Séminaire Lotharingien de Combinatoire **82B** (2019)
Article #55, 12 pp.

Proceedings of the 31st Conference on Formal Power Series and Algebraic Combinatorics (Ljubljana)

Modifying Curtiss' theorem to prove central limit theorems

Gene B. Kim*[1] and Sangchul Lee†[2]

[1] *Department of Mathematics, University of Southern California, Los Angeles, CA*
[2] *Department of Mathematics, University of California – Los Angeles, Los Angeles, CA*

Abstract. The distribution of descents in fixed conjugacy classes of S_n has been studied, and it is shown that its moments have interesting properties. Kim and Lee showed, by using Curtiss' theorem and moment generating functions, how to prove a central limit theorem for descents in arbitrary conjugacy classes of S_n. In this paper, we prove a modified version of Curtiss' theorem to shift the interval of convergence in a more convenient fashion and use this to show that the joint distribution of descents and major indices in conjugacy classes is asymptotically bivariate normal.

Keywords: central limit theorem, descent, major index, generating function

1 Introduction

The theory of descents in permutations has been studied thoroughly and is related to many questions. In [16], Knuth connected descents with the theory of sorting and the theory of runs in permutations, and in [8], Diaconis, McGrath, and Pitman studied a model of card shuffling in which descents play a central role. Bayer and Diaconis also used descents and rising sequences to give a simple expression for the chance of any arrangement after any number of shuffles and used this to give sharp bounds on the approach to randomness in [1]. Garsia and Gessel found a generating function for the joint distribution of descents, major index, and inversions in [12], and Gessel and Reutenauer showed that the number of permutations with given cycle structure and descent set is equal to the scalar product of two special characters of the symmetric group in [13]. Diaconis and Graham also explained Peirce's dyslexic principle using descents in [7]. Petersen also has an excellent and very thorough book on Eulerian numbers [17].

Definition 1.1. A permutation $\pi \in S_n$ has a *descent* at position i if $\pi(i) > \pi(i+1)$, where $i = 1, \ldots, n-1$. The *descent number* of π, denoted $d(\pi)$, is defined as the number of all descents of π plus 1. The *major index* of π, denoted $maj(\pi)$, is the sum of the positions at which π has a descent.

*genebkim@usc.edu
†sos440@math.ucla.edu

It is well known ([2] [9]) that the distribution of $d(\pi)$ in S_n is asymptotically normal with mean $\frac{n+1}{2}$ and variance $\frac{n+1}{12}$. Fulman also used Stein's method to show that the number of descents of a random permutation satisfies a central limit theorem with error rate $n^{-1/2}$ in [11]. In [18], Vatutin proved a central limit theorem for $d(\pi) + d(\pi^{-1})$, where π is a random permutation.

Fulman [10] proved that the distribution of descents in conjugacy classes with large cycles is asymptotically normal, and Kim [14] proved that descents in fixed point free involutions (matchings) is also asymptotically normal. After the latter result was proved, Diaconis [6] conjectured that there are asymptotic normality results for descents in conjugacy classes that are fixed point free. Kim and Lee, in [15], proved a more general result that descents in arbitrary conjugacy classes are asymptotically normal, where the parameters depend only on the ratio of fixed points to n.

There have been central limit theorems proven about major indices as well. In [4], Chen and Wang also use generating functions and Curtiss' theorem to prove asymptotic normality results about major indices of derangements. Billey et al, in [3], consider the distribution of major indices on standard tableaux of arbitrary straight shape and certain skew shapes.

In this paper, we show that the joint distribution of descents and major indices in any conjugacy classes of S_n is asymptotically bivariate normal. The precise formulation of the main statement is as follows.

Theorem 1.2. *For each conjugacy class $\mathcal{C}_\lambda$ of S_n, write α_1 for the density of fixed points of any permutations in $\mathcal{C}_\lambda$ and define*

$$W_\lambda = \left(\frac{d(\pi) - \frac{1-\alpha_1^2}{2} n}{n^{1/2}}, \frac{maj(\pi) - \frac{1-\alpha_1^2}{4} n^2}{n^{3/2}} \right),$$

where π is chosen uniformly at random from $\mathcal{C}_\lambda$. Then, along any sequence of $\mathcal{C}_\lambda$'s such that $n \to \infty$ and $\alpha_1 \to \alpha \in [0,1]$, W_λ converges in distribution to a bivariate normal distribution of zero mean and the covariance matrix Σ_α depending only on α.

We will need two major ingredients for this; one is a modification of Curtiss' theorem relating pointwise convergence of moment generating function (m.g.f.) to the convergence in distribution of corresponding random variables, and the other is a uniform estimate on the m.g.f. of the joint distribution of descents and major index. The asymptotic joint normality will then follow as an immediate corollary. This uniform estimate will be strong enough to prove an analogous result for a more general class of subsets of of S_n, encompassing the asymptotical joint normality for derangements.

Theorem 1.3. *Suppose that A_n is a subset of S_n which is invariant under conjugation and that all $\pi \in A_n$ have the same number of fixed points. Denote by $\alpha_{1,n}$ the common density of fixed* *points of elements in A_n, and define*

$$W_n = \left(\frac{d(\pi) - \frac{1-\alpha_{1,n}^2}{2} n}{n^{1/2}}, \frac{maj(\pi) - \frac{1-\alpha_{1,n}^2}{4} n^2}{n^{3/2}} \right),$$

where π is chosen uniformly at random from A_n. If $\alpha_{1,n} \to \alpha \in [0,1]$ as $n \to \infty$, then W_n converges in distribution to a bivariate normal distribution of zero mean and the covariance matrix Σ_α.

This paper is organized as follows: In Section 2, we modify Curtiss' theorem in the form which is applicable to our proof. In Section 3, we establish a formula for the joint generating function of $(d(\pi), maj(\pi))$ for π in the conjugacy class $\mathcal{C}_\lambda$. In Section 4, we analyze this formula analytically to provide a uniform estimate on the m.g.f. M_{W_λ} and then apply the modified Curtiss' theorem to conclude both main theorems.

2 A modification of Curtiss' theorem

For a random variable X taking values in $\mathbb{R}^d$, its moment generating function (m.g.f.) is defined as

$$M_X(s) = \mathbf{E}\left[e^{\langle s,X\rangle}\right], \qquad s \in \mathbb{R}^d.$$

In his paper [5], Curtiss showed a version of continuity theorem that, if $\{X_n\}_{n\in\mathbb{N}}$ is a sequence of random variables in $\mathbb{R}$ such that

(C) $M_{X_n}(s)$ converges pointwise on a neighborhood of $s = 0$,

then $\{X_n\}_{n\in\mathbb{N}}$ converge in distribution. This result has an advantage over Lévy's continuity theorem in that the pointwise limit of M_{X_n}'s is guaranteed to be a m.g.f. Such a stronger conclusion is possible because **(C)** guarantees the tightness of $\{X_n\}_{n\in\mathbb{N}}$.

In some applications, however, **(C)** is quite costly to verify and requires extra inputs, while the stronger part of its conclusion - that the limit is always a m.g.f. of some distribution - is not essential. For instance, in [14] and [15], the m.g.f.s of normalized descents in conjugacy classes are analyzed with their series expansions, which fail to converge for $s > 0$. In [14], Kim circumvented this technical difficulty by establishing a bijection to show the convergence for $s > 0$. In [15], Kim and Lee calculated an alternative form of the m.g.f. that is convergent for $s > 0$ by expanding the original generating function in Laurent series at ∞ rather than at 0. Moreover, in both proofs, the limit is shown explicitly to be the m.g.f. of the normal distribution. If we were to take a similar approach, we would have to show that M_{W_λ}, the m.g.f.s of the normalized descent/major-index pairs

W_λ, converge pointwise on an open set containing $(0,0)$. However, as the known series expansion of M_{W_λ} is convergent only on $\{(s,r)|s \leq 0, r \leq 0\}$, we would need to use similar methods used in [15] in order to come up with possibly several expressions for M_{W_λ} that are convergent on different regions, whose union covers an open set containing $(0,0)$.

In this section, we provide a simple result that takes cares of this situation.

Proposition 2.1. Let X_n be random vectors in $\mathbb{R}^d$ for each $n \in \mathbb{N} \cup \{\infty\}$. Suppose that there is a non-empty open subset $U \subseteq \mathbb{R}^d$ such that $\lim_{n\to\infty} M_{X_n}(s) = M_{X_\infty}(s)$ for all $s \in U$. Then, X_n converges in distribution to X_∞.

Before we prove the proposition, we introduce the following lemma:

Lemma 2.2. *Let X_n be random vectors in $\mathbb{R}^d$ for each $n \in \mathbb{N} \cup \{\infty\}$. Then the followings are equivalent:*

1. *X_n converges in distribution to X_∞.*
2. *X_n converges vaguely to X_∞, i.e., $\lim_{n\to\infty} \mathbf{E}[\varphi(X_n)] = \mathbf{E}[\varphi(X_\infty)]$ for all $\varphi \in C_c(\mathbb{R}^d)$, where $C_c(\mathbb{R}^d)$ denotes the set of all continuous compactly supported functions $\varphi : \mathbb{R}^d \to \mathbb{R}$.*

Proof of Proposition 2.1. Fix $a \in U$ and introduce random vectors Y_n in $\mathbb{R}^d$ whose laws are given by the exponential tilting

$$\mathbf{P}[Y_n \in \mathrm{d}x] = \frac{e^{\langle a,x\rangle}}{M_{X_n}(a)} \mathbf{P}[X_n \in \mathrm{d}x].$$

For each $t \in \mathbb{R}^d$, there exists $\delta > 0$ such that $a + st \in U$ for all $s \in (-\delta, \delta)$. Then, for all $s \in (-\delta, \delta)$, we have

$$\lim_{n\to\infty} M_{\langle t,Y_n\rangle}(s) = \lim_{n\to\infty} \frac{M_{X_n}(a+st)}{M_{X_n}(a)} = \frac{M_{X_\infty}(a+st)}{M_{X_\infty}(a)} = M_{\langle t,Y_\infty\rangle}(s),$$

and so, $\langle t, Y_n\rangle$ converges to $\langle t, Y_\infty\rangle$ in distribution by Curtiss' continuity theorem. By Cramer-Wold device, this implies that Y_n converges in distribution to Y_∞. Then, for each $\varphi \in C_c(\mathbb{R}^d)$, the function $\varphi(\cdot)e^{\langle a,\cdot\rangle}$ is bounded, and so,

$$M_{X_n}(a)\mathbf{E}\left[\varphi(Y_n)e^{\langle a,Y_n\rangle}\right] \xrightarrow[n\to\infty]{} M_{X_\infty}(a)\mathbf{E}\left[\varphi(Y_\infty)e^{\langle a,Y_\infty\rangle}\right].$$

In other words, $\mathbf{E}[\varphi(X_n)] \to \mathbf{E}[\varphi(X_\infty)]$, and so, by Lemma 2.2, X_n converges to X_∞ in distribution. □

3 Generating function of $(d(\pi), maj(\pi))$

For each finite set A, we write $|A|$ for the cardinality of A. For each integers $a \leq b$, the double-struck interval notation $[\![a, b]\!] = [a, b] \cap \mathbb{Z}$ denotes the set of all integers between a and b.

Throughout this article, λ will always denote an integer partition of a non-negative integer n. For each permutation $\pi \in S_n$, we write $m_k(\pi)$ for the number of k-cycles in π. Then, $\mathcal{C}_\lambda = \{\pi \in S_n : m_k(\pi) = \lambda_k \text{ for all } k \in [\![1, n]\!]\}$ denotes the conjugacy class of S_n with the cycle structure λ. Associated to each λ is the density $\alpha_1 = \alpha_1(\lambda) = \lambda_1/n$ of fixed points. We will see that α_1 is essentially the only parameter that determines the shape of the limiting distribution of the descent/major index pair.

We will set up some probability notations. For each integer partition λ, $\mathbf{P}_\lambda$ will denote the probability law under which π has the uniform distribution over $\mathcal{C}_\lambda$ and σ_i has the uniform distribution over S_{λ_i} for each $i \in [\![1, n]\!]$. We will always assume, without mentioning, that π and σ_i's have their respective, aforementioned distributions under $\mathbf{P}_\lambda$. Then, $\mathbf{E}_\lambda$ will denote the expectation corresponding to $\mathbf{P}_\lambda$.

It is convenient to consider some special functions. Let $\Gamma(s)$ denote the gamma function. By setting $x! = \Gamma(x+1)$, the factorial extends to all of $\mathbb{R} \setminus \{-1, -2, \cdots\}$. Then, we extend binomial coefficients accordingly. We also introduce the q-bracket notation $[a]_q = \frac{1-q^a}{1-q}$.

In [10, Theorem 1], Fulman derived a formula for the generating function of descent numbers in conjugacy classes. This formula was a key ingredient in [14] and [15] for establishing the central limit theorems. In this section, we would like to derive an analogous formula for the generating function of pairs of descent/major-index in conjugacy classes, which is the statement of the following proposition.

Proposition 3.1. Let $f_{i,a}(q) = \frac{1}{i} \sum_{d|i} \mu(d) [a]_{q^d}^{i/d}$. If σ_i has the uniform distribution over S_{λ_i} under $\mathbf{P}_\lambda$ for each $i \in [\![1, n]\!]$, then,

$$\frac{\sum_{\pi \in \mathcal{C}_\lambda} t^{d(\pi)} q^{maj(\pi)}}{(1-t)(1-qt)\cdots(1-q^n t)} = \sum_{a \geq 1} t^a \left(\prod_{i=1}^{n} \mathbf{E}_\lambda \left[\prod_{k \geq 1} f_{i,a}(q^k)^{m_k(\sigma_i)} \right] \right). \tag{3.1}$$

Proof. Recall that for non-negative integers $r_1, \cdots, r_n$ summing to n, the quantity

$$M(r_1, \cdots, r_a) = \frac{1}{n} \sum_{d | r_1, \cdots, r_a} \mu(d) \frac{(n/d)!}{(r_1/d)! \cdots (r_a/d)!}$$

counts the number of primitive circular words of length n from the alphabet $[\![1, a]\!]$ in which the letter i appears r_i times. Associated to this quantity, we define $J_{i,m,a}$ by

$$J_{i,m,a} = \sum_{\substack{r_1 + \cdots + r_a = i \\ 1r_1 + \cdots + ar_a = i+m}} M(r_1, \cdots, r_a).$$

Then, one of the main results in [10] is the generating function

$$\sum_{n\geq 0}\frac{y^n\sum_{\pi\in S_n}t^{d(\pi)}q^{maj(\pi)}\prod_i x_i^{m_i(\pi)}}{(1-t)(1-qt)\cdots(1-q^nt)}=\sum_{a\geq 1}t^a\prod_{\substack{i\geq 1\\ m\geq 0}}\left(\frac{1}{1-q^m x_i y^i}\right)^{J_{i,m,a}}. \tag{3.2}$$

Our strategy is to expand the huge product on the right-hand side. In the course of computations, the following lemmas will be useful.

Lemma 3.2. *For $q\in[0,1]$, we have $f_{i,a}(q)=\sum_{m\geq 0}J_{i,m,a}q^m$.*

Lemma 3.3. *Suppose that σ is uniformly distributed over S_n. Then,*

$$\mathbf{E}\left[\prod_k x_k^{m_k(\sigma)}\right]=\sum_{r\vdash n}\left[\prod_k\frac{1}{r_k!}\left(\frac{x_k}{k}\right)^{r_k}\right]. \tag{3.3}$$

Returning to the proof of Equation (3.1), we find that the coefficient of the right-hand side of (3.2) can be rearranged, by using the series expansion $-\log(1-x)=\sum_{k=1}^{\infty}\frac{1}{k}x^k$ and Lemma 3.2, as

$$\prod_{\substack{i\geq 1\\ m\geq 0}}\left(\frac{1}{1-q^m x_i y^i}\right)^{J_{i,m,a}}=\exp\left\{\sum_{i\geq 1}\sum_{k\geq 1}\frac{1}{k}f_{i,a}(q^k)x_i^k y^{ik}\right\}. \tag{$\diamond$}$$

For each given i and k, we apply the Taylor series $e^x=\sum_{s=0}^{\infty}\frac{x^s}{s!}$ to expand the factor $\exp\{\frac{1}{k}f_{i,a}(q^k)x_i^k y^{ik}\}$. Since the generic variable s needs to be distinguished for different choices of i and k, we explicate this dependence by writing the generic variables as $s_{i,k}$. The resulting expansion takes the form

$$(\diamond)=\sum_{(s_{i,k})_{i,k\in\mathbb{N}}}\prod_{i,k\geq 1}\frac{1}{s_{i,k}!}\left(\frac{f_{i,a}(q^k)x_i^k y^{ik}}{k}\right)^{s_{i,k}}.$$

Now, for each $\lambda\vdash n$, we collect all terms satisfying $\sum_k ks_{i,k}=\lambda_i$ for each $i\in[\![1,n]\!]$. Then, by Lemma 3.3, $(\diamond)$ simplifies to

$$\begin{aligned}(\diamond)&=\sum_{n\geq 0}\sum_{\lambda\vdash n}\sum_{\substack{(s_{i,k})_{i,k\in\mathbb{N}}\\ \sum_k ks_{i,k}=\lambda_i}}\left[\prod_{i,k\geq 1}\frac{1}{s_{i,k}!}\left(\frac{f_{i,a}(q^k)}{k}\right)^{s_{i,k}}\right]y^n\prod_{i\geq 1}x_i^{\lambda_i}\\ &\overset{(3.3)}{=}\sum_{n\geq 0}\sum_{\lambda\vdash n}\left(\prod_{i=1}^{n}\mathbf{E}_\lambda\left[\prod_{k\geq 1}f_{i,a}(q^k)^{m_k(\sigma_i)}\right]\right)y^n\prod_{i\geq 1}x_i^{\lambda_i}.\end{aligned}$$

Plugging this back into the generating function proves (3.1) as required. □

As a sanity check, recall that $\mathbf{E}\left[x^{\sum_i m_i(\sigma)}\right] = \binom{n+x-1}{n}$ holds if σ is uniformly distributed over S_n. Taking the limit as $q \uparrow 1$ to the key identity (3.1), we obtain

$$\frac{\sum_{\pi \in \mathcal{C}_\lambda} t^{d(\pi)}}{(1-t)^{n+1}} = \sum_{a \geq 1} t^a \prod_{i \geq 1} \binom{\lambda_i + f_{i,a}(1) - 1}{\lambda_i},$$

which is exactly the conclusion of [10, Corollary 3].

4 Main result

Let π be chosen uniformly at random from $\mathcal{C}_\lambda$. In order to establish the asymptotic normality of $(d(\pi), maj(\pi))$, we consider the following normalization

$$W_\lambda = \left(\frac{d(\pi) - \frac{1-\alpha_1^2}{2} n}{n^{1/2}}, \frac{maj(\pi) - \frac{1-\alpha_1^2}{4} n^2}{n^{3/2}} \right).$$

We aim to prove that W_λ is asymptotically normal with mean zero and covariance matrix Σ_{α_1}, where Σ_α is defined by the following 2 by 2 matrix

$$\Sigma_\alpha = \begin{pmatrix} \frac{1}{12}(1 - 4\alpha^3 + 3\alpha^4) & \frac{1}{24}(1 - 4\alpha^3 + 3\alpha^4) \\ \frac{1}{24}(1 - 4\alpha^3 + 3\alpha^4) & \frac{1}{36}(1 - \alpha^3) \end{pmatrix}.$$

Since any real symmetric matrix determines a quadratic form and vice versa, we will abuse the notation to write $\Sigma_\alpha(x) = x^\mathsf{T} \Sigma_\alpha x$ for any $x \in \mathbb{R}^2$. The goal of this section is to establish the proof of the following uniform estimate.

Theorem 4.1. *For each $s > 0$ and $r > 0$, there exists a constant $C = C(r,s) > 0$, depending only on s and r, such that*

$$\left| M_{W_\lambda}(-s,-r) - e^{\frac{1}{2}\Sigma_{\alpha_1}(s,r)} \right| \leq C(r,s) n^{-1/6} \tag{4.1}$$

holds for any $n \geq 1$ and for any conjugacy class $\mathcal{C}_\lambda$ of S_n.

4.1 Notations and conventions

For the remainder of this paper, we fix two positive reals $s, r > 0$. It will become clear that the window of scale $r/n^{3/2}$ is a natural choice for analyzing the behavior of W_λ. For brevity's sake, we write

$$\delta = r/n^{3/2}.$$

Comparing the m.g.f. of W_λ to the generating function (3.1) shows that q and t are related to n by $q = e^{-r/n^{3/2}} = e^{-\delta}$ and $t = e^{-s/n^{1/2}} = e^{-\frac{sn}{r}\delta}$, and we assume so hereafter. We also choose $\epsilon > 0$ such that $2e(s+r)\epsilon/r < 1$.

In what follows, the asymptotic notations $f(x) = \mathcal{O}_a(g(x))$ and $f(x) \lesssim_a g(x)$ will denote the fact that there exists a constant $C > 0$, depending only on s, r and the parameter a, such that $|f(x)| \leq Cg(x)$ holds for all x in the prescribed range. If no parameter a is involved, we simply write $f(x) = \mathcal{O}(g(x))$ or $f(x) \lesssim g(x)$.

Along the proof, we will encounter large chunks of expressions. Since it is counter-productive and not aesthetic to carry them all the way through the proof, we will introduce generic symbols to replace them. First, we define $\mathsf{K}_{\lambda,r,a,i}$ and $\mathsf{L}_{\lambda,s,r}$ by

$$\mathsf{K}_{\lambda,r,a,i} := \lambda_i! i^{\lambda_i} \mathbf{E}_\lambda \left[\prod_{k\geq 1} f_{i,a}(q^k)^{m_k(\sigma_i)} \right],$$

and

$$\mathsf{L}_{\lambda,s,r} := \frac{t^{\frac{\alpha_1^2}{2}n} q^{\frac{\alpha_1^2}{4}n^2}}{\frac{1}{\delta^{n+1}} \int_0^1 u^{\frac{sn}{r}-1}(1-u)^n \, \mathrm{d}u} \left(\sum_{a\geq 1} t^a \left(\prod_{i=1}^n \mathsf{K}_{\lambda,r,a,i} \right) \right).$$

For the definition of $\mathsf{K}_{\lambda,r,a,i}$, we recall that σ_i has the uniform distribution over S_{λ_i} under $\mathbf{P}_\lambda$ for each $i \in [\![1,n]\!]$. Also, for convenience, we decompose L further into

$$\mathsf{L}_{\lambda,s,r} = \mathsf{L}_{\mathrm{small},\lambda,s,r} + \mathsf{L}_{\mathrm{large},\lambda,s,r},$$

where $\mathsf{L}_{\mathrm{small},\lambda,s,r}$ (respectively, $\mathsf{L}_{\mathrm{large},\lambda,s,r}$) is the restriction of the sum in the definition of L onto the range $a < \epsilon/\delta$ (respectively, $a \geq \epsilon/\delta$). Then, we define $\mathsf{F}_{n,r,k}(u)$ and $\mathsf{G}_{\lambda,s,r}(u)$ by

$$\mathsf{F}_{n,r,k}(u) := \frac{\delta^{k-1}}{k^2} \frac{1-u^k}{(1-u)^k},$$

and

$$\mathsf{G}_{\lambda,s,r}(u) := t^{\frac{\alpha_1^2}{2}n} q^{\frac{\alpha_1^2}{4}n^2} \lambda_1! \sum_{\mu \vdash \lambda_1} \prod_{k\geq 1} \frac{\mathsf{F}_{n,r,k}(u)^{\mu_k}}{\mu_k!}.$$

The roles of these quantities will become clear as the proof proceeds. Finally, for these quantities, s, r and λ will be suppressed notationally whenever the dependence on these variables is clear from context.

4.2 Separating the contribution of fixed points

In this section, we provide a representation of the m.g.f. of W_λ which is adequate for analyzing the effect of fixed points. Rewrite M_{W_λ} as

$$M_{W_\lambda}(-s,-r) = t^{-\frac{1-\alpha_1^2}{2}n} q^{-\frac{1-\alpha_1^2}{4}n^2} \mathbf{E}_\lambda \left[t^{d(\pi)} q^{maj(\pi)} \right], \tag{4.2}$$

where we recall that π is uniformly distributed over $\mathcal{C}_\lambda$ under $\mathbf{P}_\lambda$ and that q and t are defined by $q = e^{-r/n^{3/2}}$ and $t = e^{-s/n^{1/2}}$. To prevent the reader from being distracted by the jumble of computations, we first state the main result of this subsection. This is a combination of Propositions 4.3, 4.4 and 4.6.

Theorem 4.2. *As $n \to \infty$, we have*

$$M_{W_\lambda}(-s,-r) = \left(1 + \mathcal{O}(n^{-1/2})\right) e^{\frac{1}{2}\Sigma_0(s,r)} \left(\mathsf{L}_{\text{large}} + \mathsf{L}_{\text{small}}\right),$$

where $\mathsf{L}_{\text{small}}$ decays at least exponentially fast and $\mathsf{L}_{\text{large}}$ is asymptotically the ratio of two integrals

$$\mathsf{L}_{\text{large}} = \left(1 + \mathcal{O}(n^{-1/2})\right) \frac{\int_0^{e^{-\epsilon}} u^{\frac{sn}{r}-1}(1-u)^n \mathsf{G}(u)\,\mathrm{d}u}{\int_0^1 u^{\frac{sn}{r}-1}(1-u)^n\,\mathrm{d}u}.$$

Returning to the problem of estimating (4.2), we will adopt (3.1) as our starting point. Since $|t| < 1$ and $|q| < 1$, the generating function (3.1) converges absolutely. We first analyze the asymptotic behavior of the common factor $\prod_{j=0}^n (1 - tq^j)$.

Proposition 4.3. As $n \to \infty$, we have

$$\frac{1}{n!} \prod_{j=0}^{n} (1 - tq^j) = \left(1 + \mathcal{O}(n^{-1/2})\right) \frac{t^{\frac{1}{2}n} q^{\frac{1}{4}n^2} e^{\frac{1}{2}\Sigma_0(s,r)}}{\frac{1}{\delta^{n+1}} \int_0^1 u^{\frac{sn}{r}-1}(1-u)^n\,\mathrm{d}u}. \tag{4.3}$$

Consequently,

$$M_{W_\lambda}(-s,-r) = \left(1 + \mathcal{O}(n^{-1/2})\right) e^{\frac{1}{2}\Sigma_0(s,r)} \mathsf{L}. \tag{4.4}$$

Proposition 4.3 illuminates the meaning of L as the normalized m.g.f.

Proposition 4.4. As $n \to \infty$, we have

$$\mathsf{L}_{\text{small}} \lesssim n^{1/2} \left(2e(s+r)\epsilon/r\right)^{n+1}. \tag{4.5}$$

Our next goal is to estimate $\mathsf{L}_{\text{large}}$. From now on, we focus on the case $a \geq \epsilon/\delta$. A trivial but crucial observation is that $0 \leq q^a \leq q^{\epsilon/\delta} = e^{-\epsilon}$ holds, hence q^a is uniformly away from 1. This fact will be extensively used in the sequel, and is indeed one of the main reasons for separating the small range, $a \leq \epsilon/\delta$, from our main computation. We begin by improving the estimate on $\mathsf{K}_{i,a}$'s.

Lemma 4.5. *For any integer $a \geq \epsilon/\delta$ and for any real x satisfying $x \geq \epsilon/\delta$ and $|x-a| \leq 1$, we have*

$$\prod_{i\geq 1} \mathsf{K}_{a,i} = \frac{1+\mathcal{O}(n^{-1/2})}{\delta^n} (1-q^x)^n \lambda_1! \mathbf{E}_\lambda \left[\prod_{k\geq 1} (k\mathsf{F}_k(q^x))^{m_k(\sigma_1)} \right] \tag{4.6}$$

This lemma already hints that the contribution of fixed points is the only factor that affects the asymptotic distribution of the normalized pair W_λ. This point is more clearly seen from the following result, which is also last piece of the main claim of this section.

Proposition 4.6. We have

$$\mathsf{L}_{\text{large}} = \left(1+\mathcal{O}(n^{-1/2})\right) \frac{\int_0^{e^{-\epsilon}} u^{\frac{sn}{r}-1}(1-u)^n \mathsf{G}(u)\,\mathrm{d}u}{\int_0^1 u^{\frac{sn}{r}-1}(1-u)^n\,\mathrm{d}u}. \tag{4.7}$$

This result tells us that G accounts for the perturbation caused by the presence of fixed points. Before proceeding to the general case, we rejoice this result by looking into the case where $\mathcal{C}_\lambda$ is free of fixed points, i.e. $\lambda_1 = 0$. In such case, we identically have $\mathsf{G}(u) = 1$, and hence quickly establish the following partial result.

Corollary 4.7. *For each $s > 0$ and $r > 0$, there exists a constant C, depending only on s and r, such that*

$$\left| M_{W_\lambda}(-s,-r) - e^{\frac{1}{2}\Sigma_0(s,r)} \right| \leq Cn^{-1/2}$$

holds for any $n \geq 1$ and for any fixed-point-free conjugacy class $\mathcal{C}_\lambda$ of S_n. Consequently, along any sequence of $\mathcal{C}_\lambda$'s such that $n \to \infty$ and $\alpha_1 = 0$, W_λ converges in distribution to a bivariate normal distribution of zero mean and the covariance matrix Σ_0.

Proof of Corollary 4.7. Combining all the estimates and setting $\lambda_1 = 0$, we have

$$M(-s,-r) = \mathcal{O}(n^{-1/2}) + e^{\frac{1}{2}\Sigma_0(s,r)} \left(1 - \frac{\int_{e^{-\epsilon}}^1 u^{\frac{sn}{r}-1}(1-u)^n\,\mathrm{d}u}{\int_0^1 u^{\frac{sn}{r}-1}(1-u)^n\,\mathrm{d}u} \right).$$

It is easy to check that the ratio of two integrals above vanishes exponentially fast. Therefore, the first claim follows, and the second claim is a direct application of Proposition 2.1. □

4.3 Investigating the contribution of fixed points

The aim of this subsection is to analyze the effect of the perturbation term G in Proposition 4.6 and finalize the proof of the main theorem. To do so, we prove the following estimate on $\mathsf{L}_{\text{large}}$.

Proposition 4.8. We have

$$\mathsf{L}_{\text{large}} = \left(1 + \mathcal{O}\left(n^{-1/6}\right)\right) e^{\frac{1}{2}\left(\Sigma_{\alpha_1}(s,r) - \Sigma_0(s,r)\right)}. \tag{4.8}$$

This is the last ingredient towards establishing Theorem 4.1 and consequently Theorems 1.2 and 1.3.

Proof of Theorems 1.2, 1.3 and 4.1. Combining Theorem 4.2 and Proposition 4.8, we have

$$M_{W_\lambda}(-s,-r) = \left(1 + \mathcal{O}\left(n^{-1/6}\right)\right) e^{\frac{1}{2}\Sigma_{\alpha_1}(s,r)} + \mathcal{O}(\mathsf{L}_{\text{small}}).$$

By noting that $\mathsf{L}_{\text{small}}$ decays at least exponentially fast, it can be absorbed into the error term $\mathcal{O}(n^{-1/6})$, and hence, Theorem 4.1 follows.

Since Theorem 1.2 is a special case of Theorem 1.3, it is enough to prove the latter. Fix $s, r > 0$ and let $C = C(s,r) > 0$ be as in Theorem 4.1. Let A_n and $\alpha_{1,n}$ satisfy the hypotheses of Theorem 1.3. Then, for each $n \in \mathbb{N}$, there exists a family Λ_n of integer partitions of n such that $A_n = \bigcup_{\lambda \in \Lambda_n} \mathcal{C}_\lambda$ and that $\alpha_1(\lambda) = \alpha_{1,n}$ for all $\lambda \in \Lambda_n$. Then,

$$\begin{aligned}\left| M_{W_n}(-s,-r) - e^{\frac{1}{2}\Sigma_{\alpha_{1,n}}(s,r)} \right| &\leq \sum_{\lambda \in \Lambda_n} \frac{|\mathcal{C}_\lambda|}{|A_n|} \left| M_{W_\lambda}(-s,-r) - e^{\frac{1}{2}\Sigma_{\alpha_1(\lambda)}(s,r)} \right| \\ &\leq C n^{-1/6}.\end{aligned}$$

Since $\alpha_{1,n} \to \alpha$ by the hypothesis, $M_{W_n}(-s,-r) \to e^{\frac{1}{2}\Sigma_\alpha(s,r)}$, and therefore, the conclusion follows from the modified Curtiss' theorem. □

References

[1] D. Bayer and P. Diaconis. "Trailing the dovetail shuffle to its lair". *Ann. Appl. Probab.* **2**.2 (1992), pp. 294–313. Link.

[2] E. Bender. "Central and local limit theorems applied to asymptotic enumeration". *J. Comb. Theory Ser. A* **15** (1973), pp. 91–111. Link.

[3] S. C. Billey, M. Konvalinka, and J. Swanson. "Asymptotic normality of the major index on standard tableaux". *Adv. Appl. Math.* **313** (2020), Art. 101972, 36 pp. Link.

[4] W. Y. C. Chen and D. G. L. Wang. "The limiting distribution of q-derangement numbers". *Eur. J. Combin.* **31**.8 (2010), pp. 2006–2013. Link.

[5] J. H. Curtiss. "A note on the theory of moment generating functions". *Ann. Math. Stat.* **13** (1942), pp. 430–433. Link.

[6] P. Diaconis and J. Fulman. Personal communication. 2017.

[7] P. Diaconis and R. Graham. "The magic of Charles Sanders Peirce". *The mathematics of various entertaining subjects, vol. 3*. Princeton Univ. Press, Princeton, NJ, 2019, pp. 161–203.

[8] P. Diaconis, M. McGrath, and J. Pitman. "Riffle shuffles, cycles, and descents". *Combinatorica* **15**.1 (1995), pp. 11–29. Link.

[9] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[10] J. Fulman. "The distribution of descents in fixed conjugacy classes of the symmetric group". *Journal of Combinatorial Theory Series A* **84**.2 (1998), pp. 171–180. Link.

[11] J. Fulman. "Stein's method and non-reversible Markov chains". *Stein's Method: Expository Lectures and Applications*. Beachwood, Ohio, USA: Institute of Mathematical Statistics, 2004, pp. 66–74. Link.

[12] A. M. Garsia and I. Gessel. "Permutation statistics and partitions". *Adv. in Math.* **31** (1979), pp. 288–305. Link.

[13] I. Gessel and C. Reutenauer. "Counting permutations with given cycle structure and descent set". *J. Comb. Theory Ser. A* **64**.2 (1993), pp. 189–215. Link.

[14] G. B. Kim. "Distribution of descents in matchings". *Ann. Comb.* **23** (2019), pp. 73–87. Link.

[15] G. B. Kim and S. Lee. "Central limit theorem for descents in conjugacy classes of S_n". *J. Combin. Theory Ser. A* **169** (2020), Art. 105123, 13 pp. Link.

[16] D. Knuth. *The Art of Computer Programming, Vol. 3: Sorting and Searching*. Addison-Wesley, 1973.

[17] T. K. Petersen. *Eulerian Numbers*. Birkhäuser-Springer, 2015.

[18] V. A. Vatutin. "The numbers of ascending segments in a random permutation and in the inverse to it are asymptotically independent". *Discrete Math. Appl.* **6**.1 (1996), pp. 41–52. Link.