AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **86(3)** (2023), Pages 414–422

The small cycle counts of random feedback shift registers

Dudley Stark

Queen Mary, University of London
Mile End, London E1 4NS, U.K.
d.s.stark@qmul.ac.uk

Abstract

The small cycle counts of a random feedback register are shown to be binomially distributed and mutually independent. The proof uses strings related to 2-coloured necklaces.

1 Introduction

Feedback shift registers are sequences $x_i \in \{0,1\}$, $i \geq 1$, where 0 and 1 are the elements of the finite field with two elements $\mathbb{F}$, given by a recursion

$$x_{t+n} = x_t \oplus f(x_{t+1}, x_{t+2}, \ldots, x_{t+n-1}), \quad n \geq 2,$$

for a Boolean function

$$f : \mathbb{F}_2^{n-1} \to \mathbb{F}_2.$$

We will consider all $2^{2^{n-1}}$ possible f to be equally likely. The map

$$\begin{aligned} \pi_f : \mathbb{F}_2^n &\to \mathbb{F}_2^n \\ (x_0, x_1, \ldots, x_{n-1}) &\mapsto (x_1, \ldots, x_{n-1}, x_{n+1}) \\ &= (x_1, \ldots, x_{n-1}, x_0 \oplus f(x_1, x_2, \ldots, x_{n-1})) \end{aligned}$$

is easily shown to to be a permutation of the 2^n objects in $\mathbb{F}_2^n$.

Let $S_i(n)$ be the number of cycles of length i in π_f. The distributions of $S_1(n)$ and $S_2(n)$ are straightforward to find. As is stated in [10], $(0,0,\ldots,0) \in \mathbb{F}^n$ is a fixed point of π_f if and only if $f(0,0\ldots,0) = 0$, $(0,0,\ldots,0) \in \mathbb{F}^{n-1}$. Similarly, $(1,1,\ldots,1) \in \mathbb{F}^n$ is a fixed point of π_f if and only if $f(1,1\ldots,1) = 0$. These are the only two possible fixed points, as setting $(x_1, x_2, \ldots, x_n) = (x_2, x_3, \ldots, x_{n+1})$ shows. Thus, $S_1(n) \sim \text{Binomial}(2, 1/2)$ for all $n \geq 1$. The only possible 2-cycle is given by

$$\begin{aligned} (1,0,1,0,\ldots,(1+(-1)^{n+1})/2) &\mapsto (0,1,0,1,\ldots,(1+(-1)^n)/2) \\ (0,1,0,1,\ldots,(1+(-1)^n)/2) &\mapsto (1,0,1,0,\ldots,(1+(-1)^{n+1})/2), \end{aligned}$$

ISSN: 2202-3518

for which

$$f(0,1,0,\ldots,(1+(-1)^{n+1})/2) = (1+(-1)^{n+1})/2$$

and

$$f(1,0,1,\ldots,(1+(-1)^{n})/2) = (1+(-1)^{n+1})/2)$$

are required. Therefore, $S_2(n) \sim \text{Binomial}(1,1/4)$ for $n \geq 2$.

Define

$$\eta_k = \frac{1}{k}\sum_{d|k} 2^d \mu(k/d), \tag{1}$$

where μ is the Möbius function. It is stated in [10] that η_k is the number of possible cycles of length k; we prove this fact at (11) in the proof of Theorem 1. We now present our main result.

Theorem 1 *For each $1 \leq k \leq (n-1)/3$, $S_k(n) \sim \text{Binomial}(\eta_k, 2^{-k})$ and the random variables $S_k(n)$, $1 \leq k \leq (n-1)/3$, are mutually independent.*

2 Logarithmic combinatorial objects

Let $N := 2^n$. The cycle structure $(S_1(n), S_2(n) \ldots, S_N(n))$ of π_f was suggested to be "flat", meaning close to the cycle structure of a permutation chosen at random from all $N!$ possibilities, in [10]. Theorem 1 implies that it is impossible that the cycle structure is flat or even has the distribution of the cycle structure of any random decomposable structure. To appreciate the significance of Theorem 1, we will therefore review the theory of decomposable structures and especially that of logarithmic combinatorial objects.

Let us consider the flat distribution on permutations. We denote the number of cycles of size i in a permutation on n elements chosen at uniformly at random by $C_i(n)$. Using inclusion-exclusion, [11] proved

$$\mathbb{P}(C_i(n) = k) = \frac{i^{-k}}{k!}\sum_{j=0}^{\lfloor n/i \rfloor - k} (-1)^j \frac{i^{-j}}{j!},$$

where, for any real x, $\lfloor x \rfloor$ is the largest integer less than or equal to x. The case $i = 1$ gives the solution to the following problem. Suppose that n individuals exchange hats randomly. What is the probability that none of the individuals get back their own hat? We see that

$$\lim_{n\to\infty} \mathbb{P}(C_i(n) = k) = e^{-1/i}\frac{i^{-k}}{k!} = \mathbb{P}(Z_i = k),$$

where $Z_i \sim \text{Poisson}(1/i)$. Moreover, with the Z_i mutually independent, it can be shown, for example by the method of moments, that the finite dimensional distributions of the process of $C_i(n)$ converges to those of the Z_i:

$$\lim_{n\to\infty} \mathbb{P}(C_1(n) = k_1, C_2(n) = k_2, \ldots, C_s(n) = k_s) = \mathbb{P}(Z_1 = k_1, Z_2 = k_2, \ldots, Z_s = k_s)$$

for all fixed $s \geq 1$ and $k_j \geq 0$ for all $1 \leq j \leq s$.

The connection between the $C_i(n)$ and the Z_i becomes manifest through the Conditioning Relation

$$(C_1(n), C_2(n) \ldots, C_n(n)) \stackrel{d}{=} ((Z_1, Z_2, \ldots, Z_n)|T_n = n), \tag{2}$$

where $\stackrel{d}{=}$ means equality in distribution and

$$T_n = \sum_{i=1}^{n} iZ_i.$$

The Conditioning Relation (2) can be proved for permutations by Cauchy's formula for the number of permutations having a given cycle structure; see (1.2) of [1].

The effect of the conditioning in (2) on Z_i for small i was was quantified in Theorem 2 of [4]. Given $1 \leq b \leq n$, Let $d_b(n)$ be the total variation distance

$$d_b(n) = \frac{1}{2} \sum_{\mathbf{a} \in \mathbb{Z}_+^b} \left| \mathbb{P}((C_1(n), C_2(n) \ldots, C_b(n)) = \mathbf{a}) - \mathbb{P}((Z_1, Z_2, \ldots, Z_b) = \mathbf{a}) \right|. \tag{3}$$

There is an (explicit) function satisfying $\log F(x) \sim -x \log x$ as $x \to \infty$ such that for all $1 \leq b < n$, the upper bound

$$d_b(n) \leq F(n/b)$$

holds. Thus, the distance $d_b(n)$ goes to 0 super-exponentially quickly as a function of n/b when $b = o(n)$.

The distribution of the sizes of the largest cycles is also well understood. Let

$$\Delta = \left\{ (x_1, x_2, \ldots) \in [0,1]^\infty : \sum_{i=1}^{\infty} x_i = 1 \right\}$$

and let $\sigma_1, \sigma_2, \ldots$ be the points of a Poisson process with intensity $\theta e^{-x}/x$, $x > 0$. The Poisson-Dirichlet distribution $PD(\theta)$ on Δ was defined in [12] to be the distribution of $(\sigma_1/\sigma, \sigma_2/\sigma, \ldots)$ on Δ, where $\sigma = \sigma_1 + \sigma_2 + \cdots$. Let $L_i(n)$ be the size of the ith largest cycle of a permutation of size n chosen uniformly at random, with $L_i(n) = 0$ if the permutation has fewer than i cycles. It was shown in [13, 16] that

$$n^{-1}(L_1(n), L_2(n), \ldots) \stackrel{d}{\to} \mathbf{Z} \text{ with } PD(1) \text{ distribution},$$

where $\stackrel{d}{\to}$ denotes convergence in distribution.

Theorem 1 shows that the distribution of cycles of π_f cannot actually be flat, because the small cycle counts are binomially distributed, not Poisson. Remarkably, however, [2] proves that the large cycles of π_f are asymptotically flat. Let $\mathbf{L}(n)$ denote the process of large cycles of π_f. Then,

$$N^{-1}\mathbf{L}(n) \stackrel{d}{\to} \mathbf{Z} \text{ with } PD(1) \text{ distribution}.$$

The conditioning relation (2) holds for many other decomposable combinatorial objects besides permutations, as is explained in [1, 5], which we follow below.

Decomposable combinatorial objects on a set of labelled elements are called assemblies. Examples include set partitions, graphs, 2-regular graphs, permutations, mappings, and both unrooted and rooted forests. Mappings are functions from a finite set to itself and have components which are cyles of rooted trees. Suppose that for an assembly there are m_i possible components that can be assembled from any subset of vertices of size i. There are $n!$ permutations on n elements and $m_i = (i-1)!$ cycles; there are n^n mappings on n elements and $m_i = (i-1)! \sum_{k=0}^{i} i^k/k!$ cycles of rooted trees. Let the number of components of a randomly chosen instance of an assembly of size i be denoted by $C_i(n)$ as was done above for permutations. Then, the conditioning relation (2) holds with

$$Z_i \sim \text{Poisson}\left(\frac{m_i x^i}{i!}\right)$$

independent, for any $x > 0$. The choice of x is arbitrary; however for permutations it is natural to take $x = 1$.

Multisets have a universe of objects with positive integer weights and m_i objects of weight i. The integer n is partitioned into parts, and for every part of size i one of the m_i objects of weight i is assigned. Examples include integer partitions, mappings patterns, forests of unlabeled rooted and unrooted trees, and monic polynomials over finite fields. There are q^n monic polynomials of degree n over a finite field of size q. They can be factored into irreducible monic polynomials, of which there are $m_i = \frac{1}{i}\sum_{d|i} \mu(i/d) q^d$ of degree i. Let the number of parts of weight i of a randomly chosen instance of an assembly be denoted by $C_i(n)$. Then, choosing from all instances of a multiset of total weight n, the conditioning relation (2) holds with

$$Z_i \sim \text{Negative Binomial}\left(m_i, x^i\right)$$

independent, for any $0 < x < 1$.

Selections are like multisets, except that each object can be assigned at most once. In the case of monic polynomials over finite fields, the corresponding selection is monic square free monic polynomials over finite fields. Choosing from all instances of a selection of total weight n, the conditioning relation (2) holds with

$$Z_i \sim \text{Binomial}\left(m_i, \frac{x^i}{1+x^i}\right)$$

independent, for any $x > 0$.

If permutations on n letters are chosen proportionally to $\theta^{K(n)}$ where $K(n)$ is the number of cycles in the permutation, the resulting distribution on component counts is called the Ewens sampling distribution, which arose originally in population genetics [8]. The Ewens sampling formula satisfies the conditioning relation (2) with

$$Z_i \sim \text{Poisson}(\theta/i). \tag{4}$$

In general, the distribution of component counts when an instance of a combinatorial object is chosen proportionally to $\theta^{K(n)}$, where $K(n)$ is the number of components, is called tilted.

For many random combinatorial objects, which satisfy (2), x can be chosen such that the Logarithmic Condition,

$$i\mathbb{P}(Z_i = 1) \to \theta,\ i\mathbf{E}Z_i \to \theta \ \text{ as } i \to \infty, \tag{5}$$

holds for some $\theta > 0$. We call such an object logarithmic. The Ewens sampling formula is logarithmic with Z_i defined by (4) and mappings are logarithmic with $\theta = 1/2$ (taking $x = e^{-1}$). Let $d_b(n)$ be defined by (3) for a logarithmic combinatorial object. Under certain auxiliary conditions,

$$d_b(n) = o(b/n)$$

if $\theta = 1$, and

$$d_b(n) \sim c(b)b/n$$

where

$$c(b) = \frac{|1-\theta|}{2b}\mathbf{E}|T_b - \mathbf{E}(T_b)|$$

if $\theta \neq 1$; see [1, 3, 15].

Results on the large components of logarithmic combinatorial objects may also be shown. Let $L_i(n)$ be the size of the ith largest component of an instance of a logarithmic combinatorial object of size n with $L_i(n) = 0$ if the instance has fewer than i cycles. Then,

$$n^{-1}(L_1(n), L_2(n), \ldots) \xrightarrow{d} \mathbf{Z} \text{ with } PD(\theta) \text{ distribution}$$

assuming that a certain local limit theorem holds, which does hold for logarithmic assemblies, multisets and selections.This and related results are contained in [1].

Theorem 1 shows that it is impossible for π_f to have the distribution of any random combinatorial object, for which any subset of the component counts are dependent. We note, however, that if Z_i are mutually independent, with $Z_i \sim$ Binomial$(\eta_k.2^{-k})$ as in Theorem 1, then the resulting combinatorial object with distribution given by (2) is logarithmic with $\theta = 1$. This observation follows from the estimate $\eta_k = k^{-1}2^k + O(2^{k/2})$.

3 Basic strings

It will be convenient to use the notation of strings. A string of length k (sometimes called a word) is defined to be a finite sequence $a_0a_1 \ldots a_{k-1}$, $k \geq 1$, $a_i \in \{0, 1\}$ in [9, 14], for example. We have $a_0a_1 \ldots a_{k-1} = b_0b_1 \ldots b_{k'-1}$ if and only if $k = k'$ and $a_i = b_i$ for all $0 \leq i \leq k-1$.

Definition 1 Given a string $A = a_0a_1 \dots a_{k-1}$, for each $1 \leq i \leq k-1$ we say that the string $a_i \dots a_{k-1}a_0 \dots a_{i-1}$ is a *circular shift* of A. A string not equal to any of its circular shifts is called *basic*. The set of basic strings of length k is denoted by $\mathcal{A}_k$.

Basic strings correspond to primitive necklaces [7]. In [6], basic strings are strings of beads of two colours, length k, and period k. In that paper, it is shown that

$$|\mathcal{A}_k| = k\eta_k, \tag{6}$$

with η_k given by (1), by Möbius inversion.

Lemma 1 *If A is a basic string, then so is every circular shift of A.*

Proof It is enough to observe that the inequalities $a_1 \dots a_{k-1}a_0 \neq a_i \dots a_{k-1}a_0 \dots a_{i-1}$ for all $2 \leq i \leq k$ follow immediately from $a_0 \dots a_{k-1} \in \mathcal{A}_k$ and then use induction. □

Definition 2 An integer $0 \leq p \leq t-1$, is called a *period* of string $x_0 \dots x_{t-1}$ if $x_i = x_{i+p}$ for $0 \leq i < t-p$. The smallest non-zero period of a string is called its *basic period*. If a string has no non-zero period, then its basic period is its length,

The significance of a string's period is that the string shifted right by p positions, and placed over the original copy, matches in the overlapping part.

Given y real and the set of integers $\mathbb{Z}$, define $\lfloor y \rfloor = \sup\{m \in \mathbb{Z} : m \leq y\}$. If $X = x_0 \dots x_{t-1}$ has period p, then it is easily shown that X can be written as the concatenation $X = A^lA^*$, where $A = x_0x_1 \dots x_{p-1}$, $l = \lfloor t/p \rfloor$, and $A^* = x_0x_1 \dots x_{t-lp-1}$ is a (possibly empty) prefix of A.

A connection between basic strings and basic period is contained in Lemma 2.

Lemma 2 *Let $X = A^lA^*$, where $|A| = k$, A^* is a prefix of A, and either $l \geq 2$ or $l = 1$ and $|A^*| = |A| - 1$. The string X has basic period $|A|$ if and only if A is basic.*

Proof Let $t = |X|$, $X = x_0 \dots x_{t-1}$, $A = a_0 \dots a_{k-1}$ and note that $x_s = a_{s \bmod k}$ for all $0 \leq s < t$. Suppose that A is basic. The string X has period k because $|A| = k$. If X had period $0 < p < k$, then $a_{s \bmod k} = x_s = x_{s+p} = a_{(s+p) \bmod k}$ for $0 \leq s < t-p$. We have $t-p > 2k-1-k = k-1$ and so Definition 1 is violated for $i = p$. Therefore, X has basic period k. If A is not basic, then $a_0a_1 \dots a_{k-1} = a_p \dots a_{k-1}a_0 \dots a_{p-1}$ for some $1 \leq p \leq k-1$ and so $x_{i+p} = a_{(i+p) \bmod k} = a_{i \bmod k} = x_i$ for $0 \leq i < t-p$, making p a period of X. □

To illustrate Lemma 2, let $A = 1100$ and $B = 1010$. The string A is basic and AA has basic period 4, but the string B is not basic and BB has basic period 2. The string $C = 101$ is basic and CC has basic period 3, but C has basic period 2. The string $D = 1101$ is basic, and the strings DD and 1101110 have basic period 4, but the string 11011 has basic period 3. The author has not been able to find a basic string A for which AA^*, $|A^*| = |A| - 2$, does not have basic period $|A|$.

4 Binomially distributed and independent cycle counts

In this section we will prove Theorem 1.

Consider the set $\mathcal{S}_{n,k}$ of sequences of length $k+1$, $k \leq n$, of elements of $\mathbb{F}_2^n$

$$(x_0, x_2, \ldots, x_{n-1}), (x_1, x_3, \ldots, x_n), \ldots, (x_k, x_{k+1}, \ldots, x_{k+n-1})$$

satisfying the conditions

$$(x_i, x_{i+1}, \ldots, x_{i+n-1}) \neq (x_0, x_1, \ldots, x_{n-1}) \text{ for all } 1 \leq i \leq k-1, \tag{7}$$

and

$$(x_k, x_{k+1}, \ldots, x_{k+n-1}) = (x_0, x_2, \ldots, x_{n-1}). \tag{8}$$

The conditions (7) and (8) are necessary for the sequence to be the vertices of a k-cycle. Such a sequence actually are the vertices of a k-cycle if, in addition,

$$\pi_f((x_i, x_{i+1}, \ldots, x_{i+n-1})) = (x_{i+1}, x_{i+2}, \ldots, x_{i+n}) \text{ for all } 0 \leq i \leq k-1. \tag{9}$$

Consider the string $X = x_0 x_1 \ldots x_{k+n-1}$ of length $k+n$ with x_i satisfying (7) and (8). Condition (8) says that k is a period of X. It follows that (9) can be replaced by

$$\pi_f((x_i, x_{i+1}, \ldots, x_{i+n-1})) = (x_{i+1}, x_{i+2}, \ldots, x_{i+n-1}, x_{i+n-k}) \text{ for all } 0 \leq i \leq k-1. \tag{10}$$

Moreover, $X = A^l A^*$ for $l = \lfloor (n+k)/k \rfloor \geq 2$ and $A^* = x_0 x_1 \ldots x_{n+k-lk-1}$. The string A must be basic, because otherwise by Lemma 2, X would have basic period $p < k$ and (7) would not hold for $i = p$. On the other hand, given $A \in \mathcal{A}_k$, the x_i, $0 \leq i \leq k+n-1$, given by $X = x_0 x_2 \ldots x_{k+n-1} = A^l A^*$ satisfy (7) and (8) because $n \geq k$. It follows that $|\mathcal{S}_{n,k}| = |\mathcal{A}_k| = k\eta_k$.

For a string $X = A^l A^*$ with $A \in \mathcal{A}_k$ obtained from (7) and (8), the string corresponding to the first vertex is $x_0 x_1 \ldots x_{n-1} = A^{l-1} A^*$. The string corresponding to the second vertex is $x_1 x_2 \ldots x_n = B^{l-1} B^*$, where $B = a_1 \ldots a_{k-1} a_0$ is a circular shift of A and is basic by Lemma 1. Continuing in this way, the strings corresponding to vertices in a possible k-cycle including the vertex $A^{l-1} A^*$ are of the form $B^{l-1} B^*$ where either B is a circular shift of A or $B = A$. Define an equivalence relation R on strings of the form $A^{l-1} A^*$, $A \in \mathcal{A}_k$, $|A^{l-1} A^*| = n$ by $A^{l-1} A^* \sim B^{l-1} B^*$ if either B is a circular shift of A or $B = A$. Let $\mathcal{C}_{n,k}$ denote the set of equivalence classes of such strings with respect to R. The number of equivalence classes is

$$|\mathcal{C}_{n,k}| = |\mathcal{S}_{n,k}|/k = \eta_k. \tag{11}$$

A vertex can only be in a k-cycle with the other $k-1$ vertices in their equivalence class.

It remains to examine condition (10). Choose any $\mathbf{c} \in \mathcal{C}_{n,k}$, $1 \leq k \leq (n-1)/3$, and any string $x_i \ldots x_{i+n-1} \in \mathbf{c}$, which corresponds to vertex $(x_i, \ldots, x_{i+n-1})$. In order for

(10) to be satisfied at this vertex, we must have $x_i \oplus f(x_{i+1}, x_{i+2}, \ldots, x_{i+n-1}) = x_{i+n-k}$ or, equivalently,

$$f(x_{i+1}, x_{i+2}, \ldots, x_{i+n-1}) = x_i \oplus x_{i+n-k}, \tag{12}$$

which happens with probability $1/2$. We will show that it is impossible that the vertex

$$(\overline{x_i}, x_{i+1}, \ldots, x_{i+n-1}) = (x_i \oplus 1, x_{i+1}, \ldots, x_{i+n-1}) \tag{13}$$

is in $\mathbf{c}' \in \mathcal{C}_{n,k'}$, for some $1 \le k' \le (n-1)/3$. Note that the vertex (13) is the only vertex besides $(x_i, x_{i+1}, \ldots, x_{i+n-1})$ for which $f(x_{i+1}, \ldots, x_{i+n-1})$ is relevant.

The string $B = x_{i+1}x_{i+2}\ldots x_{i+k}$ is a circular shift of $A = x_i \ldots x_{i+k-1} \in \mathcal{A}_k$ and so $B \in \mathcal{A}_k$ by Lemma 1. Thus, we have the decomposition $\overline{x_i}x_{i+1}\ldots x_{i+n-1} = \overline{x_i}B^m B^\dagger$, where $B \in \mathcal{A}_k$, $m = \lfloor (n-1)/k \rfloor \ge 3$ and $|B^\dagger| = n - 1 - mk$. If $\overline{x_i}B^m B^\dagger$ belongs to $\mathbf{c}' \in \mathcal{S}_{n,k'}$, $k' \le (n-1)/3$, then $\overline{x_i}B^m B^\dagger = D^{m'}D^*$ for $D \in \mathcal{A}_{k'}$, $m' = \lfloor n/k' \rfloor$ and so k' is a period of $\overline{x_i}B^m B^\dagger$. From $B \in \mathcal{A}_k$ and $m \ge 2$, it follows that k is the basic period of $B^m B^\dagger$. However, $x_i = x_{i+k}$ is the last letter of B and therefore, the basic period of $\overline{x_i}B^m B^\dagger$ is greater than $(m-1)k$. We now obtain the contradiction

$$k' > (m-1)k \ge \left(\frac{n-1}{k} - 2\right) k = n - 2k - 1 \ge n - \frac{2(n-1)}{3} - 1 = \frac{n-1}{3},$$

where the last inequality follows from the restriction $k \le (n-1)/3$.

We conclude that the events

$$\{f(x_{i+1}, x_{i+2}, \ldots, x_{i+n-1}) = x_i \oplus x_{i+n-k}\}, \quad x_i x_{i+1}\ldots x_{i+n-1} \in \bigcup_{\mathbf{c} \in \mathcal{C}_{n,k}} \mathbf{c}$$

are mutually independent. Given $\mathbf{c} \in \mathcal{C}_{n,k}$, let $\{\mathbf{c} \in \pi_f\}$ denote the event that the vertices corresponding to the strings in $\mathbf{c}$ belong to a k-cycle in π_f. We have shown that the events $\{\mathbf{c} \in \pi_f\}$ are mutually independent and that $\mathbb{P}(\mathbf{c} \in \pi_f) = 2^{-k}$. Moreover, in view of (11),

$$S_k(n) = \sum_{\mathbf{c} \in \mathcal{C}_{n,k}} I[\mathbf{c} \in \pi_f] \sim \text{Binomial}(\eta_k, 2^{-k}),$$

where $I[\mathbf{c} \in \pi_f]$ is the indicator random variable for $\{\mathbf{c} \in \pi_f\}$, and the $S_k(n)$, $1 \le k \le (n-1)/3$ are mutually independent.

Acknowledgements

The author thanks the anonymous referees for their helpful reports and Richard Arratia for pointing out reference [7].

References

[1] R. Arratia, A. D. Barbour and S. Tavaré, Logarithmic combinatorial structures: a probabilistic approach, EMS Monographs in Mathematics, European Mathematical Society (EMS), Zürich, 2003.

[2] R. Arratia, E. R. Canfield and A. W. Hales, Random Feedback Shift Registers, and the Limit Distribution for Largest Cycle Lengths. Preprint: https://arxiv.org/abs/1903.09183.

[3] R. Arratia, D. Stark and S. Tavaré, Total variation asymptotics for Poisson process approximations of logarithmic combinatorial assemblies, *Ann. Probab.* **23** (1995), 1347–1388.

[4] R. Arratia and S. Tavaré, The cycle structure of random permutations, *Ann. Probab.* **20** (1992), 1567–1591.

[5] R. Arratia and S. Tavaré, Independent process approximations for random combinatorial structures, *Adv. Math.* **104** (1994), 90–154.

[6] E. A. Bender and J. R. Goldman, On the applications of Möbius inversion in combinatorial analysis, *Amer. Math. Monthly* **82** (1975), 789–803.

[7] D. Coppersmith, R. C. Rhoades and J. M. Vanderkam, Counting De Bruijn sequences as perturbations of linear recursions. Preprint: https://arxiv.org/abs/1705.07835.

[8] W. J. Ewens, The sampling theory of selectively neutral alleles, *Theoret. Population. Biol.* **3** (1972), 87–112.

[9] L. J. Guibas and A. M. Odlyzko, Periods in strings, *J. Combin. Theory Ser. A* **30** (1981), 19–42.

[10] S. W. Golomb, Shift Register Sequences, Holden Day, Inc., San Francisco, Calif.-Cambridge-Amsterdam, 1967. Portions co-authored by L.R. Welch, R.M. Goldstein and A.W. Hales.

[11] V. L. Goncharov, Some facts from combinatorics, *Izvestia Akad. Nauk SSSR, Ser. Mat.* **8** (1944), 3–48. See also: On the field of combinatory analysis, *Amer. Math. Soc. Transl. (2)* **19** (1962), 1–46.

[12] J. F. C. Kingman, Random discrete distributions, *J. Roy. Stat. Soc., Ser. B* **37** (1975), 1–22.

[13] J. F. C. Kingman, The population structure associated with the Ewens sampling formula, *Theoret. Population Biol.* **11** (1977), 274–283.

[14] D. Stark, First occurrence in pairs of long words: a Penney-ante conjecture of Pevzner, *Combin. Probab. Comput.* **4** (1995), 279–285.

[15] D. Stark, Total variation asymptotics for independent process approximations of logarithmic multisets and selections, *Random Structures Algorithms* **11** (1997), 51–80.

[16] A. M. Vershik and A. A. Schmidt, Limit measures arising in the asymptotic theory of symmetric groups. I, *Theory Probab. Appl.* **22** (1977), 70–85.

(Received 25 June 2022; revised 24 Apr 2023)