Homogeneous fragmentations and coalescences

Julien Berestycki

Fragmentations et coalescences homogènes

Julien Berestycki

Fragmentations et coalescences homogènes

Fragmentations et coalescences homogènes

Julien Berestycki

5 décembre 2003

Remerciements

Mes remerciements vont avant tout à Jean Bertoin qui a dirigé cette thèse et accompagné mes premiers pas dans la recherche. Je voudrais lui exprimer ici toute ma gratitude pour sa confiance, ses conseils, et son soutien bienveillant.

Je suis reconnaissant à Quansheng Liu et Jason Schweinsberg qui ont bien voulu évaluer ce travail.

Je remercie vivement Alain Rouault, Sylvie Méléard, Marc Yor et Quansheng Liu qui me font l'honneur de faire partie de mon Jury.

Je voudrais exprimer ma reconnaissance à tous ceux qui m'ont accueilli dans le monde de la recherche. Sylvie Méléard et l'ensemble de l'équipe Modal'X m'ont offert un accueil chaleureux et amical en tant qu'ATER auquel j'ai été très sensible. J'ai souvent bénéficié de discussions toujours intéressantes et instructives avec Alain Rouault. Les cours passionnants de Jean-François Le Gall ont été une source d'inspiration et j'ai été heureux de la gentillesse avec laquelle j'ai été accueilli l'an dernier à son séminaire. J'ai eu la chance d'effectuer un séjour enrichissant au Technion sur l'invitation de Leonid Mytnik. Les conversations avec Nicolas Fournier ont toujours été stimulantes. J'aimerais aussi remercier Gregory Miermont avec qui j'ai beaucoup apprécié de pouvoir souvent discuter et parler de nos sujets.

L'ambiance chaleureuse et la solidarité de l'équipe des thésards de Paris VI a été importante pour moi au cours de cette thèse. Merci à Bénédicte, Victor, Joaquin, Karine, Sylvain, Alexis, Marc, Amaury, Ashkan, Mathilde, Jan, Hélène, Luciano et Giovanni.

Un grand merci à toute l'équipe administrative du laboratoire, Josette Saman, Nelly Lecquyer, Maryvonne de Béru, Philippe Macé, Jacques Portès, Caroline Boulic et bien sûr Geneviève Fournier, pour leur gentillesse et leur compétence.

Durant ces années de thèse, j'ai également eu le plaisir de collaborer avec Didier Faivre et plusieurs membres de son équipe sur des projets de recherche et développement, d'abord au sein de la banque CPR puis à CAI. Merci, donc, à Didier pour m'avoir fait confiance et pour m'avoir encadré avec amitié, à Antonin Chaix, Guillaume Bernis, Gaël Giraud, Shaojuan Huang, Eric Néré, Nicolas Baud. Merci également à Jean-Michel Lasry pour ses encouragements bienveillants, tant sur ma thèse qu'à la banque.

Je voudrais enfin dire toute mon affection et ma gratitude à ma famille, à maman, à papa qui m'a donné le goût des mathématiques, à mamie, à mes frères et soeurs Nathanaël, Noémie et Clara, à Marie-Anne et aux cousins, cousines, tantes. Je pense aussi à Suzanne qui n'est plus là.

Merci à tous mes amis pour leur soutien et leurs encouragements.

Enfin, je veux dire ici à Valentine combien sa présence à mes cotés tout au long de ces trois années a été essentielle et une source d'inspiration pour moi.

TABLE DES MATIÈRES

0. INTRODUCTION

Le travail présenté dans cette thèse se divise en trois chapitres, chacun correspondant à un article. Il s'agit de *Ranked fragmentations (fragmentations ordonnées)* [3], paru dans ESAIM P&S, *Multifractal Spectra of fragmentation processes (Spectre multifractal des processus de fragmentation)* [5], à paraître dans Journal of Statistical Physics, et de *Exchangeable Fragmentation Coalescence processes and their equilibrium measure (Processus échangeables de fragmentation-coalescence et leur mesure d'équilibre)* [4], qui sera prochainement soumis à revue avec comité de lecture. Comme les titres l'indiquent, ces travaux sont liés par un thème commun : les processus de fragmentation et de coalescence. Au sein de ce thème, tous trois se concentrent sur un même aspect : les processus dit *auto-similaires.*

Nous commencerons par introduire de façon volontairement heuristique les processus et les concepts communs aux trois articles. Une étude plus détaillée des approches probabilistes de ces modèles se trouve dans l'excellent article "survey" de Aldous [2].

0.1 Modèles discrets et comportements asymptotiques.

L'un des premiers modèles destinés à l'étude de systèmes de particules qui peuvent entrer en collision et fusionner fut introduit indépendamment par Marcus [18] à la fin des années 60 et Lushnikov [17] en 1975.

On peut le décrire de la façon suivante : soit une collection de n particules susceptibles de coalescer deux par deux. Une paire de particules donnée, disons une de masse x et une de masse y, peut donc coaguler pour former une particule de masse $x+y$ avec un certain taux $K(x,y)$. Si initialement seul un nombre fini de particules sont présentes et si le noyau de coagulation K est fini, ceci suffit à définir un processus de Markov de saut. Supposons que l'on ait par exemple n particules initiales, de masses respectives $x_1,x_2,...,x_n$. On se donne ensuite $n(n-1)/2$ variables $(e_{i,j})_{1\leq i<j\leq n}$ indépendantes distribuées exponentiellement et de paramètres respectifs $K(x_i,x_j)$. Soient $k,l \leq n$ tels que

$$e_{k,l} = \min_{1\leq i<j\leq n} \{e_{i,j}\}.$$

Les particules x_k et x_l fusionnent au premier temps de coalescence qui est $e_{k,l}$. On relance alors la procédure avec maintenant $n-1$ particules, indépendamment du passé.

Souvent, on s'intéresse au comportement limite lorsque le nombre initial

d'objets tend vers l'infini ($n \to \infty$). En particulier de récents progrès ont été réalisés pour montrer que le modèle de Marcus-Lushnikov convenablement renormalisé converge vers une solution des équations de Smoluchowski (on peut se reporter par exemple aux travaux de Norris [19, 20]). Les équations de Smoluchowski correspondent à une approche de champ moyen en *volume infini* des phénomènes de fragmentation coalescence à travers des systèmes d'équations déterministes. On suppose que le processus est spatialement homogène. L'état au temps t est donc entièrement caractérisé par les densités $n(x,t)$ en particules de masse x (avec $x \in \mathbb{N}$ si on suppose que les masses ne peuvent prendre que des valeurs discrètes, et $x \in \mathbb{R}+$ s'il y a un continuum de valeurs possibles). Le noyau $K(x,y)$ est le taux auquel des particules de masse x fusionnent avec des particules de masse y. Les détails précis du mécanisme de collision et de fusion des particules n'interviennent qu'à travers le noyau K. Ceci conduit naturellement à formuler un système d'équations différentielles connues sous le nom d'équations de Smoluchowski pour la coagulation.

On peut aussi définir des processus de fragmentation de la même façon : on introduit un noyau de dislocation $S(x,l)$ qui est le taux auquel des particules de taille x se subdivisent en particules de tailles xl et $x(1-l)$ (avec bien sûr $l \in]0,1[$). On peut définir des modèles de Marcus-Lushnikov où les deux phénomènes, coalescence et fragmentation, sont présents simultanément. Là encore on dispose de résultats de convergence vers une équation de Smoluchowski modifiée (voir les articles de Fournier et Giet, [13]).

Nous allons, quant à nous, nous intéresser au cas des limites en *volume fini*. On fait toujours tendre n vers l'infini, mais en gardant la masse totale du système fixée à une certaine valeur m. L'espace d'état du processus limite est donc l'ensemble des décompositions d'une masse m en un nombre dénombrable de fragments. Par changement d'échelle on peut toujours supposer que $m = 1$. La première construction rigoureuse de ce type de processus est due à Evans et Pitman dans [12] pour des noyaux de coalescence constants ($K(x,y) = k$), additifs ($K(x,y) = k(x+y)$) et plus généralement pour des noyaux vérifiant certaines conditions de régularité.

0.2 Espaces d'état

Dans la suite nous allons nous intéresser à trois façons de décrire ces processus qui correspondent à trois espaces d'état différents. Il s'agit des fragmentations/coalescences d'intervalle, des fragmentations/coalescences ordonnées et des fragmentations/coalescences de partitions.

Fragmentations/coalescences d'intervalle

Un bon exemple pour un objet de masse 1 sur lequel on désire faire agir une fragmentation ou une coalescence est l'intervalle]0,1[. On appelle $\mathcal{O}$ l'espace des ouverts de]0,1[. A chaque élément $O \in \mathcal{O}$ on peut associer sa décomposition en intervalles, c'est à dire (à la numérotation près) l'unique suite d'intervalles disjoints $I_n \subseteq]0,1[$ telle que $O = \cup_n I_n$. On interprète les I_n comme des fragments de]0,1[. On parlera de *fragmentation* pour un processus à valeurs dans $\mathcal{O}$ quand pour tous $t,s \geq 0$ les fragments à la date $t+s$ sont inclus dans un fragment à la date t et de *coalescence* dans le cas où tout fragment à la date t est inclus dans un fragment à la date $t+s$.

Des fragmentations de ce type ont par exemple été étudiées par Bertoin dans [7]. Des processus à valeurs dans $\mathcal{O}$ dont la dynamique correspond à une évolution discrète par fission - fusion des intervalles ont également été étudiées récemment, citons parmi d'autres Pitman [22], Durrett et Limic [11] ou Mayer-Wolf et. al. [10].

Fragmentations/coalescences ordonnées

En général on ne s'intéresse pas aux aspects spatiaux du problème, il suffit donc de caractériser le processus par la suite des masses des fragments. Ainsi, pour un processus à valeurs dans $\mathcal{O}$ les fragments ont une localisation précise et on dispose donc d'une information de type spatial là où l'on pourrait se contenter de la suite des longueurs des intervalles I_n. On adopte la convention de prendre la suite *ordonnée* des masses. L'espace d'état naturel est alors

$$\mathcal{S}^{\downarrow} := \{x_1 \geq x_2 \geq .... \geq 0, \sum_{i=1}^{\infty} x_i \leq 1\}. \tag{0.1}$$

On ne requiert pas que $\sum_i x_i = 1$ car, comme nous le verrons, il est possible qu'une partie de la masse se désagrège en "poussière" (ceci correspond au cas $O \in \mathcal{O}$ et $\lambda(O) < 1$ où λ est la mesure de Lebesgue).

Fragmentations/coalescences de partitions

Bien que $\mathcal{S}^{\downarrow}$ soit sans doute l'espace le plus naturel, nous devons introduire un cadre abstrait qui nous permette de définir rigoureusement les processus de fragmentation coalescence. Dans ce nouvel espace, les partitions de $\mathbb{N}$, on va bénéficier d'un outil crucial, la théorie des partitions échangeables de Kingman. Le coeur de la démonstration de Evans et Pitman dans [12] pour l'existence de processus de coalescence avec un nombre infini de fragments repose sur ce principe, de même que la construction des fragmentations homogènes de Bertoin dans [6]. Kingman lui-même avait déjà exploité cette idée dans [16, 14] pour construire rigoureusement le coalescent de Kingman (pour lequel $K \equiv 1$). Nous commençons par quelques notations et rappels sur les partitions échangeables de $\mathbb{N}$.

0.3 Partitions échangeables

On appelle $\mathcal{P}$ l'espace des partitions de $\mathbb{N}$. Un élément $\pi \in \mathcal{P}$ peut être décrit de deux façons : On peut l'identifier à une collection infinie de blocs disjoints de réunion $\mathbb{N}$. On note $\pi = (B_1,B_2,....)$ où la numérotation correspond soit à l'ordre du plus petit élément, soit à la convention suivante (qui est aussi celle adoptée dans le chapitre 1) : B_i est le bloc qui contient i lorsque ce dernier est le plus petit élément de son bloc, et sinon $B_i = \emptyset$. Dans le cas où il n'y a qu'un nombre fini de blocs non-vides, on autorise la suite à se terminer par des ensembles vides. Une autre façon de décrire une partition π de $\mathbb{N}$ est de l'identifier à une relation d'équivalence (notée $\overset{\pi}{\sim}$) sur $\mathbb{N}$ en posant

$$i \overset{\pi}{\sim} j \Leftrightarrow i \text{ et } j \text{ sont dans le même bloc de } \pi.$$

Une partition $\pi \in \mathcal{P}$ induit une partition sur $[n] = \{1,2,...,n\}$ notée $\pi_{|[n]}$ ("la restriction de π à $[n]$") en posant, pour tous $i,j \leq n$

$$i \overset{\pi_{|[n]}}{\sim} j \Leftrightarrow i \overset{\pi}{\sim} j.$$

On munit $\mathcal{P}$ de la distance suivante, qui le rend compact : pour $\pi^{(1)},\pi^{(2)} \in \mathcal{P}$ on pose

$$d(\pi^{(1)},\pi^{(2)}) = 1/n(\pi^{(1)},\pi^{(2)})$$

où $n(\pi_1,\pi_2) = \max\{n \in \mathbb{N} : \pi^{(1)}_{|[n]} = \pi^{(2)}_{|[n]}\}$.

On définit à présent la notion d'échangeabilité. On peut définir l'action d'une permutation σ de rang fini (i.e., $\exists k : \sigma(n) = n \ \forall n \geq k$) sur une partition $\pi \in \mathcal{P}$ de la façon suivante : on appelle $\sigma(\pi)$ la partition définie par $i \overset{\sigma(\pi)}{\sim} j \Leftrightarrow \sigma(i) \overset{\pi}{\sim} \sigma(j)$, c'est-à-dire que les blocs de $\sigma(\pi)$ sont les images de ceux de π par l'inverse de σ. Une variable aléatoire Π à valeurs dans $\mathcal{P}$ est dite échangeable si

$$\Pi \overset{\mathcal{L}}{=} \sigma(\Pi)$$

pour toute permutation de rang fini σ. Un processus $(\Pi(t),t \geq 0)$ est dit échangeable si Π et $\sigma(\Pi)$ ont la même loi en tant que processus pour toute permutation de rang fini σ.

Un exemple très important de partition échangeable est ce que Kingman appelle la partition "boîte de peinture". Imaginons que l'on dispose d'une collection dénombrable de couleurs : couleur 1 , couleur 2, etc... On se donne ensuite une boîte de peinture $x = (x_1,x_2,...) \in \mathcal{S}^{\downarrow}$, où x_i est la proportion de couleur i disponible. On colorie chaque entier indépendamment avec une

couleur choisie aléatoirement (couleur 1 avec probabilité x_1, couleur 2 avec probabilité x_2, etc.... et absence de coloriage avec probabilité $1 - \sum_i x_i$). La partition de $\mathbb{N}$ est définie par les blocs d'entiers qui ont la même couleur, les entiers non-coloriés restant des singletons. La loi de cette partition est notée μ_x.

L'importance de cet exemple vient de ce que Kingman [15, 16] (voir aussi [1] pour une preuve plus simple qui fait mieux apparaître le lien avec le théorème de de Finetti) a montré que toute partition échangeable est un mélange de boîtes de peinture. Plus précisément, si μ est la loi d'une partition aléatoire échangeable alors il existe une loi ν sur $\mathcal{S}^\downarrow$ telle que pour un ensemble (mesurable) $A \subseteq \mathcal{P}$

$$\mu(A) = \int_{\mathcal{S}^\downarrow} \mu_x(A)\nu(dx).$$

On dit que le mélange est dirigé par ν et nous noterons sa loi par μ_ν. Ceci signifie que pour tirer une partition aléatoire Π de loi μ_ν, tout se passe comme si on commençait par tirer une variable aléatoire X à valeurs dans $\mathcal{S}^\downarrow$ de loi ν, puis que, conditionnellement à $X = x$, on tirait ensuite Π de loi μ_x.

Pour nous, la masse d'un bloc B correspond à sa fréquence asymptotique $|B|$ qui est définie comme la limite

$$|B| := \lim_{n\to\infty} \frac{1}{n} \#\{k \leq n : k \in B\}$$

lorsqu'elle existe. Pour une partition "boîte de peinture" $x = (x_i)$, la loi des grands nombres implique que presque sûrement, simultanément pour tous les $i \in \mathbb{N}$, le bloc des entiers de couleur i a une fréquence asymptotique égale à x_i. Cette propriété est préservée par mélange et reste donc vraie pour toute partition échangeable : presque sûrement, tous les blocs d'une partition échangeable ont une fréquence asymptotique.

La même idée permet également d'associer une partition échangeable à un élément $O \in \mathcal{O}$: soit $(u_i)_{i\in\mathbb{N}}$ une suite de variables i.i.d. uniformes sur]0,1[. On définit une relation d'équivalence sur $\mathbb{N}$ en posant $i \sim j$ si et seulement si u_i et u_j sont dans la même composante connexe de O. Si on étudie une fragmentation ou une coalescence sur un objet E, cette procédure permet donc de changer d'espace pour travailler avec des processus échangeables à valeurs dans $\mathcal{P}$.

On appelle Λ l'application (non-continue) de $\mathcal{P}$ dans $\mathcal{S}^\downarrow$ défini par

$$\Lambda : \pi = (B_1,B_2,..) \to (|B_1|,|B_2|,...)^\downarrow$$

où $(|B_1|,|B_2|,...)^{\downarrow}$ est le réarrangement décroissant des fréquences asymptotiques $|B_i|$. Sous certaines hypothèses, les fréquences asymptotiques d'un processus échangeable existent pour tous les temps t simultanément et on peut passer de processus à valeurs dans $\mathcal{P}$ à des processus à valeurs dans $\mathcal{S}^{\downarrow}$ par l'application Λ.

0.4 Fragmentations homogènes et coalescents échangeables

Pour arriver aux modèles précis qui nous intéressent, deux étapes restent à franchir : nous devons généraliser les noyaux de dislocation et de fusion pour autoriser des coalescences et des divisions multiples puis imposer une condition d'*homogénéité*, ou plus généralement d'*auto-similarité*.

Dans [6, 7] Bertoin définit et étudie des fragmentations dites *homogènes*, respectivement *auto-similaires*. Il s'agit de fragmentations dans lesquelles chaque nouveau fragment démarre une nouvelle fragmentation indépendante de tout le reste (propriété de *fragmentation*) et de même loi que la fragmentation globale initiale à un changement de temps et d'échelle près (propriété d'*auto-similarité*). Pour prendre un exemple, une fragmentation binaire ne sera homogène que si son noyau de dislocation a la forme

$$S(x,l) = s(l)$$

pour une certaine fonction s (on rappelle que $S(x,l)$ est le taux auquel les particules de masse x se divisent en fragments de masses xl et $x(1-l)$). Si l'on choisit plutôt

$$S(x,l) = x^{\alpha} s(l)$$

on est alors dans le cas auto-similaire (avec indice d'auto-similarité α). Quand $\alpha < 0$ les dislocations surviennent de plus en plus rapidement à mesure que la taille des fragments diminue, au contraire lorsque $\alpha > 0$ le processus à tendance à ralentir sur les fragments de faible taille.

On donne ici la définition précise d'une fragmentation de partition auto-similaire, des définitions équivalentes peuvent être données pour les fragmentations d'intervalle ou les fragmentations ordonnées.

Définition 0.1. *Un processus de Markov échangeable* $(\Pi(t),t \geq 0)$ *à valeurs dans* $\mathcal{P}$ *est appelé une fragmentation auto-similaire d'indice* α *si*

– $\Pi(0)$ *est presque sûrement la partition triviale en un seul bloc,*

- $\Lambda(\Pi(t))$ *existe pour tous les* t *simultanément presque sûrement et est continu en probabilité,*
- *pour tous* $t,s \geq 0$, *sachant* $\Pi(t) = (B_1,B_2,...)$ *la distribution de* $\Pi(t+s)$ *est la même que celle de la partition aléatoire dont les blocs sont ceux des partitions* $(B_i \cap \Pi^{(i)}(s|B_i|^\alpha))_{i\in\mathbb{N}}$ *où les* $\Pi^{(i)}(.)$ *sont des copies i.i.d. du processus* Π.

Les divisions, dans une fragmentation auto-similaire, sont essentiellement de deux types et correspondent à deux phénomènes bien distincts :

- Les *dislocations* correspondent à des blocs B de masse positive r qui se divisent soudainement en donnant naissance à une collection de sous-blocs de masses rx_i (et à des singletons si $\sum_i x_i < 1$).
- L'*érosion* est un phénomène de perte de masse continu qui se traduit dans $\mathcal{P}$ par le fait que chaque entier $n \in \mathbb{N}$ est sélectionné avec un certain taux (noté c dans le chapitre 1 et c_e dans le chapitre 3) et est coupé de son bloc pour devenir un singleton.

Pour essayer de comprendre au moins intuitivement ce qu'est ce phénomène d'érosion on peut remarquer que pour tout $c_e \geq 0$, le processus déterministe à valeurs dans $\mathcal{S}^\downarrow$ défini par $X(t) = (e^{-c_e t},0,0,..)$ est bien une fragmentation homogène ordonnée (voir le chapitre 1). Cette fragmentation correspond au cas où il n'y a pas de dislocations mais uniquement de l'érosion.

Le taux d'arrivée des dislocations est contrôlé par une mesure ν_{Disl} sur $\mathcal{S}^\downarrow$ qui vérifie certaines propriétés d'intégrabilité et qui, dans un certain sens, généralise les noyaux de dislocation S. Plus précisément, pour une fragmentation homogène, pour toute masse $r \in]0,1[$, $\nu_{Disl}(dx)$ est le taux d'arrivée des dislocations d'une masse r en fragments de masse $(rx_1,rx_2,...)$ avec $(x_1,x_2,..) \in dx$.

Dans [6, 7] Bertoin montre que la distribution de toute fragmentation auto-similaire à valeurs dans $\mathcal{P}$ est caractérisée par la mesure ν_{Disl}, le paramètre d'érosion c_e et l'indice d'auto-similarité α. Une construction explicite en termes de processus ponctuels de Poisson est aussi proposée.

Les processus de coalescence sont plus délicats à étudier justement parce qu'ils n'ont pas cette propriété de branchement selon laquelle les fragments évoluent indépendamment les uns des autres.

Pitman et Schweinsberg proposent une approche qui repose sur l'existence de taux de coalescence homogènes : on dira qu'un processus de coalescence est *homogène* si le taux avec lequel les blocs fusionnent ne dépend pas des tailles des blocs impliqués. Le coalescent de Kingman, pour lequel chaque paire de blocs est susceptible de fusionner à taux 1 en est un exemple important.

Pitman a montré dans [21] comment construire des coalescents homogènes admettant des collisions multiples pour lesquels, lorsqu'une coagulation a lieu, tous les blocs impliqués fusionnent en un unique bloc. Schweinsberg a développé cette idée dans [23] en permettant des collisions multiples telles que plusieurs blocs fusionnent simultanément mais pour donner naissance à plusieurs nouveaux agrégats.

Dans son cours à l'IHP sur les processus de fragmentation et de coalescence [9], Bertoin a montré qu'on pouvait donner la définition suivante, plus proche de celle utilisée pour les fragmentations. On définit $Coag(\pi,\pi')$ comme la coagulation d'une partition $\pi \in \mathcal{P}$ par $\pi' \in \mathcal{P}$, c'est-à-dire la partition que l'on obtient quand on coalesce les blocs de π dont les numéros sont dans un même bloc de π'.

Définition 0.2. *On dira alors qu'un processus de Markov* $(\Pi(t),t \geq 0)$ *à valeurs dans* $\mathcal{P}$ *est une coalescence homogène si, pour tous* $t,s \geq 0$*, conditionnellement à* $\Pi(t) = \pi$ *on a*

$$\Pi(t+s) \stackrel{\mathcal{L}}{=} Coag(\pi,\Pi'(s))$$

où $\Pi'(s) \stackrel{\mathcal{L}}{=} \Pi(s)$.

Nous utiliserons dans [4] une définition équivalente qui consiste à considérer leurs restrictions :

Définition 0.3. *Un processus de Markov échangeable* $(\Pi(t),t \geq 0)$ *à valeurs dans* $\mathcal{P}$ *est une coalescence homogène si et seulement si ses restrictions* $(\Pi_{|[n]}(t),t \geq 0)$ *sont des chaînes de Markov (à espace d'état fini) dont les seules transitions possibles correspondent à des coalescences.*

Comme pour les fragmentations, les événements de coalescence possibles dans un processus homogène correspondent à deux types de phénomènes qualitativement différents. On peut avoir :

- soit des *collisions multiples* pour lesquelles une proportion strictement positive des blocs présents avant la coalescence sont impliqués. En particulier si une infinité de blocs étaient présents, un nombre infini d'entre eux sont impliqués.
- Soit des coalescents de Kingman, i.e., des coalescences au cours desquelles une paire de blocs fusionne pour n'en former qu'un seul.

Ainsi, par exemple quand on a un nombre infini de blocs, il n'y a jamais de coalescence qui implique exactement 3 blocs.

Bertoin a montré dans son cours à l'IHP (en préparation) que l'on peut exprimer les résultats de Pitman et Schweinsberg dans un formalisme très

semblable à celui utilisé pour les fragmentations. Le taux d'arrivée des collisions multiples est caractérisé par une mesure ν_{Coag} sur $\mathcal{S}^{\downarrow}$ qui vérifie certaines propriétés d'intégrabilité, et un paramètre $c_k \geq 0$ qui contrôle le taux des coalescences de Kingman. La distribution de toute coalescence homogène est complètement caractérisée par ce couple (ν_{Coag},c_k). On peut là aussi proposer une construction à partir de processus ponctuels de Poisson.

Nous présentons maintenant chacun des trois chapitres qui composent cette thèse.

0.5 Chapitre 1 : Fragmentation ordonnée

L'objet du travail présenté dans cette première section est d'étudier plus précisément les fragmentations ordonnées auto-similaires à l'aide des résultats de Bertoin sur les fragmentations à valeurs dans $\mathcal{P}$. Dans cette introduction, par souci de clarté, on se concentre surtout sur le cas des fragmentations homogènes.

On commence par les définir, sur le modèle de la définition 0.1, par leur semi-groupe.

Définition 0.4. *Un processus de Markov $X(t) = (X_1(t),X_2(t),...)$ à valeurs dans $\mathcal{S}^{\downarrow}$ est une fragmentation homogène ordonnée si*

- $X(0) = (1,0,0,..)$,
- *X est continu en probabilité,*
- *et pour tous $t,s \geq 0$, conditionnellement à $X(t)$, la loi de $X(t+s)$ est la distribution du réarrangement décroissant des termes des suites $Y^{(i)}(s)$, où les $Y^{(i)}(s)$ sont données par $Y^{(i)}(s) = X_i(t).X^{(i)}(s)$ avec $(X^{(i)}(s))_{i\in\mathbb{N}}$ une suite de variables i.i.d. à valeurs dans $\mathcal{S}^{\downarrow}$ de même loi que $X(s)$.*

Par exemple, pour toute fragmentation homogène au sens de la définition 0.1 $\Pi(t)$, à valeurs dans $\mathcal{P}$, le processus des fréquences asymptotiques $\Lambda(\Pi(t))$ est une fragmentation ordonnée homogène.

Notre premier résultat est de montrer qu'il n'y a en fait pas d'autres fragmentations auto-similaires :

Proposition 0.1. *Les fragmentations à valeurs dans $\mathcal{P}$ et celles à valeurs dans $\mathcal{S}^{\downarrow}$ sont liées par les relations suivantes :*

1. *Si Π est une fragmentation auto-similaire d'indice α à valeurs dans $\mathcal{P}$ alors $\Lambda(\Pi)$ a les distributions marginales fini-dimensionnelles d'une fragmentation ordonnée de même indice d'auto-similarité.*

2. *Si λ est une fragmentation ordonnée auto-similaire d'indice α, alors on peut construire Π_λ une fragmentation échangeable auto-similaire de même indice à valeurs dans $\mathcal{P}$ telle que $\Lambda(\Pi_\lambda) \stackrel{\mathcal{L}}{=} \lambda$.*

Ce résultat a pour conséquence que toute fragmentation ordonnée a une structure généalogique bien définie. Ceci n'est pas complètement évident dans la définition qui s'appuie sur le semi-groupe du processus. Pour reprendre les termes de Evans et Pitman [12], on a montré que les fragmentations ordonnées admettaient une famille de fonction de trace.

La proposition implique aussi que la caractérisation des $\mathcal{P}$-fragmentations auto-similaires par un triplet (ν,c,α) (Bertoin, [6]) se transpose sans difficulté dans ce nouveau cadre. Il est alors naturel de chercher une construction des $\mathcal{S}^\downarrow$-fragmentations auto-similaires par des processus ponctuels de Poisson. Seul le cas particulier des processus homogènes est traité dans le premier chapitre, la construction dans le cas auto-similaire étant renvoyée en annexe.

On commence par montrer que pour les fragmentations homogènes ordonnées l'érosion se traduit par un phénomène complètement déterministe.

Proposition 0.2. *Si $\tilde{\lambda}$ est une fragmentation homogène de paramètres (ν,0), alors pour tout $c \geq 0$ le processus $\lambda = (e^{-ct}\tilde{\lambda}(t), t \geq 0)$ est une fragmentation homogène de paramètre (ν,c).*

Il nous suffit donc de savoir construire des fragmentations sans érosion. C'est l'objet de la proposition suivante, qui est un analogue direct de celle donnée par Bertoin pour la construction des $\mathcal{P}$-fragmentations. La preuve en est cependant relativement technique. Nous utilisons ici la notation # pour désigner la mesure de comptage sur $\mathbb{N}$.

Théorème 0.1. *Soit λ une fragmentation homogène à valeurs dans $\mathcal{S}^\downarrow$ sans érosion ($c_e = 0$) et de mesure de dislocation ν (avec ν vérifiant $\nu(\{s : s_2 > 0\} = \infty)$). Alors*

1. *λ est un processus de saut pur.*
2. *Il existe un PPP $K = (S(t),k(t))_{t\geq 0}$ à valeurs dans $\mathcal{S}^\downarrow \times \mathbb{N}$ et de mesure d'intensité $\nu \otimes \#$, tel que les sauts de λ correspondent aux atomes de K. Plus précisément, λ ne saute qu'aux temps où $(S(t),k(t))$ a un atome, et à un tel temps, $\lambda(t)$ est obtenu à partir de $\lambda(t-)$ en disloquant le $k(t)$-ième fragment de $\lambda(t-)$ par $S(t)$ (i.e., on remplace $\lambda_{k(t)}(t-)$ par la suite $\lambda_{k(t)}(t-)S(t)$ et on réordonne la nouvelle suite de fragments). A l'inverse si $(S(t),k(t))$ est un atome, alors λ a un saut au temps t, i.e., λ_i saute au temps t pour $i = k(t)$.*

Ce théorème est en fait vrai sans la restriction sur ν mais cette hypothèse nous permet d'éviter des difficultés techniques et de nous concentrer sur les

cas les plus intéressants.

La dernière partie du premier chapitre exploite certaines conséquences de ces résultats pour analyser le comportement en temps petit des fragmentations ordonnées auto-similaires. Plus précisément on va s'intéresser au comportement des deux plus grands fragments en utilisant l'idée suivante : pour des temps proches de 0 le plus grand fragment d'une fragmentation à valeurs dans $\mathcal{P}$ contient presque sûrement 1, or on sait que la taille du fragment qui contient 1 peut s'exprimer à l'aide d'un subordinateur dont on connaît explicitement la loi. D'autre part, la taille de ce plus gros fragment, $X_1(t)$ tend vers 1. Le second plus gros fragment $X_2(t)$ est donc presque sûrement le plus gros fragment qui s'est détaché de X_1 avant t. Ceci nous permet d'utiliser des résultats de la théorie des processus records : soit $(s_2^{(1)}(u))$ la suite des tailles des plus gros fragments qui se détachent de X_1. On sait que c'est un PPP sur lequel on peut définir un processus record $R(t) := \sup_{u\leq t} s_2^{(1)}(u)$.

Proposition 0.3. *Soit*

$$\lambda = (\lambda(t),t \geq 0) = (\lambda_1(t),\lambda_2(t),...),t \geq 0)$$

une fragmentation homogène de $\mathcal{S}^{\downarrow}$ avec caractéristiques (ν,c). Alors :

1. *Il existe un subordinateur ξ avec dérive c et mesure de Lévy*

$$L(dx) = \nu(-\log s_1 \in dx) \ , \ x \in]0,\infty[$$

tel que p.s.

$$\lambda_1(t) = \exp(-\xi(t))$$

pour t assez petit.

2.
$$\lambda_2(t) \sim R(t), \qquad t \to 0+ \qquad p.s.$$

On peut montrer (voir annexe A) que le Théorème 1.1 ainsi que la Proposition 1.4 ci-dessus restent valable pour les fragmentations auto-similaire d'indice positif. On montre aussi que des résultats plus précis sur la loi limite du second fragment peuvent être obtenus sous des hypothèses de variation régulière pour la queue de distribution de la mesure intensité de $(s_2^{(1)}(u))$ et que ce résultat se généralise à tous les fragments dans le cas où la fragmentation est de plus binaire.

0.6 Chapitre 2 : Spectre multifractal des processus de fragmentation

On sait que pour une fragmentation homogène, la taille du plus grand fragment tend toujours vers 0, ce qui est donc le cas de tous les fragments. La vitesse de cette décroissance est la même pour presque tous les fragments. Plus précisément, Bertoin donne le résultat suivant dans [8] : soit $X(t)$ une fragmentation ordonnée homogène sans érosion. On introduit les mesures aléatoires

$$\rho_t(dy) = \sum_{i=1}^{\infty} X_i(t)\delta_{t^{-1}\log X_i(t)}(dy).$$

Il existe une valeur déterministe $v_{\text{typ}} > 0$ telle que

$$\rho_t \Rightarrow \delta_{-v_{\text{typ}}}$$

faiblement quand $t \to \infty$. Ceci implique que, pour des temps t grands, presque toute la masse est concentrée sur des fragments de taille approximativement $e^{-v_{\text{typ}}t}$. Le nombre de fragments de taille anormalement petite ou anormalement grande, peut être estimé par des techniques du type "grandes déviations".

On a vu que la loi d'une fragmentation homogène sans érosion est caractérisée par la mesure de dislocation ν_{Disl} (notée ν dans le chapitre 1 et 2). On va utiliser la fonctionnelle suivante :

$$\Phi(q) = \int_{\mathcal{S}^{\downarrow}} (1 - \sum_i x_i^{q+1})\nu_{Disl}(dx).$$

Comme nous l'avons dit plus haut, la taille du fragment qui contient 1 dans une fragmentation $\Pi(t)$ à valeurs dans $\mathcal{P}$ s'exprime à l'aide d'un subordinateur. Plus précisément, on peut montrer que le processus

$$\xi(t) := -\log(|B_1(t)|)$$

est un subordinateur. On sait que sa loi est caractérisée par sa transformée de Laplace

$$E(\exp(qX_t)) = \exp(t\Phi(q))$$

où Φ est justement la fonctionnelle définie ci-dessus. Le fragment qui contient 1 est important car, conditionnellement à la suite des masses $X(t) = \Lambda(\Pi(t))$, on a $P(|B_1(t)| = X_k(t)) = X_k(t)$ pour chaque $k \in \mathbb{N}$. La masse du premier bloc est donc tirée aléatoirement dans la suite X avec un biais par la

taille. Intuitivement, ceci est dû au fait que pour la fragmentation d'intervalle associée, B_1 correspond au bloc qui contient U_1, une variable aléatoire distribuée uniformément sur]0,1[. En ce sens on peut dire que Φ, qui en général ne permet pas de retrouver toute l'information contenue dans ν_{Disl}, caractérise le comportement d'un fragment typique. En particulier, la vitesse de fragmentation du fragment typique, v_{typ}, s'exprime très simplement à l'aide de Φ puisque

$$v_{\text{typ}} = \Phi'(0+).$$

En fait on peut lire tout le spectre des vitesses dans Φ : la vitesse de fragmentation minimale que nous noterons $v_{\min}$ est donnée par $\Phi'(\bar{p})$ où $\bar{p} > 1$ résout l'équation $p\Phi'(p-1) = \Phi(p-1)$. La vitesse de fragmentation maximale, que nous noterons $v_{\max}$, est "en général" infinie. Pour $v \in]v_{\min}, v_{\max}[$ on introduit $v \to \Upsilon_v$ la fonction inverse de Φ', i.e., $\Phi'(\Upsilon_v) = v$ et on définit

$$C(v) = (\Upsilon_v + 1)v - \Phi(\upsilon_v).$$

Le nombre de fragments de taille anormale est étroitement relié à cette fonction $C(v)$ puisque l'on a avec probabilité 1 que

$$\lim_{\epsilon \to 0} \lim_{t \to \infty} t^{-1} \log \left(\text{Card} \left\{ i \in \mathbb{N} : e^{-(v+\epsilon)t} \leq X_i(t) \leq e^{-(v-\epsilon)t} \right\} \right) = C(v), \quad (0.2)$$

En particulier la solution positive de $C(v) = 0$ est $v_{\min}$, c'est-à-dire le taux de décroissance du plus gros fragment, i.e., pour tout $\epsilon > 0$ on a $X_1(t)e^{t(v_{\min}+\epsilon)} \to \infty$ et $X_1(t)e^{t(v_{\min}-\epsilon)} \to 0$.

Considérons maintenant une fragmentation d'intervalle $(S(t), t \geq 0)$ sans érosion et sans perte de masse. Pour $x \in]0,1[$ on appelle $I_x(t)$ le fragment qui contient x au temps t et $|I_x(t)|$ sa taille. Quand elle existe on dit que la limite $-\lim_{t\to\infty} \log(|I_x(t)|)/t$ est la vitesse de fragmentation de x. On peut facilement montrer que presque sûrement un point u pris uniformément sur]0,1[a pour vitesse de fragmentation v_{typ}. Dit autrement : presque tous les points de]0,1[ont une vitesse de fragmentation qui est v_{typ}. Les points qui ont une vitesse différente, disons $v \neq v_{\text{typ}}$, forment des ensembles $\mathcal{G}_v$

$$\mathcal{G}_v := \{x \in]0,1[: \lim_{t\to\infty} -t^{-1} \log(|I_x(t)|) = v\}.$$

Il est assez facile, à l'aide de la propriété de branchement, de voir qu'ils sont soit vides, soit partout denses, presque sûrement, et toujours de mesure de Lebesgue nulle. La bonne notion pour étudier leur "taille" est la dimension de Hausdorff que nous noterons "Dim" dans la suite. On définit aussi $\overline{\mathcal{G}}_v$ et

$\underline{\mathcal{G}}_v$ comme suit :

$$\begin{aligned}\overline{\mathcal{G}}_v &:= \{x \in]0,1[: \limsup_{t\to\infty} -t^{-1}\log(|I_x(t)|) \leq v\},\\ \underline{\mathcal{G}}_v &:= \{x \in]0,1[: \liminf_{t\to\infty} -t^{-1}\log(|I_x(t)|) \geq v\}.\end{aligned}$$

Un point dans $\underline{\mathcal{G}}_v$ (resp. dans $\overline{\mathcal{G}}_v$) sera, pour t assez grand, dans un petit (resp. grand) fragment par rapport à e^{-vt}.

Le résultat de grandes déviations de Bertoin conduit naturellement pour les fragmentations d'intervalle à une analyse multifractale des vitesses de fragmentation, i.e., on peut calculer la fonction qui à v associe $\mathrm{Dim}(\mathcal{G}_v)$ (ou $\mathrm{Dim}(\underline{\mathcal{G}}_v)$ ou $\mathrm{Dim}(\overline{\mathcal{G}}_v)$).

Théorème 0.2. *Pour chaque $v \in]v_{min},v_{max}[$, presque sûrement*

$$\mathrm{Dim}(\mathcal{G}_v) = C(v)/v, \tag{0.3}$$

$$\mathrm{Dim}(\overline{\mathcal{G}}_v) = C(v)/v \text{ si } v \leq v_{typ} \text{ et } = 1 \text{ si } v \geq v_{typ}, \tag{0.4}$$

$$\mathrm{Dim}(\underline{\mathcal{G}}_v) = C(v)/v \text{ si } v \geq v_{typ} \text{ et } = 1 \text{ si } v \leq v_{typ}. \tag{0.5}$$

Les arguments utilisés dans la preuve s'appuient sur le lien étroit et déjà mentionné qui existe entre les fragmentations homogènes et les processus de branchement.

0.7 Chapitre 3 : Processus de fragmentation-coalescence échangeables et leur mesure d'équilibre.

Le dernier chapitre est consacré à la définition et à l'étude des processus de fragmentation-coalescence échangeables à valeurs dans $\mathcal{P}$ (nous les appellerons "processus EFC" dans la suite).

Définition 0.5. *Un processus de Markov échangeable $(\Pi(t),t \geq 0)$ à valeurs dans $\mathcal{P}$ est un processus EFC si et seulement si ses restrictions $(\Pi_{|[n]}(.))_{n\in\mathbb{N}}$ sont des chaînes de Markov (à espace d'états fini) dont les seuls sauts possibles correspondent soit à une coalescence, soit à la fragmentation d'un seul bloc.*

Nous montrons que chaque processus EFC correspond à un unique couple constitué d'une fragmentation homogène et d'une coalescence échangeable dont il est en fait la combinaison. On peut donc utiliser les résultats de base de [6, 7, 21, 23] pour caractériser les lois des processus EFC. Dans la proposition suivante, ν_{Disl} et ν_{Coag} sont des mesure sur $\mathcal{S}^{\downarrow}$ déjà évoquées qui vérifient certaines conditions d'intégrabilité, et c_e,c_k sont respectivement le paramètre d'érosion et le paramètre de Kingman. Plus précisément on rappelle que $\nu_{Disl}(dx)$ est le taux auquel un bloc de fréquence asymptotique r se fragmente en sous blocs de masses respectives rx_1,rx_2,.. avec $x \in dx$ tandis que $\nu_{Coag}(dx)$ est le taux auquel surviennent des coalescences telles que une proportion x_1 des blocs avant coalescence fusionnent en un unique nouveau bloc, une proportion x_2 fusionne en un autre bloc, etc... avec $x \in dx$. De plus on a

$$\int_{\mathcal{S}^{\downarrow}} (1 - x_1)\nu_{Disl}(dx) < \infty$$

et

$$\int_{\mathcal{S}^{\downarrow}} (\sum_{i=1}^{\infty} x_i^2)\nu_{Coag}(dx) < \infty.$$

Proposition 0.4. *La distribution d'un processus EFC* $\Pi(.)$ *est complètement caractérisée par la loi de* $\Pi(0)$*, par le couple de mesures* ν_{Disl},ν_{Coag} *et les paramètres* c_e,$c_k \in \mathbb{R}+$.

Les restrictions de Π étant des chaînes de Markov sur des espaces d'états finis il nous suffit de montrer qu'elles sont irréductibles et apériodiques (dans le cas où on a bien et de la coalescence et de la fragmentation) pour conclure qu'elles admettent chacune une mesure de probabilité invariante unique. On utilise alors le fait que les restrictions sont, par définition, compatibles pour conclure que pour un processus EFC Π,

Proposition 0.5. *Il existe une unique mesure de probabilité invariante* ρ *qui est en plus échangeable. On a*

$$\rho(.) = \delta_{\mathbf{0}}(.) \Leftrightarrow c_k = 0 \text{ et } \nu_{Coag}(.) \equiv 0$$

et

$$\rho(.) = \delta_{\mathbf{1}}(.) \Leftrightarrow c_e = 0 \text{ et } \nu_{Disl}(.) \equiv 0$$

où $\delta_{.}(.)$ *est la masse de Dirac.*

De plus, $\Pi(.)$ *converge en distribution vers* ρ.

On explore ensuite certaines propriétés de cette mesure d'équilibre. La première question que nous traitons est celle de savoir si ρ charge ou non les partitions avec un nombre fini de blocs. Le résultat suivant donne une

condition suffisante pour que $\rho(\{\pi \in \mathcal{P} : \#\pi < \infty\}) = 0$ où $\#\pi$ est le nombre de blocs de π.

Théorème 0.3. *Soit* $\Pi(.)$ *un processus EFC avec caractéristiques* ν_{Coag},ν_{Disl},$c_k \geq 0$ *et* $c_e \geq 0$. *Alors*

$$\nu_{Disl}(\mathcal{S}^{\downarrow}) = \infty \text{ ou } c_e > 0 \Rightarrow \rho(\{\pi \in \mathcal{P} : \#\pi < \infty\}) = 0.$$

Bien que l'on ne dispose pas d'une condition nécessaire dans le cas général, nous pouvons présenter un résultat partiel dans le cas où la partie fragmentation du processus n'a pas d'érosion et est "binaire" (chaque dislocation crée exactement deux fragments).

Proposition 0.6. *Si* $c_e = 0, \nu_{Disl}(\mathcal{S}^{\downarrow}) < \infty$ *et*

$$\nu_{Disl}(\{x \in \mathcal{S}^{\downarrow} : x_1 + x_2 < 1\}) = 0$$

(les dislocations sont binaires), alors

$$c_k > 0 \Rightarrow \rho(\{\pi \in \mathcal{P} : \#\pi < \infty) = 1.$$

La deuxième question que l'on peut se poser sur ρ est de savoir si cette mesure charge des partitions ayant de la poussière, c'est-à-dire une proportion strictement positive de singletons (on les note $dust(\pi)$) parmi ses blocs. On ne dispose là aussi que d'une condition suffisante pour l'absence de poussière :

Théorème 0.4. *Soit* $(\Pi(t), t \geq 0)$ *un processus EFC et* ρ *sa mesure de probabilité invariante.*

Alors

$$\int_{\mathcal{S}^{\downarrow}} \left(\sum_i x_i \right) \nu_{Coag}(dx) = \infty \text{ ou } c_k > 0 \Rightarrow \rho(\{\pi \in \mathcal{P} : dust(\pi) \neq \emptyset\}) = 0.$$

Finalement, on étudie quelques propriétés trajectorielles des processus EFC. Plus précisément, on aimerait savoir si un processus issu d'une partition infinie (i.e., avec un nombre infini de fragments) peut "descendre de l'infini" selon l'expression de Schweinsberg, en temps fini, ou même instantanément. A l'inverse on peut se demander si en partant d'une partition finie on peut atteindre l'infini. Nous donnons ici une condition suffisante pour que l'ensemble des temps (noté G) où le nombre de blocs non-vides est infini soit presque sûrement dense (dans ce cas le processus atteint l'infini immédiatement).

Théorème 0.5. *Soit* Π *un processus EFC tel que* $c_e > 0$ *ou* $\nu_{Disl}(\mathcal{S}^{\downarrow}) = \infty$. *Alors, p.s.* G *est partout dense.*

BIBLIOGRAPHIE

[1] D. J. Aldous. Exchangeability and related topics. In *École d'été de probabilités de Saint-Flour, XIII—1983*, volume 1117 of *Lecture Notes in Math.*, pages 1–198. Springer, Berlin, 1985.

[2] D. J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5(1):3–48, 1999.

[3] J. Berestycki. Ranked fragmentations. *ESAIM Probab. Statist.*, 6:157–175 (electronic), 2002.

[4] J. Berestycki. Exchangeable fragmentation-coalescence processes. *preprint*, 2003.

[5] J. Berestycki. Multifractal spectra of fragmentation processes. *J. Statist. Phys.*, to appear, 2003.

[6] J. Bertoin. Homogeneous fragmentation processes. *Probab. Theory Related Fields*, 121(3):301–318, 2001.

[7] J. Bertoin. Self-similar fragmentations. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(3):319–340, 2002.

[8] J. Bertoin. The asymptotic behaviour of fragmentation processes. *J. Euro. Math. Soc.*, to appear, 2003.

[9] J. Bertoin. Fragmentations et coalescences stochastiques. In preparation, 2003.

[10] P. Diaconis, E. Mayer-Wolf, O. Zeitouni, and M. P. W. Zerner. Uniqueness of invariant measures for split-merge transformations and the poisson-dirichlet law. *Ann. Probab.*

[11] R. Durrett and V. Limic. A surprising model arising from a species competition model. *Stoch. Process. Appl*, 102:301–309, 2002.

[12] S. N. Evans and J. Pitman. Construction of Markovian coalescents. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(3):339–383, 1998.

[13] N. Fournier and J.-S. Giet. On small particles in coagulation-fragmentation equations. *J. Statist. Phys.*, 111(5-6):1299–1329, 2003.

[14] J. F. C. Kingman. Random partitions in population genetics. *Proc. Roy. Soc. London Ser. A*, 361(1704):1–20, 1978.

[15] J. F. C. Kingman. The representation of partition structures. *J. London Math. Soc. (2)*, 18(2):374–380, 1978.

[16] J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13(3):235–248, 1982.

[17] A. Lushnikov. Evolution of coagulating systems. *J. Colloid Interface Sci.*, 45:549–556, 1973.

[18] A. H. Marcus. Stochastic coalescence. *Technometrics*, 10:133–143, 1968.

[19] J. R. Norris. Smoluchowski's coagulation equation: uniqueness, nonuniqueness and a hydrodynamic limit for the stochastic coalescent. *Ann. Appl. Probab.*, 9(1):78–109, 1999.

[20] J. R. Norris. Cluster coagulation. *Comm. Math. Phys.*, 209(2):407–435, 2000.

[21] J. Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27(4):1870–1902, 1999.

[22] J. Pitman. Poisson-Dirichlet and GEM invariant distributions for split-and-merge transformation of an interval partition. *Combin. Probab. Comput.*, 11(5):501–514, 2002.

[23] J. Schweinsberg. Coalescents with simultaneous multiple collisions. *Electron. J. Probab.*, 5:Paper no. 12, 50 pp. (electronic), 2000.

1. RANKED FRAGMENTATION

Abstract

In this paper we define and study self-similar ranked fragmentations. We first show that any ranked fragmentation is the image of some partition-valued fragmentation, and that there is in fact a one-to-one correspondence between the laws of these two types of fragmentations. We then give an explicit construction of homogeneous ranked fragmentations in terms of Poisson point processes. Finally we use this construction and classical results on records of Poisson point processes to study the small-time behavior of a ranked fragmentation.

1.1 Introduction

Splitting models are meant to describe an object that falls apart. Applications are numerous and may be found in various fields such as physical chemistry (aerosols, phase separation, polymerization), mathematical population genetics or astronomy (we refer to [3] for a survey on applications and motivations).

This paper focuses on self-similar ranked fragmentations. For the sake of describing our results, let us just give some heuristic descriptions while precise definitions will be given in the next sections.

Imagine a unit-mass object that fragments as time runs. We only consider the ordered sequence of the masses of the fragments so the state space is

$$\mathcal{S}^{\downarrow} := \{s = (s_1, s_2, ...), s_1 \geq s_2 \geq ... \geq 0, \sum_i s_i \leq 1\},$$

the situation where $\sum_i s_i < 1$ corresponding to the fact that a part of the initial mass has been lost, i.e. the sum of the masses of the remaining fragments is less than the original total mass.

Let $\lambda = (\lambda(t), t \geq 0)$ be a Markov process with values in $\mathcal{S}^{\downarrow}$. Call λ a self-similar ranked fragmentation if it fulfills the *scaling* and *fragmentation* properties.

The *scaling* property means that there exists a real number α, called the index of self-similarity, such that if $\mathbb{P}_r$ is the law of λ started from $(r, 0, 0, ..)$ then the distribution of $(r\lambda(r^{\alpha}t), t \geq 0)$ under $\mathbb{P}_1$ is $\mathbb{P}_r$.

The *fragmentation* property is a version of the branching property i.e. for any $u, t \geq 0$, for any $s = (s_1, s_2, ..) \in \mathcal{S}^{\downarrow}$, conditionally on $\lambda(u) = s$, $\lambda(t+u)$ has the same distribution as the variable obtained by concatenating and ordering the sequences $\lambda^{(1)}, \lambda^{(2)}, ..$ where for each i, $\lambda^{(i)}$ has the distribution of $\lambda(t)$ under $\mathbb{P}_{s_i}$.

Here is a simple prototype taken from Brennan and Durrett [10, 11] who consider the following model for polymer degradation : A particle of mass m splits with exponential rate m^α, $\alpha \in \mathbb{R}^+$, and gives rise to two particles of mass Vm and $(1-V)m$, where V is a random variable with values in $(0,1)$ independent of the past. The new particles follow the same dynamic independently. The ordered sequence of the masses of the particles is a self-similar ranked fragmentation of index α.

This example can be extended in two ways. First one can suppose that when a particle splits, it might give birth to any number of particles, possibly infinite, and not just two. Second, in the example of Brennan and Durrett, the splitting times are "discrete", the first time of splitting is almost surely strictly positive. It is natural to consider more generally the case where fragmentation may occur continuously. For instance this happens for the fragmentation process obtained by logging the continuous random tree of Aldous in [1].

A classical tool in the study of ranked fragmentations or coalescences is to use a somewhat different state-space and to introduce the so-called *partition valued fragmentation* (see for instance Kingman in [13] or Evans and Pitman in [12]). Roughly speaking a partition fragmentation, say $\Pi(t)$, is a process that lives in the space of partitions of $\mathbb{N}$, such that for any $0 < s \leq t$, $\Pi(t)$ is a refinement of $\Pi(s)$. A way to construct such a fragmentation which makes clear the connection with the above particle model is the following : imagine an object E endowed with a unit mass measure μ that falls apart as time runs. Call *object fragmentation* the process $F(t)$ with values in partitions of E that describes this fragmentation. Next, let $(u_i)_{i\in\mathbb{N}}$ be a sequence of iid E-valued variables with distribution μ and for each t let $\Pi_F(t)$ be the partition of $\mathbb{N}$ such that for all i and j in $\mathbb{N}$, i and j belong to the same block of $\Pi(t)$ iff u_i and u_j are in the same fragment of E at time t. By the SLLN we can recover the mass of a fragment as the asymptotic frequency of the corresponding block. Then Π_F is a partition fragmentation.

Using partition fragmentations to construct ranked fragmentations is typical of the existing results. These constructions benefit from two important features : there is a clear genealogical structure, and partition fragmentations, when they are self-similar, are characterized by an index of self-similarity and a so-called characteristic exchangeable measure, on which results concerning exchangeability can be usefully applied (see [2] for a survey on exchangeability).

However partition-valued fragmentations are perhaps less natural and could be less general than ranked fragmentations, precisely because we have

endowed them with this extra genealogic structure. In other words it is not clear that an arbitrary ranked fragmentation can be studied through partition fragmentations.

In section 2 we show that it is in fact the case. More precisely, if we give ourselves a ranked fragmentation λ, there is a partition fragmentation Π such that the asymptotic frequencies of Π have the distribution as λ.

In the next section we use this equivalence between ranked and partition fragmentations to give a Poisson construction of homogeneous ranked fragmentation which is an analogue of that given in [5] for partition fragmentations. The difficulty comes from the fact that we can no longer use a genealogic structure, which played a crucial role in the partition case.

In section 4, this construction allows us to tackle the study of small time behavior of a ranked fragmentation. We show that the 2nd largest fragment, correctly renormalized, behaves as the record of the size of the particles detaching from the main fragment.

1.2 Definitions and first properties

1.2.1 Ranked Fragmentations

For each l in $[0,1]$ let $P(l)$ be a probability on $\{s \in \mathcal{S}^{\downarrow} : \sum_i s_i \leq l\}$, the space of all the possible fragmentations of l. Then for $L = (l_1, l_2, ...) \in \mathcal{S}^{\downarrow}$, define $P(L)$ as the distribution on $\mathcal{S}^{\downarrow}$ of the concatenation and the decreasing rearrangement of independent $\mathcal{S}^{\downarrow}$-valued variables with respective law $P(l_i)$. Call $(P(L), L \in \mathcal{S}^{\downarrow})$ a fragmentation kernel on $\mathcal{S}^{\downarrow}$. One says that the family $(P(l), l \in [0,1])$ generates $(P(L), L \in \mathcal{S}^{\downarrow})$.

Definition 1.1. *An $\mathcal{S}^{\downarrow}$-valued process $\lambda(.)$ is called a $\mathcal{S}^{\downarrow}$-fragmentation if it is a time-homogeneous Markov process such that*

1. *λ is continuous in probability and starts from $\lambda(0) = (1, 0, 0, ...)$ a.s.*
2. *the transition semigroup $(P_t(L))$ of λ is given by fragmentation kernels.*

In words, at a given time t, each fragment of $\lambda(t) = (\lambda_1(t), \lambda_2(t), ...)$, say $\lambda_i(t)$, gives rise to an independent fragmentation process whose distribution only depends on the value $\lambda_i(t)$. λ is the concatenation and the reordering of all those processes.

For $l \in [0,1]$, let g_l be the function from $\mathcal{S}^{\downarrow} \to \mathcal{S}^{\downarrow}$ defined by

$$g_l : x = (x_1, x_2, ..) \to (lx_1, lx_2, ...).$$

Definition 1.2. *The fragmentation λ, with transition kernels generated by the family $(P_t(l); t \geq 0, l \in [0,1])$ is said to be self-similar with index $\alpha \in \mathbb{R}$ if (in the notations introduced above) for all $l \in [0,1]$ the distribution $P_t(l)$ coincides with the image of $P_{l^\alpha t}(1)$ by g_l.*

When $\alpha = 0$ the fragmentation is said to be homogeneous.

$\mathcal{S}^\downarrow$ is endowed with the uniform distance. Note that for any $s = (s_1, s_2, ...) \in \mathcal{S}^\downarrow$ we must have for every $k \in \mathbb{N}, s_k \leq \frac{1}{k}$, and thus the uniform and pointwise convergences are the same.

In [5] Bertoin showed that the semi-group of a homogeneous partition fragmentation (the definition and main properties of partition fragmentations are discussed in section 2.2) has the Feller property. As we are working with a different topology his result does not apply. However, we have the following

Proposition 1.1. *(Feller property) Under the weak topology on probability measures, the semi-group P_t of a self-similar ranked fragmentation of index α, fulfills the Feller property. That is $\forall t \geq 0$ the map*

$$L \to P_t(L)$$

is continuous on $\mathcal{S}^\downarrow$ and for each fixed $L \in \mathcal{S}^\downarrow$, $P_t(L)$ converges to the Dirac mass at L as $t \to 0$.

Proof. The second point is simply the continuity in probability of a self-similar fragmentation.

For the first point consider a sequence $(L_n, n \in \mathbb{N})$ in $\mathcal{S}^\downarrow$ which converges to $L \in \mathcal{S}^\downarrow$. Note $L_n = (l_1^{(n)}, l_2^{(n)}, ...)$, then for all k, $l_k^{(n)} \to l_k$.

Let $(Y_i(t))_{i\in\mathbb{N}}$ be a sequence of iid $\mathcal{S}^\downarrow$-fragmentations with same semi-group $(P_t(S), S \in \mathcal{S}^\downarrow, t \geq 0)$, then, by definition, for all $n \in \mathbb{N}$ the $\mathcal{S}^\downarrow$ random variable $Z^{(n)}(t)$, obtained by the decreasing rearrangement of the terms $g_{l_i^{(n)}}\left(Y_1\left(t(l_i^{(n)})^\alpha\right)\right)$ for i in $\mathbb{N}$:

$$Z^{(n)}(t) = \left(g_{l_1^{(n)}}\left(Y_1\left(t(l_1^{(n)})^\alpha\right)\right), g_{l_2^{(n)}}\left(Y_2\left(t(l_2^{(n)})^\alpha\right)\right), ...\right)^\downarrow$$

has law $P_{L^{(n)}}(t)$. In the same way

$$Z(t) = (g_{l_1}(Y_1(t(l_1)^\alpha)), g_{l_2}(Y_2(t(l_2)^\alpha)), ...)^\downarrow$$

has law $P_L(t)$. Now fix $\epsilon > 0$, and take $N > \frac{2}{\epsilon}$. Then

$$\forall k \geq N, \forall n \in \mathbb{N}, l_k^{(n)} < \epsilon/2.$$

Thus for all ω

$$\sup_{k\geq N}\left(\text{dist}\left(g_{l_k^{(n)}}(Y_k(t(l_k^{(n)})^\alpha)), g_{l_k}(Y_k(t(l_k)^\alpha))\right)\right) < \epsilon.$$

On the other hand, by the continuity in probability of the processes $(Y_i)_{i\in\{1,\dots,N-1\}}$, we have that almost surely

$$P\left[\sup_{k\in\{1,\dots,N-1\}}\left(\text{dist}\left(g_{l_k^{(n)}}(Y_k(t(l_k^{(n)})^\alpha)), g_{l_k}(Y_k(t(l_k)^\alpha))\right)\right) > \epsilon\right] \underset{n\to\infty}{\to} 0.$$

Hence dist $\left(Z^{(n)}(t), Z(t)\right)$ converges in probability, and thus in law, to 0. □

1.2.2 Partition Fragmentations

Most of the results on fragmentation available in the literature are (or can be) formulated in terms of a type of fragmentation called *partition fragmentation*, which is basically a process that can be described as a partition of $\mathbb{N}$ getting finer as time runs. We recall in this section several well-known facts that we will use repeatedly in the following. We refer to Bertoin [5], Bolthausen and Sznitman [9], Evans and Pitman [12] and Kingman [13] (among others) for results and discussions on partition valued processes and the link between partition fragmentations/coalescences and ranked fragmentations/coalescences.

More precisely, call a subset of $\mathbb{N}$, say B, a "block". When the limit

$$|B| := \lim_{n\to\infty}\frac{1}{n}\text{Card}\{0\leq k\leq n : k\in B\}$$

exists, it is called the asymptotic frequency of B. A partition of $\mathbb{N}$ can be thought of as a sequence $B_1, B_2, \dots$ of disjoint blocks whose union is $\mathbb{N}$. The labelling obeys the following rule : if B_i is not empty, then its least element is i. Call $\mathcal{P}$ the space of the partitions of $\mathbb{N}$, and recall that $\mathcal{P}$ is a compact metric space (see [13]).

A finite permutation σ (i.e. a bijection $\mathbb{N}\to\mathbb{N}$ such that $\sigma(n) = n$ for n large enough) acts on a partition π in the following way : for any i and j in $\mathbb{N}$, i and j are in the same block of $\sigma(\pi)$ iff $\sigma(i)$ and $\sigma(j)$ are in the same block of π, this equivalence relation can be identified as a partition and thus completely defines $\sigma(\pi)$.

A measure μ on $\mathcal{P}$ is said to be *exchangeable* if for any measurable set $A\subseteq\mathcal{P}$, for any finite permutation σ

$$\mu(A) = \mu(\sigma(A)),$$

where σ acts on the sets in the obvious way.

A $\mathcal{P}$-valued process Π is said to be exchangeable if the permuted process $\sigma(\Pi)$ has the same distribution as the original process Π. For instance the $\mathcal{P}$-valued process $\Pi_F(t)$ presented in the introduction is exchangeable.

For all $B \subseteq \mathbb{N}$, let P_B be a probability on the partitions of B. For all $\pi = (B_1, B_2, ..) \in \mathcal{P}$, let P_π be the distribution of the partition with blocks $B_{(1,1)}, B_{(1,2)}, ..., B_{(2,1)}, B_{(2,2)}, ...$ where $\pi^{(i)} = (B_{(i,1)}, B_{(i,2)}, ...)$ is a partition of B_i and has law P_{B_i}. The family $(P_\pi, \pi \in \mathcal{P})$ is, in the terminology of Pitman [15], a fragmentation kernel on $\mathcal{P}$.

Definition 1.3. *Call $\mathcal{P}$-fragmentation any exchangeable $\mathcal{P}$-valued Markov process, starting from the trivial partition ($\mathbb{N}$ is the only non empty block), which is continuous in probability and has fragmentation kernels as its transition semi-group.*

If π is a random exchangeable partition, by a result of Kingman [13] (see also Aldous [2] for a simpler proof), every block of π has an asymptotic frequency almost surely, i.e. $|B_i|$ exists with probability 1 for all $i = 1, ...$

We call an exchangeable $\mathcal{P}$-valued process Π *nice* if with probability 1, $\Pi(t)$ has asymptotic frequencies for all $t \geq 0$ simultaneously. Evans and Pitman [12] have shown that it is always the case when Π is an exchangeable $\mathcal{P}$-process with proper frequencies (i.e. for each $t \geq 0$, $\sum_{i\in\mathbb{N}} |B_i(t)| = 1$ almost surely), and Bertoin [5] proved that so-called homogeneous fragmentation were nice. Observe that when $\Pi(t)$ is nice, the ordered sequence of the asymptotic frequencies is well defined and is a $\mathcal{S}^\downarrow$-valued process.

As we shall construct a Markovian semi-group on $\mathcal{S}^\downarrow$, we need a notion slightly more general than the asymptotic frequency, well defined for any subset B of $\mathbb{N}$. We write

$$\Lambda(\Pi(t)) = (\Lambda_1(\Pi(t)), \Lambda_2(\Pi(t)), ...)^\downarrow = (\lambda_1(t), \lambda_2(t), ..)$$

for the decreasing rearrangement of the quantities

$$\Lambda_i(\Pi(t)) = \liminf_{n\to\infty} \frac{1}{n} \#\{k \leq n : k \in B_i(t)\}.$$

By extension we also define

$$\Lambda(B) = \liminf_{n\to\infty} \frac{1}{n} \#\{k \leq n : k \in B\}$$

for any $B \subseteq \mathbb{N}$.

Λ is a functional of $\Pi(t)$ that takes its values in $\mathcal{S}^{\downarrow}$. We stress that Λ is not continuous.

Next for every $C \subseteq \mathbb{N}$ and every $\pi = \{B_1, B_2, ...\} \in \mathcal{P}$, we define the partition of C induced[1] by π :

$$\pi \cap C = (B_1 \cap C, B_2 \cap C, ...)\,.$$

Definition 1.4. *A $\mathcal{P}$-fragmentation $\Pi = (\Pi(t), t \geq 0)$ is called self-similar with index $\alpha \in \mathbb{R}$ if :*

1. *Π starts a.s. from the trivial partition.*
2. *The ranked fragmentation $\Lambda(\Pi)$ associated with Π is continuous in probability.*
3. *For every $B \subseteq \mathbb{N}$, $\forall t \geq 0$ $P_B(t)$ (in the above notations) is the distribution of $\Pi(t\Lambda(B)^{\alpha}) \cap B$.*

When $\alpha = 0$ we will say that Π is a homogeneous fragmentation.

Following Kingman [13] (see also [2] for a survey), to each $s = (s_1, s_2, ...) \in \mathcal{S}^{\downarrow}$ one can associate a unique exchangeable probability measure μ_s on $\mathcal{P}$ such that μ_s-almost every partition has ranked asymptotic frequencies s.

This is how one proceeds : let $(X_i)_{i\in\mathbb{N}}$ a family of iid variables such that $\forall k \in \mathbb{N}, P(X_i = k) = s_k$ and $P(X_i = 0) = 1 - \sum_k s_k$, then define the s - paintbox [2] partition (or "s-paintbox process") Π by the equivalence relation

$$\forall i, j \in \mathbb{N}, i \sim j \Leftrightarrow X_i = X_j > 0.$$

We denote by μ_s the law of the s-paintbox process. It is clear by the LLN that μ_s-almost surely $\Lambda(\Pi) = s$.

[1] there is in fact another natural way of defining this partition : it is to take the image of π by the mapping that sends $\mathbb{N}$ onto $C = \{c_1, c_2, ...\}$ (where $c_1 < c_2 < ...$) :

$$\pi \circ C = (\{c_j : j \in B_i\}_{i=1,...})$$

Suppose now that π is an exchangeable random $\mathcal{P}$-valued variable, for all $k > 0$, for any finite permutation σ such that

$$\forall i \leq k; \sigma(i) = c_i,$$

π and $\sigma(\pi)$ have same law, thus in the sense of the equality of the finite-dimensional margins $\pi \circ C$ and $\pi \cap C$ have same law.

[2] The reason for the name (due to Kingman) is the following : imagine that we have a choice of colors $(c_k)_{k\in\mathbb{N}}$. Then paint each integer n independently with a randomly chosen color, c_k with probability s_k. Then the partition of $\mathbb{N}$ defined by the equivalence relation "being of the same color" is the s-paintbox process.

For each self-similar $\mathcal{P}$-fragmentation one can take the associated $\mathcal{S}^{\downarrow}$ ranked fragmentation, thus defining a map from $\mathcal{P}$-fragmentation laws into $\mathcal{S}^{\downarrow}$-fragmentation laws. Suppose now that Π_1 and Π_2 are two self-similar $\mathcal{P}$-fragmentations such that for any fixed t the $\mathcal{S}^{\downarrow}$ variables $\Lambda(\Pi_1(t))$ and $\Lambda(\Pi_2(t))$ have same law. $\Pi_1(t)$ and $\Pi_2(t)$ being exchangeable, by de Finetti's theorem (see [2]) one can show that they are mixtures of paintbox processes directed respectively by $\Lambda(\Pi_1)$ and $\Lambda(\Pi_2)$, i.e.

$$P(\Pi_{1,2} \in A) = \int_{\mathcal{S}^{\downarrow}} \mu_s(A) P(\Lambda(\Pi_{1,2}) \in ds).$$

We conclude that they have the same distribution. So to every $\mathcal{P}$-fragmentation corresponds a different $\mathcal{S}^{\downarrow}$-fragmentation. Our first result will be to show that there is in fact a one to one relation.

1.2.3 From Ranked to Partition Fragmentations

Let Π be a nice self-similar fragmentation of index α, then it is known (see [5, 13]) that its asymptotic frequencies $\Lambda(\Pi)$ form a self-similar ranked fragmentation of index α. In the converse direction, Evans and Pitman in [12] showed that given a ranked fragmentation that admits a so-called tracking function one can always see it as the image of a corresponding partition fragmentation. We shall extend this result by showing that this is the case for any $\mathcal{S}^{\downarrow}$-fragmentation, which means that any self-similar $\mathcal{S}^{\downarrow}$-fragmentation admits a family of tracking functions.

Proposition 1.2. *We have the following relations between $\mathcal{S}^{\downarrow}$ and $\mathcal{P}$ fragmentations :*

1. *If Π is a $\mathcal{P}$-fragmentation then $\Lambda(\Pi)$ has the finite-dimensional marginal distributions of an $\mathcal{S}^{\downarrow}$-fragmentation. Moreover Λ preserves self-similarity.*
2. *if λ is a $\mathcal{S}^{\downarrow}$-fragmentation, then we can construct Π_λ an exchangeable $\mathcal{P}$-fragmentation such that $\Lambda(\Pi_\lambda) \overset{\mathcal{L}}{=} \lambda$. Moreover this construction preserves self-similarity.*

The first point is clear, the difficulty here lies in the second part of this proposition. The main idea is that as $\mathcal{P}$ is a compact metric space, it is enough to construct an adequate Markovian semi-group to ensure the existence of the desired $\mathcal{P}$-process. Then the conservation of the index will be a simple consequence of our construction.

Let $(P_t(S), t \geq 0, S \in \mathcal{S}^{\downarrow})$ be a transition kernel on $\mathcal{S}^{\downarrow}$ generated, in the notation of definition (1.1), by the family $(P_t(l), t \geq 0, l \in [0,1])$.

Let $\tilde{P}_t(l)$ be the image of $P_t(l)$ by $g_{l^{-1}}$, the map $(x_1, x_2, ...) \to (x_1/l, x_2/l, ...)$. Let $(R(t,l), l \in [0,1], t \geq 0)$ be a family of probability measures on $\mathcal{P}$ where, for a fixed t, $R(t,l)$ is a mixture of s-paintbox processes directed by $\tilde{P}_t(l)$, i.e. for $A \subseteq \mathcal{P}$

$$R(t,l)(A) = \int_{\mathcal{S}^{\downarrow}} \mu_s(A) \tilde{P}_t(l, ds).$$

For $B \subseteq \mathbb{N}$ define $Q(t,B)$ the distribution of $\Pi_B \cap B$ where Π_B is a $\mathcal{P}$ valued random variable with law $R(t, \Lambda(B))$. Practically this means that one begins by drawing a variable λ_B with law $\tilde{P}_t(\Lambda(B))$, we then take the sub-partition of B induced by the λ_B-paintbox partition.

Now let $\pi = (\pi_1, \pi_2, ...) \in \mathcal{P}$ and $\forall t \geq 0$ let $(\Pi_{\pi_i}(t))_{i \in \mathbb{N}}$ be a sequence of independent variables with respective law $Q(t, \pi_i)$. Define $Q(t, \pi)$ the law of the partition whose blocks are the blocks of the $(\Pi_{\pi_i}(t), i \in \mathbb{N})$.

Our proof of Proposition 1.2 shall thus consist of showing that the family $(Q(t,\pi), \pi \in \mathcal{P}, t \geq 0)$ forms a semi-group.

Proof. From the above description it should be clear that it suffices to show

$$\forall \pi \in \mathcal{P}, Q(t+u, \pi) = \int_{\pi' \in \mathcal{P}} Q(t, \pi') Q(u, \pi)(d\pi') \tag{1.1}$$

in the obvious notation. If for any subset B of $\mathbb{N}$, we denote the set of partitions of B by $\mathcal{P}_B$, then the construction, (1.1) is equivalent to

$$\forall B \subseteq \mathbb{N}, Q(t+u, B) = \int_{\pi' \in \mathcal{P}_B} Q(t, \pi') Q(u, B)(d\pi'). \tag{1.2}$$

We can reformulate (1.2) as : $Q(t+u, B)$ is the distribution of the random partition $\Pi(t,u)$ of B (and this is what we actually shall prove) obtained by the following two-step procedure :

1. draw $\Pi(u) = (\pi_1(u), \pi_2(u), ...)$ an exchangeable partition of B with law $Q(u, B)$.
2. given $\Pi(u)$ draw a sequence $\left(\Pi_{\pi_i(u)}(t)\right)_{i \in \mathbb{N}}$ of independent $\mathcal{P}_{\pi_i(u)}$-variables with respective law $Q(t, \pi_i(u))_{i \in \mathbb{N}}$.
3. $\Pi(t,u)$ is just the collection of all the blocks of the $\Pi_{\pi_i(u)}(t)$.

We begin by proving so for $B = \mathbb{N}$. By construction we can always suppose that $\Pi(u)$ is a mixture of paintbox processes directed by $P_u((1,0,...))$, i.e. conditionally on $\lambda(u)$ (a random variable with law $P_u((1,0,...))$), $\Pi(u)$ is a $\lambda(u)$-paintbox process (resp. for each $i \in \mathbb{N}$, $\Pi_{\pi_i(u)}(t)$ is constructed by taking the intersection of $\pi_i(u)$ and a $\lambda^{(i)}$ paintbox-process where $\lambda^{(i)}$ is a $\mathcal{S}^{\downarrow}$-variable with law $\tilde{P}_t(|\pi_i(u)|)$).

This means that conditionally on $\lambda(u) = (\lambda_1(u), \lambda_2(u), ...)$, one draws an i.i.d. sequence of random variables $(X_i)_{i\in\mathbb{N}}$ with values in $\{0, 1, 2, ...\}$ such that $P(X_1 = k) = \lambda_k(u)$ for all $k \geq 1$ and $P(X_i = 0) = 1 - \sum_n \lambda_n(u)$. These random variables determine $\Pi(u)$. in the same way, for each $i \in \mathbb{N}$, conditionally on $\lambda^{(i)}$, one draws an i.i.d. sequence of random variable $(Y_k^{(i)})_{k\in\mathbb{N}}$ which determines $\Pi_{\pi_i(u)}(t)$. Note that $i \overset{\Pi(t,u)}{\sim} j$ if and only if $X_i = X_j > 0$ and $Y_i^{(X_i)} = Y_j^{(X_j)} > 0$. Since $(X_i)_{i\in\mathbb{N}}$ and $(Y_k^{(i)})_{k\in\mathbb{N}}$ are i.i.d. sequences, it follows that $\Pi(t, u)$ is exchangeable.

The law of an exchangeable random partition is completely determined by the law of its asymptotic frequencies, here the $\lambda_i(u) \times \lambda^{(i)}(t)$. As $\lambda(.)$ is a $\mathcal{S}^{\downarrow}$-fragmentation we have by construction that

$$\left((\lambda_i(u) \times \lambda^{(i)}(t))_{i\in\mathbb{N}}\right)^{\downarrow} \overset{\mathcal{L}}{=} \lambda(t + u).$$

So $\Pi(t, u)$ has law $Q(t + u, \{\mathbb{N}\})$.

Then take B a subset of $\mathbb{N}$. By construction $Q(t + u, B)$ is the law of $\tilde{\Pi}_{t+u}(B) \cap B$ where $\tilde{\Pi}_{t+u}(B)$ is a $\mathcal{P}$-variable and $\Pi(t, u) = \tilde{\Pi}(t, u) \cap B$ where $\tilde{\Pi}(t, u)$ is a $\mathcal{P}$-variable. It is clear that replacing the generating family $(R(t, l), t \geq 0, l \in [0, 1])$ by $(R'(t, l) = R(t, \Lambda(B)l), t \geq 0, l \in [0, 1])$ in the above arguments yields $\tilde{\Pi}(t, u) \overset{\mathcal{L}}{=} \tilde{\Pi}_{t+u}(B)$ and thus for all $B \subseteq \mathbb{N}$

$$\Pi(t, u) \overset{\mathcal{L}}{=} \Pi_{t+u}(B).$$

So we have proved the existence of a Markov $\mathcal{P}$-process Π with semi-group $Q(t, \pi)$, which, by construction, is a fragmentation whose asymptotic frequencies have same distribution (in the sense of finite-dimensional distributions) as λ_t our starting $\mathcal{S}^{\downarrow}$-fragmentation.

For each ranked fragmentation λ we can thus construct a partition fragmentation Π_λ such that $\Lambda(\Pi_\lambda)$ has same law as λ.

We now turn to the conservation of self-similarity : suppose λ is a self-similar $\mathcal{S}^{\downarrow}$-fragmentation with index α, so $\tilde{P}_l(t) = P_1(tl^\alpha)$, looking at the above construction of the semi-group shows that $R(t, l) = R(tl^\alpha, 1)$, so Π is also self-similar of index α. □

It is now natural to look for some explicit construction of ranked fragmentations, i.e. an equivalent of Theorem 1 in [5].

1.3 Homogeneous fragmentations

In [5] J. Bertoin shows how a homogeneous $\mathcal{P}$-fragmentation process can be decomposed into a Poisson point process of partitions, whose distribution is determined by the so-called characteristic measure. As one can always suppose that a $\mathcal{S}^{\downarrow}$-fragmentation is the image by Λ of some $\mathcal{P}$-fragmentation, one might ask if it is possible to construct an $\mathcal{S}^{\downarrow}$-fragmentation directly from some Poisson point process.

1.3.1 Lévy-Itô decomposition of homogeneous $\mathcal{P}$-fragmentations

We start by recalling some results lifted from [5]. The distribution of a homogeneous $\mathcal{P}$-fragmentation Π is determined by an exchangeable measure κ on $\mathcal{P}$, called the characteristic measure of Π, that assigns zero mass to the trivial partition and satisfies the condition $\kappa(\mathcal{P}_2^*) < \infty$ where $\mathcal{P}_2^*$ is the set of the partitions of $\mathbb{N}$ for which 1 and 2 do not belong to the same block. Given such a measure κ, one can construct a homogeneous $\mathcal{P}$-fragmentation admitting κ as its characteristic measure as follows : Let $K = ((\Delta(t), k(t)), t \geq 0)$ a Poisson point process with values in $\mathcal{P} \times \mathbb{N}$ with intensity measure $M := \kappa \otimes \#$ where $\#$ stands for the counting measure on $\mathbb{N}$. This means that for a measurable set $A \subseteq \mathcal{P} \times \mathbb{N}$ with $M(A) < \infty$, the counting process

$$N^A(t) = \mathrm{Card}(s \in [0,t] : (\Delta(s), k(s)) \in A), t \geq 0)$$

is a Poisson process with intensity $M(A)$, and to disjoint sets correspond independent processes.

Then one can construct a unique $\mathcal{P}$-valued process $\Pi_\kappa = (\Pi_\kappa(t), t \geq 0)$ started from the trivial partition, with càdlàg sample paths, such that Π_κ only jumps at time t at which K has an atom $(\Delta(t), k(t))$, and in that case $\Pi_\kappa(t)$ is the partition whose blocks are the $B_i(t-)$ (the blocks of $\Pi_\kappa(t-)$) except for $B_{k(t)}(t-)$ which is replaced by the partition of $B_{k(t)}(t-)$ induced by $\Delta(t)$ (that is $\Delta(t) \cap B_{k(t)}(t-)$).

Π_κ is a homogeneous $\mathcal{P}$-fragmentation with characteristic measure κ. Conversely, any homogeneous $\mathcal{P}$-valued fragmentation Π has the same law as Π_κ for some unique exchangeable measure κ.

As a consequence of Kingman's representation of exchangeable partitions [13], every exchangeable partition measure can be decomposed into the sum of a *dislocation* measure and an *erosion* measure :

- δ_π stands for the Dirac point mass at $\pi \in \mathcal{P}$, for all $n \in \mathbb{N}$ let ϵ_n be the partition of $\mathbb{N}$ with only two non-void blocks : $\{n\}$ and $\mathbb{N}\backslash\{n\}$, then for every $c \geq 0$, the measure
$$\mu_c = c \sum_{n=1}^{\infty} \delta_{\epsilon_n}$$
is an exchangeable measure. The μ_c's are called *erosion* measures.

- The dislocation measures are constructed from so-called Lévy measures on $\mathcal{S}^\downarrow$. We call a measure ν on $\mathcal{S}^\downarrow$ a Lévy measure if ν does not charge $(1, 0, 0, ..)$ and verifies the integral condition
$$\int_{\mathcal{S}^\downarrow} (1 - s_1)\nu(ds) < \infty$$
where $s = (s_1, s_2, ...)$ denotes a generic sequence in $\mathcal{S}^\downarrow$. The mixture of paintbox processes
$$\mu_\nu(.) = \int_{\mathcal{S}^\downarrow} \mu_s(.)\nu(ds)$$
is a measure on $\mathcal{P}$, called the *dislocation* measure directed by ν.

Then for any κ exchangeable partition measure there exists a unique $c \geq 0$ and a unique Lévy measure ν such that $\kappa = \mu_c + \mu_\nu$.

Thus the law of a homogeneous $\mathcal{P}$-fragmentation is completely characterized by the pair (ν, c). Using Proposition 1.2, we conclude that :

Corollary 1.1. *There is a bijective correspondence between the laws of homogeneous ranked fragmentations and the pairs (ν, c) where ν is a Lévy measure on $\mathcal{S}^\downarrow$ and $c \geq 0$.*

A ranked fragmentation is thus completely characterized (in terms of distribution) by the pair (ν, c) associated to its law.

We would like to transfer the Poisson point process construction of $\mathcal{P}$-fragmentations to $\mathcal{S}^\downarrow$-fragmentations. The main difficulty in doing so comes from the lack of a genealogy structure in this new setting. [3]

[3] Of course the branching character of ranked fragmentation that comes from the fragmentation property basically enforces some kind of genealogical structure. By lack of genealogy we only mean that, given $\lambda(t)$ and $\lambda(t+s)$ it is not always possible to determine for a given fragment $\lambda_i(t+s)$ from which fragment of $\lambda(t)$ it detached, whereas this is always possible for $\mathcal{P}$-fragmentations.

To illustrate this, let $K = (\Delta(t), k(t))$ a PPP with measure intensity $\mu_\nu \otimes \#$ and Π the corresponding $\mathcal{P}$ fragmentation (hence with no erosion), and suppose that at time t there is an atom (Δ, k) : the k-th block of $\Pi(t-)$ (i.e. its least element is k) fragments, or otherwise said $\Lambda(B_k(t-)) > \Lambda(B_k(t)) > 0$. Then it is clear that at time t there is also a dislocation in the associated ranked fragmentation $\lambda = \Lambda(\Pi)$. The label of the mass of $\lambda(t-)$ that fragments, noted $\Phi(t-, k)$, is an integer that depends on $\Pi(t-)$ and k and can informally be seen as the rank of the size of the k-th block of $\Pi(t-)$. In the same way that Π is constructed from K, one might hope that $\Lambda(\Pi)$ is constructed from $(\Lambda(\Delta(t)), \Phi(t-, k(t)))$ but we will still have to show that this last point process is a Poisson point process with the right intensity, then that the jump-times of λ are exactly the atom times of K and finally that λ is a pure-jump process (in a sense to be defined).

But first we show how to get rid of erosion.

1.3.2 Erosion in homogeneous ranked fragmentations

Let us first examine the trivial case when the fragmentation is pure erosion. It is then intuitively clear that the homogeneity in time and space entails that the ranked fragmentation $\lambda(t)$ with values in $S^\downarrow$ with characteristics $(0, c)$ (where the 0 means that the measure ν is trivial with mass 0) is given by

$$\lambda(t) = (e^{-ct}, 0, 0, ...).$$

To demonstrate this define

$$k = \mu_c$$

with $c > 0$, and let Π be the $\mathcal{P}$-fragmentation associated to the P.P.P. $K = (\Delta(t), k(t))_{t\geq 0}$ with intensity $\mu_c \otimes \#$ and values in $\mathcal{P} \times \mathbb{N}$. Π can be thought of as an isolation process, indeed at each jump time of K, say t, some point of $\mathbb{N}$, say n, is designated, (i.e. $\Delta(t) = \delta_{\epsilon_n}$). If the block containing n, $\beta(n, t)$, is not reduced to the singleton $\{n\}$, then it is fragmented into $\{n\}$ and $\beta(n, t)\backslash\{n\}$, "$n$ is isolated from its block", or else nothing happens. Hence, at all time there is only one block which is not a singleton, by an argument that will be established in Theorem 1.1, we can always suppose that this block also contains 1. If we consider the restriction of Π to $\{1, 2, .., n\}$, denoted by $\Pi^{(n)}$, then $\Pi^{(n)}$ only jumps at atom-times of K for which $k_t = 1$ and $\Delta_t \in \{\delta_{\epsilon_1}, \delta_{\epsilon_2}, ..., \delta_{\epsilon_n}\}$. The restriction of the Poisson process to this set is a Poisson process with intensity of finite mass and thus have discrete jump-times. The processes of the times of exclusion of each point are independent of one another. By standard calculation on Poisson processes the probability

that a given point have been excluded at time t is $\exp(-tc)$, thus the law of the number of points excluded at time t is a Bernoulli with parameter (e^{-ct}). By the law of the large numbers, at time t, the asymptotic frequency of the only block not reduced to a singleton is (e^{-ct}) almost surely. So a.s. for every $t \in \mathbb{Q}$

$$\Lambda(\Pi_t) = ((e^{-ct}), 0, 0, ...)$$

and as $\lambda_1(t)$ is monotone decreasing the relation holds almost surely for all t. This result is the key for the following.

Proposition 1.3. *If $\tilde{\lambda}$ is a homogeneous $(\nu, 0)$ ranked fragmentation, then $\lambda = (e^{-ct}\tilde{\lambda}(t), t \geq 0)$ is a homogeneous (ν, c) ranked fragmentation.*

Proof. Let $\tilde{\Pi}$ and Π be some homogeneous partition fragmentations with characteristics $(\nu, 0)$ and (ν, c) respectively. Then call $\tilde{\lambda}$ the process of the ordered asymptotic frequencies of $\tilde{\Pi}$ and λ those of Π. Suppose Π is constructed from the Poisson point process $K = (\Delta(t), k(t), t \geq 0)$ with characteristic measure $\mu_\nu + \mu_c$. Let $K_1 = (\Delta(t), k(t), t \geq 0)$ the Poisson point process with characteristic measure $\mu_\nu \otimes \#$ and $K_2 = (\Delta(t), t \geq 0)$ the Poisson point process with characteristic measure μ_c.

Thus Π appears as (i.e. is equal in law to) the intersection of Π_1 (constructed from K_1) and a pure erosion process Π_2 (constructed from K_2), i.e. $\Pi = \Pi_1(.) \cap \Pi_2(.)$ defined by the equivalence relation

$$\forall i, j \in \mathbb{N} : (i \overset{\Pi_1(.)\cap\Pi_2(.)}{\sim} j) \Leftrightarrow \left((i \overset{\Pi_1(.)}{\sim} j) \text{ and } (i \overset{\Pi_2(.)}{\sim} j)\right).$$

Given a random exchangeable subset of $\mathbb{N}$, say A, independent of $(\Pi_2(t))_{t\geq 0}$, with random asymptotic frequency l, the asymptotic frequency of the subset of A defined as the points that have not been excluded up to time t is le^{-ct} a.s. for all t.

Therefore $\Lambda(\Pi(t)) \overset{\mathcal{L}}{=} e^{-ct}\Lambda(\Pi_1(t))$. As we can always suppose that a ranked fragmentation is the associated ranked fragmentation of some partition fragmentation the result is proven. □

Thus it suffices to know how to construct a homogeneous ranked fragmentation without erosion from a PPP to know how to construct any homogeneous ranked fragmentation.

1.3.3 Construction of homogeneous ranked fragmentations with no erosion

Let λ be an $\mathcal{S}^{\downarrow}$-fragmentation, with characteristics (ν, c), then for every $k \in \mathbb{N}$ the process $\lambda_1(t) + ... + \lambda_k(t)$ is monotone decreasing. λ is said to be a pure jump process if for any k, $\lambda_1(t) + ... + \lambda_k(t)$ is a pure jump process.

In the following we shall focus on the case where for each fixed t there is a infinite number of fragments of strictly positive size almost surely. A necessary and sufficient condition for this is

$$\nu(s \in \mathcal{S}^{\downarrow} : s_2 > 0) = \infty.$$

Indeed, fix $t > 0$ and suppose that $\lambda_1(t) > 0$. Then, for any $\epsilon > 0$, during the time interval $[t - \epsilon, t]$, λ_1 has been affected by an infinite number of dislocations such that at least one small fragment detached from the main one, thus an infinite number of fragments have been created, and the lifetime of those variables form a sequence of independent identically distributed random variables, thus with probability one an infinite number of them have survived at time t. The same line of arguments also shows that $\inf\{t \geq 0 : \lambda_1(t) = 0\} = \infty$ almost surely.

If

$$\nu(s \in \mathcal{S}^{\downarrow} : s_2 > 0) < \infty,$$

the dislocation times for a tagged fragment are almost surely discrete (in particular almost surely there is a strictly positive random time before which there is only one fragment). Nevertheless, it is possible that at a positive time an infinite number of nonzero components are present with positive probability (a single dislocation could generate a infinite number of fragments).

Although the following theorem is still true for any Lévy measure ν, making this hypothesis enables us to focus on the most interesting case and to avoid some technical difficulties.

Theorem 1.1. *Let λ be a homogeneous $\mathcal{S}^{\downarrow}$-fragmentation with no erosion ($c = 0$) and Lévy measure ν as above (i.e. $\nu(\{s : s_2 > 0\} = \infty)$. Then*

1. *λ is a pure jump process.*
2. *there exists a PPP $K = (S(t), k(t))_{t \geq 0}$ with values in $\mathcal{S}^{\downarrow} \times \mathbb{N}$ and intensity measure $\nu \otimes \#$, such that the jumps of λ correspond to the atoms of K. More precisely, λ only jumps at times at which $(S(t), k(t))$ has an atom, and at such a time $\lambda(t)$ is obtained from $\lambda(t-)$ by dislocating the $k(t)$-th component of $\lambda(t-)$ by $S(t)$ (i.e. replacing $\lambda_{k(t)}(t-)$ by the*

sequence $\lambda_{k(t)}(t-)S(t)$) and reordering the new sequence of fragments. Conversely if $(S(t), k(t))$ is an atom then λ has a jump at t, i.e. λ_i jumps at t for some i.

Although this result is intuitive in regard to the equivalence between $\mathcal{P}$ and $\mathcal{S}^{\downarrow}$ fragmentation, it requires some technical work.

We give ourselves a homogeneous $(\nu, 0)$ $\mathcal{S}^{\downarrow}$-fragmentation λ with ν satisfying $\nu(s_2 > 0) = \infty$. There is no loss of generality in supposing that λ is constructed as follows : Call $K = ((\Delta(t), k(t)))_{t\geq 0}$ a PPP with intensity measure $\mu_\nu \otimes \#$ with values in $\mathcal{P} \times \mathbb{N}$. Let Π be the homogeneous $(\nu, 0)$ $\mathcal{P}$-fragmentation constructed from K, then define

$$\lambda = \Lambda(\Pi).$$

Call $\mathcal{F}_t = \sigma\{\Pi_s, s \leq t\}$ the natural filtration of the $\mathcal{P}$-fragmentation Π.

Then at any time t, call $\phi(t, .) = \phi_t(.)$ the random, $\mathcal{F}_t$ measurable function from $\mathbb{N} \to \mathbb{N} \cup \infty$ (where ∞ serves as a cemetery point) defined as

- if $|B_k(t)| > 0$ then $\phi(t, k)$ is the rank of the asymptotic frequency of $B_k(t)$ (it is well defined because the number of blocks of greater asymptotic frequencies is always finite with an upper bound of $|B_k(t)|^{-1}$, and in case two blocks have the same asymptotic frequency, they are ranked as their least element).
- if $|B_k(t)| = 0$ (with the convention $|\emptyset| = 0$) then $\phi(t, k) = \infty$

We also define $\tilde{k}(t) = \phi(t-, k(t))$. Note that under our hypothesis that there is always an infinite number of fragments $\forall t \geq 0, \mathbb{N} \subset \{\phi(t, k), k \in \mathbb{N}\}$.

We will first prove that the point process image of K, denoted $\tilde{K}$, whose atoms are the points of $(\Lambda(\Delta(t)), \tilde{k}(t))_{t\geq 0}$ such that $\tilde{k}(t) \in \mathbb{N}$, is a Poisson point process with intensity measure $\nu \otimes \#$. Then we will show that this is also the process of the jumps of $\Lambda(\Pi)$ and this last process is a pure jump process so it can wholly be recovered from $(\Lambda(\Delta(t)), \tilde{k}(t))_{t\geq 0}$. This will complete the proof of Theorem 1.1.

Lemma 1.1. *The point process $\tilde{K}(t)$ derived from $(\Lambda(\Delta(t)), \tilde{k}(t))_{t\geq 0}$ by only keeping the atoms such that $\tilde{k}(t) \neq \infty$ is a Poisson point process with intensity measure $\nu \otimes \#$.*

Proof. Let A be a subset of $\mathcal{S}^{\downarrow}$ such that $\nu(A) < \infty$. For $i = 1, ...$ let

$$N_A^{(i)}(t) = \#\{u \leq t : \Lambda(\Delta(u)) \in A, k(u) = i\}.$$

Then set

$$N_A(t) = \#\{u \leq t : \Lambda(\Delta(u)) \in A, \tilde{k}(u) = 1\}.$$

$N_A(t)$ is increasing, right-continuous with left-limits with jumps of size 1 (the $N_A^{(i)}(t)$ being independent Poisson processes they do not jump at the same time almost surely). By definition we have

$$dN_A(t) = \sum_{i=1}^{\infty} \mathbf{1}_{\{\phi(t-,i)=1\}} dN_A^{(i)}(t).$$

Define

$$d\tilde{N}_A^{(i)}(t) = \mathbf{1}_{\{\phi(t-,i)=1\}} dN_A^{(i)}(t).$$

It is clear that $\mathbf{1}_{\{\phi(t-,i)=1\}}$ is adapted and left-continuous in $(\mathcal{F}_t)$ and hence predictable. The $N_A^{(i)}(.)$ are i.i.d. Poisson processes with intensity $\nu(A)$ in $(\mathcal{F}_t)$. Thus, for each i the process

$$M_A^{(i)}(t) = \tilde{N}_A^{(i)}(t) - \nu(A) \int_0^t \mathbf{1}_{\{\phi(u-,i)=1\}} du = \int_0^t \mathbf{1}_{\{\phi(u-,i)=1\}} d(N_A^{(i)}(u) - \nu(A)u)$$

is a square integrable martingale. Then define

$$M_A(t) = \sum_{i=1}^{\infty} \int_0^t \mathbf{1}_{\{\phi(u-,i)=1\}} d(N_A^{(i)}(u) - \nu(A)u)$$

Define $f_i(t) = \mathbf{1}_{\{\phi(t-,i)=1\}}$, then, for all $i \neq j, \forall t \geq 0$, $f_i(t) f_j(t) = 0$, and $\forall t$, $\sum_{i=1}^{\infty} f_i(t) = 1$.

As the $N_A^{(i)}(t)$ are independent Poisson processes they do not jump simultaneously and so the martingales $M_A^{(i)}(t)$ do not either. They are thus orthogonal (see for example chapter 8, Theorem (43)-D in [**?**] for a proof). Moreover the oblique bracket of M is

$$\begin{aligned} \langle M_A \rangle(t) &= \sum_{i=1}^{\infty} < \int_0^t f_i(u) d(N_A^{(i)}(u) - \nu(A)u) > \\ &= \nu(A)t. \end{aligned}$$

So M_A is a L_2 martingale.

So we have demonstrated that $N_A(t)$ is increasing, right-continuous, left limited with jumps of size 1 with compensator $\nu(A)t$. Using classical results

(see for instance chapter 2.6 in [**?**], Theorem 6.2) we conclude that $N_A(t)$ is a Poisson process with intensity $\nu(A)$. Now take $B \in \mathcal{S}^{\downarrow}$ such that $A \cap B = \emptyset$, we can use the same construction as above replacing A with B and the fact that $N_A^{(i)}(t)$ and $N_B^{(i)}(t)$ are independent Poisson processes in the same filtration to see that

$$N_1(t) = \{\Lambda(\Delta(u)) : u \leq t, \tilde{k}(u) = 1\}$$

is a P.P.P. with intensity measure ν. The same arguments yield that

$$N_2(t) = \{\Lambda(\Delta(u)) : u \leq t, \tilde{k}(u) = 2\}$$

is also a P.P.P. with intensity measure ν. It is clear that N_1 and N_2 have no jumps in common because the $N_A^{(i)}(t)$'s do not, so they are independent. By iteration this shows that $(\Lambda(\Delta(t)), \tilde{k}(t))$ is a P.P.P. with intensity measure $\nu \otimes \#$. □

Introduce the blocks $\Pi = (\Pi(t), t \geq 0) = ((B_1(t), B_2(t), ...), t \geq 0)$ where Π is the $(\nu, 0)$ $\mathcal{P}$-fragmentation constructed from K, and define $\lambda = \Lambda(\Pi) = (\lambda_1(t), \lambda_2(t), ..)$ the ordered vector of asymptotic frequencies. In the case considered here Π is nice so almost surely for all t $|B_i(t)|$ exists for all $i \in \mathbb{N}$. Recall that $\phi(t, k)$ is the rank of the asymptotic frequency $|B_k(t)|$ at time t.

We now need to show that λ is a pure jump process in the sense that for each k the decreasing process $\lambda_1 + ... + \lambda_k$ is pure jump and that all its jumps are indeed images of some atoms of K (Λ being not continuous it is not *a priori* evident).

In [5] it is shown that $|B_1(t)|$, the asymptotic frequency of the block that contains $\{1\}$, is the inverse of the exponential of a subordinator with 0-drift, and so it is a pure-jump process. By the Markov and homogeneity property this implies that for all $i > 1$ the process $|B_i(t)|$, the asymptotic frequency of the block whose least element is i, is càdlàg, started at 0, such that at $\tau_i = \inf\{t \geq 0 : |B_i(t)| > 0\}$ we have that conditionally on $\tau_i < \infty$, $|B_i(\tau_i)| > 0$ (i.e. it leaves 0 with a jump), and after τ_i the process $\frac{|B_i(t-\tau_i)|}{|B_i(\tau_i)|}$ is the inverse of the exponential of a subordinator with no drift, in particular it is a pure jump process. Furthermore it is clear by construction that all the jumps of $B_i(.)$ correspond to some atom of $\tilde{K}$.

For each t define $\psi_t(.)$ the function from $\mathbb{N} \to \mathbb{N}$ inverse of $\phi(t, .)$, i.e.

$$\psi_t(\phi(t, i)) = i$$

(exists because ϕ is surjective on $\mathbb{N}$).

Lemma 1.2.
- *for all $k > 0$, $\lambda_1(t) + \lambda_2(t) + ... + \lambda_k(t)$ is a pure jump process.*
- *with probability one, for all $t \geq 0$, if λ jumps at t, then $\tilde{K}$ has an atom at t.*

Proof. We will begin by proving the result for λ_1, the size of the largest fragment and then turn our attention to the small ones.

λ_1 is a supremum of a countable family of pure jump processes (the $|B_i(.)|$). However it is easy to exhibit an example of a supremum of a countable family of pure jump processes that is not a pure jump process. So the proof will consist in showing that almost surely on a fixed time interval λ_1 is the supremum of a finite number of pure jump processes.

The key point is to show that given $t > 0$ and an arbitrary $\epsilon > 0$ we can almost surely choose a M such that

$$\forall s \leq t \text{ and } n > M, |B_n(s)| < \epsilon.$$

In particular this would mean that we only need to consider a finite number of $B_i(.)$ to be sure to "catch" a fragment whose size is greater than ϵ.

To do this , it is convenient to work with so-called *interval fragmentations*.

Interval fragmentations are a particular case of the *object fragmentations* that we presented in the introduction for which the "object" E is simply the interval $[0, 1]$ endowed with the Lebesgue measure. More precisely, call ν the space of the open subsets of $[0, 1]$. Elements of ν admit a unique decomposition in intervals (in the sense that the ordered vector of the lengths is unique). An interval decomposition is a process $F(t)$ with values in ν such that for any $0 \leq s < t$ one has $F(s) \subseteq F(t)$ i.e. $F(t)$ is finer than $F(s)$.

Take a sequence $(u_i)_{i\in\mathbb{N}}$ of i.i.d. variables uniformly distributed on $[0, 1]$. F is then transformed into a $\mathcal{P}$-process Π by the following rule

$$i \overset{\Pi(t)}{\sim} j \Leftrightarrow [u_i, u_j] \subseteq F(t).$$

This last process obviously conserves the refinement property. Moreover, if we define the interval fragmentations to have the scaling and branching properties, then Π will be a self-similar $\mathcal{P}$-fragmentation.

We refer to [6] for a precise definition of interval fragmentation and the equivalence between interval fragmentations and partition fragmentations.

There is no loss of generality in supposing that Π is constructed from an interval fragmentation $F(t)$ and a sequence $(u_i)_{i\in\mathbb{N}}$ of iid variables uniformly distributed on $[0, 1]$.

Denote by $(I_i(t), i \in \mathbb{N})$ the associated ordered lengths of the interval decomposition of F (which are also the associated ordered frequencies of $\Pi(t)$). If $I^{(i)}(t)$ denotes the length of the interval that contains u_i in the interval decomposition of $F(t)$, then

$$I^{(i)}(t) = l_i(t)$$

where $l_i(t) = |\beta(i,t)|$ is the asymptotic frequency of the block of $\Pi(t)$ that contains i.

Calling τ_n the stopping time $\inf\{t > 0, |B_n(t)| > 0\}$ we have that at τ_n

$$\forall i < n \ , \ n \overset{\Pi(\tau_n)}{\not\sim} i.$$

Thus u_n does not belong to any block of $F(\tau_n)$ that contains some u_i for any $i < n$, hence the asymptotic frequency of the block of $F(\tau_n)$ that contains u_n is bounded from above by $\sup_{i,j\in\{1,..,n\}} |u_i - u_j|$ which converges to 0 almost surely when $n \to \infty$.

Note that

$$\sup_{r\in\mathbb{R}^+} \{|B_n(r)|\} = |B_n(\tau_n)|$$

to see that

$$\lim_{n\to\infty} \left(\sup_{r>0} (|B_n(r)|) \right) = 0 \text{ a.s.}$$

Now fix $\epsilon > 0$ and n_0 and condition on the events $\{\lambda_1(T) \geq \epsilon\}$, and

$$\{\sup_{n>n_0} \{\sup_{r>0} (|B_n(r)|)\}\} < \epsilon\}.$$

Note that the probability of the second event can be taken arbitrarily close to 1 by taking n_0 sufficiently large. On this event, for all $r \in [0,T]$ we have that

$$\lambda_1(r) = \max_{i=1,\ldots,n_0} |B_i(r)|.$$

Thus $\lambda_1(.)$ is a pure jump process because all the $|B_i(.)|$ are. Moreover $\lambda_1(.)$ only jumps at times at which $\tilde{K}$ has an atom for the same reason.

Turning our attention to the other fragments, it should be clear that the above discussion also entails that for an arbitrary $\epsilon > 0$, for all $s \in [0,t]$ the process $\lambda_k(s) \bigwedge \epsilon$ can almost surely be expressed in terms of a finite number of pure jump processes (the $|B_i(s)|$ for $i = 1,..,M$) and the $\lambda_j(s)$ for $j = 1,..,k-1$. Thus inductively, the result is proven. The second point is immediate from this proof.

□

In conclusion, if we call Γ the set of times at which $(\Delta(t), k(t))$ has an atom, then writing $\lambda(t) = (\lambda_1(t), \lambda_2(t), ...)$ for $\Lambda(\Pi(t))$:

1. $\lambda(.)$ is a pure jump process, càdlàg and starts almost surely from $(1, 0, 0, ...)$
2. if $t \notin \Gamma$,
$$\lambda(t) = \lambda(t-)$$
3. if t is a jump-time for λ, then almost surely $t \in \Gamma$ and $\lambda(t)$ is the reordering of the concatenation of two sequences : $(\lambda_i(t-))_{\{i \neq k(t)\}}$ and $\lambda_{\phi_t(k(t))}(t-)\Lambda(\Delta(t))$.

As λ is a pure jump process it is completely defined by this description.

All we have to do now is collect the preceding results : let $K = (\Delta(t), k(t))$ be a Poisson point process with intensity measure $\mu_\nu \otimes \#$ and let Π the associated $(\nu, 0)$ homogeneous $\mathcal{P}$-fragmentation. Then the Poisson point process $(\Lambda(\Delta(t)), \phi_{t-}(k(t)))$ and the asymptotic frequency process $\Lambda(\Pi(t))$ have the desired properties, so Theorem 1.1 is proved.

1.4 Small time Asymptotic behavior

In [1] Aldous and Pitman consider a specific fragmentation associated with the standard additive coalescent. After renormalizing the vector $(1 - x_1(t), x_2(t), ...)$ by $\frac{1}{t^2}$ they are able to give its asymptotic distribution when $t \to 0$ in terms of a stable subordinator of index $1/2$. In [6] Bertoin showed that the characteristics of this fragmentation are $c = 0$, $\alpha = 1/2$ and the Lévy measure, whose support is included in $\{s \in \mathcal{S}^\downarrow : s_3 = 0\}$, is specified by

$$\nu(s_1 \in dx) = (2\pi x^3(1-x)^3)^{-1/2}dx, \qquad x \in [1/2, 1[\,.$$

In this section we intend to extend this type of results to all ranked fragmentations such that $\alpha \geq 0$. We shall use the Poisson construction of ranked fragmentations we just established to study their asymptotic behavior near 0. The results we give are very close in spirit to those concerning the asymptotic behavior of subordinators.

A subordinator, say ξ, is an increasing Lévy process whose distribution is specified by its Laplace exponent Ψ that is given by the identity

$$\mathbb{E}(\exp\{-q\xi_t\}) = \exp\{-t\Psi(q)\}.$$

The Lévy-Khintchine formula reads

$$\Psi(q) = k + dq + \int_{]0,\infty[} (1 - e^{-qx})\upsilon(dx)$$

where $k \geq 0$ is the so-called killing rate, $d \geq 0$ is the drift coefficient and υ a measure on $]0, \infty[$ with $\int (1 \wedge x)\upsilon(dx) < \infty$, called the Lévy measure of ξ.

The asymptotic behavior of these processes is well known, for instance we have results concerning their distribution :

$$\frac{1}{t} P(\xi(t) \in .) \underset{t \to 0+}{\to} \upsilon(.)$$

(see Corollary 8.9 in [16]).

On the other hand, under conditions of regular variation on the tail of υ, there are also results concerning the sample path behavior of the limsup and the law of the iterated logarithm (see for instance the end of chapter III in [4]).

Thus to study the asymptotic behavior of a fragmentation we may benefit from the fact that $|B_1|$ (the mass of the block that contains 1) can be described in terms of a subordinator (see [5]).

We focus on the behavior of the largest (λ_1) and of the second block (λ_2) of a ranked fragmentation even though we have a more general result in the case of so-called *binary* fragmentations.

Although the study of λ_1 is relatively straightforward, that of λ_2 requires to use some results on record-processes theory. Most of those that will be used in this section are well known or are adapted from standard facts that can be found in textbooks, see [8] for instance.

We stress that $\lambda_2(t)$ is not monotone. More precisely it decreases when the second largest fragment undergoes a dislocation and can increase when the largest fragment undergoes a dislocation and one of the new fragments created becomes the second largest.

The idea is to use the Poisson construction : near 0 the largest fragment is almost of size 1, thus the second largest fragment is always a "direct son" of the main fragment, and we shall be able to express its law in terms of the distribution of the largest fragment that has detached from the main.

For a general $\mathbb{R}$-valued P.P.P. $K = (K_t, t \geq 0)$ with intensity measure μ such that $\forall \epsilon > 0, \mu(]\epsilon, \infty]) < \infty$, it is possible to define the associated record process $R(t)$ as follows : at time t

$$R(t) = \max_{s \leq t} \{K_s\}.$$

Let λ be a homogeneous $\mathcal{S}^{\downarrow}$ fragmentation with characteristic (ν, c) constructed from the P.P.P.

$$K = (S(t), k(t))_{t\geq 0} = ((s_1(t), s_2(t), ...), k(t))_{t\geq 0}$$

of intensity measure $\nu \otimes \#$. Let $(S^{(i)}(t), t \geq 0) = (s_j^{(i)}(t), j = 1, 2, ...; t \geq 0)$ be the P.P.P. with values in $\mathcal{S}^{\downarrow}$ derived from K by keeping the points such that $k(t) = i$ (the second coordinate being always i, it is not expressed). So $s_j^{(i)}(t)$ is the relative size of the j^{th} block of the dislocation occurring at time t on the i^{th} block. $S^{(i)}$ is a P.P.P. with intensity measure ν. The $\mathbb{R}$-valued point process $(s_j^{(i)}(t))$ is thus a P.P.P. with intensity

$$\nu_j(dx) = \nu(\{s = (s_1, s_2, ...) \in \mathcal{S}^{\downarrow} : s_j \in dx\}).$$

Introduce the function

$$x \to \overline{\nu}_2(x) = \nu(s \in \mathcal{S}^{\downarrow} : s_2 \geq x)$$

from $\left]0, \frac{1}{2}\right] \to \mathbb{R}^+$, and denote by f its generalized inverse (i.e. $f(x) = \inf\{t : \overline{\nu}_2(t) \geq x\}$).

Note that $\overline{\nu}_2(.)$ is finite, i.e. for all $x > 0$ $\overline{\nu}_2(x) < \infty$. To see this, let $b \in [0, 1/2]$

$$\begin{aligned}
\int_{\mathcal{S}^{\downarrow}} (1 - s_1)\nu(ds) &\geq \int_{\mathcal{S}^{\downarrow}} s_2\nu(ds) \\
&= \int_0^{1/2} x\nu_2(dx) \\
&\geq \int_b^{1/2} x\nu_2(dx) \\
&> b\overline{\nu}_2(b)
\end{aligned}$$

Let $R(t)$ designate the record at time t of the P.P.P. $s_2^{(1)}(.)$ which is well defined according to the above argument.

Proposition 1.4. *Let*

$$\lambda = (\lambda(t), t \geq 0) = (\lambda_1(t), \lambda_2(t), ...), t \geq 0)$$

be a homogeneous $\mathcal{S}^{\downarrow}$ fragmentation with characteristic (ν, c), then

1. *there exists a subordinator ξ with drift c and Lévy measure*

$$L(dx) = \nu(-\log s_1 \in dx) \ , \ x \in]0, \infty[$$

such that

$$\lambda_1(t) = \exp(-\xi(t))$$

for t small enough a.s.

2.

$$\lambda_2(t) \sim R(t), \qquad t \to 0+ \qquad a.s.$$

Proof. (Proposition 1.4-(1)) Suppose first that $c = 0$. The general case follows easily.

Let us consider the P.P.P. $(s_1^{(1)}(t), t \geq 0)$. Its atoms are the relative sizes of the largest fragment created each time that λ_1 splits. Its intensity measure is given by $\nu_1(dx) = \nu(s : s_1 \in dx)$. Introduce the process

$$\xi(t) = -\log(\prod_{u \leq t} s_1^{(1)}(u)) \ , \ t \geq 0.$$

It is a subordinator with 0 drift and Lévy measure $\nu(-\log(s_1) \in dx)$.

Conditionally on $\lambda_1(T) > 1/2$, one has $\lambda_1(t) = e^{-\xi(t)}$ for all $t < T$. Thus $\lambda_1(t) = e^{-\xi(t)}$ for all $t \leq T_{1/2}$ where $T_{1/2} = \inf\{t : \lambda_1(t) \leq 1/2\}$. □

For the second point the idea is to describe the asymptotic behavior of λ_2 in terms of the records of $s_2^{(1)}$. We begin with the following technical lemma

Lemma 1.3. *Let*

$$\chi_t = \left(\prod_{u \in [0,t[} s_1^{(1)}(u) \right) \left(\prod_{u \in [0,t[} s_1^{(2)}(u) \right),$$

and suppose $c = 0$ (there is no erosion). Then on the event $\{\lambda_1(t) \geq 1/2\}$

$$\chi_t R(t) \leq \lambda_2(t) \leq R(t).$$

Proof. As noted earlier, one can suppose that $\lambda(.)$ is the asymptotic frequency of some $(\nu, 0)$ $\mathcal{P}$-fragmentation Π, and K is the image of the P.P.P.

$$(\Delta(.), k(.)) \to (\Lambda(\Delta(.)), \phi(., k(.)))$$

with intensity measure $(\mu_\nu \otimes \#)$. At time t we recall the notation $\psi_t(1)$ for the least element of the block of greatest asymptotic frequency in Π, which is well defined.

Define

$$f_i(u) = \mathbf{1}_{\{i \overset{\Pi(u)}{\sim} \psi_t(1)\}}$$

i.e. $f_i(u)$ is 1 if at time u the integer i is in the same block that the least element of the largest block at time t and 0 otherwise. $f_i(0) = 1$ a.s. Thus, almost surely

$$D_i(t) = \inf\{u \in [0,t] : f_i(u) = 0\} \wedge t > 0.$$

Note that as we are on $\{\lambda_1(t) \geq \frac{1}{2}\}$, $\psi_t(1)$ is always in the block of greatest asymptotic frequency of $\Pi(u)$ for any $u \in [0,t]$; so $D_i(t)$ is the *detachment time* of i from the main block (if i is still in the main block, $D_i(t)$ is taken equal to t).

Now take $k \geq 2$, and suppose that at time t there are at least k blocks (almost surely the case under our hypothesis) so $\psi_t(k)$ (the least element of the block of k-th greatest asymptotic frequency in Π at time t) is well defined, almost surely

$$D_{\psi_t(k)}(t) > 0.$$

so if we define $\beta(i,u)$ for the block of $\Pi(u)$ that contains i and $D(k,t) = D_{\psi_t(k)}(t)$ we have that

$$\begin{array}{rcl} \beta((\psi_t(k)), D(k,t-)) & = & \beta(\psi_{D(k,t-)}(1), D(k,t)-) \\ |\beta\left((\psi_t(k)\right), D(k,t))\,| & < & |\beta(\psi_{D(k,t)}(1), D(k,t))| \end{array}$$

(recall that $\beta(\psi_t(1),t)$ is the largest block at time t). Thus

$$\lambda_k(t) \leq |\beta\left((\psi_t(k)\right), D(k,t))\,| \leq \lambda_1(D(k,t-)) s_2^{(1)}(D(k,t)) \leq s_2^{(1)}(D_{\psi_t(k)}(t)).$$

As obviously

$$s_2^{(1)}(D_{\psi_t(k)}(t)) \leq R(t)$$

we conclude that

$$\lambda_k(t) \leq R(t). \tag{1.3}$$

We now prove the lower-bound part of the lemma.

Let $T(t) = \inf\{u \leq t : R(u) = R(t)\}$ (the "record-time"). Note that for all $u \in [0,t]$ at which $S^{(2)}$ has an atom,

$$\lambda_2(u-)s_1^{(2)}(u) \leq \lambda_2(u).$$

This is not an equality because the largest fragment resulting from the dislocation of $\lambda_2(u-)$ can be smaller than $\lambda_3(u-)$. Then, for all $u \in [0,t]$ not an atom for $S^{(2)}$,

$$\lambda_2(u-) \leq \lambda_2(u).$$

This is due to the fact that u could be an atom for $S^{(1)}$, for which $\lambda_1(u-)s_2^{(1)}(u) > \lambda_2(u-)$. Recalling that we are still conditioning on $\{\lambda_1(t) > \frac{1}{2}\}$ we have, using the fact that λ_2 is a pure jump process, that

$$\lambda_2(T(t))\Big(\prod_{u\in[T(t),t[} s_1^{(2)}(u)\Big) \leq \lambda_2(t) \tag{1.4}$$

and here again this is not an equality because a reordering might occur.

Then remark

$$\lambda_2(T(t)) \geq R(t)\Big(\prod_{u\in[0,T(t)[} s_1^{(1)}(u)\Big). \tag{1.5}$$

In words : at the time of the record $R(t)$, the second fragment resulting from the dislocation of λ_1, is not necessarily λ_2, but in any case it is smaller or equal.

We can combine (1.4) and (1.5) to get

$$\chi_t R(t) \leq \lambda_2(t) \tag{1.6}$$

□

We can now prove the second part of proposition 1.4 :

Proof. Proposition 1.4-(2) When $c = 0$ we now only have to show that $\chi_t \underset{t\searrow 0}{\to} 1$ almost surely. $\Big(\prod_{u\in[0,t[}(s_1^{(1)}(u))\Big)$ and $\Big(\prod_{u\in[0,t[}(s_1^{(2)}(u))\Big)$ are independent and identically distributed, and on the event $\Big(\prod_{u\in[0,t[}(s_1^{(1)}(u))\Big) \geq \frac{1}{2}$ this last quantity is exactly the $\lambda_1(t)$ of some $(\nu, 0)$ fragmentation, thus almost surely

$$\Big(\prod_{u\in[0,t[}(s_1^{(1)}(u))\Big) \underset{t\searrow 0}{\to} 1$$

which thus concludes our proof.

Finally in the case of a homogeneous (ν, c) fragmentation λ with $c \geq 0$, the effect of the erosion is just of multiplying the size of each fragment by a factor e^{-ct}. So clearly the upper bound of Lemma 1.3 is still valid, on the other hand we have

$$\xi(t)e^{-ct}R(t) \leq \lambda_2(t)$$

and only a slight modification of the proof for the case $c = 0$ is needed. □

Remark 1.1. In the case where the fragmentations considered are not homogeneous but only self-similar and without erosion, a more technical version of Theorem 1.1 should still hold. One should be able to give a Poisson point process construction of a partition-valued $(\alpha, \nu, 0)$ fragmentation, by using a thinning of the Poisson point process when $\alpha > 0$ to decrease the rate of dislocation, and by adding atoms in the case $\alpha < 0$ to increase this rate. Most of the technical work developed in section 3.3 is adaptable to the self-similar setting so one should be able to "pass" this construction from the partition to the $\mathcal{S}^{\downarrow}$ setting.

However, the results in this section on the small-time asymptotic results of homogeneous fragmentations can be extended to self-similar fragmentations of index α only when $\alpha \geq 0$. Lemma 1.3 does not holds when $\alpha < 0$, furthermore the convergence of χ_t is not true in this case (informally the reason is that too many dislocations of the second largest fragment accumulate).

When $\alpha > 0$ the record like behavior of the small fragments tends to be accentuated because they are "frozen" and do not split a lot once they are born. On the other hand, when $\alpha < 0$ small fragments vanish quickly. In [7] Bertoin shows that when $\alpha < -1$, although dislocations occur continuously and that at each splitting an infinite number of fragments are created, at any fixed time t there are almost surely a finite number of fragments of positive mass. Also in that case it holds with probability 1 that $\zeta := \inf\{t \geq 0 : X(t) = (0, 0, ...)\} < \infty$.

Remark 1.2. In the homogeneous case, if $\overline{\nu}_2(.)$ is regularly varying with index $(-a)$ in 0^+, $a \geq 0$, classical results of record-processes theory used with proposition 1.4 show that

$$\frac{\lambda_2(t)}{f(\frac{1}{t})} \xrightarrow{\mathcal{L}} L$$

when $t \searrow 0$ where L is the extreme law with distribution function $\exp(-x^{-a})$ and f is the generalized inverse of $x \to \bar{\nu}_2(x)$.

Remark 1.3. Suppose that λ is a binary fragmentation, that is ν has its support in the subset of $\mathcal{S}^{\downarrow}$ defined as $\{s \in \mathcal{S}^{\downarrow}, s_3 = s_4 = ... = 0\}$ and that $\overline{\nu}_2(x) = \nu(\{s \in \mathcal{S}^{\downarrow} : s_2 \geq x\})$ is regularly varying near 0 with index $-a$. Then using the same ideas as in the above arguments one can show that, if $R_2(k,t)$ denotes the k^{th} record of $s_2^{(1)}(.)$ up to time t, then we have the following asymptotic distributions of the renormalized λ_k for any $k > 1$:

$$\forall k > 1, a.s. \qquad \lambda_k(t) \underset{t \searrow 0^+}{\sim} R_2(k,t).$$

As a consequence

$$\frac{\lambda_k(t)}{f(1/t)} \xrightarrow{\mathcal{L}} L(k,a)$$

where $L(k,a)$ is the law with distribution function

$$F_{k,a}(x) = \left(\sum_{i=0,...,k-1} e^{-x^{-a}} \frac{(x^{-ai})}{i!} \right).$$

More generally, the convergence in law holds jointly. The limit distribution function for the N largest blocks being given by

$$f_N(x_2, x_3, ...x_N) = \left(\prod_{i=2}^{i=N-1} \left(\exp -x_i^{-a}\right) \right) \int_0^{x_N} (\exp -u^{-a}) \nu(s_2 \in du)$$

for $x_1 > x_2 > ... > x_N$ (see [14] for instance).

Acknowledgement I would like to thank two anonymous referees for their very careful reading and insightful comments.

BIBLIOGRAPHY

[1] D. Aldous and J. Pitman. The standard additive coalescent. *Ann. Probab.*, 26:1703–1726, 1998.

[2] D. J. Aldous. Exchangeability and related topics. In *École d'été de probabilités de Saint-Flour, XIII—1983*, volume 1117 of *Lecture Notes in Math.*, pages 1–198. Springer, Berlin, 1985.

[3] D. J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5(1):3–48, 1999.

[4] J. Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

[5] J. Bertoin. Homogeneous fragmentation processes. *Probab. Theory Related Fields*, 121(3):301–318, 2001.

[6] J. Bertoin. Self-similar fragmentations. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(3):319–340, 2002.

[7] J. Bertoin. The asymptotic behaviour of fragmentation processes. *J. Euro. Math. Soc.*, to appear, 2003.

[8] N. Bingham, C. Goldie, and J. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1987.

[9] E. Bolthausen and A.-S. Sznitman. On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.*, 197(2):247–276, 1998.

[10] M. D. Brennan and R. Durrett. Splitting intervals. *Ann. Probab.*, 14(3):1024–1036, 1986.

[11] M. D. Brennan and R. Durrett. Splitting intervals ii. limit laws for lengths. *Probab. Theory Related Fields*, 75(1):109–127, 1987.

[12] S. N. Evans and J. Pitman. Construction of Markovian coalescents. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(3):339–383, 1998.

[13] J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13(3):235–248, 1982.

[14] M. Perman. Order statistics for jumps of normalised subordinators. *Stochastic Process. Appl.*, 46(2):267–281, 1993.

[15] J. Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27(4):1870–1902, 1999.

[16] K. Sato. *Lévy Processes and Infinitly Divisible Distributions*, volume 68 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 1999.

2. MULTIFRACTAL SPECTRA OF FRAGMENTATION PROCESSES

Abstract

Let $(S(t), t \geq 0)$ be a homogeneous fragmentation of $]0,1[$ with no loss of mass. For $x \in]0,1[$, we say that the fragmentation speed of x is v if and only if, as time passes, the size of the fragment that contains x decays exponentially with rate v. We show that there is $v_{\text{typ}} > 0$ such that almost every point $x \in]0,1[$ has speed v_{typ} and that, nonetheless, for v in a certain range, the random set $\mathcal{G}_v$ of points of speed v, is dense in $]0,1[$, and we compute explicitly the spectrum $v \to \text{Dim}(\mathcal{G}_v)$ where Dim is the Hausdorff dimension.

2.1 Introduction

Fragmentation underlies a number of physical, chemical and geological processes, such as polymer degradation (see [24, 1]), liquid droplet breakup ([31]), crushing or grinding of rocks ([15]), atomic collisions and nuclear multifragmentation ([8]) or energy cascade in turbulence to name just a few. One can also report to the proceedings [6] for some applications in physics and to [10, 20, 19] for computer science. The fragmented quantity in such processes are diverse : mass, momentum, energy or surface. But a global characteristic feature is that each new fragment continues splitting independently. Usually the simplifying assumption that each fragment can be described by a single state variable (e.g. mass) is also made.

Informally, the purpose of this work is to investigate the set of locations having an abnormally fast (or slow) fragmentation speed in a so-called homogenous fragmentation of a one dimensional object.

2.1.1 An example

Let us first introduce *via* a simple example the ideas and notions on which this paper focuses. A fragmentation model describes an object endowed with a unit mass measure that falls apart as time runs. We consider the following example which is a continuous version of a model due to Kolmogoroff in what seems to be one of the first probabilistic work on random fragmentations (see [18]). Define $S(t)$ to be a Markov process with values in the space $\mathcal{O}$ of the open subsets of $]0,1[$, which starts from $S(0) =]0,1[$ and evolves as follows. Each segment of length m is cut in two with rate 1 (i.e after an exponential time with mean 1). It then gives rise to two new segments (which are thus included in their "father") of respective length Vm and $(1-V)m$, where V is a random variable with values in $]0,1[$ independent of the past. The new

segments follow the same dynamics independently. The process $S(t)$ is called a homogeneous *interval fragmentation*.

For a given point $x \in]0,1[$ and a realization of an interval fragmentation we will say that x has *fragmentation speed* or *rate of decrease* v if

$$-\log(|I_x(t)|)/t \to v$$

where $|I_x(t)|$ is the size of the fragment that contains x at time t. Pick $u \in]0,1[$ at random from the uniform distribution, then it is plain from our construction that $|I_u(t)|$ is a pure jump process, and that the waiting times between each jump are exponential with parameter 1. Clearly $\xi(t) = -\log(|I_u(t)|)$ is a compound Poisson process whose increment is given by the random variable $-\log(Z)$ where conditionally on V, Z is V with probability V and is $1-V$ with probability $1-V$. Thus, if we define $v_{\text{typ}} := \mathbf{E}(-\log(Z))$, then almost surely

$$-\frac{1}{t}\log|I_u(t)| \to v_{\text{typ}} \text{ as } t \to \infty.$$

Hence almost every points in $]0,1[$, in the sense of the Lebesgue measure, has rate of decrease v_{typ}, which is thus the *typical* rate of decrease.

On the other hand, we can define the process $J(t)$ of "the largest fragment followed" inductively by following the largest fragment at each dislocation. Again $-\log(J(t))/t$ is a compound Poisson process and the SLLN entails the almost sure convergence of $-\log(J(t))/t \to v'$ where clearly $v' < v_{\text{typ}}$. So there must be some points whose fragmentation speed is less than v_{typ}. The same technique also allows us to find some fast points of fragmentation by selecting the smallest fragment at each dislocation.

If we give ourselves an interval $]a,b[$ in $]0,1[$, then almost surely there exists a time t such that one of the interval components of $S(t)$ is included in $]a,b[$. As new segments evolve independently and with the same dynamics as the original fragment, the same analysis applies and $]a,b[$ contains almost surely a point with fragmentation speed v'. This shows that if there exists almost surely a point whose fragmentation speed is $v > 0$, then there is a dense subset of $]0,1[$ of points having the same property.

Although the example above is good for intuition, it has two limits. First one can suppose that when a segment splits, it might give birth to any number of sub-segments, possibly infinite, and not just two. Second, in this example, the splitting times are "discrete", the first time of splitting is almost surely strictly positive. In the sequel we consider more generally the case where fragmentation may occur continuously.

The aim of this paper is to show that, for a large class of homogeneous fragmentations, there are some points with a different fragmentation speed than the typical one and to study the sets of such points. This problem resembles by some aspects the study of the fast points of the Brownian motion of Orey and Taylor [26] (see also Davis [9], Kahane [17] and Perkins [28] for more insight on fast and slow points of the Brownian motion) or the recent works of Dembo et al. on thin and thick points of planar and d-dimensional Brownian motion (see [12, 13, 11]). There are also some natural and obvious connections with the theory of branching processes that stem from the fact that there is a genealogic structure in the interval fragmentations. Some of the techniques we use are close to those used by Shieh & Taylor [30], Shieh & Mörters [25] and Liu [21] for studying the multifractal spectra of the branching measure. This connection was already used in the analysis of self-similar recursive fractal, see Hambly and Jones [16] and the references therein. Although we do not make use of it here, there seems to be a way of doing some of the proofs using ideas of percolation on a tree (see [22, 27] or the book in preparation [23]).

We now introduce informally some notions related to fragmentations (definitions being given in section 2) and state our main results.

2.1.2 Main results

A homogeneous interval fragmentation $S(t)$ can be heuristically described as a nested family of open subsets of $]0,1[$ (i.e. $S(t) \subseteq S(s)$ whenever $t \geq s$) such that when a new fragment appears, it starts a new independent fragmentation which, up to the scale factor, has the same law as the initial one.

Define

$$\mathcal{S}^{\downarrow} := \{x = (x_1, x_2, ...), x_1 \geq x_2 \geq ... \geq 0, \sum_i x_i \leq 1\}.$$

Denote by Λ the map $\mathcal{O} \to \mathcal{S}^{\downarrow}$ that associates to an open subset of $]0,1[$ the ordered vector of the lengths of its interval decomposition. Then if $S(t)$ is an interval fragmentation, we denote by $X(t) = \Lambda(S(t))$ the associated *ranked fragmentation*. [1] The configuration space for ranked fragmentations is $\mathcal{S}^{\downarrow}$.

We suppose that at all times the Lebesgue measure of $S(t)$ is 1 (i.e. there is no loss of mass), thus the associated ranked fragmentation $X(t) = \Lambda(S(t))$

[1] Conversely, it is known that given $X(t)$ there exists $S(t)$ such that $X(t) \stackrel{\mathcal{L}}{=} \Lambda(S(t))$ in the sense of identity of the finite dimensional marginals.

takes its values in the smaller space

$$\mathcal{S}_1^{\downarrow} := \{x = (x_1, x_2, ...), x_1 \geq x_2 \geq ... \geq 0, \sum_i x_i = 1\}.$$

Bertoin has shown in [3] that the law of the process $X(t)$ is completely characterized by the so-called splitting measure $\nu(.)$ which is a measure on $\mathcal{S}_1^{\downarrow}$ such that

$$\int_{\mathcal{S}_1^{\downarrow}} (1 - x_1)\nu(dx) < \infty. \tag{2.1}$$

Roughly, the splitting measure describes the "rates" at which fragments split. Heuristically, if $A \subset \mathcal{S}^{\downarrow}$, then, for any $r \in]0,1[$, $\nu(A)$ is the rate at which a fragment of size r splits into smaller masses $(x_1, x_2, ...)$ such that $(x_1/r, x_2/r, ...) \in A$.

Some information about ν is caught by the function Φ :

$$\Phi(q) := \int_{\mathcal{S}_1^{\downarrow}} \left(1 - \sum_{i=1}^{\infty} x_i^{q+1}\right) \nu(dx), \ q > \underline{p}, \tag{2.2}$$

where

$$\underline{p} := \inf \left\{ p \in \mathbb{R} : \int_{\mathcal{S}_1^{\downarrow}} \sum_{i=2}^{\infty} x_i^{p+1} \nu(dx) < \infty \right\}.$$

From now on, we always suppose $\underline{p} < 0$.

As in [5], define two constants

$$v_{\text{typ}} := \Phi'(0+) = -\int_{\mathcal{S}_1^{\downarrow}} \left(\sum_{i=1}^{\infty} x_i \log(x_i)\right) \nu(dx)$$

which, as we shall see, is the typical rate of decrease, i.e. it is a.s. the fragmentation speed of a point picked at random in $]0,1[$ from the uniform distribution, and $v_{\min}$ as follows. Let $\overline{p} > 0$ be the unique solution of the equation

$$(p+1)\Phi'(p) = \Phi(p), p > \underline{p}.$$

The function $p \to \Phi(p)/(p+1)$ reaches its unique maximum $v_{\min}$ on $\left]\underline{p}, \infty\right[$ at $\overline{p}$ and

$$v_{\min} := \Phi'(\overline{p}) = \Phi(\overline{p})/(\overline{p}+1).$$

It is shown in [5] that $v_{\min}$ is the rate of exponential decrease of the largest fragment, i.e. almost surely

$$\lim_{t \to \infty} -t^{-1} \log X_1(t) = v_{\min}.$$

So clearly $v_{\min} \leq v_{\text{typ}}$. However this does not mean that there exists some point with rate $v_{\min}$: for that we would need the existence of a point which is often enough in the largest fragment, and such a point might well not exist. Define also

$$v_{\max} := \Phi'(\underline{p}+) \in [0, \infty] .$$

Let $\mathcal{G}_v$ be the set of points with fragmentation speed v

$$\mathcal{G}_v := \{x \in]0,1[: \lim_{t\to\infty} -t^{-1} \log(|I_x(t)|) = v\}.$$

Also define $\overline{\mathcal{G}}_v$ and $\underline{\mathcal{G}}_v$ as follows :

$$\begin{aligned} \overline{\mathcal{G}}_v &:= \{x \in]0,1[: \limsup_{t\to\infty} -t^{-1} \log(|I_x(t)|) \leq v\}, \\ \underline{\mathcal{G}}_v &:= \{x \in]0,1[: \liminf_{t\to\infty} -t^{-1} \log(|I_x(t)|) \geq v\}. \end{aligned}$$

Thus a point in $\underline{\mathcal{G}}_v$ (resp. in $\overline{\mathcal{G}}_v$) will be, for t large enough, in a small (resp. large) fragment compared to e^{-vt}.

Let Υ_v be the reciprocal of v by Φ' i.e. $\Phi'(\Upsilon_v) = v$. Define

$$C(v) := (\Upsilon_v + 1)v - \Phi(\Upsilon_v) \tag{2.3}$$

for $v \in]v_{\min}, v_{\max}[$ and $C(v) = -\infty$ elsewhere. As $\tilde{C}(v) = C(v) - v$ is thus the Legendre transform $\mathcal{L}$ of Φ, one has

$$\mathcal{L}\tilde{C}(.) = \mathcal{L}\mathcal{L}\Phi(.) = \Phi(.).$$

Thus Φ and C determine each other uniquely.

A fragmentation such that each splitting produces two fragments (i.e. ν only charges the subspace $\{x \in \mathcal{S}^{\downarrow} : x_3 = 0\}$) is called a *binary fragmentation.* In the case of a binary fragmentation, it is not hard to see that Φ determines ν, thus Φ characterizes the law of the fragmentation.

Define the *fragmentation spectra* to be, for $v_{\min} < v < v_{\max}$, the function that associates the Hausdorff dimension (denoted by Dim(.) in the sequel) of the sets $\mathcal{G}_v$, $\overline{\mathcal{G}}_v$ and $\underline{\mathcal{G}}_v$ to v (for a definition and main properties of the Hausdorff measure and dimension see for instance [14]).

The following theorem, which gives the fragmentation spectra explicitly in terms of the function $v \to C(v)$ thus entails that the law of a binary fragmentation is characterized by its spectra.

Theorem 2.1. *For each $v \in]v_{min}, v_{max}[$, almost surely*

$$\text{Dim}(\mathcal{G}_v) = C(v)/v = 1 + \Upsilon_v - \Phi(\Upsilon_v)/v, \tag{2.4}$$

$$\mathrm{Dim}(\overline{\mathcal{G}}_v) \;=\; C(v)/v \text{ if } v \le v_{typ} \text{ and } = 1 \text{ if } v \ge v_{typ}, \tag{2.5}$$

$$\mathrm{Dim}(\underline{\mathcal{G}}_v) \;=\; C(v)/v \text{ if } v \ge v_{typ} \text{ and } = 1 \text{ if } v \le v_{typ}. \tag{2.6}$$

One can easily verify that $C(v_{\mathrm{typ}})/v_{\mathrm{typ}} = 1$, $C(v_{\min})/v_{\min} = 0$, and that $C(v)/v$ is continuous and decreases as v get farther of v_{typ}.

However, we do not necessarily have $C(v)/v \to 0$ when $v \to v_{\max}$.

More precisely as

$$\lim_{v \to v_{\max}} C(v)/v = \lim_{p \searrow \underline{p}} 1 + p - \Phi(p)/\Phi'(p) \ge 1 + \underline{p},$$

as soon as $\underline{p} > -1$, for each $v > v_{\mathrm{typ}}$ one has that

$$\mathrm{Dim}(\underline{\mathcal{G}}_v) \ge 1 + \underline{p} > 0.$$

The natural question in this setting is "are there some points with a super-exponential fragmentation behavior ?" or more precisely, can we define a set of such points with a non-trivial dimension ?

Consider the set

$$\mathcal{H} = \{x \in]0,1[: \limsup_{t \to \infty} -\frac{1}{t} \log |I_x(t)| = +\infty\}.$$

Then it can be shown in the case $\Phi'(\underline{p}+) = \infty$ that

$$\mathrm{Dim}(\mathcal{H}) = 1 + \underline{p}.$$

The upper bound is established as in the case of $\underline{\mathcal{G}}_v$ (see section 2.3) and the lower bound can be obtained roughly through the same techniques employed in section 2.4.

The rest of this paper is organized as follows. The next section introduces notations, notions and definitions. Upper bounds are given in section 3. Section 4, which represents the most important part of this work, gives a lower bound for the Hausdorff dimension of $\mathcal{G}_v$ using a Galton-Watson tree that reflects the genealogical structure of the interval fragmentation.

2.2 Preliminaries

We now recall some facts mostly lifted from [2, 3, 4, 5] on homogeneous fragmentations and their asymptotic behavior.

A homogeneous interval fragmentation is a Markov process with values in $\mathcal{O}$ which enjoys two key properties : fragmentation and homogeneity. The fragmentation property states that when a new fragment (a new segment in the example) is born, it starts a new independent fragmentation of its own. This can be seen as a version of the branching property. The homogeneity property states that this new fragmentation has, up to a scaling factor, the same law as the initial one.

Specifically, if $\mathbb{P}$ stands for the law of the interval fragmentation $S(t)$ started from $]0,1[$, then for $s,t \geq 0$ conditionally on

$$S(t) = \bigcup_{i\in\mathbb{N}} J_i(t)$$

(where $J_i(t)$ is the interval decomposition of the open $S(t)$, i.e. for each i, $J_i(t)$ is an open subinterval of $S(t)$, the J_i's are disjoints and $\cup_i J_i(t) = S(t)$) $S(t+s)$ has same law as $S^{(1)}(s) \cup S^{(2)}(s) \cup ...$ where for each i, $S^{(i)}(s)$ is a subset of $J_i(t)$ and has same distribution as the image of $S(s)$ by the homothetic map $]0,1[\to J_i(t)$.

Similarly, a homogeneous ranked fragmentation is a Markov process with values in $\mathcal{S}_1^{\downarrow}$ such that if $\mathbb{P}$ stands for the law of the ranked fragmentation X started from $(1,0,0,..)$, then for $s,t \geq 0$ conditionally on $X(t) = (x_1, x_2, ...)$, $X(t+s)$ has same law as the variable obtained by reordering the elements of the random sequences $X^{(1)}(s), X^{(2)}(s), ...$, where for each i, $X^{(i)}(s) \in \mathcal{S}^{\downarrow}$ has same distribution as $X(s)$ under $\mathbb{P}_{x_i}$ where $\mathbb{P}_r$ is the image of $\mathbb{P}$ by the map

$$(x_1, x_2, ...) \to (rx_1, rx_2, ...).$$

The homogeneous fragmentations we shall consider in this work are those for which there is no loss of mass, i.e. such that almost surely, for all $t > 0$

$$\sum_i X_i(t) = 1.$$

This is why the configuration space is $\mathcal{S}_1^{\downarrow}$ and not, as usual, the more general space $\mathcal{S}^{\downarrow}$. As we have said, the law of such a ranked fragmentation is completely characterized by a so called splitting measure, ν, which is a measure on $\mathcal{S}_1^{\downarrow}$ that verifies the integral condition (2.1).

The interpretation of the function Φ given by (2.2) is the following. Suppose that at the initial time, a random point u with uniform distribution is tagged on $]0,1[$: as in the example $|I_u(t)|$ is a size biased pick from $X(t) = (X_1(t), X_2(t), ..) = \Lambda(S(t))$, i.e.

$$|I_u(t)| \stackrel{\mathcal{L}}{=} X_K(t)$$

where K is a random variable with values in $\mathbb{N}$ such that

$$P(K = k|X(t)) = X_k(t),\ k = 1, ...$$

Then the process

$$\xi(t) = -\log(|I_u(t)|) \tag{2.7}$$

is a subordinator (an increasing Lévy process) and we have

$$\mathbf{E}(|I_u(t)|^q) = \mathbf{E}(e^{-q\xi(t)}) = e^{-t\Phi(q)},\ t \geq 0,$$

where Φ is given by (2.2)(see [3] for the proof and discussion). This has direct consequences such as $\Phi : \left]\underline{p}, \infty\right[\to]-\infty, \infty[$ being the Laplace exponent of a subordinator, it is a concave increasing analytic function. Furthermore $\Phi(0+) = 0$ (this comes from the mass conservation). Remark that

$$v_{\text{typ}} := \mathbf{E}(\xi(1)).$$

Then by the L.L.N., if $v_{\text{typ}} < \infty$ (which holds whenever $\underline{p} < 0$), for Lebesgue almost every point $x \in]0,1[$

$$\lim_{t\to\infty} -t^{-1}\log(|I_x(t)|) = v_{\text{typ}} \qquad \text{a.s.}$$

which proves that $|\mathcal{G}_{v_{\text{typ}}}|$, the Lebesgue measure of $\mathcal{G}_{v_{\text{typ}}}$, is 1.

The starting point of this work is an estimate obtained by Bertoin in [5] concerning the number of abnormally "large" or "small" fragments at time t. More precisely, consider a homogeneous ranked fragmentation $(X_t)_{t\geq 0}$, then, for $v \in]v_{\min}, v_{\max}[$ one has with probability one

$$\lim_{\epsilon\to 0}\lim_{t\to\infty} t^{-1}\log\left(\text{Card}\left\{i \in \mathbb{N} : e^{-(v+\varepsilon)t} \leq X_i(t) \leq e^{-(v-\varepsilon)t}\right\}\right) = C(v), \tag{2.8}$$

where $C(v)$ is the function defined by (2.3).

We proceed in two steps to prove Theorem 2.1 : we will first give upper bounds and then a lower bound for $\mathcal{G}_v$ and use inclusions to conclude.

2.3 Upper bound

We prove the upper bound for the dimension of $\underline{\mathcal{G}}_v$ and $\overline{\mathcal{G}}_v$ and the conclusion follows for $\mathcal{G}_v$ by the inclusion

$$\mathcal{G}_v \subset \underline{\mathcal{G}}_v \cap \overline{\mathcal{G}}_v.$$

Let $(S(t))_{t\geq 0}$ be an interval fragmentation. We denote by

$$X(t) = (X_1(t), X_2(t), ...) = \Lambda(S(t))$$

the associated ranked fragmentation. We also use the notation

$$S(t) = \bigcup_{i\in\mathbb{N}} J_i(t)$$

where $(J_1(t), J_2(t), ...)$ is an interval decomposition of $S(t)$ and the labelling is size-wise, i.e. for each i, $X_i(t) = |J_i(t)|$.

2.3.1 Upper bound for (2.6)

In this section we consider the case $v > v_{\text{typ}}$. Define the collection of indices

$$\Theta_v(t) = \{i \in \mathbb{N} : X_i(t) \leq e^{-vt}\}$$

and note that $\forall N \in \mathbb{N}$, for all $w \in]v_{\text{typ}}, v[$ the set $\cup_{n\geq N} \cup_{i\in\Theta_w(n)} J_i(n)$ is a cover of $\underline{\mathcal{G}}_v$ (actually it is a cover for the larger set

$$\{x \in]0,1[: \limsup_{t\to\infty} -\frac{1}{t}\log|I_x(t)| > v\}$$

and hence the upper-bound for the dimension of $\mathcal{H}$ is also proven here). Thus we want to show that for $\alpha > C(v)/v$ as closed to $C(v)/v$ as wished, for w close enough to v

$$\sum_n \sum_{i\in\Theta_w(n)} X_i^\alpha(n) < \infty$$

Fix $\epsilon > 0$ and take $\alpha = C(v)/v + \epsilon$.

Clearly, for any $\beta \in [0, \alpha[$

$$\sum_{i\in\Theta_w(n)} X_i^\alpha(n) \leq e^{-nw(\alpha-\beta)} \sum_{i\in\mathbb{N}} X_i^\beta(n).$$

Choose $\beta = \Upsilon_v + 1$. Thus, as

$$\alpha = C(v)/v + \epsilon = \beta + \epsilon - \Phi(\Upsilon_v)/v$$

and as $\Phi(\Upsilon_v)/v < 0$ when $v > v_{\text{typ}}$ we see that $\beta < \alpha$.

Remark also that

$$(\alpha - \beta)v = -\Phi(\beta - 1) + \epsilon v.$$

Thus, if we choose w close enough to v

$$e^{-nw(\alpha-\beta)} \sum_{i\in\mathbb{N}} X_i^\beta(n) = e^{-n\epsilon'} e^{n\Phi(\beta-1)} \sum_{i\in\mathbb{N}} X_i^\beta(n),$$

where $\epsilon' > 0$. Bertoin has shown in [5] (see Theorem 2 therein) that

$$e^{n\Phi(\beta-1)} \sum_{i\in\mathbb{N}} X_i^\beta(n)$$

is a positive martingale. Hence, a.s.

$$e^{-nw(\alpha-\beta)} \sum_{i\in\mathbb{N}} X_i^\beta(n) = \mathrm{o}(e^{-n\epsilon'})$$

which concludes the proof.

2.3.2 Upper bound for (2.5)

In this section we consider the case $v < v_{\text{typ}}$. The main difference between the proofs for the upper bounds of (2.5) and (2.6) come from the fact that when $v > v_{\text{typ}}$ $\Upsilon_v < 0$ and hence $\Phi(\Upsilon_v) < 0$ whereas the converse is true when $v < v_{\text{typ}}$.

Denote by $\Theta_v^c(t) := \mathbb{N}\backslash\Theta_v(t)$ the complementary in $\mathbb{N}$ of $\Theta_v(t)$. Note that $\forall N \in \mathbb{N}$, for all $w \in]v, v_{\text{typ}}[$ the set $\cup_{n\geq N} \cup_{i\in\Theta_w^c(n)} J_i(n)$ is a cover of $\overline{\mathcal{G}}_v$.

Recall from (2.7) the notation $I_u(t)$ for the size-biased pick and $\xi_t = -\log|I_u(t)|$ for the associated subordinator. Clearly

$$\begin{aligned}\mathbf{E}\left[\sum_{i\in\Theta_w^c(t)} X_i^\alpha(t)\right] &= \mathbf{E}\left[\exp(-(\alpha-1)\xi_t), \xi_t < wt\right] \\ &\leq e^{awt}\mathbf{E}\left[\exp(-(a+\alpha-1)\xi_t)\right]\end{aligned}$$

for all $a > 0$. As $\Upsilon_v > 0$, when $\epsilon = \alpha - C(v)/v < \Phi(\Upsilon_v)/v$ one may choose $a = \Upsilon_v + 1 - \alpha = \Phi(\Upsilon_v)/v - \epsilon > 0$. Hence the right hand term becomes

$$e^{-[\Phi(\Upsilon_v)-(\Upsilon_v+1)w+\alpha w]t} = e^{-\epsilon v t} e^{(\Upsilon_v+1-\alpha)(w-v)t} = e^{-\epsilon' t}$$

for a well chosen $\epsilon' > 0$ when w is close enough to v. Hence the series $\sum_{i\in\Theta_w^c(n)} X_i^\alpha(n)$ is convergent and the upper bound for (2.5) is proven.

2.4 Lower bound

To complete the proof of Theorem 2.1, we wish to construct a subset K of $\mathcal{G}_v$ of Hausdorff dimension large enough. More precisely, we shall obtain a lower bound for $\mathrm{Dim}(K)$ by using the Hölder index of an increasing process indexed by $t \in]0,1[$ that only grows on points of K, and which can thus be seen as a local time on this set.

We obtain K by mean of a branching process $(G(n))_{n\in\mathbb{N}}$. More precisely $G(n)$ is the union of a collection $H(n)$ of some of the fragments that are present at time δn and that are included in $G(n-1)$, so $(G(n))_{n\in\mathbb{N}}$ is a nested sequence. We will then define $K = \cap_{n\in\mathbb{N}} G(n)$. We begin by a careful construction of G. We first define a somewhat "natural" branching process associated to the fragmentation and then show how to modify it to use classical results from the theory of branching processes.

2.4.1 Construction of the branching process

Remark that there is a natural notion of genealogy for interval fragmentations. Namely, the "sons" at time $t+s$ of a given fragment I of $S(s)$ are just the fragments of $S(t+s)$ that are included in I. Our strategy to find points in $\mathcal{G}_v$ will be to look at $S(t)$ at a set of times of the form $\{\delta n\}_{n\in\mathbb{N}}$, and at each step to select the sons of the preceding generation such that the ratio of the sizes father/son lies in an interval above v if the father was too large and under v in the opposite case.

More precisely, take an interval $\kappa \subset]v_{\min}, v_{\max}[$, then for all $t > 0$ define

$$\chi_\kappa(t) = \mathrm{Card}\{i : -\log(|J_i(t)|)/t \in \kappa, \{0,1\} \cap \bar{J}_i(t) = \emptyset\}, \tag{2.9}$$

i.e. the number of intervals with sizes in $\exp(-t\kappa)$ which do not touch the boundary of $]0,1[$. The reason why we impose this last condition is that we want to take the intersection of an infinite nested sequence of collections of open intervals. If the closure of each generation is in the interior of the preceding generation and is not empty, then the intersection is not empty. Remark that by homogeneity, for $s > 0$ and $j \in \mathbb{N}$, conditionally on $|J_j(s)| > 0$, $\chi_\kappa(t)$ has the same law as

$$\mathrm{Card}\left\{ i : \begin{array}{c} J_i(t+s) \subseteq J_j(s)\ , \\ -t^{-1}\log\left(\frac{|J_i(t+s)|}{|J_j(s)|}\right) \in \kappa, \\ \partial J_j(s) \cap \bar{J}_i(t+s) = \emptyset \end{array} \right\}$$

where ∂I is the boundary of I.

Take $\varepsilon, \delta > 0$ and $H_{\varepsilon,\delta}(0) := \{]0,1[\}$. Define inductively on n the sets $H_{\varepsilon,\delta}(n)$ as the collection of the the interval components of $S(n\delta)$ which fulfill the following three conditions. Firstly, every $I \in H_{\varepsilon,\delta}(n)$ must be included in some $J \in H_{\varepsilon,\delta}(n-1)$. Then there is a relative-size condition : if $I \in H_{\varepsilon,\delta}(n)$ and $J \in H_{\varepsilon,\delta}(n-1)$ are such that $I \subset J$ then

- if $|J| < e^{-v(n-1)\delta}$ then $-\delta^{-1}\log(|I|/|J|) \in [v, v+\varepsilon]$.
- if $|J| \geq e^{-v(n-1)\delta}$ then $-\delta^{-1}\log(|I|/|J|) \in [v-\varepsilon, v]$.

In both case we finally impose that $\bar{I} \cap \partial J = \emptyset$.

In some respects $H_{\varepsilon,\delta}$ is much like a multi-type branching process, with each particle corresponding to a segment and thus having some length attached. A "particle" I of the n^{th} generation (i.e. a segment of the collection $H_{\varepsilon,\delta}(n)$) will be called of

- type 1 if $|I| < e^{-vn\delta}$ and in that case its offspring has same distribution as $\chi_{[v,\varepsilon+v]}$
- and of type 2 if $|I| \geq e^{-vn\delta}$ and in that case its offspring has same distribution as $\chi_{[v-\varepsilon,v]}$.

The difference being that although here, as in the classical case, the law of the total number of children of a particle I only depends on its type, it happens that the repartition between type 1 and 2 of these children depends on the precise size of I.

However, it can easily be seen by induction that for all $n \in \mathbb{N}$, for any $I_n \in H_{\varepsilon,\delta}(n)$

$$e^{vn\delta}|I_n| \in \left[e^{-\varepsilon\delta}, e^{\varepsilon\delta}\right]. \tag{2.10}$$

Thus a.s. for any nested sequence of non-empty intervals $I_n \in H_{\varepsilon,\delta}(n)$ (implicitly, we are conditioning on non-extinction)

$$\lim_{n\to\infty} -\log(|I_n|)/n\delta = v.$$

If there exists $\{x\} = \cap I_n$, if we denote by $n_t := \sup\{n \in \mathbb{N} : n\delta < t\}$, then for all t one has the bounds

$$-\log(|I_{n_t}|)/((n_t+1)\delta) \leq -\log(|I_x(t)|)/t \leq -\log(|I_{(n_t+1)}|)/(n_t\delta).$$

Hence, one has that $-\log(|I_x(t)|)/t \to v$ almost surely and $x \in \mathcal{G}_v$.

For all n we define

$$G_{\varepsilon,\delta}(n) := \bigcup_{I \in H_{\varepsilon,\delta}(n)} I,$$

i.e. $H_\cdot(.)$ is the collection of the open interval components of $G_\cdot(.)$. Hence $\cap_n G_{\varepsilon,\delta}(n) \subseteq \mathcal{G}_v$. Note that we could not use a monotype branching process here, i.e. at each generation keep the sons such that

$$|I|/|J| \in \left[e^{-(v+\varepsilon)\delta}, e^{-(v-\varepsilon)\delta}\right]$$

because this would lead to points in $\overline{\mathcal{G}}_{v+\varepsilon} \cap \underline{\mathcal{G}}_{v-\varepsilon}$ and not necessarily in $\mathcal{G}_v$.

There are two issues we must take care of now :

- first, we must choose ε and δ such that we catch enough of $\mathcal{G}_v$, and this amounts to control the growth of the branching process
- and second, we would rather work with a true Galton-Watson tree. The branching process $H_{\varepsilon,\delta}$ is the most natural to consider, but in order to use classical results of the branching processes theory we need to cut some branches in order to obtain a true, super-critical, Galton-Watson process. Furthermore, we must do so while keeping its rate of growth close enough to its original value.

2.4.2 Rate of growth

In the right-hand side of (2.8) the interval $[v-\varepsilon, v+\varepsilon]$ is symmetric around v, but it is easy to see that one hardly needs to change the arguments used in [5] to have that a.s.

$$C(v) = \lim_{\varepsilon \to 0} \lim_{t \to \infty} t^{-1} \log \chi_{[v-\varepsilon,v]}(t) \tag{2.11}$$

and

$$C(v) = \lim_{\varepsilon \to 0} \lim_{t \to \infty} t^{-1} \log \chi_{[v,v+\varepsilon]}(t), \tag{2.12}$$

where $\chi_{[a,b]}(t)$ is defined by (2.9).

Hence, clearly if we fix $\epsilon' > 0$ and $\eta > 0$, then we may find $\varepsilon > 0$ and t_0 arbitrarily large such that

$$\forall t > t_0 : P(|t^{-1} \log(\chi_{[v-\varepsilon,v]}(t)) - C(v)| > \eta) < \epsilon' \tag{2.13}$$

and of course the same is true replacing $[v-\varepsilon, v]$ by $[v, v+\varepsilon]$.

For each $t > 0$ consider a variable $\tilde{\chi}(t)$ which law is given by

$$P(\tilde{\chi}(t) = e^{(C(v)-\eta)t}) = 1 - \epsilon'$$

and

$$P(\tilde{\chi}(t) = 0) = \epsilon'$$

where

$$\epsilon' = P(|t^{-1}\log(\chi_{[v-\varepsilon,v]}(t)) - C(v)| > \eta) \vee P(|t^{-1}\log(\chi_{[v,v+\varepsilon]}(t)) - C(v)| > \eta).$$

Note that

$$|t^{-1}\log\left(E\left[\tilde{\chi}(t)\right]\right) - C(v)| \leq \eta + t^{-1}|\log(1-\epsilon')|.$$

Fix ϵ and choose ϵ' and η such that $\eta + |\log(1-\epsilon')| < \epsilon$, then choose ε and $t_0 > 1$ by (2.13).

Plainly $\tilde{\chi}(t)$ is stochastically dominated by $\chi_{[v-\varepsilon,v]}(t)$ and $\chi_{[v,v+\varepsilon]}(t)$. Hence, we can construct a true Galton-Watson tree by thinning $H_{\varepsilon,\delta}$ where $\delta > t_0$. More precisely there exists a procedure for deciding at each node to erase or not some or all of the offspring and such that the resulting tree, denoted by $\mathbb{H}_{v,\epsilon}$ is a Galton-Watson tree, with offspring distribution given by the law of $\tilde{\chi}(\delta)$.

Thus, if we define $m := \mathbf{E}(\tilde{\chi}(\delta))$, the expectation of the number of children of a particle, one has

1.
$$\left|\delta^{-1}\log m - C(v)\right| < \epsilon \tag{2.14}$$
and thus $m > 1$ and $\mathbb{H}_{v,\epsilon}$ is super-critical.

2. for each $n \in \mathbb{N}$ the closure of $\mathbb{G}_{v,\epsilon}(n) := \cup_{I\in\mathbb{H}_{v,\epsilon}(n)} I$ is in $\mathbb{G}_{v,\epsilon}(n-1)$

3.
$$\cap_{n\in\mathbb{N}}\mathbb{G}_{v,\epsilon}(n) \subseteq \mathcal{G}_v.$$

This last point only makes sense if the tree doesn't die, so in the following we condition systematically on non-extinction.

We now show that $\mathrm{Dim}(\mathcal{G}_v) \geq C(v)/v$ for $v \in]v_{\min}, v_{\max}[$ which entails the result for $\overline{\mathcal{G}}_v$ and $\underline{\mathcal{G}}_v$ by inclusion.

2.4.3 Proof of the lower bound

Fix $v \in]v_{\min}, v_{\max}[$ and $\epsilon > 0$. Choose ε and $\delta > t_0$ as shown above and consider the tree $\mathbb{G}_{v,\epsilon}$. Let us recall the signification of the parameters : ϵ controls the precision of the growth rate, ε is the width of the window of acceptable sizes and δ is our time-step. Define

$$Z_{v,\epsilon}(n) = \mathrm{Card}\{\mathbb{H}_{v,\epsilon}(n)\}$$

the size of the nth generation.

In the following to simplify the notations we drop the subscript v, ϵ and we note $Z(n), \mathbb{G}(n)$ or $\mathbb{H}(n)$ for $Z_{v,\epsilon}(n), \mathbb{G}_{v,\epsilon}(n)$ or $\mathbb{H}_{v,\epsilon}(n)$ respectively.

Recall we are conditioning on non-extinction of the branching process $\mathbb{H}(.)$. This conditioning can be made at no cost because in the event that $\mathbb{H}(.)$ becomes extinct, one can restart a new independent tree on any fragment present at the extinction time for instance, and iterate this procedure until one find an infinite tree. Otherwise said, $\cap_n G_{\epsilon,\delta}(n)$ is nonempty only with positive probability, but it is however enough to show that its dimension is the correct one only with positive probability, for the dimension must be a constant a.s. due to the independence of fragmentation on different subsets.

It is well known that almost surely

$$m^{-n} Z(n) \to W > 0$$

(more precisely $W > 0$ on the survival set of the tree).

Let σ be a node of our tree (thus it is also a subinterval of $]0,1[$) and let $\mid \sigma \mid$ designate its height in the tree, let $Z^{(\sigma)}(n)$ be the number of its offspring in the tree at the generation $\mid \sigma \mid + n$, finally call $W(\sigma)$ the "renormalized weight" of the tree rooted at σ, i.e.

$$W(\sigma) := \lim_{n\to\infty} m^{-n} \mathrm{Card}\{\sigma' \in \mathbb{H}(\mid \sigma \mid + n) \ , \ \sigma' \subset \sigma\}.$$

Fix an interval $I \subset]0,1[$ and introduce

$$\begin{aligned} \mathbb{H}_I(n) &= \{\sigma \in \mathbb{H}(n) \ , \ \sigma \cap I \neq \emptyset\} \\ Z_I(n) &= \mathrm{Card}(\mathbb{H}_I(n)). \end{aligned}$$

Define

$$x \to L_x := \lim_n m^{-n} Z_{]0,x[}(n) \ , \ x \in]0,1[\, .$$

Lemma 2.1. *For each $\epsilon > 0$,*

1. *there exists a version $\tilde{L}$ of L which is Hölder continuous of order α for any $\alpha < (C(v) - \epsilon)/v$*

2. *L only grows on the set $\cap_n \mathbb{G}_{v,\epsilon}(n)$.*

Proof. We show the first point by verifying Kolmogoroff's criterium (see [29] Theorem 2.1 p.26).

Clearly one has that for all $x < y \in]0,1[$

$$|L_x - L_y| = \lim_n m^{-n} Z_{]x,y[}(n).$$

For any J open subinterval of $]0,1[$, define

$$\eta(J) := \sup\{n \in \mathbb{N} : e^{-v\delta n} \geq |J|\} = [-\log(|J|)/v\delta]\,.$$

This is very close to the largest n for which J can be included in an interval of the collection $\mathbb{H}(n)$, thus it is not difficult to see that at time $\eta(J)$ the number of intervals of $\mathbb{H}(\eta(J))$ which have a non empty intersection with J is bounded. More precisely, according to (2.10), for each n the size of the intervals in $\mathbb{H}(n)$ have a lower bound given by $e^{-vn\delta-\varepsilon\delta}$, so $|J|e^{-(v+\varepsilon)\delta}$ is a lower bound for the size of the intervals of $\mathbb{H}(\eta(J))$, and thus $Z_J(\eta(J)) \leq e^{(v+\varepsilon)\delta}$.

Thus, for all x, y such that $x < y$, one has by definition of L_x and using (2.14) that

$$\begin{aligned}
|L_y - L_x| &\leq m^{-\eta(]x,y[)} \sum_{\sigma \in \mathbb{H}_{]x,y[}(\eta(]x,y[))} W(\sigma) \\
&\leq e^{\log m(\frac{1}{v\delta}\log(y-x)+1)} \sum_{\sigma \in \mathbb{H}_{]x,y[}(\eta(]x,y[))} W(\sigma) \\
&\leq m|y-x|^{(C(v)-\epsilon)/v} \sum_{\sigma \in \mathbb{H}_{]x,y[}(\eta(]x,y[))} W(\sigma)
\end{aligned}$$

Remark that for all $\gamma > 1$ and all $J \subset]0,1[$ one has

$$\mathbf{E}\left[\Big(\sum_{\sigma \in \mathbb{H}_J(\eta(J))} W(\sigma)\Big)^\gamma\right] < \mathbf{E}\left[(W_1 + W_2 + ... + W_{[e^{(v+\varepsilon)\delta}]+1})^\gamma\right] < \infty$$

where the W_i are independent copies of W and the finiteness comes from the existence of finite moments of all order for W (this follows from e.g. Biggins and Bingham [7]).

We conclude that for each $\gamma > 1$ there exists a $k > 0$ such that

$$\mathbf{E}(|L_y - L_x|^\gamma) \quad \leq \quad k|y-x|^{\gamma(C(v)-\epsilon)/v}$$

which proves our first assertion.

The second part of the lemma is straightforward. The increasing function L_x only grows on the points of $]0,1[$ that correspond to the frontier of the tree, i.e. on $\cap_n \mathbb{G}(n)$. More precisely, for any interval $]a,b[\subset (\cap_n \mathbb{G}(n))^c$ one clearly has that $L_a = L_b$ by definition. On the other hand $L_{0+} = 0$ and $L_{1-} = W > 0$ so L_x can be thought of as a local time on $\cap_n \mathbb{G}(n)$. □

We conclude with the proof of Theorem 1. For a cover of $\cap_n G(n)$ of the type $\cup_{i=1}^N]l_i, r_i[$ (where the $]l_i, r_i[$ are disjoints open intervals) one has a.s.

$$\sum_{i=1}^N |\tilde{L}_{r_i} - \tilde{L}_{l_i}| = W.$$

Thus for all such cover with $\max_i(r_i - l_i)$ small enough

$$W \leq k \sum_{i=1}^N (r_i - l_i)^{(C(v)-\epsilon)/v}$$

and hence a.s.

$$\mathrm{Dim}(\mathcal{G}_v) \geq \mathrm{Dim}(\cap_n(G(n))) \geq (C(v) - \epsilon)/v.$$

To conclude simply let $\epsilon \to 0$.

Alternatively one could use the same approach as Orey and Taylor in [26]. They apply a lemma for the lower bound which is specific to self-similar sets, but as the arguments are very similar to those above we do not include this version of the proof. There is also a way of doing the proof by using some fine results of Liu [21] on the local behavior of the branching measure that does not rely on the Kolmogorov criterium.

We conclude with the three following remarks.

Remark 1: Although formally $\underline{p} \geq -1$, in the cases analogue to the introductory example (i.e. when the fragmentation is slow enough and Φ can be analytically extended beyond -1) then it can be shown that the theorem holds for some v larger than $\Phi'(\underline{p}+)$, more precisely for $v \in]v_{\min}, \Phi'(p^*)[$ where Φ is extended and

$$p^* = \inf\{p : 1 + p^* - \Phi(p^*)/\Phi'(p^*) > 0\}.$$

When it is finite, $\Phi'(p^*)$ is the maximum fragmentation speed.

Remark 2: In theorem 2.1, relations (2.5) and (2.6) hold almost surely simultaneously for all v. Indeed, call Ω_0 a set of probability 1 on which (2.4),(2.5) and (2.6) are true for each v rational. Define the set of events

$$A := \{\exists v \in]v_{\min}, v_{\max}[\text{ s.t. (2.5) or (2.6) is not true }\}.$$

As for each $v \leq w$ one has $\overline{\mathcal{G}}_v \subset \overline{\mathcal{G}}_w$ and $\underline{\mathcal{G}}_w \subset \underline{\mathcal{G}}_v$, it is clear that $A \subset \Omega_0^c$ the complementary of Ω_0, and hence $P(A) = 0$. The same arguments show that almost surely for all $v \in]v_{\min}, v_{\text{typ}}[$ one has

$$\text{Dim}(\mathcal{G}_v) \leq C(v)/v,$$

and almost surely for all $v \in]v_{\text{typ}}, v_{\max}[$ one has

$$\text{Dim}(\mathcal{G}_v) \geq C(v)/v.$$

Although it seems doubtful that there exists an exceptional set of v's for which (2.4) does not holds, the techniques employed in the present work does not allow one to conclude on that matter.

Remark 3: Define

$$\begin{aligned}\overline{\mathcal{H}}_v &:= \{x \in]0,1[: \liminf_{t\to\infty} -t^{-1}\log(|I_x(t)|) \leq v\},\\ \underline{\mathcal{H}}_v &:= \{x \in]0,1[: \limsup_{t\to\infty} -t^{-1}\log(|I_x(t)|) \geq v\}.\end{aligned}$$

Clearly $\overline{\mathcal{G}}_v \subset \overline{\mathcal{H}}_v$ and $\underline{\mathcal{G}}_v \subset \underline{\mathcal{H}}_v$, thus

$$\text{Dim}(\overline{\mathcal{H}}_v) \geq \text{Dim}(\overline{\mathcal{G}}_v)$$

and

$$\text{Dim}(\underline{\mathcal{H}}_v) \geq \text{Dim}(\underline{\mathcal{G}}_v).$$

Furthermore, it is easy to see by inspection of the proof in section 2.3 that the the same upper bound holds for $\overline{\mathcal{H}}_v$ (resp. $\underline{\mathcal{H}}_v$) and $\overline{\mathcal{G}}_v$ (resp. $\underline{\mathcal{G}}_v$). Thus

$$\text{Dim}(\overline{\mathcal{H}}_v) = \text{Dim}(\overline{\mathcal{G}}_v)$$

and

$$\text{Dim}(\underline{\mathcal{H}}_v) = \text{Dim}(\underline{\mathcal{G}}_v).$$

Acknowledgment. I would like to express my gratitude to professor J. Bertoin for suggesting the problem, for numerous discussions and for comments and to Professor A. Rouault for pointing out some references and for fruitful discussions. The author also wish to thanks two anonymous referees for their careful reading and suggesting some references.

BIBLIOGRAPHY

[1] A. Basedow, K. Ebert, and H. Ederer. Kinetic studies on the acidic hydrolysis of dextran. *Macromolecules*, 11:774–781, 1978.

[2] J. Berestycki. Ranked fragmentations. *ESAIM Probab. Statist.*, 6:157–175 (electronic), 2002.

[3] J. Bertoin. Homogeneous fragmentation processes. *Probab. Theory Related Fields*, 121(3):301–318, 2001.

[4] J. Bertoin. Self-similar fragmentations. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(3):319–340, 2002.

[5] J. Bertoin. The asymptotic behaviour of fragmentation processes. *J. Euro. Math. Soc.*, to appear, 2003.

[6] D. Beysens, X. Campi, and E. Peffekorn, editors. *Proceedings of the workshop : Fragmentation phenomena*, Les Houches Series. World Scientific, 1995.

[7] J. D. Biggins and N. H. Bingham. Large deviations in the supercritical branching process. *Adv. in Appl. Probab.*, 25(4):757–772, 1993.

[8] X. Campi. Multifragmentation: nuclei break up like percolation clusters. *J. Phys. A: Math. Gen.*, 19:917–921, 1986.

[9] B. Davis. On brownian slow points. *Z. Wahrscheinlichkeistheorie. und Verw. Gebiete*, 64(3):359–367, 1983.

[10] D. S. Dean and S. N. Majumdar. Phase transition in a random fragmentation problem with applications to computer science. *Phys. A: Math. Gen.*, 35(32):501–507, 2002.

[11] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Thin points for brownian motion. *Ann. Inst. H. Poincaré Probab. Statist*, 36(6):749–774.

[12] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Thick points for spatial brownian motion: Multifractal analysis of occupation measure. *Ann. Probab.*, 28:1—35, 2000.

[13] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni. Thick points for planar brownian motion and the erdös-taylor conjecture on random walk. *Acta Math.*, 186(2):239–270, 2001.

[14] K. J. Falconer. *The geometry of fractal sets*, volume 85 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1986.

[15] J. Gilvarry. Fracture of brittle solids. *J. Appl. Phys*, 32:391–399, 1961.

[16] B. M. Hambly and O. D. Jones. Thick and thin points for random recursive fractals. *Adv. in Appl. Probab.*, 35(1):251–277, 2003.

[17] J. Kahane. *Some random series of function*, volume 5 of *Cambridge studies in advanced mathematics*. Second edition, 1985.

[18] A. Kolmogoroff. Über das logaritmisch normale verteilungsgesetz de dimensionen de teilchen bei zerstückelung. *Acad. Sci. URSS*, 31:99–101, 1941.

[19] P. Krapivsky and S. N. Majumdar. traveling waves, front selection, and exact nontrivial exponents in a random fragmentation problem. *Phys. Rev. Lett.*, 85:5492, 2000.

[20] P. Krapivsky and S. N. Majumdar. Extreme value statistics and traveling fronts: An application to computer science. *Phys. Rev. E*, 65:036127, 2002.

[21] Q. Liu. Local dimensions of the branching measure on a Galton-Watson tree. *Ann. Inst. H. Poincaré Probab. Statist.*

[22] R. Lyons. Random walks and percolation on trees. *Ann. Probab.*, 18(3):931–958, 1990.

[23] R. Lyons and Y. Peres. *Probability on Trees and Networks*. Cambridge University Press, 1997.

[24] E. D. McGrady and R. M. Ziff. The kinetics of cluster fragmentation and depolymerisation. *J. Phys. A*, 18(15):3027–3037, 1985.

[25] P. Mörters and N.-R. Shieh. Thin and thick points for branching measure on a Galton-Watson tree. *Statist. Probab. Lett.*, 58(1):13–22, 2002.

[26] S. Orey and S. J. Taylor. How often on a Brownian path does the law of iterated logarithm fail? *Proc. London Math. Soc. (3)*, 28:174–192, 1974.

[27] Y. Peres. Probability on trees: an introductory climb. In *Lectures on probability theory and statistics (Saint-Flour, 1997)*, volume 1717 of *Lecture Notes in Math.*, pages 193–280. Springer, Berlin, 1999.

[28] E. Perkins. On the Hausdorff dimension of the Brownian slow points. *Z. Wahrsch. Verw. Gebiete*, 64(3):369–399, 1983.

[29] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[30] N.-R. Shieh and S. J. Taylor. Multifractal spectra of branching measure on a Galton-Watson tree. *J. Appl. Probab.*, 39(1):100–111, 2002.

[31] R. Shinar. On the behaviour of liquid dispersions in mixing vessels. *J. Fluid Mech.*, 10:259, 1961.

3. EXCHANGEABLE FRAGMENTATION COALESCENCE PROCESSES AND THEIR EQUILIBRIUM MEASURE

Abstract

We define and study a family of Markov processes with state space the compact set of all partitions of $\mathbb{N}$ that we call exchangeable fragmentation-coalescence processes. They can be viewed as a combination of exchangeable fragmentation as defined by Bertoin and of homogenous coalescence as defined by Pitman and Schweinsberg or Möhle and Sagitov. We show that they admit a unique invariant probability measure and we study some properties of their paths and of their equilibrium measure.

3.1 Introduction

Coalescence phenomena (coagulation, gelation, aggregation,...) and their duals fragmentation phenomena (splitting, erosion, breaks up,...), are present in a wide variety of contexts.

References as to the fields of application of coalescence and fragmentation models (physical chemistry, astronomy, biology, computer sciences...) may be found in Aldous [1] -mainly for coalescence- and in the proceedings [7] for fragmentation (some further references can be found in the introduction of [2]). Clearly, many fragmentation or coalescence phenomena are not "pure" in the sense that both are present at the same time. For instance, in the case of polymer formation there is a regime near the critical temperature where molecules break up and recombine simultaneously. Another example is given by Aldous [1], when, in his *one specific application* section, he discusses how certain liquids (e.g. olive oil and alcohol) mix at high temperature but separate under some critical level. When one lowers very slowly the temperature through this threshold, droplets of one liquid begin to form, merge and dissolve back very quickly.

It appears that coalescence-fragmentation processes are somewhat less tractable mathematically than pure fragmentation or pure coalescence. One of the reasons is that by combining these processes we lose some of the nice properties they exhibit when they stand alone, as for instance their genealogic or branching structure. Nevertheless, it is natural to investigate such processes, and particularly to look for their equilibrium measures.

In this direction Diaconis, Mayer-Wolf, Zeitouni and Zerner [8] considered a coagulation-fragmentation transformation of partitions of the interval $(0, 1)$ in which the coalescence procedure corresponds to the multiplicative coalescent while the splittings are driven by a quadratic fragmentation. By

relating it to the random transposition random walk on the group of permutations, they were able to prove a conjecture of Vershik stating that the unique invariant measure of this Markov process is the Poisson-Dirichlet law. We would also like to mention the work of Pitman [14] on a closely related split and merge transformation of partitions of $(0, 1)$ as well as Durrett and Limic [9] on another fragmentation-coalescence process of $(0, 1)$ and its equilibrium behavior. However, a common characteristic of all these models is that they only allow for binary splittings (a fragment that splits creates exactly two new fragments) and pairwise coalescences. Furthermore the rate at which a fragment splits or merges depends on its size and on the size of the other fragments.

Here, we will focus on a rather different class of coagulation-fragmentations that can be deemed *exchangeable* or *homogeneous*. More precisely, this paper deals with processes which describe the evolution of a countable collection of masses which results from the splitting of an initial object of unit mass. Each fragment can split into a countable, possibly finite, collection of sub-fragments and each collection of fragments can merge. One can have simultaneously infinitely many clusters that merge, each of them containing infinitely many masses.

We will require some homogeneity property in the sense that the rate at which fragments split or clusters merge does not depend on the fragment sizes or any other characteristic and is not time dependent.

Loosely speaking, such processes are obtained by combining the semigroups of a homogenous fragmentation and of an exchangeable coalescent. Homogeneous fragmentations were introduced and studied by Bertoin in [3, 4, 5]. Exchangeable coalescents, or rather Ξ-coalescents, where introduced independently by Schweinsberg in [15] [1] and by Möhle and Sagitov in [12] who obtained them by taking the limits of scaled ancestral processes in a population model with exchangeable family sizes.

Precise definitions and first properties are given in Section 3. Next, we prove that there is always a unique stationary probability measure for these processes and we study some of their properties. Section 5 is dedicated to the study of the paths of exchangeable fragmentation-coalescence processes.

The formalism used here and part of the following material owe much to a work in preparation by Bertoin based on a series of lectures given at the

[1] Schweinsberg was extending the work of Pitman [13] who treated a particular case, the so-called Λ-coalescent in which when a coalescence occurs, the involved fragments always merge into a single cluster.

IHP in 2003, [6].

3.2 Preliminaries

Although the most natural state space for processes such as fragmentation or coalescence might be the space of all possible ordered sequence of masses of fragments

$$\mathcal{S}^{\downarrow} = \{1 \geq x_1, \geq x_2 \geq ... \geq 0, \sum_i x_i \leq 1\},$$

as in the case of pure fragmentation or pure coalescence, we prefer to work with the space $\mathcal{P}$ of partitions of $\mathbb{N}$. An element π of $\mathcal{P}$ can be identified with an infinite collection of blocks (where a block is just a subset of $\mathbb{N}$ and can be the empty set) $\pi = (B_1, B_2, ...)$ where $\cup_i B_i = \mathbb{N}$, $B_i \cap B_j = \emptyset$ when $i \neq j$ and the labelling corresponds to the order of the least element, i.e. if w_i is the least element of B_i (with the convention $\min \emptyset = \infty$) then $i \leq j \Rightarrow w_i \leq w_j$. The reason for such a choice is that we can discretize the processes by looking at their restrictions to $[n] := \{1, ..., n\}$.

As usual, an element $\pi \in \mathcal{P}$ can be identified with an equivalence relation by setting

$$i \overset{\pi}{\sim} j \Leftrightarrow i \text{ and } j \text{ are in the same block of } \pi.$$

Let $B \subseteq B' \subseteq \mathbb{N}$ be two subsets of $\mathbb{N}$, then a partition π' of B' naturally defines a partition $\pi = \pi'_{|B}$ on B by taking $\forall i, j \in B, i \overset{\pi}{\sim} j \Leftrightarrow i \overset{\pi'}{\sim} j$, or otherwise said, if $\pi' = (B'_1, B'_2, ...)$ then $\pi = (B'_1 \cap B, B'_2 \cap B, ...)$ and the blocks are relabelled.

Let $\mathcal{P}_n$ be the set of partitions of $[n]$. For an element π of $\mathcal{P}$ the restriction of π to $[n]$ is $\pi_{|[n]}$ and we identify each $\pi \in \mathcal{P}$ with the sequence $(\pi_{|[1]}, \pi_{|[2]}, ..) \in \mathcal{P}_1 \times \mathcal{P}_2, ...$ We endow $\mathcal{P}$ with the distance

$$d(\pi^1, \pi^2) = 1/\max\{n \in \mathbb{N} : \pi^1_{|[n]} = \pi^2_{|[n]}\}.$$

The space $(\mathcal{P}, d)$ is then compact. In this setting it is clear that if a family $(\Pi^{(n)})_{n\in\mathbb{N}}$ of $\mathcal{P}_n$-valued random variable is *compatible*, i.e. if for each n

$$\Pi^{(n+1)}_{|[n]} = \Pi^{(n)},$$

i.e., the restriction to $[n]$ of $\Pi^{(n+1)}$ is $\Pi^{(n)}$, then the family $(\Pi^{(n)})_{n\in\mathbb{N}}$ uniquely determines a $\mathcal{P}$-valued variable Π such that for each n one has

$$\Pi_{|[n]} = \Pi^{(n)}.$$

Thus we may define the exchangeable fragmentation-coalescence processes by their $[n]$-restrictions.

Let us now define deterministic notions which will play a crucial role in the forthcoming constructions. We define two operators on $\mathcal{P}$, a coagulation operator, $\pi, \pi' \in \mathcal{P} \to \mathit{Coag}(\pi, \pi')$ (the coagulation of π by π') and a fragmentation operator $\pi, \pi' \in \mathcal{P}, k \in \mathbb{N} \to \mathit{Frag}(\pi, \pi', k)$ (the fragmentation of the k-th block of π by π').

- Take $\pi = (B_1, B_2, ...)$ and $\pi' = (B'_1, B'_2, ...)$. Then $\mathit{Coag}(\pi, \pi') = (B''_1, B''_2, ...)$, where $B''_1 = \cup_{i \in B'_1} B_i, B''_2 = \cup_{i \in B'_2} B_i, ...$ Observe that the labelling is consistent with our convention.
- Take $\pi = (B_1, B_2, ...)$ and $\pi' = (B'_1, B'_2, ...)$. Then $\mathit{Frag}(\pi, \pi', k)$ is the relabelled collection of blocks formed by all the B_i for $i \neq k$, plus the sub-blocks of B_k given by $\pi'_{|B_k}$.

Similarly, when $\pi \in \mathcal{P}_n$ and $\pi' \in \mathcal{P}$ or $\pi' \in \mathcal{P}_k$ for $k \geq \#\pi$ (where $\#\pi$ is the number of non-empty blocks of π) one can define $\mathit{Coag}(\pi, \pi')$ as above and when $\pi' \in \mathcal{P}$ or $\pi' \in \mathcal{P}_m$ for $m \geq \mathrm{Card}(B_k)$ one can define $\mathit{Frag}(\pi, \pi', k)$ as above.

Define $\mathbf{0} := (\{1\}, \{2\}, ...)$ the partition of $\mathbb{N}$ into singletons, $\mathbf{0}_n := \mathbf{0}_{|[n]}$, and $\mathbf{1} := (\{1, 2, ...\})$ the trivial partition of $\mathbb{N}$ in a single block, $\mathbf{1}_n := \mathbf{1}_{|[n]}$. Then $\mathbf{0}$ is the neutral element for *Coag*, i.e. for each $\pi \in \mathcal{P}$

$$\mathit{Coag}(\pi, \mathbf{0}) = \mathit{Coag}(\mathbf{0}, \pi) = \pi,$$

(for $\pi \in \cup_{n \geq 2} \mathcal{P}_n$, as $\mathit{Coag}(\mathbf{0}, \pi)$ is not defined one only has $\mathit{Coag}(\pi, \mathbf{0}) = \pi$) and $\mathbf{1}$ is the neutral element for *Frag*, i.e. for each $\pi \in \mathcal{P}$ one has

$$\mathit{Frag}(\mathbf{1}, \pi, 1) = \mathit{Frag}(\pi, \mathbf{1}, k) = \pi.$$

When $\pi \in \cup_{n \geq 2} \mathcal{P}_n$, for each $k \leq \#\pi$ one only has

$$\mathit{Frag}(\pi, \mathbf{1}, k) = \pi.$$

Note also that the coagulation and fragmentation operators are not really reciprocal because *Frag* can only split one block at a time.

Much of the power of working in $\mathcal{P}$ instead of $\mathcal{S}^{\downarrow}$ comes from Kingman's theory of exchangeable partitions. For the time being, let us just recall the basic definition. Define the action of a permutation $\sigma : \mathbb{N} \to \mathbb{N}$ on $\mathcal{P}$ by

$$i \overset{\sigma(\pi)}{\sim} j \Leftrightarrow \sigma(i) \overset{\pi}{\sim} \sigma(j).$$

A random element π of $\mathcal{P}$ or a $\mathcal{P}$ valued process Π is said to be exchangeable if for any permutation σ such that $\sigma(n) = n$ for all large enough n one has $\sigma(\pi) \overset{\mathcal{L}}{=} \pi$ or $\Pi(.) \overset{\mathcal{L}}{=} \sigma(\Pi(.))$.

3.3 Definition and construction of exchangeable fragmentation-coalescence processes

We can now define precisely the exchangeable fragmentation-coalescence processes and state some of their properties. Most of the following material is very close to the analogous definitions and arguments for pure fragmentations (see [3]) and coalescents (see [13, 15]).

Definition 3.1. *A $\mathcal{P}$-valued Markov process $(\Pi(t), t \geq 0)$, is an exchangeable fragmentation-coalescent process ("EFC process" thereafter) if it has the following properties :*

- *It is exchangeable.*
- *Its restrictions $\Pi_{|[n]}$ are finite state Markov chains which can only evolve by fragmentation of one block or by coagulation.*

More precisely, the transition rate of $\Pi_{|[n]}(.)$ from π to π', say $q_n(\pi, \pi')$, is non-zero only if $\exists \pi''$ such that $\pi' = Coag(\pi, \pi'')$ or $\exists \pi'', k \geq 1$ such that $\pi' = Frag(\pi, \pi'', k)$.

Fix n and $\pi \in \mathcal{P}_n$. For convenience we will also use the following notations for the transition rates : For $\pi' \in \mathcal{P}_m \backslash \{\mathbf{0}_m\}$ where $m = \#\pi$ the number of non-empty blocks of π, call

$$C_n(\pi, \pi') := q_n(\pi, Coag(\pi, \pi'))$$

the rate of coagulation by π'. For $k \leq \#\pi$ and $\pi' \in \mathcal{P}_{|B_k|} \backslash \{\mathbf{1}_{|B_k|}\}$ where $|B_k|$ is the cardinal of the k-th block, call

$$F_n(\pi, \pi', k) := q_n(\pi, Frag(\pi, \pi', k))$$

the rate of fragmentation of the kth block by π'.

We will say that an EFC process is non-degenerated if it has both a fragmentation and coalescence component, i.e. for each n there are some $\pi'_1 \neq \mathbf{1}_n$ and $\pi'_2 \neq \mathbf{0}_n$ such that $F_n(\mathbf{1}_n, \pi'_1, 1) > 0$ and $C_n(\mathbf{0}_n, \pi'_2) > 0$.

Of course the compatibility of the $\Pi_{|[m]}$ and the exchangeability requirement entail that not every family of transition rates is admissible. In fact, it is enough to know how $\Pi_{|[m]}$ leaves $\mathbf{1}_m$ and $\mathbf{0}_m$ for every $m \leq n$ to know all the rates $q_n(\pi, \pi')$.

Proposition 3.1. *There exists two families* $((C_n(\pi))_{\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}})_{n \in \mathbb{N}}$ *and* $((F_n(\pi))_{\pi \in \mathcal{P}_n \setminus \{\mathbf{1}_n\}})_{n \in \mathbb{N}}$ *such that for every* $m \leq n$, *for every* $\pi \in \mathcal{P}_n$ *such that* $\#\pi = m$

1. *For each* $\pi' \in \mathcal{P}_m \setminus \{\mathbf{0}_m\}$

$$q_n(\pi, Coag(\pi, \pi')) = C_n(\pi, \pi') = C_m(\pi').$$

2. *For each* $k \leq m$ *and for each* $\pi' \in \mathcal{P}_{|B_k|} \setminus \{\mathbf{1}_{|B_k|}\}$,

$$q_n(\pi, Frag(\pi, \pi', k)) = F_n(\pi, \pi', k) = F_{|B_k|}(\pi').$$

3. *All other transition rates are zero.*

Furthermore, these rates are exchangeable, i.e. for any permutation σ *of* $[n]$, *for all* $\pi \in \mathcal{P}_n$ *one has* $C_n(\pi) = C_n(\sigma(\pi))$ *and* $F_n(\pi) = F_n(\sigma(\pi))$.

Proof. The compatibility of the chains $\Pi_{|[n]}$ can be expressed in terms of transition rates as follows : For $m < n \in \mathbb{N}$ and $\pi, \pi' \in \mathcal{P}_n$ one has

$$q_m(\pi_{|[m]}, \pi'_{|[m]}) = \sum_{\pi'' \in \mathcal{P}_n : \pi''_{|[m]} = \pi'_{|[m]}} q_n(\pi, \pi'').$$

Consider $\pi \in \mathcal{P}_n$ such that $\pi = (B_1, B_2, ..., B_m, \emptyset, ...)$ has $m \leq n$ non-empty blocks. Call $w_i = \inf\{k \in B_i\}$ the least element of B_i and σ a permutation of $[n]$ that maps every $i \leq m$ on w_i. Let π' be an element of $\mathcal{P}_m$, then the restriction of the partition $\sigma(Coag(\pi, \pi'))$ to $[m]$ is given by : for $i, j \leq m$

$$\begin{aligned} i \overset{\sigma(Coag(\pi,\pi'))}{\sim} j &\Leftrightarrow \sigma(i) \overset{Coag(\pi,\pi')}{\sim} \sigma(j) \\ &\Leftrightarrow \exists k, l : \sigma(i) \in B_k, \sigma(j) \in B_l, k \overset{\pi'}{\sim} l \\ &\Leftrightarrow i \overset{\pi'}{\sim} j \end{aligned}$$

and hence

$$\sigma(Coag(\pi, \pi'))_{|[m]} = \pi'. \tag{3.1}$$

By definition $C_n(\pi,\pi')$ is the rate at which the process $\sigma(\Pi_{|[n]}(.))$ jumps from $\sigma(\pi)$ to $\sigma(Coag(\pi,\pi'))$. Hence, by exchangeability

$$C_n(\pi,\pi') = q_n(\sigma(\pi),\sigma(Coag(\pi,\pi'))).$$

Remark that $\sigma(\pi)_{|[m]} = \mathbf{0}_m$. Hence if π'' is a coalescence of $\sigma(\pi)$ it is completely determined by $\pi''_{|[m]}$. Thus, for all $\pi'' \in \mathcal{P}_n$ such that $\pi''_{|[m]} = \sigma(Coag(\pi,\pi'))_{|[m]}$ and $\pi'' \neq \sigma(Coag(\pi,\pi'))$ one has

$$q_n(\sigma(\pi),\pi'') = 0. \tag{3.2}$$

For each $\pi \in \mathcal{P}_n$ define

$$Q_n(\pi,m) := \{\pi' \in \mathcal{P}_n : \pi_{|[m]} = \pi'_{|[m]}\}$$

(for $\pi \in \mathcal{P}$ we will also need $Q(\pi,m) := \{\pi' \in \mathcal{P} : \pi_{|[m]} = \pi'_{|[m]}\}$). Clearly, (3.2) yields

$$q_n(\sigma(\pi),\sigma(Coag(\pi,\pi')) = \sum_{\pi'' \in Q_n(\sigma(Coag(\pi,\pi')),m)} q_n(\sigma(\pi),\pi'')$$

because there is only one non-zero term in the right hand-side sum. Finally recall (3.1) and use the compatibility relation to have

$$\begin{aligned}
C_n(\pi,\pi') &= q_n(\sigma(\pi),\sigma(Coag(\pi,\pi')) \\
&= \sum_{\pi'' \in Q_n(\sigma(Coag(\pi,\pi')),m)} q_n(\sigma(\pi),\pi'') \\
&= q_m(\sigma(\pi)_{|[m]},\sigma(Coag(\pi,\pi'))_{|[m]}) \\
&= q_m(\mathbf{0}_m,\pi') \\
&= C_m(\mathbf{0}_m,\pi') \\
&:= C_m(\pi').
\end{aligned}$$

Let us now take care of the fragmentation rates. The argument is essentially the same as above. Suppose $B_k = \{n_1,...,n_{|B_k|}\}$. Let σ be a permutation of $[n]$ such that of all $j \leq |B_k|$ one has $\sigma(j) = n_j$. Hence, in $\sigma(\pi)$ the first block is $[|B_k|]$. The process $\sigma(\Pi_{|[n]}(.))$ jumps from $\sigma(\pi)$ to the state $\sigma(Frag(\pi,\pi',k))$ with rate $F_n(\pi,\pi',k)$. Remark that for $i,j \leq |B_k|$

$$\begin{aligned}
i \overset{\sigma(Frag(\pi,\pi',k))}{\sim} j &\Leftrightarrow \sigma(i) \overset{Frag(\pi,\pi',k)}{\sim} \sigma(j) \\
&\Leftrightarrow n_i \overset{Frag(\pi,\pi',k)}{\sim} n_j \\
&\Leftrightarrow i \overset{\sigma(\pi')}{\sim} j
\end{aligned}$$

and hence

$$\sigma(Frag(\pi, \pi', k)) = Frag(\sigma(\pi), \sigma(\pi'), 1). \tag{3.3}$$

Thus by exchangeability $F_n(\pi, \pi', k) = F_n(\sigma(\pi), \sigma(\pi'), 1)$, and it is straightforward to see that by compatibility that

$$F_n(\sigma(\pi), \sigma(\pi'), 1) = F_{|B_k|}(\mathbf{1}_{|B_k|}, \sigma(\pi'), 1) = F_{|B_k|}(\sigma(\pi')).$$

The invariance of the rates $C_n(\mathbf{0}_n, \pi')$ and $F_n(\mathbf{1}_n, \pi', 1)$ by permutations of π' is also a direct consequence of exchangeability. In particular $F_{|B_k|}(\sigma(\pi')) = F_{|B_k|}(\pi')$ and thus we conclude that $F_n(\pi, \pi', k) = F_{|B_k|}(\pi')$.

□

Remark also that the only thing we impose on $\Pi(0)$ is that it should be exchangeable. This implies that the only possible deterministic starting points are $\mathbf{1}$ and $\mathbf{0}$ (because the measures $\delta_{\mathbf{1}}$ and $\delta_{\mathbf{0}}$ are exchangeable). If $\Pi(0) = \mathbf{0}$ we say that the process is started from dust, and if $\Pi(0) = \mathbf{1}$ we say it is started from unit mass.

Observe that, for n fixed, the finite families $(C_n(\pi))_{\pi \in \mathcal{P}_n \backslash \{\mathbf{0}_n\}}$ and $(F_n(\pi))_{\pi \in \mathcal{P}_n \backslash \{\mathbf{1}_n\}}$ may be seen as measures on $\mathcal{P}_n$. The compatibility of the $\Pi_{|[n]}(.)$ implies the same property for the (C_n, F_n), i.e., as measures, the restriction to $\mathcal{P}_n \backslash \{\mathbf{1}_n\}$ of (C_{n+1}) is C_n and the restriction to $\mathcal{P}_n \backslash \{\mathbf{0}_n\}$ of (F_{n+1}) is F_n. (see Lemma 1 in [3] for a precise demonstration in the case where there is only fragmentation ($C \equiv 0$), the general case being a simple extension). Hence, there exists a unique measure C and a unique measure F on $\mathcal{P}$ such that for each n

$$C_n = C_{|\mathcal{P}_n \backslash \{\mathbf{1}_n\}} \text{ and } F_n = F_{|\mathcal{P}_n \backslash \{\mathbf{0}_n\}}.$$

Furthermore, as we have remarked, the measures C_n and F_n are exchangeable. Hence, C and F are exchangeable measures. They must also verify some integrability conditions in order for the $\Pi_{|[n]}(.)$ to be Markov chains, i.e., to have a finite jump rate at any state. For $\pi \in \mathcal{P}$ recall that $Q(\pi, n) = \{\pi' \in \mathcal{P} : \pi'_{|[n]} = \pi_{|[n]}\}$. Then for each $n \in \mathbb{N}$ we must have

$$C(\mathcal{P} \backslash Q(\mathbf{0}, n)) < \infty$$

and

$$F(\mathcal{P} \backslash Q(\mathbf{1}, n)) < \infty.$$

It is clear that we can suppose without loss of generality that C and F assign no mass to the respective neutral elements for *Coag* and *Frag*, i.e. $C(\mathbf{0}) = 0$ and $F(\mathbf{1}) = 0$.

Here are three simple examples of exchangeable measures.

1. Let ϵ_n be the partition that has only two non empty blocks : $\mathbb{N}\backslash\{n\}$ and $\{n\}$. Then the (infinite) measure $\mathbf{e}(.) = \sum_{n\in\mathbb{N}} \delta_{\epsilon_n}(.)$ (where δ is the Dirac mass) is exchangeable. We call it the erosion measure .

2. For each $i \neq j \in \mathbb{N}$, call $\epsilon_{i,j}$ be the partition that has only one block which is not a singleton : $\{i, j\}$. Then the (infinite) measure $\kappa(.) = \sum_{i<j\in\mathbb{N}} \delta_{\epsilon_{i,j}}(.)$ is exchangeable. We call it the Kingman measure.

3. Take $x \in \mathcal{S}^{\downarrow} := \{1 \geq x_1 \geq x_2 \geq ... \geq 0; \sum_i x_i \leq 1\}$. Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of independent variables with respective law given by $P(X_i = k) = x_k$ for all $k \geq 1$ and $P(X_i = -i) = 1 - \sum_j x_j$. Define a random variable π with value in $\mathcal{P}$ by letting $i \overset{\pi}{\sim} j \Leftrightarrow X_i = X_j$. Following Kingman, we call π the x-paintbox process and denote by μ_x its distribution. Let ν be a measure on $\mathcal{S}^{\downarrow}$, then the mixture μ_ν of paintbox processes directed by ν, i.e.

$$\mu_\nu(A) = \int_{\mathcal{S}^{\downarrow}} \mu_x(A)\nu(dx),$$

is an exchangeable measure. We call it the ν-paintbox measure.

Extending seminal results of Kingman [10], Bertoin has shown in [3] and in his course at IHP that any exchangeable measure that verifies the required conditions is a combination of these three types. Hence the following proposition merely restates these results.

Proposition 3.2. *For each exchangeable measure C on $\mathcal{P}$ such that $C(\{\mathbf{0}\}) = 0$, and $\forall n \in \mathbb{N}, C(\{\pi \in \mathcal{P} : \pi_{|[n]} \neq \mathbf{0}_n\}) < \infty$ there exists a unique $c_k \geq 0$ and a unique measure ν_{Coag} on $\mathcal{S}^{\downarrow}$ such that*

$$\begin{gathered} \nu_{Coag}(\{\mathbf{0}\}) = 0, \\ \int_{\mathcal{S}^{\downarrow}} \left(\sum_{i=1}^{\infty} x_i^2\right) \nu_{Coag}(dx) < \infty, \\ \text{and } C = c_k\kappa + \mu_{\nu_{Coag}}. \end{gathered} \tag{3.4}$$

For each exchangeable measure F on $\mathcal{P}$ such that $F(\{\mathbf{1}\}) = 0$ and $F(\{\pi \in \mathcal{P} : \pi_{|[n]} \neq \mathbf{1}_n\}) < \infty, \forall n \in \mathbb{N}$ there exists a unique $c_e \geq 0$ and a unique

measure ν_{Disl} on $\mathcal{P}$ such that

$$\begin{aligned} \nu_{Disl}(\{(1,0,..)\}) &= 0, \\ \int_{\mathcal{S}^{\downarrow}} \left(1 - \sum_{i=1}^{\infty} x_i^2\right) \nu_{Disl}(dx) &< \infty \\ \text{and } F &= c_e \mathbf{e} + \mu_{\nu_{Disl}}. \end{aligned} \tag{3.5}$$

The two integrability conditions on ν_{Disl} and ν_{Coag} (3.5) and (3.4) ensure that $C(\mathcal{P}\backslash Q(\mathbf{0}, n)) < \infty$ and $F(\mathcal{P}\backslash Q(\mathbf{1}, n)) < \infty$. See [3] for the demonstration concerning F. The part that concerns C can be shown by the same arguments.

The condition on ν_{Disl} (3.5) may seem at first sight different from the condition that Bertoin imposes in [3] and which reads

$$\int_{\mathcal{S}^{\downarrow}} (1 - x_1) \nu_{Disl}(dx) < \infty$$

but they are in fact equivalent because

$$1 - \sum_i x_i^2 < 1 - x_1^2 < 2(1 - x_1)$$

and on the other hand

$$1 - \sum_i x_i^2 \geq 1 - x_1 \sum_i x_i \geq 1 - x_1.$$

Thus the above proposition implies that for each EFC process Π there is a unique exchangeable fragmentation $\Pi^{(F)}(t)$ and a unique exchangeable coalescence $\Pi^{(C)}(t)$ such that Π is a combination of $\Pi^{(F)}$ and $\Pi^{(C)}$. This was not self-evident because a compensation phenomena could have allowed weaker integrability conditions.

One can sum up the preceding analysis in the following characterization of exchangeable fragmentation-coalescence processes.

Proposition 3.3. *The distribution of an EFC process $\Pi(.)$ is completely characterized by the initial condition (i.e. the law of $\Pi(0)$), the measures ν_{Disl} and ν_{Coag} as above and the parameters $c_e, c_k \in \mathbb{R}+$.*

Remark : The above results are well known for pure fragmentation or pure coalescence. If, for instance, we impose that all transition rates $q_n(\pi, \pi')$ are 0 when π' is obtained of π by fragmentation (i.e. there is only coalescence

and no fragmentation, the EFC process is degenerated), the above proposition shows that our definition agrees with Definition 3 in Schweinsberg [15]. On the other hand if there is only fragmentation and no coalescence, our definition is equivalent with the one given by Bertoin in [3], which relies on some fundamental properties of the semi-group. There, the Markov chain property of the restrictions is deduced from the definition as well as the characterization of the distribution by c and ν_{Disl}.

Nevertheless, the formulation of Definition 3.1 is new. More precisely, it was not known that the exchangeability and Markov requirement for the restrictions to $[n]$ was enough to obtain a fragmentation procedure in which each fragment splits independently from the others (point 2 of Proposition 3.1).

As for exchangeable fragmentation or coalescence, one can construct EFC processes by using Poisson point processes (PPP in the following). More precisely let $P_C = ((t, \pi^{(C)}(t)), t \geq 0)$ and $P_F = ((t, \pi^{(F)}(t), k(t)), t \geq 0)$ be two independent PPP in the same filtration. The atoms of the PPP P_C are points in $\mathbb{R}^+ \times \mathcal{P}$ and its intensity measure is given by $dx \otimes (\mu_{\nu_{Coag}} + c_k \kappa)$. The atoms of P_F are points in $\mathbb{R}^+ \times \mathcal{P} \times \mathbb{N}$ and its intensity measure is $dx \otimes (c_e \mathbf{e} + \mu_{\nu_{Disl}}) \otimes \#$ where $\#$ is the counting measure on $\mathbb{N}$ and dx is the Lebesgue measure.

Take $\pi \in \mathcal{P}$ an exchangeable random variable and define a family of $\mathcal{P}_n$-valued processes $\Pi^n(.)$ as follows : for each n fix $\Pi^n(0) = \pi_{|[n]}$ and

- if t is not an atom time neither for P_C or P_F then $\Pi^n(t) = \Pi^n(t-)$,
- if t is an atom time for P_C such that $(\pi^{(C)}(t))_{|[n]} \neq \mathbf{0}_n$ then
$$\Pi^n(t) = \mathit{Coag}(\Pi^n(t-), \pi^{(C)}(t)),$$
- if t is an atom time for P_F such that $k(t) < n$ and $(\pi^{(F)}(t))_{|[n]} \neq \mathbf{1}_n$ then
$$\Pi^n(t) = \mathit{Frag}(\Pi^n(t-), \pi^{(F)}(t), k(t)).$$

Note that the Π^n are well defined because on any finite time interval, for each n, one only needs to consider a finite number of atoms. Furthermore P_C and P_F being independent in the same filtration, almost surely there is no t which is an atom time for both PPP's. This family is constructed to be compatible and thus defines uniquely a process Π such that $\Pi_{|[n]} = \Pi^n$ for each n. By analogy with homogeneous fragmentations ([3]) and homogeneous coalescence ([13, 15]) the following should be clear.

Proposition 3.4. *The process* Π *constructed above is an EFC process with characteristics* c_k, ν_{Coag}, c_e *and* ν_{Disl}.

Proof. It is straightforward to check that the restrictions $\Pi_{|[n]}(t)$ are Markov chains whose only jumps are either coagulations or fragmentations. The transition rates are constructed to correspond to the characteristics c_k, ν_{Coag}, c_e and ν_{Disl}. The only thing left to check is thus exchangeability. Fix $n \in \mathbb{N}$ and σ a permutation of $[n]$, then $(\sigma(\Pi^n(t)))_{t\geq 0}$ is a jump-hold Markov process. Its transition rates are given by $q_n^{(\sigma)}(\pi, \pi') = q_n(\sigma^{-1}(\pi), \sigma^{-1}(\pi'))$.

Suppose first that $\pi' = \mathit{Frag}(\pi, \pi'', k)$ for some π''. Remark that there exists a unique $l \leq \#\pi$ and a permutation σ' of $[m]$ (where $m = |\pi_k|$ is the cardinal of the k-th block of π we want to split) such that

$$\sigma^{-1}(\pi') = \mathit{Frag}(\sigma^{-1}(\pi), \sigma'(\pi''), l).$$

Using Proposition 3.1 we then obtain that

$$\begin{aligned} q_n^{(\sigma)}(\pi, \pi') &= q_n(\sigma^{-1}(\pi), \sigma^{-1}(\pi')) \\ &= q_n(\sigma^{-1}(\pi), \mathit{Frag}(\sigma^{-1}(\pi), \sigma'(\pi''), l)) \\ &= F_m(\sigma'(\pi'')) \\ &= F_m(\pi'') \\ &= q_n(\pi, \pi') \end{aligned}$$

The same type of arguments show that when $\pi' = \mathit{Coag}(\pi, \pi'')$ for some π'' we also have

$$q_n^{(\sigma)}(\pi, \pi') = q_n(\pi, \pi').$$

Thus, Π^n and $\sigma(\Pi^n)$ have the same transition rates and hence the same law.

As this is true for all n, it entails that Π and $\sigma(\Pi)$ also have the same law.

□

Let $\Pi(.)$ be an EFC process and define P_t its semi-group, i.e. for a continuous function $\phi : \mathcal{P} \to \mathbb{R}$

$$P_t\phi(\pi) := \mathbf{E}_\pi(\phi(\Pi(t)))$$

the expectation of $\phi(\Pi)$ at time t conditionally on $\Pi(0) = \pi$.

Corollary 3.1. *An EFC process* $\Pi(.)$ *has the Feller property, i.e.*

- *for each continuous function* $\phi : \mathcal{P} \to \mathbb{R}$, *for each* $\pi \in \mathcal{P}$ *one has*

$$\lim_{t\to 0+} P_t\phi(\pi) = \phi(\pi),$$

- *for all $t > 0$ the function $\pi \to P_t\phi(\pi)$ is continuous.*

Proof. Call C_f the set of functions

$$C_f = \{f : \mathcal{P} \to \mathbb{R} : \exists n \in \mathbb{N} \text{ s.t. } \pi_{|[n]} = \pi'_{|[n]} \Rightarrow f(\pi) = f(\pi')\}$$

which is dense in the space of continuous functions of $\mathcal{P} \to \mathbb{R}$. The first point is clear for a function $\Phi \in C_f$ (because the first jump-time of $\Phi(\Pi(.))$ is distributed as an exponential variable with finite mean). We conclude by density. For the second point, consider $\pi, \pi' \in \mathcal{P}$ such that $d(\pi, \pi') < 1/n$ (i.e., $\pi_{|[n]} = \pi'_{|[n]}$) then use the same PPP P_C and P_F to construct two EFC processes, $\Pi(.)$ and $\Pi'(.)$, with respective starting points $\Pi(0) = \pi$ and $\Pi'(0) = \pi'$. By construction $\Pi_{|[n]}(.) = \Pi'_{|[n]}(.)$ in the sense of the identity of the paths. Hence

$$\forall t \geq 0, d(\Pi(t), \Pi'(t)) < 1/n.$$

□

As a direct consequence, one also has the following characterization of EFC's in terms of the infinitesimal generator : Let $(\Pi(t), t \geq 0)$ be an EFC process, then the infinitesimal generator of Π, denoted by $\mathcal{A}$, acts on the functions $f \in C_f$ as follows :

$$\begin{aligned}\forall \pi \in \mathcal{P}, \mathcal{A}(f)(\pi) = \quad & \int_{\mathcal{P}} C(d\pi')(f(Coag(\pi, \pi')) - f(\pi)) \\ & + \sum_{k \in \mathbb{N}} \int_{\mathcal{P}} F(d\pi')(f(Frag(\pi, \pi', k)) - f(\pi)),\end{aligned}$$

where $F = c_e \mathbf{e} + \mu_{\nu_{Disl}}$ and $C = c_k \kappa + \mu_{\nu_{Coag}}$. Indeed, take $f \in C_f$ and n such that $\pi_{|[n]} = \pi'_{|[n]} \Rightarrow f(\pi) = f(\pi')$, then as $\Pi_{|[n]}(.)$ is a Markov chain the above formula is just the usual generator for Markov chains. The fact that this property characterizes EFC processes comes from the fact that it implies that the transition rates have the required properties.

3.4 Equilibrium measures

Consider an EFC process Π which is not trivial, i.e. $\nu_{Coag}, \nu_{Disl}, c_e$ and c_k are not zero simultaneously. Then the following is immediate :

Proposition 3.5. *There exists a unique (exchangeable) stationary probability measure ρ on $\mathcal{P}$. Furthermore one has*

$$\rho = \delta_{\mathbf{0}} \Leftrightarrow c_k = 0 \text{ and } \nu_{Coag} \equiv 0$$

and

$$\rho = \delta_{\mathbf{1}} \Leftrightarrow c_e = 0 \text{ and } \nu_{Disl} \equiv 0$$

where δ_π *is the Dirac mass at* π*.*

Furthermore, $\Pi(.)$ *converges in distribution to* ρ*.*

Proof. If the process Π is a pure coalescence process (i.e. $c_e = 0$ and $\nu_{Disl}(.) \equiv 0$) it is clear that $\mathbf{1}$ is an absorbing state towards which the process converges almost surely. In the pure fragmentation case it is $\mathbf{0}$ that is absorbing and attracting.

In the non-degenerated case, for each $n \in \mathbb{N}$, the process $\Pi_{|[n]}(.)$ is a finite state Markov chain. Let us now check the irreducibility in the non-degenerated case. Suppose first that $\nu_{Disl}(\mathcal{S}^\downarrow) > 0$. For every state $\pi \in \mathcal{P}_n$, if $\Pi_{|[n]}(t) = \pi$ there is a positive probability that the next jump of $\Pi_{|[n]}(t)$ is a coalescence. Hence, for every starting point $\Pi_{|[n]}(0) = \pi \in \mathcal{P}_n$ there is a positive probability that $\Pi_{|[n]}(.)$ reaches $\mathbf{1}_n$ in finite time T before any fragmentation has occurred. Now take $x \in \mathcal{S}^\downarrow$ such that $x_2 > 0$ and recall that μ_x is the x-paintbox distribution. Then for every $\pi \in \mathcal{P}_n$ such that $\#\pi = 2$ one has

$$\mu_x(Q(\pi, n)) > 0.$$

That is the n-restriction of the x-paintbox partition can be any partition of $[n]$ in two blocks with positive probability. More precisely if $\pi \in \mathcal{P}_n$ is such that $\pi = (B_1, B_2, \emptyset, \emptyset...)$ with $|B_1| = k$ and $|B_2| = n - k$ then

$$\mu_x(Q(\pi, n)) \geq x_1^k x_2^{n-k} + x_2^k x_1^{n-k}.$$

Hence, for any $\pi \in \mathcal{P}$ with $\#\pi = 2$, the first transition after T is $\mathbf{1}_n \to \pi$ with positive probability. As any $\pi \in \mathcal{P}_n$ can be obtained from $\mathbf{1}_n$ by a finite series of binary fragmentations we can iterate the above idea to see that with positive probability the jumps that follow T are exactly the sequence of binary splitting needed to get to π and the chain is hence irreducible.

Suppose now that $\nu_{Disl} \equiv 0$, there is only erosion $c_e > 0$, and that at least one of the following two condition holds

- for every $k \in \mathbb{N}$ one has $\nu_{Coag}(\{x \in \mathcal{S}^\downarrow : \sum_{i=1}^{i=k} x_i < 1\}) > 0$,
- there is a Kingman component, $c_k > 0$,

then almost the same demonstration applies. We first show that the state $\mathbf{0}_n$ can be reached from any starting point by a series of splittings corresponding

to erosion, and that from there any $\pi \in \mathcal{P}_n$ is reachable through binary coagulations.

In the remaining case (i.e., $c_k = 0, \nu_{Disl} \equiv 0$ and there exists $k > 0$ such that $\nu_{Coag}(\{x \in \mathcal{S}^{\downarrow} : \sum_{i=1}^{i=k} x_i < 1\}) = 0$) the situation is slightly different in that $\mathcal{P}_n$ is not the irreducible class. It is easily seen that the only partitions reachable from $\mathbf{0}_n$ are those with at most k non-singletons blocks. But for every starting point π one reaches this class in finite time almost surely. Hence there is no issues with the existence of an invariant measure for this type of $\Pi_{|[n]}$, it just does not charge partitions outside this class.

Thus there exists a unique stationary probability measure $\rho^{(n)}$ on $\mathcal{P}_n$ for the process $\Pi_{|[n]}$. Clearly by compatibility of the $\Pi_{|[n]}(.)$ one must have

$$\rho^{(n+1)}_{|\mathcal{P}_n}(.) = \rho^{(n)}(.).$$

This implies that there exists a unique probability measure ρ on $\mathcal{P}$ such that for each n one has $\rho^{(n)}(.) = \rho_{|\mathcal{P}_n}(.)$. The exchangeability of ρ is a simple consequence of the exchangeability of Π. Finally, the chain $\Pi_{|[2]}(.)$ is specified by two transition rates $\{1\}\{2\} \to \{1,2\}$ and $\{1,2\} \to \{1\}\{2\}$, which are both non-zero as soon as the EFC is non-degenerated. Hence,

$$\rho_{|\mathcal{P}_2} \notin \{\delta_{\mathbf{1}_2}(.), \delta_{\mathbf{0}_2}(.)\}.$$

Hence, when we have both coalescence and fragmentation $\rho \notin \{\delta_{\mathbf{1}}, \delta_{\mathbf{0}}\}$.

The $\Pi_{|[n]}(.)$ being finite states Markov chains, it is well known that they converge in distribution to $\rho^{(n)}$, independently of the initial state. By definition of the distribution of Π this implies that $\Pi(.)$ converges in distribution to ρ.

□

Although we cannot give an explicit expression for ρ in terms of c_k, ν_{Coag}, c_e and ν_{Disl}, we now relate certain properties of ρ to these parameters. In particular we will ask ourselves the following two natural questions :

- under what conditions does ρ charge only partitions with an infinite number of blocks, resp. a finite number of blocks, resp. both ?
- under what conditions does ρ charge partitions with dust (i.e. partitions such that $\sum_i \|B_i\| \leq 1$ where $\|B_i\|$ is the asymptotic frequency of block B_i) ?

3.4.1 Number of blocks

We first examine whether of not ρ charges partitions with a finite number of blocks.

Theorem 3.1. *Let $\Pi(.)$ be an EFC process with characteristics $\nu_{Coag}, \nu_{Disl}, c_k \geq 0$ and $c_e \geq 0$. Then*

$$\nu_{Disl}(\mathcal{S}^{\downarrow}) = \infty \text{ or } c_e > 0 \Rightarrow \rho(\{\pi \in \mathcal{P} : \#\pi < \infty\}) = 0.$$

Proof. We will prove that for each $K \in \mathbb{N}$ one has $\rho(\{\pi : \#\pi = K\}) = 0$.

Let us write the equilibrium equations for $\rho^{(n)}(.)$, the invariant measure of the Markov chain $\Pi_{|[n]}$. For each $\pi \in \mathcal{P}_n$

$$\rho^{(n)}(\pi) \sum_{\pi' \in \mathcal{P}_n \setminus \{\pi\}} q_n(\pi, \pi') = \sum_{\pi'' \in \mathcal{P}_n \setminus \{\pi\}} \rho^{(n)}(\pi'') q_n(\pi'', \pi)$$

where $q_n(\pi, \pi')$ is the rate at which $\Pi_{|[n]}$ jumps from π to π'. Fix $K \in \mathbb{N}$ and for each $n \geq K$, call $A_{n,K} := \{\pi \in \mathcal{P}_n : \#\pi \leq K\}$ and $D_{n,K} := \mathcal{P}_n \setminus A_{n,K}$ where $\#\pi$ is the number of non-empty blocks of π.

Summing over $A_{n,K}$ yields

$$\sum_{\pi \in A_{n,K}} \rho^{(n)}(\pi) \left[\sum_{\pi' \in A_{n,K} \setminus \{\pi\}} q_n(\pi, \pi') + \sum_{\pi \in D_{n,K}} q_n(\pi, \pi') \right]$$

$$= \sum_{\pi \in A_{n,K}} \left[\sum_{\pi'' \in A_{n,K} \setminus \{\pi\}} \rho^{(n)}(\pi'') q_n(\pi'', \pi) + \sum_{\pi \in D_{n,K}} \rho^{(n)}(\pi'') q_n(\pi'', \pi) \right]$$

but as

$$\sum_{\pi \in A_{n,K}} \rho^{(n)}(\pi) \left[\sum_{\pi' \in A_{n,K} \setminus \{\pi\}} q_n(\pi, \pi') \right] = \sum_{\pi \in A_{n,K}} \left[\sum_{\pi'' \in A_{n,K} \setminus \{\pi\}} \rho^{(n)}(\pi'') q_n(\pi'', \pi) \right]$$

one has

$$\sum_{\pi \in A_{n,K}} \rho^{(n)}(\pi) \left[\sum_{\pi \in D_{n,K}} q_n(\pi, \pi') \right] = \sum_{\pi \in A_{n,K}} \left[\sum_{\pi \in D_{n,K}} \rho^{(n)}(\pi'') q_n(\pi'', \pi) \right].$$

That is, if we define $q_n(\pi, C) = \sum_{\pi' \in C} q_n(\pi, \pi')$ for each $C \subseteq \mathcal{P}_n$,

$$\sum_{\pi \in A_{n,K}} \rho^{(n)}(\pi) q_n(\pi, D_{n,K}) = \sum_{\pi'' \in D_{n,K}} \rho^{(n)}(\pi'') q_n(\pi'', A_{n,K}). \tag{3.6}$$

Therefore

$$\sum_{\pi\in A_{n,K}\setminus A_{n,K-1}} \rho^{(n)}(\pi)q_n(\pi, D_{n,K}) \leq \sum_{\pi''\in D_{n,K}} \rho^{(n)}(\pi'')q_n(\pi'', A_{n,K}). \quad (3.7)$$

Hence, all we need to prove that $\rho(\{\pi \in \mathcal{P} : \#\pi_{|[n]} = K\}) \to 0$ when $n \to \infty$ is to give an upper bound for the right hand-side of (3.7) which is uniform in n and to show that

$$\min_{\pi\in A_{n,K}\setminus A_{n,K-1}} q_n(\pi, D_{n,K}) \underset{n\to\infty}{\to} \infty. \quad (3.8)$$

Let us begin with (3.8). Define

$$\Phi(q) := c_e(q+1) + \int_{\mathcal{S}^\downarrow} (1 - \sum_i x_i^{q+1})\nu_{Disl}(dx).$$

This function was introduced by Bertoin in [3], where it plays a crucial role as the Laplace exponent of a subordinator; in particular, Φ is a concave increasing function. When k is an integer greater or equal than 2, $\Phi(k-1)$ is the rate at which $\{[k]\}$ splits, i.e., it is the arrival rate of atoms $(\pi^{(F)}(t), k(t), t)$ of P_F such that $\pi^{(F)}_{|[k]}(t) \neq \mathbf{1}_k$ and $k(t) = 1$. More precisely $c_e k$ is the rate of arrival of atoms that correspond to erosion and $\int_{\mathcal{S}^\downarrow}(1 - \sum_i x_i^k)\nu_{Disl}(dx)$ is the rate of arrival of dislocations. Hence, for $\pi \in \mathcal{P}_n$ such that $\#\pi = K$, say $\pi = (B_1, B_2, ..., B_K, \emptyset, \emptyset, ...)$, one has

$$q_n(\pi, D_{n,K}) = \sum_{i:|B_i|>1} \Phi(|B_i| - 1)$$

because it only takes a fragmentation that creates at least one new block to enter $D_{n,K}$.

First remark that

$$\sum_{i:|B_i|>1} c_e|B_i| \geq c_e(n - K + 1),$$

next note that

$$q \to \int_{\mathcal{S}^\downarrow} (1 - \sum_i x_i^{q+1})\nu_{Disl}(dx)$$

is also concave and increasing for the same reason that Φ is and furthermore

$$\int_{\mathcal{S}^\downarrow} (1 - \sum_i x_i)\nu_{Disl}(dx) \geq 0.$$

Hence, for every $(B_1, ..., B_K) \in \mathcal{P}_n$ one has the lower bound

$$q_n(\pi, D_{n,K}) = \sum_{i:|B_i|>1} \Phi(|B_i|-1) \geq \int_{\mathcal{S}^\downarrow} (1-\sum_i x_i^{(n-K)+1})\nu_{Disl}(dx)+c_e(n-K+1).$$

As $\Phi(x) \underset{x\to\infty}{\to} \infty \Leftrightarrow \nu_{Disl}(\mathcal{S}^\downarrow) = \infty$ or $c_e > 0$ one has

$$c_e > 0 \text{ or } \nu_{Disl}(\mathcal{S}^\downarrow) = \infty \Rightarrow \lim_{n\to\infty} \min_{\pi:\#\pi=K} q_n(\pi, D_{n,K}) = \infty.$$

On the other hand it is clear that $q_n(\pi, A_{n,K})$ only depends on $\#\pi$ and K (by definition the precise state π and n play no role in this rate). By compatibility it is easy to see that if π, π' are such that $\#\pi' > \#\pi = K$ then

$$q_n(\pi, A_{n,K}) \geq q_n(\pi', A_{n,K}).$$

Hence, for all $\pi \in D_{n,K}$ one has

$$q_n(\pi, A_{n,K}) \leq \tau_K$$

where $\tau_K = q_n(\pi', A_{n,K})$ for all n and any $\pi' \in \mathcal{P}_n$ such that $\#\pi' = K+1$, and hence τ_K is a constant that only depends on K.

Therefore

$$\begin{aligned}
\min_{\pi\in\mathcal{P}_n:\#\pi=K} q_n(\pi, D_{n,K}) \sum_{\pi\in\mathcal{P}_n:\#\pi=K} \rho^{(n)}(\pi) &\leq \sum_{\pi\in\mathcal{P}_n:\#\pi=K} \rho^{(n)}(\pi) q_n(\pi, D_{n,K}) \\
&\leq \sum_{\pi''\in D_{n,K}} \rho^{(n)}(\pi'') q_n(\pi'', A_{n,K}) \\
&\leq \tau_K \sum_{\pi''\in D_{n,K}} \rho^{(n)}(\pi''),
\end{aligned}$$

where, on the second inequality, we used (3.7). Thus

$$\rho^{(n)}(\{\pi \in \mathcal{P}_n : \#\pi = K\}) \leq \tau_K / \min_{\pi\in\mathcal{P}_n:\#\pi=K} q_n(\pi, D_{n,K}).$$

This show that for each $K \in \mathbb{N}$, one has $\lim_{n\to\infty} \rho^{(n)}(\{\pi \in \mathcal{P}_n : \#\pi = K\}) = 0$ and thus $\rho(\#\pi < \infty) = 0$. □

Although we don't have a converse to Theorem 3.1 in the general case, we know that when there is no erosion, the fragmentation is binary (i.e., each splitting produces exactly two fragments) and there is a Kingman component in the coalescence, then the following holds :

Proposition 3.6. *If* $c_e = 0$, $\nu_{Disl}(\mathcal{S}^{\downarrow}) < \infty$ *and*

$$\nu_{Disl}(\{x \in \mathcal{S}^{\downarrow} : x_1 + x_2 < 1\}) = 0$$

(the fragmentation component is binary), then

$$c_k > 0 \Rightarrow \rho(\{\pi \in \mathcal{P} : \#\pi < \infty) = 1.$$

Proof. For each $n \in \mathbb{N}$ we define the sequence $(a_i^{(n)})_{i\in\mathbb{N}}$ by

$$a_i^{(n)} := \rho^{(n)}(A_{n,i} \backslash A_{n,i-1}) = \rho^{(n)}(\{\pi \in \mathcal{P}_n : \#\pi = i\}).$$

We also note $p := \nu_{Disl}(\mathcal{S}^{\downarrow})$ the total rate of fragmentation. The equation (3.6) becomes for each $K \in [n]$

$$\sum_{\pi:\#\pi=K} \rho^{(n)}(\{\pi\}) q_n(\pi, D_{n,K}) = \sum_{\pi'' \in D_{n,K}} \rho^{(n)}(\{\pi''\}) q_n(\pi'', A_{n,K}) \tag{3.9}$$

because the fragmentation is binary. When $\#\pi = K$ one has $q_n(\pi, D_{n,K}) \leq Kp$, thus

$$\begin{aligned} a_K^{(n)} Kp &\geq \sum_{\pi'' \in D_{n,K}} \rho^{(n)}(\{\pi''\}) q_n(\pi'', A_{n,K}) \\ &\geq \sum_{\pi'':\#\pi''=K+1} \rho^{(n)}(\{\pi''\}) q_n(\pi'', A_{n,K}) \\ &\geq \sum_{\pi'':\#\pi''=K+1} \rho^{(n)}(\{\pi''\}) c_k K(K+1)/2 \\ &\geq a_{K+1}^{(n)} c_k K(K+1)/2. \end{aligned} \tag{3.10}$$

Hence for all $K \in [n-1]$

$$a_K^{(n)} p \geq a_{K+1}^{(n)} c_k (K+1)/2$$

and thus

$$1 = \sum_{i=1}^{n} a_i^{(n)} < a_1^{(n)} (1 + \sum_{i=1}^{n-1} (p/c_k)^i 2^{i-1}/i!).$$

We conclude that $a_1^{(n)}$ is uniformly bounded from below by $(1 +$

$\sum_{i=1}^{\infty}(p/c_k)^i 2^{i-1}/i!)^{-1}$. On the other hand, as $a_1^{(n)} \leq 1$ one has

$$\begin{aligned}\lim_{n\to\infty}\sum_{i>K}^{n} a_i^{(n)} &\leq \lim_{n\to\infty} a_1^{(n)} \sum_{i>K-1}^{n-1} \frac{(2p/c_k)^i}{2i!} \\ &\leq \sum_{i>K-1}^{n-1} \frac{(2p/c_k)^i}{2i!} \\ &\leq \sum_{i>K-1}^{\infty} \frac{(2p/c_k)^i}{2i!} \to 0\end{aligned}$$

when $K \to \infty$. Hence if we define $a_i := \lim_{n\to\infty} a_i^{(n)} = \rho(\{\pi \in \mathcal{P} : \#\pi = i\})$ we have proved that the series $\sum_i a_i$ is convergent and hence $\lim_{K\to\infty}\sum_{i>K} a_i = 0$. This shows that $\rho(\{\pi \in \mathcal{P} : \#\pi = \infty\}) = 0$.

□

Thus for an EFC process with a binary fragmentation component, a Kingman coalescence component and no erosion (i.e. $c_k > 0, c_e = 0$ and $\nu_{Disl}(\{x \in \mathcal{S}^{\downarrow} : x_1 + x_2 < 1\}) = 0$) we have the equivalence

$$\rho(\{\pi \in \mathcal{P} : \#\pi = \infty\}) = 1 \Leftrightarrow \nu_{Disl}(\mathcal{S}^{\downarrow}) = \infty$$

and when $\nu_{Disl}(\mathcal{S}^{\downarrow}) < \infty$ then $\rho(\{\pi \in \mathcal{P} : \#\pi = \infty\}) = 0$.

3.4.2 Dust

For any fixed time t the partition $\Pi(t)$ is exchangeable. Hence, by Kingman's theory of exchangeable partition, its law is a mixture of paintbox processes. A direct consequence is that every block $B_i(t)$ of $\Pi(t)$ is either a singleton or an infinite block with strictly positive asymptotic frequency. Recall that the asymptotic frequency of a block $B_i(t)$ is given by

$$\|B_i(t)\| = \lim_{n\to\infty} \frac{1}{n}\text{Card}\{k \leq n : k \in B_i(t)\}$$

so part of Kingman's result is that this limit exists almost surely for all i simultaneously. The asymptotic frequency of a block corresponds to its mass, thus singletons have zero mass, they form what we call dust. More precisely, for $\pi \in \mathcal{P}$ define the set

$$\text{dust}(\pi) := \bigcup_{j:\|B_j\|=0} B_j.$$

When π is exchangeable we have almost surely

$$\text{dust}(\pi) = \{i \in \mathbb{N} : \exists j \text{ s.t. } \{i\} = B_j\}$$

and

$$\sum_i \|B_i\| + \|\text{dust}(\pi)\| = 1.$$

For fragmentation or EFC processes, dust can be created *via* two mechanisms : either from erosion (that's the atoms that correspond to the erosion measure $c_e\mathbf{e}$ when $c_e > 0$), or from sudden splitting which corresponds to atoms associated to the measure $\mu_{\nu'_{Disl}}$ where ν'_{Disl} is simply ν_{Disl} restricted to $\{s \in \mathcal{S}^\downarrow : \sum_i s_i < 1\}$. Conversely, in the coalescence context mass can condensate out of dust, thus giving an entrance law in $\mathcal{S}^\downarrow$, see [13].

More precisely if one considers a standard coalescent $\Pi(.)$, started from the partition $\mathbf{0}$, then the asymptotic frequencies process $\Lambda(\Pi(.))$ exists, lives in $\mathcal{S}^\downarrow$ and $\Lambda(\Pi(0)) = (0, 0, ..)$. If for instance Π is a Kingman coalescent, almost surely for every $t > 0$, one has $\Lambda(\Pi(t)) \in \{x \in \mathcal{S}^\downarrow : \exists k \text{ s.t. } \sum_1^k x_i = 1\}$ (see [13]). Thus, almost surely for every $t > 0$ one has $\|B_1(t)\| > 0$. In fact, for every positive time $t > 0$, there are only a finite number of non empty blocks whose asymptotic frequencies sum up to 1.

The following theorem states that when the coalescence is strong enough in an EFC process, the equilibrium measure does not charge partitions with dust.

Theorem 3.2. *Let $(\Pi(t), t \geq 0)$ be an EFC process and ρ its invariant probability measure.*

Then

$$\int_{\mathcal{S}^\downarrow} \left(\sum_i x_i\right) \nu_{Coag}(dx) = \infty \text{ or } c_k > 0 \Rightarrow \rho(\{\pi \in \mathcal{P} : dust(\pi) \neq \emptyset\}) = 0.$$

In case of no fragmentation, this follows from proposition 30 in [15].

Proof. Define $I_n := \{\pi = (B_1, B_2, ...) \in \mathcal{P} : B_1 \cap [n] = \{1\}\}$ (when no confusion is possible we sometime use $I_n := \{\pi \in \mathcal{P}_n : B_1 = \{1\}\}$) i.e. the partitions of $\mathbb{N}$ such that the only element of their first block in $[n]$ is $\{1\}$. Our proof relies on the fact that

$$\rho(\{\pi \in \mathcal{P} : \text{dust}(\pi) \neq \emptyset\}) > 0 \Rightarrow \rho(\{\pi \in \mathcal{P} : \pi \in \cap_n I_n\}) > 0.$$

As above let us write down the equilibrium equations for $\Pi_{|[n]}(.)$:

$$\sum_{\pi\in\mathcal{P}_n\cap I_n}\rho^{(n)}(\pi)q_n(\pi,I_n^c)=\sum_{\pi'\in I_n^c}\rho^{(n)}(\pi')q_n(\pi,I_n).$$

Recall that $A_{n,b}$ designates the set of partitions $\pi\in\mathcal{P}_n$ such that $\#\pi\leq b$ and $D_{n,b}=\mathcal{P}_n\backslash A_{n,b}$. For each b remark that

$$\min_{\pi\in D_{n,b}\cap I_n}\{q_n(\pi,I_n^c)\}=q_n(\pi',I_n^c)$$

where π' can be any partition in $\mathcal{P}_n$ such that $\pi'\in I_n$ and $\#\pi'=b+1$. We can thus define

$$f(b):=\min_{\pi\in D_{n,b}\cap I_n}\{q_n(\pi,I_n^c)\}.$$

If $c_k>0$ and $\pi\in D_{n,b}\cap I_n$ one can exit from I_n by a coalescence of the Kingman type. This happens with rate greater than c_kb. If $\nu_{Coag}(\mathcal{S}^{\downarrow})>0$ one can also exit *via* a coalescence with multiple collision, and this happens with rate greater than

$$\zeta(b):=\int_{\mathcal{S}^{\downarrow}}\left(\sum_i x_i\left(1-(1-x_i)^{b-1}\right)\right)\nu_{Coag}(dx).$$

This $\zeta(b)$ is the rate of arrival of atoms $\pi^{(C)}(t)$ of P_C such that $\pi^{(C)}(t)\notin I_b$ and which do no correspond to a Kingman coalescence. Thus $\sup_{b\in\mathbb{N}}\zeta(b)$ is the rate of arrival of "non-Kingman" atoms $\pi^{(C)}(t)$ of P_C such that $\pi^{(C)}(t)\notin I:=\cap_n I_n$. This rate being $\int_{\mathcal{S}^{\downarrow}}\left(\sum_i x_i\right)\nu_{Coag}(dx)$ and $\zeta(b)$ being an increasing sequence one has

$$\lim_{b\to\infty}\zeta(b)=\int_{\mathcal{S}^{\downarrow}}\left(\sum_i x_i\right)\nu_{Coag}(dx).$$

Thus it is clear that, under the conditions of the proposition, $f(b)\to\infty$ when $b\to\infty$.

On the other hand, when $\pi\in I_n^c$, the rate $q_n(\pi,I_n)$ is the speed at which 1 is isolated from all the other points, thus by compatibility it is not hard to see that

$$q_2:=\int_{\mathcal{S}^{\downarrow}}(1-\sum_i x_i^2)\nu_{Disl}(dx)\geq q_n(\pi,I_n)$$

where q_2 is the rate at which 1 is isolated from its first neighbor (the inequality comes from the inclusion of events).

Hence,

$$\begin{aligned}\sum_{\pi\in I_n\cap D_{n,b}}\rho^{(n)}(\pi)f(b) &\leq \sum_{\pi\in I_n\cap D_{n,b}}\rho^{(n)}(\pi)q_n(\pi,I_n^c)\\ &\leq \sum_{\pi'\in I_n^c}\rho^{(n)}(\pi')q_n(\pi,I_n)\\ &\leq \sum_{\pi'\in I_n^c}\rho^{(n)}(\pi')q_2\\ &\leq q_2\end{aligned}$$

which yields

$$\rho^{(n)}(I_n\cap D_{n,b})\leq q_2/f(b).$$

Now as ρ is exchangeable one has $\rho(I\cap A_b)=0$ where $I=\cap_n I_n$ and $A_b=\cap_n A_{n,b}$ (exchangeable partitions who have dust have an infinite number of singletons, and thus cannot have a finite number of blocks). Hence $\rho^{(n)}(I_n\cap A_{n,b})\to 0$.

Fix $\epsilon>0$ arbitrarily small and choose b such that $q_2/f(b)\leq\epsilon/2$. Then choose n_0 such that for all $n\geq n_0,\rho^{(n)}(I_n\cap A_{n,b})\leq\epsilon/2$. Hence

$$\forall n\geq n_0:\rho^{(n)}(I_n)=\rho^{(n)}(I_n\cap A_{n,b})+\rho^{(n)}(I_n\cap D_{n,b})\leq\epsilon/2+\epsilon/2.$$

Thus $\lim_{n\to\infty}\rho^{(n)}(I_n)=0$ which entails $\rho(B_1=\{1\})=0$. Now we use the following fact :

$$\rho(B_1=\{1\})=\int_{\mathcal{P}}(1-\sum_i\|B_i\|)\rho(d\pi)$$

to see that $\rho(\text{dust}(\pi)\neq\emptyset)=0$. □

3.5 Path properties

3.5.1 Number of blocks along the path.

One of the problem tackled by Pitman [13] and Schweinsberg [16, 15]about coalescent processes is whether or not they come down from infinity. Let us first recall some of their results. By definition if $\Pi^C(.)$ is a standard coalescent $\Pi^C(0)=\mathbf{0}$ and thus $\#\Pi^C(0)=\infty$. We say that Π^C comes down from infinity if $\#\Pi^C(t)<\infty$ a.s. for all $t>0$. We say it stays infinite if $\#\Pi^C(t)=\infty$ a.s. for all $t>0$.

Define $\Delta_f = \{x \in \mathcal{S}^{\downarrow} : \exists i \in \mathbb{N} \text{ s.t. } \sum_{j=1}^{i} x_j = 1\}$. We know by Lemma 31 in [16], which is a generalization of Proposition 23 in [13], that if $\nu_{Coag}(\Delta_f) = 0$ the coalescent either stays infinite or comes down from infinity.

For $b \geq 2$ let λ_b denote the total rate of all collisions when the coalescent has b blocks

$$\lambda_b = \mu_{\nu_{Coag}}(\mathcal{P} \backslash Q(\mathbf{1}, b)) + c_k \frac{b(b-1)}{2}.$$

Let γ_b be the total rate at which the number of blocks is decreasing when the coalescent has b blocks,

$$\gamma_b = c_k \frac{b(b-1)}{2} + \sum_{k=1}^{b-1} (b-k) \nu_{Coag}(\{\pi : \#\pi_{|[b]} = k\}).$$

If $\nu_{Coag}(\Delta_f) = \infty$ or $\sum_{b=2}^{\infty} \gamma_b^{-1} < \infty$, then the coalescent comes down from infinity. The converse is not always true but holds for instance for the important case of the Λ-coalescents (i.e., those for which many fragments can merge into a single block, but only one such merger can occur simultaneously).

This type of properties concerns the paths of the processes, and it seems that they bear no simple relations with properties of the equilibrium measure. For instance the equilibrium measure of a coalescent that stays infinite is $\delta_{\mathbf{1}}(.)$ and therefore only charges partitions with one block, but its path lays entirely in the subspace of $\mathcal{P}$ of partitions with an infinite number of blocks.

Let $\Pi(.)$ be an EFC process. Define the sets

$$G := \{t \geq 0 : \#\Pi(t) = \infty\}$$

and

$$\forall k \in \mathbb{N}, G_k := \{t \geq 0 : \#\Pi(t) > k\}.$$

Clearly every arrival time t of an atom of P_C such that $\pi^{(C)}(t) \in \Delta_f$ is in G^c the complementary of G. In the same way an arrival time t of an atom of P_F such that $\pi^{(F)}(t) \in \mathcal{S}^{\downarrow} \backslash \Delta_f$ and $B_{k(t)}(t-)$ (the fragmented block) is infinite immediately before the fragmentation, must be in G. Hence, if $\nu_{Disl}(\mathcal{S}^{\downarrow} \backslash \Delta_f) = \infty$ and $\nu_{Coag}(\Delta_f) = \infty$, then both G and G^c are everywhere dense, and this independently of the starting point which may be $\mathbf{1}$ or $\mathbf{0}$.

The following proposition shows that when the fragmentation rate is infinite, G is everywhere dense. Recall the notation $\Pi(t) = (B_1(t), B_2, (t), ...)$.

Theorem 3.3. *Let Π be an EFC process such that $c_e > 0$ or $\nu_{Disl}(\mathcal{S}^{\downarrow}) = \infty$. Then, a.s. G is everywhere dense.*

As $G = \cap G_k$ we only need to show that a.s. for each $k \in \mathbb{N}$ the set G_k is everywhere dense and open to conclude with Baire theorem. The proof relies on two lemmas.

Lemma 3.1. *Let Π be an EFC process started from $\mathbf{1}$ and such that $c_e > 0$ or $\nu_{Disl}(\mathcal{S}^{\downarrow}) = \infty$. Then, a.s. for all $k \in \mathbb{N}$*

$$\inf\{t \geq 0 : \#\Pi(t) > k\} = 0.$$

Proof. Fix $k \in \mathbb{N}$ and $\epsilon > 0$, we are going to show that there exists $t \in [0, \epsilon[$ such that

$$\exists n \in \mathbb{N} : \#\Pi_{|[n]}(t) \geq k.$$

As $\nu_{Disl}(\mathcal{S}^{\downarrow}) = \infty$ (or $c_e > 0$) it is clear that almost surely $\exists n_1 \in \mathbb{N} : \exists t_1 \in [0, \epsilon[$ such that $\Pi_{|[n_1]}(t_1-) = \mathbf{1}_{|[n_1]}$ and t_1 is a fragmentation time such that $\Pi_{|[n_1]}(t_1)$ contains at least two blocks, say $B_{i_1}(t_1) \cap [n_1]$ and $B_{i_2}(t_1) \cap [n_1]$, of which at least one is not a singleton and is thus in fact infinite when seen in $\mathbb{N}$. The time of coalescence of i_1 and i_2 (i.e. the first time at which they are in the same block again is exponentially distributed with parameter

$$\int_{\mathcal{S}^{\downarrow}} (\sum_i x_i^2) \nu_{Coag}(dx) + c_k < \infty.$$

Hence if we define

$$\tau_{i_1,i_2}(t_1) := \inf\{t \geq t_1 : i_1 \overset{\Pi(t)}{\sim} i_2\}$$

and if we call $B(i,t)$ the block of $\Pi(t)$ that contains i then almost surely we can find $n_2 > n_1$ large enough such that the first time t_2 of fragmentation of $B(i_1, t_1) \cap [n_2]$ or $B(i_2, t_1) \cap [n_2]$ is smaller than $\tau_{i_1,i_2}(t_1)$ (i.e. i_1 and i_2 have not coalesced yet) and t_2 is a fragmentation time at which $B(i_1, t_2-) \cap [n_2]$ or $B(i_2, t_2-) \cap [n_2]$ is split into two blocks. Hence at t_2 there are at least 3 non-empty blocks in $\Pi_{|[n_2]}(t_2)$, and at least one of them is not a singleton. By iteration, almost surely, $\exists n_k : \exists t_k \in [0, \epsilon[$ such that t_k is a fragmentation time and

$$\#\Pi_{|[n_k]}(t_k) \geq k.$$

□

Lemma 3.2. *Let Π be an EFC process such that $c_e > 0$ or $\nu_{Disl}(\mathcal{S}^{\downarrow}) = \infty$. Then, a.s. G_k is everywhere dense and open for each $k \in \mathbb{N}$.*

Proof. Fix $k \in \mathbb{N}$, call $\Gamma_k = \{t_1^{(k)} < t_2^{(k)} < ...\}$ the collection of atom times of P_C such that a coalescence occurs on the $k+1$ first blocks if there are more than $k+1$ blocks, i.e.,

$$\pi^{(C)}(t) \notin Q(\mathbf{0}, k+1)$$

(recall that $Q(\mathbf{0}, k+1) = \{\pi \in \mathcal{P} : \pi_{|[k+1]} = \mathbf{0}_{k+1}\}$). Suppose $t \in G_k$, then by construction $\inf\{s > t : s \in G_k^c\} \in \Gamma_k$ (because one must at least coalesce the first $k+1$ distinct blocks present at time t before having less than k blocks. As the $t_i^{(k)}$ are stopping times, the strong Markov property and the first lemma imply that $G_k^c \subseteq \Gamma_k$. Hence G_k is a dense open subset of $\mathbb{R}+$. □

We can apply Baire's theorem to conclude that $\cap_k G_k = G$ is almost surely everywhere dense in $\mathbb{R}+$.

As a corollary, we see that when the coalescence is "mostly" Kingman (i.e. $c_k > 0, \nu_{Coag}(\mathcal{S}^{\downarrow}) < \infty$) and the fragmentation rate is infinite ($c_e > 0$ or $\nu_{Disl}(\mathcal{S}^{\downarrow}) = \infty$), then we have that the sets of times G^c are exactly the atom times for P_C such that $\pi^{(C)}(.) \in \Delta_f$. Define $\Delta_f(k) := \{x \in \Delta_f : \sum_1^k x_i = 1\}$.

Corollary 3.2. *When $c_k \geq 0, \nu_{Coag}(\mathcal{S}^{\downarrow}) < \infty$ and ($c_e > 0$ or $\nu_{Disl}(\mathcal{S}^{\downarrow}) = \infty$) one has*

$$G^c = \{t : \pi^{(C)}(t) \in \Delta_f\} \tag{3.11}$$

and for all $n \geq 2$

$$G_n^c = \{t : \pi^{(C)}(t) \in \Delta_f(n-1)\}. \tag{3.12}$$

Proof. As $G^c = \cup G_n^c$, it suffices to show (3.12) for some $n \in \mathbb{N}$. Call $(t_i^{(k)})_{i\in\mathbb{N}}$ the sequence of of atom times of P_C which correspond to coalescence such that $\pi^{(c)}(t_i^{(k)}) \notin Q(\mathbf{0}, k)$ for all i. The preceding theorem implies that almost surely there exists a decreasing sequence $\tau_j \searrow 0$ such that $\#\Pi(\tau_j) = \infty$. Fix j, then at time τ_j call $b_1(\tau_j), b_2(\tau_j), ..., b_k(\tau_j)$ the least elements of the k first blocks. By taking j large enough we can always suppose $\tau_j < t_1^{(k)}$, and it is then clear that for all $s \in \left[\tau_j, t_1^{(k)}\right[$ one has $\#\Pi_{|[a_k(\tau_j)]}(s) \geq k$. Thus, one can easily see that $G_k^c \subseteq \{t_i^{(k)}, i \in \mathbb{N}\}$.

Now consider simultaneously the sequence $(t_i^{(k)})_{i\in\mathbb{N}}$ and $(t_i^{(k+1)})_{i\in\mathbb{N}}$ and call

$$t_{m(1)}^{(k+1)} = \max\{t_i^{(k+1)} \leq t_1^{(k)}\}.$$

One always has $t_{m(1)}^{(k+1)} = t_1^{(k)}$, thus there exists $r_1^{(k)} \in]t_{m(1)-1}^{(k+1)}, t_1^{(k)}[$ and $n_1 < \infty$ such that for all $s \in]r_1^{(k)}, t_1^{(k)}[$: $\#\Pi_{[n_1]}(s) \geq k+1$. Hence, a necessary condition

to have $\#\Pi_{|[n_1]}(t_1^{(k)}) < k$ is that $t_1^{(k)}$ is a multiple collision time. Hence

$$G_k^c \subseteq \{t_i^{(k)} \text{ s.t. } t_i^{(k)} \text{ is a multiple collision time }\}.$$

Fix $n \in \mathbb{N}$ and consider the first coalescence time τ such that $\pi^{(c)}(\tau) \notin Q(\mathbf{0}, k)$ (hence $\tau \in (t_i^{(k)})_{i\in\mathbb{N}}$) and $\#\pi^{(c)}_{|[n]}(\tau) \geq k$. Then it can be shown by the same arguments as above that there exists $m \in \mathbb{N}$ such that $\#\Pi_{|[m]}(\tau-) \geq n$ and hence $\#\Pi_{|[m]}(\tau) \geq k$.

Call $(T_i^{(k)})$ the subsequence of $(t_i^{(k)})$ that corresponds to the multiple collision times such that $\pi^{(c)}(T_i^{(k)}) \notin \Delta_f(k)$. Then it should be clear that

$$\{T_i^{(k)}\} = \cup_{n\in\mathbb{N}}\{t : \pi^{(c)}(t) \notin Q(\mathbf{0}, k) \text{ and } \#\pi^{(c)}_{|[k+n]}(t) > k\}.$$

Hence the only elements t of $\{t_i^{(k)}, i \in \mathbb{N}\}$ for which $\Pi(t)$ can be in G_k^c are those who corresponds to coalescence time such that $\pi^{(c)}(t) \in \Delta_f(k-1)$. □

As recently noted by Lambert [11], there is an interpretation of some EFC processes in terms of population dynamics. More precisely if we consider an EFC process $(\Pi(t), t \geq 0)$ such that $\nu_{Disl}(\mathcal{S}^{\downarrow}) < \infty$ and

$$(H) \qquad \begin{cases} \nu_{Disl}(\mathcal{S}^{\downarrow}\backslash\Delta_f) = 0 \\ c_e = 0 \\ \nu_{Coag}(\mathcal{S}^{\downarrow}) = 0 \end{cases}$$

then, if at all time all the blocks of $\Pi(t)$ are infinite we can see the number of blocks $(Z(t) = \#\Pi(t), t \geq 0)$ as the size of a population where each individuals gives rise (without dying) to a progeny of size i with rate $\nu_{Disl}(\Delta_f(i+1))$ and there is a negative density-dependence due to competition pressure. This is reflected by the Kingman coalescence phenomena which results in a quadratic death rate term. The natural death rate is set to 0, i.e. there is no linear component in the death rate. In this context, an EFC process that comes down from infinity corresponds to a population started with a very large size. Lambert has shown that a sufficient condition to be able to define what he terms a logistic branching process started from infinity is

$$(L) \qquad \sum_k p_k \log k < \infty$$

where $p_k = \nu_{Disl}(\Delta_f(k+1))$. Note that from a population dynamic point of view this corresponds to the case where the natural death rate is set to 0.

More precisely, this means that if P_n is the law of the $\mathbb{N}$-valued Markov chain $(Y(t), t \geq 0)$ started from $Y(0) = n$ with transition rates

$$\forall i \in \mathbb{N} \left\{ \begin{array}{c} i \to i + j \text{ with rate } ip_j \text{ for all } j \in \mathbb{N} \\ i \to i - 1 \text{ with rate } c_k i(i-1)/2 \text{ when } i > 1. \end{array} \right. ,$$

then P_n converge weakly to a law P_∞ which is the law of a $\mathbb{N} \cup \infty$-valued Markov process $(Z(t), t \geq 0)$ started from ∞, with same transition semigroup on $\mathbb{N}$ as Y and whose entrance law can be exhibited. Moreover, if we call $\tau = \inf\{t \geq 0 : Z(t) = 1\}$ we have that $\mathbf{E}(\tau) < \infty$. If we note $\mathbf{E}_n(.)$ the expectation under the probability P_n and if we define

$$z_i := \mathbf{E}_{i+1}(\inf\{t \geq 0 : Y(t) = i\}$$

then it is not hard to see that

$$\mathbf{E}(\tau) = \sum_{i \in \mathbb{N}} z_i$$

because one must pass through every state $i \in \mathbb{N}$ before reaching 1.

This result translates easily in our framework into

Proposition 3.7. *Let Π be an EFC process started from dust (i.e. $\Pi(0) = \mathbf{0}$) and verifying the conditions (H) and (L). Then one has*

$$\forall t > 0, \ \#\Pi(t) < \infty \ a.s.$$

Proof. Our proof relies on a careful comparison of the $(\#\Pi_{|[n]}(t))_{n \in \mathbb{N}}$ with $Z(t)$ the process studied by Lambert. For each $n \in \mathbb{N}$ and $\pi \in \mathcal{P}_n$ such that $\pi \neq \mathbf{1}_n$ we define a random variable $L^{(n)}(\pi)$ with same distribution as the random time $\inf\{t \geq 0 : \#\Pi_{|[n]}(t) = \#\pi - 1\}$ conditionally on $\Pi_{|[n]}(0) = \pi$. We call

$$y_i^{(n)} = \max_{\pi \in \mathcal{P}_n, \#\pi = i+1} \mathbf{E}(L^{(n)}(\pi)).$$

Hence if we call $T_n = \inf\{t \geq 0 : \#\Pi_{|[n]}(t) = 1\}$ we have that

$$\mathbf{E}(T_n) \leq \sum_{i=1}^{n-1} y_i^{(n)}.$$

The first step is to show recursively that for each fixed n, for each $i \leq n-1$ one has $y_i^{(n)} \leq z_i$.

Fix n, let us first show that $y_{n-1}^{(n)} \leq z_{n-1}$. One has by construction that $y_{n-1}^{(n)} = c_k \frac{n(n-1)}{2}$. From now on we will suppose that $c_k/2 = 1$ (we do not lose

any generality as this corresponds to a change of scale in time) and hence $y_{n-1}^{(n)} = n(n-1)$.

We now construct a random variable L_{n-1} which has same distribution as $\inf\{t \geq 0 : Z(t) = n-1\}$ conditionally on $Z(0) = n$ (and hence $\mathbf{E}(L_{n-1}) = z_{n-1}$). Let $(t_1^{(c)}, t_2^{(c)}, ...)$ be the jump times of a Poisson process of intensity $n(n-1)$ and let $(t_1^{(F)}, t_2^{(F)}, ...)$ be the jump times of an independent Poisson process of intensity $n\rho$. Furthermore let $(E_i)_{i\in\mathbb{N}}$ be an i.i.d. sequence of variables where E_i has the distribution of the length of an excursion of Z(.) above n, i.e. if $S = \inf\{t : Z(t) \neq n\}$ then conditionally on $Z(0) = n$ and $Z(S) > n$ one has $E_i \stackrel{\mathcal{L}}{=} (\inf\{t \geq S : Z(t) = n\} - S)$.

Then we define L_{n-1} as follows :

- if $t_1^{(c)} < t_1^{(F)}$ then a coalescence occurs before the first fragmentation and hence the first jump of Z is to $n-1$, hence $L_{n-1} = t_1^{(c)}$,
- else we jump to the right, hence we leave n at time $t_1^{(F)}$ and return to n after an excursion above n at time $t_1^{(F)} + E_1$, at which time the Markov property of the Poisson processes implies that we are in the same situation as at time 0, i.e. the time before the next $t_i^{(c)}$ is exponential with parameter $n(n-1)$ and the time before the next $t_i^{(F)}$ is exponential with parameter ρn. Hence we can start the procedure again with the shifted Poisson processes.

This implies that $L_{n-1} \in \{t_1^{(c)}, t_2^{(c)}, ...\}$ and thus $\mathbf{E}(L_{n-1}) = z_{n-1} \geq n(n-1)$.

Suppose now that for all $j \in \{k+1, ..., n-1\}$ one has $y_j^{(n)} \leq z_j$. Let us show that $y_k^{(n)} \leq z_k$. If we note $\bar{p}_i = \sum_{k\geq i} p_k$ the same technique we employed to show that $y_{n-1}^{(n)} \leq z_{n-1}$ yields

$$\begin{aligned} y_k^{(n)} \leq & \frac{1}{k(k+1)+\rho(k+1)} \frac{k(k+1)}{k(k+1)+\rho(k+1)} \\ & + \frac{\rho(k+1)}{k(k+1)+\rho(k+1)} \left(y_k^{(n)} + \bar{p}_1 y_{k+1}^{(n)} + ... + \bar{p}_{n-1-k} y_{n-1}^{(n)} \right) \\ \leq & \frac{1}{k(k+1)+\rho(k+1)} \frac{k(k+1)}{k(k+1)+\rho(k+1)} \\ & + \frac{\rho(k+1)}{k(k+1)+\rho(k+1)} \left(y_k^{(n)} + \bar{p}_1 z_{k+1} + ... + \bar{p}_{n-1-k} z_{n-1} \right). \end{aligned}$$

We only need to remark that

$$z_k = \frac{1}{k(k+1)+\rho(k+1)}\frac{k(k+1)}{k(k+1)+\rho(k+1)} + \frac{\rho(k+1)}{k(k+1)+\rho(k+1)}\left(z_k + \sum_{j=1}^{\infty}\bar{p}_j z_{k+j}\right)$$

to conclude that necessarily

$$y_k^{(n)} \leq z_k.$$

The bound being uniform in n this implies that $\sup_n \mathbf{E}(T_n) \leq \mathbf{E}(\tau) < \infty$. If $T = \inf\{t : \#\Pi(t) < \infty\}$, the monotone convergence theorem implies that $\mathbf{E}(T) < \infty$ and hence T is almost surely finite. A simple application of Proposition 23 in [13] and Lemma 31 in [15] shows that if there exists $t < \infty$ such that $\#\Pi(t) < \infty$ then $\inf\{t : \#\Pi(t) < \infty\} = 0$.

To conclude suppose that $\#\Pi(0) < \infty$, or without loss of generality that $\Pi(0) = \mathbf{1}$, then it is easily seen that $T^{(\infty)} := \inf\{t : \#\Pi(t) = \infty\} = \infty$. The idea is that the fragmentation alone can not reach infinity : Suppose Π (which still fulfills (H) and (L)) is degenerated : $c_k = 0$. Then if we define $m = \sum kp_k$ the expectation of the number of fragments created at each dislocation, it is not hard to see that the LLN implies that the sequence of the fragmentation times t_n should behave as the series

$$t_n \sim \frac{1}{\rho} + \frac{1}{\rho(1+m)} + \frac{1}{\rho(1+2m)} + ... + \frac{1}{\rho(1+nm)}$$

which diverges. This entails that when an EFC process verifying (H) and (L) reaches a finite level it cannot go back to infinity. As $\inf\{t : \#\Pi(t) < \infty\} = 0$, this means that

$$\forall t > 0, \#\Pi(t) < \infty.$$

□

Remark : Let $\Pi(.) = (B_1(.), B_2(.), ...)$ be a "(H)-(L)" EFC process started from dust, $\Pi(0) = \mathbf{0}$. Then for all $t > 0$ one has a.s. $\sum_i \|B_i(t)\| = 1$. This is clear because at all time $t > 0$ there are only a finite number of blocks.

If we drop the hypothesis $\nu_{Disl}(\mathcal{S}^{\downarrow}) < \infty$ (i.e. we drop (L) and we suppose $\nu_{Disl}(\mathcal{S}^{\downarrow}) = \infty$), the process Π stays infinite (Corollary 3.2). We now show that furthermore, for a fixed t, almost surely $\|B_1(t)\| > 0$. We define by induction a sequence of integers $(n_i)_{i\in\mathbb{N}}$ as follows : we fix $n_1 = 1$ and for each i we chose n_i such that there exists a time $t_i < t$ such that t_i is a

coalescence time at which the block 1 coalesces with the block n_i and such that $n_i > w_{n_{i-1}}(t_{i-1})$ where $w_k(t)$ is the least element of the kth block at time t. This last condition ensure that $(w_{n_i}(t_i))$ is a strictly increasing sequence because one always has $w_n(t) \geq n$. Hence at time t one knows that for each i there has been a coalescence between 1 and $w_{n_i}(t_i)$. Consider $(\Pi_t^{(F)}(s), s \in [0, t[)$ a coupled fragmentation process defined as follows : $\Pi_t^{(F)}(0)$ has only one block which is not a singleton which is

$$B_1^{(F)}(0) = \{1, w_{n_2}(t_2), w_{n_3}(t_3), ...\}.$$

The fragmentations are given by the same PPP P_F used to construct Π (and hence the processes are coupled). It should be clear that if $w_{n_i}(t_i)$ is in the same block with 1 for $\Pi^{(F)}(t)$ the same is true for $\Pi(t)$ because it means that no dislocation separates 1 from $w_{n_i}(t_i)$ during $[0, t]$ for $\Pi^{(F)}$ and hence

$$1 \overset{\Pi(t)}{\sim} w_{n_i}(t_i).$$

Using this fact and standard properties of homogeneous fragmentations one has a.s.

$$\|B_1(t)\| \geq \|B_1^{(F)}(t)\| > 0.$$

Hence for all $t > 0$ one has $P(\{1\} \subset \text{dust}(\Pi(t))) = 0$ and hence $P(\text{dust}(\Pi(t)) \neq \emptyset) = 0$. Otherwise said, when $\nu_{Disl}(\mathcal{S}^{\downarrow}) = \infty$ the fragmentation part does not let a "(H)" EFC process come down from infinity, but it let the dust condensates into mass. Note that "binary-binary" EFC processes are a particular case. The question of the case $\nu_{Disl}(\mathcal{S}^{\downarrow}) < \infty$ but (L) is not true remains open.

BIBLIOGRAPHY

[1] D. J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5(1):3–48, 1999.

[2] J. Berestycki. Multifractal spectra of fragmentation processes. *J. Statist. Phys.*, to appear, 2003.

[3] J. Bertoin. Homogeneous fragmentation processes. *Probab. Theory Related Fields*, 121(3):301–318, 2001.

[4] J. Bertoin. Self-similar fragmentations. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(3):319–340, 2002.

[5] J. Bertoin. The asymptotic behaviour of fragmentation processes. *J. Euro. Math. Soc.*, to appear, 2003.

[6] J. Bertoin. Fragmentations et coalescences stochastiques. In preparation, 2003.

[7] D. Beysens, X. Campi, and E. Peffekorn, editors. *Proceedings of the workshop : Fragmentation phenomena*, Les Houches Series. World Scientific, 1995.

[8] P. Diaconis, E. Mayer-Wolf, O. Zeitouni, and M. P. W. Zerner. Uniqueness of invariant measures for split-merge transformations and the poisson-dirichlet law. *Ann. Probab.*

[9] R. Durrett and V. Limic. A surprising model arising from a species competition model. *Stoch. Process. Appl*, 102:301–309, 2002.

[10] J. F. C. Kingman. The representation of partition structures. *J. London Math. Soc. (2)*, 18(2):374–380, 1978.

[11] A. Lambert. The branching process with logistic growth. preprint, 2003.

[12] M. Möhle and S. Sagitov. A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.*, 29(4):1547–1562, 2001.

[13] J. Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27(4):1870–1902, 1999.

[14] J. Pitman. Poisson-Dirichlet and GEM invariant distributions for split-and-merge transformation of an interval partition. *Combin. Probab. Comput.*, 11(5):501–514, 2002.

[15] J. Schweinsberg. Coalescents with simultaneous multiple collisions. *Electron. J. Probab.*, 5:Paper no. 12, 50 pp. (electronic), 2000.

[16] J. Schweinsberg. A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Comm. Probab.*, 5:1–11 (electronic), 2000.

4. ANNEXE A

4.1 Self-similar fragmentation without erosion

In this annexe we present precisely how to extend the Poisson Point Process construction given in chapter 1 for homogeneous ranked fragmentations to the self-similar case. More precisely we shall give an equivalent of theorem (1.1). In the case $\alpha > 0$ we shall use this to show that the results on the small time behavior are still true.

To construct a non-homogeneous ranked fragmentation, we have to control the rate of fragmentation in terms of the size of each fragment. Note that the behavior of a self-similar fragmentation is qualitatively different depending on the sign of α : when $\alpha < 0$ the rate of fragmentation increases when the fragments become small, at the contrary, when $\alpha > 0$ the rate decreases which means that small fragments tend to live longer. For instance Bertoin has shown in [2] that when $\alpha < 0$ the life-time of the process is almost surely finite (whereas of course it is always infinite when $\alpha \geq 0$).

In [1] it was shown that a homogeneous $\mathcal{P}$-fragmentation can be transformed into a self-similar one of index α through a change of time. Essentially the idea is to use the mechanism of the Lamperti construction of semi-stable processes. More precisely denote by $l(i,t)$ the asymptotic frequency of the block of $\Pi(t)$ that contains i. Then for an arbitrary $\beta \in \mathbb{R}$ introduce

$$C_i^{(\beta)}(t) := \int_0^t l(i,r)^{-\beta} dr$$

and its generalized inverse

$$T_i^{(\beta)}(t) = \inf\{u \geq 0 : C_i^{(\beta)}(u) > t\}, t \geq 0.$$

Consider the random partition $\Pi^{(\beta)}(t)$ such that $i, j \in \mathbb{N}$ are in the same block of $\Pi^{(\beta)}(t)$ if and only if they are in the same block of $\Pi(T_i^{(\beta)}(t))$, then if Π is a self-similar fragmentation with index α the process $\Pi^{(\beta)} = (\Pi^{(\beta)}(t), t \geq 0)$ is a self-similar $\mathcal{P}$-valued fragmentation with index $\alpha + \beta$. Moreover Π can be recovered from $\Pi^{(\beta)}$, by the operation $\Pi = (\Pi^{(\beta)})^{(-\beta)}$ in the obvious notation. Hence a self-similar fragmentations of index α can always be seen as the transformation of a homogeneous one.

Note that as expected when $\alpha > 0$ we have that $T_i^{(\alpha)}(t) < t$ because $\forall i \in \mathbb{N}, l(i,r)^{-\alpha} > 1$ so the change of time slows down the fragmentation.

4.1.1 Poisson point process description

We will now introduce a new way of describing a self-similar ranked fragmentation that does not use a change of time but rather a thinning of a PPP. We do so only in the case of a positive index of self-similarity but a similar construction for the negative case should also work. Let dx denote the Lebesgue measure on R^+.

Proposition 4.1. *Let λ be an $(\alpha, \nu, 0)$ $\mathcal{S}^\downarrow$-fragmentation with ν verifying the hypothesis $\nu(\{s \in \mathcal{S}^\downarrow : s_2 > 0\} = \infty$ and $\alpha > 0$ then :*

- *λ is a pure jump process,*
- *there exists a Poisson point process $\tilde{K} = (S(t), k(t), U(t))$ with measure intensity $\nu \otimes \# \otimes dx_{|[0,1]}$ such that λ only jumps at times t at which $\tilde{K}$ has an atom such that $U(t) < \lambda_{k(t)}(t-)^\alpha$, and at such a time $\lambda(t)$ is obtained from $\lambda(t-)$ simply by dislocating the $k(t)$-th fragment of $\lambda(t-)$ by $S(t)$ and reordering the new sequence of fragments.*

Our first step is to prove a similar result for $\mathcal{P}$-fragmentations. Let Π be a homogeneous self similar $\mathcal{P}$-fragmentation and K be the associated Poisson point process of its jumps.

Recall that until now K had its values in $\mathcal{P} \times \mathbb{N}$ and that its measure intensity was given by $\mu_\nu \otimes \#$. We now add a mark $U(t)$ uniform on $]0,1[$ to each atom of K, i.e., for each atom $(\Delta(t), k(t))$ we draw an independent variable $U(t)$ uniformly distributed on $]0,1[$. Otherwise said, K has its values in $\mathcal{P} \times \mathbb{N} \times]0,1[$ and its measure intensity is given by $\mu_\nu \otimes \# \otimes dx_{|[0,1]}$.

We adopt the following notation :

- Let $\Pi^{(\alpha)}(t)$ be the self-similar fragmentation obtained from Π by the above described change of time :

$$(i \overset{\Pi^{(\alpha)}(t)}{\sim} j) \Leftrightarrow (i \overset{\Pi(T_i^{(\alpha)}(t))}{\sim} j).$$

- We call $l^{(\alpha)}(i,t)$ the asymptotic frequency of the block of $\Pi^{(\alpha)}(t)$ that contains i (and we will note $(l^{(\alpha)}(i,t))^\alpha$ when we want to raise it to the power α). Hence, for all $t \geq 0$ one has

$$l^{(\alpha)}(i,t) = l(i, T_i^{(\alpha)}(t)).$$

- Call $B(i,t)$ (resp.$B^{(\alpha)}(i,t)$) the block of $\Pi(t)$ (resp. of $\Pi^{(\alpha)}(t)$) that contains i, then by definition

$$B(i,t) = B^{(\alpha)}(i, C_i^{(\alpha)}(t)).$$

An atom $(\Delta(t), k(t), U(t))$ such that $B_{k(t)}(t-)$ is not empty corresponds to a jump of Π at time t, the same jump occurs later for $\Pi^{(\alpha)}$, at time $C_{k(t)}^{(\alpha)}(t)$.

Hence the jump structure of $\Pi^{(\alpha)}$ corresponds to the point process obtained by moving each atom $(\Delta(t), k(t), U(t))$ of K from the date t to the date $C_{k(t)}^{(\alpha)}(t) > t$. More precisely $\Pi^{(\alpha)}$ verifies the following :

- if t is an atom time for K such that $k(t) = i$, then

$$\Pi^{(\alpha)}(C_i^{(\alpha)}(t)) = \Delta^{(i)}(t) \overset{i}{\circ} \Pi^{(\alpha)}(C_i^{(\alpha)}(t-)).$$

- For all s which is not the image of an atom time (Δ, k, U) by the corresponding map $C_k^{(\alpha)}$ one has $\Pi^{(\alpha)}(s) = \Pi^{(\alpha)}(s-)$.

We now show that the point process of the displaced atoms can be seen as a thinned Poisson point process. Call H the point process obtained by moving each atom $(\Delta(t), k(t), U(t))$ of K from the date t to the date $C_{k(t)}^{(\alpha)}(t) > t$ and by further replacing the mark $U(t)$ by $l^\alpha(k(t), t-)U(t)$. Hence H defines $\Pi^{(\alpha)}$ in the same way that K defines Π.

Let K_i be the PPP obtained from K by only keeping the atoms such that $k(t) = i$. The K_i are independent in the same filtration. Similarly, call H_i the point processes obtained from H by keeping the atoms such that the second coordinate is i. The measure intensity of K_i is $\mu_\nu \otimes \delta_i \otimes dx_{|[0,1]}$. If H_i has an atom at time $s = C_i^{(\alpha)}(t)$, its mark is

$$U(t)l^\alpha(i,t-) \;=\; U(T_i^{(\alpha)}(s))(l^{(\alpha)}(i,s-))^\alpha \tag{4.1}$$

To be able to construct H_i on a given time interval, say $[0,t]$ we need to know K_i and $l(i,s)$ for $s \in [0, T_i^{(\alpha)}(t)]$ (remark that $T_i^{(\alpha)}(t)$ is a stopping time in the filtration generated by the process $l(i,s)$). Conversely one can recover $(l(i,s), K_i(s))_{s \le T_i^{(\alpha)}(t)}$ from $(l^{(\alpha)}(i,s), H_i(s))_{s \le t}$. More precisely let $s \le T_i^{(\alpha)}(t)$, then

$$\begin{aligned} l(i,s) &= l^{(\alpha)}(i, C_i^{(\alpha)}(s)) \\ &= l^{(\alpha)}(i, \inf\{u : T_i^{(\alpha)}(u) > s\}) \\ &= l^{(\alpha)}(i, \inf\{u : \int_0^u (l^{(\alpha)}(i,r-))^\alpha dr > s\}) \end{aligned}$$

where by definition $\inf\{u : \int_0^u (l^{(\alpha)}(i, r-))^\alpha dr > s\}$ is bounded by t.

Let us call

$$\mathcal{F}_i(t) := \sigma\{l(i,u), K_i(u) : u \leq t\}$$

the filtration generated by $l(i,.)$ and K_i up to time t. If similarly we let

$$\mathcal{F}_i^{(\alpha)}(t) := \sigma\{l^{(\alpha)}(i,u), H_i(u) : u \leq t\}$$

be the filtration generated by $l^{(\alpha)}(i,.)$ and H_i up to time t, it is now clear that for all $t \geq 0$

$$\mathcal{F}_i^{(\alpha)}(t) = \mathcal{F}_i(T_i^{(\alpha)}(t))$$

because the trajectories of $l(i,.)$ and K_i can be recovered from the trajectories of $l^{(\alpha)}(i,.)$ and H_i. Remark that there is some redondant information in $\sigma\{l(i,u), K_i(u) : u \leq t\}$ but that K_i alone does not generate the correct filtration because for small times the label of $B(i,t)$ is not i, and thus the dislocations of $B(i,t)$ are not described by the atoms of K_i.

Furthermore, if we define $\mathcal{F}(t) := \sigma\{K(u), u \leq t\}$ (resp. $\mathcal{F}^{(\alpha)}(t) := \sigma\{H(u), u \leq t\}$) there is no simple ways of going from $\mathcal{F}(t)$ to $\mathcal{F}^{(\alpha)}(t)$ or the converse.

Lemma 4.1. *For each i the point process H_i has the following stochastic measure intensity :*

$$dt \otimes \mu_\nu \otimes \delta_i \otimes dx_{|[0,(l^{(\alpha)}(i,t-))^\alpha]}.$$

Proof. Consider a cylinder set of the form $\Gamma := A \times \{i\} \times]0,1[$ where the set $A \in \mathcal{P}$ is such that $\mu_\nu(A) < \infty$. We choose the whole set $]0,1[$ for the last coordinate to make the demonstration clearer.

Let $N^{(\alpha)}(]a,b[)$ (resp. $N(]a,b[)$)be the number of atoms of H_i (resp. K_i) that fall in Γ during the time interval $]a,b[$. The first step is to show that

$$Y^{(\alpha)}(t) = N^{(\alpha)}(]0,t[) - \mu_\nu(A) \int_0^t (l^{(\alpha)}(i,u-))^\alpha du$$

is a martingale in the filtration $\mathcal{F}_i^{(\alpha)}$ and in $\mathcal{F}^{(\alpha)}$. Clearly the process $Y(t) = N(]0,t[) - t\mu_\nu(A)$ is a martingale in $\mathcal{F}_i$ and in $\mathcal{F}$ (because $N(]s,t[)$ is independent of what happens before s). We will use the change of time and the martingale property for Y to have the result for $Y^{(\alpha)}$.

Remark that $l^{(\alpha)}(i,t-)$ is a càdlàg process, so almost surely for all $t \geq 0$ one has $\int_0^t (l^{(\alpha)}(i,u-))^\alpha du = \int_0^t (l^{(\alpha)}(i,u))^\alpha du$, and we do not need to worry about the left limits in the integral.

Hence, making the change of variable $v = C_i^{(\alpha)}(u)$ (although $C_i^{(\alpha)}$ is not derivable, it has right derivatives everywhere almost surely, and this right derivative is càdlàg)

$$\begin{aligned}
& \mathbf{E} \quad \left(N^{(\alpha)}(]s,t[) - \mu_\nu(A) \int_s^t (l^{(\alpha)}(i,v-))^\alpha dv | \mathcal{F}^{(\alpha)}(s) \right) \\
= \; & \mathbf{E} \quad \left(N^{(\alpha)}(]s,t[) - \mu_\nu(A) \int_{T_i^{(\alpha)}(s)}^{T_i^{(\alpha)}(t)} (l^{(\alpha)}(i, C_i^{(\alpha)}(u-)))^\alpha l^{-\alpha}(i,u-) du | \mathcal{F}^{(\alpha)}(s) \right) \\
= \; & \mathbf{E} \quad \left(N^{(\alpha)}(]s,t[) - \mu_\nu(A) \int_{T_i^{(\alpha)}(s)}^{T_i^{(\alpha)}(t)} 1 du | \mathcal{F}^{(\alpha)}(s) \right)
\end{aligned}$$

because $l(i,u-) = l^{(\alpha)}(i, C_i^{(\alpha)}(u-))$. Note that if we do not suppose that $\nu(\sum_i x_i < 1) = 0$ we can have $l(i,t) = 0$ for a finite t, but the above holds nonetheless if we adopt the convention $0 \times \infty = 0$. Conditionally on $\mathcal{F}_i^{(\alpha)}(s)$ one has that $N^{(\alpha)}(]s,t[)$ and $l^{(\alpha)}(i,u), u \geq s$ are independent of the $\mathcal{F}_j^{(\alpha)}(s)$ for $j \neq i$ (see [1], where the stopping times $T_i^{(\alpha)}(.)$ are called "frosts"). Thus one may replace the conditioning as follows

$$\begin{aligned}
& \mathbf{E} \quad \left(N^{(\alpha)}(]s,t[) - \mu_\nu(A)(T_i^{(\alpha)}(t) - T_i^{(\alpha)}(s)) | \mathcal{F}^{(\alpha)}(s) \right) \\
= \; & \mathbf{E} \quad \left(N^{(\alpha)}(]s,t[) - \mu_\nu(A)(T_i^{(\alpha)}(t) - T_i^{(\alpha)}(s)) | \mathcal{F}_i^{(\alpha)}(s) \right)
\end{aligned}$$

Furthermore it is clear that

$$N^{(\alpha)}(]s,t[) = N(]T_i^{(\alpha)}(s), T_i^{(\alpha)}(t)[).$$

Thus

$$\begin{aligned}
& \mathbf{E} \quad \left(N^{(\alpha)}(]s,t[) - \mu_\nu(A)(T_i^{(\alpha)}(t) - T_i^{(\alpha)}(s)) | \mathcal{F}_i^{(\alpha)}(s) \right) \\
= \; & \mathbf{E} \quad \left(N(]T_i^{(\alpha)}(s), T_i^{(\alpha)}(t)[) - \mu_\nu(A)(T_i^{(\alpha)}(t) - T_i^{(\alpha)}(s)) | \mathcal{F}_i(T_i^{(\alpha)}(s)) \right) \\
= \; & \mathbf{E} \quad \left(Y(T_i^{(\alpha)}(s)) - Y(T_i^{(\alpha)}(t)) | \mathcal{F}(T_i^{(\alpha)}(s)) \right) \\
= \; & 0
\end{aligned}$$

where the last equality comes from the fact that $Y(t) = N(]0,t[) - t\mu_\nu(A)$ is a martingale and $T_i^{(\alpha)}(r) \leq r$ for all $r \geq 0$. Thus we can apply the optional sampling theorem for bounded stopping-times and $Y^{(\alpha)}$ is also a martingale.

Hence the value of the intensity of $N^{(\alpha)}$ at time t is $(l^{(\alpha)}(i,t-))^\alpha \mu_\nu(A)$. Next it is not hard to see that the assertion holds by considering more general cylinders of the form $A \times \{i\} \times I$ where $I \subset]0,1[$ is a Borel set. □

Lemma 4.2. *There exists a PPP $K'(.) = (\Delta'(.), k'(.), U'(.))$ with same distribution as K such that*

- *if t is not an atom time for K' one has $\Pi^{(\alpha)}(t) = \Pi^{(\alpha)}(t-)$,*
- *if t is an atom time for K' such that $U'(t) > (l^{(\alpha)}(k'(t), t-))^\alpha$ then $\Pi^{(\alpha)}(t) = \Pi^{(\alpha)}(t-)$,*
- *and finally if t is an atom time for K' such that $U'(t) \leq (l^{(\alpha)}(k'(t), t-))^\alpha$ then*

$$\Pi^{(\alpha)}(t) = \Delta'(t) \overset{k'(t)}{\circ} \Pi^{(\alpha)}(t-).$$

Furthermore, $\Pi^{(\alpha)}$ has no other jumps.

Proof. For each i, conditionally on $l^{(\alpha)}(i, .)$, define an independent point process $\bar{H}_i(t)$, with values in $\mathcal{P} \times \{i\} \times [0, 1]$ and stochastic measure-intensity

$$m_i(t) = dt \otimes \mu_\nu \otimes \delta_i \otimes dx_{|[(l^{(\alpha)}(i,t-))^\alpha, 1]}.$$

Then by standard properties of independent point processes $K'_i := \bar{H}_i(t) + H_i(t)$ is a càdlàg point process with measure intensity $\mu_\nu \otimes \delta_i \otimes dx_{|[0,1]}$, and hence it is a Poisson point process. Call $K' = \sum K'_i$ the superposition of the K'_i, then K' is a PPP with measure intensity $dt \otimes \mu_\nu \otimes \# \otimes dx_{|[0,1]}$ such that H is exactly the point process obtained by only keeping the atoms $(\Delta'(t), k'(t), U'(t))$ of K' such that $U'(t) \leq (l^{(\alpha)}(k'(t), t-))^\alpha$. Finally, $\Pi^{(\alpha)}$ has no other jumps because all jumps of Π correspond to atoms of K. □

As in the homogeneous case we can define a random $(\mathcal{F}_t)$-adapted map $\phi : \mathbb{N} \times \mathbb{R}_+ \to \bar{\mathbb{N}}$ that associates to (n, t) the rank of the asymptotic frequency of $B_n^{(\alpha)}(t)$ in $\Pi^{(\alpha)}(t)$. Define the point process $\tilde{K}(.)$ as the image of K' by the following random $\mathcal{F}_t$-measurable map :

$$(\Delta', k', U') \to (\Lambda(\Delta'), \phi(t-, k'(t)), U')$$

where the atoms such that $\phi(t-, k'(t)) = \infty$ are discarded.

Lemma 4.3. *1. The point process $\tilde{K}(t)$ is a PPP with measure intensity $\nu \otimes \# \otimes dx_{|[0,1]}$.*

2. If t is a jump-time for $\Lambda(\Pi^{(\alpha)})$, then K' has an atom at t.

3. For ν such that $\nu(\{s \in \mathcal{S}^\downarrow : s_2 > 0\}) = \infty$, $\lambda = \Lambda(\Pi^{(\alpha)})$ is a pure jump process.

Proof. The first point is just Lemma 1.1. For the second point, as Λ is not continuous, we need to show that if t is a jump time for $\Lambda(\Pi^{(\alpha)}(.))$ it is also a jump time for $\Pi^{(\alpha)}$ and hence an atom time for K'.

If t is a jump time for $\Lambda(\Pi^{(\alpha)}(.))$, say for instance that $|B_k^{(\alpha)}(.)|$ jumps at t, then $|B_k(.)|$ jumps at time $T_k^{(\alpha)}(t)$. As Π is homogeneous we can apply Lemma 3.5 and hence $B_k(.)$ jumps at time $T_k^{(\alpha)}(t)$ and thus $B_k^{(\alpha)}(.)$ jumps at t.

Finally, for the third assertion, remark that

$$\{\sup_{n>n_0}\{\sup_{u>0}\left(|B_n^{(\alpha)}(u)|\right)\}\} = \{\sup_{n>n_0}\{\sup_{u>0}\left(|B_n(u)|\right)\}\}.$$

Thus, as $\sup_{n>n_0}\{\sup_{u>0}(|B_n(u)|)\} \to 0$ a.s. when $n_0 \to \infty$, one has that for each $\epsilon > 0$, almost surely $\exists n_0$ such that

$$\lambda_1(s) \wedge \epsilon = \max_{n \leq n_0} |B_n^{(\alpha)}(s)| \wedge \epsilon.$$

The processes $|B_n^{(\alpha)}(.)|$ are just the images of the $|B_n(.)|$ by a change of time so they are pure jump processes. Hence $\lambda_1 \wedge \epsilon$ is also a pure jump process. One can generalize to any λ_k and thus λ is a pure jump process.

□

Hence, if $\tilde{K}$ has an atom (S, k, U) at time t, then it is the image of some atom of K, therefore if $U \leq \lambda_k^{\alpha}(t-)$ the process λ has a jump at time t and $\lambda(t)$ is obtained from $\lambda(t-)$ simply by dislocating the $k(t)$-th fragment of $\lambda(t-)$ by $S(t)$ and reordering the new sequence of fragments. If $U > \lambda_k^{\alpha}(t-)$ the atom does not belong to H and thus the arguments used above for the second point show that $\lambda(t) = \lambda(t-)$.

We conclude that $\tilde{K}$ and $\lambda = \Lambda(\Pi^{(\alpha)})$ are the processes whose existences are asserted in Proposition (4.1).

4.1.2 Asymptotic behavior

This construction allows us to study the asymptotic behavior of the fragment's size near 0 as in the last section of chapter 1.

Corollary 4.1. *The proposition 1.4 is still true for an* $(\alpha, \nu, 0)$ $\mathcal{S}^{\downarrow}$*-fragmentation with* $\alpha > 0$.

It is easy to verify that the part concerning the asymptotic behavior of $1 - \lambda_1$ remains true.

For the small fragments, the idea is that taking $\alpha > 0$ reinforces their record-like behavior : as λ_1 is near 1, it is not affected by the index of self-similarity, and as the small fragments are near 0, they are slowed by the self-similarity property.

The key-point of the demonstration in the case $\alpha = 0$ for λ_2 was the Lemma 1.3, so the demonstration mainly consists in showing that it still stands when $\alpha > 0$.

Consider an $(\alpha, \nu, 0)$ fragmentation λ, with $\alpha \geq 0$, constructed with the PPP $K(.) = (S(.), k(.), U(.))$. $R(t)$ still designates the record at time t of the real-valued PPP $s_2^{(1)}(.)$, and let $\overline{R}(t)$ be the record at time t of the same point process when we only consider the atoms such that

$$U(t) \leq \lambda_1^{\alpha}(t).$$

It is clear that $\overline{R}^{(1)}(t) \leq R^{(1)}(t)$, as the arguments that led to the upper-bound of lemma (1.3) can be re-used here in

$$\lambda_2(t) \leq \overline{R}^{(1)}(t)$$

we conclude that the upper-bound of lemma (1.3) is still valid.

Let us now assert the lower-bound part. Let

$$\overline{\chi}_t = \left(\prod_{u \in [0,t[} \mathbf{1}_{\{U(u) \leq \lambda_1(u)^{\alpha}\}} s_1^{(1)}(u) \right) \left(\prod_{u \in [0,t[} \mathbf{1}_{\{U(u) \leq \lambda_2(u)^{\alpha}\}} s_1^{(2)}(u) \right)$$

with the convention that when the products are empty the are taken equal to 1. We clearly have that $\overline{\chi}_t \geq \chi(t)$. In this setting (1.6) becomes

$$\overline{\chi}_t \overline{R}^{(1)}(t) \leq \lambda_2(t).$$

So we would like to replace $\overline{R}^{(1)}(t)$ by the true record $R^{(1)}(t)$. To do so, note that the family $(U(u); u \geq 0)$ is independent of $T(t)$, the time of record, so for all $t \geq 0$, $U(T(t))$ has uniform distribution on $[0, 1]$. As

$$\{R^{(1)}(t) = \overline{R}^{(1)}(t)\} = \{U(T(t)) \leq \lambda_1(t)^{\alpha}\}$$

and almost surely $\lambda_1^{\alpha}(t) \to 1$ so this event is asymptotically almost certain. On this event the lower-bound part of lemma (1.3) is still valid, the proposition is hence demonstrated.

This says that the renormalization and the asymptotic law of $\lambda_2(t)$ does not depend on α when it is positive. The arguments are adapted in the same way for λ_k for $k \geq 2$ in the case of a binary fragmentation.

We would like to mention that Miermont and Schweinsberg proved the Corollary 4.1 by very close methods in [3].

BIBLIOGRAPHY

[1] J. Bertoin. Self-similar fragmentations. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(3):319–340, 2002.

[2] J. Bertoin. The asymptotic behaviour of fragmentation processes. *J. Euro. Math. Soc.*, to appear, 2003.

[3] G. Miermont and J. Schweinsberg. Self-similar fragmentations and stable subordinators. To appear in Seminaire de Probabilites XXXVII, 2003.