

Counting subwords in flattened partitions of sets

Toufik Mansour

Department of Mathematics, University of Haifa, 31905 Haifa, Israel
tmansour@univ.haifa.ac.il

Mark Shattuck

Department of Mathematics, University of Tennessee, Knoxville, TN 37996
shattuck@math.utk.edu

Stephan Wagner

Department of Mathematical Sciences, Stellenbosch University, Matieland 7602, South Africa
swagner@sun.ac.za

Abstract

In this paper, we consider the problem of avoidance of subword patterns in flattened partitions, which extends recent work of Callan. We determine in all cases explicit formulas and/or generating functions for the number of set partitions of size n which avoid a single subword pattern of length three. The asymptotic behavior of the resulting counting sequences turns out to depend quite heavily on the specific pattern. For the cases of 312 and 213, we make use of the kernel method to determine the generating function which counts the members of the avoidance class. Furthermore, in the cases of 132, 231, and 123, we also find formulas concerning the distribution on P_n for the statistics recording the number of occurrences of the pattern in question and some related bijective proofs are given. Finally, in each of these cases, it is shown that the number of occurrences of the pattern asymptotically follows a normal distribution.

Keywords: flattened set partitions, pattern avoidance, subword patterns, kernel method, asymptotic enumeration

2010 *Mathematics Subject Classification:* 05A18, 05A15, 05A16, 05A19

1 Introduction

Let $\pi = \pi_1\pi_2\cdots\pi_n$ and $\sigma = \sigma_1\sigma_2\cdots\sigma_d$ be permutations of length n and d , where $n \geq d$. The permutation π is said to contain σ as a *subword* if there exists a set of consecutive letters $\pi_i\pi_{i+1}\cdots\pi_{i+d-1}$ in π that is order-isomorphic to σ . Otherwise, π is said to *avoid* σ or be σ -free. In this context, σ is usually called a (subword) *pattern*. For example, $\pi = 13476528 \in \mathcal{S}_8$ contains two occurrences of the pattern 321 (corresponding to 765 and 652; note that occurrences of a given pattern need not be disjoint), but avoids the pattern 312. Subwords of the form $\pi_i\pi_{i+1}$, where $1 \leq i \leq n-1$ and $\pi_i < \pi_{i+1}$ (resp., $\pi_i > \pi_{i+1}$), are called *ascents* (resp., *descents*).

The problem of counting permutations according to the number of occurrences of a given subword has been studied from various perspectives in both enumerative and algebraic combinatorics (see, e.g., [8]). The comparable problem has also been considered on other discrete structures such as k -ary words [2], compositions [12], and set partitions [11] (see also [6] and the references contained therein).

In his study of finite set partitions, Callan [3] introduced the notion of a flattened partition and considered the problem of avoiding a single classical pattern of length three in this sense. Here,

we extend this work to the avoidance of subword patterns by flattened partitions, and in three cases, are able further to ascertain formulas which count the partitions of size n according to the number of occurrences of a pattern. See also [9] and [10] for recent related work on flattened permutations.

Let $[n] = \{1, 2, \dots, n\}$ if $n \geq 1$, with $[0] = \emptyset$. By a *partition* of $[n]$, we will mean a collection of pairwise disjoint subsets, called *blocks*, whose union is $[n]$. We will denote the set of all partitions of $[n]$ by P_n and the subset of P_n whose members contain exactly m blocks by $P_{n,m}$. Let $B_n = |P_n|$ denote the n -th Bell number and $S_{n,m} = |P_{n,m}|$ denote the Stirling number of the second kind. Here, we will find q -generalizations of B_n by counting the members of P_n according to various statistics.

Let us now recall the definition of *flattened partitions* introduced in [3]. Suppose that $\pi = B_1/B_2/\dots \in P_n$ is represented in *standard form*, i.e., elements within each block written in increasing order, with the blocks B_i arranged from left to right in increasing order of size of their smallest elements. Let $\text{Flatten}(\pi)$ be the permutation obtained by erasing the parentheses enclosing the blocks of π and considering the resulting word. For example, the partition $\pi = \{1, 3, 8\}, \{2, 5, 6\}, \{4\}, \{7, 9\} \in P_9$ is in standard form and $\text{Flatten}(\pi) = 138256479$.

One can combine the ideas in the previous paragraphs and say that a partition π *contains* the subword τ in the flattened sense if and only if $\text{Flatten}(\pi)$ contains τ (as a subword) in the usual sense and *avoids* τ otherwise. Using this definition of subword containment, we consider here the case when τ has length three. For example, the partition π considered in the preceding paragraph contains three occurrences of 123 in the flattened sense but avoids 132 as $\text{Flatten}(\pi) = 138256479$ contains three occurrences of 123 but avoids 132.

In this paper, we find explicit formulas and/or generating functions for the number of members of P_n avoiding any subword of length three. For the cases 132, 231, and 123, we also develop formulas for the distribution on P_n corresponding to the statistic recording the number of occurrences of the pattern in question. As a result, we obtain q -generalizations of the Bell numbers in these cases. In particular, we find q -analogues of the Bell number recurrence and exponential generating function (egf). From the egf, explicit formulas can then be obtained for the total number of occurrences of a pattern by differentiation. Furthermore, in the case 132, an explicit formula for the distribution on P_n (taken jointly with the number of blocks) is derived from the egf and a combinatorial proof is given. Finally, it is shown that the number of occurrences of each of these three patterns asymptotically follows a normal distribution.

In the third section, we deal with the patterns 312 and 213, which are apparently more difficult, and treat only the avoidance. We consider here the ordinary generating function (ogf), instead of the egf, to derive our results and use the *kernel method* (see [7] and the references contained therein) to solve systems of functional equations that are satisfied by related ogf's. To determine the number of members of P_n which avoid 312 or 213 in the flattened sense, we refine part of the set in question by considering two suitable statistics on the set and then finding recurrences satisfied by the refined numbers which arise. This method is somewhat reminiscent of a strategy of Zeilberger (see [15]) that he used to determine an unknown sequence. In the case of 312, we find that the avoidance class has the same cardinality as a certain class of Motzkin type paths of length $n - 1$. For the case 213, though no explicit formula is determined, we find that the corresponding ogf is the solution to a relatively simple functional equation in a single variable.

Section 4 summarizes and compares the results obtained for the different patterns and concludes the paper.

2 Counting subwords in flattened partitions

In this section, we consider the problem of counting the members of P_n according to the number of occurrences of a subword of length three. We need not consider the case of counting occurrences of 321.

Proposition 2.1. *All members of P_n avoid the subword 321 in the flattened sense.*

Proof. Suppose $\pi \in P_n$ and $\text{Flatten}(\pi) = \pi_1\pi_2 \cdots \pi_n$. If $\pi_i > \pi_{i+1}$, then the $(i+1)$ -st letter from the left must start a block of π , by the ordering of elements within blocks. But then either $i = n-1$ or $\pi_{i+2} > \pi_{i+1}$, again by the ordering of elements. \square

In the next three subsections, we count the members of P_n according to the number of occurrences of the patterns 132, 231, and 123, deriving explicit formulas for the egf's in each case.

2.1 The case 132

Let $a_n(y, q)$ denote the joint distribution on P_n for the statistics recording the number of blocks and the number of occurrences of 132. The polynomials $a_n(y, q)$ satisfy the following recurrence relation.

Lemma 2.2. *If $n \geq 1$, then*

$$a_n(y, q) = ya_{n-1}(y, q) + y \sum_{i=0}^{n-2} \left(\binom{n-1}{i} - (1-q)i \right) a_i(y, q), \quad (1)$$

with $a_0(y, q) = 1$.

Proof. Let $n \geq 2$ and let us abbreviate $a_n(1, q)$ by a_n . For $1 \leq i \leq n$, let $a_{n,i} = a_{n,i}(q)$ be the distribution for the number of occurrences of 132 on the members of P_n whose first block B has size i . Clearly, $a_{n,1} = a_{n-1}$. Let us find an expression for $a_{n,i}$ when $i \geq 2$. If $B = \{1, 2, \dots, i\}$, then the total weight of all such members of P_n is a_{n-i} . Suppose now B is of the form $B = \{1, 2, \dots, t, a, b, \dots\}$, where $1 \leq t \leq i-2$ and $a, b > t+1$. Then the second block of the partition starts with $t+1$ and there is no 132 created in going from the first to the second block since the last two elements of the first block are greater than $t+1$. This implies a contribution of

$$\sum_{t=1}^{i-2} \binom{n-t-1}{i-t} a_{n-i} = a_{n-i} \sum_{t=1}^{i-2} \binom{n-t-1}{n-i-1} = a_{n-i} \left(\binom{n-1}{n-i} - 1 - (n-i) \right),$$

where we have used the identity $\sum_{r=m}^n \binom{r}{m} = \binom{n+1}{m+1}$. Finally, if $B = \{1, 2, \dots, i-1, a\}$, where $a > i$, then the letters $i-1, a, i$ form an occurrence of 132 in the flattened sense when going from the first block to the second, which implies a contribution of $q(n-i)a_{n-i}$ in this case. Combining the three cases then gives

$$\begin{aligned} a_{n,i} &= a_{n-i} + \left(\binom{n-1}{n-i} - 1 - (n-i) \right) a_{n-i} + q(n-i)a_{n-i} \\ &= \left(\binom{n-1}{n-i} - (1-q)(n-i) \right) a_{n-i}, \quad 2 \leq i \leq n. \end{aligned}$$

Thus, we have

$$\begin{aligned} a_n &= \sum_{i=1}^n a_{n,i} = a_{n-1} + \sum_{i=2}^n \left(\binom{n-1}{n-i} - (1-q)(n-i) \right) a_{n-i} \\ &= a_{n-1} + \sum_{i=0}^{n-2} \left(\binom{n-1}{i} - (1-q)i \right) a_i. \end{aligned}$$

Adding a variable that marks the number of blocks in the preceding argument then gives (1). \square

Lemma 2.3. *We have*

$$\sum_{n \geq 0} a_n(y, q) \frac{x^n}{n!} = 1 + ye^{-y} \int_0^x e^{ye^t - (1-q)yt^2/2 + t} dt. \quad (2)$$

Proof. To get rid of the $(1-q)ia_i(y, q)$ part of the sum on the right-hand side of (1), which is inconvenient when taking exponential generating functions, we consider the difference $a_n(y, q) - a_{n-1}(y, q)$. By (1), we have

$$\begin{aligned} a_n(y, q) - a_{n-1}(y, q) &= y(a_{n-1}(y, q) - a_{n-2}(y, q)) \\ &+ y \sum_{i=0}^{n-2} \left(\binom{n-1}{i} - (1-q)i \right) a_i(y, q) - y \sum_{i=0}^{n-3} \left(\binom{n-2}{i} - (1-q)i \right) a_i(y, q), \end{aligned}$$

which is equivalent to

$$\begin{aligned} a_n(y, q) &= a_{n-1}(y, q) - y(1-q)(n-2)a_{n-2}(y, q) \\ &+ y \sum_{i=0}^{n-1} \binom{n-1}{i} a_i(y, q) - y \sum_{i=0}^{n-2} \binom{n-2}{i} a_i(y, q), \quad n \geq 2, \end{aligned} \quad (3)$$

with $a_0(y, q) = 1$ and $a_1(y, q) = y$. Let $A(x; y, q) = \sum_{n \geq 0} a_n(y, q) \frac{x^n}{n!}$ denote the egf for the sequence $a_n(y, q)$. Multiplying both sides of (3) by $x^{n-1}/(n-1)!$, and summing over $n \geq 2$, yields

$$\begin{aligned} \frac{d}{dx} A(x; y, q) - y &= A(x; y, q) - 1 - y(1-q) \left(xA(x; y, q) - \int_0^x A(t; y, q) dt \right) \\ &+ y \left(e^x A(x; y, q) - 1 - \int_0^x e^t A(t; y, q) dt \right). \end{aligned}$$

Differentiating both sides of this last equation with respect to x gives

$$\frac{d^2}{dx^2} A(x; y, q) = (1 - (1-q)xy + e^x y) \frac{d}{dx} A(x; y, q).$$

Thus we have $\ln\left(\frac{d}{dx} A(x; y, q)\right) = x - (1-q)x^2y/2 + e^x y + c$ for some constant c . Since $\frac{d}{dx} A(x; y, q) \big|_{x=0} = y$, we have $c = \ln y - y$. Therefore, $\frac{d}{dx} A(x; y, q) = ye^{y(e^x - 1) - (1-q)x^2y/2 + x}$. Using the initial condition $A(0; y, q) = 1$, we obtain (2). \square

Remark 2.4. *Integrating by parts the formula in (2) gives*

$$A(x; y, q) = e^{y(e^x - 1) - (1-q)x^2y/2} + (1-q)ye^{-y} \int_0^x te^{ye^t - (1-q)t^2y/2} dt. \quad (4)$$

Taking $q = 1$ in (4) recovers the egf for the Bell polynomial $B_n(y) = \sum_{k=0}^n S_{n,k} y^k$ (see (7.54) in [5, p. 351]).

If j is odd, then let $j!! = j(j-2) \cdots 1$, with $(-1)!! = 1$. We have the following explicit formula for the polynomial $a_n(y, q)$.

Theorem 2.5. *If $n \geq 1$, then*

$$a_n(y, q) = \sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} ((q-1)y)^i \binom{n-1}{2i} (2i-1)!! B_{n-2i}(y). \quad (5)$$

Proof. By (2), we have that $a_n(y, q)$ is the coefficient of $\frac{x^{n-1}}{(n-1)!}$ in the convolution

$$ye^{y(e^x-1)+x} \cdot e^{(q-1)x^2y/2} = \sum_{i \geq 0} B_{i+1}(y) \frac{x^i}{i!} \cdot \sum_{j \geq 0} \left(\frac{(q-1)y}{2} \right)^j (2j)^i \frac{x^{2j}}{(2j)!},$$

which is given by

$$\sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} \left(\frac{(q-1)y}{2} \right)^i \binom{n-1}{2i} (2i)^i B_{n-2i}(y).$$

Formula (5) now follows from noting that $(2i)^i = (2i)!/i! = (2i)!!/i! \cdot (2i-1)!! = 2^i(2i-1)!!$. \square

Remark 2.6. Taking $y = 1$ and $q = 0$ in (5) gives a formula for the number of members of P_n which avoid the subword 132 in the flattened sense. Furthermore, we see from (5) that the number of members of $P_{n,m}$ which avoid 132 is given by

$$[y^m]a_n(y, 0) = \sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} (-1)^i \binom{n-1}{2i} (2i-1)!! S_{n-2i, m-i}, \quad n \geq 1.$$

Differentiating both sides of (5) with respect to q , and letting $q = 1$, gives the following result.

Corollary 2.7. If $n \geq 1$ and $1 \leq m \leq n$, then the total number of occurrences of the subword 132 within all the members of $P_{n,m}$ and P_n , respectively, is given by $\binom{n-1}{2} S_{n-2, m-1}$ and $\binom{n-1}{2} B_{n-2}$.

It is instructive to provide a bijective proof of this result.

Combinatorial proof of Corollary 2.7.

We prove only the second statement. A similar argument applies to the first upon fixing the number of blocks. We will refer to an occurrence of the subword 132 in a flattened partition such that the 2 and the 3 correspond to the actual letters i and j as an (i, j) -occurrence of 132. Given $2 \leq i < j \leq n$, it suffices to show that the number of (i, j) -occurrences of 132 within all of the members of P_n is B_{n-2} . That is, we need to show that there are B_{n-2} members of P_n having an (i, j) -occurrence of 132. In order for $\pi \in P_n$ to contain such an occurrence, the element i must start a block B which directly follows a block C whose largest element is j . Note further that there is at least one other element in C , with all elements of C other than j less than i . Let π' be the partition obtained by removing the elements i and j from π and merging the blocks $B - \{i\}$ and $C - \{j\}$ into a single block D .

Note that the mapping $\pi \mapsto \pi'$ defines a bijection between the members of P_n containing an (i, j) -occurrence of 132 and partitions of $[n] - \{i, j\}$. To reverse it, given a partition ρ of the set $[n] - \{i, j\}$, identify the block of ρ whose smallest element is largest among all of the blocks of ρ whose smallest element is less than i . (Note that there is at least one such block since $i > 1$.) We then split this block into two parts, one containing all of the elements in the block that are less than i and another containing any elements greater than i , where the second part is possibly empty. We then add j to the end of the first part and i to the beginning of the second to obtain a partition of $[n]$ containing an (i, j) -occurrence of 132. (Note that if the second part is empty, then we just add the singleton block $\{i\}$.) Thus, the number of members of P_n having an (i, j) -occurrence of 132 is B_{n-2} for each i and j , as desired. \square

It is possible to extend the argument to prove (5).

Combinatorial proof of Theorem 2.5.

We first prove the $q = 0$ case of (5) and show that the members of P_n (weighted by the number of blocks) which avoid 132 in the flattened sense are enumerated by

$$\sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} (-y)^i \binom{n-1}{2i} (2i-1)!! B_{n-2i}(y).$$

To do so, let S be a subset of $\{2, 3, \dots, n\}$ of size $2i$. We partition S into i doubleton blocks $\{s_j, t_j\}$, where $s_j < t_j$ and $s_1 < s_2 < \dots < s_i$ (i.e., we form a *perfect matching* of the elements of S). Note that there are $\binom{n-1}{2i}$ choices for S and $(2i-1)!!$ ways to partition it into doubletons. Next, we form a partition π of the elements of $[n] - S$ in one of B_{n-2i} ways; the polynomial $B_{n-2i}(y)$ also takes the number of blocks into account. Now we add the elements of S to π in a systematic way. Just as we inserted two elements into the partition ρ in the second part of the preceding proof, we insert the elements s_1 and t_1 into π to obtain a partition π_1 of size $n - 2i + 2$ containing an (s_1, t_1) -occurrence of 132. In general, we insert the elements s_{j+1} and t_{j+1} , inductively, into the partition π_j to obtain π_{j+1} in the same manner for each $1 \leq j < i$. This results in a partition $\pi' = \pi_i$ of $[n]$ such that there is an (s_j, t_j) -occurrence of 132 in π' for each $1 \leq j \leq i$. Observe that the mapping $\pi \mapsto \pi'$ is a bijection, upon undoing each step sequentially starting with the last, which is accomplished by merging blocks as described in the first paragraph of the preceding proof. Note that i additional blocks are created in the transition from π to π' since a new block is created at each step. Thus, $y^i B_{n-2i}(y)$ counts all partitions of $[n]$ in which the elements within each block of the matching of S correspond to an occurrence of 132 (along with possibly some elements of $[n] - S$, depending on the choice of the partition π). By an application of the inclusion-exclusion principle, the number of members of P_n that avoid 132 is as asserted above.

For the case of general q , we make use of a characteristic function argument as follows. Suppose $\rho \in P_n$ and let T denote the set of all pairs of elements (s, t) , $s < t$, such that ρ contains an (s, t) -occurrence of 132. Note that the exponent of the y -factor within all of the terms of (5) that count ρ is the same and is given by the number of blocks of ρ . To complete the proof, we then must show that the sum of the factors of q in these terms is q^i , where i denotes the number of pairs of T . Suppose R is a subset of the pairs in T of size j . Then any partition σ , in particular ρ , in which the elements within each pair of R form an occurrence of 132 in σ is counted by a factor of $(q-1)^j$, by the argument for the $q = 0$ case above. Since the pairs comprising the complete set of 132-occurrences of ρ is precisely T , it follows that ρ is counted

$$\sum_{R \subseteq T} (q-1)^{|R|} = \sum_{j=0}^i \binom{i}{j} (q-1)^j = q^i$$

times by the sum in (5), as desired, which completes the proof. □

The combinatorial argument given in the two preceding proofs also proves the following theorem, which can also be obtained from the generating function.

Theorem 2.8. *The number of occurrences of 132 is equidistributed with the number of doubletons (blocks of cardinality 2) not containing 1. Even the joint distributions with the number of blocks are identical.*

The rest of this subsection is devoted to an asymptotic analysis. Let us first consider the number of 132-free set partitions. We make use of the following general theorem on Hayman-admissible functions (in a version suitable for our purposes).

Theorem 2.9 (cf. [4, Theorem VIII.4]). *Let $G(z) = e^{h(z)}$ be an entire function that is positive for positive real values of z . Suppose that the following conditions are satisfied:*

- $a(r) = rh'(r) \rightarrow \infty$ and $b(r) = r^2h''(r) + rh'(r) \rightarrow \infty$ as $r \rightarrow \infty$.
- For some function $\theta_0(r)$ defined on $(0, \infty)$ with $\theta_0(r) \in (0, \pi)$ for all (sufficiently large) r , one has

$$G(re^{i\theta}) \sim G(r)e^{i\theta a(r) - \theta^2 b(r)/2}$$

uniformly in $|\theta| \leq \theta_0(r)$ as $r \rightarrow \infty$.

- Uniformly in $\theta_0(r) \leq |\theta| < \pi$,

$$G(re^{i\theta}) = o\left(\frac{G(r)}{\sqrt{b(r)}}\right).$$

If $r = r(n)$ is such that $a(r) - n = o(\theta_0(r))$ (in particular, if $rh'(r) = a(r) = n$), then the coefficients of $G(z)$ satisfy, as $n \rightarrow \infty$:

$$g_n = [z^n]G(z) \sim \frac{G(r)}{r^n \sqrt{2\pi b(r)}}.$$

Write f_n for the number of 132-free set partitions of P_n . Its generating function is obtained by taking $y = 1$ and $q = 0$ in (4). To simplify things, we also take the derivative with respect to x , which merely corresponds to a coefficient shift:

$$F(x) = \sum_{n \geq 0} f_{n+1} \frac{x^n}{n!} = \sum_{n \geq 0} a_{n+1}(1, 0) \frac{x^n}{n!} = e^{e^x + x - x^2/2 - 1}.$$

This generating function is easily seen to be Hayman-admissible in the sense of Theorem 2.9 (with $h(z) = e^z - z^2/2 + z - 1$, $\theta_0(r)$ can be taken as $e^{-2r/5}$). As for the Bell numbers B_n , an asymptotic expansion is only possible for the logarithm, the asymptotic formula for the number itself involves an implicitly defined quantity.

Theorem 2.10. *The number f_n of 132-free set partitions in P_n is given by the asymptotic formula*

$$f_n \sim n! \cdot \frac{e^{e^r - r^2/2 - 1}}{r^n \sqrt{2\pi r(r+1)} e^r},$$

where r is the unique positive real solution of the equation

$$re^r = n.$$

The proportion of 132-free set partitions in P_n is

$$\frac{f_n}{B_n} \sim e^{-r^2/2}.$$

Proof. Since the generating function satisfies the conditions, Theorem 2.9 yields

$$f_{n+1} = a_{n+1}(1, 0) \sim n! \cdot \frac{e^{e^r + r - r^2/2 - 1}}{r^n \sqrt{2\pi r(r+1)} e^r},$$

where r is the unique positive real solution of the equation

$$re^r = n + 1.$$

Note here that $rh'(r) + r^2h''(r) = r(r+1)e^r - 2r^2 + r \sim r(r+1)e^r$. Since $re^r = n + 1$, we obtain

$$f_{n+1} \sim (n+1)! \cdot \frac{e^{e^r - r^2/2 - 1}}{r^{n+1} \sqrt{2\pi r(r+1)} e^r},$$

from which the first statement of the theorem follows (replacing n by $n - 1$). Since

$$B_n \sim n! \cdot \frac{e^{e^r - 1}}{r^n \sqrt{2\pi r(r+1)} e^r}$$

by the same reasoning (cf. [4, Proposition VIII.3]), the second statement follows immediately. \square

We know from Corollary 2.7 that the average number of occurrences of 132 in a randomly chosen partition of $[n]$ is $\binom{n-1}{2} B_{n-2} / B_n \sim (\log n)^2 / 2$. The distribution satisfies a central limit theorem that follows directly from the quasi-powers theorem, see [4, Theorem IX.13].

Theorem 2.11. *Suppose that the probability generating functions $p_n(u)$ of a sequence of random variables X_n whose values are non-negative integers satisfy an asymptotic formula of the form*

$$p_n(u) = e^{h_n(u)}(1 + o(1)),$$

uniformly with respect to u in some interval around 1, where each $h_n(u)$ is analytic in this interval. If the additional conditions

$$h'_n(1) + h''_n(1) \rightarrow \infty \quad \text{and} \quad \frac{h'''_n(u)}{(h'_n(1) + h''_n(1))^{3/2}} \rightarrow 0$$

are satisfied uniformly in u as $n \rightarrow \infty$, then the normalized random variable

$$Y_n = \frac{X_n - h'_n(1)}{(h'_n(1) + h''_n(1))^{1/2}}$$

converges in distribution to a standard Gaussian distribution.

In our specific case, we obtain the following result.

Theorem 2.12. *The number of occurrences of 132 in a randomly chosen set partition of $[n]$ asymptotically follows a normal distribution with mean $\mu_n \sim (\log n)^2 / 2$ and variance $\sigma_n^2 \sim (\log n)^2 / 2$.*

Proof. We use the bivariate generating function (obtained from (4) by taking the derivative with respect to x and setting $y = 1$):

$$\sum_{n \geq 0} a_{n+1}(1, q) \frac{x^n}{n!} = e^{e^x + x + (q-1)x^2/2 - 1}.$$

The conditions of Theorem 2.9 are still satisfied, and we obtain, as before,

$$a_n(1, q) \sim n! \cdot \frac{e^{e^r + (q-1)r^2/2 - 1}}{r^n \sqrt{2\pi r(r+1)} e^r},$$

where $r = r(n)$ is defined by $re^r = n$. It follows that the probability generating function satisfies

$$p_n(q) = \frac{a_n(1, q)}{B_n} \sim e^{(q-1)r^2/2},$$

uniformly in q (in any compact subinterval of $(0, \infty)$). Now we can invoke Theorem 2.11 with $h_n(q) = (q-1)r^2/2$ to complete the proof, and the mean and variance are asymptotically given by

$$\mu_n \sim h'_n(1) = \frac{r^2}{2} \sim \frac{(\log n)^2}{2}$$

and

$$\sigma_n^2 \sim h'_n(1) + h''_n(1) = \frac{r^2}{2} \sim \frac{(\log n)^2}{2}.$$

\square

2.2 The case 231

Let $b_n(y, q)$ denote the joint distribution on P_n for the statistics recording the number of blocks and the number of occurrences of 231. Modifying the proof above for Lemma 2.2 gives the following recurrence for $b_n(y, q)$.

Lemma 2.13. *If $n \geq 1$, then*

$$b_n(y, q) = yb_{n-1}(y, q) + y \sum_{i=1}^{n-1} \left(q \binom{n-1}{i-1} + (1-q)i \right) b_{i-1}(y, q), \quad (6)$$

with $b_0(y, q) = 1$.

Let u_n denote sequence A005425 in [14] defined by the recurrence

$$u_n = 2u_{n-1} + (n-1)u_{n-2}, \quad n \geq 2,$$

with initial values $u_0 = 1$ and $u_1 = 2$. Among other things, it counts the number of partitions of $[n]$ into blocks of size 1 or 2, where singleton blocks come in two colors. Considering the difference $b_n(1, 0) - b_{n-1}(1, 0)$ for $n \geq 3$ in (6) gives the following result.

Corollary 2.14. *If $n \geq 1$, then the number of members of P_n which avoid the subword 231 in the flattened sense is given by u_{n-1} .*

The asymptotic behavior of u_n is obtained by means of the saddle point method. We can apply Theorem 2.9 once again, starting from the generating function

$$\sum_{n \geq 0} \frac{u_n}{n!} x^n = e^{x^2/2+2x}.$$

This generating function satisfies the conditions with $\theta_0(r) = r^{-3/4}$ and $r = r(n) = \sqrt{n} - 1$. We obtain $b(r) = 2r(r+1) \sim 2n$ and

$$u_n \sim n! \cdot \frac{e^{r^2/2+2r}}{r^n \sqrt{4\pi n}} \sim \frac{1}{\sqrt{2}} \cdot n^{n/2} e^{2\sqrt{n}-n/2-1},$$

so that we have the following theorem (after replacing n by $n-1$).

Theorem 2.15. *The number of members of P_n which avoid the subword 231 in the flattened sense is asymptotically equal to*

$$\frac{1}{\sqrt{2n}} \cdot n^{n/2} e^{2\sqrt{n}-n/2-1}.$$

Let us now return to counting the number of occurrences of 231. We have the following explicit formula for the egf of the sequence $b_n(y, q)$.

Theorem 2.16. *We have*

$$\sum_{n \geq 0} b_n(y, q) \frac{x^n}{n!} = 1 + ye^{-qy} \int_0^x e^{qye^t + (1-q)yt^2/2 + (1+y(1-q))t} dt. \quad (7)$$

Proof. Proceeding as in the proof of (2), we replace n by $n-1$ in (6) and subtract to get

$$\begin{aligned} b_n(y, q) &= (1 + y(1-q))b_{n-1}(y, q) + y(1-q)(n-2)b_{n-2}(y, q) + qy \sum_{i=0}^{n-1} \binom{n-1}{i} b_i(y, q) \\ &\quad - qy \sum_{i=0}^{n-2} \binom{n-2}{i} b_i(y, q), \quad n \geq 2, \end{aligned}$$

with $b_0(y, q) = 1$ and $b_1(y, q) = y$. Let $B(x; y, q) = \sum_{n \geq 0} b_n(y, q) \frac{x^n}{n!}$. Multiplying this last recurrence by $\frac{x^{n-1}}{(n-1)!}$, summing over $n \geq 2$, and differentiating with respect to x gives the relation

$$\frac{d^2}{dx^2} B(x; y, q) = (1 + (1 - q)y + (1 - q)xy + qye^x) \frac{d}{dx} B(x; y, q).$$

Since $\frac{d}{dx} B(x; y, q) |_{x=0} = y$, we get $\frac{d}{dx} B(x; y, q) = ye^{qye^x + (1-q)x^2y/2 + (1+y(1-q))x - qy}$. Formula (7) now follows from noting the initial condition $B(0; y, q) = 1$. \square

Corollary 2.17. *If $n \geq 1$ and $1 \leq m \leq n$, then the total number of occurrences of the subword 231 within all the members of $P_{n,m}$ and P_n , respectively, is given by*

$$S_{n+1,m} - S_{n,m} - S_{n,m-1} - (n-1)S_{n-1,m-1} - \binom{n-1}{2} S_{n-2,m-1}$$

and

$$B_{n+1} - 2B_n - (n-1)B_{n-1} - \binom{n-1}{2} B_{n-2}.$$

Proof. Differentiating the formula in (7) with respect to q gives

$$\begin{aligned} \frac{d}{dq} B(x; y, q) |_{q=1} &= y^2 \int_0^x \left(e^t - \frac{t^2}{2} - t - 1 \right) e^{y(e^t-1)+t} dt \\ &= y \left[e^{y(e^t-1)+t} \right]_0^x - y \int_0^x e^{y(e^t-1)+t} dt \\ &\quad - y \int_0^x \sum_{n \geq 0} \left(B_{n+1}(y) \left(\frac{t^{n+2}}{2n!} + \frac{t^{n+1}}{n!} + \frac{t^n}{n!} \right) \right) dt \\ &= 1 - y + ye^{y(e^x-1)+x} - e^{y(e^x-1)} \\ &\quad - y \sum_{n \geq 1} \left(\binom{n-1}{2} B_{n-2}(y) + (n-1)B_{n-1}(y) + B_n(y) \right) \frac{x^n}{n!} \\ &= \sum_{n \geq 1} \left(B_{n+1}(y) - (1+y)B_n(y) - (n-1)yB_{n-1}(y) - \binom{n-1}{2} yB_{n-2}(y) \right) \frac{x^n}{n!}. \end{aligned}$$

Extracting the coefficient of $\frac{x^n}{n!} y^m$ now gives the result (for the total number of occurrences in P_n , we set $y = 1$ and extract the coefficient of $\frac{x^n}{n!}$ afterwards, or we sum over all m). \square

It is possible to provide a bijective proof of Corollary 2.17 in the same way as for Corollary 2.7.

Combinatorial proof of Corollary 2.17.

Suppose $\pi \in P_n$ and $\text{Flatten}(\pi) = \pi_1 \pi_2 \cdots \pi_n$. By a *block ascent* (or *descent*), we will mean an index $i \in [n-1]$ such that $\pi_i < \pi_{i+1}$ (or, $\pi_i > \pi_{i+1}$, respectively) with π_{i+1} corresponding to the smallest element of some block of π .

We prove the second statement, the first following in a similar manner upon fixing the number of blocks in a partition. We first argue that the total number of block ascents within all of the members of P_n is given by $(n-1)B_{n-1}$. To do so, we will show for each j that there are B_{n-1} members of P_n in which j corresponds to the second element of a block ascent for $2 \leq j \leq n$. Suppose σ is a partition of the set $[n] - \{j\}$. If one adds j to the block of σ whose smallest element is largest among all blocks of σ containing an element smaller than j and breaks the new block into two parts according to whether or not an element is greater than or equal j ,

then one obtains a member of P_n in which j corresponds to the second letter of a block ascent. Since this operation is seen to be reversible, it follows that there are B_{n-1} such members of P_n , as desired.

Next, observe that since each block of a partition but the first is involved in a block ascent or descent with its predecessor, we have

$$\text{tot}(\text{block descents}) + \text{tot}(\text{block ascents}) = \text{tot}(\text{blocks}) - (\# \text{ of members of } P_n),$$

where tot denotes the total number of the indicated item taken over all the members of P_n . Note that the total number of blocks in P_n is $B_{n+1} - B_n$. To see this, suppose $\pi \in P_{n+1}$, with the element $n+1$ not occurring as a singleton block. Removing $n+1$ from its block within π and then circling this block results in an arbitrary member of P_n in which one of the blocks is distinguished. We then get by subtraction that there are $B_{n+1} - 2B_n - (n-1)B_{n-1}$ block descents within all of the members of P_n , since the total number of block ascents is $(n-1)B_{n-1}$.

Finally, note that a block descent corresponds to either an occurrence of the 231 or 132 subword in the flattened sense. Thus, we have

$$\text{tot}(231) + \text{tot}(132) = \text{tot}(\text{block descents}) = B_{n+1} - 2B_n - (n-1)B_{n-1},$$

from which the result now follows by means of Corollary 2.7. \square

In analogy to Theorem 2.8, we have the following result.

Theorem 2.18. *The number of occurrences of 231 is equidistributed with the number of blocks of size greater than 2 not containing 1. Even the joint distributions with the number of blocks are identical.*

Moreover, the following central limit theorem holds in analogy to Theorem 2.12.

Theorem 2.19. *The number of occurrences of 231 in a randomly chosen set partition of $[n]$ asymptotically follows a normal distribution with mean $\mu_n \sim n/\log n$ and variance $\sigma_n^2 \sim n/(\log n)^2$.*

Proof. The generating function

$$\sum_{n \geq 0} b_{n+1}(1, q) \frac{x^n}{n!} = e^{qe^x - (q-1)x^2/2 - (q-2)x - 1}$$

satisfies the conditions of Theorem 2.9 with $\theta_0(r) = e^{-2r/5}$ and $r = r(n)$ defined by $qre^r = n+1$, and this is uniform in q (restricted to any compact subinterval of $(0, \infty)$). One obtains

$$b_{n+1}(1, q) \sim n! \cdot \frac{e^{qe^r - (q-1)r^2/2 - (q-2)r - 1}}{r^n \sqrt{2\pi qr(r+1)} e^r} = (n+1)! \cdot \frac{e^{qe^r - (q-1)r^2/2 - (q-1)r - 1}}{qr^{n+1} \sqrt{2\pi(n+1)(r+1)}},$$

and replacing n by $n-1$ we get

$$b_n(1, q) \sim n! \cdot \frac{e^{qe^r - (q-1)r^2/2 - (q-1)r - 1}}{qr^n \sqrt{2\pi n(r+1)}},$$

where r is implicitly defined by $qre^r = n$. This makes it possible to apply Theorem 2.11 again. Implicit differentiation can be used to obtain $r'(1)$ and higher-order derivatives, and the mean and variance are found to be

$$\mu_n \sim \left. \frac{n}{r} \right|_{q=1} \sim \frac{n}{\log n}$$

and

$$\sigma_n^2 \sim \left. \frac{n}{r(1+r)} \right|_{q=1} \sim \frac{n}{(\log n)^2}.$$

\square

2.3 The case 123

Let $d_n(q)$ denote the distribution on P_n for the number of occurrences of the subword 123.

Lemma 2.20. *If $n \geq 2$, then*

$$d_n(q) = 2q^{n-2} + (n-2)q^{n-3} + \sum_{i=1}^{n-2} q^{i-2} \left(q^2 - 1 + \binom{n-1}{i-1} \right) d_{n-i}(q), \quad (8)$$

with $d_0(q) = d_1(q) = 1$.

Proof. Abbreviate $d_n(q)$ by d_n . If $1 \leq i \leq n$, then let $d_{n,i} = d_{n,i}(q)$ be the distribution for the number of occurrences of 123 on the members of P_n whose first block has size i . Note first that for $n \geq 2$, we have $a_{n,n} = q^{n-2}$ and $a_{n,n-1} = q^{n-2} + (n-2)q^{n-3}$. Furthermore, if $n \geq 3$ and $1 \leq i \leq n-2$, then

$$d_{n,i} = q^i d_{n-i} + q^{i-2} \left(\binom{n-1}{i-1} - 1 \right) d_{n-i}. \quad (9)$$

To show (9), note that the first term on the right-hand side gives the contribution of those $\pi \in P_n$ whose first block B is $\{1, 2, \dots, i\}$. Since the letters directly following B in $\text{Flatten}(\pi)$ are $i+1, a$ for some $a > i+1$, we see that each letter in B starts an occurrence of a 123 subword, whence the factor of q^i . If $i \geq 2$ and B is any other i -element set containing 1, then only the first $i-2$ letters of B can start an occurrence of 123 since in this case there would be a descent in $\text{Flatten}(\pi)$ at index i . Thus, the weight in this case will be $q^{i-2} \left(\binom{n-1}{i-1} - 1 \right) d_{n-i}$, which gives (9). Recurrence (8) now follows from (9) and the fact $d_n = \sum_{i=1}^n d_{n,i}$. \square

We see from (8) that $d_n(0) = (n-2)d_{n-2}(0)$ if $n \geq 4$. Since $d_2(0) = 2$ and $d_3(0) = 1$, we get the following result.

Corollary 2.21. *If $n \geq 1$, then the number of members of P_{2n+1} which avoid the subword 123 in the flattened sense is given by $(2n-1)!!$ and the number of such members of P_{2n} is given by $2(2n-2)!!$.*

For a combinatorial explanation of this result, first note that if $\pi \in P_{2n+1}$ avoids 123 in the flattened sense, then all the blocks of π must be doubletons, with the exception of the final block, which is a singleton. If $1 \leq i \leq n$, then there are $2n+1-2i$ choices for the larger element in the i -th doubleton since it cannot be any one of the $2i-1$ elements of $[2n+1]$ that have already been used nor can it be the smallest element yet to be used (the latter requirement in order to avoid 123). Thus, there are $\prod_{i=1}^n (2n+1-2i) = (2n-1)!!$ choices regarding the elements comprising the doubleton blocks of π , which determines the element lying in the final block. If $\pi \in P_{2n}$, by the same reasoning, there are $\prod_{i=1}^{n-1} (2n-2i) = (2n-2)!!$ choices for the first $n-1$ blocks of π , which must all be doubletons, with the remaining two elements either belonging to the same block or separate blocks. \square

Let $D(x; q) = \sum_{n \geq 0} d_n(q) \frac{x^n}{n!}$ denote the egf for the sequence $d_n(q)$.

Theorem 2.22. *We have*

$$D(x; q) = 1 + \int_0^x \left(e^{2qt-t/q-1/q^2+e^{qt}/q^2} + 2(1-q)e^{2qt-t/q+e^{qt}/q^2} \int_0^t e^{-2qr+r/q-e^{qr}/q^2} dr \right) dt. \quad (10)$$

Proof. First note that (8) may be rewritten as

$$d_n(q) = 2q^{n-2} + (n-2)q^{n-3} + (1-1/q^2) \sum_{i=2}^{n-1} q^{n-i} d_i(q) + \frac{1}{q^2} \sum_{i=2}^{n-1} q^{n-i} \binom{n-1}{i} d_i(q). \quad (11)$$

To eliminate the sum $\sum_{i=2}^{n-1} q^{n-i} d_i(q)$ in (11), we consider the difference $d_n(q) - qd_{n-1}(q)$, which gives us

$$d_n(q) - qd_{n-1}(q) = q^{n-3} + (q-1/q)d_{n-1}(q) + \frac{1}{q^2} \sum_{i=2}^{n-1} q^{n-i} \binom{n-1}{i} d_i(q) - \frac{1}{q^2} \sum_{i=2}^{n-2} q^{n-i} \binom{n-2}{i} d_i(q),$$

which may be rewritten as

$$d_n(q) = (2q - 1/q)d_{n-1}(q) + \frac{1}{q} \sum_{i=0}^{n-1} q^{n-1-i} \binom{n-1}{i} d_i(q) - \sum_{i=0}^{n-2} q^{n-2-i} \binom{n-2}{i} d_i(q), \quad n \geq 3,$$

with $d_0(q) = d_1(q) = 1$ and $d_2(q) = 2$. Multiplying this last recurrence by $\frac{x^{n-1}}{(n-1)!}$, summing over $n \geq 3$, and differentiating with respect to x yields the relation

$$\frac{d^2}{dx^2} D(x; q) = (2q - 1/q + e^{qx}/q) \frac{d}{dx} D(x; q) + 2(1 - q).$$

The preceding differential equation is first order linear in the quantity $\frac{d}{dx} D(x; q)$ satisfying the condition $\frac{d}{dx} D(x; q) |_{x=0} = 1$, which implies

$$\frac{d}{dx} D(x; q) = e^{2qx-x/q-1/q^2+e^{qx}/q^2} + 2(1-q)e^{2qx-x/q+e^{qx}/q^2} \int_0^x e^{-2qt+t/q-e^{qt}/q^2} dt. \quad (12)$$

Formula (10) now follows from noting the initial condition $D(0; q) = 1$. \square

Let B_n^* denote the complementary Bell number determined by $e^{1-e^x} = \sum_{n \geq 0} B_n^* \frac{x^n}{n!}$; see the paper by Rao Uppuluri and Carpenter [13] and entry A000587 in [14].

Corollary 2.23. *If $n \geq 1$, then the total number of occurrences of the subword 123 within all the members of P_n is given by*

$$2(n-1)B_{n-1} + (n+3)B_n - 2B_{n+1} + 2 \sum_{i=1}^{n-1} \sum_{j=0}^{i-1} (-1)^{i-j} \binom{n-1}{i} \binom{i-1}{j} B_{n-i} B_j^*.$$

Proof. Differentiating under the integral sign in (10) gives

$$\frac{d}{dq} D(x; q) |_{q=1} = \int_0^x ((t-2)e^t + 3t + 2) e^{e^t+t-1} dt - 2 \int_0^x e^{e^t+t} \int_0^t e^{-e^r-r} dr dt.$$

The first integral may be computed as before in Corollary 2.17 and yields

$$\int_0^x ((t-2)e^t + 3t + 2) e^{e^t+t-1} dt = \sum_{n \geq 1} (2(n-1)B_{n-1} + (n+3)B_n - 2B_{n+1}) \frac{x^n}{n!}.$$

As for the second integral, we have

$$\begin{aligned} \int_0^x e^{e^t+t} \int_0^t e^{-e^r-r} dr dt &= \int_0^x e^{e^t+t-1} \int_0^t e^{1-e^r-r} dr dt \\ &= \int_0^x e^{e^t+t-1} \int_0^t \sum_{j \geq 0} v_j \frac{r^j}{j!} dr dt = \int_0^x \sum_{i \geq 0} B_{i+1} \frac{t^i}{i!} \cdot \sum_{j \geq 1} v_{j-1} \frac{t^j}{j!} dt \\ &= \sum_{n \geq 1} \left(\sum_{i=1}^{n-1} \binom{n-1}{i} v_{i-1} B_{n-i} \right) \frac{x^n}{n!}, \end{aligned}$$

where $v_i = \sum_{j=0}^i (-1)^{i-j} \binom{i}{j} B_j^*$. Combining this with the previous integral gives the desired result. \square

Since the exact formula for the total number of occurrences of 123 is rather complicated, it is desirable to have an asymptotic formula as well.

Corollary 2.24. *The total number of occurrences of the subword 123 within all the members of P_n is given by*

$$2nB_{n-1} + (n - A + 3)B_n - 2B_{n+1} + o(B_n),$$

where the constant A is

$$2 \int_0^\infty e^{1-e^r-r} dr \approx 0.807305.$$

Remark 2.25. *In particular, this implies that the average number of occurrences is $n - 2n/\log n + o(n/\log n)$.*

Proof. We can focus on the part

$$-2 \int_0^x e^{e^t+t} \int_0^t e^{-e^r-r} dr dt$$

of the generating function, since we have exact formulas for the rest. Integration by parts yields

$$\begin{aligned} \int_0^x e^{e^t+t} \int_0^t e^{-e^r-r} dr dt &= e^{e^x} \int_0^x e^{-e^r-r} dr - \int_0^x e^{-t} dt \\ &= e^{e^x} \int_0^x e^{-e^r-r} dr + e^{-x} - 1. \end{aligned}$$

Clearly, the part $e^{-x} - 1$ only contributes $O(1)$. For the rest, we can use a general result recently proven in [1]: namely, if g is an entire function in the complex plane with $g(z) = O(e^{\epsilon|z|})$ for some $\epsilon > 0$ as $|z| \rightarrow \infty$, then the coefficients of

$$F(x) = e^{e^x} \int_0^x e^{-e^t} g(t) dt$$

satisfy

$$[x^n]F(x) = B_n \left(C + O(e^{-\kappa n / \log^2 n}) \right),$$

where $C = \int_0^\infty e^{1-e^t} g(t) dt$ and κ is a positive constant. The function $g(x) = e^{-x}$ satisfies this condition, so we can apply this result to our situation. Putting everything together yields the desired statement. \square

We conclude this section with a central limit theorem for the number of occurrences.

Theorem 2.26. *The number of occurrences of 123 in a randomly chosen set partition of $[n]$ asymptotically follows a normal distribution with mean $\mu_n \sim n(1 - 2/\log n)$ and variance $\sigma_n^2 \sim 4n/(\log n)^2$.*

Proof. Once again, we make use of the saddle point method and the quasi-power theorem (Theorems 2.9 and 2.11). One can argue as in [1] to show that the integral in the generating function (12) can be extended to the range $(0, \infty)$ at the expense of a small error term in the coefficients (cf. the proof of Corollary 2.24 above). We are then left with the analysis of the generating function

$$e^{2qx-x/q+e^{qx}/q^2} \left(e^{-1/q^2} + 2(1-q) \int_0^\infty e^{-2qt+t/q-e^{qt}/q^2} dt \right).$$

The second factor only depends on q , and we can apply Theorem 2.9 to the first factor, with $\theta_0(r) = e^{-2r/5}$ as before and $r = r(n)$ defined implicitly by $re^{qr} = qn$. Since the remaining steps are analogous to Theorems 2.12 and 2.19, we omit the details to avoid repetition. \square

Remark 2.27. *The central limit theorems for the patterns 132, 231 and now 123 have the following intuitive interpretation: the pattern 123 occurs everywhere inside of blocks and only very rarely at the border between blocks, so the number of its occurrences is essentially n minus twice the number of blocks (which is well-known to be normally distributed with mean $B_{n+1}/B_n \sim n/\log n$ and variance $\sim n/(\log n)^2$, see [4, Proposition IX.20]). At the border between blocks, we typically have a descent, and the pattern 231 occurs far more frequently than 132, so that its number follows almost the same distribution as the number of blocks.*

3 Avoiding 312 or 213 and the kernel method

In this section, we consider the cases 312 and 213 and the problem of enumerating the members of P_n avoiding either pattern. Here, we make use of ordinary generating functions and the kernel method to deal with these apparently more difficult cases, and our arguments do not seem to extend to the more general subword counting problem for these patterns.

3.1 The case 312

We seek to enumerate the members of P_n which avoid the pattern 312 in the flattened sense. To do so, we refine part of the set in question as follows. Given $2 \leq i < j \leq n$, let $a_{n,i,j}$ count the members of P_n avoiding 312 whose first block has size at least three, with i the second element of the first block and j the last element of this block. Given $2 \leq i \leq n$, let $b_{n,i}$ count the members of P_n avoiding 312 whose first block has size two and second letter i . For example, we have $a_{5,2,4} = 3$, the enumerated partitions being 124/35, 124/3/5 and 1234/5, and $b_{6,4} = 3$, the partitions being 14/25/36, 14/25/3/6 and 14/256/3.

Let a_n denote the number of members of P_n avoiding 312 in the flattened sense. From the definitions, we have upon considering the size of the first block the relation

$$a_n = a_{n-1} + \sum_i b_{n,i} + \sum_{i,j} a_{n,i,j}, \quad n \geq 3, \quad (13)$$

with $a_1 = 1$ and $a_2 = 2$.

The arrays $a_{n,i,j}$ and $b_{n,i}$ may be determined recursively as follows.

Lemma 3.1. *The arrays $a_{n,i,j}$ and $b_{n,i}$ can assume non-zero values only when $2 \leq i < j \leq n$ and $2 \leq i \leq n$. They satisfy the recurrences*

$$a_{n,i,j} = b_{n-1,j-1} + \sum_{r=i}^{j-2} a_{n-1,r,j-1}, \quad 2 \leq i < j \leq n, \quad (14)$$

and

$$b_{n,i} = \sum_{r=i-1}^{n-2} b_{n-2,r} + \sum_{r=i-1}^{n-2} \sum_{j=r+1}^{n-2} a_{n-2,r,j}, \quad 4 \leq i \leq n, \quad (15)$$

along with the conditions $b_{2,2} = 1$ and $b_{n,2} = b_{n,3} = a_{n-2}$ if $n \geq 3$.

Proof. The first statement is clear from the definitions. Let $\mathcal{A}_{n,i,j}$ and $\mathcal{B}_{n,i}$ denote the subsets of P_n whose members are enumerated by $a_{n,i,j}$ and $b_{n,i}$, respectively. Suppose $\pi \in \mathcal{A}_{n,i,j}$. Then the second letter of the first block is extraneous concerning the avoidance of the 312 subword since it is followed by a larger letter. Deletion of this letter then leads to a member of $\mathcal{B}_{n-1,j-1}$ (upon relabeling elements) if the first block has cardinality three, or to a member of $\cup_{r=i}^{j-2} \mathcal{A}_{n-1,r,j-1}$ if it has cardinality four or more, which implies (14). Now suppose $\pi \in \mathcal{B}_{n,i}$, where $4 \leq i \leq n$. Note that (15) holds when $i = n$, by the convention for empty sums, since if $n \geq 4$ and if the

first block of π is $\{1, n\}$, then the letter n always starts an occurrence of 312 in $\text{Flatten}(\pi)$. If $4 \leq i < n$, then the first block of π being $\{1, i\}$ implies that the second block must contain at least two elements and start with $2, s$, where $s > i$ (for otherwise, there would be an occurrence of 312). Deletion of the first block of π then leads to a member of $\cup_{r=i-1}^{n-2} \mathcal{B}_{n-2,r}$ if the second block has size two, or to a member of $\cup_{r=i-1}^{n-2} \cup_{j=r+1}^{n-2} \mathcal{A}_{n-2,r,j}$ if it has size greater than two, which implies (15). Finally, if the first block of π is $\{1, 2\}$ or $\{1, 3\}$, then both letters in this block are extraneous concerning the avoidance of 312, and deletion of this block leads to an arbitrary partition of length $n - 2$ avoiding 312, which implies $b_{n,2} = b_{n,3} = a_{n-2}$ if $n \geq 3$. \square

Let $L_n, n \geq 1$, denote sequence A005773 in [14] which has generating function

$$\sum_{n \geq 1} L_n x^n = \frac{3x - 1 + \sqrt{1 - 2x - 3x^2}}{2(1 - 3x)}.$$

Note that L_n counts, among other things, the number of directed animals of size n and the number of lattice paths of length $n - 1$ using $(1, 1)$, $(1, 0)$, and $(1, -1)$ steps which start at the origin and stay above the x -axis.

The following result provides an apparently new combinatorial interpretation of this sequence.

Theorem 3.2. *If $n \geq 1$, then the number of members of P_n which avoid the subword 312 in the flattened sense is given by L_n .*

Proof. We first determine some recurrences satisfied by the generating functions for the arrays given in Lemma 3.1. Multiplying both sides of (14) by various factors and adding as the usual first step leads to complications later on. Instead, we start the problem by replacing i with $i + 1$ in (14), and then subtract to get

$$a_{n,i+1,j} - a_{n,i,j} = -a_{n-1,i,j-1}, \quad 2 \leq i < j - 1 \leq n - 1, \quad (16)$$

with $a_{n,j-1,j} = b_{n-1,j-1}$ if $3 \leq j \leq n$. We now consider the sum $a_{n,i} = \sum_{j=i+1}^n a_{n,i,j}$, where $2 \leq i \leq n - 1$, and treat its recurrence instead of dealing with (16) directly, which again seems to be too difficult. Put $a_{n,i} = 0$ if $i < 2$ or $i > n - 1$. Note that

$$a_n = a_{n-1} + \sum_{i=2}^{n-1} a_{n,i} + \sum_{i=2}^n b_{n,i}, \quad n \geq 2, \quad (17)$$

with $a_1 = 1$. We seek to find a_n .

Summing both sides of (16) over $j \geq i + 2$ yields

$$a_{n,i+1} = a_{n,i} - b_{n-1,i} - a_{n-1,i}, \quad 2 \leq i \leq n - 2, \quad (18)$$

with $a_{n,2} = a_{n-1} - a_{n-2}$, $a_{n,n-1} = b_{n-1,n-1}$, and $b_{n,2} = a_{n-2}$ if $n \geq 3$.

Replacing i with $i + 1$ in (15), and subtracting, gives

$$\begin{aligned} b_{n,i+1} - b_{n,i} &= -b_{n-2,i-1} - \sum_{j=i}^{n-2} a_{n-2,i-1,j} \\ &= -b_{n-2,i-1} - a_{n-2,i-1}, \quad 4 \leq i \leq n - 1, \end{aligned} \quad (19)$$

with $b_{n,3} = a_{n-2}$ if $n \geq 3$ and $b_{n,n} = 0$ if $n \geq 4$. Now let $a_n(u) = \sum_{i=2}^{n-1} a_{n,i} u^i$ if $n \geq 3$ and $b_n(u) = \sum_{i=2}^n b_{n,i} u^i$ if $n \geq 2$. Multiplying both sides of (18) by u^{i+1} , and summing over $2 \leq i \leq n - 2$, yields

$$a_n(u) - (a_{n-1} - a_{n-2})u^2 = u(a_n(u) - a_{n,n-1}u^{n-1}) - u(b_{n-1}(u) - b_{n-1,n-1}u^{n-1}) - ua_{n-1}(u),$$

which may be rewritten as

$$(1-u)a_n(u) - (a_{n-1} - a_{n-2})u^2 = -u(a_{n-1}(u) + b_{n-1}(u)), \quad n \geq 4. \quad (20)$$

Note that (20) is also seen to hold for $n = 3$. Multiplying both sides of (19) by u^{i+1} , and summing over $4 \leq i \leq n-1$, yields

$$\begin{aligned} & (b_n(u) - a_{n-2}u^2(1+u) - b_{n,4}u^4) - u(b_n(u) - a_{n-2}u^2(1+u)) \\ & = -u^2(b_{n-2}(u) - a_{n-4}u^2) - u^2(a_{n-2}(u) - (a_{n-3} - a_{n-4})u^2), \quad n \geq 5. \end{aligned}$$

Using (15) with $i = 4$, or by direct combinatorial reasoning, one has $b_{n,4} = a_{n-2} - 2a_{n-3}$ if $n \geq 4$ so that the last equation may be rewritten as

$$(1-u)b_n(u) - a_{n-2}u^2 + a_{n-3}u^4 = -u^2(a_{n-2}(u) + b_{n-2}(u)), \quad n \geq 5, \quad (21)$$

which is also seen to hold for $n = 4$.

Let $g(x; u) = \sum_{n \geq 3} a_n(u)x^n$ and $h(x; u) = \sum_{n \geq 2} b_n(u)x^n$. Multiplying both sides of (20) by x^n , and summing over $n \geq 3$, yields

$$(1-u)g(x; u) - u^2 \sum_{n \geq 3} (a_{n-1}(1) + b_{n-1}(1))x^n = -xu(g(x; u) + h(x; u)),$$

where we have used (17). The last equation may be rewritten as

$$(1-u+xu)g(x; u) + xuh(x; u) = xu^2(g(x; 1) + h(x; 1)). \quad (22)$$

Let $f(x) = \sum_{n \geq 1} a_n x^n$. Multiplying both sides of (21) by x^n , and summing over $n \geq 4$, yields

$$(1-u)(h(x; u) - x^2u^2 - x^3(u^2 + u^3)) - x^2u^2(f(x) - x) + x^3u^4f(x) = -x^2u^2(g(x; u) + h(x; u)).$$

Using (17), one can show that

$$f(x) = \frac{g(x; 1) + h(x; 1) + x}{1-x}. \quad (23)$$

Substituting (23) into the equation preceding it, and simplifying, gives

$$x^2u^2g(x; u) + (1-u+x^2u^2)h(x; u) = \frac{x^2u^2(1-xu^2)(g(x; 1) + h(x; 1))}{1-x} + \frac{x^2u^2(1-u+xu-xu^2)}{1-x}. \quad (24)$$

We now wish to solve for the quantity $g(x; 1) + h(x; 1)$ in the system of functional equations (22) and (24). Solving for $g(x; u)$ in (24), and substituting into (22), gives

$$\begin{aligned} (1-u+xu) \left(-\frac{1-u+x^2u^2}{x^2u^2} h(x; u) + \frac{(1-xu^2)(g(x; 1) + h(x; 1))}{1-x} + \frac{1-u+xu-xu^2}{1-x} \right) \\ + xuh(x; u) = xu^2(g(x; 1) + h(x; 1)). \end{aligned} \quad (25)$$

We now solve (25) using the *kernel method* (see [7]). Substituting

$$u_0 = u_0(x) = \frac{1-x-\sqrt{1-2x-3x^2}}{2x^2}$$

for u in (25) cancels out the coefficient of $h(x; u)$ and implies

$$g(x; 1) + h(x; 1) = \frac{(1-u_0+xu_0)^2 - xu_0^2(1-u_0+xu_0)}{xu_0^2(1-x) - (1-xu_0^2)(1-u_0+xu_0)}. \quad (26)$$

We can simplify the expression in (26). To do so, first observe that u_0 is a root of the quadratic equation $x^2u^2 + (x-1)u + 1 = 0$ and hence $1 - u_0 + xu_0 = -x^2u_0^2$. Therefore, by (26), we have

$$\begin{aligned} g(x; 1) + h(x; 1) &= \frac{x^4u_0^4 + x^3u_0^4}{xu_0^2(1-x) + x^2u_0^2(1-xu_0^2)} = \frac{(1+x)x^2u_0^2}{1-x^2u_0^2} = \frac{(1-x^2)u_0 - (1+x)}{2 - (1-x)u_0} \\ &= \frac{2x^2 - x^2(1-x)u_0}{1-3x}, \end{aligned}$$

where the last equality can be shown by cross-multiplying and using the relation $x^2u_0^2 = (1-x)u_0 - 1$. Thus, we have

$$\begin{aligned} f(x) &= \frac{g(x; 1) + h(x; 1) + x}{1-x} = \frac{2x^2 - x^2(1-x)u_0}{(1-x)(1-3x)} + \frac{x}{1-x} \\ &= \frac{1}{1-x} \left(-\frac{1+x}{2} + \frac{(1-x)(1-x-2x^2u_0)}{2(1-3x)} \right) + \frac{x}{1-x} \\ &= -\frac{1+x}{2(1-x)} + \frac{\sqrt{1-2x-3x^2}}{2(1-3x)} + \frac{x}{1-x} = \frac{3x-1 + \sqrt{1-2x-3x^2}}{2(1-3x)} = \sum_{n \geq 1} L_n x^n, \end{aligned}$$

which completes the proof. \square

Remark 3.3. A standard application of the Flajolet-Odlyzko singularity analysis (see [4, Section VI]) shows that

$$L_n \sim \frac{1}{\sqrt{3\pi n}} \cdot 3^n,$$

i.e., the number of 312-avoiding set partitions only grows exponentially, in contrast to all the cases studied in Section 2.

3.2 The case 213

Let $a_{n,i,j}$, $b_{n,i}$, and a_n be just as in the preceding subsection but with the pattern 312 replaced by 213. Note that the relation (13) continues to hold in this case. Here the arrays are determined recursively as follows.

Lemma 3.4. *The arrays $a_{n,i,j}$ and $b_{n,i}$ can assume non-zero values only when $2 \leq i < j \leq n$ and $2 \leq i \leq n$. They satisfy the recurrences*

$$a_{n,i,j} = b_{n-1,j-1} + \sum_{r=i}^{j-2} a_{n-1,r,j-1}, \quad 2 \leq i < j \leq n, \quad (27)$$

and

$$b_{n,i} = a_{n-3} + \sum_{r=2}^{i-2} b_{n-2,r} + \sum_{r=2}^{i-2} \sum_{j=r+1}^{n-2} a_{n-2,r,j}, \quad 4 \leq i \leq n, \quad (28)$$

along with the conditions $b_{n,2} = a_{n-2}$ and $b_{n,3} = \delta_{n,3}$ if $n \geq 3$ and $b_{2,2} = 1$.

Proof. The proof is similar to that of Lemma 3.1 above. Note that in this case we have $b_{n,3} = \delta_{n,3}$ since if $\{1, 3\}$ is the first block of a partition of size n with $n \geq 4$, then the 3 is always the first letter in an occurrence of 213. For (28), observe that if $4 \leq i \leq n$, then the right-hand side is seen to count partitions according to whether the second block has size one, two, or greater than two. Note that the right-hand side of (28) reduces to a_{n-2} in the case when $i = n$. \square

The generating function $f(x) = \sum_{n \geq 1} a_n x^n$ which counts all the partitions of size n avoiding 213 in the flattened sense is determined by a functional equation.

Theorem 3.5. *We have*

$$f(x) = \frac{x(1-x)u_0}{2-u_0} + \frac{xu_0}{2-u_0}f(xu_0), \quad (29)$$

where $u_0 = u_0(x) = \frac{x-1+\sqrt{1-2x+5x^2}}{2x^2}$.

Proof. We proceed as in the prior case. Replacing i with $i+1$ in (27), and subtracting, gives

$$a_{n,i+1,j} - a_{n,i,j} = -a_{n-1,i,j-1}, \quad 2 \leq i < j-1 \leq n-1,$$

which implies

$$a_{n,i+1} = a_{n,i} - a_{n-1,i} - b_{n-1,i}, \quad 2 \leq i \leq n-2, \quad (30)$$

where $a_{n,i} = \sum_{j=i+1}^n a_{n,i,j}$. Replacing i with $i+1$ in (28), and subtracting, gives

$$b_{n,i+1} = b_{n,i} + a_{n-2,i-1} + b_{n-2,i-1}, \quad 4 \leq i \leq n-1. \quad (31)$$

If $g(x; u)$ and $h(x; u)$ are defined analogously as before, then (30) and (31) lead to the functional equations

$$(1-u+xu)g(x; u) + xuh(x; u) = xu^2((1-x)f(x) - x) \quad (32)$$

and

$$-x^2u^2g(x; u) + (1-u-x^2u^2)h(x; u) = x^2u^2(1-u+xu^2)f(x) - x^2u^3f(xu) + x^2u^2(1-u+xu). \quad (33)$$

Solving for $g(x; u)$ in (33), and substituting into (32), gives a functional equation involving the quantities $h(x; u)$, $f(x)$, and $f(xu)$. Replacing u with u_0 in this equation, where u_0 is as defined above, cancels out the coefficient of $h(x; u)$ and yields

$$(1-u_0+xu_0)(u_0-1-xu_0^2)f(x) + (1-u_0+xu_0)u_0f(xu_0) - (1-u_0+xu_0)^2 = x(1-x)u_0^2f(x) - x^2u_0^2.$$

Since $x^2u_0^2 = 1-u_0+xu_0$, this last equation simplifies to

$$x(u_0-1-xu_0^2)f(x) + xu_0f(xu_0) - x^3u_0^2 = (1-x)f(x) - x,$$

which reduces further to

$$(u_0-2)f(x) + xu_0f(xu_0) = x(x-1)u_0.$$

□

4 Conclusion

Let us summarize the results obtained and compare the behavior observed for the different patterns, starting with the number of set partitions avoiding a certain pattern as a subword in the flattened sense.

Pattern	Number of set partitions of $[n]$ avoiding the pattern
123	$(n-2)!!$ (n odd), $2(n-2)!!$ (n even) – Corollary 2.21
132	$\sim B_n \exp(-(\log n)^2/2 + o((\log n)^2))$ – Theorem 2.10
213	Unknown – generating function determined by Theorem 3.5
231	$\sim \frac{1}{\sqrt{2n}} \cdot n^{n/2} e^{2\sqrt{n}-n/2-1}$ – Theorem 2.15
312	$\sim \frac{1}{\sqrt{3\pi n}} \cdot 3^n$ – Remark 3.3
321	All set partitions, $B_n \sim \exp(n \log n - n \log \log n - n + o(n))$ – Proposition 2.1

The table shows that the asymptotic behavior varies quite a lot: as n goes to infinity, we get the following chain of inequalities, where $f_n^{(\text{pat})}$ is the number of **pat**-free set partitions of $[n]$:

$$f_n^{321} > f_n^{132} > f_n^{231} > f_n^{123} > f_n^{312}.$$

It would be very interesting to determine where f_n^{213} fits in. Numerically, it seems to be between f_n^{321} and f_n^{132} .

We also obtained results on the number of occurrences of four different patterns. As Proposition 2.1 shows, 321 never occurs. For three other patterns, we were able to show that the number of occurrences asymptotically follows a normal distribution:

- 132: mean $\sim (\log n)^2/2$, variance $\sim (\log n)^2/2$ (Theorem 2.12),
- 231: mean $\sim n/\log n$, variance $\sim n/(\log n)^2$ (Theorem 2.19),
- 123: mean $\sim n(1 - 2/\log n)$, variance $\sim 4n/(\log n)^2$ (Theorem 2.26).

Similar results can be expected for the patterns 213 and 312, but it is not at all clear how to obtain them and what the order of magnitude of the mean and variance will be.

References

- [1] W. Asakly, A. Blecher, C. Brennan, A. Knopfmacher, T. Mansour, and S. Wagner, Set partition asymptotics and a conjecture of Gould and Quaintance, submitted.
- [2] A. Burstein and T. Mansour, Counting occurrences of some subword patterns, *Discrete Math. Theor. Comput. Sci.* **6** (2003) 1–11.
- [3] D. Callan, Pattern avoidance in “flattened” partitions, *Discrete Math.* **309** (2009) 4187–4191.
- [4] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.
- [5] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science* (second edition), Addison-Wesley, Boston 1994.
- [6] S. Heubach and T. Mansour, *Combinatorics of Compositions and Words*, CRC Press, Boca Raton, 2009.
- [7] Q. Hou and T. Mansour, Kernel method and linear recurrence system, *J. Comput. Appl. Math.* **261(1)** (2008) 227–242.
- [8] S. Kitaev, *Patterns in Permutations and Words*, Springer-Verlag, Heidelberg, 2011.
- [9] T. Mansour and M. Shattuck, Pattern avoidance in flattened permutations, *Pure Math. Appl. (P.U.M.A.)* **22(1)** (2011) 75–86.
- [10] T. Mansour, M. Shattuck, and D. G. L. Wang, Counting subwords in flattened permutations, *J. Comb.* **4:3** (2013) 327–356.
- [11] T. Mansour, M. Shattuck, and S. H. F. Yan, Counting subwords in a partition of a set, *Electron. J. Combin.* **17** (2010) #R19.
- [12] T. Mansour and B. Sirhan, Counting ℓ -letter subwords in compositions, *Discrete Math. Theor. Comput. Sci.* **8** (2006) 285–298.
- [13] V. R. Rao Uppuluri and J. A. Carpenter, Numbers generated by the function $\exp(1 - e^x)$, *Fibonacci Quart.* **7** (1969) 437–448.
- [14] N. J. Sloane, The On-Line Encyclopedia of Integer Sequences, <http://oeis.org>, 2010.
- [15] D. Zeilberger, Enumeration schemes, and more importantly, their automatic generation, *Ann. Comb.* **2** (1998) 185–195.