The time evolution of permutations under random stirring

Bálint Vető

Abstract. We consider permutations of $\{1,\dots,n\}$ obtained by $\lfloor\sqrt{n}t\rfloor$ independent applications of random stirring. In each step the same marked stirring element is transposed with probability $1/n$ with any one of the n elements. Normalizing by $\sqrt{n}$, we describe the asymptotic distribution of the cycle structure of these permutations, for all $t\geq 0$, as $n\to\infty$.

1. Introduction

We consider the following random stirring mechanism: n numbered balls are given in the beginning on their corresponding numbered places. In each step, independently, the first ball, which is referred to as the *stirring particle* or *stirring element*, changes place with one of the n balls or stays unchanged with probability $1/n$. We investigate that permutation which brings the balls from their initial place to their place after i steps.

Formally, let $\pi^{(n)}(i) := T_i^{(n)}\circ T_{i-1}^{(n)}\circ\dots\circ T_1^{(n)}$ be a permutation acting on the set $[n] := \{1,\dots,n\}$. The permutations $(T_i^{(n)})_{i=1}^\infty$ are chosen independently with uniform distribution from the $n-1$ transpositions moving the stirring particle and the identity permutation.

Let σ be a permutation of a finite set S, i.e. an $S\to S$ bijective function. The *cycle*s (orbits) of σ are the sets of form $\{v,\sigma(v),\sigma^2(v),\dots\}\subseteq S$ for some $v\in S$. The set S is the disjoint union of its cycles. The *cycle structure* of σ is the sequence of the cardinalities of the different cycles in non-increasing order.

In our case one of the cycles can be distinguished from the others (namely the cycle that contains the stirring element), which will be called the *active cycle*. For the complete description it is enough to determine the distribution of the cycle structure of the permutation $\pi^{(n)}(i)$ (regarding the active cycle separately). This gives the distribution of the conjugacy class of $\pi^{(n)}(i)$, with attention restricted to the conjugation with permutations that fix the stirring particle. The distribution of $\pi^{(n)}(i)$ is uniform within a fixed conjugacy class.

We encode the permutation $\pi^{(n)}(i)$ with the vector $\mathbf{C}^{(n)}(i) := (C_0^{(n)}(i), C_1^{(n)}(i), C_2^{(n)}(i),\dots)$, where $C_0^{(n)}(i)$ denotes the length of the active cycle, $C_1^{(n)}(i), C_2^{(n)}(i),\dots$ the lengths of those cycles in non-increasing order which are already moved by one of the transpositions $(T_j^{(n)})_{j=1}^i$. All other $C_j^{(n)}(i)$-s are 0. $(\mathbf{C}^{(n)}(i))_{i=0}^\infty$ is a process

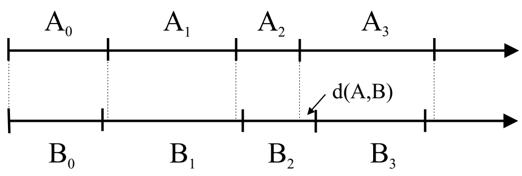

Figure 1: The metric on $\mathbf{S}$

on the state space

$$\mathbf{S} := \{(s_0, s_1, s_2, \dots) : s_n \in \mathbb{R}, \quad s_n \geq 0 \quad n = 0, 1, 2, \dots, \\ s_1 \geq s_2 \geq \dots \geq s_n \geq \dots \text{ and } s_j > 0 \text{ for finitely many } j\}$$

with the distance

$$d(A, B) := \sup \left\{ \left| \sum_{j=0}^{k} A_j - \sum_{j=0}^{k} B_j \right| : k = 0, 1, 2, \dots \right\} \tag{1}$$

where $A = (A_0, A_1, A_2, \dots)$ and $B = (B_0, B_1, B_2, \dots)$ are elements of $\mathbf{S}$. (See Figure 1.) The ranking is not a natural part of the problem, but it facilitates studying the model.

At each step after applying a random transposition two types of changes may happen in the cycle structure: merging of two distinct cycles or splitting of a cycle in two. While different transpositions $(T_j^{(n)})_{j=1}^{i}$ are applied (which means that the stirring particle chooses a new element in each step until i), the cycle decomposition of $\pi^{(n)}(i)$ contains only fixed points and the active cycle, which increases by one in each step: $\mathbf{C}^{(n)}(i) = (i + 1, 0, 0, \dots)$. If a transposition recurs, then the cycle splits in two, one of which will be the new active cycle. If there are already more than one non-trivial cycles in the decomposition, then the active cycle can merge with another cycle. (See Figure 2.) The model realizes a coagulation–fragmentation process.

A reduction of the problem is to study the coagulation and fragmentation events of the cycles together, because both of these events happen when the stirring element steps to a place already visited. We investigate this simpler question first. Then we introduce a continuous time process on $\mathbf{S}$, which turns out to be the limit process. The convergence is proved by coupling. In Section 4 we show that the stationary distribution of the underlying split-and-merge transformation is an adequate modification of the Poisson – Dirichlet distribution.

A similar model is studied by Schramm in [11]. He chooses $(T_i)_{i=1}^{\infty}$ to be independent random transpositions with uniform distribution from all possible transpositions of the set $[n]$. He identifies the limit distribution of the proportions of the giant cycles in the permutation $\Pi(t) = T_t \circ \dots \circ T_1$ after $t = cn$ steps as $n \to \infty$, where $c > 1/2$ is a constant.

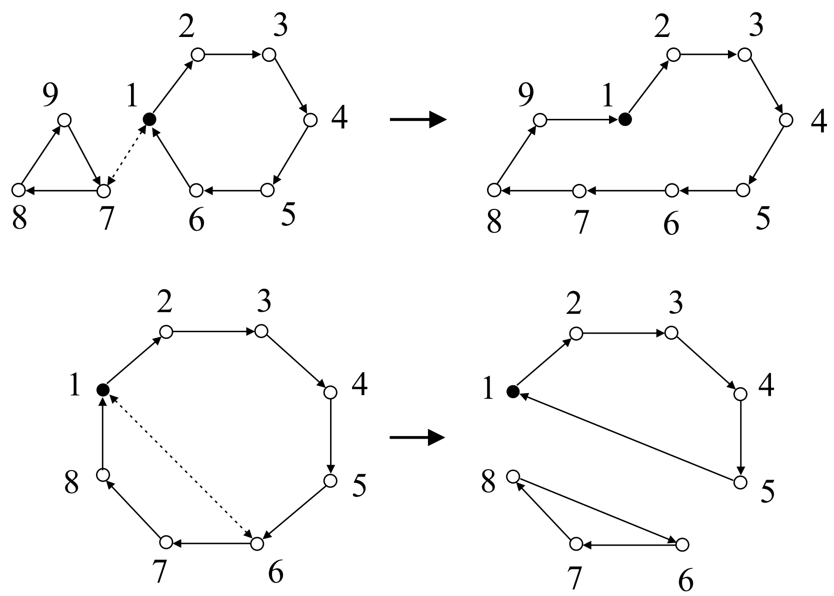

Figure 2: Coagulation and fragmentation of cycles

This result is in accordance with the classical theory of the random graphs derived from Erdős [6]. Let us consider the random graph $G(t)$ on the vertex set $[n]$ where $\{u, v\}$ is an edge in it if and only if the transposition (u, v) appears in $\{T_1, \dots, T_t\}$. By the Erdős – Rényi Theorem [7] the graph $G(t)$ has a giant connected component only in the case $t/n = c > 1/2$ similarly to the condition on the random permutations. (For random graphs and random graph processes see [8] and [12]).

In Schramm's paper the vector of the cycle sizes of $\Pi(t)$, listed in non-increasing order and normalized by the magnitude of the giant connected component of $G(t)$, converges in distribution to Poisson – Dirichlet distribution with parameter 1 after $t = cn$ steps ($c > 1/2$) as $n \to \infty$. That is the limit distribution of the relative cycle sizes in a random permutation chosen uniformly from all permutations of $[n]$ as $n \to \infty$. Thus for large n the permutation $\Pi(t)$ behaves on the giant connected component of the Erdős – Rényi graph $G(t)$ as a uniform permutation.

Our paper is motivated by Tóth [13] in connection with the quantum-physical applications of the problem. Angel analyzed Tóth's random walk model on regular trees in [1]. For similar random stirring models see also [2] and [5].

2. Return times of the stirring particle

The movement of the stirring particle is a random walk $(B_i^{(n)})_{i=0}^{\infty}$ on the set $[n]$, which is homogeneous in space and time. Let

$$V_i^{(n)} := \#\{k : k \leq i, \exists j < k : B_j^{(n)} = B_k^{(n)}\}$$

be the number of *returns* until the ith step to places already visited by the random walk $(B_j^{(n)})_{j=0}^{\infty}$. We also include those steps when the stirring particle keeps its place.

After the ith step the stirring element has already visited exactly $i + 1 - V_i^{(n)}$ places (including the starting point), so the transition probabilities of the Markov-chain $(V_i^{(n)})_{i=0}^{\infty}$ are

$$\mathbb{P}\left(V_{i+1}^{(n)} - V_i^{(n)} = 1 \;\middle|\; V_i^{(n)}\right) = 1 - \mathbb{P}\left(V_{i+1}^{(n)} - V_i^{(n)} = 0 \;\middle|\; V_i^{(n)}\right) = \frac{i+1-V_i^{(n)}}{n}. \quad (2)$$

In order to get a non-trivial limit distribution, the time of the processes should be accelerated. As opposed to Schramm [11], in Theorem 1 the scaling is $\sqrt{n}$. This means that we describe the beginning of the evolution, because after $\mathcal{O}(\sqrt{n})$ steps the bulk of the elements is still unchanged. Simultaneously we normalize the cycle sizes with $\sqrt{n}$ and we let $n \to \infty$.

From now on we investigate the asymptotic behaviour of $\left(\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)/\sqrt{n}\right)_{t\geq 0}$ as $n \to \infty$, where the division is meant coordinatewise, namely $\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)/\sqrt{n} := (C_0^{(n)}(\lfloor\sqrt{n}t\rfloor)/\sqrt{n}, C_1^{(n)}(\lfloor\sqrt{n}t\rfloor)/\sqrt{n}, \dots)$. Elementary calculations, similar to the classical birthday problem, give the following limit distribution of the returns. For limit theorems related to generalizations of the birthday problem see also [3].

Proposition 1. *Let $(V_t)_{t\geq 0}$ be an inhomogeneous Poisson point process with intensity $\rho(t) = t$. Then*

$$(V_{\lfloor\sqrt{n}t\rfloor}^{(n)})_{t\geq 0} \stackrel{\mathrm{d}}{\Longrightarrow} (V_t)_{t\geq 0} \qquad (n \to \infty)$$

in terms of finite dimensional marginal distributions.

3. Coupling

Much more can be stated for the above model. Not only $(V_{\lfloor\sqrt{n}t\rfloor}^{(n)})_{t\geq 0}$, but the sequence of the processes $\left(\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)/\sqrt{n}\right)_{t\geq 0}$ converges. Moreover, a stronger type of convergence is realized by means of coupling.

The limit process is a natural continuous extension of the discrete processes $(\mathbf{C}^{(n)}(i))_{i=0}^{\infty}$. For large n, the active coordinate $C_0^{(n)}(i)$ increases in the bulk of the steps (when no split or merge occurs). In the times of jumps of $(V_i^{(n)})_{i=0}^{\infty}$ a split or a merge happens depending on the proportions of the cycle sizes as follows. The probability of a split in the ith step, conditionally, given that the stirring particle returns to a place already visited, is

$$\frac{C_0^{(n)}(i-1)}{\sum_{m=0}^{\infty} C_m^{(n)}(i-1)}. \tag{3}$$

The conditional probability of the merge of the jth cycle and the active one is

$$\frac{C_j^{(n)}(i-1)}{\sum_{m=0}^{\infty} C_m^{(n)}(i-1)}. \tag{4}$$

We define an $\mathbf{S}$-valued continuous time stochastic process $\mathbf{C}(t) = (C_0(t), C_1(t), C_2(t), \dots)$ with càdlàg paths, which imitates the evolution of $(\mathbf{C}^{(n)}(i))_{i=0}^{\infty}$. It is built on a Poisson point process $(V_t)_{t\geq 0}$ with intensity $\rho(t) = t$. Similarly to the discrete case, at the moments of jumps of $(V_t)_{t\geq 0}$ a split or a merge event occurs with probability proportional to the coordinates of $\mathbf{C}$.

The initial state is $\mathbf{C}(0) := (0, 0, 0, \dots)$. The evolution of the process is as follows: the coordinate $C_0(t)$ increases with constant speed 1 between the jumps of $(V_t)_{t\geq 0}$. Let τ_k be the kth time of jump of $(V_t)_{t\geq 0}$, in other words $V_{\tau_k} = k$ and $V_{\tau_k -} = \lim_{\varepsilon \downarrow 0} V_{\tau_k - \varepsilon} = k - 1$. Let $(U_k)_{k=1}^{\infty}$ be i.i.d. random variables with uniform distribution on $[0,1]$, independent of $(V_t)_{t\geq 0}$. One of the following two actions occurs at time τ_k.

1. *Split:* If

$$U_k \leq \frac{C_0(\tau_k-)}{\sum_{m=0}^{\infty} C_m(\tau_k-)}, \tag{5}$$

then let $C_0(\tau_k) := U_k \sum_{m=0}^{\infty} C_m(\tau_k-)$, and the sequence $(C_m(\tau_k))_{m=1}^{\infty}$ will be the collection of $(C_m(\tau_k-))_{m=1}^{\infty}$ and $C_0(\tau_k-) - U_k \sum_{m=0}^{\infty} C_m(\tau_k-)$ rearranged in decreasing order.

2. *Merge:* Otherwise, a unique index $j \geq 1$ can be chosen a.s. via

$$\frac{\sum_{m=0}^{j-1} C_m(\tau_k-)}{\sum_{m=0}^{\infty} C_m(\tau_k-)} < U_k \leq \frac{\sum_{m=0}^{j} C_m(\tau_k-)}{\sum_{m=0}^{\infty} C_m(\tau_k-)}. \tag{6}$$

Let $C_0(\tau_k) := C_0(\tau_k-) + C_j(\tau_k-)$, and $C_m(\tau_k) := C_m(\tau_k-)$ if $1 \leq m < j$, and $C_m(\tau_k) := C_{m+1}(\tau_k-)$ if $m \geq j$, restoring the decreasing order.

Observe that $\sum_{m=0}^{\infty} C_m(t) = t$, but we did not use it to simplify the formulas (5) and (6) in the above definition because the analogous discrete assertion is not true; compare with (3) and (4).

The main result of this paper is that the normalized discrete processes converge in probability to $(\mathbf{C}(t))_{t\geq 0}$, uniformly on each finite time-interval, in terms of the distance defined by (1).

Theorem 1. *There exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, on which the discrete processes $(\mathbf{C}^{(n)}(i))_{i=0}^{\infty}$, $n = 1, 2, \dots$, and the continuous time process $(\mathbf{C}(t))_{t\geq 0}$ can be jointly realized so that if $T > 0$ is fixed and $f(n)$ is any function tending to infinity with n, then*

$$\mathbb{P}\left(\sup_{0\leq t\leq T} d\left(\mathbf{C}(t), \frac{\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}}\right) < \frac{f(n)}{\sqrt{n}}\right) \to 1 \quad \text{as } n \to \infty.$$

3.1. The convergence of the return process

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be such a probability space where a Poisson point process $(V_t)_{t\geq 0}$ with intensity $\rho(t) = t$ and the i.i.d. random variables $(U_k)_{k=1}^{\infty}$ and $(Z_i^{(n)})_{i,n=1}^{\infty}$ with uniform distribution on $[0, 1]$ are given independently of each other.

We have constructed the process $(\mathbf{C}(t))_{t\geq 0}$ from $(V_t)_{t\geq 0}$ and $(U_k)_{k=1}^{\infty}$ earlier. We first re-create the processes $(V_i^{(n)})_{i=0}^{\infty}$ with the appropriate distributions on the new probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The main idea of the construction is that we observe the process $(V_t)_{t\geq 0}$ in time intervals of length $1/\sqrt{n}$.

Let $X_i^{(n)} := \mathbb{1}\left(V_{i/\sqrt{n}} - V_{(i-1)/\sqrt{n}} \geq 1\right)$, $i = 1, 2 \dots$, $n = 1, 2, \dots$, be the indicators of the increase of the process $(V_t)_{t\geq 0}$, which are Bernoulli random variables with respective parameters

$$p_i^{(n)} = 1 - \exp\left(-\frac{2i-1}{2n}\right) = \frac{i}{n} + \mathcal{O}\left(\frac{i^2}{n^2}\right). \tag{7}$$

The required parameter for the increase of $V_i^{(n)}$ is

$$q_i^{(n)} = \frac{i}{n} - \frac{V_{i-1}^{(n)}}{n}. \tag{8}$$

We define the values of $V_i^{(n)}$ for fixed n with induction on i. Let $V_0^{(n)} := 0$, $n = 1, 2, \dots$, and

$$\begin{aligned} Y_i^{(n)} := X_i^{(n)} &- \mathbb{1}\left(p_i^{(n)} > q_i^{(n)}\right)\mathbb{1}\left(X_i^{(n)} = 1\right)\mathbb{1}\left(Z_i^{(n)} > \frac{q_i^{(n)}}{p_i^{(n)}}\right) \\ &+ \mathbb{1}\left(p_i^{(n)} < q_i^{(n)}\right)\mathbb{1}\left(X_i^{(n)} = 0\right)\mathbb{1}\left(Z_i^{(n)} < \frac{q_i^{(n)} - p_i^{(n)}}{1 - p_i^{(n)}}\right). \end{aligned} \tag{9}$$

We define $V_i^{(n)} := V_{i-1}^{(n)} + Y_i^{(n)}$.

It is easy to see that the distribution of the new $(V_i^{(n)})_{i=0}^{\infty}$ is in accordance with (2). Later on we say that a *correction* happens if the products of the indicators in (9) do not disappear. We will see that the total probability that a correction ever occurs is small if n is large enough. This gives an alternative proof of Proposition 1.

Lemma 1. *Let $T>0$ be fixed, denote by $0=\tau_0,\tau_1,\dots,\tau_\kappa$ the random times of jumps of the process $(V_t)_{0\le t\le T}$ and let $0=\tau_0^{(n)},\tau_1^{(n)},\dots,\tau_{\kappa^{(n)}}^{(n)}$ be those of the discrete process $(V^{(n)}_{\lfloor\sqrt{n}t\rfloor})_{0\le t\le T}$ defined above. Then, for sufficiently large n, with probability close to 1, the number of the jumps are equal: $\kappa=\kappa^{(n)}$. Furthermore, there exists a bijection between the jumps of the processes in such a way that*

$$|\tau_k-\tau_k^{(n)}|\le\frac{1}{\sqrt{n}},\quad k=1,\dots,\kappa,$$

hold simultaneously with probability tending to 1 as $n\to\infty$.

For technical convenience, we introduce the following events for fixed $\varepsilon,\delta>0$:

$$E_\varepsilon:=\{V^{(n)}_{\lfloor\sqrt{n}T\rfloor}\le K_\varepsilon,\ n=N_\varepsilon,N_\varepsilon+1,\dots\},$$

where K_ε is a sufficiently large constant and N_ε is a threshold satisfying $\mathbb{P}(E_\varepsilon)\ge 1-\varepsilon$. This makes sense by Proposition 1. Later we always suppose that $n>N_\varepsilon$. Let

$$M_\delta:=\left\{\min_{k:\tau_k\le T}\{\tau_k-\tau_{k-1}\}>\delta\right\}\cap\{V_T-V_{T-\delta}=0\},$$

where τ_k is the time of the kth jump of the process $(V_t)_{0\le t\le T}$ and $\tau_0=0$. It is easy to see that $\lim_{\varepsilon\downarrow 0}\mathbb{P}(E_\varepsilon)=\lim_{\delta\downarrow 0}\mathbb{P}(M_\delta)=1$.

Proof of Lemma 1: By definition, on the event M_δ the increment of the process $(V_t)_{0\le t\le T}$ on any interval $[i/\sqrt{n},(i+1)/\sqrt{n}]$ does not exceed 1 if $n>1/\delta^2$, hence $V_{i/\sqrt{n}}-V_{(i-1)/\sqrt{n}}=X_i^{(n)}$. Since $V^{(n)}_{i/\sqrt{n}}-V^{(n)}_{(i-1)/\sqrt{n}}=Y_i^{(n)}$, it is enough to prove that

$$\mathbb{P}\left(\{\exists i\le\lfloor\sqrt{n}T\rfloor : X_i^{(n)}\ne Y_i^{(n)}\}\cap E_\varepsilon\cap M_\delta\right)\to 0\quad(n\to\infty)$$

for all fixed $\varepsilon,\delta>0$.

On the event E_ε, $X_i^{(n)}=1$ can be true for at most K_ε many indices i. So the probability of the correction in the cases $p_i^{(n)}>q_i^{(n)}$ satisfies

$$1-\frac{q_i^{(n)}}{p_i^{(n)}}=\mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\quad(n\to\infty)$$

using the power series of the exponential function and equations (7) and (8) to estimate $p_i^{(n)}$ and $q_i^{(n)}$. If we add this at most K_ε many times, then the sum still goes to 0 as $n\to\infty$. A similar calculation shows that for an i, for which $p_i^{(n)}<q_i^{(n)}$ holds, the probability of the correction is at most

$$\frac{q_i^{(n)}-p_i^{(n)}}{1-p_i^{(n)}}=\mathcal{O}\left(\frac{1}{n}\right)\quad(n\to\infty).$$

Summing up for $i=1,\dots,\lfloor\sqrt{n}T\rfloor$, the total probability still tends to 0, as required.

3.2. Splits and merges

The processes $(V_i^{(n)})_{i=0}^{\infty}$ determine when splits or merges occur, and our task is now to describe how these events should happen. Similarly to the definition of the limit process $(\mathbf{C}(t))_{t\geq 0}$, we can prescribe the evolution of the discrete processes $(\mathbf{C}^{(n)}(i))_{i=0}^{\infty}$ with the use of the same independent uniform random variables $(U_k)_{k=1}^{\infty}$ as follows. Let $C_0^{(n)}(0) := 1, \quad C_m^{(n)}(0) := 0,\ m = 1, 2, \ldots$. The development of the process $\mathbf{C}^{(n)}$ in the steps $i = 1, 2, \ldots$ is described below:

- if $V_i^{(n)} - V_{i-1}^{(n)} = 0$, then $C_0^{(n)}(i) := C_0^{(n)}(i-1) + 1$ and the other coordinates are unchanged,
- if $V_i^{(n)} - V_{i-1}^{(n)} = 1$ and $V_i^{(n)} = k$, then the uniform random variable U_k determines a unique index j with probability 1 as in (6) via

$$\frac{\sum_{m=0}^{j-1} C_m^{(n)}(i-1)}{\sum_{m=0}^{\infty} C_m^{(n)}(i-1)} < U_k \leq \frac{\sum_{m=0}^{j} C_m^{(n)}(i-1)}{\sum_{m=0}^{\infty} C_m^{(n)}(i-1)}. \tag{10}$$

Similarly to the definition of the limit process

1. $j = 0$: ***split.*** If $U_k \sum_{m=0}^{\infty} C_m^{(n)}(i-1) < 1$, then let everything be unchanged: $\mathbf{C}^{(n)}(i) := \mathbf{C}^{(n)}(i-1)$; we call this case *fictitious split* (corresponding to the event that the stirring particle keeps its place). Otherwise, $C_0^{(n)}(i) := \lfloor U_k \sum_{m=0}^{\infty} C_m^{(n)}(i-1) \rfloor$, and the broken fragment $C_0^{(n)}(i-1) - \lfloor U_k \sum_{m=0}^{\infty} C_m^{(n)}(i-1) \rfloor$ is added to the collection of nonactive pieces $(C_m^{(n)}(i-1))_{m=1}^{\infty}$ to form the new ranked sequence $(C_m^{(n)}(i))_{m=1}^{\infty}$.
2. $j > 0$: ***merge.*** Let $C_0^{(n)}(i) := C_0^{(n)}(i-1) + C_j^{(n)}(i-1)$ and for the re-ranking put $C_m^{(n)}(i) := C_m^{(n)}(i-1)$ if $0 < m < j$, and $C_m^{(n)}(i) := C_{m+1}^{(n)}(i-1)$ if $m \geq j$.

It is easy to show that this new definition of $(\mathbf{C}^{(n)}(i))_{i=0}^{\infty}$ provides the same distribution as in the model generated by transpositions, so we prove the convergence for these processes.

Proof of Theorem 1: Let $\varepsilon, \delta > 0$ be fixed. Let A_n denote the event that the assertion of Lemma 1 holds for $(V_{\lfloor\sqrt{n}t\rfloor}^{(n)})_{0\leq t\leq T}$. We restrict ourselves to the events $E_\varepsilon \cap M_\delta \cap A_n$. Let us define a measure (which is not a probability measure) on the sets $B \in \mathcal{F}$:

$$\mathbb{P}_{\varepsilon,\delta,n}(B) := \mathbb{P}(B \cap E_\varepsilon \cap M_\delta \cap A_n).$$

By Lemma 1, it is enough to show that for fixed $\varepsilon, \delta > 0$ the processes $\mathbf{C}(t)$ and $\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)/\sqrt{n}$ are sufficiently close to each other for large n except a set with $\mathbb{P}_{\varepsilon,\delta,n}$-measure tending to 0 as $n \to \infty$. The proof consists of the following steps:

1. We estimate the increase of the distance between $\mathbf{C}(t)$ and $\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)/\sqrt{n}$ between two successive split or merge events.

2. We introduce those events when the distance under discussion cannot be estimated: the awkward events (defined later) and the fictitious splits. We show that they have small probability.

3. On the complementary event, which has probability tending to 1 as $n \to \infty$, we show that a merge does not increase the distance between $\mathbf{C}(t)$ and $\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)/\sqrt{n}$ very much.

4. We do this also for the splits.

5. We summarize the estimates.

STEP 1. Let

$$d_k^- := d\left(\mathbf{C}(\tau_k-), \frac{\mathbf{C}^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)}{\sqrt{n}}\right), \quad d_k^+ := d\left(\mathbf{C}(\tau_k), \frac{\mathbf{C}^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor)}{\sqrt{n}}\right)$$

denote the distance between the discrete and continuous processes before and after the time of the kth split or merge. Recall that τ_k is the time of the kth jump of $(V_t)_{0\le t\le T}$ and $\tau_k^{(n)}$ is that of $(V^{(n)}_{\lfloor\sqrt{n}t\rfloor})_{0\le t\le T}$, which are close $\mathbb{P}_{\varepsilon,\delta,n}$-almost surely by Lemma 1.

While no split or merge occurs, the distance between the processes does not increase very much. From Lemma 1, the difference between τ_k and $\tau_k^{(n)}$ can be at most $1/\sqrt{n}$. The discrete processes $(\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)/\sqrt{n})_{t\ge 0}$ change only in the times which are multiples of $1/\sqrt{n}$. Thus $\mathbb{P}_{\varepsilon,\delta,n}$-almost surely

$$d_k^- \le d_{k-1}^+ + \frac{2}{\sqrt{n}}. \tag{11}$$

STEP 2. From now on we investigate only the split or merge points of the processes. At the kth time of jump of $(V_t)_{0\le t\le T}$ and $(V^{(n)}_{\lfloor\sqrt{n}t\rfloor})_{0\le t\le T}$ we choose with the help of U_k one of the components of $\mathbf{C}(\tau_k-)$ and $\mathbf{C}^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)$ via (6) and (10). Let us call the possibility that these components are of different indices an *awkward event*. If an awkward event or a fictitious split (which means that $U_k\sum_{m=0}^{\infty} C_m^{(n)}(i-1) < 1$) occurs, then we cannot estimate $d(\mathbf{C}, \mathbf{C}^{(n)}/\sqrt{n})$. We will see that these events have probability tending to 0 as $n \to \infty$.

We can choose the components of $\mathbf{C}$ and $\mathbf{C}^{(n)}$ as follows. We construct two partitions of $[0,1]$ into subintervals. Let the partition $\mathcal{P}$ contain intervals of length equal to the coordinates of the vector $\mathbf{C}(\tau_k-)/\sum_{m=0}^{\infty} C_m(\tau_k-)$. Then the right end point of the jth subinterval in this partition is $\sum_{m=0}^{j} C_m(\tau_k-)/\sum_{m=0}^{\infty} C_m(\tau_k-)$, $j = 0,1,2,\ldots$. We do this at the same time with dividing points $\sum_{m=0}^{j} C_m^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)/\sum_{m=0}^{\infty} C_m^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)$, $j = 0,1,2,\ldots$ to obtain the partition $\mathcal{P}^{(n)}$. Let W_k denote the set of those points in $[0,1]$ which are covered by the subintervals of $\mathcal{P}$

and $\mathcal{P}^{(n)}$ of different indices. The probability of the awkward events (which is an upper estimate for their $\mathbb{P}_{\varepsilon,\delta,n}$-measure) is exactly the Lebesgue measure of W_k.

We know that $\sum_{m=0}^{\infty} C_m(t) = t$ for all $t \geq 0$. From the construction,

$$\frac{\lfloor\sqrt{n}t\rfloor - K_\varepsilon}{\sqrt{n}} \leq \frac{\sum_{m=0}^{\infty} C_m^{(n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}} \leq t \quad \text{if} \quad t \in [0,T], \tag{12}$$

because at the split or merge points (occurring at most K_ε many times) the total length of the discrete process does not increase. Hence, by Lemma 1,

$$\begin{aligned}\left|\sum_m C_m(\tau_k-) - \frac{\sum_m C_m^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)}{\sqrt{n}}\right| \\ \leq \left|\tau_k - \frac{\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1 - K_\varepsilon}{\sqrt{n}}\right| \leq \frac{K_\varepsilon}{\sqrt{n}} + \frac{3}{\sqrt{n}}.\end{aligned} \tag{13}$$

It is an elementary exercise to show from the above that the distance between the corresponding dividing points of the partitions $\mathcal{P}$ and $\mathcal{P}^{(n)}$ can be, respectively, at most

$$\frac{d_k^- + \frac{K_\varepsilon+3}{\sqrt{n}}}{\tau_k}.$$

Since the number of coordinates is at most K_ε, this implies

$$\mathrm{Leb}(W_k) \leq \frac{d_k^- + \frac{K_\varepsilon+3}{\sqrt{n}}}{\tau_k} K_\varepsilon \leq \frac{d_k^- + \frac{K_\varepsilon+3}{\sqrt{n}}}{\delta} K_\varepsilon,$$

where we used the fact that $\tau_k = \sum_{m=0}^{\infty} C_m(\tau_k-) \geq \delta$ for all $k = 1, 2, \dots$ on the event M_δ. This yields

$$\mathbb{P}_{\varepsilon,\delta,n}(\text{awkward event at } \tau_k) \leq \mathrm{Leb}(W_k) \leq \frac{d_k^- + \frac{K_\varepsilon+3}{\sqrt{n}}}{\delta} K_\varepsilon.$$

Furthermore,

$$\mathbb{P}_{\varepsilon,\delta,n}(\text{fictitious split at } \tau_k^{(n)}) \leq \frac{\frac{1}{\sqrt{n}}}{\sum_{m=0}^{\infty} C_m^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)/\sqrt{n}} \leq \frac{2}{\delta\sqrt{n}},$$

if n is large enough, by (12). So we conclude that

$$\begin{aligned}&\mathbb{P}_{\varepsilon,\delta,n}(\text{awkward event or fictitious split at the } k\text{th split or merge point}) \\ &\qquad\qquad \leq \frac{K_\varepsilon}{\delta} d_k^- + \frac{K_\varepsilon^2 + 3K_\varepsilon + 2}{\delta\sqrt{n}}.\end{aligned}$$

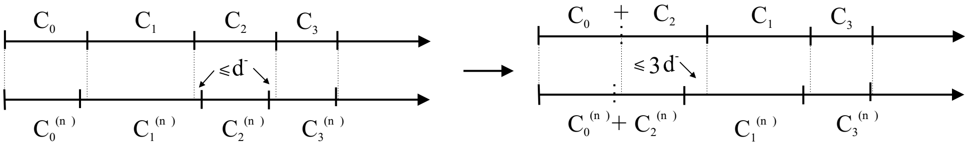

Figure 3: The piece C_2 merges C_0 parallel with the $C_2^{(n)} - C_0^{(n)}$ coagulation

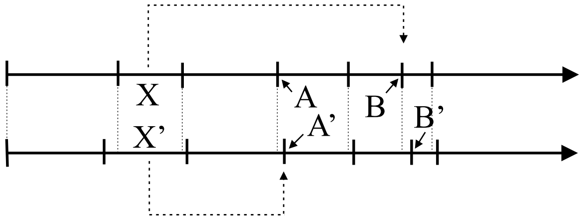

Figure 4: Split: the broken pieces from the coordinate 0 are X and X' which have to be moved to places B and A'

STEP 3. In the case when the random variable U_k chooses the same components of $\mathbf{C}$ and $\mathbf{C}^{(n)}$ and it is not the active coordinate, i.e. there is a merge in both processes (see Figure 3), then

$$d_k^+ \leq 3d_k^-. \tag{14}$$

STEP 4. If a (non-fictitious) split occurs in the discrete and continuous processes, then we have

$$\left| C_0(\tau_k) - \frac{C_0^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor)}{\sqrt{n}} \right|$$
$$\leq \left| U_k \sum_{m=0}^{\infty} C_m(\tau_k-) - \frac{\lfloor U_k \sum_{m=0}^{\infty} C_m^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)\rfloor}{\sqrt{n}} \right| \leq \frac{K_\varepsilon + 3}{\sqrt{n}} + \frac{1}{\sqrt{n}}.$$

using inequality (13). This is why the broken pieces from the components $C_0(\tau_k-)$ and $C_0^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)/\sqrt{n}$ (denoted by X and X' on Figure 4) can differ by at most $d_k^- + \frac{K_\varepsilon+4}{\sqrt{n}}$: the difference can be $\frac{K_\varepsilon+4}{\sqrt{n}}$ between the left end points and at most d_k^- between the right end points.

It is possible that the two broken pieces do not come to the same place in the decreasing order of the coordinates. This case is shown on Figure 4. Let x be the number telling how many pieces are between the present place of X and its place in the decreasing order, and define x' similarly for X' ($x = 3, x' = 1$ on the figure). Move first X and X' $\min(x, x')$ places to the right (to the places A and A' on the figure). The result is two vectors: the modifications of $\mathbf{C}(\tau_k-)$ and $\mathbf{C}^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)/\sqrt{n}$, but one of them is not necessarily in decreasing order.

Because $|A - A'| \le d_k^-$ and $\big|\,|X| - |X'|\,\big| \le d_k^- + \frac{K_\varepsilon+4}{\sqrt{n}}$, the $d(\cdot,\cdot)$-distance of these two vectors is at most $2d_k^- + \frac{K_\varepsilon+4}{\sqrt{n}}$ bigger than before this modification.

In the second step we move X from A to B. The lengths of the parts between A and B are at least $|X|$ and at most $|X'| + 2d_k^- \le |X| + 3d_k^- + \frac{K_\varepsilon+4}{\sqrt{n}}$. So any two of these parts have lengths that differ by at most $3d_k^- + \frac{K_\varepsilon+4}{\sqrt{n}}$. If we swap X always with its right neighbour until it hits place B, then we see that the number of the swaps is at most K_ε, and at each swap the distance can increase by at most $3d_k^- + \frac{K_\varepsilon+4}{\sqrt{n}}$, so we have

$$d_k^+ \le 2d_k^- + \frac{K_\varepsilon + 4}{\sqrt{n}} + K_\varepsilon\left(3d_k^- + \frac{K_\varepsilon + 4}{\sqrt{n}}\right). \tag{15}$$

STEP 5. Summing up the estimates (11), (14) and (15) we get easily the following recursive bound:

$$d_k^+ \le \max(3, 2+3K_\varepsilon)d_{k-1}^+ + \frac{K_\varepsilon^2 + 5K_\varepsilon + 4 + 2\max(3, 2+3K_\varepsilon)}{\sqrt{n}} =: ad_{k-1}^+ + \frac{b}{\sqrt{n}}.$$

Hence $\sup_{0\le k\le K_\varepsilon} d_k^\pm \le \left(\sum_{i=0}^{K_\varepsilon} ba^i\right)/\sqrt{n}$. Considering the results of steps 1 and 2, the assertion of the theorem follows.

4. Stationary distribution and generalizations

It is a natural question to identify the stationary distribution of the stirring process generated by transpositions. This means that we look at the asymptotic behaviour of the process $(\mathbf{C}^{(n)}(\lfloor nt\rfloor)/n)_{t\ge 0}$. Observe that the time scale is of order n, i.e. the time scale when the stirring element has already visited the bulk of the n places. This setup is the same as that of the problem studied by Schramm in [11], but different from the phenomenon described by Theorem 1.

In this section we consider the following split-and-merge transformation corresponding to the stirring generated by random transpositions. Let $\mathbf{C} = (C_0, C_1, C_2, \dots) \in \mathbf{S}$ be a random probability distribution, i.e. $\sum_m C_m = 1$ almost surely, where C_0 is the active component. Let U be a random variable with uniform distribution of $[0,1]$ which is independent of $\mathbf{C}$. If $U \le C_0$, then C_0 splits, i.e. the new active component will be U and $(C_0 - U, C_1, C_2, \dots)$ will be the remaining components after restoring the decreasing order. If $\sum_{m=0}^{j-1} C_m < U \le \sum_{m=0}^{j} C_m$, then C_0 merges with C_j similarly as in (3)–(6) because $\sum_m C_m = 1$.

In limit theorems of random partitions and permutations the following distribution appears often. Let the random variables $W_1, W_2, \dots$ be independent with uniform distribution on $[0,1]$. Let $(Q_1, Q_2, \dots)$ be the decreasing rearrangement of the random variables

$$(P_1, P_2, \dots) := (W_1, (1-W_1)W_2, (1-W_1)(1-W_2)W_3, \dots).$$

Then the random sequence $(P_1, P_2, \dots)$ has GEM(1) distribution after Griffiths, Engen and McCloskey and $(Q_1, Q_2, \dots)$ has *Poisson–Dirichlet* distribution with parameter 1, abbreviated PD(1). For more about this family of distributions see [9].

Let $(p_1, p_2, \dots)$ be a random probability distribution. We construct its size-biased permutation. Let $U_1, U_2, \dots$ be i.i.d. uniform random variables on $[0,1]$, independent of $(p_1, p_2, \dots)$. Let I_j be the unique index for which $\sum_{i=1}^{I_j-1} p_i \le U_j < \sum_{i=1}^{I_j} p_i$. Let J_k denote the kth smallest integer m satisfying $I_m \notin \{I_1, I_2, \dots, I_{m-1}\}$. Then the vector $(p_{J_1}, p_{J_2}, \dots)$ is called the ***size-biased permutation*** of $(p_1, p_2, \dots)$. It is well known that the size-biased permutation of a random partition with PD(1) distribution has GEM(1) distribution. See also [10].

Consider the following probability distribution on $\mathbf{S}$. Let $(Q_1, Q_2, \dots)$ have PD(1) distribution. Let C_0 be a size-biased part from $(Q_1, Q_2, \dots)$ (i.e. the first component of the size-biased permutation of $(Q_1, Q_2, \dots)$) corresponding to the active cycle and the rest $(C_1, C_2, \dots)$ is the vector of the remaining Q_j-s in non-increasing order. We denote by μ the distribution of $\mathbf{C} = (C_0, C_1, C_2, \dots)$.

Theorem 2. *The distribution μ is invariant under the above split-and-merge transformation.*

Proof: By definition, a random partition $\mathbf{C}$ with distribution μ can be considered as follows. Let $W_1, W_2, \dots$ be i.i.d. uniform random variables on $[0,1]$ as in the definition of PD(1). Because the size-biased permutation of PD(1) is GEM(1), we can suppose that for the active component $C_0 = W_1$ holds and $(C_1, C_2, \dots)$ is the decreasing rearrangement of $((1-W_1)W_2, (1-W_1)(1-W_2)W_3, \dots)$. Let ν be the distribution of the random partition obtained by the application of a stirring step to $\mathbf{C}$.

If $U < W_1$ for the $[0,1]$-uniform random variable U, then the new non-active components are $(W_1 - U, (1-W_1)W_2, (1-W_1)(1-W_2)W_3, \dots)$ in decreasing order. Conditionally on $\{U < W_1\}$ and on U, the variable W_1 is uniform on $[U, 1]$, thus the vector of the non-active components has PD(1) distribution scaled by $(1-U)$. This yields that ν, conditioned on $\{U < W_1\}$ and on U, is the same as μ conditioned on the active component having size U.

If $U > W_1$, then a coagulation occurs. Conditioned on $\{U > W_1\}$ and on the value of W_1, the size of the component which merges with C_0 has uniform distribution on $[0, 1-W_1]$ because it is a size-biased component. We get the same distribution if we choose this component merging C_0 to be of length $U - W_1$. Conditionally on $\{U > W_1\}$ and on U, the rest has PD(1) distribution scaled by $(1-U)$. Thus, a sample from ν, conditioned on $\{U > W_1\}$ and on U, has an active coordinate of size U and the remaining components with a scaled PD(1) distribution.

Hence, a vector with distribution ν can be obtained by sampling U uniformly on $[0,1]$, taking the active coordinate of length U and taking a scaled PD(1) distribution on the rest. This shows that $\nu = \mu$, as required.

Theorem 2 proves that μ is a stationary measure for our process, but it is not at all clear if this is a *unique* stationary distribution. The proof of this would be the analogue of Schramm's result in [11].

A possible generalization of the model studied in this paper is *multiple stirring.* It means that we consider more than one stirring particles. For a fixed number k of stirring elements an analogous limit theorem can be proved with a coupling similar to that in Theorem 1. The case when the number of the stirring elements depends on the size of the set $[n]$ might also be worth studying (for example with $k(n) = n^\alpha$ where $0 < \alpha < 1$). Of course, we would need different scaling of time and space in this case.

An open question is for our original model to establish after how much time a permutation can be regarded as a random permutation chosen with the uniform distribution, if it can so be regarded at all. The solution to this problem is not obvious in the least. See [4] for more about this problem in similar models .

Acknowledgement. I thank Bálint Tóth and Benedek Valkó for initiating these investigations and for their permanent support and useful comments while writing this paper. I am grateful to the referee for pointing out the proof of Theorem 2, and I thank Sándor Csörgő for his helpful remarks on this paper.

References

[1] O. Angel, Random infinite permutations and the cyclic time random walk, *Discrete Mathematics and Theoretical Computer Science*, Nancy, pages 9-16, 2003.

[2] N. Berestycki, R. Durrett, A phase transition in the random transposition random walk, To appear in *Probability Theory and Related Fields*, 2005.

[3] M. Camarri, J. Pitman, Limit distributions and random trees derived from the birthday problem with unequal probabilities, *Electronic Journal of Probability*, vol. 5, paper no. 2, pages 1-18, 2000.

[4] P. Diaconis, *Group Representations in Probability an Statistics*, IMS Lecture notes – Monograph Series, vol. 11, Institute of Mathematical Statistics, Hayward, California, 1988.

[5] E. Mayer-Wolf, O. Zeitouni, M. Zerner, Asymptotics of certain coagulation–fragmentation processes and invariant Poisson – Dirichlet measures, *Electronic Journal of Probability*, vol. 7, pages 1-25, 2002.

[6] P. Erdős, Some remarks on the theory of graphs, *Bulletin of the American Mathematical Society*, **53**, 292-294, 1947.

[7] P. Erdős, A. Rényi, On random graphs I, *Publicationes Mathematicae Debrecen*, **6**, 290-297, 1959.

[8] S. JANSON, T. ŁUCZAK, A. RUCIŃSKI, *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, New York, Wiley-Interscience, 2000.

[9] J. PITMAN, *Combinatorial Stochastic Processes*, Technical Report no. 621, Dept. Statistics, U. C. Berkeley, Lecture notes for St. Flour course, 2002.

[10] J. PITMAN, Poisson – Dirichlet and GEM invariant distributions for split-and-merge transformations of an interval partition, *Combinatorics, Probability and Computing*, **11**, 501-514, 2002.

[11] O. SCHRAMM, Compositions of random transpositions, *Israel Journal of Mathematics*, vol. **147**, 221-244, 2005.

[12] V. E. STEPANOV, The probability of connectedness of a random graph $\mathcal{G}_m(t)$, *Teor. Veroyatnost. i Primenen.* **15**, 55-68 (Russian); English translation *Theory of Probability and Its Applications*, **15**, 55-67, 1970.

[13] B. TÓTH, Improved lower bound on the thermodynamic pressure of the spin 1/2 Heisenberg ferromagnet, *Letters in Mathematical Physics*, **28**(1):75-84, 1993.

B. VETŐ, Institute of Mathematics, Technical University Budapest, Egry J. u. 1, 1111 Budapest, Hungary; *e-mail:* vetob@math.bme.hu