Large random planar maps and their scaling limits
Lectures 1,2,3: Random trees

Jean-François Le Gall
Université Paris-Sud and Institut universitaire de France

Notes for a course given at the Clay Mathematical Institute Summer School 2010

These notes correspond to the first three lectures of a course on "Large random planar maps and their scaling limits" given jointly with Grégory Miermont at the 2010 Clay Mathematical Institute Summer School. The goal is these three lectures is to discuss the results about random trees and their continuous limits that play a key role in the analysis of scaling limits of large random planar maps.

1 Discrete trees and convergence towards the Brownian excursion

1.1 Plane trees

We will be interested in (finite) rooted ordered trees, which are called plane trees in combinatorics (see e.g. [19]). We set $\mathbb{N} = \{1, 2, \ldots\}$ and by convention $\mathbb{N}^0 = \{\varnothing\}$. We introduce the set

$$\mathcal{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n.$$

An element of $\mathcal{U}$ is thus a sequence $u = (u^1, \ldots, u^n)$ of elements of $\mathbb{N}$, and we set $|u| = n$, so that $|u|$ represents the "generation" of u. If $u = (u^1, \ldots u^m)$ and $v = (v^1, \ldots, v^n)$ belong to $\mathcal{U}$, we write $uv = (u^1, \ldots u^m, v^1, \ldots, v^n)$ for the concatenation of u and v. In particular $u\varnothing = \varnothing u = u$.

The mapping $\pi : \mathcal{U}\backslash\{\varnothing\} \longrightarrow \mathcal{U}$ is defined by $\pi((u^1, \ldots, u^n)) = (u^1, \ldots, u^{n-1})$ ($\pi(u)$ is the "parent" of u).

A plane tree τ is a finite subset of $\mathcal{U}$ such that:

(i) $\varnothing \in \tau$.

(ii) $u \in \tau\backslash\{\varnothing\} \Rightarrow \pi(u) \in \tau$.

(iii) For every $u \in \tau$, there exists an integer $k_u(\tau) \geq 0$ such that, for every $j \in \mathbb{N}$, $uj \in \tau$ if and only if $1 \leq j \leq k_u(\tau)$

The number $k_u(\tau)$ is interpreted as the "number of children" of u in τ.

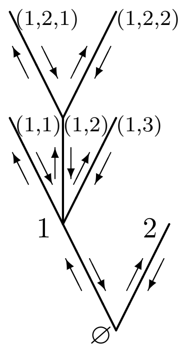
We denote by $\mathbf{A}$ the set of all plane trees. In what follows, we see each vertex of the tree τ as an individual of a population whose τ is the family tree. By definition, the size $|\tau|$ of τ is the number of edges of τ, $|\tau| = \#(\tau) - 1$. For every integer $k \geq 0$, we put

$$\mathbf{A}_k = \{\tau \in \mathbf{A} : |\tau| = k\}.$$

Exercise 1.1. *Verify that the cardinality of $\mathbf{A}_k$ is the k-th Catalan number*

$$\#(\mathbf{A}_k) = \frac{1}{k+1}\binom{2k}{k}.$$

A plane tree can be coded by its Dyck path or **contour function**. Suppose that the tree is embedded in the half-plane in such a way that edges have length one. Informally, we imagine the motion of a particle that starts at time $t = 0$ from the root of the tree and then explores the tree from the left to the right, moving continuously along the edges at unit speed (in the way explained by the arrows of Fig.1), until all edges have been explored and the particle has come back to the root. Since it is clear that each edge will be crossed twice in this evolution, the total time needed to explore the tree is $2|\tau|$. The value C_s of the contour function at time $s \in [0, 2|\tau|]$ is the distance (on the tree) between the position of the particle at time s and the root. By convention $C_s = 0$ if $s \geq 2|\tau|$. Fig.1 explains the construction of the contour function better than a formal definition.



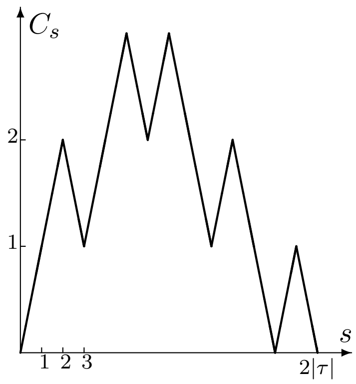


Figure 1

Let $k \geq 0$ be an integer. A Dyck path of length $2k$ is a sequence $(x_0, x_1, x_2, \ldots, x_{2k})$ of nonnegative integers such that $x_0 = x_{2k} = 0$, and $|x_i - x_{i-1}| = 1$ for every $i = 1, \ldots, 2k$. Clearly, if τ is a plane tree of size k, and $(C_s)_{s\geq 0}$ is its contour function, the sequence $(C_0, C_1, \ldots, C_{2k})$ is a Dyck path of length $2k$. More precisely, we have the following easy result.

Proposition 1.2. *The mapping $\tau \mapsto (C_0, C_1, \ldots, C_{2k})$ is a bijection from $\mathbf{A}_k$ onto the set of all Dyck paths of length $2k$.*

1.2 Galton-Watson trees

Let μ be a critical or subcritical offspring distribution. This means that μ is a probability measure on $\mathbb{Z}_+$ such that

$$\sum_{k=0}^{\infty} k\mu(k) \leq 1.$$

We exclude the trivial case where $\mu(1) = 1$.

To define Galton-Watson trees, we let $(K_u, u \in \mathcal{U})$ be a collection of independent random variables with law μ, indexed by the set $\mathcal{U}$. Denote by θ the random subset of $\mathcal{U}$ defined by

$$\theta = \{u = (u^1, \ldots, u^n) \in \mathcal{U} : u^j \leq K_{(u^1,\ldots,u^{j-1})} \text{ for every } 1 \leq j \leq n\}.$$

Proposition 1.3. *θ is a.s. a tree. Moreover, if*

$$Z_n = \#\{u \in \theta : |u| = n\},$$

$(Z_n, n \geq 0)$ is a Galton-Watson process with offspring distribution μ and initial value $Z_0 = 1$.

Remark. Clearly $k_u(\theta) = K_u$ for every $u \in \theta$.

The tree θ, or any random tree with the same distribution, will be called a Galton-Watson tree with offspring distribution μ, or in short a μ-Galton-Watson tree. We also write Π_μ for the distribution of θ on the space $\mathbf{A}$.

We leave the easy proof of the proposition to the reader. The finiteness of the tree θ comes from the fact that the Galton-Watson process with offspring distribution μ becomes extinct a.s., so that $Z_n = 0$ for n large.

If τ is a tree and $1 \leq j \leq k_\varnothing(\tau)$, we write $T_j\tau$ for the tree τ shifted at j:

$$T_j\tau = \{u \in \mathcal{U} : ju \in \tau\}.$$

Note that $T_j\tau$ is a tree.

Then Π_μ may be characterized by the following two properties (see e.g. [15] for more general statements):

(i) $\Pi_\mu(k_\varnothing = j) = \mu(j), \quad j \in \mathbb{Z}_+$.

(ii) For every $j \geq 1$ with $\mu(j) > 0$, the shifted trees $T_1\tau, \ldots, T_j\tau$ are independent under the conditional probability $\Pi_\mu(d\tau \mid k_\varnothing = j)$ and their conditional distribution is Π_μ.

Property (ii) is often called the branching property of the Galton-Watson tree.

We now give an explicit formula for Π_μ.

Proposition 1.4. *For every $\tau \in \mathbf{A}$,*

$$\Pi_\mu(\tau) = \prod_{u \in \tau} \mu(k_u(\tau)).$$

Proof. We can easily check that

$$\{\theta = \tau\} = \bigcap_{u \in \tau} \{K_u = k_u(\tau)\},$$

so that

$$\Pi_\mu(\tau) = P(\theta = \tau) = \prod_{u \in \tau} P(K_u = k_u(\tau)) = \prod_{u \in \tau} \mu(k_u(\tau)).$$

$\square$

We will be interested in the particular case when $\mu = \mu_0$ is the (critical) geometric offspring distribution, $\mu_0(k) = 2^{-k-1}$ for every $k \in \mathbb{Z}_+$. In that case, the proposition gives

$$\Pi_{\mu_0}(\tau) = 2^{-2|\tau|-1}$$

(note that $\sum_{u \in \tau} k_u(\tau) = |\tau|$ for every $\tau \in \mathbf{A}$).

In particular $\Pi_{\mu_0}(\tau)$ only depends on $|\tau|$. As a consequence, for every integer $k \geq 0$, the conditional probability distribution $\Pi_{\mu_0}(\cdot \mid |\tau| = k)$ is just the uniform probability measure on $\mathbf{A}_k$. This fact will be important later.

1.3 The contour function in the geometric case

In general, the Dyck path of a Galton-Watson tree does not have a "nice" probabilistic structure (see however Section 1 of [13]). In this section we restrict our attention to the case when $\mu = \mu_0$ is the critical geometric offspring distribution.

First recall that $(S_n)_{n\geq 0}$ is a simple random walk on $\mathbb{Z}$ (started from 0) if it can be written as

$$S_n = X_1 + X_2 + \cdots + X_n$$

where $X_1, X_2, \ldots$ are i.i.d. random variables with distribution $P(X_n = 1) = P(X_n = -1) = \frac{1}{2}$.

Set $T = \min\{n \geq 0 : S_n = -1\} < \infty$ a.s. The random finite path

$$(S_0, S_1, \ldots, S_{T-1})$$

(or any random path with the same distribution) is called an excursion of simple random walk. Obviously this random path is a random Dyck path of length $T - 1$.

Proposition 1.5. *Let θ be a μ_0-Galton-Watson tree. Then the Dyck path of θ is an excursion of simple random walk.*

Proof. Since plane trees are in one-to-one correspondence with Dyck paths (Proposition 1.2), the statement of the proposition is equivalent to saying that the random plane tree θ coded by an excursion of simple random walk is a μ_0-Galton-Watson tree. To see this, introduce the upcrossing times of the random walk S from 0 to 1:

$$U_1 = \inf\{n \geq 0 : S_n = 1\}\ ,\ V_1 = \inf\{n \geq U_1 : S_n = 0\}$$

and by induction, for every $j \geq 1$,

$$U_{j+1} = \inf\{n \geq V_j : S_n = 1\}\ ,\ V_{j+1} = \inf\{n \geq U_{j+1} : S_n = 0\}.$$

Let $K = \sup\{j : U_j \leq T\}$ ($\sup \varnothing = 0$). From the relation between a plane tree and its associated Dyck path, one easily sees that $k_\varnothing(\theta) = K$, and that for every $i = 1, \ldots, K$, the Dyck path associated with the subtree $T_i\theta$ is the path ω_i, with

$$\omega_i(n) := S_{(U_i+n)\wedge(V_i-1)} - 1 \quad ,\ 0 \leq n \leq V_i - U_i - 1.$$

A simple application of the Markov property now shows that K is distributed according to μ_0 and that conditionally on $K = k$, the paths $\omega_1, \ldots, \omega_k$ are k independent excursions of simple random walk. The characterization of Π_{μ_0} by properties (i) and (ii) listed before Proposition 1.4 now shows that θ is a μ_0-Galton-Watson-tree. $\square$

1.4 Brownian excursions

Our goal is to prove that the (suitably rescaled) contour function of a tree uniformly distributed over $\mathbf{A}_k$ converges in distribution as $k \to \infty$ towards a normalized Brownian excursion. We first need to recall some basic facts about Brownian excursions.

We consider a standard linear Brownian motion $B = (B_t)_{t\geq 0}$ starting from the origin. The process $\beta_t = |B_t|$ is called reflected Brownian motion. We denote by $(L_t^0)_{t\geq 0}$ the local time process of B (or of β) at level 0, which can be defined by the approximation

$$L_t^0 = \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon} \int_0^t ds\, \mathbf{1}_{[-\varepsilon,\varepsilon]}(B_s) = \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon} \int_0^t ds\, \mathbf{1}_{[0,\varepsilon]}(\beta_s),$$

for every $t \geq 0$, a.s.

Then $(L^0_t)_{t\geq 0}$ is a continuous increasing process, and the set of increase points of the function $t \to L^0_t$ coincides with the set

$$\mathcal{Z} = \{t \geq 0 : \beta_t = 0\}$$

of all zeros of β. Consequently, if we introduce the right-continuous inverse of the local time process,

$$\sigma_\ell := \inf\{t \geq 0 : L^0_t > \ell\}\ , \quad \text{for every } \ell \geq 0,$$

we have

$$\mathcal{Z} = \{\sigma_\ell : \ell \geq 0\} \cup \{\sigma_{\ell-} : \ell \in D\}$$

where D denotes the countable set of all discontinuity times of the mapping $\ell \to \sigma_\ell$.

We call *excursion interval* of β (away from 0) any connected component of the open set $\mathbb{R}_+ \backslash \mathcal{Z}$. The preceding discussion shows that, with probability one, the excursion intervals of β away from 0 are exactly the intervals $]\sigma_{\ell-}, \sigma_\ell[$ for $\ell \in D$. Then, for every $\ell \in D$, we define the excursion $e_\ell = (e_\ell(t))_{t\geq 0}$ associated with the interval $]\sigma_{\ell-}, \sigma_\ell[$ by setting

$$e_\ell(t) = \begin{cases} \beta_{\sigma_{\ell-}+t} & \text{if } 0 \leq t \leq \sigma_\ell - \sigma_{\ell-}\ , \\ 0 & \text{if } t > \sigma_\ell - \sigma_{\ell-}\ . \end{cases}$$

We view e_ℓ as an element of the excursion space E, which is defined by

$$E = \{e \in C(\mathbb{R}_+, \mathbb{R}_+) : e(0) = 0 \text{ and } \zeta(e) := \sup\{s > 0 : e(s) > 0\} \in]0, \infty[\},$$

where $\sup \varnothing = 0$ by convention. Note that we require $\zeta(e) > 0$, so that the zero function does not belong to E. The space E is equipped with the metric d defined by

$$d(e, e') = \sup_{t\geq 0} |e(t) - e'(t)| + |\zeta(e) - \zeta(e')|$$

and with the associated Borel σ-field. Notice that $\zeta(e_\ell) = \sigma_\ell - \sigma_{\ell-}$ for every $\ell \in D$. The following theorem is the basic result of excursion theory in our particular setting.

Theorem 1.6. *The point measure*

$$\sum_{\ell \in D} \delta_{(\ell, e_\ell)}(ds\, de)$$

is a Poisson measure on $\mathbb{R}_+ \times E$, with intensity

$$ds \otimes \mathbf{n}(de)$$

where $\mathbf{n}(de)$ is a σ-finite measure on E.

The measure $\mathbf{n}(de)$ is called the Itô measure of positive excursions of linear Brownian motion, or simply the Itô excursion measure. The next corollary follows from standard properties of Poisson measures.

Corollary 1.7. *Let A be a measurable subset of E such that $0 < \mathbf{n}(A) < \infty$, and let $T_A = \inf\{\ell \in D : e_\ell \in A\}$. Then, T_A is exponentially distributed with parameter $\mathbf{n}(A)$, and the distribution of e_{T_A} is the conditional measure*

$$\mathbf{n}(\cdot \,|\, A) = \frac{\mathbf{n}(\cdot \cap A)}{\mathbf{n}(A)}.$$

Moreover, T_A and e_{T_A} are independent.

This corollary can be used to calculate various distributions under the Itô excursion measure. The distribution of the height and the length of the excursion are given as follows: For every $\varepsilon > 0$,

$$\mathbf{n}\Big(\max_{t\geq 0} e(t) > \varepsilon\Big) = \frac{1}{2\varepsilon}$$

and

$$\mathbf{n}(\zeta(e) > \varepsilon) = \frac{1}{\sqrt{2\pi\varepsilon}}.$$

The Itô excursion measure enjoys the following scaling property. For every $\lambda > 0$, define a mapping $\Phi_\lambda : E \longrightarrow E$ by setting $\Phi_\lambda(e)(t) = \sqrt{\lambda}\, e(t/\lambda)$, for every $e \in E$ and $t \geq 0$. Then we have $\Phi_\lambda(\mathbf{n}) = \sqrt{\lambda}\,\mathbf{n}$.

This scaling property is useful when defining conditional versions of the Itô excursion measure. We discuss the conditioning of $\mathbf{n}(de)$ with respect to the length $\zeta(e)$. There exists a unique collection of probability measures $(\mathbf{n}_{(s)}, s > 0)$ on E such that the following properties hold:

(i) For every $s > 0$, $\mathbf{n}_{(s)}(\zeta = s) = 1$.

(ii) For every $\lambda > 0$ and $s > 0$, we have $\Phi_\lambda(\mathbf{n}_{(s)}) = \mathbf{n}_{(\lambda s)}$.

(iii) For every measurable subset A of E,

$$\mathbf{n}(A) = \int_0^\infty \mathbf{n}_{(s)}(A)\,\frac{ds}{2\sqrt{2\pi s^3}}.$$

We may and will write $\mathbf{n}_{(s)} = \mathbf{n}(\cdot \,|\, \zeta = s)$. The measure $\mathbf{n}_{(1)} = \mathbf{n}(\cdot \,|\, \zeta = 1)$ is called the law of the normalized Brownian excursion.

There are many different descriptions of the Itô excursion measure: See in particular [17, Chapter XII]. We state the following proposition, which emphasizes the Markovian properties of $\mathbf{n}$. For every $t > 0$ and $x > 0$, we set

$$q_t(x) = \frac{x}{\sqrt{2\pi t^3}}\exp(-\frac{x^2}{2t}).$$

Note that the function $t \to q_t(x)$ is the density of the first hitting time of x by B.

Proposition 1.8. *The Itô excursion measure* $\mathbf{n}$ *is the only* σ*-finite measure on* E *that satisfies the following two properties.*

(i) *For every* $t > 0$*, and every* $f \in C(\mathbb{R}_+, \mathbb{R}_+)$*,*

$$\mathbf{n}(f(e(t))\,\mathbf{1}_{\{\zeta > t\}}) = \int_0^\infty f(x)\,q_t(x)\,dx.$$

(ii) *Let* $t > 0$*. Under the conditional probability measure* $\mathbf{n}(\cdot \,|\, \zeta > t)$*, the process* $(e(t+r))_{r\geq 0}$ *is Markov with the transition kernels of Brownian motion stopped upon hitting* 0*.*

This proposition can be used to establish absolute continuity properties of the conditional measures $\mathbf{n}_{(s)}$ with respect to $\mathbf{n}$. For every $t \geq 0$, let $\mathcal{F}_t$ denote the σ-field on E generated by the mappings $r \longrightarrow e(r)$, for $0 \leq r \leq t$. Then, if $0 < t < 1$, the measure $\mathbf{n}_{(1)}$ is absolutely continuous with respect to $\mathbf{n}$ on the σ-field $\mathcal{F}_t$, with Radon-Nikodym density

$$\frac{d\mathbf{n}_{(1)}}{d\mathbf{n}}\Big|_{\mathcal{F}_t}(e) = 2\sqrt{2\pi}\,q_{1-t}(e(t)).$$

This formula provides a simple derivation of the finite-dimensional marginals under $\mathbf{n}_{(1)}$, noting that the finite-dimensional marginals under $\mathbf{n}$ are easily obtained from Proposition 1.8.

1.5 Convergence of contour functions to the Brownian excursion

The following theorem can be viewed as a special case of the results in Aldous [2]. The space of all continuous functions from $[0,1]$ into $\mathbb{R}_+$ is denoted by $C([0,1],\mathbb{R}_+)$, and is equipped with the topology of uniform convergence.

Theorem 1.9. *For every integer $k \geq 1$, let θ_k be a random tree that is uniformly distributed over $\mathbf{A}_k$, and let $(C_k(t))_{t\geq 0}$ be its contour function. Then*

$$\Big(\frac{1}{\sqrt{2k}}C_k(2k\,t)\Big)_{0\leq t\leq 1} \xrightarrow[k\to\infty]{\text{(d)}} (\mathbf{e}_t)_{0\leq t\leq 1}$$

where $\mathbf{e}$ is distributed according to $\mathbf{n}_{(1)}$ (i.e. $\mathbf{e}$ is a normalized Brownian excursion) and the convergence holds in the sense of weak convergence of the laws on the space $C([0,1],\mathbb{R}_+)$.

Proof. (sketch) We already noticed that $\Pi_{\mu_0}(\cdot \mid |\tau| = k)$ coincides with the uniform distribution over $\mathbf{A}_k$. By combining this with Proposition 1.5, we get that $(C_k(0), C_k(1), \ldots, C_k(2k))$ is distributed as an excursion of simple random walk conditioned to have length $2k$. To get the desired result, we thus need to verify that the law of

$$\Big(\frac{1}{\sqrt{2k}}S_{[2kt]}\Big)_{0\leq t\leq 1}$$

under $P(\cdot \mid T = 2k+1)$ converges to $\mathbf{n}_{(1)}$ as $k \to \infty$. This follows from a conditional version of Donsker's theorem: See Kaigh [11] (the technical estimates that are derived below in the proof of Lemma 3.3, see in particular (10) and (11), already give the tightness needed for the preceding convergence, and the convergence of finite-dimensional marginals can be obtained from formulas such as (12) below). $\square$

Extensions of Theorem 1.9 can be found in [7] and [6]. To illustrate the power of this theorem, let us give a typical application. The height $H(\tau)$ of a plane tree τ is the maximal generation of a vertex of τ.

Corollary 1.10. *Let θ_k be uniformly distributed over $\mathbf{A}_k$. Then*

$$\frac{1}{\sqrt{2k}}H(\theta_k) \xrightarrow[k\to\infty]{\text{(d)}} \max_{0\leq t\leq 1} \mathbf{e}(t).$$

Since

$$\frac{1}{\sqrt{2k}}H(\theta_k) = \max_{0\leq t\leq 1} \Big(\frac{1}{\sqrt{2k}}C_k(2k\,t)\Big)$$

the result of the corollary is immediate from Theorem 1.9.

The limiting distribution in Corollary 1.10 is known in the form of a series: For every $x > 0$,

$$P\Big(\max_{0\leq t\leq 1} \mathbf{e}_t > x\Big) = 2\sum_{k=1}^{\infty}(4k^2x^2 - 1)\,\exp(-2k^2x^2).$$

See Chung [5].

2 Real trees and the Gromov-Hausdorff convergence

Our main goal in this section is to interpret the convergence of contour functions in Theorem 1.9 as a convergence of discrete random trees towards a "continuous random tree" which is coded by the Brownian excursion in the same sense as a plane tree is coded by its contour function. We need to introduce a suitable notion of a continuous tree, and then to explain in which sense the convergence takes place.

2.1 Real trees

We start with a formal definition. In these notes, we consider only *compact* real trees, and so we include this compactness property in the definition.

Definition 2.1. *A compact metric space $(\mathcal{T}, d)$ is a real tree if the following two properties hold for every $a, b \in \mathcal{T}$.*

(i) There is a unique isometric map $f_{a,b}$ from $[0, d(a,b)]$ into $\mathcal{T}$ such that $f_{a,b}(0) = a$ and $f_{a,b}(d(a,b)) = b$.

(ii) If q is a continuous injective map from $[0,1]$ into $\mathcal{T}$, such that $q(0) = a$ and $q(1) = b$, we have

$$q([0,1]) = f_{a,b}([0, d(a,b)]).$$

A rooted real tree is a real tree $(\mathcal{T}, d)$ with a distinguished vertex $\rho = \rho(\mathcal{T})$ called the root. In what follows, real trees will always be rooted, even if this is not mentioned explicitly.

Informally, one should think of a (compact) $\mathbb{R}$-tree as a connected union of line segments in the plane with no loops. Asssume for simplicity that there are finitely many segments in the union. Then, for any two points a and b in the tree, there is a unique path going from a to b in the tree, which is the concatentation of finitely many line segments. The distance between a and b is then the length of this path.

Let us consider a rooted real tree $(\mathcal{T}, d)$. The range of the mapping $f_{a,b}$ in (i) is denoted by $[\![a, b]\!]$ (this is the "line segment" between a and b in the tree). In particular, $[\![\rho, a]\!]$ is the path going from the root to a, which we will interpret as the ancestral line of vertex a. More precisely we can define a partial order on the tree by setting $a \preccurlyeq b$ (a is an ancestor of b) if and only if $a \in [\![\rho, b]\!]$.

If $a, b \in \mathcal{T}$, there is a unique $c \in \mathcal{T}$ such that $[\![\rho, a]\!] \cap [\![\rho, b]\!] = [\![\rho, c]\!]$. We write $c = a \wedge b$ and call c the most recent common ancestor to a and b.

By definition, the multiplicity of a vertex $a \in \mathcal{T}$ is the number of connected components of $\mathcal{T}\backslash\{a\}$. Vertices of $\mathcal{T}\backslash\{\rho\}$ which have multiplicity 1 are called leaves.

2.2 Coding real trees

In this subsection, we describe a method for constructing real trees, which is well-suited to our forthcoming applications to random trees. This method is nothing but a continuous analogue of the coding of discrete trees by contour functions.

We consider a (deterministic) continuous function $g : [0,1] \longrightarrow [0, \infty[$ such that $g(0) = g(1) = 0$. To avoid trivialities, we will also assume that g is not identically zero. For every $s, t \geq 0$, we set

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

and

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

Clearly $d_g(s,t) = d_g(t,s)$ and it is also easy to verify the triangle inequality

$$d_g(s,u) \leq d_g(s,t) + d_g(t,u)$$

for every $s, t, u \geq 0$. We then introduce the equivalence relation $s \sim t$ iff $d_g(s,t) = 0$ (or equivalently iff $g(s) = g(t) = m_g(s,t)$). Let $\mathcal{T}_g$ be the quotient space

$$\mathcal{T}_g = [0,1]/\sim .$$

Obviously the function d_g induces a distance on $\mathcal{T}_g$, and we keep the notation d_g for this distance. We denote by $p_g : [0,1] \longrightarrow \mathcal{T}_g$ the canonical projection. Clearly p_g is continuous (when $[0,1]$ is equipped with the Euclidean metric and $\mathcal{T}_g$ with the metric d_g), and the metric space $(\mathcal{T}_g, d_g)$ is thus compact.

Theorem 2.1. *The metric space $(\mathcal{T}_g, d_g)$ is a real tree. We will view $(\mathcal{T}_g, d_g)$ as a rooted tree with root $\rho = p_g(0) = p_g(1)$.*

Remark. It is also possible to prove that any (rooted) real tree can be represented in the form $\mathcal{T}_g$. We will leave this as an exercise for the reader.

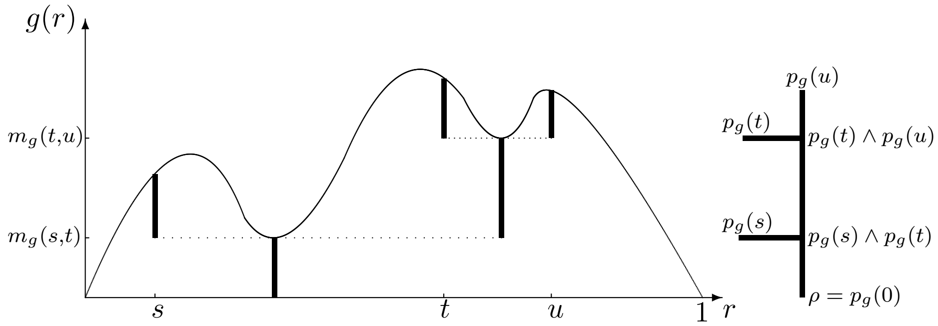


Figure 2

To get an intuitive understanding of Theorem 2.1, the reader should have a look at Figure 2. This figure shows how to construct a simple subtree of $\mathcal{T}_g$, namely the "reduced tree" consisting of the union of the ancestral lines in $\mathcal{T}_g$ of three vertices $p_g(s), p_g(t), p_g(u)$ corresponding to three (given) times $s, t, u \in [0,1]$. This reduced tree is the union of the five bold line segments that are constructed from the graph of g in the way explained on the left part of the figure. Notice that the lengths of the horizontal dotted lines play no role in the construction, and that the reduced tree should be viewed as pictured on the right part of Figure 2. The ancestral line of $p_g(s)$ (resp. $p_g(t), p_g(u)$) is a line segment of length $g(s)$ (resp. $g(t), g(u)$). The ancestral lines of $p_g(s)$ and $p_g(t)$ share a common part, which has length $m_g(s,t)$ (the line segment at the bottom in the left or the right part of Figure 2), and of course a similar property holds for the ancestral lines of $p_g(s)$ and $p_g(u)$, or of $p_g(t)$ and $p_g(u)$.

The following re-rooting lemma, which is of independent interest, is a useful ingredient of the proof of Theorem 2.1.

Lemma 2.2. *Let $s_0 \in [0,1[$. For any real $r \geq 0$, denote by $\overline{r} = r - [r]$ the fractional part of r. Set*

$$g'(s) = g(s_0) + g(\overline{s_0 + s}) - 2m_g(s_0, \overline{s_0 + s}),$$

for every $s \in [0,1]$. Then, the function g' is continuous and satisfies $g'(0) = g'(1) = 0$, so that we can define $\mathcal{T}_{g'}$. Furthermore, for every $s, t \in [0,1]$, we have

$$d_{g'}(s,t) = d_g(\overline{s_0 + s}, \overline{s_0 + t}) \tag{1}$$

and there exists a unique isometry R from $\mathcal{T}_{g'}$ onto $\mathcal{T}_g$ such that, for every $s \in [0,1]$,

$$R(p_{g'}(s)) = p_g(\overline{s_0 + s}). \tag{2}$$

Assuming that Theorem 2.1 is proved, we see that $\mathcal{T}_{g'}$ coincides with the real tree $\mathcal{T}_g$ re-rooted at $p_g(s_0)$. Thus the lemma tells us which function codes the tree $\mathcal{T}_g$ re-rooted at an arbitrary vertex.

Proof. It is immediately checked that g' satisfies the same assumptions as g, so that we can make sense of $\mathcal{T}_{g'}$. Then the key step is to verify the relation (1). Consider first the case where $s, t \in [0, 1 - s_0[$. Then two possibilities may occur.

If $m_g(s_0 + s, s_0 + t) \geq m_g(s_0, s_0 + s)$, then $m_g(s_0, s_0 + r) = m_g(s_0, s_0 + s) = m_g(s_0, s_0 + t)$ for every $r \in [s, t]$, and so

$$m_{g'}(s,t) = g(s_0) + m_g(s_0 + s, s_0 + t) - 2m_g(s_0, s_0 + s).$$

It follows that

$$\begin{aligned} d_{g'}(s,t) &= g'(s) + g'(t) - 2m_{g'}(s,t) \\ &= g(s_0 + s) - 2m_g(s_0, s_0 + s) + g(s_0 + t) \\ &\quad - 2m_g(s_0, s_0 + t) - 2(m_g(s_0 + s, s_0 + t) - 2m_g(s_0, s_0 + s)) \\ &= g(s_0 + s) + g(s_0 + t) - 2m_g(s_0 + s, s_0 + t) \\ &= d_g(s_0 + s, s_0 + t). \end{aligned}$$

If $m_g(s_0 + s, s_0 + t) < m_g(s_0, s_0 + s)$, then the minimum in the definition of $m_{g'}(s,t)$ is attained at r_1 defined as the first $r \in [s,t]$ such that $g(s_0 + r) = m_g(s_0, s_0 + s)$ (because for $r \in [r_1, t]$ we will have $g(s_0 + r) - 2m_g(s_0, s_0 + r) \geq -m_g(s_0, s_0 + r) \geq -m_g(s_0, s_0 + r_1)$). Therefore,

$$m_{g'}(s,t) = g(s_0) - m_g(s_0, s_0 + s),$$

and

$$\begin{aligned} d_{g'}(s,t) &= g(s_0 + s) - 2m_g(s_0, s_0 + s) + g(s_0 + t) - 2m_g(s_0, s_0 + t) + 2m_g(s_0, s_0 + s) \\ &= d_g(s_0 + s, s_0 + t). \end{aligned}$$

The other cases are treated in a similar way and are left to the reader.

By (1), if $s, t \in [0,1]$ are such that $d_{g'}(s,t) = 0$, we have $d_g(\overline{s_0 + s}, \overline{s_0 + t}) = 0$ so that $p_g(\overline{s_0 + s}) = p_g(\overline{s_0 + t})$. Noting that $\mathcal{T}_{g'} = p_{g'}([0,1])$, we can define R in a unique way by the relation (2). From (1), R is an isometry, and it is also immediate that R takes $\mathcal{T}_{g'}$ onto $\mathcal{T}_g$. $\square$

Thanks to the lemma, the fact that $\mathcal{T}_g$ verifies property (i) in the definition of a real tree is obtained from the particular case when $a = \rho$ and $b = p_g(s)$ for some $s \in [0,1]$. In that case however, the isometric mapping $f_{\rho,b}$ is easily constructed by setting

$$f_{\rho,b}(t) = p_g(\sup\{r \leq s : g(r) = t\}) \ , \quad \text{for every } 0 \leq t \leq g(s) = d_g(\rho, b).$$

The remaining part of the argument is straightforward: See Section 2 in [8].

Remark. A short proof of Theorem 2.1 using the characterization of real trees via the so-called four-point condition can be found in [9].

2.3 The Gromov-Hausdorff convergence

In order to make sense of the convergence of discrete trees towards real trees, we will use the Gromov-Hausdorff distance between compact metric spaces, which has been introduced by Gromov (see e.g. [10]) in view of geometric applications.

If (E,δ) is a metric space, the notation $\delta_{Haus}(K,K')$ stands for the usual Hausdorff metric between compact subsets of E :

$$\delta_{Haus}(K,K') = \inf\{\varepsilon > 0 : K \subset U_\varepsilon(K') \text{ and } K' \subset U_\varepsilon(K)\},$$

where $U_\varepsilon(K) := \{x \in E : \delta(x,K) \leq \varepsilon\}$.

A rooted metric space is just a pair consisting of a metric space E and a distinguished point ρ of $\mathcal{T}$ called the root.

Then, if E_1 and E_2 are two rooted compact metric spaces, with respective roots ρ_1 and ρ_2, we define the distance $d_{GH}(E_1,E_2)$ by

$$d_{GH}(E_1,E_2) = \inf\{\delta_{Haus}(\varphi_1(E_1),\varphi_2(E_2)) \vee \delta(\varphi_1(\rho_1),\varphi_2(\rho_2))\}$$

where the infimum is over all possible choices of the metric space (E,δ) and the isometric embeddings $\varphi_1 : E_1 \longrightarrow E$ and $\varphi_2 : E_2 \longrightarrow E$ of E_1 and E_2 into E.

Two rooted compact metric spaces E_1 and E_2 are called equivalent if there is a root-preserving isometry that maps E_1 onto E_2. Obviously $d_{GH}(E_1,E_2)$ only depends on the equivalence classes of E_1 and E_2. We denote by $\mathbb{K}$ the space of all equivalence classes of rooted compact metric spaces.

Theorem 2.3. *d_{GH} defines a metric on the set $\mathbb{K}$. Furthermore the metric space $(\mathbb{K}, d_{GH})$ is separable and complete.*

We refer to [3] for a proof of this theorem.

In our applications, it will be important to have the following alternative definition of d_{GH}. First recall that if (E_1,d_1) and (E_2,d_2) are two compact metric spaces, a correspondence between E_1 and E_2 is a subset $\mathcal{R}$ of $E_1 \times E_2$ such that for every $x_1 \in E_1$ there exists at least one $x_2 \in E_2$ such that $(x_1,x_2) \in \mathcal{R}$ and conversely for every $y_2 \in E_2$ there exists at least one $y_1 \in E_1$ such that $(y_1,y_2) \in \mathcal{R}$. The distorsion of the correspondence $\mathcal{R}$ is defined by

$$\mathrm{dis}(\mathcal{R}) = \sup\{|d_1(x_1,y_1) - d_2(x_2,y_2)| : (x_1,x_2),(y_1,y_2) \in \mathcal{R}\}.$$

Proposition 2.4. *Let E_1 and E_2 be two rooted compact metric spaces with respective roots ρ_1 and ρ_2. Then,*

$$d_{GH}(E_1,E_2) = \frac{1}{2} \inf_{\mathcal{R} \in \mathcal{C}(E_1,E_2),\, (\rho_1,\rho_2) \in \mathcal{R}} \mathrm{dis}(\mathcal{R}), \tag{3}$$

where $\mathcal{C}(E_1,E_2)$ denotes the set of all correspondences between E_1 and E_2.

See [3] for a proof of this proposition. The following consequence of Proposition 2.4 will be very useful.

Corollary 2.5. *Let g and g' be two continuous functions from $[0,1]$ into $\mathbb{R}_+$, such that $g(0) = g(1) = g'(0) = g'(1) = 0$. Then,*

$$d_{GH}(\mathcal{T}_g, \mathcal{T}_{g'}) \leq 2\|g - g'\|,$$

where $\|g - g'\| = \sup_{t \in [0,1]} |g(t) - g'(t)|$ is the supremum norm of $g - g'$.

Proof. We rely on formula (3). We can construct a correspondence between $\mathcal{T}_g$ and $\mathcal{T}_{g'}$ by setting

$$\mathcal{R} = \{(a,a') : \exists t \in [0,1] \text{ such that } a = p_g(t) \text{ and } a' = p_{g'}(t)\}.$$

Note that $(\rho, \rho') \in \mathcal{R}$, if $\rho = p_g(0)$, resp. $\rho' = p_{g'}(0)$, is the root of $\mathcal{T}_g$, resp. the root of $\mathcal{T}_{g'}$. In order to bound the distortion of $\mathcal{R}$, let $(a, a') \in \mathcal{R}$ and $(b, b') \in \mathcal{R}$. By the definition of $\mathcal{R}$ we can find $s, t \geq 0$ such that $p_g(s) = a$, $p_{g'}(s) = a'$ and $p_g(t) = b$, $p_{g'}(t) = b'$. Now recall that

$$\begin{aligned} d_g(a,b) &= g(s) + g(t) - 2m_g(s,t), \\ d_{g'}(a',b') &= g'(s) + g'(t) - 2m_{g'}(s,t), \end{aligned}$$

so that

$$|d_g(a,b) - d_{g'}(a',b')| \leq 4\|g - g'\|.$$

Thus we have $\mathrm{dis}(\mathcal{R}) \leq 4\|g - g'\|$ and the desired result follows from (3). $\square$

2.4 Convergence towards the CRT

As in subsection 1.5, we use the notation $\mathbf{e}$ for a normalized Brownian excursion. We view $\mathbf{e} = (\mathbf{e}_t)_{0\leq t\leq 1}$ as a (random) continuous function over the interval $[0,1]$, which satisfies the same assumptions as the function g in subsection 2.2

Definition 2.2. *The continuum random tree (CRT) is the random real tree $\mathcal{T}_{\mathbf{e}}$ coded by the normalized Brownian excursion.*

The CRT $\mathcal{T}_{\mathbf{e}}$ is thus a random variable taking values in the set $\mathbb{K}$. Note that the measurability of this random variable follows from Corollary 2.5.

Remark. Aldous [1],[2] uses a different method to define the CRT. The preceding definition then corresponds to Corollary 22 in [2]. Note that our normalization differs by an unimportant scaling factor 2 from the one in Aldous' papers: The CRT there is the tree $\mathcal{T}_{2\mathbf{e}}$ instead of $\mathcal{T}_{\mathbf{e}}$.

We will now restate Theorem 1.9 as a convergence in distribution of discrete random trees towards the CRT in the space $(\mathbb{K}, d_{GH})$.

Theorem 2.6. *For every $k \geq 1$, let θ_k be uniformly distributed over $\mathbf{A}_k$, and equip θ_k with the usual graph distance d_{gr}. Then*

$$(\theta_k, (2k)^{-1/2} d_{gr}) \xrightarrow[k\to\infty]{\text{(d)}} (\mathcal{T}_{\mathbf{e}}, d_{\mathbf{e}})$$

where the convergence holds in distribution for random variables with values in $(\mathbb{K}, d_{GH})$.

Proof. As in Theorem 1.9, let C_k be the contour function of θ_k, and define a rescaled version of C_k by setting

$$\widetilde{C}_k(t) = (2k)^{-1/2} C_k(2k\,t)$$

for every $t \in [0,1]$. Note that the function $\widetilde{C}_k$ is continuous and nonnegative over $[0,1]$ and vanishes at 0 and at 1. Therefore we can define the real tree $\mathcal{T}_{\widetilde{C}_k}$.

Now observe that this real tree is very closely related to the (rescaled) discrete tree θ_k. Indeed $\mathcal{T}_{\widetilde{C}_k}$ is (isometric to) a finite union of line segments of length $(2k)^{-1/2}$ in the plane, with genealogical structure prescribed by θ_k, in the way suggested in the left part of Figure 1. From this observation, and the definition of the Gromov-Hausdorff distance, we easily get

$$d_{GH}\Big((\theta_k, (2k)^{-1/2} d_{gr}), (\mathcal{T}_{\widetilde{C}_k}, d_{\widetilde{C}_k})\Big) \leq (2k)^{-1/2}. \tag{4}$$

On the other hand, by combining Theorem 1.9 and Corollary 2.5, we have

$$(\mathcal{T}_{\widetilde{C}_k}, d_{\widetilde{C}_k}) \xrightarrow[k\to\infty]{\text{(d)}} (\mathcal{T}_{\mathbf{e}}, d_{\mathbf{e}}).$$

The statement of Theorem 2.6 now follows from the latter convergence and (4). $\square$

Remark. Theorem 2.6 contains in fact less information than Theorem 1.9, because the lexicographical ordering that is inherent to the notion of a plane tree (and also to the coding of real trees by functions) disappears when we look at a plane tree as a metric space. Still, Theorem 2.6 is important from the conceptual viewpoint: It is crucial to think of the CRT as a continuous limit of rescaled discrete random trees.

We could state analogues of Theorem 2.6 for several other classes of combinatorial trees. For instance, if τ_n is distributed uniformly among all rooted Cayley trees with n vertices , then $(\tau_n, (4n)^{-1/2} d_{gr})$ converges in distribution to the CRT $\mathcal{T}_{\mathbf{e}}$, in the space $\mathbb{K}$. Similarly, discrete random trees that are uniformly distributed over binary trees with $2k$ edges converge in distribution (modulo a suitable rescaling) towards the CRT. All these results can be derived from a general statement of convergence of (conditioned) Galton-Watson trees due to Aldous [2] (see also [13]).

3 Labeled trees and the Brownian snake

3.1 Labeled trees

In view of forthcoming applications to random planar maps, we now introduce labeled trees. A labelled tree is a pair $(\tau, (\ell_v)_{v\in\tau})$ that consists of a plane tree τ (see subsection 1.1) and a collection $(\ell_v)_{v\in\tau}$ of integer labels assigned to the vertices of τ – in our formalism for plane trees, the tree τ coincides with the set of its vertices. We assume that labels satisfy the following three properties:

(i) for every $v \in \tau$, $\ell_v \in \mathbb{Z}$;

(ii) $\ell_\varnothing = 0$;

(iii) for every $v \in \tau\backslash\{\varnothing\}$, $\ell_v - \ell_{\pi(v)} = 1, 0,$ or -1,

where we recall that $\pi(v)$ denotes the parent of v. Condition (iii) just means that when crossing an edge of τ the label can change by at most 1 in absolute value.

The motivation for introducing labeled trees comes from the fact that (rooted and pointed) planar quadrangulations can be coded by such trees (see [4] and [14]). Our goal in the present section is to derive asymptotics for large labeled trees chosen uniformly at random, in the same way as Theorem 1.9, or Theorem 2.6, provides asymptotics for large plane trees. For every integer $k \geq 0$, we denote by $\mathbf{T}_k$ the set of all labeled trees with k edges. It is immediate that

$$\#(\mathbf{T}_k) = 3^k \#(\mathbf{A}_k) = \frac{3^k}{k+1}\binom{2k}{k}$$

simply because for each edge of the tree there are three possible choices for the label increment along this edge.

Let $(\tau, (\ell_v)_{v\in\tau})$ be a labeled tree with k edges. As we saw in subsection 1.1, the plane tree τ is coded by its contour function $(C_t)_{t\geq 0}$. We can similarly encode the labels by another function $(V_t)_{t\geq 0}$, which is defined as follows. If we explore the tree τ by following its contour, in the way suggested by the arrows of Fig.1, we visit successively all vertices of τ (vertices that are not leaves are visited more than once). Write $v_0 = \varnothing, v_1, v_2, \ldots, v_{2k} = \varnothing$ for the successive vertices visited in this exploration. For instance, in the particular example of Fig.1 we have

$$v_0 = \varnothing, v_1 = 1, v_2 = (1,1), v_3 = 1, v_4 = (1,2), v_5 = (1,2,1), v_6 = (1,2), \ldots$$

Notice that $C_i = |v_i|$, for every $i = 0, 1, \ldots, 2k$, by the definition of the contour function. We similarly set

$$V_i = \ell_{v_i} \quad \text{for every } i = 0, 1, \ldots, 2k.$$

To complete this definition, we set $V_t = 0$ for $t \geq 2k$ and, for every $i = 1, \ldots, 2k$, we define V_t for $t \in]i-1, i[$ by using linear interpolation. We will call $(V_t)_{t \geq 0}$ the "spatial contour function" of the labeled tree $(\tau, (\ell_v)_{v \in \tau})$ Clearly $(\tau, (\ell_v)_{v \in \tau})$ is determined by the pair $(C_t, V_t)_{t \geq 0}$.

Our goal is now to describe the scaling limit of this pair when the labeled tree $(\tau, (\ell_v)_{v \in \tau})$ is chosen uniformly at random in $\mathbf{T}_k$ and $k \to \infty$. As an immediate consequence of Theorem 1.9 (and the fact that the number of possible labelings is the same for every plane tree with k edges), the scaling limit of $(C_t)_{t \geq 0}$ is the normalized Brownian excursion. To describe the scaling limit of $(V_t)_{t \geq 0}$ we need to introduce the Brownian snake.

3.2 The snake driven by a deterministic function

Let $g : [0,1] \longrightarrow \mathbb{R}_+$ be a continuous function such that $g(0) = g(1) = 0$ (as in subsection 2.2). We also assume that g is Hölder continuous: There exist two positive constants K and γ such that, for every $s, t \in [0,1]$,

$$|g(s) - g(t)| \leq K\,|s-t|^\gamma.$$

As in subsection 2.2, we also set, for every $s, t \in [0,1]$,

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

Lemma 3.1. *The function $(m_g(s,t))_{s,t \in [0,1]}$ is nonnegative definite in the sense that, for every integer $n \geq 1$, for every $s_1, \ldots, s_n \in [0,1]$ and every $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, we have*

$$\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \, m_g(s_i, s_j) \geq 0.$$

Proof. Fix $s_1, \ldots, s_n \in [0,1]$, and let $t \geq 0$. For $i, j \in \{1, \ldots, n\}$, put $i \approx j$ if $m_g(s_i, s_j) \geq t$. Then $\approx$ is an equivalence relation on $\{i : g(s_i) \geq t\} \subset \{1, \ldots, n\}$. By summing over the different classes of this equivalence relation, we get that

$$\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \mathbf{1}_{\{t \leq m_g(s_i,s_j)\}} = \sum_{\mathcal{C} \text{ class of } \approx} \Big(\sum_{i \in \mathcal{C}} \lambda_i \Big)^2 \geq 0.$$

Now integrate with respect to dt to get the desired result. □

By Lemma 3.1 and a standard application of the Kolmogorov extension theorem, there exists a centered Gaussian process $(Z^g_s)_{s \in [0,1]}$ whose covariance is

$$E[Z^g_s Z^g_t] = m_g(s,t)$$

for every $s, t \in [0,1]$. Consequently we have

$$E[(Z^g_s - Z^g_t)^2] = E[(Z^g_s)^2] + E[(Z^g_t)^2] - 2E[Z^g_s Z^g_t] = g(s) + g(t) - 2m_g(s,t) \leq K'\,|s-t|^\gamma,$$

where the last bound follows from our Hölder continuity assumption on g (this calculation also shows that $E[(Z^g_s - Z^g_t)^2] = d_g(s,t)$, in the notation of subsection 2.2). From the previous bound and an application of the Kolmogorov continuity criterion, the process $(Z^g_s)_{s\in[0,1]}$ has a modification with continuous sample paths. This leads us to the following definition.

Definition 3.1. *The snake driven by the function g is the centered Gaussian process $(Z^g_s)_{s\in[0,1]}$ with continuous sample paths and covariance*

$$E[Z^g_s Z^g_t] = m_g(s,t) \;, \quad s,t \in [0,1].$$

Notice that we have in particular $Z^g_0 = Z^g_1 = 0$. More generally, for every $t \in [0,1]$, Z^g_t is normal with mean 0 and variance $g(t)$.

Remark. Recall from subsection 2.2 the definition of the equivalence relation $\sim$ associated with g: $s \sim t$ iff $d_g(s,t) = 0$. Since we have $E[(Z^g_s - Z^g_t)^2] = d_g(s,t)$, a simple argument shows that almost surely for every $s,t \in [0,1]$, the condition $s \sim t$ implies that $Z^g_s = Z^g_t$. In other words we may view Z^g as a process indexed by the quotient $[0,1]\,/\sim$, that is by the tree $\mathcal{T}_g$. Indeed, it is then very natural to interpret Z^g as Brownian motion indexed by the tree $\mathcal{T}_g$: In the particular case when $\mathcal{T}_g$ is a finite union of segments (which holds if g is piecewise monotone), Z^g can be constructed by running independent Brownian motions along the branches of $\mathcal{T}_g$. We preferred to view Z^g as a process indexed by $[0,1]$ because later the function g (and thus the tree $\mathcal{T}_g$) will be random and considering a random process indexed by a random set leads to certain difficulties.

3.3 Convergence towards the Brownian snake

Let $\mathbf{e}$ be as previously a normalized Brownian excursion. By standard properties of Brownian paths, the function $t \mapsto \mathbf{e}_t$ is a.s. Hölder continuous (with exponent $\frac{1}{2} - \varepsilon$ for any $\varepsilon > 0$), and so we can apply the construction of the previous subsection to (almost) every realization of $\mathbf{e}$.

In other words, we can construct a pair $(\mathbf{e}_t, Z_t)_{t\in[0,1]}$ of continuous random processes, whose distribution is characterized by the following two properties:

(i) $\mathbf{e}$ is a normalized Brownian excursion;

(ii) conditionally given $\mathbf{e}$, Z is distributed as the snake driven by $\mathbf{e}$.

The process Z will be called the Brownian snake (driven by $\mathbf{e}$). This terminology is a little different from the usual one: Usually, the Brownian snake is viewed as a path-valued process (see e.g. [12]) and Z_t would correspond only to the terminal point of the value at time t of this path-valued process.

We can now answer the question raised at the end of subsection 3.1. The following theorem is due to Chassaing and Schaeffer [4].

Theorem 3.2. *For every integer $k \geq 1$, let $(\theta_k, (\ell^k_v)_{v\in\theta_k})$ be distributed uniformly over the set $\mathbf{T}_k$ of all labeled trees with k edges. Let $(C_k(t))_{t\geq 0}$ and $(V_k(t))_{t\geq 0}$ be respectively the contour function and the spatial contour function of the labeled tree $(\theta_k, (\ell^k_v)_{v\in\theta_k})$. Then,*

$$\Big(\frac{1}{\sqrt{2k}} C_k(2k\,t), \Big(\frac{9}{8k}\Big)^{1/4} V_k(2k\,t)\Big)_{t\in[0,1]} \xrightarrow[k\to\infty]{\text{(d)}} (\mathbf{e}_t, Z_t)_{t\in[0,1]}$$

where the convergence holds in the sense of weak convergence of the laws on the space $C([0,1], \mathbb{R}^2_+)$.

Proof. From Theorem 1.9 and the Skorokhod representation theorem, we may assume without loss of generality that

$$\sup_{0\leq t\leq 1} |(2k)^{-1/2} C_k(2kt) - \mathbf{e}_t| \xrightarrow[k\to\infty]{\text{a.s.}} 0. \tag{5}$$

We first discuss the convergence of finite-dimensional marginals: We prove that for every choice of $0 \leq t_1 < t_2 < \cdots < t_p$, we have

$$\Big(\frac{1}{\sqrt{2k}} C_k(2k\,t_i), \Big(\frac{9}{8k}\Big)^{1/4} V_k(2k\,t_i)\Big)_{1\leq i\leq p} \xrightarrow[k\to\infty]{\text{(d)}} (\mathbf{e}_{t_i}, Z_{t_i})_{1\leq i\leq p}. \tag{6}$$

Since for every $i \in \{1, \dots, n\}$,

$$|C_k(2kt_i) - C_k([2kt_i])| \leq 1 \ , \quad |V_k(2kt_i) - V_k([2kt_i])| \leq 1$$

we may replace $2kt_i$ by its integer part $[2kt_i]$ in (6).

Consider the case $p = 1$. We may assume that $0 < t_1 < 1$, because otherwise the result is trivial. It is immediate that conditionally on θ_k, the label increments $\ell^k_v - \ell^k_{\pi(v)}$, $v \in \theta_k\backslash\{\emptyset\}$, are i.i.d. with uniform distribution on $\{-1, 0, 1\}$. Consequently, we may write

$$(C_k([2kt_1]), V_k([2kt_1])) \overset{\text{(d)}}{=} \Big(C_k([2kt_1]), \sum_{i=1}^{C_k([2kt_1])} \eta_i\Big)$$

where the variables $\eta_1, \eta_2, \dots$ are i.i.d. with uniform distribution on $\{-1, 0, 1\}$, and are also independent of the trees θ_k. By the central limit theorem,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_i \xrightarrow[n\to\infty]{\text{(d)}} \Big(\frac{2}{3}\Big)^{1/2} N$$

where N is a standard normal variable. Thus if we set for $\lambda \in \mathbb{R}$,

$$\Phi(n, \lambda) = E\Big[\exp\Big(i\frac{\lambda}{\sqrt{n}} \sum_{i=1}^{n} \eta_i\Big)\Big]$$

we have $\Phi(n, \lambda) \longrightarrow \exp(-\lambda^2/3)$ as $n \to \infty$.

Then, for every $\lambda, \lambda' \in \mathbb{R}$, we get by conditioning on θ_k

$$\begin{aligned} E\Big[\exp\Big(i\frac{\lambda}{\sqrt{2k}} C_k([2kt_1]) + i\frac{\lambda'}{\sqrt{C_k([2kt_1])}} \sum_{i=1}^{C_k([2kt_1])} \eta_i\Big)\Big] \\ = E\Big[\exp\Big(i\frac{\lambda}{\sqrt{2k}} C_k([2kt_1])\Big) \times \Phi(C_k([2kt_1]), \lambda')\Big] \\ \xrightarrow[k\to\infty]{} E[\exp(i\lambda\mathbf{e}_{t_1})] \times \exp(-\lambda'^2/3) \end{aligned}$$

using the (almost sure) convergence of $(2k)^{-1/2} C_k([2kt_1])$ towards $\mathbf{e}_{t_1} > 0$. In other words we have obtained the joint convergence in distribution

$$\Big(\frac{C_k([2kt_1])}{\sqrt{2k}}, \frac{1}{\sqrt{C_k([2kt_1])}} \sum_{i=1}^{C_k([2kt_1])} \eta_i\Big) \xrightarrow[k\to\infty]{\text{(d)}} (\mathbf{e}_{t_1}, (2/3)^{1/2} N), \tag{7}$$

where the normal variable N is independent of $\mathbf{e}$.

From preceding observations, we have

$$\Big(\frac{C_k([2kt_1])}{\sqrt{2k}},\Big(\frac{9}{8k}\Big)^{1/4}V_k([2kt_1])\Big)\overset{\rm (d)}{=}\Big(\frac{C_k([2kt_1])}{\sqrt{2k}},(3/2)^{1/2}\Big(\frac{C_k([2kt_1])}{\sqrt{2k}}\Big)^{1/2}\frac{1}{\sqrt{C_k([2kt_1])}}\sum_{i=1}^{C_k([2kt_1])}\eta_i\Big)$$

and from (7) we get

$$\Big(\frac{C_k([2kt_1])}{\sqrt{2k}},\Big(\frac{9}{8k}\Big)^{1/4}V_k([2kt_1])\Big)\xrightarrow[k\to\infty]{\rm (d)}(\mathbf{e}_{t_1},\sqrt{\mathbf{e}_{t_1}}\,N).$$

This gives (6) in the case $p=1$, since by construction $(\mathbf{e}_{t_1},Z_{t_1})\overset{\rm (d)}{=}(\mathbf{e}_{t_1},\sqrt{\mathbf{e}_{t_1}}\,N)$.

Let us discuss the case $p=2$ of (6). We fix t_1 and t_2 with $0<t_1<t_2<1$. To simplify notation, we set for every $i,j\in\{0,1,\ldots,2k\}$,

$$\check C_k^{i,j}=\min_{i\wedge j\leq n\leq i\vee j}C_k(n).$$

Write $v_0^k=\varnothing,v_1^k,\ldots,v_{2k}^k=\varnothing$ for the vertices visited successively in the contour of the tree θ_k (see the end of subsection 3.1). Then we know that

$$C_k([2kt_1])=|v^k_{[2kt_1]}|,\ C_k([2kt_2])=|v^k_{[2kt_2]}|,\ V_k([2kt_1])=\ell^k_{v^k_{[2kt_1]}},\ V_k([2kt_2])=\ell^k_{v^k_{[2kt_2]}},$$

and furthermore $\check C_k^{[2kt_1],[2kt_2]}$ is the generation in θ_k of the last common ancestor to $v^k_{[2kt_1]}$ and $v^k_{[2kt_2]}$. From the properties of labels on the tree θ_k, we now see that conditionally on θ_k,

$$(V_k([2kt_1]),V_k([2kt_2]))\overset{\rm (d)}{=}\Big(\sum_{i=1}^{\check C_k^{[2kt_1],[2kt_2]}}\eta_i+\sum_{i=\check C_k^{[2kt_1],[2kt_2]}+1}^{C_k([2kt_1])}\eta'_i\,,\ \sum_{i=1}^{\check C_k^{[2kt_1],[2kt_2]}}\eta_i+\sum_{i=\check C_k^{[2kt_1],[2kt_2]}+1}^{C_k([2kt_2])}\eta''_i\Big)\quad(8)$$

where the variables η_i,η'_i,η''_i are independent and uniformly distributed over $\{-1,0,1\}$.

From (5), we have

$$\Big((2k)^{-1/2}C_k([2kt_1]),(2k)^{-1/2}C_k([2kt_2]),(2k)^{-1/2}\check C_k^{[2kt_1],[2kt_2]}\Big)\xrightarrow[k\to\infty]{\rm a.s.}(\mathbf{e}_{t_1},\mathbf{e}_{t_2},m_{\mathbf{e}}(t_1,t_2)).$$

By arguing as in the case $p=1$, we now deduce from (8) that

$$\Big(\frac{C_k([2kt_1])}{\sqrt{2k}},\frac{C_k([2kt_2])}{\sqrt{2k}},\Big(\frac{9}{8k}\Big)^{1/4}V_k([2kt_1]),\Big(\frac{9}{8k}\Big)^{1/4}V_k([2kt_2])\Big)$$
$$\xrightarrow[k\to\infty]{\rm (d)}(\mathbf{e}_{t_1},\mathbf{e}_{t_2},\sqrt{m_{\mathbf{e}}(t_1,t_2)}\,N+\sqrt{\mathbf{e}_{t_1}-m_{\mathbf{e}}(t_1,t_2)}\,N'\ ,\ \sqrt{m_{\mathbf{e}}(t_1,t_2)}\,N+\sqrt{\mathbf{e}_{t_2}-m_{\mathbf{e}}(t_1,t_2)}\,N'')$$

where N,N',N'' are three independent standard normal variables, which are also independent of $\mathbf{e}$. The limiting distribution in the last display is easily identified with that of $(\mathbf{e}_{t_1},\mathbf{e}_{t_2},Z_{t_1},Z_{t_2})$, and this gives the case $p=2$ in (6). The general case is proved by similar arguments and we leave details to the reader.

To complete the proof of Theorem 3.2, we need a tightness argument. The laws of the processes

$$\Big(\frac{1}{\sqrt{2k}}C_k(2k\,t)\Big)_{t\in[0,1]}$$

are tight by Theorem 1.9, and so we need only verify the tightness of the processes

$$\Big(\Big(\frac{9}{8k}\Big)^{1/4} V_k(2k\,t)\Big)_{t\in[0,1]}.$$

This is a consequence of the following lemma, which therefore completes the proof of Theorem 3.2. □

Lemma 3.3. *For every integer $p \geq 1$, there exists a constant $K_p < \infty$ such that, for every $k \geq 1$ and every $s, t \in [0,1]$,*

$$E\Big[\Big(\frac{V_k(2kt) - V_k(2ks)}{k^{1/4}}\Big)^{4p}\Big] \leq K_p\,|t-s|^p.$$

Proof. Simple arguments show that we may restrict our attention to the case when $s = i/(2k)$, $t = j/(2k)$, with $i, j \in \{0, 1, \ldots, 2k\}$. By using the same decomposition as in (8), we have

$$V_k(j) - V_k(i) \overset{\text{(d)}}{=} \sum_{n=1}^{d_{gr}(v_i^k, v_j^k)} \eta_n \tag{9}$$

where the random variables η_n are independent and uniform over $\{-1, 0, 1\}$ (and independent of θ_k) and

$$d_{gr}(v_i^k, v_j^k) = C_k(i) + C_k(j) - 2\check{C}_k^{i,j}$$

is the graph distance in the tree θ_k between vertices v_i^k and v_j^k. From (9) and by conditioning with respect to θ_k, we get the existence of a constant K'_p such that

$$E[(V_k(i) - V_k(j))^{4p}] \leq K'_p\, E[(d_{gr}(v_i^k, v_j^k))^{2p}].$$

So the lemma will be proved if we can verify the bound

$$E[(C_k(i) + C_k(j) - 2\check{C}_k^{i,j})^{2p}] \leq K''_p\,|j-i|^p \tag{10}$$

with a constant K''_p independent of k.

It is in fact sufficient to prove that this bound holds when $i = 0$. Indeed, fix $i \in \{0, 1, \ldots, 2k\}$ and write $i \oplus j = i + j$ if $i + j \leq 2k$, and $i \oplus j = i + j - 2k$ if $i + j > 2k$. Then, the process

$$\Big(C_k(i) + C_k(i\oplus j) - 2\check{C}_k^{i,i\oplus j}\Big)_{0\leq j\leq 2k}$$

has the same distribution as

$$(C_k(j))_{0\leq j\leq 2k}.$$

This identity in distribution corresponds to a re-rooting invariance property of the tree θ_k: The discrete path $(C_k(i) + C_k(i\oplus j) - 2\check{C}_k^{i,i\oplus j})_{0\leq j\leq 2k}$ is interpreted as the Dyck path associated with the tree θ_k "re-rooted at its i-th corner" (compare with Lemma 2.2 in the continuous setting), and the fact that this re-rooted tree has the same distribution as the original one gives the desired result (more details can be found in [14])).

By the preceding considerations, the bound (10) will follow if we can prove that

$$E[(C_k(i))^{2p}] \leq K''_p\, i^p \tag{11}$$

for every $i \in \{0, 1, \ldots, p\}$, with a constant K''_p independent of k. From a simple time-reversal argument, we may assume that $1 \leq i \leq k$.

Recall our notation $(S_n)_{n\geq 0}$ for simple random walk on $\mathbb{Z}$, and write P_j for the probability measure under which $S_0 = j$ (so $P_0 = P$ in our previous notation). Also recall that $T = \min\{n \geq 0 : S_n = -1\}$. We know that $(C_k(n))_{0\leq n\leq 2k}$ follows the distribution of $(S_n)_{0\leq n\leq 2k}$ under the probability measure $P_0(\cdot \mid T = 2k+1)$. Thus, if $\ell \in \mathbb{Z}_+$,

$$P(C_k(i) = \ell) = P_0(S_i = \ell \mid T = 2k+1) = \frac{P_0(S_i = \ell, T = 2k+1)}{P_0(T = 2k+1)}.$$

Using the Markov property of the random walk S, together with a time reversal argument, we get

$$P_0(S_i = \ell, T = 2k+1) = P_\ell(T = 2k+1-i) \times P_\ell(T = i+1). \tag{12}$$

At this point we use the celebrated Kemperman formula (see e.g. [16] Chapter 6), which tells us that for every $n \geq 0$,

$$P_\ell(T = n) = \frac{\ell+1}{n}\, P_\ell(S_n = -1).$$

Therefore we get:

$$\begin{aligned} P_0(S_i = \ell, T = 2k+1) &= \frac{(\ell+1)^2}{(i+1)(2k-i+1)}\, P_\ell(S_{2k+1-i} = -1)\, P_\ell(S_{i+1} = -1) \\ &= \frac{(\ell+1)^2}{(i+1)(2k-i+1)}\, P_0(S_{2k+1-i} = \ell+1)\, P_0(S_{i+1} = \ell+1) \end{aligned}$$

where the last equality follows from obvious symmetry and translation arguments. Recall that we assume $i \leq k$, and thus $k+1 \leq 2k+1-i \leq 2k+1$. At this point we use a classical local limit theorem (see e.g. Chapter 2 of Spitzer [18]), which gives the existence of two positive constants c_0 and c_1 such that:

$$P_0(T = 2k+1) = \frac{1}{2k+1}\, P_0(S_{2k+1} = -1) \geq c_0\, k^{-3/2}\ ,\ \ P_0(S_{2k+1-i} = \ell+1) \leq c_1\, k^{-1/2}.$$

By substituting these estimates in the previous identities, we arrive at

$$\begin{aligned} P(C_k(i) = \ell) &\leq (c_0 k^{-3/2})^{-1} \times \frac{(\ell+1)^2}{(i+1)(2k-i+1)} \times (c_1\, k^{-1/2}) \times P_0(S_{i+1} = \ell+1) \\ &\leq (c_0)^{-1} c_1\, \frac{(\ell+1)^2}{i+1}\, P_0(S_{i+1} = \ell+1) \end{aligned}$$

From this bound, it immediately follows that

$$E[(C_k(i))^{2p}] \leq (c_0)^{-1} c_1\, (i+1)^{-1}\, E_0[(S_{i+1})^{2p+2}]. \tag{13}$$

However for simple random walk, the bound

$$E_0[(S_{i+1})^{2p+2}] \leq \overline{K}_p\, (i+1)^{p+1}$$

is well known and easy to prove. By substituting this bound in (13), we arrive at the desired estimate (11). $\square$

References

[1] D. Aldous. The continuum random tree I. *Ann. Probab.* **19** (1991), 1-28.

[2] D. Aldous. The continuum random tree III. *Ann. Probab.* **21** (1993), 248-289.

[3] D. Burago, Y. Burago, S. Ivanov. *A Course in Metric Geometry*. Graduate Studies in Mathematics, vol. 33. AMS, Boston, 2001.

[4] P. Chassaing, G. Schaeffer. Random planar lattices and integrated superBrownian excursion. *Probab. Th. Rel. Fields* **128** (2004), 161-212.

[5] K.L. Chung. Excursions in Brownian motion. *Ark. Mat.* **14** (1976) 155-177.

[6] T. Duquesne. A limit theorem for the contour process of conditioned Galton-Watson trees. *Ann. Probab.* **31** (2003), 996–1027.

[7] T. Duquesne, J.F. Le Gall. Random Trees, Lévy Processes and Spatial Branching Processes. *Astérisque* **281** (2002)

[8] T. Duquesne, J.F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Th. Rel. Fields* **131** (2005), 553-603.

[9] S.N. Evans, A. Winter. Subtree prune and re-graft: A reversible tree-valued Markov process. *Ann. Probab.* **34** (2006), 918-961.

[10] M. Gromov. *Metric Structures for Riemannian and Non-Riemannian Spaces.* Birkhäuser 1999.

[11] W.D. Kaigh. An invariance principle for random walk conditioned by a late return to zero. *Ann. Probab.* **4** (1976), 115-121.

[12] J.F. Le Gall. *Spatial Branching Processes, Random Snakes and Partial Differential Equations.* Birkhäuser 1999.

[13] J.F. Le Gall. Random trees and applications. *Probab. Surveys* **2**(2005), 245-311.

[14] J.F. Marckert, A. Mokkadem. Limit of normalized quadrangulations. The Brownian map. *Ann. Probab.*, 34 (2006), 2144-2202.

[15] J. Neveu. Arbres et processus de Galton-Watson. *Ann. Inst. Henri Poincaré Probab. Stat.* **22** (1986), 199-207.

[16] J.W. Pitman. Combinatorial stochastic processes. Ecole d'été de probabilités de Saint-Flour 2002. *Lecture Notes Math.* **1875** Springer 2006.

[17] D. Revuz, M. Yor. *Continuous Martingales and Brownian Motion.* Springer 1991.

[18] F. Spitzer. *Principles of Random Walk.* Van Nostrand 1963.

[19] R.P. Stanley. *Enumerative Combinatorics, Vol.2.* Cambridge University Press 1999