristic map ch is defined by $\mathrm{ch}\, S^\lambda := s_\lambda$ and is extended additively to all S_n-representations. Since Schur functions are $\mathbb{Z}$-linearly independent, computing the irreducible decomposition of an S_n-module M corresponds to computing the Schur expansion of $\mathrm{ch}\, M$.

Let V be a complex vector space of dimension m. Endow $V^{\otimes n}$ with the natural left $\mathrm{GL}(V)$-action by linear substitutions and the natural right S_n-action by permutation of indexes. Given an S_n-module M, define a corresponding $\mathrm{GL}(V)$-module

$$E(M) \coloneqq V^{\otimes n} \otimes_{\mathbb{C}S_n} M,$$

which we call the *Schur module* of M. The irreducible inequivalent polynomial representations of $\mathrm{GL}(V)$ are precisely the Schur modules $V^\lambda \coloneqq E(S^\lambda)$ where λ has at most $\dim(V)$ non-zero parts [31].

Let E be a finite-dimensional, polynomial representation of $\mathrm{GL}(V)$ and pick a basis $\{v_1, \ldots, v_m\}$ for V. The *Schur character* of E, denoted $\operatorname{ch} E$, is the function which sends $(x_1, \ldots, x_m)$ to the character of E evaluated at $\operatorname{diag}(x_1, \ldots, x_m) \in \mathrm{GL}(V)$, where the diagonal matrix is with respect to the basis $v_1, \ldots, v_m$. Polynomiality of E implies $\operatorname{ch} E \in \mathbb{C}[x_1, \ldots, x_m]$. Moreover,

$$\operatorname{ch} V^\lambda = \operatorname{ch} E(S^\lambda) = s_\lambda(x_1, \ldots, x_m, 0, 0, \ldots).$$

Thus, for any S_n-module M, we have "in the $m \to \infty$ limit" equality of Schur and Frobenius characters:

$$\operatorname{ch} E(M) = \operatorname{ch} M.$$

In light of this, we often leave dependence on m or V implicit.

4.2.4 Thrall's Problem

Next we recall the Lie modules $\mathcal{L}_\lambda$ arising from the study of free Lie algebras. We then summarize certain aspects of the determination of the multiplicity of V^μ in $\mathcal{L}_\lambda$. See [74] for more details.

The *tensor algebra* of V is $T(V) \coloneqq \oplus_{n=0}^{\infty} V^{\otimes n}$, which is naturally a graded $\mathrm{GL}(V)$-representation. Let $\mathcal{L}(V)$ be the Lie subalgebra of $T(V)$ generated by V, called the *free Lie algebra* on V. Now $\mathcal{L}(V)$ is also a graded $\mathrm{GL}(V)$-representation with graded components $\mathcal{L}_n(V) = V^{\otimes n} \cap \mathcal{L}(V)$. The *universal enveloping algebra* $\mathfrak{U}(\mathcal{L}(V))$ is isomorphic to $T(V)$ itself.

By the Poincaré–Birkhoff–Witt Theorem,

$$\mathfrak{U}(\mathcal{L}(V)) \cong \bigoplus_{\lambda=1^{m_1}2^{m_2}\cdots} \operatorname{Sym}^{m_1}(\mathcal{L}_1(V)) \otimes \operatorname{Sym}^{m_2}(\mathcal{L}_2(V)) \otimes \cdots$$

as graded $\mathrm{GL}(V)$-representations, where the sum is over all partitions of all non-negative integers and $\operatorname{Sym}^m(E)$ is the mth symmetric power of E [74, Lemma 8.22]. The *Lie module* associated to $\lambda = 1^{m_1}2^{m_2}\cdots$ is defined to be

$$\mathcal{L}_\lambda(V) \coloneqq \operatorname{Sym}^{m_1}(\mathcal{L}_1(V)) \otimes \operatorname{Sym}^{m_2}(\mathcal{L}_2(V)) \otimes \cdots . \tag{4.2}$$

The Lie modules hence yield a $\mathrm{GL}(V)$-module decomposition $T(V) \cong \oplus_\lambda \mathcal{L}_\lambda(V)$.

Thrall's problem is the determination of the multiplicity of V^μ in $\mathcal{L}_\lambda$, for instance by counting explicit combinatorial objects. The well-known Littlewood–Richardson rule solves the analogous problem for $V^\mu \otimes V^\nu$. It follows from (4.2) and the Littlewood–Richardson rule that, for the purposes of Thrall's problem, we may restrict our attention to the case when $\lambda = (a^b)$ is a rectangle. Since $\mathcal{L}_{(a^b)}(V) = \operatorname{Sym}^b \mathcal{L}_{(a)}(V)$, the single-row case is particularly fundamental.

At present, Thrall's problem has only been solved in the following cases:

- when $\lambda = (n)$ has a single part (described below in Corollary 4.2.7);
- when $\lambda = (1^n)$, $\mathcal{L}_{(1^n)}$ is the trivial representation;
- when $\lambda = (2^m)$, $\operatorname{ch} \mathcal{L}_{(2^m)} = \sum s_\mu$ where the sum is over $\mu \vdash 2m$ with even column sizes [59, Ex. I.8.6(b), p. 138].

Hall [41, Lemma 11.2.1] introduced what is now called the *Hall basis* for $\mathcal{L}_{(n)}$, which, in the $m \to \infty$ limit, is in content-preserving bijection with PN_n. Klyachko consequently observed that the Schur character of $\mathcal{L}_{(n)}$ is the corresponding content generating function. Taking symmetric powers, it follows that $\mathcal{L}_{(a^b)}$ has a basis consisting of multisets of primitive necklaces and the Schur character is the corresponding content generating function. Klyachko was further able to identify the Schur–Weyl dual of $\mathcal{L}_{(n)}$ as a certain induced representation, which we describe after summarizing the above.

Proposition 4.2.3 (See [50, Proposition 1])**.** *We have "in the $m \to \infty$ limit"*

$$\operatorname{ch} \mathcal{L}_{(n)} = \mathrm{PN}_n^{\mathrm{cont}}(\mathbf{x}) \qquad \textit{and} \qquad \operatorname{ch} \mathcal{L}_{(a^b)} = \left(\!\!\binom{\mathrm{PN}_a}{b}\!\!\right)^{\mathrm{cont}}(\mathbf{x}).$$

To solve Thrall's problem, we need to turn the expressions in Proposition 4.2.3 into Schur expansions.

Definition 4.2.4. Let $\sigma_n \coloneqq (1\,2\,\cdots\,n) \in S_n$ and $C_n \coloneqq \langle \sigma_n \rangle \leq S_n$ be the cyclic group of order n it generates. Fixing any primitive nth root of unity ω_n, write the irreducible characters of C_n as $\chi^1, \ldots, \chi^n$ where

$$\chi^r(\sigma_n) \coloneqq \omega_n^r.$$

Klyachko observed that $E(\chi^1{\uparrow}_{C_n}^{S_n})$, like $\mathcal{L}_{(n)}$, also has a basis indexed by primitive necklaces. Klyachko's argument may be easily generalized to $E(\chi^r{\uparrow}_{C_n}^{S_n})$. Since our argument is straightforward and the proof in [50] is somewhat terse, we now state and prove this generalization. Recall from Definition 4.2.1 that $\mathrm{NFD}_{n,r} = \{N \in \mathrm{N}_n : \mathrm{freq}(N) \mid r\}$ is the set of length n necklaces whose frequency divides r, so that $\mathrm{NFD}_{n,1} = \mathrm{PN}_n$.

Proposition 4.2.5. *For all $m \geq 1$, there is a basis for the Schur module $E(\chi^r{\uparrow}_{C_n}^{S_n})$ over $\mathrm{GL}(\mathbb{C}^m)$ indexed by necklaces of length n words with letters from $[m]$ and with frequency dividing r. Moreover,*

$$\operatorname{ch} \chi^r{\uparrow}_{C_n}^{S_n} = \mathrm{NFD}_{n,r}^{\mathrm{cont}}. \tag{4.3}$$

Proof. Suppose the underlying vector space V has basis $\{v_1, \ldots, v_m\}$. By a slight abuse of notation, we may view χ^r as the vector space $\mathbb{C}$ with the left C_n-action $\sigma_n \cdot 1 \coloneqq \omega_n^r$. Since

$\chi^r\uparrow_{C_n}^{S_n} \coloneqq \mathbb{C}S_n \otimes_{\mathbb{C}C_n} \chi^r$, we have

$$E(\chi^r\uparrow_{C_n}^{S_n}) = V^{\otimes n} \otimes_{\mathbb{C}S_n} \mathbb{C}S_n \otimes_{\mathbb{C}C_n} \chi^r \cong V^{\otimes n} \otimes_{\mathbb{C}C_n} \chi^r$$

where C_n acts on $V^{\otimes n}$ on the right by "rotating" the components of simple tensors. A spanning set for $V^{\otimes n} \otimes_{\mathbb{C}C_n} \chi^r$ is given by all $v_{i_1} \otimes \cdots \otimes v_{i_n} \otimes 1$, which we abbreviate as $[i_1 \ \cdots \ i_n]$. Acting by σ_n on χ^r on the left or on $V^{\otimes n}$ on the right gives the relation

$$[i_1 \ \cdots \ i_n] = \omega_n^{-r}[i_2 \ \cdots \ i_n \ i_1].$$

If the word $i_1 \cdots i_n$ has frequency f and period p, we then find

$$\begin{aligned} [i_1 \ \cdots \ i_n] &= \frac{1}{n} \sum_{j=0}^{n-1} \omega_n^{-jr} [i_{j+1} \ \cdots \ i_n \ i_1 \ \cdots \ i_j] \\ &= \frac{1}{n} \sum_{k=0}^{p-1} \left(\sum_{\ell=0}^{f-1} \omega_n^{-(\ell p+k)r} \right) [i_{k+1} \cdots i_n i_1 \cdots i_k] \\ &= \frac{1}{n} \left(\sum_{\ell=0}^{f-1} \omega_n^{-\ell pr} \right) \sum_{k=0}^{p-1} \omega_n^{-kr} [i_{k+1} \cdots i_n i_1 \cdots i_k]. \end{aligned}$$

Since ω_n^p is a primitive $n/p = f$-th root of unity, the factor $\sum_{\ell=0}^{f-1} \omega_n^{-\ell pr}$ is non-zero if and only if $\omega_n^{-pr} = 1$, so if and only if $f \mid r$. Moreover, when $f \mid r$, the above computation shows that $[i_1 \cdots i_n]$ is well-defined up to nonzero scalar multiplication on the level of necklaces, which explains our notation. It is easy to see that the spanning set just constructed is in fact a basis. It is also easy to see that $\mathrm{diag}(x_1, \ldots, x_m)$ itself acts diagonally on this basis. Thus, the resulting Schur character, in the $m \to \infty$ limit, is precisely $\mathrm{NFD}_{n,r}^{\mathrm{cont}}$. $\square$

Kraśkiewicz–Weyman related the graded pieces of the type A coinvariant algebra to the induced representations $\chi^r\uparrow_{C_n}^{S_n}$. Using the so-called Lusztig–Stanley theorem [87, Prop. 4.11], they obtained the irreducible decompositions in Theorem 4.2.6.

Theorem 4.2.6 (Kraśkiewicz–Weyman, [54, Corollary 3])**.** *For all $n \geq 1$ and $r = 1, \ldots, n$,*

$$\operatorname{ch} \chi^r \uparrow_{C_n}^{S_n} = \sum_{\lambda \vdash n} a_{\lambda,r} s_\lambda$$

where

$$a_{\lambda,r} \coloneqq \#\{Q \in \mathrm{SYT}(\lambda) : \operatorname{maj}(Q) \equiv_n r\}.$$

Combining Klyachko's observations, Proposition 4.2.3 and Proposition 4.2.5, with Kraśkiewicz–Weyman's result, Theorem 4.2.6, consequently solves Thrall's problem when $\lambda = (n)$.

Corollary 4.2.7. *For all $\lambda \vdash n \geq 1$, the multiplicity of V^λ in $\mathcal{L}_{(n)}$ is $a_{\lambda,1}$.*

Since $\chi^r \uparrow_{C_n}^{S_n}$ depends up to isomorphism only on n and $\gcd(n, r)$, we also have the following well-known enumerative symmetry.

Corollary 4.2.8. *For all $\lambda \vdash n \geq 1$ and $r \in \mathbb{Z}$, we have $a_{\lambda,r} = a_{\lambda,\gcd(n,r)}$.*

A bijective proof of this symmetry is currently unknown and would be quite interesting. Our argument in Section 4.3 proving Theorem 4.2.6 reduces the problem of finding a bijective proof of Corollary 4.2.8 to the problem of finding a bijective proof of Corollary 4.2.8.

While we will not have need of it, in light of Proposition 4.2.3, we would be remiss if we did not mention Gessel and Reutenauer's important and beautiful expansion of $\operatorname{ch} \mathcal{L}_\lambda$ in terms of Gessel's fundamental quasisymmetric functions [36, equation (2.1)]:

$$\operatorname{ch} \mathcal{L}_\lambda = \sum_{\substack{\sigma \in S_n \\ \sigma \text{ has cycle type } \lambda}} F_{n,\mathrm{Des}(\sigma)}(x), \tag{4.4}$$

where

$$F_{n,D}(x) = \sum_{\substack{i_1 \leq \cdots \leq i_n \\ i_j < i_{j+1} \text{ if } j \in D}} x_{i_1} \ldots x_{i_n}.$$

Gessel and Reutenauer gave an elegant bijective proof of (4.4) in [36].

4.2.5 Wreath Products

The Schur–Weyl duals of the higher Lie modules $\mathcal{L}_{(a^b)}$ have also been identified in terms of induced representations of certain wreath products. Here we summarize this connection as well as some related aspects of the representation theory of wreath products which will be used in Section 4.5. Our presentation largely mirrors [94].

Definition 4.2.9. Given a group G, the *wreath product* of G with S_n, denoted $G \wr S_n$, is the semidirect product explicitly described as follows. $G \wr S_n$ is the set $G^n \times S_n$ with multiplication given by

$$(g_1, \ldots, g_n, \sigma) \cdot (h_1, \ldots, h_n, \tau) \coloneqq (g_1 h_{\sigma^{-1}(1)}, \ldots, g_n h_{\sigma^{-1}(n)}, \sigma\tau)$$

for all $g_1, \ldots g_n, h_1, \ldots, h_n \in G$ and $\sigma, \tau \in S_n$. Furthermore, given $\alpha \vDash n$, set $G \wr \prod_i S_{\alpha_i} \coloneqq \prod_i (G \wr S_{\alpha_i})$, which has a natural inclusion into $G \wr S_n$.

Now suppose U is a G-set and V is an S_n-set. There is a natural notion of $U \wr V$ as a $G \wr S_n$-set. Explicitly, let $U \wr V$ be the set $U^n \times V$ with $G \wr S_n$-action given by

$$(g_1, \ldots, g_n, \sigma) \cdot (u_1, \ldots, u_n, v) \coloneqq (g_1 \cdot u_{\sigma^{-1}(1)}, \ldots, g_n \cdot u_{\sigma^{-1}(n)}, \sigma \cdot v)$$

for all $g_1, \ldots, g_n \in G, \sigma \in S_n, u_1, \ldots, u_n \in U, v \in V$. There is an analogous notion if U is a G-module and V is an S_n-module, namely $U \wr V \coloneqq U^{\otimes n} \otimes V$ with $G \wr S_n$-action

$$(g_1, \ldots, g_b, \sigma) \cdot (u_1 \otimes \cdots \otimes u_b \otimes v) \coloneqq (g_1 \cdot u_{\sigma^{-1}(1)}) \otimes \cdots \otimes (g_b \cdot u_{\sigma^{-1}(b)}) \otimes (\sigma \cdot v)$$

extended $\mathbb{C}$-linearly.

Since S_n acts naturally and faithfully on $[n]$, $[a] \wr 1$ has a natural $S_a \wr S_b$-action, where 1 denotes the trivial S_b-set. Identifying $[a] \wr 1$ with the set $[ab]$ and noting that the action remains faithful gives an inclusion $S_a \wr S_b \hookrightarrow S_{ab}$. Similarly we have an inclusion $C_a \wr S_b \hookrightarrow S_{ab}$.

Remark 4.2.10. The induction product of two symmetric group representations corresponds the product of their Frobenius characteristics, so that if U is an S_a-module and V is an

S_b-module, then [91, Prop. 7.18.2],

$$\operatorname{ch}\left(U \otimes V\uparrow_{S_a \times S_b}^{S_{a+b}}\right) = (\operatorname{ch} U)(\operatorname{ch} V). \tag{4.5}$$

The wreath product of symmetric group representations corresponds to the *plethysm* of their Frobenius characters. Given two symmetric functions f and $g = m_1 + m_2 + \cdots$ where the m_i are all monomials, their plethysm is given by [91, Def. A2.6]

$$f[g] = f(m_1, m_2, \dots), \tag{4.6}$$

which is well-defined since f is symmetric. Then, if U is an S_a-module and V is an S_b-module, we have [91, Thm. A2.8]

$$\operatorname{ch}\left((U \wr V)\uparrow_{S_a \wr S_b}^{S_{ab}}\right) = \operatorname{ch}(V)[\operatorname{ch}(U)]. \tag{4.7}$$

Definition 4.2.11. When G is a finite group, Specht [84] described the complex inequivalent irreducible representations of $G \wr S_n$ in terms of those for G and S_n, the conjugacy classes of G, and wreath products. In the case $C_a \wr S_b$, they are indexed by a-tuples of partitions whose total size is b, or equivalently by functions

$$\begin{gathered} \underline{\lambda}\colon [a] \to \mathrm{Par} \\ \text{where } |\underline{\lambda}| := \sum_{j=1}^{a} |\underline{\lambda}(j)| = b. \end{gathered}$$

The complex inequivalent irreducible representations of $C_a \wr S_b$ are given by

$$S^{\underline{\lambda}} := \left((\chi^1 \wr S^{\underline{\lambda}(1)}) \otimes \cdots \otimes (\chi^a \wr S^{\underline{\lambda}(a)})\right)\uparrow_{C_a \wr S_{\alpha(\underline{\lambda})}}^{C_a \wr S_b},$$

where

$$\alpha(\underline{\lambda}) := (|\underline{\lambda}(1)|, \ldots, |\underline{\lambda}(a)|) \vDash b,$$
$$S_{\alpha(\underline{\lambda})} := S_{|\underline{\lambda}(1)|} \times \cdots \times S_{|\underline{\lambda}(a)|},$$

and $C_a \wr S_{\alpha(\underline{\lambda})}$ is viewed naturally as a subgroup of $C_a \wr S_b$. In particular, the one-dimensional representations of $C_a \wr S_b$ for $b \geq 2$ are

$$\chi^{r,1} := \chi^r \wr 1 \qquad \text{and } \chi^{r,\epsilon} := \chi^r \wr \epsilon$$

where $r = 1, \ldots, a$ and 1 and ϵ are the trivial and sign representations of S_b, respectively. When $b = 1$, $\chi^{r,1}$ is simply χ^r.

Bergeron–Bergeron–Garsia [8] extended Klyachko's observation by showing that the Schur–Weyl dual of $\mathcal{L}_{(a^b)}$ is $\chi^{1,1}\uparrow_{C_a \wr S_b}^{S_{ab}}$. We next give a different argument for this fact which is straightforward given the preceding background and which uses a basic lemma we will require later in Section 4.5.

Theorem 4.2.12 ([8, §4.4]; see also [74, Thm. 8.23])**.** *We have*

$$\operatorname{ch} \chi^{1,1}\uparrow_{C_a \wr S_b}^{S_{ab}} = \left(\!\!\binom{\mathrm{PN}_a}{b}\!\!\right)^{\mathrm{cont}} = \operatorname{ch} \mathcal{L}_{(a^b)}.$$

Proof. The second equality is Proposition 4.2.3. For the first equality, by Lemma 4.2.13 below, we have

$$(\chi^1 \wr 1)\uparrow_{C_a \wr S_b}^{S_{ab}} \cong (\chi^1\uparrow_{C_a}^{S_a} \wr 1)\uparrow_{S_a \wr S_b}^{S_{ab}} .$$

By (4.7) and Proposition 4.2.5,

$$\begin{aligned} \operatorname{ch} \chi^{1,1}\uparrow_{C_a \wr S_b}^{S_{ab}} &= \operatorname{ch}(\chi\uparrow_{C_a}^{S_a} \wr 1)\uparrow_{S_a \wr S_b}^{S_{ab}} \\ &= (\operatorname{ch} 1)[\operatorname{ch} \chi\uparrow_{C_a}^{S_a}] \\ &= h_b[\mathrm{PN}_a^{\mathrm{cont}}] \\ &= \left(\!\!\binom{\mathrm{PN}_a}{b}\!\!\right)^{\mathrm{cont}} \end{aligned}$$

where h_b is the complete homogeneous symmetric function of degree b. The identification with the claimed multiset content generating function follows easily from (4.6) and the definition of h_b. The result will be complete once we prove Lemma 4.2.13. □

Lemma 4.2.13. *Suppose that H is a subgroup of a group G, that U is an H-module, and that V is an S_n-module. Then*

$$(U \wr V)\uparrow_{H \wr S_n}^{G \wr S_n} \cong \left(U\uparrow_H^G\right) \wr V$$

as $G \wr S_n$-modules.

Proof. As sets, we have

$$\begin{aligned} (U \wr V)\uparrow_{H \wr S_n}^{G \wr S_n} &= \mathbb{C}(G \wr S_n) \otimes_{\mathbb{C}(H \wr S_n)} (U^{\otimes n} \otimes V), \\ (U\uparrow_H^G) \wr V &= (\mathbb{C}G \otimes_{\mathbb{C}H} U)^{\otimes n} \otimes V. \end{aligned}$$

Define

$$\begin{aligned} \phi\colon (U \wr V)\uparrow_{H \wr S_n}^{G \wr S_n} &\to (U\uparrow_H^G) \wr V, \\ \psi\colon (U\uparrow_H^G) \wr V &\to (U \wr V)\uparrow_{H \wr S_n}^{G \wr S_n} \end{aligned}$$

by

$$\begin{aligned}
&\phi((h_1, \ldots, h_n, \tau) \otimes (u_1 \otimes \cdots \otimes u_n \otimes v)) \\
&\qquad \coloneqq (h_1 \otimes u_{\tau^{-1}(1)}) \otimes \cdots \otimes (h_n \otimes u_{\tau^{-1}(n)}) \otimes (\tau \cdot v), \\
&\psi((h_1 \otimes x_1) \otimes \cdots \otimes (h_n \otimes x_n) \otimes y) \\
&\qquad \coloneqq (h_1, \ldots, h_n, 1) \otimes (x_1 \otimes \cdots \otimes x_n \otimes y)
\end{aligned}$$

extended $\mathbb{C}$-linearly. It is straightforward to check directly that ϕ and ψ are well-defined, $G \wr S_n$-equivariant, and mutual inverses. $\square$

4.3 Cyclic Sieving and Kraśkiewicz–Weyman's Result

In this section, we present a new proof of Kraśkiewicz–Weyman's result, Theorem 4.2.6, which exposes an intimate connection between that result and cyclic sieving on words. We also discuss the largely bijective nature of the argument and contrast it with existing approaches.

Let $\mathrm{maj}_n \colon \mathrm{W}_n \to [n]$ denote the major index modulo n. This statistic is a universal sieving statistic for words under the natural S_n-action in the sense of the next result, which is equivalent to Reiner–Stanton–White's result, Theorem 1.2.6. A very similar observation appeared in [9, Prop. 3.1] in connection with cyclic sieving on parking functions.

Proposition 4.3.1. *Let $W \subset \mathrm{W}_n$ be a set of words closed under the S_n-action. Then, the triple*

$$(W, C_n, W^{\mathrm{maj}_n}(q))$$

exhibits the CSP, where W has the restricted $C_n = \langle (1\ 2\ \cdots\ n) \rangle$-action.

Proof. It suffices to consider the case when W is a single S_n-orbit. The S_n-orbits of W_n are precisely the sets W_α. The result with maj_n replaced by maj thus follows immediately from Theorem 1.2.6. The reduction modulo n does not affect (1.3). $\square$

Recall the *flex* statistic from Section 3.8. The following is a restatement of Lemma 3.8.3.

Lemma. *Let $W \subset \mathrm{W}_n$ be a set of words closed under the $C_n \coloneqq \langle (1\ 2\ \cdots\ n) \rangle$-action. Then, the triple*

$$(W, C_n, W^{\mathrm{flex}}(q))$$

exhibits the CSP.

Corollary 4.3.2. *For all $n \geq 1$,*

$$\mathrm{W}_n^{\mathrm{cont,flex}}(x;q) = \mathrm{W}_n^{\mathrm{cont,maj}_n}(x;q).$$

Proof. By Proposition 4.3.1 and Lemma 3.8.3, for each $\alpha \vDash n$, the triples

$$(\mathrm{W}_\alpha, C_n, \mathrm{W}_\alpha^{\mathrm{maj}_n}(q)) \qquad \text{and} \qquad (\mathrm{W}_\alpha, C_n, \mathrm{W}_\alpha^{\mathrm{flex}}(q))$$

exhibit the CSP. By (1.3), it follows that $\mathrm{W}_\alpha^{\mathrm{maj}_n}(q) \equiv \mathrm{W}_\alpha^{\mathrm{flex}}(q) \pmod{q^n - 1}$, which forces $\mathrm{W}_\alpha^{\mathrm{maj}_n}(q) = \mathrm{W}_\alpha^{\mathrm{flex}}(q)$. The result follows by summing over all α. $\square$

Lemma 4.3.3. *For all $n \geq 1$,*

$$\sum_{r=1}^{n} \mathrm{ch}\left(\chi^r \uparrow_{C_n}^{S_n}\right) q^r = \mathrm{W}_n^{\mathrm{cont,flex}}(x;q).$$

Proof. For each necklace $N \in \mathrm{N}_n$, $\mathrm{flex}(N) = \{\mathrm{freq}(N), 2\,\mathrm{freq}(N), \ldots, n\}$, so $\{w \in \mathrm{W}_n : \mathrm{flex}(w) = r\}$ contains exactly 1 word of each necklace in $\mathrm{NFD}_{n,r}$. Consequently,

$$\mathrm{NFD}_{n,r}^{\mathrm{cont}} = \{w \in \mathrm{W}_n : \mathrm{flex}(w) = r\}^{\mathrm{cont}}. \tag{4.8}$$

By Proposition 4.2.5,

$$\begin{aligned}\sum_{r=1}^{n} \mathrm{ch}\left(\chi^r\!\uparrow_{C_n}^{S_n}\right) q^r &= \sum_{r=1}^{n} \mathrm{NFD}_{n,r}^{\mathrm{cont}}\, q^r \\ &= \sum_{r=1}^{n} \{w \in \mathrm{W}_n : \mathrm{flex}(w) = r\}^{\mathrm{cont}} q^r \\ &= \mathrm{W}_n^{\mathrm{cont,flex}}(x;q).\end{aligned}$$

□

Lemma 4.3.4. *For all $n \geq 1$,*

$$\mathrm{W}_n^{\mathrm{cont,maj}_n}(x;q) = \sum_{\substack{\lambda \vdash n \\ r \in [n]}} a_{\lambda,r} s_\lambda q^r$$

where $a_{\lambda,r} \coloneqq \#\{Q \in \mathrm{SYT}(\lambda) : \mathrm{maj}_n(Q) = r\}$.

Proof. Using RSK and the facts that $\mathrm{cont}(w) = \mathrm{cont}(P(w))$, and $\mathrm{maj}_n(w) = \mathrm{maj}_n(Q(w))$, which are both immediate from (4.1), we have

$$\begin{aligned}\mathrm{W}_n^{\mathrm{cont,maj}_n}(x;q) &= \sum_{\lambda \vdash n} \sum_{P \in \mathrm{SSYT}(\lambda)} x^{\mathrm{cont}(P)} \sum_{Q \in \mathrm{SYT}(\lambda)} q^{\mathrm{maj}_n(Q)} \\ &= \sum_{\lambda \vdash n} s_\lambda \sum_{Q \in \mathrm{SYT}(\lambda)} q^{\mathrm{maj}_n(Q)},\end{aligned}$$

from which the result follows. □

Remark 4.3.5. Kraśkiewicz–Weyman's result, Theorem 4.2.6, now follows immediately from Corollary 4.3.2, Lemma 4.3.3, and Lemma 4.3.4. Intuitively, the argument may be summarized as follows. We exhibited an explicit basis for $E(\chi^r\!\uparrow_{C_n}^{S_n})$ giving $\sum_{r=1}^{n} \mathrm{ch}\,\chi^r\!\uparrow_{C_n}^{S_n} q^r = \mathrm{W}_n^{\mathrm{cont,flex}}(x;q)$. The universal S_n sieving result, Proposition 4.3.1—or equivalently Theorem 1.2.6—and the universal C_n sieving result, Lemma 3.8.3, allows us to replace flex with maj_n. Finally, the RSK bijection allows us to change from the monomial to the Schur basis. The only step of this approach which does not use an explicit bijection is the appeal to Theorem 1.2.6. This suggests the following coarsening of Open Problem 3.8.6.

Open Problem 4.3.6. *For each $\alpha \vDash n$, find an explicit bijection $\phi\colon \mathrm{W}_\alpha \to \mathrm{W}_\alpha$ such that* $\mathrm{maj}_n(w) = \mathrm{flex}(\phi(w))$.

Conversely, Kraśkiewicz–Weyman's result together with Lemma 4.3.3 and Lemma 4.3.4 immediately yields Corollary 4.3.2. Using Lemma 3.8.3 then yields Proposition 4.3.1, or equivalently Theorem 1.2.6. In this sense, Kraśkiewicz–Weyman's result is equivalent to Theorem 1.2.6.

Remark 4.3.7. The symmetry result for the coefficients $a_{\lambda,r}$, Corollary 4.2.8, also follows immediately from Corollary 4.3.2 and Lemma 4.3.4. Moreover, a solution to Open Problem 4.3.6 together with the preceding arguments would provide a fully bijective proof of this symmetry result.

For use in the next section, we include the following variation on Corollary 4.3.2. For $n \geq 1$ and $r = 1, \ldots n$, let

$$\mathrm{M}_{n,r} \coloneqq \{w \in \mathrm{W}_n : \mathrm{maj}_n(w) = r\}.$$

Corollary 4.3.8. *For all $n \geq 1$, and $r = 1, \ldots, n$,*

$$\mathrm{NFD}_{n,r}^{\mathrm{cont}} = \mathrm{M}_{n,r}^{\mathrm{cont}}$$

Proof. By (4.8) and Corollary 4.3.2,

$$\mathrm{NFD}_{n,r}^{\mathrm{cont}} = \{w \in \mathrm{W}_n : \mathrm{flex}(w) = r\}^{\mathrm{cont}} = \{w \in \mathrm{W}_n : \mathrm{maj}_n(w) = r\}^{\mathrm{cont}} = \mathrm{M}_{n,r}^{\mathrm{cont}}.$$

$\square$

4.4 *Induced Representations of Arbitrary Cyclic Subgroups of* S_n

We next generalize the discussion in Section 4.3 to branching rules for general inclusions $\langle\sigma\rangle \hookrightarrow S_n$, recovering a result of Stembridge, Theorem 4.4.11. We express the relevant characters in turn as a certain orbit generating function, a necklace generating function, and a generating function on words. Two variations on the major index, maj_ν and maj^ν, arise quite naturally from our argument. The CSP Theorem 1.2.6 again plays a decisive role.

Throughout this section, let $\sigma \in S_n$, let C be the cyclic group generated by σ, and let $\ell = \#C$. Fixing a primitive ℓ-th root of unity ω_ℓ, let $\chi^r \colon C \to \mathbb{C}$ for $r = 1, \ldots, n$ be the linear C-module given by $\chi^r(\sigma) \coloneqq \omega_\ell^r$. We begin by updating our notation for this setting and generalizing Proposition 4.2.5.

Definition 4.4.1. In analogy with Definition 4.2.1, suppose $\mathcal{O}$ is an orbit of W_n under the restricted C-action. The *period* of $\mathcal{O}$ is $\#\mathcal{O}$ and the *frequency* of $\mathcal{O}$, written $\mathrm{freq}(\mathcal{O})$, is the stabilizer-order of any element of $\mathcal{O}$, or equivalently $\mathrm{freq}(\mathcal{O}) = \frac{\ell}{\#\mathcal{O}}$. The set of orbits of words whose frequency divides r is

$$\mathrm{OFD}_{C,r} \coloneqq \{C\text{-orbits } \mathcal{O} \text{ of } \mathrm{W}_n : \mathrm{freq}(\mathcal{O}) \mid r\}.$$

Proposition 4.4.2. *For all $m \geq 1$, there is a basis for the Schur module $E(\chi^r\uparrow_C^{S_n})$ over $\mathrm{GL}(\mathbb{C}^m)$ indexed by C-orbits of length n words with letters from $[m]$ and with frequency dividing r. Moreover,*

$$\mathrm{ch}\left(\chi^r\uparrow_C^{S_n}\right) = \mathrm{OFD}_{C,r}^{\mathrm{cont}}. \tag{4.9}$$

Proof. The proof of Proposition 4.2.5 goes through verbatim with the C-action replacing the C_n-action. □

Our goal is broadly to replace $\mathrm{OFD}_{C,r}^{\mathrm{cont}}$ with a necklace generating function, apply cyclic sieving to get a major index generating function on words, and then apply RSK to get a

Schur expansion.

Notation 4.4.3. For the rest of the section, suppose that σ has disjoint cycle decomposition $\sigma = \sigma_1 \cdots \sigma_k$ with $\nu_i \coloneqq |\sigma_i|$. Consequently, $\ell = |\sigma| = \mathrm{lcm}(\nu_1, \ldots, \nu_k)$. Further, write

$$C_\nu \coloneqq \{\sigma_1^{r_1} \cdots \sigma_k^{r_k} \in S_n : r_1, \ldots, r_k \in \mathbb{Z}\} \cong C_{\nu_1} \times \cdots \times C_{\nu_k}$$

where $C_{\nu_i} \coloneqq \langle \sigma_i \rangle$.

We have $C \subset C_\nu \subset S_n$. The C_ν-orbits of W_n can be identified with products of necklaces $N_1 \times \cdots \times N_k$ with $N_j \in \mathrm{N}_{\nu_j}$, or equivalently with tuples $(N_1, \ldots, N_k)$. Since

$$\mathrm{Stab}_{C_\nu}(N_1 \times \cdots \times N_k) = \prod_{j=1}^{k} \mathrm{Stab}_{C_{\nu_j}}(N_j),$$

we may group together C_ν-orbits of W_n according to their stabilizer as follows.

Definition 4.4.4. Given $\nu = (\nu_1, \ldots, \nu_k)$ and $\rho = (\rho_1, \ldots, \rho_k)$, let

$$\begin{aligned} \mathrm{NF}_{\nu,\rho} &\coloneqq \mathrm{NF}_{\nu_1,\rho_1} \times \cdots \times \mathrm{NF}_{\nu_k,\rho_k} \\ \mathrm{NFD}_{\nu,\rho} &\coloneqq \mathrm{NFD}_{\nu_1,\rho_1} \times \cdots \times \mathrm{NFD}_{\nu_k,\rho_k}, \end{aligned}$$

where $\mathrm{NF}_{n,r} \coloneqq \{N \in \mathrm{N}_n : \mathrm{freq}(n) = r\}$ is the set of length n necklaces with frequency precisely r.

The elements of $\mathrm{NF}_{\nu,\rho}$ all have the same stabilizer, and the elements of $\mathrm{NFD}_{\nu,\rho}$ are precisely those whose stabilizer is contained in the common stabilizer of elements of $\mathrm{NF}_{\nu,\rho}$. Note that $\mathrm{NF}_{\nu,\rho} \neq \varnothing$ if and only if $\rho_j \mid \nu_j$ for all $j = 1, \ldots, k$, which we write as $\rho \mid \nu$. We may express $\mathrm{OFD}_{C,r}$ as a union of certain $\mathrm{NF}_{\nu,\rho}$'s arising from viewing C_ν-orbits as unions of C-orbits, resulting in the following.

Lemma 4.4.5. *For $r = 1, \dots, n$, we have*

$$\mathrm{OFD}^{\mathrm{cont}}_{C,r} = \sum_{\substack{\rho \mid \nu \\ \ell \mid r \cdot \mathrm{lcm}\left(\frac{\nu_1}{\rho_1}, \dots, \frac{\nu_k}{\rho_k}\right)}} \frac{\prod_{j=1}^{k} \frac{\nu_j}{\rho_j}}{\mathrm{lcm}\left(\frac{\nu_1}{\rho_1}, \dots, \frac{\nu_k}{\rho_k}\right)} \mathrm{NF}^{\mathrm{cont}}_{\nu,\rho} .$$

Proof. Let $\mathcal{O}$ be a C-orbit of W_n with $\mathrm{freq}(\mathcal{O}) \mid r$. Now $\mathcal{O}$ is a subset of some C_ν-orbit $N_1 \times \cdots \times N_k \in \mathrm{NF}_{\nu,\rho}$ for some $\rho \mid \nu$. Since $\mathrm{freq}(N_j) = \rho_j$, we have $\#N_j = \nu_j/\rho_j$ and so $\#N_1 \times \cdots \times N_k = \prod_{j=1}^{k} \nu_j/\rho_j$. Moreover, $\mathcal{O}$ is in bijection with the group generated by a permutation of cycle type $(\nu_1/\rho_1, \dots, \nu_k/\rho_k)$, so that $\#\mathcal{O} = \mathrm{lcm}\,(\nu_1/\rho_1, \dots, \nu_k/\rho_k)$.

Note that $\mathrm{freq}(\mathcal{O}) \mid r$ is equivalent to the condition $\ell \mid r \cdot \#\mathcal{O}$, which explains the restriction in the sum above. The C_ν-orbit $N_1 \times \cdots \times N_k$ is the disjoint union of those C-orbits contained in it. Since these C-orbits have the same size $\#\mathcal{O}$, it follows that $N_1 \times \cdots \times N_k$ contains

$$\frac{\#(N_1 \times \cdots \times N_k)}{\#\mathcal{O}} = \frac{\prod_{j=1}^{k} \frac{\nu_j}{\rho_j}}{\mathrm{lcm}\left(\frac{\nu_1}{\rho_1}, \dots, \frac{\nu_k}{\rho_k}\right)}$$

distinct C-orbits. Since $\mathcal{O}$ and $(N_1, \dots, N_k)$ have the same content, the result follows. $\square$

One could in principle use Möbius inversion on the lattice of stabilizers to convert from $\mathrm{NF}^{\mathrm{cont}}_{\nu,\rho}$ to $\mathrm{NFD}^{\mathrm{cont}}_{\nu,\rho}$. However, the following argument is more direct.

Lemma 4.4.6. *For $r = 1, \dots, n$,*

$$\mathrm{OFD}^{\mathrm{cont}}_{C,r} = \sum_{\substack{\tau \in [\nu_1] \times \cdots \times [\nu_k] \\ \sum_{j=1}^{k} \frac{\ell}{\nu_j} \tau_j \equiv_\ell r}} \mathrm{NFD}^{\mathrm{cont}}_{\nu,\tau} .$$

Proof. We have

$$\mathrm{NFD}^{\mathrm{cont}}_{\nu,\tau} = \sum_{\rho : \rho \mid \nu, \tau} \mathrm{NF}^{\mathrm{cont}}_{\nu,\rho} .$$

Consequently,

$$\sum_{\substack{\tau\in[\nu_1]\times\cdots\times[\nu_k]\\ \sum_{j=1}^k \frac{\ell}{\nu_j}\tau_j\equiv_\ell r}} \mathrm{NFD}^{\mathrm{cont}}_{\nu,\tau} = \sum_{\rho\mid\nu} c^r_{\nu,\rho}\,\mathrm{NF}^{\mathrm{cont}}_{\nu,\rho}$$

where

$$c^r_{\nu,\rho} \coloneqq \#\left\{\tau\in[\nu_1]\times\cdots\times[\nu_k] : \rho\mid\tau \text{ and } \sum_{j=1}^k \frac{\ell}{\nu_j}\tau_j \equiv_\ell r\right\}.$$

Since $\rho_j \mid \nu_j$ and $\rho_j \mid \tau_j$, write $\gamma_j \coloneqq \frac{\nu_j}{\rho_j}$ and $\delta_j = \frac{\tau_j}{\rho_j}$ so that $\delta_j = 1,\ldots,\gamma_j$. Then,

$$\sum_{j=1}^k \frac{\ell}{\nu_j}\tau_j = \sum_{j=1}^k \frac{\ell}{\gamma_j}\delta_j$$

so

$$c^r_{\nu,\rho} = \#\left\{\delta\in[\gamma_1]\times\cdots\times[\gamma_k] : \sum_{j=1}^k \frac{\ell}{\gamma_j}\delta_j \equiv_\ell r\right\}.$$

Defining a group homomorphism

$$\begin{aligned}\phi\colon \prod_{i=1}^k \mathbb{Z}/\gamma_j &\to \mathbb{Z}/\ell\\ (\delta_1,\ldots,\delta_k) &\mapsto \sum_{j=1}^k \frac{\ell}{\gamma_j}\delta_j,\end{aligned}$$

we now have $c^r_{\nu,\rho} = \#\phi^{-1}(r)$. Since $\frac{\ell}{\gamma_1}\mathbb{Z}+\cdots+\frac{\ell}{\gamma_k}\mathbb{Z} = \gcd\left(\frac{\ell}{\gamma_1},\ldots,\frac{\ell}{\gamma_k}\right)\mathbb{Z} = \frac{\ell}{\mathrm{lcm}(\gamma_1,\ldots,\gamma_k)}\mathbb{Z}$, it follows that

$$\begin{aligned}\operatorname{im}\phi &= \{r\in\mathbb{Z}/\ell : \ell \mid r\cdot\mathrm{lcm}\left(\gamma_1,\ldots,\gamma_k\right)\} \text{ and}\\ \#\operatorname{im}\phi &= \mathrm{lcm}(\gamma_1,\ldots,\gamma_k).\end{aligned}$$

For $r \in \operatorname{im}\phi$, we then have

$$c^r_{\nu,\rho} = \#\phi^{-1}(r) = \#\ker\phi = \frac{\gamma_1\cdots\gamma_k}{\operatorname{lcm}(\gamma_1,\ldots,\gamma_k)}.$$

The result follows from Lemma 4.4.5. □

From Corollary 4.3.8, we have

$$\mathrm{NFD}^{\mathrm{cont}}_{\nu,\tau} = \mathrm{M}^{\mathrm{cont}}_{\nu_1,\tau_1}\times\cdots\times\mathrm{M}^{\mathrm{cont}}_{\nu_k,\tau_k}.$$

Interpreting the right-hand side in terms of words and comparing with the indexing set in Lemma 4.4.6 motivates the following variations on the major index.

Definition 4.4.7. Let $\mathrm{maj}_\nu\colon \mathrm{W}_n \to [\nu_1]\times\cdots\times[\nu_k]$ be defined as follows. For $w \in \mathrm{W}_n$, write $w = w^1\cdots w^k$ where each w^j is a word in W_{ν_j}. Set

$$\mathrm{maj}_\nu(w)_j \coloneqq \mathrm{maj}_{\nu_j}(w^j).$$

Furthermore, let $\mathrm{maj}^\nu\colon \mathrm{W}_n \to [\ell]$ be defined by

$$\mathrm{maj}^\nu(w) \coloneqq \sum_{j=1}^{k}\frac{\ell}{\nu_j}\,\mathrm{maj}_\nu(w)_j.$$

Note that both maj_ν and maj^ν are functions of $\mathrm{Des}(w)$. We may thus define both maj_ν and maj^ν on $Q \in \mathrm{SYT}(n)$ using only $\mathrm{Des}(Q)$ in the same way. Equivalently, we may set $\mathrm{maj}_\nu(Q) \coloneqq \mathrm{maj}_\nu(w)$ and $\mathrm{maj}^\nu(Q) \coloneqq \mathrm{maj}^\nu(w)$ for any w such that $Q = Q(w)$.

Example 4.4.8. Let $\nu = (5,3,3)$ and $w = 44121361631$, so that $\ell = 15$, $w_1 = 44121$, $w_2 = 361$, and $w_3 = 631$. We have

$$\mathrm{maj}_\nu(w) = (\mathrm{maj}_5(w_1), \mathrm{maj}_3(w_2), \mathrm{maj}_3(w_3)) = (1,2,3),$$

and hence $\mathrm{maj}^{\nu}(w) = \frac{15}{5} \cdot 1 + \frac{15}{3} \cdot 2 + \frac{15}{3} \cdot 3 = 13 \pmod{15}$.

Proposition 4.4.9. *We have*

$$\sum_{r=1}^{\ell} \mathrm{ch}\left(\chi^r \uparrow_C^{S_n}\right) q^r = \mathrm{W}_n^{\mathrm{cont},\mathrm{maj}^{\nu}}(x, q).$$

Proof. From Corollary 4.3.8 and the definition of maj_{ν}, we have

$$\mathrm{NFD}_{\nu,\tau}^{\mathrm{cont}} = \{w \in \mathrm{W}_n : \mathrm{maj}_{\nu}(w) = \tau\}^{\mathrm{cont}}.$$

Using Proposition 4.4.2 and Lemma 4.4.6, we then have

$$\begin{aligned}
\sum_{r=1}^{\ell} \mathrm{ch}\left(\chi^r \uparrow_C^{S_n}\right) q^r &= \sum_{r=1}^{\ell} q^r \, \mathrm{OFD}_{C,r}^{\mathrm{cont}} \\
&= \sum_{r=1}^{\ell} q^r \sum_{\substack{\tau \in [\nu_1] \times \cdots \times [\nu_k] \\ \sum_{j=1}^{k} \frac{\ell}{\nu_j} \tau_j \equiv_{\ell} r}} \mathrm{NFD}_{\nu,\tau}^{\mathrm{cont}} \\
&= \sum_{r=1}^{\ell} q^r \sum_{\substack{\tau \in [\nu_1] \times \cdots \times [\nu_k] \\ \sum_{j=1}^{k} \frac{\ell}{\nu_j} \tau_j \equiv_{\ell} r}} \{w \in \mathrm{W}_n : \mathrm{maj}_{\nu}(w) = \tau\}^{\mathrm{cont}} \\
&= \sum_{r=1}^{\ell} q^r \{w \in \mathrm{W}_n : \mathrm{maj}^{\nu}(w) = r\}^{\mathrm{cont}} \\
&= \mathrm{W}_n^{\mathrm{cont},\mathrm{maj}^{\nu}}(x, q).
\end{aligned}$$

□

We have the following analogue of Lemma 4.3.4.

Lemma 4.4.10. *For all $n \geq 1$ and $\nu \vDash n$,*

$$\mathrm{W}_n^{\mathrm{cont},\mathrm{maj}^{\nu}}(x; q) = \sum_{\substack{\lambda \vdash n \\ r \in [\ell]}} a_{\lambda,r}^{\nu} s_{\lambda} q^r$$

where $a^{\nu}_{\lambda,r} \coloneqq \#\{Q \in \mathrm{SYT}(\lambda) : \mathrm{maj}^{\nu}(Q) = r\}$.

Proof. Replace maj with maj^{ν} in the proof of Lemma 4.3.4. $\square$

We may now state and prove Stembridge's result.

Theorem 4.4.11 (Stembridge, [94, Theorem 3.3])**.** *For all* $n \geq 1$ *and cyclic subgroups* C *of* S_n *generated by an element of cycle type* ν*, we have*

$$\mathrm{ch}\left(\chi^r\uparrow_C^{S_n}\right) = \sum_{\lambda \vdash n} a^{\nu}_{\lambda,r} s_\lambda$$

where $a^{\nu}_{\lambda,r} \coloneqq \#\{Q \in \mathrm{SYT}(\lambda) : \mathrm{maj}^{\nu}(Q) = r\}$.

Proof. Combine Proposition 4.4.9 and Lemma 4.4.10. $\square$

Since the isomorphism type of $\chi^r\uparrow_C^{S_n}$, or equivalently $\mathrm{OFD}^{\mathrm{cont}}_{C,r}$, depends only on $\gcd(\ell, r)$, we have the following generalization of Corollary 4.2.8.

Corollary 4.4.12. *For all* $n \geq 1$, $\lambda \vdash n$*, and* $\nu \vDash n$*, we have* $a^{\nu}_{\lambda,r} = a^{\nu}_{\lambda,\gcd(\ell,r)}$*, where* $\ell = \mathrm{lcm}(\nu_1, \nu_2, \ldots)$.

Our argument proving Schocker's formula uses the following variation on Proposition 4.4.9 and Lemma 4.4.10. There is also a corresponding symmetry result, Corollary 4.4.14.

Lemma 4.4.13. *If* $\nu \vDash n$*, then*

$$\mathrm{NFD}^{\mathrm{cont}}_{\nu,\tau} = \sum_{\lambda \vdash n} \mathrm{syt}^{\lambda}_{\nu,\tau}\, s_\lambda$$

where

$$\mathrm{syt}^{\lambda}_{\nu,\tau} \coloneqq \#\{Q \in \mathrm{SYT}(\lambda) : \mathrm{maj}_{\nu}(Q) = \tau\}.$$

Proof. Applying RSK and the facts that $\mathrm{cont}(P(w)) = \mathrm{cont}(w)$ and $\mathrm{maj}_\nu(w) = \mathrm{maj}_\nu(Q(w))$, which are both immediate from (4.1),

$$\begin{aligned}
\mathrm{NFD}^{\mathrm{cont}}_{\nu,\tau} &= \mathrm{M}^{\mathrm{cont}}_{\nu_1,\tau_1} \times \cdots \times \mathrm{M}^{\mathrm{cont}}_{\nu_k,\tau_k} \\
&= \{w \in \mathrm{W}_n : \mathrm{maj}_\nu(w) = \tau\}^{\mathrm{cont}} \\
&= \sum_{\lambda \vdash n} \sum_{\substack{Q \in \mathrm{SYT}(\lambda) \\ \mathrm{maj}_\nu(Q)=\tau}} \sum_{P \in \mathrm{SSYT}(\lambda)} x^{\mathrm{cont}(P)} \\
&= \sum_{\lambda \vdash n} \mathrm{syt}^{\lambda}_{\nu,\tau}\, s_\lambda.
\end{aligned}$$

$\square$

Corollary 4.4.14. *If $\nu = (\nu_1, \ldots, \nu_k)$ and $\tau = (\tau_1, \ldots, \tau_k)$ are compositions of n of length k, $\sigma \in S_k$, and $\lambda \vdash n$, then $\mathrm{syt}^{\lambda}_{\nu,\beta} = \mathrm{syt}^{\lambda}_{\sigma\cdot\nu,\sigma\cdot\beta}$.*

Proof. Since reordering does not affect contents, we have

$$\mathrm{NFD}^{\mathrm{cont}}_{\nu,\tau} = \mathrm{NFD}^{\mathrm{cont}}_{\sigma\cdot\nu,\sigma\cdot\tau}\,.$$

Now apply Lemma 4.4.13 and equate coefficients of s_λ. $\square$

4.5 Higher Lie Modules and Branching Rules

The argument in Section 4.3 solves Thrall's problem for $\lambda = (n)$ by considering all branching rules for $C_n \hookrightarrow S_n$ simultaneously and using cyclic sieving and RSK to convert from the monomial to the Schur basis. We now turn to analogous considerations for the higher Lie modules and more generally branching rules for $C_a \wr S_b \hookrightarrow S_{ab}$. We give an analogue of the flex statistic and the above monomial expansion, Lemma 4.3.3, for such branching rules. We then show how to convert from the monomial to the Schur basis assuming the existence of a certain statistic on words we call mash which interpolates between maj_n and the shape under RSK.

Proposition 4.5.1. *For all* $a, b \geq 1$ *and* $\underline{\lambda}\colon [a] \to \mathrm{Par}$ *with* $|\underline{\lambda}| = b$*, we have*

$$\mathrm{ch}\, S^{\underline{\lambda}}\!\uparrow_{C_a \wr S_b}^{S_{ab}} = \prod_{r=1}^{a} s_{\underline{\lambda}(r)}[\mathrm{NFD}_{a,r}^{\mathrm{cont}}].$$

Proof. We have

$$\begin{aligned}
S^{\underline{\lambda}}\!\uparrow_{C_a \wr S_b}^{S_{ab}} &\cong \left[\bigotimes_{r=1}^{a} (\chi^r \wr S^{\underline{\lambda}(r)})\right]\uparrow_{C_a \wr S_{\alpha(\underline{\lambda})}}^{S_{ab}} \\
&\cong \left[\bigotimes_{r=1}^{a} (\chi^r \wr S^{\underline{\lambda}(r)})\right]\uparrow_{C_a \wr S_{\alpha(\underline{\lambda})}}^{S_{a*\alpha(\underline{\lambda})}}\uparrow_{S_{a*\alpha(\underline{\lambda})}}^{S_{ab}} \\
&\cong \left[\bigotimes_{r=1}^{a} (\chi^r \wr S^{\underline{\lambda}(r)})\uparrow_{C_a \wr S_{|\underline{\lambda}(r)|}}^{S_{a|\underline{\lambda}(r)|}}\right]\uparrow_{S_{a*\alpha(\underline{\lambda})}}^{S_{ab}} \\
&\cong \left[\bigotimes_{r=1}^{a} (\chi^r \wr S^{\underline{\lambda}(r)})\uparrow_{C_a \wr S_{|\underline{\lambda}(r)|}}^{S_a \wr S_{|\underline{\lambda}(r)|}}\uparrow_{S_a \wr S_{|\underline{\lambda}(r)|}}^{S_{a|\underline{\lambda}(r)|}}\right]\uparrow_{S_{a*\alpha(\underline{\lambda})}}^{S_{ab}} \\
&\cong \left[\bigotimes_{r=1}^{a} (\chi^r\!\uparrow_{C_a}^{S_a} \wr S^{\underline{\lambda}(r)})\uparrow_{S_a \wr S_{|\underline{\lambda}(r)|}}^{S_{a|\underline{\lambda}(r)|}}\right]\uparrow_{S_{a*\alpha(\underline{\lambda})}}^{S_{ab}}
\end{aligned}$$

where the last isomorphism follows from Lemma 4.2.13. Consequently, using (4.5), (4.7), and Proposition 4.2.5, we have

$$\begin{aligned}
\mathrm{ch}\, S^{\underline{\lambda}}\!\uparrow_{C_a \wr S_b}^{S_{ab}} &= \prod_{r=1}^{a} \mathrm{ch}\left(\chi^r\!\uparrow_{C_a}^{S_a} \wr S^{\underline{\lambda}(r)}\right)\uparrow_{S_a \wr S_{|\underline{\lambda}(r)|}}^{S_{a|\underline{\lambda}(r)|}} \\
&= \prod_{r=1}^{a} (\mathrm{ch}\, S^{\underline{\lambda}(r)})[\mathrm{ch}\, \chi^r\!\uparrow_{C_a}^{S_a}] \\
&= \prod_{r=1}^{a} s_{\underline{\lambda}(r)}[\mathrm{NFD}_{a,r}^{\mathrm{cont}}].
\end{aligned}$$

□

Recall from Section 4.2.2 that given a word w, the shape of w, denoted $\mathrm{sh}(w)$, is the common shape of $P(w)$ and $Q(w)$ under RSK.

Definition 4.5.2. Fix $a, b \geq 1$. Construct statistics

$$\mathrm{flex}_a^b \quad \text{and} \quad \mathrm{maj}_a^b \colon \mathrm{W}_{ab} \to \{\underline{\lambda} \colon [a] \to \mathrm{Par} \mid |\underline{\lambda}| = b\}$$

as follows. Given $w \in \mathrm{W}_{ab}$, write $w = w^1 \cdots w^b$ where $w^j \in \mathrm{W}_a$. In this way, consider w as a word of size b whose letters are in W_a. For each $r \in [a]$, let $w^{(r)}$ denote the subword of w whose letters are those w^j such that $\mathrm{flex}(w^j) = r$. Totally order W_a lexicographically, so that RSK is well-defined for words with letters from W_a. Set

$$\mathrm{flex}_a^b(w) \colon r \mapsto \mathrm{sh}(w^{(r)}).$$

Define maj_a^b in the same way except for using maj_a instead of flex when constructing the subwords $w^{(r)}$.

Example 4.5.3. Let $w = 212023101241$ and suppose $a = 3$, $b = 4$. Write

$$w = (212)(023)(101)(241).$$

The parenthesized terms have flex statistics $2, 1, 2, 2$ and maj_3 statistics $1, 3, 1, 2$, respectively. When computing $\mathrm{flex}_3^4(w)$, we then have $w^{(1)} = (023), w^{(2)} = (212)(101)(241), w^{(3)} = \varnothing$. Since $(101) <_{\mathrm{lex}} (212) <_{\mathrm{lex}} (241)$, $\mathrm{sh}(w^1) = \mathrm{sh}(213) = (2, 1)$. Consequently,

$$\mathrm{flex}_3^4(212023101241) = \begin{cases} 1 & \mapsto (1) \\ 2 & \mapsto (2,1) \\ 3 & \mapsto \varnothing. \end{cases}$$

When computing $\mathrm{maj}_3^4(w)$, we have $w^{(1)} = (212)(101), w^{(2)} = (241), w^{(3)} = (023)$. Since

$(101) <_{\text{lex}} (212)$, $\mathrm{sh}(w^1) = \mathrm{sh}(21) = (1,1)$. Hence

$$\mathrm{maj}_3^4(212023101241) = \begin{cases} 1 & \mapsto (1,1) \\ 2 & \mapsto (1) \\ 3 & \mapsto (1). \end{cases}$$

In Section 4.3 we considered the graded Frobenius series tracking branching rules for the inclusion $C_n \hookrightarrow S_n$, $\sum_{r=1}^{n} \mathrm{ch}\left(\chi^r \uparrow_{C_n}^{S_n}\right) q^r$. The $q = 1$ specialization gives the Frobenius series for the regular representation of C_n. We next consider the analogous expression for the inclusion $C_a \wr S_b \hookrightarrow S_{ab}$. Since the irreducible representations here are not all one-dimensional, we must introduce scale factors for the $q = 1$ specialization to give the Frobenius series for the regular representation of $C_a \wr S_b$. We also use formal indeterminates $q^{\underline{\lambda}}$ and extend our generating function notation from Section 2.3 accordingly.

Theorem 4.5.4. *Fix $a, b \geq 1$. We have*

$$\begin{aligned} \sum_{\substack{\underline{\lambda}\colon [a] \to \mathrm{Par} \\ |\underline{\lambda}| = b}} \dim\left(S^{\underline{\lambda}} \uparrow_{C_a \wr S_b}^{S_{ab}}\right) \mathrm{ch}\left(S^{\underline{\lambda}} \uparrow_{C_a \wr S_b}^{S_{ab}}\right) q^{\underline{\lambda}} &= \mathrm{W}_{ab}^{\mathrm{cont}, \mathrm{flex}_a^b}(x, q) \\ &= \mathrm{W}_{ab}^{\mathrm{cont}, \mathrm{maj}_a^b}(x, q) \end{aligned}$$

where the $q^{\underline{\lambda}}$ are independent indeterminates.

Proof. Fix $\underline{\lambda}\colon [a] \to \mathrm{Par}$ with $|\underline{\lambda}| = b$. For the left-hand side, we first find

$$\begin{aligned} \dim\left(S^{\underline{\lambda}} \uparrow_{C_a \wr S_b}^{S_{ab}}\right) &= \dim\left((\chi^1 \wr S^{\underline{\lambda}(1)}) \otimes \cdots \otimes (\chi^a \wr S^{\underline{\lambda}(a)})\right) \uparrow_{C_a \wr S_{\alpha(\underline{\lambda})}}^{S_{ab}} \\ &= \dim\left((\chi^1 \wr S^{\underline{\lambda}(1)}) \otimes \cdots \otimes (\chi^a \wr S^{\underline{\lambda}(a)})\right) \cdot \frac{\# C_a \wr S_b}{\# C_a \wr S_{\alpha(\underline{\lambda})}} \\ &= \binom{b}{\alpha(\underline{\lambda})} \prod_{r=1}^{a} f^{\underline{\lambda}(r)} \end{aligned}$$

where $f^\mu \coloneqq \dim S^\mu = \#\,\mathrm{SYT}(\mu)$. Thus, by Proposition 4.5.1,

$$\dim\left(S^{\underline{\lambda}}\!\uparrow_{C_a\wr S_b}^{S_{ab}}\right)\mathrm{ch}\left(S^{\underline{\lambda}}\!\uparrow_{C_a\wr S_b}^{S_{ab}}\right) = \binom{b}{\alpha(\underline{\lambda})}\prod_{r=0}^{a-1} f^{\underline{\lambda}(r)} s_{\underline{\lambda}(r)}[\mathrm{NFD}_{a,r}^{\mathrm{cont}}]. \tag{4.10}$$

For the right-hand side, we have

$$\mathrm{W}_{ab}^{\mathrm{cont},\mathrm{flex}_a^b}(x,q)\Big|_{q^{\underline{\lambda}}} = \{w \in \mathrm{W}_{ab} : \mathrm{flex}_a^b(w) = \underline{\lambda}\}^{\mathrm{cont}}.$$

In order for $w \in \mathrm{W}_{ab}$ to have $\mathrm{flex}_a^b(w) = \underline{\lambda}$, we must have $\mathrm{sh}(w^{(r)}) = \underline{\lambda}(r)$ for each $r \in [a]$. Letting $F_{a,r} \coloneqq \{w \in \mathrm{W}_a : \mathrm{flex}(w) = r\}$, we may thus choose each $w^{(r)} \in (F_{a,r})^{\alpha_r}$ with $\mathrm{sh}(w^{(r)}) = \underline{\lambda}(r)$ independently and then shuffle them in $\binom{b}{\alpha(\underline{\lambda})}$ ways to form such $w \in \mathrm{W}_{ab}$. Consequently,

$$\begin{aligned}&\{w \in \mathrm{W}_{ab} : \mathrm{flex}_a^b(w) = \underline{\lambda}\}^{\mathrm{cont}} \\ &\qquad = \binom{b}{\alpha(\underline{\lambda})}\prod_{r=1}^{a}\{w^{(r)} \in (F_{a,r})^{\alpha_r} : \mathrm{sh}(w^{(r)}) = \underline{\lambda}(r)\}^{\mathrm{cont}}.\end{aligned}$$

The content generating function for words with a given shape $\mu \vdash n$ under RSK is given by

$$\{w \in \mathrm{W}_n : \mathrm{sh}(w) = \mu\}^{\mathrm{cont}} = f^\mu s_\mu.$$

Changing the alphabet from $\mathbb{Z}_{\geq 1}$ to $F_{a,r}$ and using (4.8) gives

$$\{w^{(r)} \in (F_{a,r})^{\alpha_r} : \mathrm{sh}(w^r) = \underline{\lambda}(r)\}^{\mathrm{cont}} = f^{\underline{\lambda}(r)} s_{\underline{\lambda}(r)}[F_{a,r}^{\mathrm{cont}}] = f^{\underline{\lambda}(r)} s_{\underline{\lambda}(r)}[\mathrm{NFD}_{a,r}^{\mathrm{cont}}].$$

The first equality in the theorem now follows from combining these observations with (4.10). The second equality follows in the same way, using Corollary 4.3.2. □

While Theorem 4.5.4 determines the monomial expansion of the graded Frobenius series tracking branching rules for $C_a \wr S_b \hookrightarrow S_{ab}$, we are ultimately interested in the corresponding

Schur expansion. We next describe how the approach in the preceding sections might be used to find this Schur expansion. The key properties used in the above proof of Theorem 4.2.6 converting from the monomial to the Schur basis were that maj_n is equidistributed with flex on each W_α and $\mathrm{maj}_n(w)$ depends only on $Q(w)$. In order to apply a similar argument for $\mathrm{ch}(S^{\underline{\lambda}}\uparrow_{C_a \wr S_b}^{S_{ab}})$, we need a statistic with the following properties.

Open Problem 4.5.5. *Fix $a, b \geq 1$. Find a statistic*

$$\mathrm{mash}_a^b \colon \mathrm{W}_{ab} \to \{\underline{\lambda} \colon [a] \to \mathrm{Par} \mid |\underline{\lambda}| = b\}$$

with the following properties.

(i) For all $\alpha \vDash ab$, flex_a^b, maj_a^b, and mash_a^b are equidistributed on W_α.

(ii) If $v, w \in \mathrm{W}_{ab}$ satisfy $Q(v) = Q(w)$, then $\mathrm{mash}_a^b(v) = \mathrm{mash}_a^b(w)$.

Finding such a statistic mash_a^b would determine the Schur decomposition of $\mathrm{ch}(S^{\underline{\lambda}}\uparrow_{C_a \wr S_b}^{S_{ab}})$ as follows.

Corollary 4.5.6. *Suppose $\mathrm{mash}_a^b \colon \mathrm{W}_{ab} \to \{\underline{\lambda} \colon [a] \to \mathrm{Par} \mid |\underline{\lambda}| = b\}$ satisfies properties (i) and (ii) in Open Problem 4.5.5. Then*

$$\mathrm{ch}(S^{\underline{\lambda}}\uparrow_{C_a \wr S_b}^{S_{ab}}) = \sum_{\nu \vdash ab} \frac{\#\{Q \in \mathrm{SYT}(\nu) : \mathrm{mash}_a^b(Q) = \underline{\lambda}\}}{\dim(S^{\underline{\lambda}}\uparrow_{C_a \wr S_b}^{S_{ab}})} s_\nu,$$

where $\mathrm{mash}_a^b(Q) \coloneqq \mathrm{mash}_a^b(w)$ for any $w \in \mathrm{W}_{ab}$ with $Q(w) = Q$.

Proof. We use, in order, Theorem 4.5.4, property (i), RSK, and property (ii) to compute

$$\begin{aligned}
\sum_{\substack{\underline{\lambda}\colon [a]\to \mathrm{Par} \\ |\underline{\lambda}|=b}} \dim(S^{\underline{\lambda}}\uparrow_{C_a \wr S_b}^{S_{ab}})\,\mathrm{ch}(S^{\underline{\lambda}}\uparrow_{C_a \wr S_b}^{S_{ab}}))q^{\underline{\lambda}} &= \mathrm{W}_{ab}^{\mathrm{cont},\mathrm{maj}_a^b}(x;q) \\
&= \sum_{\alpha \vDash ab} x^\alpha\, \mathrm{W}_\alpha^{\mathrm{maj}_a^b}(q) \\
&= \sum_{\alpha \vDash ab} x^\alpha\, \mathrm{W}_\alpha^{\mathrm{mash}_a^b}(q) \\
&= \mathrm{W}_{ab}^{\mathrm{cont},\mathrm{mash}_a^b}(x;q) \\
&= \sum_{\nu \vdash ab} (\mathrm{SSYT}(\nu) \times \mathrm{SYT}(\nu))^{\mathrm{cont},\mathrm{mash}_a^b}(x;q) \\
&= \sum_{\nu \vdash ab} \mathrm{SSYT}(\nu)^{\mathrm{cont}}(x)\, \mathrm{SYT}(\nu)^{\mathrm{mash}_a^b}(q) \\
&= \sum_{\nu \vdash ab} \mathrm{SYT}(\nu)^{\mathrm{mash}_a^b}(q) s_\nu.
\end{aligned}$$

The result follows by equating coefficients of $q^{\underline{\lambda}}$. □

Remark 4.5.7. When $a = 1$ and $b = n$, we may replace $\underline{\lambda}$ with $\lambda \vdash n$. Under this identification, $\mathrm{flex}_1^n(w) = \mathrm{sh}(w)$, which clearly also satisfies property (ii). When $a = n$ and $b = 1$, we may replace $\underline{\lambda}$ with an element $r \in [n]$. Under this identification, we may set $\mathrm{mash}_n^1(w) = \mathrm{maj}_n(w)$, which also satisfies (ii). In this sense mash_a^b interpolates between the major index maj_n and the shape under RSK, hence the name.

While maj_a^b trivially satisfies property (i), it fails property (ii) already when $a = b = 2$, as in the following example.

Example 4.5.8. Let $v = 2314$ and $w = 1423$. Then,

$$Q(v) = Q(w) = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1}\end{array}$$

while

$$\mathrm{maj}_2^2(v)\colon \begin{cases} 1 & \mapsto \varnothing \\ 2 & \mapsto (1,1) \end{cases}$$

$$\mathrm{maj}_2^2(w)\colon \begin{cases} 1 & \mapsto \varnothing \\ 2 & \mapsto (2) \end{cases}$$

Remark 4.5.9. When defining flex_a^b and maj_a^b, we somewhat arbitrarily chose the lexicographic order on W_a. Any other total order would work just as well. However, property (ii) continues to fail using any other total order when $a = b = 2$ in Example 4.5.8 since either $14 < 23$ or $23 < 14$.

Chapter 5

ON THE EXISTENCE OF TABLEAUX WITH GIVEN MODULAR MAJOR INDEX

This chapter has been published as [97]. Sundaram's motivating work has now been published as [95]. See Chapter 6 for a more recent alternate proof of the classification in Theorem 5.1.3.

5.1 Main Results

As in Section 1.2, in this chapter we focus on the counts

$$a_{\lambda,r} \coloneqq \#\{T \in \mathrm{SYT}(\lambda) : \operatorname{maj} T \equiv_n r\}$$

where r is taken mod n. To avoid giving undue weight to trivial cases, we take $n \geq 1$ throughout. Work due to Klyachko and, later, Kraśkiewicz–Weyman's result Theorem 1.3.1, gives the following.

Theorem 5.1.1 ([50, Proposition 2], [54])**.** *Let $\lambda \vdash n$ and $n \geq 1$. The constant $a_{\lambda,1}$ is positive except in the following cases, when it is zero:*

- *$\lambda = (2,2)$ or $\lambda = (2,2,2)$;*
- *$\lambda = (n)$ when $n > 1$; or $\lambda = (1^n)$ when $n > 2$.*

Let $\chi^r\!\uparrow_{C_n}^{S_n}$ be as in Section 4.2.4, so that the multiplicity of S^λ in $\chi^r\!\uparrow_{C_n}^{S_n}$ is $a_{\lambda,r}$. The following recent conjecture due to Sundaram was originally stated in terms of the multiplicity of S^λ in $1\!\uparrow_{C_n}^{S_n}$.

Conjecture 5.1.2 (Sundaram [95]). *Let $\lambda \vdash n$ and $n \geq 1$. Then $a_{\lambda,0}$ is positive except in the following cases, when it is zero: $n > 1$ and*

- $\lambda = (n-1, 1)$
- $\lambda = (2, 1^{n-2})$ *when n is odd*
- $\lambda = (1^n)$ *when n is even.*

Conjecture 5.1.2 is the $r = 0$ case of the following theorem, which is the main result of this chapter.

Theorem 5.1.3. *Let $\lambda \vdash n$ and $1 \leq r \leq n$. Then $a_{\lambda,r}$ is positive except in the following cases, when it is zero: $n > 1$ and*

- $\lambda = (2, 2)$, $r = 1, 3$; *or* $\lambda = (2, 2, 2)$, $r = 1, 5$; *or* $\lambda = (3, 3)$, $r = 2, 4$;
- $\lambda = (n-1, 1)$ *and* $r = 0$;
- $\lambda = (2, 1^{n-2})$, $r = \begin{cases} 0 & \text{if } n \text{ is odd} \\ \frac{n}{2} & \text{if } n \text{ is even}; \end{cases}$
- $\lambda = (n)$, $r \in \{1, \ldots, n-1\}$;
- $\lambda = (1^n)$, $r \in \begin{cases} \{1, \ldots, n-1\} & \text{if } n \text{ is odd} \\ \{0, \ldots, n-1\} - \{\frac{n}{2}\} & \text{if } n \text{ is even.} \end{cases}$

Equivalently, every irreducible representation appears in each $\chi^r\uparrow_{C_n}^{S_n}$ or $S^\lambda\downarrow_{C_n}^{S_n}$ except in the noted exceptional cases.

M. Johnson [46] gave an alternative proof of Klyachko's result, Theorem 5.1.1, involving explicit constructions with standard tableaux. Kovács–Stöhr [53] gave a different proof using the Littlewood–Richardson rule which also showed that $a_{\lambda,1} > 1$ implies $a_{\lambda,1} \geq \frac{n}{6} - 1$. Our approach is instead based on normalized symmetric group character estimates. It has the benefit of yielding both more general and vastly more precise estimates for $a_{\lambda,r}$.

Our starting point is the following character formula. See Section 5.3 for further discussion of its origins and a generalization. Let $\chi^\lambda(\mu)$ denote the character of S^λ at a permutation of cycle type μ. We write $\ell^{n/\ell}$ for the rectangular partition $(\ell, \ldots, \ell)$ with ℓ columns and n/ℓ rows. Write $f^\lambda \coloneqq \chi^\lambda(1^n) = \dim S^\lambda = \#\,\mathrm{SYT}(\lambda)$.

Theorem 5.1.4. *Let $\lambda \vdash n$ and $n \geq 1$. For all $r \in \mathbb{Z}/n$,*

$$\frac{a_{\lambda,r}}{f^\lambda} = \frac{1}{n} + \frac{1}{n} \sum_{\substack{\ell | n \\ \ell \neq 1}} \frac{\chi^\lambda(\ell^{n/\ell})}{f^\lambda} c_\ell(r)$$

where

$$c_\ell(r) \coloneqq \mu\left(\frac{\ell}{\gcd(\ell, r)}\right) \frac{\phi(\ell)}{\phi(\ell/\gcd(\ell, r))}$$

is a Ramanujan sum, μ is the classical Möbius function, and ϕ is Euler's totient function.

We estimate the quotients in the preceding formula using the following result due to Fomin and Lulov.

Theorem 5.1.5. *[27, Theorem 1.1] Let $\lambda \vdash n$ where $n = \ell s$. Then*

$$|\chi^\lambda(\ell^s)| \leq \frac{s!\ell^s}{(n!)^{1/\ell}} (f^\lambda)^{1/\ell}.$$

The character formula in Theorem 5.1.4 and the Fomin-Lulov bound are combined below to give the following asymptotic uniform distribution result.

Theorem 5.1.6. *For all $\lambda \vdash n \geq 1$ and all r,*

$$\left|\frac{a_{\lambda,r}}{f^\lambda} - \frac{1}{n}\right| \leq \frac{2n^{3/2}}{\sqrt{f^\lambda}}. \tag{5.1}$$

In Section 5.4 we use "opposite hook lengths" to give a lower bound for f^λ, Corollary 5.4.13. These bounds, together with a somewhat more careful analysis involving the character formula, Stirling's approximation, and the Fomin-Lulov bound, are used to deduce both our main result, Theorem 5.1.3, and the following more explicit uniform distribution result.

Theorem 5.1.7. *Let $\lambda \vdash n$ be a partition where $f^\lambda \geq n^5 \geq 1$. Then for all r,*

$$\left| \frac{a_{\lambda,r}}{f^\lambda} - \frac{1}{n} \right| < \frac{1}{n^2}.$$

In particular, if $n \geq 81$, $\lambda_1 < n-7$, and $\lambda_1' < n-7$, then $f^\lambda \geq n^5$ and the inequality holds.

Indeed, the upper bound in Theorem 5.1.7 is quite weak and is intended only to convey the flavor of the distribution of $(a_{\lambda,r})_{r=0}^{n-1}$ for fixed λ. One may use Roichman's asymptotic estimate [76] of $|\chi^\lambda(\ell^s)|/f^\lambda$ to prove exponential decay in many cases. Moreover, one typically expects f^λ to grow super-exponentially, i.e. like $(n!)^\epsilon$ for some $\epsilon > 0$ (see [57] for some discussion and a more recent generalization of Roichman's result), which in turn would give a super-exponential decay rate in Theorem 5.1.7. We have no need for such explicit, refined statements and so have not pursued them further.

Theorem 5.1.5 is based on the following generalization of the hook length formula (the $\ell = 1$ case), which seems less well-known than it deserves. We give an alternate proof of Theorem 5.1.8 in Section 5.5 along with further discussion. A *ribbon* is a connected skew shape with no 2×2 rectangles. For $\lambda \vdash n$, write $c \in \lambda$ to mean that c is a cell in λ. Further write h_c for the *hook length* of c and write $[n] \coloneqq \{1, 2, \ldots, n\}$.

Theorem 5.1.8 ([45, 2.7.32]; see also [27, Corollary 2.2])**.** *Let $\lambda \vdash n$ where $n = \ell s$. Then*

$$|\chi^\lambda(\ell^s)| = \frac{\prod_{\substack{i \in [n] \\ i \equiv_\ell 0}} i}{\prod_{\substack{c \in \lambda \\ h_c \equiv_\ell 0}} h_c} \tag{5.2}$$

whenever λ can be written as s successive ribbons of length ℓ (i.e. whenever the ℓ-core of λ is empty), and 0 otherwise.

Other work on q-analogues of the hook length formula has focused on algebraic generalizations and variations on the hook walk algorithm rather than evaluations of symmetric group characters. For instance, an application of Kerov's q-analogue of the hook walk algorithm [48] was to prove a recursive characterization of the right-hand side of (5.5) below. See [17, §6] for a relatively recent overview of literature in this direction.

The rest of the chapter is organized as follows. In Section 5.2, we recall earlier work. In Section 5.3 we discuss and generalize Theorem 5.1.4. In Section 5.4, we use symmetric group character estimates and a new estimate involving "opposite hook products," Proposition 5.4.5, to deduce our main results, Theorem 5.1.3 and Theorem 5.1.7. We give an alternative proof of Theorem 5.1.8 in Section 5.5. In Section 5.6, we briefly discuss unimodality of symmetric group characters in light of Proposition 5.4.5.

5.2 Cyclic Exponents and Ramanujan Sums

Here we review objects famously studied by Springer [86, (4.5)] and Stembridge [94] and give further background for use in later sections. All representations will be finite-dimensional over $\mathbb{C}$.

Let G be a finite group, $g \in G$ a fixed element of order n, M a finite dimensional G-module, and ω_n a fixed primitive nth root of unity. Suppose $\{\omega_n^{e_1}, \omega_n^{e_2}, \ldots\}$ is the multiset of eigenvalues of g acting on M. The multiset $\{e_1, e_2, \ldots\}$ lists the *cyclic exponents* of g on M; these integers are well-defined mod n. Following [94], define the corresponding "modular" generating function as

$$P_{M,g}(q) \coloneqq q^{e_1} + q^{e_2} + \cdots \qquad (\mathrm{mod}\ (q^n - 1)).$$

Write $\chi^M(g)$ to denote the character of M at g. Note that

$$P_{M,g}(\omega_n^s) = \chi^M(g^s), \tag{5.3}$$

so that for instance $P_{M,g}(q)$ depends only on the conjugacy class of g. When $G = S_n$ and $g \in S_n$ has cycle type $\mu \vdash n$, we write $P_{M,\mu}(q) \coloneqq P_{M,g}(q)$.

Theorem 5.2.1 (see [94, Theorem 3.3] and [54]). *Let $\lambda \vdash n$. The cyclic exponents of $(1\ 2\ \cdots\ n)$ on S^λ are the major indices of* $\mathrm{SYT}(\lambda)$, *mod n, and*

$$\begin{aligned} P_{S^\lambda,(n)}(q) &\equiv \sum_{T \in \mathrm{SYT}(\lambda)} q^{\operatorname{maj} T} \\ &\equiv \sum_{r \mid n} a_{\lambda,r} \left(\sum_{\substack{1 \leq i \leq n \\ \gcd(i,n)=r}} q^i \right) \qquad (\textit{mod } (q^n - 1)). \end{aligned} \tag{5.4}$$

We also recall Stanley's q-analogue of the hook length formula, stated above in Theorem 1.1.3. Using it to compute cyclotomic factorizations gives a particularly efficient method for computing the coefficients $a_{\lambda,r}$.

Theorem 5.2.2. *[91, 7.21.5] Let $\lambda \vdash n$ with $\lambda = (\lambda_1, \lambda_2, \ldots)$. Then*

$$\sum_{T \in \mathrm{SYT}(\lambda)} q^{\operatorname{maj}(T)} = \frac{q^{b(\lambda)} [n]_q!}{\prod_{c \in \lambda} [h_c]_q} \tag{5.5}$$

where $b(\lambda) \coloneqq \sum (i-1)\lambda_i$.

Finally, we have need of the so-called Ramanujan sums.

Definition 5.2.3. Given $j \in \mathbb{Z}_{>0}$ and $s \in \mathbb{Z}$, the corresponding Ramanujan sum is

$$c_j(s) \coloneqq \text{the sum of the } s\text{th powers of the primitive } j\text{th roots of unity.}$$

For instance, $c_4(2) = i^2 + (-i)^2 = -2 = \mu(4/2)\phi(4)/\phi(2)$. The equivalence of this definition of $c_j(s)$ and the formula in Theorem 5.1.4 is classical and was first given by Hölder; see [51, Lemma 7.2.5] for a more modern account. These sums satisfy the well-known relation

$$\sum_{v|n} c_v(n/s)c_r(n/v) = \begin{cases} n & r = s \\ 0 & r \neq s \end{cases} \tag{5.6}$$

for all $s, r \mid n$ [51, Lemma 7.2.2].

5.3 Generalizing the Character Formula

In this section we discuss Theorem 5.1.4 and present a straightforward generalization. We begin with a proof of Theorem 5.1.4 similar to but different from that in [20]. It is included chiefly because of its simplicity given the background in Section 5.2 and because part of the argument will be used below in Section 5.5.

Proof of Theorem 5.1.4. Pick $s \mid n$, so $(12\cdots n)^s$ has cycle type $((n/s)^s)$. Evaluating (5.4) at $q = \omega_n^s$ gives

$$\chi^\lambda((n/s)^s) = P_{S^\lambda,(n)}(\omega_n^s) = \sum_{r|n} a_{\lambda,r} c_{n/r}(s) \tag{5.7}$$

since $(\omega_n^s)^i = (\omega_n^i)^s$ and ω_n^i is a primitive $n/\gcd(i,n)$th root of unity. Equation (5.7) gives a system of linear equations, one for each s such that $s \mid n$, and with variables $a_{\lambda,r}$ for each $r \mid n$. The coefficient matrix is $C \coloneqq (c_{n/r}(s))_{s|n,r|n}$. For example, the $s = n$ linear equation reads

$$f^\lambda = \chi^\lambda(1^n) = \sum_{r|n} a_{\lambda,r}\phi(n/r),$$

which follows immediately from the fact that $f^\lambda = \sum_{r=0}^{n-1} a_{\lambda,r}$ and that $a_{\lambda,r}$ depends only on $\gcd(r, n)$.

As it happens, the coefficient matrix C is nearly its own inverse. Precisely,

$$(c_{n/r}(s))^2_{s|n,r|n} = n\,I, \tag{5.8}$$

where I is the identity matrix with as many rows as positive divisors of n. It is easy to see that (5.8) is equivalent to the identity (5.6) above. Using (5.8) to invert (5.7) gives

$$a_{\lambda,r}n = \sum_{s|n} \chi^\lambda((n/s)^s)c_{n/s}(r).$$

For the $s = n$ term, we have $c_1(r) = 1$ and $\chi^\lambda(1^n) = f^\lambda$. Tracking this term separately, dividing by n and replacing s with $\ell \coloneqq n/s$ now gives Theorem 5.1.4, completing the proof. $\square$

Variations on Theorem 5.1.4 have appeared in the literature numerous times in several guises, sometimes implicitly (see [20, Théorème 2.2], [50, (7)], or [91, 7.88(a), p. 541]). In this section we write out a precise and relatively general version of these results which explicitly connects Theorem 5.1.4 to the well-known corresponding symmetric function expansion due to H. O. Foulkes. Let ch denote the Frobenius characteristic map and let p_λ denote the power symmetric function indexed by the partition λ.

Theorem 5.3.1. *[28, Theorem 1] Suppose $\lambda \vdash n \geq 1$ and $r \in \mathbb{Z}/n$. Then*

$$\mathrm{ch}\, \chi^r\!\uparrow_{C_n}^{S_n} = \frac{1}{n} \sum_{\ell|n} c_\ell(r) p_{(\ell^{n/\ell})}. \tag{5.9}$$

The following straightforward result, essentially implicit in [91, 7.88(a), p. 541], connects and generalizes Theorem 5.3.1 and Theorem 5.1.4.

Theorem 5.3.2. *Let H be a subgroup of S_n and let M be a finite-dimensional H-module with character $\chi^M \colon H \to \mathbb{C}$. Then*

$$\mathrm{ch}\, M\!\uparrow_H^{S_n} = \frac{1}{|H|} \sum_{\mu \vdash n} c_\mu p_\mu \tag{5.10}$$

and, for all $\lambda \vdash n$,

$$\langle M\!\uparrow_H^{S_n}, S^\lambda \rangle = \frac{1}{|H|} \sum_{\mu \vdash n} c_\mu \chi^\lambda(\mu), \tag{5.11}$$

where

$$c_\mu \coloneqq \sum_{\substack{h \in H \\ \tau(h)=\mu}} \chi^M(h)$$

and $\tau(\sigma)$ denotes the cycle type of the permutation σ.

Proof. Write $N \coloneqq M{\uparrow}_H^{S_n}$. By definition (see [91, p. 351]),

$$\operatorname{ch} N = \sum_{\mu \vdash n} \frac{\chi^N(\mu)}{z_\mu} p_\mu \tag{5.12}$$

where z_μ is the order of the stabilizer of any permutation of cycle type μ under conjugation. From the induced character formula (see [82, 7.2, Prop. 20]), we have

$$\chi^N(\sigma) = \frac{1}{|H|} \sum_{\substack{a \in S_n \\ \text{s.t. } a\sigma a^{-1} \in H}} \chi^M(a\sigma a^{-1}).$$

Say $\tau(\sigma) = \mu$. Each $a\sigma a^{-1} = h \in H$ with $\tau(h) = \mu$ appears in the preceding sum z_μ times, since σ and h are conjugate and z_μ is also the number of ways to conjugate any fixed permutation with cycle type μ to any other fixed permutation with cycle type μ. Hence

$$\chi^N(\mu) = \frac{1}{|H|} \sum_{\substack{h \in H \\ \tau(h)=\mu}} z_\mu \chi^M(h). \tag{5.13}$$

Equation (5.10) now follows from (5.12) and (5.13). Equation (5.11) follows from (5.10) in the usual way using the fact (see [91, (7.76)]) that $p_\mu = \sum_\lambda \chi^\lambda(\mu) s_\lambda$. $\square$

Note that (5.10) specializes to Theorem 5.3.1 and (5.11) specializes to Theorem 5.1.4 when $M = \chi^r$. In that case, the only possibly non-zero c_μ arise from $\mu = (\ell^{n/\ell})$ for $\ell \mid n$.

One may consider analogues of the counts $a_{\lambda,r}$ obtained by inducing other one-dimensional representations of subgroups of S_n. Motivated by the study of so-called higher Lie modules, there is a natural embedding of reflection groups $C_a \wr S_b \hookrightarrow S_{ab}$. A classification analogous to Klyachko's result, Theorem 5.1.1, was asserted for $b = 2$ by Schocker [80, Theorem 3.4], though the "rather lengthy proof" making "extensive use of routine applications of the Littlewood-Richardson rule and some well-known results from the theory of plethysms" was omitted. By contrast, our approach using Theorem 5.3.2 may be pushed through in this case using an appropriate generalization of the Fomin-Lulov bound, such as [57, Theorem 1.1], resulting in analogues of Theorem 5.1.3 and Theorem 5.1.7. Our approach begins to break down when b is large relative to $n = ab$ and (5.11) has many terms. However, we have no current need for such generalizations and so have not pursued them further.

5.4 Proof of the Main Results

We now turn to the proofs of Theorem 5.1.3, Theorem 5.1.6, and Theorem 5.1.7. We begin by combining the Fomin–Lulov bound and Stirling's approximation, which quickly gives Theorem 5.1.6. We then use somewhat more careful estimates to give a sufficient condition, $f^\lambda \geq n^3$, for $a_{\lambda,r} \neq 0$. Afterwards we give an inequality between hook length products and "opposite" hook length products, Proposition 5.4.5, from which we classify λ for which $f^\lambda < n^3$. Theorem 5.1.3 follows in almost all cases, with the remainder being handled by brute force computer verification and case-by-case analysis. Theorem 5.1.7 will be similar, except the bound $f^\lambda < n^5$ will be used.

Lemma 5.4.1. *Suppose $\lambda \vdash n = \ell s$. Then*

$$\ln \frac{|\chi^\lambda(\ell^s)|}{f^\lambda} \leq \left(1 - \frac{1}{\ell}\right)\left[\frac{1}{2}\ln n - \ln f^\lambda + \ln\sqrt{2\pi}\right] + \frac{\ell}{12n} - \frac{1}{2}\ln \ell. \tag{5.14}$$

Proof. We apply the following version of Stirling's approximation [85, (1.53)]. For all $m \in \mathbb{Z}_{>0}$,

$$\left(m + \frac{1}{2}\right)\ln m - m + \ln\sqrt{2\pi} \leq \ln m! \leq \left(m + \frac{1}{2}\right)\ln m - m + \ln\sqrt{2\pi} + \frac{1}{12m}.$$

The Fomin–Lulov bound, Theorem 5.1.5, gives

$$\frac{|\chi^\lambda(\ell^s)|}{f^\lambda} \leq \frac{\frac{n}{\ell}!\,\ell^{n/\ell}}{(n!)^{1/\ell}(f^\lambda)^{1-1/\ell}}.$$

Combining these gives

$$
\begin{aligned}
\ln \frac{|\chi^\lambda(\ell^s)|}{f^\lambda} &\leq \ln \left(\frac{n}{\ell}\right)! + \frac{n}{\ell}\ln \ell - \frac{1}{\ell}\ln n! - \left(1 - \frac{1}{\ell}\right)\ln f^\lambda \\
&\leq \left(\frac{n}{\ell} + \frac{1}{2}\right)\ln \frac{n}{\ell} - \frac{n}{\ell} + \ln\sqrt{2\pi} + \frac{\ell}{12n} + \frac{n}{\ell}\ln \ell \\
&\quad - \frac{1}{\ell}\left(\left(n + \frac{1}{2}\right)\ln n - n + \ln\sqrt{2\pi}\right) - \left(1 - \frac{1}{\ell}\right)\ln f^\lambda \\
&= \frac{1}{2}\ln \frac{n}{\ell} + \ln\sqrt{2\pi} + \frac{\ell}{12n} - \frac{1}{2\ell}\ln n - \frac{\ln\sqrt{2\pi}}{\ell} - \left(1 - \frac{1}{\ell}\right)\ln f^\lambda.
\end{aligned}
$$

Rearranging this final expression gives (5.14). □

We may now prove Theorem 5.1.6.

Proof of Theorem 5.1.6. For $2 \leq \ell \leq n$, applying simple term-by-term estimates to (5.14) gives

$$\ln \frac{|\chi^\lambda(\ell^s)|}{f^\lambda} \leq \frac{1}{2}\ln n - \frac{1}{2}\ln f^\lambda + \ln\sqrt{2\pi} + \frac{1}{12} - \frac{\ln 2}{2}.$$

Consequently,

$$\frac{|\chi^\lambda(\ell^s)|}{f^\lambda} \leq C\sqrt{\frac{n}{f^\lambda}}$$

where $C = \sqrt{\pi}\exp(1/12) \approx 1.93 < 2$. The Ramanujan sums $c_\ell(r)$ have the trivial bound $|c_\ell(r)| \leq \ell \leq n$. The estimate in Theorem 5.1.6 now follows immediately from Theorem 5.1.4. □

Lemma 5.4.2. *Pick $\lambda \vdash n$ and $d \in \mathbb{R}$. Suppose for all $1 \neq \ell \mid n$ where λ may be written as $s := n/\ell$ successive ribbons each of length ℓ that*

$$\frac{|\chi^\lambda(\ell^s)|}{f^\lambda} \leq \frac{1}{n^d \phi(\ell)}. \tag{5.15}$$

Then for all $r \in \mathbb{Z}/n$,

$$\left|\frac{a_{\lambda,r}}{f^\lambda} - \frac{1}{n}\right| < \frac{1}{n^d}.$$

Proof. By Theorem 5.1.4, we must show

$$\frac{1}{n}\left|\sum_{\substack{\ell|n\\ \ell\neq 1}}\frac{\chi^\lambda(\ell^s)}{f^\lambda}c_\ell(r)\right| < \frac{1}{n^d}.$$

Using the explicit form for $c_\ell(r)$ in Theorem 5.1.4 and the fact that n has fewer than n proper divisors, it suffices to show

$$\left|\frac{\chi^\lambda(\ell^s)}{f^\lambda}\phi(\ell)\right| \leq \frac{1}{n^d}$$

for all $\ell \mid n$, $\ell \neq 1$, so the result follows from our assumption (5.15). □

Corollary 5.4.3. *Let $\lambda \vdash n$. If $f^\lambda \geq n^3 \geq 1$, then $a_{\lambda,r} \neq 0$.*

Proof. Equation (5.14) gives

$$\ln\frac{|\chi^\lambda(\ell^s)|}{f^\lambda} \leq \left(1-\frac{1}{\ell}\right)\left[-\frac{5}{2}\ln n + \ln\sqrt{2\pi}\right] + \frac{\ell}{12n} - \frac{1}{2}\ln\ell. \tag{5.16}$$

At $\ell = 2$, the right-hand side of (5.16) is less than $\ln\frac{1}{\phi(2)n}$ for $n \geq 3$. At $\ell = 3, 4, 5$, the same expression is less than $\ln\frac{1}{\phi(\ell)n}$ for $n \geq 4, 3, 5$, respectively. At $\ell \geq 6$, applying simple term-by-term estimates to (5.16) gives

$$\ln\frac{|\chi^\lambda(\ell^s)|}{f^\lambda} \leq -\left(1-\frac{1}{6}\right)\frac{5}{2}\ln n + \ln\sqrt{2\pi} + \frac{1}{12} - \frac{1}{2}\ln 6 \tag{5.17}$$

which is less than $\ln\frac{1}{n^2}$ for $n \geq 4$. Thus, Lemma 5.4.2 applies with $d = 1$ for all $n \geq 5$, so that

$$\left|\frac{a_{\lambda,r}}{f^\lambda} - \frac{1}{n}\right| < \frac{1}{n},$$

and in particular $a_{\lambda,r} \neq 0$. The cases $1 \leq n \leq 4$ remain, but they may be easily checked by hand. □

We next give techniques that are well-adapted to classifying $\lambda \vdash n$ for which $f^\lambda < n^d$ for fixed d. We begin with a curious observation, Proposition 5.4.5, which is similar in flavor to

[27, Theorem 2.3]. It was also recently discovered independently by Morales–Panova–Pak as a corollary of the Naruse hook length formula for skew shapes; see [65, Proposition 12.1]. See also [69] for further discussion and an alternate proof of a stronger result by F. Petrov.

Definition 5.4.4. Consider a partition $\lambda = (\lambda_1, \ldots, \lambda_m)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ as a set of cells (in French notation)

$$\lambda = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} : 1 \leq b \leq m, 1 \leq a \leq \lambda_b\}.$$

Given a cell $c = (a, b) \in \lambda \subset \mathbb{N} \times \mathbb{N}$, the *opposite hook length* h_c^{op} at c is $a + b - 1$. For instance, the unique cell in $\lambda = (1)$ has opposite hook length 1, and the opposite hook length increases by 1 for each north or east step.

It is easy to see that $\sum_{c\in\lambda} h_c^{\mathrm{op}} = \sum_{c\in\lambda} h_c$. On the other hand, we have the following inequality for their products.

Proposition 5.4.5. *For all partitions λ,*

$$\prod_{c\in\lambda} h_c^{\mathrm{op}} \geq \prod_{c\in\lambda} h_c.$$

Moreover, equality holds if and only if λ is a rectangle.

Proof. If λ is a rectangle, the multisets $\{h_c^{\mathrm{op}}\}$ and $\{h_c\}$ are equal, so the products agree. The converse will be established in the course of proving the inequality. For that, we begin with a simple lemma.

Lemma 5.4.6. *Let $x_1 \geq \cdots \geq x_m \geq 0$ and $y_1 \geq \cdots \geq y_m \geq 0$ be real numbers. Then*

$$\prod_{i=1}^{m}(x_i + y_i) \leq \prod_{i=1}^{m}(x_i + y_{m-i+1}).$$

Moreover, equality holds if and only if for all i either $x_i = x_{m-i+1}$ or $y_i = y_{m-i+1}$.

Proof. If $m = 1$, the result is trivial. If $m = 2$, we compute

$$(x_1 + y_2)(x_2 + y_1) - (x_1 + y_1)(x_2 + y_2) = (x_1 - x_2)(y_1 - y_2) \geq 0.$$

The result follows in general by pairing terms i and $m - i + 1$ and using these base cases. $\square$

Returning to the proof of the proposition, the strategy will be to break up h_c and h_c^{op} in terms of (co-)arm and (co-)leg lengths, and apply the lemma to each column of λ when computing $\prod h_c$, or equivalently to each row of λ when computing $\prod h_c^{\mathrm{op}}$. More precisely, let $c = (a, b) \in \lambda$. Take $\lambda = (\lambda_1, \lambda_2, \ldots)$ and $\lambda' = (\lambda'_1, \lambda'_2, \ldots)$. Define the *co-arm length* of c as a, the *co-leg length* of c as b, the *arm length* of c as $\alpha \coloneqq \alpha(a, b) \coloneqq \lambda_b - a + 1$, and the *leg length* of c as $\beta \coloneqq \beta(a, b) \coloneqq \lambda'_a - b + 1$; see Figure 5.1. With these definitions, we have

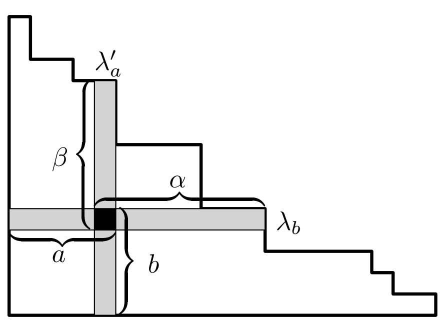


Figure 5.1: Arm length α, co-arm length a, leg length β, co-leg length b for $c = (a, b) \in \lambda$. The hook length is $h_c = \alpha + \beta - 1$ and the opposite hook length is $h_c^{\mathrm{op}} = a + b - 1$

$h_c^{\mathrm{op}} = a + b - 1$ and $h_c = \alpha + \beta - 1$. We now compute

$$\begin{aligned}
\prod_{c \in \lambda} h_c^{\mathrm{op}} &= \prod_{(a,b)\in\lambda} (a+b-1) = \prod_b \prod_{a=1}^{\lambda_b} (a+b-1) \\
&= \prod_b \prod_{a=1}^{\lambda_b} ((\lambda_b + 1 - a) + b - 1) = \prod_a \prod_{b=1}^{\lambda'_a} (\alpha + b - 1) \\
&\geq \prod_a \prod_{b=1}^{\lambda'_a} (\alpha + (\lambda'_a + 1 - b) - 1) \\
&= \prod_{(a,b)\in\lambda} (\alpha + \beta - 1) = \prod_{c\in\lambda} h_c,
\end{aligned}$$

where Lemma 5.4.6 is used for the inequality with $i \coloneqq b$, $m \coloneqq \lambda'_a$, $x_i \coloneqq \alpha - 1 = \lambda_b - a$, $y_i \coloneqq \lambda'_a + 1 - b$. Moreover, if equality occurs, then since the y_i strictly decrease, we must have $\lambda_1 = \lambda_m$ for all a, forcing λ to be a rectangle. $\square$

It would be interesting to find a bijective explanation for Proposition 5.4.5. The appearance of rectangles is particularly striking. Note, however, that $n!/\prod_{c\in\lambda} h_c^{\mathrm{op}}$ need not be an integer. In any case, we continue towards Theorem 5.1.3.

Definition 5.4.7. Define the *diagonal preorder* on partitions as follows. Declare $\lambda \lesssim^{\mathrm{diag}} \mu$ if and only if for all $i \in \mathbb{P}$,

$$\#\{c \in \lambda : h_c^{\mathrm{op}} \geq i\} \leq \#\{d \in \mu : h_d^{\mathrm{op}} \geq i\}.$$

Note that $\lesssim^{\mathrm{diag}}$ is reflexive and transitive, though not anti-symmetric, so the diagonal preorder is not a partial order. For example, the partitions $(3,1)$, $(2,2)$, and $(2,1,1)$ all have the same number of cells with each opposite hook length. A straightforward consequence of the definition is that

$$\lambda \lesssim^{\mathrm{diag}} \mu \qquad \Rightarrow \qquad \prod_{c\in\lambda} h_c^{\mathrm{op}} \leq \prod_{d\in\mu} h_d^{\mathrm{op}}. \tag{5.18}$$

Hooks are maximal elements of the diagonal preorder in a sense we next make precise.

Definition 5.4.8. Let $\lambda \vdash n$ for $n \geq 1$. The *diagonal excess* of λ is

$$N(\lambda) \coloneqq |\lambda| - \max_{c \in \lambda} h_c^{\mathrm{op}}.$$

For instance, $\lambda = (3,3)$ has opposite hook lengths ranging from 1 to 4, so $N((3,3)) = 6-4 = 2$.

The following simple observation will be used shortly.

Proposition 5.4.9. *Let $\lambda \vdash n$ for $n \geq 1$. Take $\pi \colon \lambda \to \mathbb{P}$ via $\pi(c) \coloneqq h_c^{\mathrm{op}}$. Then the fiber sizes $|\pi^{-1}(i)|$ are unimodal, and are indeed of the form*

$$1 = |\pi^{-1}(1)| < \cdots < m = |\pi^{-1}(m)| \geq |\pi^{-1}(m+1)| \geq \cdots$$

for some unique $m \geq 1$.

Proof. This follows quickly by considering the largest staircase shape contained in λ. Indeed, m is the number of rows or columns in such a staircase. $\square$

Example 5.4.10. If $\lambda \vdash n$ is a hook, the sequence of fiber sizes in Proposition 5.4.9 is

$$1 < 2 \geq 2 \geq 2 \cdots \geq 2 \geq 1 \geq \cdots \geq 1 \geq 0 \geq \cdots$$

where there are $N(\lambda)$ two's and $n - N(\lambda)$ non-zero entries. In particular, $N(\lambda)+1 \leq n-N(\lambda)$, i.e. $2N(\lambda)+1 \leq n$.

Proposition 5.4.11. *Let $\lambda \vdash n$ for $n \geq 1$. Set*

$$N \coloneqq \begin{cases} N(\lambda) & \text{if } 2N(\lambda)+1 \leq n \\ \left\lfloor \frac{n-1}{2} \right\rfloor & \text{if } 2N(\lambda)+1 > n. \end{cases} \tag{5.19}$$

Then

$$\lambda \lesssim^{\mathrm{diag}} (n-N, 1^N). \tag{5.20}$$

In particular, if $2N(\lambda)+1 \leq n$*, then the hook* $(n-N(\lambda), 1^{N(\lambda)})$ *is maximal for the diagonal preorder on partitions of size* n *with diagonal excess* $N(\lambda)$*.*

Proof. Using Proposition 5.4.9, the sequence

$$D(\lambda) \coloneqq \left(|\pi^{-1}(i)|\right)_{i\in\mathbb{P}}.$$

is of the form

$$D(\lambda) = (1, 2, \ldots, m, \ldots, 0, \ldots)$$

where the terms weakly decrease starting at m. Given a sequence $D = (D_1, D_2, \ldots) \in \mathbb{N}^{\mathbb{P}}$, define $N(D) \coloneqq \sum_{i:D_i\neq 0}(D_i - 1)$. We have $N(D(\lambda)) = N(\lambda)$. Iteratively perform the following procedure starting with $D \coloneqq D(\lambda)$ as many times as possible; see Example 5.4.12.

(i) If $2N(D)+1 > n$ and some $D_i > 2$, choose i maximal with this property. Decrease the ith entry of D by 1 and replace the first 0 term in D with 1.

(ii) If $2N(D)+1 \leq n$ and some $D_i > 2$, choose i maximal with this property. We will shortly show that there is some $j > i$ for which $D_j = 1$. Choose j minimal with this property, decrease the ith term in D by 1, and increment the jth term by 1.

Example 5.4.12. Suppose $\lambda = (4, 4, 4, 4)$, so $n = 16$ and

$$D(\lambda) = (1, 2, 3, 4, 3, 2, 1, 0, \ldots),$$

which we abbreviate as $D(\lambda) = 1234321$. Applying the procedure gives the following sequences, where modified entries are underlined:

D	$N(D)$	$2N(D)+1$
1234321	9	19
$1234\underline{2}21\underline{1}$	8	17
$123\underline{3}2211\underline{1}$	7	15
$123\underline{2}22\underline{2}11$	7	15
$12\underline{2}2222\underline{2}1$	7	15

Returning to the proof, for the claim in (ii), first note that both procedures preserve unimodality and the initial 1 in $D(\lambda)$. Hence at any intermediate step, D is of the form

$$(1, D_2, D_3, \ldots, D_k, 1, \ldots, 1, 0, \ldots)$$

where $D_2, \ldots, D_k \geq 2$ and there are $\ell \geq 0$ terminal 1's. Since $2N(D) + 1 \leq n$, we have

$$\begin{aligned} 2N(D) + 1 = 2(D_2 - 1 + \cdots + D_k - 1) + 1 \leq n = 1 + D_2 + \cdots + D_k + \ell \\ \Leftrightarrow (D_2 - 2) + \cdots + (D_k - 2) \leq \ell, \end{aligned}$$

forcing $\ell > 0$ since by assumption some $D_i > 2$, giving the claim. The procedure evidently terminates.

In applying (i), $N(D)$ decreases by 1, whereas $N(D)$ is constant in applying (ii). For the final sequence D_{fin}, it follows that $N(D_{\text{fin}}) = N$ from (5.19). Both (i) and (ii) strictly increase in the natural diagonal partial order on sequences. The final sequence will be

$$D_{\text{fin}} = (1, 2, 2, \ldots, 2, 1, 1, \ldots, 1, 0, \ldots)$$

where there are N two's and $n - N$ non-zero entries. This is precisely $D((n - N, 1^N))$ by Example 5.4.10, and the result follows. $\square$

We may now give a polynomial lower bound on f^λ.

Corollary 5.4.13. *Let $\lambda \vdash n$ for $n \geq 1$ and take N as in (5.19). For any $0 \leq M \leq N$, we have*

$$\prod_{c \in \lambda} h_c^{\mathrm{op}} \leq (n-M)!(M+1)!. \tag{5.21}$$

Moreover,

$$f^\lambda \geq \frac{1}{M+1}\binom{n}{M}. \tag{5.22}$$

Proof. Equation (5.21) in the case $M = N$ follows by combining (5.18) and (5.20). The general case follows similarly upon noting $(n-N, 1^N) \lesssim^{\mathrm{diag}} (n-M, 1^M)$ since $N \leq \left\lfloor \frac{n-1}{2} \right\rfloor$.

For (5.22), use Proposition 5.4.5 and (5.21) to compute

$$f^\lambda = \frac{n!}{\prod_{c\in\lambda} h_c} \geq \frac{n!}{\prod_{c\in\lambda} h_c^{\mathrm{op}}} \geq \frac{n!}{(n-M)!(M+1)!} = \frac{1}{M+1}\binom{n}{M}.$$

□

We now prove Theorem 5.1.3 and Theorem 5.1.7.

Proof of Theorem 5.1.3. We begin by summarizing the verification of Theorem 5.1.3 for $n \leq 33$. For $1 \leq n \leq 33$, a computer check shows that one may use Corollary 5.4.3 for all but 688 particular λ. However, the number of standard tableaux for these exceptional λ is small enough that the conclusion of the theorem may be quickly verified by computer. We now take $n \geq 34$.

Let N be as in (5.19). If $N \geq 5$, by Corollary 5.4.13,

$$f^\lambda \geq \frac{1}{6}\binom{n}{5} \geq n^3$$

for $n \geq 32$, so we may take $N \leq 4$. Since $\left\lfloor \frac{n-1}{2} \right\rfloor \geq 16 > 4 \geq N$, we must have $N = N(\lambda)$.

Write $\nu \oplus \mu$ to denote the concatenation of partitions ν and μ, where we assume the largest part of μ is no larger than the smallest part of ν. Using Proposition 5.4.9, since $n \geq 32$ and $N = N(\lambda) \leq 4$, we find that either $\lambda = (n-N) \oplus \mu$ or $\lambda' = (n-N) \oplus \mu$ for $|\mu| = N$.

To cut down on duplicate work, note that transposing $T \in \mathrm{SYT}(\lambda)$ complements the descent set of T. It follows that $b_{\lambda,i} = b_{\lambda',\binom{n}{2}-i}$, so that $a_{\lambda,r} = a_{\lambda',\binom{n}{2}-r}$. Since the statement of Theorem 5.1.3 also exhibits this symmetry, we may thus consider only the case when $\lambda = (n-N) \oplus \mu$.

There are twelve μ with $|\mu| \leq 4$. One may check that the five possible μ for $N = 4$ all result in $f^\lambda \geq n^3$ for $n \geq 34$, leaving seven remaining μ, namely

$$\mu = \varnothing, (1), (2), (1,1), (3), (2,1), (1,1,1).$$

It is straightforward (though tedious) to verify the conclusion of Theorem 5.1.3 in each of these cases. For instance, for $\mu = (1)$ and $\lambda = (n-1, 1)$, there are $n-1$ standard tableaux with major indexes $1, \ldots, n-1$ (alternatively, (5.5) results in $q[n-1]_q$). The remaining cases are omitted. $\square$

Proof of Theorem 5.1.7. If $f^\lambda \geq n^5$, then (5.14) gives

$$\ln \frac{|\chi^\lambda(\ell^s)|}{f^\lambda} \leq \left(1 - \frac{1}{\ell}\right)\left[-\frac{9}{2}\ln n + \ln\sqrt{2\pi}\right] + \frac{\ell}{12n} - \frac{1}{2}\ln \ell \tag{5.23}$$

As before one can check that the right-hand side of (5.23) is less than $\ln \frac{1}{\phi(\ell)n^2}$ for $\ell = 2, 3$ and $n \geq 3$. When $\ell \geq 4$, term-by-term estimates give

$$\ln \frac{|\chi^\lambda(\ell^s)|}{f^\lambda} \leq -\frac{9}{2}\left(1 - \frac{1}{4}\right)\ln n + \ln\sqrt{2\pi} + \frac{1}{12} - \frac{1}{2}\ln 4$$

which is less than $\ln \frac{1}{n^3}$ for $n \geq 3$. The first part of Theorem 5.1.7 now follows from Lemma 5.4.2 with $d = 2$ for $n \geq 3$. It remains true for $n = 1, 2$.

For the second part, suppose $n \geq 81$, $\lambda_1 < n - 7$, and $\lambda_1' < n - 7$. It follows from Proposition 5.4.11 that N from (5.19) satisfies $N \geq 8$. Hence by Corollary 5.4.13 we have

$$f^\lambda \geq \frac{1}{9}\binom{n}{8} \geq n^5.$$

□

5.5 Alternative Proof of the Hook Formula

The proof of Theorem 5.1.8 in [27] and [45] uses a certain decomposition of the r-rim hook partition lattice and the original hook length formula. We present an alternative proof following a different tradition, instead generalizing the approach to the original hook length formula in [91, Corollary 7.21.6]. A by-product of our proof is a particularly explicit description of the movement of hook lengths mod ℓ as length ℓ ribbons are added to a partition shape.

We are not at present aware of any other proofs or direct uses of Theorem 5.1.8, and it seems to have been neglected by the literature. Indeed, the author empirically rediscovered it and found the following proof before unearthing [27].

Proof of Theorem 5.1.8. Let $\lambda \vdash n$, $n = \ell s$. If λ cannot be written as s successive ribbons of length ℓ, then by the classical Murnaghan-Nakayama rule [91, Eq. (7.75)] we have $\chi^\lambda(\ell^s) = 0$, so assume λ can be so written.

Combining (5.4), (5.5), and (5.7) shows that we may compute $\chi^\lambda(\ell^s)$ by letting $q \to \omega_n^s$ in the right-hand side of (5.5). We may replace each q-number $[a]_q$ with $q^a - 1$ by canceling the $q-1$'s, since $\lambda \vdash n$. Since ω_n^s has order ℓ, the values of $q^a - 1$ at ω_n^s depend only on a mod ℓ. Moreover, $q^a - 1$ has only simple roots, and it has a root at ω_n^s if and only if $\ell \mid a$. The order of vanishing of the numerator at $q = \omega_n^s$ is then $\#\{i \in [n] : i \equiv_\ell 0\} = s$, and the order of vanishing of the denominator is $\#\{c \in \lambda : h_c \equiv_\ell 0\}$. The following lemma ensures these counts agree. We postpone the proof to the end of this section.

Lemma 5.5.1. *Let $\lambda \vdash n$, $n = \ell s$, and suppose λ can be written as a sequence of s successive ribbons of length ℓ. Then for any $a \in \mathbb{Z}$,*

$$\#\{c \in \lambda : h_c \equiv_\ell \pm a\} = s \cdot \#\{a, -a \ (mod\ \ell)\}.$$

Here $\#\{a, -a \pmod{\ell}\}$ is 1 if $a \equiv_\ell -a$ and 2 otherwise.

We may now compute the desired $q \to \omega_n^s$ limit by repeated applications of L'Hopital's rule. In particular, we find

$$|\chi^\lambda(\ell^s)| = \left| \lim_{q \to \omega_n^s} q^{b(\lambda)} \frac{\prod_{i \in [n]} [i]_q}{\prod_{c \in \lambda} [h_c]_q} \right| = \left| \lim_{q \to \omega_n^s} \frac{\prod\limits_{\substack{i \in [n] \\ i \not\equiv_\ell 0}} q^i - 1}{\prod\limits_{\substack{c \in \lambda \\ h_c \not\equiv_\ell 0}} q^{h_c} - 1} \right| \left| \frac{\prod\limits_{\substack{i \in [n] \\ i \equiv_\ell 0}} i\omega_n^{s(i-1)}}{\prod\limits_{\substack{c \in \lambda \\ h_c \equiv_\ell 0}} h_c \omega_n^{s(h_c - 1)}} \right| \tag{5.24}$$

The second factor in the right-hand side of (5.24) equals the right-hand side of (5.2), so we must show the first factor in the right-hand side of (5.24) is 1. For that, note that $q^a - 1$ at $q = \omega_n^s$ for $a \not\equiv_\ell 0$ is non-zero and is conjugate to $q^{-a} - 1$ at $q = \omega_n^s$. By Lemma 5.5.1, it follows that the contribution to the overall magnitude due to $\{c \in \lambda : h_c \equiv_\ell a \text{ or } -a\}$ cancels with the contribution due to $\{i \in [n] : i \equiv_\ell a \text{ or } -a\}$ for each $a \not\equiv_\ell 0$. This completes the proof of the theorem.

□

As for Lemma 5.5.1, it is an immediate consequence of the following somewhat more general result.

Lemma 5.5.2. *Suppose λ/μ is a ribbon of length ℓ. For any $a \in \mathbb{Z}$,*

$$\#\{c \in \mu : h_c \equiv_\ell \pm a\} + \#\{a, -a \ (mod\ \ell)\} = \#\{d \in \lambda : h_d \equiv_\ell \pm a\}.$$

Proof. We determine how the counts $\#\{c \in \mu : h_c \equiv_\ell \pm a\}$ change when adding a ribbon of length ℓ; see Figure 5.2. We define the following regions in λ, relying on French notation to determine the meaning of "leftmost," etc.

(I) Cells $c \in \mu$ where c is not in the same row or column as any element of λ/μ.

(II) Cells $c \in \mu$ which are in the same row as some element of λ/μ and are strictly left of the leftmost cell in λ/μ.

(III) Cells $c \in \mu$ which are in the same column as some element of λ/μ and are strictly below the bottommost cell of λ/μ.

(IV) Cells $c \in \lambda$ which are in both the same column and row as some element(s) of λ/μ. Region (IV) includes the ribbon λ/μ itself.

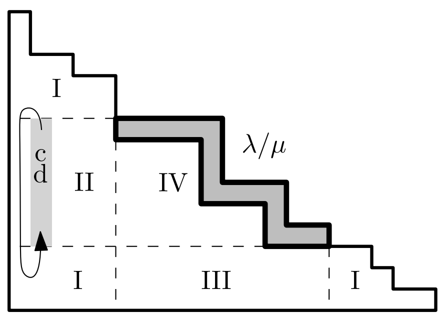


Figure 5.2: All regions of a partition λ where λ/μ is a ribbon

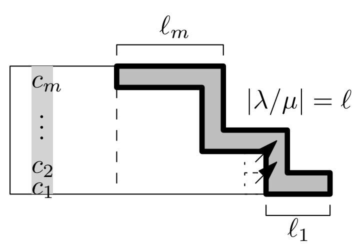


Figure 5.3: Regions (II) and (IV) up close

We now describe how hook lengths change in each region, mod the ribbon length ℓ, in going from μ to λ. They are unchanged in region (I). Regions (II) and (III) are similar, so we consider region (II). This region is a rectangle, which we imagine breaking up into columns. Write h_c^λ or h_c^μ to denote the hook length of a cell $c \in \mu$ as an element of λ or μ, respectively. For c in region (II), let d denote the cell in region (II) immediately below c, with wrap-around. We claim $h_c^\lambda \equiv_\ell h_d^\mu$. Given the claim, hook lengths mod ℓ in regions (II) and (III) are simply permuted in going from μ to λ, so changes to the counts $\#\{c \in \mu : h^\mu_c \equiv_\ell \pm a\}$ arise only from region (IV).

For the claim, let $c_1, c_2, \ldots, c_m$ be the cells of the column in region (II) containing c, listed from bottom to top; see Figure 5.3. Begin by comparing hook lengths at c_1 and c_2. Since $\lambda - \mu$ is a ribbon, the rightmost cell of μ in the same row as c_1 is directly left and below the rightmost cell of λ in the same row as c_2. It follows that $h^\mu_{c_1} = h^\lambda_{c_2}$. This procedure yields the claim except when $c = c_1$. In that case, $d = c_m$, and we further claim $h^\lambda_{c_1} = h^\mu_{c_m} + \ell$, which will finish the argument. Indeed, let ℓ_i denote the number of elements in $\lambda - \mu$ in the same row as c_i. Certainly $\ell = \ell_1 + \cdots + \ell_m$. Further, $h^\lambda_{c_i} = h^\mu_{c_i} + \ell_i$. Putting it all together, we have

$$\begin{aligned} h^\lambda_{c_1} &= h^\mu_{c_1} + \ell_1 = h^\lambda_{c_2} + \ell_1 \\ &= h^\mu_{c_2} + \ell_2 + \ell_1 = \cdots \\ &= h^\mu_{c_m} + \ell_m + \cdots + \ell_2 + \ell_1 = h^\mu_{c_m} + \ell. \end{aligned}$$

We now turn to region (IV). It suffices to consider the case depicted in Figure 5.4, where regions (I), (II), and (III) are empty. We define two more regions as follows; see Figure 5.4.

(A) Cells $c \in \lambda$ in the first row or column.

(B) Cells $c \in \lambda$ not in the first row or column.

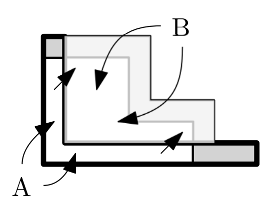


Figure 5.4: Regions (A) and (B) of a partition μ where λ/μ is a ribbon

Region (B) is precisely μ translated up and right one square. Moreover, this operation preserves hook lengths, so changes in the counts $\#\{c \in \mu : h^\mu_c \equiv_\ell \pm a\}$ arise entirely from

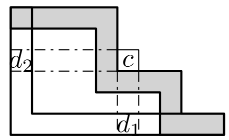


Figure 5.5: Adding a cell to region (B)

region (A). We have thus reduced the lemma to the statement

$$\#\{c \text{ in region (A) } : h_c^\lambda \equiv_\ell \pm a\} = \#\{a, -a \pmod{\ell}\}. \tag{5.25}$$

We prove (5.25) by induction on the size of region (B). In the base case, region (B) is empty, so λ is a hook, and the result is easy to see directly (for instance, negate the hook lengths in only the "vertical leg" to get entries of precisely $1, 2, \ldots, \ell$). For the inductive step, consider the effect of adding a cell c to region (B). Now c is in the same column as some cell d_1 in region (A) and c is in the same row as some cell d_2 in region (A); see Figure 5.5. Say the original hook length of d_1 is i and the original hook length of d_2 is j. It is easy to see that $i + j = \ell - 1$. Adding c to region (B) increases the hook lengths i and j each by 1, but $j + 1 \equiv_\ell -i$ and $i + 1 \equiv_\ell -j$, so the required counts remain as claimed in the inductive step. This completes the proof of the lemma and, hence, Theorem 5.1.8. □

We briefly contrast our approach with that of [27]. Let f_ℓ^λ be the number of ways to write λ as successive ribbons each of length ℓ. If $\lambda \vdash n = \ell s$, by the Murnaghan-Nakayama rule $\chi^\lambda(\ell^s)$ is a signed sum over terms counted by f_ℓ^λ. While there is typically cancellation in this sum, there is in fact none for rectangular cycle types [45, 2.7.26], i.e. $\chi^\lambda(\ell^s) = \pm f_\ell^\lambda$. Indeed, [27] proved Theorem 5.1.8 using standard rim hook tableaux instead of character evaluations, though virtually every application of their result uses the character-theoretic inequality in Theorem 5.1.5.

The sign of $\chi^\lambda(\ell^s)$ can be computed in terms of *abaci* as in [45, 2.7.23]. The sign may also be computed "greedily" by repeatedly removing ℓ-rim hooks from λ in any order whatsoever, which is a consequence of (among other things) the following corollary of Lemma 5.5.2 and Theorem 5.1.8. We have been unable to find part (iv) in the literature, though for the rest see [27, 2.5-2.7] and their references.

Corollary 5.5.3. *Let $\lambda \vdash n = \ell s$. The following are equivalent:*

(i) $\chi^\lambda(\ell^s) \neq 0$*;*

(ii) λ *can be written as successive length* ℓ *ribbons, i.e. the* ℓ*-core of* λ *is empty;*

(iii) we have

$$\#\{c \in \lambda : h_c \equiv_\ell 0\} = s;$$

(iv) for any $a \in \mathbb{Z}$*,*

$$\#\{c \in \lambda : h_c \equiv_\ell \pm a\} = s \cdot \#\{a, -a \ (\mathit{mod}\ \ell)\}.$$

Proof. (i) and (ii) are equivalent by Theorem 5.1.8. (ii) implies (iv) by Lemma 5.5.1 and (iv) implies (iii) trivially. Finally, (iii) is equivalent to (i) as follows. The expression (5.5) is a polynomial, so the order of vanishing at $q \to \omega_n^s$ of the numerator, namely s, is at most as large as the order of vanishing of the denominator, namely $\#\{c \in \lambda : h_c \equiv_\ell 0\}$. The limiting ratio is non-zero if and only if these counts agree, so (iii) is equivalent to (i). $\square$

While Corollary 5.5.3 gives equivalent conditions for $\chi^\lambda(\ell^s) \neq 0$, [89, Corollary 7.5] gives interesting and different necessary conditions for $\chi^\lambda(\nu) \neq 0$ for general shapes ν.

5.6 Unimodality and $\chi^\lambda(\mu)$

We end with a brief discussion of inequalities related to symmetric group characters. In applying Proposition 5.4.5, we essentially replaced $\frac{n!}{\prod_{c\in\lambda} h_c}$ with $\frac{n!}{\prod_{c\in\lambda} h_c^{\mathrm{op}}}$, since the latter is order-reversing with respect to the diagonal preorder by (5.18). Moreover, it is relatively straightforward to mutate partitions and predictably increase or decrease them in the diagonal preorder, as in the proof of Proposition 5.4.11. It would be desirable to instead work directly with symmetric group characters themselves and appeal to general results about how $|\chi^\lambda(\mu)|$ increases or decreases as λ is mutated and μ is held fixed, though we have found very few concrete and no conjectural results in this direction. Any progress seems both highly non-trivial and potentially useful, so in this section we record some initial observations.

We have $\chi^{(a+1,1^b)}(1^n) = \binom{n-1}{a}$ for $a+b+1=n$, so these values are unimodal in a. Using Theorem 5.1.8 shows more generally that for all $\ell \mid n$,

$$|\chi^{(a+1,1^b)}(\ell^{n/\ell})| = \binom{\frac{n}{\ell}-1}{\lfloor \frac{a}{\ell} \rfloor}$$

which is again unimodal in a. However, $|\chi^\lambda(\ell^s)|$ does not seem to respect changes in λ under dominance order in general in any suitable sense. On the other hand, if we allow the cycle type μ to vary and consider the Kostka numbers $K_{\lambda\mu}$ as a surrogate for $|\chi^\lambda(\mu)|$ (since $K_{\lambda(1^n)} = \chi^\lambda(1^n)$), we have a series of well-known and very general inequalities. We write $K_{\lambda\mu}(t)$ for the Kostka-Foulkes polynomial and $\nu \geq \mu$ for dominance order. We have:

Theorem 5.6.1 ([83], [58], [55]; [35])**.** *$K_{\lambda\nu} \leq K_{\lambda\mu}$ for all λ if and only if $\nu \geq \mu$. Indeed, $\nu \geq \mu$ implies $K_{\lambda\nu}(t) \leq K_{\lambda\mu}(t)$ (coefficient-wise) for all λ.*

Question 5.6.2. *Are there any "nice" infinite families besides hooks and rectangles for which $|\chi^\lambda(\mu)|$ is monotonic, unimodal, or suitably order-preserving as λ varies? What about as μ varies?*

Chapter 6

DISTRIBUTION OF MAJOR INDEX FOR STANDARD TABLEAUX AND ASYMPTOTIC NORMALITY

This chapter is joint work with Sara Billey and Matjaž Konvalinka. A version of it will be submitted for publication shortly [10].

6.1 Main Results

In this chapter, we study the distribution of the major index statistic generalized to standard Young tableaux of straight and skew shapes. The properties we discuss here naturally generalize known properties of the major index distribution on permutations. They have representation theory consequences in terms of the coinvariant algebras of symmetric groups. We will briefly introduce the main results. See Section 6.2 for more details on the background.

Here we are primarily interested in the major index generating function

$$\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) := \sum_{T \in \mathrm{SYT}(\lambda)} q^{\mathrm{maj}(T)} = \sum_{k \geq 0} b_{\lambda,k} q^k.$$

The polynomial $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$ has two elegant closed forms described in Corollary 1.2.1.

Let $X_\lambda[\mathrm{maj}]$ be the discrete random variable given by the maj statistic on $\mathrm{SYT}(\lambda)$ taken uniformly at random. Thus, $\mathbb{P}[X_\lambda[\mathrm{maj}] = k] = b_{\lambda,k}/\#\,\mathrm{SYT}(\lambda)$ where $b_{\lambda,k} = \#\{T \in \mathrm{SYT}(\lambda) : \mathrm{maj}(T) = k\}$. Using work of [16, 44] along with Stanley's q-analog of the hook length formula, we give exact formulas for the dth moment μ_d^λ, the dth central moment α_d^λ, and the dth cumulant κ_d^λ for $X_\lambda[\mathrm{maj}]$. The most elegant of the formulas is for the cumulants, from which the moments and central moments are all easy to compute.

Theorem 6.1.1. *Let $\lambda \vdash n$ and $d \in \mathbb{Z}_{>1}$. We have*

$$\kappa_d^\lambda = \frac{B_d}{d}\left[\sum_{j=1}^{n} j^d - \sum_{c\in\lambda} h_c^d\right] \tag{6.1}$$

where $B_0, B_1, B_2, \ldots = 1, \frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \frac{1}{42}, 0, \ldots$ are the Bernoulli numbers.

Theorem 6.1.1 generalizes the formula for the variance of $X_\lambda[\text{maj}]$ given by Adin and Roichman [2]. A similarly explicit expression holds for the mean.

With precise information about the moments and cumulants of the maj distribution on $\text{SYT}(\lambda)$, we use the method of moments to show that in a very general limiting process on partition shapes, the random variables $X_\lambda[\text{maj}]$ are asymptotically normal.

Definition 6.1.2. Given any partition $\lambda \vdash n$, the *aft* of λ is

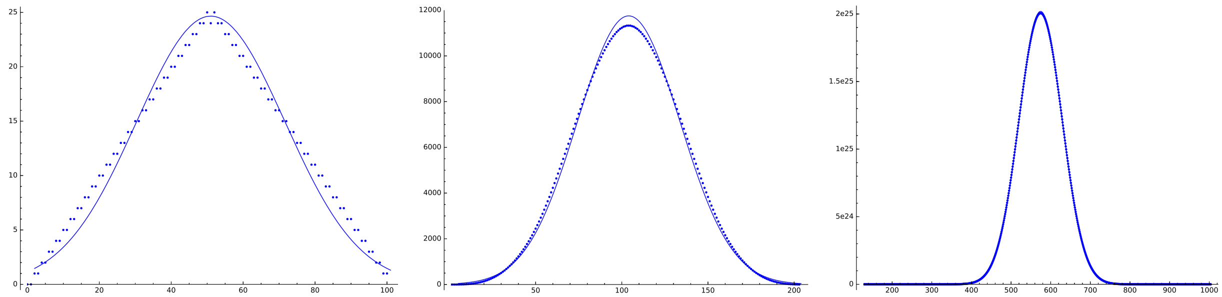
$$\text{aft}(\lambda) := n - \max\{\lambda_1, \lambda_1'\}.$$

That is, the aft of a partition whose first row is at least as long as its first column is the number of cells not in the first row. The following is our first main result in this chapter.

Theorem 6.1.3. *Suppose $\lambda^{(1)}, \lambda^{(2)}, \ldots$ is a sequence of partitions, and let $X_N := X_{\lambda^{(N)}}[\text{maj}]$ be the corresponding random variables for the* maj *statistic. Then, the sequence $X_1, X_2, \ldots$ is asymptotically normal if and only if* $\text{aft}(\lambda^{(N)}) \to \infty$ *as $N \to \infty$.*

Example 6.1.4. See Figure 6.1a and Figure 6.1b for distributions with small aft. In these cases the normal approximation differs visibly from the major index distribution. A more typical example is Figure 6.1c, where the aft is relatively large and the normal approximation is very close. When $\lambda^{(N)} := (N, N)$, Theorem 6.1.3 recovers the main result of [16], namely that q-Catalan coefficients are asymptotically normal.

Remark 6.1.5. Observe that $X_\lambda[\text{maj}]$ can be written as the sum of scaled indicator random variables $D_1, 2D_2, \ldots, (n-1)D_{n-1}$ corresponding with possible descent positions $1, 2, \ldots, n-1$



(a) $\lambda = (50, 2)$, $\mathrm{aft}(\lambda) = 2$ (b) $\lambda = (50, 3, 1)$, $\mathrm{aft}(\lambda) = 4$ (c) $\lambda = (8, 8, 7, 6, 5, 5, 5, 5, 2, 2)$, $\mathrm{aft}(\lambda) = 39$

Figure 6.1: Plots of $\#\{T \in \mathrm{SYT}(\lambda) : \mathrm{maj}(T) = k\}$ as a function of k for three partitions λ, overlaid with scaled Gaussian approximations using the same mean and variance.

respectively. While the indicator random variables D_i are equidistributed [91, Prop. 7.19.9], they are not independent. For example, if $T \in \mathrm{SYT}(\lambda)$ does not have descents in positions $1, 2, \ldots, \lambda_1 - 1$, then λ_1 must be a descent for T if λ is not a one row shape. Consequently, Theorem 6.1.3 does not follow from a standard application of a generalized central limit theorem, and in fact there are non-normal limiting distributions. The lack of independence of the D_i's likewise complicates related work by Fulman [30] and Kim–Lee [49] considering the limiting distribution of descents in certain classes of permutations.

We classify all possible limiting distributions for arbitrary sequences of partitions as follows. Given a real-valued random variable X with mean μ and variance $\sigma^2 > 0$, let

$$X^* := \frac{X - \mu}{\sigma}.$$

Let Σ_M denote the sum of M independent identically distributed uniform $[0, 1]$ random variables, known as the Irwin–Hall distribution or the *uniform sum distribution*.

Theorem 6.1.6. *Suppose $\lambda^{(1)}, \lambda^{(2)}, \ldots$ is a sequence of partitions such that $|\lambda^{(N)}| \to \infty$ and $\mathrm{aft}(\lambda^{(N)}) = M$ is constant. Let $X_N := X_{\lambda^{(N)}}[\mathrm{maj}]$. Then $X_1^*, X_2^*, \ldots$ converges in distribution to Σ_M^*.*

Combining Theorem 6.1.3 and Theorem 6.1.6 gives the following classification.

Theorem 6.1.7. *Let $\lambda^{(1)}, \lambda^{(2)}, \ldots$ be a sequence of partitions. Then $(X_{\lambda^{(N)}}[\mathrm{maj}]^*)$ converges in distribution if and only if*

(i) $\mathrm{aft}(\lambda^{(N)}) \to \infty$*; or*

(ii) $|\lambda^{(N)}| \to \infty$ *and* $\mathrm{aft}(\lambda^{(N)})$ *is eventually constant; or*

(iii) the distribution of $X^*_{\lambda^{(N)}}[\mathrm{maj}]$ *is eventually constant.*

The limit law is $\mathcal{N}(0,1)$ in case (i), Σ^*_M in case (ii), and discrete in case (iii).

Case (iii) naturally leads to the question, when does $X^*_\lambda[\mathrm{maj}] = X^*_\mu[\mathrm{maj}]$? Such a description in terms of hook lengths is given in Theorem 6.5.2.

Example 6.1.8. We illustrate each possible limit. For (i), let $\lambda^{(N)} := (N, \lfloor \ln N \rfloor)$, so that $\mathrm{aft}(\lambda^{(N)}) = \lfloor \ln N \rfloor \to \infty$ and the distributions are asymptotically normal. For (ii), fix $M \in \mathbb{Z}_{\geq 0}$ and let $\lambda^{(N)} := (N+M, M)$, so that $\mathrm{aft}(\lambda^{(N)}) = M$ is constant and the distributions converge to Σ^*_M. For (iii), let $\lambda^{(2N)} := (12, 12, 3, 3, 3, 2, 2, 1, 1)$ and $\lambda^{(2N+1)} := (15, 6, 6, 6, 4, 2)$, which have the same multisets of hook lengths despite not being transposes of each other, and consequently the same normalized maj distributions.

One motivation for the present work came from earlier work concerning the distribution of maj, or equivalently inv, on words. See [14] for further references to the probability literature including work of Diaconis, Kendall, Mann–Whitney, and others. We are able to simultaneously consider maj on words and tableaux by generalizing the preceding asymptotic results to certain skew shapes $\underline{\lambda}$. In particular we recover and refine Canfield–Janson–Zeilberger's main result in [13], quoted as Theorem 6.2.23 below.

Another motivation for the present work was consideration of the sequences

$$b_{\lambda,k} := \#\{T \in \mathrm{SYT}(\lambda) : \mathrm{maj}(T) = k\} \tag{6.2}$$

for fixed λ. These sequences have appeared in a variety of algebraic and representation-theoretic contexts, including branching rules between symmetric groups and cyclic subgroups [94], the irreducible decomposition of type A coinvariant algebras [87] (and Lusztig, unpublished), and degree polynomials of unipotent $\mathrm{GL}_n(\mathbb{F}_q)$-representations [39].

There are polynomial expressions for $b_{\lambda,k}$ in terms of H_i's, the number of cells of λ with hook-length equal to i. See Remark 6.2.6.

We consider three natural enumerative questions:

(I) which terms in (6.2) are zero?

(II) are the sequences in (6.2) unimodal?

(III) are there efficient asymptotic estimates for $b_{\lambda,k}$?

We completely settle (I) with the following result. Let $b(\lambda) := \sum_{i\geq 1}(i-1)\lambda_i$.

Theorem 6.1.9. *For every partition $\lambda \vdash n > 1$ and integer k such that $b(\lambda) \leq k \leq \binom{n}{2} - b(\lambda')$, we have $b_{\lambda,k} > 0$ except in the case when λ is a rectangle with at least 2 rows and columns and k is either $b(\lambda)+1$ or $\binom{n}{2} - b(\lambda') - 1$. We have $b_{\lambda,k} = 0$ for $k < b(\lambda)$ or $k > \binom{n}{2} - b(\lambda')$.*

As a consequence of the proof of Theorem 6.1.9, we identify a ranked poset structure on $\mathrm{SYT}(\lambda)$ where the rank function is determined by maj. Furthermore, as a corollary of Theorem 6.1.9 we have a new proof of the complete classification Theorem 5.1.3 generalizing an earlier result of Klyachko [50] for when the counts

$$a_{\lambda,r} := \{T \in \mathrm{SYT}(\lambda) : \mathrm{maj}(T) \equiv_n r\}$$

for $\lambda \vdash n$ are nonzero.

We give conjectured answers to question (II) in Section 6.7. We hope that a variation on the map used to prove Theorem 6.1.9 can be used to prove our unimodality conjecture, Conjecture 6.7.1, by constructing explicit injections where possible.

By Theorem 6.1.3 and the conjectured claim that the coefficients of $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$ are unimodal or almost unimodal for large λ, one might hope that we could approximate the number of $T \in \mathrm{SYT}(\lambda)$ with $\mathrm{maj}(T) = k$ by the density function $f(k; \kappa_1^\lambda, \kappa_2^\lambda)$ for the normal distribution with mean κ_1^λ and variance κ_2^λ. We have the following conjectured bounds on such an approximation.

Conjecture 6.1.10. *Let $\lambda \vdash n$ be any partition. Uniformly for all n, for all integers k, we have*

$$\left|\mathbb{P}[X_\lambda = k] - f(k; \kappa_1^\lambda, \kappa_2^\lambda)\right| = O\left(\frac{1}{\sigma_\lambda \, \mathrm{aft}(\lambda)}\right).$$

The conjecture has been verified for $25 < n \leq 50$ and $\mathrm{aft}(\lambda) > 1$ with a constant of $1/9$, which is tight up to reasonable limits on computation in the sense that if it is changed to $1/10$ with the other constraints the same, it fails at $n = 50$.

The rest of the chapter is organized as follows. In Section 6.2, we give background on tableaux combinatorics, combinatorial and probabilistic generating functions, and asymptotic normality. The proof of Theorem 6.1.1 follows immediately from the background material and is summarized in Remark 6.2.18. In Section 6.3, we give cumulant estimates which prove the generalization of Theorem 6.1.3 to special "block" skew shapes $\underline{\lambda}$, see Theorem 6.3.8. Section 6.4 gives similar estimates for a generalization of Theorem 6.1.6 to $\underline{\lambda}$, see Theorem 6.4.2. Section 6.5 proves the generalization of Theorem 6.1.7 to $\underline{\lambda}$, Theorem 6.5.1, and further analyzes case (iii) of Theorem 6.1.7, resulting in Theorem 6.5.2. Section 6.6 presents our combinatorial argument proving Theorem 6.1.9 and giving poset structures to sets of tableaux. Section 6.7 presents conjectures characterizing unimodality, log-concavity, and related properties of the sequences $(b_{\lambda,k})_{k\in\mathbb{Z}}$.

6.2 Combinatorial and Probabilistic Generating Functions

In this section, we review some standard terminology and results on combinatorial statistics, random variables, and asymptotic normality. An excellent source for many further details in this area can be found in [12]. See also Chapter 2.

6.2.1 Word and Tableaux Combinatorics

In addition to the notions in Section 2.2 on partitions and tableaux, we will use the following. Index the cells of a tableaux by matrix notation when we refer to their row and column. An *outer corner* of λ is any cell with hook length 1. An *inner corner* of λ is any (i, j) not in λ such that both $(i-1, j)$ and $(i, j-1)$ are both in λ. A *bijective filling* of λ is any labeling of the cells of λ by the numbers $[n] = \{1, 2, \ldots, n\}$. The symmetric group S_n acts on bijective fillings of λ by acting on the labels.

Definition 6.2.1. A *skew partition* λ/ν is a pair of partitions (ν, λ) such that the Young diagram of ν is contained in the Young diagram of λ. The cells of λ/ν are the cells in the diagram of λ which are not in the diagram of ν, written $c \in \lambda/\nu$. We identify straight partitions λ with skew partitions $\lambda/\varnothing$ where $\varnothing = (0, 0, \ldots)$ is the empty partition. The notions of bijective filling, hook lengths, inner and outer corners, standard tableaux, descent set, and major index extend verbatim to skew partitions as well.

$$\begin{array}{|c|c|c|c|c|c|} \hline 1 & 2 & 4 & 7 & 9 & 12 \\ \hline 3 & 6 & 10 \\ \cline{1-3} 5 & 8 & 11 \\ \cline{1-3} \end{array} \qquad \begin{array}{ccccccc} & & & & & \boxed{2} & \boxed{6} \\ & & & \boxed{4} & \boxed{5} & & \\ \boxed{1} & \boxed{3} & \boxed{7} & & & & \end{array}$$

Figure 6.2: On the left is a standard Young tableau of straight shape $\lambda = (6, 3, 3)$ with descent set $\{2, 4, 7, 9, 10\}$ and major index 32. On the right is a standard Young tableau of skew shape $(7, 5, 3)/(5, 3)$ corresponding to sequence of partitions $(3), (2), (2)$ with descent set $\{2, 6\}$ and major index 8.

Definition 6.2.2. Given a sequence of partitions $\underline{\lambda} = (\lambda^{(1)}, \ldots, \lambda^{(m)})$, we identify the sequence with the *block diagonal skew partition* obtained by translating the Young diagrams of the $\lambda^{(i)}$ so that the rows and columns occupied by these components are disjoint; see Figure 6.3. The actual translations used will prove to be unimportant for our purposes, though for concreteness we use the ones depicted in the figure.

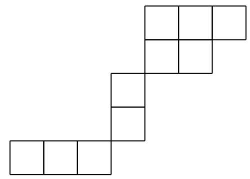


Figure 6.3: Diagram for $\underline{\lambda} = ((3), (1, 1), (3, 2))$.

The skew partitions $\underline{\lambda}$ allow us to simultaneously consider words and partitions as follows. Again let W_α be the set of all words with content α. We have a bijection

$$\phi \colon \mathrm{SYT}(((\alpha_1), (\alpha_2), \ldots)) \xrightarrow{\sim} \mathrm{W}_\alpha$$

which sends a tableau T to the word whose ith letter is the row in which i appears in T counting from the bottom up. For example, using the skew tableau T on the right of Figure 6.2, we have $\phi(T) = 1312231 \in \mathrm{W}_{(3,2,2)}$. It is easy to see that $\mathrm{Des}(\phi(T)) = \mathrm{Des}(T)$, so that $\mathrm{maj}(\phi(T)) = \mathrm{maj}(T)$.

6.2.2 Major Index Generating Functions

We next summarize some facts related to major index generating functions on words and tableaux.

Definition 6.2.3. A polynomial $P(q) = \sum_{i=0}^n c_i q^i$ is *symmetric* if $c_i = c_{n-i}$ for $0 \leq i \leq n$. We generally say $P(q)$ is *symmetric* also if there exists a integer k such that $q^k P(q)$ is symmetric. We say $P(q)$ is *unimodal* if

$$c_0 \leq c_1 \leq \cdots \leq c_j \geq c_{j+1} \geq \cdots \geq c_n$$

for some $0 \leq j \leq n$. Furthermore, $P(q)$ is *log-concave* if $c_i^2 \geq c_{i-1}c_{i+1}$ for all integers $0 < i < n$.

From Theorem 1.1.1, we see immediately that the coefficients of $\mathrm{W}_{\alpha}^{\mathrm{maj}}(q)$ are symmetric and that the leading coefficient is 1. Indeed, these polynomials are *unimodal*, generalizing the well-known case for Gaussian coefficients, [88, Thm 3.1] and [104]. The analogous expression for $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$, Theorem 5.2.2, was given by Stanley. It generalizes the famous Frame–Robinson–Thrall Hook Length Formula obtained by setting $q = 1$.

Example 6.2.4. For $\lambda = (4, 2)$, $b(\lambda) = 2$ and the multiset of hook lengths is $\{1^2, 2^2, 4, 5\}$ so $|\,\mathrm{SYT}(\lambda)| = 9$ by the Hook Length Formula. The major index generating function is given by

$$\begin{aligned}\mathrm{SYT}(4,2)^{\mathrm{maj}}(q) &= q^8 + q^7 + 2q^6 + q^5 + 2q^4 + q^3 + q^2 \\ &= q^2 \frac{[6]_q!}{[5]_q[4]_q[2]_q[2]_q} = q^2 \frac{[6]_q[3]_q}{[2]_q}.\end{aligned}$$

Note, $\mathrm{SYT}(4,2)^{\mathrm{maj}}(q)$ is symmetric but not unimodal.

For $\lambda = (4, 2, 1)$, $b(\lambda) = 4$ and the multiset of hook lengths is $\{1^3, 2, 3, 4, 6\}$ so $|\,\mathrm{SYT}(\lambda)| = 35$ by the Hook Length Formula. The major index generating function is given by

$$\begin{aligned}\mathrm{SYT}(4,2,1)^{\mathrm{maj}}(q) &= q^{14} + 2q^{13} + 3q^{12} + 4q^{11} + 5q^{10} + 5q^9 + 5q^8 + 4q^7 + 3q^6 + 2q^5 + q^4 \\ &= q^4 \frac{[7]_q!}{[6]_q[4]_q[3]_q[2]_q} = q^4 [7]_q [5]_q.\end{aligned}$$

Note, $\mathrm{SYT}(4,2,1)^{\mathrm{maj}}(q)$ is symmetric and unimodal.

Example 6.2.5. We recover q-integers, q-binomials, and q-Catalan numbers, up to q-shifts as special cases of the major index generating function for tableaux as follows:

$$\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = \begin{cases} q[n]_q & \text{if } \lambda = (n, 1), \\ q^{\binom{k+1}{2}} \binom{n}{k}_q & \text{if } \lambda = (n-k+1, 1^k), \\ q^n \frac{1}{[n+1]_q} \binom{2n}{n}_q & \text{if } \lambda = (n, n). \end{cases}$$

Remark 6.2.6. By Theorem 5.2.2, we have

$$q^{-b(\lambda)}\,\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = \frac{[n]_q!}{\prod_{c\in\lambda}[h_c]_q} = \prod_{i=1}^{n}(1-q^i)^{-(H_i-1)}, \tag{6.3}$$

where $H_i = \#\{c \in \lambda\colon h_c = i\}$. So (6.3) is equivalent to the polynomial formula for $b_{\lambda,k}$ for $k = b(\lambda) + d$ given by

$$[q^{b(\lambda)+d}]\,\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = \sum_{\substack{\tau\vdash d\\ \tau_1\le|\lambda|}}\prod_{i=1}^{|\lambda|}\binom{H_i+m_i(\tau)-2}{m_i(\tau)}, \tag{6.4}$$

where $m_i(\tau)$ is the number of parts of the partition τ equal to i. Note that if $H_i = 0$ and $m_i(\tau) = 1$, then the binomial coefficient is -1, so it is not obvious from (6.4) that the coefficients are nonnegative. The first few polynomials are given by

$$\begin{aligned}
[q^{b(\lambda)+1}]\,\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) &= H_1 - 1\\
&= \#\{c\in\lambda : c \text{ is an inner corner of } \lambda\},\\
[q^{b(\lambda)+2}]\,\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) &= \binom{H_1}{2} + H_2 - 1,\\
[q^{b(\lambda)+3}]\,\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) &= \binom{H_1+1}{3} + (H_1-1)(H_2-1) + (H_3-1)\\
[q^{b(\lambda)+4}]\,\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) &= \binom{H_1+2}{4} + \binom{H_2}{2} + \binom{H_1}{2}(H_2-1)\\
&\quad + (H_1-1)(H_3-1) + (H_4-1).
\end{aligned}$$

These exact formulas hold for all $|\lambda| \ge 4$. For smaller size partitions some terms will not appear. It is interesting to compare these polyomials to the ones described by Knuth for the number of permutations with k inversions in S_n in [52, p.16]. See also [92, Ex. 1.124].

Remark 6.2.7. Since $\#\,\mathrm{SYT}(\lambda)$ typically grows extremely quickly, Stanley's formula offers a very useful way to compute $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$ by expressing both the numerator and denominator, up to a q-shift, as a product of cyclotomic polynomials and canceling all factors in the denominator.

The following strengthening of Stanley's formula to $\underline{\lambda}$ is well known (e.g. see [94, (5.6)]), though since it is somewhat difficult to find explicitly in the literature, we include a short proof.

Theorem 6.2.8. *Let* $\underline{\lambda} = (\lambda^{(1)}, \ldots, \lambda^{(m)})$ *where* $\lambda^{(i)} \vdash n_i$ *and* $n = n_1 + \cdots + n_m$. *Then*

$$\mathrm{SYT}(\underline{\lambda})^{\mathrm{maj}}(q) = \binom{n}{n_1, \ldots, n_m}_q \prod_{i=1}^{m} \mathrm{SYT}(\lambda^{(i)})^{\mathrm{maj}}(q). \tag{6.5}$$

Proof. The stable principal specialization of skew Schur functions is given by

$$s_{\lambda/\nu}(1, q, q^2, \ldots) = \frac{\mathrm{SYT}(\lambda/\nu)^{\mathrm{maj}}(q)}{\prod_{j=1}^{|\lambda/\nu|}(1 - q^j)};$$

see [94, Lemma 3.1] or [91, Prop.7.19.11]. On the other hand, it is easy to see directly that

$$s_{\underline{\lambda}}(x_1, x_2, \ldots) = \prod_{i=1}^{m} s_{\lambda^{(i)}}(x_1, x_2, \ldots).$$

The result quickly follows. □

Theorem 5.2.2 and Theorem 6.2.8 have several immediate corollaries. First, we recover MacMahon's result, Theorem 1.1.1, from Theorem 6.2.8 when $\underline{\lambda} = ((\alpha_1), (\alpha_2), \ldots)$ by using the maj-preserving bijection ϕ above. Second, each $\mathrm{SYT}(\underline{\lambda})^{\mathrm{maj}}(q)$ is symmetric (up to a q-shift) with leading coefficient 1. In particular, there is a unique "maj-minimizer" and "maj-maximizer" tableau in each $\mathrm{SYT}(\underline{\lambda})$. Moreover,

$$\min \mathrm{maj}(\mathrm{SYT}(\underline{\lambda})) = b(\underline{\lambda}) := \sum_i b(\lambda^{(i)})$$

and

$$\max \mathrm{maj}(\mathrm{SYT}(\underline{\lambda})) = b(\underline{\lambda}) + \binom{|\underline{\lambda}| + 1}{2} - \sum_{c \in \underline{\lambda}} h_c.$$

For general skew shapes, $q^{-b(\lambda/\mu)}\,\mathrm{SYT}(\lambda/\mu)^{\mathrm{maj}}(q)$ does not factor as a product of cyclotomic polynomials. A "q-Naruse" formula due to Morales–Pak–Panova, [66, (3.4)], gives an analogue of Theorem 5.2.2 involving a sum over "excited diagrams," though for $\underline{\lambda}$ no excited moves are allowed.

6.2.3 Exponential Generating Functions

We now introduce exponential generating functions and the Bernoulli numbers, which will be used with cumulants shortly.

Definition 6.2.9. Given a rational sequence $(g_d)_{d=0}^{\infty} = (g_0, g_1, \ldots)$, the corresponding *ordinary generating function* is

$$O_g(t) := \sum_{d \geq 0} g_d t^d$$

and the corresponding *exponential generating function* is

$$E_g(t) := \sum_{d \geq 0} g_d \frac{t^d}{d!}.$$

Conversely, any rational power series

$$F(t) = \sum_{d \geq 0} f_d t^d = \sum_{d \geq 0} d! f_d \frac{t^d}{d!}$$

is the ordinary generating function of the sequence $(f_d)_{d=0}^{\infty} = (f_0, f_1, \ldots)$ and the exponential generating function of the sequence $(d! f_d)_{d=0}^{\infty}$. The exponential generating functions we will encounter will all have positive radius of convergence.

It is easy to describe products, quotients and compositions of generating functions. We recall in particular a formula for compositions of exponential generating functions for later use. Given two rational sequences $f = (f_d)_{d=0}^{\infty}$, $g = (g_d)_{d=0}^{\infty}$ such that $f_0 = 0$ and $g_0 = 1$, the composition of their exponential generating functions $E_g \circ E_f$ is again an exponential generating function for a rational sequence h, say $E_h(t) = E_g(E_f(t))$. For example, if $E_f(t) = \sum f_d t^d / d!$

and $E_g(t) = e^t$, so $g_i = 1$ for all i, then by [91, Cor. 5.1.6], the corresponding sequence $(h_d)_{d=0}^{\infty}$ is given by $h_0 = 1$ and, for $d \geq 1$,

$$h_d = \sum_{\pi \in \Pi_d} \prod_{b \in \pi} f_{|b|}, \tag{6.6}$$

where Π_d is the collection of all set partitions $\pi = \{b_1, b_2, \dots, b_k\}$ of $\{1, 2, \dots, d\}$. Collecting together S_d-orbits of Π_d in (6.6) quickly gives

$$h_d = \sum_{\lambda \vdash d} \frac{d!}{z_\lambda} \prod_i \frac{f_{\lambda_i}}{(\lambda_i - 1)!} \tag{6.7}$$

where if λ has m_i parts of length i, then $z_\lambda := 1^{m_1} 2^{m_2} \cdots m_1! m_2! \cdots$. A more computationally efficient, recursive approach to (6.6) is the formula [91, Prop. 5.1.7]

$$h_d = f_d + \sum_{m=1}^{d-1} \binom{d-1}{m-1} f_m h_{d-m}. \tag{6.8}$$

Example 6.2.10. The *Bernoulli numbers* $(B_d)_{d=0}^{\infty}$ are rational numbers determined by the exponential generating function $E_B(t) := t/(1 - e^{-t})$. The first few terms in the sequence are

$$B_0 = 1,\ B_1 = \frac{1}{2},\ B_2 = \frac{1}{6},\ B_3 = 0,\ B_4 = -\frac{1}{30},\ B_5 = 0,\ B_6 = \frac{1}{42},$$

$$B_7 = 0,\ B_8 = -\frac{1}{30},\ B_9 = 0,\ B_{10} = \frac{5}{66},\ B_{11} = 0,\ B_{12} = -\frac{691}{2730}.$$

The *divided Bernoulli numbers* are given by $\frac{B_d}{d}$ for $d \geq 1$. Their exponential generating function $E_D(t)$ satisfies $1 + t\frac{d}{dt}E_D(t) = E_B(t)$, from which it follows that

$$E_D(t) := \sum_{d \geq 1} \frac{B_d}{d} \frac{t^d}{d!} = \log\left(\frac{e^t - 1}{t}\right).$$

We caution that a common alternate convention uses $B_1 = -\frac{1}{2}$ with all other entries the same, corresponding with the exponential generating function $t/(e^t - 1)$.

The Bernoulli numbers have many interesting properties; see [64, 101] and [38, Section 6.5]. For example, they appear in the polynomial expansion of certain sums of dth powers,

$$\sum_{k=1}^{n} k^d = \frac{1}{d+1} \sum_{k=0}^{d} \binom{d+1}{k} B_k \, n^{d+1-k}. \tag{6.9}$$

Compare the formula for sums of dth powers to the Riemann zeta function $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ which can be evaluated at complex values $s \neq 1$ by analytic continuation. The divided Bernoulli numbers which appear in our formula (6.1) have the form $\frac{B_d}{d} = -\zeta(1-d)$.

6.2.4 Probabilistic Generating Functions

We next review basic vocabulary and notation for moments and cumulants of random variables. *We assume throughout the chapter that the density or mass functions of our random variables exist and decay at least exponentially in the tails.* This simple condition will be manifestly apparent in all of our examples and allows us to avoid some technical digressions. See [47] for more details.

Definition 6.2.11. Let X be a real-valued random variable where either X is continuous with probability density function $f\colon \mathbb{R} \to \mathbb{R}_{\geq 0}$ or X is discrete with probability mass function $f\colon \mathbb{Z} \to \mathbb{R}_{\geq 0}$. The *cumulative distribution function* (CDF) of X is given by

$$F(t) := \int_{-\infty}^{t} f(x)\, dx \qquad \text{or} \qquad F(t) := \sum_{k \leq t} f(k)$$

depending on whether X is continuous or discrete. For any continuous real-valued function g, there is an associated random variable $g(X)$. The *expectation* of $g(X)$ is given by

$$\mathbb{E}[g(X)] := \int_{\mathbb{R}} g(x) f(x)\, dx \qquad \text{or} \qquad \mathbb{E}[g(X)] := \sum_{k=-\infty}^{\infty} g(k) f(k).$$

The *mean* and *variance* of X are, respectively,

$$\mu := \mathbb{E}[X] \qquad \text{and} \qquad \sigma^2 := \mathbb{E}[(X-\mu)^2].$$

For $d \in \mathbb{Z}_{\geq 0}$, the *dth moment* and *dth central moment* of X are, respectively,

$$\mu_d := \mathbb{E}[X^d] \qquad \text{and} \qquad \alpha_d := \mathbb{E}[(X-\mu)^d].$$

The *moment-generating function* of X is

$$M_X(t) := \mathbb{E}[e^{tX}] = \sum_{d=0}^{\infty} \mu_d \frac{t^d}{d!},$$

which has a positive radius of convergence by our tail decay assumption. The *characteristic function* of X is

$$C_X(t) := \mathbb{E}[e^{itX}],$$

which exists for all $t \in \mathbb{R}$ and which is the Fourier transform of f, the density or mass function associated to X.

Example 6.2.12. The *probability generating function* of the random variable X associated with $\mathrm{stat}\colon W \to \mathbb{Z}_{\geq 0}$ sampled uniformly is

$$\mathbb{E}[q^X] = \frac{1}{\#W} W^{\mathrm{stat}}(q).$$

Letting $q = e^t$, the moment-generating function and characteristic function of X are

$$M_X(t) = \frac{1}{\#W} W^{\mathrm{stat}}(e^t) \qquad \text{and} \qquad C_X(t) = \frac{1}{\#W} W^{\mathrm{stat}}(e^{it}).$$

The last expression reveals an intimate connection between the study of generating functions of combinatorial statistics evaluated on the unit circle and the underlying probability distribution via the Fourier transform. In particular, the distribution determines the characteristic function and the moment-generating function, and conversely each of these determine the distribution.

Definition 6.2.13. The *cumulants* $\kappa_1, \kappa_2, \ldots$ of X are defined to be the coefficients of the exponential generating function

$$K_X(t) := \sum_{d=1}^{\infty} \kappa_d \frac{t^d}{d!} := \log M_X(t) = \log \mathbb{E}[e^{tX}].$$

While cumulants of random variables may initially be less intuitive than moments, they lead to nicer formulas in many cases, including Theorem 6.1.1, and they often have more useful properties. See [67] for some history and applications. We will use the following properties of cumulants. The proofs are straightforward from the definitions.

1. *(Familiar Values)* The first three cumulants are $\kappa_1 = \mu$, $\kappa_2 = \sigma^2$, and $\kappa_3 = \alpha_3$. The higher cumulants typically differ from the moments and central moments.

2. *(Shift Invariance)* The second and higher cumulants of X agree with those for $X - c$ for any $c \in \mathbb{R}$.

3. *(Homogeneity)* The dth cumulant of cX is $c^d \kappa_d$ for $c \in \mathbb{R}$.

4. *(Additivity)* The cumulants of the sum of *independent* random variables are the sums of the cumulants.

5. *(Polynomial Equivalence)* The cumulants, moments, and central moments are determined by polynomials in any one of these three sequences.

The polynomial equivalence property can be made explicit by the results in Section 6.2.3. Using (6.8) (or similarly (6.6) or (6.7)) allows us to express the dth moment of X as a polynomial function of the first d cumulants of X and vice versa via the recurrence

$$\mu_d = \kappa_d + \sum_{m=1}^{d-1} \binom{d-1}{m-1} \kappa_m \mu_{d-m}. \tag{6.10}$$

Using the shift invariance property of cumulants, the corresponding formula for the central moments in terms of the cumulants can be obtained from (6.10) by setting $\kappa_1 = 0$ and leaving the other cumulants alone. This gives, for $d > 1$,

$$\alpha_d = \kappa_d + \sum_{m=2}^{d-2} \binom{d-1}{m-1} \kappa_m \alpha_{d-m}. \tag{6.11}$$

For instance, at $d = 3$ we have

$$\mu_3 = \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3.$$

Setting $\kappa_1 = 0$ yields $\alpha_3 = \kappa_3$ as mentioned above.

6.2.5 Cumulant Examples

Next we describe the cumulants of some well-known distributions and use one of them to deduce a result of Hwang–Zacharovas, which immediately yields Theorem 6.1.1 as a corollary.

Example 6.2.14. Let $X = \mathcal{N}(\mu, \sigma^2)$ be the normal random variable with mean μ and variance σ^2. The density function of X is $f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$. Taking the Fourier transform gives the characteristic function $\mathbb{E}[e^{itX}] = \exp\left(i\mu t - \frac{1}{2}\sigma^2 t^2\right)$, so the moment-generating function is $\mathbb{E}[e^{tX}] = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$ and the cumulants are

$$\kappa_d = \begin{cases} \mu & d = 1, \\ \sigma^2 & d = 2, \\ 0 & d \geq 3. \end{cases} \tag{6.12}$$

Using (6.7) to compute the central moments of X from (6.12), we effectively set $\kappa_1 = 0$ and note that only $\lambda = (2, 2, \ldots, 2) = (2^{d/2})$ contributes, in which case $\alpha_d = \kappa_2^{d/2} d!/(2^{d/2}(d/2)!)$. It follows that

$$\alpha_d = \begin{cases} 0 & \text{if } d \text{ is odd,} \\ \sigma^d (d-1)!! & \text{if } d \text{ is even.} \end{cases}$$

Example 6.2.15. Let $U = U[0,1]$ be the continuous uniform random variable whose density takes the value 1 on the interval $[0,1]$ and 0 otherwise. Then the moment generating function is $M_U(t) = \int_0^1 e^{tx}dx = (e^t - 1)/t$, so the cumulant generating function $K_U(t) = \log M_U(t)$ coincides with the exponential generating function for the divided Bernoulli numbers from Section 6.2.3. That is, $\kappa_d = B_d/d$ for $d \geq 1$.

Example 6.2.16. Let $U = U_n$ be the discrete uniform random variable supported on $\{0, 1, \ldots, n-1\}$. The probability generating function for U is $[n]_q/n = (q^n - 1)/(n(q-1))$, so the cumulant generating function is

$$K_U(t) = \log\left(\frac{e^{nt} - 1}{n(e^t - 1)}\right) = \log\left(\frac{e^{nt} - 1}{nt}\right) - \log\left(\frac{e^t - 1}{t}\right).$$

It follows that for $d \geq 1$, the divided Bernoulli numbers arise again in this context,

$$\kappa_d = \frac{B_d}{d}(n^d - 1). \tag{6.13}$$

Product formulas for polynomials such as Theorem 1.1.1 and Theorem 5.2.2 give rise to explicit formulas for cumulants and moments according to the following theorem. The first part appeared in the work of Hwang–Zacharovas [44, §4.1] building on the work of Chen–Wang–Wang [16, Thm. 3.1] for q-Catalan numbers. It follows immediately from Example 6.2.16 and (6.7).

Theorem 6.2.17. *Suppose $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_m\}$ are multisets of positive integers such that*

$$P(q) = \frac{\prod_{j=1}^m [a_j]_q}{\prod_{j=1}^m [b_j]_q} = \sum c_k q^k \in \mathbb{Z}_{\geq 0}[q].$$

Let X be a discrete random variable with $\mathbb{P}[X = k] = c_k/P(1)$. Then the dth cumulant of X is

$$\kappa_d = \frac{B_d}{d} \sum_{j=1}^{m} (a_j^d - b_j^d) \tag{6.14}$$

where B_d is the dth Bernoulli number (with $B_1 = \frac{1}{2}$). Moreover, the dth central moment of X is

$$\alpha_d = \sum_{\substack{\lambda \vdash d \\ \textit{has all parts even}}} \frac{d!}{z_\lambda} \prod_{i=1}^{\ell(\lambda)} \frac{B_{\lambda_i}}{\lambda_i!} \left[\sum_{j=1}^{m} \left(a_j^d - b_j^d \right) \right]. \tag{6.15}$$

and the dth moment of X is

$$\mu_d = \sum_{\substack{\lambda \vdash d \\ \textit{has all parts either} \\ \textit{even or size } 1}} \frac{d!}{z_\lambda} \prod_{i=1}^{\ell(\lambda)} \frac{B_{\lambda_i}}{\lambda_i!} \left[\sum_{j=1}^{m} \left(a_j^d - b_j^d \right) \right]. \tag{6.16}$$

Remark 6.2.18. Theorem 6.1.1 follows immediately from Theorem 6.2.17 and Theorem 5.2.2. Moreover, using Theorem 6.2.8 the cumulants of $X_{\underline{\lambda}}[\mathrm{maj}]$ are, for $d > 1$,

$$\kappa_d^{\underline{\lambda}} = \frac{B_d}{d} \left[\sum_{j=1}^{|\underline{\lambda}|} j^d - \sum_{c \in \underline{\lambda}} h_c^d \right] \tag{6.17}$$

Adin and Roichman [2] had previously used Theorem 5.2.2 to compute the mean and variance of $X_\lambda[\mathrm{maj}]$ as

$$\mu = \frac{\binom{|\lambda|}{2} - b(\lambda') + b(\lambda)}{2} = b(\lambda) + \frac{1}{2} \left[\sum_{j=1}^{|\lambda|} j - \sum_{c \in \lambda} h_c \right],$$

and

$$\sigma^2 = \frac{1}{12} \left[\sum_{j=1}^{|\lambda|} j^2 - \sum_{c \in \lambda} h_c^2 \right].$$

6.2.6 Asymptotic Normality

Here we briefly summarize the notion of asymptotic normality and more generally convergence in distribution. Asymptotic normality lies at the intersection of probability and combinatorics. For an introduction, we recommend the chapter by Canfield in [12, Chapter 3]. We also review some of the many examples.

Definition 6.2.19. Let $X_1, X_2, \ldots$ and X be real-valued random variables with cumulative distribution functions $F_1, F_2, \ldots$ and F, respectively. We say $X_1, X_2, \ldots$ *converges in distribution* to X if for all $t \in \mathbb{R}$ at which F is continuous we have

$$\lim_{n\to\infty} F_n(t) = F(t).$$

Recall from the introduction that for a real-valued random variable X with mean μ and variance $\sigma^2 > 0$, the corresponding *normalized random variable* is

$$X^* := \frac{X - \mu}{\sigma}.$$

Observe that X^* has mean $\mu^* = 0$ and variance ${\sigma^*}^2 = 1$. The moments and central moments of X^* agree for $d \geq 2$ and are given by

$$\mu_d^* = \alpha_d^* = \alpha_d/\sigma^d.$$

Similarly, the cumulants of X^* are given by $\kappa_1^* = 0$, $\kappa_2^* = 1$, and $\kappa_d^* = \kappa_d/\sigma^d$ for $d \geq 2$.

Definition 6.2.20. Let $X_1, X_2, \ldots$ be a sequence of real-valued random variables. We say the sequence is *asymptotically normal* if $X_1^*, X_2^*, \ldots$ converges in distribution to the standard normal $\mathcal{N}(0, 1)$.

Example 6.2.21. Let $W_n := 2^{[n]}$ be the set of all subsets of $[n] := \{1, 2, \ldots, n\}$. Let X_n denote the random variable given by the size statistic on W_n taken uniformly. The following theorem is credited to de Moivre and Laplace. See [12, Theorem 3.2.1] for further discussion and references. It may be proven using Stirling's approximation.

Theorem 6.2.22 (de Moivre–Laplace)**.** *The sequence of size random variables above is asymptotically normal.*

Asymptotic normality results for combinatorial statistics are plentiful. See Table 6.1 for many more examples and further references.

Many combinatorial statistics arise from sets indexed by more complicated objects than the positive integers, in which case one can "let $n \to \infty$" in many different ways. The following result due to Canfield, Janson, and Zeilberger illustrates a more interesting limit.

Theorem 6.2.23. *[13, Theorem 1.2] Let $\alpha^{(1)}, \alpha^{(2)}, \ldots$ be a sequence of compositions, possibly of differing lengths. If $\alpha \vDash n$ has maximum m, write $s(\alpha) := n - m$. Let X_i be the inversion (or major index) statistic on words of content $\alpha^{(i)}$. Then $X_1, X_2, \ldots$ is asymptotically normal if and only if*

$$s(\alpha^{(i)}) \to \infty.$$

Remark 6.2.24. Explorations equivalent to Theorem 6.2.23 appeared significantly earlier than [13] in other contexts, for instance [21, p. 127-128] and (in the two-letter case) [63]. See [14] for further discussion and references.

6.2.7 The Method of Moments

We next describe several explicit criteria for establishing convergence in distribution or asymptotic normality of a sequence of random variables. We emphasize that the assumptions in Section 6.2.4 remain in effect throughout the chapter. Without those assumptions, more degenerate behavior is possible in Theorem 6.2.26 and Corollary 6.2.27.

Theorem 6.2.25 (Lévy's Continuity Theorem, [11, Theorem 26.3])**.** *A sequence $X_1, X_2, \ldots$ of real-valued random variables converges in distribution to a real-valued random variable X if and only if, for all $t \in \mathbb{R}$,*

$$\lim_{n \to \infty} \mathbb{E}[e^{itX_n}] = \mathbb{E}[e^{itX}].$$

Statistic	Set	Generating Function	References
# elements	subsets	$(1+q)^n$	classical
# parts	strict partitions	$\prod_{m=1}^{\infty}(1+xy^m)$	[23]
length/inversion number/major index	S_n	$[n]_q!$	[24], [37]
# cycles; # left-to-right minima	S_n	$\prod_{i=0}^{n-1}(q+i)$	[24], [37]
# descents	S_n	Eulerian polynomial	[19, pp. 150–154]
# blocks	set partitions	$\sum_k S(n,k)q^k$	[42]
# valleys	Dyck paths	$\frac{1}{[n+1]_q}\binom{2n}{n}_q$	[16, Cor. 3.3]; [32, p. 255]
length/inversion number/major index	S_n/S_J, words content α	$\binom{n}{\alpha}_q$	see Remark 6.2.24
major index	$\mathrm{SYT}(\lambda)$	$q^{b(\lambda)}\frac{[n]_q!}{\prod_{c\in\lambda}[h_c]_q}$	Theorem 6.1.3

Table 6.1: Summary of some asymptotic normality results for combinatorial statistics. See [12, Ch. 3].

Theorem 6.2.26 (Frechét-Shohat Theorem, [11, Theorem 30.2])**.** *Let $X_1, X_2, \ldots$ be a sequence of real-valued random variables and let X be a real-valued random variable. If*

$$\lim_{n\to\infty} \mu_d^{X_n} = \mu_d^X \quad \forall d \in \mathbb{Z}_{\geq 1}, \tag{6.18}$$

then $X_1, X_2, \ldots$ converges in distribution to X.

By Theorem 6.2.25, we may test for asymptotic normality by checking if the normalized characteristic functions tend pointwise to the characteristic function of the standard normal. Likewise by Theorem 6.2.26 we may instead perform the check on the level of individual normalized moments, which is often referred to as the *method of moments.* By (6.10) we may further replace the moment condition (6.18) with the cumulant condition

$$\lim_{n\to\infty} \kappa_d^{X_n} = \kappa_d^X. \tag{6.19}$$

For instance, we have the following explicit criterion.

Corollary 6.2.27. *A sequence $X_1, X_2, \ldots$ of real-valued random variables is asymptotically normal if for all $d \geq 3$ we have*

$$\lim_{n\to\infty} \frac{\kappa_d^{X_n}}{(\sigma^{X_n})^d} = 0 \tag{6.20}$$

Remark 6.2.28. In fact, the converse of Theorem 6.2.26 and Corollary 6.2.27 holds in our context, which follows from, for instance, the uniform tail decay estimate in [44, Lemma 2.8]. We do not require this implication and so do not make it precise.

6.2.8 Local Limit Theorems

Asymptotic normality concerns cumulative distribution functions, so it gives estimates for the number of combinatorial objects with a large range of statistics. However, our original motivation was to count combinatorial objects with a given statistic. Estimates of this latter form are frequently referred to as *local limit theorems.* Theorem 5.1.7 above is one such example. Further motivation was provided by the following analogue of Theorem 6.2.23.

Theorem 6.2.29. *[13, Theorem 4.5] There exists a positive constant c such that for every C, the following is true. Uniformly for all compositions $\alpha = (\alpha_1, \ldots, \alpha_m)$ such that $\max_i \alpha_i \leq Ce^{cs(\alpha)}$ and all integers k,*

$$\mathbb{P}[X_\alpha = k] = \frac{1}{\sigma\sqrt{2\pi}} \left(e^{-(k-\mu)^2/(2\sigma^2)} + O\left(\frac{1}{s(\alpha)}\right)\right)$$

where X_α denotes inversions on words of content α.

6.3 Asymptotic Normality for $X_{\underline{\lambda}}[\mathrm{maj}]$

In this section, we give asymptotic estimates for the normalized cumulants $\kappa_d^{\lambda *}$ powerful enough to prove Theorem 6.1.3 and its generalization to $\underline{\lambda}$, Theorem 6.3.8. Much of the argument applies equally well to arbitrary skew shapes, though the connection through Theorem 6.2.17 to cumulants holds only for shapes $\underline{\lambda}$.

Definition 6.3.1. A *reverse standard Young tableau* of shape λ/ν is a bijective filling of λ/ν which strictly decreases along rows and columns. The set of reverse standard Young tableaux of shape λ/ν is denoted $\mathrm{RSYT}(\lambda/\nu)$.

Lemma 6.3.2. *Let $\lambda/\nu \vdash n$ and $T \in \mathrm{RSYT}(\lambda)$. Then for all $c \in \lambda/\nu$,*

$$T_c \geq h_c. \tag{6.21}$$

Furthermore, for any positive integer d,

$$\sum_{j=1}^{n} j^d - \sum_{c\in\lambda} h_c^d = \sum_{c\in\lambda} (T_c^d - h_c^d) = \sum_{c\in\lambda} (T_c - h_c)\mathbf{h}_{d-1}(T_c, h_c), \tag{6.22}$$

where $\mathbf{h}_{d-1}$ denotes the complete homogeneous symmetric function.

Proof. For (6.21), equality holds at the outer corner c where $T_c = 1$. Removing c and subtracting 1 from each remaining entry in T allows us to induct. Equation (6.22) follows immediately by rearranging the terms and factoring $(T_c^d - h_c^d) = (T_c - h_c)\sum_{k=0}^{d-1} T_c^{d-1-k} h_c^k$. $\square$

Lemma 6.3.3. *Let* $\lambda/\nu \vdash n$ *such that* $\max_{c\in\lambda/\nu} h_c < 0.8n$*. Let* d *be any positive integer. Then*

$$\frac{n^{d+1}}{26(d+1)} - 2(0.8)^d n^d < \sum_{j=1}^{n} j^d - \sum_{c\in\lambda} h_c^d < \frac{n^{d+1}}{d+1} + n^d.$$

Proof. Using Riemmanian sums for $\int_0^n x^d dx$, we obtain the bounds

$$\frac{n^{d+1}}{d+1} < \sum_{j=1}^{n} j^d < \frac{n^{d+1}}{d+1} + n^d \tag{6.23}$$

for all positive integers d, n. The upper bound in the lemma now follows immediately.

For the lower bound, label the cells of λ by some $T \in \mathrm{RSYT}(\lambda)$. By (6.21), $h_c \leq T_c$, and by assumption we have $h_c < 0.8n$ for all $c \in \lambda/\nu$. Considering the tighter of these two bounds on each summand and using (6.23) again, we have

$$\begin{aligned}
\sum_{c\in\lambda/\nu} h_c^d &< \sum_{\substack{j\in[n]\\ j<0.8n}} j^d + \sum_{\substack{j\in[n]\\ j\geq 0.8n}} (0.8n)^d \\
&< \frac{\lfloor 0.8n\rfloor^{d+1}}{d+1} + \lfloor 0.8n\rfloor^d + (n - \lceil 0.8n\rceil + 1)(0.8n)^d \\
&\leq \frac{(0.8n)^{d+1}}{d+1} + 2(0.8n)^d + (0.2)(0.8)^d n^{d+1}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\sum_{j=1}^{n} j^d - \sum_{c\in\lambda/\nu} h_c^d &> \frac{n^{d+1}}{d+1} - \frac{(0.8n)^{d+1}}{d+1} - 2(0.8n)^d - (0.2)(0.8)^d n^{d+1} \\
&= \left(\frac{1}{d+1}(1 - (0.8)^{d+1}) - 0.2(0.8)^d\right) n^{d+1} - 2(0.8)^d n^d.
\end{aligned}$$

It is easy to check that the coefficient on n^{d+1} is bounded below by $\frac{1}{26(d+1)}$ for all positive integers d. The result follows. $\square$

Definition 6.3.4. Given any partition $\lambda/\nu \vdash n$, let the *aft* of λ/ν be

$$\operatorname{aft}(\lambda/\nu) := n - \max_{c\in\lambda/\nu}\{\operatorname{arm}(c), \operatorname{leg}(c)\}$$

where $\operatorname{arm}(c)$ is the number of cells in the same row as c to the right of c, including c itself, and $\operatorname{leg}(c)$ is the number of cells in the same column as c below c, including c. When $\nu = \varnothing$, we have $\operatorname{aft}(\lambda) = n - \max\{\lambda_1, \lambda'_1\}$ as above. When $\lambda/\nu = \underline{\lambda}$, we have $\operatorname{aft}(\underline{\lambda}) = n - \max_i\{\lambda^{(i)}_1, {\lambda^{(i)}}'_1\}$. Note that $h_c = \operatorname{arm}(c) + \operatorname{leg}(c) - 1$.

Lemma 6.3.5. *Let $\lambda/\nu \vdash n$ such that $\max_{c\in\lambda/\nu} h_c \geq 0.8n$, and let d be any positive integer. Furthermore, suppose $n \geq 10$. Then,*

$$\operatorname{aft}(\lambda/\nu)\frac{(0.1n)^d}{d} \;\leq\; \sum_{j=1}^{n} j^d - \sum_{c\in\lambda/\nu} h_c^d \;\leq\; 2\operatorname{aft}(\lambda/\nu)\left(n^d + dn^{d-1}\right). \tag{6.24}$$

Proof. The result holds trivially if $\operatorname{aft}(\lambda/\nu) = 0$ since in that case λ/ν is a single row or column, so assume $\operatorname{aft}(\lambda/\nu) > 0$. Let $m \in \lambda/\nu$ have $h_m \geq 0.8n$, where we may assume m is the first cell in its row and column. For convenience, we may further assume by symmetry that $\operatorname{arm}(m) \geq \operatorname{leg}(m)$. Since $h_m \geq 0.8n$, it also follows that $\operatorname{aft}(\lambda/\nu) = n - \operatorname{arm}(m)$.

Now let R be the set of cells in the row of m, not including m itself, which are the only cells of λ/ν in their columns. We claim that $\#R \geq 0.1n$. To see this, since $h_m \geq 0.8n$, there are at most $n - h_m \leq 0.2n$ cells of λ/ν which could possibly be in the columns of the cells of the row of m not including m. Since $\operatorname{arm}(m) \geq \operatorname{leg}(m)$ and $\operatorname{arm}(m)+\operatorname{leg}(m)-1 = h_m \geq 0.8n$, we have $\operatorname{arm}(m) \geq 0.4n$. Hence no more than $0.2n$ of the $0.4n - 1$ cells in the row of m not including m can be excluded from R, so $\#R \geq 0.4n - 1 - 0.2n \geq 0.1n$ for $n \geq 10$. Since λ/ν is a skew partition, R is connected.

Construct $T \in \operatorname{RSYT}(\lambda/\nu)$ iteratively as follows; see Figure 6.4 for an example. At each step of the iteration, we will first increment all existing labels by 1 and then label a new outer cell with 1. Begin by adding the cells of the row of m from left to right until the last cell of R has been added. Now add the remaining cells of λ/ν row by row starting at the topmost row and going from left to right. It is easy to see that the result respects the row and column conditions, so $T \in \mathrm{RSYT}(\lambda/\nu)$.

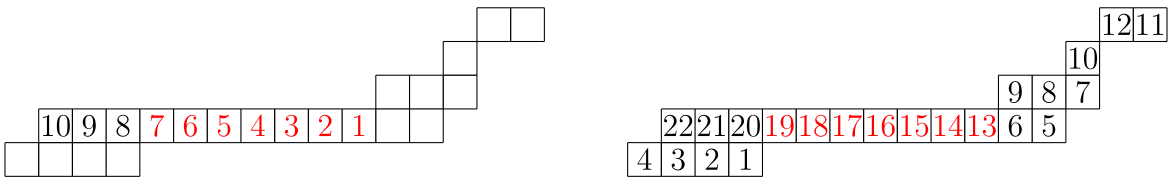


Figure 6.4: On the left, the partially constructed $T \in \mathrm{RSYT}(\lambda/\nu)$ after all the cells of R (in red) have been filled. On the right, the final $T \in \mathrm{RSYT}(\lambda/\nu)$. Here $\mathrm{aft}(\lambda/\nu) = 10$.

Consider the inequalities $T_c \geq h_c$. At every step of the iteration, a labeled cell has T_c increase by 1, while h_c increases by 1 if and only if the newly labeled cell is in the hook of c. That is, for the final filling T, $T_c - h_c$ counts the number of times after cell c was filled that the new cell was not in the same row or column as c. For each $r \in R$, it follows that $T_c - h_c = n - \mathrm{arm}(m) = \mathrm{aft}(\lambda/\nu)$.

For the lower bound, we now find

$$\begin{aligned}
\sum_{j=1}^{n} j^d - \sum_{c\in\lambda/\nu} h_c^d &\geq \sum_{c\in R} (T_c - h_c)\mathbf{h}_{d-1}(T_c, h_c) \\
&= \sum_{c\in R} \mathrm{aft}(\lambda/\nu)\mathbf{h}_{d-1}(h_c + \mathrm{aft}(\lambda/\nu), h_c) \\
&\geq \sum_{k=1}^{\lfloor 0.1n \rfloor} \mathrm{aft}(\lambda/\nu)\mathbf{h}_{d-1}(k + \mathrm{aft}(\lambda/\nu), k) \\
&\geq \mathrm{aft}(\lambda/\nu) \sum_{k=1}^{\lfloor 0.1n \rfloor} k^{d-1} \\
&\geq \mathrm{aft}(\lambda/\nu)\frac{(0.1n)^d}{d},
\end{aligned}$$

where the second inequality uses the fact that $\{h_c : c \in R\}$ has pointwise lower bounds of $\{1, 2, \ldots, \#R\}$ and the last inequality uses (6.23).

For the upper bound, we construct a new $T \in \mathrm{RSYT}(\lambda/\nu)$ as follows; see Figure 6.5 for an example. First, for each cell c in the row of m taken from left to right, add the topmost cell in the column of c. Now add the remaining cells of λ/ν exactly as before. Again consider the final differences $T_c - h_c$. For cells added in the second stage, $T_c - h_c$ could increase no more than $n - \mathrm{arm}(m) = \mathrm{aft}(\lambda/\nu)$ times, so $T_c - h_c \leq \mathrm{aft}(\lambda/\nu)$ for such c. For cells added in the first stage, we claim that $T_c - h_c \leq 2\,\mathrm{aft}(\lambda/\nu)$. For the claim, it suffices to show that after the first stage, for cells added in the first stage, $T_c - h_c \leq \mathrm{aft}(\lambda/\nu)$. During the first stage, the differences $T_c - h_c$ are zero while cells of row m are being added. Afterwards during the first phase, cells not in row m are added, of which there are no more than $n - \mathrm{arm}(m) = \mathrm{aft}(\lambda/\nu)$, so the differences $T_c - h_c$ can increase no more than $\mathrm{aft}(\lambda/\nu)$ many times during the first phase, completing the claim.

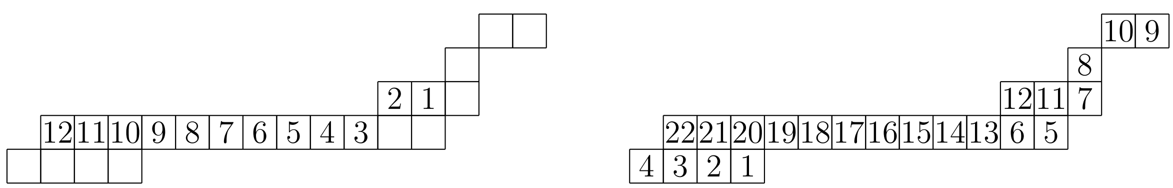


Figure 6.5: On the left, the second partially constructed $T \in \mathrm{RSYT}(\lambda/\nu)$ after the first $\mathrm{arm}(m)$ cells have been filled. On the right, the final $T \in \mathrm{RSYT}(\lambda/\nu)$.

Having established that $T_c - h_c \leq 2\,\mathrm{aft}(\lambda/\nu)$, we now find by (6.22) and (6.23),

$$\begin{aligned}
\sum_{j=1}^{n} j^d - \sum_{c\in\lambda/\nu} h_c^d &= \sum_{c\in\lambda/\nu} (T_c - h_c)\mathbf{h}_{d-1}(T_c, h_c) \\
&\leq \sum_{c\in\lambda/\nu} 2\,\mathrm{aft}(\lambda/\nu)\mathbf{h}_{d-1}(T_c, T_c) \\
&= 2\,\mathrm{aft}(\lambda/\nu) \sum_{j=1}^{n} dj^{d-1} \\
&< 2\,\mathrm{aft}(\lambda/\nu)\left(n^d + dn^{d-1}\right).
\end{aligned}$$

$\square$

Corollary 6.3.6. *For fixed $d \in \mathbb{Z}_{\geq 1}$, uniformly for all $\lambda/\nu \vdash n$,*

$$\sum_{j=1}^{n} j^d - \sum_{c \in \lambda/\nu} h_c^d = \Theta(\mathrm{aft}(\lambda/\nu) n^d).$$

That is, there are constants $c_1, c_2 > 0$ such that

$$c_1 \,\mathrm{aft}(\lambda/\nu) n^d \leq \sum_{j=1}^{n} j^d - \sum_{c \in \lambda/\nu} h_c^d \leq c_2 \,\mathrm{aft}(\lambda/\nu) n^d.$$

Proof. When $\max_{c \in \lambda/\nu} h_c \geq 0.8n$, the result follows from Lemma 6.3.5. On the other hand, when $\max_{c \in \lambda/\nu} h_c < 0.8n$, then $n \geq \mathrm{aft}(\lambda/\nu) \geq 0.2n$, and the result follows from Lemma 6.3.3. $\square$

Corollary 6.3.7. *Fix d to be an even positive integer. Uniformly for all $\underline{\lambda} \vdash n$, the normalized cumulant $|\kappa_d^{\underline{\lambda}*}|$ of $X_{\underline{\lambda}}[\mathrm{maj}]$ is $\Theta(\mathrm{aft}(\underline{\lambda})^{1-d/2})$.*

Proof. For d even, by (6.17) and Corollary 6.3.6, we have

$$|\kappa_d^{\underline{\lambda}}| = \Theta(\mathrm{aft}(\underline{\lambda}) n^d).$$

Consequently by the homogeneity of cumulants, we have

$$|\kappa_d^{\underline{\lambda}*}| = \left| \frac{\kappa_d^{\underline{\lambda}}}{(\kappa_2^{\underline{\lambda}})^{d/2}} \right| = \Theta\left(\frac{\mathrm{aft}(\underline{\lambda}) n^d}{\mathrm{aft}(\underline{\lambda})^{d/2} n^d} \right) = \Theta(\mathrm{aft}(\underline{\lambda})^{1-d/2}).$$

$\square$

We now state and prove a generalization of Theorem 6.1.3 for skew shapes of the special form $\underline{\lambda}$ as defined in Section 6.2.1.

Theorem 6.3.8. *Suppose $\underline{\lambda}^{(1)}, \underline{\lambda}^{(2)}, \ldots$ is a sequence of skew partitions, and let $X_N := X_{\underline{\lambda}^{(N)}}[\mathrm{maj}]$ be the corresponding random variables for the* maj *statistic. Then, the sequence $X_1, X_2, \ldots$ is asymptotically normal if and only if* $\mathrm{aft}(\underline{\lambda}^{(N)}) \to \infty$ *as $N \to \infty$.*

Proof. If $\mathrm{aft}(\underline{\lambda}^{(N)}) \to \infty$, the result follows immediately from Corollary 6.2.27, Corollary 6.3.7, and the fact that the odd cumulants vanish. On the other hand, if $\mathrm{aft}(\underline{\lambda}^{(N)}) \not\to \infty$, in the next section we will show that $X_1^*, X_2^*, \ldots$ has a subsequence which converges to either a discrete or uniform-sum distribution, which in either case is non-normal. □

Remark 6.3.9. Using work of Hwang–Zacharovas [44, Thm. 1.1], it would suffice to prove both directions of Theorem 6.3.8 just for the $d = 4$ case. However, the bounds we've given for κ_d^{λ} are strong enough to bound all the normalized cumulants directly, and restricting to $d = 4$ does not simplify the argument.

6.4 Uniform Sum Limit Laws for $X_{\underline{\lambda}}[\mathrm{maj}]$

The estimates from Section 6.3 apply when aft tends to ∞. We next give an analogous estimate handling the case when aft is bounded which is powerful enough to prove Theorem 6.1.6 and its generalization to block diagonal skew partitions.

Lemma 6.4.1. *Suppose $\lambda^{(N)}/\nu^{(N)} \vdash n_N$ is a sequence of skew partitions such that $\lim_{N\to\infty} n_N = \infty$ and*

$$\lim_{N\to\infty} \mathrm{aft}(\lambda^{(N)}/\nu^{(N)}) = M \in \mathbb{Z}_{\geq 0}.$$

Then for each fixed $d \in \mathbb{Z}_{\geq 1}$, we have

$$\lim_{N\to\infty} \frac{\sum_{j=1}^{n_N} j^d - \sum_{c\in\lambda^{(N)}/\nu^{(N)}} h_c^d}{M n_N^d} = 1.$$

Proof. Take N large enough so that $\mathrm{aft}(\lambda^{(N)}/\nu^{(N)}) = M$ and $n_N \gg M$. Let $m \in \lambda^{(N)}/\nu^{(N)}$ be such that $\mathrm{aft}(\lambda^{(N)}/\nu^{(N)}) = M = n_N - \mathrm{arm}(m)$ so m is the first cell in its row and column, as in the proof of Lemma 6.3.5. Consider three regions of $\lambda^{(N)}/\nu^{(N)}$:

(i) The rightmost $\mathrm{arm}(m) - M = n_N - 2M$ cells in the row of m.

(ii) The remaining leftmost M cells in the row of m.

(iii) The remaining M cells in $\lambda^{(N)}/\nu^{(N)}$.

Construct $T \in \mathrm{RSYT}(\lambda^{(N)}/\nu^{(N)})$ iteratively as in the proof of Lemma 6.3.5 as follows. First add cells in region (iii) row by row starting at the topmost row proceeding from left to right, stopping just before inserting the row of m. Next add the cells from region (ii) from left to right. Now add the remaining cells in region (iii) row by row starting at the row immediately below the row of m proceeding from left to right. Finally insert the cells from region (i) from left to right. It is easy to see that the cells in region (i) are the lowest cells in their column, from which it follows that T indeed satisfies the column and row decreasing conditions.

We now consider the contributions of regions (i)-(iii) to the quotient

$$\frac{\sum_{j=1}^{n_N} j^d - \sum_{c\in\lambda^{(N)}/\nu^{(N)}} h_c^d}{Mn_N^d}.$$

Recall that $T_c - h_c$ can be interpreted as the number of times a cell inserted after cell c was not inserted in the same hook as c. It follows that $T_c - h_c = 0$ for region (i), leaving only contributions from the $2M$ cells in regions (ii) and (iii), a bounded sum. For region (ii), we have $T_C - h_c \leq M$, so that

$$T_c^d - h_c^d = (T_c - h_c)\mathbf{h}_{d-1}(T_c, h_c) \leq (2M)dn_N^{d-1}.$$

Dividing by Mn_N^d, cells in region (ii) contribute 0 to the sum in the limit. Finally, for region (iii), we find $1 \leq h_c \leq M+1$ and $n_N - 2M + 1 \leq T_c \leq n_N$, so that for each of the M cells c in region (iii),

$$(n_N - 2M + 1)^d - (M+1)^d \leq T_c^d - h_c^d \leq n_N^d - 1^d.$$

Dividing by n_N^d, both bounds are asymptotic to 1 as $n_N \to \infty$. Adding up all M such contributions, the result follows. □

Recall from Definition 6.2.2 that we associate a block diagonal skew partition to any finite sequence of partitions $\underline{\lambda}$. The following theorem is a generalization of Theorem 6.1.6.

Theorem 6.4.2. *Let $\underline{\lambda}^{(1)}, \underline{\lambda}^{(2)}, \ldots$ be a sequence of block diagonal skew partitions such that $\lim_{N\to\infty} |\lambda^{(N)}| = \infty$ and $\operatorname{aft}(\underline{\lambda}^{(N)}/\underline{\nu}^{(N)}) = M$ is constant. Let $X_N := X_{\underline{\lambda}^{(N)}/\underline{\nu}^{(N)}}[\mathrm{maj}]$ be the corresponding random variable for the* maj *statistic. Then $X_1^*, X_2^*, \ldots$ converges in distribution to Σ_M^*.*

Proof. Using Equation (6.17) and Lemma 6.4.1, we have for $d \geq 2$ that

$$\begin{aligned}
\lim_{N\to\infty} (\kappa_d^{\underline{\lambda}^{(N)}})^* &= \lim_{N\to\infty} \frac{\kappa_d^{\underline{\lambda}^{(N)}}}{(\kappa_d^{\underline{\lambda}})^{d/2}} \\
&= \lim_{N\to\infty} \frac{(B_d/d)\left(\sum_{j=1}^{n_N} j^d - \sum_{c\in\underline{\lambda}^{(N)}} h_c^d\right)}{(B_2/2)^{d/2}\left(\sum_{j=1}^{n_N} j^2 - \sum_{c\in\underline{\lambda}^{(N)}} h_c^2\right)^{d/2}} \\
&= \lim_{N\to\infty} \frac{(B_d/d)}{(B_2/2)^{d/2}} \frac{Mn_N^d}{(Mn_N^2)^{d/2}} \\
&= \frac{(MB_d/d)}{(MB_2/2)^{d/2}}.
\end{aligned}$$

From Example 6.2.15 and the homogeneity and additivity properties of cumulants, we have

$$\begin{aligned}
(\kappa_d^{\Sigma_M})^* &= \frac{\kappa_d^{\Sigma_M}}{(\kappa_2^{\Sigma_M})^{d/2}} \\
&= \frac{(MB_d/d)}{(MB_2/2)^{d/2}}.
\end{aligned}$$

The result now follows from the equivalent formulation of Theorem 6.2.26 in terms of cumulants. □

6.5 Limiting Distribution Classification

We now give the generalization of Theorem 6.1.7 to $\underline{\lambda}$, Theorem 6.5.1, and analyze more carefully the discrete case of Theorem 6.1.7, resulting in Theorem 6.5.2.

Theorem 6.5.1. *Let $\underline{\lambda}^{(1)}, \underline{\lambda}^{(2)}, \ldots$ be a sequence of block diagonal skew partitions. Then the sequence $(X_{\underline{\lambda}^{(N)}}[\mathrm{maj}]^*)$ converges in distribution if and only if*

(i) $\mathrm{aft}(\underline{\lambda}^{(N)}) \to \infty$*; or*

(ii) $|\underline{\lambda}^{(N)}| \to \infty$ *and* $\mathrm{aft}(\underline{\lambda}^{(N)})$ *is eventually constant; or*

(iii) the distribution of $X_{\underline{\lambda}^{(N)}}[\mathrm{maj}]$ is eventually constant.

The limit law is $\mathcal{N}(0,1)$ in case (i), Σ_M^* in case (ii), and discrete in case (iii).

Proof. The backwards direction follows from Theorem 6.3.8 and Theorem 6.4.2. In the forwards direction, let $\underline{\lambda}^{(N)}$ be such a sequence where $(X_{\underline{\lambda}^{(N)}}[\mathrm{maj}]^*)$ converges in distribution. If $|\underline{\lambda}^{(N)}|$ is bounded, then there are only finitely many distinct $\underline{\lambda}^{(N)}$, forcing case (iii). If $|\underline{\lambda}^{(N)}|$ is unbounded, then we have subsequences satisfying either (i) or (ii) since the sequence converges in distribution, which from Theorem 6.3.8 and Theorem 6.4.2 gives convergence in distribution to $\mathcal{N}(0,1)$ or Σ_M^*, which are continuous, distinct distributions. The result follows. $\square$

A well-known corollary of Theorem 5.2.2 is that for partitions λ and ν of n, maj is equidistributed on $\mathrm{SYT}(\lambda)$ and $\mathrm{SYT}(\nu)$ if and only if $b(\lambda) = b(\nu)$ and the multisets $\{h_c : c \in \lambda\}$ and $\{h_d : d \in \nu\}$ are equal. These hook multisets do not entirely characterize the partition—see [43]. The following theorem gives a similar result even if we consider the standardized random variables corresponding with $X_\lambda[\mathrm{maj}]$ and $X_\nu[\mathrm{maj}]$.

Theorem 6.5.2. *Let λ and ν be partitions. Then $X_\lambda[\mathrm{maj}]^*$ and $X_\nu[\mathrm{maj}]^*$ have the same distribution if and only if* *(i) the multisets of hook lengths* $\{h_c : c \in \lambda\}$ *and* $\{h_d : d \in \nu\}$ *are equal; or*

(ii) the multisets $\{h_c : c \in \lambda\}$ *and* $\{|\lambda|\} \sqcup \{h_d : d \in \nu\}$ *are equal; or*

(iii) λ *and* ν *are each either a single row or column; or*

(iv) $\lambda, \nu \in \{(2,1),(2,2)\}$.

Moreover, case (ii) occurs if and only if, up to transposing,

(a) $\lambda = (n)$ *and* $\nu = (n-1)$ *for* $n \geq 2$*; or*

(b) $\lambda = (r+1, 1^{2r+2})$ *and* $\nu = (2^{r+1}, 1^r)$ *for* $r \geq 1$*; or*

(c) $\lambda = (s, 1^{s+2})$ *and* $\nu = (s, s, 1)$ *for* $s \geq 4$*; or*

(d) $\lambda = (3, 1^5)$ *and* $\nu = (3^2, 1)$*, or* $\lambda = (4, 1^6)$ *and* $\nu = (3^3, 1)$.

Proof. Let $n := |\lambda|$ and $m := |\nu|$. Let $f^\lambda(q) = \frac{[n]_q!}{\prod_{c\in\lambda}[h_c]}$ which is a polynomial by Theorem 5.2.2 which has constant coefficient 1. Let $f^\lambda = f^\lambda(1) = |\operatorname{SYT}(\lambda)|$. Let f^ν and $f^\nu(q)$ defined similarly.

In the backwards direction, if (i) holds, then $n = m$, $\sigma = \tau$, and $f^\lambda(q) = f^\nu(q)$, so $X_\lambda[\mathrm{maj}]^*$ and $X_\nu[\mathrm{maj}]^*$ have the same distribution. Similarly if (ii) holds $f^\lambda(q) = f^\nu(q)$, $\sigma = \tau$, and $X_\lambda[\mathrm{maj}]^*$ and $X_\nu[\mathrm{maj}]^*$ have the same distribution again. Condition (iii) holds if and only if the distributions are concentrated at a single point. For (iv), we have $f^{(2,1)}(q) = 1 + q$ and $f^{(2,2)}(q) = 1 + q^2$, so the normalized distributions are clearly equal. We will shortly see that in each of the cases (a)-(d), condition (ii) in fact holds.

In the forwards direction, suppose $X_\lambda[\mathrm{maj}]^*$ and $X_\nu[\mathrm{maj}]^*$ have the same distribution. Since $f^\lambda(q)$ has constant coefficient 1, $X_\lambda[\mathrm{maj}]$ is concentrated at a single point if and only if $f^\lambda = 1$, which occurs if and only if λ is a single row or column which is covered by case (iii). It is easy to see that $f^\lambda = 2$ if and only if $\lambda \in \{(2,1),(2,2)\}$ which is covered by case (iv).

Assume $f^\lambda, f^\nu > 2$. We claim that there are two adjacent non-zero coefficients of $f^\lambda(q)$, and simiarly for $f^\nu(q)$. To prove this, we note the constant coefficient is 1 and the linear coefficient is the number of inner corners of λ by Remark 6.2.6 which is zero if and only if λ is a rectangle with at least 2 rows and columns. Since $\lambda \neq (2,2)$, the second and third coefficient of $f^\lambda(q)$ are nonzero by Remark 6.2.6 completing the claim. Since $f^\lambda(q)$ and $f^\nu(q)$ each have constant term 1 and two adjacent non-zero coefficients, then it follows from the assumption $X_\lambda[\mathrm{maj}]^*$ and $X_\nu[\mathrm{maj}]^*$ have the same distribution that

$$f^\lambda(q) = \frac{[n]_q!}{\prod_{c\in\lambda}[h_c]_q} = \frac{[m]_q!}{\prod_{d\in\nu}[h_d]_q} = f^\nu. \tag{6.25}$$

Without loss of generality, we can assume $n \geq m$. If $n = m$, we have $\prod_{c\in\lambda}[h_c]_q = \prod_{d\in\nu}[h_d]_q$, from which it follows that the multisets of hook lengths are equal by considering multiplicities of zeroes at all primitive roots of unity as in case (i).

From here on, assume $n > m$. The multiplicity of a zero of a primitive nth root of unity in (6.25) is 0 on the right, so from the left λ must have a hook of length n so it itself a hook shape partition. Since λ is not a single row or column by the assumption $f^\lambda > 2$, we know λ does not have a cell with hook length $n-1$. Consequently, the multiplicity of a zero at a primitive $(n-1)$th root of unity in (6.25) is 1 on the left, forcing $m = n-1$ on the right. Thus (6.25) becomes

$$[m+1]_q \prod_{d\in\nu}[h_d]_q = \prod_{c\in\lambda}[h_c]_q, \tag{6.26}$$

and as before the multiset condition (ii) must hold. This completes the proof of the first statement in the theorem.

For the second statement, suppose (ii) holds, so the multisets $\{h_c : c \in \lambda\}$ and $\{|\lambda|\} \sqcup \{h_d : d \in \nu\}$ are equal. Then, $m = n-1$ and λ has a cell with hook length $|\lambda|$, so λ is a hook shape partition $(n-k, 1^k)$ for some $0 \leq k \leq n$, and

$$\{h_d : d \in \nu\} = [m-k] \sqcup [k]. \tag{6.27}$$

By transposing if necessary, we may assume $k \geq m-k$ is the maximum hook length in ν. If λ has one cell with hook length 1, then (a) holds. Otherwise, both λ and ν have precisely two cells with hook length 1, so ν is the union of two rectangles and not itself a rectangle. If ν were a hook, then it would have a hook length equal to m which would imply λ has a cell of hook length $m = n-1$ contradicting the fact that λ has two outer corners. Thus ν is not itself a hook.

Transposing ν if necessary, its first two rows are equal, say $\nu_1 = \nu_2 = s$. If $\nu'_1 = \nu'_2$, one may check that the cell furthest from the origin in the intersection of the two rectangles forming ν would be the only cell of its hook length, and that moreover its two neighbors in the intersection would each have one larger hook length, contrary to (6.27). It follows that $\nu = (s^t, 1^r)$ where $r \geq 1$, $s \geq 2$, and $t \geq 2$. We now have several cases.

- If $s = 2$, the hook lengths of ν are $\{1, \ldots, r, r+2, \ldots, r+t+1, 1, \ldots, t\}$. The "gap" between r and $r+2$ together with (6.27) forces $t = r+1$, so that $\nu = (2^{r+1}, 1^r)$ with $r \geq 1$. Here $k = r+t+1 = 2r+2$, resulting in case (b).

- If $s \geq 3$, the last two columns of ν already contain two cells with hook length 2. If $r > 1$, the first column would also have a cell with hook length 2, contradicting (6.27), so $r = 1$.

  - If $s = 3$, the hook lengths of ν are $\{1, \ldots, t, 2, \ldots, t+1, 1, 4, 5, \ldots, t+3\}$. Because of the "gap" between $t+1$ and $t+3$, this is of the form in (6.27) if and only if $t = 2$ or $t = 3$, resulting in case (d).

  - Suppose $s > 3$. If $t \geq 3$, then the final three columns of ν contain three cells with hook length 3, contradicting (6.27), so $t = 2$. The hook lengths of ν are then $\{1, 1, 2, \ldots, s-1, s+1, 2, 3, \ldots, s, s+2\}$, which is already of the form (6.27), resulting in case (c).

The reverse implications from (a)-(d) to (ii) were verified in the course of the above argument.

□

Remark 6.5.3. The proof of Theorem 6.5.2 applies more generally to arbitrary scaling factors and translations of the distributions of $X_\lambda[\text{maj}]$ and $X_\nu[\text{maj}]$, and not just those coming from means and variances.

6.6 Internal Heroes Classification

As a corollary of Stanley's formula, we know that for every partition $\lambda \vdash n \geq 1$ there is a unique tableau with minimal major index $b(\lambda)$ and a unique tableau with maximal major index $\binom{n}{2} - b(\lambda')$. These two agree for shapes consisting of one row or one column, and otherwise they are distinct. It is easy to identify these two tableaux in $\mathrm{SYT}(\lambda)$.

Definition 6.6.1.

1. The *max-maj tableau* for λ is obtained by filling the outermost, maximum length, vertical strip in λ with the largest possible numbers $|\lambda|, |\lambda| - 1, \ldots, |\lambda| - \ell(\lambda) + 1$ starting from the bottom row and going up, then filling the rightmost maximum length vertical strip containing cells not previously used with the largest remaining numbers, etc.

2. The *min-maj tableau* of λ is obtained similarly by filling the outermost, maximum length, horizontal strip in λ with the largest possible numbers $|\lambda|, |\lambda| - 1, \ldots, |\lambda| - \lambda_1 + 1$ going right to left, then filling the lowest maximum length horizontal strip containing cells not previously used with the largest remaining numbers, etc.

See Figure 6.6 for an example. Note that the max-maj tableau of λ is the transpose of the min-maj tableau of λ'.

The $q^{b(\lambda)+1}$ coefficient of $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$ can be computed as in Remark 6.2.6, resulting in the following.

Corollary 6.6.2. *We have* $[q^{b(\lambda)+1}]\,\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = 0$ *if and only if* λ *is a rectangle. If* λ *is a rectangle with more than one row and column, then* $[q^{b(\lambda)+2}]\,\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = 1$.

$$\begin{array}{llllll} 1 & 2 & 3 & 5 & 9 & \boxed{13} \\ 4 & 6 & 10 & \boxed{14} & & \\ 7 & 11 & \boxed{15} & & & \\ 8 & 12 & \boxed{16} & & & \\ \boxed{17} & & & & & \end{array}$$

(a) A max-maj tableau and its outermost vertical strip.

$$\begin{array}{llllll} 1 & 3 & 4 & 11 & \boxed{16} & \boxed{17} \\ 2 & 6 & 7 & \boxed{15} & & \\ 5 & 9 & 10 & & & \\ 8 & \boxed{13} & \boxed{14} & & & \\ \boxed{12} & & & & & \end{array}$$

(b) A min-maj tableau and its outermost horizontal strip.

Figure 6.6: Max-maj and min-maj tableau for $\lambda = (6, 4, 3, 3, 1)$.

A similar statement holds for $\mathrm{maj}(T) = \binom{n}{2} - b(\lambda') - 1$ by symmetry. Thus, $\mathrm{SYT}^{\mathrm{maj}}(q)$ has internal zeros when λ is a rectangle with at least 2 rows and columns. We will show these are the only exceptions, proving Theorem 6.1.9.

Definition 6.6.3. Let $\mathcal{E}(\lambda)$ denote the set of *exceptional* tableaux of shape λ consisting of the following elements.

(i) For all λ, the max-maj tableau for λ.

(ii) If λ is a rectangle, the min-maj tableau for λ.

(iii) If λ is a rectangle with at least two rows and columns, the unique tableau in $\mathrm{SYT}(\lambda)$ with major index equal to $\binom{n}{2} - b(\lambda') - 2$. It is obtained from the max-maj tableau of λ by applying the cycle $(2, 3, \ldots, \ell(\lambda) + 1)$, which reduces the major index by 2.

For example, $\mathcal{E}(64331)$ consists of just the max-maj tableau for 64331 in Figure 6.6a, while $\mathcal{E}(555)$ has the following three elements:

$$\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{array} \qquad \begin{array}{ccccc} 1 & 2 & 7 & 10 & 13 \\ 3 & 5 & 8 & 11 & 14 \\ 4 & 6 & 9 & 12 & 15 \end{array} \qquad \begin{array}{ccccc} 1 & 4 & 7 & 10 & 13 \\ 2 & 5 & 8 & 11 & 14. \\ 3 & 6 & 9 & 12 & 15 \end{array}$$

We prove Theorem 6.1.9 by constructing a map

$$\varphi :\ \mathrm{SYT}(\lambda) \setminus \mathcal{E}(\lambda) \longrightarrow \mathrm{SYT}(\lambda)$$

with the property

$$\mathrm{maj}(\varphi(T)) = \mathrm{maj}(T) + 1.$$

For *most* tableau T, we can find another tableau T' of the same shape such that $\mathrm{maj}(T') = \mathrm{maj}(T) + 1$ by applying some *simple cycle* to the values in T, meaning a permutation whose cycle notation is either $(i, i+1, \ldots, k-1, k)$ or $(k, k-1, \ldots, i+1, i)$ for some $i < k$. We will show there are 5 additional rules that must be added to complete the definition.

We note that technically the symmetric group S_n does not act on $\mathrm{SYT}(\lambda)$ for $\lambda \vdash n$ since this action will not generally preserve the row and column strict requirements for standard tableaux. However, S_n acts on the set of all fillings of λ using the alphabet $\{1, 2, \ldots, n\}$ by acting on the values. We will only apply permutations to tableaux after locating all values in some interval $[i, j] = \{i, i+1, \ldots, j\}$ in T. The reader is encouraged to verify that the specified permutations always maintain the row and column strict properties.

6.6.1 Rotation Rules

We next describe certain configurations in a tableau which imply that a simple cycle will increase maj by 1. Recall, the cells of a tableau are indexed by matrix notation.

Definition 6.6.4. Given $\lambda \vdash n$ and $T \in \mathrm{SYT}(\lambda)$, a *positive rotation* for T is an interval $[i, k] \subset [n]$ such that if $T' := (i, i+1, \ldots, k-1, k) \cdot T$, then $T' \in \mathrm{SYT}(\lambda)$ and there is some j for which

$$\{j\} = \mathrm{Des}(T') - \mathrm{Des}(T) \qquad \text{and} \qquad \{j-1\} = \mathrm{Des}(T) - \mathrm{Des}(T').$$

Intuitively, a positive rotation is one for which $j - 1 \in \mathrm{Des}(T)$ becomes $j \in \mathrm{Des}(T')$ and all other entries remain the same. Consequently, $\mathrm{maj}(T') = \mathrm{maj}(T) + 1$. We call j the *moving* *descent* for the positive rotation.

The positive rotations can be characterized combinatorially as follows. The proof is omitted.

Lemma 6.6.5. *An interval* $[i, k]$ *is a positive rotation for* $T \in \mathrm{SYT}(\lambda)$ *if and only if* $i < k$ *and there is some necessarily unique moving descent* j *with* $1 \le i \le j \le k \le n$ *such that*

(a) $i, \ldots, j-1$ *forms a horizontal strip,* $j-1, j$ *forms a vertical strip, and* $j, j+1, \ldots, k$ *forms a horizontal strip;*

(b) if $i < j$*, then* i *appears strictly northeast of* k *and* $i-1$ *is not in the rectangle bounding* i *and* k*;*

(c) if $i = j$*, then* $i-1$ *appears in the rectangle bounding* i *and* k*;*

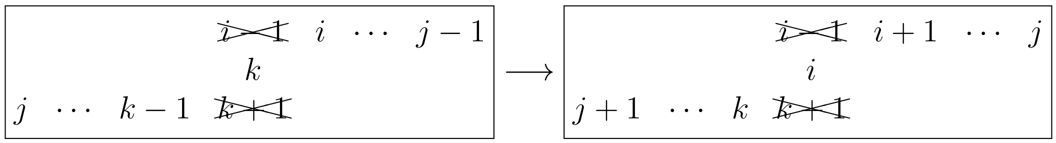
(d) if $j < k$*, then* k *appears strictly northeast of* $k-1$ *and* $k+1$ *is not in the rectangle bounding* k *and* $k-1$*; and*

(e) if $j = k$*, then* $k+1$ *appears in the rectangle bounding* k *and* $k-1$*.*

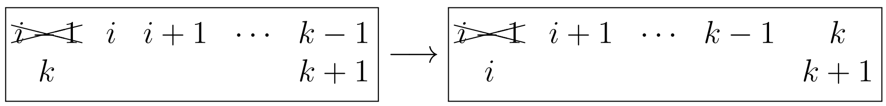
See Figure 6.7 for diagrams summarizing these conditions.

In addition to the *positive rotations* above, we can also apply *negative rotations*, which are defined exactly as in Definition 6.6.4 with $(i, i+1, \ldots, k-1, k)$ replaced by $(k, k-1, \ldots, i+1, i)$ and the rest unchanged. Combinatorially, negative rotations can be obtained from positive rotations by applying *inverse-transpose* moves, that is, by applying negative cycles $(k, k-1, \ldots, i)$ to the transpose of the configurations in Figure 6.7 and reversing the arrows. Explicitly, we have the following analogue of Lemma 6.6.5. See Figure 6.8 for the corresponding diagrams.

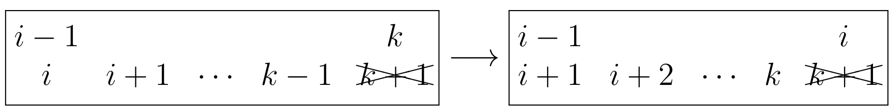
Lemma 6.6.6. *An interval* $[i, k]$ *is a negative rotation for* $T \in \mathrm{SYT}(\lambda)$ *if and only if* $i < k$ *and there is some necessarily unique moving descent* j *with* $1 \le i \le j \le k \le n$ *such that*



(a) Schematic of a positive rotation with $i < j < k$.



(b) Schematic of a positive rotation with $i < j = k$.



(c) Schematic of a positive rotation with $i = j < k$.

Figure 6.7: Summary diagrams for positive rotations.

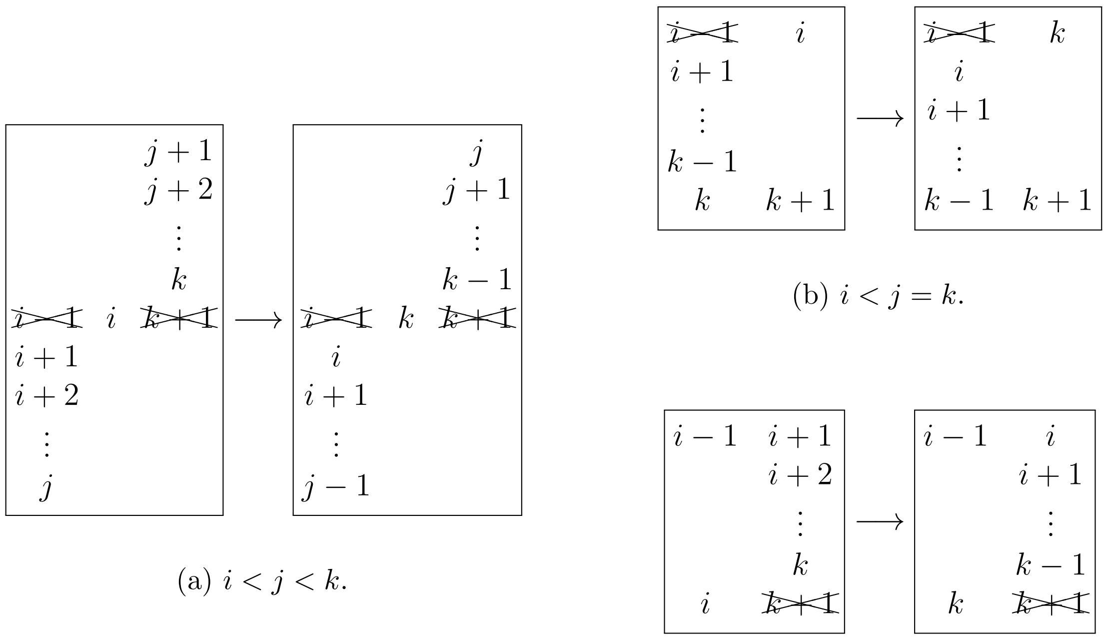
(a) *$i, \ldots, j$ forms a vertical strip, $j, j+1$ forms a horizontal strip, and $j+1, \ldots, k$ forms a vertical strip;*

(b) *if $i < j$, then $i+1$ appears strictly southwest of i and $i-1$ is not in the rectangle bounding i and $i+1$;*

(c) *if $i = j$, then $i-1$ appears in the rectangle bounding i and $i+1$;*

(d) *if $j < k$, then i appears strictly southwest of k and $k+1$ is not in the rectangle bounding i and k; and*

(e) *if $j = k$, then $k+1$ appears in the rectangle bounding i and k.*



(a) $i < j < k$.

(b) $i < j = k$.

(c) $i = j < k$.

Figure 6.8: Summary diagrams for negative rotations.

Example 6.6.7. The tableau

$$\begin{array}{ccccc} 1 & 2 & 6 & 7 & 9 \\ 3 & 4 & 8 & 13 & \\ 5 & 11 & 12 & 15 & \\ 10 & 14 & & & \end{array}$$

allows positive rotation rules with $[i, k] \in \{[5,6], [8,9], [8,10], [8,11], [9,13]\}$, and the tableau

$$\begin{array}{ccccc} 1 & 3 & 8 & 10 & 15 \\ 2 & 4 & 9 & 11 & \\ 5 & 7 & 13 & 14 & \\ 6 & 12 & & & \end{array}$$

allows negative rotation rules with $[i,k] \in \{[4,6],[6,7],[11,12]\}$.

It turns out that for the vast majority of tableaux, some rotation rule applies. For example, among the 81,081 tableaux in SYT(5442), there are only 24 (i.e., 0.03%) on which we cannot apply any positive or negative rotation rule. For example, no rotation rules can be applied to the following two tableaux:

$$\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & \\ 10 & 11 & 12 & 13 & \\ 14 & 15 & & & \end{array} \quad \text{and} \quad \begin{array}{ccccc} 1 & 2 & 3 & 8 & 12 \\ 4 & 6 & 9 & 13 & \\ 5 & 7 & 10 & 14 & \\ 11 & 15 & & & \end{array}.$$

The following lemma and its corollary give a partial explanation for why rotation rules are so common. Given a tableaux T, let $T|_{[z]}$ denote the restriction of T to those values in $[z]$.

Lemma 6.6.8. *Let $T \in \mathrm{SYT}(\lambda) \setminus \mathcal{E}(\lambda)$. Suppose z is the largest value such that $T|_{[z]}$ is contained in* maxmaj(μ) *for some μ. If $T|_{[z+1]}$ is not of the form*

$$\begin{array}{cccc} 1 & 2 & \cdots & i \\ i+1 & z+1 & & \\ i+2 & & & \\ \vdots & & & \\ z & & & \end{array}$$

then some negative rotation rule applies to T.

Proof. Since $T \notin \mathcal{E}(\lambda)$, T is not maxmaj(λ), so λ is not a one row or column shape. We have $z \geq 2$ since both two-cell tableaux are the max-maj tableau of their shape. Since maxmaj(μ) is built from successive, outermost, maximal length, vertical strips as in Figure 6.6a, the same is true of $T|_{[z]}$.

First, suppose z is not in the lowest row of $T|_{[z]}$. Let i be the value in the topmost corner cell in $T|_{[z]}$ which is strictly below z. Let $j \geq i$ be the bottommost cell in the vertical strip of $T|_{[z]}$ which contains i. See Figure 6.9a. We verify the conditions of Lemma 6.6.6, so the negative $[i, z]$-rotation rule applies with moving descent j. By construction, $i, \ldots, j$ forms a vertical strip, $j, j+1$ forms a horizontal strip, and $j+1, \ldots, z$ forms a vertical strip. If $i < j$, then since i is a corner cell, $i+1$ appears strictly southwest of i, and $i-1$ is above both i and $i+1$ so $i-1$ is not in the rectangle bounding i and $i+1$. If $i = j$, we see that $i-1$ appears in the rectangle bounded by i and $i+1$. We also see that i appears strictly southwest of z, and $z+1$ is not in the rectangle bounding i and z since i is a topmost corner and z is maximal.

Now suppose z is in the lowest row of $T|_{[z]}$. In this case, $T|_{[z]}$ is the max-maj tableau of its shape, so that $z < |\lambda|$ and $z+1$ exists in T since $T \notin \mathcal{E}(\lambda)$. By maximality of z, $z+1$ cannot be in row 1 or below z. Let $i < z$ be the value in the the rightmost cell of $T|_{[z]}$ in the row immediately above $z+1$. See Figure 6.9b. We check that the negative $[i, z]$-rotation rule applies with moving descent $j = z$ using the conditions in Lemma 6.6.6. By construction, $i, \ldots, z$ forms a vertical strip. Since $z+1$ is not below z, we see that $z, z+1$ forms a horizontal strip. Since $z+1$ is in the row below i, $i+1$ appears strictly southwest of i. We also see that $z+1$ appears in the rectangle bounded by i and z by choice of i. It remains to show that $i-1$ is not in the rectangle bounding i and $i+1$. Suppose to the contrary that $i-1$ is in the rectangle bounding i and $i+1$. Then i would have to be in row 1 by choice of $i < z$. Consequently $i+1$ is in row 2 and strictly west of i, forcing $i-1$ to be in row 1 also. It follows from the choice of z that $T|_{[i]}$ is a single row, the values $i, i+1, \ldots z$ form a vertical strip, and $T|_{[z+1]}$ is of the above forbidden form, giving a contradiction. $\square$

Corollary 6.6.9. *If $T \in \mathrm{SYT}(\lambda) \setminus \mathcal{E}(\lambda)$ and $1 \in \mathrm{Des}(T)$, then some negative rotation rule applies to T.*

Proof. Let z be as in Lemma 6.6.8. Clearly $z \geq 2$ and $T|_{[2]}$ is a single column, so $T|_{[z+1]}$ cannot possibly be of the forbidden form. $\square$

$$\begin{array}{ccc} 1 & 3 & 6 \\ 2 & 4 & 7 \\ 5 & 8 & 11 \\ 9 & & \\ 10 & & \end{array} \quad \longrightarrow \quad \begin{array}{ccc} 1 & 3 & 6 \\ 2 & 4 & 10 \\ 5 & 7 & 11 \\ 8 & & \\ 9 & & \end{array}$$

(a) For the tableau on the left above, $i = 7$ and $z = 10$ since $T|_{[10]}$ is the max-maj tableau of shape 33211, 10 is in the lowest row, 11 is in row 3, and 7 is the largest value in $T|_{[10]}$ in row 2. Apply the negative rotation $(10, 9, 8, 7)$ to get the tableau on the right, and observe maj has increased by 1. The moving descent is $j = z = 10$.

$$\begin{array}{cccc} 1 & 3 & 6 & 11 \\ 2 & 4 & 7 & 12 \\ 5 & 8 & & \\ 9 & 13 & & \\ 10 & & & \end{array} \quad \longrightarrow \quad \begin{array}{cccc} 1 & 3 & 6 & 10 \\ 2 & 4 & 7 & 11 \\ 5 & 12 & & \\ 8 & 13 & & \\ 9 & & & \end{array}$$

(b) For the tableau on the left above, $i = 8$ and $z = 12$ since $T|_{[12]}$ is contained the max-maj tableau of shape 44322, 12 is not in the lowest row, 8 is in the closet corner to 12 in $T|_{[12]}$ and below 12. Apply the negative rotation $(12, 11, 10, 9, 8)$ to get the tableau on the right, and observe maj has increased by 1. The moving descent is $j = 10$.

Figure 6.9: Examples of the negative rotations obtained from Lemma 6.6.8.

We also have the following variation on Lemma 6.6.8. It is based on finding the largest value q such that $T|_{[q]}$ is contained in an exceptional tableau of type (iii). The proof is again a straightforward verification of the conditions in Lemma 6.6.6, and is omitted.

Lemma 6.6.10. *Let $T \in \mathrm{SYT}(\lambda) \setminus \mathcal{E}(\lambda)$. Suppose the initial values of T are of the form*

$$\begin{array}{cc} 1 & 2 \\ 3 & p+1 \\ 4 & \vdots \\ \vdots & q \\ \vdots & \xcancel{q+1} \\ p & \end{array} \qquad \textit{or} \qquad \begin{array}{cccccc} 1 & 2 & \ell+1 & \cdots & \vdots & p+1 \\ 3 & z+1 & \vdots & \vdots & \vdots & \vdots \\ 4 & z+2 & \vdots & \vdots & \vdots & q \\ \vdots & \vdots & \vdots & \vdots & \vdots & \xcancel{q+1} \\ z & \ell & m & \cdots & p & \end{array} .$$

In either case, the $[p, q]$-negative rotation rule applies to T.

6.6.2 Initial Block Rules

Here we describe a collection of five additional *block rules* which may apply to a tableau that is not in the exceptional set. In each case, if the rule applies, then we specify a permutation of the entries so that we either add 1 into the descent set and leave the other descents unchanged, or we add 1 into the descent set, increase one existing descent by 1, and decrease one existing descent by 1. Thus, maj will increase by 1 in all cases. While these additional rules are certainly not uniquely determined by these criteria, they are also not arbitrary.

Example 6.6.11. For a given $T \in \mathrm{SYT}(\lambda)$, one may consider all $T' \in \mathrm{SYT}(\lambda)$ where $\mathrm{maj}(T') = \mathrm{maj}(T) + 1$. If $T' = \sigma \cdot T$ where σ is a simple cycle, then one of the rotation rules may apply to T. Table 6.2 summarizes five particular T for which *no* rotation rules apply. These examples have guided our choices in defining the block rules. In all but one of these examples, there is a unique T' with $\mathrm{maj}(T') = \mathrm{maj}(T) + 1$, though in the third case there are two such T', one of which ends up being easier to generalize.

In the remainder of this subsection, we describe the block rules, abbreviated B-rules. Then, we prove that if no rotation rules are possible for a tableau then either it is in the exceptional set or we can apply one of the B-rules. The B-rules cover disjoint cases so no tableau admits

Tableau T	Tableaux T'	σ	Block rule
$\begin{smallmatrix} 1&2&3&7 \\ 4&5&6&8 \end{smallmatrix}$	$\begin{smallmatrix} 1&3&4&6 \\ 2&5&7&8 \end{smallmatrix}$	$(2,3,4)(6,7)$	B1
$\begin{smallmatrix} 1&2&3&4 \\ 5&6&7 \end{smallmatrix}$	$\begin{smallmatrix} 1&3&4&7 \\ 2&5&6 \end{smallmatrix}$	$(2,3,4,7,6,5)$	B2
$\begin{smallmatrix} 1&2&3 \\ 4&6 \\ 5&7 \end{smallmatrix}$	$\begin{smallmatrix} 1&3&6 \\ 2&4 \\ 5&7 \end{smallmatrix}, \begin{smallmatrix} 1&4&5 \\ 2&6 \\ 3&7 \end{smallmatrix}$	$(2,3,6,4), (2,4)(3,5)$	B3, —
$\begin{smallmatrix} 1&2&7 \\ 3&5&8 \\ 4&6&9 \\ 10 \end{smallmatrix}$	$\begin{smallmatrix} 1&4&8 \\ 2&5&9 \\ 3&6&10 \\ 7 \end{smallmatrix}$	$(2,4,3)(7,8,9,10)$	B4
$\begin{smallmatrix} 1&2 \\ 3&5 \\ 4&6 \\ 7 \end{smallmatrix}$	$\begin{smallmatrix} 1&5 \\ 2&6 \\ 3&7 \\ 4 \end{smallmatrix}$	$(2,5,6,7,4,3)$	B5

Table 6.2: Some tableaux $T \in \mathrm{SYT}(\lambda)$ together with all $T' = \sigma \cdot T \in \mathrm{SYT}(\lambda)$ where $\mathrm{maj}(T') = \mathrm{maj}(T) + 1$. See Definition 6.6.13 for an explanation of the final column.

more than one block rule. To state the B-rules precisely, assume $T \in \mathrm{SYT}(\lambda) \setminus \mathcal{E}(\lambda)$ and no rotation rule applies.

Notation 6.6.12. Let c be the largest possible value such that $T|_{[c]}$ is contained in the min-maj tableau of a rectangle shape with a columns and b rows. Consequently, the first a numbers in row i, $1 \le i \le b-1$, of T are $(i-1)a+1, \dots, ia$, and row b begins with $(b-1)a+1, (b-1)a+2, \dots, c$.

Since $1 \notin D(T)$ and $T \notin \mathcal{E}(\lambda)$, we know $a, b \ge 2$. If $c+1$ is in T, then it must be either in position $(1, a+1)$ or $(b+1, 1)$. If $c = ab$, then $c < |\lambda|$ since $T \notin \mathcal{E}(\lambda)$, otherwise $c = |\lambda|$ is possible. For example, the tableaux

$$\begin{smallmatrix} 1&2&3&4&5&16 \\ 6&7&8&9&10&17 \\ 11&12&13&14&15 \end{smallmatrix}, \quad \begin{smallmatrix} 1&2&3&4&5 \\ 6&7&8&9&10 \\ 11&12&13 \end{smallmatrix}, \quad \begin{smallmatrix} 1&2&3&4 \\ 5&9 \\ 6&10 \\ 7 \\ 8 \end{smallmatrix}, \quad \begin{smallmatrix} 1&2&7&10 \\ 3&5&8&11 \\ 4&6&9&12 \\ 13 \end{smallmatrix}$$

have (a, b, c) equal to $(5,3,15)$, $(5,3,13)$, $(4,2,5)$, and $(2,2,3)$, respectively.

Definition 6.6.13. Using the notation (a, b, c), we identify the *block rules* with required assumptions as follows. See Figure 6.10 for summary diagrams.

- **Rule B1**: Assume $c = ab$, $T_{(1,a+1)} = c+1$, $T_{(2,a+1)} = c+2$, and $a < c-2$. In this case, we perform the rotations $(2,\dots,a+1)$ and $(c,c+1)$ which are sufficiently separated by hypothesis. Then, $1, a+1$ and c become descents, and a and $c+1$ are no longer descents, so the major index is increased by 1. The B1 rule is illustrated here with $a = 5$, $b = 3$:

$$\text{B1:}\quad \begin{array}{cccccc} 1 & 2 & 3 & 4 & \boxed{5} & \boxed{16} \\ 6 & 7 & 8 & 9 & 10 & 17 \\ 11 & 12 & 13 & 14 & 15 & \end{array} \qquad \begin{array}{cccccc} \boxed{1} & 3 & 4 & 5 & \boxed{6} & \boxed{15} \\ 2 & 7 & 8 & 9 & 10 & 17 \\ 11 & 12 & 13 & 14 & 16 & \end{array}$$

The boxed numbers represent descents of the tableau on the left/right that are not descents of the tableau on the right/left. The elements not shown (i.e,. $18, 19, \dots, |\lambda|$) can be in any position.

- **Rule B2**: Assume $a \geq 2$ and $c < ab$ so there exists a $1 \leq k < a$ such that $T_{(b,k)} = c$ and $T_{(b,k+1)} \neq c+1$. In this case, we perform the rotation $(2, 3, \dots, a, 2a, 3a, \dots, a(b-1), c, c-1, \dots, c-k+1 = a(b-1)+1, a(b-2)+1, \dots, 2a+1, a+1)$ around the perimeter of $T|_{[c]}$. Now 1 becomes a descent, and the other descents stay the same so the major index again increases by 1. The B2 rule is illustrated by the following (here $a = 5$, $b = 2$ and $k = 3$):

$$\text{B2:}\quad \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & \cancel{14} & \end{array} \qquad \begin{array}{ccccc} \boxed{1} & 3 & 4 & 5 & 10 \\ 2 & 7 & 8 & 9 & 13 \\ 6 & 11 & 12 & \cancel{14} & \end{array}$$

The crossed out number 14 means that 14 is not in position $(3,4)$: it can either be in positions $(1,6)$ or $(4,1)$, or it can be that $\lambda = 553$. Again, the numbers $15, \dots, |\lambda|$ can be anywhere in T.

- **Rule B3**: Assume $a \geq 3$, $c = a+1$, and there exists $k \geq 2$ such that $T_{(2,2)} = a+k+1$, $T_{(3,2)} = a+k+2$, and for all $i \in \{1, 2, \ldots, k\}$ we have $T_{(i+1,1)} = a+i$. Then we apply the rotation $(2, 3, \ldots, a, a+k+1, a+1)$. Now 1 becomes a descent, and the rest of the descent set is unchanged so the major index again increases by 1. The B3 rule is illustrated by the following (here $a = 4$, $k = 4$):

$$\text{B3:} \quad \begin{array}{llll} 1 & 2 & 3 & 4 \\ 5 & 9 & & \\ 6 & 10 & & \\ 7 & & & \\ 8 & & & \end{array} \qquad \begin{array}{llll} \boxed{1} & 3 & 4 & 9 \\ 2 & 5 & & \\ 6 & 10 & & \\ 7 & & & \\ 8 & & & \end{array}$$

- **Rule B4**: Assume that $a = 2$, $c = 3$, and there exists $k \geq 2$ such that $\{3, 4, \ldots, k+1\}$ appear in column 1 of T, $\{k+2, k+3, \ldots, 2k\}$ appear in column 2 in T. Further assume that the set $\{2k+1, 2k+2, \ldots, 3k\}$ appears in column 3, $\{3k+1, 3k+2, \ldots, 4k\}$ appears in column 4, etc., until column l for some $l > 2$ and $T_{(k+1,1)} = kl+1$ and $T_{(k+1,2)} \neq kl+2$. In this case, we can perform the two rotations $(k+1, k, \ldots, 3, 2)$ and $(k(l-1)+1, k(l-1)+2, \ldots, kl, kl+1)$. Now 1, $k+1$ and $k(l-1)$ enter the descent set, and k and $k(l-1)+1$ leave it, so the major index increases by 1. The B4 rule is illustrated by the following (here $k = 3$ and $l = 4$):

$$\text{B4:} \quad \begin{array}{llll} 1 & 2 & 7 & \boxed{10} \\ \boxed{3} & 5 & 8 & 11 \\ 4 & 6 & 9 & 12 \\ 13 & \cancel{14} & & \end{array} \qquad \begin{array}{llll} \boxed{1} & \boxed{4} & 7 & 11 \\ 2 & 5 & 8 & 12 \\ 3 & 6 & \boxed{9} & 13 \\ 10 & \cancel{14} & & \end{array}$$

- **Rule B5**: Assume that $a = 2$, $c = 3$, and there exists $k > 3$ such that $\{3, 4, \ldots, k\}$ appear in column 1 of T, $\{k+1, k+2, \ldots, 2k-2\}$ appear in column 2 in T. Furthermore, assume $T_{(k,1)} = 2k-1$ and $T_{(k,2)} \neq 2k$. Then apply the cycle $(k, k-1, \ldots, 3, 2, k+1, k+2, \ldots, 2k-1)$ to T. Now 1 becomes a descent, and the rest of the descent set remains unchanged, so the major index increases by 1. The B5 rule is illustrated by the following (here $k = 5$):

$$\text{B5:} \quad \begin{array}{cc} 1 & 2 \\ 3 & 6 \\ 4 & 7 \\ 5 & 8 \\ 9 & \cancel{10} \end{array} \quad \begin{array}{cc} \boxed{1} & 6 \\ 2 & 7 \\ 3 & 8 \\ 4 & 9 \\ 5 & \cancel{10} \end{array}$$

Lemma 6.6.14. *If $T \in \mathrm{SYT}(\lambda)$, $T \notin \mathcal{E}(\lambda)$, and $1, 2 \notin D(T)$, then either some rotation rule applies to T or a B1, B2 or B3 rule applies.*

Proof. Let c be the largest possible value such that $T|_{[c]}$ is contained in the min-maj tableau of a rectangle shape with a columns and b rows, as in the definition of the block moves. Since $1, 2 \notin D(T)$ and $T \notin \mathcal{E}(\lambda)$, we know $1, 2, 3$ are in the first row of T so $a \geq 3$, $b \geq 2$, and $a + 2 \leq |\lambda|$. By construction, we have $T_{(2,1)} = a + 1$ and $a + 2$ must appear in position $(1, a+1)$, $(2, 2)$, or $(3, 1)$ in T.

Case 1: $T_{(1,a+1)} = a + 2$**.** Observe that

$$T|_{[a+2]} = \begin{array}{cccccc} 1 & 2 & 3 & \cdots & a & a+2 \\ a+1 & & & & & \end{array}$$

and $z \geq a + 2$. Consequently, $T|_{[z+1]}$ cannot be of the form forbidden by Lemma 6.6.8, so a negative rotation rule applies.

Case 2: $T_{(2,2)} = a+2$**.** First suppose $c = ab$, then $T_{(1,a+1)} = c+1$ by choice of c. Now consider the two subcases, $T_{(2,a+1)} = c + 2$ and $T_{(2,a+1)} \neq c + 2$. In the former case, as in Figure 6.10a, the B1 rule applies to T. In the latter case, one may check that an $[i, c+1]$-positive rotation rule applies to T where $i = T_{(b,1)}$. On the other hand, if $c < ab$, then a B2 rule applies to T as in Figure 6.10b.

Case 3: $T_{(3,1)} = a + 2$**.** Let $k = \min\{j \geq a+2 \mid j \notin D(T)\}$ so $T_{(k+1,1)} = a + k$ and $T_{(k+2,1)} \neq a+k+1$, Since $T \notin \mathcal{E}(\lambda)$, we know $a+k+1$ exists in T either in position $(1, a+1)$ or $(2, 2)$, so $T|_{[a+k+1]}$ looks like

$$\begin{array}{cccccc} 1 & 2 & 3 & \cdots & a & a+k+1 \\ a+1 & & & & & \\ a+2 & & & & & \\ \vdots & & & & & \\ a+k & & & & & \end{array} \quad \text{or} \quad \begin{array}{ccccc} 1 & 2 & 3 & \cdots & a \\ a+1 & a+k+1 & & & \\ a+2 & & & & \\ \vdots & & & & \\ a+k & & & & \end{array} .$$

If $T_{(1,a+1)} = a + k + 1$, then Lemma 6.6.8 shows that a negative rotation rule applies to T. On the other hand, if $T_{(2,2)} = a + k + 1$, then observe that either a B3 move applies or the rotation $(a + k, a + k + 1)$ applies to T, depending on whether $T_{(3,2)} = a + k + 2$ or not.

$\square$

Lemma 6.6.15. *If* $T \in \mathrm{SYT}(\lambda)$*,* $T \notin \mathcal{E}(\lambda)$*,* $1 \notin D(T)$*, and* $2 \in D(T)$*, then either some rotation rule applies to* T *or a B1, B2, B4 or B5 rule applies.*

Proof. Let $k = \min\{j \geq 3 \mid j \notin D(T)\}$ so the consecutive sequence $[3, k]$ appears in the first column of T and $k + 1$ does not. By definition of k and the fact that $T \notin \mathcal{E}(\lambda)$, T must have $k + 1$ in position $(1, 3)$ or $(2, 2)$. If $T_{(1,3)} = k + 1$, then a negative rotation rule holds by Lemma 6.6.8.

Assume $T_{(2,2)} = k + 1$. Let ℓ be the maximum value such that $[k + 1, \ell]$ appears as a consecutive sequence in column 2 of T. If $\ell < 2(k - 1)$, then the negative rotation rule for $(\ell, \ell - 1, \ldots, k)$ applies to T by the first case of Lemma 6.6.10.

If $\ell = 2(k - 1)$ and $T_{(1,3)} = \ell + 1$, let m be the maximum value such that $[\ell + 1, m]$ appears as a consecutive sequence in column 3 of T. We subdivide on cases for m again. If $m < 3(k-1)$, then the negative rotation rule $(m, m-1, \ldots, \ell)$ applies to T by the second case of Lemma 6.6.10. If $m = 3(k-1)$, we consider the maximal sequence of columns containing a consecutive sequence in rows $[1, k-1]$ to the right of column 2 until one of two conditions hold

$$\boxed{\begin{array}{ccccc} 1 & 2 & \ell+1 & \cdots & \vdots \\ 3 & k+1 & \vdots & \vdots & p \\ \vdots & \vdots & \vdots & \ddots & \xcancel{p+1} \\ k & \ell & m & \cdots & \end{array}} \qquad \boxed{\begin{array}{ccccc} 1 & 2 & \ell+1 & \cdots & \\ 3 & k+1 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ k & \ell & m & \cdots & p \\ p+1 & & & & \end{array}}$$

In the first picture, $T|_{[p]}$ is not a rectangle, so we may apply a negative rotation by the second case of Lemma 6.6.10, so consider the second picture. In the second picture, $T|_{[p]}$ is a rectangle and we know $p+1$ exists in T since $T|_{[p]}$ is an exceptional tableau for a rectangle shape. If $p+2$ is in row k, column 2, a negative rotation rule applies. if $p+2$ is not in row k, column 2, then a B4-move applies.

Finally, consider the case $\ell = 2(k-1)$ and $T_{(k,1)} = \ell + 1$. If $T_{(k,2)} \neq \ell + 2$ and $k > 3$, then a B5-move applies. If $T_{(k,2)} = \ell + 2$ and $k > 3$, then the rotation $(\ell, \ell+1)$ applies to T since $\ell - 1$ is above ℓ. If $T_{(k,2)} = \ell + 2$ and $k = 3$, then $\ell = 4 = T_{(2,2)}$ and $T_{(3,1)} = 5$ so T contains

$$\begin{array}{cc} 1 & 2 \\ 3 & 4 \\ 5 & \end{array}.$$

In this case, consider the subcases $c = ab$ or $c < ab$ with $a = 2$. If $c = ab$, then $T_{(1,3)} = c + 1$ since $T \notin \mathcal{E}(\lambda)$. Either a B1-move applies if $T_{(2,3)} = c + 2$ and a $(c, c+1)$ rotation applies otherwise. On the other hand, if $c < ab$ then a B2-rule applies. □

Proof of Theorem 6.1.9. Given any $T \in \mathrm{SYT}(\lambda) \setminus \mathcal{E}(\lambda)$, we define $\varphi(T)$ with the property $\mathrm{maj}(\varphi(T)) = \mathrm{maj}(T) + 1$. If $1 \in D(T)$, define $\varphi(T) = (z, z-1, \ldots, i)T$ as in Corollary 6.6.9. If $1, 2 \notin D(T)$, then Lemma 6.6.14 applies, so define $\varphi(T)$ using the specific B1, B2, B3 or rotation rule identified in the proof of that lemma. If $1 \notin D(T)$ and $2 \in D(T)$, then

Lemma 6.6.15 applies, so define $\varphi(T)$ using the specific B1, B2, B4, B5, or negative rotation rule identified in the proof of that lemma. These rules cover all possible cases. $\square$

An *inverse-transpose* block rule is a block rule obtained from transposing the diagrams in Figure 6.10 and reversing the arrows.

Definition 6.6.16. As sets, let $P(\lambda)$ and $Q(\lambda)$ be either $\mathrm{SYT}(\lambda) \setminus \{\mathrm{minmaj}(\lambda), \mathrm{maxmaj}(\lambda)\}$ if λ is a rectangle with at least two rows and columns, or $\mathrm{SYT}(\lambda)$ otherwise.

- (Strong SYT Poset) Let $P(\lambda)$ be the partial order with covering relations given by rotations, block rules, and inverse-transpose block rules increasing maj by 1.

- (Weak SYT Poset) Let $Q(\lambda)$ be the partial order with covering relations given by $S \prec T$ if $\varphi(S) = T$ or $\varphi(T') = S'$ where S', T' are the transpose of S, T, respectively.

Corollary 6.6.17. *As posets, $P(\lambda)$ and $Q(\lambda)$ are ranked with a unique minimal and maximal element. If λ is not a rectangle, the rank function is given by $rk(T) = \mathrm{maj}(T) - b(\lambda)$. If λ is a rectangle with at least 2 rows and columns, then the rank function is given by $rk(T) = \mathrm{maj}(T) - b(\lambda) - 2$.*

Proof. By Corollary 6.6.2, $P(\lambda)$ and $Q(\lambda)$ have a single element of minimal maj and of maximal maj. Any element T besides these is covered by $\varphi(T)$ and covers $\varphi(T')'$, so is not maximal or minimal. By construction maj increases by 1 under covering relations. The result follows. $\square$

In Figure 6.11, we show an example of both the Weak SYT Poset and the Strong SYT poset for $\lambda = (3, 2, 1)$. More examples of these partial orders are given at https://sites.math.washington.edu/~billey/papers/syt.posets.

6.7 Conjectured Deviations from Unimodality and Log-Concavity

We conjecture that almost all of the polynomials of the form $\mathrm{SYT}^{\mathrm{maj}}(q)$ are unimodal and log-concave. In this section, we give specific classifications of the deviations of each of these properties. In the rare cases where unimodality or log-concavity fails, it only seems to happen the at the very beginning and end of the sequence of coefficients or near the middle coefficient.

Recall that a polynomial $P(q) = \sum_{i=0}^{n} c_i q^i$ is *unimodal* if

$$c_0 \leq c_1 \leq \cdots \leq c_j \geq c_{j+1} \geq \cdots \geq c_n$$

for some j, and $P(q)$ is *log-concave* if $c_i^2 \geq c_{i-1}c_{i+1}$ for all integers $0 < i < n$. A polynomial with nonnegative coefficients which is log-concave and has no internal zero coefficients is necessarily unimodal [90]. By Theorem 6.1.9, we know exactly where internal zeros occur so log-concavity would imply unimodality in these cases.

We say $P(q)$ is *nearly unimodal* if instead

$$c_0 \leq c_1 \leq \cdots \leq c_j, c_{j+1} = c_j - 1 < c_{j+2} \leq \cdots \leq c_{\lfloor \frac{n}{2} \rfloor}$$

for some j and $P(q)$ has symmetric coefficients. Also, a symmetric polynomial $P(q)$ is *nearly log-concave* if $c_i^2 \geq c_{i-1}c_{i+1}$ for all $1 < i < \lfloor \frac{n}{2} \rfloor - 1$.

Conjecture 6.7.1. *The polynomial* $\mathrm{SYT}^{\mathrm{maj}}(q)$ *is unimodal if* λ *has at least* 4 *corners. If* λ *has* 3 *corners or fewer, then* $\mathrm{SYT}^{\mathrm{maj}}(q)$ *is unimodal except when* λ *or* λ' *is among the following partitions:*

1. *Any partition of rectangle shape that has more than one row and column.*

2. *Any partition of the form* $(k, 2)$ *with* $k \geq 4$ *and* k *even.*

3. *Any partition of the form* $(k, 4)$ *with* $k \geq 6$ *and* k *even.*

4. *Any partition of the form* $(k, 2, 1, 1)$ *with* $k \geq 2$ *and* k *even.*

5. *Any partition of the form* $(k, 2, 2)$ *with* $k \geq 6$.

6. *Any partition on the list of 40 special exceptions:*

$$(3,3,2),(4,2,2),(4,4,2),(4,4,1,1),(5,3,3),(7,5),(6,2,1,1,1,1),$$
$$(5,5,2),(5,5,1,1),(5,3,2,2),(4,4,3,1),(4,4,2,2),(7,3,3),(8,6),(6,6,2),$$
$$(6,6,1,1),(5,5,2,2),(5,3,3,3),(4,4,4,2),(11,5),(10,6),(9,7),(7,7,2),$$
$$(7,7,1,1),(6,6,4),(6,6,1,1,1,1),(6,5,5),(5,5,3,3),(12,6),(11,7),(10,8),$$
$$(15,5),(14,6),(11,9),(16,6),(12,10),(18,6),(14,10),(20,6),(22,6).$$

Conjecture 6.7.1 was checked for all partitions up to size $n = 50$. Each of the Case 2 families $(k, 2)$, $(k, 4)$, or $(k, 2, 1, 1)$ have a relatively simple set of hook lengths so explicit formulas can be derived for the coefficients of $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$. We have found explicit proofs of near unimodality for each of these cases. They are related to known integer sequences [68, A266755] and [68, A008642] with nice generating functions. Furthermore, the Case 2 families are all nearly unimodal as well as 20 of the special exceptions. All rectangles with at least 2 rows and columns are nearly unimodal for $30 \leq n \leq 100$. We conjecture this trend also continues, hence the claim that all coefficients $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$ are close to unimodal. The Case 3 family of the form $(k, 2, 2)$ is a bit further from being unimodal. The proof of the following result is omitted.

Proposition 6.7.2. *If* $\lambda = (k, 2, 2)$ *for any positive integer* k*, then the maximal coefficient of* $q^{-b(\lambda)}\,\mathrm{SYT}^{\mathrm{maj}}(q)$*, say* c_j*, satisfies the equation* $c_j = c_{j+1} + \mathrm{floor}(k/6) + I(4 = (k \bmod 6))$ *and* $c_0 \leq c_1 \leq \cdots \leq c_j$ *and* $j+1$ *is the median nonzero coefficient. Here* I *is an indicator function which is 1 if true and 0 if false.*

Log-concavity for the polynomials $\mathrm{SYT}_{\lambda}^{\mathrm{maj}}(q)$ appears to be harder to characterize. There are examples of partitions with even 5 corners which are not log-concave. For example $f^{\lambda}(q)$ for $\lambda = (9, 9, 7, 7, 5, 5, 3, 3, 2)$ is nearly log-concave but $c_1^2 = 4^2 = 16 < 17 = c_0 c_2$. The only deviation occurs at $i = 1$. Thus, we summarize what we have observed in the following conjecture.

Conjecture 6.7.3. *The polynomials* $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$ *are almost always log-concave for partitions* $\lambda \vdash n$ *for large* n.

This conjecture is based on the fact that the normal distribution is log-concave and the following evidence. The approximate probability that a uniformly chosen partition of n has the log-concave property $\mathbb{P}(\mathrm{LC})$ and the corresponding probability for the nearly log-concave property $\mathbb{P}(\mathrm{NLC})$ is given in the following table:

n	30	40	50
$\mathbb{P}(\mathrm{LC})$	0.6734475	0.7876426	0.8753587
$\mathbb{P}(\mathrm{NLC})$	0.8003212	0.9204832	0.9688140

.

$$\boxed{\begin{array}{ccccc} 1 & 2 & \cdots & a & ab+1 \\ a+1 & a+2 & \cdots & 2a & ab+2 \\ \vdots & \vdots & \ddots & \vdots & \\ a(b-1)+1 & a(b-1)+2 & \cdots & ab & \end{array}}$$

$$\downarrow$$

$$\boxed{\begin{array}{ccccc} 1 & 3 & \cdots & a+1 & ab \\ 2 & a+2 & \cdots & 2a & ab+2 \\ \vdots & \vdots & \ddots & \vdots & \\ a(b-1)+1 & a(b-1)+2 & \cdots & ab+1 & \end{array}}$$

(a) B1.

$$\boxed{\begin{array}{ccccccc} 1 & 2 & \cdots & \cdots & \cdots & \cdots & a \\ a+1 & a+2 & \cdots & \cdots & \cdots & \cdots & 2a \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a(b-2)+1 & a(b-2)+2 & \cdots & \cdots & \cdots & \cdots & a(b-1) \\ a(b-1)+a & a(b-1)+2 & \cdots & c-1 & c & \cdots & \cancel{ab} \end{array}}$$

$$\downarrow$$

$$\boxed{\begin{array}{ccccccc} 1 & 3 & \cdots & \cdots & \cdots & \cdots & 2a \\ 2 & a+2 & \cdots & \cdots & \cdots & \cdots & 3a \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a(b-3)+1 & a(b-2)+2 & \cdots & \cdots & \cdots & \cdots & c \\ a(b-2)+1 & a(b-1)+1 & \cdots & c-2 & c-1 & \cdots & \cancel{ab} \end{array}}$$

(b) B2.

$$\boxed{\begin{array}{ccccc} 1 & 2 & \cdots & a-1 & a \\ a+1 & a+k+1 & & & \\ a+2 & a+k+2 & & & \\ \vdots & & & & \\ a+k & & & & \end{array}} \longrightarrow \boxed{\begin{array}{ccccc} 1 & 3 & \cdots & a & a+k+1 \\ 2 & a+1 & & & \\ a+2 & & & & \\ \vdots & & & & \\ a+k & & & & \end{array}}$$

(c) B3.

$$
\boxed{\begin{array}{ccccc}
1 & 2 & 2k+1 & \cdots & k(\ell-1)+1 \\
3 & k+2 & 2k+2 & \cdots & k(\ell-1)+2 \\
4 & k+3 & 2k+3 & \cdots & k(\ell-1)+3 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
k+1 & 2k & 3k & \cdots & k\ell \\
k\ell+1 & & & & \\
\xcancel{k\ell+2} & & & &
\end{array}}
$$

$$\downarrow$$

$$
\boxed{\begin{array}{ccccc}
1 & k+1 & 2k+1 & \cdots & k(\ell-1)+2 \\
2 & k+2 & 2k+2 & \cdots & k(\ell-1)+3 \\
3 & k+3 & 2k+3 & \cdots & k(\ell-1)+4 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
k & 2k & 3k & \cdots & k\ell+1 \\
k(\ell-1)+1 & & & & \\
\xcancel{k\ell+2} & & & &
\end{array}}
$$

(d) B4.

$$
\boxed{\begin{array}{cc}
1 & 2 \\
3 & k+1 \\
4 & k+2 \\
\vdots & \vdots \\
k-1 & 2k-3 \\
k & 2k-2 \\
2k-1 & \xcancel{2k}
\end{array}}
\longrightarrow
\boxed{\begin{array}{cc}
1 & k+1 \\
2 & k+2 \\
3 & k+3 \\
\vdots & \vdots \\
k-2 & 2k-2 \\
k-1 & 2k-1 \\
k & \xcancel{2k}
\end{array}}
$$

(e) B5.

Figure 6.10: Summary diagrams for block rules.

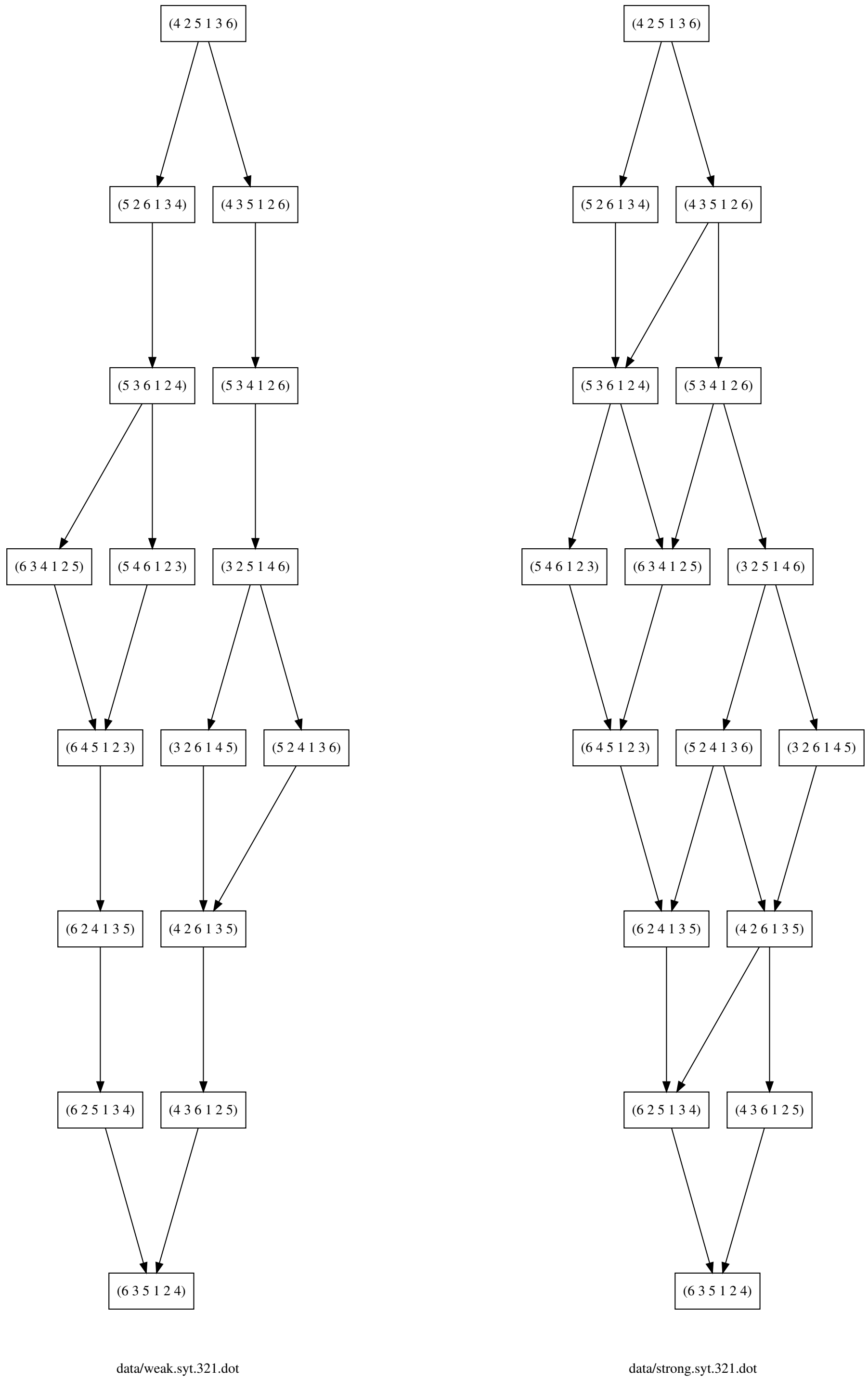


Figure 6.11: Hasse diagram of the Weak SYT Poset and the Strong SYT Poset of $\lambda = (3, 2, 1)$. Each tableau is represented by its row reading word in these pictures.

Chapter 7

ON A THEOREM OF BAXTER AND ZEILBERGER VIA A RESULT OF ROSELLE

A version of this chapter has been reviewed by Romik and Zeilberger. It will be submitted for publication in the near future [96], after the possibility of strengthening the argument to give a local limit theorem has been fully explored.

7.1 Main Results

As in Chapter 1, for a permutation $w = w_1 \cdots w_n \in S_n$, the *inversion* and *major index* statistics are given by

$$\operatorname{inv}(w) := \#\{i < j : w_i > w_j\} \qquad \text{and} \qquad \operatorname{maj}(w) := \sum_{\substack{i \in [n-1] \\ w_i > w_{i+1}}} i.$$

It is well-known that inv and maj are equidistributed on S_n with common mean and standard deviation

$$\mu_n = \frac{n(n-1)}{4} \qquad \text{and} \qquad \sigma_n^2 = \frac{2n^3 + 3n^2 - 5n}{72}.$$

(These results also follow easily from the arguments in this chapter.) In [7], Baxter and Zeilberger proved that inv and maj are jointly independently asymptotically normally distributed as $n \to \infty$. More precisely, define normalized random variables on S_n

$$X_n := \frac{\operatorname{inv} - \mu_n}{\sigma_n}, \qquad Y_n := \frac{\operatorname{maj} - \mu_n}{\sigma_n}. \tag{7.1}$$

Theorem 7.1.1 (Baxter–Zeilberger, [7])**.** *For each $u, v \in \mathbb{R}$, we have*

$$\lim_{n\to\infty} \mathbb{P}[X_n \le u, Y_n \le v] = \frac{1}{2\pi} \int_{-\infty}^{u} \int_{-\infty}^{v} e^{-x^2/2} e^{-y^2/2}\, dy\, dx.$$

See [7] for further historical background. Baxter and Zeilberger's argument involves mixed factorial moments and recurrences based on combinatorial manipulations with permutations. Romik suggested a generating function due to Roselle, quoted as Theorem 7.2.2 below, should provide another approach. Zeilberger subsequently offered a \$300 reward for such an argument. The aim of this chapter is to give such a proof. For further context, see [103] and [98].

7.2 Consequences of Roselle's Formula

Here we recall Roselle's formula, originally stated in different but equivalent terms, and derive a generating function expression which quickly motivates Theorem 7.1.1.

Definition 7.2.1. Let H_n be the bivariate $\mathrm{inv}, \mathrm{maj}$ generating function on S_n, i.e.

$$H_n(p, q) := \sum_{w \in S_n} p^{\mathrm{inv}(w)} q^{\mathrm{maj}(w)}.$$

Theorem 7.2.2 (Roselle, [77])**.** *We have*

$$\sum_{n \ge 0} \frac{H_n(p, q) z^n}{(p)_n (q)_n} = \prod_{a, b \ge 0} \frac{1}{1 - p^a q^b z} \tag{7.2}$$

where $(p)_n := (1-p)(1-p^2)\cdots(1-p^n)$.

The following is the main result of this section.

Theorem 7.2.3. *There are constants $c_\mu \in \mathbb{Z}$ indexed by integer partitions μ such that*

$$\frac{H_n(p, q)}{n!} = \frac{[n]_p! [n]_q!}{n!^2} F_n(p, q) \tag{7.3}$$

where

$$F_n(p,q) = \sum_{d=0}^{n} [(1-p)(1-q)]^d \sum_{\substack{\mu \vdash n \\ \ell(\mu)=n-d}} \frac{c_\mu}{\prod_i [\mu_i]_p [\mu_i]_q} \tag{7.4}$$

and $[n]_p! := [n]_p[n-1]_p \cdots [1]_p$, $[c]_p := 1 + p + \cdots + p^{c-1} = (1-p^c)/(1-p)$.

An explicit expression for c_μ is given below in (7.12). The rest of this section is devoted to proving Theorem 7.2.3. Straightforward manipulations with (7.2) immediately yield (7.3) where

$$F_n(p,q) := (1-p)^n(1-q)^n n! \cdot \{z^n\} \left(\prod_{a,b \geq 0} \frac{1}{1-p^a q^b z} \right) \tag{7.5}$$

and $\{z^n\}$ here refers to extracting the coefficient of z^n. Thus it suffices to show (7.5) implies (7.4). By standard arguments, the z^n coefficient of the product over a, b in (7.5) is the bivariate generating function of size-n multisets of pairs $(a,b) \in \mathbb{Z}^2_{\geq 0}$, where the weight of such a multset is its sum.

Definition 7.2.4. For $\lambda \vdash n$, let M_λ be the bivariate generating function for multisets of pairs $(a,b) \in \mathbb{Z}^n_{\geq 0}$ of type λ, i.e. some element has multiplicity λ_1, another element has multiplicity λ_2, etc.

We clearly have

$$\{z^n\} \left(\prod_{a,b \geq 0} \frac{1}{1-p^a q^b z} \right) = \sum_{\lambda \vdash n} M_\lambda(p,q), \tag{7.6}$$

though the M_λ are inconvenient to work with, so we perform a change of basis.

Definition 7.2.5. Let $P[n]$ denote the lattice of set partitions of $[n] := \{1, 2, \ldots, n\}$ with minimum $\widehat{0} = \{\{1\}, \{2\}, \ldots, \{n\}\}$ and maximum $\widehat{1} = \{\{1, 2, \ldots, n\}\}$. Here $\Lambda \leq \Pi$ means that Π can be obtained from Λ by merging blocks of Λ. The *type* of a set partition Λ is the integer partition obtained by rearranging the list of the block sizes of Λ in weakly decreasing order. For $\lambda \vdash n$, set

$$\Lambda(\lambda) := \{\{1, 2, \ldots, \lambda_1\}, \{\lambda_1 + 1, \lambda_1 + 2, \ldots, \lambda_1 + \lambda_2\}, \ldots\},$$

which has type λ.

Definition 7.2.6. For $\Pi \in P[n]$, let R_Π denote the bivariate generating function for lists $L \in (\mathbb{Z}^2_{\geq 0})^n$ where for each block of Π the entries in L from that block are all equal. Similarly, let S_Π denote the bivariate generating function of lists L where in addition to entries from the same block being equal, entries from two different blocks are not equal.

We easily see that

$$R_\Lambda(p,q) = \prod_{A \in \Lambda} \frac{1}{(1-p^{\#A})(1-q^{\#A})} \tag{7.7}$$

and that

$$R_\Lambda(p,q) = \sum_{\Pi : \Lambda \leq \Pi} S_\Pi, \tag{7.8}$$

so that, by Möbius inversion on $P[n]$,

$$S_\Pi = \sum_{\Lambda : \Pi \leq \Lambda} \mu(\Pi, \Lambda) R_\Lambda. \tag{7.9}$$

Under the "forgetful" map from lists to multisets, a multiset of type $\lambda \vdash n$ has fiber of size $\binom{n}{\lambda}$. It follows that

$$S_{\Pi(\lambda)} = \frac{n!}{\lambda!} M_\lambda \tag{7.10}$$

where $\lambda! := \lambda_1! \lambda_2! \cdots$. Combining in order (7.5), (7.6), (7.10), (7.9), and (7.7) gives

$$F_n(p,q) = \sum_{d=0}^{n} [(1-p)(1-q)]^d \sum_{\lambda \vdash n} \lambda! \sum_{\substack{\Lambda : \Pi(\lambda) \leq \Lambda \\ \#\Lambda = n-d}} \frac{\mu(\Pi(\lambda), \Lambda)}{\prod_{A \in \Lambda} [\#A]_p [\#A]_q}. \tag{7.11}$$

Now (7.4) follows from (7.11) where

$$c_\mu = \sum_{\lambda \vdash n} \lambda! \sum_{\substack{\Lambda : \Pi(\lambda) \leq \Lambda \\ \mathrm{type}(\Lambda) = \mu}} \mu(\Pi(\lambda), \Lambda). \tag{7.12}$$

This completes the proof of Theorem 7.2.3.

Remark 7.2.7. From (7.12), $c_{(1^n)} = 1$ since the sum only involves $\Lambda = \widehat{0}$. Letting $p \to 1$ in (7.4), the only surviving term is $d = 0$ and $\lambda = (1^n)$. Consequently, $H_n(1, q) = [n]_q!$, recovering a classic result of MacMahon [61, §1].

Remark 7.2.8. Using (7.3), we see that the probability generating function (discussed below in Example 7.4.3) $H_n(p, q)/n!$ differs from $[n]_p![n]_q!/n!^2$ by precisely the correction factor $F_n(p, q)$. Using (7.5), this factor has the following combinatorial interpretation:

$$F_n = \frac{n! \cdot \text{g.f. of size-}n\text{ multisets from } \mathbb{Z}^2_{\geq 0}}{\text{g.f. of size-}n\text{ lists from } \mathbb{Z}^2_{\geq 0}}.$$

Intuitively, the numerator and denominator should be the same "up to first order." Theorem 7.3.1 will give one precise sense in which they are asymptotically equal.

7.3 Estimating the Correction Factor

This section is devoted to showing that the correction factor $F_n(p, q)$ from Theorem 7.2.3 is negligible in an appropriate sense, Theorem 7.3.1. Recall that σ_n denotes the standard deviation of inv or maj on S_n.

Theorem 7.3.1. *Uniformly on compact subsets of* $\mathbb{R}^2$, *we have*

$$F_n(e^{is/\sigma_n}, e^{it/\sigma_n}) \to 1 \qquad \textit{as} \qquad n \to \infty.$$

We begin with some simple estimates starting from (7.11) which motivate the rest of the inequalities in this section. We may assume $|s|, |t| \leq M$ for some fixed M. Setting $p = e^{is/\sigma_n}, q = e^{it/\sigma_n}$, we have $|1 - p| = |1 - \exp(is/\sigma_n)| \leq |s|/\sigma_n$. For n sufficiently large compared to M, we also have $|s/\sigma_n| \ll 1$ and so, for all $c \in \mathbb{Z}_{\geq 1}$, $|[c]_p| = |[c]_{\exp(is/\sigma_n)}| \geq 1$. Thus for n sufficiently large, (7.11) gives

$$|F_n(e^{is/\sigma_n}, e^{it/\sigma_n}) - 1| \leq \sum_{d=1}^{n} \frac{|st|^d}{\sigma_n^{2d}} \sum_{\lambda \vdash n} \lambda! \sum_{\substack{\Lambda : \Pi(\lambda) \leq \Lambda \\ \#\Lambda = n-d}} |\mu(\Pi(\lambda), \Lambda)|. \tag{7.13}$$

Lemma 7.3.2. *Suppose $\lambda \vdash n$ with $\ell(\lambda) = n-k$, and fix d. Then*

$$\sum_{\substack{\Lambda:\Pi(\lambda)\leq\Lambda \\ \#\Lambda=n-d}} \mu(\Pi(\lambda),\Lambda) = (-1)^{d-k} \sum_{\substack{\Lambda\in P[n-k] \\ \#\Lambda=n-d}} \prod_{A\in\Lambda} (\#A-1)! \tag{7.14}$$

and the terms on the left all have the same sign $(-1)^{d-k}$. The sums are empty unless $n \geq d \geq k \geq 0$.

Proof. The upper order ideal $\{\Lambda \in P[n] : \Pi(\lambda) \leq \Lambda\}$ is isomorphic to $P[n-k]$ by collapsing the $n-k$ blocks of $\Pi(\lambda)$ to singletons. This isomorphism preserves the number of blocks. Furthermore, recall that in $P[n]$ we have

$$\mu(\widehat{0},\widehat{1}) = (-1)^{n-1}(n-1)!,$$

from which it follows easily that

$$\mu(\widehat{0},\Lambda) = \prod_{A\in\Lambda} (-1)^{\#A-1}(\#A-1)!. \tag{7.15}$$

The result follows immediately upon combining these observations. □

Lemma 7.3.3. *Let $\lambda \vdash n$ with $\ell(\lambda) = n-k$ and $n \geq d \geq k \geq 0$. Then*

$$\sum_{\substack{\Lambda:\Pi(\lambda)\leq\Lambda \\ \#\Lambda=n-d}} |\mu(\Pi(\lambda),\Lambda)| \leq (n-k)^{2(d-k)}. \tag{7.16}$$

Proof. Using (7.14), we can interpret the sum as the number of permutations of $[n-k]$ with $n-d$ cycles, which is a Stirling number of the first kind. There are well-known asymptotics for these numbers, though the stated elementary bound suffices for our purposes. We induct on d. At $d = k$, the result is trivial. Given a permutation of $[n-k]$ with $n-d$ cycles, choose $i, j \in [n-k]$ from different cycles. Suppose the cycles are of the form $(i' \ \cdots \ i)$ and $(j \ \cdots \ j')$.

Splice the two cycles together to obtain

$$(i' \ \cdots \ i \, j \ \cdots \ j').$$

This procedure constructs every permutation of $[n-k]$ with $n-(d+1)$ cycles and requires no more than $(n-k)^2$ choices. The result follows. $\square$

Lemma 7.3.4. *For $n \geq d \geq k \geq 0$, we have*

$$\sum_{\substack{\lambda \vdash n \\ \ell(\lambda)=n-k}} \lambda! \sum_{\substack{\Lambda : \Pi(\lambda) \leq \Lambda \\ \#\Lambda = n-d}} |\mu(\Pi(\lambda), \Lambda)| \leq (n-k)^{2d-k}(k+1)!. \tag{7.17}$$

Proof. For $\lambda \vdash n$ with $\ell(\lambda) = n-k$, $\lambda!$ can be thought of as the product of terms obtained from filling the ith row of λ with $1, 2, \ldots, \lambda_i$. Alternatively, we may fill the cells of λ as follows: put $n-k$ one's in the first column, and fill the remaining cells with the numbers $2, 3, \ldots, k+1$ starting at the largest row and proceeding left to right. It's easy to see the labels of the first filling are bounded above by the labels of the second filling, so that $\lambda! \leq (k+1)!$. Furthermore, each $\lambda \vdash n$ with $\ell(\lambda) = n-k$ can be constructed by first placing $n-k$ cells in the first column and then deciding on which of the $n-k$ rows to place each of the remaining k cells, so there are no more than $(n-k)^k$ such λ. The result follows from combining these bounds with (7.16). $\square$

Lemma 7.3.5. *For n sufficiently large, for all $0 \leq d \leq n$ we have*

$$\sum_{\lambda \vdash n} \lambda! \sum_{\substack{\Lambda : \Pi(\lambda) \leq \Lambda \\ \#\Lambda = n-d}} |\mu(\Pi(\lambda), \Lambda)| \leq 3n^{2d}.$$

Proof. For $n \geq 2$ large enough, for all $n \geq k \geq 2$ we see that $(k+1)! < n^{k-1}$. Using (7.17) gives

$$\begin{aligned}
\sum_{\lambda \vdash n} \lambda! \sum_{\substack{\Lambda : \Pi(\lambda) \le \Lambda \\ \#\Lambda = n-d}} |\mu(\Pi(\lambda), \Lambda)| &\le \sum_{k=0}^{d} (n-k)^{2d-k}(k+1)! \\
&\le n^{2d} + 2(n-1)^{2d-1} + \sum_{k=2}^{d} (n-k)^{2d-k} n^{k-1} \\
&\le n^{2d} + 2n^{2d-1} + \sum_{k=2}^{d} n^{2d-1} \\
&= n^{2d} + 2n^{2d-1} + (d-1)n^{2d-1} \le 3n^{2d}.
\end{aligned}$$

□

We may now complete the proof of Theorem 7.3.1. Combining Lemma 7.3.5 and (7.13) gives

$$|F_n(e^{is/\sigma_n}, e^{it/\sigma_n}) - 1| \le 3 \sum_{d=1}^{n} \frac{(Mn)^{2d}}{\sigma_n^{2d}}.$$

Since $\sigma_n^2 \sim n^3/36$ and M is constant, $(Mn)^{2d}/\sigma_n^{2d} \sim (36^2 M^2/n)^d$. Since M is constant, using a geometric series it follows that

$$\lim_{n\to\infty} \sum_{d=1}^{n} \frac{(Mn)^{2d}}{\sigma_n^{2d}} = 0,$$

completing the proof of Theorem 7.3.1.

Remark 7.3.6. Indeed, the argument shows that $|F_n(e^{is/\sigma_n}, e^{it/\sigma_n}) - 1| = O(1/n)$. The above estimates are particularly far from sharp for large d, though for small d they are quite accurate. Working directly with (7.11), one finds the $d = 1$ contribution to be

$$(1-p)(1-q)\frac{2 - \binom{n}{2}}{[2]_p[2]_q}.$$

Letting $p = e^{is/\sigma_n}, q = e^{it/\sigma_n}$, straightforward estimates shows that this is $\Omega(1/n)$. Consequently, the preceding arguments are strong enough to identify the leading term, and in particular

$$|F_n(e^{is/\sigma_n}, e^{it/\sigma_n}) - 1| = \Theta(1/n).$$

7.4 Deducing Baxter–Zeilberger's Result

We next summarize enough of the standard theory of characteristic functions to prove Theorem 7.1.1 using (7.3) and Theorem 7.3.1.

Definition 7.4.1. The *characteristic function* of an $\mathbb{R}^k$-valued random variable $X = (X_1, \ldots, X_k)$ is the function $\phi_X \colon \mathbb{R}^k \to \mathbb{C}$ given by

$$\phi_X(s_1, \ldots, s_k) := \mathbb{E}[\exp(i(s_1X_1 + \cdots + s_kX_k))].$$

Example 7.4.2. It is well-known that the characteristic function of the standard normal random variable with density $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ is $e^{-s^2/2}$. Similarly, the characteristic function of a bivariate jointly independent standard normal random variable with density $\frac{1}{2\pi}e^{-x^2/2-y^2/2}$ is $e^{-s^2/2-t^2/2}$.

Example 7.4.3. If W is a finite set and $\mathrm{stat} = (\mathrm{stat}_1, \ldots, \mathrm{stat}_k) \colon W \to \mathbb{Z}^k_{\geq 0}$ is some statistic, the multivariate *probability generating function* of stat on W is

$$P(x_1, \ldots, x_k) := \frac{1}{\#W} \sum_{w \in W} x_1^{\mathrm{stat}_1(w)} \cdots x_k^{\mathrm{stat}_k(w)}.$$

The characteristic function of the corresponding random variable X where the w are chosen uniformly from W is

$$\phi_X(s_1, \ldots, s_k) = P(e^{is_1}, \ldots, e^{is_k}).$$

From Example 7.4.3, Remark 7.2.7, and an easy calculation, it follows that the characteristic functions of the random variables X_n and Y_n from (7.1) are

$$\phi_{X_n}(s) = e^{-i\mu_n s/\sigma_n} \frac{[n]_{e^{is/\sigma_n}}!}{n!} \qquad \text{and} \qquad \phi_{Y_n}(t) = e^{-i\mu_n t/\sigma_n} \frac{[n]_{e^{it/\sigma_n}}!}{n!}. \tag{7.18}$$

An analogous calculation for the random variable (X_n, Y_n) together with (7.18) and (7.3) gives

$$\begin{aligned}\phi_{(X_n,Y_n)}(s,t) &= e^{-i(\mu_n s/\sigma_n + \mu_n t/\sigma_n)} \frac{H_n(e^{is/\sigma_n}, e^{it/\sigma_n})}{n!} \\ &= \phi_{X_n}(s)\phi_{Y_n}(t) F_n(e^{is/\sigma_n}, e^{it/\sigma_n}).\end{aligned} \tag{7.19}$$

Theorem 7.4.4 (Multivariate Lévy Continuity, [18, Thm. 2.6.9])**.** *Suppose that $X^{(1)}, X^{(2)}, \ldots$ is a sequence of $\mathbb{R}^k$-valued random variables and X is an $\mathbb{R}^k$-valued random variable. Then $X^{(1)}, X^{(2)}, \ldots$ converges in distribution to X if and only if $\phi_{X^{(n)}}$ converges pointwise to ϕ_X.*

If the distribution function of X is continuous everywhere, convergence in distribution means that for all $u_1, \ldots, u_k$ we have

$$\lim_{n\to\infty} \mathbb{P}[X_i^{(n)} \leq u_i, 1 \leq i \leq k] = \mathbb{P}[X_i \leq u_i, 1 \leq i \leq k].$$

Many techniques are available for proving that inv and maj on S_n are asymptotically normal. The result is typically attributed to Feller.

Theorem 7.4.5. *[25, p. 257] The sequences of random variables X_n and Y_n from (7.1) each converge in distribution to the standard normal random variable.*

We may now complete the proof of Theorem 7.1.1. From Theorem 7.4.5 and Example 7.4.2, we have for all $s, t \in \mathbb{R}$

$$\lim_{n\to\infty} \phi_{X_n}(s) = e^{-s^2/2} \qquad \text{and} \qquad \lim_{n\to\infty} \phi_{Y_n}(t) = e^{-t^2/2}. \tag{7.20}$$

Combing in order (7.20), (7.19), and Theorem 7.3.1 gives

$$\lim_{n\to\infty} \phi_{(X_n,Y_n)}(s,t) = e^{-s^2/2 - t^2/2}.$$

Theorem 7.1.1 now follows from Example 7.4.2 and Theorem 7.4.4.

BIBLIOGRAPHY

[1] Ron M. Adin, Victor Reiner, and Yuval Roichman. On cyclic descents for tableaux. arXiv:1710.06664v2, 2017.

[2] Ron M. Adin and Yuval Roichman. Descent functions and random Young tableaux. *Combin. Probab. Comput.*, 10(3):187–201, 2001.

[3] Connor Ahlbach and Joshua P. Swanson. Refined cyclic sieving. *Sém. Lothar. Combin.*, 78B:Art. 48, 12, 2017.

[4] Connor Ahlbach and Joshua P. Swanson. Cyclic sieving, necklaces, and branching rules related the Thrall's problem. In preparation, 2018.

[5] Connor Ahlbach and Joshua P. Swanson. Refined cyclic sieving on words for the major index statistic. *European Journal of Combinatorics*, 73:37 – 60, 2018.

[6] Drew Armstrong, Victor Reiner, and Brendon Rhoades. Parking spaces. *Adv. Math.*, 269:647–706, 2015.

[7] Andrew Baxter and Doron Zeilberger. The Number of Inversions and the Major Index of Permutations are Asymptotically Joint-Independently Normal (Second Edition!). arXiv:1004.1160v2, 2011.

[8] François Bergeron, Nantel Bergeron, and Adriano M. Garsia. Idempotents for the free Lie algebra and q-enumeration. In *Invariant theory and tableaux (Minneapolis, MN, 1988)*, volume 19 of *IMA Vol. Math. Appl.*, pages 166–190. Springer, New York, 1990.

[9] Andrew Berget, Sen-Peng Eu, and Victor Reiner. Constructions for cyclic sieving phenomena. *SIAM J. Discrete Math.*, 25(3):1297–1314, 2011.

[10] Sara C. Billey, Matjaž Konvalinka, and Joshua P. Swanson. Distribution of major index for standard tableaux and asymptotic normality. In preparation, 2018.

[11] Patrick Billingsley. *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, third edition, 1995.

[12] Miklós Bóna, editor. *Handbook of Enumerative Combinatorics*. Discrete Mathematics and its Applications (Boca Raton). CRC Press, Boca Raton, FL, 2015.

[13] E. Rodney Canfield, Svante Janson, and Doron Zeilberger. The Mahonian probability distribution on words is asymptotically normal. *Adv. in Appl. Math.*, 46(1-4):109–124, 2011.

[14] E. Rodney Canfield, Svante Janson, and Doron Zeilberger. Corrigendum to "The Mahonian probability distribution on words is asymptotically normal". *Adv. in Appl. Math.*, 49(1):77, 2012.

[15] Paola Cellini. Cyclic Eulerian elements. *European J. Combin.*, 19(5):545–552, 1998.

[16] William Y. C. Chen, Carol J. Wang, and Larry X. W. Wang. The limiting distribution of the coefficients of the q-Catalan numbers. *Proc. Amer. Math. Soc.*, 136(11):3759–3767, 2008.

[17] Ionuţ Ciocan-Fontanine, Matjaž Konvalinka, and Igor Pak. The weighted hook length formula. *J. Combin. Theory Ser. A*, 118(6):1703–1717, 2011.

[18] Roger Cuppens. *Decomposition of Multivariate Probabilities.* Academic Press, New York-London, 1975. Probability and Mathematical Statistics, volume 29.

[19] F. N. David and D. E. Barton. *Combinatorial Chance.* Hafner Publishing Co., New York, 1962.

[20] Jacques Désarménien. Étude modulo n des statistiques mahoniennes. *Séminaire Lotharingien de Combinatoire*, 22:27–35, 1990.

[21] Persi Diaconis. *Group Representations in Probability and Statistics*, volume 11 of *Institute of Mathematical Statistics Lecture Notes—Monograph Series.* Institute of Mathematical Statistics, Hayward, CA, 1988.

[22] Sergi Elizalde and Yuval Roichman. On rotated Schur-positive sets. *J. Combin. Theory Ser. A*, 152:121–137, 2017.

[23] Paul Erdös and Joseph Lehner. The distribution of the number of summands in the partitions of a positive integer. *Duke Math. J.*, 8:335–345, 1941.

[24] William Feller. The fundamental limit theorems in probability. *Bull. Amer. Math. Soc.*, 51:800–832, 1945.

[25] William Feller. *An Introduction to Probability Theory and Its Applications, volume I.* Third edition. John Wiley & Sons, Inc., New York-London-Sydney, 1968.

[26] Dominique Foata. On the Netto inversion number of a sequence. *Proc. Amer. Math. Soc.*, 19:236–240, 1968.

[27] Sergey Fomin and Nathan Lulov. On the number of rim hook tableaux. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 223:219–226, 1995.

[28] H. O. Foulkes. Characters of symmetric groups induced by characters of cyclic subgroups. In *Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972)*, pages 141–154. Inst. Math. Appl., Southend-on-Sea, 1972.

[29] J. S. Frame, G. de B. Robinson, and R. M. Thrall. The hook graphs of the symmetric groups. *Canadian J. Math.*, 6:316–324, 1954.

[30] Jason Fulman. The distribution of descents in fixed conjugacy classes of the symmetric groups. *J. Combin. Theory Ser. A*, 84(2):171–180, 1998.

[31] William Fulton. *Young Tableaux*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997.

[32] J. Fürlinger and J. Hofbauer. q-Catalan numbers. *J. Combin. Theory Ser. A*, 40(2):248–264, 1985.

[33] Adriano M. Garsia. Combinatorial methods in the theory of Cohen-Macaulay rings. *Adv. in Math.*, 38(3):229–266, 1980.

[34] Adriano M. Garsia. Alfred Young's construction of the irreducible representations of S_n. Unpublished notes, January 2012.

[35] Adriano M. Garsia and Claudio Procesi. On certain graded S_n-modules and the q-Kostka polynomials. *Adv. Math.*, 94(1):82–138, 1992.

[36] Ira M. Gessel and Christophe Reutenauer. Counting permutations with given cycle structure and descent set. *J. Combin. Theory Ser. A*, 64(2):189–215, 1993.

[37] V. L. Goncharov. From the realm of combinatorics. *Izv. Akad. Nauk SSSR*, 8:3–48, 1944.

[38] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics*. Addison-Wesley Publishing Company, Advanced Book Program, Reading, MA, 1989.

[39] J. A. Green. The characters of the finite general linear groups. *Trans. Amer. Math. Soc.*, 80:402–447, 1955.

[40] Hansraj Gupta. A new look at the permutations of the first n natural numbers. *Indian J. Pure Appl. Math.*, 9(6):600–631, 1978.

[41] Marshall Hall, Jr. *The Theory of Groups.* The Macmillan Co., New York, N.Y., 1959.

[42] L. H. Harper. Stirling behavior is asymptotically normal. *Ann. Math. Statist.*, 38:410–414, 1967.

[43] Joan E. Herman and Fan R. K. Chung. Some results on hook lengths. *Discrete Math.*, 20(1):33–40, 1977/78.

[44] Hsien-Kuei Hwang and Vytas Zacharovas. Limit distribution of the coefficients of polynomials with only unit roots. *Random Structures Algorithms*, 46(4):707–738, 2015.

[45] Gordon James and Adalbert Kerber. *The Representation Theory of the Symmetric Group*, volume 16 of *Encyclopedia of Mathematics and its Applications.* Addison-Wesley Publishing Co., Reading, Mass., 1981.

[46] Marianne Johnson. Standard tableaux and Klyachko's theorem on Lie representations. *J. Combin. Theory Ser. A*, 114(1):151–158, 2007.

[47] Maurice Kendall, Alan Stuart, and J. Keith Ord. *Kendall's Advanced Theory of Statistics.* The Clarendon Press, Oxford University Press, New York, fifth edition, 1987.

[48] S. Kerov. A q-analog of the hook walk algorithm for random Young tableaux. *J. Algebraic Combin.*, 2(4):383–396, 1993.

[49] Gene B. Kim and Sangchul Lee. Central limit theorem for descents in conjugacy classes of S_n. arXiv:1803.10457v1, 2018.

[50] A. A. Klyachko. Lie elements in the tensor algebra. *Siberian Mathematical Journal*, 15(6):914–920, 1974.

[51] John Knopfmacher. *Abstract Analytic Number Theory.* North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1975. North-Holland Mathematical Library, volume 12.

[52] Donald E. Knuth. *The Art of Computer Programming, volume 3.* Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1973.

[53] L. G. Kovács and Ralph Stöhr. A combinatorial proof of Klyachko's theorem on Lie representations. *J. Algebraic Combin.*, 23(3):225–230, 2006.

[54] Witold Kraśkiewicz and Jerzy Weyman. Algebra of coinvariants and the action of a Coxeter element. *Bayreuth. Math. Schr.*, (63):265–284, 2001.

[55] T. Y. Lam. Young diagrams, Schur functions, the Gale-Ryser theorem and a conjecture of Snapper. *J. Pure Appl. Algebra*, 10(1):81–94, 1977/78.

[56] Thomas Lam and Alexander Postnikov. Alcoved Polytopes II. arXiv:1202.4015v1, 2012.

[57] Michael Larsen and Aner Shalev. Characters of symmetric groups: sharp bounds and applications. *Invent. Math.*, 174(3):645–687, 2008.

[58] R. A. Liebler and M. R. Vitale. Ordering the partition characters of the symmetric group. *J. Algebra*, 25:487–489, 1973.

[59] I. G. Macdonald. *Symmetric Functions and Hall Polynomials*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, second edition, 1995.

[60] 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. *Amer. J. Math.*, 35(3):281–322, 1913.

[61] P. A. MacMahon. Two Applications of General Theorems in Combinatory Analysis: (1) To the Theory of Inversions of Permutations; (2) To the Ascertainment of the Numbers of Terms in the Development of a Determinant which has Amongst its Elements an Arbitrary Number of Zeros. *Proc. London Math. Soc.*, S2-15(1):314, 1917.

[62] P. A. MacMahon. *Combinatory analysis, volume I, II*. Dover Phoenix Editions. Dover Publications, Inc., Mineola, NY, 2004. Reprint of *An introduction to combinatory analysis* (1920) and *Combinatory analysis, volume I, II* (1915, 1916).

[63] H. B. Mann and D. R. Whitney. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statistics*, 18:50–60, 1947.

[64] Mazur, Barry. Bernoulli numbers and the unity of mathematics: Bartlett lecture notes, 2008. Online. http://www.math.harvard.edu/~mazur/papers/slides.Bartlett.pdf.

[65] Alejandro Morales, Igor Pak, and Greta Panova. Asymptotics of the number of standard Young tableaux of skew shape. arXiv:1610.07561v2, 2017.

[66] Alejandro Morales, Igor Pak, and Greta Panova. Hook formulas for skew shapes I. q-analogues and bijections. arXiv:1512.08348v4, 2017.

[67] Jonathan Novak and Piotr Śniady. What is ... a free cumulant? *Notices Amer. Math. Soc.*, 58(2):300–301, 2011.

[68] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences, 2017. Online. http://oeis.org.

[69] Igor Pak. Inequality for hook numbers in Young diagrams. MathOverflow. https://mathoverflow.net/q/243846 (2017-04-13).

[70] Oliver Pechenik. Cyclic sieving of increasing tableaux and small Schröder paths. *J. Combin. Theory Ser. A*, 125:357–378, 2014.

[71] T. Kyle Petersen. Cyclic descents and P-partitions. *J. Algebraic Combin.*, 22(3):343–375, 2005.

[72] Timothy Pressey, Anna Stokke, and Terry Visentin. Increasing tableaux, Narayana numbers and an instance of the cyclic sieving phenomenon. *Ann. Comb.*, 20(3):609–621, 2016.

[73] Victor Reiner, Dennis Stanton, and Dennis White. The cyclic sieving phenomenon. *J. Combin. Theory Ser. A*, 108(1):17–50, 2004.

[74] Christophe Reutenauer. *Free Lie algebras*, volume 7 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1993.

[75] Brendon Rhoades. Cyclic sieving, promotion, and representation theory. *J. Combin. Theory Ser. A*, 117(1):38–76, 2010.

[76] Yuval Roichman. Upper bound on the characters of the symmetric groups. *Invent. Math.*, 125(3):451–485, 1996.

[77] D. P. Roselle. Coefficients associated with the expansion of certain products. *Proc. Amer. Math. Soc.*, 45:144–150, 1974.

[78] Bruce E. Sagan. *The Symmetric Group*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001.

[79] Bruce E. Sagan. The cyclic sieving phenomenon: a survey. In *Surveys in combinatorics 2011*, volume 392 of *London Math. Soc. Lecture Note Ser.*, pages 183–233. Cambridge Univ. Press, Cambridge, 2011.

[80] Manfred Schocker. Embeddings of higher Lie modules. *J. Pure Appl. Algebra*, 185(1-3):279–288, 2003.

[81] Manfred Schocker. Multiplicities of higher Lie characters. *J. Aust. Math. Soc.*, 75(1):9–21, 2003.

[82] Jean-Pierre Serre. *Linear representations of finite groups.* Springer-Verlag, New York-Heidelberg, 1977. Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, volume 42.

[83] Ernst Snapper. Group characters and nonnegative integral matrices. *J. Algebra*, 19:520–535, 1971.

[84] Wilhelm Specht. *Eine Verallgemeinerung der symmetrischen Gruppe.* PhD thesis, Humboldt-Universitt zu Berlin, null, 1932.

[85] Joel Spencer. *Asymptopia*, volume 71 of *Student Mathematical Library.* American Mathematical Society, Providence, RI, 2014. With Laura Florescu.

[86] T. A. Springer. Regular elements of finite reflection groups. *Invent. Math.*, 25:159–198, 1974.

[87] Richard P. Stanley. Invariants of finite groups and their applications to combinatorics. *Bull. Amer. Math. Soc. (N.S.)*, 1(3):475–511, 1979.

[88] Richard P. Stanley. Weyl groups, the hard Lefschetz theorem, and the Sperner property. *SIAM J. Alg. Disc. Meth.*, 1(2):168–184, 1980.

[89] Richard P. Stanley. The stable behavior of some characters of $\mathrm{SL}(n, \mathbf{C})$. *Linear and Multilinear Algebra*, 16(1-4):3–27, 1984.

[90] Richard P. Stanley. Log-concave and unimodal sequences in algebra, combinatorics, and geometry. In *Graph theory and its applications: East and West (Jinan, 1986)*, volume 576 of *Ann. New York Acad. Sci.*, pages 500–535. New York Acad. Sci., New York, 1989.

[91] Richard P. Stanley. *Enumerative Combinatorics. Volume 2*, volume 62 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1999.

[92] Richard P. Stanley. *Enumerative Combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.

[93] R. Steinberg. A geometric approach to the representations of the full linear group over a Galois field. *Trans. Amer. Math. Soc.*, 71:274–282, 1951.

[94] John R. Stembridge. On the eigenvalues of representations of reflection groups and wreath products. *Pacific J. Math.*, 140(2):353–396, 1989.

[95] Sheila Sundaram. On conjugacy classes of S_n containing all irreducibles. *Israel Journal of Mathematics*, 225(1):321–342, 2018.

[96] Joshua P. Swanson. On a theorem of Baxter and Zeilberger via a result of Roselle. In preparation, 2018.

[97] Joshua P. Swanson. On the existence of tableaux with given modular major index. *Algebraic Combinatorics*, 1(1):3–21, 2018.

[98] Marko Thiel. The inversion number and the major index are asymptotically jointly normally distributed on words. *Combin. Probab. Comput.*, 25(3):470–483, 2016.

[99] R. M. Thrall. On symmetrized Kronecker powers and the structure of the free Lie ring. *Amer. J. Math.*, 64:371–388, 1942.

[100] Stan Wagon and Herbert S. Wilf. When are subset sums equidistributed modulo m? *Electron. J. Combin.*, 1, 1994.

[101] Wikipedia. Bernoulli number — Wikipedia, The Free Encyclopedia. 2017-07-21.

[102] Alfred Young. *The collected papers of Alfred Young (1873–1940)*. University of Toronto Press, Toronto, Ont., Buffalo, N. Y., 1977.

[103] Doron Zeilberger. The number of inversions and the major index of permutations are asymptotically joint-independently normal. Personal web page. http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/invmaj.html (2018-01-25).

[104] Doron Zeilberger. Kathy O'Hara's constructive proof of the unimodality of the Gaussian polynomials. *Amer. Math. Monthly*, 96(7):590–602, 1989.

VITA

Josh Swanson earned a B.S. from Harvey Mudd College in 2010, and is expected to earn a Ph.D. from the University of Washington in 2018. His research interests include algebraic combinatorics, probabilistic combinatorics, the representation theory of reflection groups, and major index statistics. He has published papers in Algebraic Combinatorics and Seminaire Lotharingien de Combinatoire. He has also given research talks at the University of Minnesota, the University of San Diego, the University of Michigan, the University of Washington, and several AMS Special Sessions. Josh is also interested in modernizing linear algebra education and using computers to guide pure mathematics research.