Chapter 2
Galton–Watson Trees

We recall a few elementary properties of supercritical Galton–Watson trees, and introduce the notion of size-biased trees. As an application, we give in Sect. 2.3 the beautiful conceptual proof by Lyons et al. [176] of the Kesten–Stigum theorem for the branching process.

The goal of this brief chapter is to give an *avant-goût* of the spinal decomposition theorem, in the simple setting of the Galton–Watson tree. If you are already familiar with any form of the spinal decomposition theorem, this chapter can be skipped.

2.1 The Extinction Probability

Consider a Galton–Watson process, also referred to as a Bienaymé–Galton–Watson process, with each particle (or: individual) having i children with probability p_i (for $i \ge 0$; $\sum_{j=0}^{\infty} p_j = 1$), starting with one initial ancestor. To avoid trivial discussions, we assume throughout that $p_0 + p_1 < 1$.

Let Z_n denote the number of particles in the n-th generation. By definition, if $Z_n = 0$ for a certain n, then $Z_j = 0$ for all $j \ge n$. We write

$$q := \mathbf{P}\{Z_n = 0 \text{ eventually}\}, \qquad \text{(extinction probability)}$$

$$m := \mathbf{E}(Z_1) = \sum_{i=0}^{\infty} i p_i \in (0, \infty]. \qquad \text{(mean number of offspring of each individual)}$$

Theorem 2.1

(i) *The extinction probability q is the smallest root of the equation $f(s) = s$ for $s \in [0, 1]$, where $f(s) := \sum_{i=0}^{\infty} s^i p_i$, $0^0 := 1$.*
(ii) *In particular, $q = 1$ if $m \le 1$, and $q < 1$ if $1 < m \le \infty$.*

Z. Shi, *Branching Random Walks*, Lecture Notes in Mathematics 2151,
DOI 10.1007/978-3-319-25372-5_2

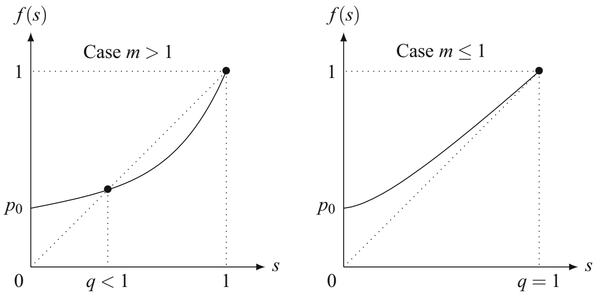

Fig. 2.1 Generating function of the reproduction law

Proof By definition, $f(s) = \mathbf{E}(s^{Z_1})$, and $\mathbf{E}(s^{Z_n} \,|\, Z_{n-1}) = f(s)^{Z_{n-1}}$. So $\mathbf{E}(s^{Z_n}) = \mathbf{E}(f(s)^{Z_{n-1}})$, which leads to $\mathbf{E}(s^{Z_n}) = f_n(s)$ for any $n \geq 1$, where f_n denotes the n-th fold composition of f. In particular, $\mathbf{P}(Z_n = 0) = f_n(0)$.

Since $\{Z_n = 0\} \subset \{Z_\ell = 0\}$ for all $n \leq \ell$, we have

$$q = \mathbf{P}\Big(\bigcup_n \{Z_n = 0\}\Big) = \lim_{n\to\infty} \mathbf{P}(Z_n = 0) = \lim_{n\to\infty} f_n(0).$$

The function $f : [0, 1] \to \mathbb{R}$ is increasing and strictly convex, with $f(0) = p_0 \geq 0$ and $f(1) = 1$. It has at most two fixed points. Note that $m = f'(1-)$. See Fig. 2.1.

If $m \leq 1$, then $p_0 > 0$, and $f(s) > s$ for all $s \in [0, 1)$. So $f_n(0) \to 1$. In other words, $q = 1$ is the unique root of $f(s) = s$.

If $m \in (1, \infty]$, then $f_n(0)$ converges increasingly to the unique root of $f(s) = s$, $s \in [0, 1)$. In particular, $q < 1$. □

It follows that in the subcritical case (i.e., $m < 1$) and in the critical case ($m = 1$), there is extinction with probability 1, whereas in the supercritical case ($m > 1$), the system survives with positive probability.

If $m < \infty$, we can define

$$M_n := \frac{Z_n}{m^n}, \qquad n \geq 0.$$

Since (M_n) is a non-negative martingale with respect to the natural filtration of (Z_n), we have $M_n \to M_\infty$ a.s., where M_∞ is a non-negative random variable. By Fatou's lemma, $\mathbf{E}(M_\infty) \leq \liminf_{n\to\infty} \mathbf{E}(M_n) = 1$. It is, however, possible that $M_\infty = 0$. So it is important to know whether $\mathbf{P}(M_\infty > 0)$ is positive.

If there is extinction, then trivially $M_\infty = 0$. In particular, by Theorem 2.1, we have $M_\infty = 0$ a.s. if $m \leq 1$. What happens if $m > 1$?

Lemma 2.2 *Assume* $m < \infty$. *Then* $\mathbf{P}(M_\infty = 0)$ *is either* q *or* 1.

Proof We already know that $M_\infty = 0$ a.s. if $m \leq 1$. So let us assume $1 < m < \infty$.

By definition, $Z_{n+1} = \sum_{i=1}^{Z_1} Z_n^{(i)}$ (notation: $\sum_{\varnothing} := 0$), where $Z_n^{(i)}$, $i \geq 1$, are copies of Z_n, independent of each other and of Z_1. Dividing both sides by m^n and letting $n \to \infty$, it follows that mM_∞ has the law of $\sum_{i=1}^{Z_1} M_\infty^{(i)}$, where $M_\infty^{(i)}$, $i \geq 1$, are copies of M_∞, independent of each other and of Z_1. Hence $\mathbf{P}(M_\infty = 0) = \mathbf{E}[\mathbf{P}(M_\infty = 0)^{Z_1}] = f(\mathbf{P}(M_\infty = 0))$, i.e., $\mathbf{P}(M_\infty = 0)$ is a root of $f(s) = s$, so $\mathbf{P}(M_\infty = 0) = q$ or 1. □

Theorem 2.3 (Kesten and Stigum [155]) *Assume* $1 < m < \infty$. *Then*

$$\mathbf{E}(M_\infty) = 1 \;\Leftrightarrow\; \mathbf{P}(M_\infty > 0 \,|\, \text{non-extinction}) = 1 \;\Leftrightarrow\; \mathbf{E}(Z_1 \ln_+ Z_1) < \infty,$$

where $\ln_+ x := \ln \max\{x,\, 1\}$.

Theorem 2.3 says that $\mathbf{E}(M_\infty) = 1 \Leftrightarrow \mathbf{P}(M_\infty = 0) = q \Leftrightarrow \sum_{i=1}^{\infty} p_i \, i \ln i < \infty$.

The proof of Theorem 2.3 is postponed to Sect. 2.3. We will see that the condition $\mathbf{E}(Z_1 \ln_+ Z_1) < \infty$, apparently technical, is quite natural.

2.2 Size-Biased Galton–Watson Trees

In order to introduce size-biased Galton–Watson trees, let us view the tree as a random element in a probability space $(\Omega, \mathscr{F}, \mathbf{P})$, using the standard formalism.

Let $\mathscr{U} := \{\varnothing\} \cup \bigcup_{k=1}^{\infty} (\mathbb{N}^*)^k$, where $\mathbb{N}^* := \{1, 2, \ldots\}$. For elements u and v of $\mathscr{U}$, let uv be the concatenated element, with $u\varnothing = \varnothing u = u$.

A tree ω is a subset of $\mathscr{U}$ satisfying the following properties: (i) $\varnothing \in \omega$; (ii) if $uj \in \omega$ for some $j \in \mathbb{N}^*$, then $u \in \omega$; (iii) if $u \in \omega$, then $uj \in \omega$ if and only if $1 \leq j \leq N_u(\omega)$ for some non-negative integer $N_u(\omega)$.

In words, $N_u(\omega)$ is the number of children of the vertex u. Vertices of ω are labeled by their line of descent: the vertex $u = i_1 \ldots i_n \in \mathscr{U}$ stands for the i_n-th child of the i_{n-1}-th child of ... of the i_1-th child of the initial ancestor $\varnothing$. See Fig. 2.2.

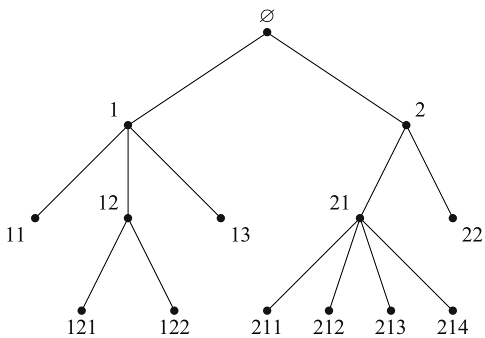

Fig. 2.2 Vertices of a tree as elements of $\mathscr{U}$

Let Ω be the space of all trees, endowed with a σ-field $\mathscr{F}$ defined as follows. For $u \in \mathscr{U}$, let $\Omega_u := \{\omega \in \Omega : u \in \omega\}$ be the subspace of Ω consisting of all the trees containing u as a vertex. [In particular, $\Omega_\varnothing = \Omega$.] Let $\mathscr{F} := \sigma\{\Omega_u, \ u \in \mathscr{U}\}$.

Let $\mathbb{T} : \Omega \to \Omega$ be the identity application.

Let $(p_k, \ k \ge 0)$ be a probability, i.e., $p_k \ge 0$ for all $k \ge 0$, and $\sum_{k=0}^{\infty} p_k = 1$. There exists a probability $\mathbf{P}$ on $(\Omega, \mathscr{F})$ [203] such that the law of $\mathbb{T}$ under $\mathbf{P}$ is the law of the Galton–Watson tree with reproduction distribution (p_k).

Let $\mathscr{F}_n := \sigma\{\Omega_u, \ u \in \mathscr{U}, \ |u| \le n\}$, where $|u|$ is the length of u (or the generation of the vertex u in the language of trees). Note that $\mathscr{F}$ is the smallest σ-field containing all the $\mathscr{F}_n$.

For any tree $\omega \in \Omega$, let $Z_n(\omega)$ be the number of individuals in the n-th generation, i.e., $Z_n(\omega) := \#\{u \in \mathscr{U} : \ u \in \omega, \ |u| = n\}$. It is easily checked that for any n, Z_n is a random variable taking values in $\mathbb{N} := \{0, 1, 2, \ldots\}$.

Assume now $m < \infty$. Since (M_n) is a non-negative martingale, we can define $\mathbf{Q}$ to be the probability on $(\Omega, \mathscr{F})$ such that for any n,

$$\mathbf{Q}_{|_{\mathscr{F}_n}} = M_n \bullet \mathbf{P}_{|_{\mathscr{F}_n}},$$

where $\mathbf{P}_{|_{\mathscr{F}_n}}$ and $\mathbf{Q}_{|_{\mathscr{F}_n}}$ are the restrictions of $\mathbf{P}$ and $\mathbf{Q}$ on $\mathscr{F}_n$, respectively.

For any n, $\mathbf{Q}(Z_n > 0) = \mathbf{E}[\mathbf{1}_{\{Z_n>0\}} M_n] = \mathbf{E}[M_n] = 1$, which yields $\mathbf{Q}(Z_n > 0, \ \forall n) = 1$: there is almost sure non-extinction of the Galton–Watson tree $\mathbb{T}$ under the new probability $\mathbf{Q}$. The Galton–Watson tree $\mathbb{T}$ under $\mathbf{Q}$ is called a size-biased Galton–Watson tree. We intend to give a description of its paths.

We start with a lemma. Let $N := N_\varnothing$. If $N \ge 1$, we write $\mathbb{T}_1, \mathbb{T}_2, \ldots, \mathbb{T}_N$ for the N subtrees rooted at each of the N individuals in the first generation.

Lemma 2.4 *Let $k \ge 1$. If $A_1, A_2, \ldots, A_k$ are elements of $\mathscr{F}$, then*

$$\mathbf{Q}(N = k, \ \mathbb{T}_1 \in A_1, \ldots, \ \mathbb{T}_k \in A_k) = \frac{kp_k}{m} \frac{1}{k} \sum_{i=1}^{k} \mathbf{P}(A_1) \cdots \mathbf{P}(A_{i-1}) \mathbf{Q}(A_i) \mathbf{P}(A_{i+1}) \cdots \mathbf{P}(A_k). \tag{2.1}$$

Proof By the monotone class theorem, we may assume, without loss of generality, that $A_1, A_2, \ldots, A_k$ are elements of $\mathscr{F}_n$, for some n. Write $\mathbf{Q}_{(2.1)}$ for $\mathbf{Q}(N = k, \ \mathbb{T}_1 \in A_1, \ldots, \ \mathbb{T}_k \in A_k)$. Then

$$\mathbf{Q}_{(2.1)} = \mathbf{E}\Big(\frac{Z_{n+1}}{m^{n+1}} \mathbf{1}_{\{N=k, \ \mathbb{T}_1 \in A_1, \ldots, \ \mathbb{T}_k \in A_k\}} \Big).$$

On $\{N = k\}$, we can write $Z_{n+1} = \sum_{i=1}^{k} Z_n^{(i)}$, where $Z_n^{(i)}$ is the number of individuals in the n-th generation of the subtree rooted at the i-th individual in the first generation. Hence

$$\mathbf{Q}_{(2.1)} = \frac{1}{m^{n+1}} \mathbf{P}(N = k) \sum_{i=1}^{k} \mathbf{E}\Big\{ Z_n^{(i)} \mathbf{1}_{\{\mathbb{T}_1 \in A_1, \ldots, \ \mathbb{T}_k \in A_k\}} \ \Big| \ N = k \Big\}.$$

We have $\mathbf{P}(N = k) = p_k$, and

$$\mathbf{E}\{Z_n^{(i)}\,\mathbf{1}_{\{\mathbb{T}_1\in A_1,\ldots,\,\mathbb{T}_k\in A_k\}}\,|\,N = k\} = \mathbf{E}[Z_n\,\mathbf{1}_{\{\mathbb{T}\in A_i\}}]\prod_{j\neq i}\mathbf{P}(A_j),$$

which is $m^n\,\mathbf{Q}(A_i)\prod_{j\neq i}\mathbf{P}(A_j)$. The lemma is proved. □

It follows from Lemma 2.4 that the root $\varnothing$ of the size-biased Galton–Watson tree has the biased distribution, i.e., having k children with probability $\frac{kp_k}{m}$; among the individuals in the first generation, one of them is chosen randomly according to the uniform distribution: the subtree rooted at this vertex is a size-biased Galton–Watson tree, whereas the subtrees rooted at all other vertices in the first generation are usual Galton–Watson trees, and all these subtrees are independent.

We iterated the procedure, and obtain a decomposition of the size-biased Galton–Watson tree into an (infinite) spine and i.i.d. copies of the usual Galton–Watson tree: The root $\varnothing =: w_0$ has the biased distribution, i.e., having k children with probability $\frac{kp_k}{m}$. Among the children of the root, one of them is chosen randomly according to the uniform distribution, as the element of the spine in the first generation; let us denote this element by w_1. We attach subtrees rooted at all other children; they are independent copies of the usual Galton–Watson tree. The vertex w_1 has the biased distribution. Among the children of w_1, we choose at random one of them as the element of the spine in the second generation, denoted by w_2. Independent copies of the usual Galton–Watson tree are attached as subtrees rooted at all other children of w_1, whereas w_2 has the biased distribution. The system goes on indefinitely. See Fig. 2.3.

Having the application of the next section in mind, let us connect the size-biased Galton–Watson tree to the branching process with immigration. The latter starts with

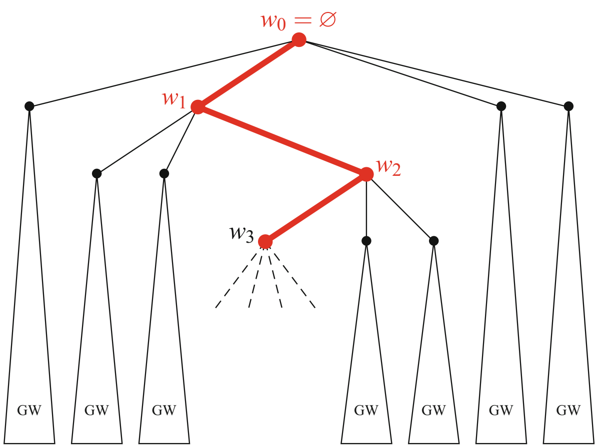

Fig. 2.3 A size-biased Galton–Watson tree

no individual (say), and is governed by a reproduction law and an immigration law. At generation n (for $n \geq 1$), Y_n new individuals are added into the system, while all individuals regenerate independently and following the same reproduction law; we assume that $(Y_n,\ n \geq 1)$ is a collection of i.i.d. random variables following the same immigration law, and independent of everything else up to that generation.

The size-biased Galton–Watson tree tells us that $(Z_n - 1,\ n \geq 0)$ under $\mathbf{Q}$ is a branching process with immigration, whose immigration law is that of $\widehat{N} - 1$, with $\mathbf{P}(\widehat{N} = k) := \frac{kp_k}{m}$, for $k \geq 1$.

2.3 Application: The Kesten–Stigum Theorem

We start with a dichotomy theorem for branching processes with immigration.

Theorem 2.5 (Seneta [219]) *Let Z_n be the number of individuals in the n-th generation of a branching process with immigration (Y_n). Assume that $1 < m < \infty$, where m denotes the expectation of the reproduction law.*

(i) *If* $\mathbf{E}(\ln_+ Y_1) < \infty$*, then* $\lim_{n\to\infty} \frac{Z_n}{m^n}$ *exists and is finite almost surely.*
(ii) *If* $\mathbf{E}(\ln_+ Y_1) = \infty$*, then* $\limsup_{n\to\infty} \frac{Z_n}{m^n} = \infty$, a.s.

Proof (ii) Assume $\mathbf{E}(\ln_+ Y_1) = \infty$. By the Borel–Cantelli lemma [102, Theorem 2.5.9], $\limsup_{n\to\infty} \frac{\ln Y_n}{n} = \infty$ a.s. Since $Z_n \geq Y_n$, it follows that for any $c > 1$, $\limsup_{n\to\infty} \frac{Z_n}{c^n} = \infty$, a.s.
(i) Assume now $\mathbf{E}(\ln_+ Y_1) < \infty$. By the law of large numbers, $\lim_{n\to\infty} \frac{\ln_+ Y_n}{n} = 0$ a.s., so for any $c > 0$, $\sum_k \frac{Y_k}{c^k} < \infty$ a.s.

Let $\mathscr{Y}$ be the σ-field generated by (Y_n). Clearly,

$$\mathbf{E}(Z_{n+1} \mid \mathscr{F}_n, \mathscr{Y}) = mZ_n + Y_{n+1} \geq mZ_n,$$

thus $(\frac{Z_n}{m^n})$ is a submartingale (conditionally on $\mathscr{Y}$), and $\mathbf{E}(\frac{Z_n}{m^n} \mid \mathscr{Y}) = \sum_{k=0}^{n} \frac{Y_k}{m^k}$. In particular, on the set $\{\sum_{k=0}^{\infty} \frac{Y_k}{m^k} < \infty\}$, we have $\sup_n \mathbf{E}(\frac{Z_n}{m^n} \mid \mathscr{Y}) < \infty$, so $\lim_{n\to\infty} \frac{Z_n}{m^n}$ exists and is finite. Since $\mathbf{P}(\sum_{k=0}^{\infty} \frac{Y_k}{m^k} < \infty) = 1$, the result follows. □

We recall an elementary result [102, Theorem 5.3.3]. Let $(\mathscr{F}_n)$ be a filtration, and let $\mathscr{F}_\infty$ be the smallest σ-field containing all $\mathscr{F}_n$. Let $\mathbf{P}$ and $\mathbf{Q}$ be probabilities on $(\Omega, \mathscr{F}_\infty)$. Assume that for any n, $\mathbf{Q}_{|\mathscr{F}_n} \ll \mathbf{P}_{|\mathscr{F}_n}$. Let $\xi_n := \frac{\mathrm{d}\mathbf{Q}_{|\mathscr{F}_n}}{\mathrm{d}\mathbf{P}_{|\mathscr{F}_n}}$, and let $\xi := \limsup_{n\to\infty} \xi_n$ which is $\mathbf{P}$-a.s. finite. Then

$$\mathbf{Q}(A) = \mathbf{E}(\xi\, \mathbf{1}_A) + \mathbf{Q}(A \cap \{\xi = \infty\}), \qquad \forall A \in \mathscr{F}_\infty.$$

It follows easily that

$$\mathbf{Q} \ll \mathbf{P} \Leftrightarrow \xi < \infty, \ \mathbf{Q}\text{-a.s.} \ \Leftrightarrow \ \mathbf{E}(\xi) = 1, \tag{2.2}$$

$$\mathbf{Q} \perp \mathbf{P} \Leftrightarrow \xi = \infty, \ \mathbf{Q}\text{-a.s.} \ \Leftrightarrow \ \mathbf{E}(\xi) = 0. \tag{2.3}$$

Proof of Theorem 2.3 If $\sum_{i=1}^{\infty} p_i \, i \ln i < \infty$, then $\mathbf{E}(\ln_+ \widehat{N}) < \infty$. By Theorem 2.5, $\lim_{n\to\infty} M_n$ exists $\mathbf{Q}$-a.s. and is finite $\mathbf{Q}$-a.s. In view of (2.2), this means $\mathbf{E}(M_\infty) = 1$; in particular, $\mathbf{P}(M_\infty = 0) < 1$, thus $\mathbf{P}(M_\infty = 0) = q$ (Lemma 2.2).

If $\sum_{i=1}^{\infty} p_i \, i \ln i = \infty$, then $\mathbf{E}(\ln_+ \widehat{N}) = \infty$. By Theorem 2.5, $\lim_{n\to\infty} M_n$ exists $\mathbf{Q}$-a.s. and is infinite $\mathbf{Q}$-a.s. Hence $\mathbf{E}(M_\infty) = 0$ (by (2.3)), i.e., $\mathbf{P}(M_\infty = 0) = 1$.

□

2.4 Notes

The material of this chapter is borrowed from Lyons et al. [176], and the presentation adapted from Chap. 1 of my lecture notes [221].

Section 2.1 collects a few elementary properties of Galton–Watson processes. For more detailed discussions, we refer to the books by Asmussen and Hering [31], Athreya and Ney [32], Harris [122].

The formalism described in Sect. 2.2 is due to Neveu [203]; the idea of viewing Galton–Watson trees as tree-valued random variables finds its root in Harris [122].

The technique of size-biased Galton–Watson trees, which goes back at least to Kahane and Peyrière [152], has been used by several authors in various contexts. Its presentation in Sect. 2.2, as well as its use to prove the Kesten–Stigum theorem, comes from Lyons et al. [176]. Size-biased Galton–Watson trees can actually be exploited to prove the corresponding results of the Kesten–Stigum theorem in the critical and subcritical cases. See [176] for more details.

Seneta's dichotomy theorem for branching processes with immigration (Theorem 2.5) was discovered by Seneta [219]; its short proof presented in Sect. 2.3 is borrowed from Asmussen and Hering [31, pp. 50–51].

http://www.springer.com/978-3-319-25371-8

Branching Random Walks
École d'Été de Probabilités de Saint-Flour XLII – 2012
Shi, Z.
2015, X, 133 p. 8 illus., 6 illus. in color., Softcover
ISBN: 978-3-319-25371-8