4.5 The critical BGW tree

4.5.1 The rooted BGW tree as a metric space

We begin by recalling that a BGW tree $\mathcal{T} \in \mathbf{T}$ with root $\varnothing$ is a graph in which the vertices are a subset of

(4.61) $\mathcal{I} = \varnothing \cup \cup_{n=1}^{\infty} \mathbb{N}_0$

satisfying conditions (3.1.2). Recall that if $x = (i_1, \dots, i_n) \in \mathcal{T}$ is said to be in generation n, denoted by $H_N(x) = n$ where $N = \#(\mathcal{T})$. The edges are given by the set of pairs of the form $((i_1, \dots, i_n), (i_1, \dots, i_n, j)$.

The *lexicographic order* is an order relation on the vertices of $\mathcal{T}$ defined as follows. We say that $x = (i_1, \dots, i_n)$ and $y = (j_1, \dots, j_m)$ have a *last common ancestor* at generation $\ell \geq 1$ if

(4.62) $(i_1, \dots, i_\ell) = (j_1, \dots, j_\ell)$ and $i_{\ell+1} \neq j_{\ell+1}($ or is empty).

We say that

(4.63) $x < y$ if x,y have a last common ancestor at ℓ and $i_{\ell+1} \neq j_{\ell+1}$.

Given $\mathcal{T}$ with $\#(\mathcal{T}) = N$ we can order the vertices in lexicographic order $\varnothing, x_1, x_2, \dots, x_{N-1}$. We can then embed it in the plane so x_i appears to the left of x_j if $i < j$.

The corresponding *height function* $H_N(k)$ of a tree of size $\#(\mathcal{T}) = N$ is defined by

(4.64) $H_N(k) := |x_k|, \quad 0 \leq k \leq N-1$

where $|x|$ denotes the generation of x.

Note that the number of visits of $H_N(k)$ to n gives the population size at generation n, that is,

(4.65) $Z_n = \sum_{k=0}^{N-1} 1_{\{n\}}(H_N(k))$

where and $1_{\{n\}}$ denotes the indicator function.

We now define a distance between the individuals in $\mathcal{T}$. If we assign length 1 to each edge then a metric $d_{\mathcal{T}}(x, y)$ can be defined on $\mathcal{T}$ by

(4.66) $d_{\mathcal{T}}(x, y) :=$ the length of the shortest path in $\mathcal{T}$ from x to y.

Since the critical BGW tree is a.s. finite this produces a compact metric space and is an example of random compact rooted real tree which we define below.

Remark 4.10 *Note that a reordering of the offspring (in the lexicographic order) defines a root preserving isometry. We can then associate to $\mathcal{T}$ the corresponding equivalence class of plane trees (modulo the family of root preserving isometries). This equivalence class is characterized by $(\#(\mathcal{T}), \varnothing, d_{\mathcal{T}}(.,.))$.*

We now briefly introduce the reduced tree at generation n. We denote the set of nth generation individuals

(4.67) $X_n = \mathcal{T} \cap \mathbb{N}_0^n$.

The *reduced tree*

(4.68)
$$\mathcal{T}_n^R := \{x \in \mathcal{T} : x = (i_1, \dots, i_r),\ r = 1, \dots, n,\ \text{such that } \exists (i_1, \dots, i_n) \in X_n\}.$$

We also define a metric on X_n by $d_n(x, y) := n - \ell$ if the last common ancestor of x, y is in generation $\ell < n$. It is easy to verify that d_n is an *ultrametric*, that is,

(4.69) $d_n(x, y) \leq \max(d_n(x, z), d_n(z, y)) \quad \text{for any } z \in X_n.$

4.5.2 The contour functions

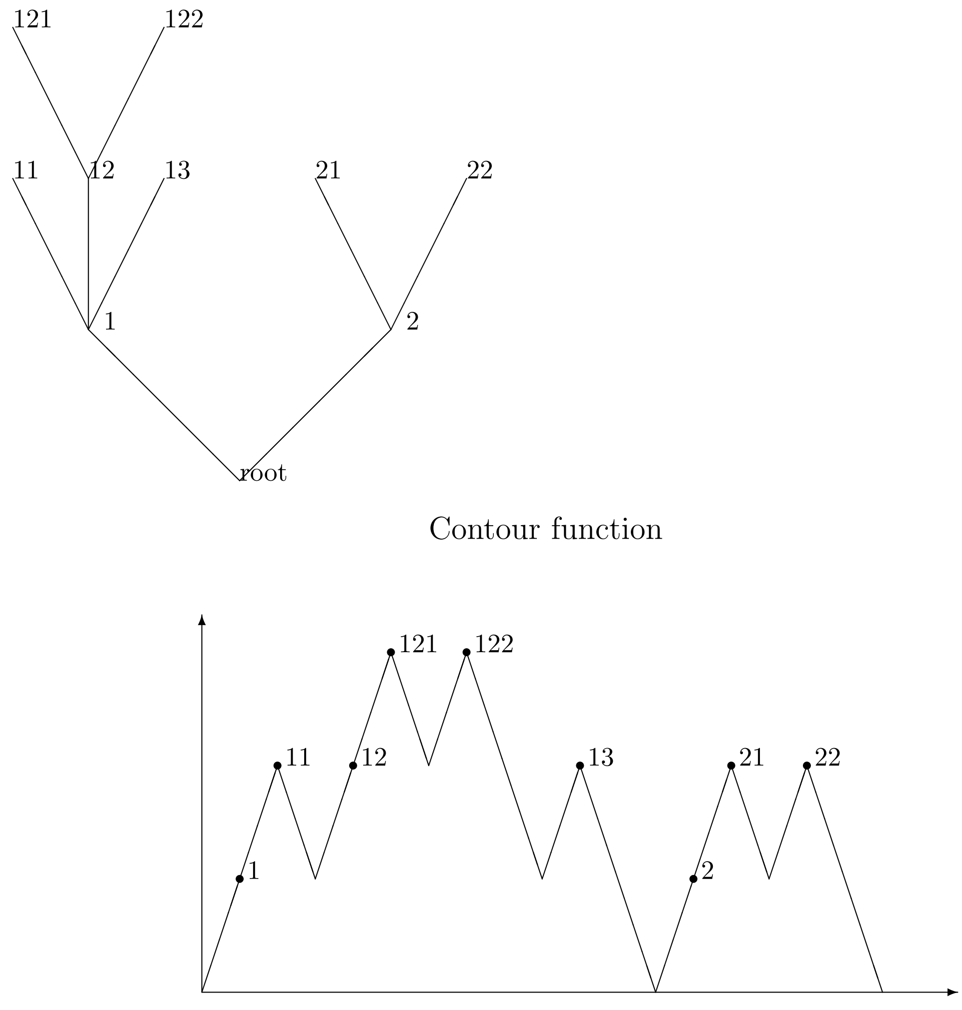
Given a tree $\mathcal{T}$ with $\#(\mathcal{T}) < \infty$ we define the *contour function*

(4.70) $C^{\mathcal{T}} = C^{\mathcal{T}}(t) : 0 \leq t \leq 2(\#(\mathcal{T}) - 1)$

which is obtained by taking a particle that starts from the root of $\mathcal{T}$ and visits continuously all edges at speed one, moving away from the root if possible otherwise going backwards along the edge leading to the root and respecting the lexicographical order of vertices. The domain of $C^{\mathcal{T}}$ can be extended to $[0, \infty)$ by setting $C^{\mathcal{T}}(t) = 0$ for $t > 2(\#(\mathcal{T} - 1))$. In other words, $C^{\mathcal{T}}$ is a piecewise linear process given by the distance from the root as we move through the tree.

We have considered above the Yaglom conditioned limit theorem (Theorem 4.1) for a critical BGW process. Similarly it is if interest to consider the conditioned BGW process conditioned on $\#(\mathcal{T})$. In order to formulate results for this we need to introduce two additional notions, real trees and the Gromov-Hausdorf metric.

Figure 4.2: BGW Tree and contour function, $N = 10$



4.5.3 Real trees

Following Evans [235] and Le Gall [425] we now introduce the notion of real trees and their coding. See Dress and Terhalle [180], [181] for general background on "tree theory".

Definition 4.11 *A metric space* $(\mathcal{T}, d)$ *is a* real tree *if the following two properties hold for every* $(x, y) \in \mathcal{T}$.

- *there is a unique isometric map* $f_{x,y}$ *from* $[0, d(x, y)]$ *into* $\mathcal{T}$ *such that* $f_{x,y}(0) =$

x and $f_{x,y}(d(x,y)) = y$

- *If q is a continuous injective map from $[0,1]$ into $\mathcal{T}$, such that $q(0) = x$, $q(1) = y$, then*

  (4.71) $q([0,1]) = f_{x,y}([0, d(x,y)])$.

  A rooted real tree is a real tree $(\mathcal{T}, d)$ with a distinguished vertex $\emptyset$ called the root.

As explained above it is natural to consider the equivalence class $\mathbb{T}$ of real trees $(\mathcal{T}, d)$ modulo the family of root preserving isometries. Since this results in a collection of compact metric spaces, it can be furnished with the Gromov-Hausdorff metric d_{GH} (see Appendix I, section 17.5).

(Recall that $d_{GH}((E_1, d_1), (E_2, d_2))$ is given by the infimum of the Hausdorff distances of the images of $(E_1, d_1), (E_2, d_2)$ under the set of isometric embeddings of $(E_1, d_1), (E_2, d_2)$, respectively, into a common compact metric space (E_0, d_0).)

Proposition 4.12 *(Evans, Pitman, Winter (2003) [236]). The space of real trees furnished with the Gromov-Hausdorff topology, $(\mathbb{T}, d_{GH})$, is Polish.*

Remark 4.13 *A metric space (E, d) can be embedded isometrically into a real tree iff the* four point condition

(4.72) $d(x,y) + d(u,v) \leq \max(d(x,u) + d(y,v), d(x,v), d(y,u))$

is satisfied for all 4-tuples u, v, x, y (Dress (1984) ,[179])

4.5.4 Excursions from zero and real trees

Consider a continuous function $g : [0,\infty) \to [0,\infty)$ with non-empty compact support such that $g(0) = 0$ and $g(s) = 0 \;\forall\; s > \inf\{t : g(t) = 0\}$ (we call this an positive excursion from 0). For $s, t \geq 0$, let

(4.73) $m_g(s,t) = \inf_{r \in [s \wedge t, s \vee t]} g(r)$,

(4.74) $d_g(s,t) = g(s) + g(t) - 2m_g(s,t)$.

It is easy to check that d_g is symmetric and satisfies the triangle inequality. Let $\mathcal{T}_g$ denote the quotient space $[0,\infty)/\equiv$ where $s \equiv t$ if $d_g(s,t) = 0$. Then it can be verified that the metric space $(\mathcal{T}_g, d_g)$ is a real tree (Le Gall (2006) [425], Theorem 2.1).

Given g the ancestral relationships can be reconstructed by noting that s is an ancestor of t, $s \prec t$ iff $g(s) = \inf_{[s,t]} g(r)$

Let $(\mathbb{C}, \|\cdot\|) := (\{(g, d_g) : g \text{ a positive excursion from } 0, d_g = \text{ sup norm metric}\})$.

It can be verified that (e.g. Le Gall (2006) [425], Lemma 2.3)) that the mapping from $(\mathbb{C}, \|\cdot\|)$ to $(\mathbb{T}, d_{GH})$ is continuous, that is, for two continuous functions g, g' such that $g(0) = g'(0) = 0$:

$$(4.75)\quad d_{GH}(\mathcal{T}_g, \mathcal{T}_{g'}) \leq 2 \parallel g - g' \parallel .$$

4.5.5 The Aldous Continuum Random Tree

Let $\{B_t\}_{t\geq 0}$ be a standard Brownian motion and

$$(4.76)\quad \tau_1 := \sup\{t \in [0,1] : B_t = 0\}, \quad \tau_2 := \inf\{t \geq 1 : B_t = 0\}.$$

Then the *Brownian excursion* is a nonhomogeneous Markov process defined as follows:

$$(4.77)\quad B^e_t := \frac{1}{\sqrt{(\tau_2 - \tau_1)}} B(\tau_1 + t(\tau_2 - \tau_1)), \quad 0 \leq t \leq 1.$$

It can be shown (see Itô-McKean) [335] that the marginal PDF is given by

$$(4.78)\quad f(t,x) = \frac{2x^2}{\sqrt{2\pi t^3(1-t)^3}} e^{-\frac{x^2}{2t(1-t)}}.$$

Definition 4.14 (Aldous continuum random tree) *Let $(B^e_t)_{0\leq t\leq 1}$ be a normalized Brownian excursion (extended to $[0,\infty)$ by setting $B^e_t = 0$ for $t > 1$). The corresponding random real tree $(\mathcal{T}^e, d^e)$ is called the* continuum random tree *(CRT). We denote by $P^{CRT} \in \mathcal{P}((\mathbb{T}, d_{GH}))$ the probability law of $\mathcal{T}^e$.*

The CRT was introduced by Aldous (1991-1993) in a series of papers [5], [6] and [7].

4.5.6 Conditioned limit theorem for the critical BGW tree

Consider the special case of a BGW process with geometric offspring distribution, that is,

$$(4.79)\quad p_k = P(\xi = k) = \frac{1}{2^{k+1}}, \quad k = 0, 1, 2, \ldots,$$

Lemma 4.15 *For the offspring distribution (4.79) the contour process $C^{\mathcal{T}}$ is given by a simple random walk $\{S_k\}$ with*

$$(4.80)\quad P(S_{k+1} - S_k = \pm 1) = \frac{1}{2}.$$

Proof. This can be verified by first noting that in this case the number of jumps from 0 to 1 corresponds to the number of offspring of the initial vertex. Now let

$\tau_1^k, \tau_2^k, \dots$ denote the times of visits to height k. Consider the mth such visit to height k, $k \geq 1$. The corresponding vertex is the offspring of a vertex at height $k-1$, say, the ℓth offspring. Then

$$C(\tau_m^k + 1) = \begin{cases} k+1, \text{ with probability } P(\xi \geq \ell+1 | \xi \geq \ell) = \dfrac{1}{2} \\ k-1 \text{ with probability } P(\xi = \ell | \xi \geq \ell) = \dfrac{1}{2} \end{cases} \tag{4.81}$$

■

Proposition 4.16 *Let $P^{BGW}(\cdot | \#(\mathcal{T}) = n) \in \mathcal{P}((\mathbb{T}, d_{GH}))$ denote the probability law of the BGW tree with offspring distribution (4.79) conditioned to have n vertices. Then*

$$P^{BGW}(\frac{\mathcal{T}}{2\sqrt{2n}} | \#(\mathcal{T} = n)) \Rightarrow P^{CRT} \tag{4.82}$$

in the sense of weak convergence in $\mathcal{P}((\mathbb{T}, d_{GH}))$.

Proof. Letting $S_0 = 0$, and $N = \min\{k > 0 : S_k = 0\}$ and conditioning on $N = 2n$ we have the contour process for this BGW process to have total population n. Note that this is simply an excursion of the simple random walk conditioned to first return to the origin at time $N = 2n$, S_k^N. But it is known that which rescaled converges as $n \to \infty$ to a Brownian excursion from 0 (see Durrett, Iglehart and Miller (1977) [184]).

$$\left(\frac{S^N(\lfloor 2nu \rfloor)}{2\sqrt{2n}} \right)_{0 \leq u \leq 1} \Rightarrow (B_u^e)_{0 \leq u \leq 1}. \tag{4.83}$$

where B^e is the standard Brownian excursion. Using (4.75) and the continuous mapping theorem ([39], Theorem 2.7) this implies that the laws of the corresponding BGW trees converge to the CRT as $n \to \infty$. ■

4.5.7 Aldous Invariance Principle for BGW trees

A remarkable result of Aldous is the *invariance principle* for scaling limit of critical BGW tree, that is, the CRT arises as the limit for the entire class of critical BGW processes with aperiodic offspring distributions having finite second moments.

Theorem 4.17 *(Invariance principle for BGW trees - Aldous (1993) [7], Theorem 23.)*
Consider the critical BGW tree with offspring distribution μ. Assume that μ is aperiodic with variance $\sigma^2 < \infty$ and aperiodic. Then the distribution of the rescaled tree

$$\frac{\sigma}{2\sqrt{n}} \mathcal{T}$$

under the probability measure $P^{BGW}(\cdot|\#(\mathcal{T})=n)$ *(i.e. conditioning that total population up to extinction is* n*) converges as* $n\to\infty$ *to the law of CRT.*

Proof. The proof is given in [7]. It is too long and complex to include here.

However some of the ideas behind the proof are as follows. Using (4.75) we see that the result would follow if the rescaled contour process

(4.84) $(\frac{\sigma}{2\sqrt{n}}C^{\mathcal{T}}(2nt):\ 0\le t\le 1)$

under the probability measure $P^{BGW}(\cdot|\#(\mathcal{T})=n)$ converges in distribution to the normalized Brownian excursion. In the general case can no longer be represented by the excursion of a simple random walk. Aldous (1993) [7] proof of the invariance result is based on a characterization of the distribution of the CRT. Marckert and Mokkadem (2003) [447] gave an alternate proof (assuming the offspring distribution has exponential moments) involving only the contour and height functions. In particular they proved that for any critical offspring distribution with variance σ^2 the weak convergence of the rescaled contour function and height processes. Their key idea is to couple the height process to the random walk ("depth-first queue process")

(4.85) $S_n(j)=\sum_{i=0}^{j-1}(\xi_i-1),\ 1\le j\le n,$

that is, with jump distribution is given by $q_i=p_{i+1},\ i=-1,0,1,2,\dots,$ conditioned by $S_n(0)=0,\ S_n(i)\ge 0,\ 1\le i\le n-1,\ S_n(n)=-1$. Then

(4.86) $H_n(\ell)=\mathrm{Card}\{\mathrm{j}:0\le \mathrm{j}\le \ell-1,\ \min_{0\le \mathrm{k}\le \ell-\mathrm{j}}\mathrm{S_n(j+k)=S_n(j)}\},\quad 0\le\ell<\mathrm{n}-1.$

They then obtain exponential bounds on deviations between the height process and the conditioned random walk S_n to prove that

(4.87) $(\frac{H_n(nt)}{\sqrt{n}})_{0\le t\le 1}\Rightarrow(\frac{2}{\sigma}B^e(t))_{0\le t\le 1}.$

■

Remark 4.18 *It has also been proved that starting the BGW process with* n *individuals then the rescaled height function*

(4.88) $\{\frac{1}{n}H_n(\lfloor n^2t\rfloor)\}_{t\ge 0}\to(H_t)_{t\ge 0}\quad \text{with } H_0=1$

where

(4.89) $H_t=(B_t-\inf_{0\le s\le t}B_s)$

where B_t is a Brownian motion, that is, H_t is reflecting Brownian motion. (See [424]).

Recall that the Ray-Knight Theorem ([522], 52.1) states that if B_t is a Brownian motion with local time $\{\ell_t^a\}$

(4.90) $T := \inf\{u : \ell_u^0 > 1\},$

then the Brownian local time $\{\ell_T^a : a \geq 0\}$ has the same law as the Feller CSB satisfying

(4.91) $dZ_t = 2\sqrt{Z_t}dW_t, \quad Z_0 = 1.$

In other words the local time of the height process is a version of the Feller CSB starting at 1. More precisely, the initial mass $Z_0 = 1$ corresponds to the local time at 0 of a reflecting Brownian motion on $[0, T]$ and for $t \geq 0$ $Z_t = \ell_T^a$, that is the occupation density of the reflecting Brownian motion.

4.6 Remark on general continuous state branching

By the basic result of Silverstein [566] the general continuous state branching process has log-Laplace equation

(4.92) $u_t(\lambda) + \int_0^t \psi(u_s(\lambda)) = \lambda,$

with

(4.93) $\psi(u) = \alpha u + \beta u^2 + \int_0^\infty (e^{-ru} - 1 + ru)\nu(dr)$

where $\alpha, \beta \geq 0$ and ν is a σ-finite measure on $(0, \infty)$ such that $\int (r \wedge r^2)\nu(dr) < \infty$. This include the class of $(1+\beta)$ CSB which arise as limits of BGW processes in which the offspring distribution has infinite second moments and are related to stable processes and other Lévy processes. The genealogical structure, stable continuum trees and convergence of the contour process in this general setting have been developed by Duquesne and LeGall [200] but we do not consider this major topic here.