SHUFFLING AND $P$-PARTITIONS

JASON FULMAN AND T. KYLE PETERSEN

Abstract. In this survey article, we highlight the direct connection between card shuffling and the functions known as $P$-partitions that come from algebraic combinatorics. While many (but not all) of the results we discuss are known, we give a unified treatment. The key idea is this: the probability of obtaining a permutation $\pi$ from shelf shuffling is the probability that a random $P$-partition is sorted by $\pi$, and the probability of obtaining $\pi$ from riffle shuffling is the probability that a random $P$-partition is sorted by $\pi^{-1}$.

1. Introduction

Since Bayer and Diaconis’ influential paper on card shuffling in 1992 [1], it has been well-understood that permutations statistics, such as the number of descents, play an important role in understanding card shuffling. In particular, a key identity used in [1] to study the Gilbert-Shannon-Reeds riffle shuffle is the following: for any permutation $\pi \in S_n$,

\[
\binom{kl + n - \text{des}(\pi) - 1}{n} = \sum_{\sigma\tau = \pi} \binom{k + n - \text{des}(\sigma) - 1}{n} \binom{l + n - \text{des}(\tau) - 1}{n},
\]

where $\text{des}(\pi) = |\{i : \pi(i) > \pi(i+1)\}|$ is the number of descents of $\pi$.

While the identity (1) may be given various proofs, one of the nicest of these follows from work of Gessel in 1984, using the theory of $P$-partitions [8]. The “$P$” in $P$-partition stands for “partially ordered set” or “poset.” Stanley defined $P$-partitions as a way to generalize integer partitions to study plane partitions, but they can also be used to give a combinatorial framework for the study of symmetric and quasi-symmetric functions, with applications to permutation enumeration. See Gessel’s survey [9], [15, Section 4.5] and [16, Section 7.19].

Identities similar to Equation (1) show up in work of Petersen [13] from 2007 using a slightly more general notion of $P$-partition that includes but also builds on Stembridge’s notion of enriched $P$-partitions [17]. In these cases, it is the number of peaks of a permutation that matter (instances in which $\pi(i - 1) < \pi(i) > \pi(i + 1)$), rather than the number of descents.

The number of peaks proved crucial to the analysis of the “shelf shuffler” machine studied by Diaconis, Fulman, and Holmes in 2013 [4]. (In fact, they used a $P$-partition argument to prove one of their main results [4, Theorem 3.2].) In this paper, we will use $P$-partitions to study a family of related shuffling schemes. Although the paper [4] mentions a connection between riffle shuffling and $P$-partitions, this does not seem to be widely known or explored (Gessel’s survey [9] of $P$-partitions says nothing about the connection to shuffling).

Date: March 14, 2020.

Fulman is supported by Simons Foundation Grant 400528. Petersen is supported by Simons Foundation Collaboration Travel Grant 353772. We thank Persi Diaconis for discussions about shuffling.
Let \( \Pr(\pi) \) denote any of these probability distributions on \( S_n \) coming from card shuffling. If we define the group algebra generating function:

\[
\phi_{Pr}(m) = \sum_{\pi \in S_n} \Pr(\pi) \cdot \pi,
\]

then repeated shuffles correspond to the multiplying \( \phi_{Pr}(m) \) by itself:

\[
\phi_{Pr}(m)^2 = \left( \sum_{\sigma \in S_n} \Pr(\sigma) \cdot \sigma \right) \left( \sum_{\tau \in S_n} \Pr(\tau) \cdot \tau \right),
\]

\[
= \sum_{\pi \in S_n} \left( \sum_{\sigma \in S_n} \Pr(\sigma) \Pr(\sigma^{-1} \pi) \right) \pi.
\]

Each shuffling scheme described generates an ergodic Markov chain, so we have convergence to uniformity in each case:

\[
\phi_{Pr}(m)^k \to \sum_{\pi \in S_n} \frac{1}{n!} \pi \quad \text{as} \quad k \to \infty.
\]

We can explain how Equation (1) is relevant now. For classic \( m \)-riffle shuffling (\( m \) is the number of piles riffled together), we have

\[
\Pr(\pi) = \frac{(m+n-\text{des}(\pi^{-1})-1)}{m^n},
\]

and thus (1) implies that \( \phi_{Pr}(m)^2 = \phi_{Pr}(m^2) \), and for larger \( k \), \( \phi_{Pr}(m^k) = \phi_{Pr}(m^k) \). This means that to analyze repeated shuffles, it suffices to study just one shuffle, but for many piles.

This line of reasoning carries through for the probability distributions coming from other shuffling schemes connected to \( P \)-partitions, as we will explain later in the article. We will reproduce key results in both the classical riffle shuffle and shelf shuffler settings, e.g.,

- probability formulas,
- convolution properties, and
- convergence estimates.

Moreover, we will give analogous new results that use left enriched \( P \)-partitions to analyze a “lazy” shelf shuffler and a corresponding riffle shuffle. In this new situation, the key permutation statistic is the number left peaks, for which we will obtain a recent enumerative result of Gessel and Zhuang [11] about the distribution of left peaks according to cycle type.

**Organization of the paper.** In Section 2 we will show how shuffling and random sampling of \( P \)-partitions are equivalent. In Section 3 we survey enumerative results in the \( P \)-partition literature and translate them into probabilistic statements about shuffling. In Section 4 we give convergence estimates for shuffling, and in Section 5 we study the distribution of cycle structure for lazy shelf shufflers.

2. Equivalence of Shelf Shuffling and \( P \)-partitions

In this section we establish the direct link between \( P \)-partitions and shuffling.
2.1. Shelf shuffling. We will now describe a new method of shuffling that we call *lazy shelf shuffling*, along with the method of shelf shuffling studied by Diaconis, Fulman, and Holmes [4] and inverse riffle shuffling studied by Bayer and Diaconis [1]. Let $n$ denote the number of cards in the deck, and suppose the cards are labeled $1, 2, \ldots, n$ from top to bottom. We quote here from the description in [4] of an actual machine with 10 shelves, manufactured for use in casinos:

A deck of cards is dropped into the top of the box. An internal elevator moves the deck up and down within the box. Cards are sequentially dealt from the bottom of the deck onto the shelves; shelves are chosen uniformly at random at the command of a random number generator. Each card is randomly placed above or below previous cards on the shelf with probability $1/2$. At the end, each shelf contains about $1/10$ of the deck. The ten piles are now assembled into one pile, in random order.

We modify this description only slightly. First, we allow any fixed number $m$ to be the number of shelves onto which we will place the cards. (The actual machine has $m = 10$.) This was of course done in [4]. Second, for convenience, we also assume the cards are sequentially dealt *from the top of the deck*, rather than from the bottom, i.e., we place card 1 first, then card 2, and so on. This choice makes some difference in the combinatorial details (allowing us to work with peaks rather than valleys), but little difference in the statistical analysis. See Remark 2.5. Third, note the final step of assembling the piles into random order is superfluous, so we put the cards on the first shelf on top, followed by the cards on the second shelf, etc.

A more significant difference from the standard shelf shuffler is in the addition of another shelf, at the top of the box, onto which cards may only be placed below previously placed cards.

To get our three card shuffling schemes from our imaginary machine, we install a control panel with buttons that can be used to direct the machine to shuffle in one of three modes. We pretend there are three buttons on the front of the machine, labeled LAZY, STANDARD, and STRICT. Here is a description of each operating mode.

- **Strict mode.** When the machine is in strict mode, it only places cards below cards that are already on a shelf. In this way, strict mode only has to choose a random shelf for each card, each with probability $1/m$. Thus each outcome occurs with probability $1/m^n$. For example, with $m = 3$ piles, and $n = 9$ cards, we might shuffle cards as shown in Table 1. We identify this shuffle with the permutation $\pi = 234569178$ by reading the card labels from the top of the top shelf to the bottom of the bottom shelf. We remark that different outcomes can produce the same permutation, e.g., we could have inserted card 4 in shelf 1 and obtained the same permutation.

- **Standard mode.** This is the card shuffling model studied in [4], except that cards are dealt from the top rather than the bottom. In this mode, for each card $i$ the machine chooses a shelf with probability $1/m$, then places card $i$ at the top of the stack of cards with probability $1/2$, and at the bottom of that stack with probability $1/2$. Alternatively, we can imagine choosing a single (shelf, side) pair with probability $1/(2m)$, where “side” can be “top” or “bottom.” Each outcome thus occurs with probability $1/(2m)^n$. For example, with $m = 2$ piles, and $n = 9$ cards, the machine might shuffle cards as shown in Table 2. We identify this shuffle with the permutation
Table 1. A strict shuffle of \( n = 9 \) cards into \( m = 3 \) shelves resulting in the permutation \( \pi = 234569178 \). Recently placed cards are shaded.

| Card | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------|---|---|---|---|---|---|---|---|---|
| Shelf | 3 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 2 |

Table 2. A standard shuffle of \( n = 9 \) cards into \( m = 2 \) shelves resulting in the permutation \( \pi = 981257436 \).

| Card | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------|---|---|---|---|---|---|---|---|---|
| Shelf | 1t | 1b | 2t | 2t | 1b | 2b | 2t | 1t | 1t |

\( \pi = 981257436 \). Again, the same permutation can result from a variety of different outcomes.
• **Lazy mode.** This mode is the same as standard mode except that there is a “Shelf 0” into which cards can only be placed at the bottom of the shelf. Thus, for each card \( i \), we choose either the lazy shelf or one of the \( 2m \) ordinary (shelf, side) pairs, each with probability \( 1/(2m+1) \). Each outcome of the machine occurs with probability \( 1/(2m+1)^n \). For example, with \( m = 2 \) piles, and \( n = 9 \) cards, we might shuffle cards as shown in Table 3. We identify this shuffle with the permutation \( \pi = 489125736 \). As in prior cases, the same permutation can result from a variety of different outcomes.

2.2. **Riffle shuffling.** Motivated by [4, Section 3.1], we will see that our 3 models of shelf shuffling are equivalent to inverse riffle shuffling for 3 models of riffle shuffling. We now describe the three types of riffle shuffling that correspond to our shelf shuffling machine. In the descriptions below, we use the notation \( A = (a_1, \ldots, a_m) \) for a weak composition of \( n \), with \( a_i \geq 0 \) and \( \sum a_i = n \), and we write the multinomial coefficient as

\[
\binom{n}{A} = \binom{n}{a_1, \ldots, a_m} = \frac{n!}{a_1! \cdots a_m!}.
\]

• **Riffle shuffle.** This is the classic Gilbert-Shannon-Reeds model of card shuffling, as analyzed in [1]. Cut the deck into \( m \) piles according to the multinomial distribution, so we expect the piles to be roughly equal in size. The piles have sizes \( A = (a_1, \ldots, a_m) \) with probability \( \binom{n}{A}/m^n \). To be clear, the first pile contains cards \( 1, \ldots, a_1 \), the second contains cards \( a_1 + 1, \ldots, a_1 + a_2 \), and so on.

We then “riffle” the cards by dropping a card from pile \( i \) with probability proportional to the size of the pile, until all piles are empty. This gives the uniform
distribution on all \( \binom{n}{A} \) interleavings of the piles. Call \( \pi \) the permutation of the cards that results.

We remark that a fixed permutation \( \pi \) can result from many different weak compositions. However given a fixed weak composition \( A \), there is at most one interleaving of the piles indexed by \( A \) that gives \( \pi \). If we keep track of the initial pile sizes as well as \( \pi \), we see the pair \((A, \pi)\) occurs with probability

\[
\frac{\binom{n}{A}}{m^n} \cdot \frac{1}{\binom{n}{A}} = 1/m^n
\]

- **Down-up riffle shuffle.** Cut the deck into \( 2m \) piles according to the multinomial distribution. The piles have sizes \( A = (b_1, a_1, \ldots, b_m, a_m) \) with probability \( \binom{n}{A}/(2m)^n \). This time we put every other pile in reverse order, beginning with the first pile. This indicates the first pile has cards \( b_1, b_1 - 1, \ldots, 1 \), the second pile has cards \( b_1 + 1, \ldots, b_1 + a_1 \), the third pile has cards \( b_1 + a_1 + b_2, b_1 + a_1 + b_2 - 1, \ldots, b_1 + a_1 + 1 \), and so on.

  We now riffle the cards as before to give the uniform distribution on all \( \binom{n}{A} \) interleavings of the piles. Call \( \pi \) the permutation of the cards that results. If we keep track of the initial pile sizes as well as \( \pi \), we see the pair \((A, \pi)\) occurs with probability \( 1/(2m)^n \).

  We note that down-up riffle shuffles were studied in [1] for \( m = 1 \) and for general \( m \) in [6].

- **Up-down riffle shuffle.** This method modifies the down-up riffle shuffle only slightly. First, cut the deck into \( 2m + 1 \) piles according to the multinomial distribution. The piles have sizes \( A = (a_0, b_1, a_1, \ldots, b_m, a_m) \) with probability \( \binom{n}{A}/(2m + 1)^n \). Every other pile is in reverse order, beginning with the second pile. This gives the first pile as cards \( 1, 2, \ldots, a_0 \), the second pile as \( a_0 + b_1, a_0 + b_1 - 1, \ldots, a_0 + 1 \), the third pile as cards \( a_0 + b_1 + 1, \ldots, a_0 + b_1 + a_1 \), and so on.

  Again we riffle the cards to give the uniform distribution on all \( \binom{n}{A} \) interleavings of the piles. Call \( \pi \) the permutation of the cards that results. If we keep track of the initial pile sizes as well as \( \pi \), we see the pair \((A, \pi)\) occurs with probability \( 1/(2m + 1)^n \).

  We note that up-down riffle shuffles were studied in [2] for \( m = 1 \) and for general \( m \) in [6].

2.3. **P-partitions.** Now we turn our attention to \( P \)-partitions, but first we first discuss background for partially ordered sets. See [15] Chapter 4 for more.

Throughout, we fix a positive integer \( n \) and let \( P \) denote a partial ordering of the set \([n] = \{1, 2, \ldots, n\}\). We write \( \prec_P \) for the order relation on \( P \), i.e., if \( i \) is below \( j \) in \( P \), we say \( i \) and \( j \) are comparable and write \( i \prec_P j \) or \( j >_P i \). If neither \( i <_P j \) nor \( i >_P j \), we say \( i \) and \( j \) are incomparable. A comparable pair \( i <_P j \) is naturally labeled if \( i <_N j \) as well. Otherwise, the pair is unnaturally labeled.

A chain is a poset in which any two elements are comparable. The \( n \) element antichain, denoted \([n]\), is the poset with no relations. We readily identify chains with permutations, via

\[
\pi(1) <_\pi \pi(2) <_\pi \cdots <_\pi \pi(n),
\]

whenever \( P = \pi \) is a chain. To say that \( i <_\pi j \) is equivalent to saying that \( \pi^{-1}(i) <_N \pi^{-1}(j) \).
We say that $Q$ refines $P$ if every relation in $P$ is a relation in $Q$. That is, $i <_P j$ implies $i <_Q j$. In this setting, chains are maximally refined posets. We define the set of linear extensions of $P$ to be the set of chains (permutations) $\pi$ such that $\pi$ refines $P$:

$$L(P) = \{ \pi \in S_n : i <_P j \Rightarrow i <_\pi j \}.$$  

In the definition of $P$-partitions, the set of integers $\mathbb{Z}$ is given the ordering $0 < Z_1 < Z_2 < \cdots$, and we define symbols “$\leq^+$” and “$\leq^-$” as follows:

$$a \leq^+ b \iff a <_Z b \text{ or } a = b \in \{0, 1, 2, 3, \ldots\},$$
$$a \leq^- b \iff a <_Z b \text{ or } a = b \in \{-1, -2, -3, \ldots\}.$$  

**Definition 2.1** ($P$-partition). A $P$-partition is an order preserving function $f : P \to \mathbb{Z}$ such that for $i <_P j$:

- $f(i) \leq^+ f(j)$ if $i <_N j$,
- $f(i) \leq^- f(j)$ if $i >_N j$.

In other words, the values of a $P$-partition on a naturally labeled pair are only allowed to agree on nonnegative values, while the values on an unnaturally labeled pair can only agree on negative values. We denote the set of $P$-partitions by $A(P)$.

We consider three subsets of $P$-partitions, characterized by restrictions on the image of $f$.

- **Positive $P$-partitions.** A $P$-partition $f$ whose image is in $\mathbb{N}$ is equivalent to the order-preserving version of Stanley’s original definition of a $P$-partition [15, Chapter 4], i.e., $i <_P j$ implies $f(i) \leq_N f(j)$ with $f(i) < f(j)$ if $i >_N j$ is an unnaturally labeled pair. We denote this subset of $P$-partitions as follows:

$$A^+(P) = \{ f \in A(P) : f(i) \in \mathbb{N} \text{ for all } i \in P \}.$$  

This set will help to encode strict shelf shuffling and classic riffle shuffling.

- **Nonzero $P$-partitions.** The $P$-partitions whose image does not contain 0 are precisely Stembridge’s enriched $P$-partitions [17]. We denote this set as:

$$A^*(P) = \{ f \in A(P) : f(i) \neq 0 \text{ for all } i \in P \}.$$  

This set will help to encode standard shelf shuffling and down-up riffle shuffling.

- **All $P$-partitions.** Without any restrictions, this is precisely the definition of left enriched $P$-partitions given by Petersen [13]. This set will help to encode lazy shelf shuffling and up-down riffle shuffling.

For example, if $n = 3$ and $P$ is the poset with $1 <_P 2$ and $3 <_P 2$, it has linear extensions $L(P) = \{132, 312\}$. We can draw the poset and its extensions with Hasse diagrams as indicated here:

\[
P: \begin{array}{c}
1 \quad 3 \\
\end{array} \quad \begin{array}{c}
L(P): \\
1 \quad 2 \\
2 \quad 3 \\
\end{array}
\]

In this case, every $P$-partition $f$ must satisfy

$$f(1) \leq^+ f(2) \geq^- f(3),$$
so
\[ A(P) = \{(a_1, a_2, a_3) \in \mathbb{Z}^3 : a_1 \leq a_2 \geq a_3\}. \]
We can write this as a disjoint union:
\[ A(P) = \{a_1 \leq a_3 \leq a_2\} \cup \{a_3 \leq a_1 \leq a_2\}. \]
But each of these smaller sets can viewed as the \( P \)-partitions for a chain:
\[ A(P) = A(132) \cup A(312). \]

By induction on the number of incomparable pairs in a general poset \( P \), we can see that, in general, the set of all \( P \)-partitions is the disjoint union of the \( P \)-partitions for its chains.

**Theorem 2.2** ([15], Lemma 4.5.3). The set of \( P \)-partitions is the disjoint union of the \( \pi \)-partitions of its linear extensions:
\[ A(P) = \bigcup_{\pi \in \mathcal{L}(P)} A(\pi). \]

The antichain \( P = [n] \) is a special case worth considering here, since every \( P \)-partition of \( n \) elements is a \( P \)-partition for the antichain. Moreover since
\[ A([n]) = \bigcup_{\pi \in S_n} A(\pi), \]
this implies that each \( P \)-partition \( f \) belongs to just one subset \( A(\pi) \). This determines a unique permutation \( \pi = \pi(f) \), which we call the sorting permutation for \( f \).

**Definition 2.3** (Sorting permutation of a \( P \)-partition). Each \( P \)-partition \( f \) determines a unique permutation \( \pi = \pi(f) \). We define \( \pi \) from \( f \) via sorting the pairs \((i, f(i))\) according to:
- if \( f(i) < f(j) \), then \( i <_\pi j \),
- if \( i <_\pi j \) and \( f(i) = f(j) \in \{0, 1, 2, \ldots\} \), then \( i <_\pi j \),
- if \( i <_\pi j \) and \( f(i) = f(j) \in \{-1, -2, -3, \ldots\} \), then \( j <_\pi i \).

We remark that this definition for a sorting permutation makes sense for nonzero and positive \( P \)-partitions as well.

For example, suppose \( n = 9 \) and we have the following \( P \)-partition \( f \), written in two-line notation with \( f(i) \) below \( i \):
\[
(2) \quad f = \begin{pmatrix}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
-1 & 0 & 0 & -2 & -1 & 1 & 0 & 2 & 2
\end{pmatrix}.
\]
The image multiset of \( f \) is \( \{0, 0, 0, -1, -1, 1, -2, 2, 2\} \), which we can denote \( \{0^3, -1^2, 1^1, -2^1, 2^2\} \) for brevity. Since \( f(2) = f(3) = f(7) = 0 \), we know 2, 3, and 7 must be the first three entries of \( \pi \) and since the image is 0, these must be in their natural order. Thus, \( \pi(1)\pi(2)\pi(3) = 237 \).

Similarly, since \( f(1) = f(5) = -1 \) is the next biggest value of \( f \), we know the next two entries of \( \pi \) are 1 and 5. Since the image of these two elements is \(-1\), they must appear in reverse order, i.e., \( \pi(4)\pi(5) = 51 \).
Continuing in this way, we can deduce the values of each entry of \( \pi = \pi(f) = 237516489 \).

\(^1\)Stanley says \( f \) is \( \pi \)-compatible in this situation ([16] Section 7.19). This perspective emphasizes \( f \) relative to \( \pi \), whereas our terminology emphasizes \( \pi \) relative to \( f \).
Thinking in terms of the two-line notation itself, we simply sort the array from left to right according to the bottom row. When there are ties, we sort in increasing order on the top row for \(f(i)\) nonnegative, and in decreasing order for \(f(i)\) negative:

\[
f \circ \pi = \begin{pmatrix} 2 & 3 & 7 & 5 & 1 & 6 & 4 & 8 & 9 \\ 0 & 0 & 0 & -1 & -1 & 1 & -2 & 2 & 2 \end{pmatrix}.
\]

We see \(\pi(f) = 237516489\) in the top line of the sorted array.

2.4. \(P\)-partitions encode shuffles. The connection between \(P\)-partitions and shelf shufflers comes from bounding \(P\)-partitions. That is, let \(\mathcal{A}(P; m)\) denote the set of \(P\)-partitions with absolute value bounded by \(m\), i.e.,

\[
\mathcal{A}(P; m) = \{f \in \mathcal{A}(P) : |f(i)| \leq m\}.
\]

When \(P = [n]\) is an antichain, we see each \(f \in \mathcal{A}([n]; m)\) is merely a record of exactly what the shelf shuffler did with each card while in lazy mode. The following proposition should be roughly self-evident.

**Proposition 2.4** (Shelf shuffling and \(P\)-partitions). There is a bijection between the set of outcomes of an \(m\)-shelf shuffler in lazy mode and \(\mathcal{A}([n]; m)\). Namely, we place cards one at a time, from \(i = 1, 2, \ldots, n\). We place card \(i\) on shelf \(|f(i)|\). If \(f(i) < 0\) the card is placed on top of the cards on the shelf, while if \(f(i) \geq 0\) the card is placed below the cards already on the shelf. The ordering of the cards after shuffling is given by \(\pi(f)\).

For example, the \(P\)-partition in (2) corresponds to the shelf shuffler taking card 1 and placing it on top of shelf 1, card 2 on the bottom of shelf 0, card 3 on the bottom of shelf 0, and so on.

This correspondence between shuffles and \(P\)-partitions restricts to other modes in the obvious way. If we want to encode the strict shuffle mode, we use the set of positive \(P\)-partitions for the antichain,

\[
\mathcal{A}^+([n]; m) = \{f \in \mathcal{A}^+([n]) : |f(i)| \leq m\},
\]

while if we want to encode the standard shuffle mode, we use nonzero \(P\)-partitions,

\[
\mathcal{A}^*([n]; m) = \{f \in \mathcal{A}^*([n]) : |f(i)| \leq m\}.
\]

**Remark 2.5.** The shelf shufflers in this paper drop cards on shelves one at a time from the top of the deck, whereas in [4], the cards are dealt into shelves one at a time from the bottom of the deck. The correspondence between \(P\)-partitions and shelf shuffling in Proposition 2.4 can be modified to match the shuffling mechanism from [4] as follows. Dealing from the bottom means that we place cards \(i = n, \ldots, 2, 1\). The \(P\)-partition example from (2),

\[
f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ -1 & 0 & 0 & -2 & -1 & 1 & 0 & 2 & 2 \end{pmatrix},
\]

is now interpreted as “place card 9 on the bottom of shelf 2, then place card 8 on the bottom of shelf 2, place card 7 on the bottom of shelf 0,” and so on. We end up sorting the array as

\[
\begin{pmatrix} 7 & 3 & 2 & 1 & 5 & 6 & 4 & 9 & 8 \\ 0 & 0 & 0 & -1 & -1 & 1 & -2 & 2 & 2 \end{pmatrix},
\]

so the permutation from the bottom-dealing shelf-shuffler gives permutation 732156498 rather than the permutation \(\pi(f) = 237516489\) we found previously. Cards on each shelf
alternately increase then decrease with the bottom-dealing mechanism, rather than decreasing then increasing with the top-dealing mechanism. (And the bottom-dealing mechanism has the has cards in decreasing order on shelf 0, rather than increasing order.) We choose the top-dealing mechanism for convenience.

From a statistical standpoint, i.e., the convergence of shuffling to uniformity, it should be clear the difference is trivial. Combinatorially, the effect of this choice will ultimately be for us to explain probabilistic results in terms of peaks rather than valleys of permutations. The translation between peaks and valleys is discussed in Section 3 of [4], particularly in the discussion around Theorem 3.2 and the proof Theorem 3.1.

We have an analogous correspondence between \( P \)-partitions and our three flavors of riffle shuffles, but it is less immediately obvious. To explain the idea for up-down riffle shuffles, suppose

\[
A = (a_0, b_1, a_1, \ldots, b_m, a_m),
\]

is a weak composition of \( n \). Define the poset \( P_A \) to be the union of chains that correspond to the piles formed in the process of an up-down riffle shuffle, with the first \( a_0 \) cards in increasing order, the next \( b_1 \) cards in decreasing order, and so on. For example, if \( A = (3, 4, 3, 4, 2) \), we have \( P_A \) is the disjoint union of the five chains below (recall we read the chains from the bottom up):

\[
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & 6 \\
7 & 8 & 9 & 10 & 11 & 12 \\
13 & 14 & 15 & 16 & & \\
\end{array}
\]

Any linear extension \( \sigma \in \mathcal{L}(P_A) \) corresponds precisely to one of the interleavings of the stacks of cards in the riffle shuffle. Moreover, we have \( 1 <_\sigma 2 <_\sigma 3 \), as well as \( 4 >_\sigma 5 >_\sigma 6 >_\sigma 7 \), and so on, or equivalently:

\[
\sigma^{-1}(1) < \sigma^{-1}(2) < \sigma^{-1}(3), \quad \sigma^{-1}(4) > \sigma^{-1}(5) > \sigma^{-1}(6) > \sigma^{-1}(7), \ldots.
\]

But this means \( \sigma^{-1} \) is the sorting permutation for the \( P \)-partition \( f \) with image multiset

\[
\{0, 0, 0, -1, -1, -1, 1, 1, 1, -2, -2, -2, 2, 2, 2\} = \{0^3, -1^4, 1^3, -2^4, 2^2\}.
\]

We summarize the general correspondence in the following proposition.

**Proposition 2.6** (Up-down riffle shuffling and \( P \)-partitions). There is a bijection between the set of outcomes of up-down \( m \)-riffle shuffling and \( \mathcal{A}([n]; m) \). Namely, if \((A, \sigma)\) is an outcome of the shuffle with \( A = (a_0, b_1, a_1, \ldots, b_m, a_m) \), then it corresponds to that \( f \in \mathcal{A}([n]; m) \) such that the image of \( f \) is \( \{0^{a_0}, -1^{b_1}, 1^{a_1}, \ldots, -m^{b_m}, m^{a_m}\} \) and \( \sigma^{-1} = \pi(f) \) is the sorting permutation of \( f \).

We can modify Proposition 2.6 to show outcomes of ordinary \( m \)-riffle shuffling are in bijection with \( \mathcal{A}^+(|[n]; m) \) and outcomes of \( m \)-down-up riffle shuffling correspond to elements of \( \mathcal{A}^*([n]; m) \).
3. Shuffling probabilities from \( P \)-partitions

In this section we will survey some enumerative results for \( P \)-partitions and use Propositions 2.4 and 2.6 to translate them into probabilistic results for shuffling.

3.1. Enumerative results for \( P \)-partitions. The order polynomial for \( P \), denoted \( \Omega_P(m) \), counts the number of \( P \)-partitions bounded by \( m \), i.e.,

\[
\Omega_P(m) = |A(P; m)| = |\{ f \in A(P) : |f(i)| \leq m \}|.
\]

We similarly define \( \Omega_P^\ast(m) = |A^\ast(P; m)| \) and \( \Omega_P^+(m) = |A^+(P; m)| \).

An immediate corollary of Theorem 2.2 is that order polynomials are sums of order polynomials for linear extensions\(^2\).

**Corollary 3.1.** The order polynomial for a poset \( P \) is the sum of the order polynomials for its linear extensions:

\[
\Omega_P(m) = \sum_{\pi \in \mathcal{L}(P)} \Omega_{\pi}(m),
\]

and similarly for \( \Omega_P^\ast(m) \) and \( \Omega_P^+(m) \).

Order polynomials for antichains are easy enough, since there are no relations to worry about.

**Observation 3.2** (Antichain order polynomials). Let \([n]\) denote the antichain on \( n \) elements. We have, for any \( m \geq 0 \),

\[
\Omega_{[n]}(m) = (2m + 1)^n,
\]

\[
\Omega_{[n]}^\ast(m) = (2m)^n,
\]

\[
\Omega_{[n]}^+(m) = m^n.
\]

The other extreme situation is the case of chains. To understand enumerative properties for order polynomials of chains, we need to discuss permutation statistics. A descent of a permutation \( \pi \) is an index \( i \) such that \( \pi(i) > \pi(i+1) \). We let des(\( \pi \)) denote the number of descents of \( \pi \). A peak of a permutation is an index \( i \) such that \( \pi(i-1) < \pi(i) > \pi(i+1) \), i.e., a descent preceded by a non-descent. The number of peaks is denoted by pk(\( \pi \)). A left peak is a peak or a descent in position 1. We let lpk(\( \pi \)) denote the number of left peaks. We have \( \text{lpk}(\pi) = \text{pk}(\pi) \) if \( \pi(1) < \pi(2) \) and \( \text{lpk}(\pi) = \text{pk}(\pi) + 1 \) if \( \pi(1) > \pi(2) \). We can now give relatively simple expressions for the order polynomials of chains.

---

\(^2\)In fact, this corollary, along with analysis of the case of a chain, gives a simple way to prove that order polynomials are actually polynomials.
Proposition 3.3 (Order polynomials for chains). Let $\pi$ denote a chain on $[n]$, i.e., a permutation in $S_n$. We have the following expressions for its order polynomials:

\[
\sum_{m \geq 0} \Omega_\pi(m) t^m = \frac{(4t)^{lpk(\pi)}(1 + t)^{n - 2lpk(\pi)}}{(1 - t)^{n+1}},
\]

\[
\sum_{m \geq 0} \Omega_\pi^*(m) t^m = \frac{(4t)^{pk(\pi)+1}(1 + t)^{n - 2pk(\pi)-1}}{2(1 - t)^{n+1}},
\]

\[
\sum_{m \geq 0} \Omega_\pi^+(m) t^m = \frac{t^{des(\pi)+1}}{(1 - t)^{n+1}}.
\]

Equivalently,

\[
\Omega_\pi(m) = 4^{lpk(\pi)} \sum_{a \geq 0} \binom{n + m - a}{n} \binom{n - 2 lpk(\pi)}{a - lpk(\pi)},
\]

\[
\Omega_\pi^*(m) = 2 \cdot 4^{pk(\pi)} \sum_{a \geq 0} \binom{n - 1 + m - a}{n} \binom{n - 1 - 2 pk(\pi)}{a - pk(\pi)}, \text{ and}
\]

\[
\Omega_\pi^+(m) = \binom{n - 1 + m - des(\pi)}{n}.
\]

Equation (5) is proved in \[15\] Theorem 4.5.14, Equation (4) is from \[17\] Theorem 4.1, and Equation (3) is from \[13\] Theorem 4.6.

The theory of $P$-partitions yields the following identities as well, which are useful for explaining repeated shuffles.

Proposition 3.4. For any integers $k$ and $l$ and any permutation $\pi \in S_n$,

\[
\Omega_\pi(2kl + k + l) = \sum_{\sigma \tau = \pi} \Omega_\sigma(k) \Omega_\tau(l),
\]

\[
\Omega_\pi^*(2kl) = \sum_{\sigma \tau = \pi} \Omega_\sigma^*(k) \Omega_\tau^*(l),
\]

\[
\Omega_\pi^+(kl) = \sum_{\sigma \tau = \pi} \Omega_\sigma^+(k) \Omega_\tau^+(l).
\]

We remark that Equation (11) is a restatement of identity (8) for the positive order polynomials $\Omega_\pi^+$. Each of these identities follows from a similar paradigm of decomposing a bipartite $P$-partition $f : \pi \to S \times T$. The image of such a function is a collection of pairs

\[
f(\pi) = \{(f_1(1), f_2(1)), (f_1(2), f_2(2)), \ldots, (f_1(n), f_2(n))\},
\]

which can be re-interpreted as a pair of $P$-partitions $f_1 : \sigma \to S$ and $f_2 : \tau \to T$, such that $\sigma \tau = \pi$. More details can be found in \[13\], which proves each of (9), (7), and (8). It should be noted, though, that the case for $\Omega_\pi^+$ (which provides the motivation for the other cases) follows from earlier work of Gessel \[8\].
When desired, we can drop the permutation from the notation for order polynomials and write only the statistic. That is, fix $n$ and let
\[
\Omega(n, k; m) = \Omega_n(m) \text{ for some } \pi \in S_n \text{ with } \text{lpk}(\pi) = k,
\]
\[
\Omega^*(n, k; m) = \Omega_n^*(m) \text{ for some } \pi \in S_n \text{ with } \text{pk}(\pi) = k, \text{ and}
\]
\[
\Omega^+(n, k; m) = \Omega_n^+(m) \text{ for some } \pi \in S_n \text{ with } \text{des}(\pi) = k.
\]

The following lemma will be useful for some of our convergence estimates later on. For strict shelf shufflers it is obvious from the explicit formula for $\Omega^+(n, k; m)$. For standard shelf shufflers it was given a probabilistic proof in [4]. We believe the result to be new for lazy shelf shufflers, and the proof method to be new for all three cases. We are able to extend the probabilistic proof of [4] to the lazy setting, but omit the details as we believe our proof to be more conceptual.

**Lemma 3.5** (Monotonicity Lemma). Order polynomials are monotone decreasing in their respective statistical indices, i.e., for any $m,n,$ and $k$ we have
\[
\Omega(n, k; m) \geq \Omega(n, k + 1; m),
\]
\[
\Omega^*(n, k; m) \geq \Omega^*(n, k + 1; m),
\]
\[
\Omega^+(n, k; m) \geq \Omega^+(n, k + 1; m).
\]

**Proof.** The argument in each case is to choose a canonical permutation $\pi$ with statistic $k+1$, another permutation $\pi'$ with statistic $k$, and construct an injection from $A(\pi) \to A(\pi')$.

We will handle $\Omega(n, k; m)$ in detail. The arguments for $\Omega^+$ and $\Omega^*$ are similar.

Fix $n$, fix $k + 1 \leq n/2$ and let $\pi \in S_n$ be the permutation that swaps $2i$ and $2i - 1$, for each $i = 1, \ldots, k + 1$. In one-line notation,
\[
\pi = 214365 \cdots (2k)(2k - 1)(2k + 2)(2k + 1)(2k + 3) \cdots n.
\]

Let $\pi'$ be similar, but with the $k + 1$st pair swapped:
\[
\pi' = 214365 \cdots (2k)(2k - 1)(2k + 1)(2k + 2) \cdots n.
\]

By construction, $\text{lpk}(\pi) = k + 1$ and $\text{lpk}(\pi') = k$.

The condition for a function $f$ in $A(\pi)$ is:
\[
f(2) \leq f(1) \leq \cdots \leq f(2k - 1) \leq^+ f(2k + 2) \leq^+ f(2k + 1) \leq^+ f(2k + 3) \leq^+ \cdots \leq^+ f(n).
\]

And the condition for a function $g$ in $A(\pi')$ is:
\[
g(2) \leq g(1) \leq^+ \cdots \leq^+ g(2k - 1) \leq^+ g(2k + 1) \leq^+ g(2k + 2) \leq^+ g(2k + 3) \leq^+ \cdots \leq^+ g(n).
\]

Now let $f \in A(\pi)$. There are two cases to consider: either $f(2k + 2) < f(2k + 1)$ or $f(2k + 2) = f(2k + 1) < 0$.

On the one hand, suppose $f(2k + 2) < f(2k + 1)$. To get a $P$-partition for $\pi'$ we define $g = f'$ by:
\[
\bullet \ f'(2k + 1) = f(2k + 2), \\
\bullet \ f'(2k + 2) = f(2k + 1), \text{ and} \\
\bullet \ f'(i) = f(i) \text{ otherwise.}
\]

On the other hand, suppose $f(2k + 2) = f(2k + 1) = -a < 0$. Now we define $g = f'$ by:
\[
\bullet \ f'(2k + 1) = f'(2k + 2) = a, \text{ and} \\
\bullet \ f'(i) = f(i) \text{ otherwise.}
\]
In each case, we can check that \( f'(2k - 1) \leq f'(2k + 1) \) and \( f'(2k + 2) \leq f'(2k + 3) \), so \( f' \) satisfies all the conditions of (12), and is indeed a \( P \)-partition for \( \pi' \).

The functions \( f' \) constructed in the first case have \( f'(2k + 1) \neq f'(2k + 2) \), so the two cases do not overlap, yielding the desired injection \( f \mapsto f' \). □

3.2. Consequences for shuffling probabilities. We now connect the results for \( P \)-partitions to shuffling probabilities. To begin, let \( m \) and \( n \) be positive integers, and let \( \pi \) be a permutation in \( S_n \). Consider an \( m \)-shelf shuffler, and define

- \( x_m(\pi) \) to be the probability of obtaining \( \pi \) in lazy mode,
- \( x_m^*(\pi) \) to be the probability of obtaining \( \pi \) in standard mode, and
- \( x_m^+(\pi) \) to be the probability of obtaining \( \pi \) in strict mode.

Similarly, we define \( y_m(\pi), y_m^*(\pi), \) and \( y_m^+(\pi) \) to be probabilities of obtaining \( \pi \) from an \( m \)-up-down riffle shuffle, an \( m \)-down-up riffle shuffle, and a classic \( m \)-riffle shuffle, respectively.

Our choice of notation is suggestive of the following exact formulas for these probabilities in terms of order polynomials. We believe this to be new in the lazy case.

**Proposition 3.6.** For each permutation \( \pi \in S_n \) and each positive integer \( m \), we have the following expressions for shuffling probabilities:

\[
\begin{align*}
    x_m(\pi) &= \frac{\Omega_\pi(m)}{(2m + 1)^n}, \\
    x_m^*(\pi) &= \frac{\Omega_\pi^*(m)}{(2m)^n}, \\
    x_m^+(\pi) &= \frac{\Omega_\pi^+(m)}{m^n}.
\end{align*}
\]

Moreover, the probabilities for shelf-shuffling and riffle shuffling are related via

\[
    y_m(\pi) = x_m(\pi^{-1}), \quad y_m^*(\pi) = x_m^*(\pi^{-1}), \quad \text{and} \quad y_m^+(\pi) = x_m^+(\pi^{-1}).
\]

In essence, this result says that the probability of obtaining \( \pi \) from shelf shuffling is the probability that a random \( P \)-partition has sorting permutation \( \pi \), and the probability of obtaining \( \pi \) from riffle shuffling is the probability that a random \( P \)-partition has sorting permutation \( \pi^{-1} \).

**Proof.** By Proposition 2.4 we know that choosing a random \( P \)-partition in \( A(P; m) \) is equivalent to a random outcome of an \( m \)-shelf shuffler in lazy mode. By Observation 3.2 there are \( \Omega_{[m]}(m) = (2m + 1)^n \) such outcomes. By Definition 2.3 and the definition of the order polynomial, precisely \( \Omega_\pi(m) \) of these correspond to the permutation \( \pi \). This proves Equation (13) for \( x_m(\pi) \).

The fact that \( x_m(\pi) = y_m(\pi^{-1}) \) is an immediate consequence of Proposition 2.6.

The arguments proving (14) and (15) (for standard and strict shuffling modes) are similar. □

We next consider, in each mode (lazy, standard, strict), a generating function for the entire probability distribution as an element in the group algebra of the symmetric group. That
is, define

\[ \phi_n(m) = \sum_{\pi \in S_n} x_m(\pi) \pi, \]
\[ \phi_n^+(m) = \sum_{\pi \in S_n} x_m^+(\pi) \pi, \]
and
\[ \phi_n^*(m) = \sum_{\pi \in S_n} x_m^*(\pi) \pi. \]

As a corollary to Proposition 3.4, we get the following identities for the distributions.

**Corollary 3.7.** For each \( n \geq 1 \) and \( k, l \geq 0 \), we have

(16) \[ \phi_n(k)\phi_n(l) = \phi_n(2kl + k + l), \]

(17) \[ \phi_n^*(k)\phi_n^*(l) = \phi_n^*(2kl), \]

(18) \[ \phi_n^+(k)\phi_n^+(l) = \phi_n^+(kl). \]

**Proof.** We handle the lazy mode case in detail. Other shuffling modes are similar.

We have

\[ \phi_n(k)\phi_n(l) = \left( \sum_{\sigma \in S_n} x_k(\sigma) \sigma \right) \left( \sum_{\tau \in S_n} x_l(\tau) \tau \right), \]

\[ = \left( \sum_{\sigma \in S_n} \frac{\Omega(\sigma)}{(2k+1)^n} \sigma \right) \left( \sum_{\tau \in S_n} \frac{\Omega(\tau)}{(2l+1)^n} \tau \right), \]

\[ = \frac{1}{(4kl+2k+2l+1)^n} \sum_{\pi \in S_n} \left( \sum_{\sigma \tau = \pi} \Omega(\sigma) \Omega(\tau) \right) \pi. \]

But equation (6) gives

\[ \sum_{\sigma \tau = \pi} \Omega(\sigma) \Omega(\tau) = \Omega_\pi(2kl + k + l), \]

so we obtain

\[ \phi_n(k)\phi_n(l) = \frac{1}{(2(2kl + k + l) + 1)^n} \sum_{\pi \in S_n} \Omega_\pi(2kl + k + l) \pi, \]

\[ = \sum_{\pi \in S_n} x_{2kl+k+l}(\pi) \pi, \]

\[ = \phi_n(2kl + k + l), \]

as claimed. \( \square \)

The immediate consequence of these identities has to do with repeated shuffles. For example, the distribution after two lazy \( m \)-shuffles is \( \phi(m)^2 \), and Corollary 3.7 tells us that \( \phi(m)^2 = \phi(2m^2 + 2m) \), so two sequential lazy \( m \)-shelf shuffles gives the same distribution as one lazy \((2m^2 + 2m)\)-shuffle. This means two sequential lazy 10-shelf shuffles give the same distribution as one pass through a lazy 220-shelf shuffler.
4. Convergence results

In both [1] and [4] we get estimates for how quickly shuffling converges to the uniform distribution on $S_n$. We follow those papers in considering the following measures for any probability distribution $P$ on $S_n$. We let $U$ be the uniform distribution, so that $U(\pi) = 1/n!$ for all $\pi$ in $S_n$, and define the total variation distance

$$\| P - U \|_{TV} = \frac{1}{2} \sum_{\pi \in S_n} | P(\pi) - U(\pi) | = \frac{1}{2} \sum_{\pi \in S_n} \left| P(\pi) - \frac{1}{n!} \right|,$$

the separation distance

$$\text{sep}(P) = \max_{\pi \in S_n} \left( 1 - \frac{P(\pi)}{U(\pi)} \right) = \max_{\pi \in S_n} (1 - nP(\pi)),$$

and $l_\infty$ distance

$$\| P - U \|_\infty = \max_{\pi \in S_n} \left| 1 - \frac{P(\pi)}{U(\pi)} \right| = \max_{\pi \in S_n} |1 - nP(\pi)|.$$

It is elementary that $\| P - U \|_{TV} \leq \text{sep}(P) \leq \| P - U \|_\infty$.

Let

$$x_m(k) = x_m(\pi)$$

for some $\pi$ with $\text{lpk}(\pi) = k$,

and similarly define $x_m^*(k)$ in terms of peaks and $x_m^+(k)$ in terms of descents.

Taking the expressions for order polynomials from Proposition 3.3, we find the following expressions for our probabilities:

$$x_m(k) = \frac{4^k}{(2m+1)^n} \sum_{a \geq 0} \binom{n-m-a}{n} \binom{n-2k}{a-k},$$

$$x_m^*(k) = \frac{4^{k+1}}{2(2m)^n} \sum_{a \geq 0} \binom{n-1+m-a}{n} \binom{n-1-2k}{a-k},$$

and

$$x_m^+(k) = \frac{1}{m^n} \binom{n-1+m-k}{n}.$$

For example in the lazy case, our total variation distance can be expressed as

$$\| x_m - U \|_{TV} = \frac{1}{2} \sum_{k=0}^{\left\lfloor n/2 \right\rfloor} l(n, k) \left| x_m(k) - \frac{1}{n!} \right|,$$

where $l(n, k) = |\{ \pi \in S_n : \text{lpk}(\pi) = k \}|$. From [12], the numbers $l(n, k)$ satisfy the recurrence

$$l(n, k) = (2k+1)l(n-1, k) + (n+1-2k)l(n-1, k-1),$$

with boundary conditions $l(n, 0) = 1$ and $l(n, k) = 0$ if $k > n/2$. The formulas make it not too difficult to use Equation (19) to compute the total variation distance for realistic values of $m$ and $n$. For example, in Table 3 we see total variation distance for various values of $m$ and $n = 52$, comparing $x_m$ (lazy), $x_m^*$ (standard), and $x_m^+$ (strict). We note also that if $m = 10$ and we pass through the lazy shuffler twice it is the same as $m = 220$. In this case $\| x_m - U \|_{TV} = .0083$.

The monotonicity lemma for order polynomials, Lemma 3.5, implies that $x_m(k) \geq x_m(k+1)$, $x_m^*(k) \geq x_m^*(k+1)$, and $x_m^+(k) \geq x_m^+(k+1)$. Thus, we see that both the $l_\infty$ and separation distances are achieved at the extremes.
Table 4. Total variation distance for shelf shufflers with \( m \) shelves and \( n = 52 \) cards, in each of the three operating modes.

| \( m \)  | 10  | 15  | 20  | 25  | 30  | 35  | 50  | 100 | 200 | 250 | 300 |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lazy   | 1   | .9372 | .7184 | .5164 | .3936 | .3003 | .1509 | .0392 | .0177 | .0100 | .0064 | .0045 |
| Standard| 1   | .9427 | .7201 | .5440 | .3910 | .2993 | .1586 | .0409 | .0183 | .0103 | .0066 | .0046 |
| Strict  | 1   | 1   | 1.9981 | 1.9825 | 1.9468 | 1.8932 | 1.7336 | 1.4199 | 1.2857 | 1.2131 | 1.1709 | 1.1438 |

**Observation 4.1.** For any distribution \( P \in \{x_m, x_m^*, x_m^+\} \), we have

\[
\text{sep}(P) = \max\{1 - n!P(0), 1 - n!P(k_{\text{max}})\},
\]

and

\[
\|P - U\|_\infty = \max\{1 - n!P(0), |1 - n!P(k_{\text{max}})|\},
\]

where \( k_{\text{max}} = \lfloor n/2 \rfloor, \lfloor (n - 1)/2 \rfloor \), and \( n - 1 \), for \( x_m, x_m^*, \) and \( x_m^+ \), respectively.

The \( l_\infty \) and separation distances are easy to study for strict shelf shufflers using the explicit formula for \( x_m^+(\pi) \). More subtle calculations are required for standard shelf shufflers \([4]\). For lazy shelf shufflers the asymptotics is the same as for standard shelf shufflers. More precisely, we have the following result, proved by the method of \([4]\).

**Theorem 4.2.** Consider the lazy shelf shuffling measure \( P_m \) with \( n \) cards and \( m \) shelves (and the additional shelf 0). Suppose \( m = cn^{3/2} \). Then as \( n \to \infty \) with \( 0 < c < \infty \) fixed,

\[
\|P_m - U\|_\infty \sim e^{1/(12c^2)} - 1
\]

\[
\text{sep}(P_m) \sim 1 - e^{-1/(24c^2)}
\]

5. **Cycle structure for lazy shelf shufflers**

The cycle structure of strict shelf shufflers is the same as the cycle structure of ordinary riffle shuffles, carefully studied in \([5]\). The cycle structure of standard shelf shufflers is studied in \([4]\). In this section we find a generating function for cycle structure of lazy shelf shufflers. Then we use it to derive the joint distribution of permutations by cycles and left peaks.

Let \( N_i(\pi) \) denote the number of \( i \)-cycles of a permutation \( \pi \), and define

\[
f_{i,m} = \frac{1}{2i} \sum_{\substack{d | i \cr d \text{ odd}}} \mu(d)(2m + 1)^{i/d} - 1,
\]

where \( \mu \) is the Moebius function of elementary number theory. Let \( x_m(\pi) \) and \( y_m(\pi) \) be as in previous sections.

**Theorem 5.1.**

\[
1 + \sum_{n \geq 1} u^n \sum_{\pi \in S_n} x_m(\pi) \prod_{i \geq 1} z_i^{N_i(\pi)} = \frac{1}{1 - z_1 u/(2m + 1) \prod_{i \geq 1} \left( 1 + z_i u/(2m + 1)^i \right)} \prod_{i \geq 1} \left( 1 - z_i u/(2m + 1)^i \right)^{f_{i,m}}.
\]
Proof. It follows from Theorem 7 of [6] that
\[
1 + \sum_{n \geq 1} u^n \sum_{\pi \in S_n} y_m(\pi) \prod_{i \geq 1} z_i^{N_i(\pi)}
\]
\[= \frac{1}{1 - z_1 u/(2m + 1)} \prod_{i \geq 1} \left( \frac{1 + z_i u^i / (2m + 1)^i}{1 - z_i u^i / (2m + 1)^i} \right)^{f_{i,m}}.\]

Note that a permutation and its inverse have the same cycle structure, and recall from Proposition 3.6 that \(x_m(\pi) = y_m(\pi^{-1})\). The result follows. \(\square\)

The generating function in Theorem 5.1 allows one to study cycle structure for lazy shelf-shufflers, in perfect analogy with the results of [5] for ordinary riffle shuffles.

The next result gives a generating function for the joint distribution of permutations by number of left peaks and cycle type. This result appeared in a very recent paper of Gessel and Zhuang ([11], Theorem 7.2), though our proof is completely different, and we discovered it independently. It is an analog of a result in [4] which gave a generating function for the joint distribution of permutations by number of peaks and cycle type, and of a result of [7] giving a generating function for the joint distribution of permutations by number of descents and cycle type.

Corollary 5.2.

\[
\frac{t}{1-t} + \sum_{n \geq 1} u^n \sum_{\pi \in S_n} \frac{(1+t)^n}{(1-t)^{n+1}} \left( \frac{4t}{(1+t)^2} \right)^{\text{lpk}(\pi)} \prod_{i \geq 1} z_i^{N_i(\pi)}
\]

\[= \sum_{m \geq 1} \frac{t^m}{1 - z_1 u} \prod_{i \geq 1} \left( \frac{1 + z_i u^i}{1 - z_i u^i} \right)^{f_{i,m}}.\]

Proof. Take the coefficient of \(t^m\) in both sides of the statement of the corollary, and replace \(u\) by \(u/(2m + 1)\). By Propositions 3.3 and 3.6

\[x_m(\pi) = \frac{1}{(2m + 1)^n [t^m]^{(1+t)^n}} \left( \frac{4t}{(1+t)^2} \right)^{\text{lpk}(\pi)},\]

where \([t^m]f(t)\) denotes the coefficient of \(t^m\) in a power series \(f(t)\). The result follows from Theorem 5.1. \(\square\)

We mention that the results in this section are closely related to work of Reiner [14].

References

[1] Bayer, D. and Diaconis, P., Trailing the dovetail shuffle to its lair, *Ann. Appl. Probab.* 2 (1992), 294-313.
[2] Bergeron, F. and Bergeron, N., Orthogonal idempotents in the descent algebra of \(B_n\) and applications, *J. Pure Appl. Algebra* 79 (1992), 109-129.
[3] Diaconis, P., Five stories for Richard, in *The mathematical legacy of Richard P. Stanley*, 131-139, Amer. Math. Soc., Providence, RI, 2016.
[4] Diaconis, P., Fulman, J., and Holmes, S., Analysis of casino shelf shuffling machines, *Ann. Appl. Probab.* 23 (2013), 1692-1720.
[5] Diaconis, P., McGrath, M., and Pitman, J., Riffle shuffles, cycles, and descents, *Combinatorica* 15 (1995), 11-29.
[6] Fulman, J., Applications of the Brauer complex: card shuffling, permutation statistics, and dynamical systems, *J. Algebra* 243 (2001), 96-122.
[7] Fulman, J., The distribution of descents in fixed conjugacy classes of the symmetric groups, *J. Combin. Theory Ser. A* 84 (1998), 171-180.

[8] Gessel, I., Multipartite $P$-partitions and inner products of skew Schur functions, *Contemporary Mathematics* 34 (1984), 289–317.

[9] Gessel, I., A historical survey of $P$-partitions, in *The mathematical legacy of Richard P. Stanley*, 169-188, Amer. Math. Soc., Providence, RI, 2016.

[10] Gessel, I. and Reutenauer, C., Counting permutations with given cycle structure and descent set, *J. Combin. Theory Ser. A* 64 (1993), 189-215.

[11] Gessel, I. and Zhuang, Y., Plethystic formulas for permutation enumeration, [arXiv:2001.00654](https://arxiv.org/abs/2001.00654) (2020).

[12] Ma, S., Derivative polynomials and enumeration of permutations by number of interior and left peaks, *Discrete Math.* 312 (2012), 405-412.

[13] Petersen, T. K., Enriched $P$-partitions and peak algebras, *Adv. Math.* 209 (2007), 561-610.

[14] Reiner, V., Signed permutation statistics and cycle type, *European J. Combin.* 14 (1993), 569-579.

[15] Stanley, R., *Enumerative Combinatorics*, Volume 1. Second edition. Cambridge Studies in Advanced Mathematics, 49. *Cambridge University Press, Cambridge*, 2012.

[16] Stanley, R., *Enumerative Combinatorics*, Volume 2. Cambridge Studies in Advanced Mathematics, 62. *Cambridge University Press, Cambridge*, 1999.

[17] Stembridge, J., Enriched $P$-partitions, *Trans. Amer. Math. Soc.* 349 (1997), 763-768.

**Department of Mathematics, USC, Los Angeles, CA**  
*E-mail address: fulman@usc.edu*

**Department of Mathematical Sciences, DePaul University, Chicago, IL**  
*E-mail address: tpeter21@depaul.edu*