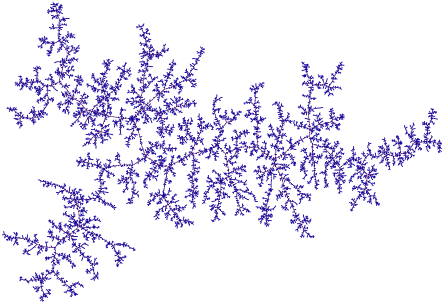
Easter Probability Meeting, Sheffield, 9th-13th April 2018

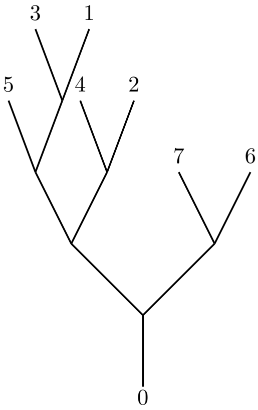
Around the Brownian continuum random tree



Christina Goldschmidt

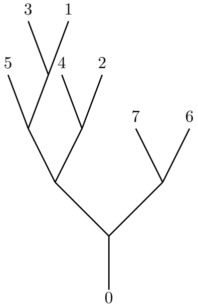
UNIVERSITY OF OXFORD

1. INTRODUCTION: BINARY TREES



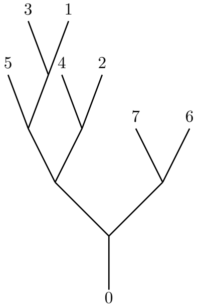
Binary leaf-labelled trees

- Let $\mathbb{T}_n$ be the set of planted binary leaf-labelled trees with n labelled leaves (note: we don't distinguish a planar ordering around each vertex).
- The root, labelled 0 is, by convention, not a leaf.
- Note that every element of $\mathbb{T}_n$ has $n-1$ internal vertices (which are not labelled) and $2n-1$ edges.



Binary leaf-labelled trees

- Let $\mathbb{T}_n$ be the set of planted binary leaf-labelled trees with n labelled leaves (note: we don't distinguish a planar ordering around each vertex).
- The root, labelled 0 is, by convention, not a leaf.
- Note that every element of $\mathbb{T}_n$ has $n-1$ internal vertices (which are not labelled) and $2n-1$ edges.



$$|\mathbb{T}_n| = \frac{1}{n}\binom{2n-2}{n-1} \quad \text{(Catalan numbers)}$$

Uniform binary leaf-labelled trees

$$|\mathbb{T}_n| = \frac{1}{n}\binom{2n-2}{n-1} \sim \frac{4^{n-1}}{n^{3/2}\sqrt{\pi}} \quad \text{as } n \to \infty.$$

Uniform binary leaf-labelled trees

$$|\mathbb{T}_n| = \frac{1}{n}\binom{2n-2}{n-1} \sim \frac{4^{n-1}}{n^{3/2}\sqrt{\pi}} \quad \text{as } n \to \infty.$$

Our first object of interest is a uniform random element of $\mathbb{T}_n$.

Uniform binary leaf-labelled trees

$$|\mathbb{T}_n| = \frac{1}{n}\binom{2n-2}{n-1} \sim \frac{4^{n-1}}{n^{3/2}\sqrt{\pi}} \quad \text{as } n \to \infty.$$

Our first object of interest is a uniform random element of $\mathbb{T}_n$.

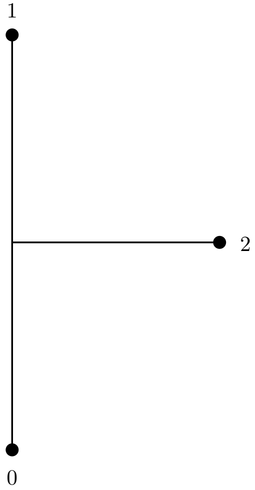
Rémy's algorithm recursively constructs a sequence $(T_n)_{n\geq 1}$ of trees such that T_n is uniform on $\mathbb{T}_n$ for each n.

Rémy's algorithm

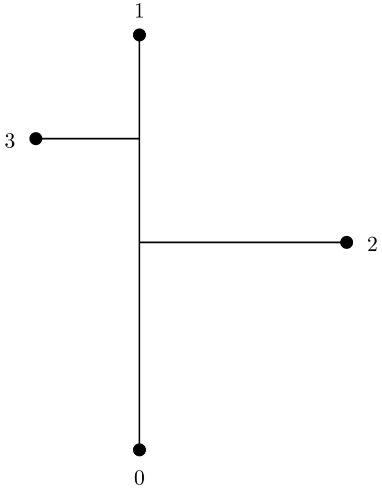
- Start from a single edge with endpoints labelled 0 and 1.
- At step $n \geq 2$, pick an edge uniformly at random, divide it into two edges, insert a new vertex in the middle and attach to that vertex a new edge with a leaf labelled n at its other end.

Rémy's algorithm

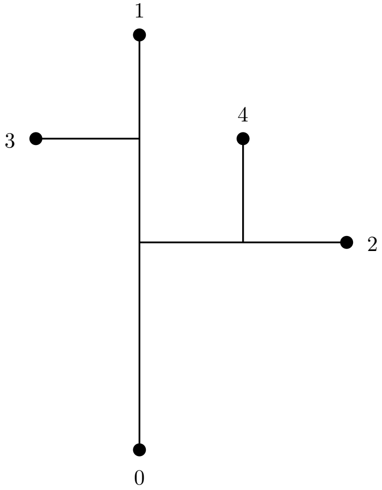
Rémy's algorithm



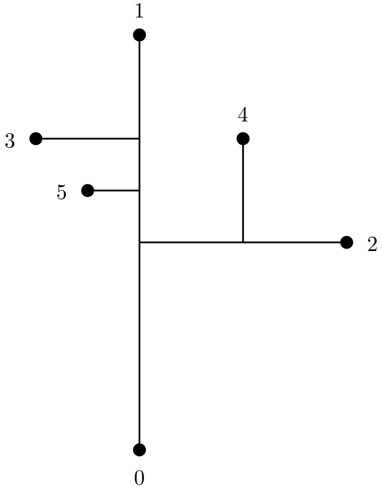
Rémy's algorithm



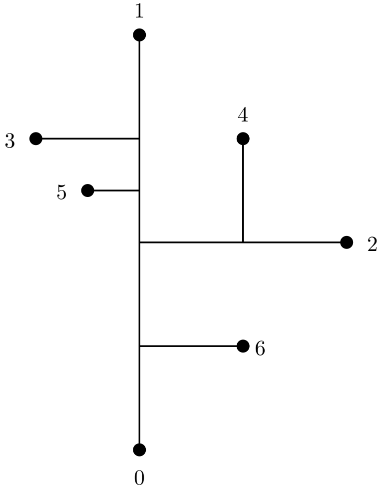
Rémy's algorithm



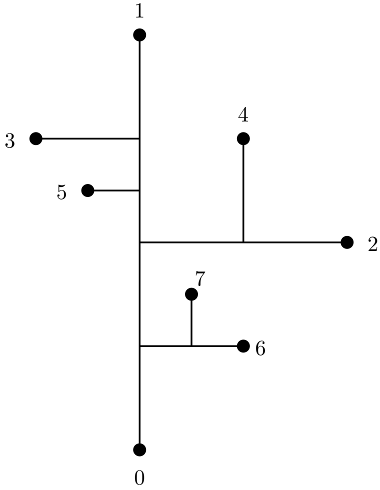
Rémy's algorithm



Rémy's algorithm

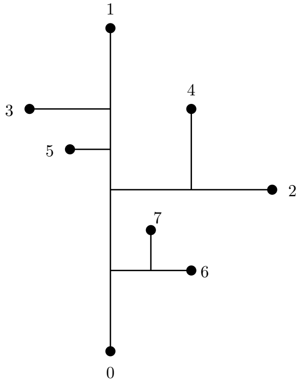


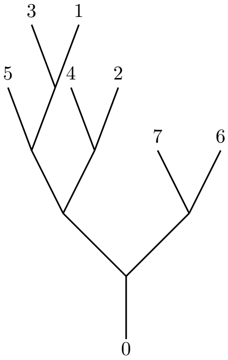
Rémy's algorithm



Rémy's algorithm

Claim: for each n, T_n is a uniform element of $\mathbb{T}_n$.





[J.-L. Rémy, **Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire**, *RAIRO. Informatique théorique* **19**:2 (1985), pp.179–195]

Taking limits

Vague question: what can we say about T_n as $n \to \infty$?

Taking limits

Vague question: what can we say about T_n as $n \to \infty$?

Concrete first question: as $n \to \infty$, how does the distance between 0 and 1 behave?

An urn in Rémy's algorithm

The total number of edges present at step n is equal to $2n - 1$.

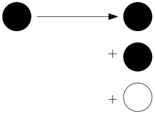
Consider the number of edges in the path between 0 and 1:

- If we add our new leaf somewhere along that path, it gets longer by 1.
- If we add our new leaf anywhere else, the length of the path remains the same.

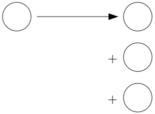
An urn in Rémy's algorithm

We have an urn process with two colours, say black and white, where each black ball represents an edge in the path between 0 and 1, and each white ball represents an edge elsewhere.

When we pick a black ball, we replace it in the urn together with one black and one white ball.



When we pick a white ball, we replace it in the urn together with two new white balls.



We start with a single black ball. At step n, we always have $2n - 1$ balls present.

An urn in Rémy's algorithm

Let B_n be the number of black balls at step n.

An urn in Rémy's algorithm

Let B_n be the number of black balls at step n.

We have $B_1 = 1$.

For $n \geq 1$,

$$\mathbb{E}\left[B_{n+1} | \mathcal{F}_n\right] = \frac{B_n}{2n-1}(B_n + 1) + \frac{2n-1-B_n}{2n-1} B_n$$

An urn in Rémy's algorithm

Let B_n be the number of black balls at step n.

We have $B_1 = 1$.

For $n \geq 1$,

$$\mathbb{E}\left[B_{n+1} | \mathcal{F}_n\right] = \frac{B_n}{2n-1}(B_n + 1) + \frac{2n-1-B_n}{2n-1} B_n = \frac{2n}{2n-1} B_n.$$

An urn in Rémy's algorithm

Let B_n be the number of black balls at step n.

We have $B_1 = 1$.

For $n \geq 1$,

$$\mathbb{E}\left[B_{n+1} | \mathcal{F}_n\right] = \frac{B_n}{2n-1}(B_n + 1) + \frac{2n-1-B_n}{2n-1} B_n = \frac{2n}{2n-1} B_n.$$

Define a sequence by $b_1 = 1$ and $b_{n+1} = \frac{2^{2n}(n!)^2}{(2n)!}$ for $n \geq 1$. Then

$$b_{n+1} = \frac{2n}{2n-1} b_n.$$

An urn in Rémy's algorithm

Let B_n be the number of black balls at step n.

We have $B_1 = 1$.

For $n \geq 1$,

$$\mathbb{E}\left[B_{n+1}|\mathcal{F}_n\right] = \frac{B_n}{2n-1}(B_n+1) + \frac{2n-1-B_n}{2n-1}B_n = \frac{2n}{2n-1}B_n.$$

Define a sequence by $b_1 = 1$ and $b_{n+1} = \frac{2^{2n}(n!)^2}{(2n)!}$ for $n \geq 1$. Then

$$b_{n+1} = \frac{2n}{2n-1}b_n.$$

Then we have that

$$\left(\frac{B_n}{b_n}\right)_{n\geq 1} \quad \text{is a non-negative martingale.}$$

Martingale limit

$(B_n/b_n)_{n\geq 1}$ is also bounded in L^2, so it has an almost sure limit.

Since

$$b_{n+1} = \frac{2^{2n}(n!)^2}{(2n)!} \sim \sqrt{\pi n},$$

we get that

$$\frac{B_n}{\sqrt{2n}} \to L \quad \text{a.s. as } n \to \infty.$$

[P. Marchal, **A note on the fragmentation of the stable tree**, *Fifth Colloquium on Mathematics and Computer Science, DMTCS* (2008), pp.489–500]

Limiting distribution for the length

It also turns out that the law of B_{n+1} is explicit:

$$\mathbb{P}\left(B_{n+1} = k\right) = \frac{k-1}{n} 2^{k-1} \frac{\binom{2n-k}{n-1}}{\binom{2n}{n}}$$

and so

$$\mathbb{P}\left(B_{n+1} = \lfloor x\sqrt{2n} \rfloor\right) \sim \frac{x}{\sqrt{2n}} e^{-x^2/2}, \quad x > 0.$$

Limiting distribution for the length

It also turns out that the law of B_{n+1} is explicit:

$$\mathbb{P}\left(B_{n+1}=k\right)=\frac{k-1}{n}2^{k-1}\frac{\binom{2n-k}{n-1}}{\binom{2n}{n}}$$

and so

$$\mathbb{P}\left(B_{n+1}=\lfloor x\sqrt{2n}\rfloor\right)\sim\frac{x}{\sqrt{2n}}e^{-x^2/2},\quad x>0.$$

In other words, we get

$$\frac{B_n}{\sqrt{2n}}\to L\quad \text{a.s. as } n\to\infty,$$

where the limit L has the Rayleigh distribution, with density $xe^{-x^2/2}$ on $\mathbb{R}_+$.

[P. Flajolet, P. Dumas and V. Puyhaubert, **Some exactly solvable models of urn process theory**, *Fourth Colloquium on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities, DMTCS* (2006), pp.59–118]

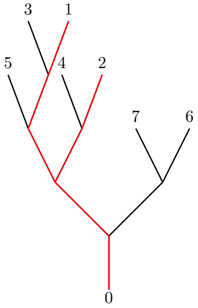
Consequences

The distance between 0 and 1 varies as $\sqrt{2n}$, with a nice almost sure limit. What can we say about the distances between the other leaves as $n \to \infty$?

Consequences

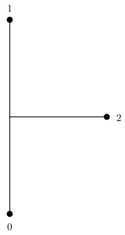
The distance between 0 and 1 varies as $\sqrt{2n}$, with a nice almost sure limit. What can we say about the distances between the other leaves as $n \to \infty$?

For example, let's think about the distance from 2 to the path between 0 and 1, and the position along that path at which it branches off.



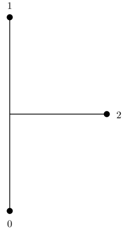
More urns: self-similarity

At step 2 of Rémy's algorithm, we have



More urns: self-similarity

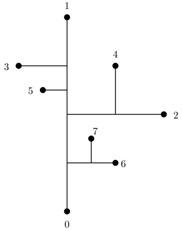
At step 2 of Rémy's algorithm, we have



Each of the three parts here behaves precisely as a little copy of Rémy's algorithm, although the numbers of leaves we add to each copy are dependent. A useful consequence is that given the three sets of leaves, these three trees are themselves uniform binary leaf-labelled trees.

More urns: self-similarity

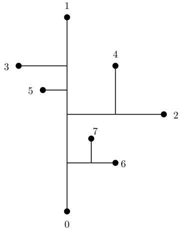
At step 2 of Rémy's algorithm, we have



Each of the three parts here behaves precisely as a little copy of Rémy's algorithm, although the numbers of leaves we add to each copy are dependent. A useful consequence is that given the three sets of leaves, these three trees are themselves uniform binary leaf-labelled trees.

More urns: self-similarity

At step 2 of Rémy's algorithm, we have

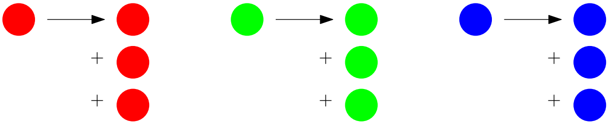


Each of the three parts here behaves precisely as a little copy of Rémy's algorithm, although the numbers of leaves we add to each copy are dependent. A useful consequence is that given the three sets of leaves, these three trees are themselves uniform binary leaf-labelled trees.

How many leaves end up in each of the three copies?

More urns: self-similarity

Consider Pólya's urn with three colours, red, green and blue. We start with one ball of each colour. We pick a ball at random and replace it in the urn with two more of the same colour. Let R_n, G_n, B_n be the numbers of red, green and blue balls respectively at step n (let us now re-number the steps from 0, so that $R_0 = G_0 = B_0 = 1$).



More urns: self-similarity

It is then standard that

$$\frac{1}{2n+3}(R_n, G_n, B_n) \to (\Delta_1, \Delta_2, \Delta_3) \quad \text{a.s. as } n \to \infty,$$

where $(\Delta_1, \Delta_2, \Delta_3) \sim \text{Dirichlet}(1/2, 1/2, 1/2)$.

More urns: self-similarity

It is then standard that

$$\frac{1}{2n+3}(R_n, G_n, B_n) \to (\Delta_1, \Delta_2, \Delta_3) \quad \text{a.s. as } n \to \infty,$$

where $(\Delta_1, \Delta_2, \Delta_3) \sim \text{Dirichlet}(1/2, 1/2, 1/2)$.

The Dirichlet distribution with parameters $\alpha_1, \alpha_2, \ldots, \alpha_k > 0$ has density

$$\frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} x_1^{\alpha_1 - 1} \ldots x_k^{\alpha_k - 1}$$

with respect to Lebesgue measure on

$$\left\{ \mathbf{x} = (x_1, \ldots, x_k) \in \mathbb{R}_+^k : \sum_{i=1}^k x_i = 1 \right\}.$$

More urns: self-similarity

It is then standard that

$$\frac{1}{2n+3}(R_n, G_n, B_n) \to (\Delta_1, \Delta_2, \Delta_3) \quad \text{a.s. as } n \to \infty,$$

where $(\Delta_1, \Delta_2, \Delta_3) \sim \text{Dirichlet}(1/2, 1/2, 1/2)$.

Let $\gamma_i \sim \text{Gamma}(\alpha_i, 1)$ for $1 \leq i \leq k$ independently. Then

$$\frac{1}{\sum_{i=1}^k \gamma_i}(\gamma_1, \gamma_2, \ldots, \gamma_k) \sim \text{Dir}(\alpha_1, \ldots, \alpha_k),$$

(and is independent of $\sum_{i=1}^k \gamma_i$).

More urns: self-similarity

The numbers of leaves in each of the three subtrees are given by

$$N_n^R = (R_n + 1)/2, \quad N_n^G = (G_n + 1)/2, \quad N_n^B = (B_n + 1)/2.$$

So we have

$$\frac{1}{n}(N_n^R, N_n^G, N_n^B) \to (\Delta_1, \Delta_2, \Delta_3) \quad \text{a.s.}$$

More urns: self-similarity

The numbers of leaves in each of the three subtrees are given by

$$N_n^R = (R_n + 1)/2, \quad N_n^G = (G_n + 1)/2, \quad N_n^B = (B_n + 1)/2.$$

So we have

$$\frac{1}{n}(N_n^R, N_n^G, N_n^B) \to (\Delta_1, \Delta_2, \Delta_3) \quad \text{a.s.}$$

Writing L_n^R, L_n^G, L_n^B for the lengths of the three paths at step n, we see that they look like small copies of the first urn model run for numbers of steps which are approximately $n\Delta_1$, $n\Delta_2$ and $n\Delta_3$.

More urns: self-similarity

The numbers of leaves in each of the three subtrees are given by

$$N_n^R = (R_n + 1)/2, \quad N_n^G = (G_n + 1)/2, \quad N_n^B = (B_n + 1)/2.$$

So we have

$$\frac{1}{n}(N_n^R, N_n^G, N_n^B) \to (\Delta_1, \Delta_2, \Delta_3) \quad \text{a.s.}$$

Writing L_n^R, L_n^G, L_n^B for the lengths of the three paths at step n, we see that they look like small copies of the first urn model run for numbers of steps which are approximately $n\Delta_1$, $n\Delta_2$ and $n\Delta_3$. It follows that

$$\frac{1}{\sqrt{2n}}(L_n^R, L_n^G, L_n^B) \to (\sqrt{\Delta_1}L_1, \sqrt{\Delta_2}L_2, \sqrt{\Delta_3}L_3) \quad \text{a.s.}$$

where L_1, L_2, L_3 are i.i.d. Rayleigh random variables, independent of $(\Delta_1, \Delta_2, \Delta_3)$.

Limiting subtree lengths

An elementary calculation yields that

$$(\sqrt{\Delta_1}L_1, \sqrt{\Delta_2}L_2, \sqrt{\Delta_3}L_3) \stackrel{d}{=} \sqrt{\Gamma_2} \times \mathrm{Dir}(1,1,1),$$

where $\Gamma_2 \sim \mathrm{Gamma}(2, 1/2)$ and the two factors are independent.

Limiting subtree lengths

An elementary calculation yields that

$$(\sqrt{\Delta_1}L_1, \sqrt{\Delta_2}L_2, \sqrt{\Delta_3}L_3) \stackrel{d}{=} \sqrt{\Gamma_2} \times \mathrm{Dir}(1,1,1),$$

where $\Gamma_2 \sim \mathrm{Gamma}(2, 1/2)$ and the two factors are independent.

More generally, if we consider the subtree spanned by 0 and the leaves labelled $1, 2, \ldots, k$, we get $2k-1$ edges whose lengths are distributed as

$$\sqrt{\Gamma_k} \times \mathrm{Dir}(\underbrace{1, 1, \ldots, 1}_{2k-1}),$$

where again $\Gamma_k \sim \mathrm{Gamma}(k, 1/2)$ and the two factors are independent.

Limiting subtree lengths

An elementary calculation yields that

$$(\sqrt{\Delta_1}L_1, \sqrt{\Delta_2}L_2, \sqrt{\Delta_3}L_3) \stackrel{d}{=} \sqrt{\Gamma_2} \times \mathrm{Dir}(1,1,1),$$

where $\Gamma_2 \sim \mathrm{Gamma}(2, 1/2)$ and the two factors are independent.

More generally, if we consider the subtree spanned by 0 and the leaves labelled $1, 2, \ldots, k$, we get $2k-1$ edges whose lengths are distributed as

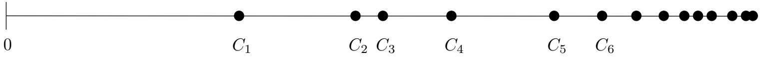
$$\sqrt{\Gamma_k} \times \mathrm{Dir}(\underbrace{1, 1, \ldots, 1}_{2k-1}),$$

where again $\Gamma_k \sim \mathrm{Gamma}(k, 1/2)$ and the two factors are independent.

Note that the $k = 1$ case fits into this pattern, since Rayleigh $\stackrel{d}{=} \sqrt{\Gamma_1}$.

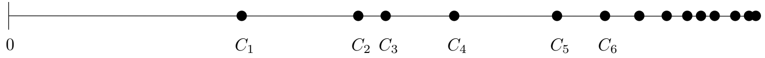
A limiting version of Rémy's algorithm: Aldous' line-breaking construction of the Brownian CRT

Take an inhomogeneous Poisson process on $\mathbb{R}_+$ of intensity t at t.



A limiting version of Rémy's algorithm: Aldous' line-breaking construction of the Brownian CRT

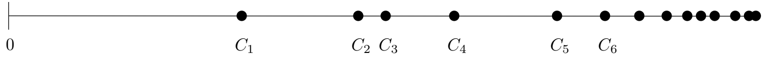
Take an inhomogeneous Poisson process on $\mathbb{R}_+$ of intensity t at t.



A useful way of constructing this is to let $E_1, E_2, \ldots$ be i.i.d. Exp$(1/2)$ and set $C_i = \sqrt{\sum_{j=1}^{i} E_j}$.

A limiting version of Rémy's algorithm: Aldous' line-breaking construction of the Brownian CRT

Take an inhomogeneous Poisson process on $\mathbb{R}_+$ of intensity t at t.

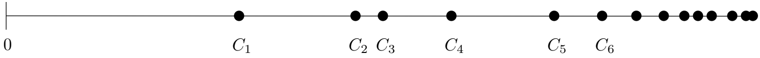


A useful way of constructing this is to let $E_1, E_2, \ldots$ be i.i.d. Exp$(1/2)$ and set $C_i = \sqrt{\sum_{j=1}^{i} E_j}$.

- Consider the line-segments $[0, C_1)$, $[C_1, C_2), \ldots$.

A limiting version of Rémy's algorithm: Aldous' line-breaking construction of the Brownian CRT

Take an inhomogeneous Poisson process on $\mathbb{R}_+$ of intensity t at t.

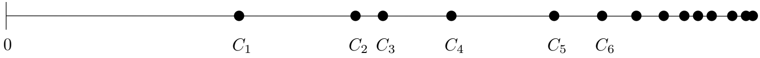


A useful way of constructing this is to let $E_1, E_2, \ldots$ be i.i.d. Exp$(1/2)$ and set $C_i = \sqrt{\sum_{j=1}^{i} E_j}$.

- Consider the line-segments $[0, C_1)$, $[C_1, C_2), \ldots$.
- Start from $[0, C_1)$ and proceed inductively.

A limiting version of Rémy's algorithm: Aldous' line-breaking construction of the Brownian CRT

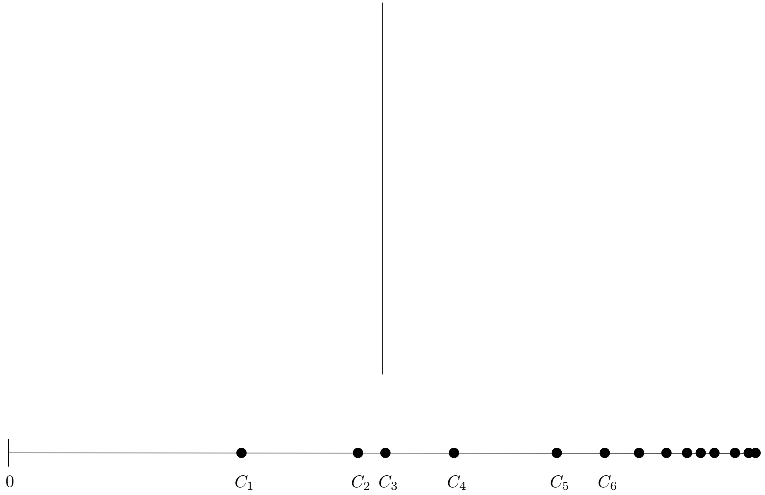
Take an inhomogeneous Poisson process on $\mathbb{R}_+$ of intensity t at t.



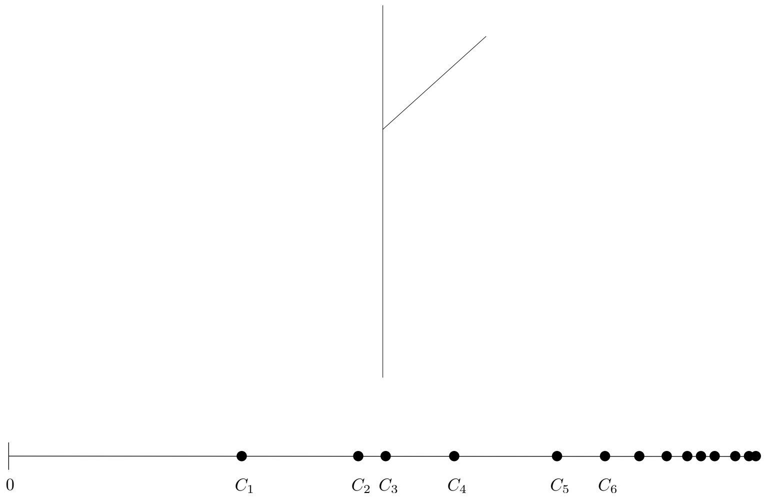
A useful way of constructing this is to let $E_1, E_2, \ldots$ be i.i.d. Exp$(1/2)$ and set $C_i = \sqrt{\sum_{j=1}^{i} E_j}$.

- Consider the line-segments $[0, C_1)$, $[C_1, C_2), \ldots$.
- Start from $[0, C_1)$ and proceed inductively.
- For $i \geq 2$, attach $[C_{i-1}, C_i)$ at a random point chosen uniformly over the existing tree.

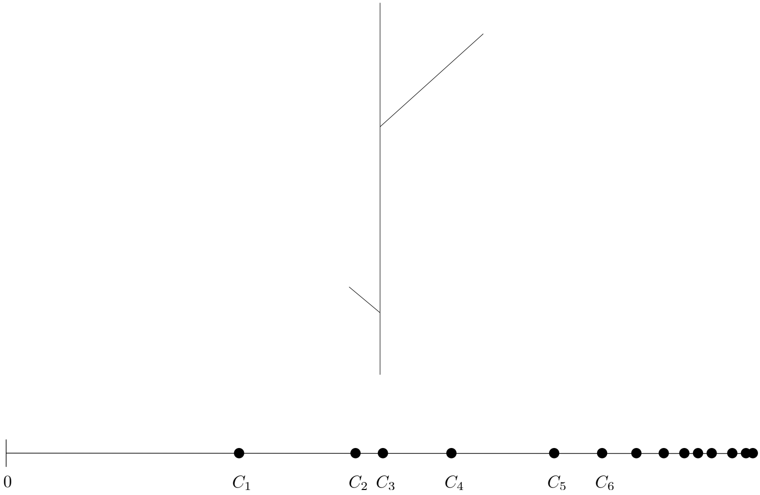
Line-breaking construction



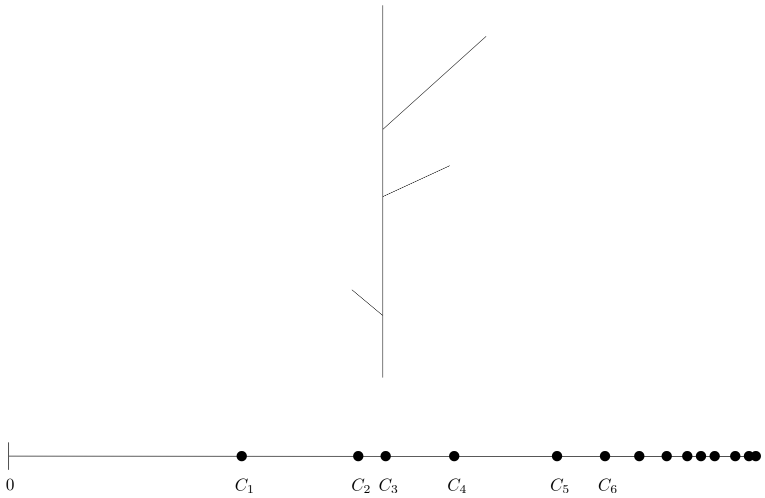
Line-breaking construction



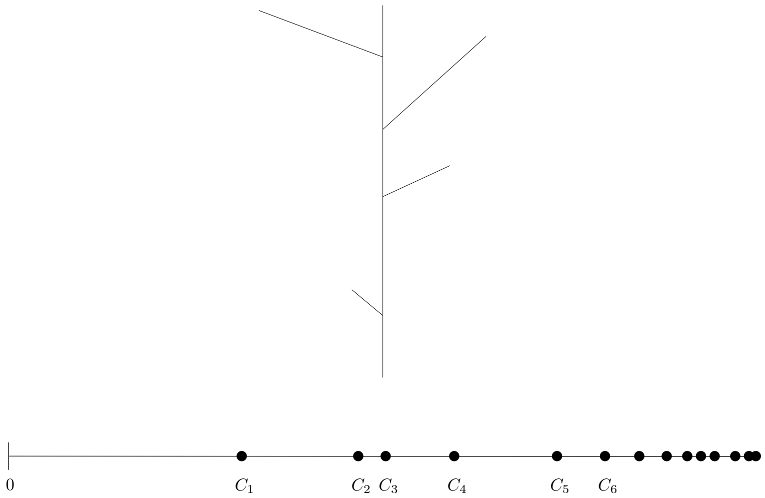
Line-breaking construction



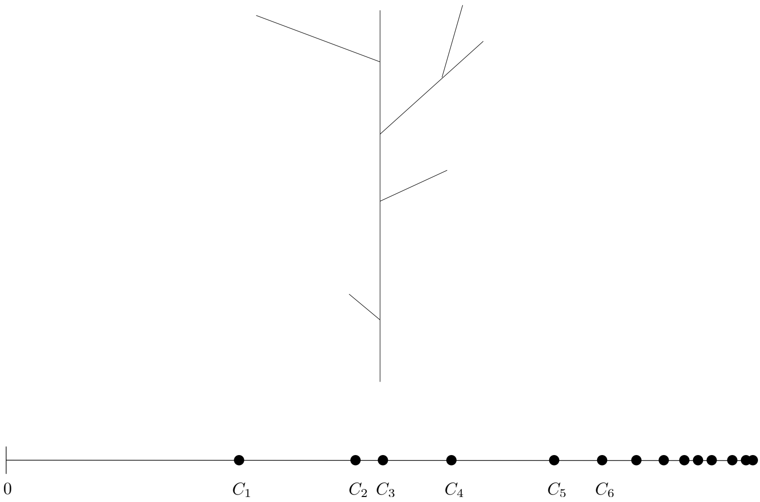
Line-breaking construction



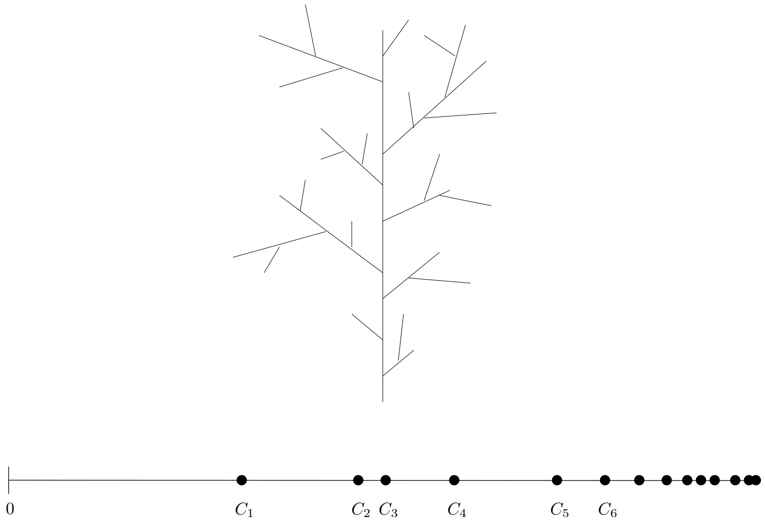
Line-breaking construction



Line-breaking construction



Line-breaking construction



Why is this the right limit?

Claim: this gives the almost sure limit of the subtree spanned by 0 and the leaves 1, 2, ..., k in the rescaled version of Rémy's algorithm.

- The tree at step $k \geq 1$ has total length

$$C_k = \sqrt{\sum_{i=1}^{k} E_i} \stackrel{d}{=} \sqrt{\mathrm{Gamma}(k, 1/2)}.$$

- The combinatorics of the attachment mechanism are exactly the same as in Rémy's algorithm – so the underlying binary leaf-labelled tree has the right distribution.
- A calculation shows that the cut-points and attachment points split up the interval $[0, C_k)$ uniformly.

The line-breaking definition of the Brownian CRT

- Start from $[0, C_1)$ and proceed inductively.
- For $i \geq 1$, sample B_i uniformly from $[0, C_i)$ and attach $[C_i, C_{i+1})$ at the corresponding point of the tree constructed so far (this is a point chosen uniformly at random over the existing tree).

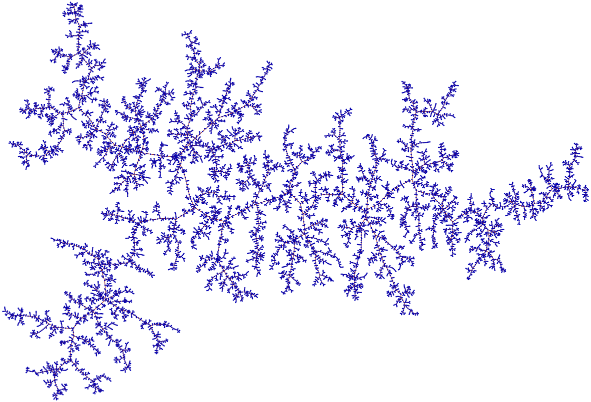
The line-breaking definition of the Brownian CRT

- Start from $[0, C_1)$ and proceed inductively.
- For $i \geq 1$, sample B_i uniformly from $[0, C_i)$ and attach $[C_i, C_{i+1})$ at the corresponding point of the tree constructed so far (this is a point chosen uniformly at random over the existing tree).

Now take the union of all the branches, thought of as a path metric space, and then take its completion.

This procedure gives (slightly informally expressed) definition of the Brownian continuum random tree (CRT) which is the key object in this minicourse.

The line-breaking definition of the Brownian CRT



[Picture by Igor Kortchemski]

The scaling limit of the uniform binary leaf-labelled tree

In the next section, we will make sense of the following statement.

Theorem. (Marchal (2003), Curien and Haas (2013))
As $n \to \infty$,

$$\frac{1}{\sqrt{2n}} T_n \to \mathcal{T} \quad \text{a.s.}$$

where $\mathcal{T}$ is the Brownian CRT.

The scaling limit of the uniform binary leaf-labelled tree

In the next section, we will make sense of the following statement.

Theorem. (Marchal (2003), Curien and Haas (2013))
As $n \to \infty$,

$$\frac{1}{\sqrt{2n}} T_n \to \mathcal{T} \quad \text{a.s.}$$

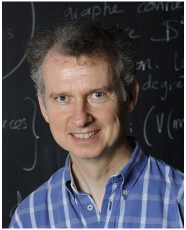
where $\mathcal{T}$ is the Brownian CRT.

We need to know what sort of objects we're really dealing with, and what is the topology in which the convergence occurs.

2. $\mathbb{R}$-TREES AND CONVERGENCE

Key reference:

Jean-François Le Gall, **Random trees and applications**, *Probability Surveys* **2** (2005) pp.245-311.



Continuous trees

We want a continuous notion of a tree. We don't really care about vertices: the important aspects are the shape of the tree and the distances. So it makes sense to think in terms of metric spaces.

$\mathbb{R}$-trees

Definition. A compact metric space $(\mathcal{T}, d)$ is an $\mathbb{R}$-tree if for all $x, y \in \mathcal{T}$,

- There exists a unique shortest path $[[x, y]]$ from x to y (of length $d(x, y)$).

- The only non-self-intersecting path from x to y is $[[x, y]]$.

$\mathbb{R}$-trees

Definition. A compact metric space $(\mathcal{T}, d)$ is an $\mathbb{R}$-tree if for all $x, y \in \mathcal{T}$,

- There exists a unique shortest path $[[x, y]]$ from x to y (of length $d(x, y)$). (There is a unique isometric map $f_{x,y}$ from $[0, d(x, y)]$ into $\mathcal{T}$ such that $f(0) = x$ and $f(d(x, y)) = y$. We write $f_{x,y}([0, d(x, y)]) = [[x, y]]$.)
- The only non-self-intersecting path from x to y is $[[x, y]]$. (If g is a continuous injective map from $[0, 1]$ into $\mathcal{T}$, such that $g(0) = x$ and $g(1) = y$, then $g([0, 1]) = [[x, y]]$.)

$\mathbb{R}$-trees

Definition. A compact metric space $(\mathcal{T}, d)$ is an $\mathbb{R}$-tree if for all $x, y \in \mathcal{T}$,

- There exists a unique shortest path $[[x, y]]$ from x to y (of length $d(x, y)$). (There is a unique isometric map $f_{x,y}$ from $[0, d(x, y)]$ into $\mathcal{T}$ such that $f(0) = x$ and $f(d(x, y)) = y$. We write $f_{x,y}([0, d(x, y)]) = [[x, y]]$.)
- The only non-self-intersecting path from x to y is $[[x, y]]$. (If g is a continuous injective map from $[0, 1]$ into $\mathcal{T}$, such that $g(0) = x$ and $g(1) = y$, then $g([0, 1]) = [[x, y]]$.)

An element $v \in \mathcal{T}$ is called a vertex.
A rooted $\mathbb{R}$-tree has a distinguished vertex ρ called the root.
The height of a vertex v is its distance $d(\rho, v)$ from the root.
A leaf is a vertex v such that $v \notin [[\rho, w]]$ for any $w \neq v$.

Coding $\mathbb{R}$-trees

Let $h : [0,1] \to \mathbb{R}^+$ be an excursion, that is a continuous function such that $h(0) = h(1) = 0$ and $h(x) > 0$ for $x \in (0,1)$.

Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...

Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...

Coding $\mathbb{R}$-trees

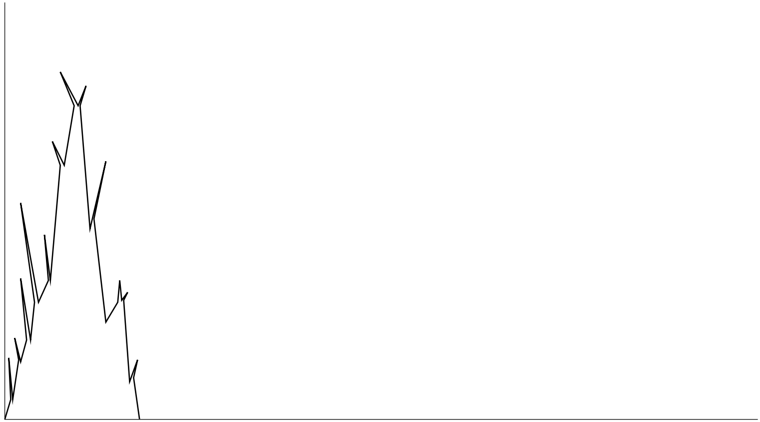
Now put glue on the underside of the excursion and push the two sides together...

Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...

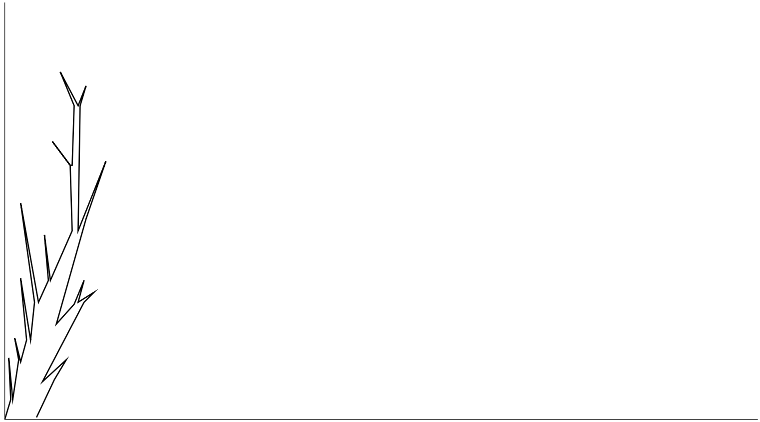
Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...



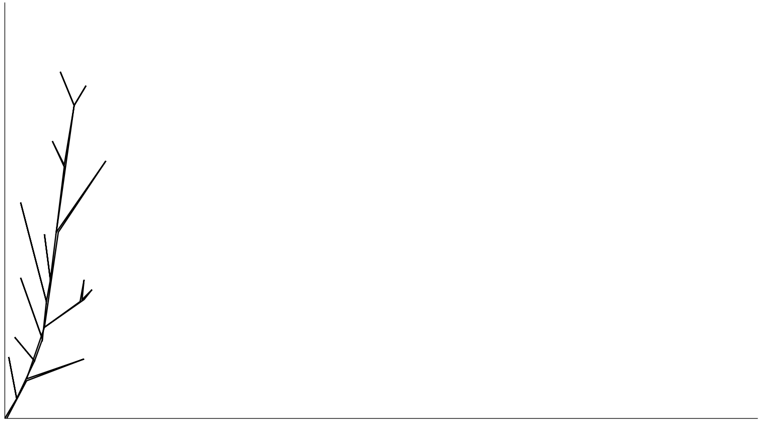
Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...



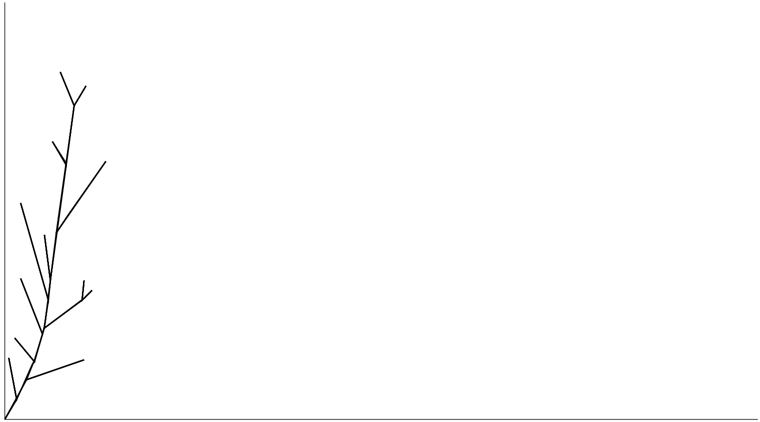
Coding $\mathbb{R}$-trees

Now put glue on the underside of the excursion and push the two sides together...



Coding $\mathbb{R}$-trees

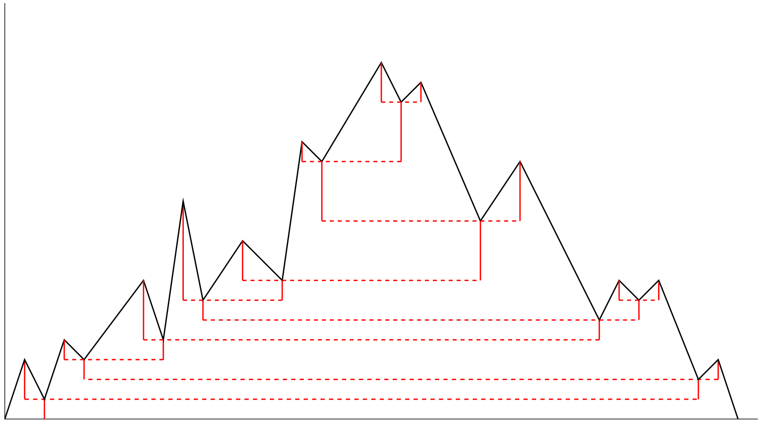
Now put glue on the underside of the excursion and push the two sides together to get a tree.



Coding $\mathbb{R}$-trees

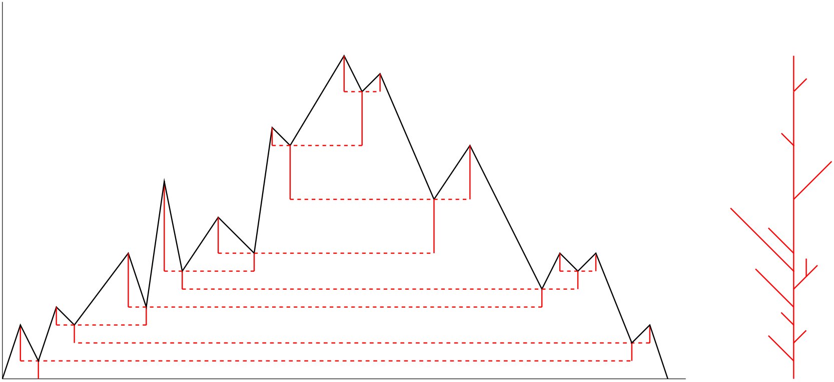
Formally, use h to define a distance:

$$d_h(x, y) = h(x) + h(y) - 2 \inf_{x \wedge y \leq z \leq x \vee y} h(z).$$



Coding $\mathbb{R}$-trees

Let $y \sim y'$ if $d_h(y, y') = 0$ and take the quotient $\mathcal{T}_h = [0,1]/\sim$.



Coding $\mathbb{R}$-trees

Theorem. For any excursion h, $(\mathcal{T}_h, d_h)$ is an $\mathbb{R}$-tree.

Write $\pi_h : [0,1] \to \mathcal{T}_h$ for the projection map.

We will often root $\mathcal{T}_h$ at $\rho = \pi_h(0) = \pi_h(1)$.

A natural measure

We will want to be able to sample random points in our trees.
There is a natural "uniform" measure μ_h which is the push-forward of the Lebesgue measure on $[0, 1]$ onto $\mathcal{T}_h$.

A natural measure

We will want to be able to sample random points in our trees. There is a natural "uniform" measure μ_h which is the push-forward of the Lebesgue measure on $[0,1]$ onto $\mathcal{T}_h$.

To pick a point of $\mathcal{T}_h$ according to μ_h, we simply sample $U \sim \mathrm{U}[0,1]$ and then take our point to be $\pi_h(U)$.

A natural measure

We will want to be able to sample random points in our trees.
There is a natural "uniform" measure μ_h which is the push-forward of the Lebesgue measure on $[0,1]$ onto $\mathcal{T}_h$.

To pick a point of $\mathcal{T}_h$ according to μ_h, we simply sample $U \sim \mathrm{U}[0,1]$ and then take our point to be $\pi_h(U)$.

We will typically think of our continuous trees as triples $(\mathcal{T}_h, d_h, \mu_h)$.

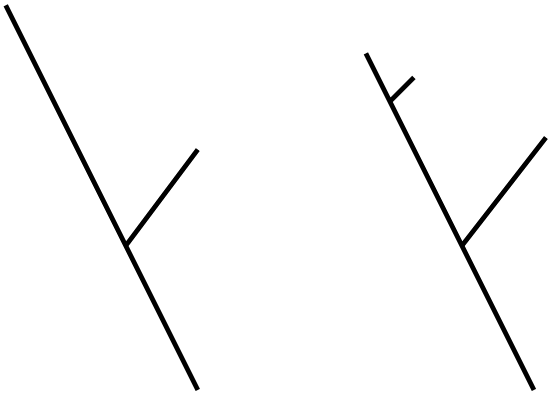
Topological considerations

Let $\mathbb{M}$ be the space of compact metric spaces endowed with a Borel probability measure, up to measure-preserving isometry.

We will define a metric d_{GHP}, the Gromov-Hausdorff-Prokhorov distance on $\mathbb{M}$.

Topological considerations

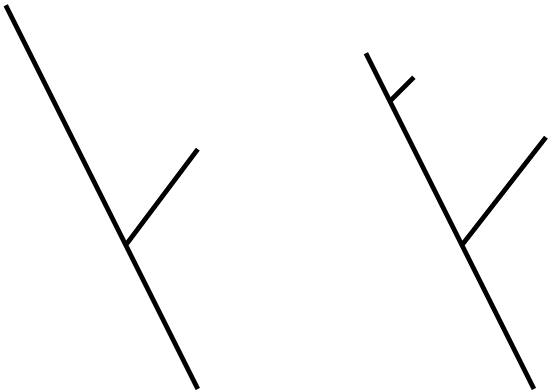
Suppose that (X, d) and (X', d') are compact metric spaces.



Topological considerations

Suppose that (X, d) and (X', d') are compact metric spaces.

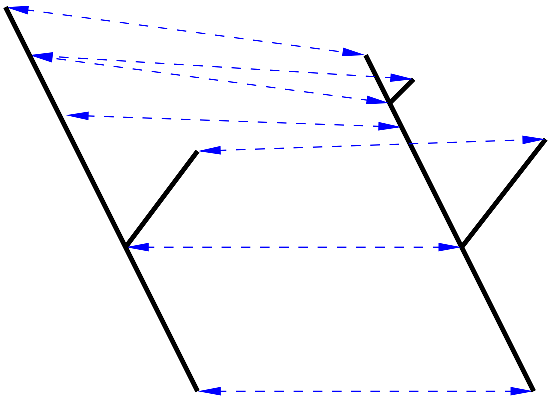
A correspondence R is a subset of $X \times X'$ such that for every $x \in X$, there exists $x' \in X'$ with $(x, x') \in R$ and vice versa.



Topological considerations

Suppose that (X, d) and (X', d') are compact metric spaces.

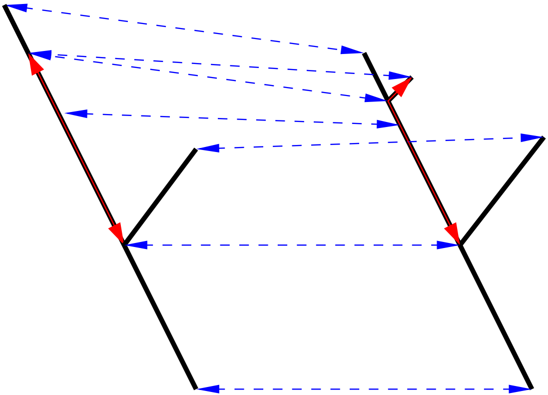
A correspondence R is a subset of $X \times X'$ such that for every $x \in X$, there exists $x' \in X'$ with $(x, x') \in R$ and vice versa.



Topological considerations

The distortion of R is

$$\mathrm{dis}(R) = \sup\{|d(x,y) - d'(x',y')| : (x,x'),(y,y') \in R\}.$$



Topological considerations

Suppose that μ is a Borel probability measure on (X, d) and that μ' is a Borel probability measure on (X', d').

A measure ν on $X \times X'$ is a coupling of μ and μ' if $\nu(\cdot, X') = \mu(\cdot)$ and $\nu(X, \cdot) = \mu'(\cdot)$.

Topological considerations

Suppose that μ is a Borel probability measure on (X, d) and that μ' is a Borel probability measure on (X', d').

A measure ν on $X \times X'$ is a coupling of μ and μ' if $\nu(\cdot, X') = \mu(\cdot)$ and $\nu(X, \cdot) = \mu'(\cdot)$.

Idea: find a correspondence and a coupling such that the correspondence has small distortion and the coupling "lines up" well with the correspondence i.e. if $(V, V') \sim \nu$ then $\mathbb{P}((V, V') \in R) = \nu(R)$ is close to 1.

Topological considerations

The Gromov-Hausdorff-Prokhorov distance between (X, d, μ) and (X', d', μ') is defined to be

$$\mathrm{d}_{\mathrm{GHP}}((X, d, \mu), (X', d', \mu')) = \frac{1}{2} \inf_{R, \nu} \max\{\mathrm{dis}(R), \nu(R^c)\}.$$

Topological considerations

The Gromov-Hausdorff-Prokhorov distance between (X, d, μ) and (X', d', μ') is defined to be

$$\mathrm{d_{GHP}}((X, d, \mu), (X', d', \mu')) = \frac{1}{2} \inf_{R,\nu} \max\{\mathrm{dis}(R), \nu(R^c)\}.$$

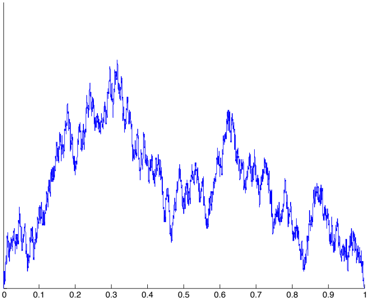
Theorem. $(\mathbb{M}, \mathrm{d_{GHP}})$ is a complete separable metric space.

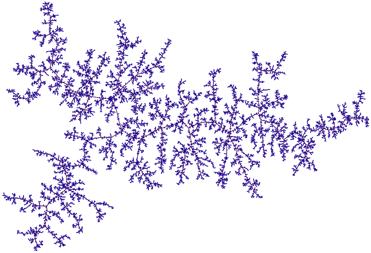
[S. Evans, J. Pitman and A. Winter, **Rayleigh processes, real trees, and root growth with re-grafting**, *Probability Theory and Related Fields* **134** (2006) pp.81-126.]

[R. Abraham, J.-F. Delmas and P. Hoscheit, **A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces**, *Electronic Journal of Probability* **18** (2013), no. 14.]

The Brownian CRT

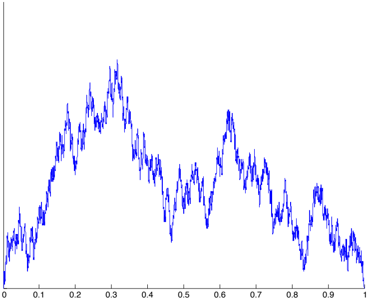
Definition. The Brownian continuum random tree is $(\mathcal{T}_{2e}, d_{2e}, \mu_{2e})$, where e is a standard Brownian excursion.

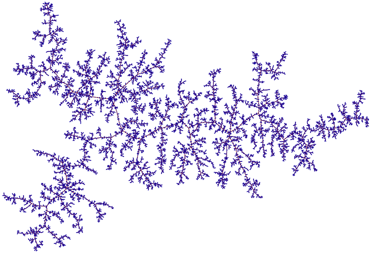




[Pictures by Igor Kortchemski]

A planar ordering



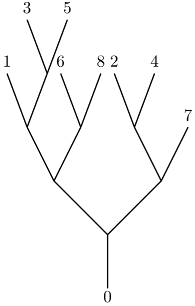


Observe that the excursion comes with slightly more information than the the tree: if $s < t$ and $\pi_{2e}(s)$ and $\pi_{2e}(t)$ are leaves, it is natural to think of $\pi_{2e}(s)$ appearing to the left of $\pi_{2e}(t)$.

Discrete trees as metric spaces

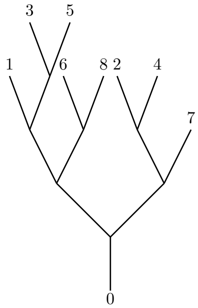
We want to think of $(T_n, n \geq 1)$ as metric spaces.

The vertices of T_n (labelled and unlabelled) come equipped with a natural metric: the graph distance d_n.



We sometimes write aT_n for the metric space (T_n, ad_n) given by the vertices of T_n with the graph distance scaled by a.

Uniform measure



We will endow T_n with μ_n, the measure which puts mass $1/(2n)$ on each of the $2n$ vertices.

Convergence

Theorem. As $n \to \infty$,

$$\left(T_n, \frac{d_n}{\sqrt{2n}}, \mu_n\right) \to (\mathcal{T}_{2e}, d_{2e}, \mu_{2e}) \quad \text{a.s.}$$

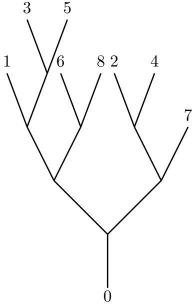
with respect to the Gromov-Hausdorff-Prokhorov topology.

[P. Marchal, **Constructing a sequence of random walks strongly converging to Brownian motion**, *Discrete Mathematics and Theoretical Computer Science*, 2003, pp.181–190.]

[N. Curien & B. Haas, **The stable trees are nested**, *Probability Theory and Related Fields* **157**, 2013, pp.847–883.]

A plane version of our binary trees

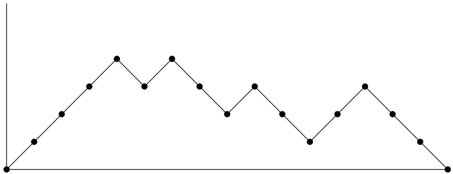
In order to see where the Brownian excursion comes from, it will be helpful for us to now give our binary trees a planar ordering. We achieve this in Rémy's algorithm by simply gluing each new branch on the left or right with equal probability.

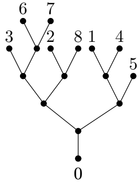


Binary trees and lattice excursions

There is a well-known bijection between planted binary plane trees with n leaves and lattice excursions with $2n$ steps.

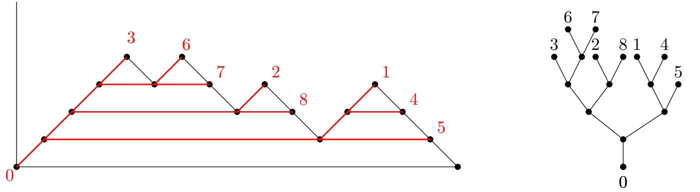
Start every excursion with a $+1$ step. Now travel round the tree from left to right, recording a step whenever you see a vertex for the first time. The step is $+1$ if the vertex is a branch-point and -1 if the vertex is a leaf.





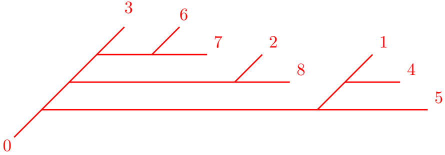
Binary trees and lattice excursions

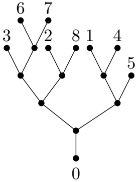
To go back the other way, it's easy to recover the tree:



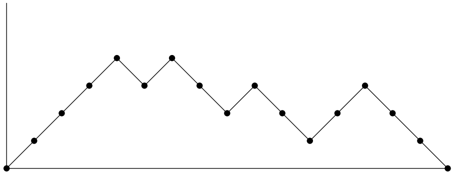
Binary trees and lattice excursions

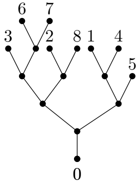
To go back the other way, it's easy to recover the tree:





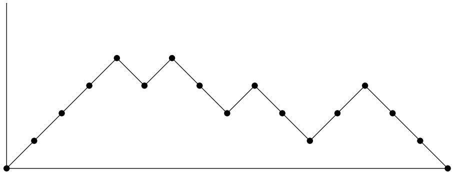
Binary trees and lattice excursions

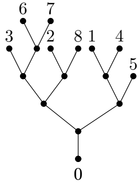




Since our trees are uniform, so are the lattice excursions. In other words, they are excursions of simple random walk away from 0.

Binary trees and lattice excursions





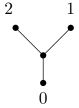
Since our trees are uniform, so are the lattice excursions. In other words, they are excursions of simple random walk away from 0. So (at least in distribution), it's clear that, suitably rescaled, they should converge to a Brownian excursion.

Binary trees and lattice excursions

Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.

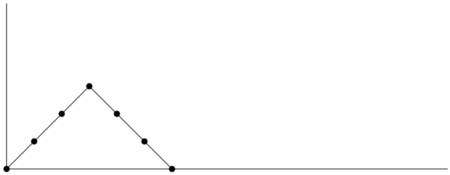
Binary trees and lattice excursions

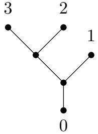
Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.



Binary trees and lattice excursions

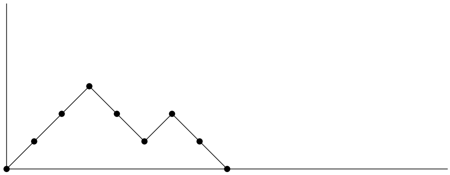
Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.

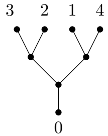




Binary trees and lattice excursions

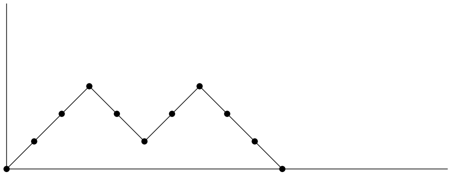
Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.

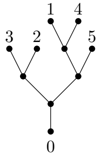




Binary trees and lattice excursions

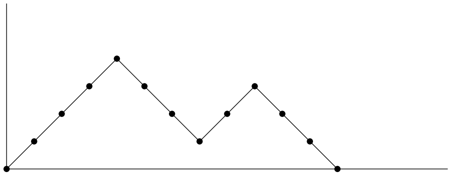
Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.

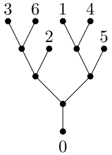




Binary trees and lattice excursions

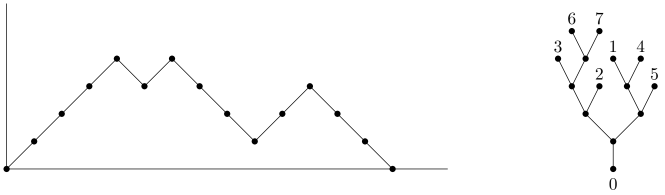
Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.





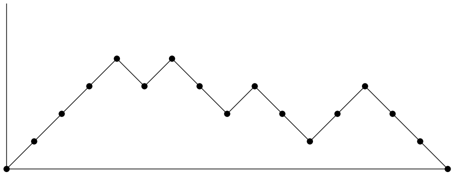
Binary trees and lattice excursions

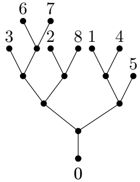
Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.



Binary trees and lattice excursions

Rémy's algorithm then corresponds to a sequence of simple operations on such lattice excursions.





Binary trees and lattice excursions

Let $(E_n)_{n\geq 1}$ be the sequence of lattice excursions.

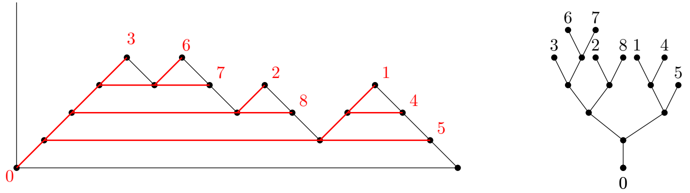
Theorem. (Marchal (2003))
As $n \to \infty$, we have

$$\frac{1}{\sqrt{2n}}(E_n(\lfloor 2nt \rfloor), 0 \leq t \leq 1) \to (e(t), 0 \leq t \leq 1)$$

uniformly on $[0, 1]$, almost surely.

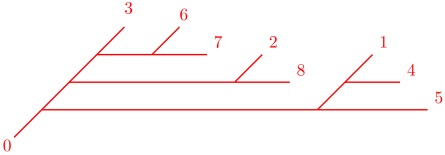
Convergence of the trees

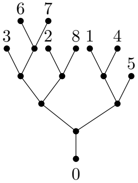
This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.



Convergence of the trees

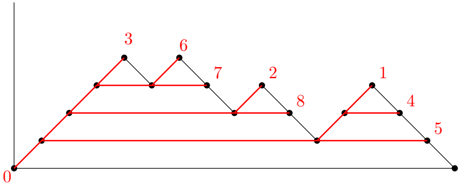
This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.

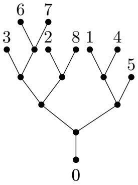




Convergence of the trees

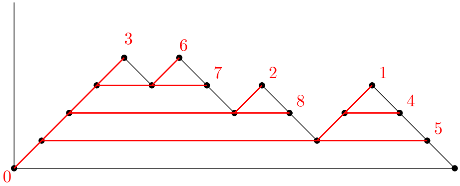
This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.

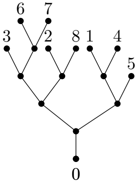




Convergence of the trees

This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.



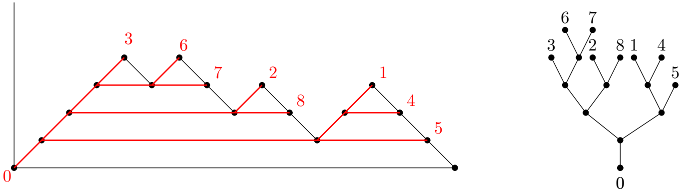


Write $H_n(k)$ for the distance from the root to the vertex visited at time k. Then

$$H_n(k) = \left| \left\{ 0 \leq i \leq k-1 : E_n(i) = \min_{i \leq j \leq k} E_n(k) \right\} \right| .$$

Convergence of the trees

This is not quite enough to conclude that the trees converge in the GHP sense. The embedding of the tree in the excursion distorts distances.



Write $H_n(k)$ for the distance from the root to the vertex visited at time k. Then

$$H_n(k) = \left| \left\{ 0 \leq i \leq k-1 : E_n(i) = \min_{i \leq j \leq k} E_n(k) \right\} \right|.$$

It turns out that $H_n(k) \approx 2E_n(k)$.

Convergence of the trees

Theorem. As $n \to \infty$,

$$\frac{1}{\sqrt{2n}}(H_n(\lfloor 2nt \rfloor), 0 \leq t \leq 1) \to (2e(t), 0 \leq t \leq 1)$$

uniformly on $[0, 1]$, almost surely.

[J.-F. Marckert & A. Mokkadem, **The depth first processes of Galton-Watson trees converge to the same Brownian excursion**, *Annals of Probability*, **31**(3), pp.1655–1678, 2003.]

Convergence of the trees

Let's call the vertices be $v_0, v_1, \ldots, v_{2n-1}$ in the order we visit them, where v_0 is the root.

Convergence of the trees

Let's call the vertices be $v_0, v_1, \ldots, v_{2n-1}$ in the order we visit them, where v_0 is the root.

By definition,

$$d_n(v_0, v_k) = H_n(k).$$

Convergence of the trees

Let's call the vertices be $v_0, v_1, \ldots, v_{2n-1}$ in the order we visit them, where v_0 is the root.

By definition,

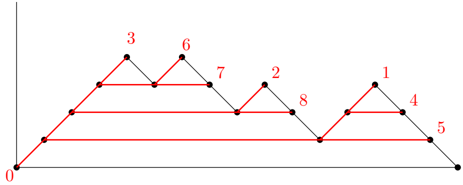
$$d_n(v_0, v_k) = H_n(k).$$

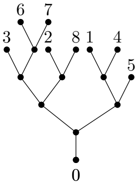
More generally, for $0 \leq i < j \leq 2n-1$, write $v_i \wedge v_j$ for the most recent common ancestor of v_i and v_j in the tree. Then

$$d_n(v_i, v_j) = d_n(v_0, v_i) + d_n(v_0, v_j) - 2d_n(v_0, v_i \wedge v_j).$$

Convergence of the trees

$$d_n(v_0, v_i \wedge v_j) = \begin{cases} \min_{i \le k \le j} H_n(k) - 1 & \text{if } v_i \text{ not an ancestor of } v_j \\ \min_{i \le k \le j} H_n(k) = H_n(i) & \text{if } v_i \text{ an ancestor of } v_j. \end{cases}$$





So

$$\left| d_n(v_0, v_i \wedge v_j) - \min_{i \le k \le j} H_n(k) \right| \le 1.$$

A correspondence

Define a correspondence R_n between $\{v_0, v_1, \ldots, v_{2n-1}\}$ and $[0,1]$ by declaring $(v_i, s) \in R_n$ if $i = \lfloor 2ns \rfloor$.

Endow $[0,1]$ with the pseudo-metric d_{2e}. We will bound $\mathrm{dis}(R_n)$.

A correspondence

Define a correspondence R_n between $\{v_0, v_1, \ldots, v_{2n-1}\}$ and $[0,1]$ by declaring $(v_i, s) \in R_n$ if $i = \lfloor 2ns \rfloor$.

Endow $[0,1]$ with the pseudo-metric d_{2e}. We will bound $\mathrm{dis}(R_n)$.

Suppose that $(v_i, s), (v_j, t) \in R_n$ with $s \leq t$. Then

$$
\begin{aligned}
&|d_n(v_i, v_j) - d_{2e}(s,t)| \\
&\leq \Bigg| \frac{1}{\sqrt{2n}} \left(H_n(\lfloor 2ns \rfloor) + H_n(\lfloor 2nt \rfloor) - 2 \min_{s \leq u \leq t} H_n(\lfloor 2nu \rfloor) \right) \\
&\qquad\qquad - \left(2e(s) + 2e(t) - 4 \min_{s \leq u \leq t} e(u) \right) \Bigg| + \frac{2}{\sqrt{2n}}.
\end{aligned}
$$

A correspondence

Define a correspondence R_n between $\{v_0, v_1, \ldots, v_{2n-1}\}$ and $[0,1]$ by declaring $(v_i, s) \in R_n$ if $i = \lfloor 2ns \rfloor$.

Endow $[0,1]$ with the pseudo-metric d_{2e}. We will bound $\mathrm{dis}(R_n)$.

Suppose that $(v_i, s), (v_j, t) \in R_n$ with $s \leq t$. Then

$$\begin{aligned}
&|d_n(v_i, v_j) - d_{2e}(s,t)| \\
&\leq \left| \frac{1}{\sqrt{2n}} \left(H_n(\lfloor 2ns \rfloor) + H_n(\lfloor 2nt \rfloor) - 2 \min_{s \leq u \leq t} H_n(\lfloor 2nu \rfloor) \right) \right. \\
&\qquad \left. - \left(2e(s) + 2e(t) - 4 \min_{s \leq u \leq t} e(u) \right) \right| + \frac{2}{\sqrt{2n}}.
\end{aligned}$$

The right-hand side converges to 0 uniformly in $s, t \in [0,1]$. So

$$\mathrm{dis}(R_n) \to 0 \quad \text{a.s.}$$

A coupling

Recall that μ_n is the measure which puts mass $1/(2n)$ on each of the vertices $v_0, v_1, \ldots, v_{2n-1}$. Then we may couple μ_n and μ_{2e} by taking $U \sim \mathrm{U}[0,1]$ and taking ν to be the law of the pair

$$(v_{\lfloor 2nU \rfloor}, \pi_{2e}(U)).$$

This is precisely the natural coupling ν_n induced by the correspondence R_n, and so $\nu_n(R_n^c) = 0$.

GHP convergence

But then

$$d_{\mathrm{GHP}}\left(\left(T_n, \frac{d_n}{\sqrt{2n}}, \mu_n\right), (\mathcal{T}_{2e}, d_{2e}, \mu_{2e})\right) \leq \frac{1}{2}\max\left\{\mathrm{dis}(R_n), \nu_n(R_n^c)\right\} \to 0,$$

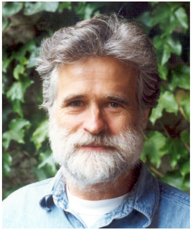
almost surely as $n \to \infty$. $\square$

3. PROPERTIES OF THE BROWNIAN CRT

Key references:

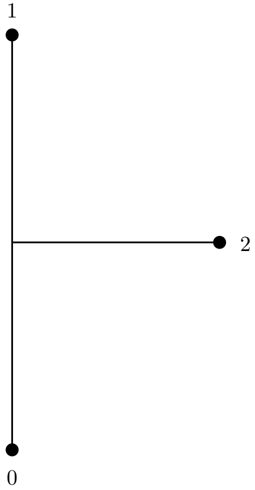
David Aldous, **The continuum random tree III**,
Annals of Probability **21** (1993) pp.248-289.

Jim Pitman, **Combinatorial stochastic processes**,
Lecture notes in mathematics **1875**, Springer-Verlag, Berlin (2006).

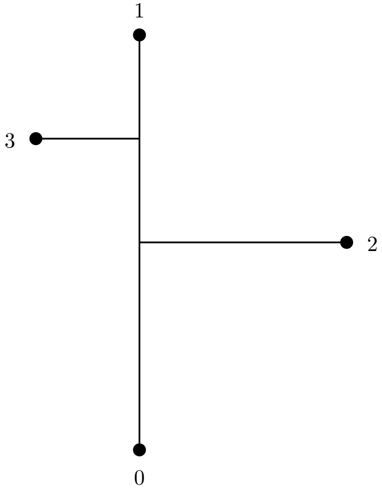


Recap: Rémy's algorithm

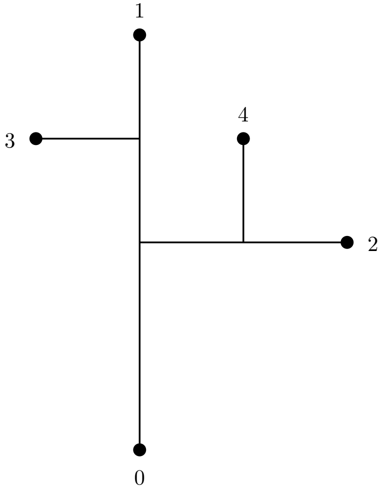
Recap: Rémy's algorithm



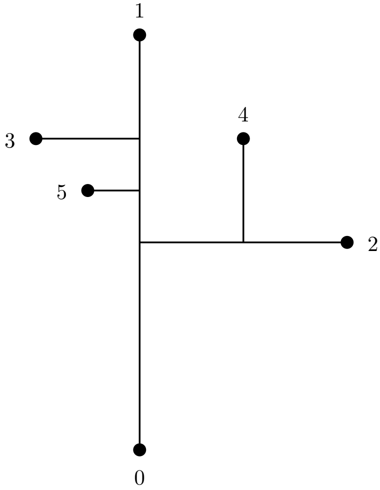
Recap: Rémy's algorithm



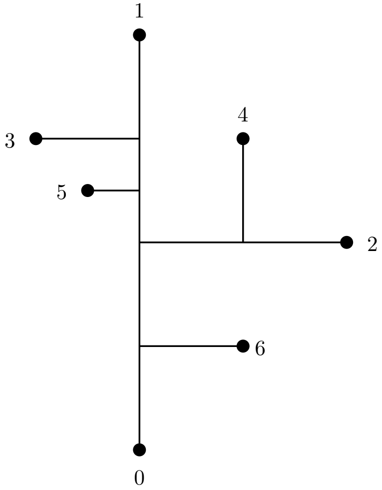
Recap: Rémy's algorithm



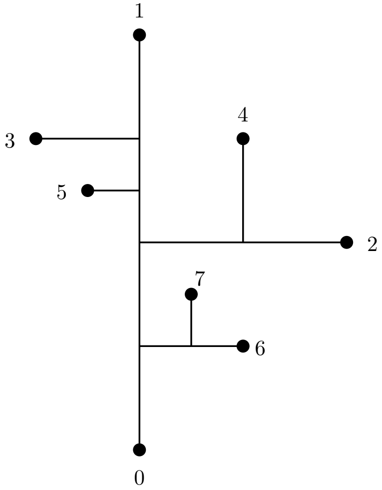
Recap: Rémy's algorithm



Recap: Rémy's algorithm

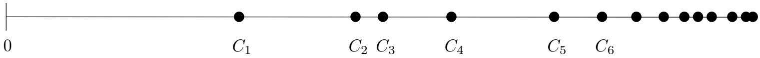


Recap: Rémy's algorithm



Recap: line-breaking construction

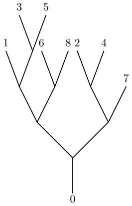
Take an inhomogeneous Poisson process on $\mathbb{R}_+$ of intensity t at t.

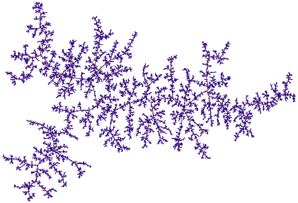


- Consider the line-segments $[0, C_1)$, $[C_1, C_2), \ldots$.
- Start from $[0, C_1)$ and proceed inductively.
- For $i \geq 2$, attach $[C_{i-1}, C_i)$ at a random point chosen uniformly over the existing tree.

Recap: convergence theorem

Recall that T_n is a uniform binary leaf-labelled tree and that $\mathcal{T}_{2e}$ is the Brownian CRT.





Theorem. As $n \to \infty$,

$$\left(T_n, \frac{d_n}{\sqrt{2n}}, \mu_n \right) \to (\mathcal{T}_{2e}, d_{2e}, \mu_{2e}) \quad \text{a.s.}$$

with respect to the Gromov-Hausdorff-Prokhorov topology.

Uniform measure on the leaves

The same is true if we take the measure to be the uniform measure just on the leaves, $\tilde{\mu}_n$.

Theorem. (Curien & Haas (2013))
As $n \to \infty$,

$$\left(T_n, \frac{d_n}{\sqrt{2n}}, \tilde{\mu}_n\right) \to (\mathcal{T}_{2e}, d_{2e}, \mu_{2e}) \quad \text{a.s.}$$

with respect to the Gromov-Hausdorff-Prokhorov topology.

Uniform measure on the leaves

The same is true if we take the measure to be the uniform measure just on the leaves, $\tilde{\mu}_n$.

Theorem. (Curien & Haas (2013))
As $n \to \infty$,

$$\left(T_n, \frac{d_n}{\sqrt{2n}}, \tilde{\mu}_n \right) \to (\mathcal{T}_{2e}, d_{2e}, \mu_{2e}) \quad \text{a.s.}$$

with respect to the Gromov-Hausdorff-Prokhorov topology.

Let $\mathcal{L}(\mathcal{T}_{2e})$ be the set of leaves of $\mathcal{T}_{2e}$. Then $\mu_{2e}(\mathcal{L}(\mathcal{T}_{2e})) = 1$.

The root

Imagine permuting the labels $0, 1, \ldots, n$ in the binary tree T_n. It's straightforward to see that this does not change its law. In particular, the root 0 acts just like a uniformly chosen leaf. So the same must also be true for $\mathcal{T}_{2e}$.

What is a continuum random tree?

A continuum tree is a triple $(\mathcal{T}, d, \mu)$ where $(\mathcal{T}, d)$ is an $\mathbb{R}$-tree with leaves $\mathcal{L}(\mathcal{T})$ and μ is a Borel probability measure on $\mathcal{T}$ which is such that

- μ is non-atomic
- $\mu(\mathcal{L}(\mathcal{T})) = 1$
- for every $v \in \mathcal{T}$ of degree $k \geq 2$, let $\mathcal{T}_1, \ldots, \mathcal{T}_k$ be the connected components of $\mathcal{T} \setminus \{v\}$. Then $\mu(\mathcal{T}_i) > 0$ for all $1 \leq i \leq k$.

What is a continuum random tree?

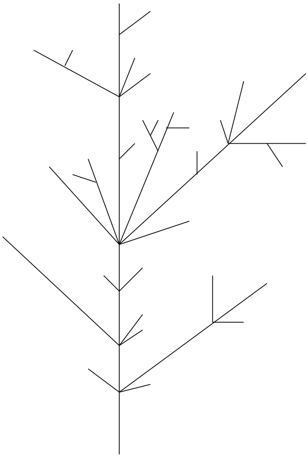
A continuum tree is a triple $(\mathcal{T}, d, \mu)$ where $(\mathcal{T}, d)$ is an $\mathbb{R}$-tree with leaves $\mathcal{L}(\mathcal{T})$ and μ is a Borel probability measure on $\mathcal{T}$ which is such that

- μ is non-atomic
- $\mu(\mathcal{L}(\mathcal{T})) = 1$
- for every $v \in \mathcal{T}$ of degree $k \geq 2$, let $\mathcal{T}_1, \ldots, \mathcal{T}_k$ be the connected components of $\mathcal{T} \setminus \{v\}$. Then $\mu(\mathcal{T}_i) > 0$ for all $1 \leq i \leq k$.

A continuum random tree (CRT) is a random variable taking values in the set of (equivalence classes of) continuum trees.

Characterising a CRT via sampling

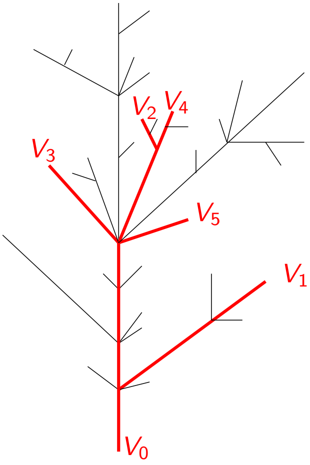
Take a CRT $(\mathcal{T}, d, \mu)$ and suppose that $V_0, V_1, \ldots$ are i.i.d. samples from the measure μ. (Note: these are a.s. leaves.) For $k \geq 1$, let $\mathcal{R}_k$ be the subtree of $\mathcal{T}$ spanned by $V_0, V_1, \ldots, V_k$.



Characterising a CRT via sampling

Take a CRT $(\mathcal{T}, d, \mu)$ and suppose that $V_0, V_1, \ldots$ are i.i.d. samples from the measure μ. (Note: these are a.s. leaves.) For $k \geq 1$, let $\mathcal{R}_k$ be the subtree of $\mathcal{T}$ spanned by $V_0, V_1, \ldots, V_k$.

$\mathcal{R}_5$



Characterising a CRT via sampling

For every $k \geq 1$, $\mathcal{R}_k$ can be regarded as a discrete tree, rooted at V_0, with edge-lengths and labelled leaves, and so its distribution is specified by its tree-shape, a rooted unordered tree with k labelled leaves, and its edge-lengths. The reduced trees are clearly consistent, in that $\mathcal{R}_k$ is a subtree of $\mathcal{R}_{k+1}$.

Characterising a CRT via sampling

For every $k \geq 1$, $\mathcal{R}_k$ can be regarded as a discrete tree, rooted at V_0, with edge-lengths and labelled leaves, and so its distribution is specified by its tree-shape, a rooted unordered tree with k labelled leaves, and its edge-lengths. The reduced trees are clearly consistent, in that $\mathcal{R}_k$ is a subtree of $\mathcal{R}_{k+1}$.

Theorem. (Aldous (1993))
The law of a continuum random tree $(\mathcal{T}, d, \mu)$ is specified by its random finite-dimensional distributions, that is the laws of $(\mathcal{R}_k, k \geq 1)$.

[D. Aldous, **The continuum random tree III**, *Annals of Probability* **21**(1), 1993, pp.248–289.]

Characterising a CRT via sampling

Moreover, if we let

$$\hat{\mu}_k = \frac{1}{k} \sum_{i=1}^{k} \delta_{V_i}$$

then

$$(\mathcal{R}_k, d|_{\mathcal{R}_k}, \hat{\mu}_k) \to (\mathcal{T}, d, \mu)$$

almost surely in d_{GHP}, as $k \to \infty$.

The random finite-dimensional distributions of the Brownian CRT

Our earlier urn results translate into facts about the Brownian CRT.

The random finite-dimensional distributions of the Brownian CRT

Our earlier urn results translate into facts about the Brownian CRT.

Recall our planted binary leaf-labelled tree T_n. As n gets large, the leaves $1, 2, \ldots, k$ behave like i.i.d. samples from $\tilde{\mu}_n$. But our urn arguments gave us the limiting distribution of the rescaled tree spanned by 0 and $1, 2, \ldots, k$.

The random finite-dimensional distributions of the Brownian CRT

Theorem. (Aldous (1993))

- $\mathcal{R}_k$ is a uniform random planted binary leaf-labelled tree, with edge-lengths distributed as

$$\sqrt{\Gamma_k} \times \mathrm{Dir}(\underbrace{1, 1, \ldots, 1}_{2k-1}),$$

  where $\Gamma_k \sim \mathrm{Gamma}(k, 1/2)$, independent of the Dirichlet vector.
- $(\mathcal{R}_k)_{k \geq 1}$ evolves according to the line-breaking construction.

[See Le Gall (2005) for a direct proof from the Brownian excursion.]

The random finite-dimensional distributions of the Brownian CRT

Note that since

$$(\mathcal{R}_k, d_{2e}|_{\mathcal{R}_k}, \hat{\mu}_k) \to (\mathcal{T}_{2e}, d_{2e}, \mu_{2e})$$

as $k \to \infty$, it follows that the Brownian CRT is binary.

Another way to see this is to observe that the local minima of a Brownian excursion are unique almost surely.

Recursive self-similarity

Consider picking three independent points V_1, V_2, V_3 from $\mathcal{T}_{2e}$ according to μ_{2e}. There is a unique branch-point between these three points, and it splits the tree into three subtrees, $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3$.

Recursive self-similarity

Consider picking three independent points V_1, V_2, V_3 from $\mathcal{T}_{2e}$ according to μ_{2e}. There is a unique branch-point between these three points, and it splits the tree into three subtrees, $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3$.

Write d_1, d_2, d_3 and μ_1, μ_2, μ_3 for the restrictions of d and μ to each of these subtrees respectively. Let

$$\Delta_1 = \mu(\mathcal{T}_1), \Delta_2 = \mu(\mathcal{T}_2), \Delta_3 = \mu(\mathcal{T}_3).$$

Recursive self-similarity

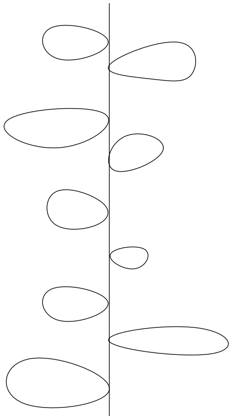
Theorem. (Aldous (1997))

- We have $(\Delta_1, \Delta_2, \Delta_3) \sim \mathrm{Dir}(1/2, 1/2, 1/2)$.
- The rescaled subtrees $(\mathcal{T}_1, d_1/\sqrt{\Delta_1}, \mu_1/\Delta_1)$, $(\mathcal{T}_2, d_2/\sqrt{\Delta_2}, \mu_2/\Delta_2)$, $(\mathcal{T}_3, d_3/\sqrt{\Delta_3}, \mu_3/\Delta_3)$ are i.i.d. Brownian CRTs, independent of $(\Delta_1, \Delta_2, \Delta_3)$.
- V_i and the original branch-point are independent samples from μ_i/Δ_i in subtree $i = 1, 2, 3$.

[D. Aldous, **Recursive self-similarity for random trees, random triangulations and Brownian excursion**, *Annals of Probability* **22**, 1997, pp.812–854.]

The spine decomposition

Now look more closely at the tree from the perspective of the path between the root and a single uniform point (the spine).



What are the masses hanging off the spine?
Where are they located?

The spine decomposition

Theorem. (Haas, Pitman & Winkel (2009))
The spinal mass partition is distributed as Poisson-Dirichlet $\mathrm{PD}(1/2, 1/2)$ and the trees corresponding to the different blocks are attached at i.i.d. uniform points along the spine. These little subtrees are randomly rescaled i.i.d. Brownian CRT's.

[B. Haas, J. Pitman & M. Winkel, **Spinal partitions and invariance under re-rooting of continuum random trees**, *Annals of Probability* **37**(4), 2009, pp.1381–1411.]

The spine decomposition

Theorem. (Haas, Pitman & Winkel (2009))
The spinal mass partition is distributed as Poisson-Dirichlet PD$(1/2, 1/2)$ and the trees corresponding to the different blocks are attached at i.i.d. uniform points along the spine. These little subtrees are randomly rescaled i.i.d. Brownian CRT's.

Fix $0 \leq \alpha < 1$ and $\theta > -\alpha$. Let $\beta_1, \beta_2, \ldots$ be independent random variables such that $\beta_i \sim \text{Beta}(1-\alpha, \theta + i\alpha)$. Let $\tilde{P}_i = \beta_i \prod_{j=1}^{i-1}(1-\beta_j)$, $i \geq 1$, and let $P_1 \geq P_2 \geq \ldots \geq 0$ be the $\tilde{P}_i$ in decreasing order. Then $(P_i)_{i \geq 1} \sim \text{PD}(\alpha, \theta)$.

[B. Haas, J. Pitman & M. Winkel, **Spinal partitions and invariance under re-rooting of continuum random trees**, *Annals of Probability* **37**(4), 2009, pp.1381–1411.]

The Chinese restaurant process

Generate an exchangeable random partition of $\mathbb{N}$ as follows.

- The first customer arrives and sits at a table.
- Suppose that after n customers have arrived, there are n_i of them sitting at table i, for $1 \leq i \leq k$.
- Customer $n+1$ arrives and sits at table i with probability $\frac{n_i - \alpha}{n+\theta}$, for $1 \leq i \leq k$, or starts a new table with probability $\frac{\theta + k\alpha}{n+\theta}$.

The Chinese restaurant process

Generate an exchangeable random partition of $\mathbb{N}$ as follows.

- The first customer arrives and sits at a table.
- Suppose that after n customers have arrived, there are n_i of them sitting at table i, for $1 \leq i \leq k$.
- Customer $n+1$ arrives and sits at table i with probability $\frac{n_i - \alpha}{n+\theta}$, for $1 \leq i \leq k$, or starts a new table with probability $\frac{\theta + k\alpha}{n+\theta}$.

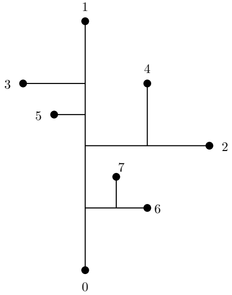
Let K_n be the number of tables occupied by the first n customers, and $\Pi_1^{(n)}, \ldots, \Pi_{K_n}^{(n)}$ the sets of customers sitting at the different tables. Then as $n \to \infty$,

$$\frac{1}{n}\left(|\Pi_1^{(n)}|, \ldots, |\Pi_{K_n}^{(n)}|\right)^{\downarrow} \to (P_i)_{i \geq 1} \quad \text{a.s.}$$

where $(P_i)_{i \geq 1} \sim \mathrm{PD}(\alpha, \theta)$.

The spine decomposition

Claim: the spinal mass partition is distributed as $\mathrm{PD}(1/2, 1/2)$ and the trees corresponding to the different blocks are attached at i.i.d. uniform points along the spine.



Every time we add a new vertex in Rémy's algorithm, we either add it to a subtree which is already hanging from the path between 0 and 1, or we create a new such subtree.

The spine decomposition

Suppose we are at step n and that the current length of the path from 0 to 1 is $k+1$, so that there are k subtrees hanging off. Suppose these subtrees contain $n_1, \dots, n_k$ leaves, listing from top to bottom.

- We add our new vertex to the ith existing subtree (containing n_i vertices) with probability $\frac{2n_i-1}{2n-1}$.
- We add a new subtree with probability proportional to $\frac{k+1}{2n-1}$.

The spine decomposition

Suppose we are at step n and that the current length of the path from 0 to 1 is $k+1$, so that there are k subtrees hanging off. Suppose these subtrees contain $n_1, \ldots, n_k$ leaves, listing from top to bottom.

- We add our new vertex to the ith existing subtree (containing n_i vertices) with probability $\frac{2n_i - 1}{2n-1}$.
- We add a new subtree with probability proportional to $\frac{k+1}{2n-1}$.

These are precisely the probabilities in the Chinese restaurant process with parameters $\alpha = 1/2, \theta = 1/2$.

The spine decomposition

Suppose we are at step n and that the current length of the path from 0 to 1 is $k+1$, so that there are k subtrees hanging off. Suppose these subtrees contain $n_1, \ldots, n_k$ leaves, listing from top to bottom.

- We add our new vertex to the ith existing subtree (containing n_i vertices) with probability $\frac{2n_i-1}{2n-1}$.
- We add a new subtree with probability proportional to $\frac{k+1}{2n-1}$.

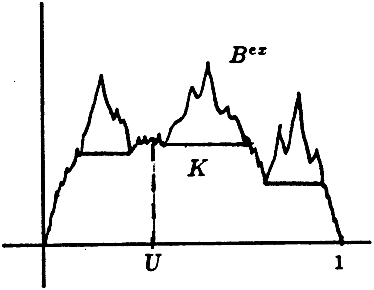
These are precisely the probabilities in the Chinese restaurant process with parameters $\alpha = 1/2, \theta = 1/2$. So the labels in the subtrees behave exactly as tables in a $(1/2, 1/2)$-Chinese restaurant process (on $\{2, 3, 4, \ldots\}$) and it follows that their ordered sizes, divided by n, converge almost surely to a PD$(1/2, 1/2)$ vector. Since each new subtree gets attached at a uniform position along the path, the same is also true in the limit.

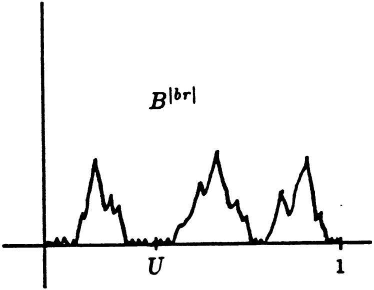
An alternative viewpoint: via a path transformation

Theorem. (Bertoin & Pitman (1994))
Let B^{ex} be a standard Brownian excursion and let $U \sim \mathrm{U}[0,1]$. Let

$$K_t = \begin{cases} \min_{t \le s \le U} B^{ex}_s & \text{for } 0 \le t \le U \\ \min_{U \le s \le t} B^{ex}_s & \text{for } U \le t \le 1. \end{cases}$$

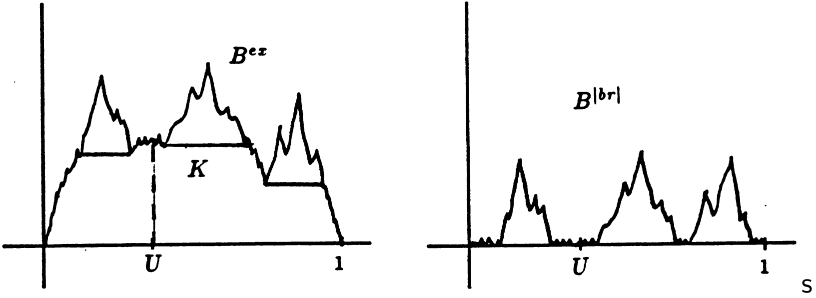
Then $B^{|br|} := B^{ex} - K$ is the modulus of a standard Brownian bridge.





[J. Bertoin & J. Pitman, **Path transformations connecting Brownian bridge, excursion and meander**, *Bull. Sci. Math.* **2**, 1994, pp.147–166.]

An alternative viewpoint: via a path transformation



The excursions of the bridge encode the little subtrees hanging off the spine. The ordered lengths of these excursions are Poisson-Dirichlet$(1/2, 1/2)$ distributed. The local time at 0 of the Brownian bridge is Rayleigh distributed, which represents the length of the path between the root and uniform leaf. U sits exactly halfway through the local time.

A random fractal

The self-similarity of the Brownian CRT tells us, in particular, that it is a random fractal. What is its dimension?

A random fractal

The self-similarity of the Brownian CRT tells us, in particular, that it is a random fractal. What is its dimension?

The Minkowski (or box-counting) dimension is defined to be $\lim_{\epsilon \downarrow 0} \frac{\log N(\mathcal{T}_{2e}, \epsilon)}{\log(1/\epsilon)}$ (if the limit exists) where $N(\mathcal{T}_{2e}, \epsilon)$ is the number of balls of radius ϵ needed to cover $\mathcal{T}_{2e}$.

A random fractal

The self-similarity of the Brownian CRT tells us, in particular, that it is a random fractal. What is its dimension?

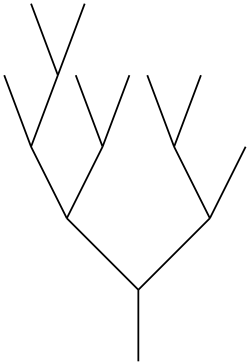
The Minkowski (or box-counting) dimension is defined to be $\lim_{\epsilon \downarrow 0} \frac{\log N(\mathcal{T}_{2e}, \epsilon)}{\log(1/\epsilon)}$ (if the limit exists) where $N(\mathcal{T}_{2e}, \epsilon)$ is the number of balls of radius ϵ needed to cover $\mathcal{T}_{2e}$.

Theorem. (Duquesne & Le Gall (2005))
The Brownian CRT has Minkowski dimension 2, almost surely.

[T. Duquesne & J.-F. Le Gall, **Probabilistic and fractal aspects of Lévy trees**, *Probability Theory and Related Fields* **131**, 2005, pp.553–603.]

Minkowski dimension: heuristic for lower bound

Consider $\mathcal{R}_k$, the tree subtree spanned by the root and k uniform leaves.



For a lower bound, we cover parts of this subtree. Recall that the lengths are

$$\sqrt{\Gamma_k} \times \mathrm{Dir}(\underbrace{1, 1, \ldots, 1}_{2k-1}),$$

where $\Gamma_k \sim \mathrm{Gamma}(k, 1/2)$.

Minkowski dimension: heuristic for lower bound

We have $\mathbb{E}\left[\sqrt{\Gamma_k}\right] = \frac{\sqrt{2}\Gamma(k+1/2)}{\Gamma(k)} \sim \sqrt{2k/e}$.

Minkowski dimension: heuristic for lower bound

We have $\mathbb{E}\left[\sqrt{\Gamma_k}\right] = \frac{\sqrt{2}\Gamma(k+1/2)}{\Gamma(k)} \sim \sqrt{2k/e}$.

For a well-chosen $\delta > 0$, there is exponentially small probability of not having at least δk of the elements of the $\mathrm{Dir}(\underbrace{1, 1, \ldots, 1}_{2k-1})$ vector larger than $1/(4k)$.

Minkowski dimension: heuristic for lower bound

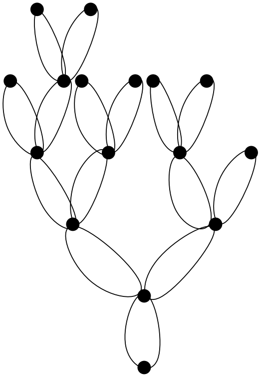
We have $\mathbb{E}\left[\sqrt{\Gamma_k}\right] = \frac{\sqrt{2}\Gamma(k+1/2)}{\Gamma(k)} \sim \sqrt{2k/e}$.

For a well-chosen $\delta > 0$, there is exponentially small probability of not having at least δk of the elements of the $\mathrm{Dir}(\underbrace{1, 1, \ldots, 1}_{2k-1})$ vector larger than $1/(4k)$.

So we can find $\Omega(k)$ disjoint balls of radius $\Theta(1/\sqrt{k})$ which cover a strict subset of $\mathcal{R}_k$.

Minkowski dimension: heuristic for upper bound

Consider again picking k uniform leaves.



We know that the $2k-1$ subtrees have masses $(\Delta_1, \ldots, \Delta_{2k-1}) \sim \mathrm{Dir}(1/2, 1/2, \ldots, 1/2)$. Inside each blob is a rescaled independent Brownian CRT. Let $R_1, \ldots, R_{2k-1}$ be i.i.d. copies of the maximum distance from the root in a Brownian CRT.

Minkowski dimension: heuristic for upper bound

R_1 has the distribution of the maximum of a standard Brownian excursion, which is such that

$$\mathbb{P}(R_1 \geq x) = \sum_{k \geq 1} (-1)^{k+1} e^{-k^2 x^2} \leq e^{-x^2}.$$

So we have covered $\mathcal{T}_{2e}$ with balls of random radius at most

$$\max_{1 \leq i \leq 2k-1} \sqrt{\Delta_i} R_i.$$

We may realise the Dirichlet vector as

$$(\Delta_1, \ldots, \Delta_{2k-1}) = \frac{1}{\sum_{i=1}^{2k-1} \gamma_i} (\gamma_1, \ldots, \gamma_{2k-1}),$$

where $\gamma_1, \ldots, \gamma_{2k-1}$ are i.i.d. Gamma$(1/2, 1)$.

Minkowski dimension: heuristic for upper bound

So

$$\max_{1 \le i \le 2k-1} \sqrt{\Delta_i} R_i = \sqrt{\frac{\max_{1 \le i \le 2k-1} \gamma_i R_i^2}{\sum_{i=1}^{2k-1} \gamma_i}}.$$

Minkowski dimension: heuristic for upper bound

So

$$\max_{1 \le i \le 2k-1} \sqrt{\Delta_i} R_i = \sqrt{\frac{\max_{1 \le i \le 2k-1} \gamma_i R_i^2}{\sum_{i=1}^{2k-1} \gamma_i}}.$$

Now

$$\mathbb{P}\left(\gamma_1 R_1^2 > x\right) \le \mathbb{E}\left[\exp(-x/\gamma_1)\right] = \exp(-2\sqrt{x}).$$

Minkowski dimension: heuristic for upper bound

So

$$\max_{1\le i\le 2k-1} \sqrt{\Delta_i} R_i = \sqrt{\frac{\max_{1\le i\le 2k-1} \gamma_i R_i^2}{\sum_{i=1}^{2k-1} \gamma_i}}.$$

Now

$$\mathbb{P}\left(\gamma_1 R_1^2 > x\right) \le \mathbb{E}\left[\exp(-x/\gamma_1)\right] = \exp(-2\sqrt{x}).$$

So $\max_{1\le i\le 2k-1} \gamma_i R_i^2 \sim (\log k)^2$. Since $\sum_{i=1}^{2k-1} \gamma_i \sim k$, we get

$$\max_{1\le i\le 2k-1} \sqrt{\Delta_i} R_i \sim \frac{\log k}{\sqrt{k}}.$$

Minkowski dimension: heuristic for upper bound

So

$$\max_{1 \le i \le 2k-1} \sqrt{\Delta_i} R_i = \sqrt{\frac{\max_{1 \le i \le 2k-1} \gamma_i R_i^2}{\sum_{i=1}^{2k-1} \gamma_i}}.$$

Now

$$\mathbb{P}\left(\gamma_1 R_1^2 > x\right) \le \mathbb{E}\left[\exp(-x/\gamma_1)\right] = \exp(-2\sqrt{x}).$$

So $\max_{1 \le i \le 2k-1} \gamma_i R_i^2 \sim (\log k)^2$. Since $\sum_{i=1}^{2k-1} \gamma_i \sim k$, we get

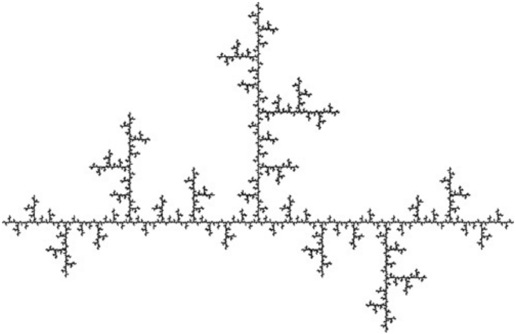
$$\max_{1 \le i \le 2k-1} \sqrt{\Delta_i} R_i \sim \frac{\log k}{\sqrt{k}}.$$

So we need $O(k)$ balls of radius approximately $k^{-1/2} \log k$ to cover $\mathcal{T}_{2e}$.

[L. Addario-Berry, N. Broutin, C. Goldschmidt & G. Miermont, **The scaling limit of the minimum-spanning tree of the complete graph**, *Annals of Probability* **45**(5), 2017, pp.3075–3144.]

A different perspective

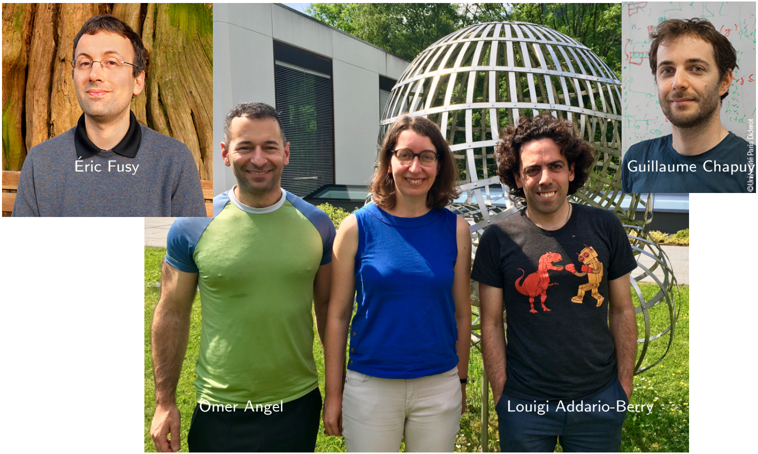
Croydon & Hambly (2008) showed that we can also view $(\mathcal{T}_{2e}, d_{2e})$ as this familiar deterministic fractal endowed with a random metric.



[D. Croydon & B. Hambly, **Self-similarity and spectral asymptotics for the continuum random tree**, *Stochastic Processes and their Applications* **11**, 2008, pp.730–754.]

4. VORONOI CELLS IN THE BROWNIAN CRT

Joint work with Louigi Addario-Berry (McGill), Omer Angel (UBC), Guillaume Chapuy (Paris 7) and Éric Fusy (École polytechnique)



[**Voronoi tessellations in the CRT and continuum random maps of finite excess**, *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2018)*, pp.933-946.]

Voronoi cells in a metric space

Let (M, d) be a metric space.

Fix $k \geq 1$ and let $S = \{x_i : 1 \leq i \leq k\}$ be a collection of points in M, the centres.

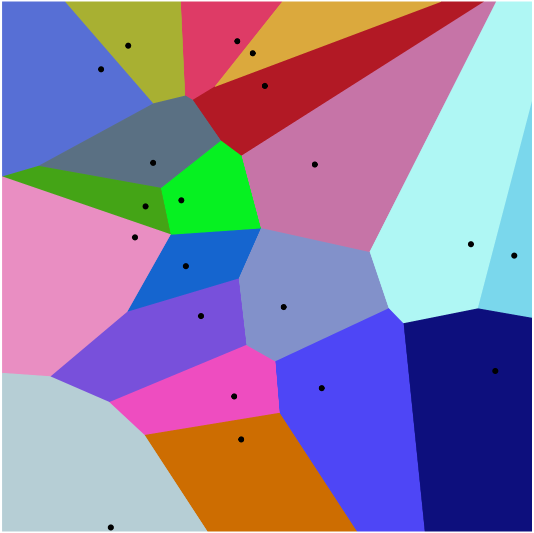
For $1 \leq i \leq k$, the Voronoi cells are

$$V_i = \{y \in M : d(y, S) = d(y, x_i)\}.$$

(Note that the Voronoi cells are not necessarily disjoint.)

Standard example: Voronoi cells in $\mathbb{R}^2$

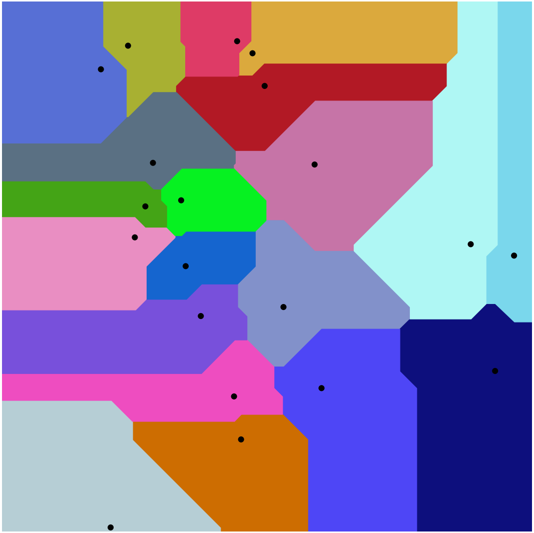
Euclidean distance



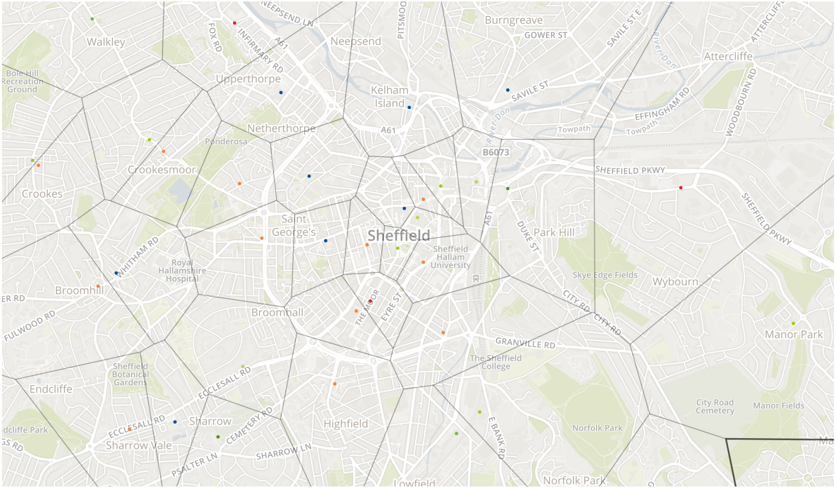
Picture by Balu Ertl (CC BY-SA 4.0, `https://commons.wikimedia.org/w/index.php?curid=38534275`)

Standard example: Voronoi cells in $\mathbb{R}^2$

Manhattan distance

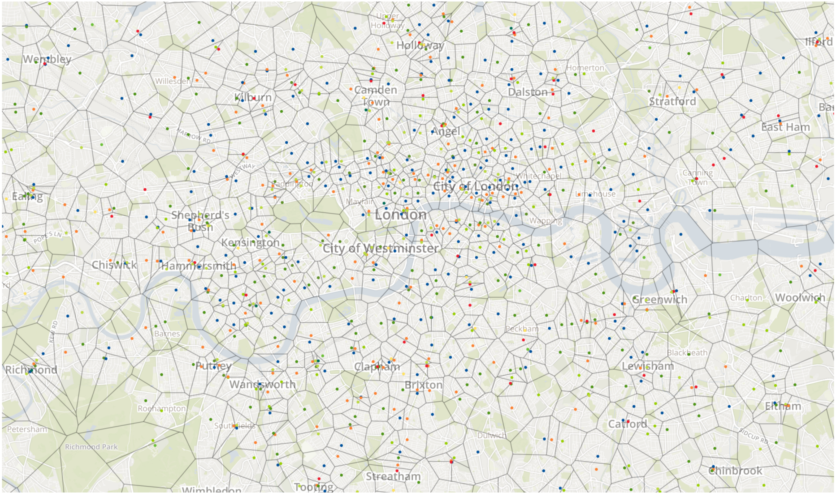


Voronoi supermarkets



See `https://chriszetter.com/voronoi-map/examples/uk-supermarkets/`

Voronoi supermarkets



See `https://chriszetter.com/voronoi-map/examples/uk-supermarkets/`

General set-up: Voronoi cells in a metric space

Let (M, d) be a metric space endowed with a Borel probability measure μ.

Fix $k \geq 1$ and let $S = \{x_i : 1 \leq i \leq k\}$ be a collection of points in M, the centres. Typically these will be random and i.i.d. samples from μ.

For $1 \leq i \leq k$, the Voronoi cells are

$$V_i = \{y \in M : d(y, S) = d(y, x_i)\}.$$

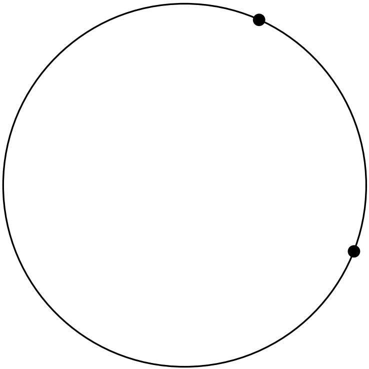
(Note that the Voronoi cells are not necessarily disjoint.)

We will be interested in the "masses" of these cells, as measured by μ, i.e.

$$(\mu(V_1), \mu(V_2), \ldots, \mu(V_k)).$$

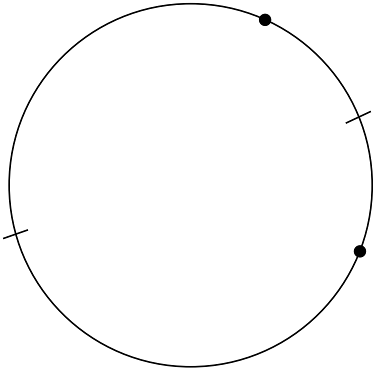
Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Any two points.



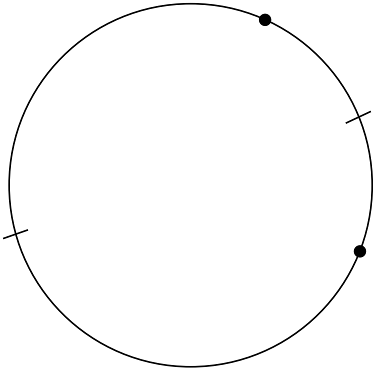
Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Any two points.



Warm-up: circle

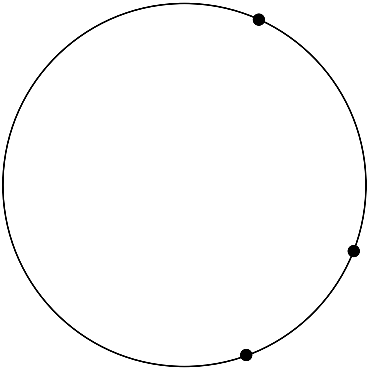
Circle of circumference 1, Euclidean distance, Lebesgue measure.
Any two points.



$(\mu(V_1), \mu(V_2)) = (1/2, 1/2)$.

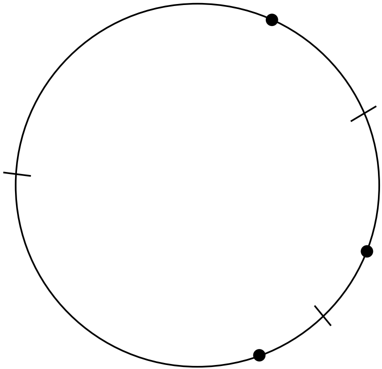
Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Three uniform points.



Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Three uniform points.



Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Three uniform points.

Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure. Three uniform points.

The lengths of these intervals are uniform on the 2-dimensional simplex i.e. have $\mathrm{Dir}(1,1,1)$ distribution.

Warm-up: circle

Circle of circumference 1, Euclidean distance, Lebesgue measure.
Three uniform points.

The lengths of these intervals are uniform on the 2-dimensional simplex i.e. have $\mathrm{Dir}(1,1,1)$ distribution.

We get that the Lebesgue measures of the Voronoi cells are

$$(\mu(V_1), \mu(V_2), \mu(V_3)) = \left(\tfrac{1}{2} U_{(2)}, \tfrac{1}{2}\left(1 - U_{(1)}\right), \tfrac{1}{2}\left(1 - U_{(1)} - U_{(2)}\right)\right)$$

(exchangeable with marginals distributed as $\frac{1}{2}\mathrm{Beta}(2,1)$).

Voronoi cells in the Brownian CRT

Question: what if we take the metric space to be the Brownian CRT?

Voronoi mass-partition in the Brownian CRT

Theorem. (Addario-Berry, Angel, Chapuy, Fusy & G. (2018))
Let $(\mathcal{T}, d, \mu)$ be the Brownian CRT. Fix $k \geq 2$ and let $X_1, X_2, \ldots, X_k$ be i.i.d. samples from μ. Let $V_1, V_2, \ldots, V_k$ be the corresponding Voronoi cells. Then

$$(\mu(V_1), \mu(V_2), \ldots, \mu(V_k)) \sim \mathrm{Dir}(1, 1, \ldots, 1).$$

Voronoi mass-partition in the Brownian CRT

Theorem. (Addario-Berry, Angel, Chapuy, Fusy & G. (2018))
Let $(\mathcal{T}, d, \mu)$ be the Brownian CRT. Fix $k \geq 2$ and let $X_1, X_2, \ldots, X_k$ be i.i.d. samples from μ. Let $V_1, V_2, \ldots, V_k$ be the corresponding Voronoi cells. Then

$$(\mu(V_1), \mu(V_2), \ldots, \mu(V_k)) \sim \mathrm{Dir}(1, 1, \ldots, 1).$$

If you want to chop up the Brownian CRT in a uniform manner, pick uniform points and find their Voronoi cells!

Voronoi mass-partition in the Brownian CRT

Theorem. (Addario-Berry, Angel, Chapuy, Fusy & G. (2018))
Let $(\mathcal{T}, d, \mu)$ be the Brownian CRT. Fix $k \geq 2$ and let $X_1, X_2, \ldots, X_k$ be i.i.d. samples from μ. Let $V_1, V_2, \ldots, V_k$ be the corresponding Voronoi cells. Then

$$(\mu(V_1), \mu(V_2), \ldots, \mu(V_k)) \sim \mathrm{Dir}(1, 1, \ldots, 1).$$

If you want to chop up the Brownian CRT in a uniform manner, pick uniform points and find their Voronoi cells!

(Compare to the $\mathrm{Dir}(1/2, 1/2, 1/2)$ mass-split we get by cutting at the branch-point between three uniform points.)

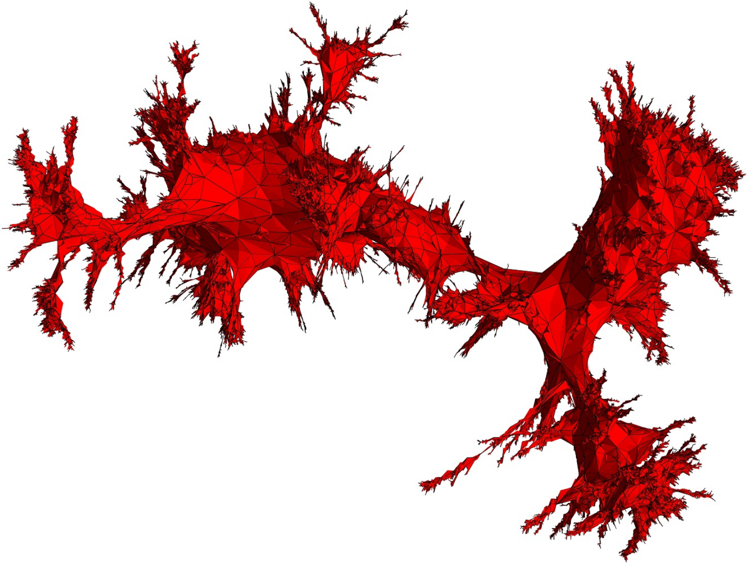
Our original motivation

Conjecture. (Chapuy (2016))
Let $(\mathcal{B}, d, \mu)$ be the Brownian map (or Brownian surface of genus $g \geq 0$). Let $X_1, X_2, \ldots, X_k$ be i.i.d. points sampled from μ and $V_1, V_2, \ldots, V_k$ be the corresponding Voronoi cells. Then

$$(\mu(V_1), \mu(V_2), \ldots, \mu(V_k)) \sim \mathrm{Dir}(1, 1, \ldots).$$

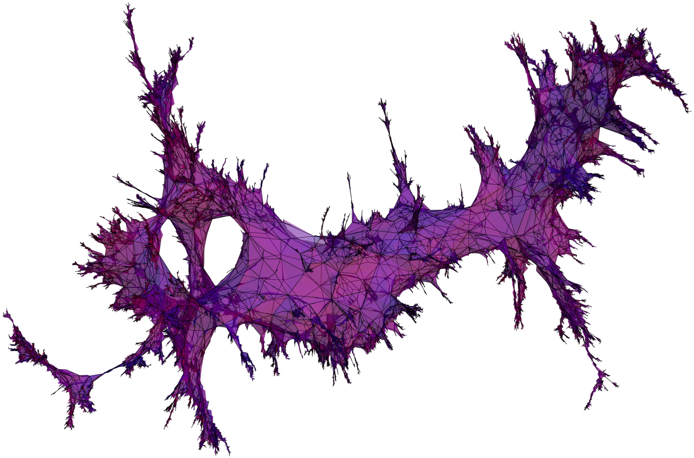
[G. Chapuy, **On tesselations of random maps and the t_g recurrence**, *Séminaire Lotharingien de Combinatoire* **78B**, 2017, no. 79, 12pp.]

The Brownian map (sphere)



[Picture by Jérémie Bettinelli]

The Brownian double torus



[Picture by Jérémie Bettinelli]

Brownian surfaces

Conjecture. (Chapuy (2016))
Let $(\mathcal{B}, d, \mu)$ be the Brownian map (or Brownian surface of genus $g \geq 0$). Let $X_1, X_2, \ldots, X_k$ be i.i.d. points sampled from μ and $V_1, V_2, \ldots, V_k$ be the corresponding Voronoi cells. Then

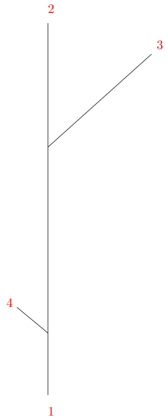
$$(\mu(V_1), \mu(V_2), \ldots, \mu(V_k)) \sim \mathrm{Dir}(1, 1, \ldots).$$

Proved for $g = 0, k = 2$ by Emmanuel Guitter (but proof does not generalise).

[E. Guitter, **A universal law for Voronoi cell volumes in infinitely large maps**, *Journal of Statistical Mechanics: Theory and Experiment*, 2018.]

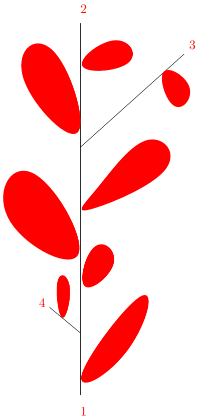
Open problem. Which properties of a random metric space give rise to uniform Voronoi mass-partitions?

Recap: reconstructing the Brownian CRT



Suppose we start from the subtree spanned by $X_1, \ldots, X_k$.

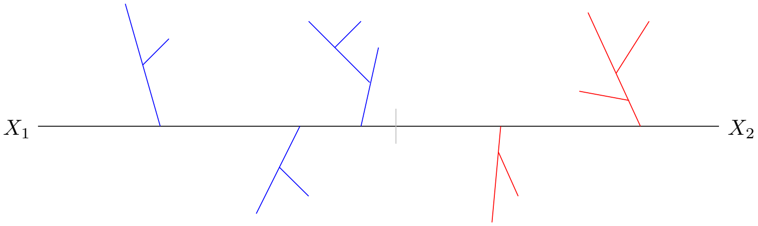
Recap: reconstructing the Brownian CRT



Suppose we start from the subtree spanned by $X_1, \ldots, X_k$. In order to get back to the whole tree, we need to take i.i.d. copies of the Brownian CRT, randomly rescaled by an exchangeable vector with sum 1, and glued onto the subtree at i.i.d. uniform positions.

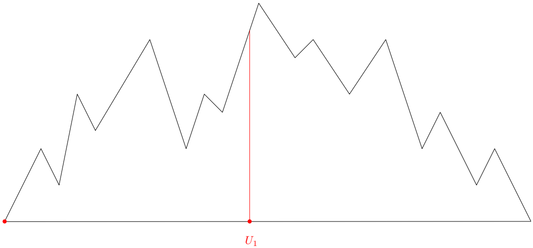
Base case: $k = 2$

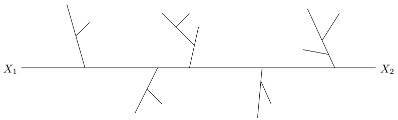
The proof goes via induction, with the base case being $k = 2$.



We wish to find the masses of the blue and red parts.

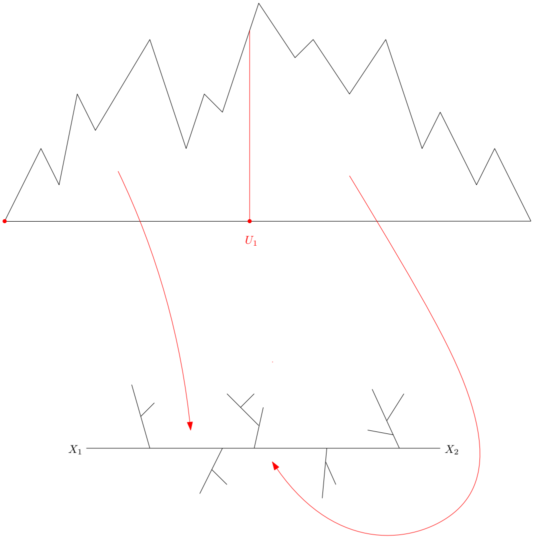
$k = 2$: an observation





Call the masses above and below the backbone the contour cells.

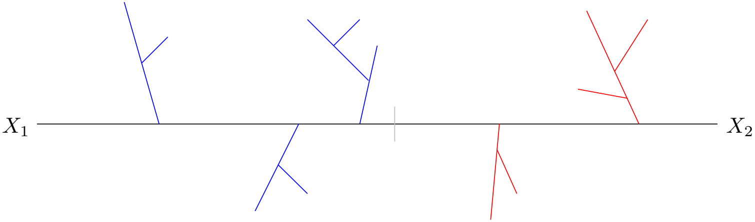
$k = 2$: an observation



Call the masses above and below the backbone the contour cells. These are equal to U_1 and $1 - U_1$, with $U_1 \sim \mathrm{U}[0, 1]$. The little trees attached to the backbone have exchangeable masses.

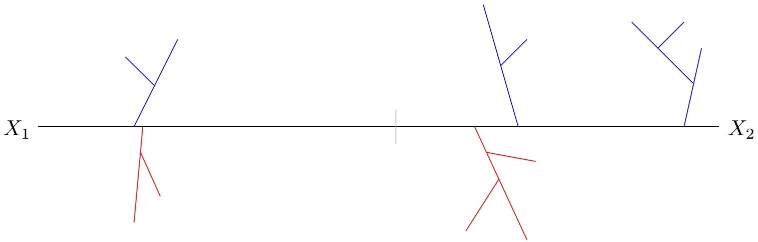
$k = 2$: a bijection

We may convert the Voronoi cells into the contour cells of a different tree:



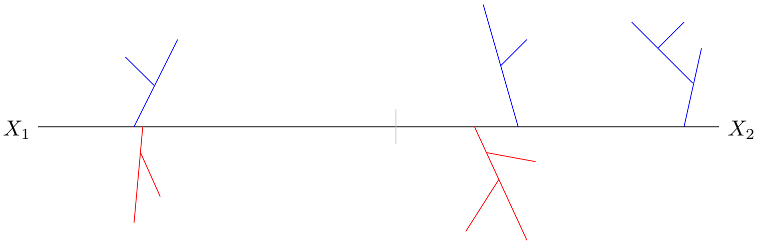
$k = 2$: a bijection

We may convert the Voronoi cells into the contour cells of a different tree:



$k = 2$: a bijection

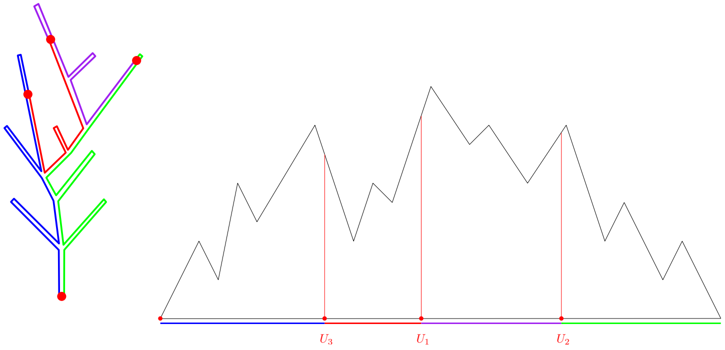
We may convert the Voronoi cells into the contour cells of a different tree:



Since the subtree masses are exchangeable, the new tree is again a Brownian CRT. But the contour cells in a Brownian CRT have $(U, 1-U)$ mass split, so the same must be true for the Voronoi cells. (This may be read off from results of Lévy (1939) or the $B^{ex} \leftrightarrow B^{|br|}$ path-transformation of Bertoin and Pitman (1994).)

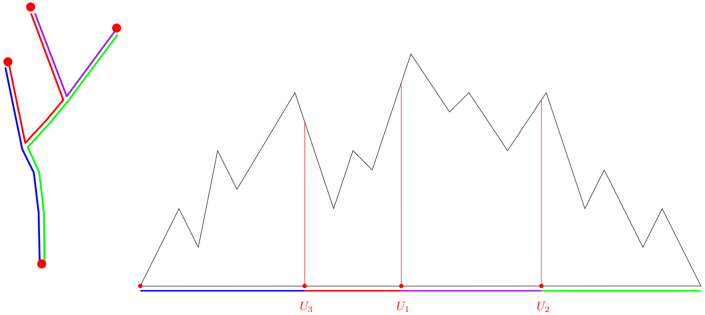
Reductions for $k \geq 3$: contour cells

Consider the subtree spanned by our uniform points.



Reductions for $k \geq 3$: contour cells

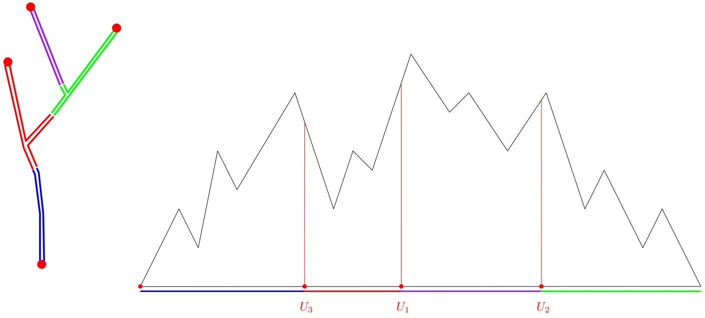
Consider the subtree spanned by our uniform points.



We will show that the lengths of the coloured intervals (the contour intervals) have the same joint law as the lengths of the Voronoi cells in the subtree.

Reductions for $k \geq 3$: contour cells

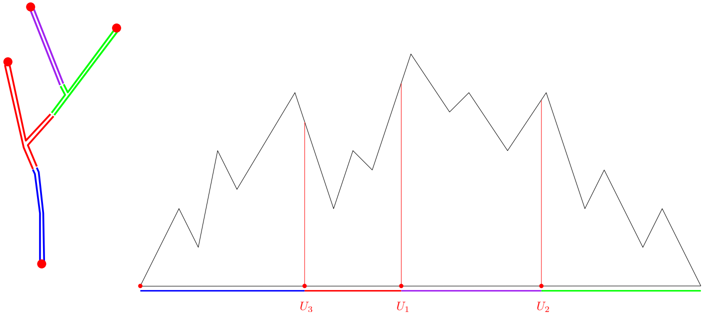
Consider the subtree spanned by our uniform points.



We will show that the lengths of the coloured intervals (the contour intervals) have the same joint law as the lengths of the Voronoi cells in the subtree.

Reductions for $k \geq 3$: contour cells

Consider the subtree spanned by our uniform points.



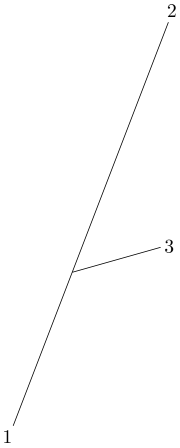
We will show that the lengths of the coloured intervals (the contour intervals) have the same joint law as the lengths of the Voronoi cells in the subtree. Since the mass attached to the contour intervals yields a uniform split of unity, the same must then be true for the Voronoi cells.

Reductions for $k \geq 3$: scaling

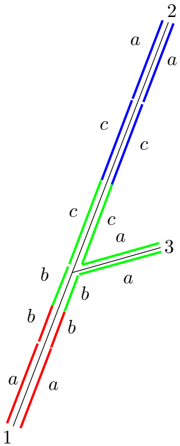
Since we're now only interested in showing that two vectors of lengths have the same distribution, it makes no difference if we rescale the whole tree.

So by the properties of the Brownian CRT, we may take the edge-lengths in the subtree spanned by our uniform points to be i.i.d. Exp(1).

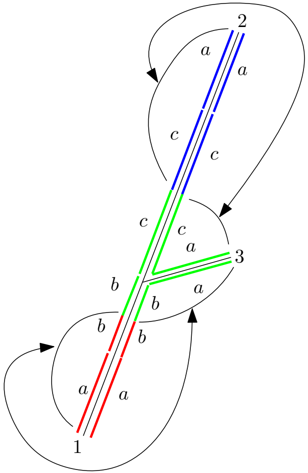
$k = 3$: contour lengths $\leftrightarrow$ Voronoi lengths



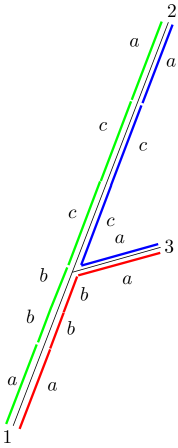
$k = 3$: contour lengths $\leftrightarrow$ Voronoi lengths



$k = 3$: contour lengths $\leftrightarrow$ Voronoi lengths



$k = 3$: contour lengths $\leftrightarrow$ Voronoi lengths

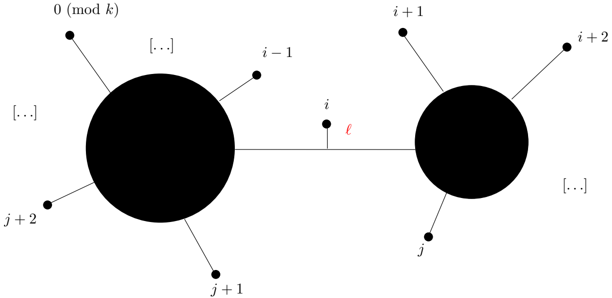


So again we have a bijection between the contour lengths and the Voronoi lengths.

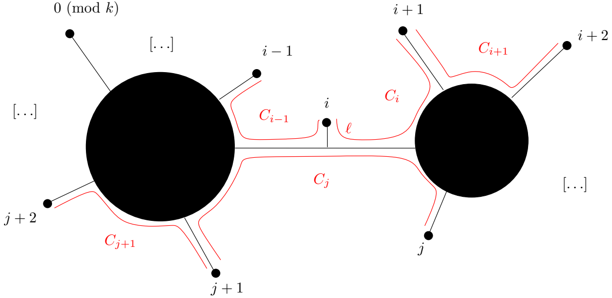
General $k \geq 3$: by induction

Suppose the result is true for all smaller k. We start with a uniform binary plane leaf-labelled tree with i.i.d. Exp(1) edge-lengths.

Start from the shortest branch incident to a leaf. This branch is uniform among all those incident to leaves. Call its leaf i and its length ℓ. Call the "opposite leaf" j.



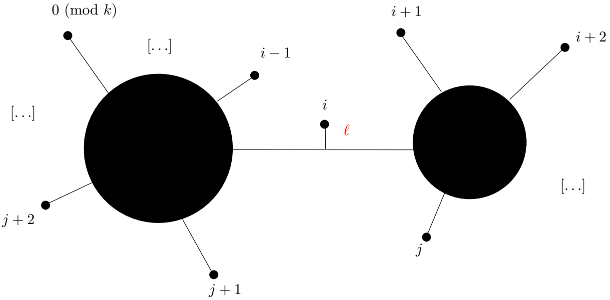
General $k \geq 3$: by induction



Voronoi lengths: $(L_0, L_1, \ldots, L_{k-1})$
Contour lengths: $(C_0, C_1, \ldots, C_{k-1})$.

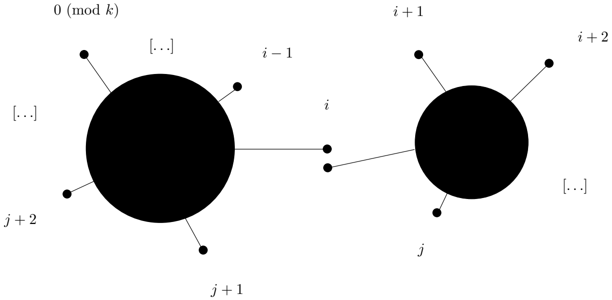
General $k \geq 3$: by induction

Now burn in from every leaf to remove length ℓ:



General $k \geq 3$: by induction

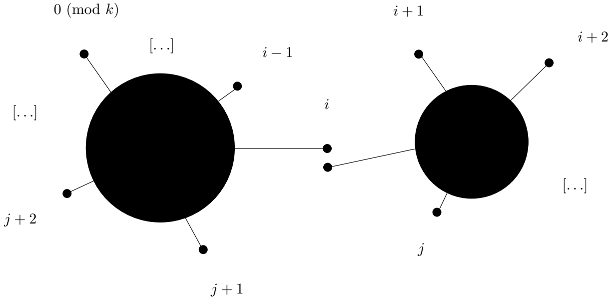
Now burn in from every leaf to remove length ℓ:



By the memoryless property of the exponential, and the uniformity of the shortest leaf, we split into two uniform binary leaf-labelled trees with i.i.d. exponential edge-lengths, each with $< k$ leaves.

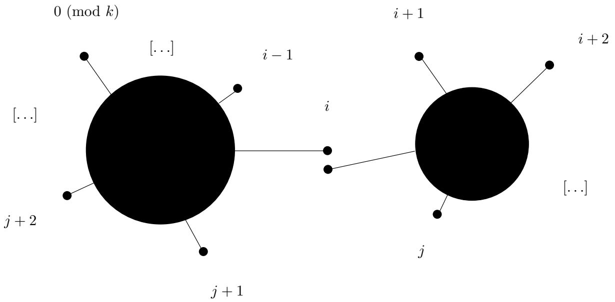
So, by the induction hypothesis, the Voronoi and contour lengths have the same laws in each of the subtrees.

General $k \geq 3$: by induction



- For each leaf other than j, we can get back the original contour length C_r from r to $r+1$ by simply adding 2ℓ to the contours in the smaller problems.
- For the contour from j to $j+1$, we must add two contours together and add 2ℓ.

General $k \geq 3$: by induction



- For the Voronoi cells, add 2ℓ to the new lengths of the cells to get L_r, $r \neq i$.
- For the cell of i, add two Voronoi cells from the smaller trees, plus 2ℓ.

By induction, the vectors of lengths therefore have the same law.

5. THE BROWNIAN CRT AS A UNIQUE FIXED POINT

Joint work with Marie Albenque (École polytechnique)

[**The Brownian continuum random tree as the unique solution to a fixed point equation**, *Electronic Communications in Probability* **20**, 2015, paper no. 61, pp.1-14.]

Recursive distributional equations

By a recursive distributional equation (RDE) for a random variable X taking values in some Polish space $\mathcal{S}$, we mean an equation of the form

$$X \stackrel{d}{=} f((\xi_i, X_i)_{i\geq 1})$$

where $X_1, X_2, \ldots$ are i.i.d. copies of X, independent of the family of random variables $(\xi_i)_{i\geq 1}$ and f is a suitable $\mathcal{S}$-valued function.

Recursive distributional equations

Example. Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. real-valued r.v.'s with finite variance such that

$$X_1 \stackrel{d}{=} \frac{X_1 + X_2 + \ldots + X_n}{\sqrt{n}}.$$

Then $X_1 \sim N(0, \sigma^2)$ for some $\sigma^2 > 0$.

Recursive distributional equations

Example. Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. real-valued r.v.'s with finite variance such that

$$X_1 \stackrel{d}{=} \frac{X_1 + X_2 + \ldots + X_n}{\sqrt{n}}.$$

Then $X_1 \sim N(0, \sigma^2)$ for some $\sigma^2 > 0$.

The centred normal distributions are the fixed points of this RDE.

Recap: recursive self-similarity

Consider picking three independent points U_1, U_2, U_3 from the Brownian CRT $\mathcal{T}$ according to μ. There is a unique branch-point between these three points, and it splits the tree into three subtrees, $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3$.

Recap: recursive self-similarity

Consider picking three independent points U_1, U_2, U_3 from the Brownian CRT $\mathcal{T}$ according to μ. There is a unique branch-point between these three points, and it splits the tree into three subtrees, $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3$.

Write d_1, d_2, d_3 and μ_1, μ_2, μ_3 for the restrictions of d and μ to each of these subtrees respectively. Let

$$\Delta_1 = \mu(\mathcal{T}_1), \Delta_2 = \mu(\mathcal{T}_2), \Delta_3 = \mu(\mathcal{T}_3).$$

Recap: recursive self-similarity

Consider picking three independent points U_1, U_2, U_3 from the Brownian CRT $\mathcal{T}$ according to μ. There is a unique branch-point between these three points, and it splits the tree into three subtrees, $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3$.

Write d_1, d_2, d_3 and μ_1, μ_2, μ_3 for the restrictions of d and μ to each of these subtrees respectively. Let

$$\Delta_1 = \mu(\mathcal{T}_1), \Delta_2 = \mu(\mathcal{T}_2), \Delta_3 = \mu(\mathcal{T}_3).$$

Theorem. (Aldous (1993))

- We have $(\Delta_1, \Delta_2, \Delta_3) \sim \mathrm{Dir}(1/2, 1/2, 1/2)$.
- The rescaled subtrees $(\mathcal{T}_1, d_1/\sqrt{\Delta_1}, \mu_1/\Delta_1)$, $(\mathcal{T}_2, d_2/\sqrt{\Delta_2}, \mu_2/\Delta_2)$, $(\mathcal{T}_3, d_3/\sqrt{\Delta_3}, \mu_3/\Delta_3)$ are i.i.d. Brownian CRTs, independent of $(\Delta_1, \Delta_2, \Delta_3)$.
- U_i and the original branch-point are independent samples from μ_i/Δ_i in subtree $i = 1, 2, 3$.

An operator on (laws of) CRTs

Let $\mathcal{M}$ be the set of probability measures on continuum trees. Define an operator $\mathfrak{F} : \mathcal{M} \to \mathcal{M}$ as follows: for $M \in \mathcal{M}$,

- Sample independent trees (T_1, d_1, μ_1), (T_2, d_2, μ_3), (T_3, d_3, μ_3) having distribution M;
- For $1 \leq i \leq 3$, sample U_i according to μ_i;
- Independently sample $(\Delta_1, \Delta_2, \Delta_3) \sim \mathrm{Dir}(1/2, 1/2, 1/2)$;
- Rescale to obtain $(T_1, \Delta_1^{1/2} d_1, \Delta_1 \mu_1)$, $(T_2, \Delta_2^{1/2} d_2, \Delta_2 \mu_2)$, $(T_3, \Delta_3^{1/2} d_3, \Delta_3 \mu_3)$.
- Identify the vertices U_1, U_2, U_3 in order to obtain a single larger tree T with a marked branch-point B; the metrics and measures naturally induce a metric d and a measure μ on T.
- Forget the branch-point in order to obtain (T, d, μ).

$\mathfrak{F}(M)$ is the distribution of (T, d, μ).

The Brownian CRT as a unique fixed point

The previous theorem told us that the law of the Brownian CRT is a fixed point of $\mathfrak{F}$.

The Brownian CRT as a unique fixed point

The previous theorem told us that the law of the Brownian CRT is a fixed point of $\mathfrak{F}$.

Theorem. (Albenque & G. (2015))

- [Unique fixed point] Suppose that M is a law on continuum trees which is a fixed point of $\mathfrak{F}$. Then there exists $\alpha > 0$ such that if $(T, d, \mu) \sim M$ then $(T, \alpha d, \mu)$ is a Brownian CRT.

The Brownian CRT as a unique fixed point

The previous theorem told us that the law of the Brownian CRT is a fixed point of $\mathfrak{F}$.

Theorem. (Albenque & G. (2015))

- [Unique fixed point] Suppose that M is a law on continuum trees which is a fixed point of $\mathfrak{F}$. Then there exists $\alpha > 0$ such that if $(T, d, \mu) \sim M$ then $(T, \alpha d, \mu)$ is a Brownian CRT.
- [Attractive] Suppose that M is a law on continuum trees such that if $(T, d, \mu) \sim M$, given (T, d, μ), V_1, V_2 are sampled independently from μ, then $\mathbb{E}\left[d(V_1, V_2)\right]$ exists and is equal to $\pi/2$. Let $M_n = \mathfrak{F}^n M$. Then M_n converges weakly to the law of the Brownian CRT in the sense of the Gromov-Prokhorov topology.

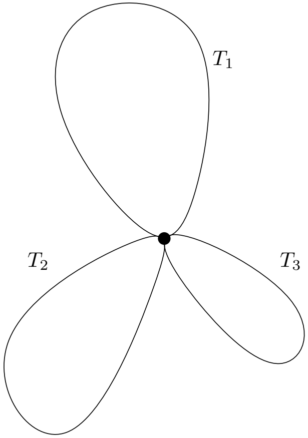
Two-point distances

Lemma. Suppose that $M \in \mathcal{M}$ is a fixed point of $\mathfrak{F}$. Let $(T, d, \mu) \sim M$ and, conditionally on (T, d, μ), let $V_1, V_2 \stackrel{\text{i.i.d.}}{\sim} \mu$. Then there exists a constant $\alpha > 0$ such that $\alpha d(V_1, V_2) \sim \text{Rayleigh}$.

Two-point distances

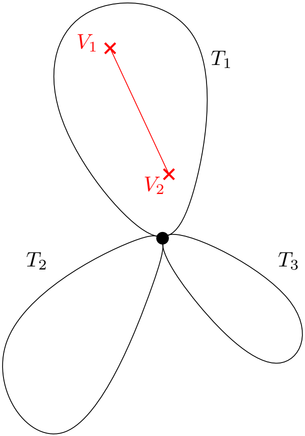
$(T, d, \mu) \sim M$ and $V_1, V_2 \overset{\text{i.i.d.}}{\sim} \mu$. Let $L = d(V_1, V_2)$.

Since M is a fixed point of $\mathfrak{F}$, we can think of (T, d, μ) as having been built out of three scaled independent copies, (T_1, d_1, μ_1), (T_2, d_2, μ_2) and (T_3, d_3, μ_3):



Two-point distances

There are two possibilities (plus symmetries) for what happens to the two points V_1 and V_2:



Two-point distances

Let P_1, P_2, P_3 denote the numbers of points falling in each of T_1, T_2, T_3 respectively.

Conditionally on $(\Delta_1, \Delta_2, \Delta_3)$, we have

$$(P_1, P_2, P_3) \sim \text{Multinomial}(2; \Delta_1, \Delta_2, \Delta_3).$$

Two-point distances

Let P_1, P_2, P_3 denote the numbers of points falling in each of T_1, T_2, T_3 respectively.

Conditionally on $(\Delta_1, \Delta_2, \Delta_3)$, we have

$$(P_1, P_2, P_3) \sim \text{Multinomial}(2; \Delta_1, \Delta_2, \Delta_3).$$

Then

$$L \stackrel{d}{=} \sqrt{\Delta_1} L_1 \mathbb{1}_{\{P_1 > 0\}} + \sqrt{\Delta_2} L_2 \mathbb{1}_{\{P_2 > 0\}} + \sqrt{\Delta_3} L_3 \mathbb{1}_{\{P_3 > 0\}},$$

where L_1, L_2, L_3 are i.i.d. copies of L, independent of $(\Delta_1, \Delta_2, \Delta_3)$ and (P_1, P_2, P_3).

Two-point distances

Let P_1, P_2, P_3 denote the numbers of points falling in each of T_1, T_2, T_3 respectively.

Conditionally on $(\Delta_1, \Delta_2, \Delta_3)$, we have

$$(P_1, P_2, P_3) \sim \text{Multinomial}(2; \Delta_1, \Delta_2, \Delta_3).$$

Then

$$L \stackrel{d}{=} \sqrt{\Delta_1} L_1 \mathbb{1}_{\{P_1 > 0\}} + \sqrt{\Delta_2} L_2 \mathbb{1}_{\{P_2 > 0\}} + \sqrt{\Delta_3} L_3 \mathbb{1}_{\{P_3 > 0\}},$$

where L_1, L_2, L_3 are i.i.d. copies of L, independent of $(\Delta_1, \Delta_2, \Delta_3)$ and (P_1, P_2, P_3).

This recursive distributional equation is an instance of the so-called smoothing transform.

The smoothing transform

Suppose that

$$X_1 \stackrel{d}{=} \sum_{i=1}^{n} W_i X_i,$$

where $X_1, X_2, \ldots, X_n$ are i.i.d. non-negative r.v.'s, independent of the non-negative r.v.'s $W_1, W_2, \ldots, W_n$ which are such that

$$\mathbb{E}\left[W_i^{\gamma}\right] < \infty$$

for all $1 \leq i \leq n$ and some $\gamma > 1$.

The smoothing transform, $X_1 \stackrel{d}{=} \sum_{i=1}^n W_i X_i$

Let $g(s) = \log\left(\sum_{i=1}^n \mathbb{E}\left[W_i^s \mathbb{1}_{\{W_i > 0\}}\right]\right)$, $s \geq 0$. Write $\mathcal{F}(\mu)$ for the distribution of $\sum_{i=1}^n W_i X_i$ when $X_1, X_2, \ldots, X_n \stackrel{\text{i.i.d.}}{\sim} \mu$.

Theorem. (Durrett & Liggett (1983))

1. Suppose that g has a unique zero $\alpha \in (0, 1]$. If $\alpha = 1$ and $g'(1) < 0$ then the RDE has a unique fixed point μ, up to a deterministic scaling factor. Write μ_m for the fixed point with mean m.
2. Suppose ν is any law on $\mathbb{R}_+$ such that $\int_0^\infty x d\nu(x) = m$. Then

$$\mathcal{F}^k(\nu) \to \mu_m$$

as $k \to \infty$.

[R. Durrett & T. Liggett, **Fixed points of the smoothing transformation**, *Zeitschrift für Warscheinlichkeitstheorie und verwandte Gebiete* **64**, 1983, pp.275–301.]

Two-point distances

For any fixed point of $\mathfrak{F}$, the two-point distances satisfy

$$L \stackrel{d}{=} \sqrt{\Delta_1} L_1 \mathbb{1}_{\{P_1 > 0\}} + \sqrt{\Delta_2} L_2 \mathbb{1}_{\{P_2 > 0\}} + \sqrt{\Delta_3} L_3 \mathbb{1}_{\{P_3 > 0\}}.$$

Two-point distances

For any fixed point of $\mathfrak{F}$, the two-point distances satisfy

$$L \stackrel{d}{=} \sqrt{\Delta_1} L_1 \mathbb{1}_{\{P_1 > 0\}} + \sqrt{\Delta_2} L_2 \mathbb{1}_{\{P_2 > 0\}} + \sqrt{\Delta_3} L_3 \mathbb{1}_{\{P_3 > 0\}}.$$

We get $g(s) = \log\left(\frac{3(s+7)}{(s+3)(s+5)}\right)$, which has its unique zero in $s \geq 0$ at $s = 1$, with $g'(1) = -7/24 < 0$. So the theorem of Durrett and Liggett applies to give a unique distributional solution, up to a deterministic constant.

Two-point distances

For any fixed point of $\mathfrak{F}$, the two-point distances satisfy

$$L \stackrel{d}{=} \sqrt{\Delta_1} L_1 \mathbb{1}_{\{P_1>0\}} + \sqrt{\Delta_2} L_2 \mathbb{1}_{\{P_2>0\}} + \sqrt{\Delta_3} L_3 \mathbb{1}_{\{P_3>0\}}.$$

We get $g(s) = \log\left(\frac{3(s+7)}{(s+3)(s+5)}\right)$, which has its unique zero in $s \geq 0$ at $s = 1$, with $g'(1) = -7/24 < 0$. So the theorem of Durrett and Liggett applies to give a unique distributional solution, up to a deterministic constant.

Moreover, since the Rayleigh distribution must be a solution, and $\mathbb{E}\left[\text{Rayleigh}\right] = \sqrt{\frac{\pi}{2}}$, the value α in the lemma is $\sqrt{\pi/2}/\mathbb{E}\left[L\right]$.

Random finite-dimensional distributions

Theorem. Suppose that $M \in \mathcal{M}$ is a fixed point of $\mathfrak{F}$. Then the random finite-dimensional distributions of M are the same as those of the Brownian CRT, up to a strictly positive scaling factor α.

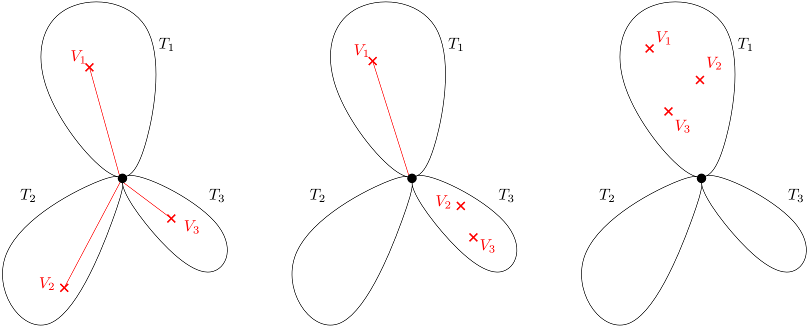
Random finite-dimensional distributions

Theorem. Suppose that $M \in \mathcal{M}$ is a fixed point of $\mathfrak{F}$. Then the random finite-dimensional distributions of M are the same as those of the Brownian CRT, up to a strictly positive scaling factor α.

Since the law of a continuum random tree is uniquely determined by its random finite-dimensional distributions, this will be enough to give the first part of our fixed point theorem.

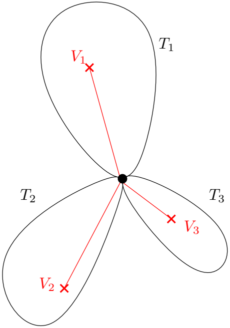
Three-point distances

There are three cases:



Three-point distances

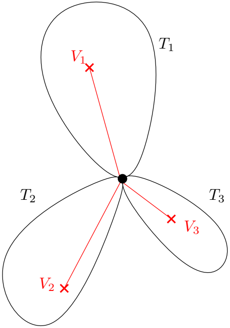
Case 1:



Accounting for symmetries, this event occurs with probability $6\mathbb{E}\left[\Delta_1\Delta_2\Delta_3\right] = 2/35$.

Three-point distances

Case 1:

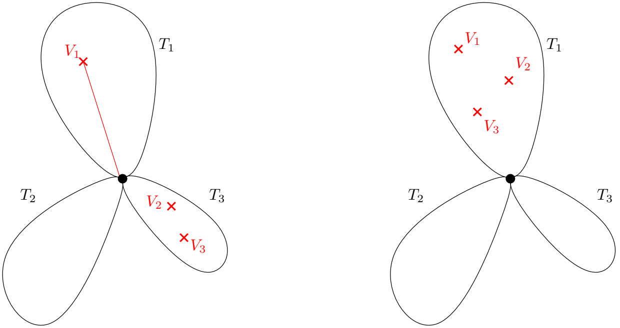


Accounting for symmetries, this event occurs with probability $6\mathbb{E}\left[\Delta_1\Delta_2\Delta_3\right] = 2/35$.

Conditioning on its occurrence yields biased sizes $(\Delta_1^*, \Delta_2^*, \Delta_3^*)$ and we get $(\sqrt{\Delta_1^*}L_1, \sqrt{\Delta_2^*}L_2, \sqrt{\Delta_3^*}L_3)$ for the three distances.

Three-point distances

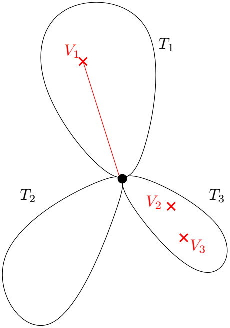
Cases 2 and 3:

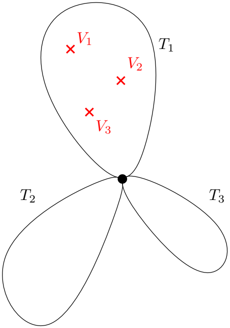


Here, in order to understand the distances, we need to figure out what's happening inside one of the level-1 subtrees. Note that, in the two cases, the problem is really the same: we have three uniform points within one of the subtrees, and want to find the distances between them – it's just that in case 2, one of the three points is the branchpoint at the centre.

Three-point distances

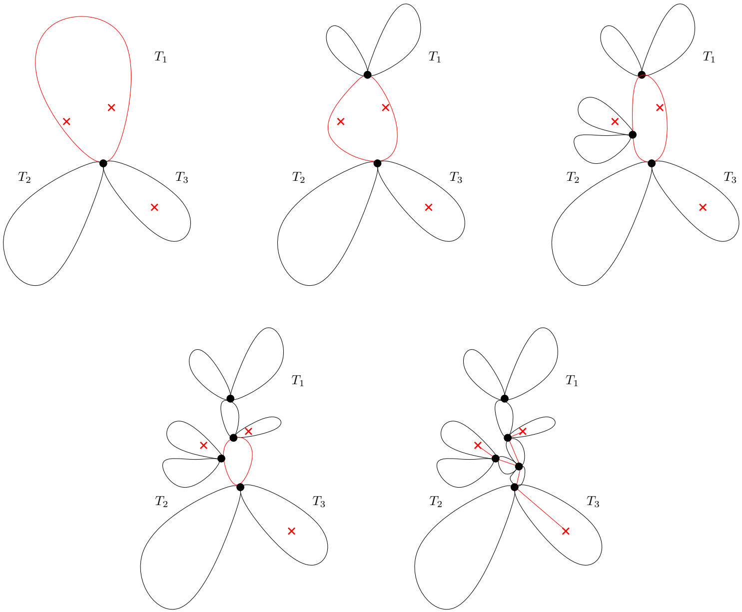
Cases 2 and 3:





Here, in order to understand the distances, we need to figure out what's happening inside one of the level-1 subtrees. Note that, in the two cases, the problem is really the same: we have three uniform points within one of the subtrees, and want to find the distances between them – it's just that in case 2, one of the three points is the branchpoint at the centre. So we really have a new copy of the original problem in this smaller subtree.

Three-point distances: an example



Three-point distances

It is always possible to split the paths between 3 points up into sums of randomly scaled copies of the path between 2 uniform points.

Three-point distances

It is always possible to split the paths between 3 points up into sums of randomly scaled copies of the path between 2 uniform points.

How deep do we need to go in order to find the decomposition? At each level, we either get that the three points are in different subtrees (which occurs with probability $2/35$) or they are not and we need to go one level deeper.

Three-point distances

It is always possible to split the paths between 3 points up into sums of randomly scaled copies of the path between 2 uniform points.

How deep do we need to go in order to find the decomposition? At each level, we either get that the three points are in different subtrees (which occurs with probability $2/35$) or they are not and we need to go one level deeper.

So the depth N to which we need to go satisfies $N \sim \text{Geometric}(2/35)$ which is, in particular, almost surely finite.

Three-point distances

It is always possible to split the paths between 3 points up into sums of randomly scaled copies of the path between 2 uniform points.

How deep do we need to go in order to find the decomposition? At each level, we either get that the three points are in different subtrees (which occurs with probability $2/35$) or they are not and we need to go one level deeper.

So the depth N to which we need to go satisfies $N \sim \text{Geometric}(2/35)$ which is, in particular, almost surely finite.

But since we know the two-point distribution is Rayleigh, the three-point distances must also be uniquely determined (and equal to their distribution for the Brownian CRT).

Random finite-dimensional distributions

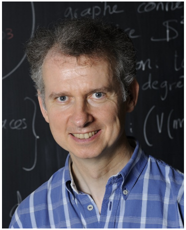
A similar inductive argument shows that the k-point distances may always be expressed in terms of sums of randomly scaled two-point distances, and so again must be uniquely fixed.

The attractiveness of the fixed point again makes use of Durrett and Liggett's theorem for the two-point distances, as well as a slightly complicated coupling. $\square$

6. UNIVERSALITY

Key reference:

Jean-François Le Gall, **Random trees and applications**, *Probability Surveys* **2** (2005) pp.245-311.



A universal scaling limit

Let T_n be the family tree of a Galton-Watson process with critical offspring distribution of variance $\sigma^2 \in (0, \infty)$, conditioned to have total progeny n. Let d_n be the graph distance on T_n and let μ_n be the uniform measure on the vertices.

Theorem. (Aldous (1993), Le Gall (2005))
As $n \to \infty$,

$$\left(T_n, \frac{\sigma}{\sqrt{n}} d_n, \mu_n \right) \xrightarrow{d} (\mathcal{T}_{2e}, d_{2e}, \mu_{2e}),$$

where convergence is in the Gromov-Hausdorff-Prokhorov sense.

Galton-Watson process

A Galton-Watson branching process $(Z_n)_{n \geq 0}$ describes the size of a population which evolves as follows:

- Start with a single individual.
- This individual has a number of children distributed according to the offspring distribution p, where $p(k)$ gives the probability of k children, $k \geq 0$.
- Each child reproduces as an independent copy of the original individual.

Z_n gives the number of individuals in generation n (in particular, $Z_0 = 1$).

Galton-Watson trees

A Galton-Watson tree is the family tree arising from a Galton-Watson branching process. We will think of this as a rooted ordered tree.

Galton-Watson trees

A Galton-Watson tree is the family tree arising from a Galton-Watson branching process. We will think of this as a rooted ordered tree.

Consider the case where the offspring distribution p is critical i.e.

$$\sum_{k=1}^{\infty} kp(k) = 1.$$

This ensures, in particular, that the resulting tree, T, is finite.

Uniform random trees

Proposition. Let T be a (rooted, ordered) Galton-Watson tree, with Poisson(1) offspring distribution and total progeny N. Assign the vertices labels uniformly at random from $\{1, 2, \ldots, N\}$ and then forget the ordering and the root. Let $\tilde{T}$ be the labelled tree obtained. Then, conditional on $N = n$, $\tilde{T}$ has the same distribution as T_n, a uniform random tree on n labelled vertices.

Other combinatorial trees in disguise

Let T be a Galton-Watson tree with offspring distribution p and total progeny N.

- If $p(k) = 2^{-k-1}$, $k \geq 0$ (i.e. Geometric(1/2) offspring distribution) then conditional on $N = n$, the tree is uniform on the set of plane trees with n vertices.
- If $p(0) = 1/2$ and $p(2) = 1/2$ then, conditional on $N = 2n$, the tree is uniform on the set of planted plane binary trees with n leaves.

Two ways of encoding a tree

As we have seen, it is convenient to encode our trees in terms of discrete functions which are easier to manipulate.

We will do this is two different ways:

- the height function
- the depth-first walk.

Height function

Suppose that our tree has n vertices. Let them be $v_0, v_1, \ldots, v_{n-1}$, listed in depth-first order.

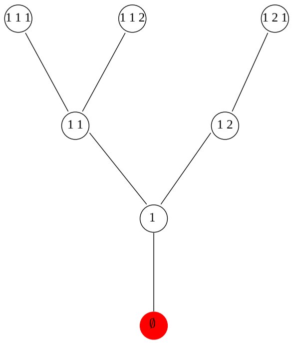
Height function

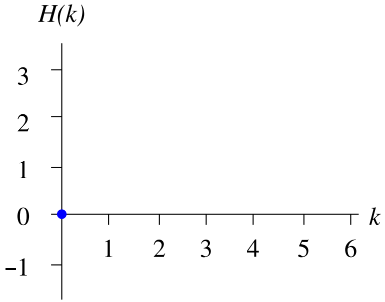
Suppose that our tree has n vertices. Let them be $v_0, v_1, \ldots, v_{n-1}$, listed in depth-first order.

Then the height function is defined by

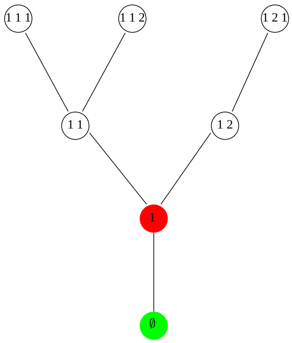
$$H(k) = d_{\mathrm{gr}}(v_0, v_k), \quad 0 \leq k \leq n-1.$$

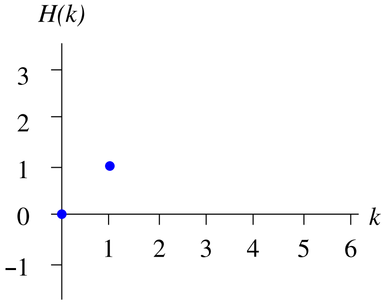
Height function



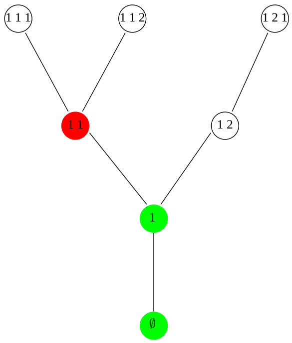


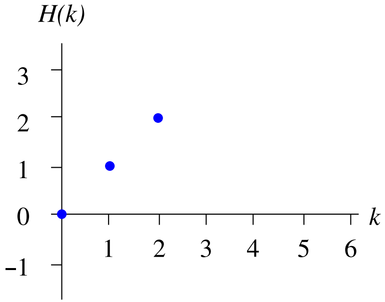
Height function



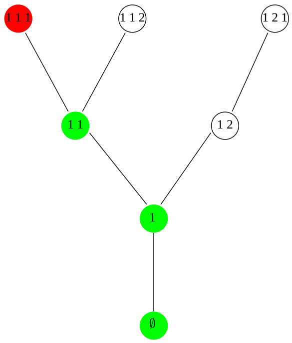


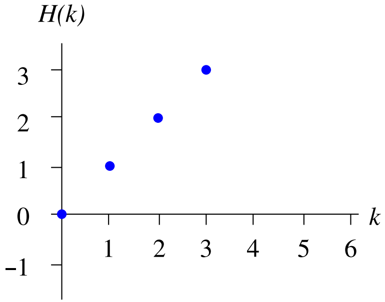
Height function



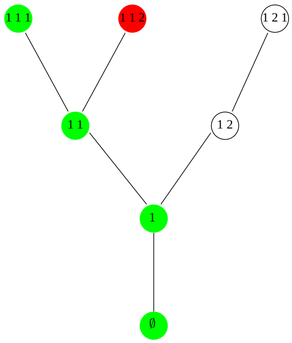


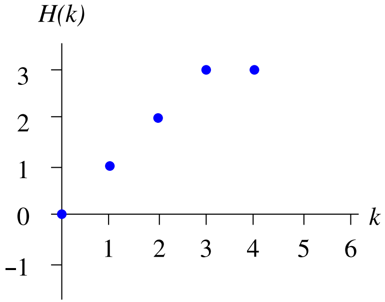
Height function



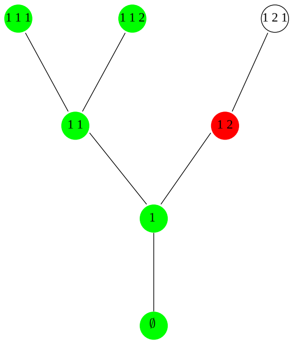


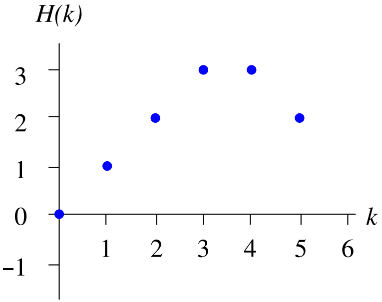
Height function



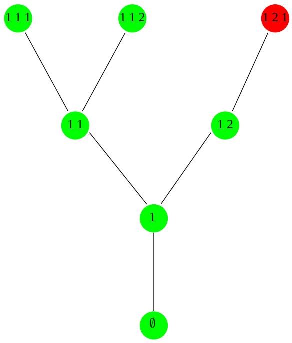


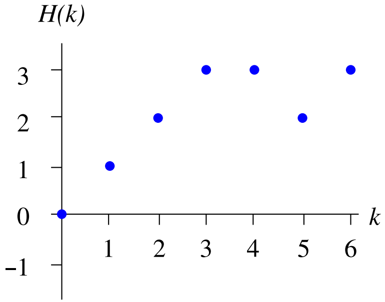
Height function



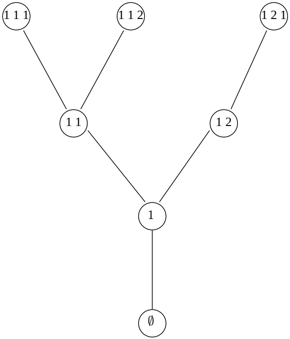


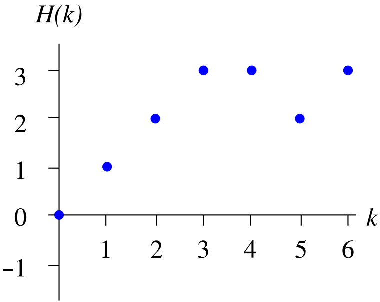
Height function





Height function





We can easily recover the tree from its height function.

Depth-first walk

Let $c(v)$ be the number of children of v, and that $v_0, v_1, \ldots, v_{n-1}$ is a list of the vertices in depth-first order.

Depth-first walk

Let $c(v)$ be the number of children of v, and that $v_0, v_1, \ldots, v_{n-1}$ is a list of the vertices in depth-first order.

Define

$$X(0) = 0,$$
$$X(i) = \sum_{j=0}^{i-1} (c(v_j) - 1), \text{ for } 1 \leq i \leq n.$$

Depth-first walk

Let $c(v)$ be the number of children of v, and that $v_0, v_1, \dots, v_{n-1}$ is a list of the vertices in depth-first order.

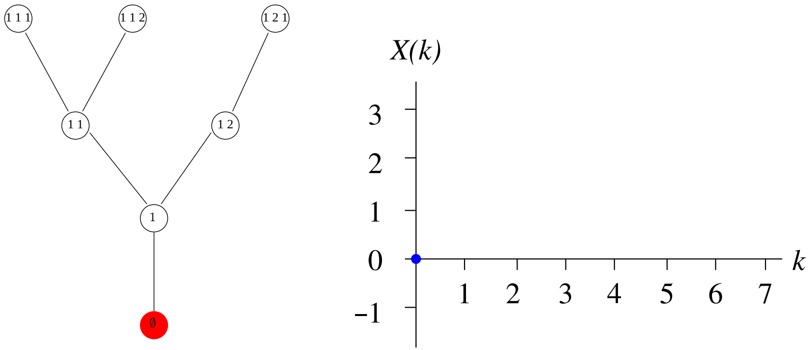
Define

$$X(0) = 0,$$
$$X(i) = \sum_{j=0}^{i-1} (c(v_j) - 1), \text{ for } 1 \leq i \leq n.$$

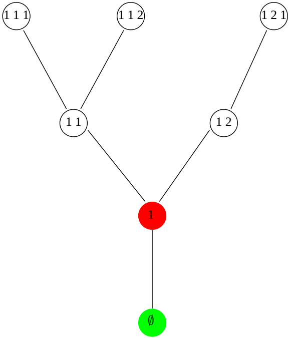
In other words,

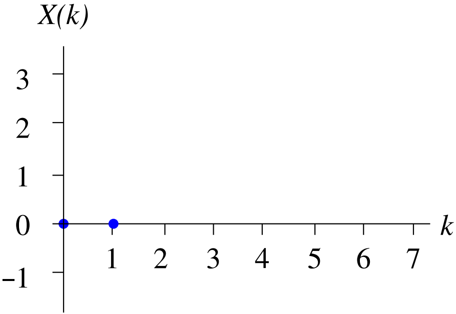
$$X(i+1) = X(i) + c(v_i) - 1, \quad 0 \leq i \leq n-1.$$

Depth-first walk

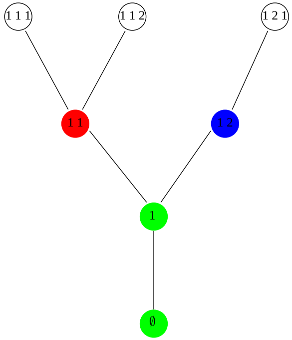


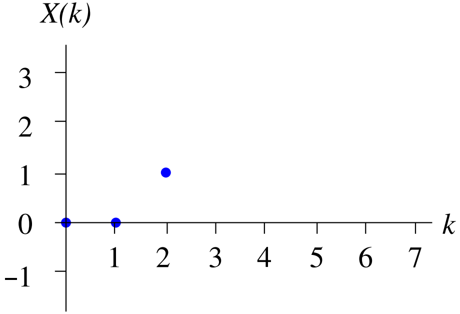
Depth-first walk



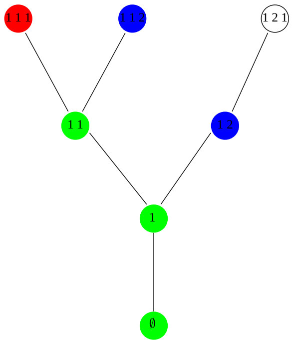


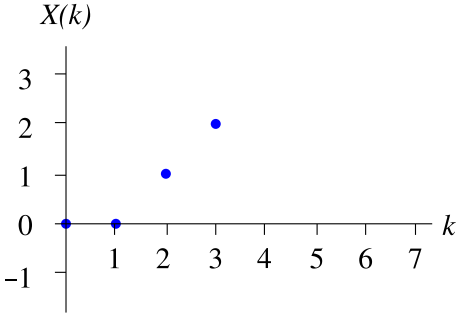
Depth-first walk



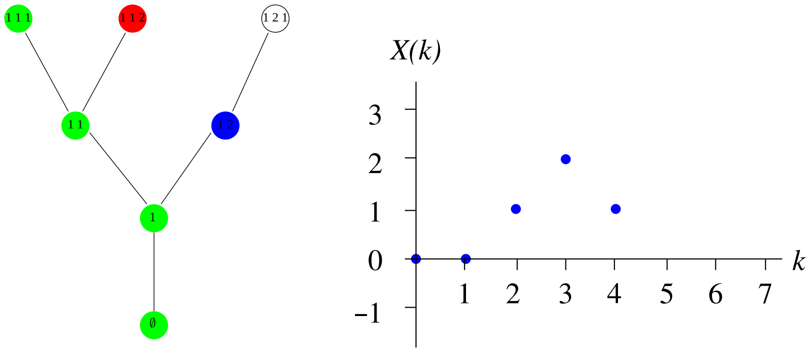


Depth-first walk

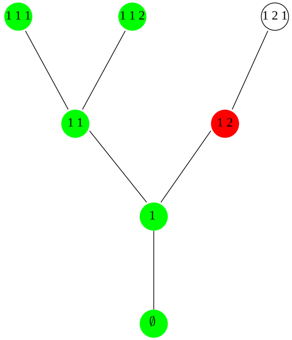


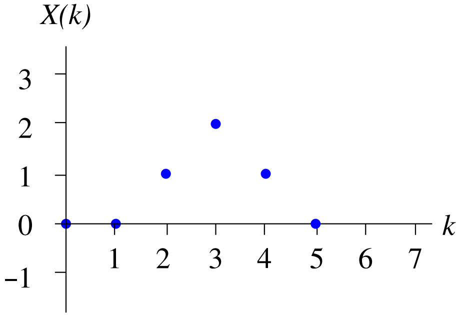


Depth-first walk

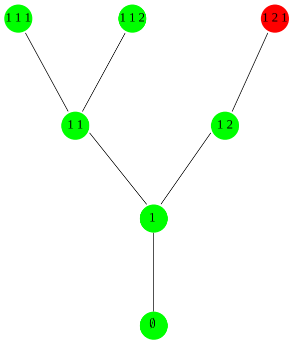


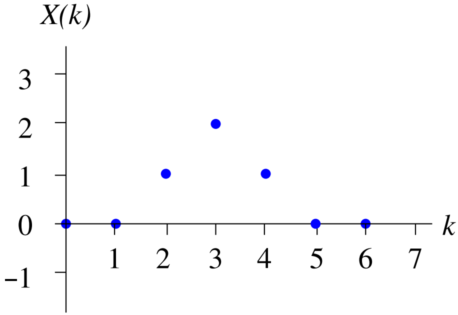
Depth-first walk



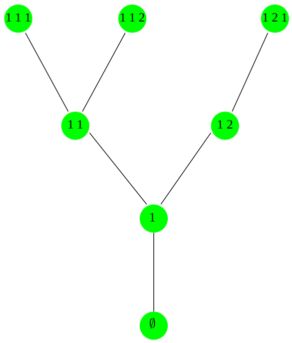


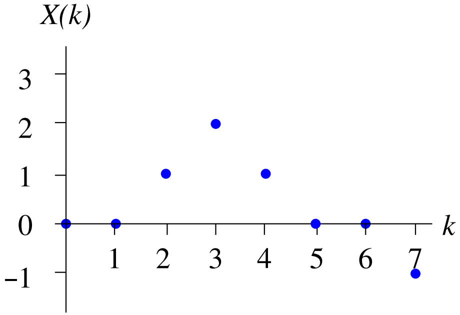
Depth-first walk



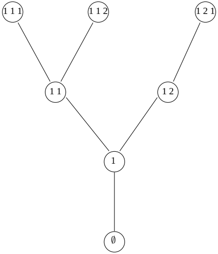


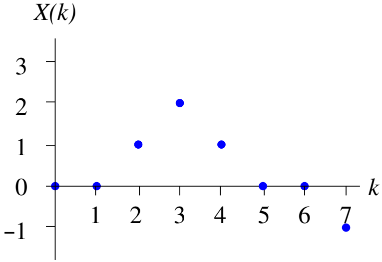
Depth-first walk





Depth-first walk





Proposition. For $0 \le i \le n-1$,

$$H(i) = \#\left\{0 \le j \le i-1 : X(j) = \min_{j \le k \le i} X(k)\right\}.$$

The depth-first walk of a Galton-Watson process is a stopped random walk

Recall that p is a distribution on $\mathbb{Z}_+$ such that $\sum_{k=1}^{\infty} kp(k) = 1$.

The depth-first walk of a Galton-Watson process is a stopped random walk

Recall that p is a distribution on $\mathbb{Z}_+$ such that $\sum_{k=1}^{\infty} kp(k) = 1$.

Proposition. Let $(R(k), k \geq 0)$ be a random walk with initial value 0 and step distribution $\nu(k) = p(k+1), k \geq -1$. Set

$$M = \inf\{k \geq 0 : R(k) = -1\}.$$

The depth-first walk of a Galton-Watson process is a stopped random walk

Recall that p is a distribution on $\mathbb{Z}_+$ such that $\sum_{k=1}^{\infty} kp(k) = 1$.

Proposition. Let $(R(k), k \geq 0)$ be a random walk with initial value 0 and step distribution $\nu(k) = p(k+1), k \geq -1$. Set

$$M = \inf\{k \geq 0 : R(k) = -1\}.$$

Now suppose that T is a Galton-Watson tree with offspring distribution p and total progeny N. Then

$$(X(k), 0 \leq k \leq N) \stackrel{d}{=} (R(k), 0 \leq k \leq M).$$

The depth-first walk of a Galton-Watson process is a stopped random walk

Recall that p is a distribution on $\mathbb{Z}_+$ such that $\sum_{k=1}^{\infty} kp(k) = 1$.

Proposition. Let $(R(k), k \geq 0)$ be a random walk with initial value 0 and step distribution $\nu(k) = p(k+1), k \geq -1$. Set

$$M = \inf\{k \geq 0 : R(k) = -1\}.$$

Now suppose that T is a Galton-Watson tree with offspring distribution p and total progeny N. Then

$$(X(k), 0 \leq k \leq N) \stackrel{d}{=} (R(k), 0 \leq k \leq M).$$

[Careful proof: see Le Gall (2005).]

Galton-Watson trees conditioned on their total progeny

Suppose now that we have offspring variance
$\sigma^2 := \sum_{k=1}^{\infty}(k-1)^2 p(k) \in (0, \infty)$.

The depth-first walk X is a random walk with step mean 0 and variance σ^2, stopped at the first time it hits -1. The underlying random walk has a Brownian motion as its scaling limit, by Donsker's theorem.

The total progeny N is equal to $\inf\{k \geq 0 : X(k) = -1\}$. We want to condition on the event $\{N = n\}$.

Galton-Watson trees conditioned on their total progeny

Suppose now that we have offspring variance $\sigma^2 := \sum_{k=1}^{\infty}(k-1)^2 p(k) \in (0, \infty)$.

The depth-first walk X is a random walk with step mean 0 and variance σ^2, stopped at the first time it hits -1. The underlying random walk has a Brownian motion as its scaling limit, by Donsker's theorem.

The total progeny N is equal to $\inf\{k \geq 0 : X(k) = -1\}$. We want to condition on the event $\{N = n\}$.

Standing assumption: $\mathbb{P}(N = n) > 0$ for all n sufficiently large.

Galton-Watson trees conditioned on their total progeny

Write $(X_n(k), 0 \leq k \leq n)$ for the depth-first walk conditioned on $\{N = n\}$. Then there is a conditional version of Donsker's theorem.

Theorem. As $n \to \infty$,

$$\frac{1}{\sigma\sqrt{n}}(X_n(\lfloor nt \rfloor), 0 \leq t \leq 1) \xrightarrow{d} (e(t), 0 \leq t \leq 1),$$

where $(e(t), 0 \leq t \leq 1)$ is a standard Brownian excursion.

[W.D. Kaigh, **An invariance principle for random walk conditioned by a late return to zero**, *Annals of Probability* **4**, 1976, pp.115-121.]

Height process

Let $(H_n(i), 0 \leq i \leq n)$ be the height process of a critical Galton-Watson tree with offspring variance $\sigma^2 \in (0, \infty)$, conditioned to have total progeny n, so that

$$H_n(i) = \# \left\{ 0 \leq j \leq i - 1 : X_n(j) = \min_{j \leq k \leq i} X_n(k) \right\}.$$

Theorem. As $n \to \infty$,

$$\frac{\sigma}{\sqrt{n}} \left(H_n(\lfloor nt \rfloor), 0 \leq t \leq 1 \right) \xrightarrow{d} 2 \left(e(t), 0 \leq t \leq 1 \right),$$

where $(e(t), 0 \leq t \leq 1)$ is a standard Brownian excursion.

Convergence to the Brownian CRT

The convergence

$$\left(T_n, \frac{\sigma}{\sqrt{n}} d_n, \mu_n\right) \xrightarrow{d} (\mathcal{T}_{2e}, d_{2e}, \mu_{2e}),$$

now follows by the same proof that we used in the case of binary trees.

Universality

The universality class of the Brownian CRT is, in fact, even larger. Some other examples of trees (and graphs!) with the Brownian CRT as their scaling limit are:

- uniform unordered unlabelled rooted trees
- uniform unordered unlabelled unrooted trees
- critical multi-type Galton-Watson trees
- random trees with a prescribed degree sequence satisfying certain conditions
- random dissections
- random graphs from subcritical classes.

A particularly useful tool: Markov branching trees.

[B. Haas & G. Miermont, **Scaling limits of Markov branching trees with applications to Galton-Watson and random unordered trees**, *Annals of Probability* **40**(6), 2012, pp.2589–2666.]

Applications

Universal scaling limits often show up in other places, and the Brownian CRT is no exception. It appears, for example, as a building block in the scaling limit of random planar maps: the Brownian map is constructed as a (complicated) quotient of the Brownian CRT.

What if the offspring variance isn't finite?

Suppose instead that the offspring distribution is critical but in the domain of attraction of an α-stable law, for $\alpha \in (1, 2)$. For example,

$$p(k) \sim Ck^{-1-\alpha} \quad \text{as } k \to \infty$$

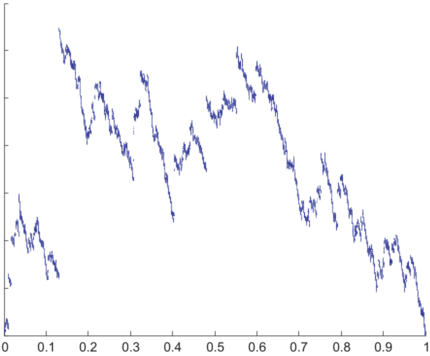
for $C > 0$.

The limiting depth-first walk

We now get

$$\frac{1}{n^{1/\alpha}}(X_n(\lfloor nt \rfloor), 0 \le t \le 1) \xrightarrow{d} (e^{(\alpha)}(t), 0 \le t \le 1),$$

where $e^{(\alpha)}$ is an excursion of a spectrally positive α-stable Lévy process.



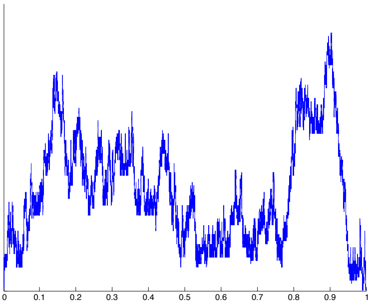
[Picture by Igor Korchemski]

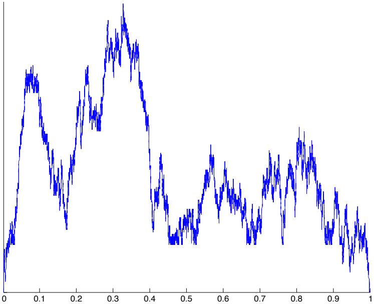
The limiting height process

It's no longer the case that H_n has the same limit. We get

$$\frac{1}{n^{1-1/\alpha}}(H_n(\lfloor nt \rfloor), 0 \leq t \leq 1) \xrightarrow{d} (h^{(\alpha)}(t), 0 \leq t \leq 1)$$

for some much more complicated continuous excursion $h^{(\alpha)}$.





[Pictures by Igor Kortchemski]

The stable trees

Theorem. (Duquesne (2003)) Suppose that $p(k) \sim ck^{-1-\alpha}$ as $k \to \infty$ for $\alpha \in (1,2)$. Then as $n \to \infty$,

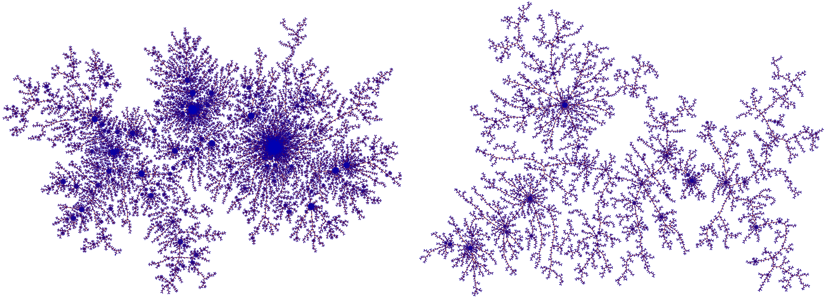
$$\frac{1}{n^{1-1/\alpha}} T_n \xrightarrow{d} c_\alpha \mathcal{T}_\alpha,$$

where $\mathcal{T}_\alpha$ is the stable tree of parameter α and c_α is a strictly positive constant.

[T. Duquesne & J.-F. Le Gall, **Random trees, Lévy processes and spatial branching processes**, *Astérisque* **281**, 2002.]

[T. Duquesne, **A limit theorem for the contour process of conditioned Galton-Watson trees**, *Annals of Probability* **31**(2), 2003, pp.996–1027.]

The stable trees



An important difference between the stable trees for $\alpha \in (1, 2)$ and the Brownian CRT is that the Brownian CRT is binary. The stable trees, on the other hand, have only branch-points of infinite degree.

[Pictures by Igor Kortchemski]

7. CONNECTED GRAPHS

Joint work with Louigi Addario-Berry (McGill) and Nicolas Broutin (Sorbonne Université Paris).

[L. Addario-Berry, N. Broutin & C. Goldschmidt, **The continuum limit of critical random graphs**, *Probability Theory and Related Fields* **152**(3-4), 2012, pp.367–406.]

[L. Addario-Berry, N. Broutin & C. Goldschmidt, **Critical random graphs: limiting constructions and distributional properties**, *Electronic Journal of Probability* **15**, 2010, paper no. 25, pp.741–775.]

Uniform connected graph with fixed surplus

Fix $k \geq 0$ and let G_n^k be a uniform connected graph with vertices labelled by $1, 2, \ldots, n$ and $n + k - 1$ edges.

Uniform connected graph with fixed surplus

Fix $k \geq 0$ and let G_n^k be a uniform connected graph with vertices labelled by $1, 2, \ldots, n$ and $n + k - 1$ edges. We say G_n^k has surplus k.

Uniform connected graph with fixed surplus

Fix $k \geq 0$ and let G_n^k be a uniform connected graph with vertices labelled by $1, 2, \ldots, n$ and $n + k - 1$ edges. We say G_n^k has surplus k.

(For $k = 0$, this is just a uniform random tree on n vertices.)

Uniform connected graph with fixed surplus

Fix $k \geq 0$ and let G_n^k be a uniform connected graph with vertices labelled by $1, 2, \ldots, n$ and $n + k - 1$ edges. We say G_n^k has surplus k.

(For $k = 0$, this is just a uniform random tree on n vertices.)

Write d_n^k for the graph distance and μ_n^k for the uniform measure on the vertices.

Uniform connected graph with fixed surplus

Fix $k \geq 0$ and let G_n^k be a uniform connected graph with vertices labelled by $1, 2, \ldots, n$ and $n + k - 1$ edges. We say G_n^k has surplus k.

(For $k = 0$, this is just a uniform random tree on n vertices.)

Write d_n^k for the graph distance and μ_n^k for the uniform measure on the vertices.

Theorem. (Addario-Berry, Broutin & G. (2012))
There exists a random compact metric measure space $(\mathcal{G}^k, d^k, \mu^k)$ such that

$$\frac{1}{\sqrt{n}}(G_n^k, d_n^k, \mu_n^k) \xrightarrow{d} (\mathcal{G}^k, d^k, \mu^k)$$

as $n \to \infty$.

Uniform connected graph with fixed surplus

Fix $k \geq 0$ and let G_n^k be a uniform connected graph with vertices labelled by $1, 2, \ldots, n$ and $n + k - 1$ edges. We say G_n^k has surplus k.

(For $k = 0$, this is just a uniform random tree on n vertices.)

Write d_n^k for the graph distance and μ_n^k for the uniform measure on the vertices.

Theorem. (Addario-Berry, Broutin & G. (2012))
There exists a random compact metric measure space $(\mathcal{G}^k, d^k, \mu^k)$ such that

$$\frac{1}{\sqrt{n}}(G_n^k, d_n^k, \mu_n^k) \xrightarrow{d} (\mathcal{G}^k, d^k, \mu^k)$$

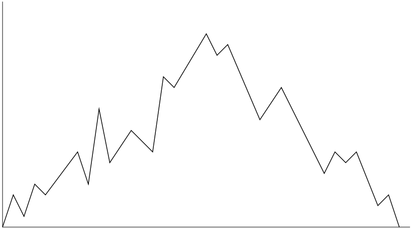
as $n \to \infty$.

We can give an explicit description for the scaling limit.

Scaling limit

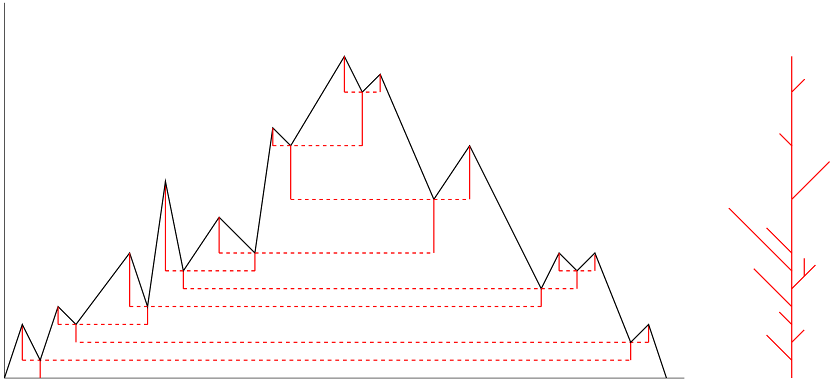
Let e be a standard Brownian excursion. Define a random excursion $\tilde{e}^k : [0,1] \to \mathbb{R}_+$ via a change of measure as follows. For any suitable test-function $f : \mathcal{C}([0,1], \mathbb{R}_+) \to \mathbb{R}$,

$$\mathbb{E}\left[f(\tilde{e}^k(t), 0 \le t \le 1)\right] = \frac{\mathbb{E}\left[f(e(t), 0 \le t \le 1)\left(\int_0^1 e(u)du\right)^k\right]}{\mathbb{E}\left[\left(\int_0^1 e(u)du\right)^k\right]}$$



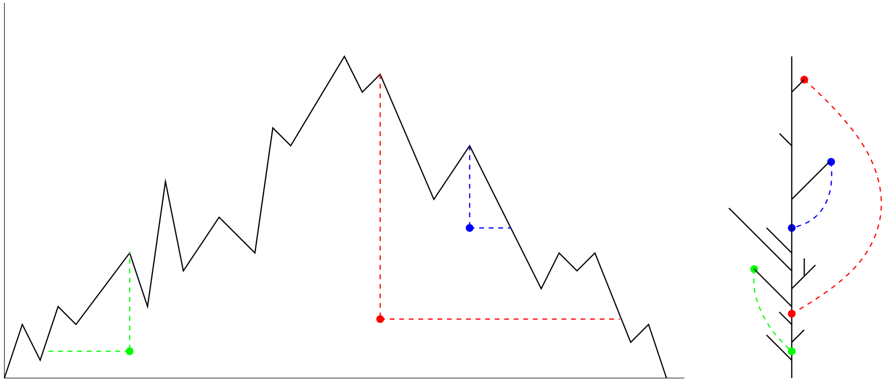
Scaling limit

Use $2\tilde{e}^k$ to encode a continuum random tree $(\tilde{\mathcal{T}}^k, \tilde{d}^k, \tilde{\mu}^k)$.



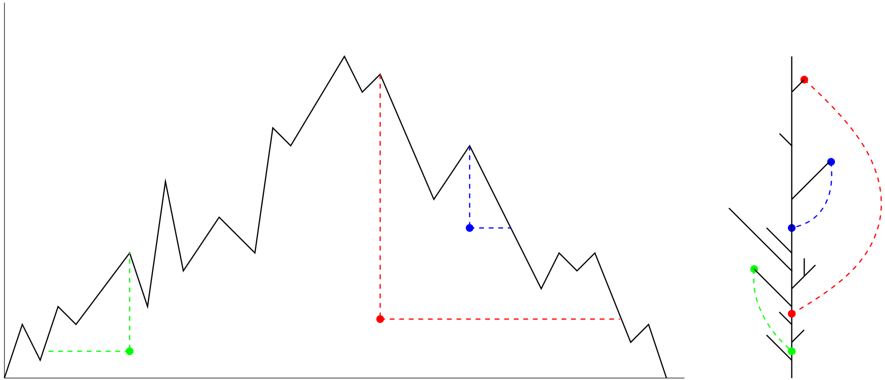
Scaling limit

Pick k independent uniform marks in the area under the curve.
Each mark picks out two points of the tree.

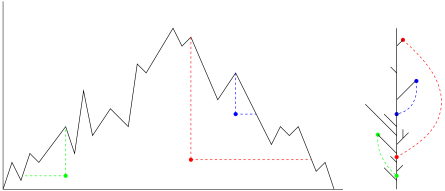


Scaling limit

Pick k independent uniform marks in the area under the curve.
Each mark picks out two points of the tree. Identify them.



Vertex identifications



Write π^k for the usual projection $[0,1] \to \tilde{\mathcal{T}}^k$.

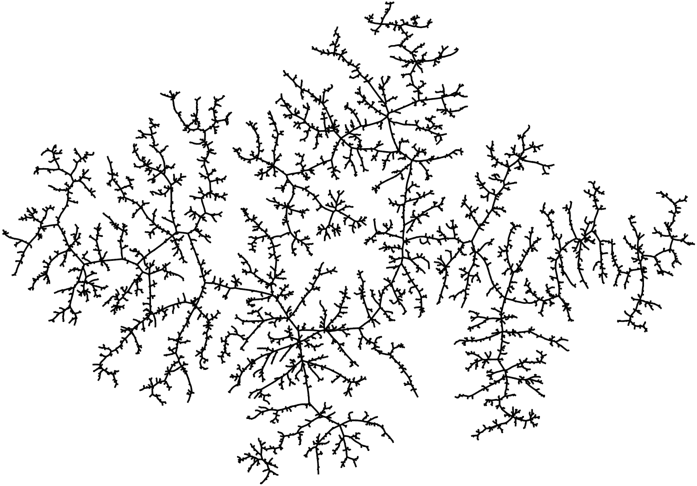
We have marks $(x_1, y_1), \ldots, (x_k, y_k)$ which are uniform in the area under the excursion. For $1 \leq i \leq k$, let

$$t_i = \inf\{t \geq x_i : 2\tilde{e}^k(t) = y_i\}.$$

Define another equivalence relation $\sim$ on $\tilde{\mathcal{T}}^k$ by declaring $\pi^k(x_i) \sim \pi^k(t_i)$ and now let $\mathcal{G}^k = \tilde{\mathcal{T}}^k / \sim$.

We have $\mathcal{G}^k = \tilde{\mathcal{T}}^k / \sim$. Let d^k be the metric and μ^k the measure induced from $\tilde{d}^k$ and $\tilde{\mu}^k$ respectively.

Scaling limit $(\mathcal{G}^k, d^k, \mu^k)$ for $k = 4$



[Picture by Nicolas Broutin]

Proof technique: depth-first exploration

For a tree of size n, we defined the depth-first walk by $X(0) = 0$ and, for $1 \leq k \leq n$,

$$X(k) = \sum_{i=0}^{k-1} (c(v_i) - 1),$$

where $c(v)$ is the number of children of vertex v and $v_0, v_1, \ldots, v_{n-1}$ are the vertices in contour order.

Proof technique: depth-first exploration

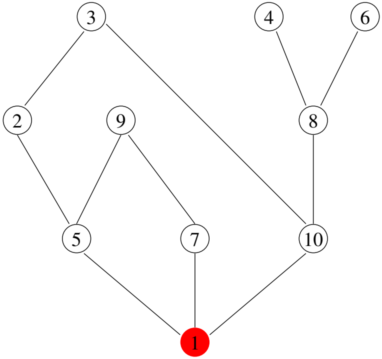
For a tree of size n, we defined the depth-first walk by $X(0) = 0$ and, for $1 \leq k \leq n$,

$$X(k) = \sum_{i=0}^{k-1} (c(v_i) - 1),$$

where $c(v)$ is the number of children of vertex v and $v_0, v_1, \ldots, v_{n-1}$ are the vertices in contour order.

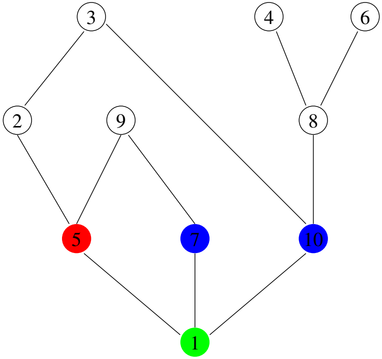
X tracks the number of vertices we have seen, but whose children we have not yet explored.

Depth-first exploration: an example



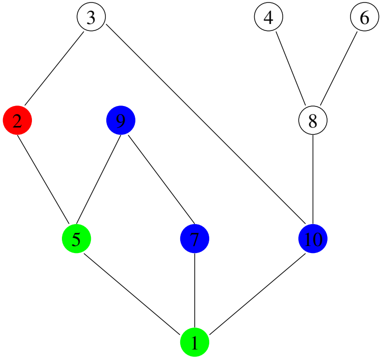
Step 0: $X(0) = 0$

Depth-first exploration: an example



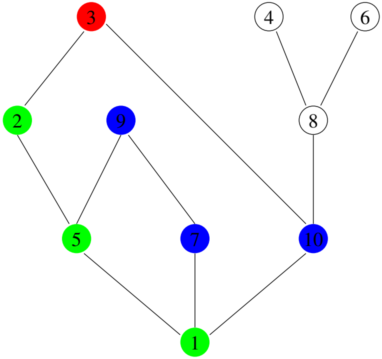
Step 1: $X(1) = 2$

Depth-first exploration: an example



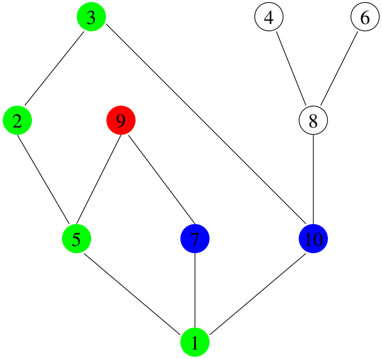
Step 2: $X(2) = 3$

Depth-first exploration: an example



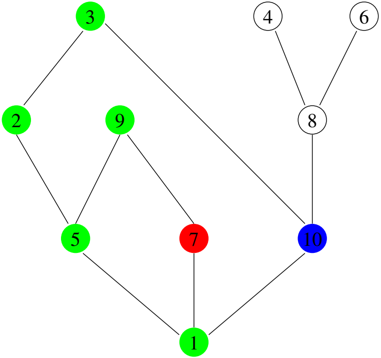
Step 3: $X(3) = 3$

Depth-first exploration: an example



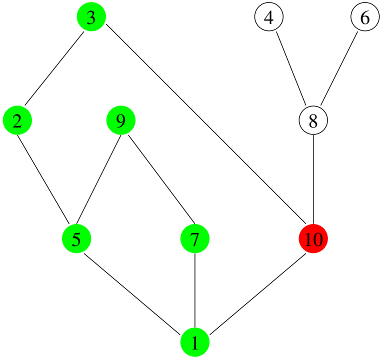
Step 4: $X(4) = 2$

Depth-first exploration: an example



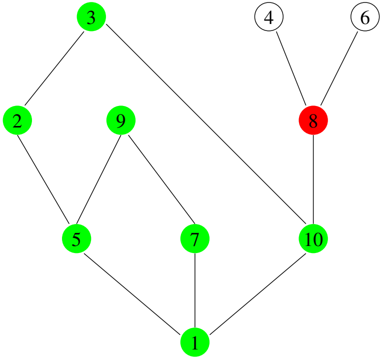
Step 5: $X(5) = 1$

Depth-first exploration: an example



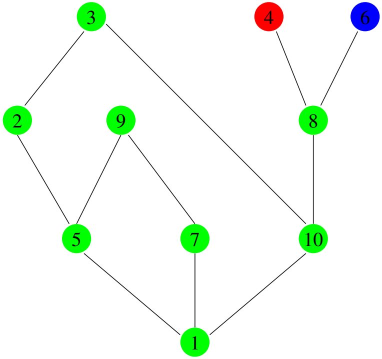
Step 6: $X(6) = 0$

Depth-first exploration: an example



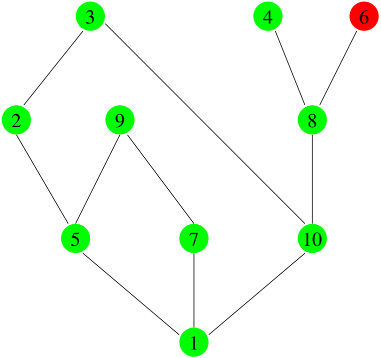
Step 7: $X(7) = 0$

Depth-first exploration: an example



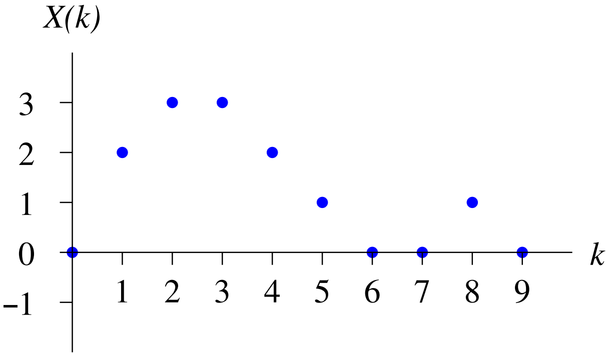
Step 9: $X(8) = 1$

Depth-first exploration: an example



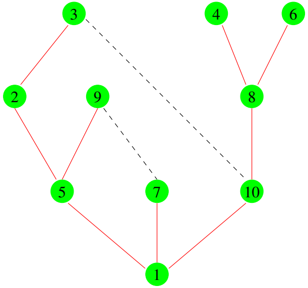
Step 10: $X(9) = 0$

Depth-first walk



Depth-first tree

In the depth-first exploration, we effectively explored this spanning tree; the dashed surplus edges made no difference.



Call the spanning tree the depth-first tree associated with the graph G, and write $T(G)$. X is also the depth-first walk of T.

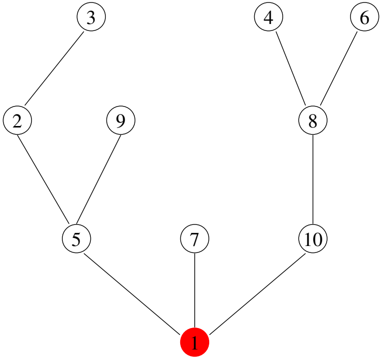
Permitted edges

Look at things the other way round: for a given tree T, which connected graphs G have depth-first tree $T(G) = T$?

In other words, where can we put surplus edges so that they don't change T?

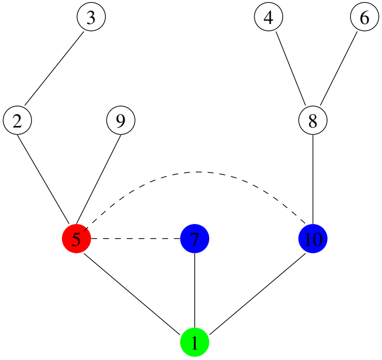
Call such edges permitted.

Depth-first walk and permitted edges



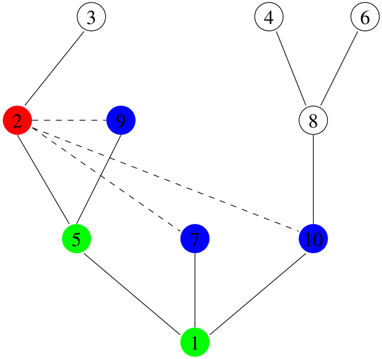
Step 0: $X(0) = 0$.

Depth-first walk and permitted edges



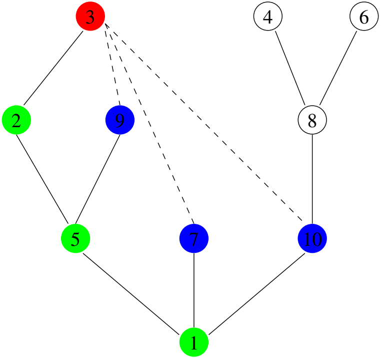
Step 1: $X(1) = 2$.

Depth-first walk and permitted edges



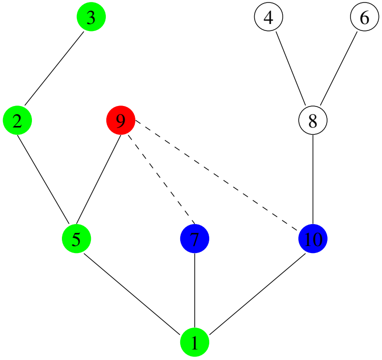
Step 2: $X(2) = 3$.

Depth-first walk and permitted edges



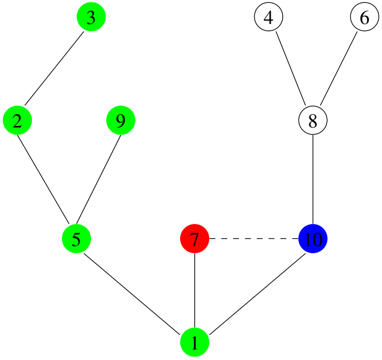
Step 3: $X(3) = 3$.

Depth-first walk and permitted edges



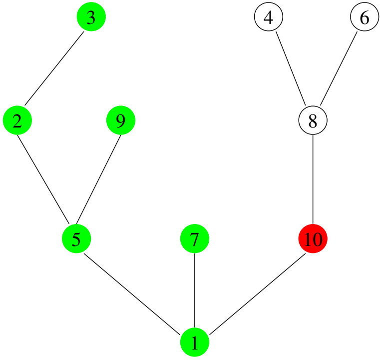
Step 4: $X(4) = 2$.

Depth-first walk and permitted edges



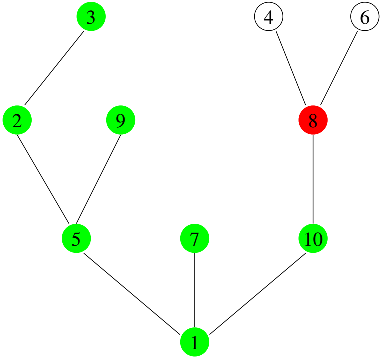
Step 5: $X(5) = 1$.

Depth-first walk and permitted edges



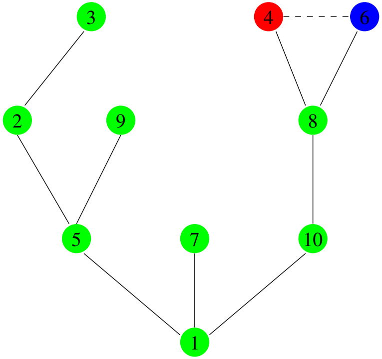
Step 6: $X(6) = 0$.

Depth-first walk and permitted edges



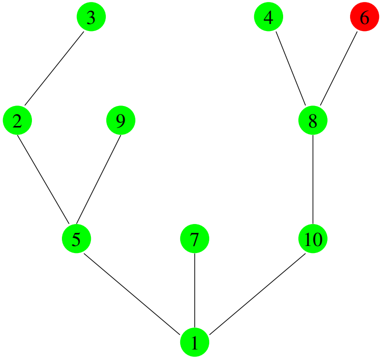
Step 7: $X(7) = 0$.

Depth-first walk and permitted edges



Step 8: $X(8) = 1$.

Depth-first walk and permitted edges



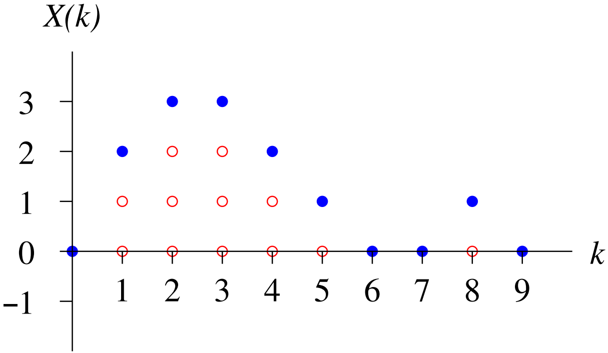
Step 10: $X(9) = 0$.

Area

At step $k \geq 0$, there are $X(k)$ permitted edges. So the total number is

$$a(T) = \sum_{k=0}^{n-1} X(k).$$

We call this the area of T.



Classifying graphs by depth-first tree

Let $\mathbb{G}_T$ be the set of graphs G such that $T(G) = T$. It follows that $|\mathbb{G}_T| = 2^{a(T)}$, since each permitted edge may either be included or not.

Classifying graphs by depth-first tree

Let $\mathbb{G}_T$ be the set of graphs G such that $T(G) = T$. It follows that $|\mathbb{G}_T| = 2^{a(T)}$, since each permitted edge may either be included or not.

Let $\mathbb{T}_{[n]}$ be the set of trees with label-set $[n] = \{1, 2, \ldots, n\}$. Then

$$\left\{\mathbb{G}_T : T \in \mathbb{T}_{[n]}\right\}$$

is a partition of the set of connected graphs on $[n]$.

Recipe for creating a uniform connected graph

Create a uniform connected graph G_n^k as follows.

Recipe for creating a uniform connected graph

Create a uniform connected graph G_n^k as follows.

- Pick a random labelled tree $\tilde{T}_n^k$ such that

$$\mathbb{P}\left(\tilde{T}_m^k = T\right) \propto \binom{a(T)}{k}, \quad T \in \mathbb{T}_{[n]}.$$

Recipe for creating a uniform connected graph

Create a uniform connected graph G_n^k as follows.

- Pick a random labelled tree $\tilde{T}_n^k$ such that

$$\mathbb{P}\left(\tilde{T}_m^k = T\right) \propto \binom{a(T)}{k}, \quad T \in \mathbb{T}_{[n]}.$$

- Choose a uniform k-set from among the $a(\tilde{T}_n^k)$ permitted edges and add them to the tree.

Taking limits

We essentially need to show

- the tree $\tilde{T}_n^k$ converges to a CRT coded by the excursion $\tilde{e}^k$;
- the locations of the surplus edges converge to the locations in the limiting picture.

Taking limits for the tree

Write $\tilde{X}_n^k$ for the depth-first walk associated with $\tilde{T}_n^k$. Then

$$a\left(\tilde{T}_n^k\right) = \sum_{i=0}^{n-1} \tilde{X}_n^k(i) = \int_0^n \tilde{X}_n^k(\lfloor s \rfloor) ds = n^{3/2} \int_0^1 n^{-1/2} \tilde{X}_n^k(\lfloor nu \rfloor) du,$$

by changing variable in the integral.

Taking limits for the tree

Write $\tilde{X}_n^k$ for the depth-first walk associated with $\tilde{T}_n^k$. Then

$$a\left(\tilde{T}_n^k\right) = \sum_{i=0}^{n-1} \tilde{X}_n^k(i) = \int_0^n \tilde{X}_n^k(\lfloor s \rfloor) ds = n^{3/2} \int_0^1 n^{-1/2} \tilde{X}_n^k(\lfloor nu \rfloor) du,$$

by changing variable in the integral.

If T_n is a uniform random tree on $[n]$ and X_n is its depth-first walk, then

$$(n^{-1/2} X_n(\lfloor nt \rfloor), 0 \leq t \leq 1) \xrightarrow{d} (e(t), 0 \leq t \leq 1).$$

Taking limits for the tree

Write $\tilde{X}_n^k$ for the depth-first walk associated with $\tilde{T}_n^k$. Then

$$a\left(\tilde{T}_n^k\right) = \sum_{i=0}^{n-1} \tilde{X}_n^k(i) = \int_0^n \tilde{X}_n^k(\lfloor s \rfloor)ds = n^{3/2} \int_0^1 n^{-1/2} \tilde{X}_n^k(\lfloor nu \rfloor)du,$$

by changing variable in the integral.

If T_n is a uniform random tree on $[n]$ and X_n is its depth-first walk, then

$$(n^{-1/2} X_n(\lfloor nt \rfloor), 0 \leq t \leq 1) \xrightarrow{d} (e(t), 0 \leq t \leq 1).$$

So by the continuous mapping theorem,

$$\int_0^1 n^{-1/2} X_n(\lfloor nu \rfloor)du \xrightarrow{d} \int_0^1 e(u)du.$$

Taking limits for the tree

Use the change of measure to get from $\tilde{X}_n^k$ to X_n: for any bounded continuous function f,

$$\mathbb{E}\left[f\left(n^{-1/2}\tilde{X}_n^k(\lfloor nt\rfloor), 0 \le t \le 1\right)\right]$$

$$= \frac{\mathbb{E}\left[f\left(n^{-1/2}X_n(\lfloor nt\rfloor), 0 \le t \le 1\right)\binom{n^{3/2}\int_0^1 n^{-1/2}X_n(\lfloor nu\rfloor)du}{k}\right]}{\mathbb{E}\left[\binom{n^{3/2}\int_0^1 n^{-1/2}X_n(\lfloor nu\rfloor)du}{k}\right]}$$

Taking limits for the tree

Use the change of measure to get from $\tilde{X}_n^k$ to X_n: for any bounded continuous function f,

$$\mathbb{E}\left[f\left(n^{-1/2}\tilde{X}_n^k(\lfloor nt\rfloor), 0 \le t \le 1\right)\right] = \frac{\mathbb{E}\left[f\left(n^{-1/2}X_n(\lfloor nt\rfloor), 0 \le t \le 1\right)\binom{n^{3/2}\int_0^1 n^{-1/2}X_n(\lfloor nu\rfloor)du}{k}\right]}{\mathbb{E}\left[\binom{n^{3/2}\int_0^1 n^{-1/2}X_n(\lfloor nu\rfloor)du}{k}\right]}$$

We have

$$n^{-3k/2}\binom{n^{3/2}\int_0^1 n^{-1/2}X_n(\lfloor nu\rfloor)du}{k} \xrightarrow{d} \frac{\left(\int_0^1 e(s)ds\right)^k}{k!} \quad \text{as } n \to \infty.$$

Taking limits for the tree

We also have uniform integrability, so we obtain

$$\mathbb{E}\left[f\left(n^{-1/2}\tilde{X}_n^k(nt), 0 \leq t \leq 1\right)\right] \to \frac{\mathbb{E}\left[f(e)\left(\int_0^1 e(u)du\right)^k\right]}{\mathbb{E}\left[\left(\int_0^1 e(u)du\right)^k\right]}$$

$$= \mathbb{E}\left[f(\tilde{e}^k)\right].$$

Taking limits for the tree

We also have uniform integrability, so we obtain

$$\mathbb{E}\left[f\left(n^{-1/2}\tilde{X}_n^k(nt), 0 \le t \le 1\right)\right] \to \frac{\mathbb{E}\left[f(e)\left(\int_0^1 e(u)du\right)^k\right]}{\mathbb{E}\left[\left(\int_0^1 e(u)du\right)^k\right]}$$

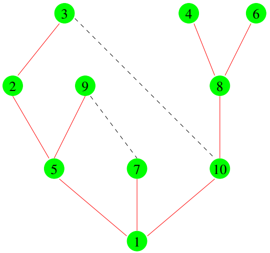
$$= \mathbb{E}\left[f(\tilde{e}^k)\right].$$

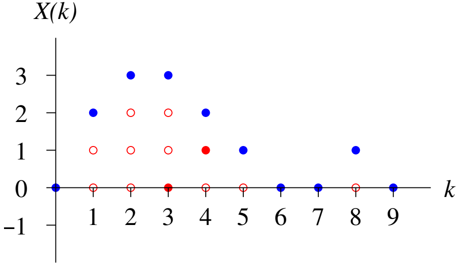
This (after converting to the height process) entails that

$$\frac{1}{\sqrt{n}}\tilde{T}_n^k \xrightarrow{d} \tilde{\mathcal{T}}^k.$$

Taking limits for the surplus edges

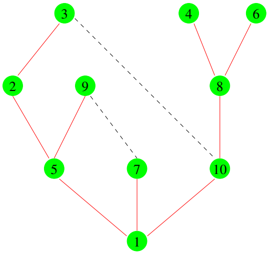
The permitted edges are in bijective correspondence with the integer points under the graph of the depth-first walk.

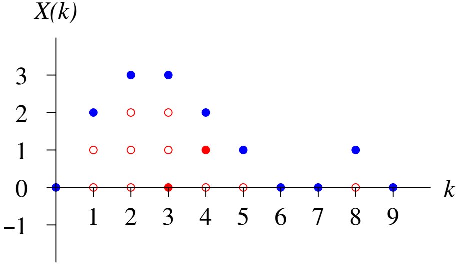




Taking limits for the surplus edges

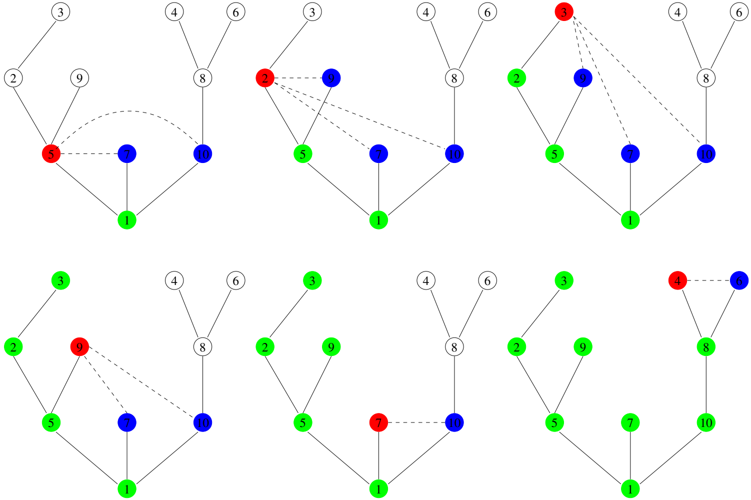
The permitted edges are in bijective correspondence with the integer points under the graph of the depth-first walk. Since we pick a uniform k-set from among these points, in the limit what we see is just k points picked independently and uniformly from the area under the limit curve.





Taking limits for the surplus edges

Surplus edges almost go to ancestors... In fact, they always go to younger children of ancestors of the current vertex.



Taking limits for the surplus edges

When we rescale, the distance between a vertex and one of its children vanishes and so, in the limit, surplus “edges” do go to ancestors of the current vertex (i.e. vertices on the path down to the root).

Taking limits for the surplus edges

When we rescale, the distance between a vertex and one of its children vanishes and so, in the limit, surplus "edges" do go to ancestors of the current vertex (i.e. vertices on the path down to the root).

The marks corresponding to the surplus edges, when rescaled, straightforwardly converge to the required independent uniform points.

Taking limits for the surplus edges

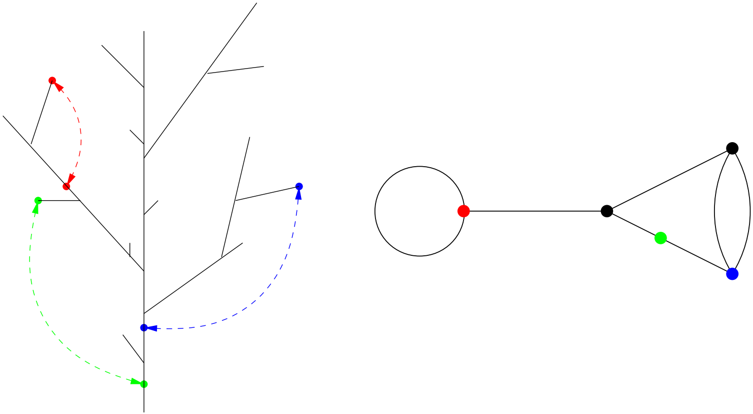
When we rescale, the distance between a vertex and one of its children vanishes and so, in the limit, surplus “edges” do go to ancestors of the current vertex (i.e. vertices on the path down to the root).

The marks corresponding to the surplus edges, when rescaled, straightforwardly converge to the required independent uniform points.

Taking care over the details, this completes the proof.

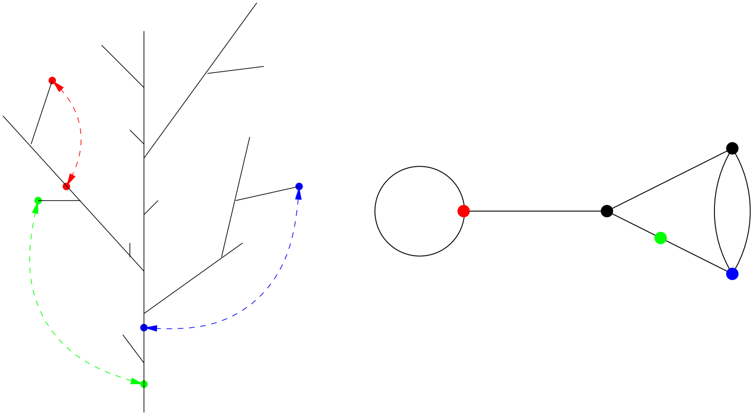
Cycle structure of a limit component

A limiting component can have quite a complicated cycle structure:



Cycle structure of a limit component

A limiting component can have quite a complicated cycle structure:



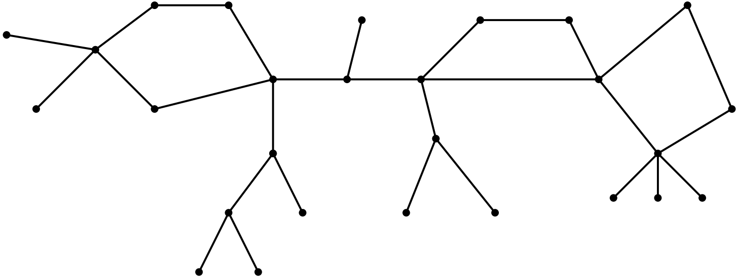
What more can we say about it?

Cycle structure of a graph

Fix a connected graph G. The core $C(G)$ consists of the edges in cycles and those joining the cycles. If G is a tree, $C(G)$ is empty.

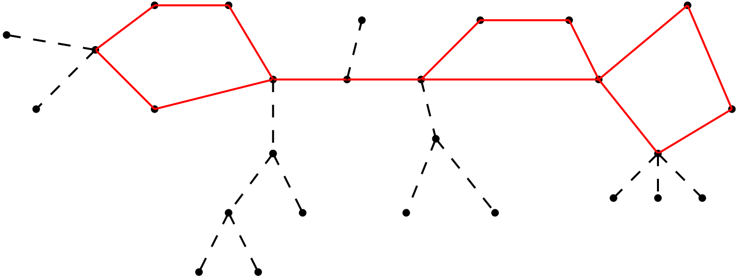
Cycle structure: an example

Graph G



Cycle structure: an example

Core $C(G)$



Cycle structure of a graph

The kernel $K(G)$ is the multigraph which gives the "shape of the core":

Cycle structure of a graph

The kernel $K(G)$ is the multigraph which gives the "shape of the core": take the vertices of the core of degree 3 or more; contract the paths between them to a single edge.

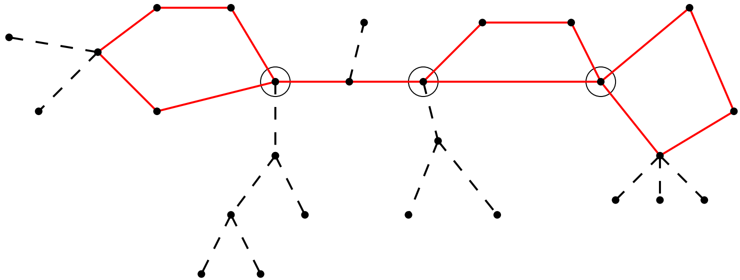
Cycle structure of a graph

The kernel $K(G)$ is the multigraph which gives the "shape of the core": take the vertices of the core of degree 3 or more; contract the paths between them to a single edge.

By convention, the kernel of a tree or unicyclic component is empty.

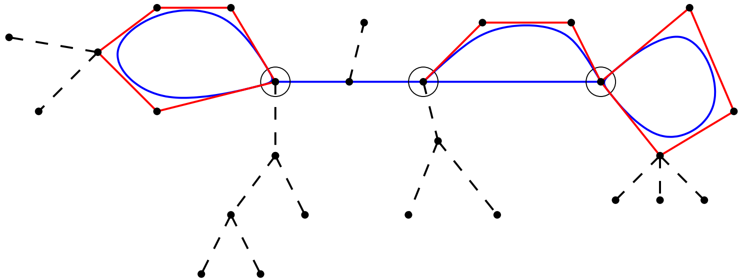
Cycle structure: an example

Vertices of degree at least 3 in the core



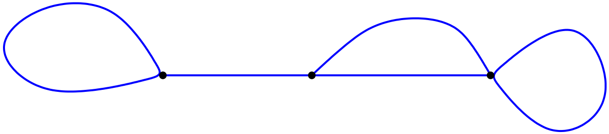
Cycle structure: an example

Contract paths between them



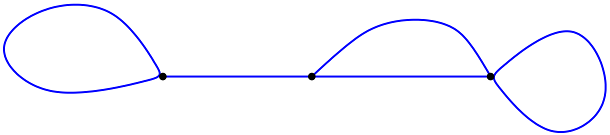
Cycle structure: an example

Kernel $K(G)$



Cycle structure: an example

Kernel $K(G)$



Note that the kernel has the same surplus as the original graph.

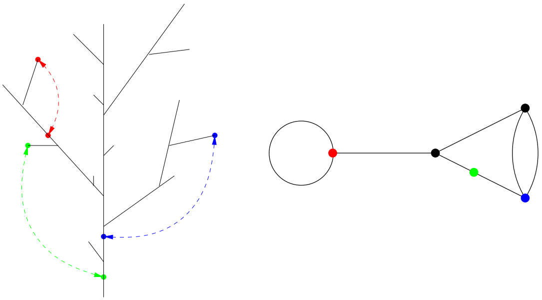
Cycle structure of a real tree with vertex identifications

It still makes sense to talk about the degree of a point in a real tree with vertex identifications.

Cycle structure of a real tree with vertex identifications

It still makes sense to talk about the degree of a point in a real tree with vertex identifications.

It's not hard to see that the core and kernel also make sense in the real tree context as a path metric space and a discrete multigraph respectively.



Law of the kernel

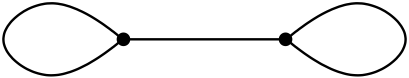
The kernel is distributed as a 3-regular multigraph sampled from the configuration model, conditioned to be connected.

In other words, take the vertices of the kernel, attach 3 half-edges to each, and take a uniformly random pairing of the half-edges to create full edges. Condition the resulting multigraph to be connected.

[S. Janson, D.E. Knuth, T. Łuczak & B. Pittel, **The birth of the giant component**, *Random Structures and Algorithms* **4**, 1993, pp.233–358]

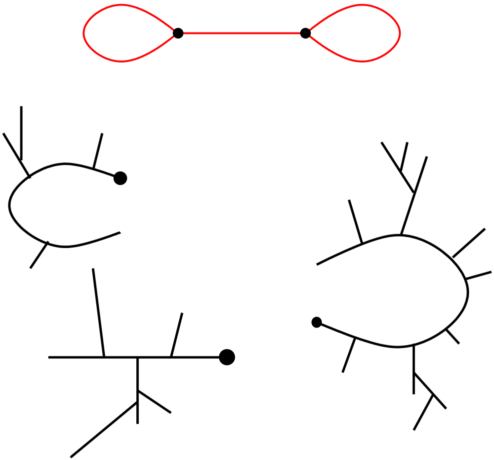
Alternative construction of a limit component

Sample a kernel according to the 3-regular configuration model, conditioned to be connected and to have surplus k. Such a kernel always has $3k - 3$ edges.



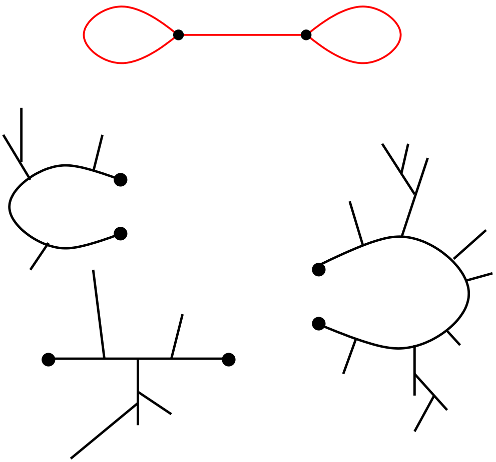
Alternative construction of a limit component

Sample independent rooted Brownian CRT's $\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_{3k-3}$.



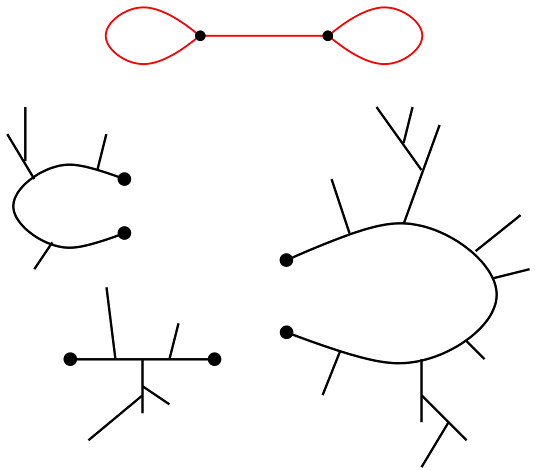
Alternative construction of a limit component

Sample a uniform point in each.



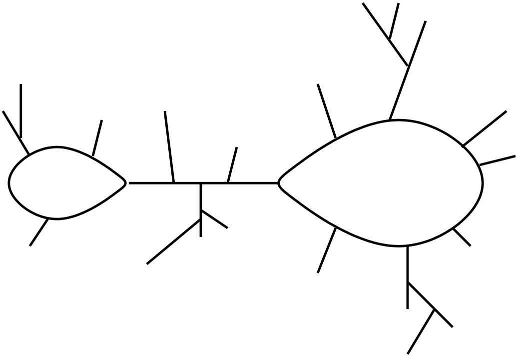
Alternative construction of a limit component

Randomly rescale so that the mass of $\mathcal{T}_i$ becomes X_i, where $(X_1, X_2, \ldots, X_{3k-3}) \sim \text{Dir}(\frac{1}{2}, \ldots, \frac{1}{2})$.



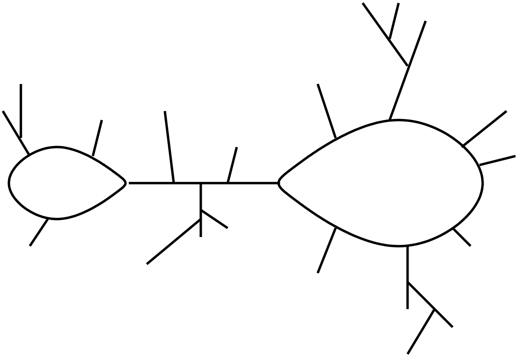
Alternative construction of a limit component

Glue the trees to the kernel.



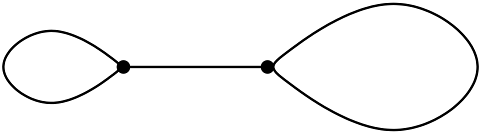
Alternative construction of a limit component

Glue the trees to the kernel.



This has the same distribution as $(\mathcal{G}^k, d^k, \mu^k)$.

The core



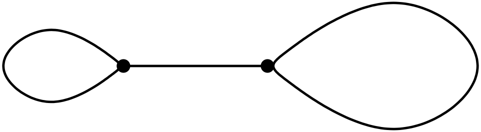
We get core paths of lengths

$$\sqrt{\Gamma} \times \mathrm{Dir}(\underbrace{1, 1, \ldots, 1}_{3k-3}),$$

where the two factors are independent and

$$\Gamma \sim \mathrm{Gamma}((3k-2)/2, 1/2).$$

The core



We get core paths of lengths

$$\sqrt{\Gamma} \times \mathrm{Dir}(\underbrace{1, 1, \ldots, 1}_{3k-3}),$$

where the two factors are independent and

$$\Gamma \sim \mathrm{Gamma}((3k-2)/2, 1/2).$$

(Compare to the random k-dimensional distributions in the Brownian CRT, where we had $\sqrt{\Gamma_k}$ with $\Gamma_k \sim \mathrm{Gamma}(k, 1/2)$, $k \geq 1$.)

Line-breaking construction

Starting from the core, it turns out we can give a line-breaking construction for the rest of the limit component.

Line-breaking construction

Starting from the core, it turns out we can give a line-breaking construction for the rest of the limit component.

Take an inhomogeneous Poisson process of rate t at time t, conditioned to have its first point at $\sqrt{\Gamma}$. Write $C_0 = \sqrt{\Gamma}$; subsequent points occur at times $C_1, C_2, \ldots$.

Line-breaking construction

Starting from the core, it turns out we can give a line-breaking construction for the rest of the limit component.

Take an inhomogeneous Poisson process of rate t at time t, conditioned to have its first point at $\sqrt{\Gamma}$. Write $C_0 = \sqrt{\Gamma}$; subsequent points occur at times $C_1, C_2, \ldots$.

Consider the line-segments $[C_0, C_1)$, $[C_1, C_2), \ldots$ and proceed inductively.

Line-breaking construction

Starting from the core, it turns out we can give a line-breaking construction for the rest of the limit component.

Take an inhomogeneous Poisson process of rate t at time t, conditioned to have its first point at $\sqrt{\Gamma}$. Write $C_0 = \sqrt{\Gamma}$; subsequent points occur at times $C_1, C_2, \ldots$.

Consider the line-segments $[C_0, C_1)$, $[C_1, C_2), \ldots$ and proceed inductively.

For $i \geq 0$, attach $[C_i, C_{i+1})$ at a random point chosen uniformly over the existing structure.

Line-breaking construction

Starting from the core, it turns out we can give a line-breaking construction for the rest of the limit component.

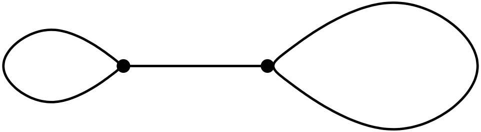
Take an inhomogeneous Poisson process of rate t at time t, conditioned to have its first point at $\sqrt{\Gamma}$. Write $C_0 = \sqrt{\Gamma}$; subsequent points occur at times $C_1, C_2, \ldots$.

Consider the line-segments $[C_0, C_1)$, $[C_1, C_2), \ldots$ and proceed inductively.

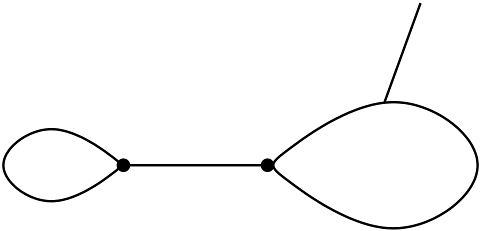
For $i \geq 0$, attach $[C_i, C_{i+1})$ at a random point chosen uniformly over the existing structure.

Take the completion of the metric space obtained.

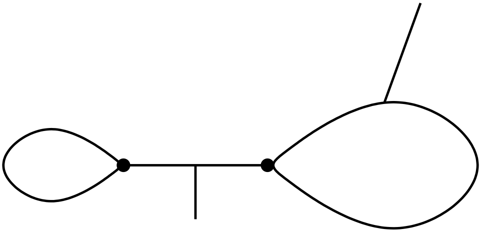
Line-breaking construction



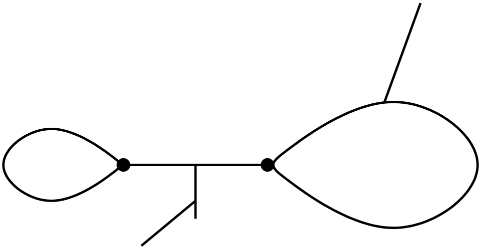
Line-breaking construction



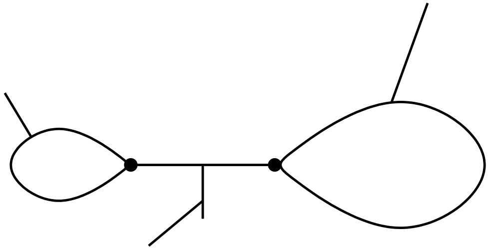
Line-breaking construction



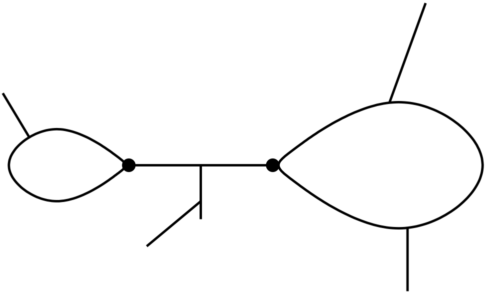
Line-breaking construction



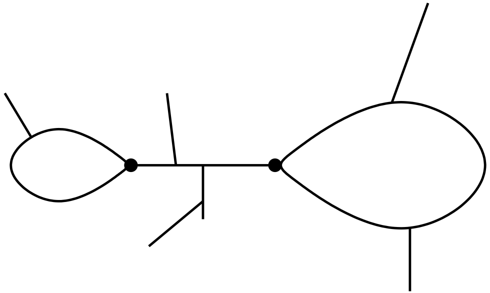
Line-breaking construction



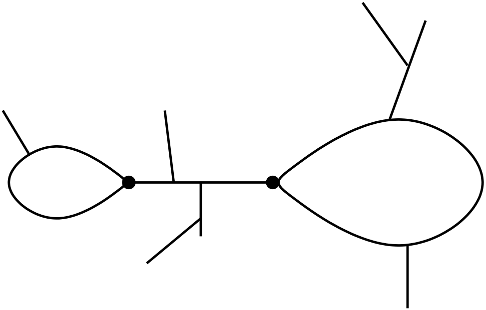
Line-breaking construction



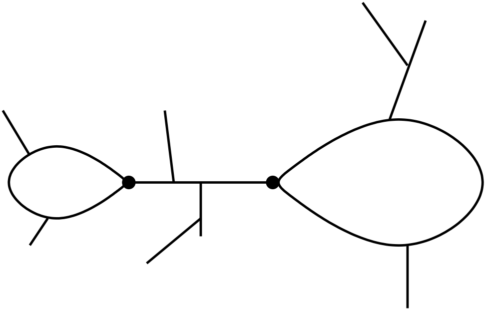
Line-breaking construction



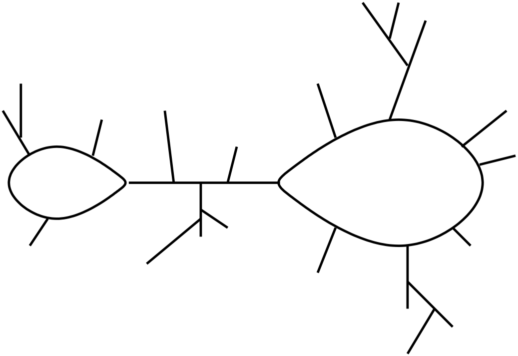
Line-breaking construction



Line-breaking construction



Line-breaking construction



8. PERSPECTIVES

(a) Unicellular random maps

Let $\mathcal{S}$ be an arbitrary compact surface without boundary. Let M_n be a uniform random map drawn on $\mathcal{S}$ with n vertices and a single face. (M_n is unicellular.)

(a) Unicellular random maps

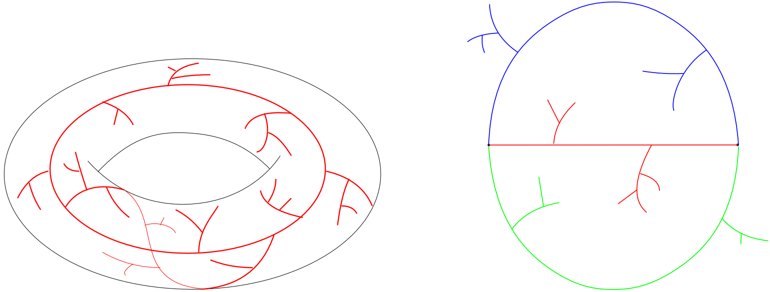
Let $\mathcal{S}$ be an arbitrary compact surface without boundary. Let M_n be a uniform random map drawn on $\mathcal{S}$ with n vertices and a single face. (M_n is unicellular.)

If $\mathcal{S}$ is the sphere then M_n is a uniform random plane tree, and has the Brownian CRT as its scaling limit.

Note: this is a very different object to the Brownian surfaces, which are the scaling limits of random maps with diverging numbers of faces!

Unicellular random maps

If $\mathcal{S}$ is a torus, we generically get:



This is an embedded version of a graph conditioned to have "theta" kernel (which is one of the two possible kernels with 2 surplus edges). The scaling limit can then be constructed out of three independent randomly rescaled Brownian CRT's.

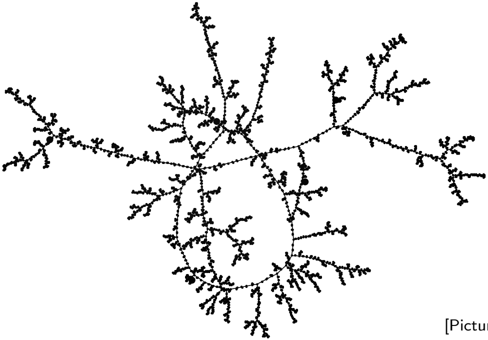
Unicellular random maps

On any surface, we get as a scaling limit (a copy of $\mathcal{G}^k$ for some k, conditioned to have one or a mixture of kernels).

$$\left(M_n, \frac{1}{\sqrt{n}} d_n, \mu_n\right) \xrightarrow{d} (\mathcal{M}, d, \mu).$$

$\mathcal{M}$ can always be constructed out of randomly rescaled independent Brownian CRT's.

Genus 2:

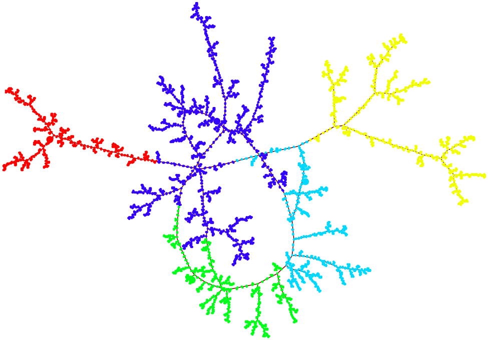


[Picture by Igor Kortchemski]

A generalisation of our Voronoi theorem

Theorem. (Addario-Berry, Angel, Chapuy, Fusy & G. (2018+))
For any compact surface $\mathcal{S}$ without boundary, the continuum random unicellular map $(\mathcal{M}, d, \mu)$ has uniform Voronoi mass-partitions.

Genus 2, $k = 5$:



[Picture by Igor Kortchemski]

(b) The critical Erdős-Rényi random graph

Consider the Erdős-Rényi random graph $G(n, p)$. There is a phase transition for the emergence of a giant component at $p = 1/n$. Aldous (1997) gives a description of the component sizes and surpluses in the critical window $p = 1/n + \lambda n^{-4/3}$. Here, the component sizes are on the order of $n^{2/3}$ and the surpluses are finite random variables.

Using the fact that components of the Erdős-Rényi random graph are uniform on their vertex-sets with the number of edges determined by the size of the vertex-set and the surplus, we can obtain a metric-space scaling limit for the whole graph.

[D. Aldous, **Brownian excursions, critical random graphs and the multiplicative coalescent**, *Annals of Probability* **25**, 1997, pp.812–854.]

[L. Addario-Berry, N. Broutin & C. Goldschmidt, **The continuum limit of critical random graphs**, *Probability Theory and Related Fields* **152**(3-4), 2012, pp.367–406.]

The critical Erdős-Rényi random graph

Let $p = 1/n + \lambda n^{-4/3}$ for fixed $\lambda \in \mathbb{R}$. Let $C_1^n, C_2^n, \ldots$ be the components of $G(n,p)$ listed in decreasing order of size, and let $d_1^n, d_2^n, \ldots$ be the graph distances and $\mu_1^n, \mu_2^n, \ldots$ be the counting measures on $C_1^n, C_2^n, \ldots$ respectively.

Theorem. (Addario-Berry, Broutin & G. (2012))
As $n \to \infty$,

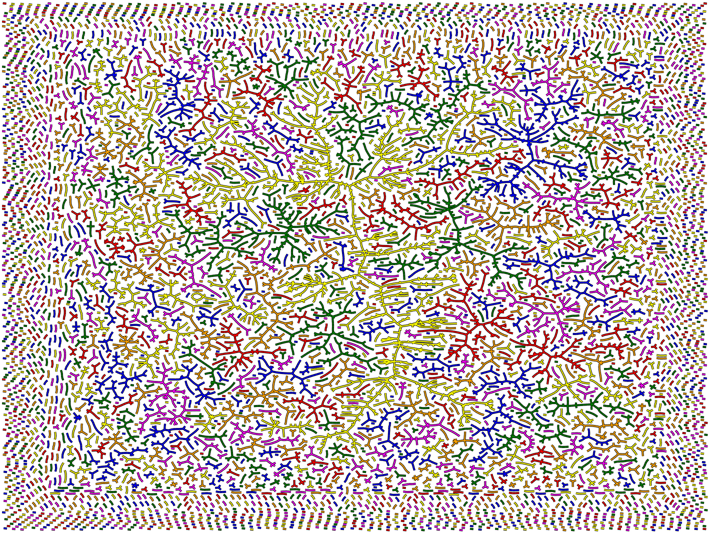
$$\left(\left(C_1^n, \frac{d_1^n}{n^{1/3}}, \frac{1}{n^{2/3}}\mu_1^n\right), \left(C_2^n, \frac{d_2^n}{n^{1/3}}, \frac{1}{n^{2/3}}\mu_2^n\right), \ldots\right)$$

$$\xrightarrow{d} ((\mathcal{C}_1, d_1, \mu_1), (\mathcal{C}_2, d_2, \mu_2), \ldots)$$

in an ℓ_4 version of GHP.

Here, the limit spaces are randomly scaled i.i.d. copies of $(\mathcal{G}^k, d^k, \mu^k)$ with a certain random surplus k.

The critical Erdős-Rényi random graph



[Picture by Nicolas Broutin]

Universality

[S. Bhamidi, N. Broutin, S. Sen & X. Wang, **Scaling limits of random graph models at criticality: Universality and the basin of attraction of the Erdős-Rényi random graph**, `arXiv:1411.3417`, 2014+.]

[S. Bhamidi & S. Sen, **Geometry of the vacant set left by random walk on random graphs, Wright's constants, and critical random graphs with prescribed degrees**, `arXiv:1608.07153`, 2016+.]

[S. Bhamidi, S. Sen & X. Wang, **Continuum limit of inhomogenous random graphs**, *Probability Theory and Related Fields*, to appear; `arXiv:1404.4118`, 2014+.]

[S. Bhamidi, R. van der Hofstad & J. van Leeuwaarden, **Scaling limits for critical inhomogeneous random graphs with finite third moments**, *Electronic Journal of Probability* **15**, 2010, paper no. 54, pp.1682–1703.]

[S. Dhara, R. van der Hofstad, J. van Leeuwaarden & S. Sen, **Critical window for the configuration model: finite third moment degrees**, `arXiv:1605.02868`, 2016+.]

(c) Critical random transposition random walk

Generate a random permutation of $\{1, 2, \ldots, n\}$ by composing i.i.d. uniform random transpositions. This gives a Markov chain on $\mathfrak{S}_n$ called the random transposition random walk (RTRW).

(c) Critical random transposition random walk

Generate a random permutation of $\{1, 2, \ldots, n\}$ by composing i.i.d. uniform random transpositions. This gives a Markov chain on $\mathfrak{S}_n$ called the random transposition random walk (RTRW).

There is a natural coupling with the Erdős-Rényi random graph process, whereby we include an edge $\{i, j\}$ in the graph iff we have multiplied by the transposition (i, j). We can use our understanding of the graph process to deduce properties of the RTRW.

[O. Schramm, **Compositions of random transpositions**, *Israel Journal of Mathematics* **147**, 2005, pp.221–243.]

[N. Berestycki, **Emergence of giant cycles and slowdown transition in random transpositions and** k**-cycles**, *Electronic Journal of Probability* **16**, paper no. 5, 2011, pp.152–173.]

Critical random transposition random walk

Each component of the graph corresponds to at least one cycle of the permutation. So, for example, there cannot be giant cycles below the point of the phase transition.

Question. What are the lengths of the cycles of the RTRW walk in the critical window?

Critical random transposition random walk

Each component of the graph corresponds to at least one cycle of the permutation. So, for example, there cannot be giant cycles below the point of the phase transition.

Question. What are the lengths of the cycles of the RTRW walk in the critical window?

There are many tree components, each of which corresponds to a single cycle of the permutation. The other components all have finite surplus. Given a graph component, notice that all edge-arrival orders are equally likely.

Critical random transposition random walk

Each component of the graph corresponds to at least one cycle of the permutation. So, for example, there cannot be giant cycles below the point of the phase transition.

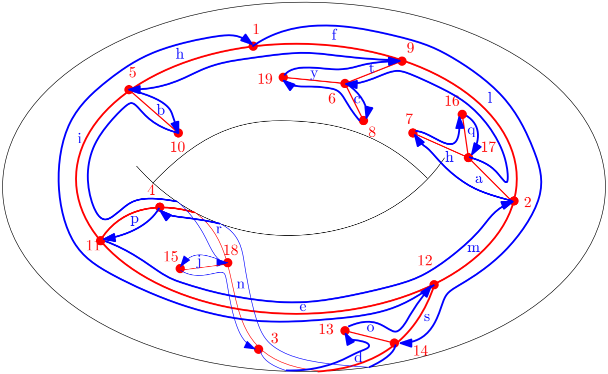
Question. What are the lengths of the cycles of the RTRW walk in the critical window?

There are many tree components, each of which corresponds to a single cycle of the permutation. The other components all have finite surplus. Given a graph component, notice that all edge-arrival orders are equally likely.

So it is sufficient to consider what happens for a uniform random connected graph on n vertices with surplus k and a uniform ordering of the edges. [→ Dominic's talk]

Critical random transposition random walk

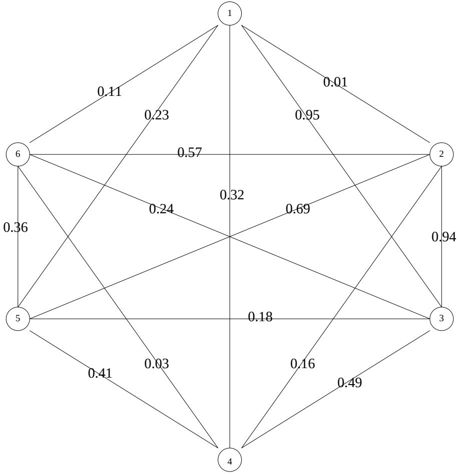
Example: $n = 19$, surplus 2, single permutation cycle.



Within a single such component, the limiting cycle-lengths can then be understood in terms of $(\mathcal{G}^k, d^k, \mu^k)$.

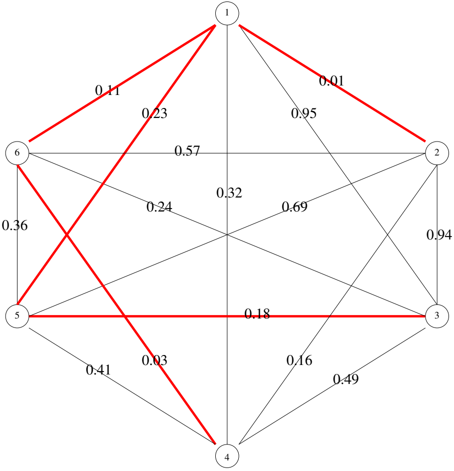
(d) The scaling limit of the minimum spanning tree of the complete graph

Consider the complete graph on n vertices with independent edge-weights which are uniformly distributed on $[0, 1]$.



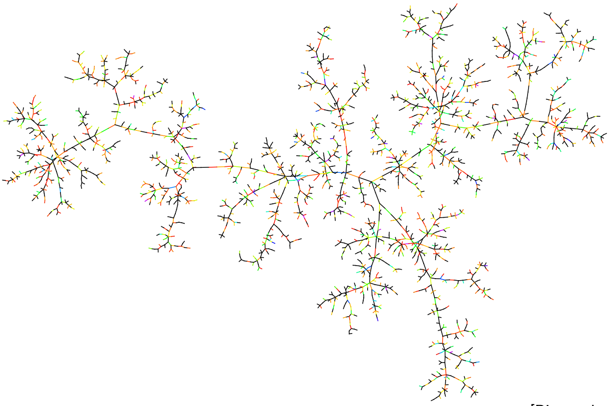
The scaling limit of the minimum spanning tree of the complete graph

Find the minimum spanning tree (MST).



The scaling limit of the minimum spanning tree of the complete graph

Question. Does the MST of the complete graph on n vertices possess a scaling limit?



[Picture by Louigi Addario-Berry]

The scaling limit of the minimum spanning tree of the complete graph

Let M_n be the MST of the complete graph on n vertices, let d_n be its graph distance, and μ_n its uniform measure.

Theorem. (Addario-Berry, Broutin, G. & Miermont (2017))
There exists a random compact measured real tree $(\mathcal{M}, d, \mu)$ such that

$$\left(M_n, \frac{d_n}{n^{1/3}}, \mu_n\right) \xrightarrow{d} (\mathcal{M}, d, \mu)$$

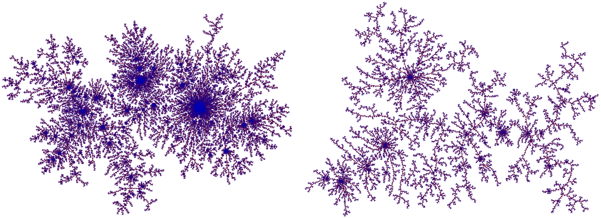
as $n \to \infty$, in GHP. $\mathcal{M}$ is binary and has Minkowski dimension 3 almost surely.

The key to understanding this result is a connection between the Erdős-Rényi random graph and Kruskal's algorithm for constructing the MST.

[L. Addario-Berry, N. Broutin, C. Goldschmidt & G. Miermont, **The scaling limit of the minimum-spanning tree of the complete graph**, *Annals of Probability* **45**(5), 2017, pp.3075–3144.]

(e) The stable trees

Recall that the α-stable tree, for $\alpha \in (1, 2)$, is the scaling limit of a Galton-Watson tree with critical offspring distribution in the domain of attraction of an α-stable law.

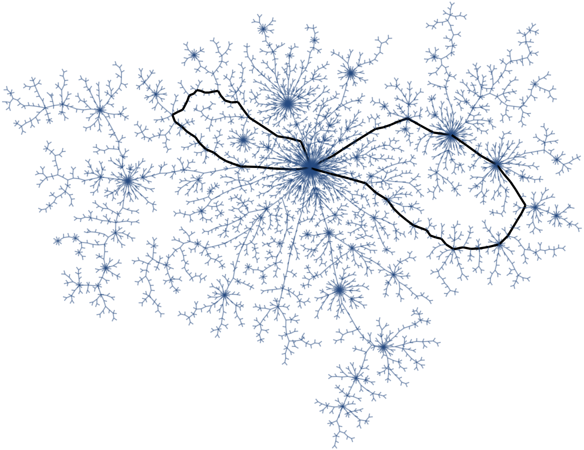


There is an analogue of Rémy's algorithm due to Marchal (2008) and there is also a (more complicated) line-breaking construction.

[P. Marchal, **A note on the fragmentation of a stable tree**, *Fifth Colloquium on Mathematics and Computer Science (DMTCS)*, 2008, pp.489–500.]

[C. Goldschmidt & B. Haas, **A line-breaking construction of the stable trees**, *Electronic Journal of Probability* **20**, 2015, Paper no. 16, pp.1–24.]

The stable graphs



[Picture by Delphin Sénizergues]

The stable graphs

The natural graph model whose scaling limit involves the stable trees is the configuration model with i.i.d. power-law degrees. Work in progress...

[G. Conchon-Kerjan & C. Goldschmidt, **The stable graph: the metric space scaling limit of a critical random graph with i.i.d. power-law degrees**, in preparation]

[C. Goldschmidt, B. Haas and D. Sénizergues, **Stable graphs: distributions and line-breaking construction**, in preparation.]

[A. Joseph, **The component sizes of a critical random graph with given degree sequence**, *Annals of Applied Probability* **24**(6), 2014, pp.2560–2594.]

Related work:
[S. Bhamidi, S. Dhara, R. van der Hofstad & S. Sen, **Universality for critical heavy-tailed network models: metric structure of maximal components**, `arXiv:1703.07145`, 2017.]

[S. Bhamidi, R. van der Hofstad & J. van Leeuwaarden, **Novel scaling limits for critical inhomogeneous random graphs**, *Annals of Probabilty* **40**(6), 2012, pp.2299–2361.]

[S. Dhara, R. van der Hofstad, J. van Leeuwaarden & S. Sen, **Heavy-tailed configuration models at criticality**, `arXiv:1612.00650`, 2016+.]

Thank you!