COMBINATORICA
Akadémiai Kiadó – Springer-Verlag

COMBINATORICA **15** (1) (1995) 11–29

RIFFLE SHUFFLES, CYCLES, AND DESCENTS

PERSI DIACONIS, MICHAEL MC GRATH, JIM PITMAN

Received July 20, 1992
Revised June 10, 1993

We derive closed form expressions and limiting formulae for a variety of functions of a permutation resulting from repeated riffle shuffles. The results allow new formulae and approximations for the number of permutations in S_n with given cycle type and number of descents. The theorems are derived from a bijection discovered by Gessel. A self-contained proof of Gessel's result is given.

1. Introduction

This paper derives theorems about permutations from the properties of a non-uniform probability on the permutation group S_n. For an integer $a \geq 1$ define an a-shuffle to be the probability measure $P_{n,a}$ on S_n defined by

$$P_{n,a}(\pi) = \frac{\binom{a+n-d(\pi)-1}{n}}{a^n} \tag{1.1}$$

for $\pi \in S_n$ with $d(\pi)$ the number of descents in π. This is the distribution of a random permutation obtained by first randomly cutting a deck of n cards into a packets (empty packets allowed) and then randomly riffling the packets together. Bayer and Diaconis [3] showed that $P_{n,2^k}(\pi)$ gives the chance that the deck is in the arrangement π after k independent 2-shuffles. Our main finding is that the distribution of the cycles structure of π under $P_{n,a}$ has a simple form.

Theorem A. *For non-negative n_j with $\Sigma j n_j = n$, the $P_{n,a}$ probability that a permutation has n_j cycles of length j, $1 \leq j \leq n$, is*

$$a^{-n} \prod_{j=1}^{n} \binom{f_{ja}+n_j-1}{n_j} \tag{1.2}$$

where f_{ja} is the number of aperiodic circular words of length j from an alphabet of a letters.

Mathematics Subject Classification (1991): 05 A 15, 60 C 05

It is well known that

$$f_{ja} = \frac{1}{j} \sum_{d|j} \mu(d) a^{j/d}. \tag{1.3}$$

The sum in (1.3) is over all divisors d of j with μ the Möbius function

$$\mu(1) = 1$$

$$\mu(d) = \begin{cases} (-1)^k & \text{if } d = p_1 p_2 \cdots p_k \text{ for distinct primes } p_1, \cdots, p_k \\ 0 & \text{otherwise.} \end{cases}$$

Theorem A is derived from a bijection discovered by Gessel. This gives a 1–1 correspondence between $\{0,1,\cdots,a-1\}^n$ and the collection of multisets of aperiodic necklaces with total length n. Gessel's result has been part of the folklore of combinatorics for several years. An extensive account in the language of representation theory and symmetric functions has recently been prepared by Gessel and Reutenauer [6]. We give a self contained treatment from first principles.

Theorem A allows us to compute and approximate the chance of various events after repeated riffle shuffles. For example, the expected number of fixed points is

$$1 + \frac{1}{a} + \frac{1}{a^2} + \cdots + \frac{1}{a^{n-1}}.$$

As a tends to infinity this expectation tends to 1, which is the expected number of fixed points of a uniformly chosen permutation. Asymptotics for fixed a as $n \to \infty$ are more interesting. The behavior of the large cycles is governed by Poisson-Dirichlet asymptotics, exactly as in the uniform case. But the limiting joint distribution of the numbers of j-cycles for a random a-shuffle, as $n \to \infty$, is the distribution of independent negative binomial variables with parameters (f_{ja}, a^{-j}). Only as $a \to \infty$ does this approach the well known limiting distribution for the uniform case defined by independent Poisson (j^{-1}) variables.

Bayer and Diaconis [3] have shown that it takes $k = \frac{3}{2} \log_2 n$ 2-shuffles to have $P_{n,2}k$ close to the uniform distribution in total variation. This corresponds to $a \gg n^{3/2}$. The present results show that features depending on cycles have the correct limiting distribution (as $n \to \infty$) after fewer shuffles. The number of fixed points has a Poisson (1) distribution to good approximation for $a \gg 1$. The distribution of the large cycles has the usual Poisson-Dirichlet asymptotics for $a = 2$.

Theorem A allows us to give closed form expressions for the number of permutations with a given cycle structure and number of descents:

Theorem B. *For non-negative n_j with $\sum j n_j = n$, the number of permutations $\pi \in S_n$ with n_j cycles of length j, $1 \le j \le n$, and $k-1$ descents is*

$$\sum_{a=1}^{k} (-1)^{k-a} \binom{n+1}{k-a} \prod_{j=1}^{n} \binom{f_{ja} + n_j - 1}{n_j}$$

with f_{ja} as in (1.3).

To illustrate Theorems A and B by example: if $n_n = 1, n_j = 0$ otherwise, Theorem A shows that the probability that an a-shuffle produces an n-cycle is

f_{na}/a^n while Theorem B shows that the number of π in S_n consisting of a single n-cycle with $k-1$ descents is

$$\sum_{a=1}^{k}(-1)^{k-a}\binom{n+1}{k-a}f_{na}.$$

In the special case $a=2$, Theorem B shows that the number of permutations π in S_n with exactly 1 descent and n_j cycles of size j is

$$\prod_{j=1}^{n}\binom{f_{j2}+n_j-1}{n_j}$$

unless $n_1=n$ when there are no such permutations.

The structure of this paper is as follows. We give basic properties of a-shuffles in Section 2, developing enough to show how Theorem A implies Theorem B. Gessel's bijection is derived in Section 3 which also contains a proof of Theorem A. Section 4 shows how cycles of permutations are related to the pieces of the Lyndon decomposition of words and thus how Theorem A allows us to derive the distribution of a variety of functionals of the uniform distribution on $\{0,1,\cdots,a-1\}^n$.

Section 5 derives closed form and asymptotic expressions for a variety of class functions on S_n. The final section develops a curious unique factorization property of derangements.

2. Preliminaries concerning a-shuffles

We begin with a careful description of an a-shuffle. Let $A=\{0,1,\cdots,a-1\}$. For $x\in A^n$ let $x^{\uparrow}$ be the non-decreasing rearrangement of x. Let $\pi_x\in S_n$ be the unique permutation such that for each $b\in A$, π_x is an increasing map from

$$\{i: x^{\uparrow}_i=b\} \quad \text{to} \quad \{j: x_j=b\}.$$

Example. Take $a=3, n=7, x=0212210$. Then $x^{\uparrow}=0011222$ and π_x is

i	1	2	3	4	5	6	7
$\pi_x(i)$	1	7	3	6	2	4	5

The shuffle in the example involves cutting a deck of 7 cards into packets of size 2, 2, and 3. After the shuffle the original top 2 cards wind up in positions 1 and 7, the next 2 cards wind up in positions 3 and 6, and the original bottom 3 cards wind up in positions 2, 4, 5.

For the general card shuffling interpretation, consider a deck of n cards initially cut into packets of n_0 cards in packet 0 on top, n_1 cards in packet 1 below packet $0,\cdots,n_{a-1}$ cards in packet $a-1$ at the bottom. (If $n_b=0$ the bth packet is empty). Now let these packets be riffled together in the order dictated by the sequence x. The order of the cards within each packet is preserved, while the cards of the bth packet appear in those places $\{j: x_j=b\}$. The map $j\longmapsto\pi_x(j)$ gives the *place* (counting down from the top) of the card initially at place j from the top, after this riffle shuffle determined by the sequence x.

If x is picked uniformly at random from A^n, the induced distribution of π_x on S_n is an a-shuffle. Put more probabilistically, a random riffling together of a packets of random sizes defines an a-shuffle iff each card in the final deck is equally likely to come from any one of the packets, independently of what packets all the other cards come from.

To be precise, we give the following definition.

Definition 2.1 For $x \in A^n$, let $n_b = \#\{i : x_i = b\}, 0 \le b \le a-1$, and let π_x be the unique permutation π such that for each b with $n_b > 0$, for j in the range

$$n_0 + \cdots + n_{b-1} < j \le n_0 + \cdots + n_b,$$

$\pi(j)$ is an increasing function of j with $x_{\pi(j)} = b$.

A permutation has a *descent* at i if $\pi(i) > \pi(i+1)$. Note that π_x has at most $a-1$ descents.

Proposition 2.1. *The map $x \longmapsto (\pi_x, x^\uparrow)$ from A^n to $S_n \times A^n$ is one to one. The range is the set*

$$\{(\pi, y) \in S_n \times A^n : y \textit{ is non-decreasing and } Desc(\pi) \subseteq Asc(y)\}$$

where

$$Desc(\pi) = \{i : \pi(i) > \pi(i+1)\}$$

is the descent set of π and

$$Asc(y) = \{i : y(i) < y(i+1)\}$$

is the ascent set of y.

Proof. With $x^\uparrow$ the non-decreasing rearrangement of x,

$$x_j^\uparrow = b \quad \text{for } n_0 + \cdots + n_{b-1} < j \le n_0 + \cdots + n_b.$$

Thus $x^\uparrow = x \circ \pi_x$, so $x = x^\uparrow \circ \pi_x^{-1}$ and the map is one to one. For the second statement, the definition of π_x yields

$$\text{if } x_i^\uparrow = x_{i+1}^\uparrow \quad \text{then } \pi_x(i) < \pi_x(i+1).$$

Equivalently

$$\text{if } \pi_x(i) > \pi_x(i+1) \text{ then } x_i^\uparrow < x_{i+1}^\uparrow.$$

So $\text{Desc}(\pi_x) \subseteq \text{Asc}(x^\uparrow)$. If $\pi \in S^n$ and $y \in A^n$ is non-decreasing with $\text{Desc}(\pi) \subseteq Asc(\pi)$, then x defined by $x_{\pi(j)} = y_j$ has $\pi_x = \pi$ and $x^\uparrow = y$. ∎

Proposition 2.2. *(Bayer-Diaconis) The range of the map $x \longmapsto \pi_x$ from $A^n \to S_n$ is the set of all permutations in S^n with at most $a-1$ descents. If π has d descents then $\pi = \pi_x$ for exactly $\binom{a+n-d-1}{n}$ distinct sequences x in A^n.*

Proof. Thanks to Proposition 2.1 it suffices to count non-decreasing sequences with ascents at every place where π has descents. A non-decreasing sequence in A^n can be coded as a string of n stars and $a-1$ bars. All the stars to the left of the first bar represent zeros. Stars between the first two bars represent ones and so on. Stars to the right of the last bar represent $a-1$. For example, with $a=n=4, |**|*|*$ represents 1123 while $|\,|\,|\,****$ represents 3333. There are $\binom{n+a-1}{a-1}$ such sequences in all. To force ascents at d prespecified positions, set aside d bars and place the remaining $a-1-d$ bars and n stars in any arrangement. Then insert the original d bars following the d stars at positions where ascents are required. ∎

Remark. Proposition 2.2 shows that $P_{n,a}$ of (1.1) is the distribution of π_x obtained by choosing a sequence $x \in A^n$ uniformly. Bayer and Diaconis, following Gilbert and Shannon [7] and Reeds [15] give a sequential "riffle shuffle" description of $P_{n,a}$ and show $P_{n,a} * P_{n,b} = P_{n,ab}$.

The following re-writing of Proposition 2.2 is useful.

Proposition 2.3. *For $n=1,2,\cdots$, and $k=1,2,\cdots,n$, let $\#_{n,k}$ be the measure on S_n defined by restriction of counting measure to the set $\{\pi \in S_n : \pi$ has $k-1$ descents$\}$. For $a=1,2,\cdots$ let $M_{n,a}$ be the distribution on S_n of π_x when x is given counting distribution on $\{0,1,\cdots,a-1\}^n$. Then*

$$M_{n,a} = \sum_{k=1}^{a} \binom{n+a-k}{n} \#_{n,k}.$$

$$\#_{n,k} = \sum_{a=1}^{k} (-1)^{k-a} \binom{n+1}{k-a} M_{n,a}.$$

Proof. The first formula restates Proposition 2.2. The second follows from the first by inversion. ∎

Remarks. The a-shuffle probability $P_{n,a}$ on S_n is $a^{-n}M_{n,a}$. When evaluated on S_n, the identities of Proposition 2.3 reduce to classical results; indeed, $\#_{n,k}(S_n) = A_{n,k}$, the Eulerian numbers (see, e.g., Stanley [18] or Foata [5]) and $M_{n,a}(S_n) = a^n$. This yields an identity first proved by Worpitsky [20]:

$$A_{n,k} = \sum_{a=1}^{h} (-1)^{k-a} \binom{n+1}{k-1} a^n.$$

The implied bijective proof is standard and has been extended in several directions, see, e.g., Stanley [17, 18] or Buhler et al. [4].

Knowledge of the $M_{n,a}$ distribution of some function Y on S_n gives a formula for the number of permutations π with $Y(\pi)=y$ and $k-1$ descents. Let $D(\pi)$ be the number of descents in π. For $1 \le k \le n$

(2.1)

$$\#\{\pi \in S_n : Y(\pi) = y \quad \text{and} \quad D(\pi) = k-1\} = \sum_{a=1}^{k} (-1)^{k-a} \binom{n+1}{n-a} M_{n,a}(Y=y).$$

For example, in Section 1, Proposition A implies Proposition B. Many further examples are given in Section 5 below.

3. Bijections

This section gives a bijection between A^n and the collection of multisets of aperiodic necklaces with total length n. Roughly, the bijection begins with $x \in A^n$, passes to π_x as in definition 2.1 and then to a set of points in the unit interval defined from the cycles of π. Permutations $\pi \in S_n$ are written in cycle notation in the standard way, $\pi = (C_1)(C_2)\cdots$ where $C_1 = (1, \pi(1), \pi^2(1), \cdots), C_2$ is the orbit of the first card not in C_1, etc.

Definition 3.1. For $x \in A^n$, *the cycle sentence* of x is the sequence of words

$$s_x = (W_{x1}), (W_{x2}), \cdots$$

obtained from the cycle notation for π_x by replacing each symbol k by $x^{\uparrow}(k), 1 \leq k \leq n$.

The *cycle words of* x are the words in the cycle sentence of x.

Example 3.1. Take $a = 2, A = \{0,1\}, n = 14$. Let $x = (x_1 \cdots x_{14})$ be the word in the second row of the following table. Then $x_i^{\uparrow}$ and $\pi_x(i)$ are given in the 3rd and 4th rows.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
x_i	0	1	1	1	0	0	1	0	1	0	0	0	0	0
$x_i^{\uparrow}$	0	0	0	0	0	0	0	0	0	1	1	1	1	1
$\pi_x(i)$	1	5	6	8	10	11	12	13	14	2	3	4	7	9

In cycle notation

$$\pi_x = (1)(2,5,10)(3,6,11)(4,8,13,7,12)(9,14).$$

Replacing each k by $x_k^{\uparrow}$ gives the cycle sentence of x as

$$s_x = (0)(001)(001)(00101)(01).$$

A word ω of length j is a *Lyndon word* if ω is lexographically strictly smaller than each of the $j-1$ cycle shifts of ω. For example, 001 is Lyndon but not 010. Note that each word in s_x above is Lyndon. A total order on the collection of all Lyndon words of any length, called here the *repeat lexographic order*, is defined as follows: $\omega < z$ if $\omega^{\bullet}$ is lexographically less than $z^{\bullet}$, where $y^{\bullet}$ is the infinite sequence obtained by indefinite repetition of y. If $y^{\bullet}$ is regarded as a point in [0,1] through base a expansion, this is just the usual order in $[0,1]$. Note that the Lyndon words in s_x above are non decreasing.

The map $x \longmapsto s_x$ has an inverse which we now describe. Given a non increasing sequence of Lyndon words, say $\omega_1 \leq \omega_2 \leq \cdots$, with lengths $j_1, j_2, \cdots$ with $\sum j_i = n$, let $u_1 \leq u_2 \leq \cdots \leq u_n$ be the non decreasing sequence of points in the unit interval formed by taking all cyclic rearrangements of ω_i and extending each periodically. For each u_i define $x_i \in A$ as the last symbol in u_i before it repeats.

Example. Suppose the sentence of non-decreasing Lyndon words is the sentence (0)(001)(001)(00101)(01). Then cyclic rearrangements become

$$0^{\bullet}, 001^{\bullet}, 010^{\bullet}, 100^{\bullet}, 001^{\bullet}, 010^{\bullet}, 100^{\bullet},$$
$$00101^{\bullet}, 01010^{\bullet}, 10100^{\bullet}, 01001^{\bullet}, 10010^{\bullet}, 01^{\bullet}, 10^{\bullet}.$$

Ordering these as points of the unit interval gives

$$0^\bullet, 001^\bullet, 001^\bullet, 00101^\bullet, 010^\bullet, 010^\bullet, 01001^\bullet, 01010^\bullet,$$
$$01^\bullet, 100^\bullet, 100^\bullet, 100^\bullet, 10010^\bullet, 10100^\bullet, 10^\bullet.$$

Read off the last letters of these words yields $x = 01110010100000$. This is the word used in Example 3.1 which shows that s_x is as given above.

The main result of this section can now be stated. It presents a version of Gessel's bijection. Gessel and Reutenauer [6] give the history and many applications to permutation enumeration.

Theorem 3.1. *The map $x \longmapsto s_x$ of Definition 3.1 defines a bijection between A^n and the set of n letter sentences s such that*

1) each word in the sentence is a Lyndon word,

2) the words in the sentence are non-decreasing in the repeat lexographic order.

The proof of this Theorem is given later in this section. We first indicate a Corollary – and then show how it implies Theorem A of Section 1.

Clearly, a sentence of non decreasing Lyndon words is uniquely determined by the number of times each Lyndon word appears in the sentence. This gives the following version of Theorem 3.1 which will be used several times in what follows.

Corollary 3.1. *Let L_j be the set of Lyndon words of length j for the alphabet A. For $x \in A^n$, let $n_x(\omega)$ be the number of times the Lyndon word ω appears in the cycle sentence s_x. The map $x \longmapsto \{n_x(\omega), \omega \in \bigcup_j L_j\}$ defines a bijection between A^n and arrays of non-negative integers $\{n(\omega), \omega \in \bigcup_j L_j : \sum_j j \sum_{\omega \in L_j} n(\omega) = n\}$.*

Proof of Theorem A. There is an obvious bijection between Lyndon words and aperiodic circular words. Thus $\#(L_j) = f_{ja}$ of (1.3). The bijection of Corollary 3.1 is such that the number of cycles of length j in π_x is $\sum_{\omega \in L_j} n_x(\omega)$. For this number to equal n_j, there are $\binom{f_{ja}+n_j-1}{n_j}$ different ways to choose the $n(\omega)$ with $\omega \in L_j$ and so $\prod_{j=1}^{n} \binom{f_{ja}+n_j-1}{n_j}$ ways to choose the entire array. ▮

We now develop some preliminaries for the proof of Theorem 3.1. The argument falls into two parts. The first part is to show that the cycle sentence of x is composed of a non-decreasing sequence of Lyndon words. The second is to show that every such sequence comes from a unique x. Consider the permutations π_x^m obtained by m iterations of π_x for a fixed x. So $\pi_x^m(k)$ is the place of card k in the deck after a sequence of m shuffles according to π_x.

Definition 3.2 For each $k = 1, \ldots, n$, the *x-signature of card k* is the infinite word u_{xk} whose letters indicate the successive packets containing card k as it moves through the deck under repetitions of π_x. Formally, u_{xk} is the infinite word whose mth letter is

$$u_{xkm} = x^\uparrow \circ \pi_x^{m-1}(k), \quad m = 1, 2 \ldots.$$

To illustrate, for the 3 packet riffle π_x induced by $x \in \{0,1,2\}^n$,

$$u_{xk} = 012001 \cdots$$

means that under iterates of π_x, card k starts in the top packet 0, after the first shuffle appears in packet 1 for the second shuffle, then in packet 2 for the third shuffle, and so on. Since π_x^m is the identity for some m, the word u_{xk} is obviously periodic. Thus, for x fixed, the function $k \to u_{xk}$ gives a map

$$u_x : \{1, \ldots, n\} \longrightarrow \widehat{A},$$

where $\widehat{A}$ is the set of all periodic infinite words with letters in the alphabet A. It is easy to see that the map u_x can be recovered from the cycle sentence of x. The argument eventually shows that even x can be recovered from u_x.

To visualize u_x it is helpful to identify $\widehat{A}$ as a subset of [0,1] via the usual expansion in base a. Give $\widehat{A}$ the lexicographical order, corresponding to the usual order on [0,1]. Let θ denote the shift map on $\widehat{A}$:

$$\theta(a_1, a_2, \ldots) = (a_2, a_3, \ldots).$$

Note that θ is invertible on $\widehat{A}$, so $\theta^m : \widehat{A} \to \widehat{A}$ is defined for every integer m.

The upshot of the following lemma is that so far as anything to do with the cycle structure of π_x is concerned, the action of the shift θ on the card signatures u_{xk} is a faithful representation of the action of π_x on the cards k.

Lemma 3.1. *For each x in A^n the map $u_x : \{1, \ldots, n\} \to \widehat{A}$ is such that*

(i) $$u_{x1} \leq u_{x2} \leq \cdots \leq u_{xn}.$$

(ii) *If $u_{xi} = u_{xk}$ for some $i < k$, then for every $i \leq j \leq k$ and every integer m,*

$$\pi_x^m(j) - \pi_x^m(i) = j - i.$$

(iii) *For all integers m*

$$\theta^m \circ u_x = u_x \circ \pi_x^m.$$

(iv) *For each $k \in \{1, \ldots, n\}$ the period of k under the action of π_x (i.e., the length of the π_x-cycle containing k) is identical to the period of u_{xk} under the action of θ.*

Remark. In terms of card shuffling, (ii) means that when the shuffle π_x is iterated, cards i to k inclusive move like a clump of $k - i + 1$ cards glued together. No matter how often the shuffle is repeated, cards in this clump never become separated, though the whole clump will typically move around between packets in the deck.

Proof of Lemma 3.1.

Proof of (i): Let $i < k$. Follow cards i and k under iterates of the shuffle π_x. If $x^{\uparrow}(i) < x^{\uparrow}(k)$, then obviously $u_{xi} < u_{xk}$, since $u_{xj1} = x^{\uparrow}(j)$. So suppose $x^{\uparrow}(i) = x^{\uparrow}(k)$, meaning i and k start in the same packet. Then $u_{xi1} = u_{xk1}$. Now the key observation is that because π_x is increasing when restricted to each packet,

$$\text{if } \ \pi_x^{m-1}(i) < \pi_x^{m-1}(k) \ \text{ and } \ u_{xim} = u_{xkm} \text{ then } \ \pi_x^m(i) < \pi_x^m(k).$$

Consequently, if

$$u_{xim} = u_{xkm} \quad \text{for} \quad m = 1, \ldots, M$$

meaning that cards i and k are in the same packet for each of the first M shuffles (through which packet may vary as the shuffles proceed), then also

$$\pi_x^M(i) < \pi_x^M(k),$$

hence,

$$u_{xi,M+1} = x^{\uparrow} \circ \pi_x^M(i) \leq x^{\uparrow} \circ \pi_x^M(k) = u_{xk,M+1}.$$

Consequently, if $M+1$ is the first m such that $u_{xim} \neq u_{xkm}$, then $u_{xi,M+1} < u_{xk,M+1}$. That is to say $u_{xi} < u_{xk}$ in lexicographical order. The only other possibility is that no such $M+1$ exists. That is to say, $u_{xi} = u_{xk}$.

Proof of (ii): If $i<k$ and $u_{xi}=u_{xk}$, then by similar reasoning to that above, the number of cards between cards i and k can never decrease as the shuffles proceed. Since π_x^m is the identity for some m, this number of cards must remain constant.

Proof of (iii): This just says that if π_x^m maps j to k, then θ^m maps u_{xj} to u_{xk}. This follows at once from the definitions, first for $m=1$, then for any integer m.

Proof of (iv): Fix k and suppose that u_{xk} has period d under the action of θ. Then the $\theta^m(u_{xk}), 0 \leq m < d$ are distinct, and $\theta^d(u_{xk}) = u_{xk}$. Due to (iii), the $\pi_x^m(k)$, $0 \leq m < d$ are distinct. To see that k has period d under π_x it only remains to show that $\pi_x^d(k)=k$. To this end let

$$B = u_x^{-1}(u_{xk}),$$

the set of all $j \in \{1,\dots,n\}$ with the same x-signature as k. By (iii) for $m=d, B$ is π_x^d-invariant. Now (ii) shows π_x^d must act as the identity on B, hence $\pi_x^d(k)=k$. ∎

Proof of Theorem 3.1, Part I. *The cycle sentence of x is a non-decreasing sequence of Lyndon words.*

Suppose w is a word of length j in the cycle sentence. Say the first place in the corresponding cycle is k. Then k has period j under π_x. By definition of $w, u_{xk} = w^\bullet$, that is, w repeated indefinitely. And by (iii) above the signatures u_{xi} of the $j-1$ other cards i in the π_x cycle starting from k are the shifts $\theta^m(w^\bullet)$ for $m = 1,\dots,j-1$. By (iv) above, these signatures are all distinct. Also, by definition of the cycle sentence, k is the least element in a π_x-cycle of length j. Consequently, by (i) $u_{xk} < u_{xi}$ for every other card i in the cycle of length j containing k. That is to say:

$$w^\bullet < \theta^m(w^\bullet) \quad \text{for} \quad m = 1,\dots,j-1.$$

Thus w is a Lyndon word. That these words are non-decreasing follows immediately from (i) and the fact that each cycle is defined to start at the least index not in any previous cycles. ∎

Proof of Theorem 3.1, Part II. *Every n-letter sentence composed of a non-decreasing sequence of Lyndon words is the cycle-sentence of x for a unique x.*

The argument produces an explicit inverse which was illustrated at the start of this section.

Proposition 3.1. *Let $w_1 \leq w_2 \leq \cdots$ be a non decreasing sequence of Lyndon words with lengths $j_1, j_2 \cdots$, where $\sum j_i = n$. Let $u_1 \leq u_2 \leq \cdots \leq u_n$ be the non-decreasing* *sequence of n infinite words obtained by putting the n infinite words $(\theta^m(w_i),\ 0\le m\le j_i-1,\ i=1,2,\cdots)$ in non decreasing order. Let*

$$x=(x_1,\cdots,x_n) \quad \textit{where } x_k \textit{ is the first letter of } \theta^{-1}(u_k).$$

Then, x is the unique word whose cycle sentence is $w_1,w_2,\cdots$.

The proof of Proposition 3.1 follows from two lemmas. To motivate the first lemma, consider the sequence $u_1,\cdots,u_n$ of Proposition 3.1. Associate with this sequence its *counting distribution*, that is, the measure on $\widehat{A}$ defined by $\mu(B)=\#\{k:1\le k\le n,\ u_k\in B\}$. Clearly, this μ is shift invariant: $\mu(B)=\mu(\theta B)$.

Lemma 3.2. *Given a sequence $u=(u_1,\cdots,u_n)$ of infinite words in $\widehat{A}$ with $u_1\le u_2\le\cdots\le u_n$ such that the counting distribution of u is θ-invariant, there is a unique $\pi\in S_n$ satisfying*

1) *$\theta^m\circ u=u\circ\pi^m$ for all integers m.*
2) *for each $y\in\widehat{A}$, the permutation π is increasing when restricted to the set $u^{-1}(y)$.*

This π is given by

$$\pi(i)=\#\{j:\theta(u_j)<\theta(u_i)\}+\#\{j:\theta(u_j)=\theta(u_i)\quad\textit{and}\quad j\le i\}. \tag{3.1}$$

Proof. This is an elementary verification from the definitions. ∎

Note. The permutation π has the following property which is used below

$$j<k \text{ and } u\circ\pi(j)\le u\circ\pi(k) \quad \text{imply } \pi(j)<\pi(k). \tag{3.2}$$

Lemma 3.3. *For u and x as in Proposition* 3.1, *let π be given by* (3.1). *Then $\pi=\pi_x$.*

Proof. Let $\xi:\widehat{A}\to A$ be the first letter map. The definition of x makes $x=\xi\circ\theta^{-1}\circ u$. Now $\theta^{-1}\circ u=u\circ\pi^{-1}$ by Lemma 3.2(1), so $x=\xi\circ u\circ\pi^{-1}$. Because π^{-1} is a permutation, the counting distribution of x is identical to the counting distribution of $\xi\circ u$. But $\xi\circ u$ is a composition of two non-decreasing maps, hence non decreasing. Thus, $x^\uparrow=\xi\circ u$, and $x=x^\uparrow\circ\pi^{-1}$. Consequently, if $x^\uparrow(j)=b$, say, then $x_{\pi(j)}=b$. Further, if $j<k$ are such that $x^\uparrow(j)=x^\uparrow(k)=b$, then $\xi\circ u_j=\xi\circ u_k=b$. Since θ is increasing on sequences with a given first letter,

$$u\circ\pi(j)=\theta\circ u_j\le\theta\circ u_k=u\circ\pi(k),$$

so $\pi(j)<\pi(k)$ by (3.2). This proves $\pi=\pi_x$. ∎

Proof of Proposition 3.1. We will argue that x defined in Proposition 3.1 makes $u_j=u_{xj}$, $1\le j\le n$ and x is unique in A^n with this property. Clearly this property is equivalent to $w_i=w_{xi}$ for every i where w_{xi} is the ith word in the cycle sentence of x.

Let ξ be the first letter map of Lemma 3.3. For $m=1,2,\cdots$, $u_{xkm}=x^\uparrow\circ\pi^{m-1}(k)$. Now Lemma 3.3 gives $x^\uparrow=\xi\circ u$, so $u_{xkm}=\xi\circ u\circ\pi^{m-1}(k)=\xi\circ\theta^{m-1}\circ u_k=u_{km}$, where the next to last equality follows from Lemma 3.2 (1).

For uniqueness, suppose $x\in A^n$ has $u_j=u_{xj}, 1\le j\le n$. Then, the $(u_{xk},\ 1\le k\le n)$ determine $x^\uparrow(k)=u_{xk}$, for $1\le k\le n$. Also, u_x and π_x satisfy the hypothesis of Lemma 3.2. So u_x determines $x^\uparrow$, hence π_x, and so $x=x^\uparrow\circ\pi_x^{-1}$. ∎

4. Word lengths in the Lyndon decomposition of a random word and irreducible polynomials

The joint distribution appearing in Theorem A turns out to be identical to the joint distribution of word lengths in the Lyndon decomposition of a random word of length n from an alphabet of a letters. This would appear to provide the most elementary proof that the formula (1.2) does define a joint probability distribution over n-tuples of non-negative integers (n_j) with $\Sigma_j j n_j = n$.

To formulate this precisely, fix the alphabet A with a letters. Let L_j be the set of Lyndon words of length j. The set of all Lyndon words is $\cup_j L_j$. Instead of the repeat lexicographic order imposed on $\cup_j L_j$ in Section 3, now give $\cup_j L_j$ the ordinary lexicographic order for words of finite length, in which $v < w$ for any word $w = vx$ obtained by concatenating v and another word x (e.g., $01 < 01001$ in this ordinary order, but $01 > 01001$ in the repeat order).

According to the fundamental result of Lyndon (see, e.g., Lothaire [12], Theorem 5.15), when $\cup_j L_j$ is given the ordinary lexicographical order, every word $x \in A^n$ can be written uniquely as the concatenation of a non-increasing sequence of Lyndon words. Call this the *Lyndon decomposition of* x. Let $M_x(w)$ denote the number of times the Lyndon word w appears in the Lyndon decomposition of x. An immediate consequence of the Lyndon decomposition is that the map

$$x \longrightarrow (M_x(w),\ w \in \cup_j L_j)$$

induces a bijection between words $x \in A^n$ and arrays of non-negative integers

$$(n(w),\ w \in \cup_j L_j : \sum_j j \sum_{w \in L_j} n(w) = n). \tag{4.1}$$

This should be compared with the different bijection between the same sets given in the proof of Theorem A (Section 3). It follows immediately that formula (1.2) gives the probability that the Lyndon decomposition of a word picked at random from A^n contains n_j Lyndon words of length j, $1 \le j \le n$. That is to say, we have the following:

Proposition 4.1. *Let $M_j = M_{xj} = \sum_{w \in L_j} M_x(w)$ be the number of Lyndon words of length j in the Lyndon decomposition of a word x picked uniformly at random from A^n. Then the joint distribution of the counts $(M_j,\ 1 \le j \le n)$ is identical to the joint distribution of the cycle counts $(N_j,\ 1 \le j \le n)$ derived from an a-shuffle of n cards, as described in Theorem A.*

As a consequence of this proposition, every result presented in the following sections concerning the distribution of N_j and asymptotic distribution of $(N_j,\ 1 \le j \le n)$ as $n \to \infty$, applies verbatim to M_j and the asymptotic distribution of $(M_j,\ 1 \le j \le n)$ as $n \to \infty$.

Remark. The composition of the two different bijections between A^n and arrays (4.1) defines a permutation of A^n which acts on each word by rearranging its letters. Thus we have another way of inducing a permutation in S_n from a word in A^n, besides the a-shuffle. This may be rather artificial, but perhaps worth studying further.

There is another parallel set up where the results of Theorem A apply. Let $\mathbb{F}$ be a finite field with q elements. Here $q = p^m$ for p a fixed prime and $m \geq 1$ an integer. There are q^n distinct monic polynomials of degree n over $\mathbb{F}$. Suppose one of these polynomials is chosen and factored into monic irreducible polynomials. The number of monic irreducibles of degree j is well known to be f_{jq} of (1.3) (see, e.g., Theorem 3.25 in Lidl and Niederreiter [11]). Thus, the chance that a randomly chosen polynomial is irreducible of degree j is $f_j q/q^n$. Arratia, Barbour, and Tavaré [1] have derived approximations for a variety of functions of the K_j. All of these are applicable to the cycles of a random permutation under $P_{n,a}$ or to the decomposition of Lyndon words.

In the language of permutations, they give approximations for the joint distribution of small cycles, for the number of medium length cycles, and for the joint distribution of the longest cycles. They also derive limit theorems for the total number of cycles. All of their approximations are accompanied by explicit error estimates.

Proposition 4.2. *Let $K_j(h)$ be the number of irreducible factors of degree j in the decomposition of h chosen uniformly at random from the set of monic polynomials of degree n over a field with q elements. Then, the joint distribution of the counts $(K_j,\ 1 \leq j \leq n)$ is identical to the joint distribution of the cycle counts $(N_j,\ 1 \leq j \leq n)$ derived from a q-shuffle of n cards as described in Theorem A.*

5. Exact and Asymptotic Distributions

In this section we derive closed form distributions for the number of i cycles after an a-shuffle of n cards. We also derive the limiting joint distribution for the number of cycles of various types as n tends to infinity and as n and a tend to infinity. For ease of reference, the results are described first. Combined with the inversion theorem of Section 2, the results give formulae and asymptotics for the number of cycles of permutations in S_n with a fixed number of descents. Using the results of Section 4 they give formulae for the word lengths in the Lyndon decomposition of a word over an arbitrary alphabet.

Proposition 5.1. *Let $P_{n,a}$ denote the distribution of an a-shuffle and $N_i(\pi)$ the number of cycles of length i in the permutation $\pi \in S_n$. Then, for $m = 0, 1, 2, \cdots, \lfloor n/i \rfloor$*

$$P_{n,a}(N_i = m) = \binom{f_{ia} + m - 1}{m} a^{-im} p_{n-im,a,i} \tag{5.0}$$

where

$$p_{n,a,i} = P_{n,a}(N_i = 0) = \sum_{k=0}^{\lfloor n/i \rfloor} \binom{f_{ia}}{k} (-1)^k a^{-ik}.$$

Note: $p_{0,a,i} = 1$ and $p_{n,a,i} = (1 - a^{-i})^{f_{ia}}$ for $n \geq i f_{ia}$. Here $f_{ia} = \frac{1}{i} \sum_{d|i} \mu(d) a^{d/i}$.

Proposition 5.2. *Let* $(n)_k = n(n-1)\cdots(n-k+1)$. *Then*

$$E_{n,a}(N_i)_k = (f_{ia})_k \sum_{m=k}^{\lfloor n/i \rfloor} \binom{m-1}{m-k} a^{-im}.$$

Remark. For example, $E_{n,a}(N_1) = 1 + \frac{1}{a} + \frac{1}{a^2} + \cdots + \frac{1}{a^{n-1}}$.

For the next result, recall that a negative binomial variable X with parameters f and p has

$$P\{X = m\} = \binom{f+m-1}{m} p^m (1-p)^f, \; m = 0, 1, 2, \cdots.$$

Proposition 5.3. *For fixed a as n tends to infinity, the $P_{n,a}$ joint distribution of the numbers of N_i of i-cycles converges to the distribution of independent negative binomial variables with parameters (f_{ia}, a^{-i}).*

The analysis of riffle shuffles given by Bayer and Diaconis [3] showed that $a \gg n^{3/2}$ is necessary and sufficient to have $P_{n,a}$ close to the uniform distribution in total variation distance. In the next two propositions, we show that features of a permutation that depend only on cycle structure have the correct distribution after $a(n)$ shuffles, where $a(n)$ tends to infinity arbitrarily slowly with n.

Proposition 5.4. *Let $a(n)$ tend to infinity with n. As $n \to \infty$, the $P_{n,a(n)}$ joint distribution of the numbers N_i of i-cycles converges to the distribution of independent Poisson variables with parameters $1/i$.*

For the next result, recall that the limiting distribution of the large cycles of a permutation chosen uniformly in S_n has been determined by Goncharov [8, 9], Shepp and Lloyd [16], Vershik and Schmidt [19], and others. For example, the mean length of the longest cycle L_1 is approximately $.63n$ and L_1/n has a known limiting distribution.

Proposition 5.5. *Fix k, and let $L_1(\pi), L_2(\pi), \cdots, L_k(\pi)$ be the lengths of the k longest cycles in π. Then, for a fixed, or growing with n, as $n \to \infty$,*

$$\left| P_{n,a}\{L_1/n \le t_1, \cdots, L_k/n \le t_k\} - P_{n,\infty}\{L_1/n \le t_1, \cdots, L_k/n \le t_k\} \right| \longrightarrow 0,$$

uniformly in $t_1, t_2, \cdots, t_k$.

The proofs of the propositions all depend on information from generating functions. We thus begin by rewriting Theorem A analytically. For a function $X(\pi)$ of a permutation π, let

$$E_{n,a}X = \sum_{\pi \in S_n} P_{n,a}(\pi) X(\pi).$$

Proposition 5.6. *Let*

$$g_{na}(x_1,\cdots,x_n) = E_{n,a}\prod_{i=1}^{n} x_i^{N_i}$$

be the generating function for the cycle counts N_i derived from an a-shuffle of n cards. Then

$$\sum_{n=0}^{\infty} z^n a^n g_{na} = \prod_{i=1}^{\infty}(1-z^i x_i)^{-f_{ia}}. \tag{5.1}$$

Proof. Theorem A translates to

$$g_{na} = a^{-n}\sum\prod_{i=1}^{n}\binom{f_{ia}+n_i-1}{n_i}x_i^{n_i}$$

where the sum is over all sequences of non-negative integers $(n_1,n_2,\cdots,n_n)$ with $\sum in_i = n$. Now

$$(1-z)^{-f} = \sum_{m=0}^{\infty}\binom{f+m-1}{m}z^m.$$

Expanding each term in the product of (5.1) and comparing coefficients of $(za)^n$ completes the proof. ∎

Remarks. The identity (5.1) is an extension of Witts cyclotomic identity

$$(1-az)^{-1} = \prod_{i=1}^{\infty}(1-z^i)^{-f_{ia}}. \tag{5.2}$$

This can be seen by setting $x_i=1$ in (5.1). The above argument shows this identity amounts to the fact that the probabilities of Theorem A sum to 1 when summed over all sequences $(n_1,\cdots,n_n)$ of non negative integers with $\sum in_i=n$. Metropolis and Rota [13, 14] give alternate proofs of (5.2). A probabilistic interpretation of (5.1) can be given as follows:

Proposition 5.7. *Fix t with $0<t<1$. Let N can be chosen at random from $\{0,1,2,\cdots\}$ according to the geometric distribution $P(N=n)=(1-t)t^n$. Given N, let π result from a random a-shuffle of N cards. Let N_i be the number of cycles of π of length i. Then, the random variables N_i, $1\le i<\infty$, are independent and N_i has a negative binomial distribution with parameters f_{ia} and $(t/a)^i$.*

Proof. Using (5.2) in (5.1)

$$\sum_{n=0}^{\infty}(1-t)t^n g_{na} = \prod_{i=1}^{\infty}\left(\frac{1-(t/a)^i}{1-(t/a)^i x_i}\right)^{f_{ia}}. \tag{5.3}$$

This is valid in the ring of formal power series. If all but a finite number of x_i are set to 1, it is valid for $0\le t<1$ and the remaining x_i in $[0,1]$, for all $a=1,2,\cdots$. This shows that any finite collection of N_i have the stated distribution since a negative binomial variable with parameters f and p has generating function

$$\left(\frac{1-p}{1-px}\right)^f. \qquad \blacksquare$$

Proof of Proposition 5.1. Consider the case $i=1$. Let $N=N_1$ and write

$$N=\sum_{k=0}^{a-1} N(k)$$

where $N(k)$ is the number of fixed points in the kth packet. A counting argument based on Corollary 3.1 shows (as in the uniform case)

$$P_{n,a}\{N=m\}=a^{-m}\binom{a+m-1}{m}P_{n-m,a}\{N=0\}.$$

In the generating function (5.1) set $x_i=1$, $2\leq i<\infty$. This gives the sum of the generating functions for N as n varies as $\left(\frac{1-z}{1-zx}\right)^a(1-az)^{-1}$. Setting $x=0$ gives the generating function $\sum(az)^n P_{n,a}(N=0)=(1-z)^a(1-az)^{-1}$. Thus

$$P_{n,a}(N=0)=\sum_{k=0}^{n}\binom{a}{k}(-1)^k a^{-k}.$$

The argument above extends to general i in a straightforward way: One shows (5.0) by a counting argument, then calculates $p_{n,a,i}$ from the generating function. $\blacksquare$

Remark. The $N(k)$ introduced in the proof of Proposition 5.1 have some nice properties. Corollary 3.1 shows that for $m_k\geq 0$, $\sum_k m_k=m$,

$$P\{N(k)=m_k,\ 0\leq k\leq a-1|N=m\}=\binom{a+m-1}{m}^{-1}.$$

Consequently, the $N(k)$, $0\leq k\leq a-1$ are exchangeable random variables. It follows easily from the case $k=0$, $P\{N(k)\geq j\}=a^{-j}$.

Proof of Proposition 5.2. Use (5.3) with $x_j=1$ for $j\neq 1$. This gives

$$\sum_{n=0}^{\infty}(1-t)t^n E_{n,a}x^{N_i}=\left(\frac{1-(t/a)^i}{1-(t/a)^i x}\right)^{f_{ia}}.$$

Differentiate k times with respect to x and set $x=1$ to get

$$\sum_{n=0}^{\infty}(1-t)t^n E_{n,a}(N_i)_k=(f_{ia})_k(t/a)^{ki}(1-(t/a)^i)^{-k}.$$

The result follows by comparing coefficients of t^n on both sides. ∎

Proof of Proposition 5.3. A negative binomial variable with parameters f and p has kth falling factorial moment $(f)_k\left(\frac{p}{1-p}\right)^k$. Letting n tend to infinity with a fixed, shows that the moments $E_{n,a}(N_i^k)$ converge to the moments of a negative binomial (f_{ia}, a^{-i}). Essentially the same computation works for any finite set of joint moments. Further details are omitted. ∎

Proof of Proposition 5.4. For notational simplicity, the argument is given just for N_1. From Proposition (5.2), the kth falling factorial moment of N_1 is

$$E_{n,a}(N_1)_k = \frac{(a)_k}{a^k}\sum_{j=0}^{n-k}\binom{k+j-1}{j}a^{-j}.$$

The sum is bounded below by 1 and above by $(1-\frac{1}{a})^{-k}$. Thus, for any fixed k, for $a(n)$ tending to infinity with n

$$E_{n,a(n)}(N_1)_k \to 1 \quad \text{as} \quad n \to \infty.$$

This is the kth falling factorial moment for a Poisson variate with parameter 1. The argument for the joint moments of $N_1, N_2, \cdots, N_i$ is essentially the same for any fixed i. Further details are omitted. ∎

Remark. The method of moments proof given above does not give a bound for finite n and $a(n)$. In the case of convergence of N_i to Poisson $(1/i)$ under the uniform distribution, sharp rates of convergence have been given by Barbour and Stein [2]. The closed form expressions of Proposition 5.1 may be used to get explicit bounds for individual N_i but more refined probabilistic representations are need to get rates for the joint distribution of $N_1, \ldots, N_i$. This is carried out by Arratia, Barbour, and Tavaré [1] in the language of polynomials over a field (cf. Proposition 4.2).

Proof of Proposition 5.5. This can be derived from the generating function. We will not give further details because they have been elegantly written out in the language of the distribution of the large irreducible factors of random polynomials with coefficients over a finite field with q elements ($q=a$) by Arratia, Barbour, and Tavaré [1] or by Hansen [10], example 3).

Remark. The results of Arratia et al. [1] and Hansen [10] can be used to show that the large cycles have what are called Poisson-Dirichlet asymptotics. These have been very thoroughly studied in statistical and population biology settings.

6. Derangements with one descent

The theory developed in Section 3 gives a curious operation with a unique decomposition theory on the set of fixed point free permutations with exactly one descent. Call such a permutation a *deriffle.* For example, 236145 $\in S_6$ is a deriffle.

Let D_m be the set of deriffles in S_m. It is easy to see that $\#\{D_m\}=2^{m-2}$. A subset A of $\{1,2,\cdots,n\}$ is π invariant if $\pi \in S_n$ maps A to A. If A has m elements, the action of π on A defines a permutation $\sigma \in S_m$ via the increasing correspondence between $\{1,2,\cdots,m\}$ and A. We will say π acts like σ on A.

Proposition 6.1. *For $\alpha \in D_m$, $\beta \in D_n$, there is a unique $\pi \in D_{m+n}$, denoted $\pi = \alpha+\beta$, which admits disjoint, invariant sets of size m and n on which π acts like α and β respectively. If $d(\alpha)$ is the unique place d with $\alpha(d)>\alpha(d+1)$, then $d(\alpha+\beta)=d(\alpha)+d(\beta)$. The operation $+$ on $D_\infty=\bigcup_{m=2}^{\infty}D_m$ is commutative and associative. If $A_1, A_2, \cdots$ are the π invariant sets defined by the cycles of $\pi \in D_n$, and π acts like σ_j on A_j, then $\pi=\sum_j \sigma_j$. This representation gives the unique decomposition of π into cyclic deriffles up to rearrangement.*

Remark. Thus the cyclic deriffles are the "primes" of D_∞. It is easy to see that there are $f_{2,m}$ cyclic deriffles in S_m.

Proof. By Proposition 2.1, for each $\beta \in D_n$, there is a unique sequence $x \in \{0,1\}^n$ such that $\beta = \pi_x$. Corollary 3.1 identifies x and hence β with an array of non-negative integers:

$$\beta \longleftrightarrow (n_\beta(w), w \in \cup_j L_j : n_\beta(w) = 0 \quad \text{for } w \in L_1, \quad \text{and} \quad \sum_j \sum_{\omega \in L_j} n_\beta(w) = n).$$

Now $\alpha \in D_m$ corresponds to a similar array of coefficients, say $n_\alpha(w)$. From the definitions, a permutation $\pi \in S_n$ acts like α and β on disjoint π-invariant subsets if and only if

$$n_\pi(w) = n_\alpha(w) + n_\beta(w) \quad \text{for} \quad w \in \bigcup_j L_j.$$

Corollary 3.1 says there is a unique corresponding such $\pi \in D_{n+m}$.

This gives a commutative, associative product. To see $d(\alpha+\beta)=d(\alpha)+d(\beta)$, simply note that for $\alpha \in D_m$, $d(\alpha)$ is the number of zeros in the sequence $x \in \{0,1\}^m$ such that $\pi_x = \alpha$. The sequences x and y corresponding to α and β are disjoint subsequences of the sequence z corresponding to $\alpha+\beta$. ∎

Example. The sum $236145+312=345791268$. Indeed, the first permutation corresponds to 100110, the second to 100. These correspond to Lyndon words 000111 and 001. Taking all cyclic rearrangements and ordering yields

$$000111^\bullet, 001^\bullet, 001110^\bullet, 010^\bullet, 011100^\bullet, 100011^\bullet, 100^\bullet, 11001^\bullet, 111000^\bullet.$$

The last letters of these are 110001010 which corresponds to the stated sum.

Note added in proof

Corollary of Proposition 5.3. *Fix a. For each Lyndon word w, let N^w denote the renadom number of times that w appears in the cycle sentence induced by an a-shuffle of n cards. As $n \to \infty$, for each $w \in L_i$ the asymptotic distribution of N^w is geometric with parameter a^{-i}; moreover the N^w are asymptotically independent for $w \in \cup_i L_i$.*

Proof. From Corollary 3.1, for fixed n and a, under $P_{n,a}$ given $N_i = n_i$ for $i = 1, 2, \ldots$, where $\Sigma_i i n_i = n$, The sequence of random variables N^w, as w ranges over some listing of words in L_i, is uniformly distributed over all sequences of non-negative integers with sum n; moreover these sequences are independent as i varies. Thus for each n and a, the $P_{n,a}$ conditional distribution of the N^w, $w \in \cup_{i=1}^{j} L_i$ given $N_1, \ldots, N_j$ is exactly as claimed in the limit. So the claimed convergence in distribution follows immediately from that of the $N_1, \ldots, N_j$. ∎

This argument also yields a similar corollary to Proposition 5.7. There the counts N^w of the various Lyndon words w are independent geometrics, with parameters $(t/a)^i$ for $w \in L_i$. So the expression $N_i = \Sigma_{w \in L_i} \; N^w$ decomposes the negative binomial variable N_i as the sum of f_{ia} i.i.d. geometric variables.

References

[1] Arratia, R., Barbour, A., and Tavaré, S.: On Random Polynomials over Finite Fields. Technical Report, Department of Statistics, U.S.C., 1992.

[2] Barbour, A., and Stein, C.: On the joint distribution of the cycles of random permutations, Technical report, Dept. of Statistics, Stanford University, 1991.

[3] Bayer, D., and Diaconis, P. Trailing the dovetail shuffle to its lair. *Ann. Appl. Prob.*, **2** (1992), 294–313.

[4] Buhler, J., Eisenbud, D., Graham, R. L., and Wright, C.: Juggling drops and descents. *Amer. Math. Monthly*, **101** (1994), 507–519.

[5] Foata, D.: Distributions euleriennes et mahoniennes sur le groupe des permutations. In: M. Aigner (ed.), *Higher Combinatorics*, Holland: Reidel, 1977.

[6] Gessel, I., and Reutenauer, C.: Counting permutations with given cycle structure and descent set. *J. Combin. Theory, A* **64** (1993), 184–215.

[7] Gilbert, E.: Theory of Shuffling. Technical memorandum, Bell Laboratories, 1955.

[8] Goncharov, V. L.: Sur la distribution des cycles dans les permutations, C. R. (Doklady). *Acad. Sci. URSS (N.S)* **35** (1942), 267–269.

[9] Goncharov, V. L.: Du domaine d'analyse combinatoric. *Bull. Acad. Sci. USSR Ser. Mat.* (Izv. Akad. Nauk SSSR) **8** (1944), 3–48. *Amer. Math. Soc. Trans.* (2) **19** (1962), 1–46.

[10] Hansen, J.: Order statistics for decomposable combinatorial structures. *Rand. Struct. Alg.* **5** (1994), 517–533.

[11] Lidl, R., and Niederreiter, H.: *Introduction to Finite Fields and their Applications*, Cambridge University Press, Cambridge, 1986.

[12] Lothaire, M.: *Combinatorics on Words*, Addison Wesley, Reading, Mass, 1983.

[13] Metropolis, N, and Rota, G. C.: Witt vectors and the algebra of necklaces, *Adv. Math.* **50** (1983), 15–125.

[14] Metropolis, N., and Rota, G. C.: The cyclotomic identity, *Contemp. Math.* **34** (1984), 19–27.

[15] Reeds, J.: Theory of Shuffling. Unpublished manuscript, 1976.

[16] Shepp, L., and Lloyd, S. P.: Ordered cycle length in a random permutation, *Trans. Amer. Math. Soc.* **121** (1966), 340–357.

[17] STANLEY, R.: *Ordered Structures and Partitions*, Memoirs of The Amer. Math. Soc., Providence, R.I. 1971.

[18] STANLEY, R.: *Enumerative Combinatorics*, Wadsworth, Belmont CA, 1986.

[19] VERSHIK, A. M., and SCHMIDT, A.: Limit measures arising in asymptotic theory of symmetric groups, *Prob. Th. Appl.* **22** (1977), 72–88, **23** (1977), 34–46.

[20] WORPITSKY, J.: Studien über die Bernoullischen und Eulerschen Zahlen, *Jour. für die reine und angewandte Math.* **94** (1881), 103–232.

Persi Diaconis

Dept. of Mathematics
Harvard University
Cambridge, MA 02138

Michael McGrath

410B2 Cirrus Logic
3100 West Warren Avenue
Fremont, CA 94358

Jim Pitman

Dept. of Statistics
University of California
Berkeley, CA 94720