ASYMPTOTICS OF FLUCTUATIONS IN CRUMP–MODE–JAGERS PROCESSES: THE LATTICE CASE

SVANTE JANSON

Abstract. Consider a supercritical Crump–Mode–Jagers process such that all births are at integer times (the lattice case). Let $\widehat{\mu}(z)$ be the generating function of the intensity of the offspring process, and consider the complex roots of $\widehat{\mu}(z) = 1$. The smallest (in absolute value) such root is $e^{-\alpha}$, where $\alpha > 0$ is the Malthusian parameter; let γ_* be the second smallest absolute value of a root.

We show, assuming some technical conditions, that there are three cases:

(i) if $\gamma_* > e^{-\alpha/2}$, then the second-order fluctuations of the age distribution are asymptotically normal;

(ii) if $\gamma_* = e^{-\alpha/2}$, then the fluctuations are still asymptotically normal, but with a larger order of the variance;

(iii) if $\gamma_* < e^{-\alpha/2}$, then the fluctuations are even larger, but will oscillate and (except in degenerate cases) not converge in distribution.

This trichotomy is similar to what has been seen in related situations, e.g. for some other branching processes, and for Pólya urns.

The results lead to a symbolic calculus describing the limits. The results extends to populations counted by a random characteristic.

1. Introduction

Consider a Crump–Mode–Jagers branching process, starting with a single individual born at time 0, where an individual has $N \leqslant \infty$ children born at the times when the parent has age $\xi_1 \leqslant \xi_2 \leqslant \dots$. Here N and $(\xi_i)_i$ are random, and different individuals have independent copies of these random variables. Technically, it is convenient to regard $\{\xi_i\}_1^N$ as a point process Ξ on $[0, \infty)$, and give each individual x an independent copy Ξ_x of Ξ. For further details, see e.g. Jagers [7].

We consider the supercritical case, when the population grows to infinity (at least with positive probability). As is well-known, under weak assumptions, the population grows exponentially, like $e^{\alpha t}$ for some constant $\alpha > 0$ known as the *Malthusian parameter*, see e.g. [7, Theorems (6.3.3) and (6.8.1)]; in particular, the population size properly normalized converges

Date: 29 May, 2018.

2010 *Mathematics Subject Classification.* 60J80; 60F05.

Key words and phrases. Crump–Mode–Jagers processes; age distribution.

Partly supported by the Knut and Alice Wallenberg Foundation.

to some positive random variable, and the age distribution stabilizes. Our purpose is to study the second-order fluctuations of the age distribution, or more generally, of the population counted with a random characteristic.

We consider in this paper the lattice case; we thus assume that the ξ_i are integer-valued and thus all births occur at integer times a.s., but there is no $d > 1$ such that all birth times a.s. are divisible by d.

Our setting can, for example, be considered as a model for the (female) population of some animal that is fertile several years and gets one or several children once every year, with the numbers of children different years random and dependent.

Our main results (Theorems 2.1–2.3) show that there are three different cases depending on properties of the intensity measure $\mathbb{E}\,\Xi$ of the offspring process: in one case fluctuations are, after proper normalization, asymptotically normal, with only a short-range dependence between different times; in another case, there is a long-range dependence and, again after proper normalization (different this time), the fluctuations are a.s. approximated by oscillating (almost periodic) random functions of $\log n$, which furthermore essentially are determined by the initial phase of the branching process, and presumably non-normal; the third case is an intermediate boundary case. See Section 2 for precise results.

A similar trichotomy has been found in several related situations. Similar results are proved for multi-type Markov branching processes by Asmussen and Hering [1, Section VIII.3]. Their type space may be very general, so this setting includes also the single-type non-Markov case studied here (also in the non-lattice case [1, Section VIII.12]), since a Crump–Mode–Jagers branching process may be seen as a Markov process where the type of an individual is its entire life history until present. However, this will in general be a large type space, and the assumptions of [1] will in general not be satisfied; in particular, their "condition (M)" [1, p. 156] is typically not satisfied, by the same argument as in [1, p. 173] for a related situation. Hence, we can not obtain our results directly from the closely related results in [1], although there is an overlap in some special cases (for example the Galton–Watson case in Example 2.5).

Another related situation is given by multi-colour Pólya urn processes, see e.g. [9] (which uses methods and results from branching process theory). The same trichotomy appears there too, with a criterion formulated in terms of eigenvalues of a matrix that can be seen as the (expected) "offspring matrix" in that setting.

It would be interesting to find more general theorems that would include these different but obviously related results together.

Remark 1.1. Our setup includes the Galton–Watson case, where all births occur when the mother has age 1 (Example 2.5), but this case is much simpler than the general case and can be treated by simpler methods; see

Jagers [7, Section 2.10], where results closely related to the ones below are given.

Remark 1.2. It would be very interesting to extend the results to the perhaps more interesting non-lattice case; we expect similar results (under suitable assumptions), but this case seems to present new technical challenges, and we leave this as an open problem.

2. Assumptions and main result

Let $\mu := \mathbb{E}\,\Xi$ be the intensity measure of the offspring process; thus $\mu := \sum_{k=0}^{\infty} \mu_k \delta_k$, where μ_k is the expected number of children that an individual bears at age k (and δ_k is the Dirac delta, i.e., a point mass at k). Let $N_k := \Xi\{k\}$ be the number of children born to an individual at age k. Thus $N = \sum_{k=1}^{\infty} N_k$ and $\mu_k = \mathbb{E}\,N_k$.

We make the following standing assumptions, valid throughout the paper. The first assumption (supercriticality) is essential; otherwise there is no asymptotic behaviour to analyse. The assumptions (A2)–(A4) are simplifying and convenient but presumably not essential. (For (A4), this is shown in Example 11.4.)

- (A1) The process is supercritical, i.e., $\mu([0,\infty]) = \sum_{k=0}^{\infty} \mu_k = \mathbb{E}\,N > 1$.
- (A2) No children are born instantaneously, i.e., $\mu_0 = 0$.
- (A3) $N \geqslant 1$ a.s. Thus the process a.s. survives.
- (A4) There are no deaths.

Define, for all complex z such that either $z \geqslant 0$ or the sums or expectations below converge absolutely,

$$\widehat{\mu}(z) := \sum_{k=0}^{\infty} \mu_k z^k = \sum_{k=0}^{\infty} \mathbb{E}[N_k] z^k = \mathbb{E} \sum_{i=1}^{N} z^{\xi_i} \tag{2.1}$$

and the complex-valued random variable

$$\widehat{\Xi}(z) := \int_0^{\infty} z^x \,\mathrm{d}\Xi(x) = \sum_{i=1}^{N} z^{\xi_i} = \sum_{k=0}^{\infty} N_k z^k. \tag{2.2}$$

Thus $\widehat{\mu}(z) = \mathbb{E}\,\widehat{\Xi}(z)$.

We make two other standing assumptions:

- (A5) $\widehat{\mu}\big(m^{-1}\big) = 1$ for some $m > 1$.

Thus $\alpha := \log m$ satisfies $\sum_{k=1}^{\infty} \mu_k e^{-k\alpha} = \widehat{\mu}(e^{-\alpha}) = 1$, so α is the Malthusian parameter, and the population grows roughly with a factor $e^{\alpha} = m$ for each generation (see e.g. (2.7) and (2.8) below).

- (A6) $\mathbb{E}[\widehat{\Xi}(r)^2] < \infty$ for some $r > m^{-1/2}$.

We fix in the sequel some $r > m^{-1/2}$ satisfying (A6). We assume for convenience $r \leqslant 1$. Note that (A6) implies

$$\widehat{\mu}(r) = \mathbb{E}\,\widehat{\Xi}(r) < \infty. \tag{2.3}$$

Hence $\widehat{\mu}(z)$ and $\widehat{\Xi}(z)$ are defined, and analytic, at least for $|z| \leqslant r$. Since $\widehat{\mu}(z)$ is a strictly increasing function on $[0, \infty)$, m^{-1} in (A5) is the unique positive root of $\widehat{\mu}(z) = 1$. However, $\widehat{\mu}(z) = 1$ may have other complex roots; we shall see that the asymptotic behaviour of the fluctuations depends crucially on the position of these roots. We define, with $D_r := \{|z| < r\}$,

$$\Gamma := \{z \in D_r : \widehat{\mu}(z) = 1\}, \qquad \Gamma_* := \Gamma \setminus \{m^{-1}\}, \tag{2.4}$$

$$\gamma_* := \inf\{|z| : z \in \Gamma_*\}, \tag{2.5}$$

$$\Gamma_{**} := \{z \in \Gamma_* : |z| = \gamma_*\}, \tag{2.6}$$

with $\gamma_* = \infty$ if $\Gamma_* = \emptyset$. (These sets may depend on the choice of r, but for our purposes this does not matter. Recall that we assume $r > m^{-1/2}$.) Since $\widehat{\mu}(z)$ is analytic, Γ is discrete and thus, if $\gamma_* < \infty$, then Γ_{**} is a finite non-empty set which we write as $\{\gamma_1, \dots, \gamma_q\}$.

Let Z_n be the total number of individuals at time n. (Which by (A2) equals the number of individuals born up to time n.) We define Z_n for all integers n by letting $Z_n := 0$ for $n < 0$. By assumption, $Z_0 = 1$. It is well-known that the number of individuals Z_n grows asymptotically like m^n as $n \to \infty$. For example, see e.g. [7, Theorem (6.3.3)] (and remember that we here consider the lattice case),

$$\mathbb{E}\, Z_n \sim c_1 m^n, \qquad \text{as } n \to \infty, \tag{2.7}$$

with some $c_1 > 0$. Moreover, if $\mathbb{E}[\widehat{\Xi}(m^{-1}) \log \widehat{\Xi}(m^{-1})] < \infty$, and in particular if $\mathbb{E}[\widehat{\Xi}(m^{-1})^2] < \infty$, which follows from our assumption (A6), then

$$Z_n/m^n \xrightarrow{\text{a.s.}} \mathcal{Z}, \qquad \text{as } n \to \infty, \tag{2.8}$$

for some random variable $\mathcal{Z} > 0$, see e.g. Nerman [10]. In particular, it follows that for any fixed $k \geqslant 1$

$$Z_{n-k}/Z_n \xrightarrow{\text{a.s.}} m^{-k}. \tag{2.9}$$

The number of individuals of age $\geqslant k$ at time n is Z_{n-k}. For large n, we expect this to be roughly $m^{-k} Z_n$, see (2.9), and to study the fluctuations, we define

$$X_{n,k} := Z_{n-k} - m^{-k} Z_n, \qquad k = 0, 1, \dots \tag{2.10}$$

Note that $X_{n,0} = 0$.

We state our main results as three separate theorems, treating the cases $\gamma_* > m^{-1/2}$, $\gamma_* = m^{-1/2}$ and $\gamma_* < m^{-1/2}$ separately. In particular, note that Theorems 2.1–2.2 yield asymptotic normality of $X_{n,k}$ when $\gamma_* \geqslant m^{-1/2}$. Proofs are given in later sections. The results are extended to random characteristics in Section 11.

By the assumption (A6) and (2.2), $\mathbb{E}\, N_k^2 < \infty$ for every $k \geqslant 1$. Define, for $j, k \geqslant 1$,

$$\sigma_{jk} := \operatorname{Cov}(N_j, N_k) \tag{2.11}$$

and, at least for $|z| < r$,

$$\Sigma(z) := \sum_{i,j} \sigma_{ij} z^i \bar{z}^j = \operatorname{Cov}\Bigl(\sum_i N_i z^i, \sum_j N_j \bar{z}^j\Bigr) = \mathbb{E}\bigl|\widehat{\Xi}(z) - \widehat{\mu}(z)\bigr|^2. \tag{2.12}$$

Let, for $R > 0$, ℓ^2_R be the Hilbert space of infinite vectors

$$\ell^2_R := \Bigl\{(a_k)_{k=0}^\infty : \|(a_k)_0^\infty\|^2_{\ell^2_R} := \sum_{k=0}^\infty R^{2k}|a_k|^2 < \infty\Bigr\}. \tag{2.13}$$

(We often simplify the notation and denote a vector in ℓ^2_R by $(a_k)_k$.)

We begin with the case $\gamma_* > m^{-1/2}$, which by (2.4)–(2.5) is equivalent to:

(B) $\widehat{\mu}(z) \neq 1$ for all complex $|z| \leqslant m^{-1/2}$ except possibly $z = m^{-1}$.

Theorem 2.1. *Assume* (A1)–(A6) *and* (B), *i.e.,* $\gamma_* > m^{-1/2}$. *Then, as* $n \to \infty$,

$$X_{n,k}/\sqrt{Z_n} \xrightarrow{\mathrm{d}} \zeta_k, \tag{2.14}$$

jointly for all $k \geqslant 0$, *for some jointly normal random variables* ζ_k *with mean* $\mathbb{E}\,\zeta_k = 0$ *and covariance matrix given by, for any finite sequence* $a_0, \dots, a_K$ *of real numbers,*

$$\operatorname{Var}\Bigl(\sum_k a_k\zeta_k\Bigr) = \frac{m-1}{m}\oint_{|z|=m^{-1/2}} \frac{\bigl|\sum_k a_k z^k - \sum_k a_k m^{-k}\bigr|^2}{|1-z|^2\,|1-\widehat{\mu}(z)|^2}\Sigma(z)\frac{|\mathrm{d}z|}{2\pi m^{-1/2}}. \tag{2.15}$$

The convergence (2.14) *holds also in the stronger sense that* $(Z_n^{-1/2}X_{n,k})_k \xrightarrow{\mathrm{d}} (\zeta_k)_k$ *in the Hilbert space* ℓ^2_R, *for any* $R < m^{1/2}$. *The limit variables* ζ_k *are non-degenerate unless* Ξ *is deterministic, i.e.,* $N_k = \mu_k$ *a.s. for each* $k \geqslant 0$.

Recall that joint convergence of an infinite number of variables means joint convergence of any finite set. (This is convergence in the product space $\mathbb{R}^\infty$, see [2].) Note that trivially $\zeta_0 = 0$ (included for completeness).

The variance formula (2.15) can be interpreted as a stochastic calculus, where the limit variables are seen as stochastic integrals (in a general sense) of certain functions on the circle $|z| = m^{-1/2}$; these functions thus represent the random variables ζ_k, and therefore asymptotically $X_{n,k}$; moreover, they can be used for convenient calculations. See Section 10 for details.

We give two proofs of Theorem 2.1. The first, in Sections 4–5, is based on the elementary central limit theorem for sums of independent variables, together with some approximations. This proof is extended to random characteristics in Section 11. The second proof is given in Sections 6–7; it is based on a martingale central limit theorem. This proof easily adapts to give a proof of Theorem 2.2 below in Section 8.

We consider next the cases $\gamma_* \leqslant m^{1/2}$. Then $\Gamma_{**} = \{\gamma_1, \dots, \gamma_q\}$ is a non-empty finite set. For simplicity, we assume the condition

$$\widehat{\mu}'(\gamma) \neq 0, \qquad \gamma \in \Gamma_{**}, \tag{2.16}$$

i.e., that the points in Γ_{**} are simple roots of $\widehat{\mu}(z)=1$; the modifications in the case with a multiple root are left to the reader. (See Remark 3.8, and note the related results for Pólya urns in [9, Theorems 3.23–3.24] and [11, Theorems 3.5–3.6].)

Theorem 2.2. *Assume* (A1)–(A6) *and* $\gamma_*=m^{-1/2}$. *Suppose further that* (2.16) *holds. Then, as* $n\to\infty$,

$$X_{n,k}/\sqrt{nZ_n}\xrightarrow{\mathrm{d}}\zeta_k, \tag{2.17}$$

jointly for all $k\geqslant 0$, *for some jointly normal random variables* ζ_k *with mean* $\mathbb{E}\,\zeta_k=0$ *and covariance matrix given by, for any finite sequence* $a_0,\dots,a_K$ *of real numbers,*

$$\operatorname{Var}\Bigl(\sum_k a_k\zeta_k\Bigr)=(m-1)\sum_{p=1}^{q}\frac{\bigl|\sum_k a_k\gamma_p^k-\sum_k a_k m^{-k}\bigr|^2}{|1-\gamma_p|^2\,|\widehat{\mu}'(\gamma_p)|^2}\Sigma(\gamma_p). \tag{2.18}$$

Moreover, the convergence (2.14) *holds also in the Hilbert space* ℓ^2_R, *for any* $R<m^{1/2}$.

The limit variables ζ_k *are non-degenerate unless* $\widehat{\widehat{\Xi}}(\gamma_p)$ *is deterministic for each* $\gamma_p\in\Gamma_{**}$.

Theorem 2.3. *Assume* (A1)–(A6) *and* $\gamma_*<m^{-1/2}$. *Suppose further that* (2.16) *holds. Then there exist complex random variables* $U_1,\dots,U_q$ *and linearly independent vectors* $\vec{u}_i:=\bigl(\gamma_i^k-m^{-k}\bigr)_k$, $i=1,\dots,q$, *such that*

$$\gamma_*^n\vec{X}_n-\sum_{i=1}^{q}\bigl(\bar{\gamma}_i/|\gamma_i|\bigr)^nU_i\vec{u}_i\to 0 \tag{2.19}$$

a.s. and in $L^2(\ell^2_R)$, *for any* $R<m^{1/2}$. *Furthermore,* $\mathbb{E}\,U_i=0$, *and* U_i *is non-degenerate unless* $\widehat{\widehat{\Xi}}(\gamma_i)$ *is degenerate.*

Theorems 2.1–2.3 exhibit several differences between the cases $\gamma_*<m^{-1/2}$, $\gamma_*=m^{-1/2}$ and $\gamma_*>m^{-1/2}$; cf. the similar results for Pólya urns in e.g. [9, Theorems 3.22–3.24].

- The fluctuations $X_{n,k}$, for a fixed k, are asymptotically normal when $\gamma_*\geqslant m^{-1/2}$, but (presumably) not when $\gamma_*<m^{-1/2}$.
- The fluctuations are typically of order $Z_n^{1/2}\asymp m^{n/2}$ when $\gamma_*>m^{-1/2}$, slightly larger (by a power of n) when $\gamma_*=m^{-1/2}$, and of the much larger order γ_*^{-n} when $\gamma_*<m^{-1/2}$.
- When $\gamma_*<m^{-1/2}$, the fluctuations exhibit oscillations that are periodic or almost periodic (see [3]) in $\log n$. (Note that $\gamma_i/|\gamma_i|\neq 1$ in (2.19), since m^{-1} is the only positive root in Γ.)
- When $\gamma_*<m^{-1/2}$, there is the a.s. approximation result (2.19), implying both long-range dependence as $n\to\infty$, and that the asymptotic behaviour essentially is determined by what happens in the first few generations. In contrast, the limits in (2.14) and (2.17)

are mixing (see the proofs), i.e., the results holds also conditioned on the life histories of the first M individuals for any fixed M, and thus also conditioned on $Z_1, \dots, Z_K$ for any fixed K; hence, when $\gamma_* \geqslant m^{-1/2}$, the initial behaviour is eventually forgotten. Moreover for $\gamma_* > m^{-1/2}$, there is only a short-range dependence, see Example 10.1, while the case $\gamma_* = m^{-1/2}$ shows an intermediate "medium-range" dependence, see Subsection 10.2.

- When $\gamma_* > m^{-1/2}$, the limit random variables ζ_k in (2.14) are linearly independent, as a consequence of (2.15). When $\gamma_* \leqslant m^{-1/2}$, the limits in (2.17), or the components of the sum in (2.19), span a (typically) q-dimensional space of random variables, and any $q+1$ of them are linearly dependent; see also Section 10.

Remark 2.4. We consider above $X_{n,k}$ for $k \geqslant 0$, i.e., the age distribution of the population at time n. We can define $X_{n,k}$ by (2.10) also for $k < 0$; this means looking into the future and can be interpreted as predicting the future population. As shown in Section 10, (2.14)–(2.15) and (2.17)–(2.18) extend to all $k \in \mathbb{Z}$ (still jointly), and, similarly, taking the kth component in (2.19) yields a result that extends to all $k \in \mathbb{Z}$.

This enables us, for example, to obtain (by standard linear algebra) the best linear predictor of Z_{n+1} based on the observed $Z_n, \dots, Z_{n-K}$ for any fixed K.

Example 2.5 (Galton–Watson)**.** The simplest example is a Galton–Watson process, where all children are born in a single litter at age 1 of the parent, so $N_k = 0$ for $k \geqslant 2$. (But all individuals live forever in our setting. In the traditional setting, only the newborns are counted, i.e., $Z_n - Z_{n-1}$; the results are easily transferred to this version.) Then $N = N_1$, $m = \mu_1$ and $\widehat{\mu}(z) = mz$. Hence $\Gamma = \{m^{-1}\}$, $\Gamma_* = \emptyset$, and $\gamma_* = \infty > m^{-1/2}$. We assume $\mathbb{E}\, N^2 < \infty$; then (A6) holds for any r; we also assume $N \geqslant 1$ a.s. and $\mathbb{P}(N > 1) > 0$; then (A1)–(A6) and (B) hold.

Thus Theorem 2.1 applies. We obtain, for example, with $\sigma^2 := \operatorname{Var}(N) = \sigma_{11}$,

$$\begin{aligned}\operatorname{Var}\bigl(\zeta_1\bigr) &= \frac{m-1}{m} \oint_{|z|=m^{-1/2}} \frac{\bigl|z - m^{-1}\bigr|^2}{|1-z|^2\,|1-mz|^2} \sigma^2 |z|^2 \frac{|\mathrm{d}z|}{2\pi m^{-1/2}} \\ &= \sigma^2 \frac{m-1}{m^4} \oint_{|z|=m^{-1/2}} \frac{1}{|1-z|^2} \frac{|\mathrm{d}z|}{2\pi m^{-1/2}} \\ &= \sigma^2 m^{-3}. \end{aligned} \tag{2.20}$$

This can be shown directly in a much simpler way; see [7, Theorem (2.10.1)], which is essentially equivalent to our Theorem 2.1 in the Galton–Watson case (but without our assumption (A3)).

Example 2.6. Suppose that all children are born when the mother has age one or two, i.e., $N_k = 0$ for $k > 2$. Then $\widehat{\mu}(z) = \mu_1 z + \mu_2 z^2$, where by assumption $\mu_1+\mu_2>1$ and $\mu_1>0$. (A5) yields $m^2=\mu_1 m+\mu_2$, and thus

$$m=\frac{\mu_1+\sqrt{\mu_1^2+4\mu_2}}{2}. \tag{2.21}$$

The equation $\widehat{\mu}(z)=1$ has one other root, viz. γ_1 with

$$\gamma_1^{-1}=-\frac{\sqrt{\mu_1^2+4\mu_2}-\mu_1}{2}. \tag{2.22}$$

The condition (B) is thus equivalent to $|\gamma_1|>m^{-1/2}$, or $\gamma_1^{-2}<m$, which after some elementary algebra is equivalent to, for example,

$$\mu_1^3+3\mu_1\mu_2+\mu_2-\mu_2^2>0. \tag{2.23}$$

Thus, Theorem 2.1 applies when (2.23) holds, Theorem 2.2 when there is equality in (2.23), and Theorem 2.3 when the left-hand side of (2.23) is negative. (In this example, (2.16) is trivial.)

For a simple numerical example with $\gamma_*=m^{-1/2}$, take $\mu_1=2$ and $\mu_2=8$. Then (2.21)–(2.22) yield $m=4$ and $\gamma_1=-\frac{1}{2}$. We obtain by (2.18), for example,

$$X_{n,1}/\sqrt{nZ_n}\xrightarrow{\mathrm{d}}\zeta_1\sim N\Bigl(0,\frac{1}{768}\operatorname{Var}(N_2-2N_1)\Bigr). \tag{2.24}$$

Suppose now instead that (2.23) holds, so Theorem 2.1 applies. Let $\lambda:=\gamma_1^{-1}$ be given by (2.22). Then $1-\widehat{\mu}(z)=(1-mz)(1-\lambda z)$, and thus (2.15) yields, for example,

$$\begin{aligned}\operatorname{Var}\bigl(\zeta_1\bigr)&=\frac{m-1}{m}\oint_{|z|=m^{-1/2}}\frac{\bigl|z-m^{-1}\bigr|^2}{|1-z|^2\,|1-\widehat{\mu}(z)|^2}\Sigma(z)\frac{|\mathrm{d}z|}{2\pi m^{-1/2}}\\&=\frac{m-1}{m^3}\oint_{|z|=m^{-1/2}}\frac{\sigma_{11}|z|^2+\sigma_{12}(z+\bar z)|z|^2+\sigma_{22}|z|^4}{|1-z|^2\,|1-\lambda z|^2}\frac{|\mathrm{d}z|}{2\pi m^{-1/2}}.\end{aligned} \tag{2.25}$$

This integral can be evaluated by expanding $(1-z)^{-1}(1-\lambda z)^{-1}$ in a Taylor series; this yields after some calculations

$$\operatorname{Var}\bigl(\zeta_1\bigr)=\frac{(m+\lambda)(\sigma_{11}+\sigma_{22}/m)+2(1+\lambda)\sigma_{12}}{m^2(m-\lambda)(m-\lambda^2)}. \tag{2.26}$$

Remark 2.7. The limit in (2.14) is by Theorem 2.1 degenerate only when the entire process is, and thus each $X_{n,k}$ is degenerate. In contrast, the limit in (2.17) or the approximation in (2.19) may be degenerate even in other (special) situations. For example, let N_1 be non-degenerate with $\mathbb{E}\,N_1=2$, let $N_2:=2N_1+4$, and let $N_k:=0$ for $k>2$. Then $\mu_1=2$ and $\mu_2=8$, and Example 2.6 shows that $\gamma_*=\frac{1}{2}=m^{-1/2}$; furthermore, (2.24) applies and yields $X_{n,k}/\sqrt{nZ_n}\xrightarrow{\mathrm{d}}0$.

We conjecture that in this case (and similar ones with $\zeta_k=0$ in Theorem 2.2), $X_{n,k}/\sqrt{Z_n}$ has a non-trivial normal limit in distribution; we leave this as an open problem. Simliarly, we conjecture that when each $\widehat{\Xi}(\gamma_i)$ is degenerate in Theorem 2.3, the distribution of $X_{n,k}$ is asymptotically determined by the next smallest roots in Γ_*.

2.1. **More notation.** For a random variable X in a Banach space $\mathcal{B}$, we define $\|X\|_{L^2(\mathcal{B})} := (\mathbb{E}\,\|X\|_{\mathcal{B}}^2)^{1/2}$, when $\mathcal{B} = \mathbb{R}$ or $\mathbb{C}$ abbreviated to $\|X\|_2$.

For infinite vectors $\vec{x} = (x_j)_{j=0}^{\infty}$ and $\vec{y} = (y_j)_{j=0}^{\infty}$, let $\langle \vec{x}, \vec{y} \rangle := \sum_{j=0}^{\infty} x_j y_j$, assuming that the sum converges absolutely.

C denotes different constants that may depend on the distribution of the branching process (i.e., on the distribution of N and (ξ_i)), but not on n and similar parameters; the constant may change from one occurrence to the next.

$O_{\text{a.s.}}(1)$ means a quantity that is bounded by a random constant that does not depend on n.

All unspecified limits are as $n \to \infty$.

3. Preliminaries

Let

$$B_n := Z_n - Z_{n-1} \tag{3.1}$$

be the number of individuals born at time n (with $B_0 = Z_0$). Thus,

$$Z_n = Z_{n-1} + B_n = \sum_{i=0}^{n} B_i, \qquad n \geqslant 0. \tag{3.2}$$

Let $B_{n,k}$ be the number of individuals born at time $n+k$ by parents that are themselves born at time n, and thus are of age k. Thus, recalling (A2),

$$B_n = \sum_{k=1}^{n} B_{n-k,k}, \qquad n \geqslant 1. \tag{3.3}$$

Let $\mathcal{F}_n$ be the σ-field generated by the life histories of all individuals born up to time n, with $\mathcal{F}_n$ trivial for $n < 0$. Then $B_{n,k}$ is $\mathcal{F}_n$-measurable, and B_n is $\mathcal{F}_{n-1}$-measurable by (3.3). Furthermore,

$$\mathbb{E}\big(B_{n,k} \mid \mathcal{F}_{n-1}\big) = \mu_k B_n, \qquad n \geqslant 0. \tag{3.4}$$

For $k \geqslant 1$, let

$$W_{n,k} := B_{n,k} - \mathbb{E}\big(B_{n,k} \mid \mathcal{F}_{n-1}\big) = B_{n,k} - \mu_k B_n. \tag{3.5}$$

(Thus $W_{n,k} = 0$ if $n < 0$.) Then $W_{n,k}$ is $\mathcal{F}_n$-measurable with

$$\mathbb{E}\big(W_{n,k} \mid \mathcal{F}_{n-1}\big) = 0. \tag{3.6}$$

Let further

$$W_n := B_n - \sum_{k=1}^{n} \mu_k B_{n-k} = B_n - \sum_{k=1}^{\infty} \mu_k B_{n-k}. \tag{3.7}$$

Thus $W_0 = B_0 = Z_0$, and for $n \geqslant 1$, by (3.7), (3.3) and (3.5),

$$W_n = \sum_{k=1}^{n} W_{n-k,k}. \tag{3.8}$$

Lemma 3.1. *Assume* (A1)–(A6)*. Then, for all $n \geqslant 1$ and $k \geqslant 1$, $\mathbb{E}[W_{n,k}^2] \leqslant Cr^{-2k}m^n$ and $\mathbb{E}[W_n^2] \leqslant Cm^n$.*

Proof. Recall that N_k is the number of children born at age k of an individual, and that $\mathbb{E}\,N_k = \mu_k$. Furthermore, by (2.2), $\widehat{\Xi}(r) \geqslant N_k r^k$ and thus

$$\operatorname{Var} N_k \leqslant \mathbb{E}\,N_k^2 \leqslant r^{-2k}\,\mathbb{E}[\widehat{\Xi}(r)^2] = Cr^{-2k}. \tag{3.9}$$

Let $n \geqslant 0$ and $k \geqslant 1$. Given $\mathcal{F}_{n-1}$, $B_{n,k}$ is the sum of B_n independent copies of N_k, and thus, see (3.5), (3.4) and (3.9),

$$\mathbb{E}\big(W_{n,k}^2 \mid \mathcal{F}_{n-1}\big) = B_n \operatorname{Var}(N_k) \leqslant Cr^{-2k}B_n. \tag{3.10}$$

Taking the expectation and using (2.7) we find

$$\mathbb{E}[W_{n,k}^2] \leqslant Cr^{-2k}\,\mathbb{E}\,B_n \leqslant Cr^{-2k}\,\mathbb{E}\,Z_n \leqslant Cr^{-2k}m^n, \tag{3.11}$$

as asserted. Consequently $\|W_{n,k}\|_2 \leqslant Cr^{-k}m^{n/2}$ and, by (3.8) and Minkowski's inequality, using $rm^{1/2} > 1$,

$$\|W_n\|_2 \leqslant \sum_{k=1}^{n} \|W_{n-k,k}\|_2 \leqslant Cm^{n/2}\sum_{k=1}^{\infty}(rm^{1/2})^{-k} \leqslant Cm^{n/2}. \tag{3.12}$$

□

For $n \geqslant 0$ and $k \geqslant 1$, by (2.10),

$$\begin{aligned} X_{n+1,k} &= Z_{n+1-k} - m^{-k}Z_{n+1} = X_{n,k-1} + m^{1-k}Z_n - m^{-k}Z_{n+1} \\ &= X_{n,k-1} + m^{-k}(mZ_n - Z_{n+1}). \end{aligned} \tag{3.13}$$

Furthermore, by (3.1) and (2.10), we have, for $k \geqslant 0$,

$$B_{n-k} = Z_{n-k} - Z_{n-k-1} = X_{n,k} - X_{n,k+1} + (m-1)m^{-k-1}Z_n. \tag{3.14}$$

By (3.2), (3.7) and (3.14), recalling that $X_{n,0} = 0$ by (2.10) and $\widehat{\mu}(m^{-1}) = 1$ by (A5), for $n \geqslant 0$,

$$\begin{aligned} mZ_n - Z_{n+1} &= (m-1)Z_n - B_{n+1} = (m-1)Z_n - \sum_{k=1}^{\infty}\mu_k B_{n+1-k} - W_{n+1} \\ &= (m-1)Z_n - \sum_{k=1}^{\infty}\mu_k\big(X_{n,k-1} - X_{n,k} + (m-1)m^{-k}Z_n\big) - W_{n+1} \\ &= (m-1)Z_n - \sum_{k=1}^{\infty}\mu_k\big(X_{n,k-1} - X_{n,k}\big) - (m-1)\widehat{\mu}(m^{-1})Z_n - W_{n+1} \\ &= \sum_{k=1}^{\infty}\mu_k\big(X_{n,k} - X_{n,k-1}\big) - W_{n+1}. \end{aligned} \tag{3.15}$$

Consequently, (3.13) yields, for $n \geqslant 0$ and $k \geqslant 1$,

$$X_{n+1,k} = X_{n,k-1} + m^{-k}\Bigl(\sum_{j=1}^{\infty} \mu_j\bigl(X_{n,j} - X_{n,j-1}\bigr) - W_{n+1}\Bigr). \tag{3.16}$$

Introduce the vector notation $\vec{X}_n := (X_{n,k})_{k=0}^{\infty}$ and

$$\vec{v} := (0, m^{-1}, m^{-2}, \dots) = \bigl(m^{-k}\mathbf{1}\{k>0\}\bigr)_{k=0}^{\infty}, \tag{3.17}$$

and for vectors $\vec{y} = (y_k)_0^{\infty}$ such that the sum converges, define

$$\Psi\bigl((y_k)_0^{\infty}\bigr) := \sum_{k=1}^{\infty} \mu_k (y_k - y_{k-1}). \tag{3.18}$$

Let S be the shift operator $S\bigl((y_k)_0^{\infty}\bigr) := (y_{k-1})_0^{\infty}$ with $y_{-1} := 0$, and let T be the linear operator

$$T(\vec{y}) := S(\vec{y}) + \Psi(\vec{y})\vec{v}. \tag{3.19}$$

Then (3.16) can be written, again recalling that $X_{n,0} = 0$,

$$\vec{X}_{n+1} = S(\vec{X}_n) + \bigl(\Psi(\vec{X}_n) - W_{n+1}\bigr)\vec{v} = T(\vec{X}_n) - W_{n+1}\vec{v}. \tag{3.20}$$

This recursion leads to the following formula.

Lemma 3.2. *For every $n \geqslant 0$,*

$$\vec{X}_n = -\sum_{k=0}^{n} W_{n-k} T^k(\vec{v}). \tag{3.21}$$

Proof. For the initial value $\vec{X}_0$, we have by (2.10) $X_{0,k} = -m^{-k}Z_0$ for $k \geqslant 1$, and thus by (3.17) $\vec{X}_0 = -Z_0\vec{v} = -W_0\vec{v}$, recalling that $W_0 = B_0 = Z_0$. This verifies (3.21) for $n = 0$. The general case follows by (3.20) and induction. □

Remark 3.3. It follows from the proofs below, that the sum in (3.21) is dominated by the first few terms in the case $\gamma_* > m^{-1/2}$, and by the last few terms in the case $\gamma_* < m^{-1/2}$, while all terms are of about the same size when $\gamma_* = m^{-1/2}$. This explains much of the different behaviours seen in Section 2.

We now consider T defined in (3.19) as an operator on the complex Hilbert space ℓ^2_R defined in (2.13) for a suitable $R > 0$. Recall that the *spectrum* $\sigma(T)$ of a linear operator in a complex Hilbert (or Banach) space is the set of complex numbers λ such that $\lambda - T$ is *not* invertible; see e.g. [4, Section VII.3].

Lemma 3.4. *Suppose that $1 \leqslant R < m$ and that $\widehat{\mu}(R^{-1}) < \infty$. Then $\vec{v} \in \ell^2_R$, Ψ is a bounded linear functional on ℓ^2_R and T is a bounded linear operator on ℓ^2_R. Furthermore, if $\lambda \in \mathbb{C}$ with $|\lambda| > R$, then $\lambda \in \sigma(T)$ if and only if $\lambda^{-1} \in \Gamma_*$, i.e., if and only if $\lambda \neq m$ and $\widehat{\mu}(\lambda^{-1}) = 1$.*

Proof. By (3.17) and (2.13), and because $R < m$,

$$\|\vec{v}\|^2_{\ell^2_R} = \sum_{k=1}^{\infty} R^{2k} m^{-2k} < \infty. \tag{3.22}$$

Next, it is clear from (2.13) that the shift operator S is bounded on ℓ^2_R (with norm R). Furthermore, by (2.1) and assumption,

$$\sum_{k=1}^{\infty} R^{-2k} \mu_k^2 \leqslant \widehat{\mu}(R^{-1})^2 < \infty \tag{3.23}$$

and it follows by the Cauchy–Schwarz inequality that $\Psi_1\big((a_k)_0^\infty\big) := \sum_{k=1}^\infty \mu_k a_k$ defines a bounded linear functional Ψ_1 on ℓ^2_R. Since Ψ can be written $\Psi = \Psi_1 - \Psi_1 S$, Ψ too is bounded. It now follows from (3.19) that T is a bounded linear operator on ℓ^2_R.

For the final statement we note that the mapping $(a_k)_0^\infty \mapsto \sum_{k=0}^\infty a_k z^k$ is an isometry of ℓ^2_R onto the Hardy space H^2_R consisting of all analytic functions $f(z)$ in the disc $\{z : |z| < R\}$ such that

$$\|f\|^2_{H^2_R} := \sup_{r<R} \frac{1}{2\pi} \int_0^{2\pi} |f\big(re^{i\theta}\big)|^2 \, d\theta < \infty. \tag{3.24}$$

(See e.g. [5].) In particular, $\vec{v}$ corresponds to the function

$$v(z) := \sum_{k=1}^{\infty} m^{-k} z^k = \frac{z/m}{1 - z/m} = \frac{z}{m - z}. \tag{3.25}$$

We use the same notations Ψ, S and T for the corresponding linear functional and operators on H^2_R, and note that the shift operator S on ℓ^2_R corresponds to the multiplication operator $Sf(z) = zf(z)$ on H^2_R. The definition (3.19) thus translates to

$$Tf(z) = zf(z) + \Psi(f)v(z). \tag{3.26}$$

Consequently, for any $h \in H^2_R$, the equation $(\lambda - T)f = h$ is equivalent to

$$(\lambda - z)f(z) - \Psi(f)v(z) = h(z). \tag{3.27}$$

Any solution to (3.27) has to be of the form

$$f(z) = c\frac{v(z)}{\lambda - z} + \frac{h(z)}{\lambda - z}, \tag{3.28}$$

where

$$c = \Psi(f) = c\Psi\Big(\frac{v(z)}{\lambda - z}\Big) + \Psi\Big(\frac{h(z)}{\lambda - z}\Big). \tag{3.29}$$

Suppose $|\lambda| > R$; then $1/(\lambda - z)$ is a bounded analytic function on the domain $\{|z| < R\}$, so it follows from (3.24) and $v, h \in H^2_R$ that $v(z)/(\lambda - z) \in H^2_R$ and $h(z)/(\lambda - z) \in H^2_R$. If $\Psi\big(v(z)/(\lambda - z)\big) \neq 1$, then (3.29) has a unique solution c for any $h \in H^2_R$, and thus (3.27) has a unique solution $f \in H^2_R$, given by (3.28). In other words, then $\lambda - T$ is invertible on H^2_R and $\lambda \notin \sigma(T)$. (Continuity of $(\lambda - T)^{-1}$ is automatic, by the closed graph theorem.)

Conversely, if $\Psi\bigl(v(z)/(\lambda-z)\bigr)=1$, then (3.27) has either no solution or infinitely many solutions f for any given $h\in H^2_R$, and thus $\lambda\in\sigma(T)$.

We have shown that for $|\lambda|>R$,

$$\lambda\in\sigma(T) \iff \Psi\Bigl(\frac{v(z)}{\lambda-z}\Bigr)=1. \tag{3.30}$$

We analyse the condition in (3.30) further. If $|\lambda|>R$ and $\lambda\neq m$, then, by (3.25),

$$\frac{v(z)}{\lambda-z}=\frac{z}{(\lambda-z)(m-z)}=\frac{1}{m-\lambda}\Bigl(\frac{\lambda}{\lambda-z}-\frac{m}{m-z}\Bigr). \tag{3.31}$$

Furthermore, $\lambda/(\lambda-z)=\sum_{k=0}^\infty \lambda^{-k}z^k$ and thus by (3.18) and (2.1),

$$\Psi\Bigl(\frac{\lambda}{\lambda-z}\Bigr)=\sum_{k=1}^\infty \mu_k\lambda^{-k}(1-\lambda)=(1-\lambda)\widehat{\mu}\bigl(\lambda^{-1}\bigr). \tag{3.32}$$

Hence, (3.31) yields, recalling $\widehat{\mu}(m^{-1})=1$ by (A5),

$$\begin{aligned}\Psi\Bigl(\frac{v(z)}{\lambda-z}\Bigr)&=\frac{1}{m-\lambda}\left(\Psi\Bigl(\frac{\lambda}{\lambda-z}\Bigr)-\Psi\Bigl(\frac{m}{m-z}\Bigr)\right)\\&=\frac{1}{m-\lambda}\bigl((1-\lambda)\widehat{\mu}(\lambda^{-1})-(1-m)\widehat{\mu}(m^{-1})\bigr)\\&=\frac{1}{m-\lambda}\bigl((1-\lambda)\widehat{\mu}(\lambda^{-1})+m-1\bigr). \end{aligned}\tag{3.33}$$

Consequently, for $|\lambda|>R$ with $\lambda\neq m$, by (3.30) and (3.33),

$$\begin{aligned}\lambda\in\sigma(T)&\iff \Psi\Bigl(\frac{v(z)}{\lambda-z}\Bigr)=1\\&\iff (1-\lambda)\widehat{\mu}(\lambda^{-1})+m-1=m-\lambda\\&\iff (1-\lambda)\widehat{\mu}(\lambda^{-1})=1-\lambda\\&\iff \widehat{\mu}(\lambda^{-1})=1.\end{aligned}\tag{3.34}$$

In the special case $\lambda=m$, we find by continuity, letting $\lambda\to m$ in (3.33),

$$\begin{aligned}\Psi\Bigl(\frac{v(z)}{m-z}\Bigr)&=\lim_{\lambda\to m}\Psi\Bigl(\frac{v(z)}{\lambda-z}\Bigr)=-\frac{d}{d\lambda}\bigl((1-\lambda)\widehat{\mu}(\lambda^{-1})\bigr)\big|_{\lambda=m}\\&=\widehat{\mu}(m^{-1})-(m-1)m^{-2}\widehat{\mu}'(m^{-1})<\widehat{\mu}(m^{-1})=1\end{aligned}\tag{3.35}$$

since $\widehat{\mu}'(x)>0$ for $x>0$. Hence $m\notin\sigma(T)$. □

Remark 3.5. It is easily seen that $\lambda\in\sigma(T)$ for every λ with $|\lambda|\leqslant R$, e.g. by taking $h=v$ in (3.27)–(3.28) and noting that $v(z)/(\lambda-z)\notin H^2_R$. Thus we have a complete description of the spectrum $\sigma(T)$ on ℓ^2_R.

Lemma 3.6. *Suppose that $1\leqslant R<m$ and that $\widehat{\mu}(R^{-1})<\infty$. Suppose furthermore that $\widehat{\mu}(z)\neq 1$ for every complex $z\neq m^{-1}$ with $|z|<R^{-1}$. Then, for every $R_1>R$, there exists $C=C(R_1)$ such that*

$$\|T^n\|_{\ell^2_R}\leqslant CR_1^n,\qquad n\geqslant 0. \tag{3.36}$$

Proof. By Lemma 3.4, T is a bounded linear operator on ℓ^2_R and if $\lambda \in \sigma(T)$ with $|\lambda| > R$, then $\widehat{\mu}(\lambda^{-1}) = 1$ and $\lambda^{-1} \neq m^{-1}$. By assumption, there is no such λ, and thus $\sigma(T) \subseteq \{\lambda : |\lambda| \leqslant R\}$. (Actually, equality holds by Remark 3.5.) In other words, the spectral radius

$$r(T) := \sup_{\lambda\in\sigma(T)} |\lambda| \leqslant R. \tag{3.37}$$

By the spectral radius formula [4, Lemma VII.3.4], $r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$ and thus (3.37) implies that, for any $R_1 > R$, $\|T^n\|^{1/n} < R_1$ for large n, which yields (3.36). □

We shall use Lemma 3.6 when $\gamma_* > m^{-1/2}$. In the case $\gamma_* \leqslant m^{-1/2}$, we use instead the following lemma, based on a more careful spectral analysis of T. Recall the definitions (2.4)–(2.6).

Lemma 3.7. *Assume that $R = r^{-1} \geqslant 1$, where $\widehat{\mu}(r) < \infty$. Suppose furthermore that $\Gamma_{**} = \{\gamma_1, \dots, \gamma_q\} \neq \emptyset$, and that (2.16) holds. Let $\lambda_i := \gamma_i^{-1}$. Then there exist eigenvectors $\vec{v}_i$ with $T\vec{v}_i = \lambda_i\vec{v}_i$ and linear projections P_i with range $\mathcal{R}(P_i) = \{c\vec{v}_i : c \in \mathbb{C}\}$ (i.e., the span of $\vec{v}_i$), $i = 1, \dots q$, and furthermore a bounded operator T_0 in ℓ^2_R and a constant $\tilde{R} < \gamma_*^{-1}$ such that, for any $n \geqslant 0$,*

$$T^n = T_0^n + \sum_{i=1}^{q} \lambda_i^n P_i \tag{3.38}$$

and

$$\left\|T_0^n\right\|_{\ell^2_R} \leqslant C\tilde{R}^n. \tag{3.39}$$

Explicitly,

$$\vec{v}_i = P_i(\vec{v}) = \frac{1}{\gamma_i(\gamma_i - 1)\widehat{\mu}'(\gamma_i)} \left(\gamma_i^k - m^{-k}\right)_k. \tag{3.40}$$

Proof. Since the points in Γ_* are isolated, there is a number $\tilde{r} > \gamma_*$ such that $|z| > \tilde{r}$ for any $z \in \Gamma_* \setminus \Gamma_{**}$. We may assume $\tilde{r} < r$. Let $\tilde{R} := \tilde{r}^{-1} > R$. By Lemma 3.4, $\lambda_i = \gamma_i^{-1} \in \sigma(T)$ with $|\lambda_i| = \gamma_*^{-1}$, and $|\lambda| < \tilde{R} < \gamma_*^{-1}$ for any $\lambda \in \sigma(T) \setminus \{\lambda_1, \dots, \lambda_q\}$.

Since $\lambda_1, \dots, \lambda_q$ thus are isolated points in $\sigma(T)$, by standard functional calculus, see e.g. [4, Section VII.3], there exist commuting projections (not necessarily orthogonal) $P_0, \dots, P_q$ in ℓ^2_R such that $\sum_{i=0}^q P_i = 1$, T maps each subspace $E_i := P_i(\ell^2_R)$ into itself, and if $\hat{T}_i$ is the restriction of T to E_i, then $\hat{T}_i$ has spectrum $\sigma(\hat{T}_i) = \{\lambda_i\}$ for $1 \leqslant i \leqslant q$ and $\sigma(\hat{T}_0) = \sigma(T) \setminus \{\lambda_i\}_1^q$. In particular, the spectral radius $r(\hat{T}_0) < \tilde{R}$, and thus, by the spectral radius formula [4, Lemma VII.3.4],

$$\|\hat{T}_0^n\| \leqslant C\tilde{R}^n, \qquad n \geqslant 0. \tag{3.41}$$

Let $T_0 := TP_0$. Then $T_0^n = T^n P_0 = \hat{T}_0^n P_0$, and (3.39) follows.

It remains to show that the spaces $E_i = \mathcal{R}(P_i)$ are one-dimensional, and spanned by the vectors $\vec{v}_i$ in (3.40).

We use, as the proof of Lemma 3.4, the isometry $(a_k)_0^\infty \mapsto \sum_{k=0}^\infty a_k z^k$ of ℓ^2_R onto H^2_R.

For each λ_i, $\widehat{\mu}(\lambda_i^{-1}) = 1$, and thus $\Psi\bigl(v(z)/(\lambda_i - z)\bigr) = 1$ by (3.33), see also (3.30). Hence, (3.27)–(3.29) show, by taking $h = 0$, that the kernel $\mathcal{N}(\lambda_i - T)$ is one-dimensional and spanned by $v(z)/(\lambda_i - z)$. Similarly, again by (3.27)–(3.29), the range $\mathcal{R}(\lambda_i - T)$ is given by

$$\mathcal{R}(\lambda_i - T) = \Bigl\{h \in \ell^2_R : \Psi\Bigl(\frac{h(z)}{\lambda_i - z}\Bigr) = 0\Bigr\}. \tag{3.42}$$

By differentiating (3.33), we find for $|\lambda| > R$ with $\lambda^{-1} \in \Gamma_*$, i.e., $\lambda \neq m$ and $\widehat{\mu}\bigl(\lambda^{-1}\bigr) = 1$,

$$\begin{aligned}\Psi\Bigl(\frac{v(z)}{(\lambda - z)^2}\Bigr) &= -\frac{\mathrm{d}}{\mathrm{d}\lambda}\Psi\Bigl(\frac{v(z)}{\lambda - z}\Bigr) = \frac{\mathrm{d}}{\mathrm{d}\lambda}\Bigl(1 - \Psi\Bigl(\frac{v(z)}{\lambda - z}\Bigr)\Bigr) \\ &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{(1-\lambda)(1-\widehat{\mu}(\lambda^{-1}))}{m - \lambda} = \frac{(1-\lambda)\widehat{\mu}'(\lambda^{-1})}{(m-\lambda)\lambda^2}.\end{aligned} \tag{3.43}$$

Thus, the assumption (2.16) implies that $\Psi\bigl(v(z)/(\lambda_i - z)^2\bigr) \neq 0$, and thus $v(z)/(\lambda_i - z) \notin \mathcal{R}(\lambda_i - T)$ by (3.42). Hence, $\mathcal{N}(\lambda_i - T) \cap \mathcal{R}(\lambda_i - T) = \{0\}$. Consequently, for every $h \in \mathcal{R}(\lambda_i - T)$, (3.27) has a unique solution $f \in \mathcal{R}(\lambda_i - T)$, i.e., the restriction of $\lambda_i - T$ to $\mathcal{R}(\lambda_i - T)$ is invertible.

It follows that the projection P_i is the projection onto $\mathcal{N}(\lambda_i - T) = \{cv(z)/(\lambda_i - z)\}$ that vanishes on $\mathcal{R}(\lambda_i - T)$, which by (3.42) is given by

$$P_i(f(z)) = \frac{\Psi\bigl(f(z)/(\lambda_i - z)\bigr)}{\Psi\bigl(v(z)/(\lambda_i - z)^2\bigr)} \cdot \frac{v(z)}{\lambda_i - z}. \tag{3.44}$$

In particular, since $\Psi\bigl(v(z)/(\lambda_i - z)\bigr) = 1 \neq 0$, $P_i(v)$ is a non-zero multiple of $v(z)/(\lambda_i - z)$. Let $\vec{v}_i := P_i(\vec{v})$. Thus $T\vec{v}_i = \lambda_i \vec{v}_i$, and, for $n \geqslant 0$,

$$T^n = T^n P_0 + \sum_{i=1}^q T^n P_i = T_0^n + \sum_{i=1}^q \lambda_i^n P_i, \tag{3.45}$$

showing (3.38).

Finally, (3.44) and (3.43) yield

$$v_i(z) := P_i\bigl(v(z)\bigr) = \frac{(m - \lambda_i)\lambda_i^2}{(1-\lambda_i)\widehat{\mu}'(\lambda_i^{-1})} \cdot \frac{v(z)}{\lambda_i - z}, \tag{3.46}$$

and (3.40) follows because $\lambda_i = \gamma_i^{-1}$ and by (3.25), for $|\lambda| > R$,

$$(m - \lambda)\frac{v(z)}{\lambda - z} = \frac{\lambda}{\lambda - z} - \frac{m}{m - z} = \sum_{k=0}^\infty (\lambda^{-k} - m^{-k}) z^k. \tag{3.47}$$

$\square$

Remark 3.8. It follows also that (2.16) implies that the points $\lambda_i \in \sigma(T)$ are simple poles of the resolvent $(\lambda - T)^{-1}$, and conversely. Lemma 3.7 can be extended without assuming (2.16); the general result is similar but more complicated, and is left to the reader. Cf. [4, Theorem VII.3.18].

We shall also use another similar calculation.

Lemma 3.9. *Suppose that* $1 \leqslant R < m$ *and that* $\widehat{\mu}(R^{-1}) < \infty$. *If* $|\lambda| > R$ *and* $\widehat{\mu}\bigl(\lambda^{-1}\bigr) \neq 1$, *then*

$$(\lambda - T)^{-1}(\vec{v}) = \frac{1}{(1-\lambda)(1-\widehat{\mu}(\lambda^{-1}))}\bigl(\lambda^{-k} - m^{-k}\bigr)_k. \tag{3.48}$$

Proof. Taking $h = v$ in (3.27)–(3.29), we find

$$(\lambda - T)^{-1} v(z) = f(z) = b\frac{v(z)}{\lambda - z} \tag{3.49}$$

for a constant b such that $b = \Psi(f) + 1$. This yields by (3.33)

$$b - 1 = \Psi(f) = \frac{b}{m-\lambda}\bigl((1-\lambda)\widehat{\mu}(\lambda^{-1}) + m - 1\bigr) \tag{3.50}$$

with the solution

$$b = \frac{m-\lambda}{(1-\lambda)(1-\widehat{\mu}(\lambda^{-1}))}. \tag{3.51}$$

Hence, using (3.47), for $|z| < R$,

$$f(z) = b\frac{v(z)}{\lambda - z} = \frac{1}{(1-\lambda)(1-\widehat{\mu}(\lambda^{-1}))}\sum_{k=0}^{\infty}(\lambda^{-k} - m^{-k})z^k. \tag{3.52}$$

□

4. A first normal convergence result

Let $\vec{\eta} := (\eta_0, \eta_1, \eta_2, \dots)$, where $(\eta_k)_0^\infty$ are jointly normal random variables with means $\mathbb{E}\,\eta_k = 0$ and covariances

$$\operatorname{Cov}(\eta_j, \eta_k) = \sigma_{jk} = \operatorname{Cov}(N_j, N_k), \tag{4.1}$$

see (2.11). Note that $\eta_0 = 0$ since $N_0 = 0$.

Lemma 4.1. *Assume* (A1)–(A6), *and let* $\vec{\eta}^{(k)} = (\eta_j^{(k)})_{j=0}^\infty$, $k = 1, 2, \dots$, *be independent copies of the random vector* η. *Then, as* $n \to \infty$,

$$Z_n^{-1/2} W_{n-k,j} \xrightarrow{\mathrm{d}} (1 - 1/m)^{1/2} m^{-k/2} \eta_j^{(k)}, \tag{4.2}$$

jointly for all (j,k) *with* $j \geqslant 0$ *and* $k \geqslant 0$.

Proof. Consider first a fixed $k \geqslant 0$. Given B_{n-k}, the vector $\vec{B}_{n-k} := (B_{n-k,j})_{j=0}^\infty$ is the sum of B_{n-k} independent copies of the random vector $\vec{N}$, and by (3.5), the vector $\vec{W}_{n-k} := (W_{n-k,j})_{j=0}^\infty$ is the sum of B_{n-k} independent copies of the centered random vector $\vec{N} - \mathbb{E}\,\vec{N}$. By (3.1) and (2.9),

$$\frac{B_n}{Z_n} = 1 - \frac{Z_{n-1}}{Z_n} \xrightarrow{\text{a.s.}} 1 - m^{-1} > 0. \tag{4.3}$$

In particular, $B_n \to \infty$ a.s., and thus $B_{n-k} \to \infty$. Consequently, by the central limit theorem for i.i.d. finite-dimensional vector-valued random variables, and the definition of η_j,

$$B_{n-k}^{-1/2} W_{n-k,j} \xrightarrow{\mathrm{d}} \eta_j \overset{\mathrm{d}}{=} \eta_j^{(k)}, \tag{4.4}$$

jointly for any finite set of $j \geqslant 0$.

Moreover, by (4.3) and (2.9),

$$B_{n-k}/Z_n \xrightarrow{\text{a.s.}} (1 - 1/m)m^{-k}, \tag{4.5}$$

and thus (4.2) for a fixed k follows from (4.3) and (4.4).

To extend this to several k, the problem is that $W_{n-k,j}$ for different k are, in general, dependent. (For example, conditioned on Z_{n-1} and B_{n-1}, $W_{n-1,1}$ determines $B_{n-1,1}$ which contributes to B_n, and thus influences $W_{n,j}$.) We therefore approximate $W_{n-k,j}$ as follows.

We may assume that for each k, we have an infinite sequence $(\vec{N}^{(k,i)})_{i \geqslant 1}$ of independent copies of $\vec{N}$, such that $\vec{W}_{n-k}$ is the sum $\sum_{i=1}^{B_{n-k}} \vec{N}^{(k,i)}$ of the first B_{n-k} vectors; furthermore, these sequences for different k are independent.

Fix $J, K \geqslant 1$ and consider only $j \leqslant J$ and $k \leqslant K$. Let, for $0 \leqslant k \leqslant K$,

$$\overline{B}_{n-k} := \lfloor m^{K-k} B_{n-K} \rfloor \tag{4.6}$$

and let

$$\overline{W}_{n-k,j} := \sum_{i=1}^{\overline{B}_{n-k}} \vec{N}_j^{(k,i)}. \tag{4.7}$$

Then by the central limit theorem, exactly as for (4.4),

$$\overline{B}_{n-k}^{-1/2} \overline{W}_{n-k,j} \xrightarrow{\mathrm{d}} \eta_j^{(k)}, \tag{4.8}$$

jointly for all $j \leqslant J$ and $k \leqslant K$; note that now, if we condition on B_{n-K}, the left-hand sides for different k are independent. Furthermore, by (4.3) and (2.9), $\overline{B}_{n-k}/B_{n-k} \xrightarrow{\text{a.s.}} 1$ for every k. Hence (4.8) yields, jointly,

$$B_{n-k}^{-1/2} \overline{W}_{n-k,j} \xrightarrow{\mathrm{d}} \eta_j^{(k)}. \tag{4.9}$$

Moreover, using (4.7),

$$\mathbb{E}\big((\overline{W}_{n-k,j} - W_{n-k,j})^2 \mid B_{n-k}, \overline{B}_{n-k}\big) = |B_{n-k} - \overline{B}_{n-k}| \operatorname{Var} N_j \tag{4.10}$$

and, consequently, for every fixed $j \geqslant 0$, $k \geqslant 0$ and $\varepsilon > 0$,

$$\mathbb{P}\big(|\overline{W}_{n-k,j} - W_{n-k,j}| > \varepsilon B_{n-k}^{1/2} \mid B_{n-k}, \overline{B}_{n-k}\big) \leqslant |1 - \overline{B}_{n-k}/B_{n-k}| \sigma_{jj} \varepsilon^{-2} \xrightarrow{\text{a.s.}} 0.$$

Taking the expectation, we obtain by dominated convergence that for every j and k, $\mathbb{P}\big(|\overline{W}_{n-k,j} - W_{n-k,j}| > \varepsilon B_{n-k}^{1/2}\big) \to 0$ for every $\varepsilon > 0$, and thus

$$B_{n-k}^{-1/2} \overline{W}_{n-k,j} - B_{n-k}^{-1/2} W_{n-k,j} \xrightarrow{\mathrm{p}} 0. \tag{4.11}$$

Combining (4.9) and (4.11) yields

$$B_{n-k}^{-1/2} W_{n-k,j} \xrightarrow{\mathrm{d}} \eta_j^{(k)}, \tag{4.12}$$

still jointly for all $j \leqslant J$ and $k \leqslant K$. The result follows by this and (4.5), since J and K are arbitrary. □

5. First proof of Theorem 2.1

In this section we assume (A1)–(A6) and also (B), i.e., $\gamma_* > m^{-1/2}$. In other words, see (2.5), each $z \in \Gamma_*$ satisfies $|z| > m^{-1/2}$. Hence, we may decrease r so that the disc D_r contains no roots of $\widehat{\mu}(z) = 1$ except m^{-1}, and still $r > m^{-1/2}$. Thus, with $R := 1/r$ and assuming (A1)–(A6), we see that $\gamma_* > m^{-1/2}$ is equivalent to:

(B′) There exists R with $1 \leqslant R < m^{1/2}$ such that $\widehat{\mu}(R^{-1}) < \infty$ and, furthermore, $\widehat{\mu}(z) \neq 1$ for every complex $z \neq m^{-1}$ with $|z| < R^{-1}$.

We fix an R such that (B′) holds, and (A6) holds with $r = 1/R$. Note that R may be chosen arbitrarily close to $m^{1/2}$. Furthermore, we fix R_1 with $R < R_1 < m^{1/2}$. Then (B′) and Lemma 3.6 show that (3.36) holds, i.e., $\|T^n\|_{\ell^2_R} = O\bigl(R_1^n\bigr)$.

Lemma 5.1. *Assume* (A1)–(A6) *and* (B). *If* $R < m^{1/2}$, *then*

$$\mathbb{E}\,\|\vec{X}_n\|^2_{\ell^2_R} \leqslant C m^n \tag{5.1}$$

and thus

$$\mathbb{E}\, X^2_{n,k} \leqslant C R^{-2k} m^n \tag{5.2}$$

for all $n, k \geqslant 0$.

Proof. By (3.21), Lemma 3.1, (3.36) and Minkowski's inequality,

$$\begin{aligned} \|\vec{X}_n\|_{L^2(\ell^2_R)} &\leqslant \sum_{k=0}^{n} \|W_{n-k}\|_{L^2} \|T^k(\vec{v})\|_{\ell^2_R} \leqslant C \sum_{k=0}^{n} m^{(n-k)/2} R_1^k \\ &= C m^{n/2} \sum_{k=0}^{\infty} (R_1/m^{1/2})^k = C m^{n/2}. \end{aligned} \tag{5.3}$$

This yields (5.1), and (5.2) follows by (2.13). □

Define for convenience $W_{n,j}$ also for $n < 0$ by $W_{-1,1} := W_0$ and $W_{n,j} = 0$ for $n \leqslant -1$ and $j \geqslant 1$ with $(n,j) \neq (-1,1)$. Then (3.8) holds also for $n \leqslant 0$, provided the sum is extended to ∞, and (3.21) can be written

$$\vec{X}_n = -\sum_{k=0}^{\infty} \sum_{j=1}^{\infty} W_{n-k-j,j} T^k(\vec{v}). \tag{5.4}$$

For each finite M define also the truncated sum

$$\vec{X}_{n,M} := -\sum_{k=0}^{M} \sum_{j=1}^{M} W_{n-k-j,j} T^k(\vec{v}). \tag{5.5}$$

Lemma 4.1 implies that for any fixed M, as $n\to\infty$,

$$Z_n^{-1/2}\vec{X}_{n,M} \xrightarrow{\mathrm{d}} -\sum_{k=0}^{M}\sum_{j=1}^{M}(1-m^{-1})^{1/2}m^{-(k+j)/2}\eta_j^{(k+j)}T^k(\vec{v}) \tag{5.6}$$

in ℓ^2_R. Furthermore, by (5.4)–(5.5), Minkowski's inequality, Lemma 3.1 and (3.36), regarding $\vec{X}_n$ and $\vec{X}_{n,M}$ as elements of $L^2(\ell^2_R)$, the space of ℓ^2_R-valued random variables with square integrable norm,

$$\begin{aligned}\|\vec{X}_n-\vec{X}_{n,M}\|_{L^2(\ell^2_R)} &\leqslant \sum_{k>M \text{ or } j>M}\|W_{n-k-j,j}\|_{L^2}\|T^k(\vec{v})\|_{\ell^2_R}\\ &\leqslant C\sum_{k>M \text{ or } j>M} r^{-j}m^{(n-k-j)/2}R_1^k\\ &= Cm^{n/2}\sum_{k>M \text{ or } j>M}(R/m^{1/2})^j(R_1/m^{1/2})^k.\end{aligned} \tag{5.7}$$

Since the sum on the right-hand side of (5.7) converges, it tends to 0 as $M\to\infty$, and thus $m^{-n/2}\bigl(\vec{X}_n-\vec{X}_{n,M}\bigr)\to 0$ in $L^2(\ell^2_R)$, and thus in probability, uniformly in n. Since $Z_n/m^n \xrightarrow{\text{a.s.}} \mathcal{Z}>0$, see (2.8), $\sup_n m^n/Z_n$ is an a.s. finite random variable; hence also

$$Z_n^{-1/2}\bigl(\vec{X}_n-\vec{X}_{n,M}\bigr) = \Bigl(\frac{m^n}{Z_n}\Bigr)^{1/2}m^{-n/2}\bigl(\vec{X}_n-\vec{X}_{n,M}\bigr)\xrightarrow{\mathrm{p}} 0 \tag{5.8}$$

as $M\to\infty$, uniformly in n.

Moreover, the right-hand side of (5.6) converges as $M\to\infty$ in $L^2(\ell^2_R)$, and thus in distribution, since by (3.9)

$$\mathbb{E}[(\eta_j^{(k)})^2] = \operatorname{Var} N_j \leqslant Cr^{-2j} = CR^{2j}, \tag{5.9}$$

and thus, using also (3.36),

$$\begin{aligned}\sum_{k=0}^{\infty}\sum_{j=1}^{\infty}m^{-(k+j)/2}\|\eta_j^{(k+j)}T^k(\vec{v})\|_{L^2(\ell^2_R)} &= \sum_{k=0}^{\infty}\sum_{j=1}^{\infty}m^{-(k+j)/2}\|\eta_j^{(k+j)}\|_{L^2}\|T^k(\vec{v})\|_{\ell^2_R}\\ &\leqslant C\sum_{k=0}^{\infty}\sum_{j=1}^{\infty}m^{-(k+j)/2}R^jR_1^k<\infty.\end{aligned} \tag{5.10}$$

It follows, see [2, Theorem 4.2], that (5.6) extends to $M=\infty$, i.e.,

$$Z_n^{-1/2}\vec{X}_n \xrightarrow{\mathrm{d}} -\sum_{k=0}^{\infty}\sum_{j=1}^{\infty}(1-m^{-1})^{1/2}m^{-(k+j)/2}\eta_j^{(k+j)}T^k(\vec{v}) \tag{5.11}$$

in ℓ^2_R as $n\to\infty$. The right-hand side is obviously a Gaussian random vector in ℓ^2_R, which we write as $\vec{\zeta}=(\zeta_0,\zeta_1,\dots)$. Then (5.11) yields (2.14).

It remains to calculate the covariances of ζ_k. Let $\vec{a} = (a_0, a_1, \dots)$ be a (real) vector with only finitely many non-zero elements. Then, by (5.11),

$$\sum_{\ell=0}^{\infty} a_\ell \zeta_\ell = \langle \vec{a}, \vec{\zeta} \rangle = -(1-m^{-1})^{1/2} \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} m^{-(k+j)/2} \eta_j^{(k+j)} \langle T^k(\vec{v}), \vec{a} \rangle \tag{5.12}$$

with the sum converging absolutely in L^2 by (5.10).

By the definition of $\eta_j^{(k)}$ in (4.1) and Lemma 4.1,

$$\operatorname{Cov}\Bigl(m^{-k/2}\eta_i^{(k)}, m^{-\ell/2}\eta_j^{(\ell)}\Bigr) = m^{-(k+\ell)/2}\delta_{k,\ell}\sigma_{ij} = \oint_{|w|=m^{-1/2}} \sigma_{ij} w^k \bar{w}^\ell \frac{|\mathrm{d}w|}{2\pi m^{-1/2}}. \tag{5.13}$$

Hence, (5.12) yields

$$\begin{aligned}
&(1-m^{-1})^{-1} \operatorname{Var}\bigl(\langle \vec{a}, \vec{\zeta} \rangle\bigr) \\
&\quad= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \langle T^k(\vec{v}), \vec{a} \rangle \langle T^\ell(\vec{v}), \vec{a} \rangle \oint_{|w|=m^{-1/2}} \sigma_{ij} w^{k+i} \bar{w}^{\ell+j} \frac{|\mathrm{d}w|}{2\pi m^{-1/2}} \\
&\quad= \oint_{|w|=m^{-1/2}} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sigma_{ij} w^i \bar{w}^j \left| \sum_{k=0}^{\infty} w^k \langle T^k(\vec{v}), \vec{a} \rangle \right|^2 \frac{|\mathrm{d}w|}{2\pi m^{-1/2}}.
\end{aligned} \tag{5.14}$$

Furthermore, if $|w| = m^{-1/2}$, then $\sum_{k=0}^{\infty} \|w^k T^k(\vec{v})\|_{\ell^2_R} < \infty$ by (3.36), and thus

$$\sum_{k=0}^{\infty} w^k T^k(\vec{v}) = (1-wT)^{-1}(\vec{v}). \tag{5.15}$$

Let $\lambda := w^{-1}$, so $|\lambda| = m^{1/2} > R$. We use as in the proof of Lemma 3.4 the standard isometry $\ell^2_R \to H^2_R$, and let $f(z) \in H^2_R$ be the function corresponding to $(1-wT)^{-1}(\vec{v}) = \lambda(\lambda - T)^{-1}(\vec{v})$. Thus, see (3.26)–(3.27),

$$(\lambda - z)f(z) - \Psi(f)v(z) = (\lambda - T)f(z) = \lambda v(z) \tag{5.16}$$

and thus, cf. (3.27)–(3.29),

$$f(z) = b\frac{v(z)}{\lambda - z} \tag{5.17}$$

for a constant b such that $b = \Psi(f) + \lambda$. This yields by (3.33)

$$b - \lambda = \Psi(f) = \frac{b}{m-\lambda}\bigl((1-\lambda)\widehat{\mu}(\lambda^{-1}) + m - 1\bigr) \tag{5.18}$$

with the solution

$$b = \frac{\lambda(m-\lambda)}{(1-\lambda)(1-\widehat{\mu}(\lambda^{-1}))}. \tag{5.19}$$

Hence, using (3.31), for $|z| \leqslant R$,

$$\begin{aligned} f(z) = b\frac{v(z)}{\lambda - z} &= \frac{\lambda}{(1-\lambda)(1-\widehat{\mu}(\lambda^{-1}))}\Bigl(\frac{\lambda}{\lambda - z} - \frac{m}{m-z}\Bigr) \\ &= \frac{\lambda}{(1-\lambda)(1-\widehat{\mu}(\lambda^{-1}))}\sum_{\ell=0}^{\infty}(\lambda^{-\ell} - m^{-\ell})z^{\ell}. \\ &= \frac{1}{(w-1)(1-\widehat{\mu}(w))}\sum_{\ell=0}^{\infty}(w^{\ell} - m^{-\ell})z^{\ell}. \end{aligned} \tag{5.20}$$

Thus, $(1-wT)^{-1}(\vec{v}) = \bigl(((w-1)(1-\widehat{\mu}(w)))^{-1}(w^{\ell}-m^{-\ell})\bigr)_{\ell}$ and, using (5.15),

$$\sum_{k=0}^{\infty} w^k \langle T^k(\vec{v}), \vec{a}\rangle = \langle (1-wT)^{-1}(\vec{v}), \vec{a}\rangle = \frac{1}{(w-1)(1-\widehat{\mu}(w))}\sum_{\ell=0}^{\infty} a_{\ell}(w^{\ell} - m^{-\ell}). \tag{5.21}$$

Hence (2.15) follows from (5.14).

Finally, by (2.15), the variable ζ_k is degenerate only if $\Sigma(z) = 0$ for every z with $|z| = m^{-1/2}$, and thus, by (2.12), $\widehat{\Xi}(z) = \widehat{\mu}(z)$ a.s. for every such z, which by (2.1)–(2.2) implies $N_k = \mu_k$ a.s. for every k. $\square$

6. A martingale

In the remaining sections, we let $R := r^{-1} < m^{1/2}$, where r is as in (A6). (We may assume that R is arbitrarily close to $m^{1/2}$ by decreasing r.) We consider as above the operator T on ℓ^2_R.

Fix a real vector $\vec{a} \in \ell^2_{R^{-1}}$ (for example any finite real vector), and write

$$\alpha_k = \alpha_k(\vec{a}) := \langle T^k(\vec{v}), \vec{a}\rangle. \tag{6.1}$$

Then (3.21) and (3.8) yield

$$\langle \vec{X}_n, \vec{a}\rangle = -\sum_{k=0}^{\infty}\sum_{j=1}^{\infty} W_{n-k-j,j}\alpha_k = -\sum_{\ell=0}^{n}\sum_{j=1}^{n-\ell} W_{\ell,j}\alpha_{n-j-\ell} \tag{6.2}$$

Define

$$\Delta M_{n,\ell} := \sum_{j=1}^{n-\ell} \alpha_{n-j-\ell} W_{\ell,j}, \tag{6.3}$$

$$M_{n,k} := \sum_{\ell=0}^{k} \Delta M_{n,\ell}. \tag{6.4}$$

Then (3.6) shows that $\mathbb{E}\bigl(\Delta M_{n,\ell} \mid \mathcal{F}_{\ell-1}\bigr) = 0$, and thus $(M_{n,k})_{k=0}^{n}$ is a martingale with respect to $(\mathcal{F}_k)_k$. Furthermore, by (6.2),

$$\langle \vec{X}_n, \vec{a}\rangle = -M_{n,n}. \tag{6.5}$$

Conditioned on $\mathcal{F}_{\ell-1}$, the vector $(W_{\ell,j})_j$ is the sum of B_ℓ independent copies of $\vec{N} - \mathbb{E}\,\vec{N}$, where $\vec{N} = (N_j)_0^\infty$, and thus, recalling (2.11),

$$\begin{aligned}
Q_{n,l} := \mathbb{E}\big((\Delta M_{n,\ell})^2 \mid \mathcal{F}_{\ell-1}\big) &= B_\ell \operatorname{Var}\Bigl(\sum_{j=1}^{n-\ell} \alpha_{n-\ell-j} N_j\Bigr) \\
&= B_\ell \sum_{i,j=1}^{n-\ell} \sigma_{ij}\alpha_{n-\ell-i}\alpha_{n-\ell-j}. \qquad (6.6)
\end{aligned}$$

The conditional quadratic variation of the martingale $(M_{n,k})_k$ is thus

$$V_n := \sum_{\ell=0}^{n} Q_{n,\ell} = \sum_{\ell=0}^{n} B_\ell \sum_{i,j=1}^{n-\ell} \sigma_{ij}\alpha_{n-\ell-i}\alpha_{n-\ell-j} = \sum_{\ell=0}^{n} B_{n-\ell} \sum_{i,j=1}^{\ell} \sigma_{ij}\alpha_{\ell-i}\alpha_{\ell-j}. \qquad (6.7)$$

By (2.2), $N_k \leqslant r^{-k}\widehat{\Xi}(r)$, and thus by (2.11) and the Cauchy–Schwarz inequality,

$$|\sigma_{ij}| \leqslant r^{-i-j}\,\mathbb{E}\,\widehat{\Xi}(r)^2 = CR^{i+j}. \qquad (6.8)$$

7. Second proof of Theorem 2.1

As said earlier, we give here another proof of Theorem 2.1, based on a martingale central limit theorem. and the martingale in Section 6. The main reason is that the new proof with small modifications also applies to Theorem 2.2, see Section 8, and we prefer to present it first for Theorem 2.1. (The proof in Section 5 does not seem to extend easily to Theorem 2.2.)

Let R and R_1 be as in Section 5. Then, (6.1) and (3.36) show that, for a fixed $\vec{a}$, with $C = C(\vec{a})$,

$$|\alpha_k| \leqslant CR_1^k. \qquad (7.1)$$

Consequently, by (6.6), (6.8) and (7.1), since $R/R_1 < 1$,

$$\frac{Q_{n,\ell}}{B_\ell} = \sum_{i,j=1}^{n-\ell} \sigma_{ij}\alpha_{n-\ell-i}\alpha_{n-\ell-j} \leqslant C \sum_{i,j=1}^{\infty} R^{i+j} R_1^{2(n-\ell)-i-j} \leqslant CR_1^{2(n-\ell)}. \qquad (7.2)$$

Hence, by (6.7), (6.6), (3.1) and (2.9), using dominated convergence justified by (7.2) and $R_1^2/m < 1$,

$$\begin{aligned}
\frac{V_n}{Z_n} &= \sum_{\ell=0}^{n} \frac{B_{n-\ell}}{Z_n}\frac{Q_{n,n-\ell}}{B_{n-\ell}} = \sum_{\ell=0}^{n} \frac{Z_{n-\ell} - Z_{n-\ell-1}}{Z_n} \sum_{i,j=1}^{\ell} \sigma_{ij}\alpha_{\ell-i}\alpha_{\ell-j} \\
&\xrightarrow{\text{a.s.}} \sigma^2(\vec{a}) := \sum_{\ell=0}^{\infty} \bigl(m^{-\ell} - m^{-\ell-1}\bigr) \sum_{i,j=1}^{\ell} \sigma_{ij}\alpha_{\ell-i}\alpha_{\ell-j} \qquad (7.3)
\end{aligned}$$

We cannot use a martingale central limit theorem directly for the martingale $(M_{n,k})_k$ defined in (6.4), because the calculations above show that most of the conditional quadratic variation V_n comes from a few terms (the last ones), cf. Remark 3.3. We thus introduce another martingale.

Number the individuals $1, 2, \dots$ in order of birth, with arbitrary order at ties, and let $\mathcal{G}_\ell$ be the σ-field generated by the life histories of individuals $1, \dots, \ell$. Each Z_n is a stopping time with respect to $(\mathcal{G}_\ell)_\ell$, and $\mathcal{G}_{Z_n} = \mathcal{F}_n$.

We refine the martingale $(M_{n,k})_k$ by adding the contribution from each individual separately. Let τ_i denote the birth time of i, and $N_{i,k}$ the copy of N_k for i (i.e., the number of children i gets at age k). Let

$$\Delta\widehat{M}_{n,i} := \sum_{j=1}^{n-\tau_i} \alpha_{n-\tau_i-j}\bigl(N_{i,j} - \mu_j\bigr), \tag{7.4}$$

$$\widehat{M}_{n,k} := \sum_{i=1}^{k} \Delta\widehat{M}_{n,i}. \tag{7.5}$$

Then $(\widehat{M}_{n,k})_k$ is a $(\mathcal{G}_k)_k$-martingale with $\widehat{M}_{n,\infty} = \widehat{M}_{n,Z_n} = M_{n,n} = -\langle \vec{X}_n, \vec{a}\rangle$, see (6.3)–(6.5), and the conditional quadratic variation

$$\widehat{V}_n := \sum_i \mathbb{E}\bigl((\Delta\widehat{M}_{n,i})^2 \mid \mathcal{G}_{i-1}\bigr) = V_n \tag{7.6}$$

given by (6.7). Moreover, by (7.4) and (7.1),

$$\bigl|\Delta\widehat{M}_{n,i}\bigr| \leqslant C \sum_{j=0}^{\infty} R_1^{n-\tau_i-j}\bigl(N_{i,j} + \mu_j\bigr) = C R_1^{n-\tau_i}\bigl(\widehat{\Xi}_i(R_1^{-1}) + \widehat{\mu}(R_1^{-1})\bigr). \tag{7.7}$$

Define the random variable $U := \widehat{\Xi}(R_1^{-1}) + \widehat{\mu}(R_1^{-1})$. Then $\mathbb{E}\, U^2 < \infty$ by (A6), since $R_1^{-1} < r$. It follows from (7.7) that for some $c > 0$ and every $\varepsilon > 0$, defining $h(x) := \mathbb{E}\bigl(U^2 \mathbf{1}\{U > cx\}\bigr)$,

$$\begin{aligned}\mathbb{E}\bigl(\bigl|\Delta\widehat{M}_{n,i}\bigr|^2 \mathbf{1}\{\bigl|\Delta\widehat{M}_{n,i}\bigr| > \varepsilon\} \mid \mathcal{G}_{i-1}\bigr) &\leqslant C R_1^{2(n-\tau_i)}\, \mathbb{E}\bigl(U^2 \mathbf{1}\{U > c\varepsilon R_1^{\tau_i - n}\}\bigr) \\ &= C R_1^{2(n-\tau_i)} h\bigl(\varepsilon R_1^{\tau_i-n}\bigr) \leqslant C R_1^{2(n-\tau_i)} h\bigl(\varepsilon R_1^{-n}\bigr),\end{aligned} \tag{7.8}$$

Thus,

$$\sum_i \mathbb{E}\bigl(\bigl|\Delta\widehat{M}_{n,i}\bigr|^2 \mathbf{1}\{\bigl|\Delta\widehat{M}_{n,i}\bigr| > \varepsilon\} \mid \mathcal{G}_{i-1}\bigr) \leqslant C \sum_{k=0}^{n} B_k R_1^{2(n-k)} h\bigl(\varepsilon R_1^{-n}\bigr). \tag{7.9}$$

Finally, we normalize $\widehat{M}_{n,k}$ and define $\widetilde{M}_{n,k} := m^{-n/2}\widehat{M}_{n,k}$; this yields a martingale $(\widetilde{M}_{n,k})_k$ with conditional quadratic variation

$$\widetilde{V}_n := \sum_i \mathbb{E}\bigl((\Delta\widetilde{M}_{n,i})^2 \mid \mathcal{G}_{i-1}\bigr) = m^{-n}\widehat{V}_n \xrightarrow{\text{a.s.}} \sigma^2(\vec{a})\mathcal{Z}, \tag{7.10}$$

by (7.6), (7.3) and (2.8). Furthermore, by (7.9),

$$\sum_i \mathbb{E}\bigl(\bigl|\Delta\widetilde{M}_{n,i}\bigr|^2 \mathbf{1}\{\bigl|\Delta\widetilde{M}_{n,i}\bigr| > \varepsilon\} \mid \mathcal{G}_{i-1}\bigr) \leqslant C h\bigl(\varepsilon m^{n/2} R_1^{-n}\bigr) m^{-n} \sum_{k=0}^{n} B_k R_1^{2(n-k)}, \tag{7.11}$$

which tends to 0 a.s. as $n \to \infty$, because $(m^{1/2}R_1^{-1})^n \to \infty$ and consequently $h\bigl(\varepsilon m^{n/2}R_1^{-n}\bigr) \to 0$, and

$$m^{-n}\sum_{k=0}^{n} B_k R_1^{2(n-k)} = m^{-n}\sum_{k=0}^{n} B_{n-k}R_1^{2k} = \sum_{k=0}^{n}\frac{B_{n-k}}{m^{n-k}}\Bigl(\frac{R_1^2}{m}\Bigr)^k = O_{\text{a.s.}}(1), \tag{7.12}$$

by (2.8) and $R_1^2 < m$.

The martingales $(\widetilde{M}_{n,i})_i$ thus satisfy a conditional Lindeberg condition, which together with (7.10) implies, by [6, Corollary 3.2], that, using (7.6),

$$M_{n,n}/V_n^{1/2} = \widehat{M}_{n,Z_n}/\widehat{V}_n^{1/2} = \widetilde{M}_{n,Z_n}/\widetilde{V}_n^{1/2} \xrightarrow{\text{d}} N(0,1) \tag{7.13}$$

as $n \to \infty$; furthermore, the limit is mixing. (The fact that we here sum the martingale differences to a stopping time Z_n instead of a deterministic k_n as in [6] makes no difference.) By (6.5) and (7.3), this yields

$$\langle \vec{X}_n, \vec{a}\rangle / Z_n^{1/2} \xrightarrow{\text{d}} N\bigl(0, \sigma^2(\vec{a})\bigr). \tag{7.14}$$

We can evaluate the asymptotic variance $\sigma^2(\vec{a})$ given in (7.3) by

$$\begin{aligned}
\frac{\sigma^2(\vec{a})}{1-m^{-1}} &= \sum_{\ell=0}^{\infty} m^{-\ell}\sum_{i,j=1}^{\ell}\sigma_{ij}\alpha_{\ell-i}\alpha_{\ell-j} \\
&= \sum_{k,p=0}^{\infty}\sum_{i,j=1}^{\ell}\sigma_{ij}\alpha_k\alpha_p \mathbf{1}\{i+k=j+p\} m^{-i-k} \\
&= \sum_{k,p,i,j}\sigma_{ij}\alpha_k\alpha_p\oint_{|z|=m^{-1/2}} z^{i+k}\bar{z}^{j+p}\frac{|\mathrm{d}z|}{2\pi m^{-1/2}} \\
&= \oint_{|z|=m^{-1/2}}\Bigl|\sum_k \alpha_k z^k\Bigr|^2 \sum_{i,j}\sigma_{ij}z^i\bar{z}^j\frac{|\mathrm{d}z|}{2\pi m^{-1/2}}.
\end{aligned} \tag{7.15}$$

Furthermore, for $|z| = m^{-1/2}$ (and any z with $|z| < R^{-1} = r$ and $\widehat{\mu}(z) \neq 1$), by (6.1) and Lemma 3.9 with $\lambda = z^{-1}$,

$$\begin{aligned}
\sum_{k=0}^{\infty}\alpha_k z^k &= \Bigl\langle \sum_{k=0}^{\infty} z^k T^k(\vec{v}), \vec{a}\Bigr\rangle = \bigl\langle (1-zT)^{-1}(\vec{v}), \vec{a}\bigr\rangle \\
&= \frac{1}{(z-1)(1-\widehat{\mu}(z))}\sum_{\ell} a_\ell\bigl(z^\ell - m^{-\ell}\bigr).
\end{aligned} \tag{7.16}$$

By (7.15)–(7.16), $\sigma^2(\vec{a})$ equals the right-hand side in (2.15). Thus, (7.14) shows convergence as in (2.14) for any finite linear combination of $Z_n^{-1/2}X_{n,k}$, and thus joint convergence in (2.14) by the Cramér–Wold device.

Convergence in $L^2(\ell^2_R)$ follows from this and Lemma 5.1 (with a slightly increased R) by a standard truncation argument; we omit the details.

By (2.15), the variable ζ_k is degenerate only if $\Sigma(z) = 0$ for every z with $|z| = m^{-1/2}$, and thus, by (2.12), $\widehat{\Xi}(z) = \widehat{\mu}(z)$ a.s. for every such z, which by (2.1)–(2.2) implies $N_k = \mu_k$ a.s. for every k. □

8. Proof of Theorem 2.2

We assume in this section that $\gamma_* = m^{-1/2}$ and that (2.16) holds. By Lemma 3.4, the spectral radius $r(T) = \gamma_*^{-1} = m^{1/2}$. Lemma 3.7 applies with $\gamma_* = m^{-1/2}$, and thus $\tilde{R} < m^{1/2}$; we may assume $\tilde{R} > R$.

Fix as in Section 6 a real vector $\vec{a} \in \ell^2_{R^{-1}}$, and define, using (3.40),

$$\beta_i = \beta_i(\vec{a}) := \langle P_i(\vec{v}), \vec{a}\rangle = \langle \vec{v}_i, \vec{a}\rangle = \frac{1}{\gamma_i(\gamma_i - 1)\widehat{\mu}'(\gamma_i)} \sum_{k=0}^{\infty} a_k\bigl(\gamma_i^k - m^{-k}\bigr). \tag{8.1}$$

Then, by (6.1) and Lemma 3.7,

$$\alpha_k = O\bigl(\tilde{R}^k\bigr) + \sum_{i=1}^{q} \lambda_i^k \langle P_i(\vec{v}), \vec{a}\rangle = \sum_{i=1}^{q} \beta_i \lambda_i^k + O\bigl(\tilde{R}^k\bigr) = O\bigl(m^{k/2}\bigr). \tag{8.2}$$

Furthermore, the O's in (8.2) hold uniformly in all $\vec{a}$ with $\|\vec{a}\|_{\ell^2_{R^{-1}}} \leqslant 1$, as does every O in this section.

Define also, for $p, t = 1, \dots, q$,

$$\sigma^*_{pt} := \sum_{i,j=1}^{\infty} \sigma_{ij} \lambda_p^{-i} \lambda_t^{-j}, \tag{8.3}$$

and note that, using (6.8), $|\lambda_p| = m^{1/2}$ and $R < m^{1/2}$,

$$\sum_{i,j=1}^{\ell} \sigma_{ij} \lambda_p^{-i} \lambda_t^{-j} = \sigma^*_{pt} + O\Bigl(\sum_{i>\ell, j\geqslant 1} R^{i+j} (m^{1/2})^{-i-j}\Bigr) = \sigma^*_{pt} + O\bigl((R/m^{1/2})^{\ell}\bigr). \tag{8.4}$$

Let

$$s_\ell := \sum_{i,j=1}^{\ell} \sigma_{ij} \alpha_{\ell-i} \alpha_{\ell-j}. \tag{8.5}$$

Then, by (8.2) and symmetry, using again (6.8) and $|\lambda_p| = m^{1/2}$, and (8.4),

$$\begin{aligned} s_\ell &:= \sum_{i,j=1}^{\ell} \sigma_{ij} \sum_{p=1}^{q} \sum_{t=1}^{q} \beta_p \lambda_p^{\ell-i} \beta_t \lambda_t^{\ell-j} + O\Bigl(\sum_{i,j=1}^{\ell} R^{i+j} m^{(\ell-i)/2} \tilde{R}^{\ell-j}\Bigr) \\ &= \sum_{p=1}^{q} \sum_{t=1}^{q} \beta_p \beta_t \lambda_p^{\ell} \lambda_t^{\ell} \sum_{i,j=1}^{\ell} \sigma_{ij} \lambda_p^{-i} \lambda_t^{-j} + O\bigl((m^{1/2}\tilde{R})^{\ell}\bigr) \\ &= \sum_{p=1}^{q} \sum_{t=1}^{q} \beta_p \beta_t \lambda_p^{\ell} \lambda_t^{\ell} \sigma^*_{pt} + O\bigl((m^{1/2}\tilde{R})^{\ell}\bigr). \end{aligned} \tag{8.6}$$

In particular,

$$s_\ell = O\bigl(m^\ell\bigr). \tag{8.7}$$

It follows by (6.7), (8.5), (2.8), (8.7) and (8.6) that, a.s.,

$$\begin{aligned}
\frac{V_n}{B_n} &= \sum_{\ell=0}^{n} \frac{B_{n-\ell}}{B_n} s_\ell = \sum_{\ell=0}^{n} m^{-\ell}\bigl(1+o(1)+O_{\mathrm{a.s.}}(1)\mathbf{1}\{n-\ell<\log n\}\bigr)s_\ell \\
&= \sum_{\ell=0}^{n} m^{-\ell}s_\ell + o(n) = \sum_{\ell=0}^{n} m^{-\ell}\sum_{p=1}^{q}\sum_{t=1}^{q}\beta_p\beta_t\lambda_p^\ell\lambda_t^\ell\sigma^*_{pt} + o(n) \\
&= \sum_{p=1}^{q}\sum_{t=1}^{q}\beta_p\beta_t\sigma^*_{pt}\sum_{\ell=0}^{n}\Bigl(\frac{\lambda_p\lambda_t}{m}\Bigr)^\ell + o(n).
\end{aligned} \tag{8.8}$$

Recall that $|\lambda_p| = |\lambda_t| = m^{1/2}$, so $|\lambda_p\lambda_t/m| = 1$. Hence, if $\lambda_t = \bar\lambda_p$, then $\sum_{\ell=0}^{n}\bigl(\lambda_p\lambda_t/m\bigr)^\ell = n+1$, while if $\lambda_t \neq \bar\lambda_p$, then $\sum_{\ell=0}^{n}\bigl(\lambda_p\lambda_t/m\bigr)^\ell = O(1)$. Consequently, (8.8) yields, since $B_n/Z_n \xrightarrow{\text{a.s.}} 1-m^{-1}$ by (3.1) and (2.8),

$$\begin{aligned}
\frac{V_n}{nZ_n} \xrightarrow{\text{a.s.}} \sigma^2(\vec a) &:= \frac{m-1}{m}\sum_{p=1}^{q}\sum_{t=1}^{q}\beta_p\beta_t\sigma^*_{pt}\mathbf{1}\{\lambda_t=\bar\lambda_p\} \\
&= \frac{m-1}{m}\sum_{p=1}^{q}|\beta_p|^2\sum_{i,j=1}^{\infty}\sigma_{ij}\lambda_p^{-i}\bar\lambda_p^{-j} \\
&= \frac{m-1}{m}\sum_{p=1}^{q}|\beta_p|^2\Sigma(\gamma_p).
\end{aligned} \tag{8.9}$$

We refine the martingale $(M_{n,k})_k$ to $(\widehat M_{n,k})_k$ as in Section 7, but this time we normalize it to $\widetilde M_{n,k} := (nm^n)^{-1/2}\widehat M_{n,k}$. It follows from (8.9) and (2.8) that the conditional quadratic variation $\widetilde V_n = V_n/(nm^n) \xrightarrow{\text{a.s.}} \sigma^2(\vec a)\mathcal Z$, i.e., (7.10) holds also in the present case. Furthermore, if we now let $R_1 := m^{1/2}$, then (7.1) and (7.7)–(7.9) hold, and it follows that (7.11) is modified to

$$\begin{aligned}
\sum_i \mathbb E\bigl(\bigl|\Delta\widetilde M_{n,i}\bigr|^2\mathbf 1\{\bigl|\Delta\widetilde M_{n,i}\bigr|>\varepsilon\}\mid\mathcal G_{i-1}\bigr) &\leqslant Ch\bigl(\varepsilon n^{1/2}\bigr)\frac{1}{nm^n}\sum_{k=0}^{n}B_km^{n-k} \\
&= O_{\mathrm{a.s.}}\bigl(h\bigl(\varepsilon n^{1/2}\bigr)\bigr) \xrightarrow{\text{a.s.}} 0.
\end{aligned} \tag{8.10}$$

Hence the conditional Lindeberg condition holds in the present case too, and (7.13) follows again by [6, Corollary 3.2], which now by (8.9) and (6.5) yields (mixing)

$$\langle\vec X_n,\vec a\rangle/(nZ_n)^{1/2} \xrightarrow{\text{d}} N\bigl(0,\sigma^2(\vec a)\bigr). \tag{8.11}$$

By (8.9) and (8.1), this proves (2.17)–(2.18).

By (2.18), the variable ζ_k is degenerate only if $\Sigma(\gamma_p)=0$ for every p, and thus, by (2.12), $\widehat\Xi(\gamma_p) = \widehat\mu(\gamma_p)$ a.s.

As in Section 7, convergence in $L^2(\ell^2_R)$ follows by a standard truncation argument, now using the following lemma (with an increased R); we omit the details. $\square$

Lemma 8.1. *Assume* (A1)–(A6), $\gamma_* = m^{-1/2}$ *and* (2.16). *If* $R < m^{1/2}$, *then*

$$\mathbb{E}\,\|\vec{X}_n\|^2_{\ell^2_R} \leqslant Cnm^n \tag{8.12}$$

and

$$\mathbb{E}\,X^2_{n,k} \leqslant Cnm^nR^{-2k} \tag{8.13}$$

for all $n, k \geqslant 0$.

Proof. By (6.5), (6.7), (8.5), (2.7) and (8.7),

$$\mathbb{E}\langle\vec{X}_n,\vec{a}\rangle^2 = \mathbb{E}\,V_n = \mathbb{E}\sum_{\ell=0}^{n} B_{n-\ell}s_\ell \leqslant Cnm^n, \tag{8.14}$$

uniformly for $\|\vec{a}\|_{\ell^2_{R^{-1}}} \leqslant 1$. Taking $\vec{a} = R^k(\delta_{kj})_j$, we obtain (8.13).

Finally, applying (8.13) with R replaced by some R' with $R < R' < m^{1/2}$,

$$\mathbb{E}\,\|\vec{X}_n\|^2_{\ell^2_R} = \sum_{k=0}^{\infty} R^{2k}\,\mathbb{E}\,X^2_{n,k} \leqslant Cnm^n\sum_{k=0}^{\infty}(R/R')^{2k} = Cnm^n. \tag{8.15}$$

$\square$

9. Proof of Theorem 2.3

Assume now that $\gamma_* < m^{-1/2}$. By Lemma 3.4, the spectral radius $r(T) = \gamma_*^{-1} \geqslant m^{1/2}$. We apply Lemma 3.7, assuming as we may that $\tilde{R} > m^{1/2}$. (Otherwise we increase $\tilde{R}$, keeping $\tilde{R} < \gamma_*^{-1}$.) Hence, by (3.38),

$$T^k(\vec{v}) = T_0^k(\vec{v}) + \sum_{i=1}^{q}\lambda_i^kP_i(\vec{v}) = T_0^k(\vec{v}) + \sum_{i=1}^{q}\lambda_i^k\vec{v}_i. \tag{9.1}$$

Thus, by (3.21),

$$\vec{X}_n = -\sum_{k=0}^{n}W_k(TP_0)^{n-k}(\vec{v}) - \sum_{i=1}^{q}\sum_{k=0}^{n}\lambda_i^{n-k}W_k\vec{v}_i. \tag{9.2}$$

Let, recalling (3.8),

$$\check{U}_i := -\sum_{k=0}^{\infty}\gamma_i^kW_k = -\sum_{k=0}^{\infty}\lambda_i^{-k}W_k = -\sum_{\ell=0}^{\infty}\sum_{j=1}^{\infty}\lambda_i^{-\ell-j}W_{\ell,j}, \tag{9.3}$$

noting that by Lemma 3.1 and $|\gamma_i| = \gamma_* < m^{-1/2}$, the sum converges in L^2 and

$$\Bigl\|\check{U}_i + \sum_{k=0}^{n}\lambda_i^{-k}W_k\Bigr\|_2 \leqslant \sum_{k=n+1}^{\infty}C|\lambda_i|^{-k}m^{k/2} \leqslant C\bigl(\gamma_*m^{1/2}\bigr)^n. \tag{9.4}$$

Furthermore, by Lemma 3.1 and (3.39), since $\tilde{R} > m^{1/2}$,

$$\Bigl\|\sum_{k=0}^{n} W_k (TP_0)^{n-k}(\vec{v})\Bigr\|_{L^2(\ell^2_R)} \leqslant \sum_{k=0}^{n} \|W_k\|_2 \cdot \|(TP_0)^{n-k}(\vec{v})\|_{\ell^2_R} \leqslant C\sum_{k=0}^{n} m^{k/2}\tilde{R}^{n-k} \leqslant C\tilde{R}^n. \tag{9.5}$$

By (9.2), (9.5), (9.4), defining $U_i := \bigl(\gamma_i(\gamma_i-1)\widehat{\mu}'(\gamma_i)\bigr)^{-1}\check{U}_i$ so $\check{U}_i\vec{v}_i = U_i\vec{u}_i$ by (3.40),

$$\Bigl\|\gamma_*^n\vec{X}_n - \sum_{i=1}^{q}\bigl(\lambda_i/|\lambda_i|\bigr)^n U_i\vec{u}_i\Bigr\|_{L^2(\ell^2_R)} \leqslant C\gamma_*^n\tilde{R}^n + \sum_{i=1}^{q}\Bigl\|\sum_{k=0}^{n}\lambda_i^{-k}W_k\vec{v}_i + \check{U}_i\vec{v}_i\Bigr\|_{L^2(\ell^2_R)} \leqslant C(\gamma_*\tilde{R})^n + C(\gamma_* m^{1/2})^n \leqslant C(\gamma_*\tilde{R})^n. \tag{9.6}$$

Since $\gamma_*\tilde{R} < 1$, this shows convergence in (2.19) in $L^2(\ell^2_R)$; furthermore, convergence a.s. follows by (9.6) and the Borel–Cantelli lemma.

We have $\mathbb{E}\,U_i = \mathbb{E}\,\check{U}_i = 0$ by (9.3) since $\mathbb{E}\,W_k = 0$ by (3.6)–(3.8). Furthermore, $W_{0,k} = B_{0,k} - \mu_k = N_k - \mu_k$, while $\mathbb{E}\bigl(W_{n,k} \mid \mathcal{F}_0\bigr) = 0$ for $n \geqslant 1$ by (3.6); hence by (3.8), $\mathbb{E}\bigl(W_n \mid \mathcal{F}_0\bigr) = W_{0,n} = N_n - \mu_n$, and thus

$$\mathbb{E}\bigl(\check{U}_i \mid \mathcal{F}_0\bigr) = -\sum_{k=0}^{\infty}\gamma_i^k(N_k - \mu_k) = -\widehat{\Xi}(\gamma_i) + \widehat{\mu}(\gamma_i). \tag{9.7}$$

Hence, U_i is degenerate only if $\widehat{\Xi}(\gamma_i)$ is so. □

10. A stochastic integral calculus

The limit variables ζ_k in Theorems 2.1 and 2.2 can be interpreted as stochastic integrals of certain functions ("*symbols*"); which gives a useful symbolic calculus. There are also some partial related results for Theorem 2.3.

We consider the three cases in Theorems 2.1–2.3 separately.

10.1. The case $\gamma_* > m^{-1/2}$.

Assume throughout this subsection that Theorem 2.1 applies; in particular that $\gamma_* > m^{-1/2}$.

Let ν be the finite measure on the circle $|z| = m^{-1/2}$ given by

$$\mathrm{d}\nu(z) := \frac{m-1}{m}|1-z|^{-2}\,|1-\widehat{\mu}(z)|^{-2}\Sigma(z)\frac{|\mathrm{d}z|}{2\pi m^{-1/2}}, \tag{10.1}$$

and consider an isomorphism $\mathcal{I} : L^2(\nu) \to \mathcal{H}$ of the Hilbert space $L^2(\nu)$ into a Gaussian Hilbert space $\mathcal{H}$, i.e., a Hilbert space of Gaussian random variables; $\mathcal{I}$ can be interpreted as a stochastic integral, see [8, Section VII.2]. We let here $L^2(\nu)$ be the space of complex square-integrable functions, but regard it as a real Hilbert space with the inner product $\langle f, g\rangle_\nu := \operatorname{Re}\int f\bar{g}\,\mathrm{d}\nu$. Then (2.14)–(2.15) can be stated as

$$Z_n^{-1/2}X_{n,k} \xrightarrow{\mathrm{d}} \zeta_k := \mathcal{I}\bigl(z^k - m^{-k}\bigr), \tag{10.2}$$

jointly for all $k \geqslant 0$. This yields a convenient calculus for joint limits.

Example 10.1. Let $k, \ell \geqslant 0$. Then, by (2.10),

$$X_{n-\ell,k} = X_{n,k+\ell} - m^{-k} X_{n,\ell} \tag{10.3}$$

and thus, recalling (2.9), jointly for all $k, \ell \geqslant 0$,

$$\begin{aligned} Z_{n-\ell}^{-1/2} X_{n-\ell,k} &\xrightarrow{\mathrm{d}} m^{\ell/2}\bigl(\zeta_{k+\ell} - m^{-k}\zeta_\ell\bigr) = m^{\ell/2}\mathcal{I}\bigl(z^{k+\ell} - m^{-k} z^\ell\bigr) \\ &= \mathcal{I}\bigl((zm^{1/2})^\ell (z^k - m^{-k})\bigr). \end{aligned} \tag{10.4}$$

Denoting this limit by $\zeta_k^{(\ell)}$, we have of course $\zeta_k^{(\ell)} \overset{\mathrm{d}}{=} \zeta_k$, which corresponds to the fact that $|zm^{1/2}|^\ell = 1$ on the support of ν. More interesting is the joint convergence $(Z_n^{-1/2} X_{n,k}, Z_{n-\ell}^{-1/2} X_{n-\ell,k}) \xrightarrow{\mathrm{d}} (\zeta_k, \zeta_k^{(\ell)})$, with covariance

$$\begin{aligned} \operatorname{Cov}\bigl(\zeta_k, \zeta_k^{(\ell)}\bigr) &= \langle z^k - m^{-k}, (zm^{1/2})^\ell (z^k - m^{-k})\rangle_\nu \\ &= \operatorname{Re} \int_{|z|=m^{-1/2}} (zm^{1/2})^\ell |z^k - m^{-k}|^2 \,\mathrm{d}\nu. \end{aligned} \tag{10.5}$$

The measure ν is by (10.1) absolutely continuous on the circle $|z| = m^{-1/2}$. With the change of variables $z = m^{-1/2} e^{\mathrm{i}\theta}$, we have $(zm^{1/2})^\ell = e^{\mathrm{i}\ell\theta}$ and the Riemann–Lebesgue lemma shows that $\operatorname{Cov}\bigl(\zeta_k, \zeta_k^{(\ell)}\bigr) \to 0$ as $\ell \to \infty$, for fixed every k. Roughly speaking, $X_{n-\ell,k}$ and $X_{n,k}$ are thus essentially uncorrelated when ℓ is large, which justifies the claim in Section 2 that there is only a short-range dependence in this case.

Example 10.2. We can define $X_{n,k}$ by (2.10) also for $k < 0$. Then, the calculations in Example 10.1 apply to any $\ell \geqslant 0$ and any $k \geqslant -\ell$. Hence, replacing n by $n + \ell$ in (10.4), for any fixed ℓ,

$$Z_n^{-1/2} X_{n,k} \xrightarrow{\mathrm{d}} \mathcal{I}\bigl((zm^{1/2})^\ell (z^k - m^{-k})\bigr) \tag{10.6}$$

jointly for all $k \geqslant -\ell$. Since the factor $(zm^{1/2})^\ell$ does not depend on k and has absolute value 1, this means (by changing the isomorphism $\mathcal{I}$) that (10.2) holds jointly for all $k \geqslant -\ell$. Since ℓ is arbitrary, this means that (10.2) holds jointly for all $k \in \mathbb{Z}$. Hence, (2.14)–(2.15) extend to all $k \in \mathbb{Z}$, as claimed in Remark 2.4.

Example 10.3. We have, by (2.10),

$$m^{-j} Z_{n+j} - m^{-j-1} Z_{n+j+1} = m^{-j} X_{n+j+1,1}. \tag{10.7}$$

Hence, by Lemma 5.1, for $j \geqslant 0$,

$$\|m^{-j} Z_{n+j} - m^{-j-1} Z_{n+j+1}\|_2 \leqslant C m^{-j+(n+j+1)/2} = C m^{n/2-j/2}. \tag{10.8}$$

Summing (10.8) for $j \geqslant \ell$ we obtain, recalling (2.8),

$$\|m^{-\ell} Z_{n+\ell} - m^n \mathcal{Z}\|_2 \leqslant C m^{n/2-\ell/2} \tag{10.9}$$

for $n \geqslant 1$ and $\ell \geqslant 0$. Hence, as $\ell \to \infty$, $m^{-n/2}\bigl(m^{-\ell} Z_{n+\ell} - m^n \mathcal{Z}\bigr) \to 0$ in L^2, and thus in probability, uniformly in n. Since $Z_n/m^n \xrightarrow{\text{a.s.}} \mathcal{Z} > 0$, and thus $\sup_n m^n/Z_n < \infty$ a.s., it follows that, still uniformly in n,

$$Z_n^{-1/2}\bigl(m^{-\ell} Z_{n+\ell} - m^n \mathcal{Z}\bigr) \xrightarrow{\text{p}} 0, \qquad \ell \to \infty. \tag{10.10}$$

Define the random variables

$$Y_{n,\ell} := Z_n^{-1/2}\bigl(Z_n - m^{-\ell} Z_{n+\ell}\bigr) = -Z_n^{-1/2} m^{-\ell} X_{n,-\ell}, \qquad \ell \geqslant 0. \tag{10.11}$$

Then, by (10.2) and Example 10.2, for every fixed ℓ,

$$Y_{n,\ell} \xrightarrow{\text{d}} -m^{-\ell}\zeta_{-\ell} = \mathcal{I}\bigl(1 - m^{-\ell} z^{-\ell}\bigr), \qquad n \to \infty. \tag{10.12}$$

Furthermore, by (10.10), $Y_{n,\ell} \xrightarrow{\text{p}} Z_n^{-1/2}\bigl(Z_n - m^n \mathcal{Z}\bigr)$ as $\ell \to \infty$, uniformly in n. Finally, $|mz| = m^{1/2} > 1$ on the support of ν, and thus $1-(mz)^{-\ell} \to 1$ in $L^2(\nu)$ as $\ell \to \infty$; hence $\mathcal{I}\bigl(1 - m^{-\ell} z^{-\ell}\bigr) \to \mathcal{I}(1)$ as $\ell \to \infty$, in L^2 and thus in distribution. It follows that we can let $\ell \to \infty$ in (10.12), see [2, Theorem 4.2], and obtain

$$Z_n^{-1/2}\bigl(Z_n - m^n \mathcal{Z}\bigr) \xrightarrow{\text{d}} \mathcal{I}\bigl(1\bigr), \qquad n \to \infty. \tag{10.13}$$

This is jointly with all (10.2), and thus, jointly for all $k \in \mathbb{Z}$,

$$Z_n^{-1/2}\bigl(Z_{n-k} - m^{n-k} \mathcal{Z}\bigr) = Z_n^{-1/2}\bigl(X_{n,k} + m^{-k}(Z_n - m^n \mathcal{Z})\bigr) \xrightarrow{\text{d}} \mathcal{I}\bigl(z^k\bigr). \tag{10.14}$$

Conversely, (10.2) follows immediately from (10.14).

In the Galton–Watson case (Example 2.5), (10.14) is equivalent to the case $q = 0$ of [7, Theorem (2.10.2)].

10.2. **The case $\gamma_* = m^{-1/2}$.** Assume now that Theorem 2.2 applies; thus $\gamma_* = m^{-1/2}$ and (2.16) holds.

In this case, let ν be the discrete measure , with support Γ_{**},

$$\nu := (m-1) \sum_{p=1}^{q} |1 - \gamma_p|^{-2}\, |\widehat{\mu}'(\gamma_p)|^{-2} \Sigma(\gamma_p) \delta_{\gamma_p}, \tag{10.15}$$

and consider an isomorphism I of $L^2(\nu)$ into a Gaussian Hilbert space as above. Then (2.17)–(2.18) can be stated as (10.2), with the normalizing factor changed from $Z_n^{-1/2}$ to $(nZ_n)^{-1/2}$.

With this change of normalization of $X_{n,k}$, all results in the preceding subsection hold, with one exception: The measure ν has finite support, and thus there exists a sequence $\ell_j \to \infty$ such that $(zm^{1/2})^{\ell_j} \to 1$ as $j \to \infty$ for every $z \in \operatorname{supp}(\nu) = \Gamma_{**}$; hence (10.5) implies $\limsup_{\ell\to\infty} \operatorname{Corr}\bigl(\zeta_k, \zeta_k^{(\ell)}\bigr) = 1$. Hence, although the convergence in (2.18) is mixing, so there is no dependence on the initial generations as in the case $\gamma_* < m^{-1/2}$, there is a dependence over longer ranges than in the case $\gamma_* > m^{-1/2}$.

Furthermore, each ζ_k now belongs to the (typically q-dimensional) space spanned by $\zeta_1, \dots, \zeta_q$, which yields the linear dependence of the limits ζ_k claimed in Section 2.

Example 10.4. In the simplest case, $\Gamma_{**} = \{-m^{1/2}\}$. (See Example 2.6 for an example.) Then $\zeta_k = \bigl((-1)^k m^{-k/2} - m^{-k}\bigr)\zeta$ for some $\zeta \sim N\bigl(0, \nu\{-m^{1/2}\}\bigr)$ and all $k \in \mathbb{Z}$.

Furthermore, $zm^{1/2} = -1$ on $\operatorname{supp}\nu$, and thus (10.4) yields $\zeta_k^{(\ell)} = (-1)^\ell \zeta_k$; in particular, $\zeta_k^{(\ell)} = \zeta_k$ for every even ℓ.

10.3. **The case** $\gamma_* < m^{-1/2}$**.** In this case, there is no limit, but we can argue with the components of the approximating sum in (2.19) in the same way as with ζ_k in Examples 10.1–10.2, and draw the conclusion that (2.19), interpreted component-wise, extends also to $k < 0$, as claimed in Remark 2.4. We omit the details.

11. Random characteristics

A random characteristic is a random function $\chi(t) : [0, \infty) \to \mathbb{R}$ defined on the same probability space as the prototype offspring process Ξ; we assume that each individual x has an independent copy (Ξ_x, χ_x) of (Ξ, χ), and interpret $\chi_x(t)$ as the characteristic of x at age t. We consider as above the lattice case, and define, denoting the birth time of x by τ_x,

$$Z_n^\chi := \sum_{x:\tau_x \leqslant n} \chi_x(n - \tau_x), \tag{11.1}$$

the total characteristic of all individuals at time n. See further Jagers [7]. We assume:

(C) There exists $R_2 < m^{1/2}$ such that $\mathbb{E}[\chi(k)^2] \leqslant C R_2^{2k}$ for some $C < \infty$ and all $k \geqslant 0$.

We define

$$\lambda_k^\chi := \mathbb{E}\,\chi(k), \qquad k \geqslant 0, \tag{11.2}$$

$$\Lambda^\chi(z) := \sum_{k=0}^\infty \lambda_k^\chi z^k, \tag{11.3}$$

$$\lambda^\chi := \bigl(1 - m^{-1}\bigr)\Lambda^\chi\bigl(m^{-1}\bigr) = \sum_{k=0}^\infty \bigl(m^{-k} - m^{-k-1}\bigr)\lambda_k^\chi, \tag{11.4}$$

$$\kappa_{j,k} := \operatorname{Cov}\bigl(\chi(j), N_k\bigr), \tag{11.5}$$

and also $\lambda_k^\chi := 0$ for $k < 0$. Note that (C) implies

$$|\lambda_k^\chi| = |\mathbb{E}\,\chi(k)| \leqslant C R_2^k. \tag{11.6}$$

Hence, the sum in (11.3) converges absolutely at least for $|z| \leqslant m^{-1/2}$; in particular, the sum in (11.4) converges absolutely.

We split the characteristic into its mean $\lambda_k^\chi = \mathbb{E}\,\chi(k)$ and the centered part

$$\tilde{\chi}(k) := \chi(k) - \mathbb{E}\,\chi(k) = \chi(k) - \lambda_k^\chi. \tag{11.7}$$

We define

$$V_{n,k}^{\chi} := \sum_{x:\tau_x=n} \tilde{\chi}_x(k) = \sum_{x:\tau_x=n} \bigl(\chi_x(k) - \lambda_k^{\chi}\bigr) = \sum_{x:\tau_x=n} \chi_x(k) - \lambda_k^{\chi} B_n. \tag{11.8}$$

Then, (11.1) implies

$$Z_n^{\tilde{\chi}} = \sum_{k=0}^{n} V_{n-k,k}^{\chi} = \sum_{k=0}^{\infty} V_{n-k,k}^{\chi} \tag{11.9}$$

and, furthermore,

$$Z_n^{\chi} = \sum_{k=0}^{n} \bigl(V_{n-k,k}^{\chi} + \lambda_k^{\chi} B_{n-k}\bigr) = Z_n^{\tilde{\chi}} + \sum_{k=0}^{\infty} \lambda_k^{\chi} B_{n-k}. \tag{11.10}$$

Hence, recalling (11.4), (3.1) and (2.10), we have the decomposition

$$\begin{aligned} Z_n^{\chi} - \lambda^{\chi} Z_n &= Z_n^{\tilde{\chi}} + \sum_{k=0}^{\infty} \lambda_k^{\chi} \bigl(B_{n-k} - (m^{-k} - m^{-k-1}) Z_n\bigr) \\ &= Z_n^{\tilde{\chi}} + \sum_{k=0}^{\infty} \lambda_k^{\chi} \bigl(X_{n,k} - X_{n,k+1}\bigr) \\ &= Z_n^{\tilde{\chi}} + \sum_{k=1}^{n} \bigl(\lambda_k^{\chi} - \lambda_{k-1}^{\chi}\bigr) X_{n,k} = Z_n^{\tilde{\chi}} + \langle \vec{X}_n, \Delta\vec{\lambda}^{\chi} \rangle, \end{aligned} \tag{11.11}$$

where $\Delta\vec{\lambda}^{\chi}$ is the vector $\bigl(\lambda_k^{\chi} - \lambda_{k-1}^{\chi}\bigr)_{k=0}^{\infty}$. Here $\Delta\vec{\lambda}^{\chi} \in \ell^2_{R^{-1}}$ by (11.6), and thus the asymptotic behaviour of $\langle \vec{X}_n, \Delta\vec{\lambda}^{\chi} \rangle$ is given by Theorems 2.1–2.3.

The term $Z_n^{\tilde{\chi}}$ in (11.11) is asymptotically normal after normalization, for any value of γ_*, as shown by the following theorem. (Note that the assumption $\mathbb{E}\,\chi(k) = 0$ is equivalent to $\chi = \tilde{\chi}$.)

Theorem 11.1. *Assume* (A1)–(A6) *and* (C). *If* $\mathbb{E}\,\chi(k) = 0$ *for every* $k \geqslant 0$, *then as* $n \to \infty$,

$$Z_n^{-1/2} Z^{\chi} \xrightarrow{\mathrm{d}} \zeta^{\chi}, \tag{11.12}$$

for some normal random variable ζ^{χ} *with mean* $\mathbb{E}\,\zeta^{\chi} = 0$ *and variance*

$$\operatorname{Var}\bigl(\zeta^{\chi}\bigr) = \frac{m-1}{m} \sum_{k=0}^{\infty} m^{-k} \operatorname{Var}\bigl(\chi(k)\bigr). \tag{11.13}$$

Before proving Theorem 11.1, we note that in the case $\gamma_* > m^{-1/2}$, Theorems 11.1 and 2.1 show that $Z_n^{\tilde{\chi}}$ and $\langle \vec{X}_n, \Delta\vec{\lambda}^{\chi} \rangle$ in (11.11) both are asymptotically normal after normalization by $Z_n^{1/2}$. In this case, as shown below, the two terms are jointly asymptotically normal, leading by (11.11) to the following extension of Theorem 2.1 (which is the deterministic case $\chi(k) = \sum_{j\leqslant k} a_j$).

Theorem 11.2. *Assume* (A1)–(A6), (B) *and* (C). *Then, as* $n \to \infty$,

$$Z_n^{-1/2}\bigl(Z^\chi - \lambda^\chi Z_n\bigr) \xrightarrow{\mathrm{d}} \zeta^\chi, \tag{11.14}$$

for some normal random variable ζ^χ *with mean* $\mathbb{E}\,\zeta^\chi = 0$ *and variance*

$$\begin{aligned}\operatorname{Var}\bigl(\zeta^\chi\bigr) = \frac{m-1}{m}\Biggl(&\sum_{k=0}^\infty m^{-k}\operatorname{Var}\bigl(\chi(k)\bigr)\\ &- 2\oint_{|z|=m^{-1/2}} \frac{(1-z)\Lambda^\chi(z)-\lambda^\chi}{(z-1)(1-\widehat{\mu}(z))}\sum_{k=0}^\infty\sum_{j=1}^\infty \kappa_{kj} z^j \bar{z}^k \frac{|\mathrm{d}z|}{2\pi m^{-1/2}}\\ &+ \oint_{|z|=m^{-1/2}} \frac{|(1-z)\Lambda^\chi(z)-\lambda^\chi|^2}{|1-z|^2\,|1-\widehat{\mu}(z)|^2}\sum_{i,j}\sigma_{ij} z^i \bar{z}^j \frac{|\mathrm{d}z|}{2\pi m^{-1/2}}.\Biggr)\end{aligned} \tag{11.15}$$

Remark 11.3. In both Theorems 11.1 and 11.2, joint asymptotic normality for several characteristics, with a corresponding formula for asymptotic covariances, follow by the proof, or by the Cramér–Wold device.

Proof of Theorems 11.1 and 11.2. We use results from Section 5, and assume as we may that R is chosen with $R_2 < R < m^{1/2}$.

Given B_{n-k}, $V^\chi_{n-k,k}$ is the sum of B_{n-k} independent copies of $\tilde{\chi}(k) = \chi(k) - \mathbb{E}\,\chi(k)$. Hence, using (C), (2.7) and $B_{n-k} \leqslant Z_{n-k}$,

$$\mathbb{E}\bigl(V^\chi_{n-k,k}\bigr)^2 = \mathbb{E}\bigl(\mathbb{E}\bigl(V^\chi_{n-k,k}\bigr)^2 \mid B_{n-k}\bigr) = \operatorname{Var}\bigl(\chi(k)\bigr)\,\mathbb{E}\,B_{n-k} \leqslant Cm^{n-k}R_2^{2k} \tag{11.16}$$

and, using (11.6) and Lemma 5.1,

$$\mathbb{E}\bigl(\lambda^\chi_k(X_{n,k} - X_{n,k+1})\bigr)^2 \leqslant CR_2^{2k}\bigl(\mathbb{E}\,X^2_{n,k} + \mathbb{E}\,X^2_{n,k+1}\bigr) \leqslant Cm^n(R_2/R)^{2k}. \tag{11.17}$$

Since we assume $R_2 < R < m^{1/2}$, it follows by standard arguments that if we replace χ by the truncated characteristic $\chi_K(k) := \chi(k)\mathbf{1}\{k \leqslant K\}$, then the error $Z_n^{-1/2}\bigl(Z_n^\chi - \lambda^\chi Z_n - (Z_n^{\chi_K} - \lambda^{\chi_K} Z_n)\bigr)$ tends to 0 in probability as $K \to \infty$, uniformly in n, and as a consequence, see [2, Theorem 4.2], it suffices to prove both theorems for the truncated characteristic χ_K. Hence we may in the sequel assume (changing notation) that $\chi(k) = 0$ for $k > K$, for some $K < \infty$.

Let $\vec{\vartheta} = (\vartheta_0, \vartheta_1, \dots)$ be a random vector such that $(\vec{\vartheta}, \vec{\eta})$ is jointly normal with mean 0 and covariances given by (4.1) and

$$\operatorname{Cov}(\vartheta_j, \vartheta_k) = \operatorname{Cov}\bigl(\chi(j), \chi(k)\bigr), \tag{11.18}$$

$$\operatorname{Cov}(\vartheta_j, \eta_k) = \kappa_{j,k} := \operatorname{Cov}\bigl(\chi(j), N_k\bigr). \tag{11.19}$$

Let $\bigl(\vec{\vartheta}^{(k)}, \vec{\eta}^{(k)}\bigr)$ be independent copies of $(\vec{\vartheta}, \vec{\eta})$.

The proof of Lemma 4.1 extends to show that (4.2) holds jointly with

$$Z_n^{-1/2} V^\chi_{n-k,k} \xrightarrow{\mathrm{d}} \bigl(1 - m^{-1}\bigr)^{1/2} m^{-k/2} \vartheta_k^{(k)}, \qquad k \geqslant 0. \tag{11.20}$$

Summing (11.20) over $k \leqslant K$, we obtain

$$Z_n^{-1/2} Z_n^{\tilde{\chi}} \xrightarrow{\mathrm{d}} \zeta^{\chi} := \bigl(1-m^{-1}\bigr)^{1/2} \sum_{k=0}^{\infty} m^{-k/2} \vartheta_k^{(k)}, \tag{11.21}$$

which yields (11.12) and (11.13) in the case $\chi = \tilde{\chi}$; recall that the terms $\vartheta_k^{(k)}$ are independent. This completes the proof of Theorem 11.1.

In the remainder of the proof, we thus consider Theorem 11.2, and thus assume that (B) holds. We have just shown that (4.2) holds jointly with (11.20). Hence, by the proof in Section 5, (5.11) holds jointly with (11.20) for all k, and thus also with (11.21). Consequently, by (11.11),

$$\begin{aligned} &\bigl(1-m^{-1}\bigr)^{-1/2} Z_n^{-1/2} \bigl(Z_n^{\chi} - \lambda^{\chi} Z_n\bigr) \\ &\qquad \xrightarrow{\mathrm{d}} \sum_{k=0}^{\infty} m^{-k/2} \vartheta_k^{(k)} - \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} m^{-(k+j)/2} \eta_j^{(k+j)} \langle T^k(\vec{v}), \Delta\vec{\lambda}^{\chi} \rangle. \end{aligned} \tag{11.22}$$

Write the right-hand side as $A_1 - A_2$, and note that A_1 and A_2 are jointly normal with means 0. It remains to calculate $\operatorname{Var}(A_1 - A_2)$.

Since the terms in the sum A_1 are independent, we have, cf. (11.21) and (11.13),

$$\operatorname{Var}(A_1) = \sum_{k=0}^{\infty} m^{-k} \operatorname{Var}\bigl(\vartheta_k\bigr) = \sum_{k=0}^{\infty} m^{-k} \operatorname{Var}\bigl(\chi(k)\bigr), \tag{11.23}$$

which yields the first term in (11.15),

$\operatorname{Var}(A_2)$ was calculated in Section 5, see (5.14) and (2.15), which yields the last term in (11.15), using $\sum_k (\lambda_k^{\chi} - \lambda_{k-1}^{\chi}) z^k = (1-z)\Lambda^{\chi}(z)$ and (11.4).

Finally, using (11.19) and (5.21),

$$\begin{aligned} \operatorname{Cov}(A_1, A_2) &= \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} m^{-(k+j)} \kappa_{k+j,j} \langle T^k(\vec{v}), \Delta\vec{\lambda}^{\chi} \rangle \\ &= \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} \langle T^k(\vec{v}), \Delta\vec{\lambda}^{\chi} \rangle \oint_{|z|=m^{-1/2}} z^{k+j} \sum_{\ell=0}^{\infty} \bar{z}^{\ell} \kappa_{\ell,j} \frac{|\mathrm{d}z|}{2\pi m^{-1/2}} \\ &= \oint_{|z|=m^{-1/2}} \langle (1-zT)^{-1}(\vec{v}), \Delta\vec{\lambda}^{\chi} \rangle \sum_{\ell=0}^{\infty} \sum_{j=1}^{\infty} z^j \bar{z}^{\ell} \kappa_{\ell,j} \frac{|\mathrm{d}z|}{2\pi m^{-1/2}} \\ &= \oint_{|z|=m^{-1/2}} \frac{(1-z)\Lambda^{\chi}(z) - (1-m^{-1})\Lambda^{\chi}(m^{-1})}{(z-1)(1-\widehat{\mu}(z))} \sum_{\ell=0}^{\infty} \sum_{j=1}^{\infty} z^j \bar{z}^{\ell} \kappa_{\ell,j} \frac{|\mathrm{d}z|}{2\pi m^{-1/2}} \end{aligned} \tag{11.24}$$

The result (11.15) follows by combining (11.23), (11.24) and (2.15), recalling (11.4). $\square$

Theorem 11.2 yields asymptotic normality of Z_n^{χ} when $\gamma_* > m^{-1/2}$, and Theorem 11.1 shows the same for any γ_* in the special case when $\mathbb{E}\,\chi(k) = 0$

for every k. It remains to consider the case when $\lambda_k^\chi = \mathbb{E}\,\chi(k) \neq 0$ for some k and $\gamma_* \leqslant m^{-1/2}$. If $\gamma_* = m^{-1/2}$ and (2.16) holds, then Theorem 2.2 shows that $\langle \vec{X}_n, \Delta\vec{\lambda}^\chi\rangle/\sqrt{nZ_n} \xrightarrow{\mathrm{d}} N(0,\sigma^2)$, where σ^2 is given by (2.18) and $\sigma^2 > 0$ except in degenerate cases. Since Theorem 11.1 implies that $Z_n^{\tilde\chi}/\sqrt{nZ_n} \xrightarrow{\mathrm{p}} 0$, it follows from (11.11) that $(Z_n^\chi - \lambda^\chi Z_n)/\sqrt{nZ_n} \xrightarrow{\mathrm{d}} N(0,\sigma^2)$. Similarly, if $\gamma_* < m^{-1/2}$, then Theorem 11.1 implies $\gamma_*^n Z_n^{\tilde\chi} \xrightarrow{\mathrm{p}} 0$, and (11.11) shows that $Z_n^\chi - \lambda^\chi Z_n$ has the same (oscillating) asymptotic behaviour as $\langle \vec{X}_n, \Delta\vec{\lambda}^\chi\rangle$, given by Theorem 2.3.

Summarizing, if $\gamma_* \leqslant m^{-1/2}$, then the randomness in the characteristic χ only gives an effect of smaller order than the mean $\mathbb{E}\,\chi$, and unless the mean vanishes (or the limits degenerate), Z_n^χ has the same asymptotic behaviour as if χ is replaced by the deterministic $\mathbb{E}\,\chi$, which is treated by Theorems 2.2 and 2.3.

Example 11.4. We have in the present paper for simplicity assumed (A4), that there are no deaths. Suppose now, more generally, that each individual has a random lifelength $\ell \leqslant \infty$, as usual with i.i.d. copies (Ξ_x, ℓ_x) for all individuals x. The results in Section 2 apply if we ignore deaths and let Z_n denote the number of individuals born up to time n, living or dead. Moreover, the number of living individuals at time n is Z_n^χ, for the characteristic $\chi(k) := \mathbf{1}\{\ell > k\}$.

Similarly, for example, the number of living individuals at time $n-j$ is $Z_n^{\chi_j}$ with $\chi_j(k) := \mathbf{1}\{\ell > k-j \geqslant 0\}$. The analogue of $X_{n,j}$ in (2.10) but counting only living individuals is thus given by $Z_n^{\chi_j - m^{-j}\chi}$, and results extending Theorems 2.1–2.3 without assuming (A4) follow. We leave the details to the reader.

Acknowledgement. I thank Peter Jagers and Olle Nerman for helpful comments.

References

[1] Søren Asmussen and Heinrich Hering, *Branching Processes.* Birkhäuser, Boston, 1983.
[2] Patrick Billingsley, *Convergence of Probability Measures.* Wiley, New York, 1968.
[3] Harald Bohr, Zur Theorie der Fastperiodischen Funktionen. I–III. *Acta Math.* **45** (1925), no. 1, 29–127; **46** (1925), no. 1-2, 101–214; **47** (1926), no. 3, 237–281.
[4] Nelson Dunford and Jacob T. Schwartz, *Linear Operators. I. General Theory.* Interscience Publishers, New York, 1958.
[5] Peter L. Duren, *Theory of H^p Spaces.* Academic Press, New York, 1970.
[6] P. Hall and C. C. Heyde, *Martingale Limit Theory and its Application.* Academic Press, New York, 1980.
[7] Peter Jagers, *Branching Processes with Biological Applications.* John Wiley & Sons, London, 1975.
[8] Svante Janson, *Gaussian Hilbert Spaces.* Cambridge Univ. Press, Cambridge, UK, 1997.
[9] Svante Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Process. Appl.* **110** (2004), 177–245.

[10] Olle Nerman, On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrsch. Verw. Gebiete* **57** (1981), no. 3, 365–395.

[11] Nicolas Pouyanne, An algebraic approach to Pólya processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** (2008), no. 2, 293–323.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

E-mail address: `svante.janson@math.uu.se`

URL: `http://www2.math.uu.se/~svante/`