Combinatorics, Probability and Computing (2008) 17, 329–337. © 2007 Cambridge University Press
doi:10.1017/S0963548307008784 First published online 16 November 2007 Printed in the United Kingdom

Two-Parameter Poisson–Dirichlet Measures and Reversible Exchangeable Fragmentation–Coalescence Processes

JEAN BERTOIN

Laboratoire de Probabilités, Université Pierre et Marie Curie
and
DMA, Ecole Normale Supérieure, 45, rue d'Ulm, F-75005 Paris, France
(e-mail: jbe@ccr.jussieu.fr)

Received 24 April 2007; revised 7 October 2007

We show that for $0 < \alpha < 1$ and $\theta > -\alpha$, the Poisson–Dirichlet distribution with parameter (α, θ) is the unique reversible distribution of a rather natural fragmentation–coalescence process. This completes earlier results in the literature for certain split-and-merge transformations and the parameter $\alpha = 0$.

1. Introduction

For every $\theta > 0$, consider a sequence $\beta_1, \beta_2, \dots$ of i.i.d. random variables with the beta$(1, \theta)$ distribution, *i.e.*, $\mathbb{P}(\beta_i \in \mathrm{d}u) = \theta(1-u)^{\theta-1}\mathrm{d}u$ with $u \in]0,1[$. The Residual Allocation Model defined by the sequence of random variables

$$\xi_1 = \beta_1,\ \xi_2 = \beta_2(1-\beta_1), \dots, \xi_{n+1} = \beta_{n+1} \times \prod_{i=1}^{n}(1-\beta_i), \dots$$

has the Griffiths–Engen–McClosey distribution with parameter θ, GEM(θ). Rearranging the terms of the sequence $(\xi_n)_{n\in\mathbb{N}}$ in decreasing order yields a random variable $\xi^{\downarrow}$ with values in the space of (proper) partitions of the unit mass

$$\mathscr{P}_1 := \left\{ \mathbf{x} = (x_1, x_2, \dots) : x_1 \geqslant x_2 \geqslant \cdots \geqslant 0 \text{ and } \sum_{n=1}^{\infty} x_n = 1 \right\}.$$

The ranked sequence $\xi^{\downarrow}$ has a Poisson–Dirichlet distribution, which is denoted here by PD$(0, \theta)$.

The Poisson–Dirichlet laws PD$(0, \theta)$ form a one-parameter family of probability measures on $\mathscr{P}_1$ that play an important role in a variety of areas. Let us just mention, firstly for $\theta = 1$, limit theorems for large combinatorial structures (*e.g.*, decomposition of permutations into cycles, or of integers into prime factors), and then for general $\theta > 0$, Bayesian statistics and population genetics. We refer to the book [1] by Arratia, Barbour and Tavaré for precise references and much

https://doi.org/10.1017/S0963548307008784 Published online by Cambridge University Press

more on this topic. The motivation for the present note stems from the fact that PD$(0,\theta)$ arises as invariant distribution for certain split-and-merge transformations; see Tsilevich [20], Gnedin and Kerov [9], Pitman [15], Diaconis, Mayer-Wolf, Zeitouni and Zerner [6], Mayer-Wolf, Zeitouni and Zerner [11], Schramm [18], among others. Specifically, consider the discrete-time Markov chain with values in $\mathscr{P}_1$ whose transitions can be described as follows. When the chain starts from some configuration $\mathbf{x} = (x_i)_{i\in\mathbb{N}}$, pick two i.i.d. random integers N and N' according to the law $\mathbb{P}(N = i) = \mathbb{P}(N' = i) = x_i$. If $N \neq N'$, then merge x_N and $x_{N'}$, otherwise split the term $x_N = x_{N'}$ uniformly at random, and in both cases, leave the other terms of $\mathbf{x}$ unchanged. Finally, reorder the resulting terms in decreasing order to get a partition of the unit mass. The Poisson–Dirichlet distribution PD$(0,1)$ is the unique invariant (in fact, reversible) probability measure for this split-and-merge evolution. More generally, any PD$(0,\theta)$ for $\theta > 0$ appears as the invariant law of a variation of the preceding dynamics. In this direction, we mention that the question of the existence of reversible laws for a large class of coagulation–fragmentation processes with values in certain *finite* subsets of $\mathscr{P}_1$ (see [10, 21] for some literature on this topic) has been solved by Durrett, Granovsky and Gueron [7] and Erlihson and Granovsky [8]. However, due to the requirement of finiteness, it does not seem that the latter results can be applied to tackle the problem of existence and uniqueness of the invariant distribution for the split-and-merge transformation above.

Perman, Pitman and Yor [13] extended the notion of Poisson–Dirichlet distributions to a two-parameter family PD(α,θ) with $0 \leqslant \alpha < 1$ and $\theta > -\alpha$. One quick way to do this is to apply the Residual Allocation Model above to the situation where $\beta_1, \beta_2, \ldots$ are again independent beta variables, but now β_n has parameter $(1-\alpha, \theta + n\alpha)$, namely,

$$\mathbb{P}(\beta_n \in \mathrm{d}u) = \mathrm{B}(1-\alpha, \theta+n\alpha)^{-1} u^{-\alpha}(1-u)^{\theta+n\alpha-1}\mathrm{d}u \quad \text{for } 0 < u < 1.$$

This two-parameter family appears naturally in connection with excursion lengths of Bessel processes and their bridges; we refer to the survey by Pitman and Yor [13] for a detailed account of this topic. The purpose of this note is to point out that for any $0 < \alpha < 1$ and $\theta > -\alpha$, the Poisson–Dirichlet distribution PD(α,θ) also arises as the unique reversible law for a certain quite natural fragmentation–coalescence process.

The fragmentation–coalescence processes that we shall deal with here are not directly related to the split-and-merge transformation, but rather belong to the family of Exchangeable Fragmentation–Coalescence (EFC) processes, which have been introduced recently by Berestycki [3]. Roughly speaking, EFC processes evolve in continuous time and combine the dynamics of coalescents with simultaneous multiple collisions (see [12, 19] or Chapter 4 in [5]) and that of homogeneous fragmentations (see [4] or Chapter 3 in [5]). We stress that, in general, they are not step Markov processes, in the sense that their first exit from an arbitrary configuration is instantaneous a.s. In this work, collisions in coagulation events are multiple (in fact, every coagulation event involves an infinite number of terms a.s.) and coagulations are simple, in the sense that all the terms involved in a coagulation event merge in a single term. More precisely, using the terminology of Pitman [14] (see also Sagitov [17]), the rates of multiple collisions are governed by the measure $\Lambda_{\theta/\alpha}(\mathrm{d}u) = (1-u)^{\theta/\alpha}\mathrm{d}u$ for $u \in]0,1[$. The rates of dislocations are in turn governed by an infinite measure of Poisson–Dirichlet type denoted by PD$(\alpha,-\alpha)$, which has been introduced recently by Basdevant [2] in a study of Ruelle's cascades (note that $(\alpha,-\alpha)$ is excluded from the range of parameters for Poisson–Dirichlet probability measures). It

https://doi.org/10.1017/S0963548307008784 Published online by Cambridge University Press

is interesting to observe that the coagulation rates only depend on the ratio θ/α, whereas those for dislocations only depend on the parameter α.

2. Some background

The purpose of this section is to provide a rigorous presentation of notions which have been discussed informally above. We therefore now introduce the material which will be needed to state and prove our main result in the next section. Exchangeability is one of the major ingredients. We first recall the key connexion between random mass-partitions and random exchangeable partitions of $\mathbb{N}$ due to Kingman, and an important sampling formula for Poisson–Dirichlet distributions due to Pitman (we refer to, *e.g.*, Section 2.3 in [5] for a complete account). Then we briefly present some main features on EFC processes which are gleaned from [3].

2.1. Exchangeable random partitions

A partition π of $\mathbb{N} := \{1, 2, \dots\}$ is an infinite sequence of pairwise disjoint blocks $\pi_1, \pi_2, \dots$ ordered according to increasing order of their least element, and such that $\bigcup_{i\in\mathbb{N}} \pi_i = \mathbb{N}$. The space of partitions of $\mathbb{N}$ is denoted by $\mathscr{P}_\mathbb{N}$. For every $n \in \mathbb{N}$, the restriction to $[n] := \{1, \dots, n\}$ induces a natural projection $\pi \to \pi_{|[n]}$ from $\mathscr{P}_\mathbb{N}$ to $\mathscr{P}_{[n]}$, the space of partitions of $[n]$. Plainly, the sequence of restricted partitions $(\pi_{|[n]}, n \in \mathbb{N})$ is compatible, in the sense that for every integer $n \geqslant 2$, $\pi_{|[n]}$ is a partition of $[n]$ whose restriction to $[n-1]$ coincides with $\pi_{|[n-1]}$. Conversely, for any sequence of compatible partitions $\pi^{[n]} \in \mathscr{P}_{[n]}$ for $n \in \mathbb{N}$, there exists a unique partition $\pi \in \mathscr{P}_\mathbb{N}$ such that $\pi_{|[n]} = \pi^{[n]}$.

A random partition of $\mathbb{N}$ is called *exchangeable* when its distribution is invariant by the action of permutations on $\mathscr{P}_\mathbb{N}$. One can associate to every partition of the unit mass $\mathbf{x} = (x_1, \dots)$ an exchangeable random partition π by the following construction, which is often referred to as the paint-box process. The partition of the unit mass $\mathbf{x}$ is viewed as a (proper) probability measure on $\mathbb{N}$. We introduce a sequence of i.i.d. discrete variables $\eta_1, \eta_2, \dots$ distributed according to $\mathbf{x}$, *i.e.*, $\mathbb{P}(\eta_n = i) = x_i$, and then define a random partition π of $\mathbb{N}$ by declaring that two integers k, ℓ belong to the same block of π if and only if $\eta_k = \eta_\ell$. Note that π contains no singletons a.s., and, by the strong law of large numbers, that every block of π, say B, possesses an asymptotic frequency $|B| := \lim_{n\to\infty} n^{-1}\mathrm{Card}(B \cap [n])$. Further, the sequence $|\pi| = (|\pi_i|, i \in \mathbb{N})$ of the asymptotic frequencies of the blocks of π is a size-biased reordering of the sequence $\mathbf{x}$. Conversely, any exchangeable random partition π which contains no singleton a.s. is distributed as a mixture of such paint-boxes. In particular, all the blocks of π possess asymptotic frequencies a.s., and their rearrangement in decreasing order yields a random (proper) partition $|\pi|^\downarrow = \mathbf{x}$ of the unit mass whose law characterizes that of π.

The distribution of any exchangeable random partition π is determined by an Exchangeable Partition Probability Function (EPPF)

$$\mathrm{p} : \bigcup_{k\in\mathbb{N}} \mathbb{N}^k \to [0, 1], \quad \mathrm{p}(n_1, \dots, n_k) = \mathbb{P}(\pi_{|[n]} = \gamma(n_1, \dots, n_k)),$$

where $n = n_1 + \cdots + n_k$ and $\gamma(n_1, \dots, n_k)$ denotes an arbitrary partition of $[n]$ which has k non-empty blocks with cardinals $n_1, \dots, n_k$ (by the exchangeability property, the preceding

https://doi.org/10.1017/S0963548307008784 Published online by Cambridge University Press

probability does not depend on the choice for such a partition $\gamma(n_1,\dots,n_k)$, and in particular p is a symmetric function of the variables $n_1,\dots,n_k$).

We next focus on Poisson–Dirichlet distributions, and fix a pair of parameters (α,θ) with $0<\alpha<1$ and $\theta>-\alpha$. Let $\mathbf{x}_{\alpha,\theta}$ be a random partition of the unit mass with law $\mathrm{PD}(\alpha,\theta)$ and $\pi_{\alpha,\theta}$ a random partition of $\mathbb{N}$ distributed as a paint-box based on $\mathbf{x}_{\alpha,\theta}$. The EPPF $\mathrm{p}_{\alpha,\theta}$ of the $\mathrm{PD}(\alpha,\theta)$ random partition $\pi_{\alpha,\theta}$ is given by Pitman's sampling formula:

$$\mathrm{p}_{\alpha,\theta}(n_1,\dots,n_k)=\frac{(\theta/\alpha)_{k\uparrow}}{(\theta)_{n\uparrow}}\prod_{i=1}^{k}-(-\alpha)_{n_i\uparrow}, \tag{2.1}$$

where, for every integer ℓ and real number $a\notin\mathbb{Z}_-$,

$$(a)_{\ell\uparrow}=\Gamma(a+\ell)/\Gamma(a)=a(a+1)\cdots(a+\ell-1),$$

and for $\theta=0$, the ratio in the right-hand side of (2.1) is understood as $\alpha^{-1}(k-1)!/(n-1)!$.

2.2. Exchangeable Fragmentation–Coagulation processes

This section essentially follows Berestycki [3], with some minor modifications which are made to adapt the presentation to our purposes.

A Markov process $\Pi=(\Pi(t),t\geqslant 0)$ with values in $\mathscr{P}_{\mathbb{N}}$ is called an EFC process if it is exchangeable (*i.e.*, its law is invariant by the action of permutations on $\mathscr{P}_{\mathbb{N}}$-valued processes), and if, for every integer n, its restriction $\Pi_{|[n]}$ to $\mathscr{P}_{[n]}$ is a continuous-time Markov chain that only evolves by fragmentation of one block or by coagulation. This means that when $\Pi_{|[n]}$ starts from some configuration, say $\pi^{[n]}\in\mathscr{P}_{[n]}$, then the only sites at which it can make its first jump are those that can be obtained from $\pi^{[n]}$ by splitting exactly one of its blocks (which thus has at least two elements), or by merging at least two of its non-empty blocks. It is important to observe that an EFC process Π can be recovered from the family of its restrictions $\Pi_{|[n]}$, and thus its evolution is entirely determined by the jump rates of a family of continuous-time Markov chains with values in some finite space, even though the process Π itself is not a step process in general. It is easily seen that, with the obvious exceptions of the cases when the fragmentation or coagulation component is degenerate, the restrictions $\Pi_{|[n]}$ are irreducible and thus possess a unique invariant distribution.

The evolution of an EFC process may be easier to understand when the latter is viewed as an infinite particle system, in which particles correspond to blocks of the random partitions $\Pi(t)$. Specifically, as Π is an exchangeable process, for every $t\geqslant 0$, all the blocks of the partition $\Pi(t)$ have asymptotic frequencies a.s.[1] We write $\mathbf{X}(t)$ for the partition of the unit mass given by the sequence ranked in decreasing order of the asymptotic frequencies of the blocks of $\Pi(t)$. We think of the process $(\mathbf{X}(t),t\geqslant 0)$ as a particle system in which particles are characterized by their masses, and may split or coagulate with other particles. According to Theorem 7 of Berestycki [3], this particle system has the Feller property, and it is easily seen that its distribution determines that of Π.

One of the basic results about EFC processes (*cf.* Propositions 4 and 5 in [3]) states that their dynamics are entirely characterized by a pair of measures on the space of mass-partitions, which

[1] In fact, Berestycki [3] showed the much stronger result that with probability one, the partitions $\Pi(t)$ have asymptotic frequencies simultaneously for all $t\geqslant 0$.

https://doi.org/10.1017/S0963548307008784 Published online by Cambridge University Press

specify, respectively, the rates of multiple coagulation and of splittings, and two non-negative coefficients, one for the rate of binary coagulations and the other for the erosion. Here, we shall only consider the situation where there are no binary coagulation and no erosion, so the aforementioned coefficients are both zero. Further, we shall focus on the case when coagulation events are always simple, *i.e.*, at any given time, all the particles involved into a coagulation event merge in a single new particle. This enables one to represent the measure characterizing coagulation rates as some finite measure Λ on $]0,1]$. The measure on $\mathscr{P}_1$ that characterizes splitting rates will be denoted by ν; it fulfils the conditions

$$\nu(\{(1,0,\dots)\}) = 0 \quad \text{and} \quad \int_{\mathscr{P}_1} (1-x_1)\nu(\mathrm{d}\mathbf{x}) < \infty. \tag{2.2}$$

Informally, the dynamics of the particle system $\mathbf{X}$ (and hence also of Π) are determined by the measures Λ and ν as follows. Coagulation events involving a proportion $u \in]0,1]$ of particles (that are chosen uniformly at random) occur with rate $u^{-2}\Lambda(\mathrm{d}u)$, and each particle, say y, splits into a sequence of smaller particles $yx_1, yx_2, \dots$ with rate $\nu(\mathrm{d}\mathbf{x})$, independently of the other particles in the system. Conversely, given an arbitrary finite measure Λ on $]0,1]$ and an arbitrary measure ν on $\mathscr{P}_1$ such that (2.2) holds, there exists an EFC process having only simple (but not binary) coagulations, with characteristics Λ and ν.

This description of the evolution of the particle system $\mathbf{X}$ may be intuitively appealing; however, it would be difficult to make it formal because coagulation (respectively, splitting) events occur immediately whenever $\int u^{-2}\Lambda(\mathrm{d}u) = \infty$ (respectively, $\nu(\mathscr{P}_1) = \infty$). Working with the $\mathscr{P}_{\mathbb{N}}$-valued process Π and its restrictions enables one to circumvent this difficulty. Indeed, the transitions of the continuous-time Markov chains $\Pi_{|[n]}$ are defined rigorously in terms of Λ and ν as follows.

We first consider coagulation events. For all integers k, ℓ, n with $2 \leqslant k \leqslant \ell \leqslant n$, pick an arbitrary partition $\gamma \in \mathscr{P}_{[n]}$ which has ℓ non-empty blocks, and $\gamma' \in \mathscr{P}_{[n]}$ another partition which is obtained from γ by a simple coagulation event involving k of its blocks. Then the jump rate of $\Pi_{|[n]}$ corresponding to the coagulation event which transforms γ into γ' is

$$\mathsf{c}_n(\gamma,\gamma') = \mathsf{c}(\ell,k) = \int_{[0,1]} u^{k-2}(1-u)^{\ell-k}\Lambda(\mathrm{d}u).$$

These are the only non-zero rates for coagulation events.

We next turn our attention to splitting events. Pick an arbitrary partition $\gamma \in \mathscr{P}_{[n]}$ which is not the partition into singletons, and select an arbitrary block $B = (i_1, \dots, i_k)$ of γ with $k \geqslant 2$ elements. Then pick an arbitrary partition $\eta \in \mathscr{P}_{[k]}$ which is not trivial (*i.e.*, $\eta \neq ([k], \varnothing, \dots)$), and split B according to η. This means that we replace the block B by smaller blocks of the type $\{i_j, j \in C\}$, where C stands for a generic block of η, and leave the other blocks of γ unchanged. We write γ' for the resulting partition of $[n]$. Then the jump rate of $\Pi_{|[n]}$ corresponding to the splitting event which transforms γ into γ' is

$$\mathsf{s}_n(\gamma,\gamma') = \mathsf{s}(\eta) = \int_{\mathscr{P}_1} \nu(\mathrm{d}\mathbf{x})\mathbb{P}(\pi^{\mathbf{x}}_{|[k]} = \eta), \tag{2.3}$$

where $\pi^{\mathbf{x}}$ stands for an exchangeable random partition of $\mathbb{N}$ constructed as a paint-box based on $\mathbf{x}$. These are the only non-zero rates for splitting events. It is often convenient to rewrite (2.3) in terms of EPPF. Specifically, for every partition $\mathbf{x}$ of a unit mass, write $\mathrm{p}_{\mathbf{x}}$ for the EPPF of a

https://doi.org/10.1017/S0963548307008784 Published online by Cambridge University Press

paint-box based on $\mathbf{x}$. Then

$$\mathrm{s}(\eta) = \mathrm{s}(k_1, \dots, k_\ell) = \int_{\mathscr{P}_1} \nu(\mathrm{d}\mathbf{x}) \mathrm{p}_{\mathbf{x}}(k_1, \dots, k_\ell), \tag{2.4}$$

where $k_1, \dots, k_\ell$ are the cardinals of the (non-empty) blocks of η.

3. Some reversible EFC processes

Here, we fix two parameters $0 < \alpha < 1$ and $\theta > -\alpha$ and consider first the finite measure on $]0, 1]$ given by

$$\Lambda_{\theta/\alpha}(\mathrm{d}u) = (1-u)^{\theta/\alpha} \mathrm{d}u.$$

Next, following Basdevant [2], we introduce an infinite measure $\mathrm{PD}(\alpha, -\alpha)$ on $\mathscr{P}_1$ as follows. We denote as usual the generic partition of the unit mass by $\mathbf{x} = (x_1, x_2, \dots)$, and recall that the limit

$$L = L(\mathbf{x}) = \lim_{n\to\infty} n x_n^\alpha \tag{3.1}$$

exists $\mathrm{PD}(\alpha, 0)$-a.s. Further, it is well known that, for every real number a,

$$\mathbb{E}_{\alpha,0}(L^a) < \infty \iff a > -1,$$

where the notation $\mathbb{E}_{\alpha,0}$ indicates that the expectation is computed under the law $\mathrm{PD}(\alpha, 0)$. In this setting, the Poisson–Dirichlet distribution $\mathrm{PD}(\alpha, \theta)$ is absolutely continuous with respect to $\mathrm{PD}(\alpha, 0)$:

$$\mathrm{PD}(\alpha, \theta) = \frac{L^{\theta/\alpha}}{\mathbb{E}_{\alpha,0}(L^{\theta/\alpha})} \mathrm{PD}(\alpha, 0).$$

Even though $\mathbb{E}_{\alpha,0}(1/L) = \infty$, one can introduce the infinite measure

$$\mathrm{PD}(\alpha, -\alpha) = L^{-1} \mathrm{PD}(\alpha, 0). \tag{3.2}$$

More precisely, Basdevant [2] constructed the measure $\mathrm{PD}(\alpha, -\alpha)$ by using a sequence of independent beta variables and a variation of the Residual Allocation Model, and showed in Theorems 4.4 and 4.6 of [2] that these two constructions are equivalent. Further, $\mathrm{PD}(\alpha, -\alpha)$ fulfils (2.2).

Let $\Pi_{\alpha,\theta}$ denote an EFC process with characteristic measures $\Lambda = \Lambda_{\theta/\alpha}$ and $\nu = \mathrm{PD}(\alpha, -\alpha)$, and let $\mathbf{X}_{\alpha,\theta}$ stand for the particle system given by the process of the asymptotic frequencies of blocks of $\Pi_{\alpha,\theta}$ ranked in decreasing order. We are now able to state the main result of this note.

Theorem 3.1. *The Poisson–Dirichlet distribution* $\mathrm{PD}(\alpha, \theta)$ *is the unique reversible probability measure for the process* $\mathbf{X}_{\alpha,\theta}$.

Proof. Fix an integer $n \geqslant 2$, and denote by $\rho_{\alpha,\theta}^{[n]}$ the probability measure on $\mathscr{P}_{[n]}$, which is given by the image of $\mathrm{PD}(\alpha, \theta)$ by the paint-box process restricted to $[n]$. Pick positive integers $k < n$ and $n_1, \dots, n_k$ such that $n_1 + \cdots + n_k = n$, and consider an arbitrary partition $\gamma \in \mathscr{P}_{[n]}$ which has k non-empty blocks with sizes $n_1, \dots, n_k$. According to Pitman's sampling formula (2.1), we

https://doi.org/10.1017/S0963548307008784 Published online by Cambridge University Press

have

$$\rho_{\alpha,\theta}^{[n]}(\{\gamma\}) = \frac{(\theta/\alpha)_{k\uparrow}}{(\theta)_{n\uparrow}} \prod_{j=1}^{k} -(-\alpha)_{n_j\uparrow}.$$

Then pick some integer $i \leqslant k$ such that $n_i \geqslant 2$. Consider some partition η of $[n_i]$ with $\ell \in \{2, \dots, n_i\}$ non-empty blocks, say with sizes $n_{i1}, \dots, n_{i\ell}$, and denote by $\gamma' \in \mathscr{P}_{[n]}$ the partition of $[n]$ which is obtained from γ by splitting its ith block according to η. In particular, γ' has $k + \ell - 1$ non-empty blocks with respective sizes $n_1, \dots, n_{i-1}, n_{i1}, \dots, n_{i\ell}, n_{i+1}, \dots, n_k$. Another application of Pitman's sampling formula (2.1) gives

$$\rho_{\alpha,\theta}^{[n]}(\{\gamma'\}) = \frac{(\theta/\alpha)_{k+\ell-1\uparrow}}{(\theta)_{n\uparrow}} \left(\prod_{j=1}^{k} -(-\alpha)_{n_j\uparrow}\right) \left(\prod_{m=1}^{\ell} -(-\alpha)_{n_{im}\uparrow}\right) (-(-\alpha)_{n_i\uparrow})^{-1}.$$

On the one hand, the sampling formula of Proposition 4.3 in Basdevant [2] states that for $\nu = \mathrm{PD}(\alpha, -\alpha)$, the splitting rate $\mathtt{s}_n(\gamma, \gamma') = \mathtt{s}(\eta)$ defined by (2.3) and (2.4) is given explicitly by

$$\mathtt{s}_n(\gamma, \gamma') = \frac{(\ell - 2)!}{-(-\alpha)_{n_i\uparrow}} \prod_{m=1}^{\ell} -(-\alpha)_{n_{im}\uparrow}.$$

On the other hand, there is a unique way of recovering γ from γ' by a coagulation event. The latter involves exactly ℓ blocks amongst the $k + \ell - 1$ blocks of γ', and thus this coagulation event occurs with rate

$$\begin{aligned}
\mathtt{c}_n(\gamma', \gamma) = \mathtt{c}(k + \ell - 1, \ell) &= \int_{[0,1]} u^{\ell-2}(1-u)^{k-1} \Lambda_{\theta/\alpha}(\mathrm{d}u) \\
&= \int_{[0,1]} u^{\ell-2}(1-u)^{k-1+\theta/\alpha} \mathrm{d}u \\
&= \frac{\Gamma(\ell - 1)\Gamma(k + \theta/\alpha)}{\Gamma(k + \ell - 1 + \theta/\alpha)} \\
&= \frac{(\ell - 2)!(\theta/\alpha)_{k\uparrow}}{(\theta/\alpha)_{k+\ell-1\uparrow}}.
\end{aligned}$$

Putting the pieces together, we realize that there exists the identity

$$\rho_{\alpha,\theta}^{[n]}(\{\gamma\})\mathtt{s}_n(\gamma, \gamma') = \rho_{\alpha,\theta}^{[n]}(\{\gamma'\})\mathtt{c}_n(\gamma', \gamma). \tag{3.3}$$

In other words, for every integer $n \geqslant 2$, the detailed balance equation holds for the Markov chain $\Pi_{|[n]}$ and the probability measure $\rho_{\alpha,\theta}^{[n]}$. It should now be plain from the results recalled in Section 2 that this proves our claim. □

Our proof of Theorem 3.1 amounts to checking the detailed balance equation (3.3) by inspection, and does not make the statement quite intuitive. So it may be interesting to explain quickly some hints on this result. First, for $\theta = 0$, Pitman [14] showed that the uniform measure Λ_0 on $[0, 1]$ governs the coagulation rates of the Bolthausen–Sznitman coalescent C^{BS}. Further, the logarithmic time-change $t \to -\ln t$ transforms the Bolthausen–Sznitman coalescent into a time-inhomogeneous fragmentation process $F^{\mathrm{R}}(t) = C^{\mathrm{BS}}(-\ln t)$, $t \in]0, 1[$, which is closely related to Ruelle's cascades. As Basdevant [2] identified the instantaneous dislocation measure of F^{R} at time t as $t^{-1}\mathrm{PD}(t, -t)$, this suggests that combining fragmentation with rates governed by

https://doi.org/10.1017/S0963548307008784 Published online by Cambridge University Press

$\mathrm{PD}(t,-t)$ and coagulation with rates governed by Λ_0 might yield a stationary Markov process whose invariant law should be given by the distribution of the Bolthausen–Sznitman coalescent at time $-\ln t$, that is, $\mathrm{PD}(t,0)$. For a general $\theta > -\alpha$, Theorem 3.1 has its roots in the remarkable duality identity between coagulation and fragmentation operators based on Poisson–Dirichlet variables due to Pitman (see Theorem 12 in [14] or Theorem 4.4 in [5]), combined with intuition gained from the preceding analysis in the case $\theta = 0$.

We now complete this work by proving the convergence to equilibrium for the EFC process $\mathbf{X}_{\alpha,\theta}$.

Corollary 3.2. *For every configuration* $\mathbf{x} \in \mathscr{P}_1$, *the distribution of* $\mathbf{X}_{\alpha,\theta}(t)$ *given* $\mathbf{X}_{\alpha,\theta}(0) = \mathbf{x}$ *converges weakly to* $\mathrm{PD}(\alpha,\theta)$ *when* $t \to \infty$.

Proof. We consider the process $\Pi_{\alpha,\theta}$ with values in $\mathscr{P}_{\mathbb{N}}$ started from an exchangeable random partition given by a paint-box process based on $\mathbf{x}$. In particular, for every $t \geqslant 0$, the law of $|\Pi_{\alpha,\theta}(t)|^{\downarrow}$, the sequence of the asymptotic frequencies of the blocks of $\Pi_{\alpha,\theta}(t)$ ranked in decreasing order, is that of $\mathbf{X}_{\alpha,\theta}(t)$ given $\mathbf{X}_{\alpha,\theta}(0) = \mathbf{x}$. According to Theorem 8 of Berestycki [3], we know that $\Pi(t)$ converges in distribution as $t \to \infty$ to the equilibrium measure of Π, say $\rho_{\alpha,\theta}$. An application of Proposition 2.9 in [5] shows that, given $\mathbf{X}_{\alpha,\theta}(0) = \mathbf{x}$, $\mathbf{X}_{\alpha,\theta}(t)$ converges in distribution as $t \to \infty$ to the image of $\rho_{\alpha,\theta}$ by the map $\pi \to |\pi|^{\downarrow}$, which is the stationary law for the EFC process $\mathbf{X}_{\alpha,\theta}$. By Theorem 3.1, we know that the latter is $\mathrm{PD}(\alpha,\theta)$. □

We stress that, more generally, this argument can be applied to general EFC processes to establish the convergence to equilibrium for the mass-partition-valued process $\mathbf{X}$.

References

[1] Arratia, R., Barbour, A. D., and Tavaré, S. (2003) *Logarithmic Combinatorial Structures: A Probabilistic Approach*, Monographs in Mathematics, European Mathematical Society, Zürich.

[2] Basdevant, A.-L. (2006) Ruelle's probability cascades seen as a fragmentation process. *Markov Process. Rel. Fields* **12** 447–474.

[3] Berestycki, J. (2004) Exchangeable fragmentation–coalescence processes and their equilibrium measures. *Electron. J. Probab.* **9** 770–824. Available via: http://www.math.washington.edu/~ejpecp/viewarticle.php?id=1477&layout=abstract.

[4] Bertoin, J. (2001) Homogeneous fragmentation processes. *Probab. Theory Rel. Fields* **121** 301–318.

[5] Bertoin, J. (2006) *Random Fragmentation and Coagulation Processes*, Cambridge University Press, Cambridge.

[6] Diaconis, P., Mayer-Wolf, E., Zeitouni, O., and Zerner, M. P. W. (2004) The Poisson–Dirichlet law is the unique invariant distribution for uniform split-merge transformations. *Ann. Probab.* **32** 915–938.

[7] Durrett, R., Granovsky, B. L. and Gueron, S. (1999) The equilibrium behavior of reversible coagulation–fragmentation processes. *J. Theoret. Probab.* **12** 447–474.

[8] Erlihson, M. M. and Granovsky, B. L.(2004) Reversible coagulation–fragmentation processes and random combinatorial structures: Asymptotics for the number of groups. *Random Struct. Alg.* **25** 227–245.

[9] Gnedin, A. and Kerov, S. (2001) A characterization of GEM distributions. *Combin. Probab. Comput.* **10** 213–217.

[10] Kelly, F. P. (1979) *Reversibility and Stochastic Networks*, Wiley Series in Probability and Mathematical Statistics, Chichester.

https://doi.org/10.1017/S0963548307008784 Published online by Cambridge University Press

[11] Mayer-Wolf, E., Zeitouni, O. and Zerner, M. P. W. (2002) Asymptotics of certain coagulation–fragmentation processes and invariant Poisson–Dirichlet measures. *Electron. J. Probab.* **7**. Available via: http://www.math.washington.edu/~ejpecp/viewarticle.php?id=1309&layout=abstract.
[12] Möhle, M. and Sagitov, S. (2001) A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.* **29** 1547–1562.
[13] Perman, M., Pitman, J. and Yor, M. (1992) Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Rel. Fields* **92** 21–39.
[14] Pitman, J. (1999) Coalescents with multiple collisions. *Ann. Probab.* **27** 1870–1902.
[15] Pitman, J. (2002) Poisson–Dirichlet and GEM invariant distributions for split-and-merge transformation of an interval-partition. *Combin. Probab. Comput.* **11** 501–514.
[16] Pitman, J. and Yor, M. (1997) The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25** 855–900.
[17] Sagitov, S. (1999) The general coalescent with asynchronous mergers of ancestor lines. *J. Appl. Probab.* **36** 1116–1125.
[18] Schramm, O. (2005) Composition of random transpositions. *Israel J. Math.* **147** 221–243.
[19] Schweinsberg, J. (2000) Coalescents with simultaneous multiple collisions. *Electron. J. Probab.* **5–12** 1–50. Available via: http://www.math.washington.edu/~ejpecp/viewarticle.php?id=1270&layout=abstract
[20] Tsilevich, N. V. (2000) Stationary random partitions of positive integers. *Theor. Probab. Appl.* **44** 60–74.
[21] Whittle, P. (1986) *Systems in Stochastic Equilibrium*, Chichester, Wiley.

https://doi.org/10.1017/S0963548307008784 Published online by Cambridge University Press