

# On kernels of descent statistics

William L. Clark<sup>a</sup>

Yan Zhuang<sup>b</sup>

Submitted: Jun 27, 2023; Accepted: May 4, 2024; Published: May 31, 2024

© The authors. Released under the CC BY-ND license (International 4.0).

## Abstract

The kernel  $\mathcal{K}^{\text{st}}$  of a descent statistic  $\text{st}$ , introduced by Grinberg, is a subspace of the algebra  $\text{QSym}$  of quasisymmetric functions defined in terms of  $\text{st}$ -equivalent compositions, and is an ideal of  $\text{QSym}$  if and only if  $\text{st}$  is shuffle-compatible. This paper continues the study of kernels of descent statistics, with emphasis on the peak set  $\text{Pk}$  and the peak number  $\text{pk}$ . The kernel  $\mathcal{K}^{\text{Pk}}$  in particular is precisely the kernel of the canonical projection from  $\text{QSym}$  to Stembridge's algebra of peak quasisymmetric functions, and is the orthogonal complement of Nyman's peak algebra. We prove necessary and sufficient conditions for obtaining spanning sets and linear bases for the kernel  $\mathcal{K}^{\text{st}}$  of any descent statistic  $\text{st}$  in terms of fundamental quasisymmetric functions, and give characterizations of  $\mathcal{K}^{\text{Pk}}$  and  $\mathcal{K}^{\text{pk}}$  in terms of the fundamental basis and the monomial basis of  $\text{QSym}$ . Our results imply that the peak set and peak number statistics are  $M$ -binomial, confirming a conjecture of Grinberg.

**Mathematics Subject Classifications:** 05E05 (Primary); 05A05, 05C50, 15A99.

## 1. Introduction

This paper studies ideals of the ring of quasisymmetric functions associated with shuffle-compatible permutation statistics. We begin by giving the relevant definitions.

We call  $\pi$  a *permutation of length  $n$*  if it is a sequence of  $n$  distinct positive integers, displayed as the word  $\pi = \pi_1\pi_2\cdots\pi_n$ . Let  $|\pi|$  denote the length of a permutation  $\pi$  and let  $\mathfrak{P}_n$  denote the set of all permutations of length  $n$ . Note that  $\mathfrak{P}_n$  contains the set  $\mathfrak{S}_n$  of permutations of  $[n] := \{1, 2, \dots, n\}$ , as every permutation in  $\mathfrak{S}_n$  can be written in one-line notation as a sequence of  $n$  distinct positive integers, but  $\mathfrak{P}_n$  and  $\mathfrak{S}_n$  are not the same. For example, 83416 is an element of  $\mathfrak{P}_5$  but not of  $\mathfrak{S}_5$ .

Given a permutation  $\pi \in \mathfrak{P}_n$ , define the *standardization*  $\text{std } \pi$  of  $\pi$  to be the unique permutation in  $\mathfrak{S}_n$  obtained by replacing the smallest letter of  $\pi$  by 1, the second smallest by 2, and so on. For example,  $\text{std } 83416 = 52314$ . A *permutation statistic* is a function  $\text{st}$

---

<sup>a</sup>Mathematical Institute, University of Bonn, Bonn, Germany (s82wclar@uni-bonn.de).

<sup>b</sup>Department of Mathematics and Computer Science, Davidson College, Davidson, North Carolina, USA (yazhuang@davidson.edu).

defined on permutations such that  $\text{st } \pi = \text{st } \sigma$  whenever  $\text{std } \pi = \text{std } \sigma$ .<sup>1</sup> Note that every permutation statistic  $\text{st}$  defined on  $\mathfrak{S}_n$  can be extended to  $\mathfrak{P}_n$  by taking  $\text{st } \pi := \text{st}(\text{std } \pi)$ .

A classical example of a permutation statistic is the descent set, defined as follows. We say that  $i \in [n - 1]$  is a *descent* of a permutation  $\pi \in \mathfrak{P}_n$  if  $\pi_i > \pi_{i+1}$ , and the *descent set*

$$\text{Des } \pi := \{ i \in [n - 1] : \pi_i > \pi_{i+1} \}$$

of  $\pi$  is its set of descents.

The information contained inside the descent set can also be encoded as an integer composition. First observe that every permutation can be uniquely decomposed into a sequence of maximal increasing consecutive subsequences called *increasing runs*. The *descent composition* of  $\pi$ , denoted  $\text{Comp } \pi$ , is the composition whose parts are the lengths of the increasing runs of  $\pi$  in the order that they appear. For example, the increasing runs of  $\pi = 379426$  are 379, 4, and 26, so  $\text{Comp } \pi = (3, 1, 2)$ . If  $\text{Comp } \pi = (j_1, j_2, \dots, j_m)$ , then the descent set of  $\pi$  is given by

$$\text{Des } \pi = \{ j_1, j_1 + j_2, \dots, j_1 + j_2 + \dots + j_{m-1} \}. \quad (1.1)$$

Conversely, if  $\text{Des } \pi = \{ i_1 < i_2 < \dots < i_m \}$ , then

$$\text{Comp } \pi = (i_1, i_2 - i_1, \dots, i_m - i_{m-1}, n - i_m) \quad (1.2)$$

where  $n$  is the length of  $\pi$ . We shall use the notations  $L \vDash n$  and  $|L| = n$  to indicate that  $L$  is a composition of  $n$ , and  $\mathcal{C}$  for the set of all compositions. By convention, we allow the “empty composition”  $\emptyset$  to be a composition of 0; this is the descent composition of the empty permutation, the sole element of  $\mathfrak{P}_0$ .

A permutation statistic  $\text{st}$  is called a *descent statistic* if  $\text{Comp } \pi = \text{Comp } \sigma$  implies  $\text{st } \pi = \text{st } \sigma$ —that is, if  $\text{st}$  depends only on the descent composition, or equivalently, on the descent set and the length. Whenever  $\text{st}$  is a descent statistic, we may write  $\text{st } L$  for the value of  $\text{st}$  on any permutation with descent composition  $L$ . Besides the descent set  $\text{Des}$ , examples of descent statistics include the descent number  $\text{des}$ , major index  $\text{maj}$ , peak set  $\text{Pk}$ , peak number  $\text{pk}$ , exterior peak set  $\text{Epk}$ , valley set  $\text{Val}$ , and valley number  $\text{val}$ ; their definitions will be given later.

For a descent statistic  $\text{st}$ , two compositions  $J$  and  $K$  are said to be *st-equivalent* if  $\text{st } J = \text{st } K$  and  $|J| = |K|$ ; when this is the case, we write  $J \sim_{\text{st}} K$ . Then the *kernel* of  $\text{st}$ , denoted  $\mathcal{K}^{\text{st}}$ , is the subspace

$$\mathcal{K}^{\text{st}} := \text{span}\{ F_J - F_K : J \sim_{\text{st}} K \}$$

of the  $\mathbb{Q}$ -algebra  $\text{QSym}$  of quasisymmetric functions, where  $F_L$  refers to the *fundamental quasisymmetric function*

$$F_L := \sum_{\substack{i_1 \leq i_2 \leq \dots \leq i_n \\ i_j < i_{j+1} \text{ if } j \in \text{Des } L}} x_{i_1} x_{i_2} \cdots x_{i_n}$$

<sup>1</sup>The condition that  $\text{std } \pi = \text{std } \sigma$  implies  $\text{st } \pi = \text{st } \sigma$  is not actually used in this paper, but it plays a role in the theory of shuffle-compatible permutation statistics.

where  $n = |L|$ .

The kernel of a descent statistic was defined by Grinberg in [8, 9],<sup>2</sup> which was a continuation of the work by Gessel and Zhuang on shuffle-compatible permutation statistics [6]. Gessel and Zhuang defined the shuffle algebra  $\mathcal{A}^{\text{st}}$  of a shuffle-compatible permutation statistic (see Section 2.3 for definitions), and showed that whenever a descent statistic  $\text{st}$  is shuffle-compatible, its shuffle algebra is isomorphic to a quotient of  $\text{QSym}$ . The kernel  $\mathcal{K}^{\text{st}}$  is an ideal of  $\text{QSym}$  if and only if  $\text{st}$  is shuffle-compatible, and is in fact the kernel of the canonical projection from  $\text{QSym}$  to  $\mathcal{A}^{\text{st}}$ .

Using their framework, Gessel and Zhuang proved that a number of descent statistics are shuffle-compatible and gave explicit descriptions for their shuffle algebras, but left the shuffle-compatibility of the exterior peak set  $\text{Epk}$  as a conjecture. Grinberg proved the shuffle-compatibility of  $\text{Epk}$  and gave two characterizations of the ideal  $\mathcal{K}^{\text{Epk}}$ : one in terms of the fundamental quasisymmetric functions and another in terms of the *monomial quasisymmetric functions*

$$M_L := \sum_{i_1 < i_2 < \dots < i_m} x_{i_1}^{j_1} x_{i_2}^{j_2} \dots x_{i_m}^{j_m}$$

where  $L = (j_1, j_2, \dots, j_m)$ . Both characterizations have nice combinatorial descriptions involving the underlying compositions. We state these two results together in the following theorem.

**Theorem 0** (Grinberg [8, Propositions 103 and 105]). *For compositions  $J = (j_1, j_2, \dots, j_m)$  and  $K$ , we write  $J \rightarrow K$  if there exists  $l \in \{2, 3, \dots, m\}$  for which  $j_l > 2$  and*

$$K = (j_1, \dots, j_{l-1}, 1, j_l - 1, j_{l+1}, \dots, j_m),$$

*and we write  $J \triangleright K$  if there exists  $l \in \{2, 3, \dots, m\}$  for which  $j_l > 2$  and*

$$K = (j_1, \dots, j_{l-1}, 2, j_l - 2, j_{l+1}, \dots, j_m).$$

*Then the ideal  $\mathcal{K}^{\text{Epk}}$  is spanned (as a  $\mathbb{Q}$ -vector space) in the following ways:*

$$\mathcal{K}^{\text{Epk}} = \text{span}\{F_J - F_K : J \rightarrow K\} = \text{span}\{M_J + M_K : J \triangleright K\}.$$

Grinberg suggests a systematic study of kernels of descent statistics. One direction of research concerns the  $M$ -binomial property: a descent statistic  $\text{st}$  is said to be  $M$ -binomial if  $\mathcal{K}^{\text{st}}$  can be spanned by elements of the form  $\lambda M_J + \mu M_K$  with  $\lambda, \mu \in \mathbb{Q}$ . For example, it follows from Theorem 0 that  $\text{Epk}$  is  $M$ -binomial, and Grinberg gave a list of other descent statistics—including the peak set  $\text{Pk}$  and the peak number  $\text{pk}$ —for which computational evidence suggests are  $M$ -binomial as well [8, Question 107].

<sup>2</sup>The paper [9] is an extended version of [8], containing additional results and more detailed proofs. We will cite the published version [8] unless referring to content which only appears in the extended version [9].

### 1.1. Overview of results

The purpose of the present work is to continue the study of kernels of descent statistics initiated by Grinberg, with emphasis on the kernels of the peak set and the peak number. Both  $\text{Pk}$  and  $\text{pk}$  are shuffle-compatible statistics, so  $\mathcal{K}^{\text{Pk}}$  and  $\mathcal{K}^{\text{pk}}$  are ideals of  $\text{QSym}$ . We will give characterizations—analogueous to the ones given by Grinberg for  $\mathcal{K}^{\text{EpK}}$ —for both ideals, and the ones given in terms of the monomial quasisymmetric functions show that  $\text{Pk}$  and  $\text{pk}$  are indeed  $M$ -binomial.

To obtain characterizations for the kernels in terms of the fundamental quasisymmetric functions, we prove necessary and sufficient conditions for a subset of  $\{F_J - F_K : J \sim_{\text{st}} K\}$  to be: (a) a spanning set of  $\mathcal{K}^{\text{st}}$ , and (b) linearly independent. Note that every subset of  $\{F_J - F_K : J \sim_{\text{st}} K\}$  can be written as

$$\mathcal{F}_S^{\text{st}} := \{F_J - F_K : (J, K) \in S\}$$

for some subset  $S$  of  $\{(J, K) : J \sim_{\text{st}} K\}$ , and we associate to  $S$  a directed graph  $G_S$  with vertex set  $\mathcal{C}$  and edge set  $S$ —i.e., there is an edge from  $J$  to  $K$  if and only if  $(J, K) \in S$ .

For us, a *connected component* of a directed graph refers to a connected component of the underlying undirected graph, and a directed graph is called a *forest* if its underlying undirected graph has no cycles.

**Theorem 1.** *Let  $\text{st}$  be a descent statistic and let  $S \subseteq \{(J, K) : J \sim_{\text{st}} K\}$ .*

- (a) *The kernel  $\mathcal{K}^{\text{st}}$  is spanned by  $\mathcal{F}_S^{\text{st}}$  if and only if the connected components of  $G_S$  are precisely the  $\text{st}$ -equivalence classes of  $\mathcal{C}$ —i.e.,  $J \sim_{\text{st}} K$  if and only if  $J$  and  $K$  are in the same connected component of  $G_S$ .*
- (b) *The set  $\mathcal{F}_S^{\text{st}}$  is linearly independent if and only if  $G_S$  is a forest.*

We will apply Theorem 1 to the peak set and peak number statistics. Given compositions  $J = (j_1, j_2, \dots, j_m)$  and  $K$ , we write:

- $J \rightarrow_1 K$  if there exists  $l \in [m]$  for which  $j_l > 2$  and

$$K = (j_1, \dots, j_{l-1}, 1, j_l - 1, j_{l+1}, \dots, j_m);$$

- $J \rightarrow_2 K$  if  $j_m = 2$  and

$$K = (j_1, \dots, j_{m-1}, 1, 1);$$

- $J \rightarrow_3 K$  if  $j_i \leq 2$  for all  $i \in [m]$ ,  $j_m = 1$ ,  $j_l = 1$  and  $j_{l+1} = 2$  for some  $l \in [m - 2]$ , and

$$K = (j_1, \dots, j_{l-1}, j_{l+1}, j_l, j_{l+2}, \dots, j_m).$$

See Figure 1.1 for all relations  $\rightarrow_1$ ,  $\rightarrow_2$ , and  $\rightarrow_3$  among compositions of at most 5. Ignoring edge labels, this is also the subgraph of  $G_S$  for  $S = \{(J, K) : J \rightarrow_1 K, J \rightarrow_2 K, \text{ or } J \rightarrow_3 K\}$  induced by the compositions of at most 5.



- $J \in \tilde{\mathcal{C}}$  if  $J = (1^{m-1}, 2) = (\underbrace{1, \dots, 1}_{m-1 \text{ parts}}, 2)$ . (Note that  $(2) \in \tilde{\mathcal{C}}$  in the case  $m = 1$ .)

See Figure 1.2 for all relations  $\triangleright_1$  and  $\triangleright_2$ , and elements of  $\tilde{\mathcal{C}}$ , among compositions of at most 5.

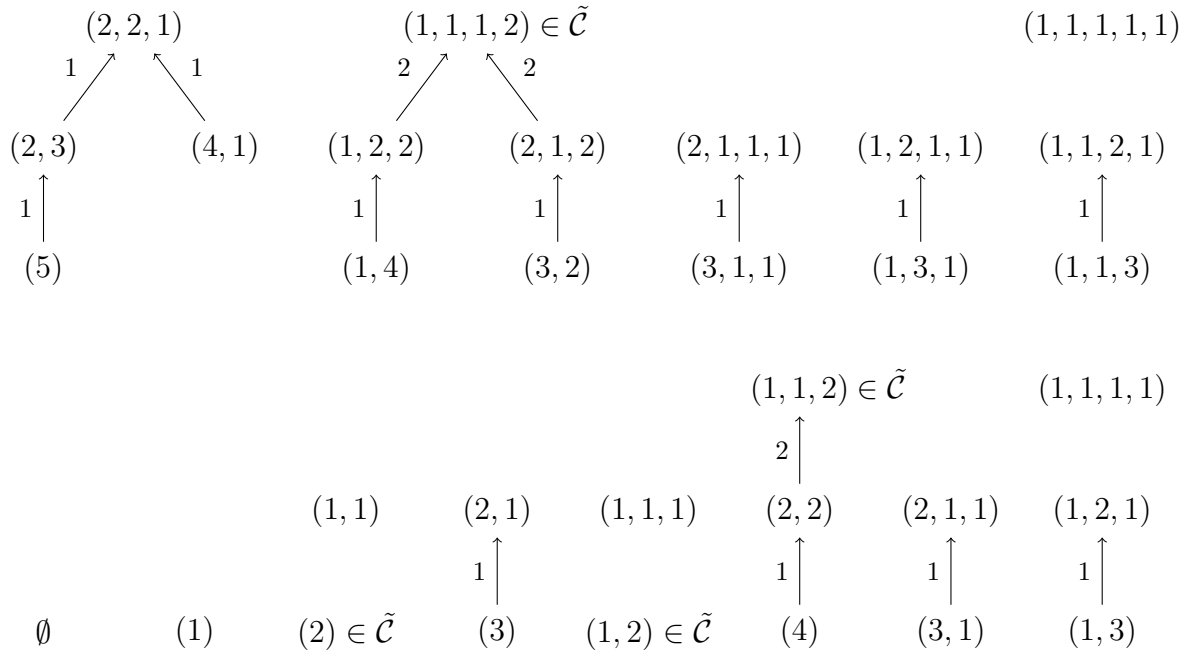


Figure 1.2: All relations  $\triangleright_1$  and  $\triangleright_2$ , and elements of  $\tilde{\mathcal{C}}$ , among compositions of at most 5.

**Theorem 3.** *The ideals  $\mathcal{K}^{\text{Pk}}$  and  $\mathcal{K}^{\text{pk}}$  are spanned (as  $\mathbb{Q}$ -vector spaces) in the following ways:*

- $\mathcal{K}^{\text{Pk}} = \text{span}\left(\{M_J + M_K : J \triangleright_1 K \text{ or } J \triangleright_2 K\} \cup \{M_J : J \in \tilde{\mathcal{C}}\}\right)$
- $\mathcal{K}^{\text{pk}} = \text{span}\left(\{M_J + M_K : J \triangleright_1 K \text{ or } J \triangleright_2 K\} \cup \{M_J : J \in \tilde{\mathcal{C}}\} \cup \{M_J - M_K : J \rightarrow_3 K\}\right)$

Therefore,  $\text{Pk}$  and  $\text{pk}$  are  $M$ -binomial.

We note that the shuffle algebra  $\mathcal{A}^{\text{Pk}}$  is isomorphic to Stembridge's algebra  $\Pi$  of peak quasisymmetric functions [17], which occupies an important position in the theory of combinatorial Hopf algebras [2]. Our results about  $\mathcal{K}^{\text{Pk}}$  can thus be translated into results about the canonical projection map from  $\text{QSym}$  to  $\Pi$ . Moreover,  $\mathcal{K}^{\text{Pk}}$  is the orthogonal complement of Nyman's peak algebra [13] (as a nonunital subalgebra of the noncommutative symmetric functions), and similarly  $\mathcal{K}^{\text{pk}}$  is the orthogonal complement of Schocker's Eulerian peak algebra [14].

Peaks and valleys in permutations are related via complementation, and this symmetry will be exploited to obtain from Theorem 2 an analogous result for the kernels of the valley set and valley number statistics. We will give the statement of this result later.

## 1.2. Outline

This paper is organized as follows. Section 2 focuses on background material, including exposition on descent statistics, the connection between quasisymmetric functions and shuffle-compatibility, and a result from linear algebra which will be used in our later proofs. In Section 3, we first prove a couple general results about spanning sets and linear independence in arbitrary vector spaces, which together imply Theorem 1. We then use Theorem 1 to produce our characterizations for the ideals  $\mathcal{K}^{\text{Pk}}$  and  $\mathcal{K}^{\text{pk}}$  given in Theorem 2, and also obtain linear bases for  $\mathcal{K}^{\text{Pk}}$  and  $\mathcal{K}^{\text{pk}}$  by trimming the spanning sets from Theorem 2 in a simple way. Section 4 will be devoted to the proof of Theorem 3, and Section 5 on the ideals  $\mathcal{K}^{\text{Val}}$  and  $\mathcal{K}^{\text{val}}$  for the valley set and valley number statistics. We end in Section 6 with a discussion of future directions of research.

All vector spaces, algebras, linear combinations, spans, and related notions are over the field  $\mathbb{Q}$  except in Section 3.1 (where we work over an arbitrary field of characteristic  $\neq 2$ ).

## 2. Preliminaries

### 2.1. Descent statistics

Recall that a permutation statistic  $\text{st}$  is a *descent statistic* if  $\text{st } \pi = \text{st } \sigma$  whenever  $\text{Comp } \pi = \text{Comp } \sigma$ . The descent set  $\text{Des}$  is clearly a descent statistic, and other examples of descent statistics include the following:

- The *descent number*  $\text{des}$  and *major index*  $\text{maj}$ . Given a permutation  $\pi$ , define

$$\text{des } \pi := |\text{Des } \pi| \quad \text{and} \quad \text{maj } \pi := \sum_{k \in \text{Des } \pi} k$$

to be its number of descents and its sum of descents, respectively.

- The *peak set*  $\text{Pk}$  and the *peak number*  $\text{pk}$ . Given  $\pi \in \mathfrak{P}_n$ , we say that  $i \in \{2, 3, \dots, n-1\}$  is a *peak* of  $\pi$  if  $\pi_{i-1} < \pi_i > \pi_{i+1}$ . Then  $\text{Pk } \pi$  is defined to be the set of peaks of  $\pi$  and  $\text{pk } \pi$  its number of peaks.
- The *exterior peak set*  $\text{Epk}$  and the *exterior peak number*  $\text{epk}$ . Given  $\pi \in \mathfrak{P}_n$ , we say that  $i \in [n]$  is an *exterior peak* of  $\pi$  if  $i$  is a peak of  $\pi$ , if  $i = 1$  and  $\pi_1 > \pi_2$ , if  $i = n$  and  $\pi_{n-1} < \pi_n$ , or if  $i = n = 1$ . Then  $\text{Epk } \pi$  is defined to be the set of exterior peaks of  $\pi$  and  $\text{epk } \pi$  its number of exterior peaks.

For example, if  $\pi = 713649$ , then  $\text{des } \pi = 2$ ,  $\text{maj } \pi = 5$ ,  $\text{Pk } \pi = \{4\}$ ,  $\text{pk } \pi = 1$ ,  $\text{Epk } \pi = \{1, 4, 6\}$ , and  $\text{epk } \pi = 3$ . Several additional descent statistics—associated with valleys, left peaks and right peaks—will be introduced later in this paper. Other descent statistics which we do not consider in this work include those based on double descents, alternating descents, and biruns; see [6, 18] for definitions.

Next, recall that the notation  $\text{Des } L$  refers to the descent set of any permutation with descent composition  $L$ , and from Equation (1.1), we have

$$\text{Des } L = \{j_1, j_1 + j_2, \dots, j_1 + j_2 + \dots + j_{m-1}\}$$

for  $L = (j_1, j_2, \dots, j_m)$ . Similarly, we can use Equation (1.2) to define  $\text{Comp}$  on subsets: given  $C = \{i_1 < i_2 < \dots < i_m\} \subseteq [n - 1]$ , let

$$\text{Comp } C := (i_1, i_2 - i_1, \dots, i_m - i_{m-1}, n - i_m).$$

Then  $\text{Des}$  and  $\text{Comp}$  are inverse bijections between compositions of  $n$  and subsets of  $[n - 1]$ .

For our characterizations of the ideals  $\mathcal{K}^{\text{Pk}}$  and  $\mathcal{K}^{\text{pk}}$ , it will be helpful to have an explicit formula for the  $\text{Pk}$  and  $\text{pk}$  statistics on compositions. The lemma below follows immediately from the fact that the peaks of a permutation occur precisely at the end of its non-final increasing runs of length at least 2.

**Lemma 2.1.** *Let  $L = (j_1, j_2, \dots, j_m)$  be a composition. Then:*

- (a)  $\text{Pk } L = \left\{ \sum_{i=1}^k j_i : j_k \geq 2 \text{ and } k \in [m - 1] \right\}$
- (b)  $\text{pk } L = |\{k \in [m - 1] : j_k \geq 2\}|$

## 2.2. Quasisymmetric functions and shuffle-compatibility

Quasisymmetric functions arose in the early work of Stanley as generating functions for  $P$ -partitions [15], were first defined and studied per se by Gessel [5], and are now ubiquitous in algebraic combinatorics. We review some elementary definitions and results surrounding quasisymmetric functions, emphasizing their role in the theory of shuffle-compatibility; see [16, Section 7.19], [10, Section 5], and [12] for further references.

Let  $x_1, x_2, \dots$  be commuting variables. A formal power series  $f \in \mathbb{Q}[[x_1, x_2, \dots]]$  of bounded degree is called a *quasisymmetric function* if for any positive integers  $a_1, a_2, \dots, a_k$ , if  $i_1 < i_2 < \dots < i_k$  and  $j_1 < j_2 < \dots < j_k$  then

$$[x_{i_1}^{a_1} x_{i_2}^{a_2} \dots x_{i_k}^{a_k}] f = [x_{j_1}^{a_1} x_{j_2}^{a_2} \dots x_{j_k}^{a_k}] f,$$

i.e., the monomials  $x_{i_1}^{a_1} x_{i_2}^{a_2} \dots x_{i_k}^{a_k}$  and  $x_{j_1}^{a_1} x_{j_2}^{a_2} \dots x_{j_k}^{a_k}$  have the same coefficients in  $f$ . Let  $\text{QSym}_n$  denote the vector space of quasisymmetric functions homogeneous of degree  $n$ , and let

$$\text{QSym} := \bigoplus_{n=0}^{\infty} \text{QSym}_n.$$

The monomial quasisymmetric functions  $\{M_L\}_{L \models n}$  and the fundamental quasisymmetric functions  $\{F_L\}_{L \models n}$  defined in the introduction are two bases of  $\text{QSym}_n$ , and since (for  $n \geq 1$ ) there are  $2^{n-1}$  compositions of  $n$ , it follows that  $\text{QSym}_n$  has dimension  $2^{n-1}$ .



Through the inverse bijections  $\text{Comp}$  and  $\text{Des}$ , we may also index the monomial and fundamental quasisymmetric functions by subsets  $C$  of  $[n - 1]$  in writing

$$M_{n,C} := M_{\text{Comp } C} \quad \text{and} \quad F_{n,C} := F_{\text{Comp } C},$$

and it will sometimes be convenient for us to do so.

Let us recall the change-of-basis formulas between the monomial and fundamental bases. First, we say that  $J \vDash n$  *refines* (or is a *refinement* of)  $K \vDash n$  if  $\text{Des } K \subseteq \text{Des } J$ . Informally, this amounts to saying that we can obtain  $K$  from  $J$  by combining some of its adjacent parts. For example, we have that  $J = (2, 1, 3, 1, 1, 2)$  refines  $K = (3, 5, 2)$  because  $\text{Des } J = \{2, 3, 6, 7, 8\}$  contains  $\text{Des } K = \{3, 8\}$ , and indeed we have  $2 + 1 = 3$  and  $3 + 1 + 1 = 5$ . Let us write  $J \leq K$  if  $J$  refines  $K$ . Then we have

$$F_L = \sum_{K \leq L} M_K, \quad \text{or equivalently,} \quad F_{n,C} = \sum_{C \subseteq B \subseteq [n-1]} M_{n,B}.$$

Part (a) of the next lemma then follows from inclusion-exclusion. Parts (b)–(c) appear as Propositions 5.10 (b)–(c) of [9].

**Lemma 2.2.** *Let  $C \subseteq [n - 1]$ .*

(a) *We have*

$$M_{n,C} = \sum_{C \subseteq B \subseteq [n-1]} (-1)^{|B \setminus C|} F_{n,B}.$$

(b) *Suppose that  $k \in [n - 1]$  and  $k \notin C$ . Then*

$$M_{n,C} + M_{n,C \cup \{k\}} = \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k \notin B}} (-1)^{|B \setminus C|} F_{n,B}.$$

(c) *Suppose that  $k \in [n - 1]$ ,  $k \notin C$ , and  $k - 1 \notin C \cup \{0\}$ . Then*

$$M_{n,C} + M_{n,C \cup \{k\}} = \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k, k-1 \notin B}} (-1)^{|B \setminus C|} (F_{n,B} - F_{n,B \cup \{k-1\}}).$$

Our present work on kernels of descent statistics only concerns the vector space structure of the quasisymmetric functions, but it is worthwhile to also discuss the ring structure of  $\text{QSym}$  as it will allow us to understand the background and motivation for studying these kernels. Let  $\pi \in \mathfrak{P}_m$  and  $\sigma \in \mathfrak{P}_n$  be *disjoint* permutations—that is, they have no letters in common. Then we say that  $\tau \in \mathfrak{P}_{m+n}$  is a *shuffle* of  $\pi$  and  $\sigma$  if both  $\pi$  and  $\sigma$  are subsequences of  $\tau$ , and we let  $\pi \sqcup \sigma$  denote the set of shuffles of  $\pi$  and  $\sigma$ . For example, given  $\pi = 13$  and  $\sigma = 42$ , we have

$$\pi \sqcup \sigma = \{1342, 1432, 1423, 4132, 4123, 4213\}.$$

The product of two fundamental quasisymmetric functions is given by

$$F_L F_K = \sum_{\tau \in \pi \sqcup \sigma} F_{\text{Comp } \tau} \tag{2.1}$$

where  $\pi$  and  $\sigma$  are any disjoint permutations satisfying  $\text{Comp } \pi = L$  and  $\text{Comp } \sigma = K$ . It follows that  $\text{QSym}$  is a graded subalgebra of  $\mathbb{Q}[[x_1, x_2, \dots]]$ .

In order for the product formula (2.1) to make sense, the multiset  $\{\text{Comp } \tau : \tau \in \pi \sqcup \sigma\}$  must only depend on the descent compositions of  $\pi$  and  $\sigma$ ; equivalently, in terms of descent sets, the multiset  $\{\text{Des } \tau : \tau \in \pi \sqcup \sigma\}$  only depends on  $\text{Des } \pi$ ,  $\text{Des } \sigma$ , and the lengths of  $\pi$  and  $\sigma$ . More generally, we say that a permutation statistic  $\text{st}$  is called *shuffle-compatible* if for any disjoint permutations  $\pi$  and  $\sigma$ , the multiset  $\{\text{st } \tau : \tau \in \pi \sqcup \sigma\}$  giving the distribution of  $\text{st}$  over  $\pi \sqcup \sigma$  depends only on  $\text{st } \pi$ ,  $\text{st } \sigma$ , and the lengths of  $\pi$  and  $\sigma$ . In other words, (2.1) implies that the descent set  $\text{Des}$  is a shuffle-compatible permutation statistic, which is implicit in Stanley’s theory of  $P$ -partitions [15]; the shuffle-compatibility of the statistics  $\text{des}$ ,  $\text{maj}$ , and  $(\text{des}, \text{maj})$  follow from Stanley’s work as well.

Before proceeding, we note that  $\text{QSym}$  additionally has the structure of a dendriform algebra [7] and a Hopf algebra (and is in fact the terminal object in the category of combinatorial Hopf algebras [2]). The dendriform structure of  $\text{QSym}$  is relevant to shuffle-compatibility and the kernels of descent statistics; this connection is not important for our present work but will be touched on in Section 6.

### 2.3. Shuffle algebras and the kernels $\mathcal{K}^{\text{st}}$

Motivated by the shuffle-compatibility results implicit in Stanley’s work on  $P$ -partitions, Gessel and Zhuang [6] formalized the notion of a shuffle-compatible permutation statistic and built a framework for investigating this phenomenon centered around the shuffle algebra of a shuffle-compatible statistic. Let us outline this construction below.

We say that permutations  $\pi$  and  $\sigma$  are *st-equivalent* if  $\text{st } \pi = \text{st } \sigma$  and  $|\pi| = |\sigma|$ . We write the  $\text{st}$ -equivalence class of  $\pi$  as  $[\pi]_{\text{st}}$ . For a shuffle-compatible statistic  $\text{st}$ , we associate to  $\text{st}$  a  $\mathbb{Q}$ -algebra in the following way. First, associate to  $\text{st}$  a  $\mathbb{Q}$ -vector space by taking as a basis the  $\text{st}$ -equivalence classes of permutations. We give this vector space a multiplication by taking

$$[\pi]_{\text{st}} [\sigma]_{\text{st}} = \sum_{\tau \in \pi \sqcup \sigma} [\tau]_{\text{st}},$$

which is well-defined if and only if  $\text{st}$  is shuffle-compatible. The resulting algebra  $\mathcal{A}^{\text{st}}$  is called the *shuffle algebra* of  $\text{st}$ . Observe that  $\mathcal{A}^{\text{st}}$  is graded by length—i.e.,  $[\pi]_{\text{st}}$  belongs to the  $n$ th graded component of  $\mathcal{A}^{\text{st}}$  if  $\pi$  has length  $n$ .

When  $\text{st}$  is a descent statistic, the notion of  $\text{st}$ -equivalence of permutations induces the notion of  $\text{st}$ -equivalence of compositions as defined in the introduction, so we can think of the basis elements of  $\mathcal{A}^{\text{st}}$  as being  $\text{st}$ -equivalence classes of compositions. From this perspective, it is evident from the product formula (2.1) that  $\mathcal{A}^{\text{Des}}$  is isomorphic to  $\text{QSym}$  with the basis of  $\text{Des}$ -equivalence classes corresponding to the fundamental quasisymmetric functions.

The following provides a necessary and sufficient condition for a descent statistic to be shuffle-compatible.

**Theorem 2.3** (Gessel–Zhuang [6, Theorem 4.3]). *A descent statistic  $\text{st}$  is shuffle-compatible if and only if there exists a  $\mathbb{Q}$ -algebra homomorphism  $\phi_{\text{st}}: \text{QSym} \rightarrow A$ , where  $A$  is a  $\mathbb{Q}$ -algebra with basis  $\{u_\alpha\}$  indexed by  $\text{st}$ -equivalence classes  $\alpha$  of compositions, such that  $\phi_{\text{st}}(F_L) = u_\alpha$  whenever  $L$  is in the  $\text{st}$ -equivalence class  $\alpha$ . When this is the case, the map  $u_\alpha \mapsto \alpha$  is a  $\mathbb{Q}$ -algebra isomorphism from  $A$  to  $\mathcal{A}^{\text{st}}$ .*

It follows from Theorem 2.3 that, whenever  $\text{st}$  is a shuffle-compatible descent statistic, the linear map  $p_{\text{st}}: \text{QSym} \rightarrow \mathcal{A}^{\text{st}}$  sending each  $F_L$  to the  $\text{st}$ -equivalence class of  $L$  is a  $\mathbb{Q}$ -algebra homomorphism with the same kernel as  $\phi_{\text{st}}$ ; this implies that  $\text{QSym} / \ker p_{\text{st}}$  is isomorphic to  $\mathcal{A}^{\text{st}}$  as algebras. Note that when  $\text{st}$  is not shuffle-compatible, it still holds that  $\text{QSym} / \ker p_{\text{st}}$  and  $\mathcal{A}^{\text{st}}$  are isomorphic as vector spaces.

Gessel and Zhuang [6] used Theorem 2.3 to give explicit descriptions of the shuffle algebras of a number of descent statistics—including  $\text{des}$ ,  $\text{maj}$ ,  $\text{Pk}$ ,  $\text{pk}$ , and  $(\text{pk}, \text{des})$ —which yield algebraic proofs for their shuffle-compatibility. In [8], Grinberg proved Gessel and Zhuang’s conjecture that the exterior peak set  $\text{Epk}$  is shuffle-compatible and gave a characterization of its shuffle algebra, introduced a strengthening of shuffle-compatibility called “LR-shuffle-compatibility” which is closely related to the dendriform algebra structure of  $\text{QSym}$ , and initiated the study of the kernels  $\mathcal{K}^{\text{st}}$ .

Recall that the *kernel*  $\mathcal{K}^{\text{st}}$  of a descent statistic  $\text{st}$  is the subspace

$$\mathcal{K}^{\text{st}} = \text{span}\{F_L - F_K : L \sim_{\text{st}} K\}$$

of  $\text{QSym}$ . Later on, we shall use the notation  $\mathcal{K}_n^{\text{st}}$  for the  $n$ th homogeneous component of  $\mathcal{K}^{\text{st}}$ , so that

$$\mathcal{K}_n^{\text{st}} = \text{span}\{F_L - F_K : L \sim_{\text{st}} K \text{ and } L, K \vDash n\}$$

and

$$\mathcal{K}^{\text{st}} = \bigoplus_{n=0}^{\infty} \mathcal{K}_n^{\text{st}}.$$

It is easy to see that  $\mathcal{K}^{\text{st}}$  is precisely the kernel of the linear map  $p_{\text{st}}$  defined above, hence the name “kernel”. The following is then a consequence of Theorem 2.3.

**Theorem 2.4** (Grinberg [8, Proposition 101]). *A descent statistic  $\text{st}$  is shuffle-compatible if and only if  $\mathcal{K}^{\text{st}}$  is an ideal of  $\text{QSym}$ . When this is the case,  $\mathcal{A}^{\text{st}}$  is isomorphic to  $\text{QSym} / \mathcal{K}^{\text{st}}$  as  $\mathbb{Q}$ -algebras.*

## 2.4. Linear expansions and triangularity

Lastly, we state a lemma concerning “invertibly triangular expansions” which will be used in our proof of Theorem 3. This lemma appears (in a slightly different yet equivalent form) in the Appendix to [10], which gives a treatment of some fundamental results from linear algebra for matrices whose rows and columns are indexed by arbitrary objects (rather

than numbers). In particular, we are interested in the case when both the rows and columns are indexed by elements of a finite poset  $S$ , such that the matrix is “invertibly triangular”—i.e., all the entries  $a_{s,s}$  on the “diagonal” are invertible and  $a_{s,t} = 0$  whenever we do not have  $t \leq s$ . (When working over a field, as we do, the condition that  $a_{s,s}$  is invertible is equivalent to  $a_{s,s} \neq 0$ .) Note that this reduces to the typical notion of an invertible lower-triangular  $n \times n$  matrix upon taking  $S = [n]$ .

A *family*  $(f_s)_{s \in S}$  indexed by a set  $S$  refers to an assignment of an object  $f_s$  to each  $s \in S$ . The objects in a family need not be distinct—i.e., we may have  $f_s = f_t$  for  $s \neq t$ . Roughly speaking, a family  $(e_s)_{s \in S}$  can be expanded invertibly triangularly in another family  $(f_s)_{s \in S}$  if we can write the  $e_s$  as linear combinations of the  $f_s$  such that the coefficients of these linear combinations form an invertibly triangular matrix in the sense described above. However, rather than giving the formal definition of an invertibly triangular expansion in terms of these generalized matrices, it is easier for our purposes to give the following equivalent definition.

Given a  $\mathbb{Q}$ -vector space  $V$ , a finite poset  $S$ , and two families  $(e_s)_{s \in S}$  and  $(f_s)_{s \in S}$  of elements of  $V$ , we say that  $(e_s)_{s \in S}$  *expands invertibly triangularly* in  $(f_s)_{s \in S}$  if, for each  $s \in S$ , we can write  $e_s$  as

$$e_s = c_s f_s + \sum_{t > s} c_t f_t \tag{2.2}$$

for some  $c_t \in \mathbb{Q}$  with  $c_s \neq 0$ . Importantly, two families have the same span if one can be expanded invertibly triangularly in another.

**Lemma 2.5** (Grinberg–Reiner [10, Corollary 11.1.19 (b)]). *Let  $V$  be a  $\mathbb{Q}$ -vector space,  $S$  a finite poset, and  $(e_s)_{s \in S}$  and  $(f_s)_{s \in S}$  two families of elements of  $V$ . If  $(e_s)_{s \in S}$  expands invertibly triangularly in  $(f_s)_{s \in S}$ , then  $\text{span}(e_s)_{s \in S} = \text{span}(f_s)_{s \in S}$ .*

We note that our definition of “expands invertibly triangularly” is the opposite of that in [10]—i.e., [10] has  $t < s$  in place of  $t > s$  in (2.2)—but it is clear that Lemma 2.5 still holds as we can simply reverse the order of the poset  $S$ .

### 3. Characterizations in terms of the fundamental basis

The purpose of this section is to provide proofs for Theorems 1 and 2. In proving Theorem 1, we will first establish a couple general results for arbitrary vector spaces.

#### 3.1. Proof of Theorem 1

Throughout this section, fix a field  $\mathbb{k}$  of characteristic  $\neq 2$  and let  $V$  be a vector space over  $\mathbb{k}$  with basis  $\{u_s\}_{s \in I}$  (where  $I$  is an index set for the basis). Let  $\sim$  be an equivalence relation on  $I$ , and let

$$R := \{(s, t) \in I^2 : s \sim t\} \quad \text{and} \quad \mathcal{U} := \{u_s - u_t : s \sim t\}.$$

Consider the subspace  $W$  of  $V$  spanned by the set  $\mathcal{U}$ . Note that every subset of  $\mathcal{U}$  can be written as

$$\mathcal{U}_S := \{u_s - u_t : (s, t) \in S\}$$

for some  $S \subseteq R$ . We associate to  $S$  a directed graph  $G_S$  with vertex set  $I$  and edge set  $S$ —i.e., there is an edge from  $s$  to  $t$  if and only if  $(s, t) \in S$ . As a special case, let  $G := G_R$ , so the connected components of  $G$  are simply the equivalence classes of  $I$  under  $\sim$ . More generally, the connected components of  $G_S$  refine the equivalence classes of  $I$ .

**Lemma 3.1.** *Let  $S \subseteq R$ . If  $s, t \in I$  are in the same connected component of  $G_S$ , then  $s \sim t$ .*

*Proof.* Let  $s, t \in I$  be in the same connected component of  $G_S$ . Then there is a sequence

$$s = s_0 \leftrightarrow s_1 \leftrightarrow \cdots \leftrightarrow s_k = t$$

where  $s_{i-1} \leftrightarrow s_i$  means that there is an edge in either direction (not necessarily both) between  $s_{i-1}$  and  $s_i$ . As a result, we have either  $(s_{i-1}, s_i) \in S$  or  $(s_i, s_{i-1}) \in S$  for each  $i \in [k]$ , which means that  $s_{i-1} \sim s_i$  for each  $i \in [k]$  and therefore  $s \sim t$  by transitivity.  $\square$

**Lemma 3.2.** *Let  $s, t \in I$  and  $S \subseteq R$ . Then  $u_s - u_t \in \text{span}\mathcal{U}_S$  if and only if  $s$  and  $t$  are in the same connected component of  $G_S$ .*

*Proof.* Suppose that  $s$  and  $t$  are in the same connected component of  $G_S$ . Then, as in the proof of Lemma 3.1, there is a sequence

$$s = s_0 \leftrightarrow s_1 \leftrightarrow \cdots \leftrightarrow s_k = t$$

where  $s_{i-1} \leftrightarrow s_i$  means that there is an edge in either direction (not necessarily both) between  $s_{i-1}$  and  $s_i$ . For each  $i \in [k]$ , we have  $s_{i-1} \leftrightarrow s_i$  which means that  $u_{s_{i-1}} - u_{s_i} \in \mathcal{U}_S$  or  $u_{s_i} - u_{s_{i-1}} \in \mathcal{U}_S$ . Thus  $u_{s_{i-1}} - u_{s_i} \in \text{span}\mathcal{U}_S$  for each  $i \in [k]$ , which implies

$$u_s - u_t = u_{s_0} - u_{s_k} = \sum_{i=1}^k (u_{s_{i-1}} - u_{s_i}) \in \text{span}\mathcal{U}_S.$$

Conversely, suppose that  $u_s - u_t \in \text{span}\mathcal{U}_S$ . Let  $f_s: V \rightarrow \mathbb{k}$  be the linear map defined by

$$f_s(u_r) = \begin{cases} 1, & \text{if } s \text{ and } r \text{ are in the same connected component of } G_S, \\ 0, & \text{otherwise.} \end{cases}$$

Recall that if  $u_p - u_q \in \mathcal{U}_S$ , then  $p$  and  $q$  are joined by an edge, which means  $f_s(u_p) = f_s(u_q)$ ; after all,  $s$  is either in the same connected component of  $G_S$  as  $p$  and  $q$  or it is not. Thus, all of the spanning vectors  $u_p - u_q$  of  $\mathcal{U}_S$  belong to the kernel of  $f_s$ , so  $\text{span}\mathcal{U}_S \subseteq \ker f_s$ . Since  $u_s - u_t \in \text{span}\mathcal{U}_S$ , it follows that  $u_s - u_t \in \ker f_s$ . Hence,  $f_s(u_t) = f_s(u_s) = 1$  which concludes the proof.  $\square$

**Theorem 3.3.** *Let  $S \subseteq R$ . Then  $W$  is spanned by  $\mathcal{U}_S$  if and only if the connected components of  $G_S$  are precisely the connected components of  $G$  (as sets of vertices, not as subgraphs).*

*Proof.* Suppose that the connected components of  $G_S$  are those of  $G$ . Fix  $s, t \in I$  satisfying  $s \sim t$ . Then  $s$  and  $t$  belong to the same connected component of  $G$  and thus the same connected component of  $G_S$ , so  $u_s - u_t \in \text{span } \mathcal{U}_S$  by Lemma 3.2. Since  $u_s - u_t$  was arbitrarily taken from  $\mathcal{U}$ , which spans  $W$ , it follows by linearity that  $W$  is spanned by  $\mathcal{U}_S$ .

Conversely, suppose that  $W$  is spanned by  $\mathcal{U}_S$ . To show that  $G_S$  and  $G$  have the same connected components, it suffices to show that for any  $s, t \in I$ , if  $s$  and  $t$  are in the same connected component of  $G$  then they are in the same connected component of  $G_S$ . As such, fix  $s, t \in I$  which belong to the same connected component of  $G$ . Then  $s \sim t$ , and so  $u_s - u_t \in \mathcal{U} \subseteq W = \text{span } \mathcal{U}_S$ . By Lemma 3.2,  $s$  and  $t$  are in the same connected component of  $G_S$ , and we are done.  $\square$

**Theorem 3.4.** *Let  $S \subseteq R$ . Then  $\mathcal{U}_S$  is linearly independent if and only if  $G_S$  is a forest.*

*Proof.* Suppose that  $G_S$  is a forest, and assume toward contradiction that  $\mathcal{U}_S$  is linearly dependent. Then there exists  $(s, t) \in S$  for which  $u_s - u_t$  is a linear combination of other elements of  $\mathcal{U}_S$ . Let  $S' := S \setminus \{(s, t)\}$ , so that  $u_s - u_t \in \text{span } \mathcal{U}_{S'}$ . Applying Lemma 3.2 to  $S'$ , it follows that there is a path from  $s$  to  $t$  in the underlying undirected graph of  $G_{S'}$ . Now, recall that there is an edge between  $s$  and  $t$  in  $G_S$  because  $(s, t) \in S$ ; combining this edge with the path from  $s$  to  $t$  in  $G_{S'}$  yields a cycle in  $G_S$ , which contradicts  $G_S$  being a forest. Therefore,  $\mathcal{U}_S$  is linearly independent.

Conversely, suppose that  $G_S$  is not a forest. Then the undirected graph  $G_S$  has a cycle  $(s_1, s_2, \dots, s_k)$ , so each of the vectors

$$u_{s_1} - u_{s_2}, u_{s_2} - u_{s_3}, \dots, u_{s_{k-1}} - u_{s_k}, u_{s_k} - u_{s_1}$$

belongs to  $\mathcal{U}_S$  up to sign (either it belongs to  $\mathcal{U}_S$  or its negative does). However, the sum of these vectors is equal to 0, so  $\mathcal{U}_S$  is linearly dependent.  $\square$

We can now recover Theorem 1 from Theorems 3.3 and 3.4.

*Proof of Theorem 1.* Take  $\mathbb{k} = \mathbb{Q}$ ,  $V = \text{QSym}$ ,  $I = \mathcal{C}$ , the basis  $\{u_s\}$  to be the fundamental basis of  $\text{QSym}$ , and  $\sim$  to be st-equivalence of compositions. Part (a) then follows from Theorem 3.3, and part (b) from Theorem 3.4.  $\square$

### 3.2. Application to the peak set—proof of Theorem 2 (a)

We shall now apply Theorem 1 to obtain spanning sets for the kernels  $\mathcal{K}^{\text{Pk}}$  and  $\mathcal{K}^{\text{pk}}$  as given in Theorem 2. Let us begin with the peak set statistic. Recall the notations  $\rightarrow_1$  and  $\rightarrow_2$  defined in the introduction; for convenience, let us write  $J \rightarrow_{\text{Pk}} K$  if  $J \rightarrow_1 K$  or  $J \rightarrow_2 K$ .

**Lemma 3.5.** *If  $J \rightarrow_{\text{Pk}} K$ , then  $J$  and  $K$  are Pk-equivalent.*

*Proof.* It is immediate from the definitions of  $\rightarrow_1$  and  $\rightarrow_2$  that  $J \rightarrow_{\text{Pk}} K$  implies  $|J| = |K|$ . So, it remains to show that  $J \rightarrow_{\text{Pk}} K$  implies  $\text{Pk } J = \text{Pk } K$ . Suppose that  $J \rightarrow_1 K$ , so that there exists  $l \in [m]$  for which  $j_l > 2$  and

$$K = (j_1, \dots, j_{l-1}, 1, j_l - 1, j_{l+1}, \dots, j_m).$$

Since  $j_l > 2$ , we have that  $j_l - 1 \geq 2$  and thus  $\text{Pk } J = \text{Pk } K$  by Lemma 2.1 (a). The proof for the case  $J \rightarrow_2 K$  is similar.  $\square$

We are now ready to prove Theorem 2 (a).

*Proof of Theorem 2 (a).* We shall apply Theorem 1 (a) to  $\text{st} = \text{Pk}$ . By Lemma 3.5, we may take  $S = \{(J, K) : J \rightarrow_{\text{Pk}} K\}$ . We seek to show that  $J \sim_{\text{Pk}} K$  if and only if  $J$  and  $K$  are in the same connected component of  $G_S$ , and in light of Lemma 3.1, it remains to prove the forward direction.

Suppose that  $J \sim_{\text{Pk}} K$ . Let  $J_\alpha$  be the composition obtained from  $J$  by replacing each part  $j_l > 2$  with the parts of the composition  $(1^{j_l-2}, 2) \vDash j_l$ . For example, if  $J = (3, 2, 4, 1)$  then  $J_\alpha = (1, 2, 2, 1, 1, 2, 1)$ . Then there is a sequence of compositions satisfying

$$J \rightarrow_1 J_1 \rightarrow_1 J_2 \rightarrow_1 \cdots \rightarrow_1 J_\alpha.$$

Continuing the example above, we have

$$J = (3, 2, 4, 1) \rightarrow_1 (1, 2, 2, 4, 1) \rightarrow_1 (1, 2, 2, 1, 3, 1) \rightarrow_1 (1, 2, 2, 1, 1, 2, 1) = J_\alpha.$$

It follows that  $J$  and  $J_\alpha$  are in the same connected component of  $G_S$ , and similarly with  $K$  and  $K_\alpha$ .

Next, if 2 is the final part of  $J_\alpha$ , then let  $J_\beta$  be the composition obtained from  $J_\alpha$  by replacing the final part 2 with two 1s. For instance, if  $J_\alpha = (1, 2, 1, 2, 2)$  then  $J_\beta = (1, 2, 1, 2, 1, 1)$ . Otherwise, if  $J_\alpha$  ends with a 1, then set  $J_\beta := J_\alpha$ . Note that either  $J_\alpha = J_\beta$  or  $J_\alpha \rightarrow_2 J_\beta$ ; either way,  $J_\alpha$  is in the same connected component as  $J_\beta$ , and thus  $J$  is as well. The same holds for  $K$  and  $K_\beta$ . We then have  $J_\beta \sim_{\text{Pk}} J \sim_{\text{Pk}} K \sim_{\text{Pk}} K_\beta$  by Lemma 3.1, so  $\text{Pk } J_\beta = \text{Pk } K_\beta$  and  $|J_\beta| = |K_\beta|$ .

We claim that  $J_\beta = K_\beta$ ; this will imply that  $J$  and  $K$  are in the same connected component. Assume by contradiction that  $J_\beta \neq K_\beta$ . Let  $l$  be the position of the first part where  $J_\beta$  and  $K_\beta$  differ. Since  $J_\beta$  and  $K_\beta$  have all parts 1 or 2, we may assume without loss of generality that the  $l$ th part of  $J_\beta$  is a 1 and that the  $l$ th part of  $K_\beta$  is a 2. Note that the  $l$ th part of  $J_\beta$  or  $K_\beta$  cannot be its final part; after all,  $K_\beta$  cannot end with a 2 by construction, and if the  $l$ th part of  $J_\beta$  is its final part then we would have  $|J_\beta| < |K_\beta|$ . Thus, Lemma 2.1 (a) implies that the sum of the first  $l$  parts of  $K_\beta$  is an element of  $\text{Pk } K_\beta$  but not of  $\text{Pk } J_\beta$ , which contradicts  $\text{Pk } J_\beta = \text{Pk } K_\beta$ . Therefore,  $J_\beta = K_\beta$ .

We have shown that the connected components of  $G_S$  are precisely the Pk-equivalence classes of compositions; hence the result follows from Theorem 1 (a).  $\square$

The spanning set for  $\mathcal{K}^{\text{Pk}}$  provided by Theorem 2 (a) is not a basis for  $\mathcal{K}^{\text{Pk}}$ , as the corresponding directed graph is not a forest. For example, upon revisiting Figure 1.1, we see that there is a cycle formed by the relations

$$(3, 2) \rightarrow_1 (1, 2, 2) \rightarrow_2 (1, 2, 1, 1) \xleftarrow{1} (3, 1, 1) \xleftarrow{2} (3, 2).$$

One way that we can obtain a basis from Theorem 2 (a) is to specify that we must “split” the first entry that we are able to. Then, for example, we no longer have  $(3, 2) \rightarrow_2 (3, 1, 1)$

because we are forced to split the first part, leading to  $(1, 2, 2)$ . More formally, we have the following: Given compositions  $J = (j_1, j_2, \dots, j_m)$  and  $K$ , let us write  $J \rightarrow_{\text{pk}} K$  if

$$l = \min\{i : j_i > 2, \text{ or } i = m \text{ and } j_m = 2\}$$

exists and

$$K = (j_1, \dots, j_{l-1}, 1, j_l - 1, j_{l+1}, \dots, j_m).$$

**Theorem 3.6.** *The set  $\{F_J - F_K : J \rightarrow_{\text{pk}} K\}$  is a linear basis of the ideal  $\mathcal{K}^{\text{Pk}}$ .*

*Proof.* Let  $S = \{(J, K) : J \rightarrow_{\text{pk}} K\}$ , so that we wish to show that  $\mathcal{F}_S^{\text{Pk}}$  is a basis of  $\mathcal{K}^{\text{Pk}}$ . We shall first argue that  $G_S$  is a forest. First, observe that  $G_S$  has no directed cycles; after all,  $J \rightarrow_{\text{pk}} K$  implies that  $K$  has more parts than  $J$ . Therefore,  $G_S$  can only fail to be a forest if it has an undirected cycle which is not a directed cycle, and it is not hard to see that this is only possible if two edges of  $G_S$  have the same tail. (Given an undirected cycle, any attempt to orient the edges of this cycle will either result in a directed cycle or a vertex with outdegree 2.) By definition of  $\rightarrow_{\text{pk}}$ , for every composition  $J$  there is at most one  $K$  for which  $J \rightarrow_{\text{pk}} K$ , which means that no vertex of  $G_S$  is the tail of more than one edge. Therefore,  $G_S$  is a forest which implies that  $\mathcal{F}_S^{\text{Pk}}$  is linearly independent by Theorem 1 (b).

The proof of Theorem 2 (a) can be easily adapted to show that  $\mathcal{F}_S^{\text{Pk}}$  is a spanning set for  $\mathcal{K}^{\text{Pk}}$  as well. Indeed, we already have  $J_\alpha \rightarrow_{\text{pk}} J_\beta$  and  $K_\alpha \rightarrow_{\text{pk}} K_\beta$ , and when forming the sequences  $J \rightarrow_1 J_1 \rightarrow_1 J_2 \rightarrow_1 \dots \rightarrow_1 J_\alpha$  and  $K \rightarrow_1 K_1 \rightarrow_1 K_2 \rightarrow_1 \dots \rightarrow_1 K_\alpha$ , we can require that the parts be split from left to right so that

$$J \rightarrow_{\text{pk}} J_1 \rightarrow_{\text{pk}} J_2 \rightarrow_{\text{pk}} \dots \rightarrow_{\text{pk}} J_\alpha \quad \text{and} \quad K \rightarrow_{\text{pk}} K_1 \rightarrow_{\text{pk}} K_2 \rightarrow_{\text{pk}} \dots \rightarrow_{\text{pk}} K_\alpha.$$

The rest of the proof proceeds in the same way. □

### 3.3. Application to the peak number—proof of Theorem 2 (b)

Similar to the notation  $\rightarrow_{\text{pk}}$  defined in Section 3.2, let us write  $J \rightarrow_{\text{pk}} K$  if  $J \rightarrow_1 K$ ,  $J \rightarrow_2 K$ , or  $J \rightarrow_3 K$ .

**Lemma 3.7.** *If  $J \rightarrow_{\text{pk}} K$ , then  $J$  and  $K$  are pk-equivalent.*

*Proof.* If  $J \rightarrow_1 K$  or  $J \rightarrow_2 K$ , then it follows from Lemma 3.5 that  $J$  and  $K$  are Pk-equivalent and thus pk-equivalent. If  $J \rightarrow_3 K$ , then  $K$  has the same parts as  $J$  but listed in a different order, which implies  $\text{pk } J = \text{pk } K$  by Lemma 2.1 (b) as well as  $|J| = |K|$ . □

We now proceed to the proof of Theorem 2 (b).

*Proof of Theorem 2 (b).* We follow the same approach taken in the proof of part (a). Let  $S = \{(J, K) : J \rightarrow_{\text{pk}} K\}$ , which is a subset of  $\{(J, K) : J \sim_{\text{pk}} K\}$  by Lemma 3.7. From Lemma 3.1, we know that if  $J$  and  $K$  are in the same connected component of  $G_S$  then  $J \sim_{\text{pk}} K$ , so it remains to prove the converse.



Suppose that  $J \sim_{\text{pk}} K$ . Define  $J_\beta$  and  $K_\beta$  in the same way as in the proof of part (a); as before,  $J$  and  $J_\beta$  are in the same connected component and the same is true for  $K$  and  $K_\beta$ . Recall that  $J_\beta$  and  $K_\beta$  have all parts 1 and 2, and end with a 1. Let  $J_\gamma$  be the composition  $(2^b, 1^a)$  where  $a > 0$  is the number of 1s in  $J_\beta$  and  $b$  is the number of 2s in  $J_\beta$ . Then there is a sequence of compositions satisfying

$$J_\beta \rightarrow_3 J_1 \rightarrow_3 J_2 \rightarrow_3 \cdots \rightarrow_3 J_\gamma.$$

For example, if  $J_\beta = (1, 2, 2, 1, 1)$ , then

$$J_\beta = (1, 2, 2, 1, 1) \rightarrow_3 (2, 1, 2, 1, 1) \rightarrow_3 (2, 2, 1, 1, 1) = J_\gamma.$$

Note that  $J_\gamma$  is in the same connected component of  $G_S$  as  $J_\beta$  and thus  $J$ , and similarly with  $K_\gamma$  and  $K$ . Therefore,  $J \sim_{\text{pk}} J_\gamma$  and  $K \sim_{\text{pk}} K_\gamma$ . Along with  $J \sim_{\text{pk}} K$ , these pk-equivalences imply that  $J_\gamma \sim_{\text{pk}} K_\gamma$ , so  $\text{pk } J_\gamma = \text{pk } K_\gamma$  and  $|J_\gamma| = |K_\gamma|$ . From Lemma 2.1 (b), we know that the number of peaks is equal to the number of non-final parts of size at least 2, and since both  $J_\gamma$  and  $K_\gamma$  consist of a sequence of 2s followed by a sequence of 1s, we conclude that  $J_\gamma = K_\gamma$ . Hence,  $J$  and  $K$  are in the same connected component of  $G_S$ .

Since the connected components of  $G_S$  are precisely the pk-equivalence classes of compositions, applying Theorem 1 (a) yields the desired result.  $\square$

Like with  $\mathcal{K}^{\text{Pk}}$ , we shall trim the spanning set for  $\mathcal{K}^{\text{pk}}$  provided by Theorem 2 (b) to yield a basis for this ideal. Given compositions  $J = (j_1, j_2, \dots, j_m)$  and  $K$ , we write  $J \rightarrow_{\text{pk}} K$  if either  $J \rightarrow_{\text{Pk}} K$ —or, in the case that  $l$  (as in the definition of  $\rightarrow_{\text{Pk}}$ ) does not exist, if

$$k = \min\{i : j_i = 1 \text{ and } j_{i+1} = 2\}$$

exists and

$$K = (j_1, \dots, j_{k-1}, j_{k+1}, j_k, j_{k+2}, \dots, j_m).$$

Informally speaking, here we are also requiring the “swapping” ( $\rightarrow_3$ ) to occur from left to right and only after all of the “splitting” ( $\rightarrow_1$  and  $\rightarrow_2$ ) has taken place.

Given a composition  $J = (j_1, j_2, \dots, j_m)$ , let us call the pair  $(j_k, j_l)$  an *inversion* of the composition  $J$  if  $1 \leq k < l \leq m$  and  $j_k > j_l$ .

**Theorem 3.8.** *The set  $\{F_J - F_K : J \rightarrow_{\text{pk}} K\}$  is a linear basis of the ideal  $\mathcal{K}^{\text{pk}}$ .*

*Proof.* Let  $S = \{(J, K) : J \rightarrow_{\text{pk}} K\}$ ; we wish to show that  $\mathcal{F}_S^{\text{pk}}$  is a basis of  $\mathcal{K}^{\text{pk}}$ . By the same reasoning as in the proof of Theorem 3.6,  $G_S$  cannot have a directed cycle containing any edges of the form  $J \rightarrow_{\text{Pk}} K$ . On the other hand, we also cannot have a directed cycle only containing edges of the form  $J \rightarrow_3 K$ , as when  $J \rightarrow_3 K$ , the number of inversions of  $K$  is greater than that of  $J$ . Therefore,  $G_S$  does not have any directed cycles. The same reasoning from the proof of Theorem 3.6 shows that  $G_S$  has no undirected cycles either. Thus,  $G_S$  is a forest, and it follows from Theorem 1 (b) that  $\mathcal{F}_S^{\text{pk}}$  is linearly independent.

The proof that  $\mathcal{F}_S^{\text{pk}}$  spans  $\mathcal{K}^{\text{pk}}$  is similar to that of the analogous result for Pk (again, see the proof of Theorem 3.6); we omit the details.  $\square$

## 4. Characterizations in terms of the monomial basis

### 4.1. The peak set—proof of Theorem 3 (a)

We now turn our attention to characterizing the ideals  $\mathcal{K}^{\text{Pk}}$  and  $\mathcal{K}^{\text{pk}}$  in terms of the monomial basis, beginning with  $\mathcal{K}^{\text{Pk}}$ . Our proofs here are based on Grinberg's proof for the analogous result on  $\mathcal{K}^{\text{Epk}}$ . Let us first introduce some notation which will simplify the presentation of our proofs.

Recall that  $\mathcal{K}_n^{\text{st}}$  denotes the  $n$ th homogeneous component of  $\mathcal{K}^{\text{st}}$ . Define

$$\begin{aligned}\mathcal{F}_n^{\text{Pk}} &:= \{F_J - F_K : J \rightarrow_{\text{Pk}} K \text{ and } J, K \vDash n\} \quad \text{and} \\ \mathcal{M}_n^{\text{Pk}} &:= \{M_J + M_K : J \triangleright_1 K \text{ or } J \triangleright_2 K, \text{ and } J, K \vDash n\} \cup \{M_{(1^{n-2}, 2)}\}.\end{aligned}$$

We know from Theorem 2 (a) that  $\mathcal{K}_n^{\text{Pk}} = \text{span } \mathcal{F}_n^{\text{Pk}}$  for all  $n$ , and our present goal is to show that  $\mathcal{K}_n^{\text{Pk}} = \text{span } \mathcal{M}_n^{\text{Pk}}$  for all  $n$ .

Let

$$\begin{aligned}\Omega_{n,1} &:= \{(C, k) : C \subsetneq [n-1], k \notin C, k-1 \notin C, k-2 \in C \cup \{0\}\}, \\ \Omega_{n,2} &:= \{(C, k) : C \subsetneq [n-2], n-2 \in C, k \notin C, k+1 \in C, k-1 \in C \cup \{0\}\}, \text{ and} \\ \Omega_{n,3} &:= \{([n-2], n-1)\}\end{aligned}$$

be subsets of  $2^{[n-1]} \times [n-1]$ , where  $2^{[n-1]}$  is the power set of  $[n-1]$ . Here we let  $[0]$  be the empty set, so that  $\Omega_{2,3} = \{(\emptyset, 1)\}$ . Note that  $\Omega_{n,1}$  is empty for  $n \leq 2$ ,  $\Omega_{n,2}$  is empty for  $n \leq 3$ , and  $\Omega_{n,3}$  is empty for  $n \leq 1$ . It is easy to check that the sets  $\Omega_{n,1}$ ,  $\Omega_{n,2}$ , and  $\Omega_{n,3}$  are disjoint for any fixed  $n$ ; let

$$\Omega_n := \Omega_{n,1} \sqcup \Omega_{n,2} \sqcup \Omega_{n,3}.$$

Given  $(C, k) \in \Omega_n$ , define the quasisymmetric functions  $\mathbf{f}_{C,k}$  and  $\mathbf{m}_{C,k}$ <sup>3</sup> by

$$\mathbf{f}_{C,k} = \begin{cases} F_{n,C} - F_{n,C \cup \{k-1\}}, & \text{if } (C, k) \in \Omega_{n,1}, \\ F_{n,C} - F_{n,C \cup \{n-1\}}, & \text{if } (C, k) \in \Omega_{n,2} \text{ or } (C, k) \in \Omega_{n,3}, \end{cases}$$

and

$$\mathbf{m}_{C,k} = \begin{cases} M_{n,C} + M_{n,C \cup \{k\}}, & \text{if } (C, k) \in \Omega_{n,1} \text{ or } (C, k) \in \Omega_{n,2}, \\ M_{n,C}, & \text{if } (C, k) \in \Omega_{n,3}. \end{cases}$$

In what follows, we will consider the families  $(\mathbf{f}_{C,k})_{(C,k) \in \Omega_n}$  and  $(\mathbf{m}_{C,k})_{(C,k) \in \Omega_n}$ . As noted in Section 2.4, a family can admit repeated members. For instance, given  $(C, k_1), (C, k_2) \in \Omega_{n,2}$  where  $k_1 \neq k_2$ , we have  $\mathbf{f}_{C,k_1} = \mathbf{f}_{C,k_2}$ . This is in contrast to  $\mathcal{F}_n^{\text{Pk}}$  and  $\mathcal{M}_n^{\text{Pk}}$ , which have no repeated elements by virtue of being sets.

In order to prove Theorem 3 (a), we first prove  $\text{span}(\mathbf{f}_{C,k})_{(C,k) \in \Omega_n} = \text{span } \mathcal{F}_n^{\text{Pk}}$  and  $\text{span}(\mathbf{m}_{C,k})_{(C,k) \in \Omega_n} = \text{span } \mathcal{M}_n^{\text{Pk}}$ , followed by  $\text{span}(\mathbf{f}_{C,k})_{(C,k) \in \Omega_n} = \text{span}(\mathbf{m}_{C,k})_{(C,k) \in \Omega_n}$ . In doing so, we will make repeated use of the correspondence between compositions of  $n$  and subsets of  $[n-1]$ , but many of the details will be omitted as they are straightforward to verify from the relevant definitions yet distract from the main essence of the proof.

<sup>3</sup>The  $\mathbf{f}_{C,k}$  and  $\mathbf{m}_{C,k}$  depend on  $n$ , but we suppress  $n$  from their notation for the sake of simplicity.

**Proposition 4.1.** For all  $n \geq 0$ , we have  $\text{span}(\mathbf{f}_{C,k})_{(C,k) \in \Omega_n} = \text{span} \mathcal{F}_n^{\text{Pk}}$ .

*Proof.* Fix  $(C, k) \in \Omega_n$ . The following are readily checked:

- (1) If  $(C, k) \in \Omega_{n,1}$ , then  $\text{Comp } C \rightarrow_1 \text{Comp}(C \cup \{k-1\})$  and so taking  $J = \text{Comp } C$  and  $K = \text{Comp}(C \cup \{k-1\})$  yields

$$\mathbf{f}_{C,k} = F_{n,C} - F_{n,C \cup \{k-1\}} = F_J - F_K \in \mathcal{F}_n^{\text{Pk}}.$$

- (2) If  $(C, k) \in \Omega_{n,2}$  or  $(C, k) \in \Omega_{n,3}$ , then  $\text{Comp } C \rightarrow_2 \text{Comp}(C \cup \{n-1\})$  and so taking  $J = \text{Comp } C$  and  $K = \text{Comp}(C \cup \{n-1\})$  implies

$$\mathbf{f}_{C,k} = F_{n,C} - F_{n,C \cup \{n-1\}} = F_J - F_K \in \mathcal{F}_n^{\text{Pk}}.$$

Thus  $\mathbf{f}_{C,k} \in \mathcal{F}_n^{\text{Pk}}$  for all  $(C, k) \in \Omega_n$ , which implies  $\text{span}(\mathbf{f}_{C,k})_{(C,k) \in \Omega_n} \subseteq \text{span} \mathcal{F}_n^{\text{Pk}}$  by linearity.

For the reverse inclusion, fix  $J, K \models n$  satisfying  $J \rightarrow_{\text{Pk}} K$ . As above, we consider separate cases:

- (1) Suppose  $J \rightarrow_1 K$ , and let us write  $J = (j_1, j_2, \dots, j_m)$  and  $K = (j_1, \dots, j_{l-1}, 1, j_l - 1, j_{l+1}, \dots, j_m)$  where  $j_l > 2$ . Then  $\text{Des } K = (\text{Des } J) \cup \{k-1\}$  where  $k-1 = j_1 + \dots + j_{l-1} + 1$ , and  $(\text{Des } J, k) \in \Omega_{n,1}$ . Hence

$$F_J - F_K = F_{n, \text{Des } J} - F_{n, (\text{Des } J) \cup \{k-1\}} = \mathbf{f}_{\text{Des } J, k}.$$

- (2) Suppose  $J \rightarrow_2 K$ , and write  $J = (j_1, j_2, \dots, j_{m-1}, 2)$  and  $K = (j_1, \dots, j_{m-1}, 1, 1)$ . Then  $\text{Des } K = (\text{Des } J) \cup \{n-1\}$ . We now consider three subcases:

- (a) Suppose that  $j_a = 2$  for some  $1 \leq a \leq m-1$ , and let  $k = j_1 + \dots + j_a - 1$ . Then  $(\text{Des } J, k) \in \Omega_{n,2}$ , so

$$F_J - F_K = F_{n, \text{Des } J} - F_{n, (\text{Des } J) \cup \{n-1\}} = \mathbf{f}_{\text{Des } J, k}.$$

- (b) Suppose that none of the parts  $j_1, \dots, j_{m-1}$  are equal to 2 but that  $j_a > 2$  for some  $a \in [m-1]$ . Let  $J_i$  be the composition obtained from  $J$  by replacing  $j_a$  with the parts of the composition  $(1^i, j_a - i)$ , and define  $K_i$  in the same way. Then we have

$$\begin{aligned} J &= J_0 \rightarrow_1 J_1 \rightarrow_1 J_2 \rightarrow_1 \dots \rightarrow_1 J_{j_a-2} \quad \text{and} \\ K &= K_0 \rightarrow_1 K_1 \rightarrow_1 K_2 \rightarrow_1 \dots \rightarrow_1 K_{j_a-2}, \end{aligned}$$

and also  $J_{j_a-2} \rightarrow_2 K_{j_a-2}$ . Furthermore, by the telescope principle, we have

$$F_J - F_K = \sum_{i=1}^{j_a-2} (F_{J_{i-1}} - F_{J_i}) + (F_{J_{j_a-2}} - F_{K_{j_a-2}}) - \sum_{i=1}^{j_a-2} (F_{K_{i-1}} - F_{K_i}).$$

From Case (1) above, we know that each of the terms  $F_{J_{i-1}} - F_{J_i}$  and  $F_{K_{i-1}} - F_{K_i}$  belongs to the family  $(\mathbf{f}_{C,k})_{(C,k) \in \Omega_n}$ , and from Case (2a), the same is true for  $F_{J_{j_a-2}} - F_{K_{j_a-2}}$ . Therefore, we have  $F_J - F_K \in \text{span}(\mathbf{f}_{C,k})_{(C,k) \in \Omega_n}$ .

- (c) Suppose that all of the parts  $j_1, \dots, j_{m-1}$  are equal to 1. Then  $\text{Des } J = [n - 2]$  and  $(\text{Des } J, n - 1) \in \Omega_{n,3}$ , so

$$F_J - F_K = F_{n,[n-2]} - F_{n,[n-2] \cup \{n-1\}} = \mathbf{f}_{\text{Des } J, n-1}.$$

In all cases we have  $F_J - F_K \in \text{span}(\mathbf{f}_{C,k})_{(C,K) \in \Omega_n}$ , so  $\text{span } \mathcal{F}_n^{\text{Pk}} \subseteq \text{span}(\mathbf{f}_{C,k})_{(C,K) \in \Omega_n}$  by linearity. We conclude that the two spans are equal.  $\square$

**Proposition 4.2.** *For all  $n \geq 0$ , we have  $\text{span}(\mathbf{m}_{C,k})_{(C,k) \in \Omega_n} = \text{span } \mathcal{M}_n^{\text{Pk}}$ .*

*Proof.* Fix  $(C, k) \in \Omega_n$ . Then one can readily verify the following:

- (1) If  $(C, k) \in \Omega_{n,1}$ , then  $\text{Comp } C \triangleright_1 \text{Comp}(C \cup \{k\})$  and so taking  $J = \text{Comp } C$  and  $K = \text{Comp}(C \cup \{k\})$  yields

$$\mathbf{m}_{C,k} = M_{n,C} + M_{n,C \cup \{k\}} = M_J + M_K \in \mathcal{M}_n^{\text{Pk}}.$$

- (2) If  $(C, k) \in \Omega_{n,2}$ , then  $\text{Comp } C \triangleright_2 \text{Comp}(C \cup \{k\})$  and so taking  $J = \text{Comp } C$  and  $K = \text{Comp}(C \cup \{k\})$  yields

$$\mathbf{m}_{C,k} = M_{n,C} + M_{n,C \cup \{k\}} = M_J + M_K \in \mathcal{M}_n^{\text{Pk}}.$$

- (3) If  $(C, k) \in \Omega_{n,3}$ , then  $\text{Comp } C = (1^{n-2}, 2)$  and so

$$\mathbf{m}_{C,k} = M_{n,C} = M_{(1^{n-2}, 2)} \in \mathcal{M}_n^{\text{Pk}}.$$

Since  $\mathbf{m}_{C,k} \in \mathcal{M}_n^{\text{Pk}}$  for all  $(C, k) \in \Omega_n$ , we have  $\text{span}(\mathbf{m}_{C,k})_{(C,k) \in \Omega_n} \subseteq \text{span } \mathcal{M}_n^{\text{Pk}}$  by linearity.

Conversely, fix  $J, K \vDash n$  satisfying  $J \triangleright_1 K$  or  $J \triangleright_2 K$ .

- (1) Suppose  $J \triangleright_1 K$ , so that we may write  $J = (j_1, j_2, \dots, j_m)$  and  $K = (j_1, \dots, j_{l-1}, 2, j_l - 2, j_{l+1}, \dots, j_m)$  where  $j_l > 2$ . Taking  $k = j_1 + \dots + j_{l-1} + 2$ , we have  $\text{Des } K = (\text{Des } J) \cup \{k\}$  and  $(\text{Des } J, k) \in \Omega_{n,1}$ , so

$$M_J + M_K = M_{n, \text{Des } J} + M_{n, (\text{Des } J) \cup \{k\}} = \mathbf{m}_{\text{Des } J, k}.$$

- (2) Suppose  $J \triangleright_2 K$ , so that we may write  $J = (j_1, j_2, \dots, j_{m-1}, 2)$  where  $j_l = 2$  for some  $l \in [m-1]$ , and  $K = (j_1, \dots, j_{l-1}, 1, 1, j_{l+1}, \dots, j_{m-1}, 2)$ . Taking  $k = j_1 + \dots + j_{l-1} + 1$ , we have  $\text{Des } K = (\text{Des } J) \cup \{k\}$  and  $(\text{Des } J, k) \in \Omega_{n,2}$ , so

$$M_J + M_K = M_{n, \text{Des } J} + M_{n, (\text{Des } J) \cup \{k\}} = \mathbf{m}_{\text{Des } J, k}.$$

Moreover, if  $J = (1^{n-2}, 2)$ , then we have  $\text{Des } J = [n - 2]$  and  $(\text{Des } J, n - 1) \in \Omega_{n,3}$ , which means  $M_J = \mathbf{m}_{\text{Des } J, n-1}$ . We thus have  $\text{span } \mathcal{M}_n^{\text{Pk}} \subseteq \text{span}(\mathbf{m}_{C,k})_{(C,k) \in \Omega_n}$  by linearity, and the proof is complete.  $\square$

**Proposition 4.3.** For all  $n \geq 0$ , we have  $\text{span}(\mathbf{f}_{C,k})_{(C,k) \in \Omega_n} = \text{span}(\mathbf{m}_{C,k})_{(C,k) \in \Omega_n}$ .

*Proof.* Let us define a partial order on each  $\Omega_{n,i}$  by setting  $(B, k) \geq (C, l)$  if  $k = l$  and  $C \subseteq B$ . Then, we endow  $\Omega_n$  with the partial order obtained by taking the disjoint union of the posets  $\Omega_{n,1}$ ,  $\Omega_{n,2}$ , and  $\Omega_{n,3}$ , so that  $(B, k)$  and  $(C, l)$  are incomparable if they are not in the same  $\Omega_{n,i}$ . By Lemma 2.5, it suffices to show that  $(\mathbf{m}_{C,k})_{(C,k) \in \Omega_n}$  expands invertibly triangularly in  $(\mathbf{f}_{C,k})_{(C,k) \in \Omega_n}$  with respect to this partial order.

First, fix  $(C, k) \in \Omega_{n,1}$ . It is straightforward to show that

$$\{(B, k) : C \subseteq B \subseteq [n-1], k-1 \notin B, k \notin B\} = \{(B, l) \in \Omega_{n,1} : (B, l) \geq (C, k)\};$$

by this set equality and Lemma 2.2 (c), we have

$$\begin{aligned} \mathbf{m}_{C,k} &= M_{n,C} + M_{n,C \cup \{k\}} \\ &= \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k-1, k \notin B}} (-1)^{|B \setminus C|} (F_{n,B} - F_{n,B \cup \{k-1\}}) \\ &= \sum_{\substack{(B,l) \in \Omega_{n,1} \\ (B,l) \geq (C,k)}} (-1)^{|B \setminus C|} \mathbf{f}_{B,l} \\ &= \mathbf{f}_{C,k} + \sum_{\substack{(B,l) \in \Omega_n \\ (B,l) > (C,k)}} (-1)^{|B \setminus C|} \mathbf{f}_{B,l}. \end{aligned}$$

Second, fix  $(C, k) \in \Omega_{n,2}$ . By Lemma 2.2 (b), we have

$$\begin{aligned} \mathbf{m}_{C,k} &= M_{n,C} + M_{n,C \cup \{k\}} \\ &= \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k \notin B}} (-1)^{|B \setminus C|} F_{n,B} \\ &= \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k, n-1 \notin B}} (-1)^{|B \setminus C|} F_{n,B} + \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k \notin B, n-1 \in B}} (-1)^{|B \setminus C|} F_{n,B}. \end{aligned}$$

Recall that that  $(C, k) \in \Omega_{n,2}$  implies  $k \neq n-1$ . By toggling the inclusion of  $n-1$  in  $B$ , we see that the sets  $B$  satisfying  $C \subseteq B \subseteq [n-1]$  and  $k, n-1 \notin B$  are in bijection with those satisfying  $C \subseteq B \subseteq [n-1]$ ,  $k \notin B$ , and  $n-1 \in B$ . Thus, we can write

$$\begin{aligned} \mathbf{m}_{C,k} &= \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k, n-1 \notin B}} (-1)^{|B \setminus C|} F_{n,B} + \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k, n-1 \notin B}} (-1)^{|(B \cup \{n-1\}) \setminus C|} F_{n, B \cup \{n-1\}} \\ &= \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k, n-1 \notin B}} (-1)^{|B \setminus C|} (F_{n,B} - F_{n, B \cup \{n-1\}}). \end{aligned}$$

A routine argument yields the set equality

$$\{(B, k) : C \subseteq B \subseteq [n-1], k \notin B, n-1 \notin B\} = \{(B, l) \in \Omega_{n,2} : (B, l) \geq (C, k)\},$$

whence

$$\mathbf{m}_{C,k} = \sum_{\substack{(B,l) \in \Omega_{n,2} \\ (B,l) \succeq (C,k)}} (-1)^{|B \setminus C|} \mathbf{f}_{B,l} = \mathbf{f}_{C,k} + \sum_{\substack{(B,l) \in \Omega_n \\ (B,l) > (C,k)}} (-1)^{|B \setminus C|} \mathbf{f}_{B,l}.$$

Lastly, fix  $(C, k) \in \Omega_{n,3}$ , so that  $C = [n-2]$  and  $k = n-1$ . Then from Lemma 2.2 (a), we obtain

$$\mathbf{m}_{C,k} = M_{n,[n-2]} = F_{n,[n-2]} - F_{n,[n-1]} = F_{n,C} - F_{n,C \cup \{n-1\}} = \mathbf{f}_{C,k}.$$

We have now shown that  $(\mathbf{m}_{C,k})_{(C,k) \in \Omega_n}$  expands invertibly triangularly in  $(\mathbf{f}_{C,k})_{(C,k) \in \Omega_n}$ , so the result follows from Lemma 2.5.  $\square$

We are now ready to give the proof of Theorem 3 (a).

*Proof of Theorem 3 (a).* We have

$$\begin{aligned} \mathcal{K}_n^{\text{Pk}} &= \text{span } \mathcal{F}_n^{\text{Pk}} && \text{(by Theorem 2 (a))} \\ &= \text{span}(\mathbf{f}_{C,k})_{(C,k) \in \Omega_n} && \text{(by Proposition 4.1)} \\ &= \text{span}(\mathbf{m}_{C,k})_{(C,k) \in \Omega_n} && \text{(by Proposition 4.3)} \\ &= \text{span } \mathcal{M}_n^{\text{Pk}} && \text{(by Proposition 4.2)} \end{aligned}$$

for all  $n \geq 0$ . Therefore,

$$\begin{aligned} \mathcal{K}^{\text{Pk}} &= \bigoplus_{n=0}^{\infty} \mathcal{K}_n^{\text{Pk}} = \bigoplus_{n=0}^{\infty} \text{span } \mathcal{M}_n^{\text{Pk}} \\ &= \text{span} \left( \{ M_J + M_K : J \triangleright_1 K \text{ or } J \triangleright_2 K \} \cup \{ M_J : J \in \tilde{\mathcal{C}} \} \right). \end{aligned} \quad \square$$

#### 4.2. The peak number—proof of Theorem 3 (b)

Let us now proceed to  $\mathcal{K}^{\text{Pk}}$ . Define

$$\begin{aligned} \mathcal{F}_n^{\text{Pk}} &:= \{ F_J - F_K : J \rightarrow_{\text{Pk}} K \text{ and } J, K \vDash n \} \quad \text{and} \\ \mathcal{M}_n^{\text{Pk}} &:= \{ M_J + M_K : J \triangleright_1 K \text{ or } J \triangleright_2 K, \text{ and } J, K \vDash n \} \cup \{ M_{(1^{n-2}, 2)} \} \\ &\quad \cup \{ M_J - M_K : J \rightarrow_3 K \text{ and } J, K \vDash n \}, \end{aligned}$$

so that we wish to prove  $\mathcal{K}_n^{\text{Pk}} = \text{span } \mathcal{M}_n^{\text{Pk}}$  for all  $n$ .

Recall the definitions of  $\Omega_{n,1}$ ,  $\Omega_{n,2}$ , and  $\Omega_{n,3}$  from Section 4.1, and define  $\Omega_{n,4} \subseteq 2^{[n-1]} \times [n-1]$  by

$$\begin{aligned} \Omega_{n,4} &:= \{ (C, k) : n-1 \in C, k-1 \in C \cup \{0\}, k \in C, k+1 \notin C, k+2 \in C, \\ &\quad \text{and } j \in (C \cup \{0\}) \setminus \{n-1\} \implies j+1 \in C \text{ or } j+2 \in C \}. \end{aligned}$$

Note that  $\Omega_{n,4}$  is empty for  $n \leq 3$ , and that  $\Omega_{n,4}$  is disjoint from  $\Omega_{n,1}$ ,  $\Omega_{n,2}$ , and  $\Omega_{n,3}$ . Write

$$\Theta_n := \Omega_{n,1} \sqcup \Omega_{n,2} \sqcup \Omega_{n,3} \sqcup \Omega_{n,4} = \Omega_n \sqcup \Omega_{n,4}.$$

Next, we shall expand the definitions of  $\mathbf{f}_{C,k}$  and  $\mathbf{m}_{C,k}$  from Section 4.1 to all  $(C, k) \in \Theta_n$ .  
Let

$$\mathbf{f}_{C,k} := \begin{cases} F_{n,C} - F_{n,C \cup \{k-1\}}, & \text{if } (C, k) \in \Omega_{n,1}, \\ F_{n,C} - F_{n,C \cup \{n-1\}}, & \text{if } (C, k) \in \Omega_{n,2} \text{ or } (C, k) \in \Omega_{n,3}, \\ F_{n,C} - F_{n,(C \cup \{k+1\}) \setminus \{k\}}, & \text{if } (C, k) \in \Omega_{n,4}, \end{cases}$$

and

$$\mathbf{m}_{C,k} := \begin{cases} M_{n,C} + M_{n,C \cup \{k\}}, & \text{if } (C, k) \in \Omega_{n,1} \text{ or } (C, k) \in \Omega_{n,2}, \\ M_{n,C}, & \text{if } (C, k) \in \Omega_{n,3}, \\ M_{n,C} - M_{n,(C \cup \{k+1\}) \setminus \{k\}}, & \text{if } (C, k) \in \Omega_{n,4}. \end{cases}$$

The proof of the next lemma is routine and so it is omitted.

**Lemma 4.4.** *For all  $n \geq 0$  and  $J, K \vDash n$ , we have  $J \rightarrow_3 K$  if and only if  $\text{Des } K = (\text{Des } J \cup \{k+1\}) \setminus \{k\}$  and  $(\text{Des } J, k) \in \Omega_{n,4}$  for some  $k \in [n-1]$ .*

**Proposition 4.5.** *For all  $n \geq 0$ , we have  $\text{span}(\mathbf{f}_{C,k})_{(C,k) \in \Theta_n} = \text{span } \mathcal{F}_n^{\text{pk}}$ .*

*Proof.* In light of Proposition 4.1, it suffices to show that

$$\text{span}(\mathbf{f}_{C,k})_{(C,k) \in \Omega_{n,4}} = \text{span}\{F_J - F_K : J \rightarrow_3 K \text{ and } J, K \vDash n\}. \quad (4.1)$$

Fix  $(C, k) \in \Omega_{n,4}$ . Let  $J = \text{Comp } C$  and  $K = \text{Comp}((C \cup \{k+1\}) \setminus \{k\})$ , so that  $\text{Des } K = (\text{Des } J \cup \{k+1\}) \setminus \{k\}$  and  $(\text{Des } J, k) \in \Omega_{n,4}$ . By Lemma 4.4, we have

$$\mathbf{f}_{C,k} = F_{n,C} - F_{n,(C \cup \{k+1\}) \setminus \{k\}} = F_J - F_K$$

where  $J \rightarrow_3 K$ , so the forward inclusion of (4.1) follows from linearity. The reverse inclusion is similar.  $\square$

**Proposition 4.6.** *For all  $n \geq 0$ , we have  $\text{span}(\mathbf{m}_{C,k})_{(C,k) \in \Theta_n} = \text{span } \mathcal{M}_n^{\text{pk}}$ .*

*Proof.* Given Proposition 4.2, it suffices to show that

$$\text{span}(\mathbf{m}_{C,k})_{(C,k) \in \Omega_{n,4}} = \text{span}\{M_J - M_K : J \rightarrow_3 K \text{ and } J, K \vDash n\},$$

but this is proved in the same way as (4.1) and so we omit the details.  $\square$

**Proposition 4.7.** *For all  $n \geq 0$ , we have  $\text{span}(\mathbf{f}_{C,k})_{(C,k) \in \Theta_n} = \text{span}(\mathbf{m}_{C,k})_{(C,k) \in \Theta_n}$ .*

*Proof.* Take the partial order defined on each  $\Omega_{n,i}$  from the proof of Proposition 4.3 and extend it to  $\Omega_{n,4}$ . Then, endow  $\Theta_n$  with the partial order obtained by taking the disjoint union of the posets  $\Omega_n$  and  $\Omega_{n,4}$ , so that the elements of  $\Omega_n$  are incomparable with those in  $\Omega_{n,4}$ . The proof of Proposition 4.3 established that  $\mathbf{m}_{C,k}$  for each  $(C, k) \in \Omega_n$  has an invertibly triangular expansion in  $(\mathbf{f}_{C,k})_{(C,k) \in \Omega_n}$  and thus  $(\mathbf{f}_{C,k})_{(C,k) \in \Theta_n}$ ; we shall show that  $\mathbf{m}_{C,k}$  for each  $(C, k) \in \Omega_{n,4}$  has such an expansion as well.

Fix  $(C, k) \in \Omega_{n,4}$ . Then

$$\mathbf{m}_{C,k} = M_{n,C} - M_{n,(C \cup \{k+1\}) \setminus \{k\}}; \quad (4.2)$$

we shall expand  $M_{n,C}$  and  $M_{n,(C \cup \{k+1\}) \setminus \{k\}}$  separately. From Lemma 2.2 (a), we have

$$\begin{aligned} M_{n,C} &= \sum_{C \subseteq B \subseteq [n-1]} (-1)^{|B \setminus C|} F_{n,B} \\ &= \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k+1 \notin B}} (-1)^{|B \setminus C|} F_{n,B} + \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k+1 \in B}} (-1)^{|B \setminus C|} F_{n,B} \end{aligned} \quad (4.3)$$

and

$$\begin{aligned} M_{n,(C \cup \{k+1\}) \setminus \{k\}} &= \sum_{(C \cup \{k+1\}) \setminus \{k\} \subseteq B \subseteq [n-1]} (-1)^{|B \setminus ((C \cup \{k+1\}) \setminus \{k\})|} F_{n,B} \\ &= \sum_{\substack{C \setminus \{k\} \subseteq B \subseteq [n-1] \\ k+1 \in B}} (-1)^{|B \setminus ((C \cup \{k+1\}) \setminus \{k\})|} F_{n,B}. \end{aligned}$$

Note that  $(C, k) \in \Omega_{n,4}$  implies  $k+1 \notin C$ , so if  $k+1 \in B$  then  $|B \setminus ((C \cup \{k+1\}) \setminus \{k\})| = |B \setminus (C \setminus \{k\})| - 1$ . Hence,

$$\begin{aligned} M_{n,(C \cup \{k+1\}) \setminus \{k\}} &= \sum_{\substack{C \setminus \{k\} \subseteq B \subseteq [n-1] \\ k+1 \in B}} (-1)^{|B \setminus (C \setminus \{k\})| - 1} F_{n,B} \\ &= - \sum_{\substack{C \setminus \{k\} \subseteq B \subseteq [n-1] \\ k+1 \in B}} (-1)^{|B \setminus (C \setminus \{k\})|} F_{n,B} \\ &= - \sum_{\substack{C \setminus \{k\} \subseteq B \subseteq [n-1] \\ k, k+1 \in B}} (-1)^{|B \setminus (C \setminus \{k\})|} F_{n,B} - \sum_{\substack{C \setminus \{k\} \subseteq B \subseteq [n-1] \\ k \notin B, k+1 \in B}} (-1)^{|B \setminus (C \setminus \{k\})|} F_{n,B}. \end{aligned}$$

Furthermore,  $(C, k) \in \Omega_{n,4}$  implies  $k \in C$ , so the two conditions  $C \setminus \{k\} \subseteq B$  and  $k \in B$  are equivalent to the single condition  $C \subseteq B$ . Also, if  $k \in B$  then  $|B \setminus (C \setminus \{k\})| = 1 + |B \setminus C|$ , whereas if  $k \notin B$  then  $|B \setminus (C \setminus \{k\})| = |B \setminus C|$ . Therefore,

$$M_{n,(C \cup \{k+1\}) \setminus \{k\}} = \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k+1 \in B}} (-1)^{|B \setminus C|} F_{n,B} - \sum_{\substack{C \setminus \{k\} \subseteq B \subseteq [n-1] \\ k \notin B, k+1 \in B}} (-1)^{|B \setminus C|} F_{n,B}. \quad (4.4)$$

Substituting (4.3) and (4.4) into (4.2) yields

$$\begin{aligned} \mathbf{m}_{C,k} &= \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k+1 \notin B}} (-1)^{|B \setminus C|} F_{n,B} + \sum_{\substack{C \setminus \{k\} \subseteq B \subseteq [n-1] \\ k \notin B, k+1 \in B}} (-1)^{|B \setminus C|} F_{n,B} \\ &= \sum_{\substack{C \setminus \{k\} \subseteq B \subseteq [n-1] \\ k \in B, k+1 \notin B}} (-1)^{|B \setminus C|} F_{n,B} + \sum_{\substack{C \setminus \{k\} \subseteq B \subseteq [n-1] \\ k \notin B, k+1 \in B}} (-1)^{|B \setminus C|} F_{n,B}. \end{aligned}$$

It is straightforward to verify that the map  $B \mapsto (B \cup \{k+1\}) \setminus \{k\}$  is a bijection from sets  $B$  satisfying  $C \setminus \{k\} \subseteq B \subseteq [n-1]$ ,  $k \in B$ , and  $k+1 \notin B$  to those satisfying



$C \setminus \{k\} \subseteq B \subseteq [n-1]$ ,  $k \notin B$ , and  $k+1 \in B$ . Moreover, for the former family of sets, we have  $|((B \cup \{k+1\}) \setminus \{k\}) \setminus C| = 1 + |(B \setminus \{k\}) \setminus C| = 1 + |B \setminus C|$ . Therefore, we have

$$\begin{aligned} \mathbf{m}_{C,k} &= \sum_{\substack{C \setminus \{k\} \subseteq B \subseteq [n-1] \\ k \in B, k+1 \notin B}} \left( (-1)^{|B \setminus C|} F_{n,B} + (-1)^{|((B \cup \{k+1\}) \setminus \{k\}) \setminus C|} F_{n,(B \cup \{k+1\}) \setminus \{k\}} \right) \\ &= \sum_{\substack{C \setminus \{k\} \subseteq B \subseteq [n-1] \\ k \in B, k+1 \notin B}} (-1)^{|B \setminus C|} (F_{n,B} - F_{n,(B \cup \{k+1\}) \setminus \{k\}}). \\ &= \sum_{\substack{C \subseteq B \subseteq [n-1] \\ k+1 \notin B}} (-1)^{|B \setminus C|} (F_{n,B} - F_{n,(B \cup \{k+1\}) \setminus \{k\}}). \end{aligned}$$

Finally, it is readily checked that

$$\{(B, k) : C \subseteq B \subseteq [n-1], k+1 \notin B\} = \{(B, l) \in \Omega_{n,4} : (B, l) \geq (C, k)\},$$

whence

$$\begin{aligned} \mathbf{m}_{C,k} &= \sum_{\substack{(B,l) \in \Omega_{n,4} \\ (B,l) \geq (C,k)}} (-1)^{|B \setminus C|} (F_{n,B} - F_{n,(B \cup \{l+1\}) \setminus \{l\}}) \\ &= \sum_{\substack{(B,l) \in \Omega_{n,4} \\ (B,l) \geq (C,k)}} (-1)^{|B \setminus C|} \mathbf{f}_{B,l} \\ &= \mathbf{f}_{C,k} + \sum_{\substack{(B,l) \in \Theta_n \\ (B,l) > (C,k)}} (-1)^{|B \setminus C|} \mathbf{f}_{B,l}. \end{aligned}$$

We have shown that  $(\mathbf{m}_{C,k})_{(C,k) \in \Theta_n}$  expands invertibly triangularly in  $(\mathbf{f}_{C,k})_{(C,k) \in \Theta_n}$ , so the desired conclusion is reached via Lemma 2.5.  $\square$

Finally, we complete the proof of Theorem 3 (b) using the preceding propositions.

*Proof of Theorem 3 (b).* We have

$$\begin{aligned} \mathcal{K}_n^{\text{pk}} &= \text{span } \mathcal{F}_n^{\text{pk}} && \text{(by Theorem 2 (b))} \\ &= \text{span}(\mathbf{f}_{C,k})_{(C,k) \in \Theta_n} && \text{(by Proposition 4.5)} \\ &= \text{span}(\mathbf{m}_{C,k})_{(C,k) \in \Theta_n} && \text{(by Proposition 4.7)} \\ &= \text{span } \mathcal{M}_n^{\text{pk}} && \text{(by Proposition 4.6)} \end{aligned}$$

for all  $n \geq 0$ , and therefore

$$\begin{aligned} \mathcal{K}^{\text{pk}} &= \bigoplus_{n=0}^{\infty} \mathcal{K}_n^{\text{pk}} = \bigoplus_{n=0}^{\infty} \text{span } \mathcal{M}_n^{\text{pk}} \\ &= \text{span} \left( \{ M_J + M_K : J \triangleright_1 K \text{ or } J \triangleright_2 K \} \cup \{ M_J : J \in \tilde{\mathcal{C}} \} \cup \{ M_J - M_K : J \rightarrow_3 K \} \right) \end{aligned}$$

as desired.  $\square$

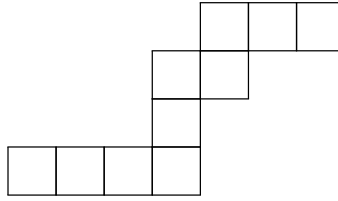


Figure 5.1: The ribbon shape of the composition  $(4, 1, 2, 3)$ .

## 5. Valleys

Given  $\pi \in \mathfrak{P}_n$ , we say that  $i \in \{2, 3, \dots, n-1\}$  is a *valley* of  $\pi$  if  $\pi_{i-1} > \pi_i < \pi_{i+1}$ . Then  $\text{Val } \pi$  is defined to be the set of valleys of  $\pi$  and  $\text{val } \pi$  the number of valleys of  $\pi$ . There is a clear symmetry relating peaks and valleys in permutations, and this section will describe the implications that this symmetry has on the statistics  $\text{Pk}$ ,  $\text{pk}$ ,  $\text{Val}$ , and  $\text{val}$ , their shuffle algebras, and their kernels. In particular, we will obtain from Theorem 2 an analogue of this theorem for  $\mathcal{K}^{\text{Val}}$  and  $\mathcal{K}^{\text{val}}$ .

The *complement*  $\pi^c$  of  $\pi \in \mathfrak{P}_n$  is the permutation obtained by (simultaneously) replacing the  $i$ th smallest letter in  $\pi$  with the  $i$ th largest letter in  $\pi$  for all  $1 \leq i \leq n$ . For example, if  $\pi = 472691$  then  $\pi^c = 627419$ . Observe that

$$\text{Pk } \pi = \text{Val } \pi^c \quad \text{and} \quad \text{pk } \pi = \text{val } \pi^c;$$

this implies that  $\text{Pk}$  and  $\text{Val}$  are *c-equivalent statistics* (see [6, Section 3.2] for the definition), and so are  $\text{pk}$  and  $\text{val}$ . Complementation is a “shuffle-compatibility-preserving” involution on permutations, and according to [6, Theorem 3.5], if two permutation statistics  $\text{st}_1$  and  $\text{st}_2$  are *f-equivalent* where  $f$  is a shuffle-compatibility-preserving involution and  $\text{st}_1$  is shuffle-compatible, then  $\text{st}_2$  is also shuffle-compatible and the map  $[\pi]_{\text{st}_1} \mapsto [\pi^f]_{\text{st}_2}$  extends to an isomorphism between their shuffle algebras. Consequently, we have  $\mathcal{A}^{\text{Pk}} \cong \mathcal{A}^{\text{Val}}$  and  $\mathcal{A}^{\text{pk}} \cong \mathcal{A}^{\text{val}}$ .

We can also define complements of compositions in the following way. Given  $L \vDash n$ , let  $L^c := \text{Comp}([n-1] \setminus \text{Des } L)$ . For example, if  $L = (4, 1, 2, 3)$  then we have  $\text{Des } L = \{4, 5, 7\}$  and thus  $L^c = \text{Comp}\{1, 2, 3, 6, 8, 9\} = (1, 1, 1, 3, 2, 1, 1)$ . A simple way to obtain  $L^c$  from  $L$  is to draw its *ribbon shape* (see Figure 5.1 for an example); reading off the columns from left to right (as opposed to the rows from bottom to top) yields  $L^c$ . It is evident that  $\text{Comp } \pi = L$  implies  $\text{Comp } \pi^c = L^c$ , so we have

$$\text{Pk } L = \text{Val } L^c \quad \text{and} \quad \text{pk } L = \text{val } L^c.$$

Hence, two compositions are  $\text{Pk}$ -equivalent (respectively,  $\text{pk}$ -equivalent) if and only if their complements are  $\text{Val}$ -equivalent (respectively,  $\text{val}$ -equivalent).

Consider the involutory automorphism  $\psi$  on  $\text{QSym}$  defined by  $\psi(F_L) = F_{L^c}$  [12, Section 3.6]. Then the following diagrams commute:

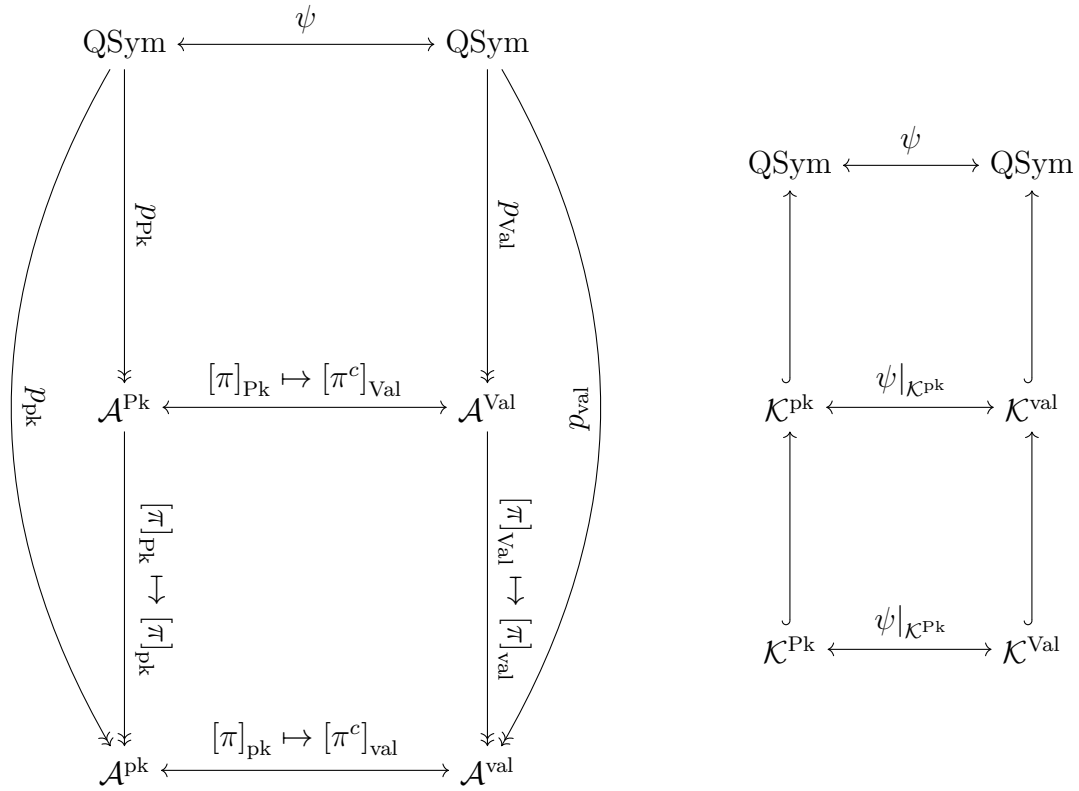


Figure 5.2: Relationships between the shuffle algebras and kernels of Pk, pk, Val, and val

The involution  $\psi$  allows us to obtain characterizations for the ideals  $\mathcal{K}^{\text{Val}}$  and  $\mathcal{K}^{\text{val}}$  in terms of the fundamental basis from the analogous results for  $\mathcal{K}^{\text{Pk}}$  and  $\mathcal{K}^{\text{pk}}$  (Theorem 2). Given compositions  $J = (j_1, j_2, \dots, j_m)$  and  $K$ , let us write:

- $J \rightarrow_1 K$  if there exists  $l \in [m - 1]$  for which  $j_l \geq 2$  and  $j_{l+1} = 1$ , and

$$K = (j_1, \dots, j_{l-1}, j_l + 1, j_{l+2}, \dots, j_m);$$

- $J \rightarrow_2 K$  if  $j_1 = j_2 = 1$  and

$$K = (2, j_3, \dots, j_m);$$

- $J \rightarrow_3 K$  if  $j_i \geq 2$  for all  $i \in \{2, 3, \dots, m\}$ , there exists  $l \in [m - 1]$  for which  $j_l \geq 2$ —and  $j_l > 2$  if  $l > 1$ —and

$$K = (j_1, \dots, j_{l-1}, j_l - 1, j_{l+1} + 1, j_{l+2}, \dots, j_m).$$

**Theorem 5.1.** *The ideals  $\mathcal{K}^{\text{Val}}$  and  $\mathcal{K}^{\text{val}}$  are spanned (as  $\mathbb{Q}$ -vector spaces) in the following ways:*

- (a)  $\mathcal{K}^{\text{Val}} = \text{span}\{F_J - F_K : J \rightarrow_1 K \text{ or } J \rightarrow_2 K\}$

$$(b) \mathcal{K}^{\text{val}} = \text{span}\{F_J - F_K : J \twoheadrightarrow_1 K, J \twoheadrightarrow_2 K, \text{ or } J \twoheadrightarrow_3 K\}$$

To prove Theorem 5.1, we will need a couple lemmas.

**Lemma 5.2.** *Let  $J$  and  $K$  be compositions.*

(a) *If  $J \rightarrow_{\text{pk}} K$ , then  $J^c \twoheadrightarrow_1 K^c$  or  $J^c \twoheadrightarrow_2 K^c$ .*

(b) *If  $J \rightarrow_{\text{pk}} K$ , then  $J^c \twoheadrightarrow_1 K^c$ ,  $J^c \twoheadrightarrow_2 K^c$ , or  $J^c \twoheadrightarrow_3 K^c$ .*

*Proof.* Let  $J = (j_1, j_2, \dots, j_m)$  and suppose that  $J \rightarrow_{\text{pk}} K$ , so either  $J \rightarrow_1 K$  or  $J \rightarrow_2 K$ .

(1) Suppose that  $J \rightarrow_1 K$ . Then there exists  $l \in [m]$  for which  $j_l > 2$  and

$$K = (j_1, \dots, j_{l-1}, 1, j_l - 1, j_{l+1}, \dots, j_m).$$

Upon drawing the ribbon diagrams of  $J$  and  $K$  and reading the columns from left to right, we see that  $J^c$  and  $K^c$  are the same except that a segment  $(\alpha, 1^{j_l-2})$  of  $J^c$  is replaced with  $(\alpha + 1, 1^{j_l-3})$  in  $K^c$ :

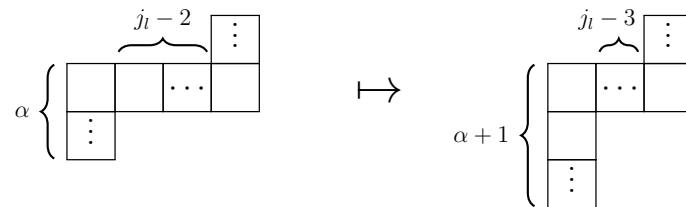


Figure 5.3: Ribbon diagram upon replacing a part  $j_l > 2$  with  $(1, j_l - 1)$ .

If  $l = 1$ , then  $\alpha = 1$  and we have  $J^c \twoheadrightarrow_2 K^c$ . Otherwise, if  $l > 1$ , then  $\alpha \geq 2$  and we have  $J^c \twoheadrightarrow_1 K^c$ .

(2) Suppose that  $J \rightarrow_2 K$ , so that  $j_m = 2$  and

$$K = (j_1, \dots, j_{m-1}, 1, 1).$$

Then  $J^c$  and  $K^c$  are the same except that  $J^c$  ends with  $(\alpha, 1)$  and  $K^c$  ends with  $\alpha + 1$ :

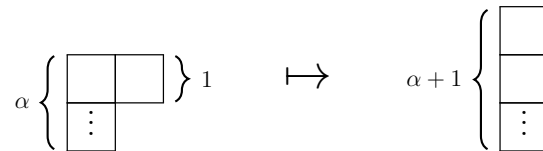


Figure 5.4: Ribbon diagram upon replacing the final part  $j_m = 2$  with  $(1, 1)$ .

If  $\alpha \geq 2$ , then  $J^c \twoheadrightarrow_1 K^c$ . Otherwise,  $\alpha = 1$  forces  $m = 1$  and thus we have  $J^c \twoheadrightarrow_2 K^c$ .

Part (a) follows from the two cases above.

In light of (a), to prove (b) it suffices to show that  $J \rightarrow_3 K$  implies  $J^c \rightarrow_3 K^c$ . To that end, let us suppose that  $J \rightarrow_3 K$ , so that  $j_i \leq 2$  for all  $i \in [m]$ ,  $j_m = 1$ ,  $j_l = 1$  and  $j_{l+1} = 2$  for some  $l \in [m-2]$ , and

$$K = (j_1, \dots, j_{l-1}, j_{l+1}, j_l, j_{l+2}, \dots, j_m).$$

Then  $J^c$  and  $K^c$  are the same except that a segment  $(\alpha, \beta)$  of  $J^c$  is replaced with  $(\alpha-1, \beta+1)$  in  $K^c$ , where  $\alpha \geq 2$  if  $\alpha$  is the first part of  $J^c$  and  $\alpha \geq 3$  otherwise:

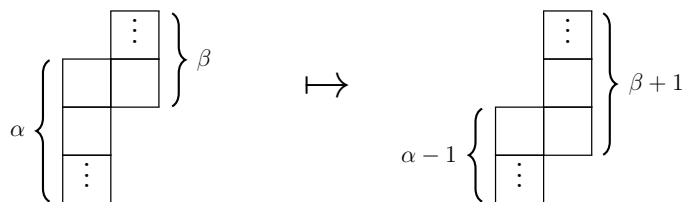


Figure 5.5: Ribbon diagram upon swapping the positions of parts (1, 2).

Furthermore, requiring  $J$  to have every part at most 2 and for  $J$  to end with a 1 implies that every part of  $J^c$ , except possibly the first one, is at least 2. Thus  $J^c \rightarrow_3 K^c$ , and the proof is complete.  $\square$

Just as the peaks of a permutation occur precisely at the end of its non-final increasing runs of length at least 2, the valleys occur precisely at the beginning of its non-initial increasing runs of length at least 2. Then the following lemma readily follows; compare with Lemma 2.1 for peaks.

**Lemma 5.3.** *Let  $L = (j_1, j_2, \dots, j_m)$  be a composition. Then:*

- (a)  $\text{Val } L = \left\{ 1 + \sum_{i=1}^{k-1} j_i : j_k \geq 2 \text{ and } k \in \{2, 3, \dots, m\} \right\}$
- (b)  $\text{val } L = |\{k \in \{2, 3, \dots, m\} : j_k \geq 2\}|$

Now we are ready to prove Theorem 5.1.

*Proof of Theorem 5.1.* In light of Theorem 2 and the involution  $\psi$ , Lemma 5.2 implies the forward inclusions

$$\begin{aligned} \mathcal{K}^{\text{Val}} &\subseteq \text{span}\{F_J - F_K : J \rightarrow_1 K \text{ or } J \rightarrow_2 K\} \quad \text{and} \\ \mathcal{K}^{\text{val}} &\subseteq \text{span}\{F_J - F_K : J \rightarrow_1 K, J \rightarrow_2 K, \text{ or } J \rightarrow_3 K\}. \end{aligned}$$

For the reverse inclusions, it suffices to show that  $J \rightarrow_1 K$  and  $J \rightarrow_2 K$  each imply  $J \sim_{\text{Val}} K$ , and that  $J \rightarrow_1 K$ ,  $J \rightarrow_2 K$ , and  $J \rightarrow_3 K$  each imply  $J \sim_{\text{val}} K$ . However, it is readily seen from Lemma 5.3 and the definitions of  $\rightarrow_1$ ,  $\rightarrow_2$ , and  $\rightarrow_3$  that these indeed hold.  $\square$

Ehrenborg's formula

$$\psi(M_L) = (-1)^{n-\ell(L)} \sum_{L \leq K} M_K$$

[4, Section 5], where  $\ell(L)$  is the number of parts of  $L$ , gives us the image of the involution  $\psi$  on the monomial basis. One can use this formula along with Theorem 3 to obtain spanning sets for  $\mathcal{K}^{\text{Val}}$  and  $\mathcal{K}^{\text{val}}$  in terms of monomial quasisymmetric functions, but more work is needed to show that Val and val are  $M$ -binomial. We will not do this here.

Finally, it is worth mentioning that the kernel  $\mathcal{K}^{\text{epk}}$  of the exterior peak number is identical to  $\mathcal{K}^{\text{val}}$ . This is because  $\text{epk } \pi = \text{val } \pi + 1$  for all (nonempty) permutations  $\pi$  [6, Lemma 2.1 (e)], and therefore  $J \sim_{\text{epk}} K$  if and only if  $J \sim_{\text{val}} K$ .

## 6. Future directions of research

We conclude this paper by surveying some directions for future research.

Given  $\pi \in \mathfrak{P}_n$ , we say that:

- $i \in [n - 1]$  is a *left peak* of  $\pi$  if  $i$  is a peak of  $\pi$ , or if  $i = 1$  and  $\pi_1 > \pi_2$ . Then  $\text{Lpk } \pi$  is defined to be the set of left peaks of  $\pi$  and  $\text{lpk } \pi$  the number of left peaks of  $\pi$ .
- $i \in \{2, 3, \dots, n\}$  is a *right peak* of  $\pi$  if  $i$  is a peak of  $\pi$ , or if  $i = n$  and  $\pi_{n-1} < \pi_n$ . Then  $\text{Rpk } \pi$  is defined to be the set of right peaks of  $\pi$  and  $\text{rpk } \pi$  the number of right peaks of  $\pi$ .

All four of these statistics are known to be shuffle-compatible [6], and they were conjectured by Grinberg (along with Pk, pk, Val, and val) to be  $M$ -binomial. While we have resolved Grinberg's conjecture for the Pk and pk statistics, it remains open for these six other statistics, so we repeat this conjecture for these remaining statistics here.

**Conjecture 1** (Grinberg [8, Question 107]). The statistics Lpk, lpk, Rpk, rpk, Val, and val are  $M$ -binomial.

Just as peaks and valleys are related by complementation, left peaks and right peaks are related by reversal. Given  $\pi = \pi_1 \pi_2 \cdots \pi_n \in \mathfrak{P}_n$ , define its *reverse*  $\pi^r$  by  $\pi^r := \pi_n \cdots \pi_2 \pi_1$ . Then

$$\text{Lpk } \pi = (n + 1) - \text{Rpk } \pi^r := \{n + 1 - i : i \in \text{Rpk } \pi^r\} \quad \text{and} \quad \text{lpk } \pi = \text{rpk } \pi^r,$$

so these pairs of statistics are *r-equivalent* (see [6, Section 3.2] for the definition). Reversal is also shuffle-compatibility-preserving, so we have the  $\mathbb{Q}$ -algebra isomorphisms  $\mathcal{A}^{\text{Lpk}} \cong \mathcal{A}^{\text{Rpk}}$  and  $\mathcal{A}^{\text{lpk}} \cong \mathcal{A}^{\text{rpk}}$ . On the level of compositions, let us define  $L^r \vDash n$  for  $L \vDash n$  by

$$L^r := \text{Comp}([n - 1] \setminus (n - \text{Des } L)), \quad \text{where} \quad n - \text{Des } L := \{n - i : i \in \text{Des } L\}.$$

Also, consider the involution  $\rho$  on  $\text{QSym}$  given by  $\rho(F_L) = F_{L^r}$ , which like  $\psi$  is a  $\mathbb{Q}$ -algebra automorphism [12, Section 3.6]. It is easily verified that  $\text{Comp } \pi = L$  implies

Comp  $\pi^r = L^r$ , and that  $\rho$  restricts to an isomorphism between  $\mathcal{K}^{\text{Lpk}}$  and  $\mathcal{K}^{\text{Rpk}}$  as well as between  $\mathcal{K}^{\text{lpk}}$  and  $\mathcal{K}^{\text{rpk}}$ . Upon obtaining an analogue of Theorem 2 for the Lpk and lpk statistics, we could then establish analogous results for Rpk and rpk using the involution  $\rho$  similar to the proof of Theorem 5.1.

Concerning these symmetries, we also pose the following question.

**Question 2.** If two statistics  $\text{st}_1$  and  $\text{st}_2$  are  $r$ -equivalent or  $c$ -equivalent, does  $\text{st}_1$  being  $M$ -binomial imply that  $\text{st}_2$  is also  $M$ -binomial?

On the other hand, the statistics  $\text{maj}$ ,  $(\text{des}, \text{maj})$ , and  $(\text{val}, \text{des})$  are not  $M$ -binomial [8, Question 107]. Are there ways that we can tell whether a descent statistic is  $M$ -binomial solely in terms of the combinatorics of the statistic?

**Problem 3.** Find necessary and/or sufficient conditions for a descent statistic to have the  $M$ -binomial property (preferably combinatorial conditions).

Grinberg [9, Section 6] considers four binary operations  $\prec$ ,  $\succeq$ ,  $\Phi$ , and  $\aleph$  on  $\text{QSym}$  and defines a notion of ideal for each of these operations. We may then naturally ask whether the kernel of a descent statistic is one of these types of ideals, which turns out to have combinatorial significance. Notably,  $\mathcal{K}^{\text{st}}$  is both an  $\prec$ -ideal and an  $\succeq$ -ideal if and only if  $\text{st}$  is LR-shuffle-compatible, in which case  $\mathcal{A}^{\text{st}}$  is canonically a dendriform algebra quotient of  $\text{QSym}$ . We restate the following question of Grinberg.

**Question 4** (Grinberg [9, Question 6.16]). Which descent statistics  $\text{st}$  have the property that  $\mathcal{K}^{\text{st}}$  is an  $\prec$ -ideal,  $\succeq$ -ideal,  $\Phi$ -ideal, and/or  $\aleph$ -ideal?

Grinberg has some results in this direction [9, Section 6.7]; for example, the kernels  $\mathcal{K}^{\text{Des}}$ ,  $\mathcal{K}^{\text{des}}$ ,  $\mathcal{K}^{(\text{des}, \text{maj})}$ , and  $\mathcal{K}^{\text{Epk}}$  are ideals with respect to all four of these operations, whereas  $\mathcal{K}^{\text{Lpk}}$  is an ideal with respect to all except  $\aleph$  (it is a left  $\aleph$ -ideal but not a right  $\aleph$ -ideal). Hence, the statistics  $\text{Des}$ ,  $\text{des}$ ,  $(\text{des}, \text{maj})$ ,  $\text{Epk}$ , and  $\text{Lpk}$  are all LR-shuffle-compatible.

Finally, we note that the notions of shuffle-compatibility and shuffle algebras have recently been extended to cyclic permutations [3, 11]. A *cyclic permutation* is an equivalence class of a (linear) permutation under cyclic rotation; for example, if  $\pi = 21948$  then

$$[\pi] = \{21948, 82194, 48219, 94821, 19482\}.$$

The cyclic shuffle algebra of any cyclic shuffle-compatible descent statistic is isomorphic to a quotient of  $\text{cQSym}^-$  [11, Theorem 3.5], the “non-Escher subalgebra” of the  $\mathbb{Q}$ -algebra  $\text{cQSym}$  of *cyclic quasisymmetric functions* introduced recently by Adin, Gessel, Reiner, and Roichman [1]. The *fundamental cyclic quasisymmetric functions*  $F_{[L]}^{\text{cyc}}$ , indexed by “non-Escher cyclic compositions”, form a basis of  $\text{cQSym}^-$ , and one can then define the kernel of a cyclic descent statistic  $\text{cst}$  to be the subspace of  $\text{cQSym}^-$  spanned by all elements of the form  $F_{[J]}^{\text{cyc}} - F_{[K]}^{\text{cyc}}$  where  $[J]$  and  $[K]$  are  $\text{cst}$ -equivalent cyclic compositions. In analogy with the linear setting, the kernel of  $\text{cst}$  is an ideal of  $\text{cQSym}^-$  if and only if  $\text{cst}$  is cyclic shuffle-compatible.

**Problem 5.** Study kernels of cyclic descent statistics.

In particular, Theorems 2.8 and 3.7 of [11] allow one (under certain conditions) to construct the cyclic shuffle algebra of a cyclic statistic from the shuffle algebra of a related shuffle-compatible (linear) statistic; this suggests an analogous result relating kernels of cyclic descent statistics and kernels of linear descent statistics. There is also an analogue of the monomial basis in  $\text{cQSym}^-$ , so one can investigate the  $M$ -binomial property in this setting.

## Acknowledgements

The work in this paper was initiated as part of the first author's honors thesis project at Davidson College (under the supervision of the second author); we thank Heather Smith Blake, Frankie Chan, Jonad Pulaj, and Carl Yerger for serving on the first author's thesis committee and giving feedback on his work. We also thank Darij Grinberg for helpful e-mail correspondence. Lastly, we thank an anonymous referee who very generously provided simpler proofs for Theorems 3.3 and 3.4, as well as a number of minor corrections and suggestions which improved the presentation of this paper.

The second author was partially supported by an AMS-Simons Travel Grant and NSF grant DMS-2316181.

## References

- [1] Ron M. Adin, Ira M. Gessel, Victor Reiner, and Yuval Roichman. Cyclic quasi-symmetric functions. *Israel J. Math.*, 243(1):437–500, 2021.
- [2] Marcelo Aguiar, Nantel Bergeron, and Frank Sottile. Combinatorial Hopf algebras and generalized Dehn-Sommerville relations. *Compos. Math.*, 142(1):1–30, 2006.
- [3] Rachel Domagalski, Jinting Liang, Quinn Minnich, Bruce E. Sagan, Jamie Schmidt, and Alexander Sietsema. Cyclic shuffle compatibility. *Sém. Lothar. Combin.*, 85:Art. B85d, 11 pp., [2020–2021].
- [4] Richard Ehrenborg. On posets and Hopf algebras. *Adv. Math.*, 119(1):1–25, 1996.
- [5] Ira M. Gessel. Multipartite  $P$ -partitions and inner products of skew Schur functions. *Contemp. Math.*, 34:289–317, 1984.
- [6] Ira M. Gessel and Yan Zhuang. Shuffle-compatible permutation statistics. *Adv. Math.*, 332:85–141, 2018.
- [7] Darij Grinberg. Dual creation operators and a dendriform algebra structure on the quasisymmetric functions. *Canad. J. Math.*, 69(1):21–53, 2017.
- [8] Darij Grinberg. Shuffle-compatible permutation statistics II: the exterior peak set. *Electron. J. Combin.*, 25(4):#P4.17, 61 pp., 2018.
- [9] Darij Grinberg. Shuffle-compatible permutation statistics II: the exterior peak set (detailed version). Available at <https://www.cip.ifi.lmu.de/~grinberg/algebra/gzshuf2-long.pdf>, 2018.



- [10] Darij Grinberg and Victor Reiner. Hopf Algebras in Combinatorics. [arXiv:1409.8356v7](https://arxiv.org/abs/1409.8356v7), 2020.
- [11] Jinting Liang, Bruce E. Sagan, and Yan Zhuang. Cyclic shuffle-compatibility via cyclic shuffle algebras. To appear in *Ann. Comb.*, 2024.
- [12] Kurt Luoto, Stefan Mykytiuk, and Stephanie van Willigenburg. *An Introduction to Quasisymmetric Schur Functions. Hopf Algebras, Quasisymmetric Functions, and Young Composition Tableaux*. Springer, 2013.
- [13] Kathryn L. Nyman. The peak algebra of the symmetric group. *J. Algebraic Combin.*, 17(3):309–322, 2003.
- [14] Manfred Schocker. The peak algebra of the symmetric group revisited. *Adv. Math.*, 192(2):259–309, 2005.
- [15] Richard P. Stanley. *Ordered Structures and Partitions*, volume 199 of *Memoirs of the American Mathematical Society*. American Mathematical Society, Providence, 1972.
- [16] Richard P. Stanley. *Enumerative Combinatorics, Vol. 2*. Cambridge University Press, 2001.
- [17] John R. Stembridge. Enriched  $P$ -partitions. *Trans. Amer. Math. Soc.*, 349(2):763–788, 1997.
- [18] Yan Zhuang. Counting permutations by runs. *J. Comb. Theory Ser. A*, 142:147–176, 2016.