What Does the Future Hold for Restricted Patterns?[1]

by Zvezdelina Stankova

Berkeley Math Circle – Advanced Group

November 26, 2013

1. An Exhaustive Survey versus Paths for Further Research

Restricted patterns made their major debut into the math arena in the 1980's, with the works of Lothaire, Lovász, Richards, Rotem, Schmidt, Simion, Wilf, and many others. In truth, they had already permeated mathematical research since the 1960's via Knuth's, Robinson-Schensted's and Stanley's earlier results. The more recent re-birth of the topic was initiated by West in 1990's, and taken up by a number of researchers, including Albert, Arratia, Backelin, Babson, Bóna, Mansour, Marcus, Stankova, and Tardos. Any time you stumble upon the Catalan, Fibonacci, or Stirling numbers, Dyck paths, Young diagrams, random matrices, generating trees or Chebychev or Kazhdan-Lusztig polynomials, restricted patterns are likely to appear in one reincarnation or another.

The topics stemming from or related to restricted patterns are so numerous that the means of a single paper are vastly insufficient to even briefly mention all of them. Thus, we are "restricted" in this paper to pursuing only several paths of pattern exploration, which we can roughly group in the following themes:

- the original *Wilf-classification* of permutations up to S_8, and necessary computer aid for further study;
- the classification of *pairs of permutations* up to S_6, and comparison with singletons, as well as triples or quadriples, of restricted permutations;
- difficulties and possibilities in *ordering* and *asymptotic ordering* of Wilf-classes of singleton permutations;
- extensions of the classical *Stanley-Wilf limits* along paths of Young diagrams and the algebraic closure of the resulting set of limits.

In view of the above, the current paper does not claim to be an exhaustive survey of restricted patterns, but rather a survey of several research paths which originated at the Research Experience for Undergraduates Program in Duluth in 1991-1992, and have been developed by Duluth alumni and other mathematicians for over a decade. The paper emphasizes the view of the author on important aspects and possibilities which will likely affect the future research of these topics. New and old conjectures and open questions will therefore permeate this study.

[1]This is a paper in progress and contains some unfinished notes.

Major contributors to the results discussed here are (in alphabetic order): Arbres, Bóna, Gire, Kremer, Le, Mansour, Shiu, Sophie, Stankova, and West. Many others will be also mentioned in due time. Because of the immense cross-referencing and intertwining of results of various people, it is practically impossible to track down and attribute the original source of every single idea in this paper; most of the unattributed ideas have originated in the course of the author's own research, and I will be grateful if further references are pointed out by the reader.

The paper strives to be written in an easy-to-follow style, intended to encourage future young, as well as "seasoned," researchers to pursue the outlined topics. Although some of the posed questions could conceivably be attacked with the existing methods and the suggested new strategies, the reader should not be deceived: many of the questions in this paper will require deeper analysis, completely new approaches and links to other math areas that have not been discovered yet.

2. Acknowledgments

Acknowledgments are usually reserved for the end. But I am compelled to make them here because of their importance on the current paper and on my own research in Enumerative Combinatorics. First and foremost, it was *Joe Gallian*, the director of the Duluth REU Program, who solicited in 1991 the original open problems on pattern avoidance; who offered both professional and personal guidance in my first research years, and whose genuine care and encouragement for me throughout the years has contributed immensely to the mathematician and the person I am now.

My first 1991-1992 projects were mentored by *David Moulton* (Center for Communications Research, New York), a graduate assistant at the Duluth REU at the time. My co-authors on later projects, *Julian West* (Malaspina University, Canada) and *Toufik Mansour* (University of Haifa, Israel), as well as *Miklos Bóna* (University of Florida), have provided further insights into the area of restricted patterns. *Herbert Wilf* (University of Pennsylvania) has also been very supportive over the years, starting from my undergraduate research attempts while at Bryn Mawr College to the very summer conference "Communicating Mathematics" in 2007, Duluth, where he took the time to study my presentation in depth and suggest the publication of the current paper.

As it will become apparent from the exposition, without computer analysis many of the conjectures and some of the classification results in the field would have been unthinkable. Another Duluth REU alumni, *David Moews* (Center for Communications Research, San Diego) wrote the computer programs used for my first 1991-92 projects, and adapted them to work faster and encompass a variety of new tasks over the ensuing 10 years of research. Finally, I would like to thank *Paulo de Souza* (UC Berkeley) for his help in implementing the necessary computer software, and *Olga Holtz* (UC Berkeley) for generously offering the capabilities of her computer in the month-long calculation of the Wilf-classification of S_8.

3. Basics on Restricted Patterns

3.1. **The primary object of study.** We agree to write a permutation τ of length k in one-row notation as $(a_1, a_2, \ldots, a_k)$ where $\tau(i) = a_i$ for $1 \le i \le k$. For $k < 10$ the commas can be supressed without causing confusion. As usual, S_n denotes the symmetric group on $[n] = \{1, 2, ..., n\}$.

The definition below is the fundamental one that underpins the whole area of restricted patterns:

Definition 1. Let τ and π be two permutations of lengths k and n, respectively. We say that π is τ-*avoiding* if there is no subsequence $i_{\tau(1)}, i_{\tau(2)}, ..., i_{\tau(k)}$ of $[n]$ such that $\pi(i_1) < \pi(i_2) < \ldots < \pi(i_k)$. If there is such a subsequence, we say that it is *of type* τ, and denote this by $\big(\pi(i_{\tau(1)}),\ \pi(i_{\tau(2)}),...,\ \pi(i_{\tau(k)})\big) \approx \tau$.

As one can see, without several examples worked out with a pencil and paper, this formal definition is not very insightful. To understand where restricted patterns really originated from, we draw on visual imagery and replace "one-dimensional" permutations by "two-dimensional" objects, matrices. In doing so, we shall violate the customary labeling of the top row as the "first" row of a matrix. Instead, we shall coordinatize our matrices from *the bottom left corner*, just like in a Cartesian coordinate system; thus, the origin will always be placed at the *bottom left corner* of the matrix (cf. Fig. 1) so that our first row will be the *bottom row*, and our first column will be (as usual) the leftmost column of a matrix. This is done in order to keep the resemblance with the "shape" of permutations, in other words, with their *graphs* as functions $\pi : [n] \to [n]$.

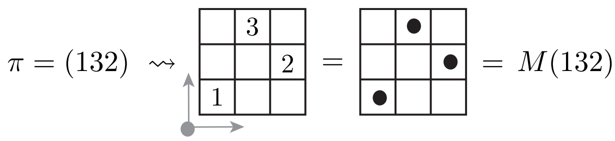


Figure 1. Permutation matrix $M(132)$

Here is a simple example to start with (cf. Fig. 1). The matrix $M(132)$ associated to $\pi = (132)$ is a 3×3 matrix with dots in cells $(1, 1)$, $(2, 3)$ and $(3, 2)$. Figure 2 displays the larger matrices $M(52687431)$, $M(3142)$ and $M(2413)$. In general,

Definition 2. Let $\pi \in S_n$. The *permutation matrix* $M(\pi)$ is the $n \times n$ matrix M_n having a 1 (or a dot) in position $(i, \pi(i))$ for $1 \le i \le n$.

As the reader has probably observed by now, a permutation matrix is nothing but an arrangement of n non-attacking rooks on an $n \times n$ board, called a *transversal* of the board with elements the "1's", or the "dots".

The original pattern-avoidance Definition 1 is designed in such a way that a permutation $\pi \in S_n$ contains a subsequence $\tau \in S_k$ exactly when the matrix $M(\pi)$ contains $M(\tau)$ as a

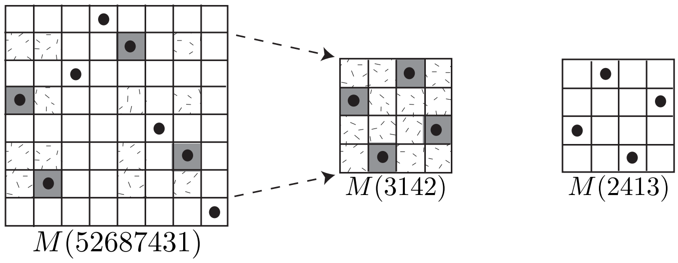


FIGURE 2. $(52687431) \notin S_8(3142)$, but $(52687431) \notin S_8(2413)$

submatrix. For instance, Figures 2a-b demonstrate that $\pi = (52687431)$ has a subsequence (6273) of type (3142) exactly because $M(52687431)$ has a 4×4 submatrix formed by the rows and columns of (6273) and identical to $M(3142)$. On the other hand, it is not hard to convince yourself that no submatrix identical to $M(2413)$ (cf. Fig. 2c) is contained in $M(52687431)$, and thus (52687431) avoids the permutation (2431).

We conclude that avoidance of *permutations* is an equivalent notion to avoidance among *permutations matrices*: permutation π avoids τ if and only if matrix $M(\pi)$ avoids $M(\tau)$, i.e. $M(\pi)$ does not contain a submatrix identical to $M(\tau)$.

3.2. **Interpretation of 231-avoidance.** We shall fix now one permutation τ and investigate the set of permutations of length n avoiding τ. This set is denoted by $S_n(\tau)$. To flesh out our understanding of pattern avoidance, let us describe one initial but nevertheless striking example: that of $S_n(231)$.

In [14] Knuth shows that $S_n(231)$ is precisely the set of *stack-sortable* permutations (cf. also [20]). To visualize the situation, imagine a train station with one main track, and one dead-end side track used for storing temporarily wagons. A cargo train is coming into the station from the right along the main track (cf. Fig. 3). Its four wagons (starting with the leading one) are numbered by "4", "1", "3" and "2". The goal is to rearrange the wagons so that the train leaves the station with wagons numbered in increasing order, "1", "2", "3", "4". We can use the side track to store as many wagons as we wish, however, at any time we can pull out only the most recently stored wagon onto the main track and we must push it immediately forward to join the sorted out train.

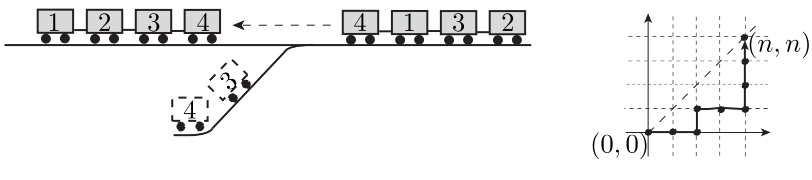


FIGURE 3. Stack-sortability, 231-avoidance, and Lattice paths

The reader can easily spot the solution to sorting out the given train "4132", and just as easily realize that this is impossible if the incoming train were labelled "3142" because of its subsequence (342) of type (231). It turns out that containing the permutation (231) is the only obstruction to sorting out our trains. Thus, the permutations that can be sorted out in the above way, the so-called *stack-sortable* permutations, are precisely those that avoid (231).

Evidently, two different such permutations have different sorting algorithms. Furthermore, the sorting algorithm for each such permutation of length n is unique, encoded by binary strings of length $2n$, where "0" stands for "move into the stack", while "1" – for "move out of the stack". Since we can't move out of the stack more wagons than what is currently stored there, these binary strings can be thought of as properly parenthesizes expressions with "(" and ")" replacing corespondingly "0" and "1". Such expressions, on the other hand, are nothing else but lattice walks from the origin $(0, 0)$ to the point (n, n) that do not cross the diagonal $y = x$ and that are composed only of unit-length steps to the right or up. For instance, our train (4132) can be encoded as 00100111="(()(()))", which in turn is the lattice path shown in Figure 3b.

But as it is well-known, the *Catalan numbers* c_n also count exactly the same lattice paths![2] Making a full circle around, we conclude that

$$|S_n(231)| = c_n \text{ for all } n \geq 1. \tag{1}$$

For instance, there are $c_4 = 14$ trains of length 4 which can be sorted out. Equation (1) is the first elementary, yet non-trivial enumerative result on restricted patterns, which should have given the reader a flavor of the rich combinatorial possibilities in this field. Even though enumerating the various sets $|S_n(\tau)|$ is a worthy and challenging problem in itself (and we shall come to it in a later part of the paper), we mainly view it as a vehicle to solving a much more enticing puzzle.

3.3. **Wilf-equivalence.** Comparing two different permutations τ and σ in our setting naturally leads to comparing their associated S_n-subsets, $S_n(\tau)$ and $S_n(\sigma)$.

Definition 3. Two permutations τ and σ are *Wilf-equivalent*, denoted by $\tau \sim \sigma$, if they are equally restrictive on any length permutations, i.e.

$$|S_n(\tau)| = |S_n(\sigma)| \text{ for all } n \in \mathbb{N}.$$

The classification of permutations up to Wilf-equivalence is the first classic and still far from resolved question in the field of restricted patterns. To get a feeling for it, let us investigate the first non-trivial situation, in S_3 (for there is nothing interesting to say in S_1 or S_2).

In Figure 4 we have grouped the six permutations of S_3 into two *symmetry classes*: $\{(321), (123)\}$ and $\{(132), (312), (213), (231)\}$. It is evident that within each such class the permutations are Wilf-equivalent; for instance, if you flip the matrix $M(321)$ across a horizontal axis, you will obtain the matrix $M(123)$; this same flip induces a bijection $S_n(321) \cong S_n(123)$, rendering (321) and (123) as Wilf-equivalent. Similarly, flipping $M(132)$

[2]The Catalan numbers are given by: $c_n = \frac{1}{n+1}\binom{2n}{n} = \sum_{i=1}^{n} c_{i-1}c_{n-i}$, $c_0 = c_1 = 1$.

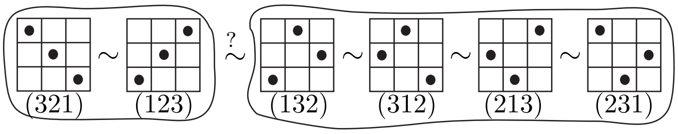


FIGURE 4. Classification of S_3 up to Wilf-equivalence

across a horizontal, vertical or diagonal axis of symmetry produces $M(312)$, $M(213)$ and $M(231)$, and implies $(132) \sim (312) \sim (213) \sim (231)$. In general, acting by the dihedral group (or the symmetries of the square) on the $n \times n$ permutation matrices produces classes of Wilf-equivalent permutations; the orbits of this action are the above-mentioned *symmetry classes* of permutations.

Thus, we have split S_3 into two symmetry classes, and the only thing left to resolve is whether permutations from different classes are *Wilf*-equivalent. At the behest of Wilf, a bijection $S_n(123) \cong S_n(132)$ was suggested by Knuth [14], and shown by Rotem [20], Richards [19], Simion-Schmidt [21], and West [29]. We do not discuss these proofs here, for we shall later, that this situation is a specific instance of a much wider phenomenon. As an exercise, the reader is encouraged to produce one such bijection, or prove that, say, $|S_n(123)| = c_n$. We conclude that

Theorem 1. *S_3 consists of a single Wilf-class, and $|S_n(\tau)| = c_n$ for all $\tau \in S_3$.*

The situation is considerably more complicated already on the level of S_4, not to mention longer permutations. We will return to the classic Wilf-classification in the main part of the paper.

3.4. **Wilf-equivalent pairs.** Nothing prevents us from "forbidding" several permutations at a time. In other words,

Definition 4. *For a collection Ω of permutations (not necessarily of the same length) we denote by $S_n(\Omega)$ the set of permutations in S_n avoiding everything in Ω: $S_n(\Omega) = \cap_{\tau \in \Omega} S_n(\tau)$. If two collections Ω and Υ are equally restrictive on any length permutations, i.e. $|S_n(\Omega)| = |S_n(\Upsilon)|$ for all n, then Ω and Υ are called Wilf-equivalent, which we denote by $\Omega \sim \Upsilon$.*

Considerable attention in the field has been devoted to studying Wilf-equivalent *pairs*. According to a classic result of Erdös and Szekeres [11], the identity permutation $I_k = (1, 2, 3, \ldots, k)$ and its reverse $J_k = (k, k-1, \ldots, 2, 1)$ impose too many conditions on large enough permutations, and hence cannot be avoided simultaneously; more precisely, $|S_n(I_k, J_l)| = 0$ for $n > (k-1)(l-1)$.

More than two decades ago, Simion-Schmidt [21] classified pairs in S_3 up to Wilf-equivalence. Figure 5 displays a representative pair for each of the 5 symmetry classes (pairs are indicated by "+"), and only 2 Wilf-equivalences between these symmetry classes.

Thus, for instance, $\{(132),(231)\} \sim \{(132),(213)\} \sim \{(123),(132)\}$ accounts for the first Wilf-class. For each of the resulting 3 Wilf-classes, a number on the side corresponds to $|S_n(\tau_1, \tau_2)|$, e.g. $|S_n((123),(231))| = \binom{n}{2} + 1$. The size 0 for the third Wilf-class is nothing but the above-mentioned Erdös-Szekeres result.

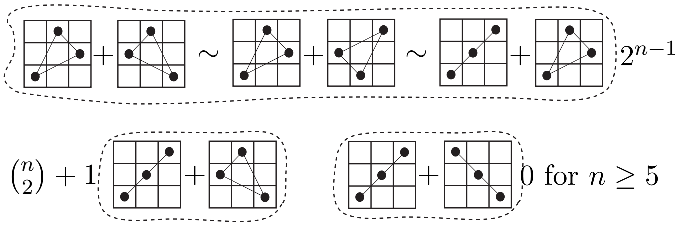


FIGURE 5. Wilf-classes of pairs in S_3

Theorem 2 (Simion-Schmidt). *There are 5 symmetry classes of pairs (τ, σ) of permutations of length 3, with representatives*[3] $(132, 231)$, $(132, 213)$, $(123, 132)$, $(123, 231)$, *and* $(123, 321)$. *There are 3 Wilf-classes of such pairs: the first three symmetry classes join to form one Wilf-class, and the last two symmetry classes stay separate to form two more Wilf-classes. Each Wilf-class produces S_n-subsets of the following sizes: $|S_n(132, 231)| = 2^{n-1}$ and $|S_n(123, 231)| = \binom{n}{2} + 1$ for all $n \geq 1$, while $|S_n(123, 321)| = 0$ for $n \geq 5$.*

The reader is again encouraged to prove these results, whether by direct enumeration of the sizes $|S_n(\tau_1, \tau_2)|$ of each symmetry class, or by finding explicit bijections between these symmetry classes.

A vast amount of research has been generated by the study of Wilf-equivalence of pairs. Not surprisingly, the classification of $(4,4)$ pairs (i.e. pairs in S_4) already demands a variety of new methods and deeper analysis. What is surprising is that non-trivial Wilf-equivalences among *pairs* seem to occur lot more frequently than among *singleton* of permutations. We shall see examples of this later, but for now it suffices to say that this phenomenon has not been yet explained at all.

3.5. **Beyond Wilf-equivalence.** Even if we manage to classify all permutations up to Wilf-equivalence, there will still be quite a few questions left to answer. For instance, if $\tau \not\sim \sigma$, then for some n one of the two sets $S_n(\tau)$ and $S_n(\sigma)$ must be smaller, say, $|S_n(\tau)| < |S_n(\sigma)|$. This would mean that τ occurs more often as a subpattern of length-n permutations and hence τ is "harder" to avoid in S_n than σ. Formally,

Definition 5. *If $|S_n(\tau)| \leq |S_n(\sigma)|$ for all $n \in \mathbb{N}$, we say that τ is more restrictive than σ, and denote this by $\tau \preceq \sigma$.*

As usual, let's examine the initial cases of Wilf-ordering. There is nothing to say for S_3, as everything there is Wilf-equivalent to anything else. To talk about Wilf-*ordering* in S_4, we must first understand the Wilf-*equivalences* in S_4.

[3]For convenience, we have dropped the parentheses around each permutations.

Theorem 3 (Stankova,West)**.** *There are* 7 *symmetry classes in* S_4*, whose representatives enter in the following Wilf-equivalences:* $(1234) \sim (1243) \sim (2143) \sim (4123)$*,* $(4132) \sim (3142)$*, and* (1324) *stays separate, for a total of* 3 *Wilf-classes.*

Representatives of each Wilf-class appear in Figure 6a. The Wilf-classification of S_4 was completed over several years by West [29] and Stankova [22, 23]. It required several new methods and is definitely not an easy exercise to offer to the reader for practice. We shall discuss it in detail in a later section. Meanwhile, let's see how the 3 Wilf-classes in S_4 measure against each other. In [5, 7], Bóna provided the only known so far result on complete Wilf-ordering of S_k:

Theorem 4 (Bóna)**.** *The three Wilf-classes in* S_4 *are ordered as* $(1342) \preceq (1234) \preceq (1324)$*:*

$$|S_n(1342)| < |S_n(1234)| < |S_n(1324)| \text{ for } n \geq 7. \tag{2}$$

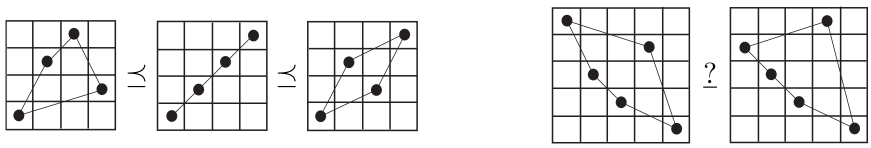


FIGURE 6. Total Wilf-ordering on S_4 and S_5

For each of the two inequalities, Bóna created essentially a new method containing a number of beautiful sophisticated ideas, to which a couple of paragraphs here will not serve justice. More recently, Stankova [25] devised another way of viewing the problem via *decomposable* permutations and generalized Bóna's result to arbitrary lengths.

The point is: even the first non-trivial case of Wilf-ordering generated a great deal of new research, and is indeed far from trivial. The ultimate success in Wilf-ordering S_4 raised hopes for an old conjecture of West, according to which any two permutations (of same length) can be ordered, and hence there is total Wilf-ordering on any S_n. Yet, after years of fruitless search for a proof, a counterexample was found by Stankova already in S_5:

$$S_7(53241) < S_7(43251) \text{ but } S_{13}(53241) > S_{13}(43251),$$

so that these two permutations *cannot* be Wilf-ordered (cf. Fig. 6b). A number of counterexamples were further traced in S_6 and S_7 (cf. [24]), and this completely obviated any hopes for a total Wilf-ordering of a general S_k. What next?

3.6. **Stanley-Wilf Limits.** Naturally, we look at asymptotic behavior of permutations and hope for *asymptotic ordering* of S_n.

Definition 6. For two permutations τ and σ, we say that τ is *asymptotically more restrictive* than σ, denoted by $\tau \preceq_a \sigma$, if $|S_n(\tau)| \leq |S_n(\sigma)|$ for all $n \gg 1$.

Again, the results are not plentiful. But interestingly enough, most of them revolve around the famous Stanley-Wilf conjecture from 1980, which was proven only recently in 2004 by Marcus and Tardos [16].

Theorem 5 (Stanley-Wilf, Marcus-Tardos)**.** *For any permutation τ there is a constant c_τ so that $|S_n(\tau)| \le c_\tau^n$ for all $n \ge 1$. Equivalently, for any τ the limit $L(\tau) = \lim_{n\to\infty} \sqrt[n]{|S_n(\tau)|}$ exists.*

Let's verify the theorem for any $\tau \in S_3$ using the properties of Catalan numbers:

$$|S_n(\tau)| = c_n = \frac{1}{n+1}\binom{2n}{n} \;\Rightarrow\; \frac{c_{n+1}}{c_n} = \frac{4n+2}{n+2} < 4 \;\Rightarrow\; c_n < 4^n \text{ for all } n \ge 1.$$

Hence, $c_\tau = 4$ is one possible Stanley-Wilf constant. That we cannot improve this constant is evident from a similar calculation:

$$\frac{c_{n+1}}{c_n} > \frac{4n}{n+2} \;\Rightarrow\; c_n > \frac{4^n}{2n(n+1)} \;\Rightarrow\; \sqrt[n]{c_n} > \frac{4}{\sqrt[n]{2n(n+1)}} \stackrel{n\to\infty}{\to} 4.$$

Therefore, $L(\tau) = 4$ for any $\tau \in S_3$.

As long as we know the Stanley-Wilf limits in S_n, inequalities among them will certainly suggest asymptotic ordering on S_n. In this vein, $L(I_n) = (n-1)^2 \le L(\tau)$ for any *layered* pattern[4] $\tau \in S_n$ is strong evidence that the identity pattern I_n is more restrictive than all layered patterns in S_n (cf. Bóna [6, 8] and Regev [18]). But in case these two limits *coincide*, even this beautiful result will *not* guarantee asymptotic ordering between I_n and τ.

To make matters more challenging, calculating Stanley-Wilf limits in general turns out to be far from the "straightforward" case of S_3. Can you calculate, for instance, $L(1324)$? Even though the other two limits in S_4 are known: $L(1342) = 8$ and $L(1234) = 9$, apart from the recent lower bound $L(1324) > 9.35$ [1], finding the exact value of $L(1324)$ is still an open question. In a later section we will explore more cases of Stanley-Wilf limits, some methods and ideas created within the proof of Theorem 5, as well as possible generalizations of Stanley-Wilf limits to paths of Young diagrams.

4. Wilf-Classification of S_n: How Much can We Hope for?

4.1. **Permutations of Length 4 Demand New Methods.** We discussed earlier that S_3 consists of one Wilf-class, and suggested some ways to show it. Theorem 3, on the other hand, only listed the 7 symmetry classes on S_4 and grouped them in 3 Wilf-classes, without indicating as to how the theorem was proven. It took several years, one Ph.D. thesis (West's [29]) and two papers written at the Duluth REU'91-'92 (Stankova's [22, 23]), to complete this project.

Let us visualize the situation as in Figure 7. If you take a closer look at the Wilf-equivalences $(2143) \sim (1243) \sim (1234)$ in the 3rd Wilf class, you will notice that representatives of the symmetry classes are obtained from one another either by switching the *two*

[4] A *layered* pattern is composed of increasing blocks of decreasing subpatterns, e.g. (3217654) is layered.

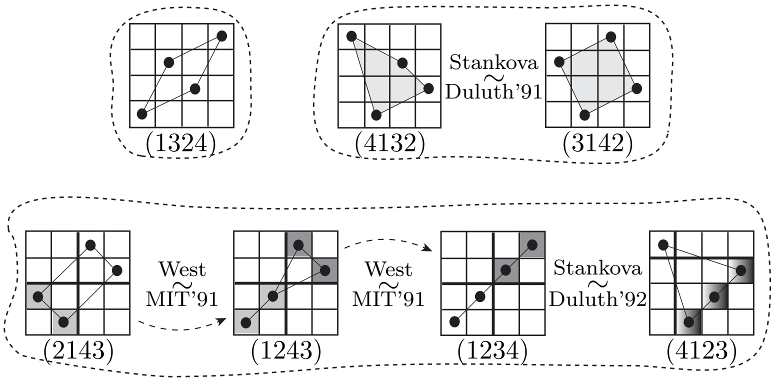


FIGURE 7. Classification of S_4 up to Wilf-equivalence

smallest elements 1 and 2, which "happen" to be in the beginning, or by switching the *two largest* elements 3 and 4, which are positioned at the end of the permutations. This is no coincidence. After applying appropriate symmetries to the permutation matrices, *both* of these Wilf-equivalences become special cases of the following theorem:

Theorem 6 (Babson-West'00)**.** *If the largest two elements of a permutation lie at the end of it, switching them preserves the Wilf-equivalency class:*

$$(a_1, a_2, \ldots, a_{n-2}, n-1, n) \sim (a_1, a_2, \ldots, a_{n-2}, n, n-1)$$

for any $n \geq 2$ *and any permutation* $(a_1, a_2, \ldots, a_{n-2}) \in S_{n-2}$.

Even though West conjectured the statement and proved convincingly a number of cases already in 1991 [27], it was a while before new tools for studying restricted patterns became available and the complete and rigorous proof of Theorem 6 was published in 2000 [3].

Instead, let us backtrack a little and concentrate on the available methods at the time when the classification of S_4 was completed.

4.2. **Generating Trees.** The method of *generating trees* was introduced in 1978 by Chung, Graham, Hoggatt, and Kleiman [10] in relation to *Baxter* permutations and has been a common tool for studying permutations ever since. The idea of using this method specifically in the context of restricted patterns was taken up by West in the early 1990's.

Definition 7. *The generating tree* $T(\tau)$ *of a permutation* τ *is defined as follows:*

- *the nodes on level* n *are the permutations of length* n *avoiding* τ*, i.e. all* $\pi \in S_n(\tau)$*;*
- *the children of a node* π *on level* n *are obtained by inserting* $n+1$ *in appropriate places in* π *so as to still avoid* τ *on level* $n+1$*.*

Figure 8a displays the tree $T(123)$; note that the shown 4 levels of the trees contain correspondingly 1, 2, 5 and 14 nodes, the Catalan numbers. On level 4 we have named only the 4 children of the most "prolific" node of length 3, (321).

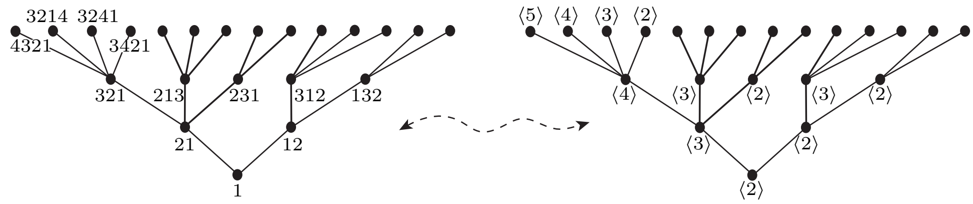


FIGURE 8. The tree $T(123)$ and its labeling

If two trees are isomorphic, then in particular their corresponding levels have the same number of nodes, which means that the two restricted permutations are Wilf-equivalent: $T(\tau) \cong T(\sigma) \ \Rightarrow\ \tau \sim \sigma$. Isomorphisms between generating trees are usually established by first *labeling* the nodes of each tree, and then constructing a 1-1 correspondence between these labelings. The label of node, hence, must contain all information about the subtree generated by this node.

It is easy to see how we should label the tree $T(123)$. If $\pi = (a_1, a_2, \ldots, a_n) \in S_n(123)$ has initial decreasing subsequence $(a_1, a_2, \ldots, a_k)$, then $n+1$ can be inserted in π everywhere in the first $k+1$ slots, but not afterwards. Thus, the terms $a_{k+1}, a_{k+2}, \ldots, a_n$ are irrelevant from now on and can be erased because π will generate the same subtree of $T(123)$ as the decreasing sequence $J_k = (k, k-1, \ldots, 1)$. In conclusion, we could label π by the length k of its initial decreasing subsequence, or equivalently, by the number of its children, $k+1$. In Figure 8b we have opted for the second approach. Note that the $k+1$ children of $\pi = \langle k+1 \rangle$ will have labels $\langle 2 \rangle$, $\langle 3 \rangle$, ..., $\langle k+2 \rangle$; thus, the label of a node completely determines the labels of its children.

The reader is invited to try to label $T(132)$ in a similar fashion and establish the isomorphism $T(123) \cong T(132)$, which confirms once again that $(123) \sim (132)$ in S_3. From the viewpoint of S_4 however, we are interested in potential trees isomorphisms for the permutations claimed to be Wilf-equivalent in Figure 7. West found two such isomorphisms in S_4 [28], and managed to extend one of them to a wider class of permutations [27].

Theorem 7 (West'91). $T(2143) \cong T(1243) \cong T(1234)$, *and the isomorphisms are unique. More generally, switching the largest two elements of the identity permutation I_n creates two isomorphic trees:*

$$T(\underbrace{1, 2, \ldots, n-2, n-1, n}_{I_n}) \cong T(\underbrace{1, 2, \ldots, n-2}_{I_{n-2}}, n, n-1).$$

As Theorem 6 suggest, the ensuing Wilf-equivalences in S_4, namely, $(2143) \sim (1243) \sim (1234)$ are part of a much wider class of Wilf-equivalences. The used labelings of the trees in the proof of Theorem 7 are easy to follow and reasonable to come up with.

But now let's take a look at the first proposed Wilf-equivalence in Figure 7: $(4132) \overset{?}{\sim} (3142)$ in S_4. It says that the *least symmetric* pattern in S_4, the quadrilateral (4132), and the *most symmetric* pattern in S_4, the square (3142), are equally restrictive?! Not only that, but West conjectured further that their permutation trees were isomorphic, and the present author proved it in [22].

Theorem 8 (Stankova'91)**.** $T(4132) \cong T(3142)$.

The proof is based on a detailed study of the "parent-children relationships" in each of the two trees, and on creating appropriate labelings to capture these relationships. To give you a sense of the involved complications, here is the labeling for $T(4132)$. If a node $\tau \in T(4132)$ has k children, its label will have k entries, $(i_1, i_2, \ldots, i_k)$; then for $s = 1, 2, \ldots, k$, its s^{th} child will have label $(\underbrace{1, 1, \ldots, 1}_{i_s}, \hat{i}_s, \hat{i}_{s+1}, \ldots, \hat{i}_k)$, where for $j = s, s+1, \ldots, k$:

$$\hat{i}_j = \begin{cases} i_j & \text{if } i_j \leq j - s, \\ i_j + 1 & \text{if } j - s < i_j \leq i_s + j - s, \\ i_s + 1 + j - s & \text{if } i_s + j - s < i_j. \end{cases}$$

$T(3142)$ has a similarly involved labeling. As one can imagine, to come up with the two labelings, to show that they work and that they are in 1-1 correspondence, takes some effort.

But here is the most striking fact: the resulting Wilf-equivalence $(4132) \sim (3142)$ does *not* fit in any larger class of equivalences. It is *the only sporadic case* of Wilf-equivalences known so far, and there is no known proof of it other than the above tree isomorphism.

Question 1. *Is there a deeper reason, beyond the tree isomorphism, for why the two apparently "extreme" patterns,* 4132 *and* 3142*, should be equally restrictive?*

4.3. **Embeddings and isomorphic subtrees.** We can now take a look at the last proposed Wilf-equivalence in S_4: $(1234) \overset{?}{\sim} (4123)$. If we flip (4123) across a horizontal axis, we obtain $(1234) \overset{?}{\sim} (1432)$, where the last *three and largest* elements (234) are flipped to (432). This looks awfully like Theorem 7, and indeed it is true, but its proof does not come up until much later (cf. [3, 4]).

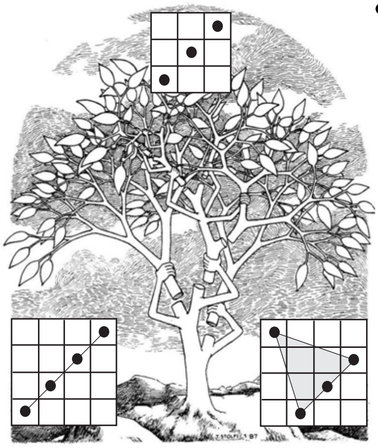
On the other hand, a survey of the trees $T(1234)$ and $T(4123)$, or any $T(\sigma)$ and $T(\tau)$ where σ and τ are in the corresponding symmetry classes, does not yield possible isomorphisms. One can be even more inventive and use other definitions for generating trees, e.g. instead of adding $n+1$ to nodes π of length n, first increase the largest entry n of π by 1 and then add n in appropriate places. Unfortunately, these tree variations also do *not* suggest any potential isomorphisms.

The classic method of generating trees, therefore, needed to be adjusted here for the purposes of the intended Wilf-equivalence:

Theorem 9 (Stankova'92)**.** $|S_n(1234)| = |S_n(4123)|$.

The proof of Theorem 9 in [23] can be roughly described by embedding infinitely many times the well-known subtree $T(123)$ inside the trees $T(1234)$ and $T(4123)$. Not surprisingly, this is done with the help of labellings:

- the standard labeling for $T(123)$ (as discussed earlier), assigning to every node α the length k of its initial decreasing subsequence;
- analogous labeling for $T(1234)$ (used already by West in [27]), assigning to every node β a pair of integers $\langle x, y\rangle$, where x is the length of the initial decreasing subsequence of β and y is the number of children of β;



- a new labeling for $T(4123)$ (created for this purpose in [23]). The *decreasing subsequence structure* of a permutation π is the sequence of lengths of the maximal consecutive subsequences in π. If $\gamma \in T(4123)$ contains a subsequence of type (123), $n+1$ cannot be added before it in γ, so we may disregard all numbers before such (123)-subsequences in γ, including the first number of the subsequence itself. If the resulting sequence is γ', and the permutation of type γ' is γ'', the label of γ is the decreasing subsequence structure of γ''. For example, $\gamma = (64\underline{1}7\underline{5}32\underline{8}) \mapsto \gamma' = (7532|8) \mapsto \gamma'' = (4321|5)$, and γ's label is $\langle 4, 1\rangle$.

The process of obtaining γ'' from γ in $T(4123)$ is called *reduction* of γ to γ''. The above is a *good* labeling because two nodes in $T(4123)$ have the same label if and only if they are *isomorphic* in $T(4123)$, in other words, they generate isomorphic subtrees and hence one can switch them without changing the overall tree. (This idea is reflected in an artistic way in the above picture.) Note that, in effect, the reduction replaces $\gamma \in T(4123)$ with an isomorphic node γ'' that avoids (123) and hence occurs in $T(123)$.

Analogous reduction processes are defined for the nodes of $T(1234)$ – the cut is done at the end of the permutations – and no reduction is necessary for $T(123)$; and analogous statements about isomorphism of nodes in $T(123)$ and in $T(1234)$ are established, involving (naturally) the decreasing sequence structure. For example, if two permutations in $T(123)$ have the same decreasing sequence structures, they are isomorphic in $T(123)$ (because they have the same length of initial decreasing subsequences, and hence same labels). Thus, $(132) \cong (231) = \langle 1\rangle$ and $(213) \cong (312) = \langle 2\rangle$ in $T(123)$.[5]

The last simple observation in $T(123)$ underpins the rest of the proof. If the numbers $k-1, k$ and $k+1$ are *not* in a monotone subsequence of $\alpha \in T(123)$, we allow switching the two occurring first and last in α, and denote this transformation by $\alpha \to \beta$. For instance, $(132) \to (231)$ and $(3214) \to (4213) \to (4312)$. It turns out that the two switching operations are well-defined on $T(123)$, and if $\alpha \to \beta$ in $T(123)$, then not only α and β are isomorphic nodes in $T(123)$, but also in $T(1234)$ and $T(4123)$.

The orbits of the switching operations (taken together) are *equivalence classes* of isomorphic nodes on every level of $T(123)$. A *partial ordering* within each equivalence class (that

[5] $(321) = \langle 3\rangle$ is not isomorphic to any other node in $S_3(123)$, and (123) itself is missing from $T(123)$.

carries over to the larger trees $T(1234)$ and $T(4123)$), is created by *orienting* the switches above: a *positive* switch to $\alpha \rightarrow^{+} \beta$ in $T(123)$ moves a larger element in front of a smaller one, e.g. $(3214) \rightarrow^{+} (4213) \rightarrow^{+} (4312)$, while $(231) \rightarrow^{-} (132)$ is a *negative* switch.

Let a *backtrack* be a number k that comes before $k+1$ in a permutation α. In $T(123)$, the set of *maximal* elements of length n with s backtracks is denoted by $\mathcal{C}_{n,s}$; note that $s \leq n/2$. For example, the backtracks in the non-maximal $\alpha = (543612)$ are four: 5, 4, 3 and 1, but maximizing α to $\beta = (643512)$ yields only three backtracks: 4, 3 and 1.

The sets $\mathcal{C}_{n,s}$ are the key to proving that $1234 \sim 4123$ and to enumerating $S_n(1234)$. Amazingly enough, we can calculate precisely $|\mathcal{C}_{n,s}|$ and represent the whole set $|S_n(123)|$ in a nice combinatorial pattern, determined by the $\mathcal{C}_{n,s}$–sets.[6] To this end, we extend the sequence of Catalan numbers to a triangular *Catalan table*: its entries are the numbers $c_{n,s}$ with with $c_{n-1,s} + c_{n-1,s-1} = c_{n,s}$, where the n-axis is the usual horizontal x–axis, while the s–axis makes a 135°–angle with the n-axis (compare Fig. 9 with Fig. 3b). In particular, $c_{2n,n} = c_n$ and $s \leq n/2$.

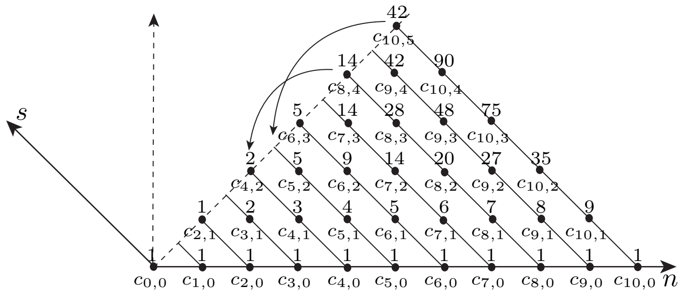


FIGURE 9. Catalan table

Lemma 1. *In $S_n(123)$, $|\mathcal{C}_{n,s}| = c_{n,s} = \binom{n}{s} - \binom{n}{s-1}$ for $s = 0, 1, ..., [n/2]$. Further, place all maximal elements in a row, reorder them within the sets $\mathcal{C}_{n,s}$ according to their backtrack number s, and underneath each maximal element list its equivalence class in a column. The resulting configuration will be a disjoint union of squares, one for each $\mathcal{C}_{n,s}$–set. In particular, the size of an equivalence class with maximal element of backtrack number s equals $|\mathcal{C}_{n,s}|$, and the total number of elements in $S_n(123)$ are $c_n = \sum_{s=0}^{[n/2]} c_{n,s}^2$.*

Figure 9 points to two examples: $c_4 = 14 = 1^2 + 3^2 + 2^2$ and $c_5 = 42 = 1^2 + 4^2 + 5^2$. To get the trees $T(1234)$ and $T(4123)$ into the picture, we investigate how their nodes on level n recursively relate to nodes on level $n+1$.

We say that a permutation in $T(1234)$ or $T(4123)$ is *full* or *irreducible* if it avoids (123), i.e. no reduction is necessary on it. On the other hand, a *reduced* permutation α'' in $T(1234)$ or $T(4123)$ is obtained from a permutation α which does indeed need reduction; α is called *reducible*. Note that a full permutation on level n (of whichever tree) has length n, while a reduced permutation has length k smaller than its level n.

[6] We changed the original notation in [23] from $\mathcal{B}_{n,s}$ and $b_{n,s}$ to $\mathcal{C}_{n,s}$ and $c_{n,s}$ to reflect the fact that we are dealing with the Catalan numbers.

Lemma 2. *Consider the set of* irreducible *permutations* $\mathcal{C}_{n,s}$ *in* $S_n(1234)$ (*or* $S_n(4123)$). *The set of all children of* $\mathcal{C}_{n,s}$, $\mathcal{C}'_{n,s}$, *after possible reduction and maximization* (*in* $T(123)$), *can be partitioned as follows:*

- *the* irreducible *children of* $\mathcal{C}_{n,s}$, *after possible maximization, form the disjoint union* $\mathcal{C}_{n+1,s} \sqcup \mathcal{C}_{n+1,s+1}$;
- *the* reducible *children of* $\mathcal{C}_{n,s}$, $\mathcal{C}^r_{n,s}$, *can be thought of as the disjoint union of* $\mathcal{C}_{n,s}$ *and all* reducible *children of* $\mathcal{C}_{n-1,s}$ *and of* $\mathcal{C}_{n-1,s-1}$.

Lemma 2 partitions all children $\mathcal{C}'_{n,s}$ (for $s \leq \lfloor n/2 \rfloor$) as follows:

$$\begin{aligned} \mathcal{C}'_{n,s} &\cong \left(\mathcal{C}_{n+1,s} \sqcup \mathcal{C}_{n+1,s+1}\right) \sqcup \mathcal{C}^r_{n,s}, \\ \mathcal{C}^r_{n,s} &\overset{s>0}{\cong} \mathcal{C}_{n,s} \sqcup \left(\mathcal{C}^r_{n-1,s} \sqcup \mathcal{C}^r_{n-1,s-1}.\right) \end{aligned}$$

Note that $\mathcal{C}^r_{n,0} = \emptyset$ for $s = 0$ since $\mathcal{C}_{n,0} = \{J_n = (n, n-1, \ldots, 1)\}$ has only irreducible children ($n+1$ in number, to be exact). Further, $\mathcal{C}_{n,s} = \mathcal{C}^r_{n,s} = \emptyset$ when $s > \lfloor n/2 \rfloor$.

To finish off the proof that $|S_n(1234)| = |S_n(4123)|$, we reduce and maximize all of the permutations in both trees, partition them canonically into $\mathcal{C}_{k,l}$–sets and compare. Let us see how this works on level $n = 3$. In the case of (1234) we have $S_3(1234) = S_3$, only one permutation requires reduction: $(12|3) \mapsto (12)$, and two permutations require maximizing: $(132) \to^+ (231)$ and $(213) \to^+ (312)$. The partitioning is as follows:

$$(3) \qquad S_3(1234) = \left| \begin{matrix} 231\ 132 \\ 132\ 213 \end{matrix} \right| \, 321 \, \left| \, 12 \, \right| = 2\mathcal{C}_{3,1} \sqcup \mathcal{C}_{3,0} \sqcup \mathcal{C}_{2,1}.$$

The coefficient 2 in front of $\mathcal{C}_{3,1}$ signifies that two (disjoint) copies of the set appear in the partition. The only difference for $S_3(4123)$ is in the unique reduction $(1|23) \mapsto (23) \mapsto (12)$, which leads anyways to an identical partition of $S_3(4123)$. Applying Lemma 2 to each term in the above decomposition yields:

$$(4) \qquad S_4(1234) \cong 2\mathcal{C}_{4,2} \sqcup 3\mathcal{C}_{4,1} \sqcup \mathcal{C}_{4,0} \sqcup 3\mathcal{C}_{3,1} \sqcup 3\mathcal{C}_{2,1} \cong S_4(4123).$$

Applying Lemma 2 once again, the reader can verify that:

$$S_5(1234) \cong 5\mathcal{C}_{5,2} \sqcup 4\mathcal{C}_{5,1} \sqcup \mathcal{C}_{5,0} \sqcup 5\mathcal{C}_{4,2} \sqcup 6\mathcal{C}_{4,1} \sqcup 11\mathcal{C}_{3,1} \sqcup 11\mathcal{C}_{2,1} \cong S_5(4123).$$

For completeness, note that $S_1(1234) = S_1(4123) = \{(1)\} = \mathcal{C}_{1,0}$, and $S_2(1234) = S_2(4123) = \{(12), (21)\} = \mathcal{C}_{2,1} \sqcup \mathcal{C}_{2,0}$, and check that Lemma 2 recursively transforms the decompositions from level 1 to level 2 to level 3.

By induction, each level of the two trees $T(1234)$ and $T(4123)$ can be partitioned into the same sets $\mathcal{C}_{k,l}$, hence $|S_n(1234)| = |S_n(4123)|$ for all n. □

An outline of the above discussion and a proof was included in [23], but the calculations and detailed examples are done here for the first time. Meanwhile, the Wilf–classification of S_4 is completed: three distinct Wilf-classes have emerged, with representatives (1324), (1234) and (4132).

4.4. **Enumerations on trees.** Ever since the initial calculation $|S_n(\tau)| = c_n$ for all $\tau \in S_3$, the question of enumerating $S_n(\tau)$–sets for longer permutations τ has fascinated reseachers; yet, it has been resolved only in ... 2 cases in S_4. In [?] Ira Gessel uses techniques outside of the scope of this article to enumerate $S_n(1234)$, and later to reduce his formula to

$$S_n(1234) = \frac{1}{(n+1)^2(n+2)} \sum_{k=0}^{n} \binom{2k}{k} \binom{n+1}{k+1} \binom{n+2}{k+1}. \tag{5}$$

As Bóna remarks in his book [7, ?], even though "all terms are non-negative, there is still a division, suggesting that a direct combinatorial proof is probably difficult to find." However, our previous detailed analysis of $T(1234)$ does yield an *explicit recursive algorithm* for moving from $S_n(1234)$ to $S_{n+1}(1234)$, and hence a possible direct combinatorial route of proving formula (5).

To enumerate $|S_n(1234)|$, we need to find the precise canonical decomposition of $S_n(1234)$ into $\mathcal{C}_{k,l}$-sets. So, let

$$S_n(1234) \cong \bigsqcup_{(k,l)} x^n_{k,l} \mathcal{C}_{k,l},$$

where $x^n_{k,l}$ stands for the number of copies of $\mathcal{C}_{k,l}$ appearing in the decomposition of $S_n(1234)$, $k \le n$ and $l \le k/2$. For easy visualization, Figure 10 displays all cases for $n = 1, 2, \ldots, 7$, the first five of which were calculated earlier. The coefficients $x^n_{k,l}$ have been placed in the positions of the generalized Catalan numbers $c_{k,l}$:

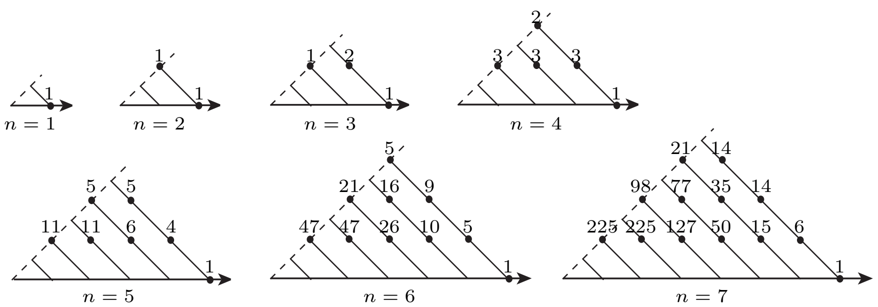


FIGURE 10. Canonical decomposition of $S_n(1234)$ for $n = 1, 2, \ldots, 7$

Let $f(n,s)$, $f^i(n,s)$ and $f^r(n,s)$ be the number of all, irreducible and reducible children of $\mathcal{C}_{n,s}$, respectively. From Lemma 2, $f(n,s) = f^i(n,s) + f^r(n,s)$, and $f^i(n,s) = c_{n+1,s} + c_{n+1,s+1} = c_{n+2,s+1}$, so we can eliminate $f^i(n,s)$:

$$\begin{aligned} f(n,s) &= c_{n+2,s+1} + f^r(n,s) \\ f^r(n,s) &= c_{n,s} + f^r(n-1,s) + f^r(n-1,s-1) \end{aligned}$$

with initial conditions $f(n,s) = f^r(n,s) = 0$ for $s > \lfloor n/2 \rfloor$ and $f^r(n,0) = 0$. One can further eliminate all $f^r(k,l)$ by plugging the first formula into the second one, paying special attention to the boundary cases of $k = 2l$ and $l = 0$:

$$f(n,s) = f(n-1,s) + f(n-1,s-1) + c_{n,s}. \tag{6}$$

From here, one can show by induction an unexpected simplification: the total number of children of each set $\mathcal{C}_{n,s}$ is given by

$$|\mathcal{C}'_{n,s}| = f(n,s) = (n+1)c_{n,s}. \tag{7}$$

One can use this formula for an alternative way of calculating $|S_{n+1}(1234)|$, knowing only the decomposition of $S_n(1234)$. For example, from the decomposition $S_3(1234) \cong 2\mathcal{C}_{3,1} \sqcup \mathcal{C}_{3,0} \sqcup \mathcal{C}_{2,1}$ in (3):

$$|S_4(1234)| = 2f(3,1) + f(3,0) + f(2,1) = 2 \cdot 4c_{3,1} + 4c_{3,0} + 3c_{2,1} = 23.$$

This calculation matches what we would obtain directly from the decomposition of $S_4(1234)$ in (4):

$$|S_4(1234)| = 2c_{4,2} + 3c_{4,1} + c_{4,0} + 3c_{3,1} + 3c_{2,1} = 23.$$

Of course, we know that 23 is the correct answer, as there is only 1 in 24 permutations of length 4 which doesn't avoid (1234), namely, (1234) itself. In general,

$$\sum_{(k,l)} x^{n+1}_{k,l} c_{k,l} = |S_{n+1}(1234)| = |S'_n(1234)| = \sum_{(k,l)} (k+1) x^n_{k,l} c_{k,l}. \tag{8}$$

Unfortunately, this type of reasoning does not carry us very far, since we still need to know the decomposition of the previous set $S_n(1234)$. We can determine, though, the precise decomposition of the set $\mathcal{C}'_{k,l}$ of children of $\mathcal{C}_{k,l}$.

Definition 8. Let (k,l) and (p,q) be two lattice points in the Catalan triangle. We say that $(p,q) \leq (k,l)$ if (p,q) is situated southwest from (k,l) in the usual Cartesian sense (possibly coinciding with (k,l)); i.e. if we draw a vertical line down from (k,l) and a horizontal line to the left from (k,l), (p,q) lies in the resulting right trapezoid. In such a case, we denote the number of Catalan paths from (p,q) to (k,l) (consisting of unit segments to the right or up) by $P^{k,l}_{p,q}$.

Lemma 3. *For $l = 0$, $\mathcal{C}'_{k,0} \cong \mathcal{C}_{k+1,0} \sqcup \mathcal{C}_{k+1,1}$. For $l > 0$:*

$$\mathcal{C}'_{k,l} \cong \mathcal{C}_{k+1,l} \sqcup \mathcal{C}_{k+1,l+1} \bigsqcup_{0<q,(p,q)\leq(k,l)} P^{k,l}_{p,q} \cdot \mathcal{C}_{p,q}. \tag{9}$$

The distinction between the cases $l > 0$ and $l = 0$ is the reason for the zero coefficients $x^n_{k,0}$ for $k = 1, 2, \dots n-1$ (cf. Fig. 10). The disjoint union over (p, q) in (9) can be interpreted combinatorially as the Catalan paths from $(0, 0)$ to (k, l), counted once for each of the points (p, q) along them with $q > 0$, i.e. each counted exactly k times. The fact that we don't count the points $(p, 0)$ along these paths is compensated "perfectly" by the two initial extra components $\mathcal{C}_{k+1,l} \sqcup \mathcal{C}_{k+1,l+1}$, so that in summary,

Lemma 4. *The children of $\mathcal{C}_{k,l}$ are in 1–1 correspondence with pairs $(\mathcal{P}, (p, q))$ of Catalan paths $\mathcal{P}$ from $(0, 0)$ to (k, l) and a point (p, q) on $\mathcal{P}$; that is, each path is counted exactly $k + 1$ times. Consequently, we again obtain that $f(k, l) = (k + 1)c_{k,l}$.*

As an example, let's decompose the set of children of $\mathcal{C}_{5,2}$:

$$\mathcal{C}'_{5,2} \cong \mathcal{C}_{6,2} \cup \mathcal{C}_{6,3} \cup \mathcal{C}_{5,2} \cup \mathcal{C}_{4,2} \cup 2\mathcal{C}_{3,1} \cup 2\mathcal{C}_{2,1}.$$

The coefficients 2 in front of $\mathcal{C}_{3,1}$ and $\mathcal{C}_{2,1}$ represent the two possible paths from $(3, 1)$ to $(5, 2)$, and from $(2, 1)$ to $(5, 2)$. Taking moduli everywhere, we obtain

$$f(5, 2) = c_{6,2} + c_{6,3} + c_{5,2} + c_{4,2} + 2c_{3,1} + 2c_{2,1} = 30 = 6 \cdot c_{5,2}.$$

Thus, starting with the decomposition $S_n(1234) \cong \bigsqcup x^n_{k,l}\mathcal{C}_{k,l}$, one can apply Lemma 3 to each component $\mathcal{C}_{k,l}$ to compute the decomposition $S_{n+1}(1234) = \bigsqcup x^{n+1}_{k,l}\mathcal{C}_{k,l}$, and then finally to compute the size $|S_{n+1}(1234)| = \sum x^{n+1}_{k,l} c_{k,l}$.

However, the inductive process $S_n(1234) \to S_{n+1}(1234)$ is far from trivial from a computational point of view. One can attempt different grouping and counting strategies, in order to avoid finding the exact decomposition of $S_{n+1}(1234)$, but still be able to compute the size $|S_{n+1}(1234)|$. The fact that we managed to compute $f(n, s)$ before we knew the exact decomposition of $\mathcal{C}'_{k,l}$ is an indication that the above may be possible. For instance, we can group the components $\mathcal{C}_{k,l}$ of $S_n(1234)$ according to the diagonal they lie on (i.e. k is fixed), and calculate the contribution $\sum_l x^n_{k,l}c_{k,l}$ of this diagonal towards the total size $|S_n(1234)|$.

The last diagonal $k = n$ is fairly easy to determine: the coefficients there coincide with the generalized Catalan numbers placed in these positions, e.g. for $n = 6$, the numbers 1, 5, 9 and 5 appear in the same positions in Figure 10 and in Figure 9, and their contribution towards $|S_6(1234)|$ is $1^2 + 5^2 + 9^2 + 5^2 = 132 = c_6$.

Lemma 5. *In $S_n(1234)$, $x^n_{n,l} = c_{n,l}$ for all l, so that the contribution of the n^{th}–diagonal towards $|S_n(1234)|$ is*

$$\sum_{l=0}^{n} x^n_{n,l}|\mathcal{C}_{n,l}| = \sum_{l=0}^{n} c^2_{n,l} = c_{2n,n} = c_n.$$

The individual coeffiecients $x^{n-1}_{k,l}$ on the $(n-1)^{\text{st}}$–diagonal are not so obvious but their total contribution towards $|S_n(1234)|$ can be found. For instance, observe for $n = 6$ that

$x^6_{5,2} \cdot c_{5,2} + x^6_{5,1} \cdot c_{5,1} = 16 \cdot 5 + 10 \cdot 4 = 120 = 6 \cdot 42 - 132 = 6 \cdot c_5 - c_6$. In fact, equating the two ways of calculating $|S_{n+1}(1234)|$ in (8) and appropriate cancellation leads to

Lemma 6. *The contribution of the $(n-1)^{st}$–diagonal towards $|S_n(1234)|$ is $n \cdot c_{n-1} - c_n$.*

A natural question arises:

Question 2. *Is the contribution of the $(n-k)^{th}$–diagonal towards $|S_n(1234)|$ expressible as a linear combination of c_n, c_{n-1}, $\ldots$, c_2, c_1 whose coefficients are some convenient functions of n and k?*

If the answer is yes, then combining all diagonals' contributions yields

$$|S_n(1234)| = x_n c_n + x_{n-1} c_{n-1} + \cdots + x_1 c_1$$

for some convenient x_k. Using the c_n's as a basis of such a linear combination is not surprising. Gessel's formula (5) already does this: just extract all $c_k = \frac{1}{k+1}\binom{2k}{k}$. For instance, the coefficients of c_n and c_{n-1} are $1/(n+1)$ and $n/2$, respectively. The interesting question is how to *combinatorially* find these coefficients using the above analysis of $S_n(1234)$. To this end, we need a combinatorial description of all components $\mathcal{C}_{k,l}$ in $S_n(1234)$'s decomposition, along with their multiplicities $x^n_{k,l}$.

Definition 9. Fix some integers n and $k \in [0, n]$. Consider any path $\mathcal{P}$ in the Catalan lattice, with a choice of (not necessarily distinct) points $\{P_0, P_1, \ldots, P_{k-1}\}$ on it and above the 0^{th} row, so that $\mathcal{P}$ consists of the following two parts:

- The first part $\mathcal{P}_1$ starts at $(0, 0)$, contains all points P_j, ends at some point P_k on the $(n-k)^{\text{th}}$–diagonal, and consists only of unit segments to the right or up, except for those unit segments to the left or down that connect two *consecutive* points P_j and P_{j+1}.
- The second part $\mathcal{P}_2$ picks up where $\mathcal{P}_1$ ended on $(n-k)^{\text{th}}$–diagonal and ends at $(2k, k)$, and consists only of unit segments to the right or up.

If P_j does not belong to the $(n-j)^{\text{th}}$–diagonal for $j \le k-1$, we say that the $(k+1)$–tuple $(\mathcal{P}, P_0, P_1, \ldots, P_{k-1})$ is a *full path of order* k, while the $(k+1)$–tuple $(\mathcal{P}_1, P_0, P_1, \ldots, P_{k-1})$ is a *partial path of order* k.

The paths of order 0 are simply all Catalan paths from $(0, 0)$ to $(2n, n)$, c_n in number For the paths $(\mathcal{P}, P_0)$ of order 1, we know that P_0 cannot be on the n^{th}–diagonal while P_1 must be on the $(n-1)^{\text{st}}$–diagonal; hence, $P_0 \le P_1$ and these paths are the usual Catalan paths from $(0, 0)$ to $(2n-2, n-1)$ with an extra choice of a point P_0 before or on $(n-1)^{\text{st}}$–diagonal, but not on the 0^{th}–row. A quick count similar to what we had earlier, shows that these paths are exactly $nc_{n-1} - c_n$ in number. The paths of order $(n-k)$ for $k \ge 2$ involve the occasional unit backtrack moves to the left or down from P_j to P_{j+1}, and their counting is considerably more complex.

Proposition 1. *The coefficient* $x^n_{k,l}$ *counts the number of partial paths* $\mathcal{P}_1$ *of order* k *which end on* $P_k = (k, l)$*, while the term* $x^n_{k,l}\mathcal{C}_{k,l}$ *corresponds to the full paths* $\mathcal{P}$ *of order* k *passing through* $P_k = (k, l)$*. Consequently, the contribution of the* $(n-k)^{th}$*–diagonal towards* $|S_n(1234)|$ *is the number of full paths of order* k*, while* $|S_n(1234)|$ *itself is the number of full paths of any order* $k = 0, 1, \dots, n-1$.

The only other exact formula for a pattern longer than 3 is for (4132) by Bóna in [7]:

$$S_n(4132) = (-1)^{n-1}\frac{7n^2-3n-2}{2} + \sum_{i=2}^{n}(-1)^{n-i}2^{i+1}\frac{(2i-4)!}{i!(i-2)!}\binom{n-i+2}{2}.$$

Question 3. *Could the labeling of* $T(4132)$ *or of* $T(3142)$ *be used to directly enumerate* $S_n(4132)$*?*

Question 4. *Could the methods in* [7] *or in* [25] *be used to enumerate* $S_n(1324)$*? No formula is known about this sequence.*

4.5. **Shape-Wilf Equivalence.** Now let's move to Wilf-classification of S_5 and beyond. There are no sporadic cases, and all Wilf-equivalencies fits nicely into Babson-West's Theorem about flipping the last k largest elements... However, the proof of the Theorem itself requires the new stronger notion of *shape–Wilf–equivalent*, which utilizes Young diagrams instead of square matrices and replaces the classic Wilf-equivalence.

4.6. S_8 **and beyond: are computer capabilities the only obstacle?** A month-long calculations yielded no new Wilf-equivalences in S_8.

Question 5. *How long will it take to convince ourselves that* S_9 *and beyond are already classified up to Wilf-equivalence? Are we dependent here only on the speed of computers, or is there an underlying reason for the apparent lack of further new Wilf-equivalences on longer patterns?*

Conclusion: of all questions on restricted patterns, the classification up to Wilf-equivalence still carries the flag of the hardest question, whose answer is not likely to come up soon.

5. On the Stanley-Wilf Limits

That the existence of the limit $L(\tau)$ implies the exponential bound c^n_τ for $|S_n(\tau)|$ is an obvious exercise in limits. However, the opposite implication is not immediate, and it was first published by Arratia [2]. We present its proof here for the benefit of the reader, as it demonstrates how properties of restricted patterns imply the (roughly) exponential growth of $|S_n(\tau)|$, and the congergence of $\sqrt[n]{|S_n(\tau)|}$.

Lemma 7. *For any pattern* τ *and any* $m, n \geq 1$, $|S_m(\tau)| \cdot |S_n(\tau)| \leq |S_{m+n}(\tau)|$.

Proof. We construct an injective map $S_m(\tau) \times S_n(\tau) \hookrightarrow S_{m+n}(\tau)$ by creating an element $\gamma \in S_{n+m}(\tau)$ from permutations $\alpha \in S_m(\tau)$ and $\beta \in S_n(\tau)$. To this end, place the permutation matrices $M(\alpha)$ and $M(\beta)$ along the diagonal of a larger empty $(n+m) \times (n+m)$ matrix, to arrive at a permutation matrix $M(\gamma)$. The permutation γ itself is in effect obtained by adding n to each entry of α and then concatinating the result with β. For instance, if $\tau = (4213)$, $\alpha = (231)$ and $\beta = (2134)$, then $\gamma = (6752134)$ (cf. Fig 11a). Unfortunately, $\gamma \notin S_7(\tau)$, as γ contains the subsequence $(5213) \approx \tau$.

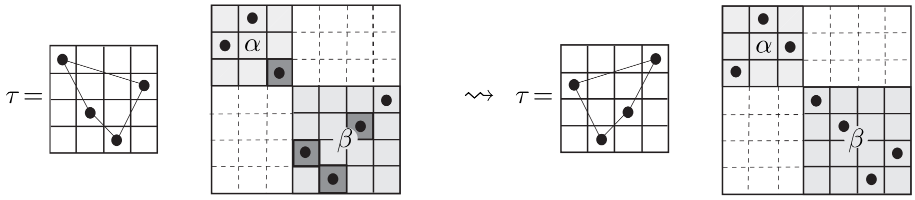


FIGURE 11. $(\alpha, \beta) \in S_n(\tau) \times S_n(\tau) \hookrightarrow S_{m+n}(\tau)$

There is an easy fix for this. From the very beginning we can assume that the smallest element of τ *preceeds* the largest one: if not, simply flip $M(\gamma)$ across a vertical axis, and along with it, the matrices corresponding to anything in $S_m(\tau)$, $S_n(\tau)$ and $S_{n+m}(\tau)$. As we have seen before, this operation does not change the cardinalities of the involved sets, but simply replaces τ by another permutation in its symmetry class.

In our example, we will have instead $\tau = (3124)$, $\alpha = (132)$, $\beta = (4312)$, and the newly created $\gamma = (5764312)$ will indeed avoid τ (cf. Fig 11b). In general, suppose that γ contains a subsequence $\delta \approx \tau$, i.e. $M(\gamma)$ contains the submatrix $M(\delta) \approx M(\tau)$. Since $M(\gamma)$ is made of two pieces $M(\alpha)$ and $M(\beta)$, each of which avoids $M(\tau)$, this can happen only if $M(\delta)$ involves elements from both $M(\alpha)$ and $M(\beta)$. But since all of $M(\alpha)$ comes before and is higher than all of $M(\beta)$, this means that the largest element A and smallest element B of $M(\delta)$ must come from $M(\alpha)$ and $M(\beta)$, respectively. In other words, A comes before B in $\delta \approx \tau$, thereby contradicting our assumption about τ in the previous paragraph.

Thus, $\gamma \in S_{n+m}(\tau)$, and our construction indeed induces an injective map $S_m(\tau) \times S_n(\tau) \hookrightarrow S_{m+n}(\tau)$. □

Such maps between pattern-avoiding sets, and especially the "concatination" of the two permutation matrices $M(\alpha)$ and $M(\beta)$, are typical constructions in the field of restricted patterns. Now, one more ingredient is necessary to imply the existence of the Stanley-Wilf limit $L(\tau)$. Recall

Lemma 8 (Fekete). *If $\{a_n\}$ is a superadditive sequence, i.e. $a_n + a_m \leq a_{n+m}$ for all $n, m \geq 1$, then $\lim_{n\to\infty} \frac{a_n}{n}$ exists and equals $\sup \frac{a_n}{n}$. (The limit may be $+\infty$.)*

Unless $\tau = (1)$ and clearly $L(\tau) = 0$, all $|S_n(\tau)| > 0$ so that we can apply Fekete's lemma (cf. [26]) to the superadditive sequence $a_n = \ln |S_n(\tau)|$ to derive that $\lim_{n\to\infty} \ln \sqrt[n]{|S_n(\tau)|}$ exists. Since $\sqrt[n]{|S_n(\tau)|}$ is bounded above by the constant c_τ, this limit is finite. We conclude that $L(\tau) = \sup \sqrt[n]{|S_n(\tau)|} \leq c_\tau$, thereby completing Arratia's equivalence argument.

The proof of the Stanley-Wilf conjecture is a different matter altogether. In short, one needs to define first a generalization of permutation matrices, the so-called $0-1$ matrices whose entries are simply 0's and 1's without any further conditions. Containment of a $0-1$ matrix A in a $0-1$ matrix M is given by a submatrix B of M such that B has the same size as A, and B has 1's at least in the cells where A has 1's. A conjecture of Füredi-Hajnal proven by Marcus and Tardos in [16] sounds very similar to Stanley-Wilf conjecture:

Theorem 10 (Marcus-Tardos)**.** *If A is a permutation matrix, let $f(n, A)$ be the maximum number of* 1*'s that can serve as entries of an A-avoiding $0-1$ matrix M of size $n \times n$. Then $f(n, A)$ is bounded above by an exponential function, i.e. there is some constant c_A such that $f(n, A) \leq c_A^n$ for all $n \geq 1$.*

Question: Is Klazar's use of bipartite containments in any relation to Jelinek's construction, and hence to shape-Wilf equivalences?

References

[1] M. H. Albert, M. Elder, A. Rechnitzer, P. Westcott, M. Zabrocki, On the Wilf-Stanley limit of 4321-avoiding permutations and a conjecture of Arratia, *Adv. in Appl. Math. 36* (2006), no. 2, 96-105.

[2] R. Arratia, On the Stanley-Wilf Conjecture for the Number of Permutations Avoiding a Given Pattern, *Electron. J. Combin. 6* (1999), no. 1, #1.

[3] E. Babson, J. West, The permutations $123p_4...p_t$ and $321p_4...p_t$ are Wilf-equivalent, *Graphs Comb. 16* (2000) 4, 373-380.

[4] J. Backelin, J. West, G. Xin, Wilf-equivalence for singleton classes, *Proceedings of the 13th Conference on Formal Power Series and Algebraic Combinatorics*, Tempe, AZ, 2001.

[5] M. Bóna, Permutations avoiding certain patterns. The case of length 4 and some generalizations, *Disc. Math. 175* (1997) 55-67.

[6] M. Bóna, The Solution of a Conjecture of Wilf and Stanley for all layered patterns, *J. Combin. Theory, Series A, 85* (1999) 96-104.

[7] M. Bóna, *Combinatorics of Permutations*, Chapman & Hall/CRC, 2004, 135-159.

[8] M. Bóna, The Limit of a Stanley-Wilf sequence is not always rational, and layered patterns beat monotone patterns, *J. Combin. Theory Ser. A 110 (2)* (2005), 223-235.

[9] M. Bóna, "New Records in Stanley-Wilf Limits", submitted to *Electron. J. Combin.*, 2006.

[10] F. Chung, , R. Graham, V. Hoggatt, and M. Kleiman, The number of Baxter permutations, *J. Combin. Theory Ser. A 24, 3* (1978), 382-394.

[11] P. Erdös and G. Szekeres, A combinatorial problem in geometry, *Compos. Math. 2* (1935), 463-470.

[12] I. Gessel, Symmetric functions and P-recursiveness, *J. Combin. Theory Ser. A, 53* (1990), no. 2, 257-285.

[13] V. Jelínek, Dyck paths and pattern-avoiding matchings, *Europ. J. Combin. 28* (2007), no. 1, 202-213.

[14] D. Knuth, *The Art of Computer Programming*, Vol.3, Addison-Wesley, Reading, MA, 1973.

[15] D. Knuth, Permutations, matrices, and generalized Young tableaux, *Pacific J. of Math. 34* (1970) 709-727.

[16] A. Marcus and J. Tardos, Excluded Permutation Matrices and the Stanley-Wilf conjecture, *J. Combin. Theory Ser. A 107* (1) (2004), 153-160.

[17] A. Mier, k-noncrossing and k-nonnesting graphs and fillings of Ferrers diagrams, to appear in Combinatorica.

[18] A. Regev, Asymptotic values for degrees associated with strips of Young diagrams, *Advances in Mathematics 41* (1981), 115-136.

[19] D. Richards, Ballot sequences and restricted permutations, *Ars Combinatoria 25* (1988) 83-86.

[20] D. Rotem, On correspondence between binary trees and a certain type of permutation, *Information Processing Letters 4* (1975), 58-61.

[21] R. Simion and F. Schmidt, Restricted permutations, *Europ. J. Combin. 6* (1985) 383-406.

[22] Z. Stankova, Forbidden subsequences, *Disc. Math. 132* (1994) 291-316.

[23] Z. Stankova, Classification of forbidden subsequences of length 4, *Europ. J. Combin. 17* (1996), 501-517.

[24] Z. Stankova and J. West, A new class of Wilf-equivalent permutations, *J. Algebraic Combin. 15* (2002), no. 3, 271-290.

[25] Z. Stankova, Shape-Wilf-Ordering on Permutations of Length 3, *Electron. J. Combin. 14 (2007), #R56.*

[26] J. Steele, *Probability theory and combinatorial optimization*, *SIAM*, Philadelphia (1997).

[27] J. West, Permutations with forbidden subsequences and stack-sortable permutations, Ph.D. Thesis, M.I.T., Cambridge, MA, 1990.

[28] J. West, Generating trees and the Catalan and Schröder numbers, *Disc. Math. 146* (1995) 247-262.

[29] J. West, Generating trees and forbidden subsequences, *Disc. Math. 157* (1996) 363-374.