DIPLOMA THESIS

The time evolution of permutations under random stirring

BÁLINT VETŐ

Supervisor: PROF. BÁLINT TÓTH

Department of Stochastics
Institute of Mathematics
Budapest University of
Technology and Economics

BME
2006

Contents

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 $\pi^{(n)}(i) = T_i^{(n)} \circ T_{i-1}^{(n)} \circ \cdots \circ T_1^{(n)}$ be a permutation acting on the set $[n] := \{1, \ldots, n\}$. 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 $\{v, \sigma(v), \sigma^2(v), \ldots\} \subseteq S$ 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 which 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 $\pi^{(n)}(i)$ restricting ourself to the conjugation with permutations fixing 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

$$\mathbf{C}^{(n)}(i) := (C_0^{(n)}(i), C_1^{(n)}(i), C_2^{(n)}(i), \ldots) \tag{1}$$

where $C_0^{(n)}(i)$ denotes the length of the active cycle, $C_1^{(n)}(i), C_2^{(n)}(i), \ldots$ the lengths of those cycles in non-increasing order which are already moved by one of the transpositions $(T_j^{(n)})_{j=1}^{i}$. Other $C_j^{(n)}(i)$-s are 0.

Note that with this definition we omit $n - \mathcal{O}(i)$ fixed points of the permutation $\pi^{(n)}(i)$, which are not yet involved to the stirring mechanism. These singletons can be considered as dust, which builds up the cycles. We will see

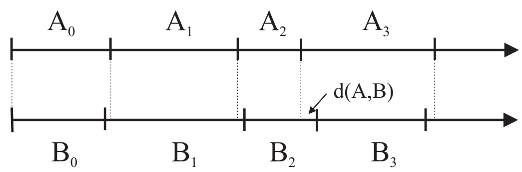

Figure 1: The metric on $\mathbf{S}$

that discarding components with value 1 does not change our limit because of the normalization. Our benefit is that $\sum_{j=0}^{\infty} C_j^{(n)}(i)$ gives the number of elements already moved.

$(\mathbf{C}^{(n)}(i))_{i=0}^{\infty}$ is a process on the state space

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

with the distance

$$d(A, B) := \sup \left\{ \left| \sum_{j=0}^{k} A_j - \sum_{j=0}^{k} B_j \right| : k = 0, 1, 2, \dots \right\} \quad (3)$$

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 (meaning 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

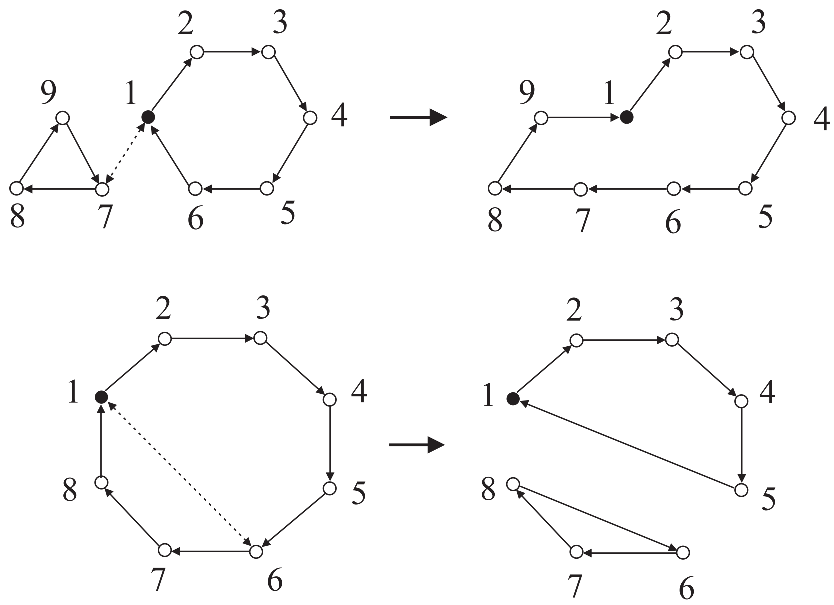

Figure 2: Coagulation and fragmentation of cycles

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.

A natural modification of the model is to consider more than one stirring particles. It is called multiple stirring, a limit theorem is stated in Section 4, which is similar to the case of one stirring element.

Section 5 discusses general coagulation-fragmentation processes. Some basic features of partition-valued distributions (e.g. the Poisson – Dirichlet distribution) are presented. Finally we introduce a split-and-merge transformation of partitions of $[0,1]$, which is in a sense the limit of the coagulation-fragmentation mechanism generated by random transpositions. The stationary measure of this random transformation is the adequate modification of the Poisson – Dirichlet distribution.

2 Related results

The issue of random graphs was first brought up by Erdős [7] in the late 1940's. The random graph $G(n,p)$ is a random variable which is a graph on the vertex set $[n]$, and each pair of points is connected by an edge with probability p, independently.

The classical Erdős – Rényi Theorem [8] describes the asymptotic behaviour of the largest connected component of the random graph $G(n, c/n)$ as $n \to \infty$ depending on the constant c. At the parameter value $c = 1$ we find a phase transition in the limit. If $c < 1$, then the size of the largest connected component is $\mathcal{O}(\log n)$ asymptotically almost surely. In $G(n, 1/n)$ we find components with size $\mathcal{O}(f(n))$ for any function $f(n)$ growing faster than $n^{2/3}$. If $c > 1$, then asymptotically almost surely there is a giant component with asymptotic size βn where β is the survival probability of a Galton – Watson branching process which has Poisson offspring distribution with mean c. In other words $\beta \in (0,1)$ is the solution of the equation $1 - \beta = e^{-\beta c}$. For more information about random graphs see e.g. [9]. A possible generalization is the random graph process in [14].

The random permutation model described by Schramm in [13] is similar to the one defined in the introduction. In that paper $(T_i)_{i=1}^{\infty}$ are independent random transpositions with uniform distribution from all possible transpositions of the set $[n]$. The limit distribution of the proportions of the giant cycles in the permutation $\Pi(cn) = T_{cn} \circ \cdots \circ T_1$ is identified as $n \to \infty$ where $c > 1/2$ is a constant.

The following informal argument shows that the phase transition in the Erdős – Rényi graph model at $c = 1$ is the same as the phase transition here at $c = 1/2$. In the critical random graph $G(n, 1/n)$ the number of edges has binomial distribution with parameters $\binom{n}{2} \sim n^2/2$ and $1/n$. The expectation is asymptotically $n/2$, which yield that if we have a phase transition, then it can be found at the parameter value $c = 1/2$.

This result is in accordance with the theory of random graphs. Let us consider the random graph $G(n)$ on the vertex set $[n]$ where $\{u, v\}$ is an edge if and only if the transposition (u, v) appears among $T_1, \ldots, T_{cn}$. By the Erdős –

Rényi Theorem the graph $G(n)$ has a giant connected component only in the case $c > 1/2$ similarly to the condition on the random permutations.

The random transposition model is motivated by Tóth in connection with the quantum-physical applications of the problem. He introduces the *random stirring process* on the lattice graph $\mathbb{Z}^d$ in [15]. The trajectory of a given particle is then a random walk with long memory, and we may ask the condition of recurrence. The model can be considered on different graphs as well: the complete graph at Schramm, the star graph in our case. Angel analyzed Tóth's random walk model on regular trees and proved the appearance of phase transition in [1].

Schramm proves that the large cycles of $\Pi(cn)$ normalized by the sum of their lengths have a distribution that is close to the distribution of the large cycles of a uniform random permutation. More precisely the vector of the cycle sizes of $\Pi(cn)$ in non-increasing order normalized by the magnitude of the giant connected component of $G(n)$ converges in distribution to Poisson – Dirichlet distribution with parameter 1 as $n \to \infty$, if $c > 1/2$ (cf. Proposition 4 in this thesis).

The proof relies on a coupling, which can be applied to justify the uniqueness of the stationary distribution for the continuous uniform coagulation-fragmentation process. This is a discrete time Markov-chain on the random probability distributions. Given $Y = (Y_1, Y_2, \dots)$ with $Y_1 \geq Y_2 \geq \dots \geq 0$ and $\sum_i Y_i = 1$ almost surely, we choose two independent random indices $I, J \in \mathbb{N}$ with $\mathbb{P}(I = k \mid Y) = \mathbb{P}(J = k \mid Y) = Y_k$. If $I \neq J$, then let Y' be obtained from Y by replacing the two entries Y_I and Y_J with the single entry $Y_I + Y_J$ and resorting. If $I = J$, then given (Y, I, J) a random variable U is selected uniformly in $[0, Y_I]$, and Y' is obtained by splitting the entry Y_I into the two entries U and $Y_I - U$ and resorting. Then Y' is the new state of the Markov-chain.

Mayer-Wolf, Zeitouni and Zerner in [6] give a characterization of invariant measures for a wider class of continuous coagulation-fragmentation processes.

Berestycki and Durrett [2] also study the random transposition walk on the group of permutations of $[n]$. Contrary to Schramm they consider this walk in continuous time starting at the identity. Let D_t be the minimum number of transpositions needed to go back to the identity from the location at time t. The main result is the phase transition of this quantity: $D_{cn} \sim cn$ if $c \leq 1/2$ and $D_{cn} \sim u(c)n$ with $u(c) < c$ if $c > 1/2$ as $n \to \infty$.

The results of Sections 3 and 5.3 are published in [16].

3 Limit theorem

First we consider the coagulation and fragmentation events together. We present a limit theorem for their occurrence. Then a stronger theorem is stated, which will be proved with coupling, and it incorporates also our first assertion.

Coagulations and fragmentations of non-trivial cycles occur exactly in those steps when the stirring element returns to a place already visited. It motivates the following definitions. 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

$$V_i^{(n)} := \#\{k : k \leq i, \exists j < k : B_j^{(n)} = B_k^{(n)}\} \tag{4}$$

be the number of the *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

$$\mathbb{P}\left(V_{i+1}^{(n)} - V_i^{(n)} = 1 | V_i^{(n)}\right) = 1 - \mathbb{P}\left(V_{i+1}^{(n)} - V_i^{(n)} = 0 | V_i^{(n)}\right) = \frac{i + 1 - V_i^{(n)}}{n}. \tag{5}$$

In order to get a non-trivial limit distribution, the time of the processes should be accelerated. As opposed to Schramm [13], in Theorem 1 the scaling will be $\sqrt{n}$. 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 limit of the vectors $\left(\frac{\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}}\right)_{t\geq 0}$ as $n\to\infty$, where the division is meant coordinatewise, namely $\frac{\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}} := (\frac{C_0^{(n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}}, \frac{C_1^{(n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}}, \dots)$. 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 [4].

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

$$(V^{(n)}_{\lfloor\sqrt{n}t\rfloor})_{t\geq 0} \stackrel{\mathrm{d}}{\Longrightarrow} (V_t)_{t\geq 0} \qquad (n\to\infty) \tag{6}$$

in terms of the finite dimensional marginal distributions.

3.1 Coupling

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

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

$$\frac{C_0^{(n)}(i-1)}{\sum_{m=0}^{\infty} C_m^{(n)}(i-1)}. \tag{7}$$

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

$$\frac{C_j^{(n)}(i-1)}{\sum_{m=0}^{\infty} C_m^{(n)}(i-1)}. \tag{8}$$

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 above process. It is built on a Poisson point process $(V_t)_{t\geq 0}$ with intensity $\rho(t) = t$. Similarly to the discrete processes $(\mathbf{C}^{(n)}(i))_{i=0}^{\infty}$ at the times 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 the following: 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 $V_{\tau_k-} = \lim_{\varepsilon\to 0} V_{\tau_k-\varepsilon} = k-1$. 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 next two actions occurs at time τ_k.

1. *Split:* If
$$U_k \leq \frac{C_0(\tau_k-)}{\sum_{m=0}^{\infty} C_m(\tau_k-)}, \tag{9}$$
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
$$\frac{\sum_{m=0}^{j-1} C_m(\tau_k-)}{\sum_{m=0}^{\infty} C_m(\tau_k-)} < U_k \leq \frac{\sum_{m=0}^{j} C_m(\tau_k-)}{\sum_{m=0}^{\infty} C_m(\tau_k-)}. \tag{10}$$
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 (9) and (10) in the above definition because the analogous discrete assertion is not true, compare with (7) and (8).

The normalized discrete processes converge in probability to $(\mathbf{C}(t))_{t\geq 0}$ in the following uniform sense in terms of the distance defined by (3).

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*
$$\mathbb{P}\left(\sup_{0\leq t\leq T} d\left(\mathbf{C}(t), \frac{\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}}\right) < \frac{f(n)}{\sqrt{n}}\right) \to 1 \quad \text{as } n\to\infty. \tag{11}$$

Proof. 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 $\frac{1}{\sqrt{n}}$ long time intervals.

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

$$p_i^{(n)} = 1 - \exp\left(-\frac{2i-1}{2n}\right) = \frac{i}{n} + \mathcal{O}\left(\left(\frac{i}{n}\right)^2\right). \tag{12}$$

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

$$q_i^{(n)} = \frac{i}{n} - \frac{V_{i-1}^{(n)}}{n}. \tag{13}$$

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

$$\begin{aligned} Y_i^{(n)} := X_i^{(n)} &- \mathbb{1}\left(p_i^{(n)} > q_i^{(n)}\right) \mathbb{1}\left(X_i^{(n)} = 1\right) \mathbb{1}\left(Z_i^{(n)} > \frac{q_i^{(n)}}{p_i^{(n)}}\right) \\ &+ \mathbb{1}\left(p_i^{(n)} < q_i^{(n)}\right) \mathbb{1}\left(X_i^{(n)} = 0\right) \mathbb{1}\left(Z_i^{(n)} < \frac{q_i^{(n)} - p_i^{(n)}}{1 - p_i^{(n)}}\right). \end{aligned} \tag{14}$$

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 (5). Later on we say that a *correction* happens if the products of the indicators in (14) do not vanish. In Subsection 3.2 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. Afterwards we show that for large n after splits and merges the distance between the continuous and discrete processes remains small with large probability, which completes the proof of Theorem 1.

3.2 The convergence of the return process

Lemma 1. *Let $T > 0$ be fixed and denote $0 = \tau_0, \tau_1, \dots, \tau_\kappa$ the random times of jumps of the process $(V_t)_{0\le t\le T}$ and denote $0 = \tau_0^{(n)}, \tau_1^{(n)}, \dots, \tau_{\kappa^{(n)}}^{(n)}$ that 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 numbers of the jumps are equal: $\kappa = \kappa^{(n)}$. Furthermore, there exists a bijection between the jumps of the processes in such a way that*

$$|\tau_k - \tau_k^{(n)}| \le \frac{1}{\sqrt{n}} \quad k = 1, \dots, \kappa \tag{15}$$

holds with probability tending to 1 as $n \to \infty$.

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

$$E_\varepsilon := \{V^{(n)}_{\lfloor\sqrt{n}T\rfloor} \le K_\varepsilon \quad n = N_\varepsilon, N_\varepsilon + 1, \dots\}, \tag{16}$$

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

$$M_\delta := \left\{ \min_{j:\tau_j \le T} \{\tau_j - \tau_{j-1}\} > \delta \right\} \cap \{V_T - V_{T-\delta} = 0\}, \tag{17}$$

where τ_j is the time of the jth jump of the process $(V_t)_{0\le t\le T}$ and $\tau_0 = 0$. $\lim_{\varepsilon\to 0} \mathbb{P}(E_\varepsilon) = 1$ by definition and

Proposition 2.

$$\lim_{\delta\to 0} \mathbb{P}(M_\delta) = 1. \tag{18}$$

Proof. Since the estimate

$$\mathbb{P}(M_\delta) \ge \mathbb{P}\left(\min_{j:\tau_j \le T} \{\tau_j - \tau_{j-1}\} > \delta \right) - \mathbb{P}(V_T - V_{T-\delta} > 0) \tag{19}$$

and $\lim_{\delta\to 0} \mathbb{P}(V_T - V_{T-\delta} > 0) = 0$ trivially holds, we will show for fixed $k \in \mathbb{N}$ that

$$\lim_{\delta\to 0} \mathbb{P}\left(\min_{j:\tau_j \le T} \{\tau_j - \tau_{j-1}\} > \delta \,\middle|\, V_T = k \right) = 1. \tag{20}$$

We give a lower bound for $\mathbb{P}\left(\min_{j:\tau_j\le T}\{\tau_j-\tau_{j-1}\}>\delta \;\middle|\; V_T=k\right)$ as follows. We partition the interval $[0,T]$ into subintervals of different lengths so that each subinterval has a weight $T\delta$ computed by integration of the intensity function $\rho(t)=t$ corresponding to the Poisson point process $(V_t)_{t\ge 0}$. Then each subinterval has length at least δ, and the number of subintervals is $\lfloor\frac{t^2/2}{t\delta}\rfloor=\lfloor t/2\delta\rfloor$ considering only subintervals with entire weight (probably we omit a remaining interval at the right end of $[0,T]$ with weight less than $T\delta$).

It is a subset of the event M_δ, if all the k times of jumps of the process $(V_t)_{0\le t\le T}$ fall to different subintervals which are not even neighbouring, and the first subinterval is empty. We can indicate in $\binom{\lfloor t/2\delta\rfloor-k}{k}$ different ways those subintervals which will contain a time of jump corresponding to the previous condition. The distribution of the number of τ_j-s falling to certain subintervals, conditioned on the event $\{V_T=k\}$, is multinomial. Hence the conditional probability that a given configuration of the times of jumps is realized equals $k!(t/2\delta)^{-k}$. It yields

$$\begin{aligned}\lim_{\delta\to 0}\mathbb{P}\left(\min_{j:\tau_j\le T}\{\tau_j-\tau_{j-1}\}>\delta \;\middle|\; V_T=k\right)&\ge\binom{\lfloor t/2\delta\rfloor-k}{k}k!\left(\frac{t}{2\delta}\right)^{-k}\\&=\prod_{i=0}^{k-1}\frac{\lfloor t/2\delta\rfloor-k-i}{t/2\delta}\to 1\qquad\text{as}\quad\delta\to 0.\qquad(21)\end{aligned}$$

For arbitrary $\varepsilon>0$

$$\begin{aligned}&\mathbb{P}\left(\min_{j:\tau_j\le T}\{\tau_j-\tau_{j-1}\}>\delta\right)\\&\qquad\ge\sum_{k=0}^{K_\varepsilon}\mathbb{P}\left(\min_{j:\tau_j\le T}\{\tau_j-\tau_{j-1}\}>\delta \;\middle|\; V_T=k\right)\mathbb{P}(V_T=k)\\&\qquad\qquad\to\sum_{k=0}^{K_\varepsilon}\mathbb{P}(V_T=k)\ge 1-\varepsilon\qquad\text{as}\quad\delta\to 0,\qquad(22)\end{aligned}$$

which gives the assertion as $\varepsilon\to 0$. □

Proof of Lemma 1. By (17), on the event M_δ the increment of the process $(V_t)_{0\le t\le T}$ on any interval $\left[\frac{i}{\sqrt{n}},\frac{i+1}{\sqrt{n}}\right]$ does not exceed 1 if $n>\frac{1}{\delta^2}$, hence $V_{\frac{i}{\sqrt{n}}}-$

$V_{\frac{i-1}{\sqrt{n}}} = X_i^{(n)}$. Since $V^{(n)}_{\frac{i}{\sqrt{n}}} - V^{(n)}_{\frac{i-1}{\sqrt{n}}} = Y_i^{(n)}$, it is enough to prove that

$$\mathbb{P}(\{\exists i \le \lfloor \sqrt{n}T \rfloor : X_i^{(n)} \ne Y_i^{(n)}\} \cap E_\varepsilon \cap M_\delta) \to 0 \quad (n \to \infty) \tag{23}$$

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

$$1 - \frac{q_i^{(n)}}{p_i^{(n)}} = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \quad (n \to \infty) \tag{24}$$

using the power series of the exponential function and equations (12) and (13) 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

$$\frac{q_i^{(n)} - p_i^{(n)}}{1 - p_i^{(n)}} = \mathcal{O}\left(\frac{1}{n}\right) \quad (n \to \infty). \tag{25}$$

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

3.3 Splits and merges

With the processes $(V_i^{(n)})_{i=0}^\infty$ we have determined when a split or a merge occurs, our task is now to define how it should happen. Similarly to the definition of the limit process $(\mathbf{C}(t))_{t\ge 0}$, 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 $C_0^{(n)}(0) := 1, \quad C_m^{(n)}(0) := 0 \quad m = 1, 2, \dots$. The evolution of the process $\mathbf{C}^{(n)}$ in the steps $i = 1, 2, \dots$ 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 stay 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 (10), via

$$\frac{\sum_{m=0}^{j-1} C_m^{(n)}(i-1)}{\sum_{m=0}^{\infty} C_m^{(n)}(i-1)} < U_k \le \frac{\sum_{m=0}^{j} C_m^{(n)}(i-1)}{\sum_{m=0}^{\infty} C_m^{(n)}(i-1)}. \tag{26}$$

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: $\mathbf{C}^{(n)}(i) := \mathbf{C}^{(n)}(i-1)$, let us 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 add the broken fragment $C_0^{(n)}(i-1) - \lfloor U_k \sum_{m=0}^{\infty} C_m^{(n)}(i-1) \rfloor$ 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 reranking $C_m^{(n)}(i) := C_m^{(n)}(i-1)$ 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.

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 $B \in \mathcal{F}$:

$$\mathbb{P}_{\varepsilon,\delta,n}(B) := \mathbb{P}(B \cap E_\varepsilon \cap M_\delta \cap A_n). \tag{27}$$

By Lemma 1 it is enough to show that for fixed $\varepsilon, \delta > 0$ the processes $\mathbf{C}(t)$ and $\frac{\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}}$ 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 $\frac{\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}}$ 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 $\frac{\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}}$ very much.

4. We do this also for the splits.

5. We summarize the estimates.

STEP 1. Let

$$d_k^- := d\left(\mathbf{C}(\tau_k-), \frac{\mathbf{C}^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)}{\sqrt{n}}\right), \quad d_k^+ := d\left(\mathbf{C}(\tau_k), \frac{\mathbf{C}^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor)}{\sqrt{n}}\right) \tag{28}$$

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 $\frac{1}{\sqrt{n}}$. The discrete processes $(\frac{\mathbf{C}^{(n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}})_{t\ge 0}$ change only in the times which are multiples of $\frac{1}{\sqrt{n}}$. Thus $\mathbb{P}_{\varepsilon,\delta,n}$-almost surely

$$d_k^- \le d_{k-1}^+ + \frac{2}{\sqrt{n}}. \tag{29}$$

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 (10) and (26). Let us call the possibility that these components are of different indices an *awkward event*. If an awkward event or a fictitious split (meaning 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 $\mathbf{C}(\tau_k-)/\sum_{m=0}^{\infty} C_m(\tau_k-)$, which gives indices corresponding to the subintervals. Then the dividing points are

$$\frac{\sum_{m=0}^{j} C_m(\tau_k-)}{\sum_{m=0}^{\infty} C_m(\tau_k-)} \qquad j = 0, 1, 2, \ldots. \tag{30}$$

We do this at the same time with dividing points

$$\frac{\sum_{m=0}^{j} C_m^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)}{\sum_{m=0}^{\infty} C_m^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)} \qquad j = 0, 1, 2, \dots \tag{31}$$

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

$$\frac{\lfloor\sqrt{n}t\rfloor - K_\varepsilon}{\sqrt{n}} \leq \frac{\sum_{m=0}^{\infty} C_m^{(n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}} \leq t \quad \text{if} \quad t \in [0, T], \tag{32}$$

because at the split or merge points (occurring at most K_ε many times) the total length of the discrete process does not increase. From this, using Lemma 1, it follows that

$$\left| \sum_m C_m(\tau_k-) - \frac{\sum_m C_m^{(n)}(\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1)}{\sqrt{n}} \right| \\ \leq \left| \tau_k - \frac{\lfloor\sqrt{n}\tau_k^{(n)}\rfloor - 1 - K_\varepsilon}{\sqrt{n}} \right| \leq \frac{K_\varepsilon}{\sqrt{n}} + \frac{3}{\sqrt{n}}. \tag{33}$$

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

$$\frac{d_k^- + \frac{K_\varepsilon+3}{\sqrt{n}}}{\tau_k}. \tag{34}$$

Since the number of coordinates is at most K_ε, this provides the following upper bound:

$$\operatorname{Leb}(W_k) \leq \frac{d_k^- + \frac{K_\varepsilon+3}{\sqrt{n}}}{\tau_k} K_\varepsilon \leq \frac{d_k^- + \frac{K_\varepsilon+3}{\sqrt{n}}}{\delta} K_\varepsilon, \tag{35}$$

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

$$\mathbb{P}_{\varepsilon,\delta,n}(\text{awkward event at } \tau_k) \leq \operatorname{Leb}(W_k) \leq \frac{d_k^- + \frac{K_\varepsilon+3}{\sqrt{n}}}{\delta} K_\varepsilon. \tag{36}$$

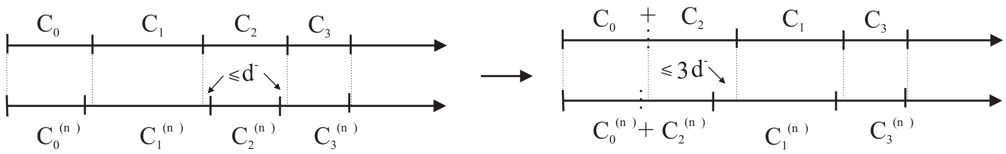

Figure 3: The piece C_2 merges C_0 parallel with the $C_2^{(n)} - C_0^{(n)}$ coagulation

Furthermore

$$\mathbb{P}_{\varepsilon,\delta,n}(\text{fictitious split at } \tau_k^{(n)}) \le \frac{\frac{1}{\sqrt{n}}}{\sum_{m=0}^{\infty} C_m^{(n)}(\lfloor \sqrt{n}\tau_k^{(n)} \rfloor - 1)/\sqrt{n}} \le \frac{2}{\delta\sqrt{n}}, \tag{37}$$

if n is large enough by (32). So we conclude that

$$\mathbb{P}_{\varepsilon,\delta,n}(\text{awkward event or fictitious split at the } k\text{th split or merge point}) \le \frac{K_\varepsilon}{\delta} d_k^- + \frac{K_\varepsilon^2 + 3K_\varepsilon + 2}{\delta\sqrt{n}}. \tag{38}$$

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

$$d_k^+ \le 3d_k^-. \tag{39}$$

Step 4. If a (non-fictitious) split occurs in the discrete and continuous processes, then using inequality (33) we have

$$\left| C_0(\tau_k) - \frac{C_0^{(n)}(\lfloor \sqrt{n}\tau_k^{(n)} \rfloor)}{\sqrt{n}} \right| \le \left| U_k \sum_{m=0}^{\infty} C_m(\tau_k -) - \frac{\lfloor U_k \sum_{m=0}^{\infty} C_m^{(n)}(\lfloor \sqrt{n}\tau_k^{(n)} \rfloor - 1) \rfloor}{\sqrt{n}} \right| \le \frac{K_\varepsilon + 3}{\sqrt{n}} + \frac{1}{\sqrt{n}}. \tag{40}$$

This is why the broken pieces from $C_0(\tau_k -)$ 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.

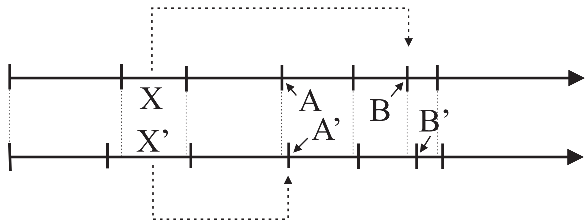

Figure 4: Split: the broken pieces from the coordinate 0 are X and X', which have to be moved to places B and A'

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 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 $|A - A'| \leq d_k^-$, 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 (see Figure 4). The lengths of the parts between A and B are at least $|X|$ and at most $|X'| + 2d_k^- \leq |X| + 3d_k^- + \frac{K_\varepsilon+4}{\sqrt{n}}$. So any two of these parts have lengths differing by at most $3d_k^- + \frac{K_\varepsilon+4}{\sqrt{n}}$. Swapping X always with its right neighbour until hitting place B, 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

$$d_k^+ \leq 2d_k^- + \frac{K_\varepsilon + 4}{\sqrt{n}} + K_\varepsilon\left(3d_k^- + \frac{K_\varepsilon + 4}{\sqrt{n}}\right). \tag{41}$$

Step 5. Summing up the estimates (29), (39) and (41) we get easily the following recursive bound:

$$d_k^+ \leq \max(3, 2 + 3K_\varepsilon)d_{k-1}^+ + \frac{K_\varepsilon^2 + 5K_\varepsilon + 4 + 2\max(3, 2 + 3K_\varepsilon)}{\sqrt{n}}$$
$$=: ad_{k-1}^+ + \frac{b}{\sqrt{n}}. \tag{42}$$

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

4 Multiple stirring

From the idea of the stirring model studied in the previous sections we can obtain a more general construction, if we allow more than one stirring particles. Let k be a fixed integer, and the elements $1, 2, \ldots, k \in [n]$ will be the stirring particles. We associate to them independent sequences of random transpositions $(T_{i,j}^{(n)})_{i=1}^{\infty}$ $\quad j = 1, \ldots, k$ in such a way that the permutation $T_{i,j}^{(n)}$ is chosen with uniform distribution from all transpositions moving the stirring element j and the identity permutation. Let

$$P_i^{(k,n)} := T_{i,k}^{(n)} \circ T_{i,k-1}^{(n)} \circ \cdots \circ T_{i,1}^{(n)} \quad i = 1, 2, \ldots.$$

The *random permutation generated by k stirring particles* is

$$\pi^{(k,n)}(i) := P_i^{(k,n)} \circ P_{i-1}^{(k,n)} \circ \cdots \circ P_1^{(k,n)}. \tag{43}$$

We mention that another definition of $\pi^{(k,n)}(i)$ could be possible: the different stirring elements could apply their transpositions in a random order. If $\xi_1, \xi_2, \ldots$ are i.i.d. random variables with uniform distribution on the discrete set $[k]$ independently from all transpositions, then the alternative definition is

$$\pi^{(k,n)}(i) := T_{i,\xi_{ki}}^{(n)} \circ T_{i-1,\xi_{i-1}}^{(n)} \circ \cdots \circ T_{1,\xi_1}^{(n)}, \tag{44}$$

but we investigate the case (43), because both processes provide the same limit distribution apart from a k-times acceleration of the time.

The evolution of the permutation $(\pi^{(k,n)}(i))_{i=0}^{\infty}$ is quite similar to that of $(\pi^{(n)}(i))_{i=0}^{\infty}$ as long as the different stirring particles choose different elements to swap places with. If two transposed elements were present in different active cycles, then these two cycles coagulate. Furthermore in a general situation more stirring elements join in a common cycle in some cyclic order.

This order determines a permutation $\Pi^{(k,n)}(i)$ on the set of the stirring particles via its cyclic decomposition. The orientation of the cycles in $\Pi^{(k,n)}(i)$ is induced by that of the cycles in $\pi^{(k,n)}(i)$.

We can associate to each stirring particle its *active cycle segment* which is the number of the elements between the current and the next stirring element in the adequate cycle with the suitable orientation including exactly one stirring particle. So the following state space could be appropriate to describe the behaviour of our processes:

$$\begin{aligned}\mathbf{S}^{(k)} := \{(\Pi, (a_1, \dots, a_k), (s_1, s_2, \dots)) : \Pi \in \mathrm{Perm}(\{1, \dots, k\}),\\ a_j \in \mathbb{R}, \quad a_j \geq 0 \quad j = 1, \dots, k, \quad s_m \in \mathbb{R} \quad m = 1, 2, \dots,\\ s_1 \geq s_2 \geq \cdots \geq 0 \text{ and } s_m > 0 \text{ for finitely many } m\}. \quad (45)\end{aligned}$$

In similar way as in the case of $\mathbf{S}$ we can define a metric on $\mathbf{S}^{(k)}$ (see (3)) provided that the Π components of the two vectors in $\mathbf{S}^{(k)}$ are equal, otherwise let the distance be ∞.

The encoding of the permutation $\pi^{(k,n)}(i)$ is

$$\mathbf{C}^{(k,n)}(t) := (\Pi^{(k,n)}(t), (A_1^{(k,n)}(t), \dots, A_k^{(k,n)}(t)), (C_1^{(k,n)}(t), C_2^{(k,n)}(t), \dots)) \in \mathbf{S}^{(k)}, \quad (46)$$

where $\Pi^{(k,n)}(i)$ is the permutation induced by the stirring elements, $A_j^{(k,n)}(i)$ is the active cycle segment of the jth stirring particle, and the numbers $C_1^{(k,n)}(i) \geq C_2^{(k,n)}(i) \geq \dots$ are the lengths of the non-active cycles. See Figure 5 for an example.

Let $(B_j^{(k,n)}(i))_{i=0}^{\infty}$ denote the random walk of the jth stirring particle ($j = 1, \dots, k$). Similarly to (4) let

$$\begin{aligned}V_j^{(k,n)}(i) := \#\{l : l \leq i, \exists m < l, \exists r \in [k] : B_r^{(k,n)}(m) = B_j^{(k,n)}(l)\}\\ + \mathbb{1}\left(\exists l : 1 \leq l < j : B_l^{(k,n)}(i) = B_j^{(k,n)}(i)\right) \quad (47)\end{aligned}$$

be the number of returns of the stirring element j until the time i. The total number of returns is

$$V^{(k,n)}(i) := \sum_{j=1}^{k} V_j^{(k,n)}(i). \quad (48)$$

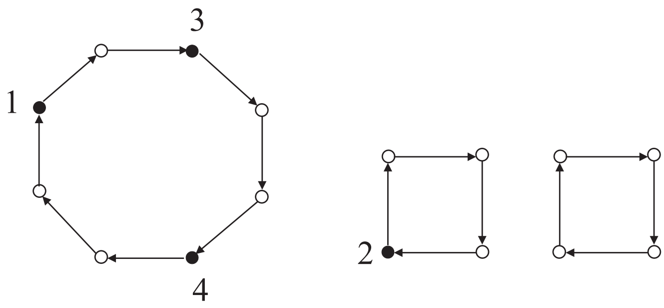

Figure 5: An example $\mathbf{C}^{(k,n)}(t)$: $\mathbf{C}^{(4,16)} = (\Pi^{(4,16)} = ((1,3,4)(2)), (A_1^{(4,16)} = 2, A_2^{(4,16)} = 4, A_3 = 3, A_4^{(4,16)} = 3), (4,0,0,\dots))$. Only the stirring particles numbered.

Proposition 3. *Let $(V_j^{(k)}(t))_{t\geq 0}$ be independent Poisson point processes with common intensity $\rho(t) = kt \quad j = 1, 2, \dots$. Then*

$$\left(V_j^{(k,n)}(\lfloor\sqrt{n}t\rfloor)\right)_{j=1}^{k} \overset{\mathrm{d}}{\Longrightarrow} \left(V_j^{(k)}(t)\right)_{j=1}^{k} \qquad (n \to \infty) \tag{49}$$

in terms of the finite dimensional marginal distributions.

The limit process in the general case is the continuous time stochastic process

$$\mathbf{C}^{(k)}(t) = (\Pi^{(k)}(t), (A_1^{(k)}(t), \dots, A_k^{(k)}(t)), (C_1^{(k)}(t), C_2^{(k)}(t), \dots)) \in \mathbf{S}^{(k)}$$

with the following transitions:

- The coordinates $A_j^{(k)}(t)$ increase with unit speed $j = 1, \dots, k$.
- In each time of jump τ of the jth Poisson point process $(V_j^{(k)}(t))_{t\geq 0}$ with intensity $\rho(t) = kt$ we choose one of the coordinates A_l or C_m with size-biased distribution (i.e. with probability proportional to the numbers $A_1^{(k)}(\tau-), \dots, A_k^{(k)}(\tau-)$ and $C_1^{(k)}(\tau-), C_2^{(k)}(\tau-), \dots$) independently from the past.

Depending on this choice the following can happen:

1. If C_m is chosen, then $C_m(\tau-)$ merges with $A_j(\tau-)$ as in the case of one stirring particle and the coordinates are re-ranked.

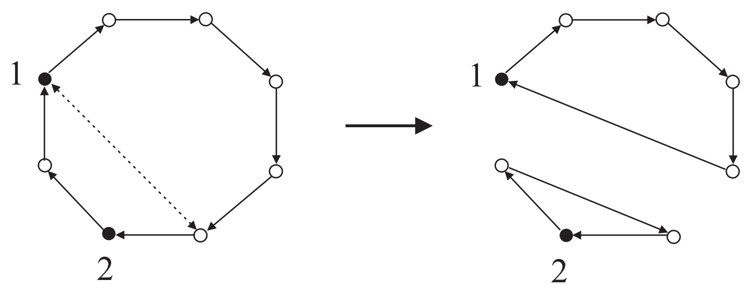

Figure 6: Stirring particle 1 leaves its cycle in Π ($j = 1$)

2. If A_j is chosen and j is a fixed point of the permutation $\Pi^{(k)}(\tau-)$, then a piece of $A_j^{(k)}(\tau-)$ splits with uniform distribution and gets to the adequate place in the decreasing order of the coordinates $C^{(k)}$. (This case corresponds to a simple split in the original model.)

   If j is not a fixed point, then it leaves its earlier cycle in $\Pi^{(k)}(\tau-)$ as a fixed point. In addition to this the active cycle segment of the stirring element preceding j in this cycle of $\Pi^{(k)}(\tau-)$ merges with a piece split with uniform distribution from $A_j^{(k)}(\tau-)$ (see Figure 6).

3. If we choose A_l, but $l \neq j$ and j and l are in the same cycle in the permutation $\Pi^{(k)}(\tau-)$, then this cycle splits in two: the "arc" starting with j ending with l and the remaining arc. Further the active cycle segment of the stirring particle preceding j in its cycle of $\Pi^{(k)}(\tau-)$ merges with a piece of $A_l^{(k)}(\tau-)$ split with uniform distribution (see Figure 7).

4. If A_l is chosen, but $l \neq j$ and these stirring elements are in different cycles of $\Pi^{(k)}(\tau-)$, then these cycles coagulate: the cycles broken before j and after l should be joined. Besides a piece split from $A_l^{(k)}(\tau-)$ with uniform distribution is added to the active cycle segment of the stirring particle preceding j in the appropriate cycle in $\Pi^{(k)}(\tau-)$ (see Figure 8).

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

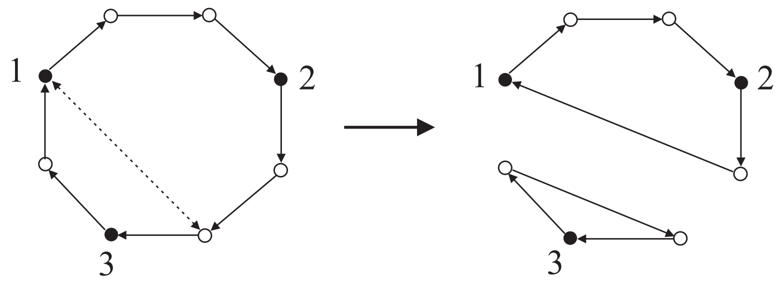

Figure 7: The cycle of stirring element 1 splits in two ($j = 1, l = 2$)

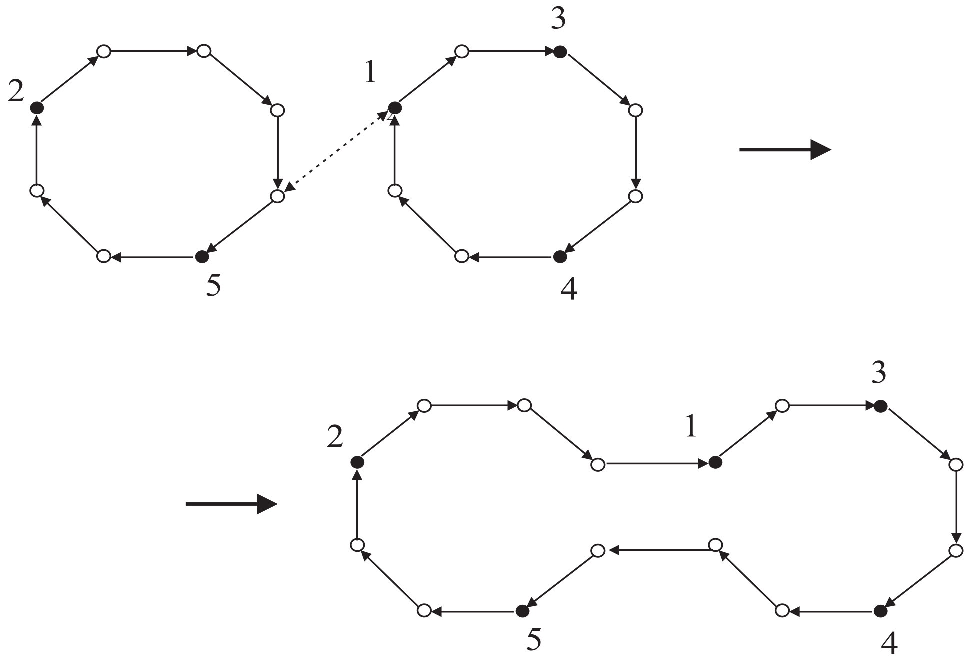

Figure 8: The cycles of stirring particles 1 and 2 coagulate ($j = 1, l = 2$)

function tending to infinity with n, then

$$\mathbb{P}\left(\sup_{0\le t\le T} d^{(k)}\left(\mathbf{C}^{(k)}(t), \frac{\mathbf{C}^{(k,n)}(\lfloor\sqrt{n}t\rfloor)}{\sqrt{n}}\right) < \frac{f(n)}{\sqrt{n}}\right) \to 1 \quad \textit{as } n\to\infty \tag{50}$$

where $d^{(k)}(\cdot,\cdot)$ is the appropriate sup-*distance on $\mathbf{S}^{(k)}$.*

We omit the proof, because it can be done in a similar fashion as that of Theorem 1, but the notation is much more cumbersome in this case.

Our multiple stirring dynamics can be generalized further: the number of the stirring elements can depend on the size of the set $[n]$. E.g. we can consider the permutation generated by $k(n)$ stirring elements, where $k(n) = n^\alpha$ with some $0<\alpha<1$, but other functions $k(n)$ satisfying the condition $\lim_{n\to\infty} k(n) = \infty$ might be worth studying. Of course, in this case we need different scaling of time and space.

5 Stationary distribution

It is a natural question to identify the stationary distribution of the stirring process. We do it for the model generated by one stirring element. First we need some basic definitions and properties of partition valued random processes, which will be used in our proof.

5.1 Poisson – Dirichlet and GEM distributions

Let $W_1, W_2, \dots$ be independent random variables with common beta$(1,\theta)$ distribution (i.e. their density is $\theta(1-x)^{\theta-1}\mathrm{d}x$ if $0<x<1$) where $\theta>0$ is a fixed parameter. Then, by definition the random sequence

$$(P_1, P_2, \dots) := (W_1, (1-W_1)W_2, (1-W_1)(1-W_2)W_3, \dots) \tag{51}$$

has GEM(θ) distribution named after Griffiths, Engen and McCloskey. The distribution of the decreasing rearrangement of $(P_1, P_2, \dots)$ is called Poisson – Dirichlet, denoted PD(θ). See also [10].

These distributions appear in our model only with the parameter value $\theta = 1$, in which case $W_1, W_2, \dots$ have uniform distribution on $[0,1]$. For general θ the GEM(θ) distribution arises as the limit frequencies of the Blackwell – MacQueen urn scheme [3]: let (X_n), with values in $[0,1]$, be governed by the following rule for $n \geq 0$:

$$\mathbb{P}\left(X_{n+1} \in \cdot \bigm| X_1, \dots, X_n\right) = \frac{\theta \operatorname{Leb}(\cdot) + \sum_{i=1}^{n} \mathbb{1}(X_i \in \cdot)}{\theta + n} \tag{52}$$

where $\operatorname{Leb}(\cdot)$ is the Lebesgue measure on $[0,1]$.

Let $\eta_1 := 1$ and $\eta_k := \min\left\{i : X_i \notin \{X_{\eta_1}, X_{\eta_2}, \dots, X_{\eta_{k-1}}\}\right\}$ be the kth value appearing in the sequence $X_1, X_2, \dots$, and define $N_k^{(n)} := \#\{i \in [n] : X_i = X_{\eta_k}\}$. Then

$$\left(\frac{N_1^{(n)}}{n}, \frac{N_2^{(n)}}{n}, \dots\right) \stackrel{\mathrm{d}}{\Longrightarrow} \operatorname{GEM}(\theta) \tag{53}$$

in distribution as $n \to \infty$.

Another formulation of this process is the so called Chinese restaurant model, see [11] for the details. In the special case $\theta = 1$ the statement about the limit of the Blackwell – MacQueen urn scheme is equivalent to the following

Proposition 4. *Let σ_n be a random permutation of $[n]$, chosen uniformly from all the $n!$ permutations. Denote by $\mathbf{C}(\sigma_n)$ the cycle structure of σ_n, i.e. the non-increasing sequence of the lengths of all cycles in σ_n. Then*

$$\frac{\mathbf{C}(\sigma_n)}{n} \stackrel{\mathrm{d}}{\Longrightarrow} \operatorname{PD}(1) \qquad (n \to \infty) \tag{54}$$

in distribution.

The sequence (51) can be regarded as the resulting pieces of a stick breaking process where in each step we break our stick with the same beta distribution scaled by the length of the current piece independently from the past. In this case it is natural to consider $(P_1, P_2, \dots)$ as a partition of $[0,1]$, namely the intervals $[0, P_1], [P_1, P_1 + P_2], \dots$. On the other hand it is also a random probability distribution, because $\sum_{i=1}^{\infty} P_i = 1$. Later on we refer to $(P_1, P_2, \dots)$ both as a partition and a probability distribution.

We introduce the *size-biased permutation* of a random probability distribution $(p_1, p_2, \dots)$. Let $U_1, U_2, \dots$ be i.i.d. uniform random variables on $[0,1]$, independently of $(p_1, p_2, \dots)$. The unique index I_1 is defined by $\sum_{i=1}^{I_1-1} p_i \le U_1 < \sum_{i=1}^{I_1} p_i$. With a piecewise linear transformation of U_2 we can easily construct a uniform random variable $\tilde{U}_2$ on $[0,1] \setminus [\sum_{i=1}^{I_1-1} p_i, \sum_{i=1}^{I_1} p_i]$. The index I_2 is determined by $\sum_{i=1}^{I_2-1} p_i \le \tilde{U}_2 < \sum_{i=1}^{I_2} p_i$. Given the values $I_1, \dots, I_{k-1}$ let U_k be transformed to a uniform random variable $\tilde{U}_k$ on

$$[0,1] \setminus \bigcup_{j=1}^{k-1} \left[\sum_{i=1}^{I_j-1} p_i, \sum_{i=1}^{I_j} p_i \right], \tag{55}$$

and let I_k be such that $\sum_{i=1}^{I_k-1} p_i \le \tilde{U}_k < \sum_{i=1}^{I_k} p_i$. The vector $(p_{I_1}, p_{I_2}, \dots)$ is called the size-biased permutation of $(p_1, p_2, \dots)$.

Theorem 3. *The* GEM(θ) *distribution is invariant under the size-biased permutation.*

The theorem clears up the relation of the GEM and the PD-family of distributions. PD(θ) is obtained by ranking of a sample with GEM(θ) distribution; the size-biased permutation of the PD(θ) is GEM(θ).

Proof. First we prove the theorem for $\theta = 1$, because only this part is needed later. After this we show how to handle the extra difficulties arising in the general case.

Let $(P_1, P_2, \dots)$ be as in (51) with $\theta = 1$ (recall that $W_1, W_2, \dots$ are i.i.d. $[0,1]$-uniform r.v.-s). Denote the size-biased permutation of $(P_1, P_2, \dots)$ by $(X_1, X_2, \dots)$, which is constructed with the use of $U_1, \tilde{U}_2, \tilde{U}_3, \dots$. We identify the density function $f_n(x_1, x_2, \dots, x_n)$ of the joint distribution of $(X_1, X_2, \dots, X_n)$.

For $n = 1$ let $0 < x < 1$. Then we have

$$\begin{aligned} \mathbb{P}(X_1 \in (x, x+\delta x)) &= \int_0^1 \mathbb{P}\left(X_1 \in (x, x+\delta x), U_1 < W_1 \,\middle|\, W_1 = w\right) \mathrm{d}w \\ &+ \int_0^1 \mathbb{P}\left(X_1 \in (x, x+\delta x), U_1 > W_1 \,\middle|\, W_1 = w\right) \mathrm{d}w =: A_1 + B_1. \end{aligned} \tag{56}$$

The first term on the right hand side of (56) is obviously

$$\begin{aligned}
A_1 &= \int_0^1 \mathbb{P}\left(X_1 \in (x, x+\delta x) \;\middle|\; U_1 < W_1, W_1 = w\right) \mathbb{P}\left(U_1 < W_1 \;\middle|\; W_1 = w\right) \mathrm{d}w \\
&= \int_0^1 \mathbb{1}\left(w \in (x, x+\delta x)\right) w \, \mathrm{d}w = \int_x^{x+\delta x} w \, \mathrm{d}w = x \, \delta x + o(\delta x).
\end{aligned} \tag{57}$$

For the second term

$$\begin{aligned}
B_1 &= \int_0^1 \mathbb{P}\left(X_1 \in (x, x+\delta x) \;\middle|\; U_1 > W_1, W_1 = w\right) \mathbb{P}\left(U_1 > W_1 \;\middle|\; W_1 = w\right) \mathrm{d}w \\
&= \int_0^{1-x} \mathbb{P}\left(\tilde{X}_1 \in \left(\frac{x}{1-w}, \frac{x}{1-w} + \frac{\delta x}{1-w}\right)\right) (1-w) \, \mathrm{d}w \\
&= \int_0^{1-x} f_1\left(\frac{x}{1-w}\right) \frac{\delta x}{1-w} (1-w) \, \mathrm{d}w + o(\delta x) \\
&= \int_0^{1-x} f_1\left(\frac{x}{1-w}\right) \delta x \, \mathrm{d}w + o(\delta x)
\end{aligned} \tag{58}$$

where $\tilde{X}_1$ has the same distribution as X_1, because the distribution of X_1 conditioned on the event $\{U_1 > W_1, W_1 = w\}$ is a size-biased sample from a GEM(θ) partition scaled by $(1-w)$. The upper limit of the integral is $1-x$, because if $w > 1-x$, then $\mathbb{P}\left(X_1 \in (x, x+\delta x) \;\middle|\; U_1 > W_1, W_1 = w\right) = 0$.

Substituting (57) and (58) to (56), dividing by δx and letting $\delta x \to 0$ we get the integral equation

$$f_1(x) = x + \int_0^{1-x} f_1\left(\frac{x}{1-w}\right) \mathrm{d}w = x\left(1 + \int_x^1 \frac{f_1(u)}{u^2} \mathrm{d}u\right) \tag{59}$$

after changing the variable under the integral sign. Then we differentiate and substitute the value of the integral expressed from (59), and we get

$$f_1'(x) = 1 + \int_x^1 \frac{f_1(u)}{u^2} \mathrm{d}u - x \frac{f_1(x)}{x^2} = 0 \qquad \forall\, 0 < x < 1, \tag{60}$$

which yields $f_1(x) = 1$ if $0 < x < 1$.

If $n = 2$, then with $x_1, x_2 \geq 0$ and $x_1 + x_2 \leq 1$

$$
\begin{aligned}
&\mathbb{P}(X_1 \in \delta x_1, X_2 \in \delta x_2) \\
&\quad = \int_0^1 \mathbb{P}\left(X_1 \in \delta x_1, X_2 \in \delta x_2, U_1 < W_1 \;\middle|\; W_1 = w\right) \mathrm{d}w \\
&\qquad + \int_0^1 \mathbb{P}\left(X_1 \in \delta x_1, X_2 \in \delta x_2, U_1 > W_1, \tilde{U}_2 < W_1 \;\middle|\; W_1 = w\right) \mathrm{d}w \\
&\qquad + \int_0^1 \mathbb{P}\left(X_1 \in \delta x_1, X_2 \in \delta x_2, U_1 > W_1, \tilde{U}_2 > W_1 \;\middle|\; W_1 = w\right) \mathrm{d}w \\
&\quad =: A_2 + B_2 + C_2
\end{aligned} \tag{61}
$$

where $(X_i \in \delta x_i)$ stands for $(X_i \in (x_i, x_i + \delta x_i)) \quad i = 1, 2$.

For the first term

$$
\begin{aligned}
A_2 &= \int_0^1 \mathbb{P}\left(X_1 \in \delta x_1, X_2 \in \delta x_2 \;\middle|\; U_1 < W_1, W_1 = w\right) \\
&\quad \times \mathbb{P}\left(U_1 < W_1 \;\middle|\; W_1 = w\right) \mathrm{d}w \\
&= \int_0^1 \mathbb{P}\left(X_2 \in \delta x_2 \;\middle|\; X_1 \in \delta x_1, U_1 < W_1, W_1 = w\right) \\
&\quad \times \mathbb{P}\left(X_1 \in \delta x_1 \;\middle|\; U_1 < W_1, W_1 = w\right) w \, \mathrm{d}w \\
&= \int_0^1 \mathbb{P}\left(\tilde{X}_1 \in \left(\frac{x_2}{1-w}, \frac{x_2}{1-w} + \frac{\delta x_2}{1-w}\right)\right) \mathbb{1}\left(w \in \delta x_1\right) w \, \mathrm{d}w \\
&= x_1 \, \delta x_1 f_1\left(\frac{x_2}{1-x_1}\right) \frac{\delta x_2}{1-x_1} + o(\delta x_1 \delta x_2) = \frac{x_1}{1-x_1} \delta x_1 \delta x_2 + o(\delta x_1 \delta x_2)
\end{aligned} \tag{62}
$$

where $\tilde{X}_1$ has the same distribution as X_1, which is already computed in the $n = 1$ case.

Very similarly

$$
\begin{aligned}
B_2 &= \int_0^1 \mathbb{P}\left(X_2 \in \delta x_2 \;\middle|\; X_1 \in \delta x_1, U_1 > W_1, \tilde{U}_2 < W_1, W_1 = w\right) \\
&\quad \times \mathbb{P}\left(\tilde{U}_2 < W_1 \;\middle|\; X_1 \in \delta x_1, U_1 > W_1, W_1 = w\right) \\
&\quad \times \mathbb{P}\left(X_1 \in \delta x_1 \;\middle|\; U_1 > W_1, W_1 = w\right) \mathbb{P}\left(U_1 > W_1 \;\middle|\; W_1 = w\right) \mathrm{d}w \\
&= \int_0^1 \mathbb{1}(w \in \delta x_2) \frac{w}{1 - x_1 + \mathcal{O}(\delta x_1)} f_1\left(\frac{x_1}{1-w}\right) \frac{\delta x_1}{1-w}(1-w)\,\mathrm{d}w \\
&= f_1\left(\frac{x_1}{1-x_2}\right) \frac{\delta x_1}{1-x_2}(1-x_2)\frac{x_2}{1-x_1} + o(\delta x_1 \delta x_2) \\
&= \frac{x_2}{1-x_1}\delta x_1 \delta x_2 + o(\delta x_1 \delta x_2).
\end{aligned}
\tag{63}
$$

Finally

$$
\begin{aligned}
C_2 &= \int_0^1 \mathbb{P}\left(X_2 \in \delta x_2 \;\middle|\; \tilde{U}_2 > W_1, X_1 \in \delta x_1, U_1 > W_1, W_1 = w\right) \\
&\quad \times \mathbb{P}\left(\tilde{U}_2 > W_1 \;\middle|\; X_1 \in \delta x_1, U_1 > W_1, W_1 = w\right) \\
&\quad \times \mathbb{P}\left(X_1 \in \delta x_1 \;\middle|\; U_1 > W_1, W_1 = w\right) \mathbb{P}\left(U_1 > W_1 \;\middle|\; W_1 = w\right) \mathrm{d}w \\
&= \int_0^{1-x_1-x_2} \mathbb{P}\left(\tilde{X}_1 \in \frac{\delta x_1}{1-w}, \tilde{X}_2 \in \frac{\delta x_2}{1-w}\right)\left(1 - \frac{w}{1-x_1+\mathcal{O}(\delta x_1)}\right) \\
&\quad \times (1-w)\,\mathrm{d}w \\
&= \int_0^{1-x_1-x_2} f_2\left(\frac{x_1}{1-w}, \frac{x_2}{1-w}\right)\frac{\delta x_1}{1-w}\frac{\delta x_2}{1-w}(1-w)\frac{1-w-x_1}{1-x_1}\,\mathrm{d}w \\
&\quad + o(\delta x_1 \delta x_2) \\
&= \delta x_1 \delta x_2 \int_{x_1}^{\frac{x_1}{x_1+x_2}} f_2\left(u, \frac{x_2}{x_1}u\right)(1-u)\frac{x_1}{u^2}\frac{1}{1-x_1}\,\mathrm{d}u + o(\delta x_1 \delta x_2)
\end{aligned}
\tag{64}
$$

where the distribution of $(\tilde{X}_1, \tilde{X}_2)$ is equal to that of (X_1, X_2), and we change the variable under the integral sign as in (59), namely $u = x_1/(1-w)$.

If we return to (61), divide by $\delta x_1 \delta x_2$ and let them tend to zero, then we get

$$
f_2(x_1, x_2) = \frac{x_1 + x_2}{1 - x_1} + \frac{x_1}{1-x_1}\int_{x_1}^{\frac{x_1}{x_1+x_2}} f_2\left(u, \frac{x_2}{x_1}u\right)(1-u)\frac{1}{u^2}\,\mathrm{d}u. \tag{65}
$$

Keeping in mind the desired result we introduce $g_2(x_1, x_2) = (1-x_1)f_2(x_1, x_2)$, and we rewrite (65) as

$$g_2(x_1, x_2) = x_1 + x_2 + x_1 \int_{x_1}^{\frac{x_1}{x_1+x_2}} g_2\left(u, \frac{x_2}{x_1}u\right) \frac{1}{u^2}\,\mathrm{d}u. \tag{66}$$

Let $\frac{x_2}{x_1}$ be fixed, $\lambda_2 = 1 + \frac{x_2}{x_1}$, and define $h_2(x_1) = g_2(x_1, \frac{x_2}{x_1}x_1)$. Then (66) transforms to

$$h_2(x_1) = x_1 \left(\lambda_2 + \int_{x_1}^{\frac{1}{\lambda_2}} \frac{h_2(u)}{u^2}\,\mathrm{d}u\right). \tag{67}$$

Because λ_2 does not depend on x_1, the differentiation yields

$$h_2'(x_1) = \underbrace{\lambda_2 + \int_{x_1}^{\frac{1}{\lambda_2}} \frac{h_2(u)}{u^2}\,\mathrm{d}u}_{h_2(x_1)/x_1} + x_1\left(-\frac{h_2(x_1)}{x_1^2}\right) = 0. \tag{68}$$

For the pair (x_1, x_2) choose $\epsilon \le \lambda_2^{-1}$. Since $g_2(\cdot,\cdot)$ is constant along every straight line starting from the origin to the first sector of the coordinate system, using (66) we get

$$\begin{aligned} g_2(x_1, x_2) = g_2\left(\epsilon, \frac{x_2}{x_1}\epsilon\right) &= x_1 + x_2 + x_1 \int_{x_1}^{\frac{x_1}{x_1+x_2}} g_2\left(\epsilon, \frac{x_2}{x_1}\epsilon\right)\frac{1}{u^2}\,\mathrm{d}u \\ &= x_1 + x_2 + (1 - x_1 - x_2)g_2\left(\epsilon, \frac{x_2}{x_1}\epsilon\right), \end{aligned} \tag{69}$$

which means $g_2(x_1, x_2) = g_2\left(\epsilon, \frac{x_2}{x_1}\epsilon\right) = 1$ and $f_2(x_1, x_2) = \frac{1}{1-x_1}$ if $x_1, x_2 \ge 0$ and $x_1 + x_2 \le 1$.

To complete the proof for $\theta = 1$ we use induction for n. We assume that the density function of $(X_1, \ldots, X_{n-1})$ is

$$f_{n-1}(x_1, \ldots, x_{n-1}) = \frac{1}{(1-x_1)(1-(x_1+x_2))\ldots(1-\sum_{i=1}^{n-2} x_i)} \tag{70}$$

if $x_i \ge 0$ and $\sum_{i=1}^{n-1} x_i \le 1$. The key step of the induction is that we decompose the event $E := \{X_1 \in \delta x_1, \ldots, X_n \in \delta x_n\}$ to the disjoint union of $n+1$ events depending on which of $X_1, \ldots, X_n$ will be equal to W_1. It is easy to see from the construction that the events $E \cap \{X_i = W_1\}$ for $i = 1, \ldots, n$

and $E \cap \{X_1 \neq W_1, \ldots, X_n \neq W_1\}$ are almost surely disjoint and their union equals E. So we can deduce the integral equation

$$f_n(x_1, \ldots, x_n) = \sum_{i=1}^{n} \frac{x_i}{(1-x_1)(1-(x_1+x_2)\ldots(1-\sum_{j=1}^{n-1} x_j)}$$
$$+\int_0^{1-\sum_{i=1}^n x_i} f_n\left(\frac{x_1}{1-w}, \ldots, \frac{x_n}{1-w}\right)\left(1-\frac{x_1}{1-w}\right)\ldots\left(1-\frac{x_1+\cdots+x_{n-1}}{1-w}\right)$$
$$\times \frac{1}{1-x_1}\frac{1}{1-(x_1+x_2)}\cdots\frac{1}{1-\sum_{j=1}^{n-1} x_j}\,\mathrm{d}w$$
$$= \frac{1}{(1-x_1)\ldots(1-\sum_{j=1}^{n-1} x_j)}\left(\sum_{i=1}^{n} x_i + \int_{x_1}^{\frac{x_1}{\sum_{i=1}^n x_i}} f_n\left(u, \frac{x_2}{x_1}u, \ldots, \frac{x_n}{x_1}u\right)\right.$$
$$\left.\times(1-u)\left(1-\frac{x_1+x_2}{x_1}u\right)\ldots\left(1-\frac{x_1+\cdots+x_{n-1}}{x_1}u\right)\frac{x_1}{u^2}\,\mathrm{d}u\right) \quad (71)$$

similarly to the $n=2$ case, where the first n summands come from the possibilities $\{X_i = W_1\}$ $\quad i = 1, \ldots, n$, the integral term corresponds to the event when all the X_i-s are chosen from $[W_1, 1]$. We introduce

$$g_n(x_1, \ldots, x_n) = (1-x_1)(1-(x_1+x_2))\ldots(1-\sum_{j=1}^{n-1} x_j)f_n(x_1, \ldots, x_n), \quad (72)$$

to obtain

$$g_n(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i + x_1 \int_{x_1}^{\frac{x_1}{\sum_{i=1}^n x_i}} g_n\left(u, \frac{x_2}{x_1}u, \ldots, \frac{x_n}{x_1}u\right)\frac{\mathrm{d}u}{u^2} \quad (73)$$

Let the ratios $\frac{x_2}{x_1}, \frac{x_3}{x_1}, \ldots, \frac{x_n}{x_1}$ be fixed and let $\lambda_n = 1 + \frac{x_2}{x_1} + \cdots + \frac{x_n}{x_1}$. With the function $h_n(x_1) = g_n(x_1, \frac{x_2}{x_1}x_1, \ldots, \frac{x_n}{x_1}x_1)$ we have

$$h_n(x_1) = x_1\left(\lambda_n + \int_{x_1}^{\frac{1}{\lambda_n}} \frac{h_n(u)}{u^2}\,\mathrm{d}u\right). \quad (74)$$

From here the computation is essentially the same as after (67), the result is $g_n(x_1, \ldots, x_n) = 1$ and

$$f_n(x_1, \ldots, x_n) = \frac{1}{(1-x_1)(1-(x_1+x_2))\ldots(1-\sum_{i=1}^{n-1} x_i)} \quad (75)$$

if $x_i \geq 0$ and $\sum_{i=1}^n x_i \leq 1$, as required.

The case of general $\theta > 0$ and $n = 1$ is not complicated: we obtain the equation

$$f_1(x) = \theta x(1-x)^{\theta-1} + \theta x^\theta \int_x^1 \frac{f_1(u)}{u^{\theta+1}} \mathrm{d}u \tag{76}$$

in the same way as in (56)-(59). After differentiation one gets immediately $f_1(x) = \theta(1-x)^{\theta-1}$ if $0 < x < 1$.

If $n = 2$, then we can deduce

$$\begin{aligned} f_2(x_1, x_2) = {} & \theta x_1 (1-x_1)^{\theta-1} \theta \left(1 - \frac{x_2}{1-x_1}\right)^{\theta-1} \frac{1}{1-x_2} \\ & + \theta (1-x_2)^{\theta-1} \theta \left(1 - \frac{x_1}{1-x_2}\right)^{\theta-1} \frac{x_2}{1-x_1} \\ & + \int_{x_1}^{\frac{x_1}{x_1+x_2}} f_2\left(u, \frac{x_2, x_1}{u}\right)(1-u)\theta \frac{x^\theta}{u^{\theta+1}} \frac{1}{1-x_1} \mathrm{d}u, \end{aligned} \tag{77}$$

which simplifies to

$$g_2(x_1, x_2) = x_1 + x_2 + \theta x_1^\theta \int_{x_1}^{\frac{x_1}{x_1+x_2}} \left(\frac{1 - u - \frac{x_2}{x_1}u}{1 - x_1 - x_2}\right)^{\theta-1} \frac{g_2\left(u, \frac{x_2}{x_1}u\right)}{u^{\theta+1}} \mathrm{d}u \tag{78}$$

with the notation $g_2(x_1, x_2) = \frac{(1-x) f_2(x_1, x_2)}{\theta^2 (1-x_1-x_2)^{\theta-1}}$. For fixed $\frac{x_2}{x_1}$ and $\lambda_2 = 1 + \frac{x_2}{x_1}$ with $h_2(x_1) = g_2(x_1, \frac{x_2}{x_1} x_1)$

$$h_2(x_1) = \lambda_2 x_1 + \theta \frac{x_1^\theta}{(1-\lambda_2 x_1)^{\theta-1}} \int_{x_1}^{\frac{1}{\lambda_2}} (1-\lambda_2 u)^{\theta-1} \frac{h_2(u)}{u^{\theta+1}} \mathrm{d}u. \tag{79}$$

If we differentiate and rewrite the derivative with the value of the integral expressed from (79), then we have the solution

$$h_2(x_1) = 1 + c(1-\lambda_2 x_1)^{-(\theta-1)}. \tag{80}$$

If one writes it back to (79), the only solution will be $h_2(x_1) = 1$ corresponding to $c = 0$. It yields $g_2(x_1, x_2) = 1$ and $f_2(x_1, x_2) = \frac{\theta^2 (1-x_1-x_2)^{\theta-1}}{1-x_1}$ if $x_1, x_2 \geq 0$ and $x_1 + x_2 \leq 1$, as required. The statement for general n is verified with exactly the same computations as in (78)-(80). □

5.2 Alternative definition

We present a new definition of the $\mathrm{PD}(\theta)$ distribution. Let $\rho(x) = \theta x^{-1}e^{-x}$ if $x > 0$ be the intensity of a Poisson point process on $(0, \infty)$ with random lengths denoted by $J_1 \geq J_2 \geq \cdots \geq 0$ in non-increasing order. The total length is $T = \sum_{i=1}^{\infty} J_i$.

It is well known that with this chose of the intensity ρ the distribution of T is gamma with parameter θ. It can be easily verified with the method of the characteristic function.

Theorem 4. *The random partition*

$$\left(\frac{J_1}{T}, \frac{J_2}{T}, \dots\right) \tag{81}$$

is independent of T and has $\mathrm{PD}(\theta)$ *distribution.*

Proof. We apply a more general proof from Pitman [11]. We only assume that $(J_1, J_2, \dots)$ is the decreasing sequence of the points of a Poisson point process with intensity ρ and that the density function of the total length $f(t) = \mathbb{P}(T \in \mathrm{d}t)/\mathrm{d}t$ exists ($T = \sum_{i=1}^{\infty} J_i$). Let $(\tilde{J}_1, \tilde{J}_2, \dots)$ be the size-biased permutation os $(J_1, J_2, \dots)$ and we introduce the notation $P_i = J_i/T$ and $\tilde{P}_i = \tilde{J}_i/T$ for $i = 1, 2, \dots$.

Let $T_1(v, \mathrm{d}v) := \sum_i J_i \mathbb{1}(J_i \in (v, v + \mathrm{d}v))$ and $T_2(v, \mathrm{d}v) := T - T_1(v, \mathrm{d}v)$. Note that for fixed v

$$\mathbb{P}(T_1(v, \mathrm{d}v) \neq 0) = \rho(v)\mathrm{d}v + o(\mathrm{d}v) = \mathcal{O}(\mathrm{d}v) \quad \text{as} \quad \mathrm{d}v \to 0 \tag{82}$$

It is obvious that

$$\begin{aligned} &\mathbb{P}(\tilde{J}_1 \in (v, v + \mathrm{d}v), T \in (t, t + \mathrm{d}t)) \\ &= \mathbb{P}(T_1(v, \mathrm{d}v) \in (v, v + \mathrm{d}v), T_2(v, \mathrm{d}v) \in (t - v, t - v + \mathrm{d}t), \tilde{J}_1 = T_1(v, \mathrm{d}v)) \\ &\qquad = \mathbb{P}\left(\tilde{J}_1 = T_1 \,\middle|\, T_1 \in (v, v + \mathrm{d}v), T_2 \in (t - v, t - v + \mathrm{d}t)\right) \\ &\qquad\qquad \times \mathbb{P}(T_1 \in (v, v + \mathrm{d}v), T_2 \in (t - v, t - v + \mathrm{d}t)). \end{aligned} \tag{83}$$

The conditional probability above equals $v/t + \mathcal{O}(\mathrm{d}v + \mathrm{d}t)$ because of the size-biased sampling. The other factor in (83) is

$$\int_v^{v+\mathrm{d}v} \mathbb{P}(T_2(v, \mathrm{d}v) \in (t-x, t-x+\mathrm{d}t))\rho(x)\mathrm{d}x$$
$$= \int_v^{v+\mathrm{d}v} (f(t-x)\mathrm{d}t + \mathcal{O}(\mathrm{d}v))\rho(x)\mathrm{d}x = f(t-v)\rho(v)\mathrm{d}v\mathrm{d}t + \mathcal{O}((\mathrm{d}v)^2) \quad (84)$$

using (82). Briefly, apart from a term tending to zero, we got the equation

$$\mathbb{P}(\tilde{J}_1 \in \mathrm{d}v, T \in \mathrm{d}t) = \rho(v)\mathrm{d}v f(t-v)\mathrm{d}t \frac{v}{t}, \quad (85)$$

which can be utilized in the computation below.

$$\frac{\mathbb{P}\left(\tilde{P}_1 \in \mathrm{d}p \mid T = t\right)}{\mathrm{d}p} = \frac{\mathbb{P}(\tilde{J}_1 \in t\mathrm{d}p, T \in \mathrm{d}t)}{\mathrm{d}p\mathbb{P}(T \in \mathrm{d}t)} = \frac{\rho(pt)t\mathrm{d}p f(\bar{p}t)\mathrm{d}t}{\mathrm{d}p f(t)\mathrm{d}t}$$
$$= pt\rho(pt)\frac{f(\bar{p}t)}{f(t)} = \rho_*(pt)\frac{f(\bar{p}t)}{f(t)} \quad (86)$$

where $\bar{p} = 1-p$ and $\rho_*(x) = x\rho(x)$.

Let $X_j = T - \sum_{k=1}^{j} \tilde{J}_k = \sum_{k=j+1}^{\infty} \tilde{J}_k$. Then the sequence $(X_0, X_1, X_2, \dots)$ is a Markov-chain since the distribution of X_{j+1} depends only on the value of X_j. From (86) the transition probabilities are

$$\mathbb{P}\left(X_{j+1} \in \mathrm{d}t_1 \mid X_j = t\right) = \rho_*(t-t_1)\frac{f(t_1)}{f(t)}\mathrm{d}t_1. \quad (87)$$

It results the representation

$$\tilde{P}_j = W_j \prod_{i=1}^{j-1}(1-W_i) \quad \text{where} \quad 1 - W_i = \frac{X_i}{X_{i-1}} \quad (88)$$

and $W_1, W_2, \dots$ are independent.

In our case $\rho_*(x) = \theta e^{-x}$ and $f(t) = t^{\theta-1}e^{-t}/\Gamma(\theta)$ from the first part of this proof. Substituting to (88) we have

$$\mathbb{P}\left(X_{j+1} \in \mathrm{d}t_1 \mid X_j = t\right) = \theta e^{-(t-t_1)}\frac{t_1^{\theta-1}e^{-t_1}}{t^{\theta-1}e^{-t}}\mathrm{d}t_1 = \theta\left(\frac{t_1}{t}\right)^{\theta-1}\mathrm{d}t_1, \quad (89)$$

which means that the distribution of $W_{j+1} = 1 - X_{j+1}/X_j$ is beta$(1,\theta)$ as in the original definition of the PD(θ) distribution. □

The family of PD distributions is closely connected to the gamma processes. Following Pitman [11] a *subordinator* $(T_s, s \geq 0)$ is an increasing process with right continuous paths, stationary independent increments, and $T_0 = 0$.

Every such process can be represented as

$$T_t = ct + \sum_{s:0<s\leq t} \Delta_s \tag{90}$$

if $t \geq 0$ for some $c \geq 0$ where $\Delta_s := T_s - T_{s-}$ and $\{(s, \Delta_s) : s > 0, \Delta_s > 0)\}$ is the set of points of a Poisson point process on $(0,\infty)^2$ with intensity measure $\text{Leb} \times \Lambda$ satisfying the following regularity conditions.

The measure Λ (called the *Lévy measure* of the process) has to fulfil

$$\begin{aligned} &\int_0^1 x \, \mathrm{d}\Lambda(x) < \infty \\ &\int_1^\infty \mathrm{d}\Lambda(x) < \infty, \end{aligned} \tag{91}$$

which is equivalent to the finiteness of of the *Laplace exponent*

$$\psi(u) = cu + \int_0^\infty (1 - e^{-ux}) \, \mathrm{d}\Lambda(x) \tag{92}$$

for some (hence all) $u > 0$.

The *standard gamma process* $(\Gamma_s, s \geq 0)$ is the subordinator with marginal densities

$$\mathbb{P}(\Gamma_s \in \mathrm{d}x)/\mathrm{d}x = \frac{1}{\Gamma(s)} x^{s-1} e^{-x} \tag{93}$$

if $x > 0$. In this case the Lévy measure of $(\Gamma_s, s \geq 0)$ is the measure with density $\rho(x) = x^{-1}e^{-x}$. Using Theorem 4 the magnitudes of jumps $\{\Delta_s, 0 < s < \theta\}$ in non-increasing order normalized by the value Γ_θ is a vector with $\text{PD}(\theta)$ distribution.

5.3 Stationary distribution of the stirring process

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. 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: 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 to (7-10) because $\sum_m C_m = 1$.

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 5. *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). By Theorem 3 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 $(1-U)$. It 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. It shows that $\nu = \mu$, as required. □

Theorem 5 proves that μ is a stationary measure for our process, but it is not at all clear if this is the *unique* stationary measure. The proof of this would be the analogue of Schramm's result in [13].

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 uniform distribution, if it can be regarded at all. The solution of the problem in this simply describable model is not obvious in the least. For more about this problem in similar models see [5].

Acknowledgement

I am grateful to Bálint Tóth, my supervisor for making me acquainted with a lot of interesting parts of probability theory and that he always had time to discuss with me. I thank Benedek Valkó for so much practical advice that he gave me during my work on this topic.

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] D. Blackwell, J. B. MacQueen, Ferguson distribution via Pólya urn schemes, *Annals of Statistics*, **1**, 353-355, 1973.

[4] 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.

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

[6] 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.

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

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

[9] S. Janson, T. Łuczak, A. Ruciński, *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, New York, Wiley-Interscience, 2000.

[10] J. F. C. Kingman, Random discrete distributions, *Journal of the Royal Statistical Society, Series B*, **37**, 1-22, 1975.

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

[12] 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.

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

[14] 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.

[15] 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.

[16] B. Vető, The time evolution of permutations under random stirring, Submitted to *Acta Scientiarum Mathematicarum*, 2005.