Shape-Wilf-Ordering on Permutations of Length 3

Zvezdelina Stankova

Dept. of Mathematics and Computer Science
Mills College, Oakland, California, USA

stankova@mills.edu

Submitted: Sep 19, 2006; Accepted: Aug 9, 2007; Published: Aug 20, 2007
Mathematics Subject Classification: 05A20

Abstract

The research on pattern-avoidance has yielded so far limited knowledge on Wilf-ordering of permutations. The Stanley-Wilf limits $\lim_{n\to\infty} \sqrt[n]{|S_n(\tau)|}$ and further works suggest asymptotic ordering of layered versus monotone patterns. Yet, Bóna has provided essentially the only known up to now result of its type on complete ordering of S_k for $k=4$: $|S_n(1342)| < |S_n(1234)| < |S_n(1324)|$ for $n \geq 7$, along with some other sporadic examples in Wilf-ordering. We give a different proof of this result by ordering S_3 up to the stronger *shape-Wilf-order*: $|S_Y(213)| \leq |S_Y(123)| \leq |S_Y(312)|$ for any Young diagram Y, derive as a consequence that $|S_Y(k+2,k+1,k+3,\tau)| \leq |S_Y(k+1,k+2,k+3,\tau)| \leq |S_Y(k+3,k+1,k+2,\tau)|$ for any $\tau \in S_k$, and find out when equalities are obtained. (In particular, for specific Y's we find out that $|S_Y(123)| = |S_Y(312)|$ coincide with every other Fibonacci term.) This strengthens and generalizes Bóna's result to arbitrary length permutations. While all length-3 permutations have been shown in numerous ways to be Wilf-equivalent, the current paper *distinguishes* between and orders these permutations by employing all *Young diagrams*. This opens up the question of whether shape-Wilf-ordering of permutations, or some generalization of it, is not the "true" way of approaching pattern-avoidance ordering.

1 Introduction

We review first basic concepts and results that are crucial to the present paper, and direct the reader to [18, 19, 22, 23] for further introductory definitions and examples on pattern-avoidance.

A permutation τ of length k is written as $(a_1, a_2, \ldots, a_k)$ where $\tau(i) = a_i$, $1 \leq i \leq k$. For $k < 10$ we suppress the commas without causing confusion. As usual, S_n denotes the symmetric group on $[n] = \{1, 2, ..., n\}$.

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 $\big(\pi(i_{\tau(1)}), \pi(i_{\tau(2)}),..., \pi(i_{\tau(k)})\big) \approx \tau$.

The following reformulation in terms of matrices is probably more insightful. In it, and throughout the paper, we coordinatize all matrices from *the bottom left corner* in order to keep the resemblance with the "shape" of permutations.

Definition 2. Let $\pi \in S_n$. The *permutation matrix* $M(\pi)$ is the $n \times n$ matrix M_n having a 1 in position $(i, \pi(i))$ for $1 \leq i \leq n$. Given two permutation matrices M and N, we say that M *avoids* N if no submatrix of M is identical to N.

A permutation matrix is simply an arrangement, called a *transversal*, of n non-attacking rooks on an $n \times n$ board. We refer to the elements of a transversal also as "1's" and "dots". Clearly, a permutation $\pi \in S_n$ contains a subsequence $\tau \in S_k$ if and only if $M(\pi)$ contains $M(\tau)$ as a submatrix. Thus, from the viewpoint of pattern avoidance, permutations and permutation matrices are interchangeable notions.

Definition 3. Let $S_n(\tau)$ denote the set of τ-avoiding permutations in S_n. Two permutations τ and σ are *Wilf-equivalent*, denoted by $\tau \sim \sigma$, if they are equally restrictive: $|S_n(\tau)| = |S_n(\sigma)|$ for all $n \in \mathbb{N}$. 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 $\tau \preceq \sigma$.

The classification of permutations in S_k for $k \geq 7$ up to Wilf-equivalence was completed over the last two decades by a number of people. We refer the reader to Simion-Schmidt [18], Rotem [17], Richards [16], and Knuth [11, 12] for length $k = 3$; to West [23] and Stankova [19, 20] for $k = 4$; to Babson-West [3] for $k = 5$; and to Backelin-West-Xin [4] and Stankova-West [21] for $k = 6, 7$.

However, total Wilf-*ordering* does not exist for a general S_k. The first counterexample occurs in S_5 (cf. [21]): if $\tau = (53241)$ and $\sigma = (43251)$, then $S_7(\tau) < S_7(\sigma)$ but $S_{13}(\tau) > S_{13}(\sigma)$, and hence τ and σ cannot be Wilf-ordered. This phenomenon prompts

Definition 4. 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 $n \gg 1$.

Stanley-Wilf Theorem (cf. Marcus and Tardos [13], Arratia [2]) gives some insight into the asymptotic ordering of permutations. Inequalities between the Stanley-Wilf limits $L(\tau) = \lim_{n\to\infty} \sqrt[n]{|S_n(\tau)|}$ suggest asymptotic comparisons between the corresponding permutations. For instance, works of Bóna [6, 8] and Regev [15] show that $L(I_k) = (k-1)^2 \leq L(\tau)$, where $I_k = (12...k)$ is the *identity* pattern and τ is any *layered* pattern in S_k (cf. Definition 7), providing strong evidence that the identity pattern is more restrictive than all layered patterns in S_k. Yet, this result will not imply asymptotic ordering between the above types of patterns if it happens that $L(I_k) = L(\tau)$ for some layered τ.

In [5, 7], Bóna provides essentially the only known so far result on complete Wilf-ordering of S_k for $k = 4$:

$$|S_n(1342)| < |S_n(1234)| < |S_n(1324)| \text{ for } n \geq 7, \tag{1}$$

along with some sporadic examples in asymptotic Wilf-ordering, e.g. $I_k \preceq_a \tau_k$ for certain $\tau_k \in S_k$ and others examples (cf. Exer. 4.25 in [7] and [9]). Since S_2 and S_3 are each a single Wilf-equivalence class (cf. [18]), the first possibility of nontrivial Wilf-ordering arises in S_4. A representative of each of the 3 Wilf-equivalence classes in S_4 appears in (1) (cf. [23, 19, 20].).

In order to prove differently and extend result (1) to Wilf-ordering of certain permutations of arbitrary lengths, we shall use the concept of a stronger Wilf-equivalence relation, called *shape-Wilf-equivalence*. The latter was introduced in [3], and further explored in consequent papers [4, 21].

Definition 5. A *transversal* T of a Young diagram Y, denoted $T \in S_Y$, is an arrangement of 1's such that every row and every column of Y has exactly one 1 in it. A subset of 1's in T forms a *submatrix* of Y if all columns and rows of Y passing through these 1's intersect *inside* Y. For a permutation $\tau \in S_k$, T *contains the pattern* τ (in Y) if some k 1's of T form a submatrix of Y identical to $M(\tau)$. Denote by $S_Y(\tau)$ the set of all transversals of Y which avoid τ.

Now, suppose $T \in S_Y$ has a *subsequence* $\mathcal{L} = (\alpha_1\alpha_2...\alpha_k) \approx \tau \in S_k$. From the above definition, in order for $T \in S_Y$ to *contain the pattern* τ in Y, it is necessary and sufficient that the column of the rightmost element of $\mathcal{L}$ and the row of the smallest element of $\mathcal{L}$ intersect inside Y. In such a case, we say that the subsequence $\mathcal{L}$ *lands* inside Y. For example, Figure 1a shows the transversal $T \in S_Y$ representing the permutation (51324). Note that T contains the patterns (312) and (321) because its subsequences (513) and (532) land inside Y. However, T's subsequence $(324) \approx (213)$ does not land in Y, and in fact, T does not contain the *pattern* (213); symbolically, $T \in S_Y(213)$.

When Y is a square diagram of size n, $S_n(\tau) \equiv S_Y(\tau)$. Let $Y(a_1, a_2, ..., a_n)$ denote the Young diagram Y whose i-th row has a_i cells, for $1 \leq i \leq n$. In order for Y to have any transversals at all, it must be *proper*: Y must have the same number of rows and columns and must contain the *staircase* diagram $St^1 = Y(n, n-1, ..., 2, 1)$; equivalently, Y must contain its southwest-northeast 45° *diagonal* $d(Y)$ which connects Y's bottom left and top right corners. If not specified otherwise, a Young diagram is always proper in this paper.

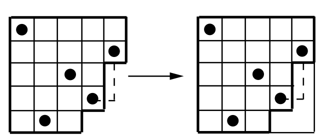

Figure 1: $T \in S_Y$ versus $T' \in S_5$

Young diagrams are traditionally coordinatized from *the top left corner*, meaning that their first (and largest) row and column are the top, respectively, leftmost ones. To avoid possible confusion with the matrix "bottom-left-corner" coordinatization used in this paper, one can think of a transversal $T \in S_Y$ by first completing the (proper) Young diagram Y to a square matrix M_n, and then taking a transversal T of M_n all of whose 1's are in the original cells of Y. Thus, whether using a matrix or a Young diagram, all transversals resemble the "shape" of permutations. For instance, in Fig. 1, the proper Young diagram $Y(5,5,4,4,3)$ is completed to the square matrix M_5, and the transversal $T \in S_Y$ induces a transversal $T' \in S_5$. As observed above, $T \in S_Y(213)$, but $T' \notin S_5(213)$ because the subsequence $(324) \approx (213)$ of T' does land in M_5.

Definition 6. Two permutations τ and σ are called *shape-Wilf-equivalent (SWE)*, denoted by $\tau \sim_s \sigma$, if $|S_Y(\tau)| = |S_Y(\sigma)|$ for all Young diagrams Y. If $|S_Y(\tau)| \leq |S_Y(\sigma)|$ for all such Y, we say that τ is *more shape-restrictive* than σ, and denote this by $\tau \preceq_s \sigma$.

Clearly, $\tau \sim_s \sigma$ ($\tau \preceq_s \sigma$) imply $\tau \sim \sigma$ ($\tau \preceq \sigma$, respectively), but the converses are false. Babson-West showed in [3] that SWE is useful in establishing more Wilf-equivalences. To the best of our knowledge, this idea of Young diagrams has not been yet been modified or used to prove Wilf-*ordering*, which the present paper will accomplish. To this end, we include below an extension of Babson-West's proposition, replacing shape-Wilf-*equivalences* "$\sim_s$" with shape-Wilf-*ordering* "$\preceq_s$". Section 2 presents a modification and extension of their original proof, and introduces along the way new notation necessary for the completion of our Wilf-ordering results.

Proposition 1. *Let $A \preceq_s B$ for some permutation matrices A and B. Then for any permutation matrix C:*

$$\left(\begin{array}{c|c} A & 0 \\ \hline 0 & C \end{array}\right) \preceq_s \left(\begin{array}{c|c} B & 0 \\ \hline 0 & C \end{array}\right).$$

If we shape-Wilf-order permutations in S_k for a small k, Proposition 1 will enable us to shape-Wilf-order some permutations in S_n for larger n. Since $(12) \sim_s (21)$ in S_2, Proposition 1 can imply in this case only shape-Wilf-*equivalences*.

The first non-trivial shape-Wilf-*ordering* can occur in S_3, since the latter splits into three distinct shape-Wilf-equivalence classes: $\{(213) \sim_s (132)\}$, $\{(123) \sim_s (231) \sim_s (321)\}$, and $\{(312)\}$. The first SWE-class was proven by Stankova-West in [21], and the second class was proven by Babson-Backelin-West-Xin in [3, 4]. The smallest Young diagram for which all three classes differ from each other is $Y = Y(5,5,5,5,4)$: $|S_Y(213)| = 37 < |S_Y(123)| = 41 < |S_Y(312)| = 42$. Numerical evidence suggests that such inequalities hold for *all* Young diagrams Y, and indeed this is true:

Theorem 1 (Main Theorem). *For all Young diagrams Y:*

$$|S_Y(213)| \leq |S_Y(123)| \leq |S_Y(312)|.$$

Figure 2 with $\tau = \emptyset$ illustrates the Main Theorem.[1]

[1] The referee of the current paper has kindly pointed out that an equivalent description of shape-Wilf equivalence has emerged recently. According to Mier [14], two permutations are shape-Wilf equivalent if and only if their *matching* graphs are *equirestrictive* among partition graphs, counted by the so-called *left-right degree sequences*. (She actually shows a more general correspondence between pattern-avoiding fillings of diagrams and pattern-avoiding graphs with prescribed degrees.) Using Mier's correspondence, one can translate a recent result of Jelinek's [10] as equivalent to the first inequality in the Main Theorem 1. Both papers [10, 14] will be published soon.

Let $Y_n = Y(n, n, n, ..., n, n-1)$ be the Young diagram obtained by removing the right bottom cell from the square M_n. Section 9 shows

$$|S_{Y_n}(213)| < |S_{Y_n}(123)| < |S_{Y_n}(312)| \text{ for } n \geq 5.$$

These strict inequalities preclude the possibility of the three permutations (213), (123), (312) to be asymptotically SWE, even though they are Wilf-equivalent. More precisely,

Theorem 2. $|S_Y(213)| < |S_Y(123)|$ *if and only if* Y *contains an* i*-critical point with* $i \geq 2$*, and* $|S_Y(123)| < |S_Y(312)|$ *if and only if* Y *contains an* i*-critical point with* $i \geq 3$.

The definition and a discussion of *critical points* can be found in Subsection 3.2. While for any $\tau \in S_3$ the "Wilf-numbers" $|S_n(\tau)|$ equal the Catalan numbers $c_n = \frac{1}{n+1}\binom{2n}{n}$, the "shape-Wilf-numbers" $|S_Y(\tau)|$ naturally vary a lot more. In particular, for the staircases $Y = St_n^3$, $|S_Y(\tau)|$ coincide with the odd-indexed Fibonacci terms f_{2n-1}, and hence involve the *golden ratio* $\phi = (1+\sqrt{5})/2$ (cf. Definition 13 and Section 9.)

Definition 7. We say that a permutation $\tau \in S_n$ is *decomposable* into blocks A_1 and A_2 if for some $k < n$, τ can be partitioned into two subpatterns $A_1 = (\tau_1, \tau_2, ..., \tau_k)$ and $A_2 = (\tau_{k+1}, \tau_{k+2}, ..., \tau_n)$ such that all entries of A_1 are *bigger* than (and a priori come before) all entries of A_2. We denote this by $\tau = (A_1|A_2)$. If there is no such decomposition into two blocks, we say that τ is *indecomposable*. In particular, a *reverse-layered* pattern τ is a permutation decomposable into increasing blocks.

For example, $(4132) = (4|132)$ is decomposable, while (3142) and (1432) are indecomposable; $(4123) = (4|123)$ is reverse-layered, while (4132) is not reverse-layered. Without confusion, we can also write $(213|1)$ instead of (3241). In this notation, Proposition 1 can be rewritten as $A \preceq_s B \;\Rightarrow\; (A|C) \preceq_s (B|C)$.

Corollary 1. *For any permutation* $\tau \in S_k$*,* $(213|\tau) \preceq_s (123|\tau) \preceq_s (312|\tau)$*, and strict asymptotic Wilf-ordering* $|S_n(213|\tau)| < |S_n(123|\tau)| < |S_n(312|\tau)|$ *occurs for* $n \geq 2k+5$.

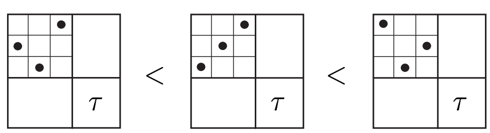

Figure 2: Corollary 1

In particular, when $\tau = (1)$ Corollary 1 reduces to:

$$\begin{aligned} &|S_n(213|1)| < |S_n(123|1)| < |S_n(312|1)| \text{ for } n \geq 7 \\ \Rightarrow\quad &|S_n(3241)| < |S_n(2341)| < |S_n(4231)| \text{ for } n \geq 7. \end{aligned}$$

Note that $(3241) \sim (1342)$ and $(4231) \sim (1324)$ (cf. Fig. 3a-c) since the two permutation matrices in each Wilf-equivalence pair can be obtained from each other by applying symmetry operations of flipping along vertical, horizontal and/or diagonal axes (cf. [23, 19]). Further, $(2341) \sim (1234)$ by the SWE-relations in [4], or by an earlier work [20]. Thus, choosing the second representatives of the three Wilf-equivalence classes in S_3, we obtain Bóna's (1) inequality as a special case of Corollary 1.

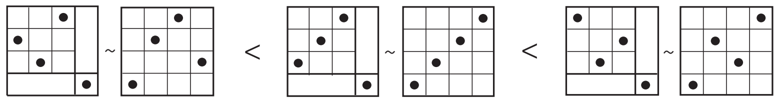

Figure 3: Wilf-Ordering of S_4

Some of the implied new shape-Wilf-orderings by Corollary 1 in S_5 and S_6 are:

$$\begin{array}{lll} (43521) \prec_s^* (54321) \prec_s (53421) & \quad & (546231) \prec_s^* (654231) \prec_s^* (645231), \\ (546321) \prec_s^* (654321) \prec_s (645321) & \quad & (546213) \prec_s^* (654213) \prec_s^* (645213). \end{array}$$

These shape-Wilf ordering inequalities imply Wilf-orderings, of which the ones corresponding to $*$'s are new. Note that the two $*$ in the left column are not surprising, since it is known that $L(43521) < L(54321)$ and $L(546321) < L(654321)$ (cf. [8, 9].)

The paper is organized as follows. Section 2 presents the proof of Proposition 1, along with a strategy for establishing strict asymptotic Wilf-orderings. In Section 3, we introduce *critical points*, provide the 0- and 1-*splittings* $S_Y(\sigma) \cong S_{Y^R}(\sigma) \times S_{{}_QY}(\sigma)$ in Proposition 2, and a 2-*critical splitting* in Lemma 6. Subsection 3.5 defines the $\sigma \to \tau$ *moves* on transversals in Y, and opens up the discussion of the induced maps $\phi : S_Y(\tau) \to S_Y(\sigma)$. Sections 4-6 contain the proof of the inequalities $|S_Y(312)| \geq |S_Y(321)|$ and $|S_Y(213)| \leq |S_Y(123)|$; a description of the structures of $T \in S_Y(321)$ and $T \in S_Y(312)$ can be found in Subsections 4.1-4.2. Using critical points, necessary and sufficient conditions for strict inequalities $|S_Y(312)| > |S_Y(321)|$ and $|S_Y(213)| < |S_Y(123)|$ are established in Sections 5-7. Section 8 provides the proof of the strict Wilf-orderings $|S_n(213|\tau)| < |S_n(123|\tau)| < |S_n(312|\tau)|$ for $n \geq 2k+5$. Finally, in Section 9 we calculate $|S_Y(\tau)|$ for $\tau \in S_3$ and Young diagrams Y which are extreme with respect to their critical points. The paper ends with a generalization of the Stanley-Wilf limits and the fact that ϕ^2 is such a limit.

2 Proof of Proposition 1

In this section we present a modified and extended version of the original proof of Babson-West to address our new setting of shape-Wilf ordering. Let the permutation matrices A, B and C in the statement of Proposition 1 represent permutations α, β and γ, respectively. Before we proceed with the proof, we need to introduce some definitions and notation.

2.1 Various subboards of Y

Let Y be a Young diagram, and let c be a cell in Y. Denote by ${}^{\bar{c}}Y$ the subboard of Y to the right and below c, not including c's row and column; and by Y_c the subboard of Y to the left and above c, including the corresponding cells in c's row and column. Since Y is a Young diagram, ${}^{\bar{c}}Y$ is also a Young diagram (not necessarily proper), and Y_c is a rectangle whose right bottom cell is c (cf. Fig. 4). This notation is created so as to match the relative positions of c and the corresponding subboard of Y, where exclusion of c's

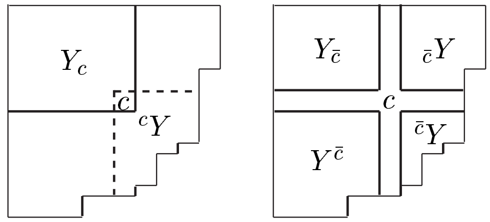

Figure 4: Notation Y_c and ^cY versus $Y_{\bar{c}}$, $^{\bar{c}}Y$, etc.

row and column is denoted by $\bar{c}$. In the same vein, we define Y^c, $_{\bar{c}}Y$, etc. We also extend the notation to (full or partial) *transversals* T of Y, to elements $a \in T$, and to grid points P of Y; for instance, $_{\bar{a}}T = T|_{_{\bar{a}}Y}$ is the restriction of T onto the subboard $_{\bar{a}}Y$, while Y^P is the subboard Y^c where P is the top right corner of cell c.

We use the symbols $\nearrow$ and $\searrow$ instead of the words "increasing" and "decreasing". Thus, $I_k \nearrow$, and its transpose $J_k \searrow$.

Definition 8. Let $T \in S_Y$, and $a, b \in T$. We say that a (21)-*dominates* b if $(ab) \searrow$. Similarly, a (12)-*dominates* b if $(ab) \nearrow$ and lands in Y. We extend these definitions to any cells of and dots in Y.

Note that, while in (21)-domination the decrease $(ab) \searrow$ automatically implies that (ab) lands in Y, the definition of (12)-domination requires this "landing" property separately. For example, recall Fig. 1a, which depicts the permutation (51324) in the Young diagram $Y(5, 5, 4, 4, 3)$. Here $a = 5$ (21)-dominates $b = 2$ and $a = 1$ (12)-dominates 3, but $a = 1$ does **not** (12)-dominate $b = 4$ since (14) does not land inside Y.

2.2 Coloring of Y with respect to T and γ

Fix a transversal $T \in S_Y$. With respect to the pattern γ, T induces a white/blue coloring on Y's cells as follows. Color a cell c in Y *white* if $^{\bar{c}}Y$ contains C as a submatrix; otherwise, color c *blue* (recall that C is the permutation matrix of γ). Clearly, for every white cell w, the rectangle Y_w is also entirely white. Hence, the white *subboard* W' of Y is a Young subdiagram of Y (not necessarily proper), and T induces a *partial* transversal $T|_{W'}$ of W'.

In order for T to avoid $(\alpha|\gamma)$, it is necessary and sufficient that $T|_{W'}$ avoids α. However, some rows and columns of W' cannot participate in any undesirable α-patterns since the 1's in them are in blue cells: recolor these white rows and columns of W' to blue. After deletion of the newly blue rows and columns of W', the latter is reduced to a white proper Young *subdiagram* W of Y, while $T|_{W'}$ is reduced to a *full* transversal $T|_W$ of W.

Definition 9. We say that the transversal T of Y *induces with respect to* γ the white subdiagram W of Y and the (full) transversal $T|_W$ of W. Let $S_Y^W(\alpha|\gamma)$ denote the set of all transversals $T \in S_Y(\alpha|\gamma)$ which induce W with respect to γ.

For example, Figure 5a shows a transversal $T \in S_Y$ and the induced white subboard W' with respect to $\gamma = (213)$: the blue subboard of Y is depicted with its grid lines, while

W' is depicted without them; the dashed lines pass through some of the blue 1's and indicate that these rows and columns of Y will be deleted from W'. Figure 5c shows the final white subdiagram $W(4,4,3,3)$ and its transversal $T|_W = (2134)$. Figure 5a-c also illustrates that $T = (7,6,9,2,10,1,4,5,3,8) \in S_Y$ avoids $(123|213)$ because $T|_W = (2134)$ avoids (123) on W, but it contains $(213|213)$ because $T|_W$ contains pattern (213) on W.

We summarize the observations in this subsection in the following

Lemma 1. *Let W be any Young subdiagram of Y. Then*

1. $T \in S_Y(\alpha|\gamma) \;\Leftrightarrow\; T|_W \in S_W(\alpha)$.
2. $S_Y(\alpha|\gamma) = \bigsqcup_{W\subset Y} S_Y^W(\alpha|\gamma)$.

2.3 Splitting of transversals $T \in S_Y$ with respect to γ

Fix now a (white) Young subdiagram W of Y, and let $T \in S_Y^W(\alpha|\gamma)$. By construction of W, T splits itself into two disjoint subsets: the induced transversal $T|_W$ of W consisting of all "white" 1's, and the remainder $T_\gamma = T\backslash T|_W$ consisting of all "blue" 1's. We denote this by

$$T = T|_W \oplus T_\gamma, \text{ where } T|_W \in S_W(\alpha).$$

A *key observation* is that, if T'_W is another transversal in $S_W(\alpha)$, then $T' = T'_W \oplus T_\gamma \in S_Y^W(\alpha|\gamma)$. This is true because fixing T_γ preserves the white cells of W, and replacing $T|_W$ with any other transversal of W certainly does not affect the blue colored cells in $Y\backslash W$. For example, Figure 5 shows $T \in S_Y^W(123|213)$ with $W = (4,4,3,3)$, $T|_W = (2134) \in S_W(123)$, and $T_{(213)} = (214538) \approx (214536)$. If we keep $T_{(213)}$ and replace $T|_W$ with another $T'_W = (3214) \in S_W(123)$ (shown in Fig. 5d), we obtain the transversal in Fig. 5e:

$$T' = (9,7,6,2,10,1,4,5,3,8) = (9,7,6,10) \oplus (2,1,4,5,3,8) \in S_Y^W(123|213).$$

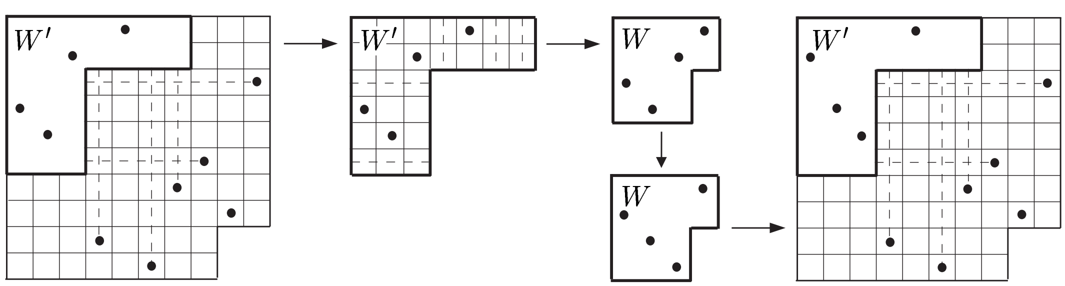

Figure 5: $T = T|_W \oplus T_{(213)} \to T' = T'_W \oplus T_{(213)}$ in $S_Y^W(123|213)$

We conclude that all transversals $T \in S_Y^W(\alpha|\gamma)$ whose second component is a fixed T_γ are obtained by adding an arbitrary transversal $T'_W \in S_W(\alpha)$ to T_γ:

$$T = T'_W \oplus T_\gamma \in S_Y^W(\alpha|\gamma) \text{ for any } T'_W \in S_W(\alpha).$$

2.4 Description of the T_γ-component of $T \in S_Y^W(\alpha|\gamma)$

We can extend the definitions of the white/blue coloring of Y above to *partial* transversals T' of Y: a *blue* cell b in Y is such that ${}^{\bar{b}}Y$ does **not** contain a γ-subpattern of T', while a *white* cell w in Y is such that ${}^{\bar{w}}Y$ does contain a γ-subpattern of T'.

Recall the notion of *reduction of* Y *along a subset* X of Y's cells, introduced in [21]: $Y/_X$ is the Young subdiagram obtained from Y by deleting all rows and columns of Y which intersect X. This notation should not be confused with $Y\backslash X$ - the subboard obtained from Y by removing the cells in X, or with $T|_W$ - the restriction of T on W.

Definition 10. Let W be a proper subdiagram of a Young diagram Y. A partial transversal T' of Y *saturates* W *with respect to* γ if the induced by T' blue/white coloring on Y with respect to γ satisfies:

(1) T''s elements are all placed in blue cells;

(2) Reducing Y along T' and removing any leftover blue cells results in W; and

(3) $|W| + |T'| = |Y|$, where $|U|$ is the size of a proper Young diagram U and $|T'|$ counts the number of elements in T'.

Since a blue cell cannot (21)-dominate a white cell, no matter which transversal of W we choose to complete T' to a (full) transversal of Y, the blue/white coloring of Y will remain the same (cf. Fig. 6.) Condition (3) ensures that there is no entirely blue row or column without an element of T'; in fact, (3) matches the sizes of W and T' so that any transversal of W will indeed complete T' to a *full* transversal of Y.

According to Definition 10, for a transversal $T \in S_Y^W(\alpha|\gamma)$ with splitting $T = T|_W \oplus T_\gamma$, the partial transversal T_γ of Y saturates W with respect to γ.

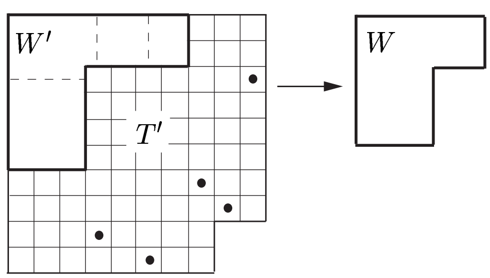

Figure 6: T' saturates W with respect to (213)

Definition 11. Given a subdiagram W of the Young diagram Y, let $\bar{S}_{Y\backslash W}(\gamma)$ denote the set of partial transversals T' of Y which saturate W with respect to γ.

2.5 Splitting Formula for $|S_Y(\alpha|\gamma)|$

We have seen that any transversal $T \in S_Y^W(\alpha|\gamma)$ splits uniquely as $T = T|_W \oplus T_\gamma$, where $T|_W$ avoids α on W and T_γ saturates W in Y with respect to γ. This defines an injective map $S_Y^W(\alpha|\gamma) \hookrightarrow S_W(\alpha) \times \bar{S}_{Y\backslash W}(\gamma)$. The key observation in Subsection 2.3 shows that this map is surjective. Therefore,

Lemma 2 (Splitting Formula for $|S_Y(\alpha|\gamma)|$). *For any subdiagram W of the Young diagram Y, the isomorphism of sets $S_Y^W(\alpha|\gamma) \cong S_W(\alpha) \times \bar{S}_{Y\setminus W}(\gamma)$ holds true. Consequently,*

$$|S_Y(\alpha|\gamma)| = \sum_{W\subset Y} |S_W(\alpha)| \cdot |\bar{S}_{Y\setminus W}(\gamma)|,$$

where the sum is taken over all Young subdiagrams W of Y.

Since the components $\bar{S}_{Y\setminus W}(\gamma)$ depend only on γ and W (but not on α), this allows for direct comparisons between $S_Y(\alpha|\gamma)$ and $S_Y(\beta|\gamma)$. In particular, if $\alpha \preceq_s \beta$, then $|S_W(\alpha)| \leq |S_W(\beta)|$ for any Young diagram W, and the splitting formulas for α and β imply $|S_Y(\alpha|\gamma)| \leq |S_Y(\beta|\gamma)|$. This completes the Proof of Proposition 1. □

2.6 Strategy for proving strict Wilf-ordering

When $\alpha \preceq \beta$, the Splitting Formula can be used to prove a strict asymptotic Wilf-ordering of the form $|S_n(\alpha|\gamma)| \lneqq |S_n(\beta|\gamma)|$, provided that for $n \gg 1$:

(SF1) there is a Young diagram W_n with $|S_{W_n}(\alpha)| \lneqq |S_{W_n}(\beta)|$; and

(SF2) there is a partial transversal T_n of M_n saturating W_n with respect to γ.

The existence of W_n and T_n ensures that $|S_{W_n}(\beta)| > 0$ and $|\bar{S}_{M_n\setminus W_n}(\gamma)| > 0$, so that

$$|S_{W_n}(\alpha)| \cdot |\bar{S}_{M_n\setminus W_n}(\gamma)| \lneqq |S_{W_n}(\beta)| \cdot |\bar{S}_{M_n\setminus W_n}(\gamma)|.$$

We shall employ this strategy in Section 8 to show strict asymptotic Wilf-ordering between the permutations $(213|\tau), (123|\tau)$ and $(312|\tau)$ of Corollary 1.

3 Critical Splittings of Diagrams and Transversals

3.1 First and second subsequences of $T \in S_Y$.

Recall that $\alpha \in T$ is a *left-to-right maximum* of T if α is **not** (21)-dominated by any other element of T, i.e. $T_{\bar{\alpha}} = \emptyset$.

Definition 12. Let $T \in S_Y$. The subsequence T^1 of all left-to-right maxima α_i of T is called the *first subsequence* of T. The *second subsequence* T^2 of T consists of all elements $\beta_j \in T\setminus T^1$ for which $Y_{\bar{\beta}_j}$ contains only elements of T^1, i.e. β_j is (21)-dominated only by (a non-empty set of) elements of T^1.

Observe that T^1 and T^2 are increasing subsequences of T. Figure 7a depicts T^1 and T^2 (via dashed lines) and three instances of $\alpha_i \in T^1$ (21)-dominating $\beta_j \in T^2$ (via solid arrows).

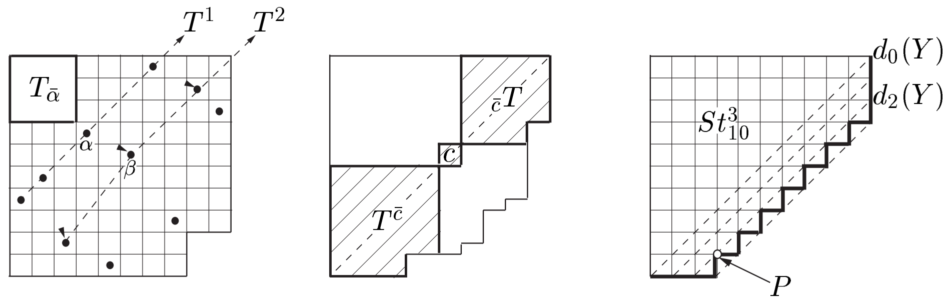

Figure 7: (a) T^1 and T^2 (b) Lemma 3 (c) 2-critical P in St^3_{10}

3.2 Diagonal Properties and Critical Points

We address now the relative positioning of an arbitrary transversal within its Young diagram.

Lemma 3. *Let $T \in S_Y$ and let c be a cell on the diagonal $d(Y)$. Then the rectangle Y_c contains some element of T^1. Consequently, all elements of the first subsequence T^1 are on or above $d(Y)$.*

Proof: Suppose Y_c contains no elements of T. But there is no transversal of Y to sustain such a big empty rectangle. Indeed, since $c \in d(Y)$, $Y^{\bar c}$ is a proper Young subdiagram of Y, say of size k, and there are no elements of T above $Y^{\bar c}$. Thus, the first k columns of Y must have their 1's within $Y^{\bar c}$, and T induces a transversal $T^{\bar c}$ of $Y^{\bar c}$. Analogously, T induces a transversal $_{\bar c}T$ of $_{\bar c}Y$. Hence, T must split into $T = T^{\bar c} \oplus T(c) \oplus {}_{\bar c}T$, where $T(c)$ is a transversal of the cell c (cf. Fig. 7b, where T is concentrated in the 3 shaded subboards). But cell c is empty by the supposition, a contradiction. Therefore, Y_c does contain some element $\gamma \in T$. Since either $\gamma \in T^1$ or γ is (21)-dominated by some $\alpha \in T^1$, we conclude that Y_c contains an element of T^1.

If some $\alpha_i \in T^1$ is below the diagonal $d(Y)$, then the rectangle $Y_{\bar\alpha_i}$ contains a cell c on $d(Y)$, and Y_c is empty, a contradiction with the previous paragraph. Therefore, T^1's elements are on or above $d(Y)$.

By the *border* of a Young diagram Y we mean the path that starts at the bottom left corner of Y, follows Y's outline *below and to the right* of $d(Y)$, and ends at the top right corner of Y.

Definition 13. For a Young diagram Y, define the *i-th diagonal* $d_i(Y)$ as follows: starting from the bottom left corner of Y, move i cells to the right, draw a parallel line to $d(Y)$ until it goes through the rightmost column of Y; the resulting segment is $d_i(Y)$. For $i \geq 1$, denote by St^i_n the *i-th Staircase* Young diagram of size n whose border is the stepwise path from the bottom left corner to the top right corner of Y that zigzags between $d_{i-1}(Y)$ and $d_i(Y)$ (cf. Fig. 7c for $d_i(Y)$ with $0 \leq i \leq 3$, and St^3_{10}.)

We distinguish between $d_0(Y)$, which is a segment going through Y's diagonal *grid points*, and $d(Y)$, which is the union of all diagonal *cells* of Y.

Definition 14. A grid point P on Y's border is called a *critical point of* Y if Y's border goes upwards to enter P and then goes to the right to leave P. If in addition $P \in d_i(Y)$, then P is called an i-critical point of Y.

Figure 7c shows the *bottom* 2-critical point P of St^3_{10}. Note that $St^n_n = M_n$ is the only Young diagram of size n with no critical points, while St^1_n has the largest number of critical points. Also, for any critical point P, the subboard ${}^PY$ has no cells and consists only of the point P, while Y_P is a rectangle.

Lemma 4. *If P is an i-critical point of Y and $T \in S_Y$, then the rectangle Y_P contains exactly i elements of T.*

Proof: Let Y have exactly k rows above P. Since $P \in d_i(Y)$, the subboard ${}_PY$ has k rows and $k-i$ columns; the latter are in fact all columns of Y which are to the right of P, and therefore each of these $k-i$ columns contains exactly 1 element of T. Hence $k-i$ of ${}_PY$'s rows contain an element of T, while i rows of ${}_PY$ are empty (cf. Fig. 8a-b for $i = 0, 1$ and Fig. 9a for $i = 2$.)

On the other hand, each of the top k rows of Y is split between the rectangle Y_P and the subboard ${}_PY$. From the viewpoint of Y_P, the above observations mean that $k-i$ rows of Y_P are empty, while exactly i rows of Y_P contain an element of T. Thus, $|T_P| = i$. □

3.3 Definition of the map ζ_P

For an i-critical point P in Y, let $Q, R \in d_0(Y)$ be the diagonal grid points of Y to the left of, respectively above, P. Then ${}_QY$ and Y^R are proper Young subdiagrams (cf. Fig. 8a-b and Fig. 9a.)

Fix $T \in S_Y$. Lemma 4 ensures that rectangle Y_P contains exactly i elements of T, which form some subsequence $\alpha = (\alpha_1, \alpha_2, ..., \alpha_i)$. While preserving the pattern α, we can simultaneously pull downward all α_i's until they become the top i elements in a transversal T_1 of Y^R, and we can also push all α_i's to the right until they become the i leftmost elements of a transversal T_2 of ${}_QY$. These operations define an injective map

$$\zeta_P : S_Y \hookrightarrow S_{Y^R} \times S_{{}_QY} \text{ where } \zeta_P(T) = (T_1, T_2).$$

For example, Fig. 8b-c show $\zeta_P(31628547) = (3142, 35214)$ with $i = 1$ and $\alpha_1 = 6$, while Fig. 9 shows $\zeta_P(831629547) = (53142, 536214)$ with $i = 2$ and $\alpha_1 = \alpha = 8$ and $\alpha_2 = \beta = 6$. Since ${}^PY$ has no cells, any subsequence of T landing inside Y must be contained either entirely in the rows of Y above P, or entirely in the columns of Y to the left of P. Consequently,

Lemma 5. *For any pattern σ, T avoids σ on Y if and only if the components T_1 and T_2 of $\zeta_P(T)$ avoid σ on Y^R and ${}_QY$, respectively. In particular, ζ_P respects pattern-avoidance and we can restrict $\zeta_P : S_Y(\sigma) \hookrightarrow S_{Y^R}(\sigma) \times S_{{}_QY}(\sigma)$.*

3.4 Critical Splittings induced by ζ_P

Proposition 2. *If P is a 0- or 1-critical point of Y, then $S_Y(\sigma) \overset{\zeta_P}{\cong} S_{Y^R}(\sigma) \times S_{{}_QY}(\sigma)$ for any $\sigma \in S_k$.*

Proof: Fix $T \in S_Y$ and let σ be any permutation. A 0-critical point P coincides with the points Q and R in the definition of ζ_P, and the rectangle Y_P has no elements of T by Lemma 4 (cf. Fig. 8a.) Thus, $\zeta_P : S_Y(\sigma) \hookrightarrow S_{Y^R}(\sigma) \times S_{{}_QY}(\sigma)$ simply restricts $T|_{Y^P} = T_1$ and $T|_{{}_PY} = T_2$; combined with Lemma 5, this yields invertibility of ζ_P. In this case, we say that ζ_P induces the 0-*splitting* $T = T|_{Y^P} \oplus T|_{{}_PY}$.

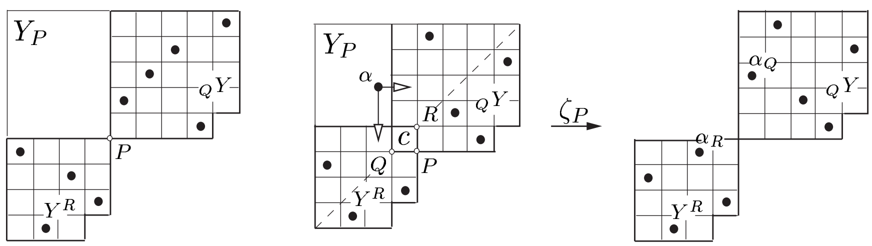

Figure 8: (a) 0-splitting (b)-(c) 1-splitting

Now, consider the case of a 1-critical point P (cf. Fig. 8b-c.) Let c be the cell whose bottom right corner is P. Then c lies on the diagonal $d(Y)$, and Q and R are also respective corners of c. Let $c = (k, m)$ where k is c's row and m is c's column in Y. By Lemma 4, the rectangle Y_P has exactly one element of T: call it α, and let it be in position (i, j) in Y. To form transversals $T_1 \in S_{Y^R}$ and $T_2 \in S_{{}_QY}$, ζ_P replaces α by α_R in position (k, j) and α_Q in position (i, m), respectively.

It is not hard to see that ζ_P is surjective. Indeed, start with $(T_1, T_2) \in S_{Y^R} \times S_{{}_QY}$. If T_1 has its top element α_Q in its j-th column, and T_2 has its leftmost element α_R in its i-th row, we can reconstruct the unique $\alpha \in Y_P$ by replacing (α_R, α_Q) by an element in position (i, j) and leaving the rest of T_1 and T_2 fixed. Combining this with Lemma 5 yields the wanted isomorphism ζ_P on $S_Y(\sigma)$. In this case, we say that ζ_P induces the 1-*splitting* $T = T|_{Y^P} \oplus_1 T|_{{}_PY}$. $\square$

As expected, i-critical points for larger i complicate matters, and in general, it is not possible to derive such nice splittings of transversals. Below we describe the image $\zeta_P(S_Y(\sigma))$ for a 2-critical point P.

Definition 15. Let $S_{\nearrow Y}(\sigma)$, respectively $S_{Y}^{\nearrow}(\sigma)$, be the set of transversals T in $S_Y(\sigma)$ whose two leftmost, respectively two top, elements form an increasing subsequence of T. Define analogously $S_{Y}^{\searrow}(\sigma)$ and $S_{\searrow Y}(\sigma)$ with appropriate replacement of $\nearrow$ by $\searrow$.

We will also need the notation $S_{\nearrow Y}^{\searrow}(\sigma) = S_{Y}^{\searrow}(\sigma) \cap S_{\nearrow Y}(\sigma)$. As with previous notation, this one preserves the relative position of the involved objects, in this case – Y and its two (top and/or leftmost) subsequences of length 2. The $\nearrow$ and $\searrow$ arrows can be arbitrarily switched to denote the corresponding other subsets of transversals.

Lemma 6. *If P is a 2-critical point of Y, then for any $\sigma \in S_k$:*

$$S_Y(\sigma) \overset{\zeta_P}{\cong} S^{\nearrow}_{Y^R}(\sigma) \times S_{\nearrow_Q Y}(\sigma) \sqcup S^{\searrow}_{Y^R}(\sigma) \times S_{\searrow_Q Y}(\sigma). \tag{2}$$

Proof: Start with $T \in S_Y$. By Lemma 4, we may assume that α and β are the only elements of T in rectangle Y_P, with α to the left of β. Depending on whether $(\alpha\beta)\nearrow$ or

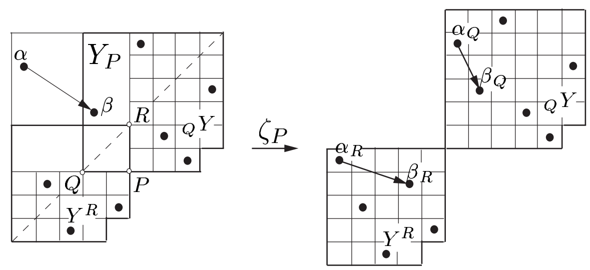

Figure 9: $\zeta_P(T) = (T_1, T_2)$ on $Y^R \times {}_Q Y$ with $(\alpha\beta)\searrow$ in Y_P

$\searrow$, either component $T_1 \in S_{Y^R}(\sigma)$ has its top two elements $(\alpha_R, \beta_R)\nearrow$ *and* component $T_2 \in S_{_QY}(\sigma)$ has its two leftmost elements $(\alpha_Q, \beta_Q)\nearrow$, *or* both of these subsequences are decreasing. For instance, Figure 9 depicts the case $(\alpha\beta)\searrow$.

Conversely, start with $(T_1, T_2) \in S^{\nearrow}_{Y^R}(\sigma) \times S_{\nearrow_Q Y}(\sigma) \sqcup S^{\searrow}_{Y^R}(\sigma) \times S_{\searrow_Q Y}(\sigma)$. If (α_R, β_R) and (α_Q, β_Q) are the top two, respectively, the leftmost two, elements of T_1 and T_2, they form the same length-2 pattern, say, they are both decreasing. This makes it possible to pull back α_R and α_Q to an element α in rectangle Y_P, and pull back β_R and β_Q to an element β in rectangle Y_P, so that (α, β) is also decreasing and $\zeta_P(\alpha, \beta) = (\alpha_R, \beta_R) \times (\alpha_Q, \beta_Q)$ in $Y^R \times {}_Q Y$. This discussion establishes the two isomorphisms $S^{\nearrow}_Y(\sigma) \cong S^{\nearrow}_{Y^R}(\sigma) \times S_{\nearrow_Q Y}(\sigma)$ and $S^{\searrow}_Y(\sigma) \cong S^{\searrow}_{Y^R}(\sigma) \times S_{\searrow_Q Y}(\sigma)$, and since $S_Y(\sigma) = S^{\nearrow}_Y(\sigma) \sqcup S^{\searrow}_Y(\sigma)$, we deduce (2). $\square$

The reader can prove a similar splitting for an i-critical point P with $i \geq 3$ and $\sigma \in S_k$:

$$\zeta_P : S_Y(\sigma) \cong \bigsqcup_{\tau \in S_i} S^{\tau}_{Y^R}(\sigma) \times S_{\tau_Q Y}(\sigma),$$

where in the notations $S^{\tau}_{Y^R}(\sigma)$ and $S_{\tau_Q Y}(\sigma)$ the patterns $\tau \in S_i$ have replaced the previously used $\nearrow = (12)$ and $\searrow = (21)$ in S_2. In order for this isomorphism to be useful, one should be able to enumerate the components $S^{\tau}_{Y^R}(\sigma)$ and $S_{\tau_Q Y}(\sigma)$; however, for a general pattern σ and high critical index i, this question acquires a level of difficulty at least comparable to that of Wilf-enumeration $|S_n(\sigma)|$. Fortunately, when $i = 2$ and $\sigma = (312)$ or (321), this enumeration is possible and is carried out in Section 5.

3.5 The $\sigma \to \tau$ moves

Let $T \in S_Y$. For any two permutations $\sigma, \tau \in S_k$ we define a $\sigma \to \tau$ *move* on T as follows: if $(\alpha_1 \alpha_2 \cdots \alpha_k)$ is a σ-subpattern of T in Y, we rearrange the α_i's within the $k \times k$ matrix they generate so as to obtain a τ-subpattern $(\beta_1\beta_2\cdots\beta_k)$ in Y. The inverse operation is obviously a $\tau \to \sigma$ move. A sequence of $\sigma \to \tau$ moves that starts with a transversal T is called "a sequence of $\sigma \to \tau$ moves *on* T".

For example, if $(\alpha\beta\gamma)$ is a (213)-pattern in T landing in Y, a $(213) \to (123)$ move switches the places of α and β to obtain $(\beta\alpha\gamma) \approx (123)$ in Y. Throughout the paper, we will use two instances of $\sigma \to \tau$ moves: $(213) \to (123)$ and $(312) \to (321)$ moves, along with their inverses. In particular, we will construct maps

$$S_Y(213) \hookrightarrow S_Y(123) \cong S_Y(321) \twoheadleftarrow S_Y(312),$$

and pose questions about the general maps $\phi : S_Y(\tau) \to S_Y(\sigma)$ that are induced under certain circumstances by a sequence of $\sigma \to \tau$ moves in Y.

4 Proof of the Inequality $S_Y(312) \geq S_Y(321)$

In this section we prove that $(321) \preceq_s (312)$. Since $(321) \sim_s (123)$, this will establish the required in Theorem 1 inequalities $|S_Y(123)| \leq |S_Y(312)|$ for all Young diagrams Y. The strategy is to describe the structures of each set $S_Y(321)$ and $S_Y(312)$, use this information to define a canonical map $\phi : S_Y(312) \to S_Y(321)$, and finally prove that ϕ is surjective.

4.1 The structure of $T \in S_Y(321)$

T is the disjoint union of its first and second subsequences: $T = T^1 \sqcup T^2$. Indeed, if there were some $\gamma \in T\backslash\{T^1 \cup T^2\}$, then $Y_{\bar{\gamma}}$ would contain some element $\beta \in T^2$, and hence $Y_{\bar{\beta}}$ would contain some element $\alpha \in T^1$, so that $(\alpha\beta\gamma) \approx (321)$ in T and lands in Y, a contradiction.

4.2 The structure of $T \in S_Y(312)$

Compared to the previous paragraph, the structure here is considerably more complex. We shall not need all of it in the proof of the inequality $S_Y(312) \geq S_Y(321)$. Yet, it is enlightening as to why the proof works and why strict inequalities $S_Y(312) > S_Y(321)$ occur for some Y. For the remainder of this subsection, we fix some transversal $T \in S_Y(312)$.

Definition 16. For any $\beta \in T^2$, define a *directed graph* G_β on the elements of ${}^\beta T$ as follows: connect by a directed edge $\overrightarrow{\delta_1\delta_2}$ any two elements δ_1 and δ_2 of ${}^\beta T$ such that $(\delta_1\delta_2)\searrow$ and there is no "intermediate" $\delta_3 \in {}^\beta T$ with $(\delta_1\delta_3\delta_2)\searrow$ (cf. Fig. 10a.)

Lemma 7. For any $\beta \in T^2$, ${}^\beta T$ avoids (12) in Y. Further, G_β is connected and, stripping off the orientation of its edges, cycle-free.

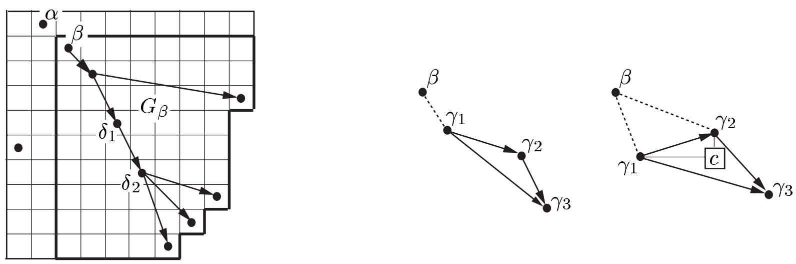

Figure 10: (a) Graph G_β for $\beta \in T^2$; (b)-(c) Lemma 7

Proof: For the first part, by definition of $\beta \in T^2$, there is some $\alpha \in T^1 \cap T_{\bar{\beta}}$ which (21)-dominates β. To avoid the possibility of α playing the role of a "3" in a (312)-pattern in T, ${}^{\beta}T$ must avoid (12) in Y.

For the second part, β (21)-dominates any $\gamma \in {}^{\bar{\beta}}T$ so that γ is connected to at least one other vertex in ${}^{\beta}T \cap T_{\bar{\gamma}}$, and eventually, there is a path starting from β and leading to γ. Thus, G_β is connected.

Suppose that there is an (undirected) cycle $\mathcal{C}$ in G_β. If we start at an arbitrary vertex $\delta \in \mathcal{C}$ and follow $\mathcal{C}$ along the orientation of its edges, we cannot come back to δ, or else we will have a decreasing sequence $(\delta, \delta_1, \delta_2, ..., \delta_k, \delta)$, which is absurd.

Therefore, going around $\mathcal{C}$ along the edge orientation leads to a *smallest* vertex γ_3 in $\mathcal{C}$, at which two edges $\overrightarrow{\gamma_1\gamma_3}$ and $\overrightarrow{\gamma_2\gamma_3}$ terminate (with, say, γ_1 before γ_2.) If $(\gamma_1\gamma_2)\searrow$, then $(\gamma_1\gamma_2\gamma_3)\searrow$, contradicting the construction of G_β without intermediate vertices (cf. Fig. 10b.) Thus, $(\gamma_1\gamma_2)\nearrow$. Since $\gamma_3 \in {}^{\bar{\gamma}_1}T \cap {}^{\bar{\gamma}_2}T$, the triangle $\gamma_1\gamma_2\gamma_3$ contains the cell c onto which $(\gamma_1\gamma_2)$ lands as a (12)-pattern, and hence ${}^{\beta}T$ also contains c (cf. Fig. 10c.) Yet, by the first part of this Lemma, ${}^{\beta}T$ avoids (12) in Y, a contradiction. Therefore, G_β has no (undirected) cycles. $\square$

Lemma 7 allows us to think of G_β as an *oriented tree rooted at* β. Now consider all trees G_{β_i}, where $T^2 = (\beta_1, \beta_2, \cdots, \beta_k)\nearrow$. For $i < j$, if $\gamma \in G_{\beta_i} \cap G_{\beta_j}$, then $\gamma \neq \beta_1, \beta_2$ and $(\beta_i\gamma) \approx (\beta_j\gamma)\searrow$. Evidently, if m is between i and j, then $(\beta_m\gamma)\searrow$, so that γ is also in G_{β_m} (cf. Fig. 11a.) In other words,

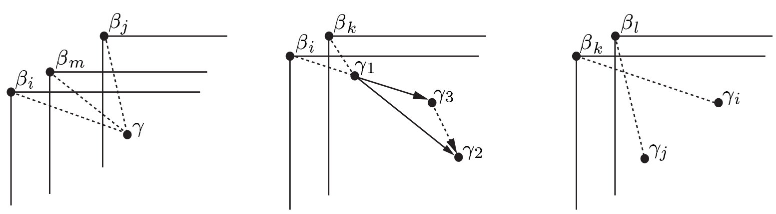

Figure 11: Lemmas 8, 9, 10

Lemma 8. *Let $G = \cup_{i=1}^{k} G_{\beta_i}$ be the union of all trees. Then each connected component C_j of G is the union of several consecutive trees: $C_j = G_{\beta_{k_j}} \cup G_{\beta_{k_j+1}} \cup G_{\beta_{k_j+2}} \cup \ldots \cup G_{\beta_{k_{j+1}-1}}$.*

By construction, each edge $\overrightarrow{\gamma_1\gamma_2}$ of a connected component C_j is entirely contained in some tree G_{β_i}. If γ_1 and γ_2 also belong to another tree G_{β_k}, then the edge $\overrightarrow{\gamma_1\gamma_2}$ must also belong to G_{β_k}. Indeed, if not, the (21)-pattern $(\gamma_1\gamma_2)$ requires at least one intermediate vertex γ_3 in G_{β_k}: $(\gamma_1\gamma_3\gamma_2)\searrow$ (cf. Fig. 11b.) But then γ_3 is also an intermediate vertex in G_{β_i}, hence the edge $\overrightarrow{\gamma_1\gamma_2}$ does not exist in G_{β_i}, a contradiction. We conclude that

Lemma 9. *Any tree G_{β_i} is a full subgraph of its connected component C_j.*

Using Lemma 9, we can augment the proof in Lemma 7 to derive in an almost identical way that each connected component C_j has no (undirected) cycles. Thus, we can think of each C_j as an oriented "tree" rooted at all of its the maximal elements, i.e. all $\beta_i \in T^2 \cap C_j$.

Lemma 10. *The connected components of G are arranged in an increasing pattern according to the β_i's they contain. More precisely, choose some $\beta_k \in C_i$ and $\beta_l \in C_j$ such that $k < l$, i.e. $(\beta_k\beta_l)\nearrow$. Then C_i is entirely to the left and below C_j.*

Proof: Consider any $\gamma_i \in C_i$ and $\gamma_j \in C_j$. If $(\gamma_i\gamma_j)\searrow$ or $(\gamma_j\gamma_i)\searrow$, Lemma 9 guarantees a path between γ_i and γ_j, contradicting $C_i \cap C_j = \emptyset$. Thus, γ_i and γ_j form (in some order) an increasing sequence. To complete the proof, we need to show $(\gamma_i\gamma_j)\nearrow$.

To the contrary, suppose $(\gamma_j\gamma_i)\nearrow$. Because of Lemma 8 and the arbitrary choice of $\beta_k \in C_i$ and $\beta_l \in C_j$, we may assume that $\gamma_i \in G_{\beta_k} \subset C_i$ and $\gamma_j \in G_{\beta_l} \subset C_j$, i.e. $(\beta_k\gamma_i)\searrow$ and $(\beta_l\gamma_j)\searrow$ (cf. Fig. 11c.) Putting together all four elements, we arrive at the subsequence $(\beta_k\beta_l\gamma_j\gamma_i) \approx (3412)$, which does not necessarily land in Y. Then $(\beta_k\gamma_j)\searrow$ so that $\gamma_j \in G_{\beta_k} \subset C_i$. Thus, $\gamma_j \in C_i \cap C_j = \emptyset$, a contradiction. If it happens that $\gamma_i = \beta_k$, or $\gamma_j = \beta_l$, or both, immediate contradictions in the overall arrangement arise.

We conclude that $(\gamma_i\gamma_j)\nearrow$, so that C_i is entirely to the left and below C_j. □

Thus, the connected components of G are arranged in a increasing diagonal fashion, symbolically, $G = (C_1, C_2, ..., C_k)\nearrow$. Correspondingly, the whole transversal $T \in S_Y(312)$ is the disjoint union of the increasing subsequence T^1 and all the vertices $|C_i|$ of the C_i's:

$$T = T^1 \sqcup |G| = T^1 \sqcup_i |C_i|. \tag{3}$$

This description of a (312)-avoiding transversal in Y is only *partial* (transversals satisfying it do not necessarily avoid (312)), but sufficient for our purpose to explain why (312) is easier to avoid than (321) on Young diagrams Y (cf. also Section 5.) In particular, the description involves only the elements of the transversal T, while it is possible to extend it to the whole Young diagram Y. To this end, let Y_j be the Young subdiagram of Y obtained after reducing Y along all elements of T not in C_j; one can think of Y_j as the Young subdiagram induced by the elements of C_j. Since the C_j's are disjoint, the Y_j's are disjoint, and we leave it to the reader to deduce in a similar fashion as above:

Lemma 11. *The Young subdiagrams Y_i are arranged in a increasing diagonal fashion: $Y = (Y_1, Y_2, ..., Y_k)\nearrow$.*

4.3 Definition of the map $\phi : S_Y(312) \to S_Y(321)$.

Fix a transversal $T \in S_Y(312)$, and decompose $T = T^1 \sqcup |G|$ as in (3) (cf. Fig. 12.) Reducing Y along T^1 leaves the pattern of $|G|$ in a Young subdiagram $Y_0 = Y/_{T^1}$. Since $|G|$ represents a transversal of Y_0, then Y_0 is proper, with diagonal $d(Y_0)$. Replacing $|G|$ by the increasing pattern $I_s = (123...s)$ along $d(Y_0)$ produces another transversal of Y_0. We reintroduce the rows and columns of the previously reduced subsequence T^1 to obtain our original Young diagram Y with a new transversal $\phi(T) = T^1 \sqcup I_s$. Since $\phi(T)$ is partitioned into two increasing subsequences, $\phi(T)$ avoids (321) and thus $\phi : S_Y(312) \to S_Y(321)$ is well-defined.

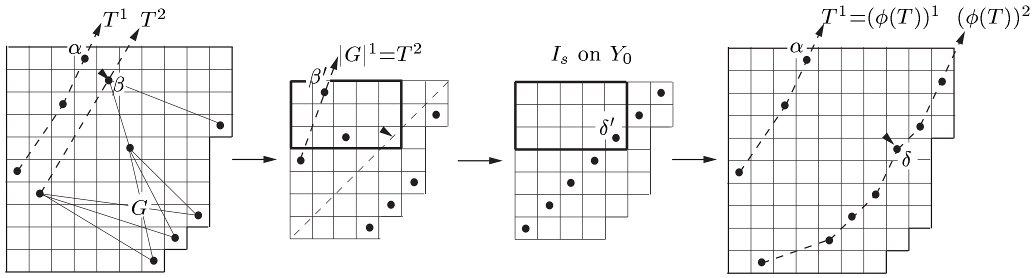

Figure 12: $T \in S_Y(312) \;\to\; T|_{Y_0} \;\to\; I_s \;\to\; \phi(T) \in S_Y(321)$

4.4 Surjectivity of ϕ.

To show that ϕ is surjective, we will first show

Lemma 12. *ϕ preserves T^1, i.e.* $(\phi(T))^1 = T^1$.

PROOF: Since the elements of T^1 are fixed by ϕ, it suffices to show that any other element $\delta \in \phi(T) \backslash T^1$, is (21)-dominated by some $\alpha \in T^1$, implying $\delta \notin (\phi(T))^1$.

Thus, start with $\delta \in \phi(T) \backslash T^1$ and pull it back to $\delta' \in I_s$ on Y_0 (cf. Fig. 12d-c.) Consider the rectangle $(Y_0)_{\delta'}$: since the cell of δ' is on the diagonal $d(Y_0)$, the proof of Lemma 3 implies that the transversal $T|_{Y_0}$ cannot sustain such a big empty rectangle. On the other hand, in the reduction $Y_0 = Y/_{T^1}$, the first sequence of the transversal $|G|$ coincides with the original second sequence T^2 in Y: $|G|^1 = T^2$ (cf. Fig. 12b.) Putting together these considerations implies the existence of some $\beta' \in |G|^1$ in the rectangle $(Y_0)_{\delta'}$. Pulling β' to $\beta \in T^2$ on Y, we deduce that some $\alpha \in T^1$ (21)-dominates β (cf. Fig. 12a.) Comparing the relative positions of α, β and δ in Y, we conclude that α (21)-dominates δ in $\phi(T)$. Therefore, $\delta \notin (\phi(T))^1$, and as noted above, this means $(\phi(T))^1 = T^1$. □

We can also think of ϕ in terms of the canonical decomposition in (3) of $T \in S_Y(312)$: replace every connected component C_i in G by the increasing sequence I_i in the Young subdiagram Y_i. Then $I_s = I_1 \sqcup I_2 \sqcup \ldots \sqcup I_k$. This works since the C_i's and the Y_i's are independent of each other and arranged in an increasing sequence in Y.

Lemma 13. *Given a fixed increasing sequence L of dots in Y, there is at most one transversal $T \in S_Y(321)$ for which $T^1 = L$.*

Proof: If $T \in S_Y(321)$ is such a transversal, then $T = T^1 \sqcup T^2$ with $T^1 = L$. Reducing $T/_L$ leaves T^2, which must be an increasing sequence in and a transversal of the resulting Young diagram $Y/_L$; yet, there is only one such sequence in $Y/_L$, namely, its diagonal sequence I_s. This uniquely defines T^2, and since the rest of T is the fixed L, it uniquely defines $T := L \sqcup I_s$ too. Of course, after putting back L and $I_s = T^2$ to Y, it may turn out that the newly added points of T^2 violate the definition of L by participating in T^1, so in this case there would be no $T \in S_Y(321)$ with $T^1 = L$. □

Proposition 3. *The map $\phi : S_Y(312) \to S_Y(321)$ is surjective.*

Proof: Let $Q \in S_Y(321)$, and decompose $Q = Q^1 \sqcup Q^2$. We will construct $T \in S_Y(312)$ such that $T^1 = Q^1$. For that, start with Q and apply any sequence of $(312) \to (321)$ moves on Q until there are no more (312)-patterns in Y. Denote the final transversal of Y by T. As an example, reverse the arrow ϕ in Figure 13 in Section 5: depending on the order of picking the (312)-patterns, one can get from $Q = (31524) \in S_Y(321)$ to $T_1 = (31542)$ or $T_2 = (32514)$ in $S_Y(312)$.

Each move replaces a (312)-pattern in Y with a (321)-pattern in Y by fixing the element playing the role of "3", and switching the other two elements as in $(12) \mapsto (21)$, and thereby increasing the number of *inversions* in the total transversal. Hence the number of moves cannot exceed $\binom{n}{2}$ and the sequence of moves eventually terminates with some $T \in S_Y(312)$.

The *first subsequences* of the original and of the final permutation coincide: $T^1 = Q^1$. Indeed, none of the moves $(\alpha_1\alpha_2\alpha_3) \approx (312) \mapsto (\alpha_1\alpha_3\alpha_2) \approx (321)$ changes the first subsequence, because α_1 (21)-dominates the other two elements, whether before or after the move. Hence α_2 and α_3 are not in and cannot land in the first subsequence via the moves, and their switch certainly does not affect in any way the existing first subsequence elements. We conclude that $T^1 = Q^1$.

By Lemma 12, ϕ preserves the first subsequence, so that applying ϕ to T yields $\phi(T) \in S_Y(321)$ with $(\phi(T))^1 = T^1 = Q^1$. But by Lemma 13, there is at most one transversal in $S_Y(321)$ with first subsequence Q^1, namely, Q. Thus, $\phi(T) = Q$ and ϕ is surjective. □

4.5 Conclusions

Proposition 3 implies that for all Young diagrams Y:

$$|S_Y(312)| \geq |S_Y(321)|,$$

which is one of the two inequalities in Theorem 1. Therefore, $(312) \succeq_s (321)$. Now Proposition 1 implies that $(312|\tau) \succeq_s (321|\tau)$ for any permutation τ; equivalently, for any Young diagram Y we have $|S_Y(312|\tau) \geq |S_Y(321|\tau)|$. Consequently, for all n:

$$|S_n(312|\tau)| \geq |S_n(321|\tau)|,$$

which completes half of Corollary 1. □

5 Strict Inequalities $S_Y(312) > S_Y(321)$

5.1 Examples of Strict Inequalities

Since $\phi : S_Y(312) \twoheadrightarrow S_Y(321)$, a strict inequality $|S_Y(312)| > |S_Y(321)|$ occurs exactly when for some $Q \in S_Y(321)$ the fiber $\phi^{-1}(Q) \subset S_Y(312)$ has more than 1 element. From the proof of Proposition 3, this happens exactly when two distinct $T_1, T_2 \in S_Y(312)$ have the same first subsequences: $(T_1)^1 = (T_2)^1$.

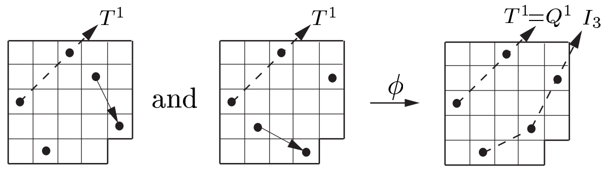

Figure 13: $T_1 = (31542),\ T_2 = (32514) \in S_{Y_5}(312) \xrightarrow{\phi} Q = (31524) \in S_{Y_5}(321)$

Example 1. We revisit the Young diagram $Y_5 = (5, 5, 5, 5, 4)$, mentioned in the Introduction. It is the smallest Young diagram on which (312) is less restrictive than (321): $|S_Y(312)| = 42 > 41 = |S_Y(321)|$. The two sets intersect in a large subset: $|S_Y(312, 321)| = 21$, and $\phi : S_Y(321) \twoheadrightarrow S_Y(312)$ acts as the identity map on this intersection. Indeed, if $T \in S_Y(312, 321)$, then $T = T^1 \sqcup T^2$, so that $I_s \equiv T^2 \nearrow$ and $\phi(T) = T^1 \sqcup I_s = T$. In addition, there are 19 transversals $U \in S_Y(321)$ whose preimages in $S_Y(312)$ consist of single elements $\phi^{-1}(U) \neq U$.

As expected, the map ϕ is non-invertible only on the remaining one transversal $Q \in S_Y(321)$, namely, $Q = (31524)$ (cf. Fig. 13, where all first subsequences are denoted by T^1.) Its preimage is $\phi^{-1}(Q) = \{T_1, T_2\}$ where $T_1 = (31542)$ and $T_2 = (32514)$. Note that $(T_1)^1 = (T_2)^1 (= \{3, 5\})$, which ensures that $\phi(T_1) = \phi(T_2) (= Q)$. Yet, the canonical decompositions of T_1 and T_2 into connected components differ: $T_1 = T^1 \sqcup \{1\} \sqcup \{4, 2\}$ and $T_2 = T^1 \sqcup \{2, 1\} \sqcup \{4\}$, causing two preimages of Q.

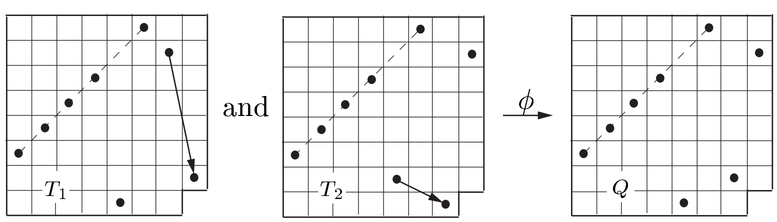

Figure 14: $T_1,\ T_2 \in S_{Y_n}(312) \xrightarrow{\phi} Q \in S_{Y_n}(321)$

Example 2. We extend Example 1 to all Y_n with $n \geq 5$. Let $T_1^n = (3, 4, ..., n-2, 1, n, n-1, 2)$ and $T_2^n = (3, 4, ..., n-2, 2, n, 1, n-1)$ (cf. Fig. 14.) It is easy to verify that T_1^n and T_2^n are (312)-avoiding on Y_n with the same first subsequence $(T_1^n)^1 =$

$(T_2^n)^1 = (3, 4, ..., n-2, n)$, and as such, they have the same image $Q = \phi(T_1^n) = \phi(T_2^n) = (3, 4, ..., n-2, 1, n, 2, n-1) \in S_{Y_n}(321)$. Hence $|S_{Y_n}(312)| > |S_{Y_n}(321)|$. Non-surprisingly, reducing Y_n along most of the first subsequence: $Y_n \big/ \{3, 4, ..., n-3\}$, we recover the permutations in Y_5 of Example 1.

5.2 Sufficient condition for strict inequality

Proposition 4. *If Y has an i-critical point with $i \geq 3$, then $|S_Y(312)| > |S_Y(321)|$.*

Proof: As in Example 2, for strict inequality it is necessary and sufficient to exhibit two distinct transversals $\bar{T}_1, \bar{T}_2 \in S_Y(312)$ with $(\bar{T}_1)^1 = (\bar{T}_2)^1$. Let P be an i-critical point of Y with $i \geq 3$. Starting from P, go down (resp. right) one cell and go left (resp. up) till hitting $d_0(Y)$: call this point S_1 (resp. S_2). With S_1S_2 as diagonal, we construct a subdiagram $Y(P)$ of Y such that $Y(P) \cong Y_{i+2}$ and P is the i-critical point of $Y(P)$. For example, in Figure 15a the subdiagram $Y(P) \cong Y_6$ is generated by the 4-critical point P; the dashed lines represent the diagonals $d_i(Y_6)$ for $0 \leq i \leq 4$.

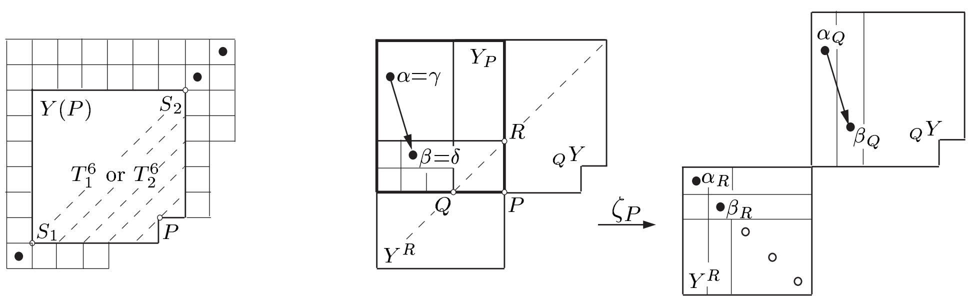

Figure 15: (a) $\bar{T}_j = (1, T_j^{i+2}, 8, 9)$ (b)-(c) $(\gamma\delta) \subset Y_P$

Now, put dots everywhere along $d(Y)$ outside of $Y(P)$. (In Fig. 15a, these dots represent 1, 8 and 9.) For $j = 1, 2$, insert T_j^{i+2} from Example 2 inside $Y(P)$ in order to obtain $\bar{T}_j$ on Y. It is immediate that $\bar{T}_1, \bar{T}_2 \in S_Y(312)$ and they have the same first subsequence, so that $\phi(\bar{T}_1) = \phi(\bar{T}_2)$, and hence $|S_Y(312)| > |S_Y(321)|$. □

5.3 Necessary condition for strict inequality $S_Y(312) > S_Y(321)$

We shall prove that strict inequalities are obtained, as Theorem 2 claims, only when Y has higher critical points. To this end, we first need to establish two technical recursive formulas for 2-critical points when the avoided pattern is $\sigma = (312)$ or (321).

5.3.1 Recursions for 2-critical points

Recall the points R and Q associated to P in the definition of the map ζ_P. When P is the *bottom* critical point of Y, Y^R and Y^Q are both *squares*, which makes the calculations below possible (cf. Fig. 15b.) Recursion (4) in Lemma 14 below reduces calculations from the larger Young diagram Y to the smaller ${}_QY$; yet, it is not very useful on its own since it also introduces the new sets $S_{\searrow_Q Y}(\sigma)$ and $S_{\nearrow_Q Y}(\sigma)$. Hence the necessity to prove recursion (5). Note the apparent similarity between these recursive formulas for $|S_Y(\sigma)|$ and $|S_{\searrow Y}(\sigma)|$.

Lemma 14. *Let Y be a Young diagram whose bottom critical point P is 2-critical. If there are k rows of Y below P, for $\sigma = (312)$ or (321) we have:*

$$\begin{aligned} |S_Y(\sigma)| &= c_{k+1} \cdot |S_{\searrow_Q Y}(\sigma)| + (c_{k+2} - c_{k+1}) \cdot |S_{\nearrow_Q Y}(\sigma)| \qquad (4) \\ |S_{\searrow Y}(\sigma)| &= c_k \cdot |S_{\searrow_Q Y}(\sigma)| + (c_{k+1} - c_k) \cdot |S_{\nearrow_Q Y}(\sigma)| \qquad (5) \end{aligned}$$

Proof: The 2-critical splitting from Lemma 6 implies:

$$|S_Y(\sigma)| = |S^{\nearrow}_{Y^R}(\sigma)| \cdot |S_{\nearrow_Q Y}(\sigma)| + |S^{\searrow}_{Y^R}(\sigma)| \cdot |S_{\searrow_Q Y}(\sigma)|.$$

Claim 1 below treats the special case of the square Y^R of size $k+2$. Substituting its results $|S^{\searrow}_{Y^R}(\sigma)| = c_{k+1}$ and $|S^{\nearrow}_{Y^R}(\sigma)| = c_{k+2} - c_{k+1}$, we readily arrive at the wanted recursion (4).

To prove (5), we restrict the map ζ_P to $S_{\searrow Y}(\sigma)$ in the 2-splitting isomorphism in (2):

$$\zeta_P(S_{\searrow Y}(\sigma)) \subset S^{\nearrow}_{\searrow Y^R}(\sigma) \times S_{\nearrow_Q Y}(\sigma) \sqcup S^{\searrow}_{\searrow Y^R}(\sigma) \times S_{\searrow_Q Y}(\sigma). \qquad (6)$$

As in the definition of ζ_P, we write $(\alpha\beta)$ for the 2-element subsequence of T inside Y_P. There are three possibilities for the initial decreasing subsequence $(\gamma\delta)$ of $T \in S_{\searrow Y}(\sigma)$.

Case 1. $(\gamma\delta)\searrow$ *is entirely in the rectangle* Y_P. Then $(\gamma\delta) = (\alpha\beta)\searrow$ and

$$\zeta_P(T) \in S^{\searrow}_{\searrow Y^R}(\sigma) \times S_{\searrow_Q Y}(\sigma),$$

with the extra condition that α_R occupies cell (1,1) and β_R occupies cell (2,2) of square Y^R (cf. Fig. 15b-c.) If avoiding $\sigma = (312)$, the remainder of the transversal in Y^R is completely determined as a decreasing subsequence (depicted in Fig. 15c via "$\circ$"), while avoiding (321) yields no possible completions in Y^R. Thus, the images $\zeta_P(T)$ are in 1-1 correspondence with $\{J_{k+2}\} \times S_{\searrow_Q Y}(\sigma)$ if $\sigma = (312)$, and there are 0 such if $\sigma = (321)$.

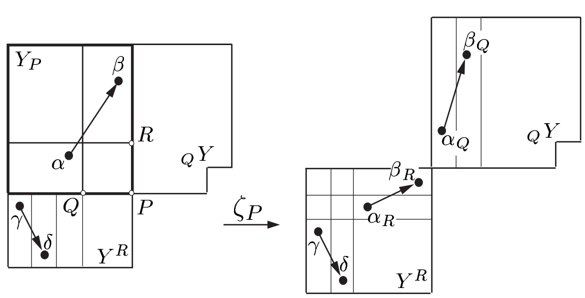

Figure 16: Case 2

Case 2. $(\gamma\delta)\searrow$ *is entirely in the square* Y^Q (cf. Fig. 16.) Then $(\gamma\delta) \cap (\alpha\beta) = \emptyset$, and hence $(\alpha\beta)$ can be $\nearrow$ or $\searrow$. In either case, the four elements $(\gamma\delta\alpha_R\beta_R)$ occupy the two leftmost columns and two top rows of Y^R. In the sub-factor $S^{\nearrow}_{\searrow Y^R}(\sigma)$ of (6),

$(\gamma\delta\alpha_R\beta_R) \approx (2134)$, while in the sub-factor $S^{\searrow}_{\searrow Y^R}(\sigma)$, $(\gamma\delta\alpha_R\beta_R) \approx (2143)$. Claim 2a-b implies that the number of images $\zeta_P(T)$ in these two subcases equals respectively $(c_{k+1} - c_k - k) \cdot |S_{\nearrow_Q Y}(\sigma)|$ or $(c_k - 1) \cdot |S_{\searrow_Q Y}(\sigma)|$.

Case 3. $\gamma \in Y_P$ *and* $\delta \in Y_Q$ (cf. Fig. 17.) Then $\gamma = \alpha$, and $(\alpha\beta)$ can be $\nearrow$ or $\searrow$. In the sub-factor $S^{\nearrow}_{\searrow Y^R}(\sigma)$, we have $(\alpha_R\delta\beta_R) \approx (213)$ where α_R occupies cell (2,1). Claim 2c implies that the number of images in this subcase is $k \cdot S_{\nearrow_Q Y}(\sigma)$. In the sub-factor $S^{\searrow}_{\searrow Y^R}(\sigma)$, $(\alpha_R\delta\beta_R) \approx (312)$ where α_R occupies cell (1,1). Thus, avoiding $\sigma = (312)$ yields 0 transversals in this subcase. For $\sigma = (321)$, the position of α_R allows for only one transversal on Y^R, namely, $T_1 = (k+2, 1, 2, ..., k+1)$ (depicted in Fig. 17b via "$\circ$"), and hence the images $\zeta_P(T)$ here are in 1-1 correspondence with $\{T_1\} \times S_{\searrow_Q Y}(\sigma)$.

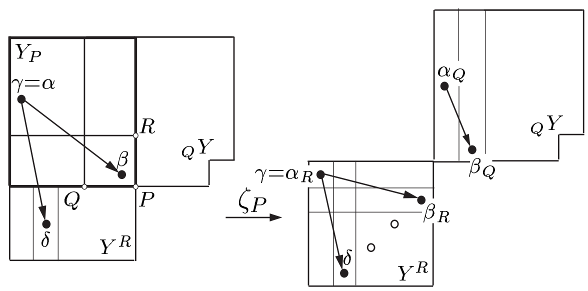

Figure 17: Case 3

Adding up the results in all three Cases, we obtain for each $\sigma = (312)$ and $\sigma = (321)$:

$$|\zeta_P(S_{\searrow Y}(\sigma))| = (0 + 1 + c_k - 1) \cdot |S_{\searrow_Q Y}(\sigma)| + (c_{k+1} - c_k - k + k) \cdot |S_{\nearrow_Q Y}(\sigma)|$$

Since ζ_P is injective, we derive the wanted recursion (5):

$$|S_{\searrow Y}(\sigma)| = |\zeta_P(S_{\searrow Y}(\sigma))| = c_k \cdot |S_{\searrow_Q Y}(\sigma)| + (c_{k+1} - c_k) \cdot |S_{\nearrow_Q Y}(\sigma)|. \quad \square$$

5.3.2 Calculations on the square Y^R

We show here all Claims from the proof of Lemma 14: they involve specific calculations on the square Y^R of size $k+2$. To simplify notation, we shall write α for α_R and β for β_R. Thus, $(\alpha\beta)$ and $(\gamma\delta)$ are the subsequences of T in the top two rows, respectively leftmost two columns, of Y^R.

Claim 1. *For* $\sigma = (312)$ *or* (321), $|S^{\searrow}_{Y^R}(\sigma)| = c_{k+1}$, *and hence* $|S^{\nearrow}_{Y^R}(\sigma)| = c_{k+2} - c_{k+1}$.

Proof: We calculate first $|S^{\searrow}_{Y^R}(\sigma)|$, so we assume that $(\alpha\beta)\searrow$ in Y^R. If $\sigma = (321)$, to avoid the situation of α and β playing the roles of "3" and "2" in a (321)-pattern in Y^R, β must be in the last column (and the second row) of Y^R. As such, β cannot participate in any (321)-pattern on Y^R, so that reducing along β we obtain a (321)-avoiding transversal T' on the rectangle $Y^R/_{\{\beta\}} = M_{k+1}$, without any further restrictions (cf. Fig. 18a-b.) The original transversal of Y^R can be reconstructed from T' by reinserting β in the last row and second column of Y^R. We have established that $S^{\searrow}_{Y^R}(321) \cong S_{k+1}(321)$, and hence $|S^{\searrow}_{Y^R}(321)| = c_{k+1}$.

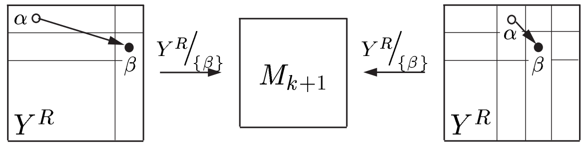

Figure 18: Claim 1

Similarly, if $\sigma = (312)$, α and β must be in adjacent columns in Y^R in order to avoid (312). But they are already in the top two rows of Y, so they are situated in diagonally-adjacent cells. Again, reducing along β we obtain a (312)-avoiding transversal T'' on the rectangle $Y^R/_{\{\beta\}} = M_{k+1}$, without any further restrictions (cf. Fig. 18c-b.) The original transversal of Y^R can be reconstructed from T'' by reinserting β in the second row and the column on the right of α's column in Y^R. We have established that $S^{\searrow}_{Y^R}(312) \cong S_{k+1}(312)$, and hence $|S^{\searrow}_{Y^R}(312)| = c_{k+1}$.

To finish the argument, we note that $S^{\nearrow}_{Y^R}(\sigma)$ is the complement of $S^{\searrow}_{Y^R}(\sigma)$ in $S_{Y^R}(\sigma)$, so that for $\sigma = (321)$ of (312) we have $|S^{\nearrow}_{Y^R}(\sigma)| = |S_{Y^R}(\sigma)| - |S^{\searrow}_{Y^R}(\sigma)| = c_{k+2} - c_{k+1}$. $\square$

Claim 2. *(a) There are $c_{k+1} - c_k - k$ transversals $T \in S^{\nearrow}_{\searrow Y^R}(\sigma)$ with $\{\gamma\delta\} \cap \{\alpha\beta\} = \emptyset$.*
(b) There are $c_k - 1$ transversals $T \in S^{\searrow}_{\searrow Y^R}(\sigma)$ with $\{\gamma\delta\} \cap \{\alpha\beta\} = \emptyset$.
(c) There are k transversals $T \in S^{\nearrow}_{\searrow Y^R}(\sigma)$ with $\alpha = \gamma$.

Proof: All transversals in question are elements of $S_{\searrow Y^R}(\sigma)$, i.e. $(\gamma\delta)\searrow$. As in the proof of Claim 1, either δ is in the last row (and second column) of Y^R ($\sigma = (321)$), or neighboring γ southeast-diagonally ($\sigma = (312)$). In either case, we reduce Y^R along δ to obtain equinumerant subsets of $Y^R/_{\{\delta\}} = M_{k+1}$ with the following restrictions (cf. Fig. 19):

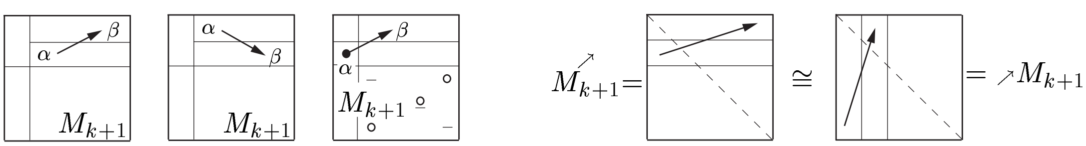

Figure 19: Claim 2. (a1)-(b1)-(c1)

$(\overset{\nearrow}{M}_{k+1})^t \cong \nearrow M_{k+1}$

(a1) all transversals of $S^{\nearrow}_{M_{k+1}}(\sigma)$ for which the top two elements $(\alpha\beta)\nearrow$ do not lie in the first column of M_{k+1} (occupied by γ);

(b1) all transversals of $S^{\searrow}_{M_{k+1}}(\sigma)$ for which the top two elements $(\alpha\beta)\searrow$ do not lie in the first column of M_{k+1} (occupied by γ);

(c1) all transversals of $S_{\overset{\nearrow}{M}_{k+1}}(\sigma)$, for which one of the top two elements $(\alpha\beta)\nearrow$ does lie in the first column of M_{k+1} ($\alpha = \gamma$ is that element.)

Let's start with case (c1). Since α is in position (2,1), the rows below α are filled either with an increasing (for $\sigma = (321)$) or with a decreasing (for $\sigma = (312)$) subsequence. In Fig. 19c, $\circ$ and $-$ denote, respectively, these increasing and decreasing subsequences. At the same time, β can be reinserted in any of the k possible cells of the top row of Y^R without creating any σ-patterns. Thus, the number of transversals in (c) is k. $\square$

Case (a1) is the complement of (c1) inside $S_{M_{k+1}}^{\nearrow}(\sigma)$. Since (321), (312) and Y^R are symmetric with respect to transposing across the northwest/southeast diagonal, and since $(S_{M_{k+1}}^{\nearrow}(\sigma))^t = S_{\nearrow M_{k+1}}(\sigma)$, we can use Claim 1 for $Y^R = M_{k+1}$ to calculate:

$$|S_{M_{k+1}}^{\nearrow}(\sigma)| = |S_{\nearrow M_{k+1}}(\sigma)| = c_{k+1} - c_k.$$

Therefore, the number of transversals in (a) equals $c_{k+1} - c_k - k$. $\square$

Finally, case (b1) misses only one transversal of the set $S_{M_{k+1}}^{\searrow}(\sigma)$: namely, when α is in position (1,1), without any more restrictions (cf. Fig 19b.) In such a situation, the rest of M_{k+1} is again filled either with an increasing or with a decreasing subsequence (respectively, for $\sigma = (321)$ and (312)). Thus, case (b1) counts 1 fewer transversals than $S_{M_{k+1}}^{\searrow}(\sigma)$. Using again the transposing argument and Claim 1 for $Y^R = M_{k+1}$, we conclude that the number of transversals in (b) equals

$$|S_{M_{k+1}}^{\searrow}(\sigma)| - 1 = |S_{\searrow M_{k+1}}(\sigma)| - 1 = c_k - 1. \quad \square$$

5.3.3 Conclusions for low-rank critical points

Lemma 15. *If Y has only 2-critical points, then $|S_{\searrow Y}(321)| = |S_{\searrow Y}(312)|$, $|S_{\nearrow Y}(321)| = |S_{\nearrow Y}(312)|$, and hence $|S_Y(321)| = |S_Y(312)|$.*

Proof: For the special case of a square $Y = M_{k+2}$ (which has no critical points), the first two equalities were proven in Claim 1 for the square Y_R, while the third equality is the well-known Wilf-equivalence $(312) \sim (321)$.

For the general case, we proceed by induction on the size n of Y. For $n \le 3$ there are no 2-critical points. Suppose that Y is of size $n \ge 4$, not a square, and has only 2-critical points. Let Y's bottom (2-)critical point be P. The Young subdiagram ${}_QY$ from Lemma 14 is of smaller size, and by construction, its critical points are all of Y's critical points, short of P. Applying induction to ${}_QY$, we have $|S_{\searrow_Q Y}(321)| = |S_{\searrow_Q Y}(312)|$ and $|S_{\nearrow_Q Y}(321)| = |S_{\nearrow_Q Y}(312)|$. Recursions (4)-(5) then imply $|S_Y(321)| = |S_Y(312)|$ and $|S_{\searrow Y}(321)| = |S_{\searrow Y}(312)|$. Since $S_{\nearrow Y}(\sigma)$ is the complement of $S_{\nearrow Y}(\sigma)$ in $S_Y(\sigma)$ for any σ, it also follows that $|S_{\nearrow Y}(321)| = |S_{\nearrow Y}(312)|$. $\square$

Proposition 5. $|S_Y(312)| = |S_Y(321)|$ *if Y has only i-critical points with $i \le 2$.*

Proof: If Y has some 0- or 1-critical point P, Proposition 2 implies that there is a 0- or 1-splitting for any permutation σ:

$$|S_Y(\sigma)| = |S_U(\sigma)| \cdot |S_V(\sigma)|,$$

where U and V are some Young subdiagrams of Y of smaller sizes. Since by construction the diagonals $d(U)$ and $d(V)$ lie on $d(Y)$, the set of critical points of U and V is the same as the set of critical points of Y, short of P. In other words, U and V again have only 0-, 1- or 2-critical points. Continuing the splitting process for every 0- or 1-critical point of the smaller diagrams, we arrive eventually at a splitting

$$|S_Y(\sigma)| = |S_{U_1}(\sigma)| \cdot |S_{U_2}(\sigma)| \cdots |S_{U_k}(\sigma)|,$$

where each subdiagram U_i has only 2-critical points (or no critical points at all). Lemma 15 guarantees that $|S_{U_i}(321)| = |S_{U_i}(321)|$ for all i, so that the products $|S_Y(321)|$ and $|S_Y(312)|$ are also equal. □

Finally, combining the results of Propositions 4-5, we derive the second necessary and sufficient condition in Theorem 2: $|S_Y(312)| > |S_Y(321)|$ if and only if Y contains an i-critical point with $i \geq 3$. □

6 Proof of the Inequality $S_Y(213) \leq S_Y(123)$

6.1 The (213)-decomposition

In [21], Stankova-West show $(213) \sim_s (132)$. Their proof introduces a special decomposition of the (213)-avoiding transversals on any Young diagram. Here we modify and extend this decomposition for our purposes, and use it later for comparing $S_Y(213)$ and $S_Y(123)$.

Definition 17. Let Y be a Young diagram and let c be a cell in the bottom row of Y. Start from the bottom left corner of c, draw a 45° ray in north-east direction until the ray intersects for the first time the border of Y, and use the resulting segment as the diagonal of a smaller subdiagram $\mathcal{A}_c$ of Y. Reducing Y along $\mathcal{A}_c$ leaves a subdiagram $\mathcal{B}_c = Y/_{\mathcal{A}_c}$. Thus, c determines a pair $(\mathcal{A}_c, \mathcal{B}_c)$ of Young subdiagrams of Y, called the (213)-*decomposition of* Y *induced by* c and denoted by $Y_{213}(c) = \mathcal{A}_c \otimes \mathcal{B}_c$. If a transversal $T \in S_Y$ is concentrated in $\mathcal{A}_c$ and $\mathcal{B}_c$, we say that T respects this (213)-decomposition of Y and we write $T = T|_{\mathcal{A}_c} \otimes T|_{\mathcal{B}_c}$.

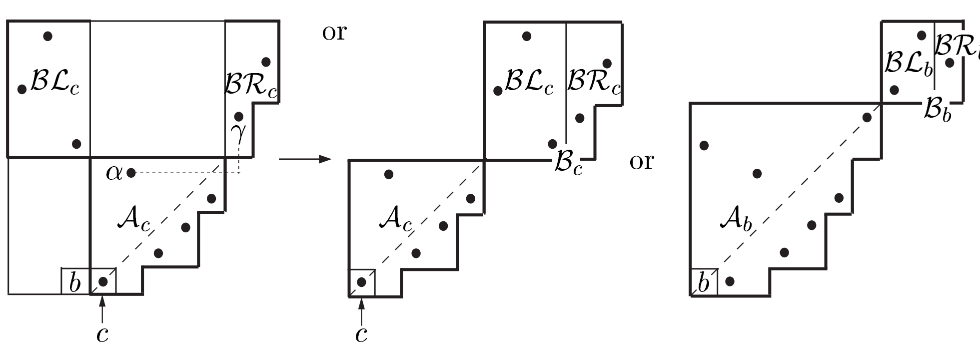

Figure 20: Minimal and non-minimal (213)-decompositions

Because of the 45° angle of the diagonal $d(\mathcal{A}_c)$, the Young subdiagram $\mathcal{A}_c$ is proper, and hence the reduction along it, $\mathcal{B}_c$, is also proper. The smallest $\mathcal{A}_c$ can be is the cell c: this happens when c is the rightmost cell in the bottom row of Y. The decomposition $Y_{213}(c)$ is trivial exactly when $\mathcal{B}_c = \emptyset$: this happens when c is the bottom left corner of Y and Y has no 0-critical points. In such a case, $Y_{213}(c) = \mathcal{A}_c$. While $\mathcal{A}_c$'s rows and columns are **not** interspersed with "outside" rows or columns from $Y \backslash \mathcal{A}_c$, in general, $\mathcal{B}_c$ splits into two parts: $\mathcal{B}_c = \mathcal{BR}_c + \mathcal{BL}_c$ where $\mathcal{BR}_c$ is to the left and above $\mathcal{A}_c$ and $\mathcal{BL}_c$ is to the right and above $\mathcal{A}_c$ (cf. Fig. 20.)

The name "(213)-decomposition" comes from the fact that all (213)-avoiding transversals T respect at least one (213)-decomposition of Y.

Proposition 6 (Stankova-West). *Let c be a bottom cell in Y, and let $T \in S_Y(213)$ have its bottom element in cell c. Then T respects the (213)-decomposition $Y_{213}(c) = \mathcal{A}_c \otimes \mathcal{B}_c$, and hence it decomposes as $T = T|_{\mathcal{A}_c} \otimes T|_{\mathcal{B}_c}$. Conversely, if $T \in S_Y$ respects this (213)-decomposition, and if the restrictions $T|_{\mathcal{A}_c}$ and $T|_{\mathcal{B}_c}$ are each (213)-avoiding on $\mathcal{A}_c$ and $\mathcal{B}_c$, respectively, then T is (213)-avoiding on all of Y.*

For some $T \in S_Y(213)$, it is possible that different bottom cells c_i's induce different (213)-decompositions $T = T|_{\mathcal{A}_{c_i}} \otimes T|_{\mathcal{B}_{c_i}}$. However, for only one (213)-decomposition T's bottom element $\hbar$ is in $\mathcal{A}_c$'s bottom left corner; we call this *the minimal* (213)-*decomposition of* T since all other (213)-decompositions will have $\hbar$ somewhere further to the left and hence their components $\mathcal{A}_{c_i}$ will contain properly $\mathcal{A}_c$. For example, Figure 20 shows the minimal $T_{213}(c) = (15234) \otimes (35124)$ and a non-minimal $T_{213}(b) = (6152347) \otimes (132)$ decompositions of $T = (8, 10, 6, 1, 5, 2, 3, 4, 7, 9)$. Proposition 6 implies that every $T \in S_Y(213)$ respects its minimal (213)-decomposition. More generally,

Definition 18. For any transversal $T \in S_Y$, the (213)-decomposition of T whose $\mathcal{A}_c$-component is contained properly in the $\mathcal{A}$-components of any other (213)-decomposition of T is called the *minimal* (213)-*decomposition of* T. If all (213)-decompositions of T are trivial, i.e. $T = T|_{\mathcal{A}_c}$, we say that T is (213)-*indecomposable.*

When it is irrelevant which bottom cell c induces some (213)-decomposition of T, we shall drop c from the notation, e.g. $T = T|_{\mathcal{A}} \otimes (T|_{\mathcal{BL}} + T|_{\mathcal{BR}})$.

For a general transversal $T \in S_Y(\sigma)$ which (213)-decomposes as $T = T|_{\mathcal{A}} \otimes T|_{\mathcal{B}}$, it is evidently true that $T|_{\mathcal{A}}$ and $T|_{\mathcal{B}}$ each avoid σ on $\mathcal{A}$ and $\mathcal{B}$, respectively. The converse is false in general: $T|_{\mathcal{A}} \times T|_{\mathcal{B}} \in S_{\mathcal{A}}(\sigma) \times S_{\mathcal{B}}(\sigma)$ does not imply $T \in S_Y(\sigma)$. Yet, for special cases of σ, the converse is true. We have seen in Proposition 6 that $\sigma = (213)$ is such a special case. Another one is $\sigma = (123)$.

Lemma 16. *If $T \in S_Y$ has a (213)-decomposition $T = T|_{\mathcal{A}} \otimes T|_{\mathcal{B}}$ on Y, and if each subtransversal avoids (123): $T|_{\mathcal{A}} \in S_{\mathcal{A}}(123)$ and $T|_{\mathcal{B}} \in S_{\mathcal{B}}(123)$, then $T \in S_Y(123)$.*

Proof: Suppose that T has a (123)-subsequence $(\alpha\beta\gamma)$ landing inside Y. Because of the hypothesis on the two subtransversals, this pattern must involve elements from both subdiagrams $\mathcal{A}$ and $\mathcal{B}$. Since $\mathcal{A}$ is entirely below $\mathcal{B}$, the element α (which plays the role of "1") must come from $\mathcal{A}$. But since α is also the leftmost element of the pattern, it eliminates any participation coming from $\mathcal{BL}$. This forces the last element γ to come from $\mathcal{BR}$ (cf. Fig. 20a.) Yet, no two elements of $\mathcal{A}$ and $\mathcal{BR}$ can participate in any pattern in Y: by construction of the 45° diagonal of $\mathcal{A}$, $\mathcal{BR}$ is entirely to the right and above $\mathcal{A}$, forcing the rows of $\mathcal{A}$ and the columns of $\mathcal{BR}$, to intersect outside Y.

We conclude that a (123)-pattern is impossible, so that $T \in S_Y(123)$. □

We shall see below that the (213)-decompositions of any $T \in S_Y$ on Y are preserved by $(213) \to (123)$- and $(123) \to (213)$ moves on T, which will allow us to prove eventually the desired inequality $|S_Y(123)| \geq |S_Y(213)|$.

6.2 Special (213)-decompositions

Fix a (213)-decomposition $T = T|_{\mathcal{A}_c} \otimes T|_{\mathcal{B}_c}$ of a transversal $T \in S_Y$. If $\mathcal{BR}_c = \emptyset$, we obtain a generalization of the decomposability Definition 7 of a permutation $\sigma \in S_n$. To keep up with the previous conventions, we say in this case that the transversal T is *decomposable*, and also write $T = (T|_{\mathcal{BL}_c} \big| T|_{\mathcal{A}_c})$. The two blocks of T are arranged in a northwest/southeast fashion.

On the other hand, if $\mathcal{BL}_c = \emptyset$, not only the given transversal decomposes as $T = T|_{\mathcal{A}_c} \otimes T|_{\mathcal{BR}_c}$, but **any** transversals T' of Y respects this decomposition. Indeed, in this case, c is the bottom left corner cell of Y and Y contains a 0-critical point P, and hence any transversal $T' \in S_Y$ has this 0-splitting with respect to P. Here the two blocks of T' are arranged in a southwest/northeast fashion.

Thus, all decompositions in this paper are (213)-decompositions or special cases of it.

6.3 $\sigma \to \tau$ moves on $T \in S_Y$

Since each $(213) \to (123)$ move decreases the number of inversions in T, any sequence of $(213) \to (123)$ moves eventually *terminates* with some $T' \in S_Y(213)$. Similarly, a sequence of $(123) \to (213)$ moves terminates with some $T'' \in S_Y(123)$.

Conjecture 1. *Starting with a transversal $T \in S_Y$, all sequences of $(213) \to (123)$ moves terminate in the same transversal $T' \in S_Y(213)$.*

The conjecture, if proven, would give a well-defined map $S_Y \to S_Y(213)$, which could be restricted to a map $\xi : S_Y(123) \to S_Y(213)$. To show then that ξ is surjective, we would start with any $T' \in S_Y(213)$ and apply any sequence of $(123) \to (213)$ moves on T' until it terminates with some $T \in S_Y(123)$. Reversing the sequence of moves would yield a sequence of $(213) \to (123)$ moves on T that terminates in T'. The definition of ξ would then give $\xi(T) = T'$ and hence ξ would be surjective, from where $|S_Y(123)| \geq |S_Y(213)|$.

In the next Subsection 6.4 we proceed in a different way by defining a *section* of the conjectured map ξ, i.e. a map $\psi : S_Y(213) \to S_Y(123)$ such that $\xi \circ \psi = \mathrm{id}_{S_Y(213)}$. As opposed to ξ, the map ψ will be given by a *specific* sequence of $(123) \to (213)$ moves which can be retraced back. The latter will then readily imply $|S_Y(213)| \leq |S_Y(123)|$.

Analogously to the *first* and *second* subsequences of $T \in S_Y$ in Definition 12, in working with (123)-avoidance we will need the following terminology.

Definition 19. Let $T \in S_Y$. The *primary* subsequence $\dot{T}$ of T consists of all elements which are **not** (12)-dominated in T. The *secondary* subsequence $\ddot{T}$ of T consists of all elements of T which are (12)-dominated by something in $\dot{T}$ and by nothing in $T \backslash \dot{T}$.

In particular, if $T \in S_Y(123)$ where Y is a *square*, then T is the disjoint union of its two decreasing subsequences $\dot{T}$ and $\ddot{T}$: $T = \dot{T} \sqcup \ddot{T}$.

6.4 Definition of $\psi : S_Y(213) \to S_Y(123)$

We define ψ by *induction on the size of* Y. When Y is a single cell, ψ is the identity map. Suppose we have defined ψ for all Young diagrams of size $< n$. Fix a Young diagram Y of size $n \geq 2$ and $T \in S_Y(213)$. Throughout this subsection, we shall refer to the element in the bottom row of Y as $\hbar$ and denote by $Y/_{\hbar}$ the reduction along it. There are two cases to discuss, depending on whether T is (213)-decomposable or not.

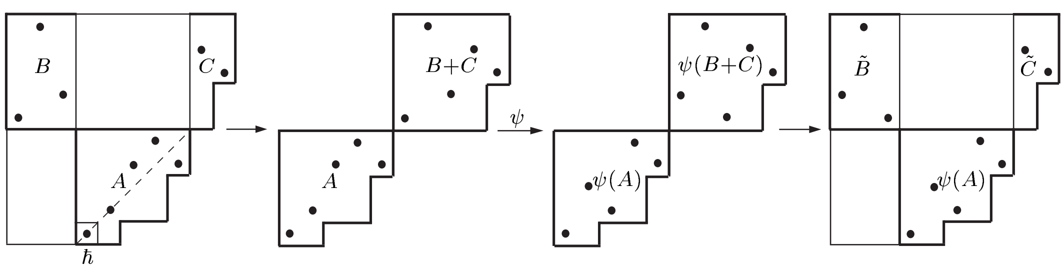

Figure 21: $\psi(A \otimes (B+C)) = \psi(A) \otimes \psi(B+C) = \psi(A) \otimes (\tilde{B} + \tilde{C})$

6.4.1 Case 1

T has a *non-trivial* (213)-decomposition; so consider its minimal (213)-decomposition $T = T|_{\mathcal{A}_\hbar} \otimes (T|_{\mathcal{BL}_\hbar} + T|_{\mathcal{BR}_\hbar})$, denoted for simplicity as $T = A \otimes (B + C)$. Since both $\mathcal{A}_\hbar$ and $\mathcal{B}_\hbar$ are of sizes $< n$, $\psi(A)$ and $\psi(B+C)$ are well-defined by induction. We can further split $\psi(B+C) = \tilde{B} + \tilde{C}$ where $\tilde{B} = \psi(B+C)|_{\mathcal{BL}_\hbar}$ and $\tilde{C} = \psi(B+C)|_{\mathcal{BR}_\hbar}$ occupy respectively the same columns as the original B and C. We define $\psi(T)$ to be the (213)-decomposable transversal of Y given by $\psi(T) = \psi(A) \otimes (\tilde{B} + \tilde{C})$. For example, the first arrow in Figure 21 signifies the (213)-decomposition $Y_{213}(\hbar)$ which combines subboards B and C; the second arrow applies ψ to each of A and $B + C$, and the third arrow splits $\psi(B+C)$ as $\tilde{B}$ and $\tilde{C}$ back into the original Young diagram Y.

6.4.2 Case 2

T is (213)-indecomposable. As remarked earlier, this can happen only if $\hbar$ is in the bottom left corner cell of Y and Y has no 0-critical points. Consider the reduction $Y/_{\hbar}$, and let $D = T|_{Y/_\hbar}$ be its transversal obtained from T by removing $\hbar$. $Y/_{\hbar}$ breaks up into two parts:

the rectangle $(Y/\hbar)'$ which lies over the bottom row of Y, and the remaining subboard $(Y/\hbar)''$ (cf. Fig. 22a.) We define $\psi(T)$ in two steps: $T \to T_1 \to \psi(T)$.

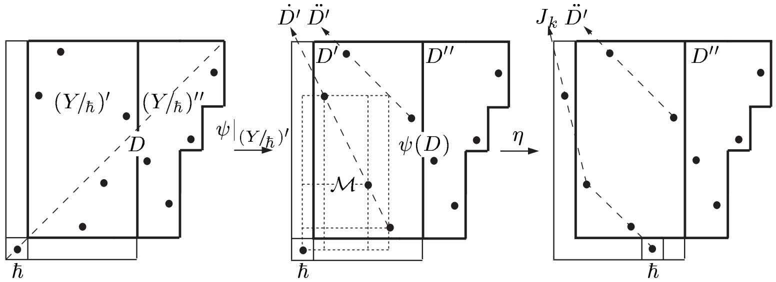

Figure 22: Definition of ψ in Case 2: $T \to T_1 \to \psi(T) = \eta(T_1)$

Since $Y/\hbar$ is of size $n-1$ and $D \in S_{Y/\hbar}(213)$, by induction $\psi(D)$ is defined as an element of $S_{Y/\hbar}(123)$. Let T_1 be the transversal of Y obtained from $\psi(D)$ by prepending $\hbar$ in its bottom left corner; symbolically, $T_1 = {}_\hbar[\psi(D)]$ (cf. Fig. 22b.)

The partition $Y/\hbar = (Y/\hbar)' + (Y/\hbar)''$ induces a partition of the transversal $\psi(D) = D' + D''$. Since D' is a (123)-avoiding partial transversal of the *rectangle* $(Y/\hbar)'$, then D' splits into its *primary* $\dot{D}'$ and *secondary* $\ddot{D}'$ decreasing subsequences, as in Fig. 22b. Let $\mathcal{M}$ be the (square) submatrix of Y induced by $\hbar$ and $\ddot{D}'$, and let $T_1|_{\mathcal{M}} = {}_\hbar[\ddot{D}']$ be $\mathcal{M}$'s transversal induced by T_1. For instance, Figure 22b shows $T_1|_{\mathcal{M}} = (1432)$ and depicts $\mathcal{M} \cong M_4$ via dotted lines. If $k-1$ is the length of $\ddot{D}'$ $(k \geq 1)$, then $\mathcal{M} \cong M_k$ and $T_1|_{\mathcal{M}} \approx (1, k, k-1, \ldots, 3, 2)$. Let $\eta(T_1)$ be the transversal of Y obtained from T_1 by replacing $T_1|_{\mathcal{M}} \mapsto J_k$ (cf. Fig. 22c.)

Set $\psi(T) := \eta(T_1)$ as the desired transversal in S_Y, and define the map $\psi : S_Y(213) \to S_Y$ as the composition $\psi = \eta \circ ({}_\hbar[\psi|_{Y/\hbar}])$.

6.5 Properties of $\psi : S_Y(213) \to S_Y(123)$

Proposition 7. *For any Y the map $\psi : S_Y(213) \to S_Y$ satisfies:*

1. *ψ is a sequence of $(123) \to (213)$ moves;*
2. *ψ maps to $S_Y(213)$;*
3. *ψ is injective.*

A key idea in the proof of Proposition 7 is the following lemma:

Lemma 17. *Given $T \in S_Y$, all $(213) \to (123)$ and $(123) \to (213)$ moves on T respect any (213)-decomposition of T. Consequently, a sequence of such moves preserves (213)-decomposability and (213)-indecomposability of transversals.*

Proof: Suppose $T \in S_Y$ has a (213)-decomposition $T = A \otimes (B + C)$. The proof of Proposition 6 implies that a (213)-pattern inside Y can involve elements only inside A or only inside $B + C$. Thus, a $(213) \to (123)$ move occurs entirely in A or in $B + C$, and hence it respects any (213)-decomposition of T. Using the proof of Lemma 16, the same reasoning shows that $(123) \to (213)$ moves also respect any (213)-decomposition of T.

It remains to prove that the two types of moves map indecomposable to indecomposable transversals. To the contrary, if, say, a $(123) \to (213)$ move maps an indecomposable T to a decomposable T', then the reverse $(213) \to (123)$ move would map $T' \to T$ and hence violate the preservation of (213)-decompositions proved above. Thus, the two types of moves preserve (213)-indecomposability too. $\square$

The properties of ψ claimed in Proposition 7 are trivial for size 1 Young diagrams. In the proof of Proposition 7, we assume by induction that the map ψ satisfies the three required properties on all Young diagrams of size smaller than n, and we fix a transversal $T \in S_Y(213)$ for some Y of size n.

6.5.1 Proof of Proposition 7, Part (1):

Suppose T has a non-trivial (213)-decomposition, so take the minimal such decomposition $T = A \otimes (B + C)$. This is Case 1 of ψ's definition ψ, where ψ consists of a move inside A and a move inside $B + C$. Hence ψ respects this minimal (213)-decomposition. By induction, $\psi(A)$ and $\psi(B+C)$ are each obtained (independently) by $(123) \to (213)$ moves. Consequently, $\psi(T)$ is obtained by the composition of all of these $(123) \to (213)$ moves.

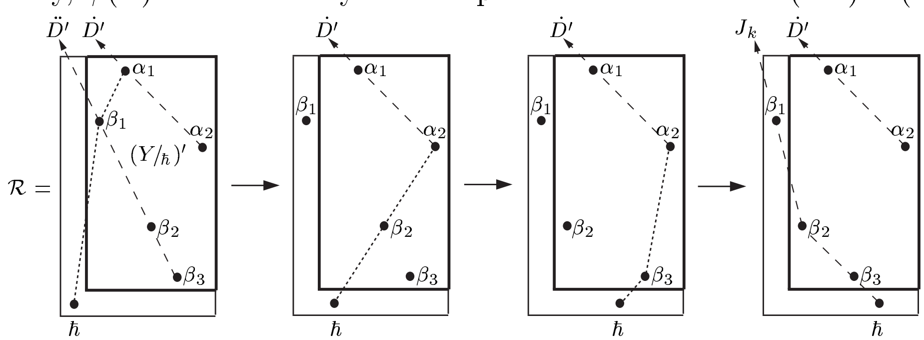

Figure 23: $(\hbar, \beta_1, \beta_2, \beta_3) \to (\beta_1, \hbar, \beta_2, \beta_3) \to (\beta_1, \beta_2, \hbar, \beta_3) \to (\beta_1, \beta_2, \beta_3, \hbar)$ in $\mathcal{R}$

Suppose now that T is (213)-indecomposable. This is Case 2 of ψ's definition. By induction on D's size, $\psi(D)$ and therefore T_1 are obtained by $(123) \to (213)$ moves inside D. It remains to show that $\psi(T_1|_{\mathcal{M}}) = J_k$ can be obtained from $T_1|_{\mathcal{M}}$ via such moves too. Recall that each $\beta \in \ddot{D}'$ is (12)-dominated in the rectangle $(Y/\hbar)'$ by some $\alpha \in \dot{D}'$. Thus, as long as $\hbar$ is before β ($\hbar$ still in the bottom row of Y), then $(\hbar\beta\alpha)$ is a (123)-pattern inside $(Y/\hbar)'$, and hence in Y. So the move $(\hbar\beta\alpha) \mapsto (\beta\hbar\alpha)$ is a $(123) \to (213)$ move which leaves $\hbar$ still in the bottom row of Y. Let $\ddot{D}' = (\beta_1, \beta_2, ..., \beta_{k-1})\searrow$. Then $(\hbar, \beta_1, \beta_2, ..., \beta_{k-1}) \approx (1, k, k-1, ..., 3, 2)$. Using the above reasoning, we can switch $\hbar$ consecutively with each of the β_i's via some $(123) \to (213)$ move. Figure 23 depicts this situation in the rectangular part $\mathcal{R} = {}_{\hbar}[(Y/\hbar)']$ of Y projecting onto Y's bottom row, where each $(123) \to (213)$ move is marked with a dotted line. Hence, the sequence

$$(\hbar, \beta_1, \beta_2, ..., \beta_{k-1}) \to (\beta_1, \hbar, \beta_2, ..., \beta_{k-1}) \to (\beta_1, \beta_2, \hbar, ..., \beta_{k-1}) \to \cdots \to (\beta_1, \beta_2, ..., \beta_{k-1}, \hbar)$$

is a composition of $(123) \to (213)$ moves, and so is ψ. $\square$

6.5.2 Proof of Proposition 7, Part (2):

In Case 1 of ψ's definition, by induction $\psi(A)$ and $\psi(B+C)$ are both (123)-avoiding. Lemma 16 implies that no new (123)-pattern can be introduced in the (213)-decomposition $\psi(T) = \psi(A) \otimes (\tilde{B} + \tilde{C})$. We conclude that $\psi(T) \in S_Y(123)$.

In Case 2 of ψ's definition, by induction $\psi(D)$ avoids (123). The only (123)-patterns in Y before applying η can occur because of $\hbar$ being prepended to D's bottom left corner, and hence any such pattern can appear only in the rectangle $\mathcal{R} = {}_{\hbar}[(Y/_{\hbar})']$ (cf. Fig. 23.) A (123)-pattern in $\mathcal{R}$ is of the form $(\hbar\beta\alpha)$ with $\alpha \in \dot{D}'$, $\beta \in \ddot{D}'$ and α (12)-dominates β in $\mathcal{R}$. The map η eliminates all these (123)-patterns by shifting $\ddot{D}'$ horizontally to the left and $\hbar$ to the right until $\hbar$ is after all of $\ddot{D}'$. Thus, after η is applied, $\hbar$ cannot participate in any more (123)-patterns in $\psi(T)$.

It remains to show that η has not created any new (123)-patterns $(\alpha\beta\gamma)$ which do not involve $\hbar$. Since η preserves $\dot{D}'$ and D'', such (123)-pattern must involve an element of $\eta(\ddot{D}')$; in fact, $\alpha \in \eta(\ddot{D}')$ since the elements of $\eta(\ddot{D}')$ do not (12)-dominate anything and hence they can play only the role of "1" in a (123)-pattern. Since $\eta(\ddot{D}')\searrow$ and $\dot{D}'\searrow$, at most one element of each can participate in this (123)-pattern. Finally, at most one element of D'' can participate too; indeed, suppose two elements β and γ of D'' participate in the (123)-pattern $(\alpha\beta\gamma)$ with $\alpha \in \eta(\ddot{D}')$ (cf. Fig. 24a.) If $\alpha_1 = \eta^{-1}(\alpha)$, then $(\alpha_1\beta\gamma)$ would also be a (123)-pattern in $\psi(D)$ since α_1 is a horizontal shift of α to the right but still inside $(Y|_{\hbar})'$ and before $\beta, \gamma \in (Y|_{\hbar})''$. This is a contradiction with the inductive assumption that $\psi(D)$ is (123)-avoiding. We conclude that at most one element of D'' can participate in the (123)-pattern, and therefore $(\alpha\beta\gamma)$ is formed by $\alpha \in \eta(\ddot{D}')$, $\beta \in \dot{D}'$, and $\gamma \in D''$.

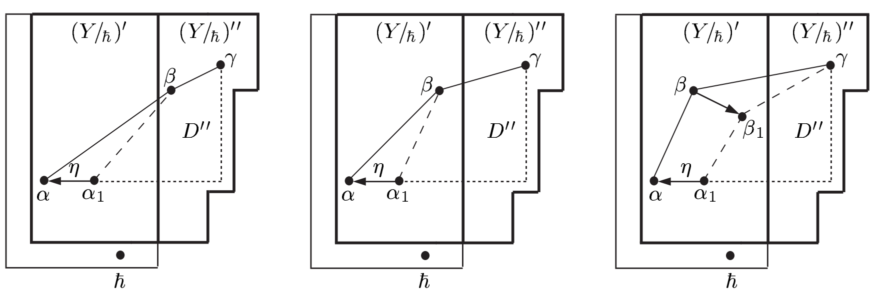

Figure 24: $\psi : S_Y(213) \to\in S_Y(123)$

If $\alpha_1 = \eta^{-1}(\alpha) \in \ddot{D}'$ is *before* β, then α_1 is (12)-dominated by β (cf. Fig. 24b) and so $(\alpha_1\beta\gamma) \approx (123)$ in $\psi(D)$, a contradiction. If α_1 is *after* β, then α_1 must be (12)-dominated by some $\beta_1 \in \dot{D}'$, $\beta_1 \neq \beta$ (cf. Fig. 24c.) Since α_1 is after β, and β_1 is after α, then β_1 is after β. Because $\beta, \beta_1 \in \dot{D}'\searrow$, it follows that $(\beta\beta_1)\searrow$. Further, $\gamma \in D''$, hence γ comes after both β and β_1. Finally, $\beta_1 < \beta < \gamma$ implies that $(\alpha_1\beta_1\gamma)\nearrow$ is a (123)-subpattern of $\psi(D)$ landing in Y: indeed, the intersection of γ's column and α_1's row is the same as the intersection of γ's column and α's row, and the latter is inside Y by the assumption that $(\alpha\beta\gamma)$ is a (123)-pattern in Y. The existence of such $(\alpha_1\beta_1\gamma)$ contradicts (123)-avoidance of $\psi(D)$.

Therefore, η gets rid of all (123)-patterns involving $\hbar$ and does not introduce any new (123)-patterns, so that $\eta(T_1)$ is (123)-avoiding and $\psi(T) \in S_Y(123)$. $\square$

6.5.3 Proof of Proposition 7, Part (3):

Let $T_2 \in \psi(S_Y(213)) \subset S_Y(123)$ for some Young diagram Y of size n, and let $T \in S_Y(213)$ be any preimage of T_2, i.e. $\psi(T) = T_2$. We will show that T can be recovered uniquely from T_2.

Case 1. Suppose that T_2 is (213)-*decomposable*, and let $T_2 = \tilde{A} \otimes (\tilde{B} + \tilde{C})$ be T_2's minimal (213)-decomposition. By Property 1 of ψ, $T \overset{\psi}{\mapsto} T_2$ is a sequence of $(123) \to (213)$ moves; inverting each of these moves, we obtain a sequence of $(213) \to (123)$ moves that takes $T_2 \mapsto T$. By Lemma 17, T respects the (213)-decomposition of T_2; moreover, the induced (213)-decomposition $T = A \otimes (B + C)$ is also minimal, i.e. $\hbar \in T$ is in the bottom left corner of A (cf. Fig. 21.) Thus, $\psi(A \otimes (B + C)) = \tilde{A} \otimes (\tilde{B} + \tilde{C})$ where $\psi(A) = \tilde{A}$ and $\psi(B + C) = \tilde{B} + \tilde{C}$ by the ψ's definition in Case 1. By induction, ψ is injective on smaller size Young diagrams, so that A and $B + C$ can be recovered from $\tilde{A}$ and $\tilde{B} + \tilde{C}$. Finally, since the decomposition of T is determined by T_2, B and C themselves can be recovered uniquely from $B + C$. We conclude that T can be recovered uniquely from T_2.

Case 2. Suppose that T_2 is (213)-indecomposable. As in Case 1, ψ must have preserved this property, i.e. T is also (213)-indecomposable. Since $T \in S_Y(213)$, this implies that $\hbar$ in T is in the bottom left corner of Y. By ψ's definition in Case 2, there is an intermediate $T_1 = {}_\hbar[\psi(D)]$ such that $\psi(T) = \eta(T_1) = T_2$ (cf. Fig. 22.) We will first show that T_1 is recoverable from T_2.

In T_2 we can uniquely determine D'' as the subtransversal in the part of Y that does not project on the bottom row of Y. The remainder $\tilde{D}$ of T_2 projects onto the bottom row of Y and lies in the rectangle $\mathcal{R}$. Since T_2 avoids (123), $\tilde{D}$ splits into its primary and secondary subsequences, $\tilde{D}_1$ and $\tilde{D}_2$, respectively. Note that $\hbar$ in $\tilde{D}$ is (12)-dominated: being in the bottom row of $\mathcal{R}$, the only way for $\hbar$ not to be (12)-dominated is to be in the rightmost (bottom) cell of $\mathcal{R}$; but then T_2 would be decomposable, contradicting our assumption in this case. Thus, $\hbar \in \tilde{D}_2$.

Now consider T_1. By η's definition in Case 2, $T_1 = {}_\hbar[\ddot{D}' + \dot{D}' + D'']$, where η fixes $\dot{D}'$ and D'', slides $\ddot{D}'$ to the left, and slides $\hbar$ to the right until $\hbar$ is after $\ddot{D}'$. Sliding $\ddot{D}'$ to the left leaves all of its elements (12)-dominated by some elements in $\dot{D}'$, and as we argued above, it makes $\eta(\hbar)$ also (12)-dominated in T_2. In other words, $\eta(\dot{D}') = \dot{D}' = \tilde{D}_1$ and $\eta({}_\hbar[\ddot{D}']) = \tilde{D}_2$. Thus, to recover T_1 from T, we keep D'' and $\tilde{D}_1$, and switch horizontally the places of $\tilde{D}_2 \backslash \hbar$ and $\hbar$. Note that at this point $\hbar \in T_1$ must be in the bottom left corner of Y by ψ's definition in Case 2.

To recover T from T_1, note that by induction ψ is injective on $Y/_\hbar$, so that $\psi(D)$ in T_1 could have come only from one transversal D; appending $\hbar$ at the bottom left corner of D gives the unique preimage $T = {}_\hbar[D] \in S_Y(213)$. $\square$

6.6 Conclusions

Subsubsections 6.5.1-3 complete inductively the proof of Proposition 7. The latter implies that $\psi : S_Y(213) \hookrightarrow S_Y(123)$ is injective for all Young diagrams Y. Thus, $|S_Y(213)| \leq |S_Y(123)|$ and Theorem 1 is proven. $\square$

We leave the following questions to the reader for further study. For which pairs of permutations σ and τ in S_k can a map $\psi_Y : S_Y(\sigma) \to S_Y(\tau)$ be well-defined via $\sigma \to \tau$ moves? What properties does ψ possess in such cases?

7 Strict Inequalities $|S_Y(213)| < |S_Y(123)|$

Below we refer to the notation from the definition of the map $\psi : S_Y(213) \hookrightarrow S_Y(123)$ in Section 6; in particular, $\psi(T) = \eta(T_1) = T_2$ for any (213)-indecomposable $T \in S_Y(213)$.

Lemma 18. *If Y has an i-critical point with $i \geq 2$ and no 0- and 1-critical points, some (213)-indecomposable $T_2 \in S_Y(123)$ is not invertible under η and hence under ψ.*

Proof: Since $i \geq 2$, the size of Y is $n \geq 4$. Place α in position $(2,1)$, $\hbar$ in $(n,2)$ and β in $(1,n)$, and set $\overline{Y} = Y/\{\alpha, \hbar, \beta\}$. The hypotheses on Y imply that $\overline{Y}$ is non-empty and has no 0-critical points. Since $(12) \sim_s (21)$ and since there is obviously exactly 1 transversal of $\overline{Y}$ that avoids (21) (namely, the diagonal transversal), there is also exactly 1 transversal $\overline{T}$ of $\overline{Y}$ that avoids (12) (cf. Fig. 25a.) Thus, $T_2 = \{\alpha, \hbar, \overline{T}, \beta\}$ is a transversal of Y. We claim that $T_2 \in S_Y(123) \backslash \psi(S_Y(213))$.

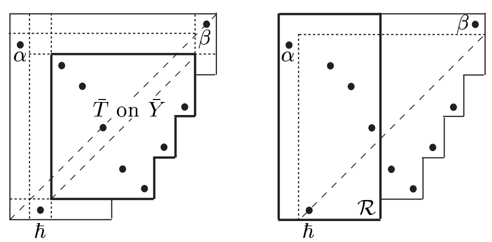

Figure 25: $T_2 \in S_Y(123) \backslash \psi(S_Y(213))$

To show that T_2 avoids (123), note that the position of α in Y precludes it from participating in any such pattern. Moreover, $\hbar$ and β cannot simultaneously participate in a (123)-pattern since β's column and $\hbar$'s row do not intersect inside the non-square Y. Yet, at most 1 element from $\overline{T}$ can participate in a (123)-pattern due to the (12)-avoidance of $\overline{T}$. This does not leave enough elements of T_2 to participate in a (123)-pattern in Y.

Next, in any (213)-decomposition of $T_2 = A \otimes (B + C)$, A contains $\hbar$ and hence the 1-diagonal $d_1(Y)$, which starts from $\hbar$ (cf. Fig. 25b.) The hypotheses on Y and the position of $\hbar$ imply that $d_1(Y)$ does not intersect the border of Y until goes through the rightmost column of Y and stops underneath β's cell. This forces the subtransversal A to involve the second row of Y and hence to contain α, as well as the rightmost column of Y and hence to contain β, i.e. $A = T_2$ and the (213)-decomposition of T_2 is trivial. Therefore, T_2 is (213)-indecomposable.

From Lemma 17, if a preimage $T \in S_Y(213)$ of T_2 existed under ψ, then T would also be (213)-indecomposable and by Case 2 of ψ's definition: $T \xrightarrow{\psi} T_1 \xrightarrow{\eta} T_2$. In particular, T and T_1 would have $\hbar$ is their bottom left corners. But $\alpha \in T_2$ is (12)-dominated only by β, and β does not project on the bottom row of Y, hence α is **not** (12)-dominated in the rectangle $\mathcal{R}$, hence $\alpha \in \dot{D}'$ (cf. Fig. 23.) Since η fixes $\dot{D}'$, inverting η would leave α fixed in the first column of T_1. This precludes $\hbar$ from occupying the bottom left corner in T_1, a contradiction. We conclude that T_2 is not invertible under η and ψ, and hence $T_2 \notin \psi(S_Y(213))$. $\square$

Proposition 8. $|S_Y(213)| < |S_Y(123)|$ *if and only if* Y *has an* i*-critical point with* $i \geq 2$.

PROOF: As in the proof of Proposition 5, for any permutation σ we can split Y and its transversals with respect to any 0- and 1-critical points:

$$|S_Y(\sigma)| = |S_{U_1}(\sigma)| \cdot |S_{U_2}(\sigma)| \cdots |S_{U_k}(\sigma)|,$$

where each U_j is either square or contains only i-critical points with $i \geq 2$. If the original Y contains only 0- or 1-critical points, then all U_j's are square with $|S_{U_j}(213)| = |S_{U_j}(123)|$, so that $|S_Y(213)| = |S_Y(123)|$.

If Y does contain some high i-critical points with $i \geq 2$, in addition to the square U_j's, there will be at least one other U_m with such a high critical point. Lemma 18 implies strict inequalities for all non-square U_j's in our decomposition. In particular, $|S_{U_m}(213)| < |S_{U_m}(123)|$ and therefore $|S_Y(213)| < |S_Y(123)|$. $\square$

This completes the proof of Theorem 2. $\square$

8 Strict Wilf-ordering for $(213|\tau)$, $(123|\tau)$ and $(312|\tau)$

Subsection 2.6 gives a strategy for proving that for any permutation τ:

$$|S_n(213|\tau)| \lneqq |S_n(123|\tau)| \lneqq |S_n(312|\tau)| \text{ for } n \gg 1.$$

Since each Young diagram Y_m has an $(m-2)$-critical point, Theorem 2 implies that $|S_{Y_m}(213)| \lneqq |S_{Y_m}(123)| \lneqq |S_{Y_m}(312)|$ for $m \geq 5$. This fulfills the first step (SF1) of the strategy. The other step (SF2) is provided by the following construction.

Lemma 19. *Given a permutation* $\tau \in S_k$*, for any* $n \geq 2k+2$ *there is a partial transversal* T_n *of* M_n *which saturates* Y_{n-2k} *with respect to* τ.

PROOF: Take two copies τ_1 and τ_2 of τ and arrange them in a southwest/northeast diagonal fashion within a square matrix M_{2k} (cf. Fig. 26a.) Insert a row and column through the middle of M_{2k} so that the resulting $\mathcal{M} \cong M_{2k+1}$ has an empty separating row and column between τ_1 and τ_2. Place $\mathcal{M}$ in the bottom right corner of M_n for $n \geq 2k+2$ (cf. Fig. 26b.) We claim that the partial transversal T_n of M_n produced by the two copies τ_1 and τ_2 in $\mathcal{M}$ saturates Y_{n-2k} with respect to τ.

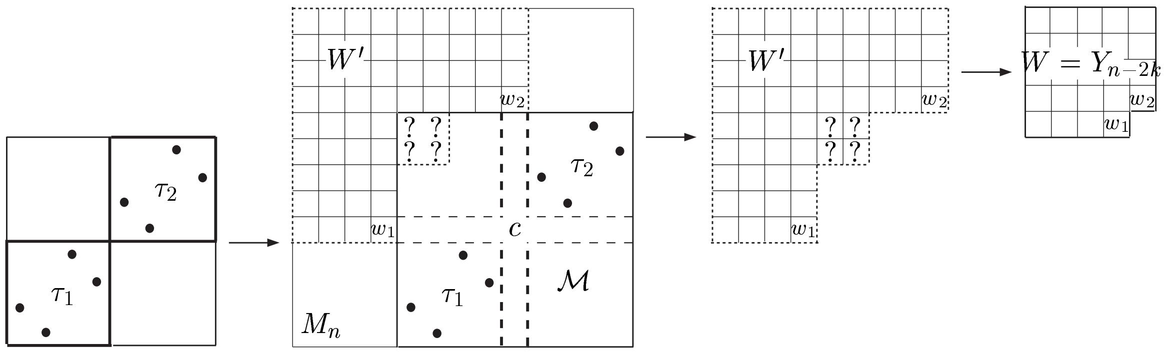

Figure 26: $T_n = \tau_1 \oplus \tau_2$ saturates $W = Y_{n-2k}$ with respect to τ

To see this, denote by w_1 and w_2 the cells of Y diagonally to the left and above the $k \times k$ matrices of τ_1 and τ_2. Then w_1 and w_2 are white cells with respect to τ and the partial transversal T_n. The existence of w_1 and w_2 gives $n \geq 2k+2$. If c is the central cell of $\mathcal{M}$, the initial white subboard W' is of the union $Y_{w_1} \cup Y_{w_2}$ of two rectangles, plus possibly some more white cells within the rectangle $\mathcal{M}_{\bar{c}}$ (these cells are depicted by "?" in Fig. 26b-c). However, the reduction of W' along $\tau_1 \cup \tau_2$ deletes *all* cells in $\mathcal{M}_{\bar{c}}$, and the consequent removal of the (blue) central row and column of $\mathcal{M}$ leaves the white diagram $W = Y_{n-2k}$ (cf. Fig. 26d). By definition, T_n saturates Y_{n-2k} with respect to τ in M_n. $\square$

When $n \geq 2k+5$, then the saturated Y_{n-2k} satisfies (SF1). Combining with (SF2),

$$|S_{Y_{n-2k}}(213)| \cdot |\bar{S}_{M_n \setminus Y_{n-2k}}(\tau)| \lneqq |S_{Y_{n-2k}}(123)| \cdot |\bar{S}_{M_n \setminus Y_{n-2k}}(\tau)| \lneqq |S_{Y_{n-2k}}(312)| \cdot |\bar{S}_{M_n \setminus Y_{n-2k}}(\tau)|$$

$$\Rightarrow \ |S_n(213|\tau)| \lneqq |S_n(123|\tau)| \lneqq |S_n(312|\tau)| \text{ for } n \geq 2k+5.$$

This completes the proof of Corollary 1. $\square$

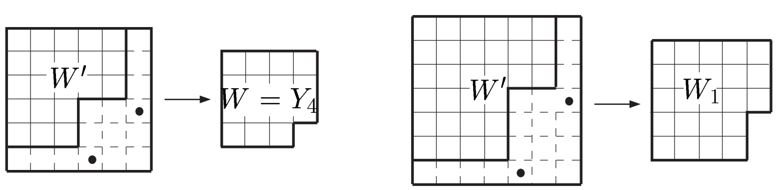

Figure 27: T' saturates $W = Y_4$ with respect to $\tau = (1)$

Example 3. As an illustration of the above inequalities, let $\tau = (1)$ and consider $(213|1) \preceq (123|1) \preceq (312|1)$. When $n = 6, 7$:

$$|S_6(3241)| = 512 < |S_6(2341)| = |S_6(4231)| = 513, \tag{7}$$
$$|S_7(3241)| = 2740 < |S_7(2341)| = 2761 < |S_7(4231)| = 2762. \tag{8}$$

Let T' be a partial transversal of M_6 that saturates a $W \subset M_6$ with respect to $\tau = (1)$. Then $|W| \leq 5$ with $|W| = 5$ if and only if T' consists of a single element in the bottom right corner of M_6; in such a case $W = M_5$. Thus, $|S_W(213)| = |S_W(123)| = |S_W(312)|$

for all $W \subset M_6$ except $W = Y_4$, where the 2-critical point of Y_4 implies the inequality $|S_{Y_4}(213)| = 12 < |S_{Y_4}(123)| = |S_{Y_4}(312)| = 13$. On the other hand, it is easy to verify that the only T' that saturates Y_4 in M_6 consists of two elements placed in positions $(4,6)$ and $(6,4)$ (cf. Fig. 27a-b.) Thus, the Splitting Formulas for $S_6(3241)$, $S_6(2341)$ and $S_6(4231)$ have all but one equal summands:

$$12 \cdot 1 = |S_{Y_4}(213)| \cdot |S| \lneqq |S_{Y_4}(123)| \cdot |S| = S_{Y_4}(312)| \cdot |S| = 13 \cdot 1,$$

where $S = \bar{S}_{M_6 \backslash Y_4}(1)$. This explains the difference of 1 between the quantities in (7).

The analogous partial transversal T'' in M_7 (whose two elements are placed in $(5,7)$ and $(7,5)$) saturates Y_5 with respect to $\tau = (1)$. The 3-critical point of Y_5 implies the following inequalities, where $S = \bar{S}_{M_7 \backslash Y_5}(1)$:

$$37 \cdot 1 = |S_{Y_5}(213)| \cdot |S| \lneqq 41 \cdot 1 = |S_{Y_5}(123)| \cdot |S| \lneqq 42 \cdot 1 = |S_{Y_5}(312)| \cdot |S|.$$

This explains the difference of 1 between $|S_7(2341)|$ and $|S_7(4231)|$ in (8). Further, Y_5, $W_1 \cong Y(5,5,5,4,4)$, its transpose $W_1^t = W_2 \cong Y(5,5,5,5,3)$ and Y_4 are saturated in M_7 by correspondingly 1, 1, 1, and 9 partial transversals of M_7. (W_1 is depicted in Fig. 27d). On all other induced Young subdiagrams of M_7, (213) and (123) are equally restrictive. Therefore, the Splitting Formulas give the remaining difference of 21 in (8):

$$\begin{aligned} |S_7(2341)| - |S_7(3241)| &= \sum_{W \in \{Y_5, W_1, W_2, Y_4\}} \big(|S_W(123)| - |S_W(213)|\big) \cdot |\bar{S}_{M_7 \backslash W}(1)| \\ &= (41-37) \cdot 1 + (37-33) \cdot 1 + (37-33) \cdot 1 + (13-12) \cdot 9 = 21. \end{aligned}$$

9 Young Diagrams with Extreme Critical Indices

9.1 The sets $|S_{Y_n}(\sigma)|$ and the Catalan numbers.

For Young diagrams Y with higher i-critical points, it is interesting to find out by how much (312) and (321) are less restrictive than (321) and (213), respectively. Below we answer this for the diagram $Y = Y_n$ with *highest* critical index $i = n-2$, and leave the general question to the reader.

Proposition 9. *$|S_{Y_n}(213)| = c_n - c_{n-2}$ for $n \geq 2$.*

Proof: This follows from Corollary 1 of the Row-Decomposition in Stankova-West [21]. Paraphrasing into the notation in the current paper, let a, b, c be the three bottom right corner cells of M_n as in Fig. 28. Then $M_n \backslash \{b\} = Y_n$. On the other hand, reducing M_n along a gives $M_n/_a = M_{n-1}$ whose right bottom cell is c. The minimal non-trivial (213)-decomposition of this M_{n-1} is obtained with respect to c: $(M_{n-1})_{213}(c) = \mathcal{A}_c \times \mathcal{B}_c = \{c\} \times M_{n-2}$. Thus, the *row-decomposition* formula for $S_{M_n}(213)$ in [21] reads: $|S_{M_n}(213)| = |S_{Y_n}(213)| + |S_{\mathcal{A}_c \times \mathcal{B}_c}(213)|$, from where $|S_{Y_n}(213)| = |S_n(213)| - |S_{n-2}(213)| = c_n - c_{n-2}$. $\square$

In the following, we keep the notation b for the bottom right cell of M_n, which is missing from Y_n. We shall enumerate $S_{Y_n}(321)$ and $S_{Y_n}(312)$ by finding out how each differs as a set from $S_n(321)$ and $S_n(312)$, respectively.

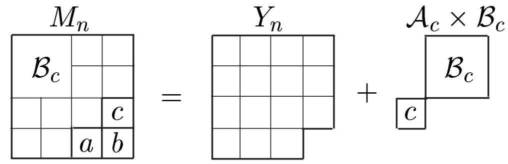

Figure 28: $|S_n(213)| = |S_{Y_n}(213)| + |S_{n-2}(213)|$

Proposition 10. $|S_{Y_n}(321)| = c_n - 1$ *for* $n \geq 2$.

Proof: Fix $T \in S_{Y_n}(321)$. Adding the cell b to Y_n induces a transversal T' on M_n, which also avoids (321) on M_n. Indeed, if $(\alpha\beta\gamma)$ were a (321)-pattern of T' in M_n, then $(\alpha\beta\gamma)$ lands on γ's cell d. Since d is dotted in T' on M_n, it is also dotted in T on Y_n, i.e. $d \neq b$. But then $(\alpha\beta\gamma)$ is a (321)-pattern of T landing on d in Y_n, a contradiction with $T \in S_{Y_n}(321)$.

Thus, we have a natural inclusion map $\iota : S_{Y_n}(321) \hookrightarrow S_n(321)$. The reasoning above also shows that the only transversals $T' \in S_n(321)$ not hit by ι are those with dotted b. However, in order to avoid (321) on M_n, a dot in b implies that the rest of T' is increasing, and there is only one such transversal, namely, $T' = (2, 3, ..., n, 1)$. We conclude that $S_n(321) = \iota(S_{Y_n}(321)) \sqcup \{T'\}$.

$$\Rightarrow \; |S_{Y_n}(321)| = |S_n(321)| - 1 = c_n - 1 \text{ for any } n \geq 2. \quad \square$$

Proposition 11. $|S_{Y_n}(312)| = 2c_n - 3c_{n-1}$ *for* $n \geq 2$.

Proof: As indicated above, we describe how $S_{Y_n}(312)$ differs as a set from $S_n(312)$.

On the one hand, $S_n(312)$ contains transversals of M_n with a dotted b. Since b cannot participate in any (312)-pattern, we can reduce M_n along b to obtain M_{n-1} without any further restrictions, and hence the number of transversals in question equals $|S_{n-1}(312)|$. None of these transversals is in $S_{Y_n}(312)$ because Y_n cannot have a dot in the missing b (cf. Fig. 29a.) Thus, $|S_n(312) \backslash S_{Y_n}(312)| = c_{n-1}$. On the other hand, $S_{Y_n}(312)$

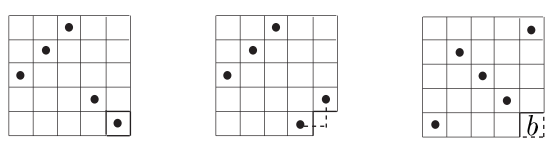

Figure 29: Examples of the difference between $S_n(312)$ and $S_{Y_n}(312)$

contains transversals of Y_n for which a (312)-subsequence lands outside Y_n (on b). As we shall see below in Lemma 20, the number of these transversals is $c_n - 2c_{n-1}$, and none of them is in $S_n(312)$ because of the (312)-pattern in M_n (cf. Fig. 29b.) Thus, $|S_{Y_n}(312) \backslash S_n(312)| = c_n - 2c_{n-1}$.

All other transversals of $S_n(312)$ and $S_{Y_n}(312)$ are identical: they don't have an element in b, and they don't have a (312)-pattern landing on b (cf. Fig. 29c.) Summarizing,

$$|S_{Y_n}(312)| = |S_n(312)| - c_{n-1} + (c_n - 2c_{n-1}) = 2c_n - 3c_{n-1} \text{ for any } n \geq 2. \quad \square$$

Incidentally, we have shown the strict inequality $|S_{Y_n}(312)| > |S_{Y_n}(321)|$ for $n \geq 5$ (proven in an indirect way in Example 2). Indeed, from Propositions 10-11, for $n \geq 5$:

$$|S_{Y_n}(312)| - |S_{Y_n}(321)| \;=\; (2c_n - 3c_{n-1}) - (c_n - 1) = \frac{(n-5)n(2n-2)!}{(n+1)!(n-2)!} + 1 \geq 1.$$

9.1.1 Claims in the Proof of Proposition 11

Lemma 20. *The number of all transversals in $S_{Y_n}(312)$ with a (312)-subsequence landing outside Y_n (on b) is $c_n - 2c_{n-1}$.*

Proof: Let $T \in S_{Y_n}(312)$, and let α and γ denote the elements of T in the bottom row and in the rightmost column of Y_n, respectively. Because Y_n misses b, $\alpha \neq \gamma$ (cf. Fig. 30a).

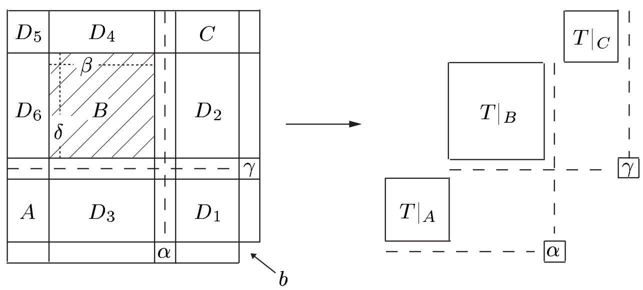

Figure 30: Splitting of $T \in S_{Y_n}(312)$

Suppose T contains a (312)-subsequence which doesn't land in Y_n, hence lands on b. Thus, for some $\beta \in T$, $(\beta\alpha\gamma) \approx (312)$. Since β is before α and above γ, without loss of generality, we can replace β by the largest element of T before α; symbolically, $\beta := \max\{t \in T_{\bar{\alpha}}\}$. Symmetrically, let δ be the leftmost element of T higher than γ. It is possible that $\delta = \beta$; if not, $(\delta\beta\alpha\gamma) \approx (3412)$ is a subsequence of T not landing in Y_n. Let B be the rectangle in Y_n defined by β's and γ's rows, and δ's and α's columns such that B includes β and δ, but excludes α and γ. Let A be the rectangle below and to the left of B, excluding α's and γ's rows; and symmetrically, let C be the rectangle to the right and above B, excluding α's and γ's columns.

Claim 3. *Save for α and γ, the transversal T is concentrated in rectangles A, B and C.*

Proof: Y_n splits as a disjoint union of 9 rectangles, plus α's and γ's rows and columns. Figure 30a depicts all these rectangles. The definitions of β and δ imply that rectangles D_4, D_5 and D_6 are empty. In order for the pair $(\beta\alpha)$ not to be completed to a (312)-pattern in Y_n, rectangles D_1 and D_2 must also be empty. Symmetrically, in order for the pair $(\delta\gamma)$ not to be completed to a (312)-pattern in Y_n, rectangles D_3 and D_1 must be empty. Thus, $T\backslash\{\alpha, \gamma\}$ is concentrated in A, B and C. □

We conclude that T induces transversals on the rectangles A, B and C, and since the latter are disjoint, they must be *squares*. Thus, T splits into an increasing sequence of

3 independent subtransversal $T|_A$, $T|_B$ and $T|_C$, with α inserted in the bottom row of Y_n so that its column is between B and C, and γ is inserted in the rightmost column of Y_n so that its row is between A and B. Finally, the assumption that T contains a (312)-subsequence not landing inside Y_n was translated above in the existence $\beta \in B$, i.e. the square B is of size at least 1. Conversely,

Claim 4. *If T is a transversal of Y_n satisfying the above description (depicted also in Fig. 30b), and such that the 3 subtransversals $T|_A$, $T|_B$ and $T|_C$ each avoid* (312) *on the respective squares A, B and C, then the whole transversal T avoids* (312) *on Y_n, and has a* (312)*-subsequence not landing in Y_n.*

PROOF: Consider the reduction $Y_n/_{\{\alpha,\gamma\}} = M_{n-2}$, along whose diagonal the squares A, B and C are arranged (in increasing order). It is evident that there can be no (312)-pattern in M_{n-2} containing elements from different squares. Since $T|_A$, $T|_B$ and $T|_C$ each avoid (312), any (312)-pattern in T on Y_n must contain α and/or γ. But α and γ cannot participate simultaneously in any pattern landing inside Y_n because of the missing cell b. Hence, only one of α and γ can participate in a (312)-pattern in Y_n.

Since α can play only the role of "1", it can participate only in a (312)-pattern of the form $(\xi\alpha\nu)$, where $(\xi\nu)\searrow$, ξ is before α and ν is after α. Yet, this arrangement is not possible since everything before α is smaller than everything after α: $A \oplus B < C$, with the exception of γ, so no such pattern is possible. "Transposing" this argument, one concludes that γ cannot participate in a (312)-pattern in Y_n either.

Therefore, $T \in S_{Y_n}(312)$. Finally, since B is of size at least 1, let $\beta \in T|_B$. Then $(\beta\alpha\gamma) \approx (312)$ landing on b. □

Claims 3-4 establish a 1-1 correspondence between the transversals $T \in S_{Y_n}(312)$ that do **not** induce transversals in $S_n(312)$ due to their (312)-subsequence landing on b, and the diagrams in Figure 30b. Therefore, each element of $S_{Y_n}(312)\backslash S_n(312)$ is uniquely determined by the size of the squares A, B and C, and the choice of (312)-avoiding transversals $T|_A, T|_B$ and $T|_C$. Below, the sum of sizes $|A| + |B| + |C| = n - 2$ accounts for $\alpha, \gamma \notin A \cup B \cup C$.

$$\begin{aligned} S_{Y_n}(312)\backslash S_n(312) &\cong \bigsqcup_{\substack{|A|+|B|+|C|=n-2 \\ |B|\geq 1}} S_A(312) \times S_B(312) \times S_C(312) \\ \Rightarrow\ |S_{Y_n}(312)\backslash S_n(312)| &= \sum_{\substack{i+j+k=n-2 \\ j\geq 1}} c_i c_j c_k = c_n - 2c_{n-1}. \end{aligned}$$

The last equality was obtained using the well-known relation $c_k = \sum_{l+m=k-1} c_l c_m$ for the Catalan numbers. This completes the proof of Lemma 20. □

9.2 The sets $|S_{St_n^3}(\tau)|$ and the Fibonacci Numbers

In this subsection, we consider the other extreme situation of a *non-decomposable* Young diagram Y: having *lowest* critical indices $i = 2$. This is $Y = St_n^3$ for $n \geq 4$, which is the *smallest non-decomposable* Young diagram of size n. The last description is also satisfied by the squares M_n with $n \leq 3$, and we set $St_n^3 := M_n$ for $n \leq 3$. This new notation and Theorem 2 imply that (123) and (312) are equinumerant on St_n^3 for all $n \geq 1$, so that we can state the following

Proposition 12. $|S_{St_n^3}(213)| = 2^{n-3}(n+2)$ for $n \geq 2$ and

$$|S_{St_n^3}(123)| = |S_{St_n^3}(312)| = f_{2n-1} = \frac{1}{\sqrt{5}}\big(\psi^{2n-1} - \psi^{-(2n-1)}\big) \text{ for } n \geq 1,$$

where f_n is the n-th Fibonacci number ($f_1 = f_2 = 1$) and $\psi = (1+\sqrt{5})/2$.

Proof: Let $a_n = |S_{St_n^3}(213)|$. From the row-decomposition formula in [21] on St_n^3:

$$|S_{St_n^3}(213)| = 2\,|S_{St_{n-1}^3}(213)| + |S_{St_{n-2}^2}(213)| \text{ (cf. Fig. 31).}$$

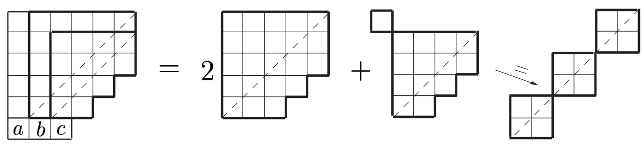

Figure 31: Row-decomposition of $S_{St_n^3}(213)$

Since St_{n-2}^2 1-decomposes as a product of $(n-3)$ squares M_2 (cf. Fig. 31d), we have $|S_{St_{n-2}^2}(213)| = 2^{n-3}$. Thus, $a_n = 2a_{n-1} + 2^{n-3}$, i.e. $a_n = 2^{n-3}(n+2)$ for $n \geq 2$.

Consider now $b_n = |S_{St_n^3}(312)|$. Let a, b and c be the bottom cells of St_n^3, as in Fig. 32a. Placing 1 in a or c does not affect the (312)-avoidance in the reduction $St_n^3/_{\{u\}} \cong St_{n-1}^3$ for $u = a$ or c (cf. Fig. 32b), and thus yields overall $2b_{n-1}$ transversals. However, placing 1 in b forces the elements in the first two columns of the reduction $St_n^3/_{\{b\}} \cong St_{n-1}^3$ to form an increasing sequence (depicted by $\nearrow$ above cell d in Fig. 32c.) In accordance with previous notation, we denote the number of such (312)-avoiding transversals of St_{n-1}^3 by $b_{n-1}^{\nearrow}$. Therefore, $b_n = 2b_{n-1} + b_{n-1}^{\nearrow}$.

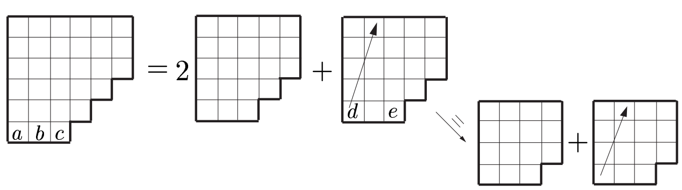

Figure 32: Row-decomposition of $|S_{St_n^3}(312)|$

To calculate $b_{n-1}^{\nearrow}$, note that 1 can be placed only in the first cell d or in the third cell e of the bottom row of St_{n-1}^3. The first case does not cause any restrictions on the reduction $St_{n-1}^3/_{\{d\}} \cong St_{n-2}^3$ (cf. Fig. 32d) and hence it produces b_{n-2} transversals. Placing 1 in e reduces to $b_{n-2}^{\nearrow}$ on $St_{n-1}^3/_{\{e\}} \cong St_{n-2}^3$ (cf. Fig. 32e). Summarizing, $b_{n-1}^{\nearrow} = b_{n-2} + b_{n-2}^{\nearrow}$.

Combining the two newly derived formulas, we obtain $b_n = 3b_{n-1} - b_{n-2}$, with $b_1 = 1$ and $b_2 = 2$. It is a standard exercise to check that the odd-indexed terms in the Fibonacci sequence satisfy the same recursive relation, and hence the desired formula for b_n involving the golden ratio ψ follows. $\square$

9.3 Generalization of Stanley-Wilf limits.

Recall the Stanley-Wilf limits $L(\tau) = \lim_{n\to\infty} \sqrt[n]{|S_n(\tau)|}$ for any $\tau \in S_k$. From works of Regev [15] and Bóna [8], it follows that $L(J_k) = (k-1)^2$ and $L(213|J_k) = (k-1+\sqrt{8})^2$ for $k \geq 1$.

From the viewpoint of the current paper, Corollary 1 has established in particular the strict inequalities $|S_n(213|J_k)| < |S_n(123|J_k)| < |S_n(312|J_k)|$ for any $k \geq 1$ and $n \geq 2k+5$. Hence the Stanley-Wilf limits follow suit for $k \geq 1$:

$$L(213|J_k) = (k-1+\sqrt{8})^2 < L(123|J_k) = L(321|J_k) = L(J_{k+3}) = (k+2)^2 \leq L(312|J_k).$$

It has been shown recently that $L(312|1) > 9.35$ (cf. [1]). In order to complete the above picture, it would certainly be nice to find the exact value of $L(312|J_k)$. Conceivably, the Splitting Formula for $(312|J_k)$ from Subsection 2.5 and other observations in this paper might be helpful towards calculating $L(312|J_k)$.

With the methods so far, all known $L(\tau)$ belong to $\mathbb{Z}[\sqrt{2}]$. However, if we generalize the definition of Stanley-Wilf limits from the square matrices M_n to using any (proper) Young diagrams Y of size n, we can obtain presumably a much greater variety of limits. To this end, consider the set $\mathcal{Y} = \cup_{n=0}^{\infty}\mathcal{Y}^n$ of all proper Young diagrams, *graded* by the size n of the diagrams. Let $\vec{Y} = \{Y^n\}$ be a sequence of (proper) Young diagrams, one per each graded piece of $\mathcal{Y}$; we can think of $\vec{Y}$ as a *path* in $\mathcal{Y}$. Define the *generalized Stanley-Wilf limit* of $\tau \in S_k$ along the path $\vec{Y}$ as

$$L_{\vec{Y}}(\tau) = \lim_{n\to\infty} \sqrt[n]{|S_{Y^n}(\tau)|}.$$

Except for the case $\vec{Y} = \{M_n\}$ where the limits $L_{\vec{Y}}(\tau) = L(\tau)$ are guaranteed by Stanley-Wilf Theorem, for all other paths in $\mathcal{Y}$ the existence of $L_{\vec{Y}}(\tau)$ must be verified.

A worthwhile consequence of Proposition 12 is the following

Corollary 2. *For* $\vec{Y} = \{St_n^3\}$, $L_{\vec{Y}}(321) = L_{\vec{Y}}(312) = \psi^2 = \frac{3+\sqrt{5}}{2}$.

Two natural questions arise: for which pairs $(\vec{Y}, \tau)$ do the limits $L_{\vec{Y}}(\tau)$ exist, and what is the algebraic closure $\overline{\mathcal{L}}$ of the set of generalized limits $\mathcal{L} = \{L_{\vec{Y}}(\tau)\}$. As of now, we have shown that $\overline{\mathcal{L}} \supset \mathbb{Q}(\sqrt{2}, \sqrt{5})$; but are there any other irrational or transcendental generalized Stanley-Wilf limits $L_{\vec{Y}}(\tau)$? We leave these questions to the reader for further study.

Acknowledgments

The author would like to thank Miklos Bóna (University of Florida) for supplying a number of useful references and discussing his and related works in relation to the present paper; David Moews (Center for Communications Research, San Diego) for writing a computer program used in this project; and Paulo de Souza (UC Berkeley) for his help in implementing the necessary computer software.

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, Electronic J. Combin. 6 (1999), no. 1, N1.

[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 Electronic J. Combin., 2006.

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

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

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

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

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

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

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

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

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

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

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

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

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

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