

# Habilitation à diriger des recherches



Spécialité : Mathématiques

## ARBRES, CARTES, FRAGMENTATION ET COALESCENCE ALÉATOIRES

Grégory Miermont

Rapporteurs :

M. Steven N. Evans	University of California, Berkeley
M. Jean-François Le Gall	Université Paris-Sud 11
M. James R. Norris	University of Cambridge

Soutenu le 27 novembre 2008 devant un jury composé de :

M. Jean Bertoin	Université Paris 6
M. Geoffrey R. Grimmett	University of Cambridge
M. Jean-François Le Gall	Université Paris-Sud 11
M. Yves Le Jan	Université Paris-Sud 11
M. James R. Norris	University of Cambridge
M. Gilles Schaeffer	École Polytechnique



# Remerciements

En tout premier lieu, je remercie très chaleureusement Jean Bertoin et Jean-François Le Gall, qui ont été mes mentors depuis le début de mes études en théorie des probabilités, et qui continuent aujourd'hui à être pour moi des modèles.

Je suis très reconnaissant envers les rapporteurs de cette habilitation à diriger des recherches. Je n'oublie pas, du reste, que c'est entre autres grâce à Steven Evans et James Norris que mes séjours au Department of Statistics de Berkeley et au Statistical Laboratory de Cambridge ont été si profitables.

Merci à Geoffrey Grimmett, Yves Le Jan et Gilles Schaeffer de me faire l'honneur de participer au jury.

Mes remerciements sincères aux mémorables enseignants qui ont guidé mes premiers pas en recherche, en particulier, et hormis ceux mentionnés ci-dessus, David Aldous, Jim Pitman et Marc Yor.

Merci à mes coauteurs, David Aldous, Jean Bertoin, Bénédicte Haas, Jean-François Marckert, Jim Pitman, Jason Schweinsberg, Mathilde Weill et Matthias Winkel, pour tout ce que notre travail m'a apporté.

Merci à tous mes collègues et amis de l'École Normale Supérieure, d'Orsay, de Paris 6, de Berkeley, et de Cambridge. Merci à la Fondation des Sciences Mathématiques de Paris, pour m'avoir donné d'excellentes conditions de travail pendant l'année 2007–2008.

Merci à mes amis, à ma famille, et à Anne, pour leur soutien constant, par temps couvert comme par temps clair.



# Table des matières

<b>1</b>	<b>Introduction et tour d'horizon</b>	<b>6</b>
1.1	Arbres de Galton-Watson spatiaux . . . . .	6
1.2	Arbres continus aléatoires . . . . .	10
1.3	Cartes aléatoires . . . . .	20
1.4	Fragmentations auto-similaires . . . . .	24
1.5	Coalescence additive . . . . .	27
<b>2</b>	<b>Principes d'invariance pour les arbres de Galton-Watson spatiaux multitypes</b>	<b>30</b>
2.1	Processus de Galton-Watson multitypes . . . . .	30
2.2	Arbres de Galton-Watson multitypes [P11,P15] . . . . .	32
<b>3</b>	<b>Cartes aléatoires</b>	<b>35</b>
3.1	Principes d'invariance pour les cartes aléatoires [P10,P11,P12] . . . . .	35
3.2	Sphéricité de la limite d'échelle des cartes [P13] . . . . .	42
3.3	Bijection à plusieurs points et géodésiques [P16] . . . . .	44
<b>4</b>	<b>Fragmentations auto-similaires et arbres de fragmentation</b>	<b>48</b>
4.1	Fragmentations auto-similaires et subordinateurs stables [P2] . . . . .	48
4.2	Deux fragmentations de l'arbre stable [P3,P7] . . . . .	49
4.3	Généalogie des fragmentations auto-similaires d'indice négatif [P5] . . . . .	51
4.4	Cladogrammes markoviens et leurs limites continues [P14] . . . . .	54
<b>5</b>	<b>p-arbres, applications aléatoires et leurs limites continues</b>	<b>58</b>
5.1	Hauteur et largeur des arbres continus inhomogènes [P6] . . . . .	58
5.2	Applications aléatoires [P4,P8] . . . . .	61
<b>6</b>	<b>Coalescence additive et parking de Knuth</b>	<b>66</b>
6.1	Coalescents ordonnés associés aux processus de Lévy [P1] . . . . .	66
6.2	Caravanes dans le problème de parking de Knuth [P9] . . . . .	68
	<b>Liste des travaux</b>	<b>72</b>

# 1 Introduction et tour d’horizon

Ce document est un guide de lecture pour les travaux [P1–P16] énumérés pages 72 et suivante. Ces derniers se classent assez naturellement en cinq thèmes, qui font l’objet des chapitres 2 à 6 :

- Les arbres de Galton-Watson spatiaux à plusieurs types et leurs limites d’échelle
- Les cartes aléatoires et leurs limites d’échelle
- Les processus de fragmentation stochastiques
- Les  $\mathbf{p}$ -arbres de Cayley, les applications aléatoires et leurs limites d’échelle
- La coalescence stochastique additive et le problème de parking de Knuth

Une des difficultés rencontrées à la rédaction du présent texte est que ces thèmes, en apparence bien séparés, sont en fait reliés entre eux par de nombreux aspects. En particulier, un thème récurrent dans mes travaux concerne l’étude de grands arbres aléatoires et de leurs limites d’échelle, dites « continues ».

Par conséquent, plutôt que d’entamer chacun des chapitres de ce mémoire par une introduction spécifique, en présentant en premier lieu un résumé global de nos travaux, nous proposons de commencer l’exposé par une introduction étendue de nos thèmes de recherche, et des concepts qu’ils sous-tendent, en gageant que la présentation gagnera en cohérence ce qu’elle peut éventuellement perdre en concision.

Ce faisant, et profitant du recul que nous avons pu acquérir depuis le début de nos recherches, nous nous sommes livrés à un léger exercice d’harmonisation des notations et des objets considérés dans différents travaux. Cette harmonisation a pu nous amener à employer des notations différentes de celles figurant dans les articles publiés, voire de reformuler certains de nos résultats, pour favoriser leur mise en perspective.

## 1.1 Arbres de Galton-Watson spatiaux

### 1.1.1 Arbres de Galton-Watson critiques

La première famille d’arbres que nous considérons est celle des arbres plans, c’est-à-dire des arbres tracés dans le plan sans croisement d’arêtes. Ces derniers admettent une description combinatoire simple. Notons

$$\mathcal{U} = \bigsqcup_{n \geq 0} \mathbb{N}^n,$$

l’ensemble des mots d’entiers, où  $\mathbb{N} = \{1, 2, \dots\}$  et  $\mathbb{N}^0 = \{\emptyset\}$  est constitué du seul mot vide. Si  $u \in \mathcal{U}$ , on note  $|u|$  sa longueur, et on note en général  $u = u_1 \dots u_n$  avec  $n = |u|$ . La concaténation des mots  $u$  et  $v$  est notée  $uv$ , et on note  $u \preceq v$  si  $u$  est un préfixe de  $v$ , c’est-à-dire s’il existe  $w \in \mathcal{U}$  tel que  $uw = v$ , définissant un ordre partiel sur  $\mathcal{U}$ . Si

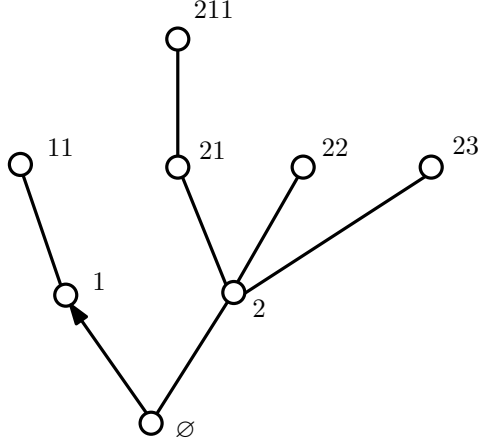


FIG. 1.1: Représentation graphique de l'arbre plan enraciné  $\{\emptyset, 1, 11, 2, 21, 211, 22, 23\}$

$u, v \in \mathcal{U}$ , on note  $u \wedge v$  le plus long préfixe commun à  $u$  et  $v$ . L'ensemble  $\mathcal{U}$  est également muni de l'ordre lexicographique : on note  $u \leq v$  si  $u \preceq v$  ou si  $u \wedge v$  est un préfixe strict de  $u$  et  $v$  tel que  $u_{|u \wedge v|+1} < v_{|u \wedge v|+1}$ . L'ordre  $(\mathcal{U}, \leq)$  est total.

**Définition 1.1** *Un arbre plan enraciné est un sous-ensemble fini  $\mathbf{t} \subset \mathcal{U}$  contenant  $\emptyset$ , tel que si  $ui \in \mathbf{t}$  avec  $u \in \mathcal{U}$  et  $i \in \mathbb{N}$ , alors*

- $uj \in \mathbf{t}$  pour  $1 \leq j \leq i$ , et
- $u \in \mathbf{t}$ .

*Les éléments de  $\mathbf{t}$  sont appelés sommets de  $\mathbf{t}$ , et  $\emptyset$  est appelé la racine.*

*Un sommet  $ui \in \mathbf{t}$ , avec  $u \in \mathcal{U}$  et  $i \in \mathbb{N}$ , est appelé un enfant de  $u$ . Leur nombre est noté  $c_u(\mathbf{t}) = \sup\{i \geq 1 : ui \in \mathbf{t}\} \in \mathbb{Z}_+$ . La longueur  $|u|$  est appelée hauteur de  $u$ .*

*On note  $\mathbf{T}$  l'ensemble des arbres plans enracinés, et  $\mathbf{T}_n \subset \mathbf{T}$  le sous-ensemble des arbres ayant  $n$  sommets.*

Un arbre plan admet une représentation comme graphe plan, où chaque sommet est relié à ses enfants par une arête, et où l'on ordonne de gauche à droite les enfants de chaque sommet par ordre lexicographique. Dans la figure 1.1, on a également distingué l'orientation de l'arête menant de  $\emptyset$  à 1, pour coïncider avec la notion d'enracinement des cartes définie plus loin.

Si  $\mathbf{t} \in \mathbf{T}$ , on note  $\emptyset = u_{\mathbf{t}}(0) \leq u_{\mathbf{t}}(1) \leq \dots \leq u_{\mathbf{t}}(n-1)$  les éléments de  $\mathbf{t}$  classés par ordre lexicographique, où  $n = \#\mathbf{t}$ . La fonction de hauteur associée à  $\mathbf{t}$  est donnée par  $(H_k^{\mathbf{t}} = |u_{\mathbf{t}}(k)|, 0 \leq k \leq \#\mathbf{t} - 1)$ . Par convention, on note  $H_{\#\mathbf{t}}^{\mathbf{t}} = H_{\#\mathbf{t}-1}^{\mathbf{t}}$ . Il est facile de voir que la fonction des hauteurs détermine l'arbre associé de façon univoque.

La première famille d'arbres aléatoires que nous considérons est celle des arbres de Galton-Watson de loi de reproduction critique. Soit  $(\mu(k), k \geq 0)$  une loi de probabilités sur  $\mathbb{Z}_+$ , que l'on suppose critique et non-dégénérée, au sens où

$$\sum_{k \geq 0} k\mu(k) = 1, \quad \mu(1) < 1.$$

## 1 Introduction et tour d'horizon

La loi de l'arbre de Galton-Watson de loi de reproduction  $\mu$  est la mesure de probabilités  $P_\mu$  sur  $\mathbf{T}$  définie par les formules

$$P_\mu(\{\mathbf{t}\}) = \prod_{u \in \mathbf{t}} \mu(c_u(\mathbf{t})), \quad \mathbf{t} \in \mathbf{T},$$

où la somme sur tous les arbres  $\mathbf{t}$  donne bien 1 du fait que  $\mu$  est supposée critique (le processus de Galton-Watson associé  $\mathbf{t} \mapsto (\#\{u \in \mathbf{t} : |u| = n\}, n \geq 0)$  sous  $P_\mu$ , s'éteint presque-sûrement).

### 1.1.2 Limite d'échelle

Nous décrivons maintenant les objets qui apparaissent lorsque l'on considère la limite d'échelle des arbres de Galton-Watson, dans le cas où  $\mu$  a une variance finie. Soit  $(e_t, 0 \leq t \leq 1)$  l'excursion brownienne normalisée standard ([79, Chapitre XII]).

On suppose que

$$\text{Var}(\mu) = \sum_{k \geq 0} k^2 \mu(k) - 1 = \sigma^2 < \infty.$$

Soit alors, pour chaque  $n$  dans  $A_\mu = \{N : P_\mu(\mathbf{T}_N) > 0\}$ , une variable aléatoire  $T_n$  de loi  $P_\mu(\cdot | \mathbf{T}_n)$ . Si  $\mathbb{D}([0, 1])$  est l'espace des fonctions càdlàg muni de la topologie de Skorokhod, on a alors la convergence en loi (où  $n$  varie dans l'ensemble autorisé  $A_\mu$ ) :

$$\left( \frac{H_{[nt]}^{T_n}}{\sqrt{n}} \right)_{0 \leq t \leq 1} \xrightarrow[n \rightarrow \infty]{(d)} \frac{2}{\sigma} e_t. \quad (1.1)$$

Ce résultat est essentiellement dû à Aldous [3], bien que ce dernier considère un processus de codage des arbres différent de la fonction de hauteur (en l'occurrence, la fonction de contour, ou chemin de Dyck associé). La différence est néanmoins mineure, comme justifié dans Duquesne [33] et Marckert & Mokkadem [71], qui ont donné des nouvelles preuves de tels résultats. Une telle convergence permet de donner une information sur de nombreux aspects de grands arbres de Galton-Watson conditionnés par leur taille. Par exemple, on obtient que

$$n^{-1/2} \max\{|u| : u \in T_n\} \xrightarrow[n \rightarrow \infty]{(d)} \frac{2}{\sigma} \sup_{0 \leq t \leq 1} e_t,$$

dont la loi est connue [38]. De même, si  $U_n$  est un sommet de  $T_n$  choisi uniformément au hasard, on a

$$n^{-1/2} |U_n| \xrightarrow[n \rightarrow \infty]{(d)} \frac{2}{\sigma} e_U,$$

où  $U$  est une variable uniforme sur  $[0, 1]$  indépendante de  $e$ . Il est connu que la variable  $2e_U$  suit une loi de Rayleigh  $P(2e_U > x) = \exp(-x^2/2)$ .



### 1.1.3 Arbres de Galton-Watson spatiaux

Nous couplons maintenant les arbres avec une marche aléatoire de branchement de la façon suivante. On note

$$\bar{\mathbf{T}} = \{(\mathbf{t}, \mathbf{l}) : \mathbf{t} \in \mathbf{T}, \mathbf{l} = (\mathbf{l}(u), u \in \mathbf{t}) \in \mathbb{R}^{\mathbf{t}}\},$$

ensemble des *arbres spatiaux*. On appellera  $\mathbf{l}(u)$  la *marque* du sommet  $u$ . Si  $(\mathbf{t}, \mathbf{l}) \in \bar{\mathbf{T}}$ , on définit

$$S_k^{\mathbf{t}, \mathbf{l}} = \mathbf{l}(u_{\mathbf{t}}(k)), \quad 0 \leq k \leq \#\mathbf{t} - 1,$$

et par convention on pose  $S_{\#\mathbf{t}}^{\mathbf{t}, \mathbf{l}} = S_{\#\mathbf{t}-1}^{\mathbf{t}, \mathbf{l}}$ . On appelle  $S^{\mathbf{t}, \mathbf{l}}$  la fonction des marques.

Soit maintenant  $\nu$  une mesure de probabilités sur  $\mathbb{R}$ . Si  $\mathbf{t} \in \mathbf{T}$ , on se donne une famille  $(Y_u, u \in \mathbf{t})$ , où l'on pose  $Y_{\emptyset} = 0$ , et où  $(Y_u, u \in \mathbf{t} \setminus \{\emptyset\})$  est i.i.d., distribuée selon  $\nu$ . On utilise ces variables aléatoires pour construire une marche branchant le long de  $\mathbf{t}$ , où  $Y_u$  est l'accroissement de la marche le long de l'arête montant à  $u$ . Précisément, on pose

$$L_u = \sum_{v \preceq u} Y_v,$$

et l'on note  $\nu_{\mathbf{t}}$  la loi de  $(L_u, u \in \mathbf{t})$ . On note enfin

$$d\mathbb{P}_{\mu, \nu}(\mathbf{t}, \mathbf{l}) = dP_{\mu}(\mathbf{t})d\nu_{\mathbf{t}}(\mathbf{l}),$$

définissant ainsi une loi de probabilités sur l'ensemble  $\bar{\mathbf{T}}$ .

On suppose maintenant que  $\nu$  est centrée et admet un moment d'ordre 4 :

$$\langle \nu, y \rangle = 0, \quad \langle \nu, |y|^4 \rangle < \infty, \quad (1.2)$$

et nous notons  $\varsigma = \text{Var}(\nu)$ . Par ailleurs, on va supposer que  $\mu$  admet des petits moments exponentiels, c'est-à-dire qu'il existe  $\varepsilon > 0$  tel que

$$\sum_{k \geq 0} \mu(k) e^{\varepsilon k} < \infty.$$

Rappelons que  $e$  est une excursion brownienne standard. Conditionnellement à  $e$ , on considère un processus gaussien  $(Z_t, 0 \leq t \leq 1)$  centré dont la covariance est donnée par

$$\text{Cov}(Z_s, Z_t) = \inf_{s \wedge t \leq u \leq s \vee t} e_u.$$

Le processus  $(e, Z)$  est parfois appelé « tête du serpent brownien » conduit par l'excursion brownienne, en référence au serpent brownien, un processus à valeurs trajectoires introduit par Le Gall, permettant de décrire le processus  $Z$  (voir [63]). Il est aisé de voir que  $Z$  admet une version continue, qui est celle avec laquelle nous travaillons. Sous les hypothèses précédentes, on se donne pour tout  $n \in A_{\mu}$  une variable aléatoire  $(T_n, L_n)$

dans  $\overline{\mathbf{T}}$  de loi  $\mathbb{P}_{\mu,\nu}(\cdot | \overline{\mathbf{T}}_n)$ , où bien sûr  $\overline{\mathbf{T}}_n = \{(\mathbf{t}, \mathbf{l}) \in \overline{\mathbf{T}} : \#\mathbf{t} = n\}$ . On peut alors étendre la convergence (1.1) à la convergence en loi dans  $\mathbb{D}([0, 1]^2)$ .

$$\left( \left( \frac{H_{[nt]}^{T_n}}{\sqrt{n}} \right)_{0 \leq t \leq 1}, \left( \frac{S_{[nt]}^{T_n, L_n}}{n^{1/4}} \right)_{0 \leq t \leq 1} \right) \xrightarrow[n \rightarrow \infty]{(d)} \left( \frac{2}{\sigma} \mathfrak{e}, \varsigma \sqrt{\frac{2}{\sigma}} Z \right). \quad (1.3)$$

La renormalisation par  $n^{1/4}$  n'est pas surprenante : considérant que  $S^{T_n, Y_n}$  décrit une famille de marches centrées le long de l'arbre  $T_n$ , et comme la hauteur typique d'un sommet de  $T_n$  est de l'ordre de  $\sqrt{n}$ , on voit que la déviation standard de  $S^{T_n, Y_n}$  est bien de l'ordre de  $\sqrt{\sqrt{n}}$ . Ce résultat, dû à Janson & Marckert [52], est un exemple de convergence d'une marche de branchement discrète, dite « serpent discret », vers le serpent brownien. C'est un tel résultat que nous généralisons dans le cadre d'arbres de Galton-Watson spatiaux multitypes dans [P11,P15], voir le chapitre 2.

## 1.2 Arbres continus aléatoires

Dans les résultats précédents (1.1,1.3), les limites des fonctions de hauteur et des marques sont des processus aléatoires et ne font plus intervenir d'arbres. Il existe cependant une façon de considérer le processus limite  $\mathfrak{e}$  comme « processus des hauteurs » d'un objet arborescent, que l'on construit de la façon suivante.

### 1.2.1 Arbres réels et codage par des fonctions continues

Nous considérons en toute généralité une fonction continue  $f : [0, 1] \rightarrow \mathbb{R}_+$ , telle que  $f(0) = f(1) = 0$ . Pour tout  $s, t \in [0, 1]$ , on note

$$d_f(s, t) = f_s + f_t - 2\check{f}_{s,t},$$

où par définition, on note

$$\check{f}_{s,t} = \inf_{s \wedge t \leq u \leq s \vee t} f_u.$$

Il n'est pas difficile de voir que  $d_f$  est une pseudo-distance sur  $[0, 1]$ , c'est-à-dire que la seule propriété qui la distingue d'une distance est que l'on peut trouver  $s \neq t$  tels que  $d_f(s, t) = 0$ . On pose donc  $s \sim_f t$  si  $d_f(s, t) = 0$ , et on munit le quotient  $\mathcal{T}_f = [0, 1] / \sim_f$  de la distance (encore notée  $d_f$ ) induite par  $d_f$ . On montre alors [36] que l'espace métrique  $(\mathcal{T}_f, d_f)$  est un *arbre réel* (compact), selon la définition suivante :

**Définition 1.2** *Un espace métrique  $(\mathcal{T}, d)$  est un arbre réel si*

1. *pour tout  $x, y \in \mathcal{T}$ , il existe une isométrie  $\varphi_{x,y} : [0, d(x, y)] \rightarrow \mathcal{T}$  telle que*

$$\varphi_{x,y}(0) = x \quad , \quad \varphi_{x,y}(d(x, y)) = y$$

2. *pour tout chemin injectif  $q : [0, 1] \rightarrow \mathcal{T}$ , on a*

$$q([0, 1]) = \varphi_{q(0), q(1)}([0, d(q(0), q(1))]).$$

Un espace métrique vérifiant la première propriété est appelé *espace métrique géodésique*. La seconde propriété est une propriété d'arbre. On note en général  $[[x, y]]$  l'unique chemin injectif menant de  $x$  à  $y$  dans un arbre réel. On peut interpréter un arbre réel comme des segments réels branchés les uns aux autres de façon à former un arbre. On dit que  $x$  est une feuille de  $\mathcal{T}$  si  $\mathcal{T} \setminus \{x\}$  est connexe, et que  $x$  est un point du squelette sinon. Un point du squelette est un point de branchement si  $\mathcal{T} \setminus \{x\}$  a au moins trois composantes connexes. On note  $L(\mathcal{T})$ ,  $\text{Sk}(\mathcal{T})$  et  $\text{Br}(\mathcal{T})$  les ensembles correspondants (ces définitions sont topologiques et ne dépendent pas de la distance).

Noter que dans la construction de  $(\mathcal{T}_f, d_f)$ , la fonction  $f$  d'origine induit plus d'information que la structure d'arbre réel. Par exemple, la classe de 0 pour  $\sim_f$  est un point distingué, appelé la racine de  $\mathcal{T}_f$ .

Plus généralement, un arbre réel pointé (on dit aussi *enraciné*) est un arbre réel dans lequel on a distingué un point particulier, appelé la racine. Si  $(\mathcal{T}, d, \rho)$  est un arbre réel pointé, on appelle *hauteur* de  $x \in \mathcal{T}$  la quantité  $\text{ht}(x) = d(\rho, x)$ . Le sous-arbre de  $\mathcal{T}$  enraciné en  $x \in \text{Sk}(\mathcal{T})$  est l'ensemble

$$\text{Desc}(\mathcal{T}, x) = \{y \in \mathcal{T} : x \in [[\rho, y]]\}. \quad (1.4)$$

Par ailleurs, la mesure de Lebesgue sur  $[0, 1]$  induit une mesure de probabilité

$$\mu_f := (p_f)_* \text{Leb}_{[0,1]} \quad (1.5)$$

où  $p_f : [0, 1] \rightarrow \mathcal{T}_f$  est la projection canonique. Enfin, l'ordre naturel sur  $[0, 1]$  induit un ordre total sur l'arbre. Ces considérations sur le codage des arbres par des fonctions sont traitées en profondeur dans Duquesne [34].

### 1.2.2 Arbre continu brownien

Si l'on applique la construction précédente à la fonction aléatoire  $f = \mathfrak{e}$ , on obtient l'objet habituellement appelé *l'arbre continu aléatoire brownien* [2, 3, 62]. Cet arbre réel (aléatoire) possède des propriétés surprenantes. Par exemple, la mesure  $\mu_{\mathfrak{e}}$  héritée de la mesure de Lebesgue est diffuse et supportée par l'ensemble des feuilles  $L(\mathcal{T}_{\mathfrak{e}})$ , presque-sûrement, c'est-à-dire que l'on a

$$\mu_{\mathfrak{e}}(\text{Sk}(\mathcal{T}_{\mathfrak{e}})) = 0, \quad \text{p.s.}$$

Ainsi, les feuilles de  $\mathcal{T}_{\mathfrak{e}}$  forment un ensemble dense dans  $\mathcal{T}_{\mathfrak{e}}$ , qui est extrêmement « touffu ». Par ailleurs, la propriété bien connue d'unicité des minima locaux de  $\mathfrak{e}$  implique que l'arbre  $\mathcal{T}_{\mathfrak{e}}$  est presque-sûrement *binnaire*, c'est-à-dire que si  $x \in \text{Br}(\mathcal{T}_{\mathfrak{e}})$  alors  $\mathcal{T}_{\mathfrak{e}} \setminus \{x\}$  a exactement trois composantes connexes.

### 1.2.3 Topologie de Gromov-Hausdorff

Une question reste néanmoins en suspens : pour pouvoir considérer  $(\mathcal{T}_{\mathfrak{e}}, d_{\mathfrak{e}})$  comme une variable aléatoire, nous devons munir l'ensemble des arbres réels (compacts) d'une topologie, ou en tout cas d'une tribu adaptée.

## 1 Introduction et tour d'horizon

De telles topologies ont été développées en Géométrie depuis une trentaine d'années, d'après des idées de Gromov [48]. Leur utilisation en Probabilités dans le contexte d'arbres réels aléatoires a été popularisée par Evans et ses coauteurs [39, 41, 42], et a appliquée avec succès dans diverses situations, par exemple Duquesne & Le Gall [36], Le Gall [65]. Si  $(X, d)$  est un espace métrique, on note  $[X, d]$  sa classe d'isométrie, c'est-à-dire la collection des espaces isométriques à  $(X, d)$ .

Dans la suite, il nous arrivera fréquemment de commettre l'abus de notation consistant à prendre un point  $x$  dans une classe d'isométrie  $[X, d]$ , pour dire que l'on considère un point dans un représentant, nous prendrons garde à ce que les propriétés alors mises en jeu soient toujours des propriétés de classe. Par exemple, « l'ensemble des points de branchement de  $\mathcal{T}$  », où  $\mathcal{T}$  est une classe d'isométrie d'arbre réel compact, a un sens non-ambigu.

Nous notons  $\mathbb{M}$  l'ensemble des classes d'isométrie d'espaces métriques compacts, et nous le munissons de la *distance de Gromov-Hausdorff* définie par

$$d_{\text{GH}}([X, d], [X', d']) = \inf_{\phi, \phi'} \delta_H(\phi(X), \phi'(X')), \quad (1.6)$$

où l'infimum est pris sur l'ensemble des plongements isométriques  $\phi, \phi'$  de  $(X, d), (X', d')$  dans un espace métrique commun  $(Z, \delta)$ , et où  $\delta_H$  est la distance de Hausdorff usuelle entre les compacts de  $Z$ . On pourra consulter [24, Chapter 7] pour les premières propriétés de cette distance, qui fait de  $\mathbb{M}$  un espace métrique polonais (c'est une extension immédiate de la preuve de [41, Theorem 1]).

Au sein de  $\mathbb{M}$ , le sous-espace  $\mathbb{T}$  formé des (classes d'isométrie d') arbres réels compacts est fermé [41], et  $(\mathbb{T}, d_{\text{GH}})$  est donc lui-même un espace métrique polonais. On vérifie que la fonction de  $\{f \in \mathcal{C}([0, 1], \mathbb{R}_+) : f(0) = f(1) = 0\}$  dans  $\mathbb{T}$  qui à  $f$  associe  $[\mathcal{T}_f, d_f]$  est continue, si l'on munit le premier espace de la norme uniforme. De ce fait, la classe d'isométrie de l'arbre brownien, notée

$$\mathcal{T}^{(2)} := [\mathcal{T}_e, d_e],$$

est bien une variable aléatoire à valeurs dans  $\mathbb{T}$ . On l'appellera aussi, avec un petit abus de notation, l'arbre continu brownien.

Avec ces considérations, on peut alors interpréter la convergence (1.1) comme une véritable convergence dans un espace d'arbres réels. A tout arbre plan enraciné  $\mathbf{t} \in \mathbb{T}$ , on associe naturellement un élément de  $\mathbb{T}$ , noté  $\tilde{\mathbf{t}}$ , en conférant à chaque arête de  $\mathbf{t}$  une longueur unité. Plus précisément, on peut construire un représentant de  $\tilde{\mathbf{t}}$  comme sous-ensemble de l'espace  $\mathbb{R}^{\mathbf{t}}$ , muni de la norme  $\|\mathbf{x}\|_1 = \sum_{u \in \mathbf{t}} x_u$ , de la façon suivante. Pour tout  $u \in \mathbf{t} \setminus \{\emptyset\}$ , on considère le segment unité  $s_u = [0, e_u] \subset \mathbb{R}^{\mathbf{t}}$ , où  $e_u$  est le vecteur de la base canonique de  $\mathbb{R}^{\mathbf{t}}$  indicé par  $u$ . L'union

$$\bigcup_{u \in \mathbf{t} \setminus \{\emptyset\}} \left( \sum_{v \prec u} e_v + s_u \right), \quad (1.7)$$

munie de la métrique induite de  $\|\cdot\|_1$ , est le représentant recherché. La convergence (1.1)

implique alors la convergence en loi pour la topologie de  $(\mathbb{T}, d_{\text{GH}})$  :

$$\frac{1}{\sqrt{n}} \tilde{T}_n \xrightarrow[n \rightarrow \infty]{(d)} \frac{2}{\sigma} \mathcal{T}^{(2)}, \quad (1.8)$$

où par définition on note  $a[X, d] = [X, ad]$ . Cette dernière convergence ne fait plus référence qu'aux arbres réels.

Nous avons déjà vu que les arbres réels de la forme  $(\mathcal{T}_f, d_f)$  sont naturellement enracinés en la classe de 0 pour  $\sim_f$ . De ce fait, il est utile de considérer une extension de la métrique de Gromov-Hausdorff à des classes d'isométrie d'espaces métriques pointés  $(X, d, x)$ , où  $x \in X$  est un point distingué. Deux tels espaces sont isométriques s'il existe une isométrie entre les deux qui fait correspondre les points distingués. On pose, en généralisant (1.6) de façon évidente,

$$d_{\text{GH}}([X, d, x], [X', d', x']) = \inf_{\phi, \phi'} (\delta(\phi(x), \phi(x')) \vee \delta_H(\phi(X), \phi'(X'))) , \quad (1.9)$$

munissant l'espace  $\mathbb{M}^*$  des classes d'isométrie d'espaces pointés d'une métrique séparable complète. La convergence (1.8) peut être étendue en une convergence pour cette métrique, où  $\tilde{T}_n$  est naturellement pointé en  $\mathbf{0} \in \mathbb{R}^{T_n}$  dans la construction (1.7) du représentant de  $\tilde{T}_n$ , et où l'arbre brownien  $\mathcal{T}_e$  est enraciné en la classe  $[0]_{\sim_e}$  de 0 pour  $\sim_e$  (on notera indifféremment  $\mathcal{T}^{(2)}$  pour la classe de  $[\mathcal{T}_e, d_e, [0]_{\sim_f}]$ ).

### 1.2.4 Arbres stables

L'arbre continu brownien  $\mathcal{T}^{(2)}$  est un objet universel, au sens où il est limite de tout arbre de Galton-Watson critique conditionné, dont la loi de reproduction a une variance finie. Lorsque l'on ne suppose plus que la variance de  $\mu$  est finie, une grande famille de limites possibles apparaît, et porte le nom *d'arbres de Lévy*. Ces derniers ont été introduits dans Le Gall & Le Jan [67] et Duquesne & Le Gall [63], dans le but de coder la généalogie des processus de branchement continus, qui apparaissent comme limites de processus de Galton-Watson. Parmi les arbres de Lévy se distingue une sous-classe importante, dite des *arbres stables*. Nous les introduisons en suivant Duquesne [33] et Duquesne & Le Gall [36].

Supposons que la loi critique  $\mu$  satisfasse

$$\mu(k) \underset{k \rightarrow \infty}{\sim} c k^{-1-\alpha},$$

où  $c > 0$  est une constante et  $\alpha \in ]1, 2[$ . Dans ce cas,  $\mu$  est dans le domaine d'attraction d'une loi stable d'indice  $\alpha$  spectralement positive. Plus précisément, si  $X_1, \dots, X_n$  sont indépendantes de loi  $\mu$ , alors on a la convergence en loi

$$\frac{\sum_{i=1}^n X_i - n}{n^{1/\alpha}} \xrightarrow[n \rightarrow \infty]{(d)} X_1^{(\alpha)},$$

où  $X_1^{(\alpha)}$  est une variable aléatoire à valeurs dans  $\mathbb{R}$  de transformée de Laplace

$$E[\exp(-\lambda X_1^{(\alpha)})] = \exp\left(c \frac{\Gamma(2-\alpha)}{\alpha(\alpha-1)} \lambda^\alpha\right). \quad (1.10)$$

Soit alors, pour chaque  $n$  suffisamment grand pour que  $P_\mu(\mathbf{T}_n) > 0$ , un arbre aléatoire  $T_n$  de loi  $P_\mu(\cdot | \mathbf{T}_n)$ . Avec les notations de 1.2.3, on a alors la convergence en loi [33, 36] dans  $(\mathbb{T}, d_{\text{GH}})$  :

$$\frac{1}{n^{1-1/\alpha}} \tilde{T}_n \xrightarrow[n \rightarrow \infty]{(d)} \left( c \frac{\Gamma(2-\alpha)}{\alpha(\alpha-1)} \right)^{1-1/\alpha} \mathcal{T}^{(\alpha)}, \quad (1.11)$$

où la limite  $\mathcal{T}^{(\alpha)}$  est une variable aléatoire dans  $\mathbb{T}$  appelée *arbre stable d'indice  $\alpha$* . On peut étendre la convergence à des arbres enracinés comme à la section précédente, et l'arbre  $\mathcal{T}^{(\alpha)}$  est naturellement enraciné. Par convention, l'arbre stable d'indice 2 est l'arbre continu brownien,  $\mathcal{T}^{(2)}$ , et la notation est cohérente.

Par ailleurs, l'arbre stable peut s'écrire sous la forme  $[\mathcal{T}_{H^{(\alpha)}}, d_{H^{(\alpha)}}]$  pour un certain processus aléatoire  $H^{(\alpha)}$  à valeurs dans  $\mathcal{C}([0, 1], \mathbb{R}_+)$  appelé processus des hauteurs [35, 33]. Nous n'en détaillerons pas la construction et considérerons ce fait pour acquis.

Les arbres stables d'indice  $\alpha \in ]1, 2[$  ont une structure très différente de l'arbre brownien. Tandis que ce dernier est un arbre binaire p.s., on montre [36] que presque-sûrement, si  $x \in \text{Br}(\mathcal{T}^{(\alpha)})$ , alors  $\mathcal{T}^{(\alpha)} \setminus \{x\}$  a une infinité de composantes connexes. Tout point de branchement est donc d'ordre infini.

### 1.2.5 Arbres continus aléatoires mesurés

Du fait des précédentes considérations, il est naturel d'appeler un arbre tel que  $\mathcal{T}^{(2)}$  un arbre continu aléatoire. Dans l'article [3], Aldous développe des résultats généraux sur des familles d'arbres continus aléatoires (CRT), en considérant plutôt des couples  $(\mathcal{T}, \mu)$  où  $\omega \mapsto (\mathcal{T}(\omega), \mu(\omega))$  est une « variable aléatoire » telle que  $\mathcal{T}$  est un arbre réel enraciné et  $\mu$  est une mesure de probabilités borélienne sur  $\mathcal{T}$  vérifiant quelques propriétés naturelles.

En fait, Aldous ne travaille généralement pas directement sur ce couple et considère plutôt des *arbres réduits*, appelés encore les *marginales* de  $(\mathcal{T}, \mu)$ . Intuitivement, ce sont les sous-arbres  $\mathcal{T}_{[k]}$  de  $\mathcal{T}$  engendrés<sup>1</sup> par la racine et les  $k$  premiers termes d'une suite échangeable de points de  $\mathcal{T}$  de mesure directrice  $\mu$ , c'est-à-dire une suite i.i.d. conditionnellement à  $(\mathcal{T}, \mu)$ , de loi  $\mu$ . La définition de ces arbres réduits en tant que variables aléatoires ne pose pas de problème, puisque ces arbres peuvent être essentiellement décrits par une suite de longueurs d'arêtes paramétrées par une structure d'arbre discret, et leurs propriétés d'échangeabilité permettent d'obtenir, *via* une *construction spéciale* du type de (1.7), un représentant particulier de la classe d'isométrie de  $(\mathcal{T}, \mu)$  comme sous-ensemble de  $(l^1(\mathbb{N}), |\cdot|_1)$  muni d'une mesure aléatoire.

Cette approche évite, entre autres, d'avoir à définir une topologie sur l'espace des arbres réels mesurés, un arbre continu étant, *in fine*, une famille d'arbres réels enracinés  $(\mathcal{T}_{[k]}, k \geq 1)$  où  $\mathcal{T}_{[k]}$  a exactement  $k$  feuilles distinctes de la racine, étiquetées de 1 à  $k$ , et vérifiant la *propriété de cohérence* suivante :  $\mathcal{T}_{[k]}$  a même loi que le sous-arbre de  $\mathcal{T}_{[k+1]}$  engendré par la racine et  $k$  de ses feuilles prises au hasard et renumérotées uniformément au hasard.

<sup>1</sup>Le sous-arbre de  $\mathcal{T}$  engendré par les points  $x_1, \dots, x_k$  est le sous-ensemble  $\cup_{1 \leq i, j \leq k} [[x_i, x_j]]$  de  $\mathcal{T}$ , muni de la métrique induite

Par ailleurs, on peut munir l'espace des arbres réels mesurés de diverses topologies. Celles-ci permettent d'éviter les nuances ci-dessus, qui obligent à considérer certains représentants particuliers de la classe d'isométrie de  $\mathcal{T}$ . Voir [45, 42, 47] et [P16].

**Définition 1.3** *Un espace métrique mesuré est un triplet  $(X, d, \mu)$  où  $(X, d)$  est un espace métrique et  $\mu$  est une mesure de probabilités borélienne sur  $X$ .*

*Deux espaces métriques mesurés  $(X, d, \mu), (X', d', \mu')$  sont isométriques s'il existe une isométrie bijective entre les deux, qui fait correspondre les mesures.*

On notera  $[X, d, \mu]$  la classe d'isométrie de  $(X, d, \mu)$ . La topologie dite *de Gromov-Prokhorov*, considérée par Greven, Pfaffelhuber & Winter [47] (certaines idées apparaissent aussi dans [48]) est la plus proche en esprit de l'approche initiale d'Aldous<sup>2</sup>. On définit ainsi, pour deux classes d'isométrie  $\mathcal{X} = [X, d, \mu], \mathcal{X}' = [X', d', \mu']$  d'espaces métriques mesurés séparables complets (et plus nécessairement compacts),

$$d_{\text{GP}}(\mathcal{X}, \mathcal{X}') = \inf_{\phi, \phi'} \delta_P(\phi_*\mu, \phi'_*\mu'),$$

où comme avant  $\phi, \phi'$  sont des isométries de  $X, X'$  dans un espace commun  $(Z, \delta)$ , et où  $\delta_P$  est la distance de Prokhorov sur les mesures de probabilité sur  $Z$  (dont la topologie coïncide avec celle de la convergence étroite) :

$$\delta_P(\nu, \nu') = \inf\{\varepsilon > 0 : \nu(C) \leq \nu'(C^\varepsilon) + \varepsilon, \forall C \text{ borélien}\},$$

où  $C^\varepsilon$  est le  $\varepsilon$ -voisinage de  $C$ . Ceci munit l'espace des classes d'isométries d'espaces polonais mesurés  $[X, d, \mu]$  de support plein (c'est-à-dire telles que  $X = \text{supp } \mu$ ) d'une métrique complète, séparable.

Si les arbres continus considérés sont compacts (Aldous [3] donne un critère pour cela en terme des arbres réduits), la topologie décrite dans [42] ou [P16] convient aussi, nous reviendrons sur cette dernière en 3.3.

Si  $(\mathcal{T}_o, d_o, \mu_o)$  est un arbre réel mesuré, il nous arrivera de considérer de façon abusive que sa classe d'isométrie  $[\mathcal{T}_o, d_o, \mu_o]$  donne lieu à une « mesure borélienne »  $\mu$  sur  $\mathcal{T} = [\mathcal{T}_o, d_o]$  : on prendra garde dans l'écriture à ce que les quantités mesurées soient toujours non-ambigües. Par exemple, si  $\mathcal{T}_o$  est un arbre pointé,  $\mu(\{x \in \mathcal{T} : \text{ht}(x) > t\})$  sera une notation pour  $\mu_o(\{x \in \mathcal{T}_o : \text{ht}(x) > t\})$ . Cette écriture est similaire à celle qui consiste à considérer des points dans un élément de  $\mathbb{M}$ .

Ainsi, pour  $\alpha \in ]1, 2]$ , nous noterons  $\mu^{(\alpha)}$  la « mesure » sur l'arbre stable  $\mathcal{T}^{(\alpha)}$  correspondant à la classe  $[\mathcal{T}_{H^{(\alpha)}}, d_{H^{(\alpha)}}, \mu_{H^{(\alpha)}}]$ , avec les notations (1.5), où  $H^{(\alpha)}$  est le processus des hauteurs de l'arbre stable.

### 1.2.6 p-arbres d'anniversaire

Nous introduisons maintenant une seconde famille d'arbres aléatoires discrets et leurs limites continues, qui jouent un rôle important dans nos travaux. Ces arbres ne sont pas

---

<sup>2</sup>Il y a néanmoins certaines différences, comme le fait que la mesure  $\mu$  sur l'arbre continu peut ne pas être de support plein dans les travaux d'Aldous, mais nous ne rentrerons pas dans ces détails

des arbres plans, mais des arbres étiquetés à  $n$  sommets, c'est-à-dire des arbres couvrants du graphe complet de sommets  $\{1, 2, \dots, n\} =: [n]$ . On les appelle *arbres de Cayley*, et leur nombre est  $n^{n-2}$ . On dit qu'un tel arbre est enraciné si l'on a distingué un de ses sommets. L'enracinement permet d'orienter les arêtes de l'arbre de façon canonique, de façon à ce qu'elles pointent toutes dans la direction opposée à la racine. Si  $\hat{\mathbf{t}}$  est un arbre de Cayley, on note  $c_i(\hat{\mathbf{t}})$ ,  $1 \leq i \leq n$  le nombre d'arêtes qui pointent depuis le sommet  $i$ , interprété comme le nombre d'enfants de  $i$ . Soit  $\hat{\mathbf{T}}_n$  l'ensemble des  $n^{n-1}$  arbres de Cayley enracinés à  $n$  sommets. On peut considérer un élément de  $\hat{\mathbf{T}}_n$  comme un arbre réel en faisant de chaque arête un segment de longueur 1, par une construction du type de (1.7), que nous ne détaillerons pas. On notera  $(\tilde{\mathbf{t}}, d)$  l'arbre réel associé à l'arbre de Cayley  $\hat{\mathbf{t}}$ .

On se donne une probabilité  $\mathbf{p} = (p_1, \dots, p_n)$  sur  $[n]$ , telle que  $\min p_i > 0$ , et on construit un élément de  $\hat{\mathbf{T}}_n$  de la façon suivante. Sur un espace probabilisé  $(\Omega, \mathcal{F}, P)$ , soit  $(X_i, i \geq 0)$  une suite i.i.d. de variables aléatoires de loi  $\mathbf{p}$ . À l'étape  $i \geq 0$ , on trace une arête allant du sommet d'étiquette  $X_i$  au sommet d'étiquette  $X_{i+1}$  à la condition que cette arête ne crée pas de cycle dans le graphe tracé lors des étapes antérieures. Cette procédure (l'algorithme d'Aldous-Broder) termine en temps fini, et donne un arbre de Cayley  $\hat{T}^{\mathbf{p}}$  qui est naturellement enraciné en  $X_0$ . Cet arbre est appelé  *$\mathbf{p}$ -arbre d'anniversaire*, ou plus simplement  *$\mathbf{p}$ -arbre*. On montre que la loi de  $\hat{T}^{\mathbf{p}}$  est donnée par les formules

$$P(\hat{T}^{\mathbf{p}} = \hat{\mathbf{t}}) = \prod_{i \in [n]} p_i^{c_i(\hat{\mathbf{t}})}, \quad \hat{\mathbf{t}} \in \hat{\mathbf{T}}.$$

Les  $\mathbf{p}$ -arbres d'anniversaire sont naturellement associés au problème (« paradoxe ») bien connu des anniversaires, selon lequel la probabilité que deux individus d'un groupe aient la même date d'anniversaire est supérieure à  $1/2$  dès que le groupe compte au moins 23 membres. Dans ce problème, on suppose généralement que les dates d'anniversaires sont équiprobables parmi les  $n = 365$  jours des années non-bissextiles, mais une situation plus réaliste est d'associer à chaque jour une probabilité spécifique — ainsi, on peut par exemple prendre en compte le 29 février en lui attribuant un poids 4 fois moins important que les autres jours. La probabilité que tous les membres du groupe aient des dates d'anniversaire différentes est alors la probabilité que  $\hat{T}^{\mathbf{p}}$  soit le graphe-ligne pour lequel  $c_i(\hat{\mathbf{t}}) \in \{0, 1\}$  pour tout  $i \in [n]$ .

Les  $\mathbf{p}$ -arbres sont des structures aléatoires naturelles qui apparaissent dans différents contextes. On évoque leur application à la coalescence additive en section 1.5. À la section 5.2, on les utilisera pour décrire, suivant [P4,P8] un modèle naturel d'application aléatoire de  $[n]$  dans lui même, et nous tirerons parti de l'étude des limites des  $\mathbf{p}$ -arbres discutée dans le paragraphe suivant et à la section 5.1.3 (suivant [P6]) pour décrire le comportement lorsque  $n \rightarrow \infty$  de ces applications.

### 1.2.7 Arbres continus inhomogènes

Les limites d'échelle possibles des  $\mathbf{p}$ -arbres ont été décrites par Camarri & Pitman [25], dans l'esprit du paragraphe précédent plutôt que pour une métrique de type Gromov-



Hausdorff. Ces limites sont paramétrées par l'ensemble

$$\Theta := \left\{ \theta = (\theta_1, \theta_2, \dots) : \begin{array}{l} \theta_1 \geq \theta_2 \geq \dots \geq 0, \\ \sum_{i \geq 1} \theta_i^2 \leq 1, \\ \sum_{i \geq 1} \theta_i^2 < 1 \text{ ou } \sum_{i \geq 1} \theta_i = \infty \end{array} \right\} \quad (1.12)$$

On notera, pour  $\theta \in \Theta$ ,  $\theta_0 = (1 - \sum_{i \geq 1} \theta_i^2)^{1/2}$ . La dernière condition dans la définition de  $\Theta$  assure que les limites  $\mathcal{T}^\theta$  considérées plus bas n'ont pas une structure « singulière » en un certain sens, qui rend leur étude moins intéressante.

Soit  $\theta \in \Theta$ . On considère un processus de Poisson ponctuel  $\{(U_i, V_i), i \geq 1\}$  sur le premier octant

$$\text{Oct} = \{(x, y) \in \mathbb{R}^2 : 0 < y < x\},$$

dont l'intensité est  $\theta_0 dx dy \mathbb{1}_{\text{Oct}}(x, y)$ . Notons que la projection  $(U_i, i \geq 1)$  est de ce fait un processus de Poisson ponctuel sur  $\mathbb{R}_+$  d'intensité  $\theta_0 x dx$ . Par ailleurs, on se donne une famille de processus de Poisson ponctuels indépendants (et indépendants du premier)  $(\xi_{i,j}, j \geq 1)$  pour  $i \geq 1$ , d'intensité  $\theta_i dx$  sur  $\mathbb{R}_+$ . Les points  $(V_i, i \geq 1)$  ainsi que les points  $(\xi_{i,1}, i \geq 1)$  sont distingués des autres et sont appelés *points de jonction*, tandis que les points  $U_i, i \geq 1, \xi_{i,j}, i \geq 1, j \geq 2$  sont appelés *points de coupure*. Si  $\eta$  est un point de coupure, on notera  $\eta^* < \eta$  le point de jonction associé, c'est-à-dire  $U_i^* = V_i$  et  $\xi_{i,j}^* = \xi_{i,1}$ .

On montre grâce à l'hypothèse faite sur  $\theta$  que l'on peut presque-sûrement ordonner les points de coupure en  $0 < \eta_1 < \eta_2 < \dots$ . Soit  $e_i$  l'élément de  $l^1(\mathbb{N})$  défini par  $e_i(j) = \delta_{ij}$ . On construit une fonction  $f : \mathbb{R}_+ \rightarrow l^1(\mathbb{N})$  par  $f(0) = 0$ , et récursivement, si l'on connaît la restriction de  $f$  à  $[0, \eta_k]$ , on pose

$$f(\eta_k + s) = f(\eta_k^*) + (s - \eta_k) e_{k+1}, \quad 0 < s \leq \eta_{k+1} - \eta_k.$$

L'espace  $(f([0, \eta_k]), |\cdot|_1)$  est alors un arbre réel, noté  $\mathcal{T}_{[k]}$ , et naturellement enraciné en  $0 \in l^1(\mathbb{N})$ . Intuitivement, il est obtenu en branchant l'extrémité gauche du segment réel  $[\eta_{k-1}, \eta_k]$  au point de branchement  $\eta_k^*$ , dont la position sur  $\mathcal{T}_{[k-1]}$  est déterminée par les étapes précédentes. Les  $\mathcal{T}_{[k]}$  sont emboîtés, et l'on pose

$$\mathcal{T}_\circ^\theta = \overline{\bigcup_{k \geq 1} \mathcal{T}_{[k]}},$$

où  $\bar{A}$  est l'adhérence de  $A$ . La classe<sup>3</sup> d'arbres réels

$$\mathcal{T}^\theta = [\mathcal{T}_\circ^\theta, |\cdot|_1]$$

<sup>3</sup>Il y a une subtilité, qui tient au fait que pour certaines valeurs de  $\theta$ , l'arbre  $\mathcal{T}_\circ^\theta$  n'est p.s. pas compact, et l'on doit faire appel à d'autres topologies que celle de  $(\mathbb{T}, d_{\text{GH}})$ . On a recours par exemple à des topologies sur les espaces métriques localement compacts, comme dans [37]. Comme suggéré en 1.2.5, le plus adapté est de considérer la métrique de Gromov-Prokhorov sur les espaces métriques polonais mesurés [47], et de considérer la classe  $[\mathcal{T}_\circ^\theta, |\cdot|_1, \mu_\circ^\theta]$  comme l'arbre continu inhomogène de paramètre  $\theta$ .

## 1 Introduction et tour d'horizon

est appelée *arbre continu inhomogène* de paramètre  $\theta$ . On définit enfin

$$\mu_{\circ}^{\theta} = \lim_{k \rightarrow \infty} \frac{1}{k+1} \sum_{x \in \mathbf{L}(\mathcal{T}_{[k]})} \delta_x \quad (1.13)$$

comme la limite (p.s. pour la convergence faible des mesures sur  $l^1(\mathbb{N})$ ) de la mesure empirique sur les feuilles de  $\mathcal{T}_{[k]}$ , et on note  $\mu^{\theta}$  la « mesure » induite sur  $\mathcal{T}^{\theta}$ . Ceci définit un arbre continu aléatoire au sens discuté dans 1.2.5, les théorèmes d'Aldous [3] montrant que si  $X_1, X_2, \dots$  est une suite i.i.d. de loi  $\mu^{\theta}$  conditionnellement à  $(\mathcal{T}^{\theta}, \mu^{\theta})$ , alors le sous-arbre de  $\mathcal{T}^{\theta}$  engendré par la racine et les points  $X_i, 1 \leq i \leq k$ , noté  $\mathcal{T}'_k$ , est tel que  $\mathcal{T}'_k$  a même loi que  $[\mathcal{T}_{[k]}, |\cdot|_1]$ .

Nous verrons en section 5.1, suivant [P6], comment on peut définir  $\mathcal{T}^{\theta}$  sous la forme  $[\mathcal{T}_H, d_H]$  où  $H$  est un « processus des hauteurs » idoine.

Qualitativement, un arbre continu inhomogène possède deux types de branchements : des branchements binaires, correspondant aux points de jonction  $f(V_i)$  qui ne servent à planter qu'une seule branche, et des branchements infinis, comme pour les arbres stables, qui correspondent aux points de jonction  $f(\xi_{i,1})$ . Il y a donc autant de points de branchement infinis que de  $i \geq 1$  tels que  $\theta_i > 0$ .

Lorsque  $\theta_0 = 1$  ( $\theta = \mathbf{0} \in \Theta$ ), l'arbre  $\mathcal{T}^{\mathbf{0}}$  ainsi construit n'est autre que (deux fois) l'arbre brownien  $2\mathcal{T}^{(2)}$ . De plus, la mesure  $\mu^{\mathbf{0}}$  construite en (1.13) est égale à  $\mu^{(2)}$ , c'est-à-dire que  $[\mathcal{T}_{\mathbf{e}}, 2d_{\mathbf{e}}, \mu_{\mathbf{e}}] =_{(d)} [\mathcal{T}_{\circ}^{\mathbf{0}}, |\cdot|_1, \mu_{\circ}^{\mathbf{0}}]$ . On peut déduire ces dernières propriétés de [3] ou [62]. Cette approche est la construction originale de l'arbre brownien, due à Aldous [2], et en permet une simulation aisée, voir la figure 1.2. Cette figure montre bien le caractère « touffu » de l'arbre, qui semble dessiner presque une surface dans  $\mathbb{R}^3$ . Ceci est lié au fait que la dimension de Hausdorff de  $\mathcal{T}^{(2)}$  est égale à 2 presque-sûrement [36] : nous reviendrons incidemment sur ce calcul de dimension en section 4.3.

### 1.2.8 Limites des p-arbres

On montre que les arbres continus inhomogènes  $\mathcal{T}^{\theta}$  sont limites des **p**-arbres introduits plus haut, au sens suivant. Si **p** est une loi de probabilités sur  $[n]$ , on note

$$\sigma(\mathbf{p}) := \sqrt{\sum_{i \geq 1} p_i^2}.$$

Soit  $\mathbf{p}_n = (p_{ni}, 1 \leq i \leq n)$  une suite de probabilités sur  $[n]$  (respectivement) avec  $p_{n1} \geq p_{n2} \geq \dots \geq p_{nn} > 0$ . On les étend de façon naturelle en probabilités sur  $\mathbb{N}$ . On suppose qu'il existe  $\theta \in \Theta$  tel que  $(\mathbf{p}_n, n \geq 1)$  remplit la condition

$$\max_{i \in \mathbb{N}} p_{ni} \xrightarrow{n \rightarrow \infty} 0 \quad \text{et} \quad \frac{p_{ni}}{\sigma(\mathbf{p}_n)} \rightarrow \theta_i, \quad i \geq 1. \quad (1.14)$$

Pour chaque  $n \geq 1$ , on se donne un  $\mathbf{p}_n$ -arbre  $\hat{T}^n$ . On fixe  $k \geq 1$  et on se donne des variables aléatoires indépendantes  $X_1^n, \dots, X_k^n$  de loi  $\mathbf{p}_n$ , indépendantes de  $\hat{T}^n$ . Ces sommets aléatoires correspondent naturellement à des sommets de l'arbre réel  $\tilde{T}^n$  associé à

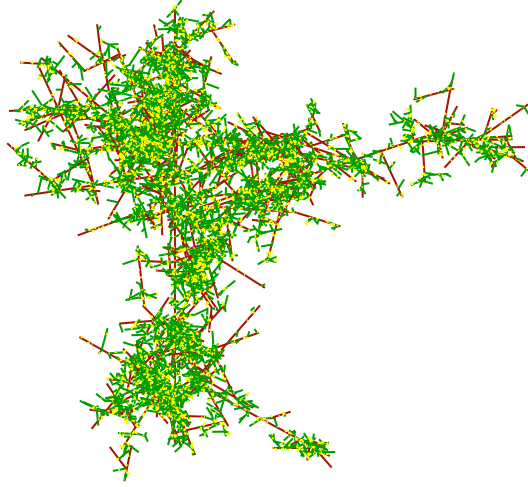


FIG. 1.2: Les 5000 premiers segments de l'arbre continu brownien, représentés dans  $\mathbb{R}^3$ , en donnant à chaque branche une orientation uniforme sur la sphère unité. Le lecteur possédant une version couleur du présent document constatera que les branches sont en marron, vert ou jaune selon leurs longueurs, des plus grandes aux plus petites.

$\hat{T}^n$ , on les note encore  $X_1^n, \dots, X_k^n$ . Soit  $\tilde{T}_k^n$  le sous-arbre engendré par la racine et ces points, qui est à son tour un arbre réel enraciné avec  $k$  points numérotés. Alors, lorsque  $n \rightarrow \infty$ , on a la convergence en loi pour la topologie de  $(\mathbb{T}, d_{\text{GH}})$  :

$$\sigma(\mathbf{p}_n)\tilde{T}_k^n \xrightarrow[n \rightarrow \infty]{(d)} [\mathcal{T}_{[k]}^\theta, |\cdot|_1].$$

Cette dernière convergence est plus faible que la convergence des arbres  $\sigma(\mathbf{p}_n)\tilde{T}^n$  vers  $\mathcal{T}^\theta$  au sens de Gromov-Hausdorff, on donnera en section 5.1 une conditions assurant cette convergence plus forte. On montre que la convergence ci-dessus est équivalente à la convergence de l'arbre mesuré  $(\tilde{T}^n, \tilde{\mathbf{p}}_n)$  vers  $(\mathcal{T}^\theta, \mu^\theta)$  pour la métrique de Gromov-Prokhorov  $d_{\text{GP}}$  définie en 1.2.5, où  $\tilde{\mathbf{p}}_n$  est la mesure image de  $\mathbf{p}_n$  par l'application  $\hat{T}^n \mapsto \tilde{T}^n$ .

Noter que si l'on prend  $\mathbf{p}_n = (1/n, \dots, 1/n)$ , la condition (1.14) est vérifiée avec  $\theta = \mathbf{0}$ , de sorte que l'arbre limite est l'arbre continu brownien. Cela peut être retrouvé en partant du fait que le  $\mathbf{p}_n$ -arbre a même loi qu'un arbre de Galton-Watson de loi de reproduction Poisson de moyenne 1, conditionné à avoir  $n$  sommets, que l'on a étiqueté uniformément au hasard avant « d'oublier » la structure planaire de l'arbre. Ce résultat est donc étroitement lié à (1.8).

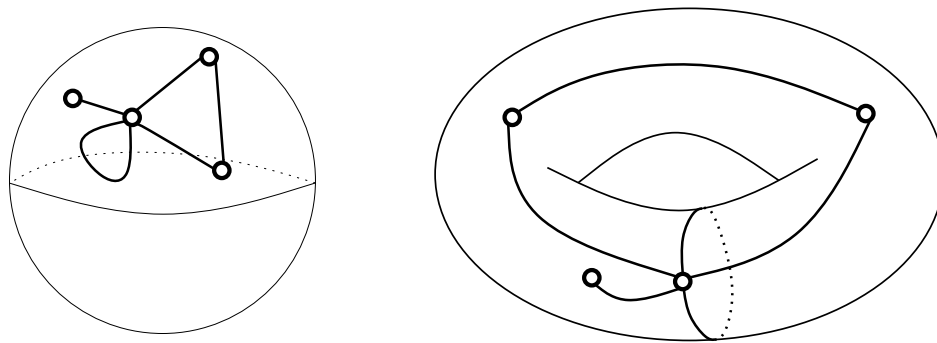


FIG. 1.3: Deux cartes, en genre 0 et 1

### 1.3 Cartes aléatoires

Le thème sur lequel j'ai concentré la plus grande part de mes efforts depuis la fin de ma thèse est l'étude des limites d'échelle de cartes aléatoires. Nous donnons ici les premières définitions ainsi que quelques uns des outils de base que nous utilisons.

On appelle *graphe plongé*, un graphe<sup>4</sup> tracé sur une surface de sorte que les arêtes ne sont autorisées à se rencontrer, éventuellement, qu'en leurs extrémités. Les surfaces considérées sont orientées, compactes, connexes et sans bord. Un théorème bien connu de classification dit que ces surfaces sont déterminées, à homéomorphisme près, par la somme connexe de  $g$  tores, où le nombre entier  $g \geq 0$ , appelé genre, est interprété comme le nombre d'anses de la surface. En sus de la structure de graphe sous-jacente, c'est-à-dire des relations d'incidence des arêtes et des sommets, un graphe plongé détermine des *faces*, qui sont les composantes connexes du complémentaire du plongement. Le plongement est dit *cellulaire* si les faces sont toutes homéomorphes à des disques ; un graphe plongé de façon cellulaire est nécessairement connexe. Dans un tel plongement, le *degré* d'une face est le nombre d'arêtes orientées que l'on rencontre en faisant le tour du bord de la face.

Une *carte* est une classe d'équivalence de graphes plongés de façon cellulaire, où l'on identifie deux graphes plongés  $G, G'$  dans les surfaces  $S, S'$  s'il existe un homéomorphisme direct  $S \rightarrow S'$ , qui fait correspondre les arêtes de  $G$  et  $G'$ . Si  $\mathbf{m}$  est une carte, on notera  $V(\mathbf{m}), E(\mathbf{m}), F(\mathbf{m})$  les ensembles de ses sommets, arêtes et faces.

Avec cette définition, les cartes forment alors une famille dénombrable. Se pose donc la question de leur énumération, qui donne lieu à des problèmes difficiles et importants, quarante ans après les premiers calculs de Tutte [82] et les développements de techniques variées, allant du comptage par séries génératrices aux intégrales matricielles et à la géométrie algébrique [86, 80, 61]. Dans les problèmes d'énumération, on suppose en général les cartes *enracinées*, c'est-à-dire que l'on distingue l'une des arêtes orientées du graphe, que nous noterons généralement  $e_*$ . L'enracinement a pour effet de rendre trivial le groupe des symétries de la carte. On note  $\mathcal{M}$  l'ensemble des cartes enracinées. Il est

<sup>4</sup>boucles et arêtes multiples sont autorisées

utile de considérer, par convention, que la carte-sommet  $\bullet$ , formée d'un unique sommet bordant une face de degré 0, est un élément de  $\mathcal{M}$ , bien qu'il n'ait aucune arête.

Les cartes sont essentiellement connues pour leurs aspects les plus combinatoires, en particulier à travers le fameux théorème des quatre couleurs. Depuis une vingtaine d'années, elles apparaissent également en physique, dans le but de construire la « gravité quantique en dimension 2 ». Dans cette théorie, on considère des intégrales par rapport à des « mesures » portées par l'ensemble des variétés Riemanniennes de dimension 2. Une approche possible pour calculer avec cette mesure est de remplacer l'intégrale sur les surfaces par des sommes sur les cartes de grande taille, de topologie fixée ou non. Il est naturel de considérer les cartes comme des espaces métriques discrets, en munissant l'ensemble de leurs sommets de la métrique de graphe usuelle, que nous noterons toujours  $d_{\text{gr}}$ . Ainsi,  $d_{\text{gr}}(u, v)$  est la longueur d'une chaîne minimale d'arêtes reliant les sommets  $u$  et  $v$ . Hormis en section 3.3, nous ne parlerons à partir de maintenant que des cartes planaires, dessinées sur la sphère, surface de genre nul.

### 1.3.1 Bijection de Schaeffer

Pour son versant mathématique, l'approche évoquée au paragraphe précédent est encore à ses débuts. Des progrès ont été rendus possibles depuis l'introduction par Schaeffer [80] de bijections astucieuses entre cartes et arbres décorés, poursuivant une approche initiée par Cori & Vauquelin [32].

L'une de ces bijections fait correspondre les quadrangulations (cartes dont les faces sont toutes de degré 4) enracinées avec une famille d'arbres planaires marqués. L'avantage de cette approche est que ces arbres marqués donnent une information sur la distance  $d_{\text{gr}}$  des cartes qu'ils codent. Pour être plus complet, nous expliquons rapidement comment cette bijection agit.

On note  $\mathbf{Q}_n$  l'ensemble des quadrangulations (planaires) enracinées, à  $n$  faces<sup>5</sup>. On donne  $\mathbf{q} \in \mathbf{Q}_n$ , on note  $e_*$  la racine, et  $e_*^- \in V(\mathbf{q})$  son origine. Chaque sommet  $v \in V(\mathbf{q})$  reçoit alors une marque  $\ell(v) = d_{\text{gr}}(e_*^-, v)$ . Le caractère biparti d'une quadrangulation planaire, c'est-à-dire le fait que  $\mathbf{q}$  n'a pas de cycle de longueur impaire, entraîne que la suite des marques des quatre sommets bordant une face, parcourue dans le sens des trigonométrie en partant d'un sommet de marque minimale dans cette face, est forcément de la forme  $l, l + 1, l + 2, l + 1$  ou  $l, l + 1, l, l + 1$ . Selon le cas, on dit que la face est *simple* ou *confluente*.

On augmente alors le graphe de  $\mathbf{q}$  en ajoutant, à l'intérieur des faces simples, une arête joignant le sommet de marque  $l + 2$  (maximale) au sommet suivant dans le sens trigonométrique, et, à l'intérieur d'une face simple, une arête diagonale joignant les deux sommets de marque  $l + 1$  (maximale). Voir la figure 1.4. On vérifie alors que le graphe formé par les arêtes nouvellement adjointes, en effaçant les intérieurs des arêtes de  $\mathbf{q}$ , est un graphe sans cycle dessiné dans le plan. Ce dernier est naturellement enraciné selon une convention simple dépendant de la racine de  $\mathbf{q}$ , et peut être vu comme un arbre plan avec les définitions données en section 1.1.1. On le note  $\mathbf{t}$ .

<sup>5</sup>un élément de  $\mathbf{Q}_n$  a donc  $n + 2$  sommets et  $2n$  arêtes par la formule d'Euler, selon laquelle  $\#V(\mathbf{m}) - \#E(\mathbf{m}) + \#F(\mathbf{m}) = 2$  pour une carte planaire  $\mathbf{m}$

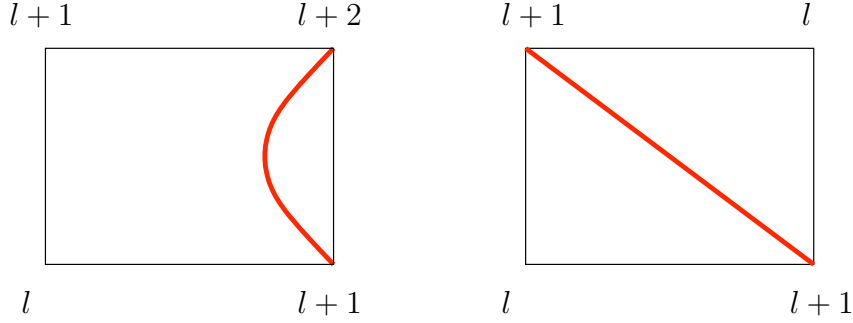


FIG. 1.4: Les arêtes supplémentaires ajoutées aux faces d'une quadrangulation

Par ailleurs, les sommets de cet arbre sont identifiés avec des sommets de  $\mathbf{q}$ . Plus précisément, tout sommet de  $\mathbf{q}$  distinct de  $e_*^-$  est un sommet de  $\mathbf{t}$ . De ce fait, les sommets de  $\mathbf{t}$  sont naturellement marqués par le truchement de la fonction  $\ell$  définie sur  $V(\mathbf{q})$ . On note  $\mathbf{l} = (\mathbf{l}(u), u \in \mathbf{t})$  la fonction de marque correspondante. Alors  $(\mathbf{t}, \mathbf{l})$  est un élément de

$$\overline{\mathbf{T}}_{n+1}^+ := \left\{ (\mathbf{t}, \mathbf{l}) \in \overline{\mathbf{T}}_{n+1} : \begin{array}{l} \mathbf{l}(\emptyset) = 1, \mathbf{l} \geq 1, \\ |\mathbf{l}(ui) - \mathbf{l}(u)| \leq 1 \forall u \in \mathbf{t}, 1 \leq i \leq c_u(\mathbf{t}) \end{array} \right\}.$$

### 1.3.2 Rayon et profil d'une quadrangulation

Une étude précise de la bijection précédente a permis à Chassaing & Schaeffer [29] de donner les comportements limite suivants. Soit  $\mathbf{q}_n$  un élément uniforme dans  $\mathbf{Q}_n$ . Soit  $u \in V(\mathbf{q}_n)$ , et

$$R(\mathbf{q}_n, u) = \max_{v \in V(\mathbf{q}_n)} d_{\text{gr}}(v, u),$$

appelé *rayon* de la carte vu du sommet  $u$ . Soit également pour  $k \geq 0$ ,

$$\mathcal{I}_{\mathbf{q}_n, u}(k) = \#\{v \in V(\mathbf{q}_n) : d_{\text{gr}}(v, u) = k\},$$

que l'on appelle *profil* de la carte vu de  $u$ . On considère  $\mathcal{I}_{\mathbf{q}_n, u}$  comme une mesure sur  $\mathbb{Z}_+$  de masse  $n+2$ . Rappelons aussi la définition de la tête du serpent brownien  $(e, Z)$ , donnée en section 1.1.3. Chassaing & Schaeffer [29] montrent alors les convergences en loi suivantes :

$$\frac{R(\mathbf{q}_n, e_*^-)}{n^{1/4}} \xrightarrow[n \rightarrow \infty]{(d)} \left(\frac{8}{9}\right)^{1/4} (\sup Z - \inf Z), \quad (1.15)$$

et

$$\frac{\mathcal{I}_{\mathbf{q}_n, e_*^-} \left( \left(\frac{8n}{9}\right)^{1/4} dx \right)}{n+2} \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{I}_\infty, \quad (1.16)$$

pour la topologie faible sur les mesures de probabilités sur  $\mathbb{R}_+$ , où  $\mathcal{I}_\infty$  est la mesure d'occupation de  $Z$  au-dessus de son infimum, définie par

$$\langle \mathcal{I}_\infty, g \rangle = \int_0^1 dt g(Z_t - \inf Z), \quad (1.17)$$

pour toute fonction mesurable positive  $g$ . Comme on peut le deviner, ces théorèmes limites sont intimement liés à des énoncés tels que (1.3). L'image de  $Q_n$  par la bijection de Schaeffer est en effet un élément  $(T_n, L_n)$  uniforme dans  $\overline{\mathbf{T}}_{n+1}^+$ , et il est bien connu qu'un arbre plan uniforme peut être vu comme un arbre de Galton-Watson conditionné. Un problème important est néanmoins que, du fait de la contrainte de positivité des marques dans la définition de  $\overline{\mathbf{T}}_{n+1}^+$ , la loi de  $T_n$  n'est pas uniforme sur  $\mathbf{T}_{n+1}$ , et  $(T_n, L_n)$  ne suit pas une loi de type  $\mathbb{P}_{\mu,\nu}(\cdot | \overline{\mathbf{T}}_n)$ . Au moins deux méthodes ont été proposées pour contourner ce problème, une dans [29] et une autre dans Le Gall [64].

Une autre possibilité est de considérer le rayon et le profil vu d'un sommet aléatoire de la carte, plutôt que de l'origine de la racine. Nous avons tiré parti de cette approche dans [P10,P11,P12], où nous nous sommes intéressés à la robustesse des résultats précédents, lorsqu'on remplace les quadrangulations par des cartes plus générales. Ceci sera traité en section 3.1.

### 1.3.3 Au-delà du rayon et du profil

Depuis l'article [29], des progrès notables ont été obtenus dans la compréhension de la structure des limites d'échelle des quadrangulations aléatoires, en particulier *via* les travaux de Le Gall et coauteurs. Comme pour les théorèmes limite sur les arbres, l'idée de considérer les quadrangulations comme des éléments de  $(\mathbb{M}, d_{\text{GH}})$  est particulièrement porteuse. En particulier, on montre que les lois des espaces métriques aléatoires  $([V(\mathbf{q}_n), n^{-1/4}d_{\text{gr}}], n \geq 1)$  forment [65] une famille tendue de l'espace des mesures de probabilités sur  $(\mathbb{M}, d_{\text{GH}})$ . Bien qu'il soit encore un problème ouvert de savoir si la limite en loi de ces espaces est unique, on peut donc considérer des limites le long de sous-suites. Le Gall [65] montre que la topologie de ces espaces limites est uniquement déterminée, comme celle d'un quotient aléatoire de l'arbre continu brownien  $(\mathcal{T}_e, d_e)$  défini en termes du processus  $Z$  (une version similaire de ce quotient apparaît déjà dans Marckert & Mokkadem [72]). Le Gall & Paulin [68] ont identifié cette topologie, qui n'est autre que celle de la sphère  $\mathbb{S}^2$  de dimension 2. Ceci donne un contenu mathématique à l'assertion selon laquelle les grandes quadrangulations aléatoires approchent une surface aléatoire. Nous reviendrons sur le théorème de Le Gall & Paulin en section 3.2, dont nous présenterons la preuve alternative de [P13]. Par ailleurs, en 3.3, nous présenterons, selon [P16], une extension de la bijection de Schaeffer permettant d'obtenir des propriétés d'unicité de la géodésique typique dans les limites de cartes, également en genre supérieur  $g \geq 1$ . Voir aussi Le Gall [66] pour d'autres résultats sur les géodésiques.

Enfin, nous mentionnons une approche des cartes aléatoires assez différente, mais présentant certains rapports, qui est l'étude des limites locales des cartes. Dans ce contexte, on n'effectue pas de renormalisation des graphes, et on étudie la structure locale de grandes cartes autour d'un sommet marqué. Différentes approches sont possibles, certaines utilisant les techniques bijectives. On consultera les travaux d'Angel & Schramm, Chassaing & Durhuus, et Krikun [9, 8, 27].

## 1.4 Fragmentations auto-similaires

Une partie importante de ma thèse a porté sur les processus de fragmentations auto-similaires, introduits par Bertoin dans [17, 19]. Voir [20] pour une exposition récente et complète. Ces processus décrivent l'évolution d'un système d'objets massifs qui se disloquent au cours du temps. Pour décrire ce système, on se contente de donner la suite des masses des objets présents dans le système à chaque instant, ignorant par exemple les aspects géométriques du système. Un processus de fragmentation prendra donc ses valeurs dans l'espace

$$\mathcal{S}^\downarrow = \left\{ \mathbf{s} = (s_1, s_2, \dots) : s_1 \geq s_2 \geq \dots \geq 0, \sum_{i \geq 1} s_i \leq 1 \right\},$$

que l'on munit de la topologie produit (de la convergence terme à terme). On notera en général  $s_0 = 1 - \sum_{i \geq 1} s_i$  le « défaut » de  $\mathbf{s} \in \mathcal{S}^\downarrow$ .

L'hypothèse de base que l'on fait est qu'un processus de fragmentation est un processus markovien satisfaisant une propriété de branchement : deux objets distincts connaissent des évolutions indépendantes. Un processus de fragmentation auto-similaire d'indice  $\beta \in \mathbb{R}$  est un processus fellerien  $(F(t), t \geq 0)$  à valeurs dans  $\mathcal{S}^\downarrow$  dont le semigroupe vérifie la propriété suivante. Sachant  $F(t) = \mathbf{s}$ , la variable aléatoire  $F(t + t')$  a même loi que le réarrangement décroissant des suites

$$s_i F^{(i)}(s_i^\beta t'), \quad i \geq 1,$$

où les processus  $F^{(i)}$  sont des copies indépendantes de  $F$ . Lorsque  $\beta = 0$ , le processus est dit homogène.

Les résultats conjoints de Bertoin [19] et Berestycki [13] montrent que les lois des processus de fragmentation auto-similaires sont entièrement déterminées par un *triplet caractéristique*  $(\beta, c, \nu)$ , où  $\beta$  est l'indice d'auto-similarité,  $c \geq 0$  est un paramètre dit d'*érosion* et  $\nu$  est une mesure  $\sigma$ -finie sur  $\mathcal{S}^\downarrow$ , appelée *mesure de dislocation*, vérifiant  $\nu(\{(1, 0, 0, \dots)\}) = 0$  et

$$\int_{\mathcal{S}^\downarrow} (1 - s_1) \nu(d\mathbf{s}) < \infty.$$

Le triplet caractéristique régit l'évolution locale du processus. De façon informelle, chaque fragment de taille  $x$  se disloque en fragments de tailles  $xs$  à taux  $x^\beta \nu(d\mathbf{s})$ , et subit une « fonte » différentiable de pente  $-cx^\beta$ . Cette caractérisation rappelle fortement celle des processus de Lévy.

Une des propriétés importantes des fragmentations est la possibilité de *formation de poussière*, c'est-à-dire qu'il se peut que la somme des masses des objets à un instant donné soit strictement inférieure à 1. Ce phénomène peut arriver de trois façons différentes :

- de la poussière se forme brusquement si  $\nu$  charge  $\{\mathbf{s} : \sum_{i \geq 1} s_i < 1\}$
- de la poussière se forme continûment si  $c > 0$
- enfin, de la poussière se forme si  $\beta < 0$ .



Le troisième phénomène est le plus intéressant. Il est dû au fait que si  $\beta < 0$ , les petits fragments tendent à se disloquer plus vite, provoquant un emballement du processus, qui s'éteint (atteint l'état  $(0, 0, \dots)$ ) en temps fini. La formation de poussière a été étudiée en détail par Haas [49, 50].

### 1.4.1 Fragmentations à valeurs partitions

L'étude des fragmentations auto-similaires par Bertoin utilise de façon clé un procédé de discrétisation de l'espace, qui permet de ramener l'étude des fragmentations auto-similaires à valeurs dans  $\mathcal{S}^\downarrow$  à des processus à valeurs dans l'espace  $\mathcal{P}_\mathbb{N}$  des partitions de  $\mathbb{N}$ . Ces processus sont échangeables, ce qui permet de tirer parti de l'étude des partitions échangeables effectuée par Kingman [56]. Rappelons que l'on dit que la partition  $\pi$  est plus fine que  $\pi'$  si tout bloc de  $\pi$  est inclus dans un bloc de  $\pi'$ , et l'on note  $\pi \geq \pi'$ . L'ensemble ordonné  $(\mathcal{P}_\mathbb{N}, \leq)$  est réticulé, admet pour borne inférieure  $\mathbf{0}_\mathbb{N} = \{\mathbb{N}\}$  et pour borne supérieure  $\mathbf{1}_\mathbb{N} = \{\{1\}, \{2\}, \dots\}$ . L'ensemble  $\mathfrak{S}_\mathbb{N}$  des permutations des entiers agit sur  $\mathcal{P}_\mathbb{N}$  par  $(\sigma, \pi) \mapsto \sigma\pi$ , où

$$i \sim_{\sigma\pi} j \iff \sigma(i) \sim_\pi \sigma(j),$$

si l'on note  $i \sim_\pi j$  pour dire que  $i$  et  $j$  sont dans le même bloc de  $\pi$ .

À chaque triplet caractéristique  $(\beta, c, \nu)$ , on peut alors associer un processus fellerien  $(\Pi(t), t \geq 0)$  à valeurs dans  $\mathcal{P}_\mathbb{N}$ , croissant, issu de la partition  $\Pi(0) = \mathbf{0}_\mathbb{N}$ , et échangeable, c'est-à-dire que  $\sigma\Pi$  a même loi que  $\Pi$  pour toute permutation  $\sigma \in \mathfrak{S}_\mathbb{N}$ . Enfin, p.s. pour tout  $t \geq 0$ , tous les blocs de  $\Pi(t)$  admettent des fréquences asymptotiques<sup>6</sup>, et si l'on note  $|\Pi(t)|^\downarrow \in \mathcal{S}^\downarrow$  la suite décroissante de ces fréquences, alors  $(|\Pi(t)|^\downarrow, t \geq 0)$  est un processus de fragmentation à valeurs dans  $\mathcal{S}^\downarrow$  de triplet caractéristique  $(\beta, c, \nu)$ .

### 1.4.2 Exemples

Nous donnons deux exemples de processus de fragmentation auto-similaires, qui sont loin d'être les plus simples, mais qui motivent nombre de mes travaux. Les deux sont obtenus à partir de l'arbre continu brownien.

Le premier exemple est dû à Aldous & Pitman [5]. Le plus simple pour le construire est d'utiliser la construction des arbres continus inhomogènes, et de voir donc l'arbre brownien comme l'adhérence  $\overline{f(\mathbb{R}_+)}$  de l'image d'une fonction  $f : \mathbb{R}_+ \rightarrow l^1(\mathbb{N})$ , obtenue par la construction de 1.2.7 avec  $\theta = \mathbf{0}$ . On se donne alors une mesure de Poisson  $\mathcal{N}(dx ds)$  d'intensité la mesure de Lebesgue sur  $\mathbb{R}_+^2$ . Pour chaque  $t \geq 0$  on note

$$C_t = \{x \in \mathbb{R}_+ : \exists s \in [0, t], \mathcal{N}(\{(x, s)\}) = 1\}.$$

Presque-sûrement, pour tout  $t$ ,  $f(C_t)$  est un ensemble de points de  $\text{Sk}(\mathcal{T}^0)$ , et ces ensembles croissent avec  $t$ . Ces points séparent l'arbre en une famille de composantes

<sup>6</sup>rappelons que si  $A \subseteq \mathbb{N}$ , on appelle fréquence asymptotique de  $A$  la limite  $|A| = \lim_{n \rightarrow \infty} n^{-1} \#(A \cap [n])$ , si elle existe

## 1 Introduction et tour d'horizon

connexes, et l'on note  $F_+^{(2)}(t) \in \mathcal{S}^\downarrow$  la suite décroissante des  $\mu^0$ -masses de ces composantes connexes.

On montre alors que  $(F_+^{(2)}(t), t \geq 0)$  est une fragmentation auto-similaire d'indice positif  $\beta = 1/2$ , d'érosion  $c = 0$ , et de mesure de dislocation  $\nu^{(2)}$  déterminée par la formule

$$\nu^{(2)}(s_1 \in dx) = \frac{dx \mathbb{1}_{\{1/2 < x < 1\}}}{(2\pi x^3(1-x)^3)^{1/2}}, \quad (1.18)$$

et la propriété que  $\nu^{(2)}(\{\mathbf{s} : s_1 + s_2 < 1\}) = 0$  (on dit que la mesure est *binnaire*).

Malgré le caractère élaboré de la construction de  $F_+^{(2)}$ , il se trouve que les marginales de dimension 1 (dont on peut déduire le semigroupe) ont une expression explicite étonnamment simple. Soit  $\sigma^{(1/2)}$  un subordonateur (processus de Lévy croissant) d'indice  $1/2$ , de loi caractérisée par la transformée de Laplace de sa valeur au temps 1 :

$$E[\exp(-\lambda \sigma_1^{(1/2)})] = \exp(-\lambda^{1/2}).$$

Alors pour tout  $t \geq 0$ ,  $F_+^{(2)}(t)$  a même loi que la famille rangée par ordre décroissant des sauts de  $\sigma^{(1/2)}$  effectués entre 0 et  $t$ , conditionnellement à l'événement  $\{\sigma_t^{(1/2)} = 1\}$ . Bien que ce dernier conditionnement soit singulier, il est possible de lui donner un sens, voir Perman, Pitman & Yor [76], ou Pitman [78]. On verra en 4.1, suivant [P2] que cette propriété marquante est très spécifique et se généralise très mal à des subordonateurs stables d'indice différent de  $1/2$  : aucune autre fragmentation auto-similaire que  $F_+^{(2)}$  n'admet pour marginales de dimension 1 de telles lois de sauts de subordonateurs stables conditionnés

Le second exemple se décrit plus facilement à l'aide de l'excursion brownienne  $e$ . Pour chaque  $t \geq 0$ , on note  $F_-^{(2)}(t) \in \mathcal{S}^\downarrow$  la suite décroissante des mesures de Lebesgue des composantes connexes de l'ouvert de  $[0, 1]$  :

$$\{s \in [0, 1] : 2e_s > t\}.$$

On peut également voir ceci comme une fragmentation de l'arbre brownien :  $F_-^{(2)}(t)$  est la suite décroissante des  $\mu^{(2)}$ -masses des composantes connexes de  $\{x \in \mathcal{T}^{(2)} : \text{ht}(x) > t/2\}$ . Le processus  $(F_-^{(2)}(t), t \geq 0)$  est à son tour [19] une fragmentation auto-similaire. Son indice d'auto-similarité est  $-1/2$ , son érosion est nulle, et, de façon un peu surprenante au premier abord, sa mesure de dislocation est également  $\nu^{(2)}$ . On remarquera que la fragmentation  $F_-^{(2)}$  s'éteint en temps fini, comme il se doit du fait que son indice est négatif. Par contraste, la somme des composantes de  $F_+^{(2)}(t)$  est 1 p.s. pour tout  $t$ .

L'objet de la section 4.2, qui suit [P3,P7], est de généraliser la construction de ces deux fragmentations « duales » (même mesure de dislocation, mais indices d'auto-similarité opposés) de l'arbre brownien aux arbres stables décrits en 1.2.4. Par ailleurs, à la section 4.3, nous décrirons les résultats de [P5], suivant lesquels toute fragmentation auto-similaire d'indice négatif sans érosion et sans perte de masse soudaine est d'une forme similaire à  $F_-^{(2)}$ , au sens où l'on peut l'interpréter comme une fragmentation par la hauteur d'un arbre continu mesuré, appelé *arbre de fragmentation*. Enfin, décrivant les

résultats de [P14] en 4.4, nous verrons que certains de ces arbres de fragmentation sont les limites d'échelle naturelles de certaines fragmentations discrètes, codées par des *cladogrammes markoviens*.

## 1.5 Coalescence additive

De façon informelle, un processus de coalescence est un processus aléatoire qui décrit l'évolution au cours du temps d'un ensemble d'objets susceptibles de fusionner. Comme pour les fragmentations, on décrira l'état du système par l'espace d'états  $\mathcal{S}^\downarrow$  (plus généralement, on peut considérer des suites de somme finie quelconque, ce que nous ferons en section 6.1, cette modification entraîne des changements mineurs, tels des changements de temps linéaires des processus décrits ici). Si  $\mathbf{s} \in \mathcal{S}^\downarrow$ , et  $i, j \geq 1$  sont deux entiers non nuls distincts, on note

$$\mathbf{s}^{\oplus(i,j)} = (s_i + s_j, s_k : k \in \mathbb{N} \setminus \{i, j\})^\downarrow$$

l'élément de  $\mathcal{S}^\downarrow$  obtenu en fusionnant le  $i$ -ième et le  $j$ -ième terme de  $\mathbf{s}$  et en réordonnant par ordre décroissant.

**Définition 1.4** *Un processus  $(C(t), t \geq 0) = ((C_1(t), C_2(t), \dots), t \geq 0)$  à valeurs dans  $\mathcal{S}^\downarrow$  est un coalescent additif si c'est un processus de Markov homogène tel que le taux de saut d'un élément  $\mathbf{s}$  à un élément  $\mathbf{s}'$  est donné par*

$$q(\mathbf{s}, \mathbf{s}') = \begin{cases} s_i + s_j & \text{s'il existe } i \neq j \text{ tels que } \mathbf{s}' = \mathbf{s}^{\oplus(i,j)} \\ 0 & \text{sinon} \end{cases}.$$

Dans le cas où l'état initial  $C(0)$  du processus est une partition de masse « finie » (au sens où  $s_i = 0$  à partir d'un certain rang),  $(C(t), t \geq 0)$  est une chaîne de Markov à espace d'états fini. Lorsque l'état initial du processus est « infini » ( $C_i(0) > 0$  pour tout  $i$ ), la définition ci-dessus devient problématique, mais peut être rendue rigoureuse, comme l'ont montré Evans et Pitman [40].

L'un des aspects frappants dans l'étude du coalescent additif est qu'il dispose d'une frontière d'entrée très riche. Plus précisément, il existe une infinité de processus *éternels*, c'est-à-dire définis sur l'intervalle de temps  $\mathbb{R}$ . Tous ces processus sont issus au temps  $-\infty$  d'une « poussière » formée de particules de masses infinitésimales. L'exemple le plus notable est celui du coalescent *standard* [40, 5], obtenu comme limite en loi du processus issu au temps  $-\log n/2$  de l'état  $(1/n, \dots, 1/n, 0, \dots)$  ( $n$  objets de masse  $1/n$ ).

En partant d'une description par Pitman du coalescent additif en termes de processus de forêts coalescentes, étroitement lié aux  $\mathbf{p}$ -arbres [77], Aldous & Pitman [6] ont donné une construction de ces versions éternelles, en termes d'un processus de fragmentation associé aux arbres continus inhomogènes décrits en 1.2.7. En particulier, le coalescent éternel standard n'est autre que le processus  $(F_+^{(2)}(e^{-t}), -\infty < t < \infty)$ , où  $F_+^{(2)}$  est la fragmentation décrite à la section 1.4.2. Les autres versions du coalescent sont essentiellement obtenus en copiant la construction de  $F_+^{(2)}$ , mais en remplaçant l'arbre brownien par un arbre continu inhomogène.

## 1 Introduction et tour d'horizon

Nous donnons néanmoins une construction différente, relativement plus simple, donnée par Bertoin [16]. Soit  $\theta \in \Theta$  défini en (1.12), et soit  $\mathbb{b}$  un pont brownien standard. On se donne une suite  $(U_i, i \geq 1)$  de variables aléatoires uniformes sur  $[0, 1]$ , indépendantes, et indépendantes de  $\mathbb{b}$ . On définit un *pont à accroissements échangeables*

$$\mathbb{b}_s^\theta = \theta_0 \mathbb{b}_s + \sum_{i \geq 1} \theta_i (\mathbb{1}_{\{U_i \leq s\}} - s), \quad 0 \leq s \leq 1. \quad (1.19)$$

Un théorème de Kallenberg [55] montre que  $\mathbb{b}^\theta$  est bien défini et est càdlàg, c'est un *pont* au sens où  $\mathbb{b}_0^\theta = \mathbb{b}_1^\theta = 0$  et  $\mathbb{b}^\theta$  est continu en 0 et 1. Plus précisément, Kallenberg montre que tout pont à accroissements échangeables càdlàg, à sauts positifs et de variation infinie p.s. est de la forme  $\lambda_0 \mathbb{b}^\theta$ , où  $\lambda_0 > 0, \theta \in \Theta$  sont des variables aléatoires, indépendantes des  $U_i, i \geq 1$  (mais pas nécessairement indépendantes entre elles).

Un résultat de Knight [57] montre qu'il existe p.s. un unique temps  $s_* \in [0, 1]$  où  $\mathbb{b}^\theta$  atteint son minimum, de plus Bertoin [16] a montré que  $\mathbb{b}^\theta$  y est continu. Nous notons

$$\mathfrak{e}_s^\theta = \mathbb{b}_{s+s_*}^\theta - \mathbb{b}_{s_*}^\theta, \quad (1.20)$$

où l'addition  $s + s_*$  est considéré modulo 1. Une telle translation cyclique au minimum est appelée *transformée de Vervaat*.

Pour  $t \geq 0$ , nous notons

$$\mathfrak{e}_s^{\theta, (t)} = \mathfrak{e}_s^\theta - st, \quad 0 \leq s \leq 1,$$

et  $\underline{\mathfrak{e}}_s^{\theta, (t)} = \inf_{0 \leq u \leq s} \mathfrak{e}_u^{\theta, (t)}$ . On note enfin  $F^\theta(t)$  la variable aléatoire de  $\mathcal{S}^\downarrow$  définie comme la suite décroissante des intervalles de constance de  $\underline{\mathfrak{e}}^{\theta, (t)}$ , c'est-à-dire des composantes connexes du complémentaire du support de  $d\underline{\mathfrak{e}}^{\theta, (t)}$ .

Lorsque  $\theta = \mathbf{0}$ , la fragmentation  $F^\mathbf{0}$  obtenue a la même loi que  $F_+^{(2)}$  [16]. Cette coïncidence très surprenante est restée relativement mystérieuse jusqu'aux récents travaux de Bertoin [21]. Notons que dans ce cas, l'excursion  $\mathfrak{e}^\theta$  a même loi que l'excursion brownienne  $\mathfrak{e}$ , par un théorème célèbre de Vervaat [84], ce qui est bien sûr une piste pour retrouver l'arbre continu brownien. On notera néanmoins que la construction de  $F_+^{(2)}$  nécessite l'introduction d'un aléa supplémentaire *via* un processus de Poisson, tandis que la construction de  $F^\mathbf{0}$  ne nécessite que la trajectoire de  $\mathfrak{e}$ .

Les résultats conjoints d'Aldous-Pitman et de Bertoin [16, 6] entraînent que le processus

$$(\mathbf{C}^\theta(t) = F^\theta(e^{-t}), -\infty < t < \infty)$$

est un coalescent additif éternel. De plus, tout coalescent éternel s'obtient comme mélange en  $\theta, t_0$  des lois des processus  $(\mathbf{C}^\theta(t-t_0), -\infty < t < \infty)$ . Autrement dit, du fait de la discussion juste après (1.19), nous pouvons refaire la construction ci-dessus en remplaçant  $\mathbb{b}^\theta$  par n'importe quel pont à accroissements échangeables de longueur 1, à sauts positifs et à variation totale infinie p.s.

En section 6.1, suivant [P1], nous considérerons certains processus éternels associés naturellement à des excursions de processus de Lévy spectralement positifs, pour lesquels

## *1 Introduction et tour d'horizon*

des calculs de semigroupe deviennent possibles. Par ailleurs, en 6.2, suivant [P9], nous verrons comment différents coalescents éternels apparaissent dans une version continue du « problème de parking » de Knuth, problème pour lequel Chassaing & Louchard [28] avaient mis en évidence des liens avec le coalescent additif standard.

## 2 Principes d'invariance pour les arbres de Galton-Watson spatiaux multitypes

Nous donnons ici une définition formelle de ces arbres, ainsi que les principes d'invariance nécessaires à la déduction des résultats du chapitre 3.1. Les résultats ci-dessous proviennent de [P11,P15].

### 2.1 Processus de Galton-Watson multitypes

Commençons par décrire brièvement les processus de Galton-Watson à plusieurs types. On se donne  $K \geq 1$  et  $\boldsymbol{\mu} = (\mu_i, 1 \leq i \leq K)$  une suite de mesures de probabilités sur  $\mathbb{Z}_+^K$ . Un processus de Galton-Watson à  $K$  types est un processus de Markov  $(\mathbf{Z}_n, n \geq 0)$  à valeurs dans  $\mathbb{Z}_+^K$ , dont les probabilités de transition sont spécifiées par la propriété suivante : sachant que  $\mathbf{Z}_n = \mathbf{z} = (z_1, \dots, z_K)$ ,  $\mathbf{Z}_{n+1}$  a même loi que

$$\sum_{i=1}^K \sum_{j=1}^{z_i} \mathbf{X}_j^{(i)},$$

où les variables  $(\mathbf{X}_j^{(i)}, i, j \geq 1)$  sont indépendantes, avec  $\mathbf{X}_j^{(i)}$  de loi  $\mu_i$  pour tout  $i, j$ . Ce processus correspond à un modèle de population où chaque individu est doté d'un type, et donne naissance indépendamment des autres à une famille d'individus selon une loi de reproduction propre à son type.

Ces processus ont été introduits par Kolmogorov, et sont étudiés dans les références classiques sur les processus de branchement [51, 10]. En particulier, il est connu que le comportement des processus à plusieurs types est similaire à celui des processus de branchement habituels, sous une hypothèse d'irréductibilité qui se formule ainsi. On note

$$m_{ij} = \sum_{\mathbf{z} \in \mathbb{Z}_+^K} z_j \mu_i(\{\mathbf{z}\})$$

le nombre moyen d'individus de type  $j$  nés d'un individu de type  $i$ . On dit que le processus de branchement associé à  $\boldsymbol{\mu}$  est *irréductible* si la matrice  $M_{\boldsymbol{\mu}} = (m_{ij})$  est irréductible, c'est-à-dire que pour tout  $i, j$  on peut trouver  $n$  tel que  $m_{ij}^{(n)} > 0$ , où ce dernier nombre est le coefficient en  $i, j$  de  $M_{\boldsymbol{\mu}}^n$ . On dit également que le processus est *non-dégénéré* s'il existe  $i$  tel que  $\mu_i(|\mathbf{z}| = 1) < 1$ , où  $|\mathbf{z}| = z_1 + \dots + z_K$ . Ainsi, il existe  $n$  tel que  $|\mathbf{Z}_n| \neq 1$  avec probabilité positive, et le comportement de  $\mathbf{Z}$  ne se réduit pas à celui d'une chaîne de Markov à valeurs dans  $[K]$ .

## 2 Principes d'invariance pour les arbres de Galton-Watson spatiaux multitypes

Sous les hypothèses que le processus est irréductible et non-dégénéré, son comportement en temps grand est déterminé par la valeur propre de Perron de la matrice  $M_{\boldsymbol{\mu}}$ . Plus exactement, si  $\varrho_{\boldsymbol{\mu}}$  est la plus grande valeur propre de  $M_{\boldsymbol{\mu}}$ , alors le processus de branchement s'éteint presque-sûrement (atteint l'état  $\mathbf{0} \in \mathbb{Z}_+^K$ ) si et seulement si  $\varrho_{\boldsymbol{\mu}} \leq 1$ . On dit que le processus est sous-critique, et qu'il est critique si  $\varrho_{\boldsymbol{\mu}} = 1$ . A partir de maintenant,  $\boldsymbol{\mu}$  désigne une loi de reproduction à  $K$  types, de processus associé non-dégénéré et critique. Plus précisément, soit

$$\phi_i(\mathbf{s}) = \sum_{\mathbf{z} \in \mathbb{Z}_+^K} \mu_i(\{\mathbf{z}\}) \mathbf{s}^{\mathbf{z}}, \quad \mathbf{s} \in [0, 1]^K, \quad i \in [K],$$

où par définition  $\mathbf{s}^{\mathbf{z}} = \prod_{i \in [K]} s_i^{z_i}$ . Nous posons

$$Q_{jk}^{(i)} = \frac{\partial^2 \phi_i}{\partial x_j \partial x_k}(1, \dots, 1), \quad i, j, k \in [K],$$

et disons que  $\boldsymbol{\mu}$  a une variance finie si toutes ces quantités sont finies. Les matrices  $(Q_{jk}^{(i)}, j, k \in [K])$  sont les matrices de formes quadratiques positives sur  $\mathbb{R}^K$ , notées  $Q^{(i)}(\mathbf{s}), \mathbf{s} \in \mathbb{R}^K$ . On pose aussi  $\mathbf{Q}(\mathbf{s}) = (Q^{(i)}(\mathbf{s}), i \in [k]) \in \mathbb{R}^K$ .

Soient alors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^k$  les vecteurs propres de  $M_{\boldsymbol{\mu}}$ , respectivement à gauche et à droite, de valeur propre associée  $\varrho_{\boldsymbol{\mu}} = 1$  et choisis de telle sorte que

$$a_1 + \dots + a_K = a_1 b_1 + \dots + a_K b_K = 1.$$

On pose  $\sigma = \sqrt{\mathbf{a} \cdot \mathbf{Q}(\mathbf{b})}$ , où  $\cdot$  est le produit scalaire usuel dans  $\mathbb{R}^K$ . Alors  $\sigma$  tient le rôle d'une « variance moyenne » dans le processus de branchement, comme le montre le résultat suivant :

**Proposition 2.1** *Supposons  $\boldsymbol{\mu}$  critique, non-dégénéré, de variance finie. Soit  $(\mathbf{Z}_n, n \geq 0)$  un processus de Galton-Watson multitype de loi de reproduction  $\boldsymbol{\mu}$ , défini dans l'espace probabilisé  $(\Omega, \mathcal{F}, P)$ . Alors la probabilité de non-extinction au temps  $n$  vérifie*

$$P(\mathbf{Z}_n \neq \mathbf{0} | \mathbf{Z}_0 = (\delta_{ik}, k \in [K])) \underset{n \rightarrow \infty}{\sim} \frac{2b_i}{n\sigma^2} \quad i \in [K].$$

*En particulier, si l'état initial est distribué selon le vecteur de probabilité  $\mathbf{a}$ , alors la probabilité de non-extinction est équivalente à  $2/(n\sigma^2)$ .*

Ce résultat, dû à Vatutin & Dyakonova [83] lorsque  $M_{\boldsymbol{\mu}}$  est également *apériodique*, est une généralisation de l'énoncé bien connu que l'on obtient lorsque  $K = 1$ , dû à Kolmogorov, sur la probabilité d'extinction d'un processus de Galton-Watson critique. Nous avons proposé dans [P15] une preuve entièrement différente de cette proposition, qui repose sur le codage des arbres de Galton-Watson associés aux processus de branchement.

## 2.2 Arbres de Galton-Watson multitypes [P11,P15]

Nous ne décrivons dans cette section que les versions conditionnées de nos résultats, qui sont nécessaires à la preuve des théorèmes sur les cartes exposés au chapitre 3.1. Pour ce faire, nous allons imposer l'hypothèse supplémentaire que  $\boldsymbol{\mu}$  admet des (petits) moments exponentiels, c'est-à-dire qu'il existe  $\varepsilon > 0$  tel que

$$\sup_{i \in [K]} \sum_{\mathbf{z} \in \mathbb{Z}_+^K} \exp(\varepsilon(z_1 + \dots + z_K)) \mu_i(\{\mathbf{z}\}) < \infty.$$

On veut donc coder la généalogie des processus de Galton-Watson multitypes à l'aide d'arbres aléatoires plans. Pour ce faire, on se donne une suite de mesures  $\boldsymbol{\zeta} = (\zeta_i, 1 \leq i \leq K)$  sur l'ensemble  $\mathcal{W}_K$  des suites finies (éventuellement vides) à valeurs dans  $[K]$ . On interprète  $\zeta_i$  comme la loi des types d'une suite ordonnée d'individus. Soit  $p : \mathcal{W}_K \rightarrow \mathbb{Z}_+^K$  la projection définie par

$$p(\mathbf{e})_i = \#\{j \geq 1 : w_j = i\}.$$

On suppose que  $p_* \zeta_i = \mu_i$  pour tout  $i$ , ce que l'on note  $p_* \boldsymbol{\zeta} = \boldsymbol{\mu}$  pour plus de commodité.

Un arbre à  $K$  types est un couple  $(\mathbf{t}, e)$  où  $\mathbf{t} \in \mathbf{T}$  et  $e : \mathbf{t} \rightarrow [K]$  est une fonction qui à un sommet associe son type. On note  $\mathbf{T}^{(K)}$  l'ensemble des arbres à  $K$  types. En général, on omettra de mentionner  $e$ , qui sera implicite, pour des notations plus aisées. Si  $\mathbf{t} \in \mathbf{T}^{(K)}$ , on note  $\mathbf{t}^{(i)} = \{u \in \mathbf{t} : e(u) = i\}$ . Si  $u \in \mathbf{t}$ , on note  $\mathbf{e}_u(\mathbf{t}) = (e(u1), e(u2), \dots, e(uc_u(\mathbf{t}))) \in \mathcal{W}_K$  la suite des types des enfants de  $u$ .

Soit  $i \in [K]$ . Les formules

$$P_{\boldsymbol{\zeta}}^{(i)}(\{\mathbf{t}\}) = \prod_{u \in \mathbf{t}} \zeta_{e(u)}(\mathbf{e}_u(\mathbf{t})), \quad \mathbf{t} \in \mathbf{T}^{(K)}, e(\emptyset) = i$$

définissent alors une mesure sur  $\mathbf{T}^{(K)}$ , appelée loi d'un arbre de Galton-Watson à  $K$  types, de loi de reproduction  $\boldsymbol{\zeta}$ , et issu d'un individu de type  $i$ . Si l'on note  $\mathbf{Z}_n(\mathbf{t}) = (\#\{u \in \mathbf{t} : |u| = n, e(u) = i\}, i \in [K])$ , on voit aisément que la loi de  $(\mathbf{Z}_n(\mathbf{t}), n \geq 0)$  sous  $P_{\boldsymbol{\zeta}}^{(i)}$  est celle d'un processus de Galton-Watson de loi de reproduction  $\boldsymbol{\mu}$ , issu d'un unique individu de type  $i$ .

Nous couplons à présent l'arbre de Galton-Watson avec un déplacement spatial, que l'on peut interpréter, sur un plan heuristique, comme une marche de branchement. Néanmoins, du fait de besoins dictés par les problèmes sur les cartes, nous considérons une situation un peu plus élaborée où le déplacement spatial de la marche peut dépendre localement de la structure de l'arbre, d'où l'aspect relativement technique des notations à suivre. Pour chaque  $i \in [K]$ ,  $\mathbf{e} \in \mathcal{W}_K$ , nous considérons une loi de probabilités  $\nu_{i,\mathbf{e}}$  sur  $\mathbb{R}^{|\mathbf{e}|}$ , où  $|\mathbf{e}| \geq 0$  est la longueur de  $\mathbf{e}$ . On note  $\boldsymbol{\nu} = (\nu_{i,\mathbf{e}}, i \in [K], \mathbf{e} \in \mathcal{W}_K)$ .

Si  $\mathbf{t} \in \mathbf{T}^{(K)}$ , on considère, pour chaque  $u \in \mathbf{t}$ , un vecteur aléatoire  $(Y_{u1}, \dots, Y_{uc_u(\mathbf{t})})$  de loi  $\nu_{e(u), \mathbf{e}_u(\mathbf{t})}$ , ce indépendamment sur les différents  $u \in \mathbf{t}$  (en revanche rien n'empêche les  $Y_{ui}, 1 \leq i \leq c_u(\mathbf{t})$  d'être dépendants, contrairement à la situation de la section 1.1). Par convention, on pose  $Y_{\emptyset} = 0$ , et définissons, comme en section 1.1.3,

$$L_u = \sum_{v \preceq u} Y_v, \quad u \in \mathbf{t}.$$



## 2 Principes d'invariance pour les arbres de Galton-Watson spatiaux multitypes

On note  $\nu_{\mathbf{t}}$  la loi de  $(L_u, u \in \mathbf{t})$ , et

$$d\mathbb{P}_{\zeta, \nu}^{(i)}(\mathbf{t}, \mathbf{y}) = dP_{\zeta}^{(i)}(\mathbf{t})d\nu_{\mathbf{t}}(\mathbf{l}),$$

définissant ainsi une loi de probabilités sur l'ensemble

$$\overline{\mathbf{T}}^{(K)} = \{(\mathbf{t}, \mathbf{l}) : \mathbf{t} \in \mathbf{T}^{(K)}, \mathbf{l} \in \mathbb{R}^{\mathbf{t}}\}$$

des *arbres spatiaux à  $K$  types*. Dans cette définition, on voit que seules les mesures  $\nu_{i, \mathbf{e}}$  telles que  $\zeta_i(\{\mathbf{e}\}) > 0$  interviennent, on adopte donc la convention que  $\nu_{i, \mathbf{e}} = \delta_{\mathbf{0}}, \mathbf{0} \in \mathbb{R}^{|\mathbf{e}|}$  pour les indices  $(i, \mathbf{e})$  tels que  $\zeta_i(\{\mathbf{e}\}) = 0$ .

Si  $(\mathbf{t}, \mathbf{l}) \in \overline{\mathbf{T}}^{(K)}$ , on définit

$$S_n^{\mathbf{t}, \mathbf{l}} = \mathbf{l}(u_{\mathbf{t}}(k)), \quad 0 \leq k \leq \#\mathbf{t} - 1,$$

et par convention, on note  $S_{\#\mathbf{t}}^{\mathbf{t}, \mathbf{l}} = S_{\#\mathbf{t}-1}^{\mathbf{t}, \mathbf{l}}$ . Enfin, on note

$$|\mathbf{y}|_2 = \sqrt{y_1^2 + \dots + y_k^2}, \quad \mathbf{y} \in \mathbb{R}^k,$$

et nous rappelons la définition de la tête du serpent brownien  $(\mathbf{e}, Z)$ . Nous avons alors :

**Théorème 2.1** *Soit  $\zeta$  une loi de reproduction telle que  $\boldsymbol{\mu} = p_*\zeta$  est irréductible, critique, non-dégénérée, et admet des moments exponentiels. On rappelle les notations  $\mathbf{a}, \mathbf{b}, \sigma$  ci-dessus.*

*Supposons également que toutes les lois de la famille  $\nu$  soient centrées, et qu'il existe  $\xi, D > 0$  tels que*

$$M_{i, \mathbf{e}} = \langle \nu_{i, \mathbf{e}}, |\mathbf{y}|_2^{8+\xi} \rangle < \infty \quad i \in [K], \mathbf{e} \in \mathcal{W}_K,$$

et

$$\sup_{i \in [K]} M_{i, \mathbf{e}} = O(|\mathbf{e}|^D).$$

Nous notons

$$\varsigma = \sqrt{\sum_{i \in [K]} a_i \sum_{\mathbf{e} \in \mathcal{W}_K} \zeta_i(\{\mathbf{e}\}) \sum_{j=1}^{|\mathbf{e}|} b_{e_j} \langle \nu_{i, \mathbf{e}}, y_j^2 \rangle}.$$

Soit  $i, j \in [K]$ , et pour chaque  $n$ , soit  $(T_n, Y_n)$  une variables aléatoires dans  $\overline{\mathbf{T}}^{(K)}$  de loi  $\mathbb{P}_{\zeta, \nu}^{(i)}(\cdot | \{\#\mathbf{t}^{(j)} = n\})$ , où  $n$  n'est considéré que le long des valeurs où  $P_{\zeta}^{(i)}(\{\mathbf{t}^{(j)} = n\}) > 0$ . Alors on a la convergence en loi dans  $\mathbb{D}([0, 1])^2$  :

$$\left( \left( \frac{H_{\lfloor \#T_n t \rfloor}^{T_n}}{\sqrt{\#T_n}} \right)_{0 \leq t \leq 1}, \left( \frac{S_{\lfloor \#T_n t \rfloor}^{T_n, Y_n}}{\#T_n^{1/4}} \right)_{0 \leq t \leq 1} \right) \xrightarrow[n \rightarrow \infty]{(d)} \left( \frac{2}{\sigma} \mathbf{e}, \varsigma \sqrt{\frac{2}{\sigma}} Z \right).$$

De plus, on a  $\#T_n/n \rightarrow a_j^{-1}$  en probabilité.

## 2 Principes d'invariance pour les arbres de Galton-Watson spatiaux multitypes

Ce résultat généralise donc ceux de Marckert et Janson (1.3) aux cas multitypes, et aux situations où les accroissements des déplacements spatiaux peuvent dépendre de l'environnement immédiat des sommets. Les hypothèses faites sont également légèrement plus contraignantes, et sans doute non-optimales. Quelques améliorations possibles sont proposées dans [P11,P15]. On notera qu'un des outils importants dans les preuves est l'utilisation de la mesure de l'arbre biaisé par la taille, telle que considérée dans Kurtz, Lyons, Pemantle & Peres [59].

## 3 Cartes aléatoires

### 3.1 Principes d'invariance pour les cartes aléatoires [P10,P11,P12]

Une question importante est de savoir si les propriétés des limites d'échelle des cartes sont *universelles*, c'est-à-dire si le choix de la famille de cartes que l'on fait est déterminant ou pas. Par exemple, les résultats de Chassaing & Schaeffer (1.15,1.16) portent sur les quadrangulations. Peut-on remplacer ces dernières par les triangulations, peut-être plus familières, et obtenir les mêmes résultats, à des facteurs d'échelle près ?

La résolution de ce problème est rendue possible par des travaux de Bouttier, Di Francesco & Guitter [23], généralisant la bijection de Schaeffer à des cartes quelconques. La *bijection BDFG* permet de coder ces dernières par des arbres spatiaux possédant diverses décorations. Ces derniers peuvent s'interpréter comme des arbres à plusieurs types, au sens du chapitre 2. Nous ne détaillerons pas la construction de cette bijection, donnant juste un exemple avec la figure 3.1

Cette bijection permet le calcul de fonctions génératrices des cartes, où les faces de degré  $k \geq 1$  sont pondérées par un poids  $X_k$  (pour l'instant une variable formelle) :

$$Z_{\mathbf{X}} = \sum_{\mathbf{m} \in \mathcal{M}} \prod_{f \in F(\mathbf{m})} X_{\deg f},$$

où par convention, la carte  $\bullet$  donne lieu à un poids  $X_0 = 1$ . La contrepartie probabiliste de ces séries génératrices sont les *mesures de Boltzmann* sur  $\mathcal{M}$ . On se donne une suite  $\mathbf{w} = (w_1, w_2, \dots)$  positive, et on définit une mesure  $\sigma$ -finie sur  $\mathcal{M}$  par

$$\mathbb{B}_{\mathbf{w}}(\{\mathbf{m}\}) = \prod_{f \in F(\mathbf{m})} w_{\deg f}.$$

Si la masse totale de cette mesure  $Z_{\mathbf{w}} := \mathbb{B}_{\mathbf{w}}(\mathcal{M})$  est finie, on peut définir la loi de Boltzmann

$$\mathbb{P}_{\mathbf{w}} = \frac{\mathbb{B}_{\mathbf{w}}}{Z_{\mathbf{w}}},$$

et l'on dit que la suite  $\mathbf{w}$  est  $\mathcal{M}$ -admissible.

Pour donner un exemple concret, supposons qu'il existe  $\kappa > 2$  tel que  $w_{\kappa} = w > 0$  et  $w_i = 0$  pour tout  $i \neq \kappa$ . Dans ce cas,  $Z_{\mathbf{w}} = \sum_{\mathbf{m} \in \mathcal{M}} w^{\#F(\mathbf{m})}$  est la série génératrice des  $\kappa$ -angulations, c'est-à-dire des cartes dont toutes les faces ont pour degré  $\kappa$ , comptées avec un poids  $w$  par face. La série est admissible si  $w$  est strictement inférieur au rayon de convergence de cette série, auquel cas la mesure conditionnée

$$\mathbb{B}_{\mathbf{w}}(\cdot | \mathcal{M}_n^f),$$

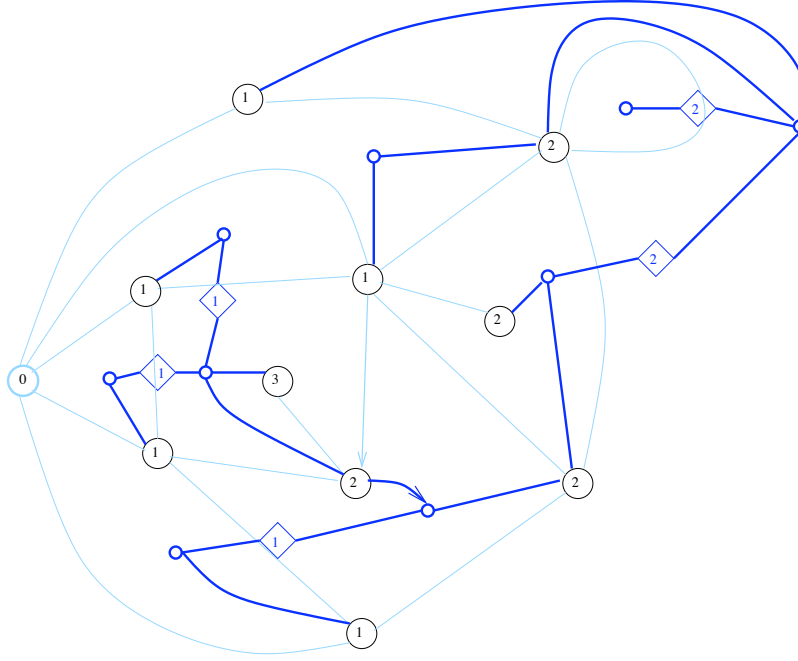


FIG. 3.1: Une carte et l'arbre spatial associé par la bijection BDFG

où  $\mathcal{M}_n^f \subset \mathcal{M}$  est l'ensemble des cartes ayant  $n$  faces, est la mesure uniforme sur l'ensemble  $\mathcal{M}_n^f(\kappa)$  des  $\kappa$ -angulations à  $n$  faces.

Il est également bon de noter que si  $\mathbf{w}$  est une suite donnée, la suite  $a\mathbf{w} = (aw_i, i \geq 1)$  est telle que

$$\mathbb{B}_{a\mathbf{w}}(\{\mathbf{m}\}) = a^{\#F(\mathbf{m})} \mathbb{B}_{\mathbf{w}}(\{\mathbf{m}\}), \quad \mathbf{m} \in \mathcal{M},$$

d'où l'on tire que

$$\mathbb{B}_{a\mathbf{w}}(\cdot | \mathcal{M}_n^f) = \mathbb{B}_{\mathbf{w}}(\cdot | \mathcal{M}_n^f).$$

Si  $a\mathbf{w}$  est admissible, cette dernière mesure est aussi  $\mathbb{P}_{a\mathbf{w}}(\cdot | \mathcal{M}_n^f)$ .

En réalité, il est plus aisé de calculer, *via* la bijection BDFG, des séries génératrices de cartes enracinées pointées, c'est-à-dire de paires  $(\mathbf{m}, u)$ , où  $\mathbf{m} \in \mathcal{M}$  et  $u \in V(\mathbf{m})$  est un sommet distingué. La raison principale est que, lorsque l'on mesure les distances dans la carte à partir de ce sommet distingué et non à partir de l'origine de la racine, comme pour la bijection détaillée en 1.3.1, il est possible de se débarrasser de la contrainte que les marques spatiales des arbres considérés sont positives.

On note  $\mathcal{M}^*$  l'ensemble des cartes pointées et enracinées, et comme auparavant on suppose que  $\bullet$  en est un élément (pointé en son unique sommet). On définit alors

$$\begin{aligned} \mathcal{M}^+ &= \{(\mathbf{m}, u) \in \mathcal{M}^* : d_{\text{gr}}(e_*^-, u) + 1 = d_{\text{gr}}(e_*^+, u)\} \cup \{\bullet\}, \\ \mathcal{M}^- &= \{(\mathbf{m}, u) \in \mathcal{M}^* : d_{\text{gr}}(e_*^-, u) = d_{\text{gr}}(e_*^+, u) + 1\}, \quad \text{et} \\ \mathcal{M}^0 &= \{(\mathbf{m}, u) \in \mathcal{M}^* : d_{\text{gr}}(e_*^-, u) = d_{\text{gr}}(e_*^+, u)\}, \end{aligned}$$

où  $e_*^-, e_*^+$  sont l'origine et le but de  $e_*$ , obtenant ainsi une tripartition de  $\mathcal{M}^*$ . On notera que tous les éléments de  $\mathcal{M}^-$  s'obtiennent en renversant l'orientation de la racine d'un élément de  $\mathcal{M}^+$ .

On définit par les mêmes formules que ci-dessus, en remplaçant  $\mathcal{M}$  respectivement par  $\mathcal{M}^*, \mathcal{M}^+, \mathcal{M}^0$ , les mesures  $\mathbb{B}_{\mathbf{w}}^*, \mathbb{B}_{\mathbf{w}}^+, \mathbb{B}_{\mathbf{w}}^0$  de masses totales  $Z_{\mathbf{w}}^*, Z_{\mathbf{w}}^+, Z_{\mathbf{w}}^0$ , et les mesures de probabilités  $\mathbb{P}_{\mathbf{w}}^*, \mathbb{P}_{\mathbf{w}}^+, \mathbb{P}_{\mathbf{w}}^0$  si ces masses sont finies. On dit alors que  $\mathbf{w}$  est  $\mathcal{M}^*, \mathcal{M}^+, \mathcal{M}^0$ -admissible.

### 3.1.1 Cartes biparties [P11]

Dans ce travail avec Jean-François Marckert, nous nous sommes intéressés au cas où  $w_i = 0$  pour tout  $i$  impair. Dans ce cas, la mesure  $\mathbb{B}_{\mathbf{w}}$  ne charge que les cartes dont toutes les faces ont pour degré pair. En genre 0, cela équivaut au fait que les cartes sont biparties. En particulier, les mesures  $\mathbb{B}_{\mathbf{w}}$  ne chargent pas  $\mathcal{M}^0$ . Ainsi, un poids  $\mathbf{w}$  est  $\mathcal{M}^*$ -admissible si et seulement s'il est  $\mathcal{M}^+$ -admissible.

Dans ce cas, la bijection BDFG est grandement simplifiée, il était donc naturel de commencer par l'étude des cartes biparties pour généraliser le résultat de Chassaing et Schaeffer.

Nous notons

$$f_{\mathbf{w}}(x) = \sum_{k \geq 0} \binom{2k+1}{k} w_{2k+2} x^k,$$

définissant une série entière  $[0, \infty) \rightarrow [0, \infty]$ , qui est une fonction convexe sur l'intervalle  $[0, r_{\mathbf{w}}]$ , où  $r_{\mathbf{w}} \geq 0$  est le rayon de convergence. Nous obtenons le critère suivant pour l'admissibilité de  $\mathbf{w}$  :

**Proposition 3.1** *La suite  $\mathbf{w}$  est  $\mathcal{M}^*$ -admissible si et seulement si l'équation*

$$f_{\mathbf{w}}(x) = 1 - 1/x$$

*admet une solution, auquel cas  $Z_{\mathbf{w}}^+$  est la solution  $x$  satisfaisant  $x^2 f'_{\mathbf{w}}(x) \leq 1$ .*

Comme  $x \mapsto 1 - 1/x$  est concave, il est aisé que l'une des solutions (la plus petite) vérifie cette inégalité — s'il se trouve que  $x = r_{\mathbf{w}}$  est solution, alors  $f'_{\mathbf{w}}(r_{\mathbf{w}})$  doit être compris comme la dérivée à gauche de  $f_{\mathbf{w}}$  en  $r_{\mathbf{w}}$ , qui égale  $\lim_{r \uparrow r_{\mathbf{w}}} f'_{\mathbf{w}}(r)$  par convergence monotone.

**Définition 3.1** *On dit que la suite  $\mathbf{w}$  est critique si elle est  $\mathcal{M}^*$ -admissible et si*

$$(Z_{\mathbf{w}}^+)^2 f'_{\mathbf{w}}(Z_{\mathbf{w}}^+) = 1,$$

*auquel cas, en particulier, l'équation  $f_{\mathbf{w}}(x) = 1 - 1/x$  admet une unique solution. Enfin, on dit que  $\mathbf{w}$  est critique régulière si  $Z_{\mathbf{w}}^+ < r_{\mathbf{w}}$ .*

On rappelle les définitions de  $(e, Z)$  (section 1.1) et  $\mathcal{I}_{\infty}$  (1.17). haut. Nous sommes maintenant prêts à énoncer le résultat principal de [P11].

**Théorème 3.1** *Soit  $\mathbf{w}$  une suite critique régulière. Soit*

$$A_{\mathbf{w}} := \left( \frac{8 + 4(Z_{\mathbf{w}}^+)^3 f_{\mathbf{w}}''(Z_{\mathbf{w}}^+)}{9(Z_{\mathbf{w}}^+ - 1)} \right)^{1/4}.$$

*Soit  $(\mathbf{m}_n, u_n)$  une variable aléatoire de loi  $\mathbb{P}_{\mathbf{w}}^*(\cdot | \mathcal{M}_n^f)$ , et soit  $v_n$  un sommet de  $\mathbf{m}_n$  choisi uniformément conditionnellement à  $\mathbf{m}_n$ . Alors on a les convergences en loi :*

$$\begin{aligned} \frac{R(\mathbf{m}_n, u_n)}{A_{\mathbf{w}} n^{1/4}} &\xrightarrow[n \rightarrow \infty]{(d)} \sup Z - \inf Z, \\ \frac{d_{\text{gr}}(u_n, v_n)}{A_{\mathbf{w}} n^{1/4}} &\xrightarrow[n \rightarrow \infty]{(d)} \sup Z, \\ \frac{\mathcal{I}_{\mathbf{m}_n, u_n}(A_{\mathbf{w}} n^{1/4} dx)}{\#V(\mathbf{m}_n)} &\xrightarrow[n \rightarrow \infty]{(d)} \mathcal{I}_{\infty}, \end{aligned}$$

la dernière ayant lieu pour la topologie faible sur les mesures de probabilités sur  $\mathbb{R}_+$ .

**Remarques.**

1. L'énoncé de [P11] est donné en termes des probabilités  $\mathbb{P}_{\mathbf{w}}^+(\cdot | \mathcal{M}_n^f)$ , mais on peut sans difficulté supplémentaire les remplacer par  $\mathbb{P}_{\mathbf{w}}^*(\cdot | \mathcal{M}_n^f)$ , comme ci-dessus.
2. Pour le cas des  $2\kappa$ -angulations avec  $\kappa \geq 2$ , les résultats précédents montrent que  $w_{\kappa} = (\kappa - 1)^{\kappa-1} / (\kappa^{\kappa} \binom{2\kappa-1}{\kappa-1})$  est tel que la suite  $w_{\kappa} \delta_{2\kappa}$  est critique régulière, et donnent la valeur explicite  $A_{\kappa} = (4\kappa(\kappa - 1)/9)^{1/4}$  pour la constante d'échelle, en accord avec le cas  $\kappa = 2$  des quadrangulations.
3. Nous montrons également un résultat où l'on conditionne non pas par le nombre de faces, mais par le nombre de sommets. Dans ce cas, il faut être attentif au fait qu'il se peut que  $\mathbb{B}_{\mathbf{w}}(\mathcal{M}_n^s)$  soit nul, où l'on a noté  $\mathcal{M}_n^s = \{\mathbf{m} \in \mathcal{M} : \#V(\mathbf{m}) = n\}$ . Le théorème reste alors vrai à condition que  $n$  varie le long d'une sous-suite où cette dernière quantité est non nulle. De plus, il convient alors de substituer à la constante d'échelle  $A_{\mathbf{w}}$  la quantité

$$C_{\mathbf{w}} := \left( \frac{8 + 4(Z_{\mathbf{w}}^+)^3 f_{\mathbf{w}}''(Z_{\mathbf{w}}^+)}{9} \right)^{1/4}.$$

4. Un simple calcul montre qu'une carte pointée de loi  $\mathbb{P}_{\mathbf{w}}^+(\cdot | \mathcal{M}_n^s)$  s'obtient en prenant une carte de loi  $\mathbb{P}_{\mathbf{w}}(\cdot | \mathcal{M}_n^s)$ , puis en distinguant l'un de ses  $n$  sommets uniformément au hasard. Ceci n'est pas vrai en général pour des cartes de loi  $\mathbb{P}_{\mathbf{w}}^+$  ou  $\mathbb{P}_{\mathbf{w}}^+(\cdot | \mathcal{M}_n^f)$ .
5. Du fait des remarques faites à la fin de la section précédente, le théorème reste vrai si l'on se donne  $\mathbf{w}$  telle qu'il existe  $a_c > 0$  avec  $a_c \mathbf{w}$  critique régulière, en prenant alors  $(\mathbf{m}_n, u_n)$  de loi  $\mathbb{B}_{\mathbf{w}}^*(\cdot | \mathcal{M}_n^f)$ . En revanche, la constante d'échelle est  $A_{a_c \mathbf{w}}$ .

La preuve repose sur le fait que l'on peut décrire des cartes de loi  $\mathbb{P}_{\mathbf{w}}^+$  à l'aide d'arbres de Galton-Watson à deux types alternants, *via* la bijection BDFG. La première propriété-clé est la propriété suivante de la bijection BDFG :

### 3 Cartes aléatoires

**Proposition 3.2** *La bijection BDFG associe à toute carte bipartie  $(\mathbf{m}, u) \in \mathcal{M}^+$ , un arbre spatial à deux types  $(\mathbf{t}, \mathbf{l})$  au sens de la section 2.2. De plus, pour  $d \geq 1$ , on a*

$$\#\{v \in \mathbf{m} : d_{\text{gr}}(u, v) = d\} = \#\{v \in \mathbf{t}^{(1)} : \mathbf{l}(v) - \min \mathbf{l} + 1 = d\}.$$

La seconde propriété-clef est celle qui donne la loi de la mesure-image de  $\mathbb{P}_{\mathbf{w}}^+$  par la bijection BDFG. Précisément, si  $\mathbf{w}$  est  $\mathcal{M}^*$ -admissible, nous considérons des lois de probabilités

$$\begin{aligned} \mu_1(k) &= (Z_{\mathbf{w}}^+)^{-1} f_{\mathbf{w}}(Z_{\mathbf{w}}^+)^k, & k \geq 0, \\ \mu_2(k) &= \frac{Z_{\mathbf{w}}^+ \binom{2k+1}{k} w_{k+1}}{f_{\mathbf{w}}(Z_{\mathbf{w}}^+)}, & k \geq 0. \end{aligned}$$

On considère alors deux lois  $\zeta = (\zeta_1, \zeta_2)$  sur  $\mathcal{W}_2$ , avec les notations de 2.2, de sorte que la variable canonique sous  $\zeta_i$  soit un mot de  $k$  lettres  $i + 1 \pmod{2}$  avec probabilité  $\mu_i(k)$ . De ce fait  $\zeta$  est une loi de reproduction d'arbre de Galton-Watson à deux types alternants, où un sommet de type  $i$  ne donne naissance qu'à des individus de type  $i + 1$ .

Nous définissons également, pour  $\mathbf{e} \in \mathcal{W}_2$  un mot formé sur la lettre  $i + 1 \pmod{2}$  dans  $\{1, 2\}$ ,  $\nu_{i, \mathbf{e}}$  la mesure de Dirac en  $0 \in \mathbb{R}^{|\mathbf{e}|}$  si  $i = 1$ , et, si  $i = 2$ , la loi sur  $\mathbb{Z}^{|\mathbf{e}|}$  de  $(X_1, X_1 + X_2, \dots, X_1 + \dots + X_{|\mathbf{e}|})$ , où  $(X_1, \dots, X_{|\mathbf{e}|+1})$  est uniforme dans l'ensemble

$$\{(x_1, \dots, x_{|\mathbf{e}|+1}) \in (\mathbb{Z}_+ \cup \{-1\})^{|\mathbf{e}|+1} : x_1 + \dots + x_{|\mathbf{e}|+1} = 0\}.$$

**Proposition 3.3** *Sous ces hypothèses et avec ces notations, la mesure image de  $\mathbb{P}_{\mathbf{w}}^+$  par la bijection BDFG est un arbre de Galton-Watson spatial à deux types de loi  $\mathbb{P}_{\zeta, \nu}^{(1)}$ , avec les notations de la section 2.2.*

#### 3.1.2 Cartes générales [P10]

Les résultats précédents se généralisent au cas de cartes sans contrainte sur les degrés des faces, au prix de complications techniques. Ces dernières proviennent du fait que la bijection BDFG permet de coder les cartes distribuées selon  $\mathbb{P}_{\mathbf{w}}^+, \mathbb{P}_{\mathbf{w}}^0$  à l'aide d'arbres de Galton-Watson à trois (ou quatre) types, et que la propriété d'alternance des types est perdue. Hormis cela, l'esprit des démonstrations est le même.

Soit donc  $\mathbf{w}$  une suite positive, nous supposons qu'il existe au moins un  $i \geq 3$  impair tel que  $w_i > 0$ . Nous définissons donc les fonctions de  $\mathbb{R}_+^2$  dans  $\mathbb{R}_+ \sqcup \{\infty\}$

$$\begin{aligned} f_{\mathbf{w}}^{\bullet}(x, y) &= \sum_{k, l \geq 0} x^k y^l \frac{(2k + l + 1)!}{k! (k + 1)! l!} w_{2+2k+l} \\ f_{\mathbf{w}}^{\diamond}(x, y) &= \sum_{k, l \geq 0} x^k y^l \frac{(2k + l)!}{k!^2 l!} w_{1+2k+l}. \end{aligned}$$

L'analogie de la proposition 3.1 est l'énoncé suivant

**Proposition 3.4** *La suite  $\mathbf{w}$  est  $\mathcal{M}^*$ -admissible si et seulement si le système*

$$\begin{aligned} \frac{z^+ - 1}{z^+} &= f_{\mathbf{w}}^{\bullet}(z^+, z^{\diamond}) \\ z^{\diamond} &= f_{\mathbf{w}}^{\diamond}(z^+, z^{\diamond}) \end{aligned}$$

*admet une solution  $(z^+, z^{\diamond}) \in ]0, \infty[^2$  telle que la matrice*

$$M_{(\mathbf{w})}(z^+, z^{\diamond}) := \begin{pmatrix} 0 & 0 & z^+ - 1 \\ \frac{z^+}{z^{\diamond}} \partial_x f_{\mathbf{w}}^{\diamond}(z^+, z^{\diamond}) & \partial_y f_{\mathbf{w}}^{\diamond}(z^+, z^{\diamond}) & 0 \\ \frac{(z^+)^2}{z^+ - 1} \partial_x f_{\mathbf{w}}^{\bullet}(z^+, z^{\diamond}) & \frac{z^+ z^{\diamond}}{z^+ - 1} \partial_y f_{\mathbf{w}}^{\bullet}(z^+, z^{\diamond}) & 0 \end{pmatrix}$$

*ait un rayon spectral  $\varrho_{(\mathbf{w})}$  dans  $[0, 1]$ . Si c'est le cas, cette solution est unique, et nous avons les formules  $Z_{\mathbf{w}}^+ = z^+$ ,  $Z_{\mathbf{w}}^{\diamond} = (z^{\diamond})^2$  et  $Z_{\mathbf{w}}^* = 2z^+ - 1 + (z^{\diamond})^2$ .*

Dans l'énoncé précédent, la matrice  $M_{(\mathbf{w})}$  est obtenue comme une matrice des moyennes de lois de reproduction d'un processus de Galton-Watson multitype, comme en 2.1. Nous n'en détaillerons pas ces lois de reproduction, qui s'obtiennent explicitement en termes de  $\mathbf{w}$  via les fonctions  $f_{\mathbf{w}}^{\bullet}, f_{\mathbf{w}}^{\diamond}$ .

On pose  $Z_{\mathbf{w}}^{\diamond} = z^{\diamond}$  pour une suite  $\mathcal{M}^*$ -admissible  $\mathbf{w}$ , si  $(z^+, z^{\diamond})$  est la « bonne » solution du système de la proposition 3.4. L'analogie de la définition 3.1 est la suivante.

**Définition 3.2** *On dit que la suite  $\mathbf{w}$  est critique si elle est  $\mathcal{M}^*$ -admissible et si  $\varrho_{(\mathbf{w})} = 1$ , c'est-à-dire si*

$$(Z_{\mathbf{w}}^+)^2 J_f(Z_{\mathbf{w}}^+, Z_{\mathbf{w}}^{\diamond}) + 1 = (Z_{\mathbf{w}}^+)^2 \partial_x f_{\mathbf{w}}^{\bullet}(Z_{\mathbf{w}}^+, Z_{\mathbf{w}}^{\diamond}) + \partial_y f_{\mathbf{w}}^{\diamond}(Z_{\mathbf{w}}^+, Z_{\mathbf{w}}^{\diamond}),$$

*où  $J_f$  est le jacobien de la fonction  $(f_{\mathbf{w}}^{\bullet}, f_{\mathbf{w}}^{\diamond})$ . On dit de plus que  $\mathbf{w}$  est critique régulière s'il existe  $\varepsilon > 0$  tel que  $f_{\mathbf{w}}^{\bullet}(Z_{\mathbf{w}}^+ + \varepsilon, Z_{\mathbf{w}}^{\diamond} + \varepsilon) < \infty$ .*

Nous pouvons maintenant énoncer notre résultat principal. On notera que dans l'énoncé suivant, nous nous restreignons à un conditionnement par rapport au nombre de sommets. Cela est lié au fait que nous n'avons autorisé, dans le théorème 2.1, que des conditionnements par rapport au nombre de sommets de type déterminé dans les arbres multitype. Bien que nous nous attendions à pouvoir généraliser le résultat suivant à des conditionnements plus généraux, nous n'en avons pas de preuve à l'heure actuelle.

**Théorème 3.2** *Soit  $\mathbf{w}$  une suite critique régulière, et soit  $(\mathbf{m}_n, u_n)$  une variable aléatoire de loi  $\mathbb{P}_{\mathbf{w}}^*(\cdot | \mathcal{M}_n^s)$ , et soit  $v_n$  un sommet de  $\mathbf{m}_n$  choisi uniformément conditionnellement à  $\mathbf{m}_n$ . Alors il existe une constante  $C_{\mathbf{w}} > 0$  telle que l'on ait les convergences en loi :*

$$\begin{aligned} \frac{R(\mathbf{m}_n, u_n)}{C_{\mathbf{w}} n^{1/4}} &\xrightarrow[n \rightarrow \infty]{(d)} \sup Z - \inf Z, \\ \frac{d_{\text{gr}}(u_n, v_n)}{C_{\mathbf{w}} n^{1/4}} &\xrightarrow[n \rightarrow \infty]{(d)} \sup Z, \\ \frac{\mathcal{I}_{\mathbf{m}_n, u_n}(C_{\mathbf{w}} n^{1/4} dx)}{\#V(\mathbf{m}_n)} &\xrightarrow[n \rightarrow \infty]{(d)} \mathcal{I}_{\infty}, \end{aligned}$$



### 3 Cartes aléatoires

la dernière ayant lieu pour la topologie faible sur les mesures de probabilités sur  $\mathbb{R}_+$ . Dans cet énoncé, il faut comprendre que  $n$  varie dans l'ensemble  $\{N : \mathbb{B}_{\mathbf{w}}(\mathcal{M}_N^s) > 0\}$ .

La valeur de la constante  $C_{\mathbf{w}}$  est définie implicitement par un système d'équations, et semble beaucoup plus ardue à calculer dans le cas général que dans les cas bipartis. Le calcul dans le cas le plus simple des triangulations à  $n$  sommets uniformes ( $\mathbf{w} = (0, 0, w, 0, \dots)$  avec  $w > 0$  bien choisi pour rendre  $\mathbf{w}$  critique) est déjà ardu. On obtient dans ce cas la valeur

$$C_{\text{triangulations}} = \left(\frac{4}{3}\right)^{1/4}.$$

Il est intéressant de noter que dans la preuve de notre résultat, il est nécessaire d'effectuer des opérations de permutations des différents niveaux des arbres de Galton-Watson multitypes qui codent les cartes de loi  $\mathbb{P}_{\mathbf{w}}^*$ , via la bijection BDFG. Ces opérations sont effectuées pour obtenir des déplacements spatiaux de lois centrées, et ainsi pouvoir en appeler au théorème 2.1. En quelque sorte, ces permutations brisent la structure planaire des arbres. On s'attend néanmoins à ce que cette opération ne soit pas nécessaire, voir les travaux récents de Marckert [70] allant dans ce sens.

#### 3.1.3 Cartes enracinées [P12]

Dans l'énoncé précédent, une des remarques faites après l'énoncé du théorème 3.1 s'applique, c'est-à-dire que l'on peut considérer que la carte aléatoire considérée est choisie selon  $\mathbb{P}_{\mathbf{w}}(\cdot | \mathcal{M}_n^s)$ , et que le sommet distingué, à partir duquel sont mesurés rayon, profil et distance à un second sommet marqué, est choisi uniformément parmi les  $n$  sommets de la carte. Dans ce travail avec Mathilde Weill, nous avons fourni un énoncé dans l'esprit direct des résultats initiaux de Chassaing & Schaeffer, utilisant l'origine de la racine  $e_*$  comme sommet de référence. Comme discuté plus haut, cette approche a pour inconvénient de devoir considérer des arbres avec des marques positives, ce qui induit des conditionnements des arbres de Galton-Watson spatiaux multitypes considérés plus haut.

De tels arbres conditionnés, et leurs limites d'échelle, ont été décrits et étudiés en profondeur dans les articles de Le Gall & Weill [69] et Le Gall [64]. En se basant sur ces travaux ainsi que sur [P11], Weill [85] a déterminé les limites du rayon et du profil des cartes biparties enracinées, où les distances sont mesurées par rapport à l'origine de l'arête. Notre travail est donc une généralisation de ce dernier travail. Nous généralisons l'approche de [64], introduisant un lemme de ré-enracinement pour des arbres de Galton-Watson à plusieurs types. Nous passerons sur les détails techniques, et nous bornerons à donner le résultat :

**Théorème 3.3** *Soit  $\mathbf{w}$  une suite  $\mathcal{M}^*$ -admissible, critique régulière. Elle est alors  $\mathcal{M}$ -admissible. Soit  $\mathbf{m}_n$  une variable aléatoire de loi  $\mathbb{P}_{\mathbf{w}}(\cdot | \mathcal{M}_n^s)$ , on note  $e_*$  sa racine. Soit  $v_n$  un sommet de  $\mathbf{m}_n$  choisi uniformément conditionnellement à  $\mathbf{m}_n$ . Alors, pour la*

même constante  $C_{\mathbf{w}}$  que dans le théorème 3.2, on a

$$\begin{aligned} \frac{R(\mathbf{m}_n, e_*^-)}{C_{\mathbf{w}} n^{1/4}} &\xrightarrow[n \rightarrow \infty]{(d)} \sup Z - \inf Z, \\ \frac{d_{\text{gr}}(e_*^-, v_n)}{C_{\mathbf{w}} n^{1/4}} &\xrightarrow[n \rightarrow \infty]{(d)} \sup Z, \\ \frac{\mathcal{I}_{\mathbf{m}_n, e_*^-} (C_{\mathbf{w}} n^{1/4} dx)}{\#V(\mathbf{m}_n)} &\xrightarrow[n \rightarrow \infty]{(d)} \mathcal{I}_{\infty}, \end{aligned}$$

la dernière convergence ayant lieu pour la topologie faible sur les mesures de probabilités sur  $\mathbb{R}_+$ . Dans cet énoncé,  $n$  varie dans l'ensemble  $\{N : \mathbb{B}_{\mathbf{w}}(\mathcal{M}_N^s) > 0\}$ .

### 3.2 Sphéricité de la limite d'échelle des cartes [P13]

Dans cet article, nous donnons une nouvelle preuve du théorème suivant dû à Le Gall & Paulin [68], dont le contexte est rappelé en section 1.3.3.

**Théorème 3.4** *Soit  $\mathbf{q}_n$  une quadrangulation aléatoire prise uniformément dans  $\mathbf{Q}_n$ , pour chaque  $n \geq 1$ . Soit  $\mathcal{S}$  une limite en loi des espaces métriques  $[V(\mathbf{q}_n), n^{-1/4} d_{\text{gr}}]$  pour la topologie de Gromov-Hausdorff, le long d'une sous-suite. Alors, presque-sûrement,  $\mathcal{S}$  est homéomorphe à la sphère  $\mathbb{S}^2$ .*

Notre preuve est très complémentaire de [68]. Elle fait appel à un renforcement de la convergence de Gromov-Hausdorff, ayant la vertu de préserver la nature topologique des surfaces. Par conséquent, nous raisonnons surtout sur les espaces métriques discrets  $[V(\mathbf{q}_n), n^{-1/4} d_{\text{gr}}]$ , tandis que [68] raisonne entièrement sur le quotient de l'arbre brownien défini dans [65]. Certains des résultats de [65] nous sont néanmoins nécessaires. La définition et le théorème qui suivent sont des adaptations au contexte de la métrique de Gromov-Hausdorff de résultats dûs à Whyburn. Voir par exemple Begle [11]. On notera  $\text{PMI}$  le sous-ensemble de  $\mathbb{M}$  constitué des espaces géodésiques, qui forme un sous-espace fermé de  $(\mathbb{M}, d_{\text{GH}})$ .

**Définition 3.3** *Soit  $(\mathcal{X}_n = [X_n, d_n], n \geq 0)$  une suite d'éléments de  $\text{PMI}$  qui converge vers  $\mathcal{X}$ . On dit que la convergence est 1-régulière si pour tout  $\varepsilon > 0$ , il existe  $\delta, N > 0$  tels que pour tout  $n \geq 0$ , tout lacet  $\gamma : [0, 1] \rightarrow X_n$  de diamètre au plus  $\delta$  est homotope au lacet trivial dans son  $\varepsilon$ -voisinage  $\gamma^\varepsilon := \{x : d_n(x, \gamma([0, 1])) \leq \varepsilon\}$ .*

La propriété clef que nous utilisons est le théorème suivant [11] : soit  $(\mathcal{X}_n, n \geq 0)$  une suite de  $\text{PMI}$  qui converge vers  $\mathcal{X}$ , non réduit à un point. On suppose que tous les  $\mathcal{X}_n$  sont homéomorphes à  $\mathbb{S}^2$  et que la convergence est 1-régulière. Alors  $\mathcal{X}$  est également homéomorphe à  $\mathbb{S}^2$ .

Pour pouvoir appliquer ce théorème aux quadrangulations, un lemme crucial consiste à exhiber, pour chaque quadrangulation  $\mathbf{q}$ , un élément de  $\text{PMI}$  homéomorphe à  $\mathbb{S}^2$  qui soit « proche » de  $(V(\mathbf{q}), d_{\text{gr}})$  pour la distance de Gromov-Hausdorff. Pour construire

### 3 Cartes aléatoires

un tel espace, on se donne, pour chaque face  $f \in F(\mathbf{q})$ , une copie  $X_f$  du cube unité évidé  $[0, 1]^3 \setminus ([0, 1]^2 \times [0, 1])$ , muni par exemple de la métrique  $\ell^1$ . Ensuite, on identifie par paires les quatre arêtes du bord  $\partial X_f := \{(x, y, 0) \in X_f\}$  de façon compatible avec la structure de la carte  $\mathbf{q}$ . L'espace métrique quotient [24]  $(\mathcal{S}_{\mathbf{q}}, d_{\mathbf{q}})$  ainsi obtenu est un espace métrique géodésique homéomorphe à la sphère, tel que

$$d_{\text{GH}}([\mathcal{S}_{\mathbf{q}}, d_{\mathbf{q}}], [V(\mathbf{q}), d_{\text{gr}}]) \leq 3.$$

La preuve se fait alors par l'absurde, nous en donnons une idée très heuristique. On se donne, par un argument reposant sur le théorème de représentation de Skorokhod, une suite de quadrangulations aléatoires  $\mathbf{q}_n$  respectivement uniformes dans  $\mathbf{Q}_n$ , telles que  $[V(\mathbf{q}_n), n^{-1/4}d_{\text{gr}}]$  converge vers une limite  $\mathcal{S}$ , presque-sûrement dans l'espace  $(\mathbb{M}, d_{\text{GH}})$ . On remplace alors espaces discrets par les surfaces  $\mathcal{S}_{\mathbf{q}_n}$  précédemment définies. En supposant alors qu'il existe des « goulets d'étranglement » pour les cartes  $\mathbf{q}_n$ , tels que dans la figure 3.2, nous montrons que l'arbre de la bijection de Schaeffer associé à  $\mathbf{q}_n$ , et tracé simultanément sur la surface  $\mathcal{S}_{\mathbf{q}_n}$ , est contraint d'effectuer des aller-retours par le goulet. Ceci s'obtient par les propriétés du serpent brownien positif [69] établies dans [65]. Si une telle propriété était vraie, alors deux sommets éloignés de l'arbre, dont l'un est un ancêtre de l'autre, devraient coder le même individu dans la carte limite. Ceci est exclu par les résultats de Le Gall [65].

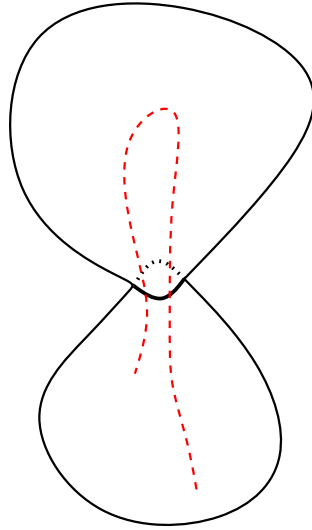


FIG. 3.2: Illustration de la preuve du théorème 3.4. En supposant l'existence de goulets d'étranglement dans les surfaces  $\mathcal{S}_{\mathbf{q}_n}$ , des branches de l'arbre de la bijection de Schaeffer (tirets) sont contraintes d'effectuer des aller-retours.

### 3.3 Bijection à plusieurs points et géodésiques [P16]

Dans ce travail, nous nous sommes intéressés à diverses propriétés combinatoires et géométriques des cartes en genre supérieur. Il a été remarqué par Marcus & Schaeffer [73] que la bijection décrite en section 1.3.1 peut se transcrire presque *verbatim* pour les surfaces de genre  $g \geq 0$  quelconque, en remplaçant, d'un côté, les quadrangulations par les quadrangulations biparties, et d'autre part, les arbres à marques positives par des cartes unicellulaires de genre  $g$  à marques positives. De ce fait, les problèmes de dénombrement des quadrangulations biparties revient à celui des cartes unicellulaires marquées, avec des marques positives.

Ce dernier problème est néanmoins difficile, et ce n'est que récemment que Chapuy, Marcus & Schaeffer [26] ont proposé une résolution du problème d'énumération bijective des cartes en genre supérieur, à l'aide d'une astuce de type « enracinement-pointage » que nous avons déjà décrite dans les sections 3.1. Cela leur a permis de retrouver les formules de dénombrement asymptotiques pour le cardinal de l'ensemble  $\mathbf{Q}_n^g$  des quadrangulations biparties de genre  $g$  à  $n$  faces :

$$\#\mathbf{Q}_n^g \underset{n \rightarrow \infty}{\sim} t_g 12^n n^{\frac{5(g-1)}{2}}, \quad (3.1)$$

établies par Bender & Canfield [12].

Nous proposons une généralisation naturelle de la bijection de Chapuy, Marcus et Schaeffer pour des cartes munies de plusieurs sommets distingués. Intuitivement, chacun de ces sommets est une source indépendante, autour de laquelle la bijection est construite de manière locale jusqu'à ce que deux sources entrent en conflit. Cette construction est directement inspirée des mosaïques de Voronoï dans le plan. Nous adjoignons à ces sources des quantités réelles, appelées *délais*, qui autorisent des retards relatifs dans la construction de la bijection autour de chaque sommet.

Plus précisément, si  $\mathbf{q}$  est une quadrangulation bipartie et  $\mathbf{x} = (x_1, \dots, x_k)$  une suite de sommets de  $\mathbf{q}$ , on considère la classe  $D = [d_1, \dots, d_k]$  d'un vecteur  $(d_1, \dots, d_k) \in \mathbb{Z}^k$  à addition d'une même constante près, et on dit que  $D$  est un délai pour  $(\mathbf{q}, \mathbf{x})$  si pour tout  $i, j \in [k]$ ,

- l'on a  $|d_i - d_j| < d_{\text{gr}}(x_i, x_j)$
- et l'on a  $d_{\text{gr}}(x_i, x_j) + d_i - d_j \in 2\mathbb{N}$ .

On notera que, si les  $d_i$  ne sont définis qu'à l'addition d'une constante commune près, les différences  $d_i - d_j$  ne sont pas ambiguës. On note  $\mathbf{Q}^{g,k}$  l'ensemble des triplets  $(\mathbf{q}, \mathbf{x}, D)$  où  $\mathbf{q}$  est une quadrangulation bipartie de genre  $g$ ,  $\mathbf{x} \in V(\mathbf{q})^k$  et  $D$  est un délai sur  $(\mathbf{q}, \mathbf{x})$ . Pour un tel triplet, on peut étiqueter les sommets de  $\mathbf{q}$  par la formule

$$\mathbf{l}(u) = \min_{1 \leq i \leq k} (d_{\text{gr}}(u, x_i) + d_i), \quad (3.2)$$

ou plus exactement, on définit ainsi un étiquetage des sommets, noté  $[\mathbf{l}]$ , à une constante additive près. La bipartition de  $\mathbf{q}$  et les conditions de parité des délais, font que les étiquettes de deux sommets voisins de la carte diffèrent toujours exactement de 1. Ainsi, si l'on oriente les arêtes de la carte de façon à ce que  $\mathbf{l}$  décroisse le long de ces arêtes,

### 3 Cartes aléatoires

un chemin orienté finira toujours en un point de  $\{x_1, \dots, x_k\}$ . On dit que l'arête  $e$  ainsi orientée est dans la  $i$ -ème cellule de sources  $\mathbf{x}$  et de délais  $D$ , si le chemin orienté commençant par  $e$  et tournant le plus à gauche possible à chaque étape, termine en  $x_i$ . Ceci définit une partition des arêtes (orientées)  $E_i(\mathbf{q}), 1 \leq i \leq k$ .

D'autre part, une carte marquée à  $k$  faces est un couple  $(\mathbf{m}, [\mathbf{l}])$  où  $\mathbf{m}$  est une carte en genre  $g$ , dont les  $k$  faces sont étiquetées  $f_1(\mathbf{m}), \dots, f_k(\mathbf{m})$ , et  $\mathbf{l} = (\mathbf{l}(u), u \in V(\mathbf{m}))$  est une fonction de  $V(\mathbf{m})$  dans  $\mathbb{Z}$  vérifiant

$$|\mathbf{l}(u) - \mathbf{l}(v)| \leq 1, \quad u \text{ voisin de } v \text{ dans } \mathbf{m},$$

et  $[\mathbf{l}]$  est sa classe à une constante additive entière près. On note  $\mathbf{LM}^{g,k}$  l'ensemble des cartes marquées à  $k$  faces en genre  $g$ . On écrira aussi  $e \in f_i(\mathbf{m})$  si l'arête orientée  $e$  est telle que la face  $f_i(\mathbf{m})$  est à sa droite. Nous établissons le résultat suivant.

**Théorème 3.5** *Il existe une bijection entre les ensembles  $\mathbf{Q}^{g,k}$  et  $\mathbf{LM}^{g,k} \times \{0, 1\}$ , telle que, si  $(\mathbf{q}, \mathbf{x}, D)$  et  $((\mathbf{m}, [\mathbf{l}]), I)$  sont associées par cette bijection ( $I \in \{0, 1\}$ ), on a pour tout  $i$  et  $d \geq 0$  :*

$$\#\{e \in E_i(\mathbf{q}) : d_{\text{gr}}(e^-, x_i) = d\} = \#\left\{e \in f_i(\mathbf{m}) : \mathbf{l}(e^-) - \min_{e_0 \in f_i(\mathbf{m})} \mathbf{l}(e_0^-) + 1 = d\right\}.$$

Dans cette bijection, l'indice  $I$  est juste présent pour déterminer une convention dans l'enracinement de  $\mathbf{q}$  : comme pour la bijection BDFG, un cadre plus naturel au théorème précédent est celui des cartes non-enracinées. La bijection se construit simplement en appliquant la construction de Schaeffer de la figure 1.4, à la carte étiquetée  $(\mathbf{q}, \mathbf{l})$  définie par (3.2). Ainsi, on effectue la bijection de Schaeffer « localement » à l'intérieur des cellules  $E_i(\mathbf{q})$ , dépliant cette dernière en une face de la carte finale  $\mathbf{m}$ .

Cette bijection étant construite, nous obtenons les limites d'échelle des cartes marquées de  $\mathbf{LM}^{g,k}$ , lorsque l'on fait tendre le nombre d'arêtes vers  $\infty$ , tout en gardant  $k$  fixé. Plutôt que de considérer des mesures de probabilités, le cadre dans lequel nous nous plaçons<sup>1</sup> est de considérer les lois de Boltzmann sur ces cartes, où l'on met un poids  $\alpha^{\#E(\mathbf{m})}$  sur la carte  $\mathbf{m}$  pour un  $\alpha > 0$ , le poids « critique » naturel étant  $\alpha = 12^{-1}$ , comme suggéré par la formule (3.1). Pour décrire ces limites d'échelle, nous procédons à une décomposition des cartes marquées  $(\mathbf{m}, [\mathbf{l}])$  en éléments plus simples, dans l'esprit de [26] : essentiellement, une carte marquée est constituée d'une carte à  $k$  faces dont tous les sommets sont de degré au moins trois, et dont on remplace chaque arête par une forêt étiquetée.

Les limites d'échelle sont proches en esprit des *mesures de Hurwicz* introduites par Okounkov & Pandharipande [75], bien que la présence des marques induisent des conditions de compatibilité supplémentaires, qui en rendent la description plus compliquée.

Parmi les applications de ces résultats, nous donnons une preuve « probabiliste » de l'équivalent (3.1), avec une expression intégrale de  $t_g$ , pour  $g \geq 1$ , différente de [26] :

$$t_g = 16 \left(\frac{3}{64}\right)^g \frac{\Gamma(5g-3)}{(6g-3)\Gamma\left(\frac{5g-3}{2}\right)} \sum_{\mathbf{m} \in \mathbf{M}_{g,1}^3} \int_{\Delta_E} \frac{\lambda(d\mathbf{u})}{\sqrt{\mathcal{Z}^{\mathbf{m}}(\mathbf{u})}}.$$

<sup>1</sup>Dans la dénomination habituelle en physique, ceci correspond au formalisme « grand-canonique »

### 3 Cartes aléatoires

Ici, la somme est sur les cartes unicellulaires dont les sommets sont tous de degré 3,  $\Delta_E$  est le simplexe  $\{\mathbf{u} = (u_e, e \in E(\mathbf{m})) \in \mathbb{R}_+^{E(\mathbf{m})} : \sum_{e \in E(\mathbf{m})} u_e = 1\}$  et  $\lambda_{\Delta}$  est la mesure de Lebesgue sur  $\Delta_E$ , et enfin,

$$\mathcal{Z}^{\mathbf{m}}(\mathbf{u}) = \sum_{\mathbf{a} \in \text{ST}(\mathbf{m})} \prod_{e \notin E(\mathbf{a})} 2\pi u_e,$$

où  $\text{ST}(\mathbf{m})$  est l'ensemble des arbres couvrants de  $\mathbf{m}$ .

La deuxième application de la bijection du théorème 3.5 et une propriété d'unicité de la géodésique typique entre deux points des espaces limites des cartes, en genre  $g$ . Nous sommes amenés à considérer de nouvelles topologies sur les classes d'isométries d'espaces métriques. Nous introduisons donc, dans l'esprit de la construction de la métrique de Gromov-Prokhorov de [47] rappelée en 1.2.5, une métrique dite de Gromov-Hausdorff-Prokhorov  $d_{\text{GHP}}$  sur l'espace  $\mathbb{M}_w$  des classes d'isométries  $[X, d, \mu]$  d'espaces métriques *compacts* mesurés. La métrique  $d_{\text{GHP}}$  est plus « forte » que  $d_{\text{GP}}$ , au prix de ne considérer que les espaces compacts plutôt que les espaces complets, par ailleurs, on ne se restreint pas au cas des mesures de support plein. On définit ainsi

$$d_{\text{GP}}([X, d, \mu], [X', d', \mu']) = \inf_{\phi, \phi'} \delta_H(\phi(X), \phi'(X')) \vee \delta_P(\phi_*\mu, \phi'_*\mu'),$$

où l'infimum est pris, encore une fois, sur les isométries  $\phi, \phi'$  de  $X, X'$  dans un espace commun  $(Z, \delta)$ . Ceci définit une distance complète sur  $\mathbb{M}_w$ , qui est induit la même topologie que celle considérée dans l'article d'Evans & Winter [42] et Fukaya [45].

Revenons maintenant aux cartes. Ici encore, nous nous plaçons dans le cadre d'une loi de Boltzmann plutôt que de considérer des cartes avec un nombre fixé de faces. Ainsi, nous notons  $\mathbf{Q}^g$  la mesure  $\sigma$ -finie sur l'ensemble  $\mathbf{Q}^g$  des quadrangulations biparties en genre  $g$ , donnée par

$$\mathbf{Q}^g(\{\mathbf{q}\}) = 12^{-\#F(\mathbf{q})}, \quad \mathbf{q} \in \mathbf{Q}^g.$$

A partir de celle mesure, on construit, pour chaque  $\beta > 0$ , une mesure de probabilités sur  $\mathbf{Q}^g$  par la formule

$$\mathbf{Q}_\beta^g(d\mathbf{q}) = \frac{\mathbf{Q}^g(e^{-\beta\#F(\mathbf{q})} d\mathbf{q})}{\mathbf{Q}^g(e^{-\beta\#F(\mathbf{q})})}, \quad g \geq 1,$$

et

$$\mathbf{Q}_\beta^0(d\mathbf{q}) = \frac{\mathbf{Q}^g(V_{\mathbf{q}}^2 e^{-\beta\#F(\mathbf{q})} d\mathbf{q})}{\mathbf{Q}^g(V_{\mathbf{q}}^2 e^{-\beta\#F(\mathbf{q})})},$$

où l'on a noté  $V_{\mathbf{q}} = \#V(\mathbf{q})$ , et où la présence du facteur  $V_{\mathbf{q}}^2$  dans le cas planaire amène certaines simplifications techniques dans les passages à la limite.

À chaque  $\mathbf{q} \in \mathbf{Q}^g$  on associe la classe d'espaces métrique mesurés  $\mathcal{X}_{\mathbf{q}} = [V(\mathbf{q}), d_{\text{gr}}, \mu_{\mathbf{q}}]$ , où

$$\mu_{\mathbf{q}} = \frac{1}{V_{\mathbf{q}}} \sum_{v \in V(\mathbf{q})} \delta_v$$

est la probabilité uniforme sur les sommets de  $\mathbf{q}$ . Nous montrons alors les résultats suivants.

### 3 Cartes aléatoires

**Théorème 3.6** *Soit  $g \geq 0$  et  $\beta > 0$  fixés. Alors les mesures de probabilités*

$$\mathbb{Q}_{\beta/a}^g(\{a^{-1/4}\mathcal{X}_{\mathbf{q}} \in \cdot\}), \quad a > 1,$$

*sont tendues dans l'ensemble des probabilités sur  $(\mathbb{M}_w, d_{\text{GHP}})$  muni de la topologie faible.*

*De plus, toute valeur d'adhérence de ces mesures, lorsque  $a \rightarrow \infty$ , est portée par les espaces  $[X, d, \mu]$  tels que*

1.  *$(X, d)$  est un espace métrique géodésique,*
2.  *$\mu$  est une mesure diffuse de support plein,*
3. *pour  $\mu \otimes \mu$  presque tout  $x, y \in X$ , il existe une unique géodésique liant  $x$  à  $y$ .*

Jusqu'au point 2., on peut rapprocher ces résultats de ceux obtenus par Le Gall en genre 0 dans [65], où la construction de  $\mu$  est donnée de façon différente, plus directe, sans faire référence à la convergence d'espaces métriques mesurés. Des résultats similaires d'unicité de la géodésique typique ont également été obtenus par Le Gall dans [66], en genre 0. Ici, nos résultats portent sur des cartes dont le nombre de faces est aléatoire et non fixé comme dans ces dernières références. La renormalisation du paramètre  $\beta/a$  est conçue pour faire tendre ce nombre de faces vers  $\infty$  lorsque  $a \rightarrow \infty$ .

Nous obtenons le troisième point de ce théorème, qui est le plus important, en utilisant la bijection du théorème 3.5 à deux points marqués. Ceci permet intuitivement de tracer le lieu géométrique  $\mathcal{L} = \{u : d(x_1, u) = d(x_2, u) + D\}$  entre deux points marqués  $x_1, x_2$  de l'espace limite, et pour un réel  $|D| < d(x_1, x_2)$ . On montre que les distances à  $x_1$  le long de ce lieu suivent un processus de type « pont brownien », qui atteint donc son minimum une unique fois, ce que l'on peut relier à la localisation de la géodésique entre  $x_1$  et  $x_2$  autour d'un unique point du lieu géométrique  $\mathcal{L}$ .

## 4 Fragmentations auto-similaires et arbres de fragmentation

Dans cette partie, si  $(f_t, t \geq 0)$  est une fonction càdlàg, on notera, pour  $t \geq 0$ ,  $\Delta f_t = f_t - f_{t-}$  le saut de  $f$  à l'instant  $t$ , et  $\Delta f_{[0,t]}$  la suite des sauts de  $f$  effectués avant l'instant  $t$  (compris), et rangés par ordre décroissant.

### 4.1 Fragmentations auto-similaires et subordinateurs stables [P2]

On a vu à la section 1.4 que les fragmentations auto-similaires sont classifiées en fonction de leurs caractéristiques locales  $(\beta, c, \nu)$ . Il est nettement plus difficile d'obtenir une information sur le semigroupe de ces processus, une exception notable étant la fragmentation  $F_+^{(2)}$  décrite en 1.4.2. Avec Jason Schweinsberg, nous nous sommes intéressés dans [P2] à la question de Pitman qui suit, où  $\sigma^{(\beta)}$  est un subordinateur stable d'indice  $\beta$  :

Pour quelles valeurs de  $\beta \in (0, 1)$  peut-on trouver un processus de fragmentation auto-similaire  $(F(t), t \geq 0)$  tel que la loi de  $F(t)$  soit égale à celle de  $\Delta\sigma_{[0,t]}^{(\beta)}$ , conditionnellement à  $\sigma_t^{(\beta)} = 1$  ?

Nous avons vu en 1.4.2 que  $\beta = 1/2$  est une valeur pour laquelle une telle fragmentation existe bien. Nous montrons que c'est la seule :

**Théorème 4.1** *Soit  $\beta \in ]0, 1[$  et  $\sigma^{(\beta)}$  un subordinateur stable d'indice  $\beta$ . Soit  $(F(t), t \geq 0)$  une fragmentation auto-similaire de caractéristiques  $(\gamma, 0, \nu)$ . On suppose que pour tout  $t \geq 0$ , la loi de  $F(t)$  est celle de  $\Delta\sigma_{[0,t]}^{(\beta)}$  sachant  $\sigma_t^{(\beta)} = 1$ . Alors  $\beta = \gamma = 1/2$ , et il existe  $C > 0$  telle que  $(F(t), t \geq 0)$  a même loi que  $(F_+^{(2)}(Ct), t \geq 0)$ .*

Au cours de la preuve, nous obtenons une construction « à la Lévy » de la mesure de dislocation d'une fragmentation auto-similaire d'indice positif *via* la proposition suivante.

**Proposition 4.1** *Soit  $(F(t), t \geq 0)$  une fragmentation auto-similaire de caractéristiques  $(\gamma, 0, \nu)$  avec  $\gamma \geq 0$ . Alors pour toute fonction  $G$  sur  $\mathcal{S}^\downarrow$  positive continue, nulle sur un voisinage de  $(1, 0, \dots)$ , on a*

$$\frac{1}{t} E[G(F(t))] \xrightarrow[t \rightarrow 0]{} \langle \nu, G \rangle.$$



## 4.2 Deux fragmentations de l'arbre stable [P3,P7]

Dans ces deux articles, nous avons cherché à généraliser la construction des deux fragmentations « duales » de l'arbre continu brownien, décrites en 1.4.2, au cas des arbres stables introduits à la section 1.2.4. Du fait des différences de structure des points de branchement déjà mentionnées plus haut entre les arbres stables d'indice  $\alpha \in ]1, 2[$  et l'arbre brownien, les constructions et les propriétés de ces fragmentations sont assez différentes de  $F_+^{(2)}, F_-^{(2)}$ .

### 4.2.1 Fragmentation par la hauteur [P3]

Nous commençons par la fragmentation la plus simple à construire, qui, elle, est une généralisation immédiate de la construction de  $F_-^{(2)}$ . Soit  $\alpha \in ]1, 2[$ , et  $H^{(\alpha)}$  le processus des hauteurs de l'arbre stable  $\mathcal{T}^{(\alpha)}$ . Pour chaque  $t \geq 0$ , on note  $F_-^{(\alpha)}(t) \in \mathcal{S}^\downarrow$  la suite décroissante des longueurs des composantes connexes de l'ouvert de  $[0, 1]$  suivant :

$$\{s \in [0, 1] : H_s^{(\alpha)} > t\}.$$

Autrement dit,  $F_-^{(\alpha)}(t)$  est la suite des  $\mu^{(\alpha)}$ -masses des composantes connexes de  $\{x \in \mathcal{T}^{(\alpha)} : \text{ht}(x) > t\}$  classées par ordre décroissant (rappelons la définition de  $\mu^{(\alpha)}$  à la fin de 1.2.5).

**Théorème 4.2** *Le processus  $(F_-^{(\alpha)}(t), t \geq 0)$  est une fragmentation auto-similaire d'indice  $1/\alpha - 1 \in ]-1/2, 0[$ , de coefficient d'érosion  $c = 0$ , et de mesure de dislocation  $\nu^{(\alpha)}$  caractérisée par : pour toute fonction  $G$  positive mesurable sur  $\mathcal{S}^\downarrow$ ,*

$$\langle \nu^{(\alpha)}, G \rangle = D_\alpha E \left[ \sigma_1^{(1/\alpha)} G \left( \Delta \sigma_{[0,1]}^{(1/\alpha)} \right) \right], \quad (4.1)$$

où dans un espace probabilisé  $(\Omega, \mathcal{F}, P)$ , le processus  $\sigma^{(1/\alpha)}$  est un subordinateur stable d'exposant de Laplace  $\lambda^{1/\alpha}$  et  $D_\alpha = \alpha(\alpha - 1)\Gamma(1 - 1/\alpha)/\Gamma(2 - \alpha) > 0$ .

Pour démontrer ce théorème, nous prouvons tout d'abord la propriété d'auto-similarité du processus de hauteur qui suit : sachant que  $F_-^{(\alpha)}(t) = (s_1, s_2, \dots)$ , les excursions du processus de hauteur  $H^{(\alpha)}$  au-dessus du niveau  $t$  sont de la forme

$$(s_i^{1-1/\alpha} H^{(\alpha),i}(u/s_i), 0 \leq u \leq s_i), \quad i \geq 1,$$

où les processus  $H^{(\alpha),i}, i \geq 1$  sont des copies indépendantes de  $H^{(\alpha)}$ . Ceci s'obtient grâce à la théorie des excursions d'Itô et en s'appuyant sur des résultats de Duquesne et Le Gall [35] sur le comportement du processus de hauteur au-dessus d'un niveau fixé. On peut également en déduire les marginales de dimension 1 (donc le semigroupe) de  $F_-^{(\alpha)}$  en considérant le temps local du processus de hauteur au niveau  $t$ .

Notons que la mesure  $\nu^{(\alpha)}$  ressemble de très près aux mesures de Poisson-Dirichlet à deux paramètres et aux partitions (« restaurants chinois ») associées, qui ont été introduites par Pitman [78]. De façon abusive, on peut considérer que  $\nu^{(\alpha)}$  est une loi de

Poisson-Dirichlet de paramètres  $(1/\alpha, -1)$ , bien que ces paramètres soient incompatibles selon les notations de [78] (d'ailleurs, la mesure n'est pas une mesure de probabilité, elle est infinie).

On notera enfin que, contrairement à  $\nu^{(2)}$ , les mesures  $\nu^{(\alpha)}$  pour  $\alpha \in ]1, 2[$  ne sont pas binaires. Chaque dislocation soudaine casse un fragment en une infinité d'autres, dont la somme des masses égale celle du fragment initial. Cette propriété est bien sûr reliée au fait que les points de branchement de  $\mathcal{T}^{(\alpha)}$  sont d'ordre infini.

#### 4.2.2 Fragmentations aux points de branchement [P7]

Pour construire une fragmentation d'indice positif de l'arbre stable, et de mesure de dislocation  $\nu^{(\alpha)}$ , on voit que le procédé de construction de  $F_+^{(2)}$  pour l'arbre brownien est voué à l'échec, puisqu'une telle construction ne marque que des points du squelette qui ne sont pas des points de branchement (en quantité dénombrable), ce qui signifie que la construction donne nécessairement une fragmentation binaire. On considère donc un procédé de marquage des points de branchement de l'arbre.

Pour ce faire, on utilise le fait que l'on peut définir des mesures de temps local sur l'arbre stable [35, 36]. Spécifiquement, on peut presque-sûrement associer à chaque élément  $x \in \text{Br}(\mathcal{T}^{(\alpha)})$  une quantité strictement positive, appelée temps local de  $x$ . On définit, en rappelant la notation (1.4)

$$\ell(x) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \mu^{(\alpha)}(\{y \in \text{Desc}(\mathcal{T}^{(\alpha)}, x) : \text{ht}(y) - \text{ht}(x) < \varepsilon\}) > 0 \quad p.s. \quad (4.2)$$

La construction est alors la suivante : on se donne une mesure de Poisson  $\mathcal{N}(dsdx)$  sur  $\mathbb{R}_+ \times \mathcal{T}^{(\alpha)}$ , dont l'intensité est

$$ds \mathbb{1}_{\{s \geq 0\}} \sum_{x \in \text{Br}(T)} \ell(x) \delta_x.$$

On note

$$C_t = \{x \in \text{Br}(\mathcal{T}^{(\alpha)}) : \exists s \in [0, t], \mathcal{N}(\{(x, s)\}) > 0\}.$$

De cette façon,  $C_t$  correspond à un marquage indépendant des points de branchement de  $\mathcal{T}^{(\alpha)}$  avec une probabilité égale à  $1 - \exp(-t\ell(x))$  pour le point  $x$ . On note alors  $F_+^{(\alpha)}(t)$  la suite décroissante des  $\mu^{(\alpha)}$ -masses des composantes connexes de  $\mathcal{T}^{(\alpha)} \setminus C_t$ .

**Théorème 4.3** *Le processus  $(F_+^{(\alpha)}(t), t \geq 0)$  est une fragmentation auto-similaire d'indice  $1/\alpha$ , de coefficient d'érosion  $c = 0$ , et de mesure de dislocation  $\nu^{(\alpha)}$ .*

La preuve de ce résultat repose sur la construction du processus des hauteurs de  $\mathcal{T}^{(\alpha)}$  à partir de la trajectoire d'une excursion de processus de Lévy stable d'indice  $\alpha$  spectralement positif. On traduit le processus de fragmentation ci-dessus en termes d'une décomposition trajectorielle de cette excursion, où le marquage des points de branchement se ramène à un marquage des sauts, dont les amplitudes correspondent précisément aux temps locaux.

Une autre propriété notable de la fragmentation  $F^+$  est qu'elle admet une autre représentation, plus simple et permettant un calcul de son semigroupe. On l'obtient également en marquant les sauts d'une excursion de processus de Lévy exactement de la même façon, mais en remplaçant la décomposition trajectorielle par un simple effacement des sauts marqués. Cette représentation rappelle la construction alternative de  $F_+^{(2)}$  décrite en 1.5, où l'on ajoute une dérive à l'excursion brownienne.

Nous décrivons brièvement cette seconde construction. Soit  $(e_u^{(\alpha)}, 0 \leq u \leq 1)$  l'excursion normalisée du processus de Lévy stable d'indice  $\alpha$ , d'exposant de Laplace  $\lambda^\alpha$ . Pour une construction, on pourra se rapporter au dernier chapitre de [14]. On considère alors une mesure de Poisson  $\mathcal{N}(dsdu)$  sur  $\mathbb{R}_+ \times [0, 1]$ , d'intensité

$$ds \mathbb{1}_{\{s \geq 0\}} \sum_{\substack{u \in [0, 1] \\ \Delta e_u^{(\alpha)} > 0}} \Delta e_u^{(\alpha)} \delta_u.$$

On pose

$$C_t = \{u \in [0, 1] : \exists s \in [0, t], \mathcal{N}(\{(s, u)\}) > 0\}.$$

Soit alors

$$J_u^{(t)} = \sum_{r \in [0, u] \cap C_t} \Delta e_r^{(\alpha)},$$

et  $e_u^{(\alpha), (t)} = e_u^{(\alpha)} - J_u^{(t)}$ ,  $0 \leq u \leq 1$ . Enfin, on note  $\underline{e}^{(\alpha), (t)}$  le processus de l'infimum passé correspondant, et  $F_{\natural}^{(\alpha)}(t)$  la suite décroissante des longueurs des composantes connexes du complémentaire du support de la mesure de Stieltjes  $-d\underline{e}^{(\alpha), (t)}$ .

**Théorème 4.4** *Le processus  $(F_{\natural}^{(\alpha)}(t), t \geq 0)$  a même loi que  $F^+$ .*

Nous montrons ce résultat par le calcul du semigroupe de  $F_{\natural}^{(\alpha)}$ , ce qui donne donc au passage le semigroupe de  $F_+^{(\alpha)}$ . Ceci est rendu possible par des relations d'absolue continuité à la Girsanov, entre le processus de Lévy totalement asymétrique d'indice  $\alpha$  et le processus obtenu par effacement de ses sauts marqués, obtenu d'une façon analogue à la construction ci-dessus pour l'excursion. Le théorème de Vervaat pour les processus stables et une décomposition du pont « à la Williams » en son minimum, dus à Chaumont [30], sont deux autres outils-clé. On retrouve ensuite la mesure de dislocation par la Proposition 4.1, ce qui permet de conclure.

### 4.3 Généalogie des fragmentations auto-similaires d'indice négatif [P5]

La motivation de ce travail avec Bénédicte Haas est de déterminer dans quelle mesure une fragmentation auto-similaire admet une représentation du type de la fragmentation  $F_-^{(\alpha)}$  étudiée ci-dessus. En d'autres termes, une fragmentation auto-similaire  $F$  est-elle une fragmentation d'un arbre continu  $(\mathcal{T}, \mu)$ , au sens où  $F$  a la même loi que

la suite décroissante des masses des composantes connexes de  $\{x \in \mathcal{T} : \text{ht}(x) > t\}$ ? Les arbres continus que nous avons considérés jusqu'ici sont compacts, et on voit donc que l'ensemble précédent est vide pour  $t$  assez grand. On s'intéresse donc *a priori* à des fragmentations d'indice strictement négatif, puisqu'elles s'éteignent en temps fini.

Il est alors assez naturel de construire un arbre généalogique de la fragmentation  $F$  de la façon qui suit. On suppose que  $F$  est d'indice  $\beta < 0$ , n'a pas d'érosion (ceci correspond au fait que la mesure de masse d'un arbre continu ne doit pas charger le squelette) et que la mesure de dislocation  $\nu \neq 0$  vérifie  $\nu(\{\mathbf{s} \in \mathcal{S}^\downarrow : s_0 > 0\}) = 0$ , de sorte qu'il n'y a pas création soudaine de poussière. Dans ce dernier cas, on dit que  $\nu$  est *conservative*. On se donne une représentation « partitions »  $(\Pi(t), t \geq 0)$  de  $F$ , comme discuté en 1.4.1. L'extinction signifie que  $\Pi$  atteint l'état  $\mathbf{1}_\mathbb{N}$  en temps fini.

À chaque  $i \in \mathbb{N}$ , on associe un « temps de vie »  $D_i = \inf\{t \geq 0 : \{i\} \in \Pi(t)\}$  qui marque la fin de la lignée de l'individu  $i$ , et la structure généalogique associée à  $\Pi$  est la structure naturelle : deux blocs  $B_1$  et  $B_2$  au temps  $t + t'$  qui faisaient partie du même bloc au temps  $t$  sont des fils de ce bloc.

Plus précisément, pour chaque  $k$ , on associe à  $\Pi$  une classe d'isométrie d'arbres réels  $\mathcal{T}_{[k]}^{\beta, \nu}$  enraciné à  $k$  feuilles distinguées, dont voici un représentant. Les sommets de l'arbre sont les couples  $(B, t)$ , où  $B$  est une intersection d'un bloc de  $\Pi(t-)$  avec  $[k]$ , et où de plus  $t \leq D_i$  si  $B = \{i\}$ . Si  $(B, t), (B', t') \in \mathcal{T}_{[k]}$ , on considère le plus grand temps  $t'' \leq t \wedge t'$  tel que  $B$  et  $B'$  sont inclus dans un même bloc  $B''$  de  $\Pi(t'')$ , et on définit

$$d_{[k]}((B, t), (B', t')) = t + t' - 2t'' .$$

L'arbre est enraciné en  $\rho = ([k], 0)$  et les  $k$  feuilles marquées sont  $(\{i\}, D_i), 1 \leq i \leq k$ . De cette façon, l'ancêtre commun à  $(B, t)$  et  $(B', t')$  de hauteur maximale est donné par  $(B'', t'')$ , c'est-à-dire que  $[[\rho, (B, t)] \cap [[\rho, (B', t')]] = [[\rho, (B'', t'')]]$ . Cet arbre réel correspond bien à la notion intuitive de généalogie associée à un processus croissant à valeurs dans les partitions de  $[k]$ . Voir la figure 4.1.

Par les propriétés d'échangeabilité de  $\Pi$ , on construit ainsi une famille cohérente d'arbres réels enracinés à  $k$  feuilles numérotées, au sens discuté en section 1.2.5. Le théorème d'Aldous [3] que nous avons mentionné alors permet de conclure à l'existence d'un(e classe d'isométrie d') arbre continu aléatoire mesuré, que nous notons  $(\mathcal{T}^{\beta, \nu}, \mu^{\beta, \nu})$ , tel que le sous-arbre engendré par la racine et une suite échangeable de  $k$  termes dirigée par la mesure  $\mu^{\beta, \nu}$  a même loi que  $\mathcal{T}_{[k]}^{\beta, \nu}$ . Cet arbre répond à la question posée :

**Théorème 4.5** *Soit  $F$  une fragmentation auto-similaire de caractéristiques  $(\beta, 0, \nu)$ , où  $\beta < 0$  et  $\nu \neq 0$  est conservative. Alors le processus  $(F'(t), t \geq 0)$ , où  $F'(t)$  est la suite décroissante des  $\mu^{\beta, \nu}$ -masses des composantes connexes de l'ouvert  $\{x \in \mathcal{T}^{\beta, \nu} : \text{ht}(x) > t\}$ , a même loi que  $F$ .*

Ce résultat admet une forme de réciproque, dans le sens où tout arbre continu aléatoire possédant une certaine forme simple d'auto-similarité est de la forme  $\mathcal{T}^{\beta, \nu}$  pour une fragmentation auto-similaire  $F$  satisfaisant aux hypothèses ci-dessus.

Nous montrons par ailleurs un résultat sur la dimension de Hausdorff de l'arbre  $\mathcal{T}^{\beta, \nu}$ .

#### 4 Fragmentations auto-similaires et arbres de fragmentation

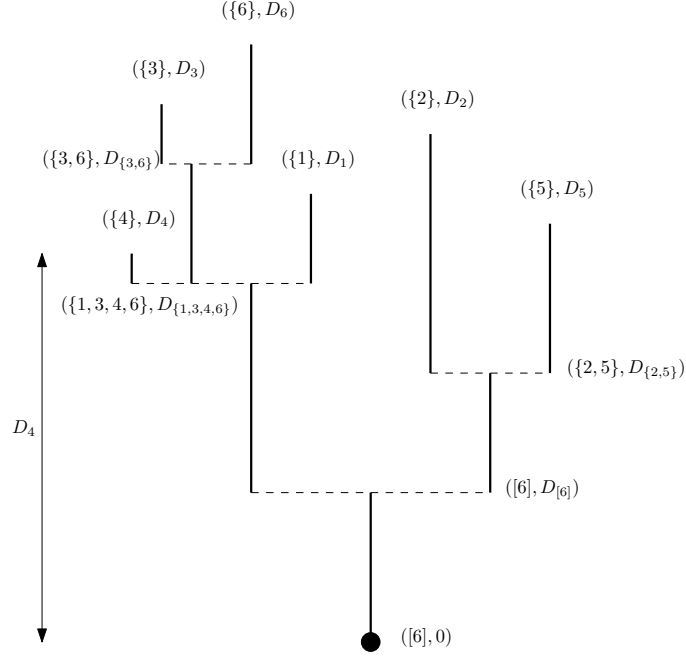


FIG. 4.1: L'arbre  $\mathcal{T}_{[6]}$  associé à un processus de fragmentation  $\Pi$ . Les lignes pointillées doivent être vues comme ayant une longueur 0, et n'existent pas dans l'arbre. On a noté ici  $D_B$  le supremum des instants  $t$  où  $B$  est strictement inclus dans un bloc de  $\Pi(t)$ . Ainsi, le point  $(\{2, 5\}, D_{\{2,5\}})$  est le point de branchement entre les feuilles  $(\{2\}, D_2)$  et  $(\{5\}, D_5)$ .

**Théorème 4.6** Soit  $F$  une fragmentation auto-similaire de caractéristiques  $(\beta, 0, \nu)$ , avec  $\beta < 0, \nu \neq 0$  et  $\nu(\{\mathbf{s} : s_0 > 0\}) = 0$ . On suppose que pour un (et donc tout)  $\varepsilon \in ]0, 1[$ , on a

$$\int_{\mathcal{S}^\downarrow} s_1^{-1} \mathbb{1}_{\{s_1 < 1-\varepsilon\}} \nu(ds) < \infty.$$

Alors la dimension de Hausdorff de l'arbre  $\mathcal{T}^{\beta, \nu}$  est, presque-sûrement,  $\dim_{\mathcal{H}}(\mathcal{T}^{\beta, \nu}) = (1/|\beta|) \wedge 1$ .

En particulier, notons qu'il est trivial que  $\dim_{\mathcal{H}}(\mathcal{T}^{\beta, \nu}) \geq 1$  puisque l'arbre contient toujours au moins une copie isométrique à un segment réel. Par ailleurs, l'hypothèse sur  $\nu$  est toujours vérifiée dans le cas où la fragmentation est binaire, ou plus généralement dans le cas où elle est *au plus n-aire*, au sens où  $\nu(\{\mathbf{s} : s_1 + \dots + s_n < 1\}) = 0$  pour un  $n \geq 2$ .

L'un des intérêts du Théorème 4.6 est qu'il peut s'appliquer au cas de l'arbre stable ( $\alpha \in ]1, 2[$ ), puisqu'on déduit des considérations précédentes que l'arbre stable  $\mathcal{T}^{(\alpha)}$  a même loi que l'arbre  $\mathcal{T}^{1/\alpha-1, \nu^{(\alpha)}}$  associé à la fragmentation  $F_-^{(\alpha)}$ , avec les notations de 4.2.1. On peut vérifier les hypothèses du Théorème 4.6 et en déduire

**Corollaire 4.1** *La dimension de Hausdorff de l'arbre stable  $\mathcal{T}^{(\alpha)}$  est égale à  $\alpha/(\alpha - 1)$  presque-sûrement.*

Ce résultat a été obtenu indépendamment par Duquesne et Le Gall [36]. Enfin, nous discutons également dans [P5] l'existence de processus de hauteur pour les arbres  $\mathcal{T}^{\beta,\nu}$ , c'est-à-dire de processus  $(H_t^{\beta,\nu}, 0 \leq t \leq 1)$  tels que (la classe d'isométrie de)  $\mathcal{T}^{\beta,\nu}$  a même loi que (la classe de)  $(\mathcal{T}_{H^{\beta,\nu}}, d_{H^{\beta,\nu}})$ . Ceci se fait en considérant des versions ordonnées (planaires) des arbres réduits  $\mathcal{T}_{[k]}^{\beta,\nu}$  construits plus haut. Des résultats sur la régularité Hölder de  $H^{\beta,\nu}$  sont également donnés.

## 4.4 Cladogrammes markoviens et leurs limites continues [P14]

Dans ce travail en commun avec Bénédicte Haas, Jim Pitman et Matthias Winkel, nous nous sommes intéressés à des modèles d'arbres ayant des motivations biologiques, appelés cladogrammes, ou arbres phylogénétiques. On renvoie à Aldous [1] pour une introduction au sujet, avec les motivations biologiques.

D'un point de vue mathématique, un cladogramme est un arbre (graphe sans cycle) enraciné, dont les sommets internes sont tous de degré au moins 3, et dont les feuilles distinctes de la racine sont numérotées. Plus formellement :

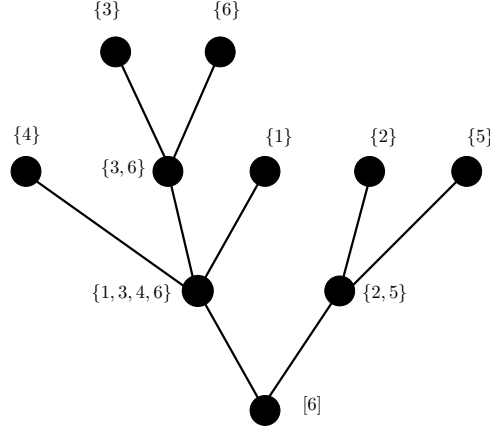
**Définition 4.1** *Un cladogramme à  $n$  feuilles est une collection  $\check{\mathbf{t}}$  de sous-ensembles de  $[n]$ , telle que*

- $\check{\mathbf{t}}$  contient  $[n]$  (la racine),
- $\check{\mathbf{t}}$  contient  $\{i\}, 1 \leq i \leq n$  (les feuilles), et
- si  $B, B' \in \check{\mathbf{t}}$ , on a  $B \cap B' = \emptyset$  ou  $B \subseteq B'$  ou  $B' \subseteq B$ .

On note  $\check{\mathbf{T}}_n$  l'ensemble des cladogrammes à  $n$  feuilles.

Pour les représenter graphiquement, un sommet correspond à un élément de  $\check{\mathbf{t}}$ , et on dit que  $B$  est le père de  $B'$  si et seulement si  $B$  est inclus strictement dans  $B'$ , et aucun sommet  $B'' \in \check{\mathbf{t}} \setminus \{B, B'\}$  ne vérifie  $B \subseteq B'' \subseteq B'$ . On peut aussi naturellement associer un cladogramme (mais pas de façon univoque) à une suite  $(\pi_1, \dots, \pi_r)$  de partitions de  $[n]$ , strictement croissante, issue de  $\pi_1 = \{\{[n]\}$  et finissant en  $\pi_r = \{\{1\}, \dots, \{n\}\}$ . Dans la figure 4.2, il faut comprendre que la structure planaire n'intervient pas dans la définition du cladogramme. Comme pour les éléments de  $\mathbf{T}_n$  (arbres plans) et  $\hat{\mathbf{T}}_n$  (arbres de Cayley), on associe naturellement à un élément  $\check{\mathbf{t}} \in \check{\mathbf{T}}_n$  un arbre réel  $\tilde{\mathbf{t}} \in \mathbb{T}$ , par exemple *via* une construction inspirée de (1.7), en faisant de chaque arête un segment de longueur 1. Cet arbre réel est naturellement enraciné.

La famille de cladogrammes aléatoires que nous considérons est la suivante. On suppose que pour tout  $n \geq 2$ , on a une loi de probabilité  $(q_n(\lambda), \lambda \vdash n)$  sur les partitions  $\lambda = (\lambda_1, \dots, \lambda_r)$  de l'entier  $n$ , non-triviales ( $r \geq 2$ ), c'est-à-dire les suites d'entiers strictement positifs décroissantes, de somme  $n$ . On notera  $m_i, 1 \leq i \leq n$  le nombre de parts de  $\lambda$  égales à  $i$ . Pour chaque  $n$ , on construit un processus de fragmentation discret de la façon suivante. On part de  $n$  boules numérotées de 1 à  $n$ . On choisit une partition


 FIG. 4.2: Le cladogramme  $\{[6], \{1, 3, 4, 6\}, \{2, 5\}, \{3, 6\}, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}$ 

$\lambda = (\lambda_1, \dots, \lambda_r) \vdash n$  avec probabilité  $q_n(\lambda)$ , et conditionnellement à cette dernière, on prend une partition de  $[n]$  uniformément parmi les

$$C_\lambda = \frac{n!}{\prod_{i=1}^r \lambda_i! \prod_{i=1}^n m_i!}$$

partitions dont les tailles des blocs donnent  $\lambda$ . On recommence une telle opération indépendamment avec chaque groupe, en remplaçant les partitions de  $[n]$  par des partitions de chacun des groupes. La procédure termine lorsqu'on a isolé les boules. Cette procédure détermine une suite aléatoire croissante de partitions de  $[n]$  issue de  $[n]$  et finissant en  $\{\{1\}, \dots, \{n\}\}$ , et donc un élément aléatoire  $\check{T}_n \in \check{\mathbf{T}}_n$ .

On demande la propriété de cohérence suivante : si l'on part de  $\check{T}_{n+1}$  et qu'on enlève la feuille d'étiquette  $n+1$  et l'arête attenante (ainsi que le sommet de degré 2 que cette opération a pu éventuellement engendrer), on obtient un nouvel arbre de même loi que  $\check{T}_n$ . Cette propriété de cohérence impose des conditions entre les probabilités  $q_n(\lambda)$ .

Notre premier résultat est de caractériser tous les modèles de cladogrammes markoviens cohérents. Nous montrons en fait que ceux-ci sont en correspondance avec les fragmentations homogènes, c'est-à-dire les fragmentations auto-similaires d'indice d'auto-similarité nul.

**Théorème 4.7** *Soit  $(q_n, n \geq 2)$  une famille de probabilités sur les partitions de  $n$  engendrant une famille cohérente de cladogrammes markoviens. Alors il existe un triplet caractéristique de fragmentation de la forme  $(0, c, \nu)$ , non-trivial ( $c \neq 0$  ou  $\nu \neq 0$ ), tel que pour tout  $n \geq 2$  et  $\lambda = (\lambda_1, \dots, \lambda_r) \vdash n$  à au moins deux parts,  $q_n(\lambda)$  s'écrit*

$$q_n^{(c, \nu)}(\lambda) = \frac{C_\lambda}{Z_n} \left( nc \mathbb{1}_{r=2, \lambda_2=1} + \int_{S^1} \nu(ds) \sum_{l=0}^{m_1} \binom{m_1}{l} \sum_{\substack{i_1, \dots, i_{r-l} \geq 1 \\ \text{distincts}}} s_0^l \prod_{j=1}^{r-l} s_{i_j}^{\lambda_j} \right), \quad (4.3)$$

#### 4 Fragmentations auto-similaires et arbres de fragmentation

où  $Z_n$  est une constante de renormalisation donnée par

$$Z_n = nc + \int_{\mathcal{S}^\downarrow} \nu(ds) \left( 1 - \sum_{i \geq 1} s_i^n \right).$$

Enfin, on a  $q_n^{(c,\nu)} = q_n^{(c',\nu')}$  pour tout  $n \geq 2$  si et seulement s'il existe  $K > 0$  tel que  $(c', \nu') = (Kc, K\nu)$ .

La formule (4.3) se simplifie grandement dans le cas où  $c = 0$  et où  $\nu$  est conservatrice : dans ce cas,

$$q_n^{(0,\nu)}(\lambda_1, \dots, \lambda_r) = \frac{C_\lambda}{Z_n} \int_{\mathcal{S}^\downarrow} \nu(ds) \sum_{\substack{i_1, \dots, i_r \geq 1 \\ \text{distincts}}} \prod_{j=1}^r s_{i_j}^{\lambda_j}.$$

Intuitivement, la formule (4.3) correspond à prendre une suite  $\mathbf{s}$  selon  $\nu$ , puis à considérer la partition de  $n$  induite par le procédé de la *boîte de peinture* de Kingman : on se donne  $n$  variables i.i.d.  $(I_1, \dots, I_n)$  dans  $\mathbb{Z}_+$  de loi  $(s_i, i \geq 0)$ , et on pose  $i \sim j$  si  $I_i = I_j \geq 1$ , définissant une partition de  $[n]$ . On notera que le terme  $s_0$  a un rôle spécial, puisque les termes de  $I_i$  qui sont nuls participent à des singletons de la partition. La suite décroissante des tailles de ses blocs donne la partition de  $n$  voulue, et l'on conditionne cette partition à avoir au moins deux parts. Par ailleurs, le terme faisant intervenir l'érosion correspond à l'isolation d'un singleton à taux  $c$ .

Un second aspect de notre travail concerne plus spécifiquement l'étude des cladogrammes, à travers celle de leurs limites d'échelle. On fait l'hypothèse de base que  $c = 0$  et que la mesure  $\nu$  est *conservatrice*, c'est-à-dire que  $\nu(\{\mathbf{s} : \sum_i s_i < 1\}) = 0$ . On fait maintenant l'hypothèse que  $\nu$  vérifie une propriété de variation régulière, donnée par

$$\nu(\{\mathbf{s} : s_1 \leq 1 - \varepsilon\}) = \varepsilon^{-\gamma_\nu} \ell(\varepsilon^{-1}), \quad 0 < \varepsilon < 1, \quad (4.4)$$

où  $\gamma_\nu \in ]0, 1[$  et  $\ell$  est une fonction à variation lente au voisinage de  $+\infty$ . On suppose enfin l'existence d'un  $\rho > 0$  tel que

$$\int_{\mathcal{S}^\downarrow} \nu(ds) \sum_{i \geq 2} s_i |\ln s_i|^\rho < \infty, \quad (4.5)$$

hypothèse satisfaite par exemple si  $\nu$  est au plus  $n$ -aire pour un  $n \geq 2$ . Le résultat principal de [P14] est :

**Théorème 4.8** *Sous les hypothèses (4.4, 4.5), soit  $\tilde{T}_n$  un cladogramme aléatoire associé à la famille de lois  $q_n^{(0,\nu)}$ , ce pour tout  $n \geq 0$ . On a alors la convergence en loi pour la topologie de  $(\mathbb{T}, d_{\text{GH}})$  (ou pour la topologie pointée (1.9) si l'on considère que les arbres sont enracinés) :*

$$\frac{1}{n^{\gamma_\nu} \ell(n) \Gamma(1 - \gamma_\nu)} \tilde{T}_n \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{T}^{-\gamma_\nu, \nu},$$

où  $\mathcal{T}^{-\gamma_\nu, \nu}$  est l'arbre de fragmentation associé à la fragmentation de triplet caractéristique  $(-\gamma_\nu, 0, \nu)$ .



#### 4 Fragmentations auto-similaires et arbres de fragmentation

Un des aspects notables de ce résultat est que le coefficient d'auto-similarité  $\gamma_\nu$  est, en quelque sorte, « forcé » par la mesure  $\nu$ , et montre que l'arbre  $\mathcal{T}^{-\gamma_\nu, \nu}$  joue un rôle particulier parmi les arbres  $\mathcal{T}^{\beta, \nu}$  avec  $\beta < 0$ . Par contraste, les objets naturels sous-jacents au Théorème 4.7 sont plutôt les fragmentations homogènes.

On peut appliquer ce résultat aux lois  $q_n^{(0, \nu^{(\alpha)})}$  associées aux mesures de dislocation  $\nu^{(\alpha)}$  définies en (1.18, 4.1) et associées à l'arbre brownien et aux arbres stables. Comme on peut le deviner, les arbres  $\mathcal{T}^{-\gamma_\nu, \nu}$  correspondants sont bien l'arbre brownien et l'arbre stable d'indice  $\alpha$ , respectivement. Nous discutons dans [P14] d'autres exemples naturels de lois markoviennes cohérentes sur les cladogrammes, issus des travaux d'Aldous [1] et Ford [44], avec des motivations à la biologie.

La preuve du théorème 4.8 repose sur un couplage naturel entre les cladogrammes  $\check{T}_n$  et l'arbre limite  $\mathcal{T}^{-\gamma_\nu, \nu}$ , un outil essentiel étant des résultats sur le nombre de blocs d'une composition régénérative, issus des travaux de Gnedin, Pitman & Yor [46].

## 5 p-arbres, applications aléatoires et leurs limites continues

### 5.1 Hauteur et largeur des arbres continus inhomogènes [P6]

Dans ce travail avec David Aldous et Jim Pitman, nous nous intéressons à différents processus associés aux arbres continus inhomogènes  $\mathcal{T}^\theta$  définis à la section 1.2.7, afin d'en comprendre plus précisément la structure.

#### 5.1.1 Processus des hauteurs

Notre résultat principal est de déterminer « le » processus des hauteurs de  $\mathcal{T}^\theta$ , au moins pour  $\theta \in \Theta_{\text{fini}}$ , où

$$\Theta_{\text{fini}} = \{\theta \in \Theta : \exists I \geq 0 : \theta_{I+1} = 0\}.$$

Ainsi, seul un nombre fini de points de branchement de  $\mathcal{T}^\theta$  sont d'ordre infini. Rappelons la définition (1.19) de  $\mathfrak{b}^\theta$ , pont à accroissements échangeables, effectuant un saut d'amplitude  $\theta_i$  à l'instant  $U_i$ , et de sa transformée de Vervaat  $e^\theta$ , obtenue par un décalage cyclique à l'instant  $s_*$  où  $\mathfrak{b}^\theta$  atteint son infimum. Nous définissons un processus  $(H_s^\theta, 0 \leq s \leq 1)$  à partir de  $e^\theta$ . Intuitivement, en notant  $t_i = U_i - s_*$  (modulo 1) le temps du saut d'amplitude  $\theta_i$  pour  $e^\theta$ , et  $T_i = \inf\{s \geq t_i : e_s^\theta = e_{t_i-}^\theta\}$ , on remplace sur l'intervalle  $[t_i, T_i]$  le processus  $e^\theta$  par le processus réfléchi  $e_s^\theta - \inf_{t_i \leq u \leq s} e_u^\theta$ . Précisément, on note  $R_i^\theta$  le processus

$$R_i^\theta(s) = \begin{cases} \inf_{t_i \leq u \leq s} e_u^\theta - e_{t_i-}^\theta & \text{si } s \in [t_i, T_i] \\ 0 & \text{sinon} \end{cases},$$

et l'on pose

$$Y^\theta = e^\theta - \sum_{i \geq 1} R_i^\theta.$$

Noter en particulier que que  $Y^0 = e$ , l'excursion brownienne.

**Théorème 5.1** *Soit  $\theta \in \Theta_{\text{fini}}$ . Alors le processus  $H^\theta = 2\theta_0^{-2}Y^\theta$  est le processus des hauteurs de l'arbre continu inhomogène  $\mathcal{T}^\theta$ , c'est-à-dire que  $[\mathcal{T}_{H^\theta}, d_{H^\theta}]$  a même loi que  $\mathcal{T}^\theta$ . On a également que la mesure  $\mu_{H^\theta}$  induite sur  $\mathcal{T}_{H^\theta}$  correspond bien à la mesure  $\mu^\theta$  sur  $\mathcal{T}^\theta$ , définie après (1.13).*

Nous montrons en fait le théorème pour tous les  $\theta \in \Theta$  tels que  $\sum_{i \geq 1} \theta_i < \infty$ , mais la question pour un  $\theta$  général reste ouverte. On notera que dans ces cas, l'arbre  $\mathcal{T}^\theta$  est nécessairement compact, ce qui n'est pas vrai pour une valeur de  $\theta$  quelconque.

### 5.1.2 Processus de largeur

On s'intéresse également au « processus de largeur » de l'arbre  $\mathcal{T}^\theta$ . On note

$$\bar{W}_s^\theta := \mu(\{x \in \mathcal{T}^\theta : \text{ht}(x) \leq s\}), \quad s \geq 0,$$

le processus de largeur cumulatif, et  $(\bar{W}^\theta)^{-1}$  son inverse continu à droite.

**Théorème 5.2** *Pour tout  $\theta \in \Theta$ , la mesure de Stieltjes  $d\bar{W}^\theta$  est absolument continue, et admet une densité  $W^\theta$  par rapport à la mesure de Lebesgue sur  $\mathbb{R}_+$ , appelé processus de largeur. De plus, on a l'identité en loi*

$$(W^\theta((\bar{W}^\theta)^{-1}(s)), 0 \leq s \leq 1) \stackrel{(d)}{=} (\mathfrak{e}_s^\theta, 0 \leq s \leq 1).$$

En d'autres termes,  $W^\theta$  est, à un changement près de type de ceux considérés par Lamperti [60], égal en loi au processus  $\mathfrak{e}^\theta$ . On note en particulier que  $W^\theta$  a des sauts de taille  $\theta_i$ , ce qui correspond à de brusques densifications des couches de l'arbre. Ces densifications correspondent bien sûr aux points de branchement de degré infini, et on peut rapprocher la présente discussion à celle du temps local des points de branchement de l'arbre stable (4.2). La combinaison des deux derniers théorèmes donne un résultat qui ne fait plus intervenir d'arbres, et qui évoque un théorème de Jeulin [53] sur le temps local d'une excursion brownienne.

**Corollaire 5.1** *Si  $\theta \in \Theta_{\text{fini}}$ , la mesure d'occupation du processus  $H^\theta$  admet une densité par rapport à la mesure de Lebesgue : pour toute fonction  $f$  positive mesurable,*

$$\int_0^1 f(H_s^\theta) ds = \int_0^\infty f(u) W_u^\theta du,$$

où  $W^\theta$  a même loi que  $(\mathfrak{e}_{\tau(u)}^\theta, u \geq 0)$ , le changement de temps  $\tau$  étant défini par

$$\tau(u) = \inf \left\{ r \geq 0 : \int_0^r \frac{ds}{\mathfrak{e}_s^\theta} > u \right\}, \quad u \geq 0.$$

Le théorème de Jeulin d'origine s'obtient dans le cas où  $\theta_0 = 1$ , où  $\mathfrak{e}^\theta = Y^\theta = \mathfrak{e}$ , et  $W^\theta$  est le temps local de  $2\mathfrak{e}$ . On peut un peu abusivement interpréter ce résultat comme un théorème de Ray-Knight conditionné<sup>1</sup>.

Les démonstrations des théorèmes 5.1 et 5.2 utilisent toutes deux une approximation par des processus associés aux  $\mathbf{p}$ -arbres. Ces processus sont proches (mais cependant différents) des fonctions de parking utilisées par Chassaing et Louchard [28], qui codent l'arbre uniforme à  $n$  sommets, et de la file d'attente LIFO utilisée dans [67] pour coder

<sup>1</sup>Dans le cas  $\theta_0 = 1$  on a l'interprétation suivante en termes de processus de branchement continu : le temps local de l'excursion brownienne est un processus de branchement continu quadratique, issu de 0 mais conditionné à avoir une population totale de taille 1

les versions « à variation finie » des arbres de Lévy. On se donne  $\mathbf{p}$  une probabilité sur  $[n]$  avec  $p_i > 0, 1 \leq i \leq n$ . Nous construisons un pont à accroissements échangeables

$$F^{\mathbf{P}}(s) = \sum_{i=1}^n p_i (\mathbb{1}_{\{U_i \leq s\}} - s), \quad 0 \leq s \leq 1,$$

où  $U_1, \dots, U_n$  sont indépendantes uniformes sur  $[0, 1]$ . Alors il existe un  $s_* \in [0, 1]$ , unique presque-sûrement, tel que  $F^{\mathbf{P}}(s_* -) = \inf_{s \in [0, 1]} F^{\mathbf{P}}(s)$ , de plus  $F^{\mathbf{P}}$  est discontinue en  $s_*$ . On note donc  $F^{\text{exc}, \mathbf{P}}(s) = F^{\mathbf{P}}(s + s_*) - F^{\mathbf{P}}(s_* -), 0 \leq s \leq 1$  la transformée de Vervaat de  $F^{\mathbf{P}}$ . On note  $V_i = U_i - s_*$  modulo 1. à partir de  $F^{\text{exc}, \mathbf{P}}$ , on construit de deux façons différentes un arbre  $\mathcal{T}^{\mathbf{P}}$  dont la loi est celle du **p**-arbre, ces deux constructions correspondant respectivement à un parcours en largeur ou en profondeur de l'arbre. Dans le cas du processus de hauteur, correspondant à un parcours en profondeur, la construction rappelle fortement celle de [67].

Sous les hypothèses de convergence du **p**-arbre vers l'arbre  $\mathcal{T}^{\boldsymbol{\theta}}$  (1.14), on montre la convergence en loi dans l'espace de Skorokhod

$$(\sigma(\mathbf{p}_n)^{-1} F^{\text{exc}, \mathbf{p}_n}(s), 0 \leq s \leq 1) \xrightarrow{d} (\mathfrak{e}_s^{\boldsymbol{\theta}}, 0 \leq s \leq 1). \quad (5.1)$$

Cette convergence sert alors de base à la construction des processus de largeur et de hauteur de l'arbre limite.

### 5.1.3 Convergence forte des **p**-arbres

Au cours de la démonstration, nous prouvons un résultat intéressant sur la convergence du processus de contour des **p**-arbres. Pour chaque  $n \geq 0$  (ou éventuellement le long d'une sous-suite), on se donne une loi de probabilité  $\mathbf{p}_n = (p_{n1}, \dots, p_{nn})$  sur  $[n]$ , avec  $p_{n1} \geq \dots \geq p_{nn} > 0$ , et nous considérons un **p**<sub>n</sub>-arbre  $\hat{T}^n$ . Avec un aléa supplémentaire, nous pouvons construire une version plane de cet arbre, appelée  $T^n$ , obtenue intuitivement en mettant les ensembles d'enfants des différents sommets en ordre échangeable, sachant  $\hat{T}^n$ . À cet arbre plan, on associe sa fonction de hauteur  $H^{T^n}$  comme en 1.1.

Nous introduisons deux hypothèses techniques. Premièrement, nous interdisons les très petites valeurs dans  $\mathbf{p}_n$ , en demandant que pour tout  $\alpha > 0$ ,

$$p_{nn}^{-1} \underset{n \rightarrow \infty}{=} o(\exp(\alpha/\sigma(\mathbf{p}_n))). \quad (5.2)$$

Deuxièmement, soit  $\boldsymbol{\theta} \in \boldsymbol{\Theta}_{\text{fmi}}$ , et  $I$  tel que  $\theta_I > \theta_{I+1} = 0$ . On demande que les « petites » valeurs  $p_{ni}, I + 1 \leq i \leq n$  soient d'ordre  $\sigma(\mathbf{p}_n)^2$ , de façon bien contrôlée. Pour cela, soit supposons qu'il existe une variable aléatoire  $Q \geq 0$  définie sur un espace  $(\Omega, \mathcal{F}, P)$  telle que, pour tout  $\lambda$  dans un voisinage de 0,

$$\lim_{n \rightarrow \infty} \sum_{i=I+1}^n p_{ni} \exp\left(\frac{\lambda p_{ni}}{\sigma(\mathbf{p}_n)^2}\right) = E[\exp(\lambda Q)] < \infty. \quad (5.3)$$

**Théorème 5.3** *Supposons que  $\mathbf{p}_n$  vérifie (1.14) avec  $\theta_I > 0, \theta_{I+1} = 0$  pour un  $I \geq 0$ . Sous les hypothèses supplémentaires (5.2) et (5.3), on a la convergence en loi suivante pour la topologie de Skorokhod*

$$\left( \sigma(\mathbf{p}_n) H_{[nt]}^{T^n} \right)_{0 \leq t \leq 1} \xrightarrow[n \rightarrow \infty]{(d)} H^\theta.$$

*En particulier, dans ce cas,  $\sigma(\mathbf{p}_n) \tilde{T}^n$  converge en distribution pour la topologie de  $(\mathbb{T}, d_{\text{GH}})$ , vers  $\mathcal{T}^\theta$ .*

Il est à noter que ce théorème n'est pas vrai en toute généralité, c'est-à-dire qu'on peut trouver  $\mathbf{p}_n$  satisfaisant (1.14) avec  $\theta \in \Theta_{\text{fini}}$  mais tels que l'on n'ait pas convergence de  $\sigma(\mathbf{p}_n) H^{T^n}$  vers  $H^\theta$  dans l'espace de Skorokhod (voir [7]), du fait de la présence de trop nombreuses valeurs de  $\mathbf{p}_n$  « minuscules », qui provoquent des empilements et forment des pics, d'aire négligeable devant  $n$ , mais de hauteur non-négligeable devant  $\sigma(\mathbf{p}_n)^{-1}$ . On montre en revanche que l'on a toujours convergence en loi pour une topologie un peu plus forte que la convergence  $L^1$  des processus, considérée dans [7].

## 5.2 Applications aléatoires [P4,P8]

Ces travaux avec Aldous et Pitman mettent en relation les résultats de convergence des  $\mathbf{p}$ -arbres vers les ICRT avec les propriétés asymptotiques quand  $n \rightarrow \infty$  d'une application de  $[n]$  dans lui-même, prise au hasard avec une certaine probabilité. Nous prouvons et étendons ainsi de façon « conceptuelle » un résultat dû à Aldous et Pitman [4], datant de 1994, mais dont la preuve est relativement technique et spécifique au cas des applications uniformes de  $[n]$  dans lui-même. Pour ces dernières, une abondante littérature existe [58]. En revanche, peu de choses sont connues lorsqu'on relâche cette uniformité. Une motivation pour généraliser les résultats de convergence connus aux  $\mathbf{p}$ -applications décrites ci-dessous provient de l'article de O'Cinneide et Pokrovski [74]. Outre le fait que nous obtenons de nouvelles preuves de résultats connus en inscrivant les applications aléatoires dans un cadre bien connu d'arbres aléatoires, cette approche est très robuste, et notre travail est une contribution à de telles généralisations.

À toute application  $m : [n] \rightarrow [n]$  est associé son graphe orienté, dont l'ensemble des sommets est  $[n]$  et dont les arêtes pointent de  $i$  vers  $m(i)$ . Un point  $i$  est dit *cyclique* s'il existe  $k \geq 1$  tel que  $m^k(i) = i$ , où  $m^k$  désigne la  $k$ -ième itérée de  $m$ . Soit  $\mathcal{C}(m)$  l'ensemble des points cycliques de  $m$ , notons que pour tout  $i$ ,  $m^k(i) \in \mathcal{C}(m)$  pour  $k$  assez grand. Si l'on efface les arêtes reliant les points cycliques entre eux, le graphe obtenu est une forêt aux sommets étiquetés par  $[n]$ , dont les arbres (composantes connexes) sont naturellement enracinés en aux points cycliques. Voir la figure 5.1 pour un exemple.

On définit de plus une relation d'équivalence par  $i \sim j$  si et seulement s'il existe  $k, k'$  tels que  $m^k(i) = m^{k'}(j)$ . Les classes associées à cette relation sont appelées *bassins d'attraction* de  $m$ . Si on ordonne ces bassins sous la forme  $\mathcal{B}_1(m), \dots, \mathcal{B}_k(m)$  selon une convention, on découpe alors  $\mathcal{C}(m)$  en cycles disjoints  $\mathcal{C}_i(m) = \mathcal{B}_i(m) \cap \mathcal{C}(m)$ . On va s'intéresser aux propriétés d'applications aléatoires de  $[n]$  dans  $[n]$  avec  $n$  grand, les

quantités d'intérêt étant la taille des bassins, des cycles, diamètre  $\sup_{j \in [n]} \inf\{k \geq 1 : m^k(j) \in \{j, m(j), \dots, m^{k-1}(j)\}\}$ .

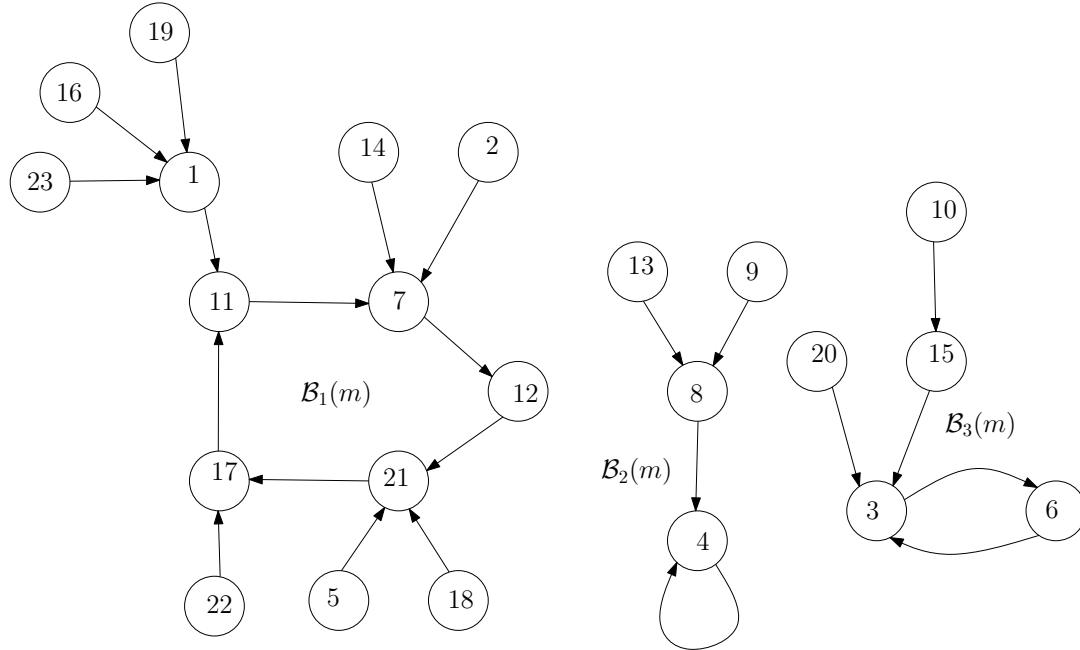


FIG. 5.1: Le graphe orienté d'une application de [23] sur lui-même

Nous nous intéressons au modèle aléatoire suivant. Soit  $\mathbf{p}$  une loi de probabilités sur  $[n]$  avec  $p_i > 0$  pour  $1 \leq i \leq n$ . Soit  $X_1, \dots, X_n$  une suite i.i.d. de loi  $\mathbf{p}$ . On définit  $M^{\mathbf{p}} : [n] \rightarrow [n]$  par  $M^{\mathbf{p}}(i) = X_i$  pour  $1 \leq i \leq n$ . On appelle  $M^{\mathbf{p}}$  la  $\mathbf{p}$ -application. Lorsque  $\mathbf{p} = (1/n, \dots, 1/n)$  est la probabilité uniforme,  $M^{\mathbf{p}}$  est une application uniforme de  $[n]$  dans  $[n]$  parmi les  $n^n$  possibles.

L'argument central repose sur des variantes d'une bijection due à Joyal [54] entre les applications de  $[n]$  dans  $[n]$  et les *vertébrés* sur  $[n]$ . Par définition, un *vertébré* est un arbre de Cayley deux fois pointé, c'est-à-dire un élément de  $\hat{\mathbf{T}}_n \times [n]$ . L'idée de base est simple : le double pointage détermine un unique chemin dans le vertébré, appelé *colonne vertébrale*. On efface alors les arêtes de la colonne vertébrale, que l'on remplace par le graphe orienté d'une permutation des sommets de la colonne vertébrale. On obtient ainsi le graphe d'une application dont les points cycliques sont précisément les sommets de la colonne vertébrale du vertébré initial.

L'avantage de cette bijection est son extrême souplesse dans le choix des permutations utilisées pour lier les sommets de la colonne vertébrale. Introduisons quelques notations. Si  $\hat{\mathbf{t}} \in \hat{\mathbf{T}}_n$ , on notera  $x_0$  sa racine. Si  $x \in [n]$ , on note  $\text{ht}(x)$  sa distance de graphe dans  $\hat{\mathbf{t}}$  à  $x_0$ , où la référence à l'arbre  $\hat{\mathbf{t}}$  devrait être non-ambigüe en fonction du contexte. Par ailleurs, on choisit d'orienter les arêtes de  $\hat{\mathbf{t}}$  de sorte qu'elles pointent vers  $x_0$ . Ainsi,  $\hat{\mathbf{t}}$  peut être vu comme le graphe orienté de l'application  $[n] \setminus \{x_0\} \rightarrow [n]$ , qui à  $x$  on associe son

père dans  $\hat{\mathbf{t}}$ . On note  $\hat{\mathbf{t}}(x)$  l'image de  $x$ . On étend cette application en  $\hat{\mathbf{t}} : [n] \rightarrow [n] \sqcup \{*\}$ , où  $*$  est un point supplémentaire, par  $\hat{\mathbf{t}}(x_0) = *$ .

Enfin, si  $(\hat{\mathbf{t}}, \bar{x}_0)$  est un vertexé sur  $[n]$  et  $x \in [n]$ , on note  $\text{Col}_{\hat{\mathbf{t}}, \bar{x}_0}$  la colonne vertébrale, et  $\text{vert}(x)$  (la mention de  $\hat{\mathbf{t}}, \bar{x}_0$  est implicite) la première des itérées successives de  $x$  par  $\hat{\mathbf{t}}$  qui appartient à  $\text{Col}_{\hat{\mathbf{t}}, \bar{x}_0}$ .

On se donne alors  $\mathbf{x} = (x_1, x_2, \dots)$  une suite à valeurs dans  $[n]$ , telle que  $\{x_i, i \geq 1\} = [n]$ . Soit  $(\hat{\mathbf{t}}, \bar{x}_0)$  un vertexé. On note  $1 = r_1 < \dots < r_k$  les instants de records stricts de la suite

$$\text{ht}(\text{vert}(x_i)), i \geq 1.$$

Noter que nécessairement, on a  $k \leq \text{ht}(\bar{x}_0)$  puisque  $\bar{x}_0$  est le sommet le plus haut de la colonne vertébrale, par définition. Notons  $y_i = \text{vert}(x_{r_i}), 1 \leq i \leq k$ , et par convention,  $y_0 = *$ . Puisque l'on a supposé que la suite  $(x_i, i \geq 1)$  épuise  $[n]$ , on a forcément que  $y_k = \bar{x}_0$ . On définit alors  $\mathcal{J}_{\mathbf{x}}(\hat{\mathbf{t}}, \bar{x}_0) = m \in [n]^{[n]}$  de la façon suivante :

$$\begin{aligned} m(x) &= \hat{\mathbf{t}}(x), & \text{si } x \notin \text{Col}_{\hat{\mathbf{t}}, \bar{x}_0} \text{ ou } \hat{\mathbf{t}}(x) \notin \{y_i, 0 \leq i \leq k-1\} \\ m(x) &= y_{i+1}, & \text{si } x \in \text{Col}_{\hat{\mathbf{t}}, \bar{x}_0} \text{ et } \hat{\mathbf{t}}(x) = y_i, \quad 0 \leq i \leq k-1. \end{aligned}$$

Intuitivement, le graphe de  $m$  s'obtient à partir de  $(\hat{\mathbf{t}}, \bar{x}_0)$  en enlevant l'arête de la colonne vertébrale qui pointe vers le sommet  $y_i, 0 \leq i \leq k-1$ , et en la remplaçant par une arête vers  $y_{i+1}$ . L'application  $m$  ainsi définie a  $k$  bassins d'attraction, qui sont naturellement ordonnés en  $\mathcal{B}_1(m), \dots, \mathcal{B}_k(m)$ , par hauteur croissante des sommets de la colonne vertébrale de  $(\hat{\mathbf{t}}, \bar{x}_0)$  correspondant aux points cycliques des bassins.

Voir la figure 5.2 pour un exemple concret. On se convainc facilement que  $\mathcal{J}_{\mathbf{x}}$  est une bijection entre  $\hat{\mathbf{T}}_n \times [n]$  et  $[n]^{[n]}$ , pour chaque choix de  $\mathbf{x}$ .

La propriété-clef est la suivante :

**Proposition 5.1** *Soit  $\hat{T}^{\mathbf{p}}$  un  $\mathbf{p}$ -arbre (section 1.2.6), et soit  $\bar{X}_0$  un élément de  $[n]$  de loi  $\mathbf{p}$ , indépendant de  $\hat{T}^{\mathbf{p}}$ . Alors pour tout  $\mathbf{x} = (x_1, \dots)$  qui épuise  $[n]$ , l'application aléatoire  $\mathcal{J}_{\mathbf{x}}(\hat{T}^{\mathbf{p}}, \bar{X}_0)$  a même loi que  $M^{\mathbf{p}}$ .*

L'idée est alors d'appliquer cette transformation en prenant  $\mathbf{x} = \mathbf{X}$  elle-même aléatoire, indépendante de  $(\hat{T}^{\mathbf{p}}, \bar{X}_0)$  : elle sera alors indépendante de  $\mathcal{J}_{\mathbf{X}}(\hat{T}^{\mathbf{p}}, \bar{X}_0)$  par la proposition précédente, d'une façon un peu surprenante au premier abord. On supposera que  $\mathbf{X}$  est i.i.d. de loi  $\mathbf{p}'$  sur  $[n]$ , qui charge tous les points. Avec probabilité 1,  $\mathbf{X}$  épuise alors les points de  $[n]$ . L'ordre  $\mathcal{B}_1(M^{\mathbf{p}}), \dots, \mathcal{B}_k(M^{\mathbf{p}})$  induit sur les bassins de la  $\mathbf{p}$ -application  $M^{\mathbf{p}} = \mathcal{J}_{\mathbf{X}}(\hat{T}^{\mathbf{p}}, \bar{X}_0)$  est appelé *l'ordre  $\mathbf{p}'$ -biaisé*, il correspond à l'ordre de découverte de ces bassins par la suite  $\mathbf{X}$ , indépendante de  $M^{\mathbf{p}}$ . Par exemple, si  $\mathbf{p}'$  est uniforme, il s'agit d'un biais par la taille des bassins.

Si l'on prend  $\mathbf{p}_n$  satisfaisant (1.14), on va alors pouvoir traduire les propriétés asymptotiques des  $\mathbf{p}_n$ -applications en termes des arbres continus inhomogènes  $\mathcal{T}^{\theta}$ , et d'une suite de points choisis au hasard selon  $\mu^{\theta}$ . Intuitivement, on peut appliquer essentiellement la même construction que celle de  $\mathcal{J}_{\mathbf{x}}$  dans le cadre continu, donnant lieu à une « application aléatoire continue ». Nous nous contenterons de citer quelques exemples. Le premier est donné par la proposition suivante, qui est un corollaire immédiat de la construction et de la convergence des  $\mathbf{p}$ -arbres exposée en 1.2.8 :

**Proposition 5.2** *Soit  $M_n$  une  $\mathbf{p}_n$ -application, où  $\mathbf{p}_n$  vérifie le régime (1.14). Soit  $X^n$  un élément de loi  $\mathbf{p}_n$  indépendant de  $M_n$ . On note  $c(X^n) = \#\{M_n^i(X^n) : i \geq 0\}$  le nombre d'images successives distinctes de  $X^n$ . Alors*

$$\sigma(\mathbf{p}_n)c(X^n) \xrightarrow[n \rightarrow \infty]{(d)} \text{ht}(x),$$

où  $x$  est un point aléatoire de  $\mathcal{T}^\theta$  choisi selon  $\mu^\theta$ .

On remarquera que cette dernière variable aléatoire a une loi explicite, donnée par Camarri & Pitman [25, Theorem 4] et que l'on peut déduire de la construction Poissonnienne de  $\mathcal{T}^\theta$  :

$$P(\text{ht}(x) \geq r) = \exp\left(-\frac{\theta_0^2}{2}r^2\right) \prod_{i \geq 1} (1 + \theta_i r) e^{-\theta_i r}.$$

Nous donnons un second résultat, portant sur la structure de tous les bassins de  $M_n$  et de leurs points cycliques. Celui-ci repose sur un raffinement de la convergence 1.2.8 des  $\mathbf{p}$ -arbres [25, Corollary 15]. Soit donc  $\bar{x}_0, x_1, x_2, \dots$  une suite échangeable de points de  $\mathcal{T}^\theta$ , de mesure directrice  $\mu^\theta$ . On note  $x_0$  la racine de  $\mathcal{T}^\theta$ , et  $[[x_0, \bar{x}_0]]$  est la colonne vertébrale de l'arbre réel deux fois pointé  $(\mathcal{T}^\theta, x_0, \bar{x}_0)$ . On note  $\text{vert}(x_i)$  le point de  $[[x_0, \bar{x}_0]]$  de hauteur maximale qui appartient à  $[[x_0, x_i]]$ . Soit  $1 = r_1 < r_2 < \dots$  les instants de record stricts de la suite  $(\text{ht}(\text{vert}(x_i)), i \geq 1)$ . On note  $y_i = \text{vert}(x_{r_i})$ , et  $y_0 = x_0$  par convention. Enfin, pour  $i \geq 1$ , notons

$$\mathcal{T}_i^\theta = \{x \in \mathcal{T}^\theta : \text{vert}(x) \in [[y_{i-1}, y_i]] \setminus \{y_{i-1}\}\}.$$

**Théorème 5.4** *On suppose que  $\mathbf{p}_n$  satisfait (1.14) avec  $\theta \in \Theta_{\text{fini}}$ . Soit  $M_n$  une  $\mathbf{p}_n$ -application dont les bassins  $\mathcal{B}_1(M_n), \dots, \mathcal{B}_{k_n}(M_n)$  sont classés par ordre  $\mathbf{p}_n$ -biaisé. Alors, on a la convergence en loi :*

$$(\mathbf{p}_n(\mathcal{B}_i(M_n)), \sigma(\mathbf{p}_n)\#\mathcal{C}_i(M_n), i \geq 1) \xrightarrow[n \rightarrow \infty]{(d)} \left(\mu^\theta(\mathcal{T}_i^\theta), \text{ht}(y_i) - \text{ht}(y_{i-1}) \mid i \geq 1\right).$$

Nous obtenons ces résultats à partir d'un codage des applications aléatoires par des chemins marqués, ces derniers convergeant dans le cas uniforme vers un pont brownien réfléchi [4] marqué. Nous généralisons ces résultats aux  $\mathbf{p}$ -applications et obtenons des limites plus générales que le pont brownien, faisant intervenir le processus des hauteurs  $H^\theta$  construit en 5.1 dans le cas  $\theta \in \Theta_{\text{fini}}$ . L'idée est que les bijections de Joyal se traduisent en des transformations trajectorielles permettant d'obtenir la marche associée à l'application à partir de la fonction des hauteurs du  $\mathbf{p}$ -arbre qui la code, la difficulté étant de voir que ces transformations « passent à la limite ». Cette approche permet également une grande souplesse dans le choix de l'ordre des bassins, qui n'est pas obligé d'être l'ordre  $\mathbf{p}_n$ -biaisé. De même, on peut considérer des limites de poids des bassins  $\mathbf{p}'_n(\mathcal{B}_i(M_n))$  pour des probabilités éventuellement différentes de  $\mathbf{p}_n$ .

D'un autre côté, cette approche impose  $\theta \in \Theta_{\text{fini}}$ . Une approche directement basée sur les arbres permet également d'obtenir le théorème précédent sans l'hypothèse  $\theta \in \Theta_{\text{fini}}$ .



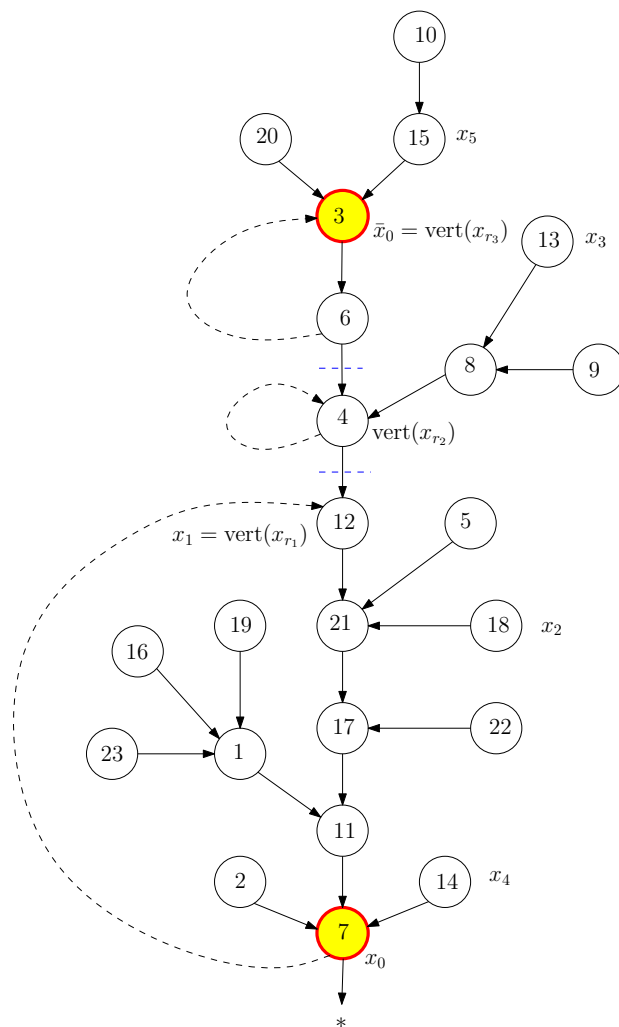


FIG. 5.2: Un vertébré et une suite  $x_1, x_2, \dots$  donnant l'application de la figure 5.1 par  $\mathcal{J}_x$  (suggérée en pointillés). On a ajouté une arête de  $x_0$  à  $*$ , qui n'est pas une arête de l'arbre. Ici,  $r_1 = 1, r_2 = 3, r_3 = 5$ , et il y a trois bassins d'attraction

## 6 Coalescence additive et parking de Knuth

### 6.1 Coalescents ordonnés associés aux processus de Lévy [P1]

Les résultats rappelés dans l'introduction (section 1.5) donnent une classification de tous les coalescents additifs éternels. Néanmoins, les situations où l'on peut effectuer des calculs de semigroupe sont rares. Dans notre premier travail, nous avons exhibé de telles situations. On note  $\mathbb{D}$  l'espace des fonctions  $(X_t, 0 \leq t \leq \zeta)$  càdlàg définies sur un intervalle fermé  $[0, \zeta]$ , où  $\zeta \in [0, \infty]$  est l'instant de mort de la trajectoire. Par convention, on prolonge les trajectoires par leur valeur terminale  $X_\zeta$  au-delà de  $\zeta$ , et on munit  $\mathbb{D}$  de la distance  $d(X, X') = d_{\mathbb{D}}(X, X') + |\zeta(X) - \zeta(X')|$ , où  $d_{\mathbb{D}}$  est la distance de Skorokhod sur les fonctions càdlàg de  $\mathbb{R}_+$  dans  $\mathbb{R}$ .

Soit  $X$  le processus canonique sur  $\mathbb{D}$ , et  $(\mathcal{F}_t, t \geq 0)$  la filtration naturelle. On note  $P$  la loi d'un processus de Lévy sans sauts négatifs (les conventions sont inversées par rapport à [P1]), de variation infinie, et ne dérivant pas vers  $+\infty$ . Les lois de tels processus sont déterminées par trois paramètres  $(\mathfrak{d}, q, \pi)$ , où  $\mathfrak{d}, q \geq 0$  et  $\pi$  est une mesure positive  $\sigma$ -finie sur  $\mathbb{R}_+$  ne chargeant pas  $\{0\}$ , telle que

$$\int_{\mathbb{R}_+} (x \wedge x^2) \pi(dx) < \infty,$$

et vérifiant de plus

$$q > 0 \quad \text{ou} \quad \int_{\mathbb{R}_+} x \pi(dx) = \infty.$$

La loi du processus de Lévy correspondant est caractérisée sa loi marginale de dimension 1, donnée par la formule de Lévy-Khintchine :

$$E[\exp(-\lambda X_1)] = \exp\left(-\mathfrak{d}\lambda + \frac{q}{2}\lambda^2 + \int_{\mathbb{R}_+} (e^{-\lambda x} - 1 + \lambda x)\pi(dx)\right), \quad \lambda \geq 0.$$

On fait l'hypothèse de base que la loi de  $X_t$  sous  $P$  admet pour tout  $t > 0$  une densité  $(p_t(x), x \in \mathbb{R})$  par rapport à la mesure de Lebesgue, bi-continue en  $x$  et  $t$ .

Nous définissons un processus de fragmentation, dans le même esprit qu'en 1.5. On note  $X_s^{(t)} = X_s - ts, 0 \leq s \leq \zeta$ , et  $\underline{X}^{(t)}$  son processus de l'infimum passé. Soit  $F(t)$  la suite décroissante des longueurs des composantes connexes du complémentaire de la mesure  $d\underline{X}^{(t)}$ . On définit ainsi une suite de somme au plus  $\zeta$ .

**Proposition 6.1** *On note  $P^{(-1, \infty)}$  la loi du processus  $X$  sous  $P$ , tué en son premier temps d'atteinte de  $-1$ . Alors sous  $P^{(-1, \infty)}$ , le processus  $F$  est un processus de fragmentation, c'est-à-dire un processus de Markov vérifiant la propriété de branchement suivante.*

## 6 Coalescence additive et parking de Knuth

Il existe des noyaux de transition  $\kappa_{t,t'}(l, ds), 0 \leq t \leq t', l > 0$  respectivement sur  $\{ls : s \in \mathcal{S}^\downarrow\}$ , tels que conditionnellement à  $F(t) = (l_1, l_2, \dots)$ , la loi de  $F(t')$  est celle du réarrangement décroissant de suites indépendantes de lois respectives  $\kappa_{t,t'}(l_i, ds), i \geq 1$ .

On notera que dans cet énoncé, la fragmentation  $F$  sous  $P^{(-1, \infty)}$  est à valeurs dans des suites de somme totale aléatoire. Nous montrons cette propriété selon un plan similaire à l'article de Bertoin [15], qui traite du cas particulier où  $P$  est la loi du mouvement brownien.

Nous décrivons alors les noyaux de transition  $\kappa_{t,t'}$ . Pour ce faire, nous construisons des processus de fragmentation davantage dans l'esprit de la construction de la section 1.5, à partir des excursions de  $X$  au-dessus de son infimum. Nous définissons la loi  $P_0^l$  du pont de longueur  $l$  de  $X$  sous  $P$ . Une façon standard de la définir [43] est de la relier à la loi  $P^l$  du processus tué au temps  $l$  par la formule

$$P_0^l(\Lambda) = E^l \left[ \mathbb{1}_\Lambda \frac{p_{l-t}(-X_t)}{p_l(0)} \right], \quad t < l, \Lambda \in \mathcal{F}_t.$$

On vérifie que ces formules définissent bien une unique mesure sur  $\mathbb{D}$ , portée par les processus de temps de vie  $l$ . Sous  $P_0^l$ , le processus  $X$  est un pont à accroissements échangeables de longueur  $l$ , et rentre donc dans la description donnée par Kallenberg et rappelée en 1.5, avec a modification mineure que l'intervalle de temps est  $[0, l]$  au lieu de  $[0, 1]$ . Les lois  $P_0^l, l > 0$  sont faiblement continues en  $l$ .

On définit alors  $n^{(l)}$  comme la loi de la transformée de Vervaat de  $X$  sous  $P_0^l$ , telle que définie en (1.20) et après une modification triviale pour adapter à l'intervalle de définition  $[0, l]$ . Nous montrons une généralisation du théorème de Vervaat, en adaptant une preuve de Biane [22] pour le mouvement brownien, et Chaumont [31] pour les processus stables. Soit  $n$  la mesure d'Itô des excursions du processus  $X$  réfléchi au-dessus de son infimum  $\underline{X}$ . La mesure  $n$  est une mesure  $\sigma$ -finie sur  $\mathbb{D}$ , portée par les excursions, c'est-à-dire les processus positifs, nuls en leur instant de mort.

**Proposition 6.2** *Les mesures  $n^{(l)}$  donnent une désintégration de  $n$  par rapport à la loi du temps de vie :*

$$n = \int_0^\infty \frac{dl}{l} p_l(0) n^{(l)}.$$

La théorie des excursions montre alors que, conditionnellement à  $F(0) = (l_1, l_2, \dots)$  sous  $P^{(-1, \infty)}$ , la variable aléatoire  $F(t)$  est le réarrangement décroissant de variables aléatoires indépendantes, distribuées respectivement comme  $F(t)$  sous  $n_{l_i}, i \geq 1$ . Ainsi, la description des noyaux  $\kappa_{t,t'}$  revient essentiellement à celle de la loi de  $F(t)$  sous  $n^{(l)}$ .

On note  $\tau^{(t)}$  le processus des temps d'atteinte de niveaux négatifs par  $X^{(t)}$ , c'est-à-dire l'inverse continu à droite de  $-\underline{X}^{(t)}$ . Ceci définit un subordonateur, dont on peut calculer explicitement la mesure de Lévy et le semigroupe en termes des densités  $p_t(x)$  (ceci repose sur une version continue du théorème du ballottage, voir [14, Corollary VII.3]). Nous obtenons alors la description suivante, très proche en esprit de celle du semigroupe de la fragmentation  $F_+^{(2)}$  en termes de sauts d'un subordonateur conditionné.

**Théorème 6.1** *La loi de  $F(t)$  sous  $n^{(l)}$  égale celle de  $\Delta\tau_{[0,t]}^{(t)}$  sous  $P$ , conditionnellement à  $\tau_{tl}^{(t)} = l$ .*

*Par conséquent, le noyau de fragmentation  $\kappa_{t,t'}(l, ds)$  est la loi de  $\Delta\tau_{[0,(t'-t)l]}^{(t')}$  conditionnellement à  $\tau_{(t'-t)l}^{(t)} = l$ .*

Au vu des résultats de 1.5, et après quelques changements mineurs, on voit que le processus  $(F(e^{-t}), -\infty < t < \infty)$  sous  $n^{(l)}$  est un coalescent additif, dont on peut écrire explicitement comme mélange des processus  $\mathbf{C}^\theta$ . Il est surprenant, et tout à fait particulier aux fragmentations ainsi décrites, que les retournés en temps de ces coalescents donnent bien des processus vérifiant la propriété de branchement de la proposition 6.1. On remarquera que ces processus interviennent précisément dans la construction de solutions éternelles à l'équation de Smoluchowski additive [18], qui est l'une des motivations de l'étude du coalescent additif.

Enfin, nous décrivons d'autres aspects spécifiques à la construction de  $F$  sous  $n^{(l)}$ , comme la description de la loi du fragment « le plus à gauche ». Nous montrons qu'il suit une loi d'un échantillon biaisé par la taille, généralisant un résultat de Schweinsberg [81]. Nous nous intéressons plus généralement à l'ordre induit sur les fragments par cette construction, ce qui nous amène à la définition d'un coalescent additif *ordonné*, que l'on peut interpréter comme ajoutant une information spatiale au système d'objets massifs décrit par le coalescent additif.

## 6.2 Caravanes dans le problème de parking de Knuth [P9]

Ce travail, avec Jean Bertoin, trouve sa motivation dans l'article de Chassaing & Louchard [28], dans lequel le coalescent additif apparaît à travers l'étude de la transition de phase dans le modèle de parking de Knuth. Le modèle, et les résultats de [28], peuvent se décrire ainsi. Considérons un parking à  $n$  places, que nous identifions avec le groupe cyclique  $\mathbb{Z}/n\mathbb{Z}$ . Au temps initial, le parking est vide, et l'on fait arriver successivement  $m \leq n$  voitures. La  $i$ -ième voiture tente de se garer en une place  $L_i$  choisie uniformément au hasard parmi les  $n$  places, indépendamment des autres et sans se soucier de savoir si la place choisie est occupée ou non. Si elle est occupée, la voiture essaie successivement les places  $L_i + 1, L_i + 2, \dots$  jusqu'à trouver un emplacement libre. Au fur et à mesure des arrivées, des blocs de voitures consécutives se forment. Ce modèle trouve son origine dans l'étude des algorithmes de hachage, où le parking correspond à un espace mémoire et les voitures sont des éléments de mémoire à stocker.

On montre qu'une transition de phase apparaît au moment où le parking est presque plein. Précisément, le plus grand bloc de voitures consécutives est de taille  $O(\log m)$  avec grande probabilité tant que  $\sqrt{m} = o(n - m)$ , tandis qu'il est de l'ordre de  $n$  (et tous les autres sont de taille  $O(\log n)$ ) avec grande probabilité si  $n - m = o(\sqrt{m})$ . Entre ces deux phénomènes, apparaît un phénomène d'agrégation lorsque  $n - m$  est de l'ordre de  $\lambda\sqrt{m}$ , où  $\lambda > 0$ . Si nous notons  $F_n(\lambda)$  les tailles des blocs de voitures consécutives divisées par

## 6 Coalescence additive et parking de Knuth

$n$  lorsque  $m = n - \lfloor \lambda\sqrt{n} \rfloor$  voitures sont arrivées sur le parking, on a

$$(F_n(\lambda), \lambda \geq 0) \xrightarrow[n \rightarrow \infty]{} (F_+^{(2)}(\lambda), \lambda \geq 0),$$

au sens de la convergence des marginales de dimension finie. En adoptant un retournement de temps exponentiel,  $\lambda = e^{-t}$ , on montre donc que l'agrégation des blocs de voitures suit le processus de coalescence additive standard (voir la section 1.5).

Nous explorons la situation plus générale où l'on tente de stocker à chaque étape des fichiers de taille plus grande, chacun nécessitant un espace de stockage fait de plusieurs unités élémentaires. Nous remplaçons donc les voitures par des *caravanes* de voitures, c'est-à-dire que plusieurs voitures sont autorisées à arriver au même moment au même endroit. Pour ce faire, nous considérons une version continue de ce problème, dont nous déduisons les résultats pour le modèle discret, et sur lequel nous nous concentrons maintenant.

Soit  $\mathbf{p} = (p_1, \dots, p_m)$  une suite de somme 1 à termes strictement positifs, et  $s_1, \dots, s_m$ ,  $m$  points distincts sur le cercle unité.  $\mathbb{U} := \mathbb{R}/\mathbb{Z}$ . Nous imaginons que  $m$  gouttes de peinture, de masses  $p_1, \dots, p_m$ , tombent successivement aux points  $s_1, \dots, s_m$ . À chaque fois qu'une goutte de peinture tombe, nous l'étalons au pinceau dans le sens des aiguilles d'une montre, de sorte que la masse nécessaire à peigner un arc est égale à sa longueur. Si nous rencontrons une partie déjà peinte, nous reportons notre pinceau au premier espace blanc que nous trouvons dans le sens des aiguilles d'une montre, autant de fois que nécessaire. À l'étape  $i$ , nous avons donc recouvert une longueur  $p_1 + \dots + p_i$  du cercle. Pour relier cette description picturale au problème de parking, la  $i$ -ième goutte de peinture prend bien sûr le rôle de la  $i$ -ème caravane, de longueur  $p_i$ , et le point  $s_i$  est l'endroit du parking « continu » où la caravane commence à se garer. Voir la figure 6.1 pour un exemple.

Nous en resterons dans ce mémoire à cette description informelle. À l'étape  $i$ , le cercle unité est couvert par une réunion d'intervalles fermés disjoints  $A_i$ , de mesure de Lebesgue  $\text{Leb}(A_i) = p_1 + \dots + p_i$ . Soit  $\Lambda^{\mathbf{P}}(i) (= \Lambda(p_1, \dots, p_m, s_1, \dots, s_m, i))$  la suite décroissante des mesures de Lebesgue des composantes connexes de  $A_i$ . Nous la complétons par une suite infinie de zéros, obtenant un élément de  $\mathcal{S}^\downarrow$ .

Nous rendons maintenant ce modèle aléatoire, de la façon suivante. Soit  $\ell$  une variable aléatoire de moyenne finie  $\mu_1 = E[\ell]$ . On dit que  $\ell \in \mathcal{D}_2$  si  $\mu_2 := E[\ell^2] < \infty$ , et pour  $\alpha \in ]0, 1[$ , on dit que  $\ell \in \mathcal{D}_\alpha$  si

$$\mathbb{P}(\ell > x) \underset{x \rightarrow \infty}{\sim} cx^{-\alpha} \tag{6.1}$$

pour une constante  $0 < c < \infty$ . Dans ce cas,  $\ell$  est dans le domaine d'attraction d'une variable aléatoire stable d'indice  $\alpha$  spectralement positive. On suppose maintenant que  $\ell \in \mathcal{D}_\alpha$  pour un  $\alpha \in ]1, 2]$ , et on se donne une suite i.i.d.  $\ell_1, \ell_2, \dots$  de variables aléatoires de même loi que  $\ell$ , ainsi qu'une suite, indépendante de cette dernière, de variables i.i.d.  $U_1, U_2, \dots$  uniformes sur  $\mathbb{U}$ . Pour chaque  $\varepsilon > 0$ , on note

$$\tau_\varepsilon = \inf\{i : \ell_1 + \dots + \ell_i \geq 1/\varepsilon\},$$

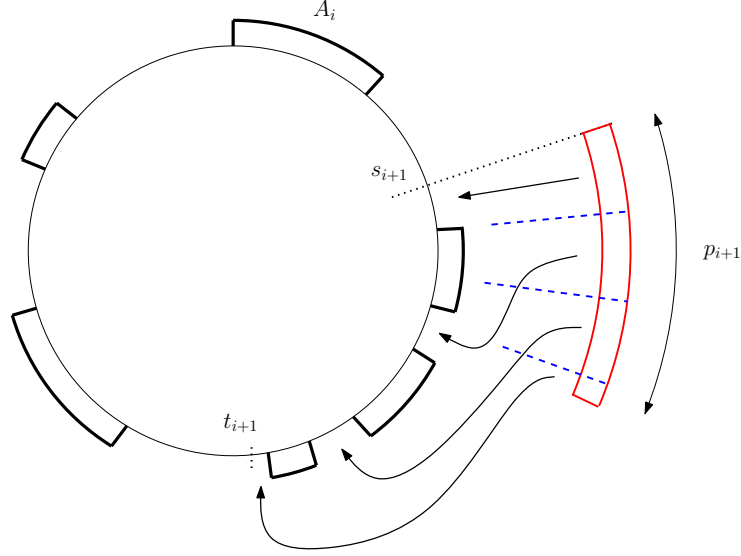


FIG. 6.1: Dans cet exemple, la  $i + 1$ -ième caravane, de longueur  $p_{i+1}$  (dans cette représentation du parking en cercle, cette longueur est angulaire, la circonférence du cercle étant supposée égale à 1), commence à se garer au point  $s_{i+1}$  du cercle unité. Interrompue successivement par trois blocs de  $A_i$ , elle se scinde en 4 morceaux, terminant de se garer en  $t_{i+1}$ .

de sorte que  $\tau_\varepsilon \sim 1/(\varepsilon\mu_1)$ , et on introduit la suite  $(\ell_i^*, 1 \leq i \leq \tau_\varepsilon)$  définie par

$$\ell_i^* = \ell_i \text{ pour } 1 \leq i \leq \tau_\varepsilon - 1 \text{ et } \ell_{\tau_\varepsilon}^* = \varepsilon^{-1} - (\ell_1 + \dots + \ell_{\tau_\varepsilon - 1}),$$

dont la somme des termes est  $1/\varepsilon$ .

En nous inspirant de [28], nous étudions le moment où le parking est presque plein, en posant

$$\mathbf{X}^{(\varepsilon)}(t) = \Lambda^{\mathbf{P}}(\tau_\varepsilon - \lfloor t\varepsilon^{-1/\alpha} \rfloor), \quad t \geq 0,$$

où  $\Lambda^{\mathbf{P}}$  est défini comme ci-dessus à partir des données  $m = \tau_\varepsilon, p_i = \varepsilon\ell_i^*, s_i = U_i$ .

La limite du processus  $\mathbf{X}^{(\varepsilon)}$  va être décrite en termes de processus de coalescence associés à des ponts à accroissements échangeables, comme en 1.5. Le pont obtenu pour  $\alpha \in ]1, 2[$  est un peu atypique, et nous le présentons maintenant. Soit  $(X_t^{(\alpha)}, t \geq 0)$  un processus de Lévy stable spectralement positif d'indice  $\alpha$ , déterminé par sa transformée de Laplace à l'instant 1 :

$$\mathbb{E} \left[ \exp(-\lambda X_1^{(\alpha)}) \right] = \exp(t\lambda^\alpha), \quad \lambda \geq 0.$$

Nous appelons *boucle stable standard* d'indice  $\alpha$ , le processus

$$\mathfrak{b}^{(\alpha)}(x) = X_u^{(\alpha)} - uX_1^{(\alpha)}, \quad 0 \leq x \leq 1. \quad (6.2)$$

Nous utilisons le terme *boucle* plutôt que pont, car  $\mathfrak{b}^{(\alpha)}$  n'est *pas* le pont de longueur 1 du processus de Lévy  $X^{(\alpha)}$ , bien qu'il commence et termine en 0. Si  $X^{(\alpha)}$  était remplacé par le mouvement brownien, nous obtiendrions bien le pont brownien, selon une construction classique, mais spécifique au cas brownien (avec dérive).

Le processus  $\mathfrak{b}^{(\alpha)}$  est néanmoins bien un pont à accroissements échangeables, et nous pouvons considérer la construction de 1.5. On pose donc  $e^{(\alpha)}$  la transformée de Vervaat de  $\mathfrak{b}^{(\alpha)}$ , puis on pose  $e_s^{(\alpha),(t)} = e^{(\alpha)} - ts, 0 \leq s \leq 1$ . On note  $F^{(\alpha)}(t)$  la suite des longueurs des composantes connexes du complémentaire de  $e^{(\alpha),(t)}$ . On notera bien que ce processus n'est pas la fragmentation associée au processus de Lévy  $X^{(\alpha)}$  comme en 6.1, puisque  $\mathfrak{b}^{(\alpha)}$  n'est pas le pont de  $X^{(\alpha)}$  de longueur 1, et  $e^{(\alpha)}$  n'est pas l'excursion de longueur 1 de  $X^{(\alpha)}$  réfléchi au-dessus de son infimum. Enfin, on pose  $\mathbf{C}^{(\alpha)}(t) = F^{(\alpha)}(e^{-t}), t \in \mathbb{R}$ , définissant ainsi un processus de coalescence additive éternel. Notre résultat principal est le suivant :

**Théorème 6.2** *Le processus  $(\mathbf{X}^{(\varepsilon)}(t), 0 \leq t < \tau_\varepsilon)$  converge lorsque  $\varepsilon \downarrow 0$ , au sens de la convergence en loi des marginales de dimension finie, vers un processus  $(\mathbf{X}(t), 0 \leq t < \infty)$ . Le renversé en temps  $(\mathbf{X}(e^{-t}), -\infty < t < \infty)$  est un coalescent additif éternel, et plus précisément :*

(i) *Si  $\alpha = 2$ ,  $(\mathbf{X}(e^{-t}), -\infty < t < \infty)$  a même loi que (rappelons la définition en 1.5)*

$$\mathbf{C}^{(2)} \left( t + \frac{1}{2} \log(\mu_2/\mu_1) - \log \mu_1 \right), \quad -\infty < t < \infty.$$

(ii) *Si  $1 < \alpha < 2$ ,  $(\mathbf{X}(e^{-t}), -\infty < t < \infty)$  a même loi que*

$$\mathbf{C}^{(\alpha)} \left( t + \frac{1}{\alpha} \log \left( \frac{\Gamma(2-\alpha)c}{(\alpha-1)\mu_1} \right) - \log \mu_1 \right), \quad -\infty < t < \infty.$$

On notera bien que, si les marginales de dimension 1 des processus limites ci-dessus sont différentes, leur noyau de transition est toujours le même. On remarquera aussi que lorsque  $\alpha$  augmente, la transition de phase a tendance à avoir lieu à un stade plus tardif dans le processus du parking.

Par ailleurs, nous montrons que ces résultats restent vrais dans un cadre discret, où  $\varepsilon = 1/n$  est l'inverse d'un entier, et où les localisations  $s_i$  sont des multiples de  $1/n$ . Ces résultats discrets sont davantage dans l'esprit initial du problème de Knuth.

# Liste des travaux

Ces articles et prépublications sont tous disponibles sur ma page web :  
<http://www.dma.ens.fr/~miermont>

- [P1] G. MIERMONT, Ordered additive coalescent and fragmentations associated to Lévy processes with no positive jumps, *Electron. J. Probab.*, **6** (2001), pp. no. 14, 33 pp. (electronic).
- [P2] G. MIERMONT AND J. SCHWEINSBERG, Self-similar fragmentations and stable subordinators, in *Séminaire de Probabilités XXXVII*, vol. 1832 of Lecture Notes in Math., Springer, Berlin, 2003, pp. 333–359.
- [P3] G. MIERMONT, Self-similar fragmentations derived from the stable tree. I. Splitting at heights, *Probab. Theory Related Fields*, **127** (2003), pp. 423–454.
- [P4] D. ALDOUS, G. MIERMONT, AND J. PITMAN, Brownian bridge asymptotics for random p-mappings, *Electron. J. Probab.*, **9** (2004), pp. no. 3, 37–56 (electronic).
- [P5] B. HAAS AND G. MIERMONT, The genealogy of self-similar fragmentations with negative index as a continuum random tree, *Electron. J. Probab.*, **9** (2004), pp. no. 4, 57–97 (electronic).
- [P6] D. ALDOUS, G. MIERMONT, AND J. PITMAN, The exploration process of inhomogeneous continuum random trees, and an extension of Jeulin’s local time identity, *Probab. Theory Related Fields*, **129** (2004), pp. 182–218.
- [P7] G. MIERMONT, Self-similar fragmentations derived from the stable tree. II. Splitting at nodes, *Probab. Theory Related Fields*, **131** (2005), pp. 341–375.
- [P8] D. ALDOUS, G. MIERMONT, AND J. PITMAN, Weak convergence of random p-mappings and the exploration process of inhomogeneous continuum random trees, *Probab. Theory Related Fields*, **133** (2005), pp. 1–17.
- [P9] J. BERTOIN AND G. MIERMONT, Asymptotics in Knuth’s parking problem for caravans, *Random Structures Algorithms* **29** (2006), pp. 38–55.
- [P10] G. MIERMONT, An invariance principle for random planar maps, (2006), DMTCS Proc. **AG**, 4th Colloquium on Mathematics and Computer Sciences, Nancy. Erratum disponible sur ma page web.
- [P11] J.-F. MARCKERT AND G. MIERMONT, Invariance principles for random bipartite planar maps, *Ann. Probab.* **35**, n.5, 1642–1705 (2007)
- [P12] G. MIERMONT, M. WEILL, Radius and profile of random planar maps with faces of arbitrary degrees, *Electron. J. Probab.* **13**, 79–106 (2008)



*Liste des travaux*

- [P13]** G. MIERMONT, On the sphericity of scaling limits of random planar maps, *Elect. Comm. Probab.* **13**, 248–257 (2008)
- [P14]** B. HAAS, G. MIERMONT, M. WINKEL, AND J. PITMAN, Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models, *Ann. Probab.* **36**, 1790–1837 (2008)
- [P15]** G. MIERMONT, Invariance principles for spatial multitype Galton-Watson trees, à paraître dans *Ann. Inst. H. Poincaré (B)* (2008)
- [P16]** G. MIERMONT, Tessellations of random maps of arbitrary genus, [arXiv : 0712.3688](#), soumis pour publication (2008)

# Bibliographie

- [1] D. Aldous. Probability distributions on cladograms. In *Random discrete structures (Minneapolis, MN, 1993)*, volume 76 of *IMA Vol. Math. Appl.*, pages 1–18. Springer, New York, 1996.
- [2] D. J. Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1):1–28, 1991.
- [3] D. J. Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993.
- [4] D. J. Aldous and J. Pitman. Brownian bridge asymptotics for random mappings. *Random Structures Algorithms*, 5(4):487–512, 1994.
- [5] D. J. Aldous and J. Pitman. The standard additive coalescent. *Ann. Probab.*, 26(4):1703–1726, 1998.
- [6] D. J. Aldous and J. Pitman. Inhomogeneous continuum random trees and the entrance boundary of the additive coalescent. *Probab. Theory Relat. Fields*, 118(4):455–482, 2000.
- [7] D. J. Aldous and J. Pitman. Invariance principles for non-uniform random mappings and trees. In V. Malyshev and A. Vershik, editors, *Asymptotic Combinatorics with Applications in Mathematical Physics*, pages 113–147. Kluwer Academic Publishers, 2002.
- [8] O. Angel. Growth and percolation on the uniform infinite planar triangulation. *Geom. Funct. Anal.*, 13(5):935–974, 2003.
- [9] O. Angel and O. Schramm. Uniform infinite planar triangulations. *Comm. Math. Phys.*, 241(2-3):191–213, 2003.
- [10] K. B. Athreya and P. E. Ney. *Branching processes*. Springer-Verlag, New York, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.
- [11] E. G. Begle. Regular convergence. *Duke Math. J.*, 11:441–450, 1944.
- [12] E. A. Bender and E. R. Canfield. The asymptotic number of rooted maps on a surface. *J. Combin. Theory Ser. A*, 43(2):244–257, 1986.
- [13] J. Berestycki. Ranked fragmentations. *ESAIM Probab. Statist.*, 6:157–175 (electronic), 2002. Disponible sur <http://www.emath.fr/ps/>.
- [14] J. Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.
- [15] J. Bertoin. A fragmentation process connected to Brownian motion. *Probab. Theory Relat. Fields*, 117(2):289–301, 2000.
- [16] J. Bertoin. Eternal additive coalescents and certain bridges with exchangeable increments. *Ann. Probab.*, 29(1):344–360, 2001.

## Bibliographie

- [17] J. Bertoin. Homogeneous fragmentation processes. *Probab. Theory Relat. Fields*, 121(3):301–318, 2001.
- [18] J. Bertoin. Eternal solutions to Smoluchowski’s coagulation equation with additive kernel and their probabilistic interpretations. *Ann. Appl. Probab.*, 12(2):547–564, 2002.
- [19] J. Bertoin. Self-similar fragmentations. *Ann. Inst. Henri Poincaré Probab. Stat.*, 38(3):319–340, 2002.
- [20] J. Bertoin. *Random fragmentation and coagulation processes*, volume 102 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.
- [21] J. Bertoin. The structure of the allelic partition of the total population for Galton-Watson processes with neutral mutations. À paraître aux *Ann. Probab.*, 2008. [arxiv:07113852](https://arxiv.org/abs/07113852).
- [22] P. Biane. Relations entre pont et excursion du mouvement brownien réel. *Ann. Inst. Henri Poincaré Probab. Stat.*, 22(1):1–7, 1986.
- [23] J. Bouttier, P. Di Francesco, and E. Guitter. Planar maps as labeled mobiles. *Electron. J. Combin.*, 11:Research Paper 69, 27 pp. (electronic), 2004.
- [24] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.
- [25] M. Camarri and J. Pitman. Limit distributions and random trees derived from the birthday problem with unequal probabilities. *Electron. J. Probab.*, 5:no. 1, 18 pp. (electronic), 2000.
- [26] G. Chapuy, M. Marcus, and G. Schaeffer. A bijection for rooted maps on orientable surfaces. 2007. [arXiv:0712.3649](https://arxiv.org/abs/0712.3649).
- [27] P. Chassaing and B. Durhuus. Local limit of labeled trees and expected volume growth in a random quadrangulation. *Ann. Probab.*, 34(3):879–917, 2006.
- [28] P. Chassaing and G. Louchard. Phase transition for parking blocks, Brownian excursion and coalescence. *Random Structures Algorithms*, 21(1):76–119, 2002.
- [29] P. Chassaing and G. Schaeffer. Random planar lattices and integrated superBrownian excursion. *Probab. Theory Related Fields*, 128(2):161–212, 2004.
- [30] L. Chaumont. Sur certains processus de Lévy conditionnés à rester positifs. *Stochastics Stochastics Rep.*, 47(1-2):1–20, 1994.
- [31] L. Chaumont. Excursion normalisée, méandre et pont pour les processus de Lévy stables. *Bull. Sci. Math.*, 121(5):377–403, 1997.
- [32] R. Cori and B. Vauquelin. Planar maps are well labeled trees. *Canad. J. Math.*, 33(5):1023–1042, 1981.
- [33] T. Duquesne. A limit theorem for the contour process of conditioned Galton-Watson trees. *Ann. Probab.*, 31(2):996–1027, 2003.

## Bibliographie

- [34] T. Duquesne. The coding of compact real trees by real valued functions. 2006. arXiv:math/0604106.
- [35] T. Duquesne and J.-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, 281:vi+147, 2002.
- [36] T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields*, 131(4):553–603, 2005.
- [37] T. Duquesne and M. Winkel. Growth of Lévy trees. *Probab. Theory Related Fields*, 139(3-4):313–371, 2007.
- [38] R. T. Durrett and D. L. Iglehart. Functionals of Brownian meander and Brownian excursion. *Ann. Probability*, 5(1):130–135, 1977.
- [39] S. N. Evans. Snakes and spiders: Brownian motion on  $\mathbf{R}$ -trees. *Probab. Theory Related Fields*, 117(3):361–386, 2000.
- [40] S. N. Evans and J. Pitman. Construction of Markovian coalescents. *Ann. Inst. Henri Poincaré Probab. Stat.*, 34(3):339–383, 1998.
- [41] S. N. Evans, J. Pitman, and A. Winter. Rayleigh processes, real trees, and root growth with re-grafting. *Probab. Theory Related Fields*, 134(1):81–126, 2006.
- [42] S. N. Evans and A. Winter. Subtree prune and regraft: a reversible real tree-valued Markov process. *Ann. Probab.*, 34(3):918–961, 2006.
- [43] P. Fitzsimmons, J. Pitman, and M. Yor. Markovian bridges: construction, Palm interpretation, and splicing. In *Seminar on Stochastic Processes, 1992 (Seattle, WA, 1992)*, volume 33 of *Progr. Probab.*, pages 101–134. Birkhäuser Boston, Boston, MA, 1993.
- [44] D. J. Ford. Probabilities on cladograms: introduction to the alpha model. 2005.
- [45] K. Fukaya. Collapsing of Riemannian manifolds and eigenvalues of Laplace operator. *Invent. Math.*, 87(3):517–547, 1987.
- [46] A. Gnedin, J. Pitman, and M. Yor. Asymptotic laws for compositions derived from transformed subordinators. *Ann. Probab.*, 34(2):468–492, 2006.
- [47] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces ( $\Lambda$ -coalescent measure trees). 2006.
- [48] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1999.
- [49] B. Haas. Loss of mass in deterministic and random fragmentations. *Stochastic Process. Appl.*, 106(2):245–277, 2003.
- [50] B. Haas. Regularity of formation of dust in self-similar fragmentations. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(4):411–438, 2004.
- [51] T. E. Harris. *The theory of branching processes*. Dover Phoenix Editions. Dover Publications Inc., Mineola, NY, 2002. Corrected reprint of the 1963 original [Springer, Berlin; MR0163361 (29 #664)].
- [52] S. Janson and J.-F. Marckert. Convergence of discrete snakes. *J. Theor. Probab.*, 18(3):615–645, 2005.

## Bibliographie

- [53] T. Jeulin. Application de la théorie du grossissement à l'étude des temps locaux browniens. In T. Jeulin and M. Yor, editors, *Grossissements de filtrations: exemples et applications*, volume 1118 of *Lecture Notes in Mathematics*, pages vi+315. Springer-Verlag, Berlin, 1985. Papers from the seminar on stochastic calculus held at the Université de Paris VI, Paris, 1982/1983.
- [54] A. Joyal. Une théorie combinatoire des séries formelles. *Adv. in Math.*, 42(1):1–82, 1981.
- [55] O. Kallenberg. Canonical representations and convergence criteria for processes with interchangeable increments. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 27:23–36, 1973.
- [56] J. F. C. Kingman. The representation of partition structures. *J. London Math. Soc. (2)*, 18(2):374–380, 1978.
- [57] F. B. Knight. The uniform law for exchangeable and Lévy process bridges. *Astérisque*, (236):171–188, 1996. Hommage à P. A. Meyer et J. Neveu.
- [58] V. F. Kolchin. *Random mappings*. Translation Series in Mathematics and Engineering. Optimization Software Inc. Publications Division, New York, 1986. Translated from the Russian, With a foreword by S. R. S. Varadhan.
- [59] T. Kurtz, R. Lyons, R. Pemantle, and Y. Peres. A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. In *Classical and modern branching processes (Minneapolis, MN, 1994)*, volume 84 of *IMA Vol. Math. Appl.*, pages 181–185. Springer, New York, 1997.
- [60] J. Lamperti. Continuous state branching processes. *Bull. Amer. Math. Soc.*, 73:382–386, 1967.
- [61] S. K. Lando and A. K. Zvonkin. *Graphs on surfaces and their applications*, volume 141 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004.
- [62] J.-F. Le Gall. The uniform random tree in a Brownian excursion. *Probab. Theory Relat. Fields*, 96(3):369–383, 1993.
- [63] J.-F. Le Gall. *Spatial branching processes, random snakes and partial differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1999.
- [64] J.-F. Le Gall. A conditional limit theorem for tree-indexed random walk. *Stochastic Process. Appl.*, 116(4):539–567, 2006.
- [65] J.-F. Le Gall. The topological structure of scaling limits of large planar maps. *Invent. Math.*, 169(3):621–670, 2007.
- [66] J.-F. Le Gall. Geodesics in large planar maps and in the Brownian map. 2008. [arXiv:0804.3012](https://arxiv.org/abs/0804.3012).
- [67] J.-F. Le Gall and Y. Le Jan. Branching processes in Lévy processes: the exploration process. *Ann. Probab.*, 26(1):213–252, 1998.
- [68] J.-F. Le Gall and F. Paulin. Scaling limits of bipartite planar maps are homeomorphic to the 2-sphere. *Geom. Funct. Anal.*, 18:893–918, 2008.

## Bibliographie

- [69] J.-F. Le Gall and M. Weill. Conditioned Brownian trees. *Ann. Inst. H. Poincaré Probab. Statist.*, 42(4):455–489, 2006.
- [70] J.-F. Marckert. The lineage process in Galton-Watson trees and globally centered discrete snakes. *Ann. Appl. Probab.*, 18(1):209–244, 2008.
- [71] J.-F. Marckert and A. Mokkadem. The depth first processes of Galton-Watson trees converge to the same Brownian excursion. *Ann. Probab.*, 31(3):1655–1678, 2003.
- [72] J.-F. Marckert and A. Mokkadem. Limit of normalized random quadrangulations: the Brownian map. *Ann. Probab.*, 34(6):2144–2202, 2006.
- [73] M. Marcus and G. Schaeffer. Une bijection simple pour les cartes orientables. 2001. Disponible sur [www.lix.polytechnique.fr/Labo/Gilles.Schaeffer/](http://www.lix.polytechnique.fr/Labo/Gilles.Schaeffer/).
- [74] C. A. O’Cinneide and A. V. Pokrovskii. Nonuniform random transformations. *Ann. Appl. Probab.*, 10(4):1151–1181, 2000.
- [75] A. Okounkov and R. Pandharipande. Gromov-Witten theory, Hurwitz numbers, and matrix models, I. 2001. [arXiv:math/0101147](https://arxiv.org/abs/math/0101147)
- [76] M. Perman, J. Pitman, and M. Yor. Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Relat. Fields*, 92(1):21–39, 1992.
- [77] J. Pitman. Coalescent random forests. *J. Combin. Theory Ser. A*, 85(2):165–193, 1999.
- [78] J. Pitman. Poisson-Kingman partitions. In *Statistics and science: a Festschrift for Terry Speed*, volume 40 of *IMS Lecture Notes Monogr. Ser.*, pages 1–34. Inst. Math. Statist., Beachwood, OH, 2003.
- [79] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, third edition, 1999.
- [80] G. Schaeffer. *Conjugaison d’arbres et cartes combinatoires aléatoires*. PhD thesis, Université Bordeaux I, 1998.
- [81] J. Schweinsberg. Applications of the continuous-time ballot theorem to Brownian motion and related processes. *Stochastic Process. Appl.*, 95(1):151–176, 2001.
- [82] W. T. Tutte. A census of planar maps. *Canad. J. Math.*, 15:249–271, 1963.
- [83] V. A. Vatutin and E. E. Dyakonova. The survival probability of a critical multitype Galton-Watson branching process. In *Proceedings of the Seminar on Stability Problems for Stochastic Models, Part II (Nalęczow, 1999)*, volume 106, pages 2752–2759, 2001.
- [84] W. Vervaat. A relation between Brownian bridge and Brownian excursion. *Ann. Probab.*, 7(1):143–149, 1979.
- [85] M. Weill. Asymptotics for rooted planar maps and scaling limits of two-type spatial trees. *Electron. J. Probab.*, 12:Paper no. 31, 862–925 (electronic), 2007.
- [86] A. Zvonkin. Matrix integrals and map enumeration: an accessible introduction. *Math. Comput. Modelling*, 26(8-10):281–304, 1997. *Combinatorics and physics* (Marseille, 1995).