GREEK LETTERS IN RANDOM STAIRCASE TABLEAUX*

SANDRINE DASSE-HARTAUT[†] AND PAWEŁ HITCZENKO[‡]

ABSTRACT. In this paper we study a relatively new combinatorial object called staircase tableaux. Staircase tableaux were introduced by Corteel and Williams in the connection with Asymmetric Exclusion Process and has since found interesting connections with Askey–Wilson polynomials. We develop a probabilistic approach that allows us to analyze several parameters of a randomly chosen staircase tableau of a given size. In particular, we obtain limiting distributions for statistics associated with appearances of Greek letters in staircase tableaux.

Key words and phrases: asymmetric exclusion process, asymptotic normality, staircase tableaux.

1. INTRODUCTION

An interesting combinatorial structure, called staircase tableaux, was introduced in recent work of Corteel and Williams [11, 12]. Staircase tableaux are related to the asymmetric exclusion process on an one-dimensional lattice with open boundaries, the ASEP. This is an important and heavily studied particle model in statistical mechanics (we refer to [12] for some background information on several versions of that model and their applications and connections to other branches of science). The study of the generating function of the staircase tableau has given a combinatorial formula for the steady state probability of the ASEP. Explicit expressions for the steady state probabilities were first given in [15]. In their work [12, 11] Corteel and Williams used staircase tableaux to give a combinatorial formula for the moments of the (weight function of the) Askey-Wilson polynomials; for a follow–up work see [7].

[†] LIAFA, Université Paris Diderot-Paris 7, F-75205 Paris, France, dasse@liafa.jussieu.fr.

[‡] Institute of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, 00-956 Warszawa, Poland and Department of Mathematics and Information Science, Warsaw University of Technology, Pl. Politechniki 1, 00-661 Warszawa, Poland, phitczenko@math.drexel.edu.

^{*} The work of the first author was carried out while she held an ANR Gamma internship at LIPN, Université Paris Nord, under the direction of Frédérique Bassino (LIPN) and Sylvie Corteel (LIAFA). She would like to thank both of them for their guidance, the members of LIPN for their hospitality, and ANR Gamma for the support. The second author was partially supported by NSA Grant #H98230-09-1-0062. Most of his work was done during his stay at LIPN in July 2010 and completed while he was at the Institute of Mathematics of the Polish Academy of Sciences and the Technical University of Warsaw in the Fall of 2010. He would like to thank Frédérique Bassino for the invitation and acknowledge the hospitality of these institutions.

The authors of [12] called for further investigation of the staircase tableaux because of their combinatorial interest and their potential connection to geometry. In this paper we take up that issue and study some basic properties of staircase tableaux. More precisely, we analyze the distribution of various parameters associated with appearances of Greek letters α , β , δ , and γ in randomly chosen staircase tableau of size n (see the next section or, e.g. [12, Section 2], for the definitions and the meaning of these symbols).

Staircase tableaux are generalizations of permutation tableaux (see e.g. [6, 9, 10, 20] and references therein for more information on these objects and their connection to a version of ASEP referred to as the partially asymmetric exclusion process; PASEP). For permutation tableaux, the authors of [6] developed a probabilistic approach that later allowed the derivation of the limiting (and even exact) distributions of various parameters of the permutation tableaux. Our goal here is the same: in Section 3 we develop a probabilistic approach parallel to that of [6] that allows us to compute generating functions of various quantities associated with staircase tableaux (see Corollary 6 in Section 4 and Proposition 11 in Subsection 6.1.3). As a consequence, we obtain the exact or limiting distributions of the parameters we study.

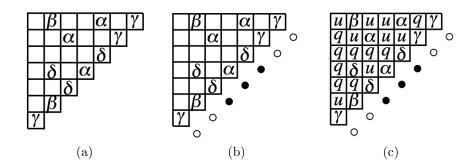
Our main results are gathered in Section 5 and may be summarized by stating that the five parameters we consider have asymptotically normal distribution when normalized in a usual way (i.e. centered by the mean and scaled by the square root of the variance). Thus, for example, Theorem 9 asserts that the number A_n of α or γ on the diagonal of a randomly chosen staircase tableau of size n has expected value n/2, variance (n+1)/12 and that $(A_n - n/2)/\sqrt{n/12}$ converges in distribution to the standard normal random variable. We refer to Theorems 8 and 9 for precise statements.

In Section 6 below (see comments at the beginning of Subsection 6.1.3 and a remark at the end of it) we find that one of the parameters we study coincides with a generalization of Eulerian numbers (see [23, sequence A060187]) related to Whitney numbers of Dowling lattices (see [23, sequences A145901, A039775] and [16, 2, 3, 5] for definitions and further information on these numbers). This rather unexpected and intriguing connection has not been explained and merits, perhaps, further studies. One consequence of our work is that the triangle of numbers [23, sequence A060187], when suitably normalized, satisfies the central limit theorem. As far as we can tell this result is new (although it is an easy consequence of a general theorem of Bender [1]). Limit theorems for a related sequence [23, A145901] are established in [5]. This link to the Whitney numbers of Dowling lattices may have unraveled connection to geometry alluded to in [12] as the sequences A060187, A145901, and A039775 from [23] all have a very strong geometrical flavor.

2. Definitions and notation

We recall the following concept first introduced in [11, 12]: A staircase tableau of size n is a Young diagram of shape (n, n - 1, ..., 2, 1) whose boxes are filled according to the following rules:

- each box is either empty or contains one of the letters α , β , δ , or γ ;
- no box on the diagonal is empty;
- all boxes in the same row and to the left of a β or a δ are empty;
- all boxes in the same column and above an α or a γ are empty.



An example of a staircase tableau is given in Fig. 1(a).

FIGURE 1. (a) A staircase tableau of size is 7. Its top row is indexed by β , the next one by α . (b) Its type ($\circ \circ \bullet \bullet \circ \circ \circ$). (c) The same tableau with *u*'s and *q*'s.

As we mentioned, staircase (or earlier permutation) tableaux were studied in the connection with ASEP. Because of the importance of this connection we briefly recall its nature. The ASEP is a Markov chain on words of size n on an alphabet $A = \{\circ, \bullet\}$ consisting of two letters. Each such word represents an one-dimensional lattice of length n with some sites occupied by particles (represented by \bullet), and others not (represented by \circ). A particle can only hop to the right or the left (with the probabilities u and q, respectively), provided that the adjacent site is unoccupied, or enter or quit the lattice. Entering from the left (right) happens with the probability α (resp., δ) if the first (last) site is unoccupied. Exiting to the left (right) happens with the probability γ (resp., β) if the first (last) site is occupied. At a given time one of the n+1 possible locations for a move is selected (uniformly at random) and, if possible, a transition described above is performed with the given probability. We refer to [14, 15, 17] or [12] for more detailed description and further references.

To describe the connection to staircase tableaux, define the *type* of a staircase tableau S of size n to be a word of the same size on the alphabet $\{\circ, \bullet\}$ obtained by reading the diagonal boxes from northeast (NE) to southwest (SW) and writing \bullet for each α or δ , and \circ for each β or γ . (Thus a type of a tableau is a possible state for the ASEP.) Fig. 1(b) shows a tableau with its type. We also need a *weight* of a tableau S. To compute it, we first label the empty boxes of S with u's and q's as follows: first, we fill all the boxes to the left of a β with u's, and all the boxes to the left of a δ with q's. Then, we fill the boxes above an α or a δ with u's, and the boxes above a β or a γ with q's. When the tableau is filled, its weight, wt(S), is a monomial of degree n(n+1)/2 in α , β , γ , δ , u and q, which is the product of labels of the boxes of S. Fig. 1(c) shows a tableau filled with u's and q's. Its weight is $\alpha^3 \beta^2 \delta^3 \gamma^3 u^8 q^9$.

Corteel and Williams [12, 11] have shown that the steady state probability that the ASEP is in state σ is

where
$$Z_n(\alpha, \beta, \gamma, \delta, q, u) = \sum_{\substack{S \text{ of size } n}} \frac{Z_{\sigma}(\alpha, \beta, \gamma, \delta, q, u)}{Z_n(\alpha, \beta, \gamma, \delta, q, u)}$$
,

We denote the set of all staircase tableaux of size n by S_n , $n \ge 1$. It is known that the cardinality of S_n is $4^n n!$. There are several proofs of this fact (c.f. [7] for one of them and for references to further proofs). All these proofs are based on combinatorial approaches and we wish to mention that a probabilistic technique that we develop in this paper provides a yet another proof of that fact. We present it in Section 4 below as an illustration of how our method works.

We now define some parameters that are the objects of our study. Let * be a subset of the set of symbols $\{\alpha, \beta, \delta, \gamma\}$. We say that a row of a staircase tableau is indexed by * if its leftmost entry is a member of *. For the sake of brevity we will refer to rows indexed by * simply as * rows. Thus, for example, the number of α/γ rows is the number of rows indexed by α or γ . The tableau in Figure 1(a) has two α/γ rows, the second from the top (indexed by α) and the bottom (indexed by γ). For a given staircase tableau $S \in \mathcal{S}_n$ we denote this quantity by $r_n(S)$ and we occasionally skip the subscript n if there is no risk of confusion. As we demonstrate in Section 3 below this parameter plays a fundamental role in our approach. Other parameters we consider are: the total number of entries β or δ (β/δ for short), the total number of entries α or γ (α/γ), the number of entries β/δ on the diagonal of the tableau, and the number of entries α/γ on the diagonal. For a given tableau $S \in \mathcal{S}_n$ these parameters will be denoted by $\Delta_n(S)$, $\Gamma_n(S)$, $B_n(S)$, and $A_n(S)$, respectively. We gather the names and the notation we use for the parameters we consider along with their values for a tableau given in Figure 1(a) in the following table

parameter	notation	value in Fig. $1(a)$
α/γ rows	r_n	2
total # of β/δ	Δ_n	5
total # of α/γ	Г	6
# of β/δ on diagonal	B_n	3
# of α/γ on diagonal	A_n	4

As we mentioned earlier our viewpoint is probabilistic. Thus, we equip the set S_n with the uniform probability measure denoted by P_n . This means that for each $S \in S_n$ we have

$$\mathsf{P}_n(S) = \frac{1}{4^n n!}.$$

As is customary we refer to a tableau chosen according to that measure as a random tableau of size n. We denote the integration with respect to the measure P_n by E_n . Our goal is to analyze probabilistic properties, like expected values, variances, and exact or limiting distributions, of random variables (or statistics) r_n , Δ_n , Γ_n , B_n , and A_n . We follow the probability theory conventions of [22], and the reader is pointed there for any unexplained terms.

3. Preliminaries and outline of the argument

In this section we detail main ideas beyond our approach and also derive basic properties of the fundamental parameter, i.e. the number of α/γ rows.

Our method is analogous to what has been done in the case of permutation tableaux (see [6] or [20]). Let us recall at this point that permutation tableaux have been used to give a combinatorial description of a stationary distribution

for the PASEP. We refer to e.g. [9, 10, 6, 20] for the definition, connections to PASEP, further properties and details. Just as PASEP is a particular case of ASEP, permutation tableaux of size n are in bijection with a subset of staircase tableaux of that size corresponding to the case $\gamma = \delta = 0$.

The approach used in [6, 20] for the permutation tableaux was to identify a fundamental parameter, trace its evolution as the size of a tableau is increased by 1, and then use successively conditioning to reduce the size of a tableau. We refer the reader to either [6] or [20] for more details, here we only recall that this fundamental parameter was the number of unrestricted rows, U_n , in a permutation tableau, and that its conditional distribution was given by $1 + \text{Bin}(U_{n-1}, 1/2)$ (here Bin(n, p) denotes a binomial random variable with parameters n and p). This is to mean that if a size of a permutation tableau with U_{n-1} unrestricted rows was increased from n-1 to n, then the number of unrestricted rows in this extension had (the conditional) distribution $1 + \text{Bin}(U_{n-1}, 1/2)$). As it turns out, in the case of staircase tableaux the role of a fundamental parameter is played by the number of α/γ rows.

To make our approach work we need to know two things. One is the knowledge of the conditional distribution of statistics of interest in extension of a given tableau of size n-1. This would allow us to use the so-called tower property of the conditional expectation to reduce the size of staircase tableaux by 1. Roughly speaking, this reducing would amount to relating a function of a statistic on, say, S_n to another function of the same statistics on S_{n-1} . Keeping track of how the functions are related would in principle allow for an iteration of this process to reduce the size of the tableaux to 1. There is, however, another difficulty. Namely, when passing from S_n to S_{n-1} , not only a function of statistic changes but also a measure over which we integrate. More specifically, the image of the uniform probability measure P_n on S_n under a natural mapping associated with our reduction does not yield the uniform probability measure P_{n-1} on S_{n-1} . To handle this difficulty we need a second ingredient which we refer to as the change of measure. It simply describes the relation between the two probability measures and allows to return to the uniform probability measure on S_{n-1} after the reduction (see Section 3.2 for more details).

In the reminder of this section we develop the details for the number of α/γ rows, a key statistic in further considerations. We do it in two separate subsections. In the first we analyze the evolution of the number of α/γ rows as the size of the tableaux increases. The key step is the computation of the conditional generating function of r_n as the size of the tableau is extended from n - 1 to n (Lemma 3 below). In the second subsection we discuss the change of measure.

3.1. The conditional generating function for r_n . Our analysis progresses in steps. We first find the number of extensions of a given tableau of size n - 1. We then find the conditional distribution of the number of α/γ rows. This enables us to compute the conditional generating function of the number of such rows in an extension of a given tableau of size n - 1. The formula expresses this generating function in terms of the number of α/γ rows of the tableau being extended and is of crucial importance for our approach.

To begin the analysis we need to briefly recall the evolution process of staircase tableaux described by Corteel, Stanton, and Williams in [8]. Let S be a staircase tableau of size n - 1. To extend its size by 1, we add a new column of length n at the left end and we fill it according to the rules. Any such filling gives a tableau

of size n which is an extension of S. The number of extensions clearly depends on the number of α/γ rows of S and is as follows:

Lemma 1. Let $S \in S_{n-1}$ be a tableau with $r_{n-1} = r_{n-1}(S) \alpha / \gamma$ rows. Then there are $4 \cdot 3^{r_{n-1}}$ different staircase tableaux of size n that are extensions of S.

Proof. We consider two cases:

1. If $r_{n-1} = 0$ then all n-1 rows of S are β/δ rows and hence the top n-1 boxes of the new column have to be empty since no entries are allowed to the left of a β/δ in the same row. Thus, if $r_{n-1} = 0$ we obtain four different tableaux of size n by putting one of the four symbols in the bottom box of the new column.

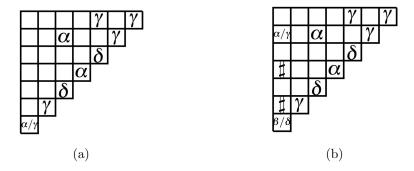


FIGURE 2. Extension of a tableau of size 6: (a) α/γ is put in the bottom corner of the 7th column. All boxes above it are to be empty. (b) β/δ is put in the bottom box and α/γ in the 3rd of the four α/γ rows in the 7th column. A box in an α/γ row above is to be empty; \sharp 's in two α/γ rows below it indicate that these boxes can be filled with β/δ or left empty.

2. If $r_{n-1} \geq 1$ we count the number of extensions as follows (see Figure 2 for an illustration): we can either put one of the symbols α/γ in the bottom corner (and then we are forced to leave all other boxes in that column empty), or we can put a β/δ in the bottom box of the new column. In the latter case, we need to fill the r_{n-1} boxes in the new column corresponding to the α/γ rows of S. According to the rules, if we put an α/γ in any of them, then we need to leave all boxes above it empty, otherwise we have a complete freedom. It follows from these considerations that any tableau of size n with $r_{n-1} \alpha/\gamma$ rows gives $4 \cdot 3^{r_{n-1}}$ different staircase tableaux of size n. Indeed, if $r_{n-1} = 0$ then there are 4 extensions and if $r_{n-1} \ge 1$ then there are

$$2 + 2(2(1 + 3 + \dots + 3^{r_{n-1}-1}) + 3^{r_{n-1}})$$

extensions. Here, the first 2 is from putting an α/γ in the bottom box of the new column, the next 2 is from putting a β/δ in that box, the term $2 \cdot 3^{i-1}$, $1 \le i \le r_{n-1}$, is from putting the first α/γ in a box of the new column corresponding to the *i*th (counting from bottom) α/γ row of S as then there are 3^{i-1} ways of filling the earlier i-1 boxes with symbols β , δ , or leaving them empty, and finally, the $3^{r_{n-1}}$ term comes from not putting an α/γ in any of the r_{n-1} boxes and thus filling them with β/δ 's or leaving them empty.

Summing the above gives

$$2 + 2\left(2\frac{3^{r_{n-1}} - 1}{2} + 3^{r_{n-1}}\right) = 2 + 2(3^{r_{n-1}} - 1 + 3^{r_{n-1}}) = 4 \cdot 3^{r_{n-1}},$$

as claimed.

As the next step we find the conditional distribution of the number of α/γ rows over the extensions of a given tableau of size n-1. To do that we phrase the preceding discussion in a more probabilistic language using what Shiryaev [22, Chapter I, §8] refers to as "decompositions" (which is just a special case of conditioning with respect to σ -algebra).

Note that every tableau from S_n is an extension of a unique tableau S from S_{n-1} . Therefore, denoting by D_S the set of all tableaux from S_n which are obtained from S by the process described above (as we just discussed, there are $4 \cdot 3^{r_{n-1}(S)}$ such tableaux), we can write

$$\mathcal{S}_n = \bigcup_{S \in \mathcal{S}_{n-1}} D_S,$$

where $(D_S)_{S \in S_{n-1}}$ are pairwise disjoint, non–empty subsets of S_n . We denote this decomposition of S_n by \mathcal{D}_{n-1} . We wish to compute the conditional probabilities $\mathsf{P}(\ \cdot |\mathcal{D}_{n-1})$ and the conditional expectations $\mathsf{E}(\ \cdot |\mathcal{D}_{n-1})$ with respect to this decomposition. We have

Lemma 2. Let $S \in S_{n-1}$ be a tableau with $r_{n-1} = r_{n-1}(S) \alpha / \gamma$ rows and let r be the number of such rows in any of its extensions to a tableau of size n. Then

(1)
$$P(r = r_{n-1} + 1|D_S) = \frac{1}{2 \cdot 3^{r_{n-1}}},$$

and, for $k = 0, 1, \ldots, r_{n-1}$

(2)
$$P(r = r_{n-1} - k | D_S) = \frac{2^{k+1} (2\binom{r_{n-1}}{k+1} + \binom{r_{n-1}}{k})}{4 \cdot 3^{r_{n-1}}}.$$

Proof. Let S and r be as in the statement. Clearly, the possible values for r are $r_{n-1} + 1, r_{n-1}, r_{n-1} - 1, \ldots, 1, 0$ and we need to know the probability of each of these possibilities. First, $r = r_{n-1} + 1$ means that we must have placed an α/γ in the bottom box of the new column since this is the only way we can increase the number of α/γ rows. Consequently, all other boxes in the new column are to remain empty. Obviously there are two such extensions which in view of Lemma 2 gives

$$\mathsf{P}(r = r_{n-1} + 1|D_S) = \frac{2}{4 \cdot 3^{r_{n-1}}} = \frac{1}{2 \cdot 3^{r_{n-1}}},$$

which justifies (1).

Next, for $k = 0, 1, \ldots, r_{n-1}$ we compute $\mathsf{P}(r = r_{n-1} - k | D_S)$. Since k is the number of β/δ 's that we put in the r_{n-1} allowable (i.e. corresponding to α/γ rows of S) boxes, $r = r_{n-1}-k$ means that we put a β/δ in the bottom box and additional $k \beta/\delta$'s in the r_{n-1} allowable boxes above it. If we do not put an α/γ in any of those k boxes, we have $2 \cdot 2^k \binom{r_{n-1}}{k}$ possibilities (2 is for putting either a β or a δ at the bottom, and the rest accounts for putting β/δ 's in any k of the allowable r_{n-1} boxes). If we do put an α/γ in one of the boxes we need to pick k + 1 of the allowable r_{n-1} boxes, put an α/γ in the topmost of them and put β/δ 's in the remaining k of them (and at the bottom). This gives $2^{k+2} \binom{r_{n-1}}{k+1}$ possibilities.

Adding up the two pieces and dividing by the total number of extensions given in Lemma 2 we obtain,

$$\mathsf{P}(r = r_{n-1} - k | D_S) = \frac{2 \cdot 2^k \binom{r_{n-1}}{k} + 2^{k+2} \binom{r_{n-1}}{k+1}}{4 \cdot 3^{r_{n-1}}} = \frac{2^{k+1} (2\binom{r_{n-1}}{k+1} + \binom{r_{n-1}}{k})}{4 \cdot 3^{r_{n-1}}}$$
required.

as required.

Lemma 2 completely describes the conditional distribution of r_n given \mathcal{D}_{n-1} and leads, in particular, to the following basic relation:

Lemma 3. For a complex number z and $n \ge 1$ (with the understanding that $r_0 \equiv 0$)

(3)
$$E(z^{r_n}|\mathcal{D}_{n-1}) = \frac{z+1}{2} \left(\frac{z+2}{3}\right)^{r_{n-1}}.$$

Proof. Recall that as we extend a tableau with $r_{n-1} \alpha / \gamma$ rows, we have $r_n = r_{n-1} + 1$ if and only if we put an α/γ at the bottom of the new column and $r_n = r_{n-1} - k$ if and only if we put a β/δ there and additional $k \beta/\delta$'s in the r_{n-1} boxes above it. Therefore, letting C_* denote the event that we put the symbol * at the bottom of the new column we have

$$\begin{split} \mathsf{E}(z^{r_n}|\mathcal{D}_{n-1}) &= \mathsf{E}(z^{r_n}I_{C_{\alpha/\gamma}}|\mathcal{D}_{n-1}) + \mathsf{E}(z^{r_n}I_{C_{\beta/\delta}}|\mathcal{D}_{n-1}) \\ &= z^{r_{n-1}+1}\frac{1}{2\cdot 3^{r_{n-1}}} + \sum_{k=0}^{r_{n-1}} z^{r_{n-1}-k}\mathsf{P}(r_n = r_{n-1} - k|\mathcal{D}_{n-1}) \\ &= \frac{z}{2}\left(\frac{z}{3}\right)^{r_{n-1}} + \frac{2}{4\cdot 3^{r_{n-1}}}\sum_{k=0}^{r_{n-1}} z^{r_n-k}2^k \left(2\binom{r_{n-1}}{k+1} + \binom{r_{n-1}}{k}\right) \right) \\ &= \frac{z}{2}\left(\frac{z}{3}\right)^{r_{n-1}} + \frac{1}{2}\left(\frac{2+z}{3}\right)^{r_{n-1}} + \frac{1}{3^{r_{n-1}}}\sum_{k=0}^{r_{n-1}-1} z^{r_n-k}2^k \binom{r_{n-1}}{k+1} \right\end{split}$$

The last term is

$$\frac{z}{2\cdot 3^{r_{n-1}}} \sum_{k=0}^{r_{n-1}-1} 2^{k+1} z^{r_n-k-1} \binom{r_{n-1}}{k+1} = \frac{z}{2\cdot 3^{r_{n-1}}} \left((2+z)^{r_{n-1}} - z^{r_{n-1}} \right),$$

so that we obtain

$$\mathsf{E}(z^{r_n}|\mathcal{D}_{n-1}) = \frac{z}{2} \left(\frac{z}{3}\right)^{r_{n-1}} + \frac{1}{2} \left(\frac{2+z}{3}\right)^{r_{n-1}} + \frac{z}{2} \left(\frac{2+z}{3}\right)^{r_{n-1}} - \frac{z}{2} \left(\frac{z}{3}\right)^{r_{n-1}}$$
$$= \frac{1+z}{2} \left(\frac{2+z}{3}\right)^{r_{n-1}}.$$
This proves (3).

This proves (3).

As we already mentioned, Lemma 3 (or its versions) plays a crucial role in our approach.

3.2. The change of measure. In this subsection we describe the second ingredient, namely the change of measure. It it is necessitated by the fact that there are two different probability measures on \mathcal{S}_{n-1} that naturally appear in our considerations. The first is, of course, the uniform measure P_{n-1} . We discuss the second one and the relation between them below.

Consider S_{n-1} , the set of all staircase tableaux of size n-1. The second probability measure on \mathcal{S}_{n-1} is obtained from the uniform probability measure on \mathcal{S}_n by "collapsing" all the elements of S_n that are extensions of the same element $S \in S_{n-1}$. More formally, consider a map $f : S_n \to S_{n-1}$ defined by f(T) = S if and only if T is an extension of S. The measure of interest is the image of P_n under f. We denote this measure $\mathsf{P}_n(S)$ (there is an apparent ambiguity of notation here, however, it disappears once we remember whether S is in S_n or S_{n-1}). Since both of these measures appear in the course of our argument, it is important to find the relationship between them. But this is straightforward: since a tableau $S \in S_{n-1}$ with $r_{n-1} \alpha / \gamma$ rows gives $4 \cdot 3^{r_{n-1}}$ tableaux in S_n we have

(4)
$$\mathsf{P}_{n}(S) = \frac{4 \cdot 3^{r_{n-1}}}{|\mathcal{S}_{n}|} = \frac{4 \cdot 3^{r_{n-1}}|\mathcal{S}_{n-1}|}{|\mathcal{S}_{n}|} \frac{1}{|S_{n-1}|} = 4 \cdot 3^{r_{n-1}} \frac{|\mathcal{S}_{n-1}|}{|\mathcal{S}_{n}|} \mathsf{P}_{n-1}(S).$$

Consequently, for any random variable X_{n-1} on S_{n-1} we have

(5)
$$\mathsf{E}_{n}X_{n-1} = \mathsf{E}_{n-1}4 \cdot 3^{r_{n-1}} \frac{|\mathcal{S}_{n-1}|}{|\mathcal{S}_{n}|} X_{n-1} = 4 \frac{|\mathcal{S}_{n-1}|}{|\mathcal{S}_{n}|} \mathsf{E}_{n-1}3^{r_{n-1}} X_{n-1}.$$

Here we have used the same convention as above; for a random variable X on S_{n-1} , $\mathsf{E}_{n-1}X$ denotes the expectation with respect to the uniform measure on S_{n-1} while $\mathsf{E}_n X$ denotes the expectation with respect to the measure that is induced on S_{n-1} by the uniform measure on S_n .

The relations (3) and (5) are key and will allow us to analyze the distributions of the various statistics on S_n . Note that (4) and (5) are true regardless of whether we know the cardinalities of S_{n-1} and S_n or not. As a matter of fact, one can use (5) to provide a yet another argument that $|S_n| = 4^n n!$ as we will see in the next section.

4. Generating function for the number of α/γ rows and some consequences

In this section we demonstrate how we intend to apply (3) and (5) to derive recurrences for generating functions that upon solving yield information on the corresponding statistics. We focus here on the number of α/γ rows which, on one hand is the easiest to analyse, and on the other is central in the analysis of other statistics.

Proposition 4. For every complex number z we have

(6)
$$E_n z^{r_n} = \frac{2^n}{|\mathcal{S}_n|} \prod_{k=1}^n (z+2k-1).$$

Proof. By the basic properties of the conditional expectation (see e.g. [22, Formula (16), p. 79]) the expectation on the right-hand side is equal to $\mathsf{E}_n\mathsf{E}(z^{r_n}|\mathcal{D}_{n-1})$. Using first (3) and then (5) we further get

$$\mathsf{E}_{n} z^{r_{n}} = \frac{z+1}{2} \mathsf{E}_{n} \left(\frac{z+2}{3}\right)^{r_{n-1}} = \frac{(z+1)}{2} \frac{4|\mathcal{S}_{n-1}|}{|\mathcal{S}_{n}|} \mathsf{E}_{n-1} 3^{r_{n-1}} \left(\frac{z+2}{3}\right)^{r_{n-1}}$$
$$= 2(z+1) \frac{|\mathcal{S}_{n-1}|}{|\mathcal{S}_{n}|} E_{n-1} (z+2)^{r_{n-1}}.$$

Applying the same procedure with z replaced by z + 2 and n by n - 1 we obtain

$$E_n z^{r_n} = 2^2 (z+1)(z+3) \frac{|\mathcal{S}_{n-2}|}{|\mathcal{S}_n|} E_{n-2} (z+4)^{r_{n-2}}.$$

Further iteration yields

$$\mathsf{E}_{n} z^{r_{n}} = 2^{n-1} (z+1)(z+3) \dots (z+2n-3) \frac{|\mathcal{S}_{1}|}{|\mathcal{S}_{n}|} \mathsf{E}_{1} (z+2(n-1))^{r_{1}}.$$

Since $|\mathcal{S}_1| = 4$ and $\mathsf{E}_1(z+2(n-1))^{r_1} = \frac{1}{2}(z+2(n-1)) + \frac{1}{2}$ the above can be written as

$$\mathsf{E}_{n} z^{r_{n}} = \frac{2^{n}}{|\mathcal{S}_{n}|} (z+1)(z+3) \dots (z+2n-1),$$

which is precisely (6).

Proposition 4 has a number of consequences. First, by putting z = 1 in (6) we obtain an independent confirmation of the count of staircase tableau of a given size.

Corollary 5. Let S_n be the set of all staircase tableaux of size $n \ge 1$. Then

$$|\mathcal{S}_n| = 4^n n!$$

Note that once this corollary is known (4) and (5) simplify to

(7)
$$\mathsf{P}_n(S) = \frac{3^{r_{n-1}}}{n} \mathsf{P}_{n-1}(S) \text{ and } \mathsf{E}_n X_{n-1} = \frac{1}{n} \mathsf{E}_{n-1} 3^{r_{n-1}} X_{n-1},$$

respectively, and this is the form we will be using from now on.

Next, by combining this corollary with (6) we obtain

Corollary 6. The probability generating function of the number of α/γ rows in a random staircase tableau of size n is given by

$$E_n z^{r_n} = \frac{2^n}{4^n n!} \prod_{k=1}^n (z+2k-1) = \prod_{k=1}^n \frac{z+2k-1}{2k}.$$

The last corollary gives, in turn, a complete information on the distribution of r_n .

Corollary 7. For every $n \ge 1$ we have

(8)
$$r_n \stackrel{d}{=} \sum_{k=1}^n J_k,$$

where J_k 's are independent and J_k is a random variable which is 1 with probability 1/(2k) and 0 with the remaining probability. In particular,

(9)
$$E_n r_n = \sum_{k=1}^n \frac{1}{2k} = \frac{H_n}{2}, \quad \operatorname{var}(r_n) = \sum_{k=1}^n \frac{1}{2k} \left(1 - \frac{1}{2k} \right) = \frac{H_n}{2} - \frac{H_n^{(2)}}{4}$$

where $H_n = \sum_{k=1}^n \frac{1}{k}$ and $H_n^{(2)} = \sum_{k=1}^n \frac{1}{k^2}$ are harmonic numbers of the first and second order, respectively. Furthermore,

(10)
$$\frac{r_n - \frac{\ln n}{2}}{\sqrt{\frac{\ln n}{2}}} \xrightarrow{d} N(0, 1).$$

Proof. Note that a factor

$$\frac{z+2k-1}{2k} = \frac{z}{2k} + 1(1-\frac{1}{2k}),$$

given in the previous corollary is the probability generating function of a random variable J_k which is 1 with probability 1/(2k) and 0 with the remaining probability.

$$\square$$

Since the product of the probability generating functions corresponds to adding independent random variables, we obtain (8) and thus also (9). Finally, since J_k are uniformly bounded and variances of their partial sums go to infinity, the Lindeberg condition for the central limit theorem (see e.g. [22, Chapter III §4]) holds trivially. Since as $n \to \infty$, $\mathsf{E}_n r_n \sim \operatorname{var}(r_n) \sim \frac{\ln n}{2}$, (10) holds.

5. Main results

Our technique allows us to obtain further results concerning the distributions (sometimes exact, sometimes only asymptotic) of the statistics discussed above. We gather our results in the following two statements, concerning the total number of entries and the number of entries on the diagonal, respectively. Recall that Δ_n and Γ_n denote the total number of β/δ and α/γ in the tableau of size n, respectively.

Theorem 8. Consider the set S_n with the uniform probability measure P_n . Then: (i) For every $n \ge 1$ we have

(11)
$$\Delta_n \stackrel{d}{=} \sum_{k=1}^n I_k,$$

where (I_k) are independent and ${\it P}(I_k=1)=1-\frac{1}{2k}, \ {\it P}(I_k=0)=\frac{1}{2k}$. In particular,

(12)
$$E_n \Delta_n = n - \frac{H_n}{2}, \quad \operatorname{var}(\Delta_n) = \frac{H_n}{2} - \frac{H_n^{(2)}}{4},$$

and, as $n \to \infty$,

(13)
$$\frac{\Delta_n - n + \frac{1}{2}\ln n}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} N(0, 1)$$

(ii) For every $n \ge 1$

 $\Gamma_n \stackrel{d}{=} \Delta_n.$ In particular, (11), (12), and (13) hold with Δ_n replaced by Γ_n .

Our second result concerns the entries on the diagonal. Recall that A_n (resp. B_n) denote the number of α/γ (resp. β/δ) on the diagonal of a tableau of size n. For these parameters we get

Theorem 9. The expected value and the variance of the number A_n of α/γ on the diagonal of a random staircase tableau of size n are, respectively,

(14)
$$E_n A_n = \frac{n}{2} \quad and \quad \operatorname{var}(A_n) = \frac{n+1}{12}.$$

Moreover,

(15)
$$\frac{A_n - n/2}{\sqrt{n/12}} \xrightarrow{d} N(0, 1).$$

Furthermore, for every $n \ge 1$ we have

$$B_n \stackrel{d}{=} A_n.$$

In particular, (14) and (15) hold for B_n in place of A_n .

Remark: While it seems intuitively clear that the expected number of letters α/γ on the diagonal is n/2 as (14) asserts, the expression for the variance is much less intuitive. It implies, in particular, that the entries α/γ and β/δ along the diagonal are not chosen independently from one another with equal probabilities as one might have hoped (if that were the case the variance would be n/4).

6. PROOFS

We begin by observing that, as is obvious from the definition, staircase tableaux are symmetric under taking the transpose and exchanging the α 's with the β 's and the γ 's with the δ 's. It therefore is immediate that parts (ii) of both theorems follow from the respective parts (i). Furthermore, the proof of Theorem 8 (i) may be completed by using Corollary 7 and the relation

(17)
$$r_n(S) + \Delta_n(S) = n$$

since then

$$\Delta_n = n - r_n = \sum_{k=1}^n (1 - J_k) \stackrel{d}{=} \sum_{k=1}^n I_k.$$

But (17) is clear once we notice that for any row of a staircase tableau exactly one of the following statements is true

- it contains a β/δ
- it is indexed by α/γ .

This proves Theorem 8 and we now turn our attention to a considerably more involved proof of Theorem 9 (i).

6.1. Proof of Theorem 9 (i). The idea is the same as for Corollary 7 except that this time we will actually need the joint probability generating function of A_n and r_n . The final expression will turn out to be substantially more complicated than what we encountered earlier and thus harder to analyze. Nonetheless, the situation is quite analogous to the case of the number of rows in permutation tableaux (see [20, Section 4]). For the reader's convenience we break up our proof into several steps each of them discussed in a separate section below. We now briefly outline the major steps in the proof indicating the section in which they are treated. We begin in the forthcoming section with the derivation of the probability generating function. Its coefficients satisfy certain recurrences. This, in particular, enables us to derive the exact formulas for the expected value and the variance of A_n (see Subsection 6.1.2 below). Furthermore, the nature of these recurrences suggests that the coefficients are related to the classical Eulerian number associated with the number of rises in random permutations. In fact, our coefficients exactly match the numbers often called the "Eulerian numbers of type B". We establish and discuss further this connection in Subsection 6.1.3. We think it is of independent interest and perhaps worthy of further exploration.

Once this connection is made it is then expected that the the coefficients follow the normal law (just as the classical Eulerian numbers do). As a matter of fact, one of the proofs (although not the first) of the asymptotic normality of the classical Eulerian numbers is via a fairly general device due to Bender ([1]) and nowadays referred to as Bender's theorem. This is, indeed, the approach we take and in the subsection 6.1.4 we verify the conditions of Bender's theorem to conclude our proof. 6.1.1. Bivariate generating function. We give the formula for the joint generating function of r_n and A_n in the following statement.

Proposition 10. Let z, t be complex numbers. Then we have

(18)
$$E_n t^{A_n} z^{r_n} = \frac{1}{2^n n!} \sum_{k=0}^n (t-1)^{n-k} c_{n,k}(z),$$

where the coefficients $\{c_{m,\ell}(z): 0 \leq \ell \leq m \leq n\}$ satisfy the following recurrence

$$c_{m+1,\ell}(z) = (z+2\ell)c_{m,\ell}(z) + (z+2(\ell-1)+1)c_{m,\ell-1}(z), \quad 1 \le \ell \le m,$$

with the following boundary conditions:

$$c_{0,0}(z) = 1$$
, $c_{m+1,m+1}(z) = (z + 2m + 1)c_{m,m}(z)$, $c_{m+1,0}(z) = zc_{m,0}(z)$.

Proof. If n = 0 and we set $c_{0,0}(z) = 1$ then both sides of (18) are 1. Otherwise, let I_j indicate the event that we put an α/γ in the *j*th box on the diagonal (counting from NE to SW). Then $A_n = \sum_{j=1}^n I_j$ and we have

$$\mathsf{E}_{n}t^{A_{n}}z^{r_{n}} = \mathsf{E}_{n}\mathsf{E}(t^{A_{n-1}+I_{n}}z^{r_{n}}|\mathcal{D}_{n-1}) = \mathsf{E}_{n}t^{A_{n-1}}\mathsf{E}(t^{I_{n}}z^{r_{n}}|\mathcal{D}_{n-1}),$$

where we have used the basic properties of the conditional expectations (see [22, Formulas (16), p. 79 and (17), p. 80]). Now, $I_n = 1$ means that we put an α/γ in the SW corner. In that case we have $r_n = r_{n-1} + 1$ and since this happens with probability $1/(2 \cdot 3^{r_{n-1}})$ we get

$$\mathsf{E}(t^{I_n} z^{r_n} | \mathcal{D}_{n-1}) = t z^{r_{n-1}+1} \frac{1}{2 \cdot 3^{r_{n-1}}} + \mathsf{E}(t^{I_n} z^{r_n} I_{I_n=0} | \mathcal{D}_{n-1}).$$

The second term is equal to

$$\begin{split} \mathsf{E}(z^{r_n}I_{I_n=0}|\mathcal{D}_{n-1}) &= & \mathsf{E}(z^{r_n}|\mathcal{D}_{n-1}) - \mathsf{E}(z^{r_n}I_{I_n=1}|\mathcal{D}_{n-1}) \\ &= & \frac{z+1}{2}\left(\frac{z+2}{3}\right)^{r_{n-1}} - \frac{z^{r_{n-1}+1}}{2\cdot 3^{r_{n-1}}}. \end{split}$$

Combining we obtain

$$\mathsf{E}(t^{A_n} z^{r_n} | \mathcal{D}_{n-1}) = \frac{z(t-1)}{2} \left(\frac{z}{3}\right)^{r_{n-1}} + \frac{z+1}{2} \left(\frac{z+2}{3}\right)^{r_{n-1}}$$

Using the second part of (7) leads to the basic recurrence

$$\mathsf{E}_{n} t^{A_{n}} z^{r_{n}} = \frac{z(t-1)}{2} \mathsf{E}_{n} t^{A_{n-1}} \left(\frac{z}{3}\right)^{r_{n-1}} + \frac{z+1}{2} \mathsf{E}_{n} t^{A_{n-1}} \left(\frac{z+2}{3}\right)^{r_{n-1}} \\ = \frac{1}{2n} \left\{ z(t-1) \mathsf{E}_{n-1} t^{A_{n-1}} z^{r_{n-1}} + (z+1) \mathsf{E}_{n-1} t^{A_{n-1}} (z+2)^{r_{n-1}} \right\}.$$

Upon further iteration of this relation we obtain for any $0 \le m < n$:

$$\mathsf{E}_{n}t^{A_{n}}z^{r_{n}} = \frac{1}{2^{m}n(n-1)\dots(n-m+1)}\sum_{\ell=0}^{m}(t-1)^{m-\ell}c_{m,\ell}\mathsf{E}_{n-m}t^{A_{n-m}}(z+2\ell)^{r_{n-m}}z^{r_{n-m}$$

for some coefficients $c_{m,\ell} = c_{m,\ell}(z)$, $0 \le \ell \le m$. To see that they satisfy the stated recurrence, we apply the basic recurrence (with n-m instead of n and $z+2\ell$ instead of z) to the expectation on the right-hand side above. We get that it is equal to

$$\frac{1}{2(n-m)} \Big\{ (z+2\ell)(t-1) \mathsf{E}_{n-m-1} t^{A_{n-m-1}} (z+2\ell)^{r_{n-m-1}} \\ + (z+2\ell+1) \mathsf{E}_{n-m-1} t^{A_{n-m-1}} (z+2(\ell+1))^{r_{n-m-1}} \Big\}.$$

Substituting this into the above formula for $\mathsf{E}_n t^{A_n} z^{r_n}$ and multiplying both sides by $2^{m+1}n(n-1)\cdots(n-m)$ (to avoid writing a denominator on the right-hand side) we get

$$\sum_{\ell=0}^{m} (t-1)^{m-\ell} c_{m,\ell} \Big\{ (z+2\ell)(t-1) \mathsf{E}_{n-m-1} t^{A_{n-m-1}} (z+2\ell)^{r_{n-m-1}} + (z+2\ell+1) \mathsf{E}_{n-m-1} t^{A_{n-m-1}} (z+2(\ell+1))^{r_{n-m-1}} \Big\}.$$

Splitting this in two sums, isolating the $\ell = 0$ term in the first, and the $\ell = m$ in the second, and then shifting the index in the second sum, we further obtain

$$\begin{split} &(t-1)^{m}c_{m,0}z(t-1)\mathsf{E}_{n-(m+1)}t^{A_{n-(m+1)}}z^{r_{n-(m+1)}}\\ &+\sum_{\ell=1}^{m}(t-1)^{m-\ell}c_{m,\ell}(z+2\ell)(t-1)\mathsf{E}_{n-m-1}t^{A_{n-m-1}}(z+2\ell)^{r_{n-m-1}}\\ &+\sum_{\ell=0}^{m-1}(t-1)^{m-\ell}c_{m,\ell}(z+2\ell+1)\mathsf{E}_{n-(m+1)}t^{A_{n-(m+1)}}(z+2(\ell+1))^{r_{n-(m+1)}}\\ &+c_{m,m}(z+2m+1)\mathsf{E}_{n-(m+1)}t^{A_{n-(m+1)}}(z+2(m+1))^{r_{n-(m+1)}}\\ &=(t-1)^{m+1}c_{m,0}z\mathsf{E}_{n-(m+1)}t^{A_{n-(m+1)}}z^{r_{n-(m+1)}}\\ &+\sum_{\ell=1}^{m}(t-1)^{m+1-\ell}c_{m,\ell}(z+2\ell)\mathsf{E}_{n-(m+1)}t^{A_{n-(m+1)}}(z+2\ell)^{r_{n-(m+1)}}\\ &+\sum_{\ell=1}^{m}(t-1)^{m-(\ell-1)}c_{m,\ell-1}(z+2(\ell-1)+1))\mathsf{E}_{n-(m+1)}t^{A_{n-(m+1)}}(z+2\ell)^{r_{n-(m+1)}}. \end{split}$$

So, if we write it as

$$\sum_{\ell=0}^{m+1} (t-1)^{m+1-j} c_{m+1,\ell} \mathsf{E}_{n-(m+1)} t^{A_{n-(m+1)}} (z+2\ell)^{r_{n-(m+1)}},$$

as postulated above, we see that

$$c_{m+1,0} = zc_{m,0}, \quad c_{m+1,m+1} = (z+2m+1)c_{m,m}$$

and, upon combining the two middle sums, that

$$c_{m+1,\ell} = (z+2\ell)c_{m,\ell} + (z+2(\ell-1)+1)c_{m,\ell-1}, \text{ for } 1 \le \ell \le m$$

Taking m = n - 1 we get

$$\mathsf{E}_{n}t^{A_{n}}z^{r_{n}} = \frac{1}{2^{n-1}n!}\sum_{\ell=0}^{n-1}(t-1)^{n-1-\ell}c_{n-1,\ell}\mathsf{E}_{1}t^{A_{1}}(z+2\ell)^{r_{1}},$$

and since

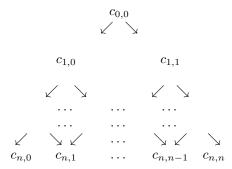
$$\mathsf{E}_1 t^{A_1}(z+2\ell)^{r_1} = \frac{1}{2}t(z+2\ell) + \frac{1}{2} = \frac{1}{2}(t-1)(z+2\ell) + \frac{1}{2}(z+2\ell+1),$$

we can write

$$\mathsf{E}_n t^{A_n} z^{r_n} = \frac{1}{2^n n!} \sum_{k=0}^n (t-1)^{n-k} c_{n,k},$$

where the coefficients $\{c_{m,\ell}: 0 \leq \ell \leq m \leq n\}$ satisfy the stated recurrence and the boundary conditions. The proof is complete.

6.1.2. Expected value and the variance. A number of properties of (A_n) can be deduced from Proposition 10. We illustrate this by obtaining the exact expression for the expected value and for the variance of A_n . To help facilitate that we put the coefficients $(c_{n,k})$ in a Pascal type triangle.



The SW move from a coefficient $c_{m,\ell}$ has weight $z + 2\ell$ and a SE move has weight $z + 2\ell + 1$. The value of a given coefficient is obtained by summing over all possible paths leading to it from the root $c_{0,0}$ the products of weighs corresponding to the moves along the path. For example, there is only one path leading to $c_{n,n}$ (all moves are SE) and hence

$$c_{n,n} = \prod_{j=0}^{n-1} (z+2j+1).$$

Likewise, paths leading to $c_{n,n-1}$ have exactly one SW move; thus there are n of them and if the sole SW move is from $c_{k,k}$, $0 \le k \le n-1$, then the weight of that path is

$$\left(\prod_{j=0}^{k-1} (z+2j+1)\right) (z+2k) \left(\prod_{j=k}^{n-2} (z+2j+1)\right).$$

Consequently,

$$c_{n,n-1}(z) = \sum_{k=0}^{n-1} \left(\prod_{j=0}^{k-1} (z+2j+1) \right) (z+2k) \left(\prod_{j=k}^{n-2} (z+2j+1) \right).$$

The significance of this is that

$$\begin{aligned} \mathsf{E}_{n}A_{n} &= \frac{\partial}{\partial t}\mathsf{E}_{n}t^{A_{n}}z_{\mid t=1, z=1}^{r_{n}} = \frac{1}{2^{n}n!}c_{n, n-1}(1) \\ &= \frac{1}{2^{n}n!}\sum_{k=0}^{n-1}\left(\prod_{j=0}^{k-1}2(j+1)\right)(2k+1)\left(\prod_{j=k}^{n-2}2(j+1)\right). \\ &= \frac{1}{2^{n}n!}2^{n-1}(n-1)!\sum_{k=0}^{n-1}(2k+1) = \frac{n}{2}. \end{aligned}$$

This proves the first part of (14). Similarly, to compute the variance we use $var(A_n) = \mathsf{E}A_n(A_n-1) + \mathsf{E}A_n - (\mathsf{E}A_n)^2$ and the fact that $\mathsf{E}A_n(A_n-1) =$

 $\frac{\partial^2}{\partial t^2}(\mathsf{E}t^{A_n})_{|t=1} = \frac{2}{2^n n!} c_{n,n-2}(1)$. Paths leading to $c_{n,n-2}(1)$ have exactly two SW moves, the first could be from any $c_{k,k}$, $0 \le k \le n-2$ and the second from $c_{\ell,\ell-1}$ for some $k < \ell \le n-1$. These two moves have weights 2k+1 and $2\ell-1$ respectively, and the remaining SE moves have weights

 $2 \cdot 1, 2 \cdot 2, \dots, 2k, 2(k+1), 2(k+2) \dots, 2(\ell-1), 2\ell, \dots, 2(n-1).$

Therefore,

$$c_{n,n-2}(1) = \sum_{k=0}^{n-2} \sum_{\ell=k+1}^{n-1} (2k+1)(2\ell-1)2^{n-2}(n-2)!$$

= $2^{n-2}(n-2)! \sum_{\ell=1}^{n-1} (2\ell-1) \sum_{k=0}^{\ell-1} (2k+1)$
= $2^{n-2}(n-2)! \sum_{\ell=1}^{n-1} (2\ell-1)\ell^2$
= $2^{n-2}(n-2)! \frac{1}{6}n(n-1)(3n^2-5n+1)$
= $2^{n-2}n!(\frac{1}{2}n^2-\frac{5}{6}n+\frac{1}{6}).$

Hence

$$\operatorname{var}(A_n) = 2\frac{2^{n-2}n!}{2^n n!} \left(\frac{1}{2}n^2 - \frac{5}{6}n + \frac{1}{6}\right) + \frac{n}{2} - \frac{n^2}{4} = \frac{n^2}{4} - \frac{5}{12}n + \frac{1}{12} + \frac{6n}{12} - \frac{n^2}{4} = \frac{n+1}{12}$$

Thus we have proved the second part of (14) as well

Thus we have proved the second part of (14) as well.

6.1.3. Connections to generalized Eulerian numbers. Before establishing the asymptotic normality of (A_n) we take a closer look at the doubly indexed sequence $\{c_{n,k}\}$ since it has intriguing connections that may be of interest in their own right. It is featured as entry A145901 in [23] and is closely related to another sequence from [23], namely A039755. More precisely, $c_{n,k} = 2^k k! W_2(n,k)$ where $W_m(n,k)$ are the Whitney numbers of the second kind satisfying the recurrence

$$W_m(n,k) = (mk+1)W_m(n-1,k) + W_m(n-1,k-1)$$

The numbers $(W_m(n,k))$ were introduced in [16] and their properties were studied in [2, 3, 4]. Since we are dealing exclusively with the case m = 2 we drop the subscript and we write W(n,k) for $W_2(n,k)$. Thus, the generating function of A_n may be written as

(19)
$$\psi_n(t) = \mathsf{E}_n t^{A_n} = \frac{1}{2^n n!} \sum_{k=0}^n 2^k k! W(n,k) (t-1)^{n-k}.$$

It is perhaps of interest to mention that the numbers $(2^k k! W(n, k))$ themselves satisfy the central (and local) limit theorem as was shown in [5]. However, this is not exactly what we want since the generating function above is in powers of t-1rather than t. In terms of powers of $t \psi_n(t)$ has the following form.

Proposition 11. The probability generating function of the number of α/γ entries on the diagonal of a random staircase tableau of size n has the form

(20)
$$\psi_n(t) = \frac{1}{2^n n!} \sum_{m=0}^n V(n,m) t^m,$$

where the numbers $\{V(n,m), 0 \le m \le n\}$ satisfy the boundary condition V(n,0) = 1, the symmetry relation V(n,m) = V(n,n-m), and the recurrence

(21)
$$V(n,m) = (2m+1)V(n-1,m) + (2(n-m)+1)V(n-1,m-1).$$

The explicit expression for V(n,m) is given by

(22)
$$V(n,m) = \sum_{k=0}^{n-m} 2^k k! W(n,k) \binom{n-k}{m} (-1)^{n-k-m}, \quad 0 \le m \le n.$$

Proof. We start out with (19). To rewrite it in powers of t, note that the mth derivative of ψ_n at t = 0 is

$$\psi_n^{(m)}(0) = \frac{1}{2^n n!} \sum_{k=0}^{n-m} 2^k k! W(n,k)(n-k) \cdots (n-k-(m-1))(-1)^{n-k-m}$$
$$= \frac{1}{2^n n!} \sum_{k=0}^{n-m} 2^k k! W(n,k) \frac{(n-k)!}{(n-k-m)!} (-1)^{n-k-m}.$$

Therefore,

$$\psi_n(t) = \sum_{m=0}^n \frac{\psi_n^{(m)}(0)}{m!} t^m = \frac{1}{2^n n!} \sum_{m=0}^n V(n,m) t^m$$

which shows that (20) holds with V(n,m) given by (22).

It remains to verify the claimed properties of the numbers V(n,m), $0 \le m \le n$. Since the symmetry condition follows by induction from the recurrence it suffices to verify the recurrence and the boundary condition. For the boundary condition, we see immediately that for $n \ge 0$ V(n,n) = W(n,0) = 1, so that once we verify the recurrence (and thus also the symmetry) we will have that V(n,0) = 1 for all $n \ge 0$. To verify (21), we use the basic recurrence for W(n,k)'s to write the left-hand side of (21) as

$$\begin{split} V(n,m) &= \sum_{k=0}^{n-m} 2^k k! \Big((2k+1)W(n-1,k) + W(n-1,k-1) \Big) \binom{n-k}{m} (-1)^{n-m-k} \\ &= \sum_{k=0}^{n-m} 2^k k! (2k+1)W(n-1,k) \binom{n-k}{m} (-1)^{n-m-k} \\ &+ \sum_{k=1}^{n-m} 2^k k! W(n-1,k-1) \binom{n-k}{m} (-1)^{n-m-k} \\ &= \sum_{k=0}^{n-m} 2^k k! (2k+1)W(n-1,k) \binom{n-k}{m} (-1)^{n-m-k} \\ &+ \sum_{k=0}^{n-m-1} 2^{k+1} (k+1)! W(n-1,k) \binom{n-k-1}{m} (-1)^{n-m-k-1} \\ &= 2^{n-m} (n-m)! (2(n-m)+1)W(n-1,n-m) \\ &+ \sum_{k=0}^{n-m-1} 2^k k! (2k+1)W(n-1,k) \binom{n-k}{m} (-1)^{n-m-k} \\ &- \sum_{k=0}^{n-m-1} 2^k k! (2(k+1))W(n-1,k) \binom{n-k-1}{m} (-1)^{n-m-k}. \end{split}$$

On the other hand, the right-hand side of (21) is

$$\begin{split} &(2m+1)\sum_{k=0}^{n-m-1}2^kk!W(n-1,k)\binom{n-1-k}{m}(-1)^{n-m-k-1}\\ &+(2(n-m)+1)\sum_{k=0}^{n-m}2^kk!W(n-1,k)\binom{n-k-1}{m-1}(-1)^{n-m-k}\\ &= &(2m+1)\sum_{k=0}^{n-m-1}2^kk!W(n-1,k)\binom{n-k-1}{m}(-1)^{n-m-k-1}\\ &+(2(n-m)+1)\sum_{k=0}^{n-m-1}2^kk!W(n-1,k)\binom{n-k-1}{m-1}(-1)^{n-m-k}\\ &+(2(n-m)+1)2^{n-m}(n-m)!W(n-1,n-m). \end{split}$$

So, we see that the coefficients in front of W(n-1, n-m) in both expressions are the same, and to complete the verification of (21) we need to see that the coefficients are the same for the remaining values of k, $0 \le k \le n-m-1$. Cancelling the common factors, we need to see that

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

= $(2(n-m)+1)\binom{n-k-1}{m-1} - (2m+1)\binom{n-k-1}{m}$.

18

But that is straightforward: using $\binom{n-k}{m} = \binom{n-k-1}{m} + \binom{n-k-1}{m-1}$ and grouping the terms this boils down to verifying that

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

or, equivalently, that

$$m\binom{n-k-1}{m} = (n-m-k)\binom{n-k-1}{m-1},$$

which follows immediately from the defining property of the binomial coefficients. $\hfill \square$

Remark: Triangle of numbers V(n,m), $0 \le m \le n$ is featured in the Online Encyclopedia of Integer Sequences [23] as a sequence A060187 (with the shift in indexing: V(n,m) = T(n+1,m+1)) and is called "Eulerian numbers of type B". This sequence can be traced back in the literature to MacMahon's paper [21] and it was subsequently studied in more detail in [19, Sec. 3.2] as a sequence $B_{n,k}(1)$. In particular, it appears that the expression for a bivariate generating function of (V(n,k)) was derived for the first time in [19]. We use this expression in the next section to derive the asymptotic normality of (A_n) .

6.1.4. Conclusion of the proof by Bender's theorem. Using the properties of the numbers V(n, m) given in Proposition 11 (and the identification with the sequence A060187 from [23]) we can complete the proof of Theorem 9 by establishing (15). To do that we will rely on a general theorem due to Bender [1, Theorem 1]. (Bender result is also described in [18]; see Section IX.6 in general, and Theorem IX.9, Example IX.12, and Proposition IX.9 in particular). Recall from [23] or [19, Formula (3.23)] (and see Section 3 of [19] for a proof) that the bivariate generating function of the numbers (V(n, k)), called in [19] $(B_{n,k}(1))$, is

$$\sum_{n \ge 0} \sum_{k=0}^{n} V(n,k) \frac{w^k z^n}{n!} = \frac{(1-w)e^{(1-w)z}}{1-we^{2(1-w)z}}.$$

Therefore, it follows from (20) that the bivariate probability generating function of the sequence (A_n) is

$$f(z,w) := \sum_{n \ge 0} \psi_n(z) z^n = \sum_{n \ge 0} \sum_{k=0}^n V(n,k) \frac{w^k}{n!} \left(\frac{z}{2}\right)^n = \frac{(1-w)e^{(1-w)z/2}}{1-we^{(1-w)z}},$$

where we define f(z, 1) = 1/(1-z). We now closely follow the way Bender applied his result. First,

$$f(z, e^s) = \frac{(1 - e^s)e^{(1 - e^s)z/2}}{1 - e^s e^{(1 - e^s)z}}$$

has a simple pole at $z = r(s) = s/(e^s - 1)$. Furthermore,

$$e^{(1-e^s)z} = e^{(1-e^s)(z-r(s))}e^{(1-e^s)r(s)} = e^{(1-e^s)(z-r(s))}e^{-s},$$

so that

$$f(z, e^s) = \frac{(1 - e^s)e^{-s/2}e^{(1 - e^s)(z - r(s))/2}}{1 - e^{(1 - e^s)(z - r(s))}} = \frac{(1 - e^s)e^{-s/2}}{e^{(e^s - 1)(z - r(s))/2} - e^{-(e^s - 1)(z - r(s))/2}}$$
Since for bounded u
$$\frac{1}{e^u - e^{-u}} = \frac{1}{2u} + O(1),$$

we get

$$f(z, e^s) = \frac{(1 - e^s)e^{-s/2}}{(e^s - 1)(z - r(s))} + O(1),$$

with the constant in O(1) bounded when both s and z - r(s) are close to 0. Thus, by a comment at the very beginning of Section 3 of [1], the conditions of Theorem 1 of that paper are satisfied, and hence the central limit theorem holds (with centering by $\mathsf{E}A_n \sim n/2$ and scaling by $\sqrt{\operatorname{var}(A_n)} \sim \sqrt{n/12}$ as implied by (14)). Alternatively, we see that

$$r(0) = 1$$
, $r'(0) = -\frac{1}{2}$, and $r''(0) = \frac{1}{6}$,

so that by [1, Theorem 1] $\mathsf{E}A_n \sim \frac{n}{2}$ and $\operatorname{var}(A_n) \sim \frac{n}{12}$ which conforms to what we have found in (14). In any event, (15) follows.

7. Conclusion and further remarks

In this paper we have developed a probabilistic approach to the analysis of properties of random staircase tableaux. Using this approach we established the asymptotic normality of several parameters associated with appearances of Greek letters α , β , γ , and δ in a randomly chosen tableau. We certainly hope that this approach will be useful in the analysis of other properties of staircase tableaux. From the combinatorial point of view, it would be of interest to analyse the number of appearances of the letter q in such tableaux. It is not clear at this point, that the method we develop is adequate to address that question, and if so, how difficult it would be to achieve. This is probably an issue worth resolving in the future. The fact that our method can be used to give new and rather complete results concerning Greek letters and the fact that it gives an easy way of enumerating of staircase tableaux of a given size makes us cautiously optimistic.

It is perhaps worth making the following point. In our arguments we relied on observations like (17) and symmetries of staircase tableaux to minimize the amount of work. We wish to emphasize, however, that our probabilistic approach does provide a unified and systematic way of analyzing each of the statistics we considered in a self-contained manner (i.e. not relying on relations between various parameters). In fact, a direct proof that Γ_n satisfies (11), (12), and (13) was given in [13]. This is partly a reason we believe that our approach has potential of being useful in addressing other questions concerning statistics of staircase tableaux. To reiterate that point we wish to briefly sketch a proof of (16) in Theorem 9 not relying on symmetries of the parameters.

7.1. The number of β/δ on the diagonal. We show that A_n and B_n have identical probability generating functions which, of course, implies (16). Write $B_n = \sum_{j=1}^n b_j$, where b_j is 1 if we put a β or δ in the *j*th box on the diagonal (counting SW from the top) and is 0 otherwise. We first derive the expression for

the conditional probability generating function: For z, t complex,

$$\begin{split} \mathsf{E}(t^{b_n} z^{r_n} | \mathcal{D}_{n-1}) &= \mathsf{E}(t^{b_n} z^{r_n} I_{b_n=1} | \mathcal{D}_{n-1}) + \mathsf{E}(t^{b_n} z^{r_n} I_{b_n=0} | \mathcal{D}_{n-1}) \\ &= t\mathsf{E}(z^{r_n} I_{b_n=1} | \mathcal{D}_{n-1}) + z^{r_{n-1}+1} P(b_n = 0 | \mathcal{D}_{n-1}) \\ &= t\mathsf{E}(z^{r_n} | \mathcal{D}_{n-1}) - t\mathsf{E}(z^{r_n} I_{b_n=0} | \mathcal{D}_{n-1}) + z^{r_{n-1}+1} \frac{1}{2 \cdot 3^{r_{n-1}}} \\ &= t \frac{z+1}{2} \left(\frac{z+2}{3}\right)^{r_{n-1}} - tz^{r_{n-1}} \frac{1}{2 \cdot 3^{r_{n-1}}} + \frac{z}{2} \left(\frac{z}{3}\right)^{r_{n-1}} \\ &= t \frac{z+1}{2} \left(\frac{z+2}{3}\right)^{r_{n-1}} + (1-t) \frac{z}{2} \left(\frac{z}{3}\right)^{r_{n-1}}. \end{split}$$

This gives an expression for the bivariate probability generating function for r_n and B_n and a recurrence for the coefficients:

$$\mathsf{E}_{n}t^{B_{n}}z^{r_{n}} = \mathsf{E}_{n}t^{B_{n-1}}\left\{t\frac{z+1}{2}\left(\frac{z+2}{3}\right)^{r_{n-1}} + (1-t)\frac{z}{2}\left(\frac{z}{3}\right)^{r_{n-1}}\right\}$$

$$= \frac{1}{2n}\left\{t(z+1)\mathsf{E}_{n-1}t^{B_{n-1}}(z+2)^{r_{n-1}} + (1-t)z\mathsf{E}_{n-1}t^{B_{n-1}}z^{r_{n-1}}\right\}$$

$$= \frac{1}{2n2(n-1)\cdots 2(n-m+1)}\left\{\sum_{\ell=0}^{m}b_{m,\ell}\mathsf{E}_{n-m}t^{B_{n-m}}(z+2\ell)^{r_{n-m}}\right\}.$$

Now,

$$\mathsf{E}_{n-m} t^{B_{n-m}} (z+2\ell)^{r_{n-m}} = \frac{1}{2(n-m)} \Big\{ t(z+2\ell+1) \mathsf{E}_{n-m-1} t^{B_{n-m-1}} (z+2\ell+1)^{r_{n-m-1}} + (1-t)(z+2\ell) \mathsf{E}_{n-m-1} t^{B_{n-m-1}} (z+2\ell)^{r_{n-m-1}} \Big\},$$

so that

$$b_{m,\ell}\mathsf{E}_{n-m}t^{B_{n-m}}(z+2\ell)^{r_{n-m}} = t(z+2\ell+1)b_{m,\ell}\mathsf{E}_{n-m-1}t^{B_{n-m-1}}(z+2(\ell+1))^{r_{n-m-1}} + (1-t)(z+2\ell)b_{m,\ell}E_{n-m-1}t^{B_{n-m-1}}(z+2\ell)^{r_{n-m-1}},$$

which means that the coefficient in front of

$$\mathsf{E}_{n-(m+1)}t^{B_{n-(m+1)}}(z+2\ell)^{r_{n-(m+1)}}$$

is

$$(1-t)(z+2\ell)b_{m,\ell} + t(z+2(\ell-1)+1)b_{m,\ell-1}.$$

So, if we write $b_{m,\ell} = a_{m,\ell} t^\ell (1-t)^{m-\ell}$ we see that the coefficients $a_{m,\ell} = a_{m,\ell}(z)$ satisfy the recurrence

$$a_{m+1,\ell} = (z+2\ell)a_{m,\ell} + (z+2\ell-1)a_{m,\ell-1},$$

with the initial condition

$$a_{0,0} = 1$$
, and $a_{m,\ell} = 0$, for $\ell > m$.

This is exactly the same recurrence as for the sequence $(c_{n,k})$ defined in the last section and thus

(23)
$$\mathsf{E}_{n}t^{B_{n}}z^{r_{n}} = \frac{1}{2^{n}n!}\sum_{k=0}^{n}c_{n,k}t^{k}(1-t)^{n-k}.$$

In particular, putting z = 1, differentiating with respect to t, evaluating at t = 1, and using $c_{n,n} = 2^n n!$ and $c_{n,n-1} = 2^{n-1} n! n$ we confirm that

$$\mathsf{E}_{n}B_{n} = \frac{1}{2^{n}n!}(-c_{n,n-1} + nc_{n,n}) = \frac{1}{2^{n}n!}(-2^{n-1}n!n + n2^{n}n!) = \frac{n}{2}.$$

Finally to see that B_n and A_n have, in fact, the same distribution, put z = 1 in (23) and expand $(1-t)^{n-k}$ to get

$$\sum_{k=0}^{n} c_{n,k} t^{k} (1-t)^{n-k} = \sum_{k=0}^{n} c_{n,k} t^{k} \sum_{j=0}^{n-k} \binom{n-k}{j} (-1)^{n-k-j} t^{n-k-j}$$
$$= \sum_{k=0}^{n} \sum_{j=0}^{n-k} t^{n-j} c_{n,k} \binom{n-k}{j} (-1)^{n-k-j}$$
$$= \sum_{m=0}^{n} t^{m} \left\{ \sum_{k=0}^{m} c_{n,k} \binom{n-k}{n-m} (-1)^{m-k} \right\}$$
$$= \sum_{m=0}^{n} t^{m} \left\{ \sum_{k=0}^{m} 2^{k} k! W(n,k) \binom{n-k}{n-m} (-1)^{m-k} \right\}$$

Now recall that by symmetry

$$V(n,m) = V(n,n-m) = \sum_{k=0}^{n-(n-m)} 2^k k! W(n,k) {\binom{n-k}{n-m}} (-1)^{n-k-(n-m)}$$
$$= \sum_{k=0}^m 2^k k! W(n,k) {\binom{n-k}{n-m}} (-1)^{m-k},$$

so that

$$\mathsf{E}_{n}t^{B_{n}} = \frac{1}{2^{n}n!}\sum_{k=0}^{n}c_{n,k}t^{k}(1-t)^{n-k} = \frac{1}{2^{n}n!}\sum_{m=0}^{n}V(n,m)t^{m},$$

which is exactly the same as the generating function of A_n as claimed.

References

- E. A. Bender. Central and local limit theorems applied to asymptotic enumeration. J. Combin. Theory Ser. A, 15:91–111, 1973.
- [2] M. Benoumhani. On Whitney numbers of Dowling lattices. Discrete Math., 159(1-3):13-33, 1996.
- [3] M. Benoumhani. On some numbers related to Whitney numbers of Dowling lattices. Adv. in Appl. Math., 19(1):106-116, 1997.
- M. Benoumhani. Log-concavity of Whitney numbers of Dowling lattices. Adv. in Appl. Math., 22(2):186–189, 1999.
- [5] L. Clark. Limit theorems for associated Whitney numbers of Dowling lattices. J. Combin. Math. Combin. Comput., 50:105–113, 2004.
- [6] S. Corteel and P. Hitczenko. Expected values of statistics on permutation tableaux. In 2007 Conference on Analysis of Algorithms, AofA 07, Discrete Math. Theor. Comput. Sci. Proc., AH, pages 325–339. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2007.
- [7] S. Corteel, D. Stanton, and R. Stanley L. K. Williams. Formulae for Askey–Wilson moments and enumeration of staircase tableaux. arXiv:1007.5174, to appear in *Trans. Amer. Math. Soc.*
- [8] S. Corteel, D. Stanton, and L. Williams. Enumeration of staircase tableaux, 2007. Preprint.
- [9] S. Corteel and L. K. Williams. A Markov chain on permutations which projects to the PASEP. Int. Math. Res. Notes, 17, 2007. Art. ID rnm055, 27pp.

- [10] S. Corteel and L. K. Williams. Tableaux combinatorics for the asymmetric exclusion process. Adv. Appl. Math., 39:293–310, 2007.
- [11] S. Corteel and L. K. Williams. Staircase tableaux, the asymmetric exclusion process, and Askey- Wilson polynomials. Proc. Natl. Acad. Sci, 107(15):6726–6730, 2010.
- [12] S. Corteel and L. K. Williams. Tableaux combinatorics for the asymmetric exclusion process and Askey–Wilson polynomials. Duke Math. J., 159:385–415, 2011.
- [13] S. Dasse-Hartaut and P. Hitczenko. Some properties of random staircase tableaux. Proceedings of the ANALCO 2011 meeting, pp. 58-66, SIAM 2011. http://www.siam.org/proceedings/analco/2011/analco11.php.
- [14] B. Derrida, E. Domany, and D. Mukamel. An exact solution of a one dimensional asymmetric exclusion process with open boundaries. J. Statist. Phys., 69(3-4):667–687, 1992.
- [15] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier. Exact solution of a 1D asymmetric exclusion model using a matrix formulation. J. Phys. A, 26(7):1493–1517, 1993.
- [16] T. A. Dowling. A class of geometric lattices based on finite groups. J. Combinatorial Theory Ser. B, 14:61–86, 1973. (Erratum: J. Combinatorial Theory Ser. B, 15:211, 1973).
- [17] E. Duchi and G. Schaeffer. A combinatorial approach to jumping particles. J. Combin. Theory Ser. A, 110(1):1–29, 2005.
- [18] P. Flajolet and R. Sedgewick. Analytic combinatorics. Cambridge University Press, Cambridge, 2009.
- [19] G. R. Franssens. On a number pyramid related to the binomial, Deleham, Eulerian, MacMahon and Stirling number triangles. J. Integer Seq., 9(4):Article 06.4.1, 34 pp. (electronic), 2006.
- [20] P. Hitczenko and S. Janson. Asymptotic normality of statistics on permutation tableaux. Contemporary Math., 520:83–104, 2010.
- [21] P.A. MacMahon. The divisors of numbers. Proc. London Math. Soc., 19:305–340, 1920.
- [22] A. N. Shiryaev. Probability, volume 95 of Graduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1996.
- [23] N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences, 2006. www.research.att.com/~njas/sequences/.