Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2014, Vol. 50, No. 2, 628–654
DOI: 10.1214/12-AIHP518

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Limit distributions for multitype branching processes of m-ary search trees

Brigitte Chauvin[a], Quansheng Liu[b] and Nicolas Pouyanne[a]

[a] *Université de Versailles–St-Quentin, Laboratoire de Mathématiques de Versailles, CNRS UMR 8100, 45 avenue des Etats-Unis, 78035 Versailles CEDEX, France*

[b] *Université de Bretagne-Sud, Laboratoire de Mathématiques de Bretagne Atlantique, CNRS UMR 6205, Campus de Tohannic, BP573, 56017 Vannes, France*

Received 1 December 2011; revised 20 July 2012; accepted 20 July 2012

Abstract. Let $m \geq 3$ be an integer. The so-called *m-ary search tree* is a discrete time Markov chain which is very popular in theoretical computer science, modelling famous algorithms used in searching and sorting. This random process satisfies a well-known phase transition: when $m \leq 26$, the asymptotic behavior of the process is Gaussian, but for $m \geq 27$ it is no longer Gaussian and a limit W^{DT} of a complex-valued martingale arises.

In this paper, we consider the multitype branching process which is the continuous time version of the m-ary search tree. This process satisfies a phase transition of the same kind. In particular, when $m \geq 27$, a limit W of a complex-valued martingale intervenes in its asymptotics. Thanks to the branching property, the law of W satisfies a *smoothing* equation of the type $Z \stackrel{\mathcal{L}}{=} \mathrm{e}^{-\lambda T}(Z^{(1)} + \cdots + Z^{(m)})$, where λ is a particular complex number, $Z^{(k)}$ are independent complex-valued random variables having the same law as Z, T is a $\mathbb{R}_+$-valued random variable independent of the $Z^{(k)}$, and $\stackrel{\mathcal{L}}{=}$ denotes equality in law. This distributional equation is extensively studied by various approaches. The existence and uniqueness of solution of the equation are proved by contraction methods. The fact that the distribution of W is absolutely continuous and that its support is the whole complex plane is shown via Fourier analysis. Finally, the existence of exponential moments of W is obtained by considering W as the limit of a complex Mandelbrot cascade.

Résumé. Soit $m \geq 3$ un entier. Très populaire en informatique fondamentale, l'*arbre m-aire de recherche* est une chaîne de Markov à temps discret qui modélise de célèbres algorithmes de tri et de recherche de données. Ce processus aléatoire vérifie une transition de phase bien connue : lorsque $m \leq 26$, le comportement asymptotique du processus est gaussien. En revanche, lorsque $m \geq 27$, il n'est plus gaussien et fait apparaître la limite W^{DT} d'une martingale à valeurs complexes.

Dans cet article, on considère le processus de branchement multitype qui est le plongement en temps continu de l'arbre m-aire de recherche. Ce processus fait l'objet d'une transition de phase du même type. En particulier, lorsque $m \geq 27$, son asymptotique s'exprime à l'aide de la limite W d'une martingale complexe. Grâce à la propriété de branchement, la loi de W est solution d'une équation en distribution du type $Z \stackrel{\mathcal{L}}{=} \mathrm{e}^{-\lambda T}(Z^{(1)} + \cdots + Z^{(m)})$ où λ est un nombre complexe particulier, les $Z^{(k)}$ sont des variables aléatoires complexes indépendantes dont la loi est celle de Z, T est une variable aléatoire réelle positive indépendante des $Z^{(k)}$, et $\stackrel{\mathcal{L}}{=}$ désigne l'égalité en distribution. On étudie cette équation en loi par des approches variées. L'existence et l'unicité de solutions sont prouvées par des méthodes de contraction. L'absolue continuité de W et le fait que son support soit le plan complexe tout entier sont démontrés par analyse de Fourier. Enfin, on obtient l'existence de moments exponentiels en considérant W comme la limite d'une cascade de Mandelbrot à valeurs complexes.

MSC: Primary 60C05; secondary 60J80; 05D40

Keywords: Martingale; Characteristic function; Embedding in continuous time; Multitype branching process; Smoothing transformation; Absolute continuity; Support; Exponential moments

1. Introduction

Consider a continuous time multitype branching process $(X(t), t \geq 0)$. Types are seen as colors of particles and there are $m-1$ colors, where $m \geq 3$ is an integer. The reproduction of the process is given by a particular matrix R (written in (2.1)), and any particle of colour j lives a random time of exponential distribution with parameter j. Such a classical process is considered for example in Athreya and Ney [1] or Janson [12] and it is precisely defined in Section 2.

When it is stopped at the nth jump time, this process is nothing but the composition vector process $(X_n^{DT}, n \geq 0)$ say, of the so-called m-ary search tree, which is an important algorithmic structure in computer science.

The continuous time random process $(X(t), t \geq 0)$ exhibits a phase transition: when $m \leq 26$, the random vector $X(t)$ has a Gaussian behaviour when t tends to infinity. This fact, a consequence of classical results on branching processes, has been known for a long time. Details are recalled in the beginning of Section 4.

When $m \geq 27$, inspired by the methods used for a two-color Pólya urn in [5], we first prove in Section 4.1 that $X(t)$ admits the following asymptotic expansion:

$$X(t) = \mathrm{e}^t \xi v_1\bigl(1 + \mathrm{o}(1)\bigr) + 2\Re\bigl(\mathrm{e}^{\lambda_2 t} W v_2\bigr)\bigl(1 + \mathrm{o}(1)\bigr) + \mathrm{o}\bigl(\mathrm{e}^{\sigma_2 t}\bigr) \quad \text{a.s.},$$

where λ_2 is a particular complex number having a real part σ_2 in $]\frac{1}{2}, 1[$, ξ is a Gamma distributed random variable and W a $\mathbb{C}$-valued one and v_1, v_2 are linearly independent vectors defined in (2.4).

We are interested in the limit random variables $W_k, k = 1, \dots, m-1$, each corresponding to $X_k(t)$ which denotes the process $X(t)$ when it starts from one particle of color k. Using the branching property, a system of dislocation equations is written for the random vectors $X_k(t)$ in Section 5.1. A system of fixed point equations satisfied by the corresponding limit laws is then derived in Section 5.2. In particular, the complex-valued random variable W_1 is a solution of the fixed point equation

$$Z \stackrel{\mathcal{L}}{=} \mathrm{e}^{-\lambda_2 T}\bigl(Z^{(1)} + \cdots + Z^{(m)}\bigr), \tag{1.1}$$

where $\{Z^{(k)}: k \geq 1\}$ are independent copies of Z, $T = \tau_{(1)} + \cdots + \tau_{(m-1)}$, $\{\tau_{(j)}: j \geq 1\}$ are random variables independent of each other and independent of $\{Z^{(k)}\}$, each $\tau_{(j)}$ has distribution $\mathcal{E}\mathrm{xp}(j)$ (we denote by $\mathcal{E}\mathrm{xp}(j)$ the exponential distribution of parameter j: it has density $x \mapsto j\mathrm{e}^{-jx}$ on $]0, \infty[$).

Further properties of W_1 are derived from a thorough study of Eq. (1.1). We first show in Theorems 6.2 and 6.4 that Eq. (1.1) admits a unique square-integrable solution having a given mean. In particular, this implies that Eq. (1.1) characterizes the distribution of W_1. This result is proven by two contraction methods applied to the corresponding smoothing transformations. The first one deals with suitable spaces of probability measures where the classical Wasserstein metric is adapted to the complex field; it leads to Theorem 6.2. The second contraction method, that gives a proof for Theorem 6.4, consists in working on Fourier transforms of solutions and provides a somewhat simpler proof. Furthermore, this second method gives a result of existence and uniqueness for solutions of the convolution equation

$$\varPhi(t) = \int_0^{+\infty} \varPhi^m(t-u) f_T(u)\,\mathrm{d}u, \quad t \in \mathbb{C},$$

in a convenient space of functions, where f_T denotes the density of T (see Remark 6.6).

Once the characterization of W_1 by Eq. (1.1) is proven, it suffices to derive properties of solutions of this distributional equation. We show in this way the following results on the law of W_1.

Theorem 1.1. *When $m \geq 27$, the complex-valued random variable W_1 admits a density and its support is the whole complex plane. Its Fourier transform satisfies*

$$\mathbb{E}\mathrm{e}^{\mathrm{i}\langle t, W_1\rangle} = \mathrm{O}\bigl(|t|^{-a}\bigr)$$

when $|t| \to +\infty$, for some $a > 1$.

This theorem is a direct consequence of Theorem 7.1 that provides such properties for solutions of (1.1) admitting a nonzero mean. Our proof consists in showing successively that the characteristic function of any solution has modulus equal to 1 only at the origin, that it tends to zero at infinity, and finally that it is of order $\mathrm{O}(|t|^{-a})$ as $|t| \to \infty$ for some $a > 1$, so that it is square-integrable on $\mathbb{C}$. In the approach we need to prove a nonlattice property of Eq. (1.1) *via* Gelfand–Schneider theorem, using the algebraicity of λ_2 (see proof of Lemma 7.4).

Theorem 1.2. *When* $m \ge 27$, *the random variable* W_1 *admits exponential moments in a neighbourhood of the origin of the complex plane. If* $L_1(z) = \mathbb{E}e^{zW_1}$ *denotes its Laplace series, then* L_1 *is holomorphic near* 0 *and, after a change of variable, the function* $z \mapsto -\frac{\rho}{z} L_1(z^{-\lambda_2})$ *is a solution of the differential equation*

$$y^{(m-1)} = y^m.$$

Theorem 1.2 is immediately derived from Theorems 8.1 and 8.4 just as Theorem 1.1 was derived from Theorem 7.1. To prove Theorems 8.1 and 8.4, we consider a solution of (1.1) as the limit of a complex Mandelbrot cascade. The results are a consequence of fine analytical properties of the Fourier transform of the limit variable.

The paper is organized as follows. The continuous time multitype branching process is defined in Section 2. Its relation with the m-ary search tree is detailed in Section 3, while Section 4 is devoted to the asymptotics of $X(t)$ and to its connection to the corresponding discrete time process. In Section 5, we use the branching property of the process to show that the martingale limits of the continuous time process are related by a system of equations in law so that the fixed point equation (1.1) emerges. These first four sections constitute the first part of the paper.

The second part of the paper consists in putting the focus on Eq. (1.1) that turns out to characterize the distribution of W_1 so that all results on solutions provide results on W_1. In Section 6 we define the natural smoothing transform associated with Eq. (1.1) and we show that it defines a contraction in the space of square-integrable probability measures with given mean. Results on the support and on absolute continuity of solutions are obtained in Section 7. Finally, Section 8 is devoted to the exponential moments and the Laplace series of solutions.

2. Definition of the branching process

In this section we introduce the definition of the continuous time multitype branching process $(X(t))$, and present the spectral decomposition of its transition matrix.

2.1. *Infinitesimal generator*

In the whole paper, the underlying vector space is $\mathbb{R}^{m-1}$ or sometimes $\mathbb{C}^{m-1}$. Let R be the following square matrix of order $m-1$:

$$R = \begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & \ddots & \\ & & & \ddots & 1 \\ m & & & & -1 \end{pmatrix}, \tag{2.1}$$

and for $k = 1, \dots, m-1$, let w_k be the kth row vector of R: when $1 \le k \le m-2$, the kth coordinate of w_k equals -1, the $(k+1)$th equals 1 and all the others are 0; w_{m-1} has m as first coordinate, -1 as last one, and 0 for all others.

Let G be the operator defined on functions f from $\mathbb{C}^{m-1}$ to any real or complex vector space by the following formula: for any vector v in $\mathbb{C}^{m-1}$,

$$G(f)(v) = \sum_{k=1}^{m-1} k l_k(v) \big[f(v + w_k) - f(v) \big], \tag{2.2}$$

where l_k are the coordinate forms: $l_k(x_1, \dots, x_{m-1}) = x_k$.

Definition 2.1. *The right-continuous process $X = (X(t), t \geq 0)$ is the only continuous time Markov process with state space $\mathbb{R}^{m-1}$ having G as infinitesimal generator.*

Equivalently, X is a continuous time multitype branching process with $m-1$ types (or colors), having R as reproduction matrix. The kth coordinate of the vector $X(t)$, namely $l_k(X(t))$, is the number of particles of color k at time t. A particle of color k lives a random exponential time with parameter k; when it dies, it reproduces one particle of color $k+1$ if $k = 1, \dots, m-2$, and m particles of color 1 if $k = m-1$.

This branching continuous time process can be thought as the embedded process of a discrete Markov chain $X^{DT} = (X_n^{DT})_{n\in\mathbb{N}}$ which is a Pólya-type discrete Markov chain associated with the node process of an m-ary search tree, an important algorithmic structure in computer science. This connection is detailed in Section 3.

2.2. *Spectral decomposition*

Let R_G be the matrix of G's restriction to linear forms in the canonical basis $(l_k)_{1\leq k\leq m-1}$. One immediately checks that

$$R_G = \begin{pmatrix} -1 & 1 & & & \\ & -2 & 2 & & \\ & & -3 & \ddots & \\ & & & \ddots & m-2 \\ m(m-1) & & & & -(m-1) \end{pmatrix},$$

where an empty entry means a zero entry. It has been established in many papers – see for example Mahmoud [16], Chern and Hwang [6] or [4] – and it can be easily checked that R_G's (unitary) characteristic polynomial is

$$\chi_{R_G}(\lambda) = \prod_{k=1}^{m-1}(\lambda + k) - m! = \frac{\Gamma(\lambda+m)}{\Gamma(\lambda+1)} - m!, \tag{2.3}$$

where Γ denotes Euler's Gamma function. All eigenvalues are simple, 1 being the one having the largest real part. If $m = 3$, the second eigenvalue is -3. When $m \geq 4$, all eigenvalues different from 1 are nonreal, except $-(m+1)$ when m is odd.

When $m \geq 4$, in the whole paper, λ_2 will denote χ_{R_G}'s root having the second largest real part named σ_2 and a positive imaginary part named τ_2.

The famous phase transition on m-ary search trees already mentioned in the introduction is due to the fact that

$$\Re(\lambda_2) > 1/2 \quad \text{if and only if} \quad m \geq 27.$$

See for example [4]. The assumption $\Re(\lambda_2) > 1/2$ will be frequently used in the sequel.

We adopt the following notations:

$$\begin{cases} \forall n \in \mathbb{Z}_{\geq 0}, \quad \dbinom{z}{n} = \dfrac{\Gamma(z+1)}{n!\Gamma(z-n+1)} = \dfrac{z(z-1)\cdots(z-n+1)}{n!}; \\ H_m(z) = \displaystyle\sum_{1\leq k\leq m-1} \frac{1}{z+k}; \\ u_1(x_1,\dots,x_{m-1}) = \displaystyle\sum_{1\leq k\leq m-1} kx_k; \\ u_2(x_1,\dots,x_{m-1}) = \displaystyle\sum_{1\leq k\leq m-1} \binom{\lambda_2+k-1}{k-1} x_k; \\ v_1 = \dfrac{1}{H_m(1)}\left(\dfrac{1}{k(k+1)}\right)_{1\leq k\leq m-1}; \\ v_2 = \dfrac{1}{H_m(\lambda_2)}\left(\dfrac{1}{k\binom{\lambda_2+k}{k}}\right)_{1\leq k\leq m-1}. \end{cases} \tag{2.4}$$

The linear forms u_1 and u_2 are eigenvectors of G, namely $G(u_1) = u_1$ and $G(u_2) = \lambda_2 u_2$. The vectors v_1 and v_2 are left eigenvectors of R_G, respectively associated with the eigenvalues 1 and λ_2. They satisfy $u_1(v_1) = u_2(v_2) = 1$ and $u_1(v_2) = u_2(v_1) = 0$. These spectral data had already been essentially computed in [4] and [18]. Note that, since λ_2 is not real, u_2 and v_2 have nonreal coordinates. Note also that $\overline{v_2}$ (for a complex vector v we denote by $\overline{v}$ its conjugate vector composed of the complex conjugates of the components of v) is an eigenvector of R_G linearly $\mathbb{C}$-independent from v_2, that $(v_1, v_2, \overline{v_2})$ can be completed to provide a basis of complex eigenvectors of R_G, and that its dual basis is of the form $(u_1, u_2, \overline{u_2}, \ldots)$. In particular, $u_2(\overline{v_2}) = 0$.

3. *m*-ary search trees and embedding

In this section we present the connection between m-ary search trees and the multitype branching process defined in Section 2. This connection is the classical embedding of a discrete time Markov chain into a continuous time Markov process.

3.1. *m-ary search trees*

We define here a discrete time Markov chain $X^{DT} = (X_n^{DT}, n \geq 0)$ with values in $\mathbb{N}^{m-1} \setminus \{0\}$. The ith coordinate of X_n^{DT} is denoted by $X_n^{(i)}$ and has a "physical" meaning detailed hereafter. The Markov chain X^{DT} is a random walk defined by an initial vector X_0^{DT} in $\mathbb{N}^{m-1} \setminus \{0\}$ and by the following transition probabilities: $\forall v \in \mathbb{N}^{m-1} \setminus \{0\}$, $\forall k = 1, \ldots, m-1$,

$$q(v, v + w_k) = \frac{k l_k(v)}{\sum_{j=1}^{m-1} j l_j(v)}, \tag{3.1}$$

where the increment vectors w_k are given in Section 2.1 and $l_k(v)$ denotes the kth coordinate of the vector v.

Classically (see Norris [17] and for a synthetic exposition Bertoin [3]), this discrete time Markov chain is embedded in continuous time using a "Poissonization" of the time: given X^{DT}, one can recover $X = (X(t), t \geq 0)$ as follows. At time 0, $X(0) = X_0^{DT}$. For any vector $v \in \mathbb{R}^{m-1}$, define[1]

$$q(v) := \sum_{k=1}^{m-1} k l_k(v).$$

Let τ_1 be a random time exponentially distributed with parameter $q(X_0^{DT})$. For any time $t \in [0, \tau_1[$, let $X(t) = X(0) = X_0^{DT}$. At time τ_1, X jumps from $v = X(0)$ to $v + w_k$ with probability given by formula (3.1). More generally, let $\tau_0 = 0$ and for any $n \geq 1$, define the nth jumping time τ_n by

$$\tau_n = \sum_{i=0}^{n-1} \frac{\epsilon_i}{q(X_i^{DT})},$$

where ϵ_i are independent random variables having the same exponential distribution with parameter 1. Let

$$X(t) = X(\tau_n) = X_n^{DT} \quad \forall t \in [\tau_n, \tau_{n+1}[.$$

At time τ_{n+1}, X jumps from $v = X(\tau_n)$ to $v + w_k$ with probability given by formula (3.1). It is easy to see that this embedded process $X(t)$ is the same one as the branching process defined in Section 2.1.

When $X_0^{DT} = (1, 0, \ldots, 0)$, each $X_n^{(i)}$, $i = 1, \ldots, m-1$, can be seen as the number of nodes of type i in a tree T_n: the sequence $(T_n, n \geq 0)$ is a sequence of random m-ary trees which grow by successive insertions of keys in their leaves. Each node of these trees contains at most $m-1$ keys. Keys are i.i.d. random variables x_i, $i \geq 1$, with any diffusive distribution on the interval $[0, 1]$. The tree T_n, $n \geq 0$, is recursively defined as follows: T_0 is reduced to an empty node-root; T_1 is reduced to a node-root which contains x_1, T_2 is reduced to a node-root which contains x_1 and $x_2, \ldots, T_{m-1}$ has a node-root containing $x_1, \ldots, x_{m-1}$. As soon as the $(m-1)$th key is inserted in the root, m empty

[1] Note that $q = u_1$ where u_1 was defined by (2.4).

subtrees of the root are created, corresponding from left to right to the m ordered intervals $I_1 =]0, x_{(1)}[, \dots, I_m =]x_{(m-1)}, 1[$, where $0 < x_{(1)} < \cdots < x_{(m-1)} < 1$ are the ordered $m-1$ first keys. Each following key $x_m, \dots,$ is recursively inserted in the subtree corresponding to the unique interval I_j to which it belongs. As soon as a node is saturated, m empty subtrees of this node are created.

For each $i = \{1, \dots, m-1\}$ and $n \geq 1$, $X_n^{(i)}$ is the number of nodes in T_n which contain $i-1$ keys (and i gaps or free places) after insertion of the nth key; such nodes are named nodes of type i. We don't worry about the number of saturated nodes. The vector X_n^{DT} is called the composition vector of the m-ary search tree. It provides a model for the space requirement of the algorithm. One can refer to Mahmoud's book [16] for further details on search trees.

Notice that, in this dynamics, the insertion of a new key is *uniform* on the *gaps*, as can be seen on the transition probabilities (3.1).

3.2. *Embedding*

The embedding properties are summarized in the following lemma.

Lemma 3.1.

(1) *For any $n \geq 1$, the distribution of $\tau_n - \tau_{n-1}$ is $\mathcal{E}\mathrm{xp}(n-1+N_0)$, where N_0 is the number of free places in $X(0)$: $N_0 = u_1(X(0))$.*
(2) *the processes $(\tau_n)_{n\geq 1}$ and $(X(\tau_n))_{n\geq 1}$ are independent.*
(3) *the processes $(X(\tau_n))_{n\geq 1}$ and $(X_n^{DT})_{n\geq 1}$ have the same distribution.*

Proof. Part (1) is a consequence of the fact that the minimum of k independent $\mathcal{E}\mathrm{xp}(1)$-distributed random variables is $\mathcal{E}\mathrm{xp}(k)$-distributed, and that the total number of free places at time τ_n equals $n-1+N_0$.

Part (2) is the classical independence between the jump chain and the jump times in such Markov processes. The initial states and evolution rules of both Markov chains in discrete time and in continuous time are the same ones, so that Part (3) holds. □

Convention. *From now on, thanks to Part (3) of Lemma 3.1, we will as usual suppose that the discrete time process and the continuous time process are built on the* same *probability space on which*

$$\big(X(\tau_n)\big)_{n\geq 1} = \big(X_n^{DT}\big)_{n\geq 1} \quad \text{a.s.} \tag{3.2}$$

Remark. *The important benefit we get with the embedding is the independence in the continuous time process. This independence is the key point for the dislocation equations later on.*

4. Asymptotics and martingale connection

In this section we present the asymptotic behaviour of the continuous time multitype branching process $(X(t))_t$ in three principal directions and its connection with the discrete time process (X_n^{DT}) defined in Section 3.1.

4.1. *Asymptotics of the continuous time branching process*

When $m \leq 26$, the random vector $X(t)$ satisfies a Gaussian asymptotics when t tends to infinity: firstly, the random vector $\mathrm{e}^{-t}X(t)$ converges almost surely to ξv_1 where ξ is a positive random variable and v_1 a deterministic vector (in fact, v_1 is defined by formula (2.4) and the proof of Theorem 4.1 shows that ξ is *Gamma*-distributed and that this convergence is valid in any L^p, $p \geq 1$). Secondly, $X(t)$ can be decomposed as the sum $X(t) = X_1(t) + X_I(t)$ of two random vectors, where X_1 is proportional to v_1, X_I belongs to some fixed supplementary subspace, $\mathrm{e}^{-t}X_1(t) \to \xi v_1$ almost surely and $\mathrm{e}^{-t/2}X_I(t)$ converges *in distribution* to $\sqrt{\xi}N$ where N is a centered Gaussian vector independent of ξ. For a statement and a proof of these facts, one can refer to Theorem 3.1 and Example 7.8 in Janson [12].

When $m \geq 27$, using the notations of Section 2 and especially the formula (2.4), the random vector $X(t)$ admits a 3-dimensional almost sure expansion as t goes to infinity, described in Theorem 4.1. Denote by $\Re(v_2)$ (*resp.* $\Im(v_2)$)

the vector of $\mathbb{R}^{m-1}$ made of the real (*resp.* imaginary) parts of v_2's coordinates. Then, geometrically speaking, Theorem 4.1 gives the principal part of $X(t)$'s coordinates along the three linearly independent vectors v_1, $\Re(v_2)$ and $\Im(v_2)$, the projection of $X(t)$ on a supplementary subspace being almost surely negligible.

Notice that these results had been partially proven in a more general frame by Athreya and Ney [1] and adapted to this particular branching process by Janson [12]. More precisely, since R_G is diagonalizable on $\mathbb{C}$, choose a basis $(v_\lambda)_{\lambda\in\mathrm{Sp}(R_G)}$ of complex eigenvectors of R_G and name its dual basis $(u_\lambda)_{\lambda\in\mathrm{Sp}(R_G)}$. Then, the spectral decomposition $\mathbb{C}^{m-1} = \bigoplus \mathbb{C}v_\lambda$ gives rise to the corresponding projections $u_\lambda v_\lambda$ on all eigenlines. With this material, Janson's result can be stated in the following way: for any $\lambda \in \mathrm{Sp}(R_G)$,

(i) if $\Re(\lambda) > 1/2$, $\mathrm{e}^{-\lambda t}u_\lambda(X(t))$ converges almost surely to some (nonnormal) random variable. In particular, let $\xi = \lim_{t\to+\infty} \mathrm{e}^{-t}u_1(X(t))$;
(ii) if $\Re(\lambda) < 1/2$, $\mathrm{e}^{-t/2}u_\lambda(X(t))$ converges in law to some product $\sqrt{\xi}N$ where N is a centered normal distribution independent of ξ (note that $\Re(\lambda) = 1/2$ never happens).

Nevertheless, the global almost sure and L^p remainder $\boldsymbol{\epsilon}_3(t)$ in Theorem 4.1 is a new result.

Theorem 4.1 (Asymptotics of continuous time process). *Suppose that $m \geq 27$. Then, with the notations of Section 2 and especially the formulae (2.4), as t tends to infinity,*

$$X(t) = \mathrm{e}^t\xi v_1\big(1+\varepsilon_1(t)\big) + 2\Re\big(\mathrm{e}^{\lambda_2 t}Wv_2\big)\big(1+\varepsilon_2(t)\big) + \mathrm{e}^{\lambda_2 t}\boldsymbol{\epsilon}_3(t), \tag{4.1}$$

where

- *ξ is a positive* Gamma-*distributed random variable with expectation $N_0 = u_1(X(0))$ (total weighted number of particles at time* 0),
- *W is a complex-valued random variable that admits moments of all orders $p \geq 1$ and whose expectation equals $u_2(X(0))$,*
- *the real-valued random variables $\varepsilon_1(t)$ and $\varepsilon_2(t)$ and the real random* vector *$\boldsymbol{\epsilon}_3(t)$ tend to* 0 *as t tends to $+\infty$, almost surely and in any* L^p*-space, $p \geq 1$.*

In other words, if one denotes by φ any argument of the complex number W, the trajectory of the random vector $X(t)$, projected in the 3-dimensional real vector space spanned by the vectors $(\Re(v_2), \Im(v_2), v_1)$ is almost surely asymptotic to the (random) spiral

$$\begin{cases} x(t) = 2|W|\mathrm{e}^{\sigma_2 t}\cos(\tau_2 t + \varphi), \\ y(t) = -2|W|\mathrm{e}^{\sigma_2 t}\sin(\tau_2 t + \varphi), \\ z(t) = \xi\mathrm{e}^t, \end{cases}$$

drawn on the (random) revolution surface

$$4|W|^2 z^{2\sigma_2} = \xi^{2\sigma_2}\big(x^2+y^2\big),$$

when t tends to infinity. See Fig. 1

In the whole paper, W denotes our hero, namely the limit complex-valued random variable that appears in $X(t)$'s expansion, as in Theorem 4.1.

Proof of Theorem 4.1. Denote by $\mathcal{A}$ the endomorphism of $\mathbb{R}^{m-1}$ having ${}^tR_G$ as matrix in the canonical basis. Let also $M(t) = \exp(-t\mathcal{A})X(t)$, for any $t \geq 0$. By standard arguments from multitype branching process theory, $(M(t))_{t\geq 0}$ is a vector-valued martingale. Since $m \geq 27$, the real part of λ_2 belongs to $]1/2, 1[$ so that the projected martingales $u_1(M(t))$ and $u_2(M(t))$ converge in L^p for any $p \geq 1$. For proofs of these results, see for example Athreya and Ney [1] or Janson [12] (especially Lemma 10.2 of Janson's paper for the L^p-boundedness, X being here an *irreducible* process in the sense of [12]). The random variables ξ and W are respectively defined by

$$\begin{cases} \xi = \lim_{t\to+\infty} u_1\big(\mathrm{e}^{-t\mathcal{A}}X(t)\big) = \lim_{t\to+\infty} \mathrm{e}^{-t}u_1\big(X(t)\big), \\ W = \lim_{t\to+\infty} u_2\big(\mathrm{e}^{-t\mathcal{A}}X(t)\big) = \lim_{t\to+\infty} \mathrm{e}^{-\lambda_2 t}u_2\big(X(t)\big). \end{cases} \tag{4.2}$$

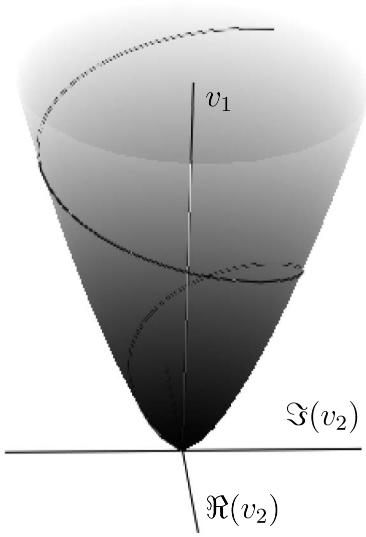

Fig. 1. Spiral in the vector space spanned by $(\Re(v_2), \Im(v_2), v_1)$ to which $X(t)$ is a.s. asymptotic.

An alternative proof of the L^p convergence can be made using the techniques of [19], as developed in [5] for two-colour urn processes. In particular, ξ's distribution is attained by explicit computation of its moments: for any non-negative integer p, an elementary computation shows directly from (2.2) that the (so-called *reduced*) polynomial

$$Q := u_1(u_1+1)(u_1+2)\cdots(u_1+p-1)$$

is an eigenvector for X's infinitesimal generator G, associated with the eigenvalue p. Thus $\mathbb{E}Q(X(t)) = \mathrm{e}^{pt}Q(X(0))$ for any t. Besides, because of (4.2), $Q(X(t)) = \mathrm{e}^{pt}\xi^p(1+\mathrm{o}(1))$ as t tends to infinity, almost surely and in L^1. Finally, the last two equalities provide

$$\mathbb{E}\xi^p = Q\big(X(0)\big) = \frac{\Gamma(N_0+p)}{\Gamma(N_0)}.$$

This shows that the law of ξ is a Gamma distribution with parameter N_0 since a Gamma distribution is completely determined by its moments. The matrix R_G is diagonalizable on $\mathbb{C}$ since all roots of its characteristic polynomial are simple (see (2.3)). Extending notations (2.4), let $(u_\lambda)_{\lambda\in\mathrm{Sp}(\mathcal{A})}$ be a basis of linear forms, each u_λ being an eigenvector of G associated with the (complex) eigenvalue λ. Let also $(v_\lambda)_{\lambda\in\mathrm{Sp}(\mathcal{A})}$ be the dual basis of $(u_\lambda)_{\lambda\in\mathrm{Sp}(\mathcal{A})}$, each v_λ being thus a vector that satisfies $u_\lambda(v_\mu) = \delta_{\lambda,\mu}$ (Kronecker's notation). Note that one can choose $u_{\lambda_2} = u_2$ and, consequently, $v_{\lambda_2} = v_2$ (cf. notations (2.4)).

For any $t \geq 0$, split the spectral decomposition of the vector $X(t)$ with respect to G into four terms:

$$X(t) = \sum_{\lambda\in\mathrm{Sp}(\mathcal{A})} u_\lambda\big(X(t)\big)\cdot v_\lambda = X_1(t) + X_2(t) + X_3(t) + X_4(t),$$

where

$$\begin{cases} X_1(t) = u_1\big(X(t)\big)v_1, \\ X_2(t) = u_{\lambda_2}\big(X(t)\big)v_{\lambda_2} + u_{\overline{\lambda_2}}\big(X(t)\big)v_{\overline{\lambda_2}}, \\ X_3(t) = \sum_{1/2<\Re\lambda<\Re\lambda_2} u_\lambda\big(X(t)\big)v_\lambda, \\ X_4(t) = \sum_{\Re\lambda<1/2} u_\lambda\big(X(t)\big)v_\lambda. \end{cases}$$

Note that this partition of $\mathrm{Sp}(\mathcal{A})$ is valid because $\frac{1}{2}$ is not an eigenvalue of $\mathcal{A}$ as can be checked from (2.3). We deal separately with these four components of $X(t)$. Define $\boldsymbol{\varepsilon}_3$ by $\boldsymbol{\varepsilon}_3(t) = X_3(t) + X_4(t)$, for any $t \geq 0$.

• The formulae (4.2) provide directly the asymptotics

$$\begin{cases} X_1(t) = \big(\mathrm{e}^t\xi + \mathrm{o}(\mathrm{e}^t)\big)v_1, \\ X_2(t) = \big(\mathrm{e}^{\lambda_2 t}W + \mathrm{o}(\mathrm{e}^{\lambda_2 t})\big)v_2 + \overline{\big(\mathrm{e}^{\lambda_2 t}W + \mathrm{o}(\mathrm{e}^{\lambda_2 t})\big)v_2} \\ \qquad\;\; = 2\Re\big(\big(\mathrm{e}^{\lambda_2 t}W + \mathrm{o}(\mathrm{e}^{\lambda_2 t})\big)v_2\big), \end{cases}$$

leading to the first two terms of the expansion (4.1).

• Suppose that λ is an eigenvalue of $\mathcal{A}$ such that $\frac{1}{2} < \Re\lambda < \Re\lambda_2$. Then, with the same general arguments as in the very beginning of the proof, it can be seen that

$$u_\lambda\big(M(t)\big) = \mathrm{e}^{-t\lambda}u_\lambda\big(X(t)\big)$$

and that $(u_\lambda(M(t)))_{t\geq 0}$ is a convergent martingale, bounded in any L^p, $p \geq 1$. In particular, $u_\lambda(X(t)) = \mathrm{o}(\mathrm{e}^{\lambda_2 t})$ as t tends to infinity, almost surely and in any L^p, $p \geq 1$. This shows that $X_3(t)$ is $\mathrm{o}(\mathrm{e}^{\lambda_2 t})$ when $t \to +\infty$.

• It remains to deal with the small eigenvalues, namely with all λ such that $\Re\lambda < \frac{1}{2}$.

Lemma 4.2. *Suppose that λ is an eigenvalue such that $\Re\lambda < \frac{1}{2}$ and let $\eta > 0$. Then, $\mathrm{e}^{-(1/2+\eta)t}u_\lambda(X(t))$ is bounded almost surely and in any L^p-space, $p \geq 1$.*

The proof of this lemma is given just hereafter. Therefore, if $\Re\lambda < \frac{1}{2}$, then

$$\mathrm{e}^{-\lambda_2 t}u_\lambda\big(X(t)\big) = \mathrm{e}^{(1/2+\eta-\lambda_2)t}\big[\mathrm{e}^{-(1/2+\eta)t}u_\lambda\big(X(t)\big)\big] \underset{t\to\infty}{\longrightarrow} 0$$

almost surely as soon as $0 < \eta < \Re\lambda_2 - \frac{1}{2}$. Such η exist because $\Re\lambda_2 > \frac{1}{2}$. This shows that $X_4(t)$ is $\mathrm{o}(\mathrm{e}^{\lambda_2 t})$ when $t \to +\infty$. The same argument holds for the L^p convergence, making the proof complete. □

Proof of Lemma 4.2. The main idea consists in taking advantage of the following fact: when t belongs to the interval $[\tau_n, \tau_{n+1}[$, the vector $X(t)$ remains equal to X_n^{DT}. This being considered, we make use of the moment bounds of the discrete time process that can be found in [19] (Theorem 3.4(1)): when $\Re\lambda < \frac{1}{2}$,

$$\forall p \geq 1, \forall \varepsilon > 0, \quad \mathbb{E}\big|u_\lambda\big(X_n^{DT}\big)\big|^p = \mathrm{O}\big(n^{p(1/2+\varepsilon)}\big), \quad n \to +\infty. \tag{4.3}$$

• Almost sure bound: we prove that

$$\lim_{C\to+\infty} \mathbb{P}\big(\exists t > 0, \mathrm{e}^{-(1/2+\eta)t}\big|u_\lambda\big(X(t)\big)\big| > C\big) = 0, \tag{4.4}$$

which suffices to get the almost sure boundedness. Let $C > 0$, $\eta > 0$ and let λ be an eigenvalue such that $\Re\lambda < \frac{1}{2}$. The jump time τ_n tends almost surely to $+\infty$ which is a classical result that can be deduced from Lemma 3.1, so that

$$\begin{aligned} &\mathbb{P}\big(\exists t > 0, \mathrm{e}^{-(1/2+\eta)t}\big|u_\lambda\big(X(t)\big)\big| > C\big) \\ &\quad\leq \sum_{n\geq 0} \mathbb{P}\big(\exists t \in [\tau_n, \tau_{n+1}[, \mathrm{e}^{-(1/2+\eta)t}\big|u_\lambda\big(X(t)\big)\big| > C\big). \end{aligned}$$

Since $X(t) = X_n^{DT}$ for any $t \in [\tau_n, \tau_{n+1}[$, this leads to

$$\begin{aligned} &\mathbb{P}\big(\exists t > 0, \mathrm{e}^{-(1/2+\eta)t}\big|u_\lambda\big(X(t)\big)\big| > C\big) \\ &\quad\leq \sum_{n\geq 0} \mathbb{P}\big(\big|u_\lambda\big(X_n^{DT}\big)\big| > C\mathrm{e}^{(1/2+\eta)\tau_n}\big). \end{aligned}$$

Conditioning with respect to τ_n, using Markov inequality and the fact that τ_n and X_n^{DT} are independent, one gets successively, for any $p \geq 1$:

$$\begin{aligned}\mathbb{P}\big(\exists t > 0, \mathrm{e}^{-(1/2+\eta)t}\big|u_\lambda\big(X(t)\big)\big| > C\big) &\leq \sum_{n\geq 0}\mathbb{E}\big(\mathbb{P}\big(\big|u_\lambda\big(X_n^{DT}\big)\big| > C\mathrm{e}^{(1/2+\eta)\tau_n}|\tau_n\big)\big)\\ &\leq \sum_{n\geq 0}\mathbb{E}\bigg(\frac{\mathbb{E}|u_\lambda(X_n^{DT})|^p}{C^p\mathrm{e}^{p(1/2+\eta)\tau_n}}\bigg) = \frac{1}{C^p}\sum_{n\geq 0}\mathbb{E}\big|u_\lambda\big(X_n^{DT}\big)\big|^p\mathbb{E}\big(\mathrm{e}^{-p(1/2+\eta)\tau_n}\big).\end{aligned}$$

The density of the nth jump time τ_n is the function

$$u \in \mathbb{R} \longmapsto n\mathrm{e}^{-u}\big(1-\mathrm{e}^{-u}\big)^{n-1}\mathbf{1}_{\mathbb{R}_+}(u),$$

so that its Laplace transform can be elementarily computed: for any $s \geq 0$,

$$\mathbb{E}\big(\mathrm{e}^{-s\tau_n}\big) = \frac{n!\Gamma(s+1)}{\Gamma(s+1+n)} \sim \Gamma(s+1)n^{-s}, \quad n \to +\infty.$$

Together with (4.3), this leads to: $\forall\eta > 0, \forall\varepsilon > 0, \forall p \geq 1$,

$$\mathbb{E}\big|u_\lambda\big(X_n^{DT}\big)\big|^p\mathbb{E}\big(\mathrm{e}^{-p(1/2+\eta)\tau_n}\big) = \mathrm{O}\bigg(\frac{1}{n^{p(\eta-\varepsilon)}}\bigg), \quad n \to +\infty,$$

which is the general term of a convergent series as soon as one takes $\varepsilon < \eta$ and $p > \frac{1}{\eta-\varepsilon}$. Finally, letting C tend to infinity shows (4.4).

- Bound in L^p-space: let $p \geq 1$ and $t > 0$. Then,

$$\big\|\mathrm{e}^{-(1/2+\eta)t}u_\lambda\big(X(t)\big)\big\|_p^p = \mathrm{e}^{-(1/2+\eta)pt}\mathbb{E}\big|u_\lambda\big(X(t)\big)\big|^p.$$

Using the relation with the discrete time process $(X_n^{DT})_n$, one has successively

$$\begin{aligned}\big\|\mathrm{e}^{-(1/2+\eta)t}u_\lambda\big(X(t)\big)\big\|_p^p &= \mathrm{e}^{-(1/2+\eta)pt}\sum_{n\geq 0}\mathbb{E}\big(\mathbf{1}_{\tau_n\leq t<\tau_{n+1}}\big|u_\lambda\big(X(t)\big)\big|^p\big)\\ &= \mathrm{e}^{-(1/2+\eta)pt}\sum_{n\geq 0}\mathbb{E}\big(\mathbf{1}_{\tau_n\leq t<\tau_{n+1}}\big|u_\lambda\big(X_n^{DT}\big)\big|^p\big)\\ &= \mathrm{e}^{-(1/2+\eta)pt}\sum_{n\geq 0}\mathbb{E}(\mathbf{1}_{\tau_n\leq t<\tau_{n+1}})\mathbb{E}\big(\big|u_\lambda\big(X_n^{DT}\big)\big|^p\big),\end{aligned}$$

where the last equality holds due to the independence between τ_n and X_n^{DT}. Besides, τ_n and $\tau_{n+1}-\tau_n$ are independent and $\tau_{n+1}-\tau_n$ is $\mathcal{E}\mathrm{xp}(n+N_0)$-distributed (see (3.1)), so that, using the density of τ_n written above, one gets

$$\begin{aligned}\mathbb{E}(\mathbf{1}_{\tau_n\leq t<\tau_{n+1}}) &= \mathbb{E}\big(\mathbf{1}_{t\geq\tau_n}\mathbb{E}(\mathbf{1}_{\tau_{n+1}-\tau_n\geq t-\tau_n}|\tau_n)\big)\\ &= \mathbb{E}\big(\mathbf{1}_{t\geq\tau_n}\mathrm{e}^{-(n+N_0)(t-\tau_n)}\big)\\ &= \int_0^t \mathrm{e}^{-(n+N_0)(t-u)}n\mathrm{e}^{-u}\big(1-\mathrm{e}^{-u}\big)^{n-1}\,\mathrm{d}u\\ &\leq n\mathrm{e}^{-(n+1)t}\int_0^t\big(\mathrm{e}^u-1\big)^{n-1}\mathrm{e}^u\,\mathrm{d}u = \big(1-\mathrm{e}^{-t}\big)^n\mathrm{e}^{-t}.\end{aligned}$$

Thus,

$$\big\|\mathrm{e}^{-(1/2+\eta)t}u_\lambda\big(X(t)\big)\big\|_p^p \leq \mathrm{e}^{-t}\mathrm{e}^{-(1/2+\eta)pt}\sum_{n\geq 0}\big(1-\mathrm{e}^{-t}\big)^n\mathbb{E}\big(\big|u_\lambda\big(X_n^{DT}\big)\big|^p\big).$$

Let now $\varepsilon > 0$. On one hand, (4.3) implies that

$$\mathbb{E}\big(\big|u_\lambda\big(X_n^{DT}\big)\big|^p\big) = \mathrm{O}\big(n^{p(1/2+\varepsilon)}\big).$$

On the other hand, Stirling's formula applied to generalized binomial coefficients yields classically that for any $\alpha \in \mathbb{C}$,

$$\big[z^n\big](1-z)^{-\alpha-1} = \frac{n^\alpha}{\Gamma(\alpha+1)}\left(1+\mathrm{O}\left(\frac{1}{n}\right)\right),$$

where the notation $[z^n]A(z)$ means the coefficient of z^n in the power expansion of $A(z)$ at the origin. Consequently,

$$\mathbb{E}\big(\big|u_\lambda\big(X_n^{DT}\big)\big|^p\big) = \mathrm{O}\big(\big[z^n\big](1-z)^{-1-p(1/2+\varepsilon)}\big).$$

This implies that for any $\varepsilon > 0$, there exists a constant C_ε such that for any $t > 0$,

$$\big\|\mathrm{e}^{-(1/2+\eta)t}u_\lambda\big(X(t)\big)\big\|_p^p \le C_\varepsilon \mathrm{e}^{-t}\mathrm{e}^{-(1/2+\eta)pt}\big(1-\big(1-\mathrm{e}^{-t}\big)\big)^{-1-p(1/2+\varepsilon)} = C_\varepsilon \mathrm{e}^{-pt(\eta-\varepsilon)}.$$

It suffices to take $\varepsilon = \eta/2$ to conclude that the L^p-norm of $\mathrm{e}^{-(1/2+\eta)t}u_\lambda(X(t))$ is bounded above. □

Remark 4.3. *The distribution of W is infinitely divisible, because it is the limit of infinitely divisible ones, obtained by scaling and projection of infinitely divisible ones. Indeed, in finite time, for any $x_0 \in \mathbb{R}^{m-1}$, denote by $(X_{x_0}(t), t \ge 0)$ the process $(X(t), t \ge 0)$ defined in Section 2.1 starting from initial state x_0. By the branching property*

$$X_{x_0}(t) \overset{\mathcal{L}}{=} [n]X_{x_0/n}(t),$$

where the notation $[n]X$ denotes the sum of n independent copies of the random variable X. The infinite divisibility of W had already been noticed by Janson ([12], *proof of Theorem* 3.9).

4.2. *Martingale connection*

In this subsection, we use the embedding equality (3.2) to deduce connections between the asymptotic behaviours of X_n^{DT} when $n \to +\infty$ and $X(t)$ when $t \to +\infty$.

The Markov chain $(X_n^{DT})_n$ exhibits a well-known phase transition of the same kind as the continous time process: when $m \le 26$, with notations of Section 2.4, $n^{-1/2}(X_n^{DT} - nv_1)$ converges in law to a centered Gaussian vector (see Mahmoud's book [16]). For $m \ge 27$, it has been proved in [4] and [18] that

$$X_n^{DT} = nv_1 + 2\Re\big(n^{\lambda_2}W^{DT}v_2\big) + \mathrm{o}\big(n^{\sigma_2}\big) \quad \text{a.s. and in } L^p, \forall p \ge 1, \tag{4.5}$$

where v_1, v_2 are the deterministic vectors defined in (2.4), W^{DT} is the limit of a complex-valued martingale. Moreover, W^{DT} admits moments of all orders that can be recursively calculated and satifies $W^{DT} = \lim_{n\to\infty} n^{-\lambda_2}u_2(X_n^{DT})$ almost surely.

Proposition 4.4. *The following two assertions hold*:

$$\lim_{n\to+\infty} n\mathrm{e}^{-\tau_n} = \xi \quad \textit{a.s. and in } L^p, \forall p \ge 1, \tag{4.6}$$

$$W = \xi^{\lambda_2}W^{DT} \quad \textit{a.s. with } \xi \textit{ and } W^{DT} \textit{ independent.} \tag{4.7}$$

The equality (4.7), commonly referred to as *martingale connection*, establishes the link between W and W^{DT}. In this way, the results on W in the present paper can be seen as a first step to a better knowledge of W^{DT} distribution.

Proof of Proposition 4.4. We first prove (4.6). Applying the first projection to the embedding equality (3.2), we obtain that

$$u_1\big(X(\tau_n)\big) = u_1\big(X_n^{DT}\big) \quad \text{a.s.},$$

where u_1 has been defined in (2.4). This is the total number of free places at time τ_n, and is equal to $n - 1 + N_0 =$

$n(1+o(1))$. Therefore, by (4.2) and the fact that the splitting times τ_n tend almost surely to $+\infty$ when n goes to $+\infty$, we have

$$\xi = \lim_{t\to+\infty} \mathrm{e}^{-t} u_1\big(X(t)\big) = \lim_{n\to+\infty} n\mathrm{e}^{-\tau_n} \quad \text{a.s.}$$

This gives (4.6).

We then prove (4.7). Applying the second projection to the embedding equality (3.2) we obtain

$$u_2\big(X(\tau_n)\big) = u_2\big(X_n^{DT}\big) \quad \text{a.s.},$$

where u_2 has been defined in (2.4). Using again (4.2) and the fact that τ_n goes to $+\infty$ when n goes to $+\infty$, we get

$$W = \lim_{t\to+\infty} \mathrm{e}^{-\lambda_2 t} u_2\big(X(t)\big) = \lim_{n\to+\infty} \mathrm{e}^{-\lambda_2 \tau_n} u_2\big(X_n^{DT}\big).$$

Therefore, (4.7) follows from (4.6) and from the asymptotics in discrete time recalled in (4.5) and the following lines. Since $\xi = \lim_n n\mathrm{e}^{-\tau_n}$, $W^{DT} = \lim_n n^{-\lambda_2} u_2(X_n^{DT})$ and since $(\tau_n)_n$ and $(X_n^{DT})_n$ are independent, the variables ξ and W^{DT} are independent as well. Note that, in using (4.5) for the above, we have also used the fact (from the end of Section 2) that $u_2(\overline{v_2}) = 0$. □

Remark 4.5. *Fill and Kapur* [9] *proved that W^{DT} is the unique solution in the space of probability distributions with a given mean and finite second absolute moment of the fixed point equation*

$$Z \stackrel{\mathcal{L}}{=} \sum_{k=1}^{m} (V_k)^{\lambda_2} Z^{(k)}, \tag{4.8}$$

where the V_k are the spaces in the statistical order of $(m-1)$ i.i.d. random variables uniformly distributed on $[0,1]$. *Because $\mathbb{R}$-valued stable distributions are solutions of the fixed point equation* (4.8) *when λ is a real number, it is somehow natural to ask whether a λ-stable distribution is a solution of Eq.* (4.8) *for a complex number λ. By λ-stable we mean operator-stable when the operator is given by a two dimensional matrix $\lambda = \sigma + i\tau = \left(\begin{smallmatrix}\sigma & -\tau\\ \tau & \sigma\end{smallmatrix}\right)$ as introduced by Sharpe* [22]. *It is known since Hudson et al.* [11] *that a λ-stable distribution has infinite moments of order p for $p > 1/\Re(\lambda)$. Consequently, neither W nor W^{DT}* (*which have moments of any order*) *can be stable distributions.*

5. A distributional equation

In this section we derive a distributional equation satisfied by the limit variable of the continuous time branching process with an appropriate norming. We shall see that this equation characterizes the limit distribution.

5.1. *Vectorial finite time dislocation equations*

Let us write dislocation equations for the continuous time branching process at finite time t. We write $X_j(t)$ for $X(t)$ when the process starts from $X(0) = e_j$, where e_j denotes the jth vector of the canonical basis of $\mathbb{R}^{m-1}$ (whose jth component is 1 and all the others are 0). This means that the process starts from an ancestor of type j.

Notice that the distribution of the first splitting time τ_1 depends on the ancestor's type; denote by $\tau_{(j)}$, $j = 1,\dots,m-1$, the first splitting time when the process starts from $X(0) = e_j$. Thus $\tau_{(j)}$ is $\mathcal{E}\mathrm{xp}(j)$ distributed.

The branching property applied at the first splitting time gives:

$$\forall t > \tau_1, \quad \begin{cases} X_1(t) \stackrel{\mathcal{L}}{=} X_2(t-\tau_{(1)}),\\ X_2(t) \stackrel{\mathcal{L}}{=} X_3(t-\tau_{(2)}),\\ \dots\\ X_{m-2}(t) \stackrel{\mathcal{L}}{=} X_{m-1}(t-\tau_{(m-2)}),\\ X_{m-1}(t) \stackrel{\mathcal{L}}{=} [m]X_1(t-\tau_{(m-1)}), \end{cases} \tag{5.1}$$

where the notation $[m]X$ denotes the sum of m independent copies of the random variable X.

In the following, let

$$T = \tau_{(1)} + \cdots + \tau_{(m-1)}, \tag{5.2}$$

where the $\tau_{(j)}$ are independent of each other and each $\tau_{(j)}$ is $\mathcal{E}\mathrm{xp}(j)$ distributed. Let us give some elementary properties of T that we shall need. It is easy to see that T has density

$$f_T(u) = (m-1)\mathrm{e}^{-u}\big(1-\mathrm{e}^{-u}\big)^{m-2}\mathbf{1}_{\mathbb{R}_+}(u), \quad u \in \mathbb{R}, \tag{5.3}$$

so that e^{-T} has a Beta distribution with parameters 1 and $m-1$. A straightforward change of variable under the integral shows that for any complex number λ such that $\Re(\lambda) > 0$,

$$\mathbb{E}\mathrm{e}^{-\lambda T} = \int_0^{+\infty} \mathrm{e}^{-\lambda u} f_T(u)\,\mathrm{d}u = (m-1)B(1+\lambda, m-1) \tag{5.4}$$

$$= \frac{(m-1)!}{\prod_{k=1}^{m-1}(\lambda+k)}, \tag{5.5}$$

where B denotes Euler's Beta function:

$$B(x,y) = \int_0^1 u^{x-1}(1-u)^{y-1}\,\mathrm{d}u = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}, \quad \Re x > 0, \Re y > 0.$$

In particular,

$$m\mathbb{E}\big|\mathrm{e}^{-\lambda T}\big| \begin{cases} < 1 & \text{if } \Re(\lambda) > 1, \\ = 1 & \text{if } \Re(\lambda) = 1, \\ > 1 & \text{if } \Re(\lambda) < 1. \end{cases} \tag{5.6}$$

5.2. *Distributional equation satisfied by the limit variable*

After projections of variables $X_j(t)$ (the process starting from $X(0) = e_j$) with u_2, scaling with $\mathrm{e}^{-\lambda_2 t}$ and taking the limit when t goes to infinity, we get the variables

$$W_j := \lim_{t \to +\infty} \mathrm{e}^{-\lambda_2 t} u_2\big(X_j(t)\big),$$

so that the system (5.1) on $X_j(t)$ leads to the following system of distributional equations on W_j:

$$\begin{cases} W_1 \stackrel{\mathcal{L}}{=} \mathrm{e}^{-\lambda_2 \tau_{(1)}} W_2, \\ W_2 \stackrel{\mathcal{L}}{=} \mathrm{e}^{-\lambda_2 \tau_{(2)}} W_3, \\ \cdots \\ W_{m-2} \stackrel{\mathcal{L}}{=} \mathrm{e}^{-\lambda_2 \tau_{(m-2)}} W_{m-1}, \\ W_{m-1} \stackrel{\mathcal{L}}{=} \mathrm{e}^{-\lambda_2 \tau_{(m-1)}} [m] W_1. \end{cases} \tag{5.7}$$

This shows that W_1 is a solution of the following fixed point equation:

$$Z \stackrel{\mathcal{L}}{=} \mathrm{e}^{-\lambda_2 T}\big(Z^{(1)} + \cdots + Z^{(m)}\big), \tag{5.8}$$

where $Z^{(i)}$ are independent copies of Z, which are also independent of T.

In terms of the Fourier transform

$$\varphi(t) := \mathbb{E}\exp\big\{\mathrm{i}\langle t, Z\rangle\big\} = \mathbb{E}\exp\big\{\mathrm{i}\Re(\bar{t}Z)\big\}, \quad t \in \mathbb{C},$$

where $\langle x, y\rangle = \Re(\overline{x}y) = \Re(x)\Re(y) + \Im(x)\Im(y)$, Eq. (5.8) reads

$$\varphi(t) = \int_0^{+\infty} \varphi^m\big(t\mathrm{e}^{-\overline{\lambda_2}u}\big) f_T(u)\,\mathrm{d}u, \quad t \in \mathbb{C}, \tag{5.9}$$

where f_T is defined by (5.3). Notice that this functional equation can also be written in a convolution form: if $\Phi(t) := \varphi(\mathrm{e}^{\overline{\lambda_2}t})$ for any $t \in \mathbb{C}$, then Φ satisfies the following functional equation:

$$\Phi(t) = \int_0^{+\infty} \Phi^m(t-u) f_T(u)\,\mathrm{d}u, \quad t \in \mathbb{C}. \tag{5.10}$$

In the following sections, we prove that the distributional equation (5.8) characterizes the law of W_1 and we get several results on W_1: for example W_1 has a density on the whole complex plane, and admits exponential moments. All these results appear as a particular case of a slightly more general situation given hereafter. From now on, for any complex number λ, consider the distributional equation

$$Z \stackrel{\mathcal{L}}{=} \mathrm{e}^{-\lambda T}\big(Z^{(1)} + \cdots + Z^{(m)}\big), \tag{5.11}$$

where $Z^{(i)}$ are independent copies of Z, which are also independent of T. In terms of Fourier transforms, it reads

$$\varphi(t) = \int_0^{+\infty} \varphi^m\big(t\mathrm{e}^{-\overline{\lambda}u}\big) f_T(u)\,\mathrm{d}u, \quad t \in \mathbb{C}, \tag{5.12}$$

where f_T is defined by (5.3).

Notice that when Z is a solution of the distributional equation (5.11), with finite and nonzero first moment, then λ is a root of the polynomial function (2.3). In particular, λ is an algebraic number.

6. The smoothing transformation

A solution of the distributional equation (5.11) is a fixed point of the associated smoothing transformation defined hereafter by (6.1). Endowing a suitable space of probability measures with two distances, we prove that the smoothing transformation is a contraction for both metrics. This provides two alternative approaches for the study of Eq. (5.11) by the contraction method. Using the Wasserstein distance as a first metric, we adapt the classical contraction method developed in [10,20] and [21]. The second metric is defined in terms of Fourier transforms of measures; it provides a short proof of our result.

For any complex number C, let $\mathcal{M}_2(C)$ be the space of probability distributions on $\mathbb{C}$ admitting a second absolute moment and having C as expectation.

Let λ be a complex number. For any probability measure μ on $\mathbb{C}$, let

$$K\mu := \mathcal{L}\big(\mathrm{e}^{-\lambda T}\big(Z^{(1)} + \cdots + Z^{(m)}\big)\big), \tag{6.1}$$

where T is given by (5.2), $Z^{(i)}$ are independent random variables of law μ, which are also independent of T. Following Durrett and Liggett [8] who considered the case of real random variables, we call K the *smoothing transformation*. Note that K depends on m and λ.

Lemma 6.1. *If λ is a root of the characteristic polynomial (2.3) such that $\Re(\lambda) > -\frac{1}{2}$ and if C is any complex number, then K maps $\mathcal{M}_2(C)$ into itself.*

Proof. Since $\Re(\lambda) > -1$, the random variable $\mathrm{e}^{-\lambda T}$ has an expectation. Furthermore, by (5.4), $m\mathbb{E}\mathrm{e}^{-\lambda T} = 1$ as λ is a root of (2.3). This ensures the conservation of the expectation by K. Since $\Re(\lambda) > -\frac{1}{2}$, then $\mathbb{E}|\mathrm{e}^{-\lambda T}|^2 < \infty$ and $K\mu$ admits a second absolute moment whenever μ does. Therefore $K\mu \in \mathcal{M}_2(C)$ whenever $\mu \in \mathcal{M}_2(C)$. □

Notice that a solution of Eq. (5.11) is a fixed point of K. We shall use the Banach fixed point theorem for two different metrics on $\mathcal{M}_2(C)$ to study the existence and uniqueness of solutions of Eq. (5.11).

6.1. *Wasserstein distance*

Let d_2 be the Wasserstein distance on $\mathcal{M}_2(C)$ (see for instance Dudley [7]): for $\mu, \nu \in \mathcal{M}_2(C)$,

$$d_2(\mu, \nu) = \Bigl(\min_{(X,Y)} \mathbb{E}\bigl(|X - Y|^2\bigr)\Bigr)^{1/2},$$

where the minimum is taken over couples of random variables (X, Y) having respective marginal distributions μ and ν; the minimum is attained by the Kantorovich–Rubinstein Theorem – see for instance Dudley [7], p. 421. With this distance d_2, $\mathcal{M}_2(C)$ is a complete metric space.

Theorem 6.2. *Let $\lambda \in \mathbb{C}$ be a root of the characteristic polynomial (2.3) such that $\Re(\lambda) > \frac{1}{2}$, and let $C \in \mathbb{C}$. Then K is a contraction on the complete metric space $(\mathcal{M}_2(C), d_2)$, and the fixed point equation (5.11) has a unique solution Z in $\mathcal{M}_2(C)$.*

We now come back to the limit variable W_1 of m-ary search trees. Since $\mathbb{E}W_1 = 1$ and $\mathbb{E}|W_1|^2 < \infty$, the following corollary is a direct consequence of Theorem 6.2, applied for $\lambda = \lambda_2$.

Corollary 6.3. *The distribution of the limit complex random variable W_1 is the unique solution in the space $\mathcal{M}_2(1)$ of the fixed point equation (5.11).*

Proof of Theorem 6.2. We argue as in [10,20] and [21] where real random variables were considered.

By the Banach fixed point theorem, it suffices to show the contraction property. Let $\mu, \nu \in \mathcal{M}_2(C)$. Let (X, Y) be a couple of complex-valued random variables such that $\mathcal{L}(X) = \mu$, $\mathcal{L}(Y) = \nu$ and $d_2(\mu, \nu) = \sqrt{\mathbb{E}|X - Y|^2}$. Let $(X_i, Y_i), i = 1, \dots, m$ be m independent copies of the d_2-optimal couple (X, Y), and T be a real random variable with density f_T defined by (5.3), independent from all (X_i, Y_i). Then,

$$\mathcal{L}\left(\mathrm{e}^{-\lambda T} \sum_{i=1}^{m} X_i\right) = K\mu \quad \text{and} \quad \mathcal{L}\left(\mathrm{e}^{-\lambda T} \sum_{i=1}^{m} Y_i\right) = K\nu,$$

so that

$$\begin{aligned}
d_2(K\mu, K\nu)^2 &\le \mathbb{E}\left|\left(\mathrm{e}^{-\lambda T} \sum_{i=1}^{m} X_i\right) - \left(\mathrm{e}^{-\lambda T} \sum_{i=1}^{m} Y_i\right)\right|^2 \\
&= \mathbb{E}\left|\mathrm{e}^{-\lambda T} \sum_{i=1}^{m} (X_i - Y_i)\right|^2 \\
&= \mathbb{E}\left|\mathrm{e}^{-\lambda T}\right|^2 \mathbb{E}\left|\sum_{i=1}^{m} (X_i - Y_i)\right|^2 \\
&= \mathbb{E}\left|\mathrm{e}^{-\lambda T}\right|^2 \left(\sum_{i=1}^{m} \mathbb{E}|X_i - Y_i|^2 + \sum_{i \ne j} \mathbb{E}(X_i - Y_i)(\overline{X_j} - \overline{Y_j})\right) \\
&= m\mathbb{E}\left|\mathrm{e}^{-2\lambda T}\right| d_2(\mu, \nu)^2.
\end{aligned}$$

Since $2\Re(\lambda) > 1$, we have $m\mathbb{E}|\mathrm{e}^{-2\lambda T}| < 1$ (see (5.6)). Therefore K is a contraction on $\mathcal{M}_2(C)$ and the proof is complete. $\square$

6.2. *Distance defined with Fourier transforms*

We now give an alternative approach for the characterization of the limit distribution via Fourier analysis. We define another distance d_2^* on $\mathcal{M}_2(C)$ as follows. Take $\mu, \nu \in \mathcal{M}_2(C)$ and denote respectively φ and ψ their characteristic functions. By definition of $\mathcal{M}_2(C)$, both φ and ψ admit the expansion $\varphi(t) = 1 + \mathrm{i}\langle t, C\rangle + \mathrm{O}(|t|^2)$ when t tends to 0. Therefore, one can define $d_2^*(\mu, \nu)$ by

$$d_2^*(\mu, \nu) = \sup_{t\in\mathbb{C}\setminus\{0\}} \frac{|\varphi(t) - \psi(t)|}{|t|^2}.$$

Clearly, $d_2^*(\mu, \nu) < \infty$, and d_2^* is a distance on $\mathcal{M}_2(C)$. It can be easily checked that $(\mathcal{M}_2(C), d_2^*)$ is a complete metric space.

The following result is a counterpart of Theorem 6.2. It gives an alternative proof for the existence and uniqueness of the solution of Eq. (5.11) in the class of probability measures on $\mathbb{C}$ with a given mean and finite second moments.

Theorem 6.4. *Let $\lambda \in \mathbb{C}$ be a root of the characteristic polynomial (2.3) such that $\Re(\lambda) > \frac{1}{2}$, and let $C \in \mathbb{C}$. Then K is a contraction on the complete metric space $(\mathcal{M}_2(C), d_2^*)$, and the fixed point equation (5.11) has a unique solution Z in $\mathcal{M}_2(C)$.*

Proof. Thanks to Banach fixed point theorem, it suffices to prove that K is a contraction on $\mathcal{M}_2(C)$ equipped with the metric d_2^*. Let $\mu, \nu \in \mathcal{M}_2(C)$ and let φ and ψ be their respective characteristic functions. An elementary computation shows that the Fourier transform of $K\mu$ is $t \mapsto \mathbb{E}\varphi^m(\mathrm{e}^{-\overline{\lambda}T}t)$ with a corresponding formula for ν. We have $|\varphi| \le 1$, $|\psi| \le 1$, so that

$$\mathbb{E}\left|\varphi^m\left(\mathrm{e}^{-\overline{\lambda}T}t\right) - \psi^m\left(\mathrm{e}^{-\overline{\lambda}T}t\right)\right| \le m\mathbb{E}\left|\varphi\left(\mathrm{e}^{-\overline{\lambda}T}t\right) - \psi\left(\mathrm{e}^{-\overline{\lambda}T}t\right)\right|.$$

Together with the inequality $|\varphi(z) - \psi(z)| \le d_2^*(\mu, \nu)|z|^2$ applied to $z = \mathrm{e}^{-\overline{\lambda}T}t$, this implies that

$$d_2^*(K\mu, K\nu) \le m\mathbb{E}\left(\mathrm{e}^{-2\Re(\lambda)T}\right)d_2^*(\mu, \nu).$$

Since $2\Re(\lambda) > 1$, we have $m\mathbb{E}(\mathrm{e}^{-2\Re(\lambda)T}) < 1$ (see (5.6)). Therefore the above inequality shows that K is a contraction on $(\mathcal{M}_2(C), d_2^*)$. $\square$

Remark 6.5. *Denote $\mathcal{F}_2(C)$ the space of Fourier transforms of elements of $\mathcal{M}_2(C)$. When λ is a root of the characteristic polynomial (2.3) such that $\Re(\lambda) > \frac{1}{2}$, the smoothing transformation K can be identified as a map (also denoted by K) on $\mathcal{F}_2(C)$ given by*

$$(K\varphi)(t) := \mathbb{E}\varphi^m\left(\mathrm{e}^{-\overline{\lambda}T}t\right), \quad t \in \mathbb{C}. \tag{6.2}$$

The proof of Theorem 6.4 also shows that K is a contraction of $\mathcal{F}_2(C)$ for the metric (also denoted by d_2^*) defined on $\mathcal{F}_2(C)$ by

$$d_2^*(\varphi, \psi) := \sup_{t\neq 0} \frac{|\varphi(t) - \psi(t)|}{|t|^2}. \tag{6.3}$$

Let $\mathcal{D}_2(C)$ be the space of all continuous functions $\varphi : \mathbb{C} \to \mathbb{C}$ that admit an expansion $\varphi(t) = 1 + \mathrm{i}\langle t, C\rangle + \mathrm{O}(|t|^2)$ at 0 and such that $\|\varphi\|_\infty \le 1$. Clearly, $\mathcal{D}_2(C)$ contains $\mathcal{F}_2(C)$. One can show that formula (6.2) defines a mapping from $\mathcal{D}_2(C)$ into itself and that K is a contraction for the metrics defined by (6.3). This provides a proof of existence and uniqueness of solutions of (5.12) on $\mathcal{D}_2(C)$.

Remark 6.6. *One can deal with the convolution equation (5.10) by arguments in the same vein. Similar computations show that this equation has a unique solution in the space $\mathcal{E}_2(\lambda, C)$ of continuous functions $\Phi : \mathbb{C} \to \mathbb{C}$ that admit an* *expansion* $\Phi(z) = 1 + \mathrm{i}\langle \mathrm{e}^{\lambda z}, C\rangle + \mathrm{O}(\mathrm{e}^{2\lambda z})$ *when* $|z|$ *tends to* $+\infty$ *and such that* $\|\Phi\|_\infty \le 1$. *This result appears once again as a consequence of Banach theorem on* $\mathcal{E}_2(\lambda, C)$ *for the metric*

$$d(\Phi, \Psi) = \sup_{z\in\mathbb{C}} \left| \frac{\Phi(z) - \Psi(z)}{\mathrm{e}^{2\lambda z}} \right|.$$

As a consequence, this shows in particular that the Fourier (complex) transform φ *of* W_1 *satisfies: for any* $w \in \mathbb{C}^*$ *and for any branch of the logarithm,*

$$\varphi(w) = \Phi\left(\frac{\log w}{\overline{\lambda_2}}\right),$$

where Φ *is the unique solution in* $\mathcal{E}_2(\overline{\lambda_2}, 1)$ *of Eq.* (5.10). *This result is the reversed version of the change of variable* $\Phi(z) = \varphi(\mathrm{e}^{\overline{\lambda_2} z})$ *that led from* (5.9) *to* (5.10).

7. Density and support

In this section we prove results on the absolute continuity and on the support of solutions of the distributional equation (5.11) via Fourier analysis. As applications, we show that the distribution of the limit variable W_1 coming from the multitype branching process $(X(t))$ always has a density and that its support is the whole complex plane.

Theorem 7.1. *Let* λ *be a complex number such that* $\lambda \neq 1$ *and* $\Re(\lambda) > 0$. *Let* Z *be a complex-valued random variable solution of the distributional equation* (5.11)

$$Z \stackrel{\mathcal{L}}{=} \mathrm{e}^{-\lambda T}\bigl(Z^{(1)} + \cdots + Z^{(m)}\bigr),$$

with $\mathbb{E}|Z| < \infty$ *and* $\mathbb{E}Z \neq 0$. *Then the following assertions hold:*

(i) *the support of* Z *is the whole complex plane* $\mathbb{C}$;
(ii) *as* $|t| \to \infty$, $\mathbb{E}\mathrm{e}^{\mathrm{i}\langle t, Z\rangle} = \mathrm{O}(|t|^{-a})$, *for any* $a \in]0, \frac{1}{\Re(\lambda)}[$;
(iii) *the distribution of* Z *has a density with respect to the Lebesgue measure on* $\mathbb{C}$.

Remark 7.2. *When* $\lambda = 1$, *the distributional equation* (5.11) *becomes*

$$X \stackrel{\mathcal{L}}{=} \mathrm{e}^{-T}\bigl(X^{(1)} + \cdots + X^{(m)}\bigr). \tag{7.1}$$

By Section 6, *it admits a unique solution in the space* $\mathcal{M}_2(C)$ *of probability measures on* $\mathbb{C}$, *with a given mean* C. *Moreover, from the dislocation equations* (5.1), *a similar argument shows that*

$$\xi := \lim_{t\to+\infty} \mathrm{e}^{-t} u_1\bigl(X(t)\bigr)$$

is a solution of this equation. By Theorem 4.1, ξ *is Gamma-distributed. Therefore the unique solution of* (7.1) *in* $\mathcal{M}_2(C)$ *is* $C\gamma$ *where* γ *is Gamma*(1)*-distributed, and its support is the half line* $C\mathbb{R}_+$.

The following corollary gives the main result for the limit variable W_1 of the multitype branching process. It is a direct consequence of Theorem 7.1 since $\mathbb{E}W_1 = 1$.

Corollary 7.3. *The distribution of* W_1 *admits a density with respect to the Lebesgue measure on* $\mathbb{C}$, *and its support is the whole complex plane* $\mathbb{C}$. *Moreover, as* $|t| \to \infty$, $\mathbb{E}\mathrm{e}^{\mathrm{i}\langle t, W_1\rangle} = \mathrm{O}(|t|^{-a})$ *for each* $a \in]0, \frac{1}{\Re(\lambda_2)}[$.

The proof of Theorem 7.1 runs along the following lines. Let φ be the Fourier transform of any solution Z of (5.11). We prove that φ is in $L^2(\mathbb{C})$ because it is dominated by $|t|^{-\delta}$ for some $\delta > 1$ so that the inverse Fourier–Plancherel transform provides a square integrable density for Z. The guiding idea consists in adapting usual methods (developed in [13] and [14]) used for positive real-valued random variables to our complex-valued case thanks to the function defined for $r \geq 0$ by

$$\psi(r) = \max_{|t|=r} \left|\varphi(t)\right|.$$

From now on,

$$A = \mathrm{e}^{-\lambda T}.$$

We proceed by a series of lemmas. The first lemma concerns a property of the support of Z. For a complex-valued random variable Z and for a complex number z, by definition,

$$z \in \mathrm{Supp}(Z) \quad \Longleftrightarrow \quad \forall \varepsilon > 0, \quad \mathbb{P}\big(|Z - z| \leq \varepsilon\big) > 0.$$

Lemma 7.4. *Let $z \in \mathbb{C}$. Then*

$$z \in \mathrm{Supp}(Z) \quad \Longrightarrow \quad D\big(0, |z|\big) \subseteq \mathrm{Supp}(Z),$$

where $D(0, |z|)$ denotes the open disc with center 0 and radius $|z|$.

Proof. We first prove the following implication:

$$\big[a \in \mathrm{Supp}(A) \text{ and } z \in \mathrm{Supp}(Z)\big] \quad \Longrightarrow \quad maz \in \mathrm{Supp}(Z).$$

Indeed, let $\varepsilon > 0$, $a \in \mathrm{Supp}(A)$ and $z \in \mathrm{Supp}(Z)$. Let also $Z^{(1)}, \ldots, Z^{(m)}$ be independent copies of Z. Then, with positive probability, $|A - a| \leq \varepsilon$ and $|Z^{(k)} - z| \leq \varepsilon$ for any k. Therefore, with positive probability,

$$\begin{aligned}\big|A\big(Z^{(1)} + \cdots + Z^{(m)}\big) - maz\big| &= \left|mz(A - a) + A\sum_{k=1}^{m}\big(Z^{(k)} - z\big)\right| \\ &\leq m\varepsilon|z| + \big(|a| + \varepsilon\big)m\varepsilon.\end{aligned}$$

The positive real ε being arbitrary, this shows that $maz \in \mathrm{Supp}\, A(Z^{(1)} + \cdots + Z^{(m)})$ which implies that $maz \in \mathrm{Supp}(Z)$ by Eq. (5.11), proving the claim.

Let $z \in \mathrm{Supp}(Z)$. Since $\mathrm{Supp}(T) = \mathbb{R}_+$ (see (5.3)), the claim implies that for any $t \geq 0$, $m\mathrm{e}^{-\lambda t}z \in \mathrm{Supp}(Z)$. Iterating this property of $\mathrm{Supp}(Z)$ shows that

$$\big\{m^n \mathrm{e}^{-\lambda t} z, n \in \mathbb{N}, t \in \mathbb{R}_+\big\} \subseteq \mathrm{Supp}(Z). \tag{7.2}$$

Since the support of a probability measure is a closed set, to show that $D(0, |z|) \subseteq \mathrm{Supp}(Z)$ it suffices to prove that $\{m^n \mathrm{e}^{-\lambda t}, n \in \mathbb{N}, t \in \mathbb{R}_+\}$ is everywhere dense in the unit disc. Taking logarithm, we show hereafter that

$$\mathcal{G} := \mathbb{N}\log m + 2\mathrm{i}\pi\mathbb{N} - \lambda\mathbb{R}_+$$

is everywhere dense in the half-strip

$$\mathcal{B} := \{x + \mathrm{i}y, x < 0, -2\pi < y \leq 0\}$$

which implies the desired result.

Let σ and τ denote respectively the real and imaginary parts of λ. Remember that, as soon as Z is a solution of Eq. (5.11) such that $\mathbb{E}|Z| < \infty$ and $\mathbb{E}Z \neq 0$, then λ is a root of χ_{R_G} defined by (2.3); this implies that λ is an algebraic number, and so is σ/τ.

Let us prove that $\rho := \mathrm{i}\pi/\log m$ is a transcendental number. In fact, if ρ were algebraic, then by the Gelfond–Schneider theorem,[2] m^ρ would be a transcendental number; but this is impossible because $m^\rho = \exp(\mathrm{i}\pi) = -1$. Therefore ρ is a transcendental number.

It follows that $2\pi\sigma/(\tau \log m)$ is not an algebraic number, hence not a rational number. So by a classical result,

$$\mathbb{N}\log m - \mathbb{N}\frac{2\pi\sigma}{\tau} \text{ is a dense subset of } \mathbb{R}.$$

Let $b = x + \mathrm{i}y \in \mathcal{B}$ with $x < 0$ and $-2\pi < y \le 0$, and let $\varepsilon > 0$. Approximating the real number $x - \frac{y\sigma}{\tau}$ by an element of $\mathbb{N}\log m - \mathbb{N}\frac{2\pi\sigma}{\tau}$, we take $n, k \in \mathbb{N}$ such that

$$\left|\left(n\log m - k\frac{2\pi\sigma}{\tau}\right) - \left(x - \frac{y\sigma}{\tau}\right)\right| \le \varepsilon.$$

Therefore

$$|b - g| \le \varepsilon, \quad \text{where } g = n\log m + 2\mathrm{i}k\pi - t\lambda \in \mathcal{G} \text{ with } t = \frac{2k\pi - y}{\tau} \ge 0.$$

This completes the proof of Lemma 7.4. □

Lemma 7.4 leads to the following key property of ψ, which will imply that the characteristic function φ of Z satisfies the Cramér condition and then the Riemann–Lebesgue condition.

Lemma 7.5. $\forall r > 0$, $\psi(r) < 1$.

Proof. Obviously, $\psi(0) = 1$ and $\psi(r) \le 1$ for any $r \ge 0$. Suppose that $r_0 > 0$ is such that $\psi(r_0) = 1$. Take thus $z_0 \in \mathbb{C}$ and $\theta_0 \in \mathbb{R}$ such that

$$|z_0| = r_0 \quad \text{and} \quad \mathbb{E}\mathrm{e}^{\mathrm{i}\langle z_0, Z\rangle} = \mathrm{e}^{\mathrm{i}\theta_0}.$$

The complex random variable $\mathrm{e}^{\mathrm{i}(\langle z_0, Z\rangle - \theta_0)}$ is of mean 1 and takes its values on the unit disc, so that it is almost surely equal to 1. This implies that $\mathrm{Supp}(Z)$ is contained in a set of countably many parallel lines of the complex plane. This contradicts Lemma 7.4 since such a set of lines is negligible with respect to the Lebesgue measure on $\mathbb{C}$. □

Remark 7.6. *The preceding arguments show the following assertion*: *for any complex-valued random variable* Z, *if* $|\mathbb{E}(\mathrm{e}^{\mathrm{i}\langle z_0, Z\rangle})| = 1$ *for some* $z_0 \in \mathbb{C}\setminus\{0\}$, *then* $\mathrm{Supp}(Z) \subseteq a + b\mathbb{Z} + c\mathbb{R}$ *for some* $a, b, c \in \mathbb{C}$ (*a set of countably many parallel lines*). *The algebraicity of* λ *that leads to the proof of Lemma* 7.4 *can thus be seen as a nonlattice assumption on the fixed point equation* (5.11).

We now prove that the characteristic function φ satisfies the Riemann–Lesbesgue condition.

Lemma 7.7. $\lim_{r\to+\infty} \psi(r) = 0$.

Proof. We argue as in the proof of Theorem 3.1 of [13] or Lemma 3.1 of [14]. Notice that from the distributional equation (5.12) we have

$$\psi(r) \le \mathbb{E}\big(\psi^m\big(r|A|\big)\big). \tag{7.3}$$

- We first prove that $\limsup_{r\to+\infty} \psi(r) = 0$ or 1. By Fatou's lemma,

$$\limsup_{r\to+\infty} \psi(r) \le \mathbb{E}\limsup_{r\to+\infty} \psi^m\big(r|A|\big) = \Big(\limsup_{r\to+\infty} \psi(r)\Big)^m,$$

[2] The Gelfond–Schneider theorem states that if a and b are algebraic numbers with $a \ne 0, 1$ and if b is not a rational number, then any value of $a^b = \exp(b\log a)$ is a transcendental number.

the last equality coming from $\mathbb{P}(|A|>0)=1$. So the real number $l:=\limsup_r \psi(r)$ satisfies both $l\le 1$ and $l\le l^m$; this implies that $l=0$ or 1.

• Suppose that $\limsup_r \psi(r)=1$. By Lemma 7.5, $\psi(1)<1$. For any $\varepsilon\in]0,1-\psi(1)[$, define

$$\begin{cases} r_1(\varepsilon)=\max\{r\in]0,1[,\ \psi(r)=1-\varepsilon\},\\ r_2(\varepsilon)=\min\{r>1,\ \psi(r)=1-\varepsilon\}.\end{cases}$$

These quantities are well defined because $\psi(0)=1$ and ψ is continuous. Then $\psi(r_1(\varepsilon))=\psi(r_2(\varepsilon))=1-\varepsilon$ and for any $r\in[r_1(\varepsilon),r_2(\varepsilon)]$, $\psi(r)\le 1-\varepsilon$.

Let us prove that $r_1(\varepsilon)$ goes to 0 when ε tends to 0. Take any limit point ρ of $r_1(\varepsilon)$. Since ψ is continuous, $\psi(\rho)=1-\varepsilon$ which implies by Lemma 7.5 that $\rho=0$: the only possible limit point is 0.

By (7.3), we have

$$\psi(r)\le \mathbb{E}\psi\bigl(r|A|\bigr).$$

Iterating this inequality we see that for all $n\ge 1$,

$$\psi(r)\le \mathbb{E}\psi\bigl(r|A_1|\cdots|A_n|\bigr),$$

where $(|A_i|)_{i\ge 1}$ are independent copies of $|A|$. With the notation

$$\lambda_n(r,\varepsilon):=\mathbb{P}\bigl(r_1(\varepsilon)<r|A_1|\cdots|A_n|\le r_2(\varepsilon)\bigr),$$

we have for any $r>0$

$$\psi(r)\le(1-\varepsilon)\lambda_n(r,\varepsilon)+1-\lambda_n(r,\varepsilon)=1-\varepsilon\lambda_n(r,\varepsilon).$$

Again by (5.12),

$$1-\varepsilon=\psi\bigl(r_2(\varepsilon)\bigr)\le \mathbb{E}\psi^m\bigl(r_2(\varepsilon)|A|\bigr)\le \mathbb{E}\bigl(1-\varepsilon\lambda_n\bigl(r_2(\varepsilon)|A|,\varepsilon\bigr)\bigr)^m.$$

In other words

$$\frac{\mathbb{E}(1-(1-\varepsilon\lambda_n(r_2(\varepsilon)|A|,\varepsilon))^m)}{\varepsilon}\le 1. \tag{7.4}$$

We are going to pass to the limit in the above ratio when ε tends to 0. Rewrite

$$\lambda_n\bigl(r_2(\varepsilon)|A|,\varepsilon\bigr)=\mathbb{P}\left(\frac{r_1(\varepsilon)}{r_2(\varepsilon)}<|A||A_1|\cdots|A_n|\le 1\right),$$

and remember that $r_1(\varepsilon)\le 1\le r_2(\varepsilon)$ and that $r_1(\varepsilon)$ goes to 0 when ε tends to 0, so that $\frac{r_1(\varepsilon)}{r_2(\varepsilon)}\le\frac{r_1(\varepsilon)}{1}$ goes to 0 when ε tends to 0. Consequently,

$$\lambda_n\bigl(r_2(\varepsilon)|A|,\varepsilon\bigr)\underset{\varepsilon\to 0}{\longrightarrow}\mathbb{P}\bigl(0\le|A||A_1|\cdots|A_n|\le 1\bigr)=\mu_n\bigl(|A|\bigr)\quad\text{a.s.},$$

where, for any $x>0$,

$$\mu_n(x):=\mathbb{P}\bigl(x|A_1|\cdots|A_n|\le 1\bigr).$$

Therefore

$$\frac{1-(1-\varepsilon\lambda_n(r_2(\varepsilon)|A|,\varepsilon))^m}{\varepsilon}\underset{\varepsilon\to 0}{\longrightarrow}m\mu_n\bigl(|A|\bigr)\quad\text{a.s.}$$

The above ratio is a function of ε, uniformly bounded on the compact set $[0, 1-\psi(1)]$, so that by dominated convergence and (7.4),

$$\frac{\mathbb{E}(1-(1-\varepsilon\lambda_n(r_2(\varepsilon)|A|,\varepsilon))^m)}{\varepsilon} \underset{\varepsilon\to 0}{\longrightarrow} m\mathbb{E}\mu_n\bigl(|A|\bigr) \le 1. \tag{7.5}$$

Besides, by Markov's inequality

$$1-\mu_n(x) \le x\mathbb{E}\bigl(|A_1|\cdots|A_n|\bigr) = x\bigl(\mathbb{E}|A|\bigr)^n.$$

Since $\Re(\lambda) > 0$, $\mathbb{E}|A| = \mathbb{E}|\mathrm{e}^{-\lambda T}| < 1$ (see (5.4)), which implies that $\lim_{n\to\infty}\mu_n(x) = 1$, so that

$$\lim_{n\to\infty}\mathbb{E}\mu_n\bigl(|A|\bigr) = 1$$

by dominated convergence. This contradicts (7.5) because $m \ge 2$. □

We need an information about the decay rate of $\varphi(t)$, of the form $\varphi(t) = \mathrm{O}(|t|^{-\delta})$ for some $\delta > 0$ when $|t| \to \infty$. To this end, we shall use the following Gronwall-type technical Lemma of [13] (see also Lemma 3.2 in [14]).

Lemma 7.8 [13]. *Let $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ be a bounded function and let B be a positive random variable such that for some constants $p \in]0,1[$, $a > 0$, $C \ge 0$ and for all $r > 0$,*

$$\psi(r) \le p\mathbb{E}\psi(Br) + Cr^{-a}.$$

If $p\mathbb{E}(B^{-a}) < 1$, then $\psi(r) = \mathrm{O}(r^{-a})$ as $r \to \infty$.

This is Lemma 4.1 of [13]. It can be proved as follows.

Let $\{B_i\}$ be independent copies of B. Then by induction, for all $n \ge 1$ and all $r > 0$,

$$\psi(r) \le p^n\mathbb{E}\psi(B_1\cdots B_n r) + Cr^{-a}\bigl[1 + p\mathbb{E}\bigl(B^{-a}\bigr) + \cdots + \bigl(p\mathbb{E}\bigl(B^{-a}\bigr)\bigr)^{n-1}\bigr].$$

Letting $n \to \infty$ we see that for all $r > 0$,

$$\psi(r) \le Cr^{-a}/\bigl[1 - p\mathbb{E}\bigl(B^{-a}\bigr)\bigr].$$

Lemma 7.9. *For all $a \in]0, \frac{1}{\Re(\lambda)}[$, as $r \to \infty$,*

$$\psi(r) = \mathrm{O}\bigl(r^{-a}\bigr).$$

Proof. We have already seen from the distributional equation (5.12) that

$$\psi(r) \le \mathbb{E}\bigl(\psi^m\bigl(r|A|\bigr)\bigr),$$

where $A = \mathrm{e}^{-\lambda T}$. By Lemma 7.7, for any $\varepsilon > 0$, there is some $r_\varepsilon > 0$ such that $\forall r \ge r_\varepsilon$, $\psi(r) \le \varepsilon$. So

$$\psi(r) \le \varepsilon^{m-1}\mathbb{E}\psi\bigl(r|A|\bigr) + \mathbb{P}\bigl(r|A| \le r_\varepsilon\bigr).$$

Therefore by Markov inequality, for $a \in]0, \frac{1}{\Re(\lambda)}[$,

$$\psi(r) \le \varepsilon^{m-1}\mathbb{E}\psi\bigl(r|A|\bigr) + r^{-a}(r_\varepsilon)^a\mathbb{E}\bigl(|A|^{-a}\bigr).$$

By (5.4), $\mathbb{E}(|A|^{-a}) = (m-1)B(1 - a\Re(\lambda), m-1) < \infty$. Taking $\varepsilon > 0$ small enough such that $\varepsilon^{m-1}\mathbb{E}(|A|^{-a}) < 1$, we see that the desired result follows from Lemma 7.8 and the preceding inequality on $\psi(r)$. □

We can now finish the proof of Theorem 7.1.

Proof of Theorem 7.1. Part (i) of the theorem comes from two facts as shown in the following.

On the one hand, by Lemma 7.4, as soon as $z \in \mathbb{C}$ is a point in the support of Z, we have $D(0, |z|) \subseteq \mathrm{Supp}(Z)$, where $D(0, |z|)$ denotes the open disc with center 0 and radius $|z|$.

On the other hand, the support of Z is unbounded. Indeed, as in (7.2), at the beginning of the proof of Lemma 7.4, as soon as $z \in \mathbb{C} \setminus \{0\}$ is a point in the support of Z, for any $t > 0$ and for any $n \in \mathbb{N}$, $m^n \mathrm{e}^{-\lambda t} z$ is in the support of Z.

For Part (ii), notice that by Lemma 7.9, for all $a \in]0, \frac{1}{\Re(\lambda)}[$,

$$\varphi(t) = \mathrm{O}\big(t^{-a}\big) \quad \text{as } |t| \to \infty. \tag{7.6}$$

Since $\mathbb{E}Z \neq 0$, by Eq. (5.12) we obtain $m\mathbb{E}\mathrm{e}^{-\lambda T} = 1$, hence $m\mathbb{E}|\mathrm{e}^{-\lambda T}| = m\mathbb{E}\mathrm{e}^{-\Re(\lambda)T} > 1$ as soon as $\Im(\lambda) \neq 0$. Notice that if $\Im(\lambda) = 0$, then $\lambda = 1$ by the equation $m\mathbb{E}\mathrm{e}^{-\lambda T} = 1$. So the hypotheses $\lambda \neq 1$ and $\mathbb{E}Z \neq 0$ imply that $\Im(\lambda) \neq 0$ and $\Re(\lambda) < 1$ (cf. (5.6)). It follows that (7.6) holds for some $a > 1$, so that the Fourier transform φ of Z is in L^2. Therefore by the inversion formula of Fourier–Plancherel transform, the distribution of Z has a density in L^2 with respect to the Lebesgue measure on $\mathbb{C}$. This ends the proof of Theorem 7.1. □

Remark 7.10. *In fact we have the following more general result. Let λ be a complex number with $\sigma := \Re(\lambda) > 0$, $\tau := \Im(\lambda) \neq 0$ and satisfying the arithmetical condition*:

$$\frac{\pi\sigma}{\tau \log m} \notin \mathbb{Q}$$

and let Z be a nontrivial solution of Eq. (5.11) (*with or without first moment*). *Then the distribution of Z is absolutely continuous with respect to the Lebesgue measure on $\mathbb{C}$, and its support is the whole complex plane $\mathbb{C}$.*

To see the conclusions of Remark 7.10, we can argue as follows. In the general case where the expectation of Z may not exist, Lemma 7.4 still holds thanks to the arithmetical condition. The rest of the proof is the same, except at the end, where $\Re(\lambda) > 1$ is no more ensured. Nevertheless we have an additional argument by iteration. Iterating the distributional equation (5.11), we obtain for $n \geq 1$,

$$Z \stackrel{\mathcal{L}}{=} \sum_{u_1 \cdots u_n \in \{1,\ldots,m\}^n} A A_{u_1} \cdots A_{u_1 \cdots u_{n-1}} Z^{(u_1 \cdots u_n)},$$

where $A = \mathrm{e}^{-\lambda T}$, A_u are independent copies of A (indexed by finite sequences of integers u), $Z^{(u)}$ are independent copies of Z, the two families $\{A_u\}$ and $\{Z^{(u)}\}$ are also independent of each other; by convention, $A_{u_1} \cdots A_{u_1 \cdots u_{n-1}}$ is taken to be 1 when $n = 1$. It is convenient to rewrite this equation in the form

$$Z \stackrel{\mathcal{L}}{=} \sum_{j=1}^{m^n} Y_j Z^{(j)}, \tag{7.7}$$

where $Z^{(j)}$ are independent copies of Z which are also independent of $\{Y_j\}$. For fixed $y = (y_j : 1 \leq j \leq m^n)$ with $\prod_{j=1}^{m^n} y_j \neq 0$, by Lemma 7.9, for $a \in]0, 1/\Re(\lambda)[$ and some constant $c > 0$,

$$\left| \mathbb{E} \exp\left(\mathrm{i} \left\langle t, \sum_{j=1}^{m^n} y_j Z^{(j)} \right\rangle \right) \right| \leq \prod_{j=1}^{m^n} c|t y_j|^{-a} = C(y)|t|^{-m^n a},$$

where $C(y) = \prod_{j=1}^{m^n} c|y_j|^{-a} > 0$ does not depend on t. Let $n \geq 1$ be large enough such that $m^n a > 1$. Then $\sum_{j=1}^{m^n} y_j Z^{(j)}$ is absolute continuous (with respect to the Lebesgue measure on $\mathbb{C}$) as its Fourier transform is square integrable on $\mathbb{C}$. This implies that for each Borel set B of $\mathbb{C}$ with Lebesgue measure 0, we have

$$\mathbb{P}\left(\sum_{j=1}^{m^n} y_j Z^{(j)} \in B \right) = 0.$$

It follows from Eq. 7.7 (by conditioning on (Y_j)) that $\mathbb{P}(Z \in B) = 0$.

8. Exponential moments and Laplace series

In this section we consider a solution Z of Eq. (5.11) and we show that its exponential moments exist in a neighborhood of 0, so that the moment exponential generating series of Z defines an analytic function in a neighbourhood of the origin. We show that this function satisfies a very simple differential equation.

Theorem 8.1. *Let $\lambda \in \mathbb{C}$ be a root of the characteristic polynomial (2.3) with $\Re(\lambda) > 1/2$ and let Z be a solution of Eq. (5.11). There exist some constants $C > 0$ and $\varepsilon > 0$ such that for all $t \in \mathbb{C}$ with $|t| \le \varepsilon$,*

$$\mathbb{E}e^{\langle t,Z\rangle} \le e^{\Re(t)+C|t|^2} \quad \textit{and} \quad \mathbb{E}e^{|tZ|} \le 4e^{|t|+2C|t|^2}. \tag{8.1}$$

To prove this theorem, we use Mandelbrot's cascades in the complex setting (see Barral et al. [2] for independent interest about complex Mandelbrot's cascades). We use the notation $A = e^{-\lambda T}$. Then $m\mathbb{E}A = 1$ because λ is a root of the characteristic polynomial (2.3) and $m\mathbb{E}|A|^2 < 1$ because $\Re(\lambda) > 1/2$ (see (5.6)). Let $A_u, u \in U$ be independent copies of A, indexed by all finite sequences of integers

$$u = u_1 \cdots u_n \in U := \bigcup_{n\ge 1} \{1, 2, \dots, m\}^n$$

and set $Y_0 = 1$, $Y_1 = mA$ and for $n \ge 2$,

$$Y_n = \sum_{u_1\cdots u_{n-1} \in \{1,\dots,m\}^{n-1}} mAA_{u_1}A_{u_1u_2}\cdots A_{u_1\cdots u_{n-1}}.$$

As $m\mathbb{E}A = 1$, $(Y_n)_n$ is a martingale. This martingale has been studied by many authors in the real random variable case, especially in the context of Mandelbrot's cascades, see for example [14] and the references therein. It can be easily seen that

$$Y_{n+1} = A\sum_{i=1}^{m} Y_{n,i}, \tag{8.2}$$

where $Y_{n,i}$ for $1 \le i \le m$ are independent of each other and independent of A and each has the same distribution as Y_n. Therefore for $n \ge 1$, Y_n is square-integrable and

$$\operatorname{Var} Y_{n+1} = \bigl(\mathbb{E}|A|^2m^2 - 1\bigr) + m\mathbb{E}|A|^2 \operatorname{Var} Y_n,$$

where $\operatorname{Var} X = \mathbb{E}(|X - \mathbb{E}X|^2)$ denotes the variance of X. Since $m\mathbb{E}|A|^2 < 1$, the martingale $(Y_n)_n$ is bounded in L^2, so that the following result holds.

Lemma 8.2. *Let λ be a root of the characteristic polynomial (2.3) with $\Re(\lambda) > 1/2$. Then, when $n \to +\infty$,*

$$Y_n \to Y_\infty \quad \textit{a.s. and in } L^2,$$

where Y_∞ is a (complex-valued) random variable with variance

$$\operatorname{Var}(Y_\infty) = \frac{\mathbb{E}|A|^2m^2 - 1}{1 - m\mathbb{E}|A|^2}.$$

Notice that, passing to the limit in (8.2) gives a new proof of the existence of a solution Z of Eq. (5.11) with $\mathbb{E}Z = 1$ and finite second moment whenever $\Re(\lambda) > 1/2$. From Section 6, we have the uniqueness of solution of this equation so that Theorem 8.1 is proved as soon as it holds for Y_∞.

Lemma 8.3. *Under the condition of Lemma 8.2, there exist some constants $C > 0$ and $\varepsilon > 0$ such that for all $t \in \mathbb{C}$ with $|t| \le \varepsilon$, we have*

$$\mathbb{E}\mathrm{e}^{\langle t, Y_\infty \rangle} \le \mathrm{e}^{\Re(t)+C|t|^2}. \tag{8.3}$$

Proof. As in [20] and [15] (where a similar problem for real random variables was considered), we use an induction argument. Notice that by Eq. (8.2), writing

$$\varphi_n(t) := \mathbb{E}\mathrm{e}^{\langle t, Y_n \rangle}, \quad t \in \mathbb{C}, n \ge 0,$$

we have

$$\varphi_{n+1}(t) = \mathbb{E}\varphi_n^m(\overline{A}t), \quad t \in \mathbb{C}. \tag{8.4}$$

We shall prove that there exist some constants $C > 0$ and $\varepsilon > 0$ such that for all $n \ge 0$ and all $t \in \mathbb{C}$ with $|t| \le \varepsilon$, we have

$$\varphi_n(t) \le \mathrm{e}^{\Re(t)+C|t|^2}. \tag{8.5}$$

Let us prove (8.5) by induction. The inequality holds clearly for $n = 0$ since $\varphi_0(t) = \mathrm{e}^{\Re(t)}$. Assume that it holds for some $n \ge 0$ and all $t \in \mathbb{C}$ with $|t| \le \varepsilon$. Then writing $A = A_1 + \mathrm{i}A_2$ ($A_i \in \mathbb{R}$), using $|A| \le 1$ and Eq. (8.4), we have for $t = t_1 + \mathrm{i}t_2$ ($t_i \in \mathbb{R}$) with $|t| \le \varepsilon$,

$$\begin{aligned}\varphi_{n+1}(t) &\le \mathbb{E}\exp\bigl\{m\bigl(A_1t_1 + A_2t_2 + C|A|^2\bigl(t_1^2 + t_2^2\bigr)\bigr)\bigr\} \\ &= \mathrm{e}^{\Re(t)+C|t|^2} g(t_1, t_2),\end{aligned} \tag{8.6}$$

where $g(t_1, t_2) = \mathbb{E}\mathrm{e}^{h(t_1,t_2)}$ with

$$h(t_1, t_2) = (mA_1 - 1)t_1 + mA_2t_2 + C\bigl(m|A|^2 - 1\bigr)\bigl(t_1^2 + t_2^2\bigr).$$

Notice that $g(0, 0) = 1$. It remains to prove that $(0, 0)$ is a local maximum of g. Clearly,

$$\begin{aligned}\frac{\partial g}{\partial t_i} &= \mathbb{E}\mathrm{e}^h\left[\frac{\partial h}{\partial t_i}\right], \quad i = 1, 2, \\ \frac{\partial^2 g}{\partial t_i^2} &= \mathbb{E}\mathrm{e}^h\left[\left(\frac{\partial h}{\partial t_i}\right)^2 + \frac{\partial^2 h}{\partial t_i^2}\right], \quad i = 1, 2, \\ \frac{\partial^2 g}{\partial t_1 \partial t_2} &= \mathbb{E}\mathrm{e}^h\left[\frac{\partial h}{\partial t_1}\frac{\partial h}{\partial t_1} + \frac{\partial^2 h}{\partial t_1 \partial t_2}\right].\end{aligned}$$

Notice that, a.s.

$$\begin{aligned}&\frac{\partial h}{\partial t_1}(0, 0) = (mA_1 - 1), \qquad \frac{\partial h}{\partial t_2}(0, 0) = mA_2, \\ &\frac{\partial^2 h}{\partial t_1 \partial t_2}(0, 0) = 0, \qquad \frac{\partial^2 h}{\partial t_i^2}(0, 0) = 2C\bigl(m|A|^2 - 1\bigr), \quad i = 1, 2.\end{aligned}$$

Recall that $m\mathbb{E}A = 1$, so that $m\mathbb{E}A_1 = 1$ and $m\mathbb{E}A_2 = 0$; hence

$$\frac{\partial g}{\partial t_1}(0, 0) = \mathbb{E}(mA_1 - 1) = 0, \qquad \frac{\partial g}{\partial t_2}(0, 0) = \mathbb{E}(mA_2) = 0,$$

so that $(0,0)$ is a critical point of g. Moreover,

$$\frac{\partial^2 g}{\partial t_1^2}(0,0) = \mathbb{E}\big[(mA_1-1)^2 + 2C\big(m|A|^2-1\big)\big],$$

$$\frac{\partial^2 g}{\partial t_2^2}(0,0) = \mathbb{E}\big[(mA_2)^2 + 2C\big(m|A|^2-1\big)\big],$$

$$\frac{\partial^2 g}{\partial t_1 \partial t_2}(0,0) = \mathbb{E}(mA_1-1)(mA_2).$$

As $\mathbb{E}(m|A|^2-1) < 0$ (recall that $\Re(\lambda) > 1/2$), it follows that the Hessian matrix at $(0,0)$ is definite negative for $C > 0$ large enough which implies that $g(0,0)$ is a local maximum of g. So for $\varepsilon > 0$ small enough, $g(t_1,t_2) \le g(0,0) = 1$ for all $t = t_1 + \mathrm{i}t_2$ with $|t| \le \varepsilon$. Hence by (8.6), for such ε and C which do not depend on n, (8.5) holds for $n+1$. Therefore, by induction, it holds for all $n \ge 0$.

Letting $n \to \infty$ in (8.5), we see that inequality (8.3) holds by Fatou's lemma. □

Proof of Theorem 8.1. By the uniqueness of solution of Eq. (5.11), $\mathcal{L}(Z) = \mathcal{L}(Y_\infty)$. So by Lemma 8.3, there are some constants $C > 0$ and $\varepsilon > 0$ such that the first inequality of (8.1) holds. To show the second one, notice that $|t||\Re(Z)| + |t||\Im(Z)|$ takes one of the four values $\pm|t|\Re(Z) \pm |t|\Im(Z)$ (according to the signs of $\Re(Z)$ and $\Im(Z)$), so that a.s.

$$\begin{aligned} \mathrm{e}^{|tZ|} &\le \mathrm{e}^{|t||\Re(Z)|+|t||\Im(Z)|} \\ &\le \mathrm{e}^{|t|\Re(Z)+|t|\Im(Z)} + \mathrm{e}^{|t|\Re(Z)-|t|\Im(Z)} + \mathrm{e}^{-|t|\Re(Z)+|t|\Im(Z)} + \mathrm{e}^{-|t|\Re(Z)-|t|\Im(Z)}. \end{aligned}$$

Taking expectation in both sides, and noticing that $\pm|t|\Re(Z) \pm |t|\Im(Z) = \langle(\pm 1 \pm i)|t|, Z\rangle$, we see that the second inequality in (8.1) follows from the first one. □

Suppose that Z is any solution of Eq. (5.11) under the assumptions of Theorem 8.1. The second inequality (8.1) shows that the exponential generating series of absolute moments of Z has a positive radius of convergence so that the formal Laplace series

$$L(z) := \sum_{p \ge 0} \frac{\mathbb{E}Z^p}{p!} z^p$$

defines an analytic function in a neighbourhood of the origin. One can also write $L(z) = \mathbb{E}\mathrm{e}^{zZ}$ when $|z|$ is sufficiently small.

Let's come back to the dislocation equations (5.7) satisfied by the limit variables $W_1, \ldots, W_m$. These variables admit finite (absolute) moments at any order. For any $k \in \{1, \ldots, m\}$, let L_k be the formal Laplace series defined by

$$L_k(z) := \sum_{p \ge 0} \frac{\mathbb{E}(W_k^p)}{p!} z^p.$$

The dislocation equations (5.7) imply recursive relations on W_k's moments. Developing these relations with the multinomial formula implies that L_k satisfy the formal differential system

$$\begin{cases} \forall k \in \{1, \ldots, m-2\}, \quad L_k(z) + \frac{\lambda_2}{k} z L_k'(z) = L_{k+1}(z), \\ L_{m-1}(z) + \frac{\lambda_2}{m-1} z L_{m-1}'(z) = \big(L_1(z)\big)^m, \end{cases} \tag{8.7}$$

with boundary conditions

$$\begin{cases} L_k(0) = 1, \quad 1 \le k \le m-1, \\ L_k'(0) = \mathbb{E}(W_k) = u_2\big(X_k(0)\big) = \binom{\lambda_2+k-1}{k-1}. \end{cases}$$

Since W_1 satisfies the assumptions of Theorem 8.1, the series L_1 has a positive radius of convergence as shown above. Therefore, the same holds for all L_k because of the system (8.7) so that the L_k define, near the origin, analytic functions related by (8.7).

Let ρ be any complex $(m-1)$th root of $(-1)^m(m-1)!$. For any $k \in \{1, \dots, m\}$, define

$$G_k(z) := (-1)^k \rho (k-1)! \frac{L_k(z^{-\lambda_2})}{z^k},$$

where $z^{-\lambda_2}$ denotes any determination of the logarithm. For sufficiently large $|z|$, this formula defines an analytic function on a slit plane. Reporting in formula (8.7) shows that the functions G_k satisfy the simple differential system

$$\begin{cases} \forall k \in \{1, \dots, m-2\}, \quad G'_k = G_{k+1}, \\ G'_{m-1} = G_1^m. \end{cases}$$

In particular, G_1 is solution of the differential equation $y^{(m-1)} = y^m$. We sum up these results in the following statement.

Theorem 8.4. *Let W_1 be the complex-valued limit distribution for the multitype branching process of m-ary search trees as defined in Section 5.2. Then*:

(i) *the Laplace series $L_1(z) = \mathbb{E}(\mathrm{e}^{zW_1})$ has a positive radius of convergence*;
(ii) *for any determination of the logarithm*, *the function*

$$z \mapsto -\frac{\rho}{z} L_1\big(z^{-\lambda_2}\big),$$

is a solution of the differential equation

$$y^{(m-1)} = y^m. \tag{8.8}$$

Remark 8.5. *As can be straightforwardly checked*, *the function $y_\kappa(z) := \frac{\kappa}{1-z}$ is a solution of Eq*. (8.8) *when the complex number κ satisfies $\kappa^{m-1} = (m-1)!$. Nonetheless*, *G_1 is not a function of this form.*

Indeed, *since $L_1(w) = 1 + w + \mathrm{o}(w)$ in a neighbourhood of the origin*, *G_1 admits the expansion*

$$G_1(z) = -\frac{\rho}{z} - \frac{\rho}{z^{1+\lambda_2}} + \mathrm{o}\left(\frac{1}{z^{1+\lambda_2}}\right),$$

while y_κ satisfies

$$\frac{\kappa}{1-z} = -\frac{\kappa}{z} - \frac{\kappa}{z^2} + \mathrm{o}\left(\frac{1}{z^2}\right).$$

*One conclude*s *by uniqueness of* (*complex*) *power expansions*, *because $\lambda_2 \neq 1$.*

Acknowledgments

The authors owe much to Philippe Flajolet, especially some crucial arguments and many enthusiastic discussions. They are also very grateful to an associate editor and to the referee for helpful comments and remarks. The work has been partially supported by the GDR 3475 and the National Natural Science Foundation of China, Grants no. 11101039 and no. 11171044.

References

[1] K. B. Athreya and P. Ney. *Branching Processes*. Springer, New York, 1972. MR0373040

[2] J. Barral, X. Jin and B. Mandelbrot. Convergence of complex multiplicative cascades. *Ann. Appl. Probab.* **20** (2010) 1219–1252. MR2676938

[3] J. Bertoin. *Random Fragmentation and Coagulation Processes*. *Cambridge Studies in Advanced Mathematics*. Cambridge Univ. Press, Cambridge, 2006. MR2253162

[4] B. Chauvin and N. Pouyanne. m-ary search trees when $m > 26$: A strong asymptotics for the space requirements. *Random Structures Algorithms* **24** (2004) 133–154. MR2035872

[5] B. Chauvin, N. Pouyanne and R. Sahnoun. Limit distributions for large Pólya urns. *Ann. Appl. Probab.* 21 (2011) 1–32. MR2759195

[6] H.-H. Chern and H.-K. Hwang. Phase changes in random m-ary search trees and generalized quicksort. *Random Structures Algorithms* **19** (2001) 316–358. MR1871558

[7] R. M. Dudley. *Real Analysis and Probability*. Cambridge Univ. Press, Cambridge, 2002. MR1932358

[8] R. Durrett and T. Liggett. Fixed points of the smoothing transformation. *Z. Wahrsch. verw. Gebiete* **64** (1983) 275–301. MR0716487

[9] J. A. Fill and N. Kapur. The space requirement of m-ary search trees: Distributional asymptotics for $m \geq 27$. In *Proceedings of the 7th Iranian Conference, Tehran*, 2004. Available at ArXiv:math.PR/0405144.

[10] Y. Guivarc'h. Sur une extension de la notion de loi semi-stable. *Ann. Inst. Henri Poincaré Probab. Stat.* **26** (1990) 261–285. MR1063751

[11] W. N. Hudson, J. A. Veeh and D. C. Weiner. Moments of distributions attracted to operator-stable laws. *J. Multivariate Anal.* **24** (1988) 1–10. MR0925125

[12] S. Janson. Functional limit theorem for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.* **110** (2004) 177–245. MR2040966

[13] Q. Liu. Asymptotic properties of supercritical age-dependent branching processes and homogeneous branching random walks. *Stochastic Process. Appl.* **82** (1999) 61–87. MR1695070

[14] Q. Liu. Asymptotic properties and absolute continuity of laws stable by random weighted mean. *Stochastic Process. Appl.* **95** (2001) 83–107. MR1847093

[15] Q. Liu and A. Rouault. Limit theorems for Mandelbrot's multiplicative cascades. *Ann. Appl. Probab.* **10** (2000) 218–239. MR1765209

[16] H. M. Mahmoud. *Evolution of Random Search Trees*. Wiley, New York, 1992. MR1140708

[17] J. R. Norris. *Markov Chains. Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge Univ. Press, Cambridge, 1997. MR1600720

[18] N. Pouyanne. Classification of large Pólya-Eggenberger urns with regard to their asymptotics. In *2005 International Conference on Analysis of Algorithms. Discrete Math. Theor. Comput. Sci. Proc., AD. Assoc. Discrete Math. Theor. Comput. Sci., Nancy* 275–285, 2005 (electronic). MR2193125

[19] N. Pouyanne. An algebraic approach to Pólya processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** (2008) 293–323. MR2446325

[20] U. Rösler. A fixed point theorem for distributions. *Stochastic Process. Appl.* **42** (1992) 195–214. MR1176497

[21] U. Rösler and L. Rüschendorf. The contraction method for recursive algorithms. *Algorithmica* **29** (2001) 3–33. MR1887296

[22] M. Sharpe. Operator-stable probability distribution on vector groups. *Trans. Amer. Math. Soc.* **136** (1969) 51–65. MR0238365