The Annals of Probability
1996, Vol. 24, No. 1, 49–73

CYCLE STRUCTURE OF RIFFLE SHUFFLES[1]

BY STEVEN P. LALLEY

Purdue University

A class of models for riffle shuffles ("f-shuffles") related to certain expansive mappings of the unit interval is studied. The main result concerns the cycle structure of the resulting random permutations in $\mathscr{S}_n$ when n is large. It describes the asymptotic distribution of the number of cycles of a given length, relating this distribution to dynamical properties of the associated mapping. This result generalizes a recent result of Diaconis, McGrath and Pitman.

1. Introduction. The cycle structure of a random permutation chosen from the uniform distribution on the permutation group $\mathscr{S}_n$ is reasonably well understood. When $n \to \infty$, the joint distribution of the "large cycles" is governed by "Poisson–Dirichlet" asymptotics (see [7] and [9]), and the number of "short cycles" of a given length j is approximately Poisson. Recently, Diaconis, McGrath and Pitman [3] obtained analogous results for random permutations from a class of nonuniform distributions, the so-called a-shuffles. These are random permutations obtained by cutting a deck of n cards into a packets and then randomly riffling the packets together. Using a bijection discovered by Gessel [4], they obtained an exact formula for the joint distribution of the numbers N_j, $j = 1, 2, \ldots, n$, of cycles of length j and used this formula to describe the asymptotics of both the large cycles and the small cycles as $n \to \infty$. The behavior of the large cycles is governed by the same Poisson–Dirichlet asymptotics that apply for random permutations from the uniform distribution. The asymptotic behavior of the short cycles, however, is markedly different than that of uniform random permutations: for large n and fixed j, the distribution of N_j is approximately the negative binomial distribution with parameters (f_{ja}, a^{-j}), where f_{ja} is the number of aperiodic "necklaces" of length j from an alphabet with a letters.

The purpose of this paper is to shed light on this last result by generalizing it to a larger class of random permutations, which we dub "f-shuffles." Here f is an expanding, piecewise C^2-mapping of the unit interval onto itself; the a-shuffle (where a is an integer greater than or equal to 2) is the special case where $f(x) = \{ax\}$ and $\{\ \}$ denotes fractional part. The random permutation is constructed as follows: n points are dropped randomly in the unit interval, according to the uniform distribution. These are then labelled $1, 2, \ldots, n$ in accordance with their relative order in $[0, 1]$. The n points $x_{(i)}$ are then mapped to the points $f(x_{(i)})$, and these are labelled $1, 2, \ldots, n$ in accordance

Received October 1994; revised February 1995.

[1] Supported by NSF Grant DMS-93-07855.

AMS 1991 *subject classifications*. 05A15, 60C05.

Key words and phrases. Riffle shuffle, random permutation, interval mapping.

with *their* relative order in $[0,1]$. The f-shuffle is the (random) permutation that takes each label i to the label of $f(x_{(i)})$ in the transformed sample. This construction was suggested by the dynamical representation of the a-shuffle in [1]. Another interesting special case (a biased riffle shuffle) is discussed in Example 2 in Section 2.2.

Our main result relates the cycle structure of the f-shuffle Π to the dynamics of the mapping f. We will show that, for large n, cycles of Π may be associated with periodic orbits of f in a natural way; in particular, the length of each cycle will agree with the minimal period of the associated periodic orbit. Then we will show the following: that, for each periodic orbit ω of f, the number of cycles of Π associated with ω is, for large n, approximately a geometric random variable with parameter $\sigma(\omega)$, where $\sigma(\omega)$, the "weight" of orbit ω, is the inverse of the product of the values of f' at the points of the orbit; and that these counts are approximately independent for different periodic orbits. The result of Diaconis, McGrath and Pitman [3] concerning the asymptotics of the "short" cycles follows as an immediate corollary.

Our proof, like that of [3], uses a form of "symbolic dynamics," using sequences from the alphabet $\{1, 2, \dots, a\}$ to represent the orbits of individual cards. The similarity ends there, however. For the a-shuffles studied in [3], the induced probability measure on $\mathscr{S}_n$ is uniform on a subset of $\mathscr{S}_n$, namely, the set of all permutations with no more than $a-1$ descents; consequently, questions about the distribution of cycles may be attacked by essentially combinatorial methods. In contrast, the probability measure induced by the f-shuffle for nonlinear f is *not* uniform (even approximately so) on a subset of $\mathscr{S}_n$.

Unfortunately, our methods do not lead to definitive results concerning the asymptotic distribution of large cycles. We conjecture that the joint distribution of the large cycles follows the same asymptotic behavior as for a-shuffles and completely random permutations. By [5], this would follow from the following apparently weaker conjecture.

CONJECTURE. *If k points are picked by random sampling without replacement from $\{1, \dots, n\}$, then for large n, with high probability, the permutation on S_k derived from the action of* Π *on the k sample points has a distribution that is close to uniform.*

Although we have not been able to establish this, we have proved a similar statement for the action of Π on k randomly chosen *neighboring* points.

2. *f*-shuffles.

2.1. *Definition.* Let $f\colon [0,1] \to [0,1]$ be a piecewise C^2, measure-preserving transformation of the unit interval. (Throughout the paper, the term *measure-preserving* refers to the Lebesgue measure on $[0,1]$. Sufficient conditions for a mapping f to be measure-preserving are given below.) The *f-shuffle* of a deck of n cards is the random permutation $\Pi \in \mathscr{S}_n$ obtained as follows: let $\xi_1, \xi_2, \dots, \xi_n$ be a random sample of size n from the uniform distribution on $[0,1]$, with order statistics $\xi_{(1)}, \xi_{(2)}, \dots, \xi_{(n)}$; let $\zeta_i = f(\xi_i)$ for $i = 1, 2, \dots, n$; and let $\zeta_{(1)}, \zeta_{(2)}, \dots, \zeta_{(n)}$ be the order statistics of the sample $\zeta_1, \zeta_2, \dots, \zeta_n$. Define Π to be the permutation such that

$$\zeta_{(\Pi(i))} = f(\xi_{(i)}) \qquad \forall\, i = 1, 2, \dots, n. \tag{1}$$

See Figure 1 for an example.

NOTE 1. By "piecewise C^2" we mean that the left and right derivatives exist everywhere, are finite and C^1, and agree except at those points where f is discontinuous.

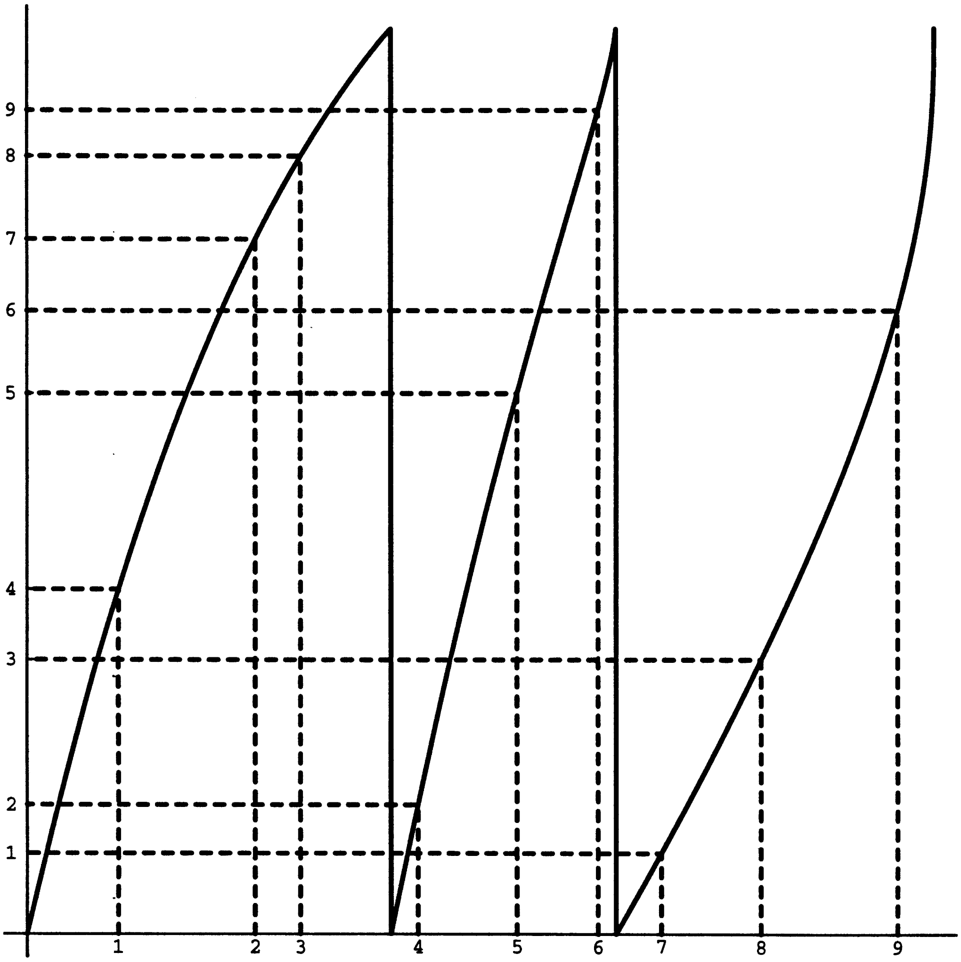

FIG. 1.

NOTE 2. There is no reason why the f-shuffle could not be similarly defined for mappings f that are *not* measure-preserving, provided, for example, that if U is a uniform-$[0,1]$ random variable, then $f(U)$ has a positive density. However, for any such f, there is a (unique) monotone increasing homeomorphism T of $[0,1]$ such that $T\circ f$ is measure-preserving. (This is an elementary consequence of the "quantile transform." In order that $T\circ f$ be measure-preserving, it is necessary and sufficient that $T(f(U))$ be uniformly distributed on $[0,1]$.) For this homeomorphism T, the $T\circ f$-shuffle is the same as the f-shuffle. Hence, there is no loss of generality in restricting attention to f-shuffles with f measure-preserving. Observe that when f is taken to be measure-preserving, the transformed sample $\zeta_1, \zeta_2, \dots, \zeta_n$ is distributed as a sample of n iid uniform random variables.

NOTE 3. In general, performing an f-shuffle and then an *independent* g-shuffle is *not* the same as performing a single $(g\circ f)$-shuffle, even if $g=f$, unless both f and g are piecewise-linear functions. Nevertheless, the dynamical system consisting of the iterates $f^{(n)}$ of f does contain information about the behavior of the random walk obtained by performing independent f-shuffles repeatedly (see [6]).

2.2. *Examples.*

EXAMPLE 1. For $f(x) = \{ax\}$, where $a \geq 2$ is an integer and $\{\ \}$ denotes fractional part, the f-shuffle coincides with the a-shuffle studied in [3]. This admits the following alternative description. Break the deck into a piles, with the vector of pile sizes having the multinomial distribution $\mathscr{M}(n, a^{-1}, a^{-1}, \dots, a^{-1})$. Then "riffle" the a piles together in an unbiased manner: drop cards one at a time from the bottoms of the piles, with the (conditional) probability that the next card comes from the ith pile being proportional to the number of cards remaining in that pile.

EXAMPLE 2. Let

$$f(x) = \begin{cases} H_p \circ M_{1/v} \circ L\left(\dfrac{p-x}{p}\right), & \text{for } x \in [0,p], \\[2ex] H_p \circ M_{1/u} \circ L\left(\dfrac{1-x}{1-p}\right), & \text{for } x \in (p,1], \end{cases}$$

where

$$\begin{aligned} L(y) &= -\log y, \\ M_\alpha(y) &= \alpha y, \\ H_\alpha(y) &= 1 - \alpha \exp(-vy) - (1-\alpha)\exp(-uy). \end{aligned}$$

To see that f is a measure-preserving mapping, consider its action on a uniform-$[0,1]$ random variable U. Conditional on $U \in [0,p]$, U is uniformly distributed on $[0,p]$, so $L(U)$ is a unit exponential random variable, and hence $M_{1/v}(L(U))$ is an exponential with mean $1/v$. Similarly, conditional on $U \in (p,1]$, $M_{1/u}(L(U))$ is an exponential with mean $1/u$. However, the mapping H_p is chosen precisely to take the (p,q)-mixture of mean $1/v$ and mean $1/u$ exponentials to a uniform-$[0,1]$. Thus, $f(U)$ has the uniform-$[0,1]$ distribution; that is, f is measure-preserving.

For this mapping f, the f-shuffle is the "(u,v)-weighted" riffle shuffle defined in [6]. In this shuffle the deck is divided into two stacks which are then riffled; however, this riffle is biased so that cards from one of the stacks are more likely to drop than cards from the other. Specifically, if at some stage of the riffle the top stack has A cards remaining and the bottom stack has B cards remaining, then the probability that the next card dropped comes from the top stack is $uA/(uA+vB)$. The division of the deck into top and bottom stacks is such that the number of cards in the top stack has the binomial-(n,p) distribution. See [6] for a proof that the (u,v)-weighted riffle shuffle is an f-shuffle with the mapping f given above.

2.3. *The inverse construction.* Our main result and the arguments required to prove it hold generally for f-shuffles such that f has a *Markov partition*. That is, there exists a finite partition $\mathscr{J}$ of $[0,1]$ into intervals $J_1, J_2, \ldots, J_a$ such that f maps each J_i monotonically onto the union of the intervals in some subset of $\mathscr{J}$ (possibly with some of the endpoints removed). For simplicity, we shall restrict our discussion to transformations f satisfying the following assumption.

ASSUMPTION 1. *The mapping $f\colon [0,1] \to [0,1]$ is a piecewise C^2 (Lebesgue) measure-preserving mapping for which there exists a partition $\mathscr{J}$ of the unit interval into a intervals $J_i = (x_i, x_{i+1}]$, with $0 = x_1 < x_2 < \cdots < x_a < x_{a+1} = 1$, such that f maps each J_i monotonically $\uparrow$ onto the unit interval, and such that f is C^2 in each (x_i, x_{i+1}).*

Assumption 1 implies that f is a-to-1 and that f^{-1} has a distinct branches f_i^{-1}, with f_i^{-1} mapping the open unit interval $(0,1)$ monotonically onto the interior of J_i. By our convention concerning C^2-mappings f of the unit interval, the derivative f' is well defined and continuous at every x except those x such that $f(x)=0$ or $f(x)=1$. A necessary and sufficient condition for a piecewise C^2 function f satisfying Assumption 1 to be measure-preserving is that, for almost every $y \in [0,1]$,

$$\sum_{x \in f^{-1}(y)} \frac{1}{f'(x)} = 1. \tag{2}$$

Observe that since the inverse image $f^{-1}(y)$ contains $a > 1$ points and since f' is piecewise continuous and bounded above, (2) implies that $\inf_{x\in[0,1]} f'(x) > 1$, that is, that f is an *expanding* map of the unit interval.

The function f has a multivalued inverse function f^{-1} with a distinct branches f_i^{-1}, $i \in [a]$. Here and throughout the paper $[a] = \{1,2,\ldots,a\}$.

Each branch f_i^{-1} maps the unit interval homeomorphically onto the corresponding interval J_i. By (2), for almost every y the numbers $1/f'(f_i^{-1}(y))$ define a probability distribution on $[a]$. These distributions allow one to "invert" the measure-preserving transformation f, in the following sense. Let Y be uniformly distributed on $[0,1]$ and, conditional on $Y = y$, let $W = f_I^{-1}(y)$, where $I \in [a]$ is chosen at random according to the distribution $1/f'(f_i^{-1}(y))$, $i \in [a]$. Then W is also uniformly distributed on $[0,1]$, and $f(W) = Y$. The proof is straightforward.

THE INVERSE CONSTRUCTION. A version of the random permutation Π may be constructed as follows. Let $\zeta_1, \zeta_2, \ldots, \zeta_n$ be a random sample of size n from the uniform distribution on $[0,1]$. For each i, choose one of the a branches of f^{-1} according to the discrete probability measure (2), with $y = \zeta_i$ (conditional on the values $\zeta_1, \zeta_2, \ldots, \zeta_n$, these choices are to be made independently). Let the resulting collection of points $f_j^{-1}(\zeta_i)$ be $\xi_{(1)} \leq \xi_{(2)} \leq \cdots \leq \xi_{(n)}$. Define Π to be the unique permutation satisfying (1).

That Π is a version of the f-shuffle defined earlier follows easily from the fact that for Y uniformly distributed on $[0,1]$, $W = f_I^{-1}(Y)$ is also uniform-$[0,1]$. The independence of the branch choices β_j (conditional on the sample $\zeta_1, \zeta_2, \ldots, \zeta_n$) in this construction will be of crucial importance.

Although (1) involves the order statistics $\zeta_{(j)}$, the permutation Π is actually a function only of the branch choices β_j. The inverse permutation Π^{-1} may be described directly in terms of the sequence $\beta = \beta_1, \beta_2, \ldots, \beta_n$ as follows (this was known to Reeds [8] in the case of the a-shuffle). Think of the integers $1, 2, \ldots, n$ as the labels of cards, stacked in increasing order (top card equals 1, bottom card equals n). Permute the cards as follows: first, remove all the cards j such that $\beta_j = a - 1$ and move them, *in order*, to the top; then remove all cards j such that $\beta_j = a - 2$, and move them, in order, to the top (just above the stack of cards with $\beta_j = a - 1$) and so on; then, finally, remove all cards labelled $\beta_j = 1$ and move them to the top. The resulting permutation of the cards $1, 2, \ldots, n$ is Π^{-1}.

In the arguments below, it will be necessary to know the effect on Π of switching two adjacent entries of the sequence $\beta = \beta_1, \beta_2, \ldots, \beta_n$.

LEMMA 1. *Let $j' = j+1$ for some $1 \leq j < n$, and suppose that $\beta_j \neq \beta_{j'}$. Let $\beta' = (\beta'_1, \beta'_2, \ldots, \beta'_n)$ be the sequence obtained from β by switching the entries β_j and $\beta_{j'}$, and let Π' be the permutation determined by the sequence β'. Then*

$$\Pi' = \Pi \sigma_j, \tag{3}$$

where σ_j is the permutation that transposes j and j'.

NOTE 4. In particular, (3) implies the following relation between Π and Π': (a) if j and j' are in different cycles $C = (j, \ldots, j_*)$ and $C' = (j', \ldots, j'_*)$ of Π, then the switch merges them into a single cycle $C'' = (j, \ldots, j_*, j', \ldots, j'_*)$ of

Π' and leaves all other cycles unchanged; and (b) if j and j' are in the same cycle $C'' = (j, \dots, j_*, j', \dots, j'_*)$ of Π, then the switch splits it into the two cycles $C = (j, \dots, j_*)$ and $C' = (j', \dots, j'_*)$ of Π' and leaves all other cycles unchanged.

PROOF. Compare the steps of the construction of Π^{-1} and $(\Pi')^{-1}$. Cards labelled $a-1$ are removed and moved to the top; then cards labelled $a-2$ are removed and moved to the top and so on. The *same* cards are removed at each step for the label sequences β and β' *except* for cards j and j'. Since $j' = j+1$, the final positions of cards j and j' are switched, and all other cards are moved to the same positions. □

3. Cycles and periodic orbits of *f*.

3.1. *First-order asymptotics.* We will need to have a rough idea of what an f-shuffle "looks like" for large n. Let Π be the random permutation generated by the f-shuffle of a deck of size n, and let $F = F_\Pi$ be the right-continuous step function with jumps at the points i/n such that

$$F_\Pi(i/n) = \Pi(i)/n. \tag{4}$$

Note that the action of Π is mirrored in the dynamics of F_Π. In particular, the cycles of Π correspond to the periodic orbits of F_Π: $(j_1, j_2, \dots, j_l)$ is a cycle of Π iff $(n^{-1}j_1, n^{-1}j_2, \dots, n^{-1}j_l)$ is a periodic orbit of F_Π with minimal period l.

LEMMA 2. *As $n \to \infty$, the random function F_Π converges in probability to f in the Skorohod topology.*

PROOF. This is a routine consequence of the Glivenko–Cantelli theorem and the fact that f is measure-preserving and continuous at all but those points of $[0,1]$ where $f(x) = 0$ or 1. (The reason for using the Skorohod topology rather than the sup norm topology is that f has points of discontinuity.) □

The following corollary is an immediate consequence.

COROLLARY 1. *For each $j \geq 1$, the jth iterate $F_\Pi^{(j)}$ of F_Π converges in probability to the jth iterate $f^{(j)}$ of f in the Skorohod topology.*

It is clear that f and F_Π may differ by $O(1)$ near the endpoints of the intervals J_i (the points of discontinuity of f). In the interiors of these intervals, however, f and F_Π are uniformly close, with high probability as $n \to \infty$.

LEMMA 3. *Let I be a closed interval contained in the interior of some J_i. Then for any $\varepsilon > 0$ and any $p > 0$,*

$$P\Big\{\sup_{x\in I} |F_\Pi(x) - f(x)| > n^{1/2+\varepsilon}\Big\} < n^{-p},$$

for sufficiently large n.

PROOF. This follows by standard estimates on large-deviations probabilities for sums of iid Bernoulli r.v.'s. In fact the probability in question decays exponentially in $n^{1/2-\varepsilon}$. □

Now consider the lth iterates f^l and F^l_Π of f and F_Π. The function f^l has a^l intervals of monotonicity, each of which it maps homeomorphically onto the whole unit interval. These intervals may be labelled

$$J_{i_1 i_2 \cdots i_l} = f^{-l}_{i_1 i_2 \cdots i_l}((0,1)),$$

where $f^{-l}_{i_1 i_2 \cdots i_l} = f^{-1}_{i_1} \circ \cdots \circ f^{-1}_{i_l}$.

COROLLARY 2. *Let I be a closed interval contained in the interior of some $J_{i_1 i_2 \cdots i_l}$. Then, for any $\varepsilon > 0$ and any $p > 0$,*

$$P\Bigl\{\sup_{x\in I} |F^l_\Pi(x) - f^l(x)| > n^{1/2+\varepsilon}\Bigr\} < n^{-p},$$

for sufficiently large n.

This follows directly from the preceding lemma.

NOTE 5. It may also be shown by standard methods that, for points of continuity $t \in (0,1)$ of f and $f^{(j)}$, $\sqrt{n}(F_\Pi(t) - f(t))$ and $\sqrt{n}(F^{(j)}_\Pi(t) - f^{(j)}(t))$ are asymptotically normal. This will not be needed, however.

NOTE 6. For large n, the increments of nF_Π for arguments in a neighborhood [distance $o(1)$] of t, where t is a point of continuity of f, are approximately independent geometric-plus-1 r.v.'s with parameter $1/f'(t)$. This follows from the "inverse construction": if $t = f^{-1}_i(y)$, then the increments of nF_Π near t are just the numbers of successive order statistics $\zeta_{(j)}$ between "successes," a success being a choice of branch i of the inverse function f^{-1}; in a neighborhood of y, the probability of a success for any $\zeta_{(j)}$ is approximately $1/f'(t)$.

3.2. *Cycles and signatures.* As explained in [3, Section 2], there is a natural mapping from the set $[a]^n$ of sequences with values in $[a] = \{1, 2, \ldots, a\}$ onto the set of permutations in $\mathscr{S}_n$ with at most $a - 1$ "descents" (π is said to have a descent at i if $\pi(i) > \pi(i+1)$: see [3]). Observe that any permutation arising from the f-shuffle of a deck of n cards has at most $a - 1$ descents, because the "cards" corresponding to the points ξ_j falling in any one of the intervals J_i must remain in order, as f is increasing in J_i. For a given sequence $x = x_1 x_2 \cdots x_n \in [a]$, the *inverse* of the corresponding permutation $\pi = \pi_x$ is obtained as follows. Start with an ordered deck of n cards, with the cards marked $1, 2, \ldots, n$. Put an additional mark on each card; card i is marked with the symbol x_i. Now remove all cards marked 1 and put them *in order* at the top; then remove all cards marked 2 and put them in order below the cards marked 1 and so on. For a permutation π with *exactly* $a - 1$ descents, there is a *unique* sequence $x \in [a]^n$ mapped to π.

LEMMA 4. *As the deck size $n \to \infty$, the probability that Π has fewer than $a - 1$ descents converges to zero.*

PROOF. In order that the f-shuffle Π have fewer than $a - 1$ descents, it is necessary that for some pair of intervals $J_j, J_{j'}$ in the partition $\mathscr{J}$,

$$\max_{\xi_i \in J_j} f(\xi_i) < \min_{\xi_i \in J_{j'}} f(\xi_i). \tag{5}$$

However, f maps each of $J_j, J_{j'}$ onto the whole unit interval, so by Lemma 2, (5) can occur only with vanishingly small probability as $n \to \infty$. □

Any permutation $\pi \in \mathscr{S}_n$ can be written as a product of *cycles*: $\pi = (C_1)(C_2)\cdots(C_k)$, where C_1 is the orbit of the first card $[C_1 = (1, \pi(1), \pi(\pi(1)), \ldots)]$, C_2 is the orbit of the first card not in cycle C_1 and so on. For those permutations π with exactly $a-1$ descents, we will assign to each cycle C_i an aperiodic sequence of length $|C_i|$ from the alphabet $[a]$; this sequence will be called the *signature* of the cycle. The assignment of signatures to cycles is as in [3, Section 3] (there they are called cycle words). In geometric terms, the cycle signatures of a permutation π with exactly $a - 1$ descents may be defined as follows:

Let F_π be as in (4). If π has exactly $a - 1$ descents, then F_π is piecewise increasing with a intervals $J_1^\pi, J_2^\pi, \ldots, J_a^\pi$ of monotonicity. (For $\Pi = \pi$ these intervals will closely approximate the intervals J_i of the partition $\mathscr{J}$, by Lemma 2.) For a given cycle C_i, let j be the smallest integer of the cycle; the signature of C_i is the sequence in $[a] = \{1, 2, \ldots, a\}$ recording the indices of the intervals J_i^π visited by $j/n, F_\pi(j/n), F_\pi(F_\pi(j/n)), \ldots$.

The signature of a cycle is always a *Lyndon word* (a Lyndon word ω of length l is an aperiodic sequence in $[a]^l$ that is lexicographically smaller than each of its $l - 1$ cycle shifts), and all Lyndon words of length less than or equal to n may occur as cycle signatures in Π (see [3], Section 3). Observe that Lyndon words of length l correspond naturally to aperiodic necklaces of length l: for each Lyndon word, one may obtain an aperiodic necklace by "fastening" the ends of the word; for each aperiodic necklace, one may recover a Lyndon word by "unfastening" it at the unique point where the resulting word is lexicographically smaller than each of the other $l - 1$ possible words obtained by unfastening the necklace at some other point.

3.3. *Signatures*; *periodic orbits*; *and weights.* Let f satisfy Assumption 1. Every periodic point $x \in [0, 1]$ of f of (minimal) period $l \geq 1$ has an *itinerary*, namely, the sequence in $[a]^l$ giving the indices of the intervals J_j visited by $x, f(x), f(f(x)), \ldots, f^{(l)}(x)$. The itineraries of the other points $f^{(j)}(x)$ in the orbit are just the cyclic shifts of the itinerary of x, and the lexicographic order of these itineraries agrees with the usual $[0, 1]$-order of the points in the orbit. (For an elementary discussion of the "symbolic dynamics" of orbits of mappings f of the unit interval, see [2], Chapter 1.) Each periodic orbit of f has a minimal point; the itinerary of this point will be called the *signature* of the orbit.

LEMMA 5. *Every Lyndon word is the signature of a unique periodic orbit of f.*

PROOF. This is a special case of a standard result for expansive mappings of the unit interval with a Markov partition. The argument, in brief, is as follows:

For each finite sequence $\mathbf{x} = x_1x_2\cdots x_n \in [a]^n$, define $J_{\mathbf{x}}$ to be the set of all points $y \in [0,1]$ whose "itinerary" begins $x_1x_2\cdots x_n$ [the *itinerary* of a point y is the sequence of indices j of the intervals J_j visited by $y, f(y), f(f(y)), \ldots$]. Then, for each $\mathbf{x}$, $J_{\mathbf{x}}$ is an interval of positive length which is mapped monotonically onto $[0,1]$ by $f^{(n)}$, where n is the length of $\mathbf{x}$. This may be seen by induction on the length of $\mathbf{x}$, using the fact that f maps each of the intervals J_i in the partition $\mathscr{J}$ *onto* the entire unit interval. Moreover, the intervals $J_{\mathbf{x}}$ are nested; that is, if $\mathbf{x}'$ is an extension of $\mathbf{x}$, then $J_{\mathbf{x}}$ contains $J_{\mathbf{x}'}$. Finally, for any infinite sequence $x_1x_2\cdots$ in $[a]^\infty$, the lengths of the intervals $J_{x_1x_2\cdots x_n}$ shrink to 0 as $n \to \infty$ because the derivative of $f^{(n)}$ converges uniformly to ∞ as $n \to \infty$, as f' is bounded away from 1 on $[0,1]$. Thus, for each itinerary $\mathbf{x} = x_1x_2\cdots$, there is a *unique* point $y \in [0,1]$ with itinerary $\mathbf{x}$.

Now let $\mathbf{w} = w_1w_2\cdots w_n$ be a Lyndon word of length n, and let $\mathbf{x} = x_1x_2\cdots$ be the infinite sequence obtained by concatenating $\mathbf{w}$ with itself infinitely many times. By the preceding paragraph, there is a unique point $y \in [0,1]$ with itinerary $\mathbf{x}$. Since $\mathbf{x}$ is periodic with minimal period equal to the length of $\mathbf{w}$, y is a periodic point of f, and its orbit has signature $\mathbf{w}$. Clearly, the orbit of y is the *only* periodic orbit with signature $\mathbf{w}$, because y is the only point with itinerary $\mathbf{x}$. □

For a given Lyndon word (cycle signature) σ, let $\mathscr{O}_\sigma = (x_1, x_2, \ldots, x_k)$ be the unique periodic orbit of f with signature σ, written so that x_1 is the minimal point of the orbit. For a cycle $C = (i_1, i_2, \ldots, i_k)$ of Π, written so that i_1 is the minimal entry, let $\mathscr{O}^*_C = (i_1/n, i_2/n, \ldots, i_k/n)$ be the corresponding periodic orbit of F_Π. If $\mathscr{O}_\sigma$ and $\mathscr{O}^*_C$ have the same length, define the distance

$$d(\mathscr{O}_\sigma, \mathscr{O}^*_C) = \max_{1\le j\le k} |x_j - i_j/n|.$$

An immediate consequence of Lemma 2 is the following corollary.

COROLLARY 3. *Fix a Lyndon word σ, and let $\mathscr{O}_\sigma$ be the unique periodic orbit with signature σ. For each $\varepsilon > 0$, the probability that there is a cycle C of Π with signature σ satisfying $d(\mathscr{O}_\sigma, \mathscr{O}^*_C) \ge \varepsilon$ converges to 0 as $n \to \infty$.*

Let $\mathbf{x} = x_1, x_2, \ldots, x_l$ be a periodic orbit of f $[f(x_i) = x_{i+1}, f(x_l) = x_1]$ with minimal period l. Define the *weight* of $\mathbf{x}$ by

$$\omega = \omega(\mathbf{x}) = \left\{\prod_{i=1}^{l} f'(x_i)\right\}^{-1}. \tag{6}$$

[The derivative f' exists at every $x \in [0,1]$ except those x such that $f(x) = 0$ or 1, so $\omega(x)$ is well defined for every periodic orbit $\mathbf{x}$ except the orbits $\{0\}$ and $\{1\}$. At $x = 0$ use the left derivative, and at $x = 1$ use the right derivative.] The weight (more precisely, its inverse) of a periodic orbit is a natural and important "dynamical" quantity, measuring the degree of expansion in the vicinity of a point of the orbit after one cycle.

Since there is a one-to-one correspondence between periodic orbits and cycle signatures (by the previous lemma) we may define the *weight* $\omega(\sigma)$ of a cycle signature σ to be the weight of the unique periodic orbit with signature σ.

EXAMPLE 3. Consider the 2-shuffle. The associated mapping f of $[0,1]$ is $x \to 2x \bmod 1$. For each Lyndon word x of length l, there is a periodic orbit of f with signature x (and, consequently, period l). The weight of this orbit is 2^{-l}.

EXAMPLE 4. Consider the (u,v)-weighted riffle shuffle with $p = 0.3$, $v = 1$ and $u = 2.5$. The associated mapping f of $[0,1]$ is given in Section 2.2. There are two periodic orbits with period 3: one with signature 001, the other with signature 011. The orbit with signature 001 is

$$0.007437 \to 0.050019 \to 0.306346 \to 0.007437$$

(accurate to five decimal places); its weight is 0.0250127. The orbit with signature 011 is

$$0.122844 \to 0.635265 \to 0.404126 \to 0.122844;$$

its weight is 0.231448. This shows that in general the weight of a periodic orbit is *not* a function of just its period.

4. Short cycle asymptotics: the main result. For a given cycle signature σ, define $N_\sigma = N_\sigma(\Pi)$ to be the number of cycles of Π with signature σ. Note that as $n \to \infty$ the probability that $N_\sigma(\Pi)$ is not defined converges to 0, by Lemma 4.

THEOREM 1. *For each Lyndon word (cycle signature) σ, the distribution of $N_\sigma(\Pi)$ converges weakly as $n \to \infty$ to the geometric distribution with parameter $\omega = \omega(\sigma)$; that is, for each $k = 0, 1, 2, \dots$,*

$$\lim_{n\to\infty} P\{N_\sigma(\Pi) = k\} = (1-\omega)\omega^k. \tag{7}$$

Moreover, for any fixed finite collection of Lyndon words $\sigma_1, \sigma_2, \dots, \sigma_r$, the joint distribution of the counts N_{σ_i} converges as $n \to \infty$ to the product of the appropriate geometric distributions; that is, the random variables N_{σ_i} are asymptotically independent.

NOTE 7. In the special case of an a-shuffle, this was proved in [3].

COROLLARY 4. *For each* $j = 1, 2, \ldots,$ *let* M_j *be the number of cycles of* Π *with length* j. *Then, for each* $k \geq 1$, *as* $n \to \infty$ *the random variables* $M_1, M_2, \ldots, M_k$ *are asymptotically independent, and, for each* $j \geq 1$, *the asymptotic distribution of* M_j *is that of a sum of independent geometric r.v.'s with parameters* $\omega(\sigma)$, *one for each Lyndon word* σ *of length* j.

This is an immediate consequence of Theorem 1.

The rest of the paper will be devoted to the proof of Theorem 1. The cycle signatures $\sigma = 1$ and $\sigma = a$ must be treated separately, because the mapping f is discontinuous at the periodic points 0 and 1 with these signatures.

PROOF OF THEOREM 1 (for $\sigma = 1, a$). Consider the case $\sigma = 1$; this is the signature of the periodic orbit 0. In order that Π have a cycle of signature 1, the top card of the deck must remain on top; that is, the preimage of the first order statistic $\zeta_{(1)}$ must be in J_1^{Π}. The number N_1 of cycles with signature 1 is the number of cards at the top of the deck that retain their positions after the shuffle. Equivalently, N_1 is the maximum integer k such that the preimages of the first k order statistics $\zeta_{(1)}, \zeta_{(2)}, \ldots, \zeta_{(k)}$ all lie in J_1^{Π}. When n is large, the intervals J_1^{Π} and J_1 are with high probability nearly the same, by Lemma 2. Consequently, with high probability, N_1 is the same as the maximum integer k such that the preimages of the first k order statistics $\zeta_{(1)}, \zeta_{(2)}, \ldots, \zeta_{(k)}$ all lie in J_1.

Conditional on the values of the first k order statistics $\zeta_{(i)}$, the probability that their preimages all lie in J_1 is $\prod_{i=1}^{k}(1/f'(f_1^{-1}(\zeta_{(i)})))$, where f_1 denotes the first branch of the inverse function f^{-1}. With high probability, all of the first k order statistics $\zeta_{(i)}$ are very close to 0, so all of the preimages $f_1^{-1}(\zeta_{(i)})$ are also close to 0. Hence, by the continuity of f', $\prod_{i=1}^{k}(1/f'(f_1^{-1}(\zeta_{(i)}))) \approx f'(0)^{-k}$. However, $\omega(\sigma) = f'(0)^{-1}$. Therefore,

$$\lim_{n \to \infty} P\{N_1 \geq k\} = \omega(1)^k.$$

A similar argument applies for the signature $\sigma = a$. □

The proof of Theorem 1 will be given in Section 6, following some auxiliary results concerning inhomogeneous multinomial processes in Section 5. In the remainder of this section, we will give an asymptotic characterization of the random variable N_σ and a heuristic argument for (7). Let σ be an arbitrary cycle signature of length l, but $\sigma \neq 1, a$. Recall (Lemma 5) that there is a unique periodic orbit $x = x_1 x_2 \cdots x_l$ of f with signature σ. Here $x_1, x_2, \ldots, x_l$ are the points of the orbit, ordered so that x_1 is the least element and so that $f(x_i) = x_{i+1}$ for each $i = 1, 2, \ldots, l-1$. Thus, x_1 is a fixed point of the mapping $f^{(l)}$, but is not fixed by $f^{(j)}$ for any $j < l$. Fix $\varepsilon > 0$ (small), and define $N_\sigma^\varepsilon = N_\sigma^\varepsilon(\Pi)$ to be the number of solutions $x_1' = j/n$, $j \in \mathbb{Z}$, of

$$F_{\Pi}^{(l)}(x_1') = x_1' \tag{8}$$

such that $|x_1' - x_1| < \varepsilon$.

LEMMA 6. *For each cycle signature $\sigma = r_1 r_2 \cdots r_l$, there exists $\varepsilon = \varepsilon_\sigma > 0$ sufficiently small that*

$$\lim_{n\to\infty} P\{N_\sigma \neq N_\sigma^\varepsilon\} = 0.$$

PROOF. Consider $C = (j_1, j_2, \ldots, j_l)$, where $j_1 < j_i$ for $i = 2, 3, \ldots, l$; set $x'_k = j_k/n$ and $x'_{l+1} = x'_1$. Then C is a cycle of Π with signature σ iff

$$(9) \qquad F_\Pi(x'_k) = x'_{k+1} \qquad \forall\, 1 \le k \le l$$

and

$$(10) \qquad x'_k \in J^\Pi_{r_k} \qquad \forall\, 1 \le k \le l.$$

By Corollary 3, if C is a cycle of Π with signature σ, then for *any* $\varepsilon > 0$ the probability that $\max_{1\le i\le l} |x_i - x'_i| \ge \varepsilon$ converges to 0 as $n \to \infty$; consequently, $x'_1 = j_1/n$ is counted in N^ε_σ.

It remains to be shown that, for sufficiently small $\varepsilon > 0$, if j/n is a solution of (8) such that $|x_1 - j/n| < \varepsilon$, then j is the smallest element of a cycle C of Π with signature σ. Clearly, if (8) holds, then j is contained in a cycle whose length divides l, so (9) will hold. Take $\delta > 0$ so small that, for each point x_i of the periodic orbit x, the distance to the nearest endpoint of one of the intervals J_j is at least 2δ. (Recall that J_j are the intervals of monotonicity of f.) By Corollary 1 and the continuity of f at $x_1, x_2, \ldots, x_l$, there exists $\varepsilon > 0$ sufficiently small that if $|j/n - x_1| < \varepsilon$, then $|F^{(i)}_\Pi(j/n) - x_i| < \delta$, for all $j = 2, 3, \ldots, l$, with probability approaching 1 as $n \to \infty$, and consequently,

$$F^{(k)}_\Pi(j/n) \in J_{r_k},$$

for all $k = 1, 2, \ldots, l$. Since, by Lemma 2, the intervals J^Π_k closely approximate the corresponding intervals J_k with high probability as $n \to \infty$, it follows that if $|j/n - x_1| < \varepsilon$, then with high probability (10) will hold. Consequently, if j/n is a solution of (8) such that $|x_1 - j/n| < \varepsilon$, then with high probability j is the smallest element of a cycle of Π with signature σ. □

Using the characterization in Lemma 6, we may now give a heuristic argument for (7). For simplicity, consider the case of a cycle signature $\sigma = i$ of length 1 for some $i \neq 1, a$. (Note that this case does not occur when $a = 2$.) By Lemma 5, there is a unique fixed point x_* of f in the interior of J_i. By Lemma 6, N_σ is with high probability equal to the number N^ε_σ of solutions $j \in \mathscr{Z}$ of

$$(11) \qquad F_\Pi(j/n) = j/n$$

such that $|x_* - j/n| < \varepsilon$ for some sufficiently small $\varepsilon > 0$. Recall that in a neighborhood of x_* the increments of nF_Π are approximately iid geometric (plus 1) r.v.'s with parameter $1/f'(x_*) = \omega(\sigma)$ (see Note 6, in Section 3.1). Condition on the value $nF_\Pi(x_-)$ at a point $x_- < x_*$ within ε of x_* but distant enough that $nx_- - nF_\Pi(x_-) \overset{\Delta}{=} K > 0$ with high probability; then, conditional on K, the process

$$S_j \overset{\Delta}{=} n(F_\Pi(j/n) - F_\Pi(x_-)) - j, \qquad nx_- < j < nx_- + o(n),$$

is approximately a random walk with geometric-(ω) increments. Now the number of solutions N_σ^ε of (11) near x_* is just the number of visits to the site K by the random walk S_j, $j > nx_-$. Since S_j is approximately a random walk with geometric increments, the number of such visits is approximately geometric with parameter ω, by standard renewal theory for random walks with iid geometric increments.

This argument is not rigorous, because we have not shown that the increments of nF_Π are (approximately) iid geometric *conditional on* $F_\Pi(x_-)$. The difficulty is that conditioning on $F_\Pi(x_-)$ places constraints on the number(s) of branch choices of the different types $i = 1, 2, \dots, a$ at the various order statistics $\zeta_{(j)}$. A rigorous argument requires an extension of the standard "renewal theory" to pinned, inhomogeneous multinomial processes. This is done in the following section.

5. A renewal theorem for nearly exchangeable processes.

In this section we shall prove a renewal-type theorem for multinomial processes that are "nearly exchangeable," in a sense to be made precise below. Our results will be formulated as limit theorems for a triangular array of multinomial processes whose rows are indexed by $n = 1, 2, \dots$. The dependence of the various parameters and random variables on n will be suppressed. Let $N = N_n$ be integers such that $N \to \infty$ as $n \to \infty$; for each n, let

$$T = (T_1, T_2, \dots, T_N)$$

be random variables valued in the finite set $[\tau] = \{1, 2, \dots, \tau\}$. Here $\tau \geq 2$ is a fixed integer not depending on n. We will sometimes refer to the elements of $[\tau]$ as *types*. We do not assume that the random variables T_i in a given row are independent or identically distributed.

For each $i \in [\tau]$ and $m = 1, 2, \dots, N$, define

$$\begin{aligned} S_m^i &= \#\{j \leq m\colon\ T_j = i\}, \\ S_m &= (S_m^1, S_m^2, \dots, S_m^\tau). \end{aligned} \tag{12}$$

Here # denotes cardinality.

HYPOTHESIS 1. There exists a probability distribution $\pi = (\pi_j)_{j \in [\tau]}$ on $[\tau]$ (*not* depending on n) such that $\pi_j > 0$ for each $j \in [\tau]$ and such that, as $n \to \infty$,

$$\frac{S_N}{N} \to_P \pi.$$

Elements $t \in [\tau]^N$ will be called *configurations*. Say that configurations t and t' are *neighbors* if t' can be obtained from t by switching two entries t_i and t_j. Let $\mathscr{N}$ be the set of all pairs (t, t') that are neighbors. For each configuration t, define

$$\lambda(t) = P\{T = t\}.$$

HYPOTHESIS 2. As $n \to \infty$,

$$\max_{\mathscr{N}} \left| \frac{\lambda(t)}{\lambda(t')} - 1 \right| \to 0.$$

This is a weaker hypothesis than exchangeability; we will refer to it as *near exchangeability*. Observe that it does not require that the random variables $T_1, T_2, \ldots, T_N$ be identically distributed. Moreover, it does not require that the ratios $\lambda(t')/\lambda(t)$ be close to 1 for arbitrary pairs of configurations.

Let $Y_1, Y_2, \ldots$ be independent, identically distributed $[\tau]$-valued random variables each with distribution π. For any random vector V, let $\mathscr{D}(V)$ denote its distribution, and let $\mathscr{D}(V|F)$ denote its conditional distribution given F. Let d_{TV} denote the total variation distance between probability distributions. For any sequence $W_1, W_2, \ldots$ of random variables let $W_{[m,m']}$ denote the vector $(W_m, W_{m+1}, \ldots, W_{m'})$.

PROPOSITION 1. *Under Hypotheses* 1 *and* 2, *for each* $\varepsilon > 0$ *and each* $k \geq 1$, *as* $n \to \infty$,

$$\max_{1 \leq m \leq N-k} d_{\text{TV}}(\mathscr{D}(T_{[m+1,m+k]} \mid S_N), \mathscr{D}(Y_{[1,k]})) \to_P 0 \tag{13}$$

and

$$\max_{1 \leq m \leq (1-\varepsilon)N} d_{\text{TV}}(\mathscr{D}(T_{[m+1,m+k]} \mid S_N; T_{[1,m]}), \mathscr{D}(Y_{[1,k]})) \to_P 0. \tag{14}$$

PROOF. Fix integers $j, j' \in \{1, 2, \ldots, N\}$, and define a new configuration T' by switching the j, j' entries in T with conditional probability (given $T = t$)

$$\min\left(1, \frac{\lambda(t')}{\lambda(t)}\right),$$

where t' is the configuration obtained by switching the jth and j'th entries in t. It is easily seen that T' has the same distribution as T. Moreover, since the ratios $\lambda(t')/\lambda(t)$ are close to 1 uniformly for all pairs t, t' of neighboring configurations, the conditional probability that the switch actually takes place is close to 1. Also, since this construction is valid for each *fixed* pair of indices j, j', it is also valid if j' is replaced by a randomly selected index ν from any subset of $\{1, 2, \ldots, N\}$ (provided, of course that ν is independent of T).

Now fix $m \in \{1, 2, \ldots, N-k\}$ and define a new configuration T^* as follows. Perform the switching operation described in the previous paragraph k times: first for the pair $m+1, \nu_1$; then for the pair $m+2$, ν_2 and so on; and finally for the pair $m + k, \nu_k$, where $\nu_1, \nu_2, \ldots$ are iid random variables uniformly distributed on the set $\{1, 2, \ldots, N\}$. Then T^* also has the distribution $\mathscr{D}(T)$, by the argument of the preceding paragraph, and with probability approaching 1 as $n \to \infty$, all j switches are actually performed. Since the probability that two of the random indices $\nu_1, \nu_2, \ldots, \nu_k$ are the same is vanishingly small as $n \to \infty$ (recall that $N \to \infty$ as $n \to \infty$), the joint distribution of $(T^*_{m+1}, \ldots, T^*_{m+k})$ is nearly the same as that of k iid random variables from the multinomial distribution $\hat{\pi} = S_N/N$. Since with high probability $S_N/N \approx \pi$ when n is large, by Hypothesis 1, this proves (13).

The argument for (14) is similar, but now, since the values of $T_1, T_2, \ldots, T_m$ are conditioned, the switching operation must be performed with indices randomly chosen from a subset of $1, 2, \ldots, N$. Thus, it is necessary to have a "local" weak law of large numbers.

LEMMA 7. *Suppose that Hypotheses* 1 *and* 2 *are both satisfied. Fix* $\varepsilon > 0$ *and* $i \in [\tau]$. *Then, as* $n \to \infty$, *the relative frequency of type* i *in any block of* εN *consecutive* T_j*'s converges in probability to* π_i, *uniformly over all such blocks. More precisely, for any* $\delta > 0$, *if* N^* *is the greatest integer in* εN, *then, as* $n \to \infty$,

$$\max_{1 \le k \le N - \varepsilon N} P\left(\left|\frac{(S^i_{k+N^*} - S^i_k)}{N^*} - \pi_i\right| > \delta\right) \longrightarrow 0. \tag{15}$$

PROOF. Fix j. Let T' be the random configuration obtained from T by switching the entry T_j with the entry T_ν at a randomly chosen $\nu \in \{1, 2, \ldots, N\}$. Then, by Hypothesis 2 (see the argument above), the distribution of T' differs from that of T by a negligible amount as $n \to \infty$. Moreover, by Hypothesis 1, the distribution of a randomly chosen entry of T is increasingly close to π as $n \to \infty$. Consequently, for each j and each type i,

$$P\{T_j = i\} \longrightarrow \pi_i, \tag{16}$$

and this holds uniformly in j.

Now fix any pair j, j' of indices. Let T'' be the configuration obtained from T by switching entries $T_j, T_{j'}$ with entries $T_\nu, T_{\nu'}$, respectively, where ν and ν' are independent, randomly chosen indices from $1, 2, \ldots, N$. Again, by Assumption 2, the distribution of T'' is nearly identical to that of T. For any index k, let Z_k be the indicator of the event $\{T_k = i\}$. Then since ν and ν' are independent, each with the uniform distribution on $\{1, 2, \ldots, N\}$,

$$E(Z_\nu Z_{\nu'} \mid S_N) = \hat{\pi}_i^2,$$

where $\hat{\pi} = S_N/N$ is the empirical distribution of types in T. By Hypothesis 1, $\hat{\pi} \approx \pi$ with high probability. Consequently,

$$EZ_j Z_{j'} \approx \pi_i^2 \tag{17}$$

uniformly in j, j' such that $j \neq j'$.

The weak law (15) now follows from Chebyshev's inequality, as (16) and (17) give asymptotic convergence of the mean and the right order-of-magnitude bound on the variance of $\sum_{j=k+1}^{k+N^*} Z_j/N^*$. □

The proof of (14) may now be completed. Construct a new configuration T^{**} from T, with the same distribution as T, by switching pairs of entries in T as follows. Let $\nu_1', \nu_2', \dots$ be independent, identically distributed random indices uniformly distributed on $\{j\colon\ N(1-\varepsilon) \le j \le N\}$. Perform the switching operation described in the proof of (13) k times: first for the pair $m+1, \nu_1'$; then for $m+2, \nu_2'$ and so on; and finally for $m+k, \nu_k'$. That the resulting configuration has the same distribution as T and that all k switches are actually performed (with probability approaching 1) follows from Hypothesis 2, as above. However, the joint distribution of $T_{\nu_1'}, \dots, T_{\nu_k'}$, conditional on T, is approximately that of an iid sample of k items taken from the last εN entries of T. Since for each $i \in [\tau]$ the relative frequency of type i in this block is with high probability close to π_i, by the preceding lemma, (14) now follows. □

Proposition 1 implies that for each type $i \in [\tau]$ the "renewal" process consisting of successive indices j at which $T_j = i$ is, at least locally, close to a Bernoulli process with success probability parameter π_i. The following corollary makes this precise. For any $k = 1, 2, \dots$, let

$$N_k^i = \#\{1 \le m \le N\colon\ m = k + S_m^i \text{ and } T_m = i\}.$$

COROLLARY 5. *Under Hypotheses* 1 *and* 2, *for each* $\varepsilon > 0$, *each type* $i \in [\tau]$ *and each* $j = 0, 1, 2, \dots$,

$$\max_{1 \le k \le N(1-\pi_i-\varepsilon)} \left| P(N_k^i = j \mid S_N) - (1-\pi_i)\pi_i^j \right| \longrightarrow 0 \tag{18}$$

in probability as $n \to \infty$.

PROOF. Let $G_m = \{m = k + S_m^i\}$. When $k \le N(1-\varepsilon)$ and N is large (as it is when n is large), $P(\bigcup_{m > N(1-\varepsilon)} G_m)$ is negligible, since, by Lemma 7, $S_m^i/m \approx \pi_i$ with high probability for large m. Consequently, to prove (18), it suffices to show that, for any k,

$$P(T_{m+1} = \cdots = T_{m+j} = i \mid S_N;\ 1_{G_m} = 1;\ T_{[1,m-1]};\ T_m = 0) \to_P \pi_i^j$$

as $n \to \infty$, *uniformly* for $k \le N(1 - \pi_i - \varepsilon)$ and $1 \le m \le N(1-\varepsilon)$. However, this is an immediate consequence of (14). □

6. Proof of Theorem 1. Theorem 1 will be deduced from Corollary 5. This will be done in three steps: first, (7) will be proved for cycle signatures of length 1; second, for cycle signatures of length greater than or equal to 2; and finally it will be proved that, for distinct cycle signatures $\sigma, \sigma', \dots$, the random variables $N_\sigma, N_{\sigma'}, \dots$ are asymptotically independent. In using the machinery of the previous section, the set of types will be $[a]^l$, where l is the length of the cycle signature being considered. Each type represents a sequence of possible branch choices; to each $j = 1, 2, \dots, n$ is attached the type

$$\beta_j^l \overset{\Delta}{=} (\beta_j, \beta_{\Pi^{-1}(j)}, \dots, \beta_{\Pi^{-l+1}(j)})$$

that describes the sequence of branch choices at the order statistics $\zeta_{(j)}$, $\zeta_{(\Pi^{-1}(j))}, \dots$ along the backward Π-orbit of j. The random variables T_j will be the types attached to the indices j in various subsets of $\{1, 2, \dots, n\}$. The connection with Theorem 1, specifically, with (8), is that in the interval $J^{\Pi}_{i_1 i_2 \cdots i_l}$, the increments of nF^l_{Π} are precisely the differences $j' - j$ between successive indices j, j' with type $t = i_1 i_2 \cdots i_l$.

To apply the results of the preceding section, we will have to verify Hypotheses 1 and 2.

6.1. *Cycle signatures of length* 1. Let $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ be the sequence of branch choices, and let $\zeta = (\zeta_{(1)}, \dots, \zeta_{(n)})$ be the vector of order statistics in the inverse construction of Π.

LEMMA 8. *For each configuration $t \in [a]^n$ let $\lambda(t) = P(\beta = t \,|\, \zeta)$ be the conditional probability that the sequence β of branch choices takes the configuration t given the values of the order statistics ζ. There exists a constant $C < \infty$, independent of n, with the following property: for any pair t, t' of configurations such that t' is obtained from t by switching the jth and the j'th entries,*

$$\left| \frac{\lambda(t)}{\lambda(t')} - 1 \right| \le C |\zeta_{(j)} - \zeta_{(j')}|. \tag{19}$$

NOTE 8. This is where the hypothesis that $f \in C^2$ is used; see also Lemma 11.

PROOF OF LEMMA 8. Conditional on ζ, the branch choices β are independent and, for each j, the distribution of β_j is $\{1/f'(f_i^{-1}(\zeta_{(j)}))\}_{1 \le i \le a}$. Consequently, the likelihoods $\lambda(t)$ and $\lambda(t')$ are products, with all factors equal except for the jth and the j'th; in particular, if the jth and the j'th entries of t are i and i', respectively, then

$$\frac{\lambda(t)}{\lambda(t')} = \frac{f'(f_i^{-1}(\zeta_{(j')}))}{f'(f_i^{-1}(\zeta_{(j)}))} \frac{f'(f_{i'}^{-1}(\zeta_{(j)}))}{f'(f_{i'}^{-1}(\zeta_{(j')}))}.$$

The result (19) therefore follows from the standing assumptions about the mapping f, in particular, that f'' is continuous and bounded and that f' is bounded away from zero and infinity. □

PROOF OF (7) WHEN $|\sigma| = 1$. Let $\sigma = i_*$, where $1 < i_* < a$ (recall that the cycle signatures $\sigma = 1, a$ have already been disposed of). By Lemma 5, there is a unique fixed point of f in (the interior of) J_{i^*}; call this fixed point x_*. By

Lemma 6, to prove (7) it suffices to prove (7) with N_σ replaced by N_σ^ε for some small $\varepsilon > 0$, where N_σ^ε is the number of integers j such that

$$F_\Pi\left(\frac{j}{n}\right) = \frac{j}{n} \quad\text{and}\quad \left|x_* - \frac{j}{n}\right| < \varepsilon. \tag{20}$$

We will show that, conditional on certain random variables to be specified shortly, the distribution of N_σ^ε is the same as that of the random variable $N_k^{i^*}$ in Corollary 5, for a suitable k.

Fix (nonrandom) integers K_- and K_+, depending on n, so that $K_- < nx_* < K_+$ and $nx_* - K_- \approx K_+ - nx_* \approx n^{2/3}$. Then, with probability approaching 1 (as $n \to \infty$), the order statistics $\zeta_{(K_-)}$ and $\zeta_{(K_+)}$ are within distance $n^{-1/4}$ of x_*. Hence, all the order statistics $\zeta_{(j)}$ indexed by $K_- \le j \le K_+$ are within distance $n^{-1/4}$ of x_*. Moreover, with probability approaching 1, $F_\Pi(K_-/n) < K_-/n$ and $F_\Pi(K_+/n) > K_+/n$. This follows from Lemma 3, because the graph of f crosses the straight line $y = x$ at (x_*, x_*) from below, transversely, since $f'(x_*) > 1$. Consequently, with high probability, any solution j of (20) is between K_- and K_+.

Let $N = K_+ - K_- \approx 2n^{2/3}$, and, for $1 \le j \le N$, let $T_j = \beta_{j+K_-}$. Let S_r be defined in terms of the random variables T_j as in (12). Observe that Hypothesis 1 holds; this is because, conditional on ζ, the random variables $T_1, T_2, \ldots, T_N$ are independent, with each T_j having a distribution $\{1/f'(f_i^{-1}(\zeta_{(j+K_-)}))\}_{1\le i\le a}$ that is increasingly close (with high probability) to the distribution π on $[a]$ given by

$$\pi_i = \frac{1}{f'(f_i^{-1}(x_*))}.$$

[By the previous paragraph, all of the order statistics $\zeta_{(j')}$ indexed by $K_- \le j' \le K_+$ are within distance $o(1)$ of the fixed point x_*, with high probability.] Observe also that Hypothesis 2 holds, by Lemma 8.

Now consider the effect of conditioning on the values of all the branch choices β_j indexed by $j \le K_-$ and $j > K_+$ (and also on ζ). Since the random variables β_j are conditionally independent given ζ, this extra conditioning does not change the distribution of $T = (T_1, T_2, \ldots, T_N)$, and so Hypotheses 1 and 2 remain valid. However, the branch choices β_j, for $j \le K_-$ and $j > K_+$, together with the value of S_N, determine the value $k = K_- - nF_\Pi(K_-/n)$. Since the number N_σ^ε of solutions to (20) equals

$$\#\{m \in [1, J]\colon\ m = k + S_m^{i_*} \text{ and } T_m = 1\},$$

(7) follows from Corollary 5. □

6.2. *Cycle signatures of length greater than or equal to* 2. This case is somewhat more difficult. When $l = 1$, the types (branch choices) β_j are conditionally independent given the order statistics ζ; however, when $l \ge 2$, this is no longer true. In fact, the sequence of types $\beta^l = (\beta_1^l, \beta_2^l, \ldots, \beta_n^l)$ is completely determined by the sequence β of branch choices, and, in particular, the sequence of rth entries is a permutation of the sequence of first entries. This makes the verification of Hypothesis 1 slightly trickier, as the weak law of large numbers for sums of independent random variables is no longer directly applicable, and it also complicates the verification of Hypothesis 2. More important, the final step in the preceding proof, in which we conditioned on the branch choices β_j for $j \notin (K_-, K_+]$, would not be valid for $l \geq 2$, because this conditioning would change the distribution of T.

Fix $l \geq 2$ and $x_1 \in (0,1)$. Let K_- and K_+ be nonrandom integers, depending on n, such that $nx_1 - K_- \approx K_+ - nx_1 \approx n^{2/3}$. Define $T_j = \beta^l_{j+K_-}$ for $1 \leq j \leq N$, where $N = K_+ - K_-$, and let S_m be defined as in (12). For each $b = i_1 i_2 \cdots i_l \in [a]^l$, define

$$\pi_b = \prod_{r=1}^{l} (f'(f^{-r}_{i_1 i_2 \cdots i_r}(x_1)))^{-1},$$

where $f^{-r}_{i_1 i_2 \cdots i_r}$ is the $i_1 i_2 \cdots i_r$ branch of f^{-r}, that is, $f^{-r}_{i_1 i_2 \cdots i_r} = f^{-1}_{i_r} \circ \cdots \circ f^{-1}_{i_1}$. The following lemma implies that Hypothesis 1, with π as defined above, is valid for the triangular array T_j.

LEMMA 9. *For each type $b \in [a]^l$, as $n \to \infty$,*

$$\frac{S^b_N}{N} \to_P \pi_b.$$

PROOF. Let $b = i_1 i_2 \cdots i_l$. Consider the value of S^b_N; this is the number of indices $j \in [1, N]$ for which $T_j = b$, equivalently, the number of indices $j \in (K_-, K_+]$ such that the first l branch choices along the backward orbit of j are $i_1, i_2, \ldots, i_l$. We will give an alternative representation in terms of F^l_Π (the lth iterate of F_Π). By Corollary 1, when n is large, F^l_Π is close in the Skorohod topology to f^l (with high probability) and hence has a^l distinct intervals of monotonicity, each mapped onto $[0,1]$ by F^l_Π. (Note that this is an abuse of terminology, since F^l_Π is really a step function. We mean that on each of these intervals F^l_Π is monotone when restricted to the points j/n.) The intervals $J^{\Pi^l}_{b'}$ of monotonicity are indexed naturally by the elements b' of the type space $[a]^l$. It is easily seen that S^b_N/n is the length of the interval in $J^{\Pi^l}_b$ mapped onto $(K_-/n, K_+/n]$ by F^l_Π (provided that all of the intervals $J^{\Pi^l}_{b'}$ are nonempty, which is the case with probability approaching 1 as $n \to \infty$). However, $(K_-/n, K_+/n]$ is an interval of length $2n^{2/3}$ centered at x_1, so, by Corollary 2 (with $\varepsilon < 1/6$), the length of the interval in $J^{\Pi^l}_b$ mapped onto $(K_-/n, K_+/n]$ by F^l_Π is to first order of approximation the same as the length of the interval in $J^{\Pi^l}_b$ mapped onto $(K_-/n, K_+/n]$ by f^l. This interval is of length $o(1)$ and approximately centered at $f^{-l}_b(x_1)$. Consequently, f^l is close to linear in this interval, with derivative π^{-1}_b. Therefore, with probability approaching 1, $S^b_N/N \approx \pi_b$. □

We turn next to Hypothesis 2. Verification of this assumption is complicated by the fact that not every possible configuration necessarily occurs with positive probability, because the complete vector β^l of types is determined by the vector β of initial entries. In fact, if $b = (b_1, b_2, \ldots, b_n)$ is a configuration (vector of types) such that $P\{\beta^l = b\} > 0$ and if b' is the vector obtained by switching the jth and j'th entries of b, then it is not necessarily the case that $P\{\beta^l = b'\} > 0$.

LEMMA 10. *Let $b = (b_1, b_2, \ldots, b_n) \in ([a]^l)^n$ be a configuration such that $P\{\beta^l = b\} > 0$, and let b' be the configuration resulting from a switch of the branch choices at neighboring indices $j, j' = j+1$. Then $b_i = b'_i$ for all indices $i = 1, 2, \ldots, n$ except for the $2l$ indices $i \in \mathscr{I}$, where $\mathscr{I} = \{\Pi^r(j)\}_{0 \le r < l} \cup \{\Pi^r(j')\}_{0 \le r < l}$. Moreover, if the switch of branch choices is nontrivial (i.e., if the branch choices at indices j and j' are different) and if the branch choices at $\Pi^r(j)$ and $\Pi^r(j')$ are the same for all $1 \le r < l$, then, for all $1 \le r < l$,*

$$b'_{\Pi^r(j)} = b_{\Pi^r(j')} \quad and \quad b'_{\Pi^r(j')} = b_{\Pi^r(j)}. \tag{21}$$

NOTE 9. The last condition, that the branch choices at $\Pi^r(j)$ and $\Pi^r(j')$ are the same for all $1 \le r \le l$, will be true provided that (i) F_Π has a distinct nonempty intervals of monotonicity (equivalently, Π has $a-1$ descents) and (ii) for each $0 \le r < l$, $\Pi^r(j)$ and $\Pi^r(j')$ are in the same interval of monotonicity of F_Π [equivalently, $\Pi^r(j) < \Pi^r(j') \ \forall\ 1 \le r \le l$]. Recall that when n is large, condition (i) holds with probability approaching 1. In the application of Lemma 10 in Lemma 11 (and Section 6.3) below, the indices j and j' will be such that j/n and j'/n are near the first point x_1 of a periodic orbit of f. Consequently, with high probability $\Pi^r(j)/n$ and $\Pi^r(j')/n$ will remain close to the points of the periodic orbit for $1 \le r \le l$ and, in particular, will remain in the same interval of monotonicity of F_Π.

PROOF OF LEMMA 10. If the branch choices at j and j' are the same, then switching them does not change the configuration, so there is nothing to prove. Suppose, then, that they are different, and consider the effect of switching them. Since $j' = j+1$, the switch changes Π to $\Pi\sigma$, where σ is the permutation that transposes j and j', by Lemma 1. Thus, the switch has no effect on the orbits (nor, therefore, on the types) at any indices j'' outside the orbits of j and j'. Moreover, the change of branch choices at j and j' can only be "seen" in the types at indices along the forward orbits of j and j' for a distance at most l along these orbits, so the types are unchanged at all $j'' \notin \mathscr{I}$.

Consider, then, the effect of the switch on the types at the indices $j'' \in \mathscr{I}$. Label the indices in the Π-orbits of j and j' as follows:

$$\cdots \to j_{-1} \to j = j_0 \to j_1 \to j_2 \to \cdots;$$
$$\cdots \to j'_{-1} \to j' = j'_0 \to j'_1 \to j'_2 \to \cdots \to j'_l.$$

Switching the branch choices at j and j' changes the backward orbit of j_r from

$$j_r \to j_{r-1} \to \cdots j_1 \to j = j_0 \to j_{-1} \to \cdots$$

to

$$j_r \to j_{r-1} \to \cdots j_1 \to j' = j'_0 \to j'_{-1} \to \cdots$$

and the backward orbit of j'_r from

$$j'_r \to j'_{r-1} \to \cdots j'_1 \to j' = j'_0 \to j'_{-1} \to \cdots$$

to

$$j'_r \to j'_{r-1} \to \cdots j'_1 \to j = j_0 \to j_{-1} \to \cdots.$$

By hypothesis, $j_r < j'_r$ for each $1 \le r \le l$, and the branch choices at j_r and j'_r are the same for each $1 \le r \le l$. Hence, the effect of the switch of branch choices at j, j' is to transpose the *types* at j_r and j'_r for each $1 \le r \le l$. □

For any interval J of integers contained in $[1, n]$ and any type $b \in [a]^l$, define M^b_J to be the number of indices $j \in J$ such that $\beta^l_j = b$, and let $M_J = (M^b_J)_{b \in [a]^l}$ be the vector of type-counts in J. Let x_1, K_-, K_+ be as in Lemma 9, and let $T = (T_1, T_2, \dots, T_N)$, where $T_j = \beta^l_{j+K_-}$ as above. For a given configuration $t \in ([a]^l)^N$ of types, define

$$\lambda(t) = P(T = t \mid \zeta; M_{[1,K_-]}, M_{(K_-,K_+]}, M_{(K_+,n]}).$$

Note that conditioning on $M_{[1,K_-]}$, $M_{(K_-,K_+]}$ and $M_{(K_+,n]}$ determines the value of $K_- - F^l_\Pi(K_-)$.

LEMMA 11. *Assume that x_1 does not lie in a periodic orbit of f of period less than l and that $f^r(x_1) \neq 0, 1$ for any $1 \le r \le l+1$. Then there exists a constant $C < \infty$, independent of n, with the following property: for any pair t, t' of configurations such that t' is obtained from t by switching the jth and the j'th entry (for any pair $1 \le j, j' \le N$ of indices),*

$$\left| \frac{\lambda(t)}{\lambda(t')} - 1 \right| \le C n^{-1/4}. \tag{22}$$

PROOF. Since x_1 does not lie in a periodic orbit of period less than l, there is a neighborhood of x_1 containing none of $f^r(x_1)$, $1 \le r < l$, and none of the points in $f^{-r}(x_1)$, $1 \le r < l$. Consequently, there is a neighborhood $\mathscr{N} = (x_1 - \varepsilon, x_1 + \varepsilon)$ of x_1 such that, for each $1 \le r < l$ and each branch $f^{-r}_{i_1 i_2 \cdots i_r}$ of the inverse function f^{-r}, the intervals $f^r(\mathscr{N})$ and $f^{-r}_{i_1 i_2 \cdots i_r}(\mathscr{N})$ are contained entirely either in $(0, x_1 - \varepsilon)$ or in $(x_1 + \varepsilon, 1)$. By Corollary 1, it follows that, with probability approaching 1 as $n \to \infty$, the images $F^r_\Pi([K_-/n, K_+/n])$ and $(F_\Pi)^{-r}_{i_1 i_2 \cdots i_r}([K_-/n, K_+/n])$ for $0 \le r < l$, are pairwise disjoint intervals, each entirely contained in either $[0, K_-/n]$ or $(K_+/n, 1]$. Moreover, since the points $f^r(x_1)$ remain bounded away from the endpoints of the intervals J_i for $1 \le r \le l$, if j and j' are integers such that j/n and j'/n are close to x_1, then with high probability $F^r_\Pi(j/n)$ and $F^r_\Pi(j'/n)$ are in the same interval of monotonicity of F_Π for all $1 \le r \le l$, and hence the branch choices at indices $\Pi^r(j)$ and $\Pi^r(j')$ are the same for all $1 \le r \le l$.

Now suppose that $j' = j + 1$ and that $T_j \neq T_{j'}$. Let $r + 1$ be the first coordinate in which T_j and $T_{j'}$ differ. Then the Π^{-1}-orbits of j and j' remain adjacent for r steps, then diverge, in particular, for each $0 \leq r' \leq r$, $\Pi^{-r'}(j') = \Pi^{-r'}(j) + 1$, and the branch choices at $\Pi^{-r}(j)$ and $\Pi^{-r}(j')$ are different. Consider the effect of switching the branch choices at $\Pi^{-r}(j)$ and $\Pi^{-r}(j')$; by Lemma 10, the types are unchanged at all indices except at the indices in the forward orbits of $\Pi^{-r}(j)$ and $\Pi^{-r}(j')$, which are switched in pairs. By the preceding paragraph, all pairs of indices whose types are switched are both in $[1, K_-]$, both in $(K_+, n]$, or both in $(K_-, K_+]$. Consequently, switching the branch choices at $\Pi^{-r}(j)$ and $\Pi^{-r}(j')$ does not affect the totals $M_{[1,K_-]}$, $M_{(K_-,K_+]}$ and $M_{(K_+,n]}$; furthermore, it has the effect of switching the *types* T_j and $T_{j'}$, and leaving all other $T_{j''}$ unchanged. Note that the indices $\Pi^{-r}(j)$ and $\Pi^{-r}(j')$ where the switch of branch choices takes place are nearest neighbors: $\Pi^{-r}(j') = \Pi^{-r}(j) + 1$.

This shows that the *types* at nearest neighbors $j, j' \in [1, N]$ can be transposed by a single switch of branch choices at neighboring indices $\Pi^{-r}(j)$ and $\Pi^{-r}(j')$. Consequently, the types at *non*-nearest neighbors $j, j' \in [1, N]$ can be transposed by making a sequence of branch choice switches at neighboring indices j'' and j'''. The indices j'' and j''' at which these switches occur are all in $\bigcup_{r=0}^{l-1} \Pi^{-r}((K_-, K_+])$, and for any such pair j'', j''' there at most two branch choice switches. Moreover, in this sequence of branch choice switches, the totals $M_{[1,K_-]}$, $M_{(K_-,K_+]}$ and $M_{(K_+,n]}$ are unchanged.

Finally, consider the change in likelihood caused by any one of these nearest neighbor branch choice switches. If the switch occurs at neighboring indices j'' and j''', if the branches switched are i and i' and if the switch changes the original configuration $b \in [a]^n$ of branch choices to the new configuration b', then

$$\frac{P(\beta = b \mid \zeta)}{P(\beta = b' \mid \zeta)} = \frac{f'(f_i^{-1}(\zeta_{(j''')}))}{f'(f_i^{-1}(\zeta_{(j'')}))} \frac{f'(f_{i'}^{-1}(\zeta_{(j''')}))}{f'(f_{i'}^{-1}(\zeta_{(j'')}))}.$$

As in the proof of Lemma 8, the right-hand side differs from 1 by an amount bounded by $C|\zeta_{(j'')} - \zeta_{(j''')}|$. Consequently, for any pair $j, j' \in [1, N]$, if $t \in ([a]^l)^N$ is a configuration of types such that $P\{T = t\} > 0$ and if t' is the configuration obtained by switching the types at j, j', then

$$\begin{aligned}\left|\frac{\lambda(t)}{\lambda(t')} - 1\right| &\leq 2C \sum_{r=1}^{l} \sum_{\nu=j}^{j'} |\zeta_{(\Pi^{-r+1}(\nu+K_-))} - \zeta_{(\Pi^{-r+1}(\nu+1+K_-))}| \\ &\leq 2C \sum_{r=1}^{l} |\zeta_{(\Pi^{-r+1}(K_+))} - \zeta_{(\Pi^{-r+1}(K_-))}|.\end{aligned}$$

That the last sum is of smaller order of magnitude than $n^{-1/4}$ follows from the assumption that $K_+ - K_- = N \approx n^{2/3}$ and Corollary 2. □

PROOF OF (7) WHEN $|\sigma| \geq 2$. By Lemma 5, f has a unique periodic orbit $x = x_1 x_2 \cdots x_l$ of signature σ, with $x_1 < x_r$ for all $2 \leq r \leq l$ and $f(x_r) =$

x_{r+1} ∀ r. This orbit has *minimal* period l, so its elements x_r are distinct. By Lemma 6, with probability approaching 1 as $n \to \infty$, $N_\sigma = N_\sigma^\varepsilon$ for sufficiently small ε, where N_σ^ε is the number of solutions of (8).

As in Lemmas 9 and 11, fix (nonrandom) integers K_- and K_+, depending on n, so that $K_- < nx_1 < K_+$ and $nx_1 - K_- \approx K_+ - nx_1 \approx n^{2/3}$. With probability approaching 1, the order statistics $\zeta_{(K_-)}$ and $\zeta_{(K_+)}$ are within distance $n^{-1/4}$ of x_1, and hence the order statistics $\zeta_{(j)}$ indexed by $K_- \le j \le K_+$ are within distance $n^{-1/4}$ of x_1. Also, with probability approaching 1, $F_\Pi^{(l)}(K_-/n) < K_-/n$ and $F_\Pi^{(l)}(K_+/n) > K_+/n$. (This follows from Lemma 3.) Consequently, any j for which (8) holds must be between K_- and K_+ (with high probability).

Lemma 11 implies that the *conditional* distribution of T given the order statistics ζ and the totals $M_{[1,K_-]}$, $M_{(K_-,K_+]}$ and $M_{(K_+,n]}$ satisfies Hypothesis 2 of Section 5. Lemma 9 implies that Hypothesis 1 is satisfied. Consequently, the conclusion of Corollary 5 holds. Now consider the increments of nF_Π^l near x_1: they are precisely the differences $j' - j$ between successive indices $j < j'$ where $T_j = T_{j'} = t$, where $t = t_1 t_2 \cdots t_l$ is the signature σ "run backward," that is, $\sigma = t_l t_{l-1} \cdots t_1$. Consequently, the number of solutions of (8) is just

$$\#\{1 \le m \le N\colon\ m = k + S_m^t \text{ and } T_m = t\},$$

where $k = K_- - nF_\Pi^l(K_-)$. Since the conditioning (on the totals $M_{[1,K_-]}$, $M_{(K_-,K_+]}$ and $M_{(K_+,n]}$) determines the value of k, (7) now follows from Corollary 5. [Note that since with high probability both $nF_\Pi^l(K_-) - K_-$ and $nF_\Pi^l(K_+) - K_+$ are of size approximately $\frac{1}{2}(N - S_N^t)$, the condition $k < N(1 - \pi_t - \varepsilon)$ will be met with probability approaching 1.] □

6.3. *Asymptotic independence of* $N_\sigma, N_{\sigma'}, \ldots.$ The argument is an augmentation of the proof for cycle signatures of length greater than or equal to 2. We shall give only a brief sketch, considering only the case of two distinct cycle signatures σ and σ'. A straightforward induction extends the result to an arbitrary finite set of distinct cycle signatures.

Let $x = (x_1, x_2, \ldots, x_l)$ and $y = (y_1, y_2, \ldots, y_{l'})$ be the unique periodic orbits of f with signatures σ and σ', respectively. Observe that these orbits do not intersect and, furthermore, that the inverse images of x_1 and y_1 under f^{-r}, for $1 \le r \le \max(l, l')$ do not intersect. Take integers K_-, K_+ and K'_-, K'_+, depending on n, so that

$$nx_1 - K_- \approx n^{2/3} \approx K_+ - nx_1,$$
$$ny_1 - K'_- \approx n^{2/3} \approx K'_+ - ny_1.$$

For definiteness suppose that $K_- < K_+ < K'_- < K'_+$. Condition on (i) the order statistics ζ, (ii) the totals $M_{[0,K_-]}$, $M_{(K_-,K_+]}$, $M_{(K_+,K'_-]}$, $M_{(K'_-,K'_+]}$ and $M_{(K'_+,n]}$ and (iii) the values β_j^l of the l-types for $K_- < j \le K_+$. [Note that the totals M_J should be computed for types $b \in [a]^{l''}$, where $l'' = \max(l, l')$.] Observe that (i)–(iii) determine the value N_σ^ε and consequently (with high probability) the value of N_σ. Now repeat the argument of Section 6.2 for $N_{\sigma'}$, conditional on (i)–(iii). The only part of the argument that may be affected by the extra condition (iii) is the argument in the proof of Lemma 11 that the type-totals (ii) are unchanged after switching l'-types at indices $j, j' \in (K'_-, K'_+]$. However, this may be redone along the same lines, using the fact that the inverse images $f^{-r}(x_1)$ and $f^{-r}(y_1)$ are distinct. The upshot is that

$$P(N_{\sigma'} = k \mid N_\sigma) \to (1 - \omega(\sigma'))\omega(\sigma')^k,$$

proving that N_σ and $N_{\sigma'}$ are asymptotically independent as $n \to \infty$. □

REFERENCES

[1] BAYER, D. and DIACONIS, P. (1992). Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.* **2** 294–313.

[2] DEVANEY, R. L. (1989). *An Introduction to Chaotic Dynamical Systems*. Addison-Wesley, Reading, MA.

[3] DIACONIS, P., MCGRATH, M. and PITMAN, J. (1995). Riffle shuffles, cycles, and descents. *Combinatorica* **15** 11–29.

[4] GESSEL, I. and REUTENAUER, C. (1993). Counting permutations with given cycle structure and descent set. *J. Combin. Theory Ser. A* **64** 189–215.

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

[6] LALLEY, S. (1994). Riffle shuffles and dynamical systems on the unit interval. Technical report, Dept. Statistics, Purdue Univ.

[7] LLOYD, S. P. and SHEPP, L. (1966). Ordered cycle lengths in a random permutation. *Trans. Amer. Math. Soc.* **121** 340–357.

[8] REEDS, J. (1981). Unpublished manuscript.

[9] VERSHIK, A. M. and SCHMIDT, A. (1977). Limit measures arising in asymptotic theory of symmetric groups. *Probability Theory and Applications* **22** 72–88; **23** 34–46.

DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
MATHEMATICAL SCIENCES BUILDING
WEST LAFAYETTE, INDIANA 47907
E-MAIL: lalley@stat.purdue.edu