5

Root Growth with Re-Grafting

5.1 Background and Motivation

Recall the special case of the tree-valued Markov chain that was used in the proof of the Markov chain tree theorem, Theorem 2.1, when the underlying Markov chain is the process on $\{1, 2, \ldots, n\}$ that picks a new state uniformly at each stage.

Algorithm 5.1.

- *Start with a rooted (combinatorial) tree on n labeled vertices $\{1, 2, \ldots, n\}$.*
- *Pick a vertex v uniformly from $\{1, 2, \ldots, n\} \backslash \{current\ root\}$.*
- *Erase the edge leading from v towards the current root.*
- *Insert an edge from the current root to v and make v the new root.*
- *Repeat.*

We know that this chain converges in distribution to the uniform distribution on rooted trees with n labeled vertices.

Imagine that we do the following.

- Start with a rooted subtree (that is, one with the same root as the "big" tree).
- At each step of the chain, update the subtree by removing and adding edges as they are removed and added in the big tree and adjoining the new root of the big tree to the subtree if it isn't in the current subtree.

The subtree will evolve via two mechanisms that we might call *root growth* and *re-grafting*. Root growth occurs when the new root isn't in the current subtree, and so the new tree has an extra vertex, the new root, that is connected to the old root by a new edge. Re-grafting occurs when the new root is in the current subtree: it has the effect of severing the edge leading to a subtree of the current subtree and re-attaching it to the current root by a new edge. See Figure 5.1.

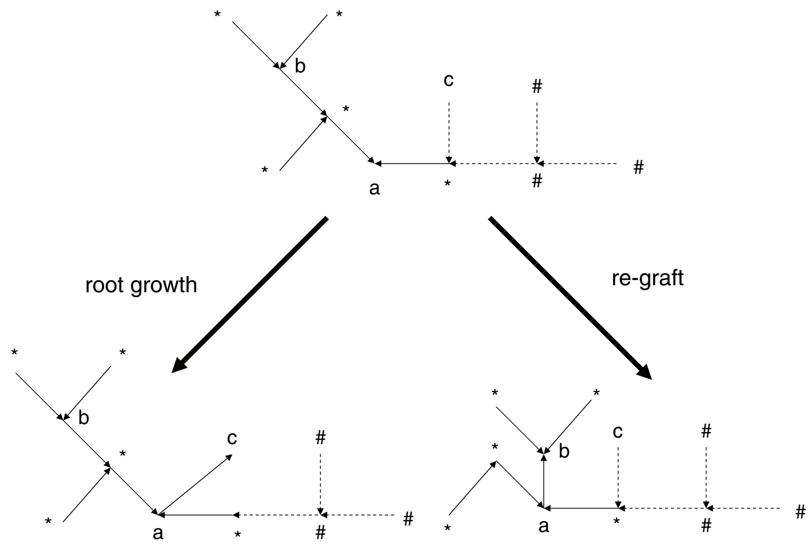

Fig. 5.1. Root growth and re-graft moves. The big tree with $n = 11$ vertices consists of the solid and dashed edges in all three diagrams. In the top diagram, the current subtree has the solid edges and the vertices marked a, b, $*$. The vertices marked c and $\#$ are in the big tree but not the current subtree. The big tree and the current subtree are rooted at a. The bottom left diagram shows the result of a root growth move: the vertex c now belongs to the new subtree, it is the root of the new big tree and the new subtree, and is connected to the old root a by an edge. The vertices marked $\#$ are not in the new subtree. The bottom right diagram shows the result of a re-graft move: the vertex b is the root of the new big tree and the new subtree, and it is connected to the old root a by an edge. The vertices marked c and $\#$ are not in the new subtree.

Now consider what happens as n becomes large and we follow a rooted subtree that originally has $\approx \sqrt{n}$ vertices. Replace edges of length 1 with edges of length $\frac{1}{\sqrt{n}}$ and speed up time by $\sqrt{n}$.

In the limit as $n \to \infty$, it seems reasonable that we have a $\mathbb{R}$-tree-valued process with the following *root growth with re-grafting* dynamics.

- The edge leading to the root of the evolving tree grows at unit speed.
- Cuts rain down on the tree at unit rate per length$\times$time, and the subtree above each cut is pruned off and re-attached at the root.

We will establish a closely related result in Section 5.4. Namely, we will show that if we have a sequence of chains following the dynamics of

Algorithm 5.1 such that the initial combinatorial tree of the n^{th} chain re-scaled by $\sqrt{n}$ converges in the Gromov–Hausdorff distance to some compact $\mathbb{R}$-tree, then if we re-scale space and time by $\sqrt{n}$ in the n^{th} chain we get weak convergence to a process with the root growth with re-grafting dynamics.

This latter result might seem counter-intuitive, because now we are working with the whole tree with n vertices rather than a subtree with $\approx \sqrt{n}$ vertices. However, the assumption that the initial condition scaled by $\sqrt{n}$ converges to some compact $\mathbb{R}$-tree means that asymptotically most vertices are close to the leaves and re-arranging the subtrees above such vertices has a negligible effect in the limit.

Before we can establish such a convergence result, we need to show that the root growth with re-grafting dynamics make sense even for compact trees with infinite total length. Such trees are the sort that will typically arise in the limit when we re-scale trees with n vertices by $\sqrt{n}$. This is not a trivial matter, as the set of times at which cuts appear will be dense and so the intuitive description of the dynamics does not make rigorous sense. See Theorem 5.5 for the details.

Given that the chain of Algorithm 5.1 converges at large times to the uniform rooted tree on n labeled vertices and that the uniform tree on n labeled vertices converges after suitable re-scaling to the Brownian continuum random tree as $n \to \infty$, it seems reasonable that the root growth with re-grafting process should converge at large times to the Brownian continuum random tree and that the Brownian continuum random tree should be the unique stationary distribution. We establish that this is indeed the case in Section 5.3. An important ingredient in the proofs of these facts will be Proposition 5.7, which says that the root growth with re-grafting process started from the trivial tree consisting of a single point is related to the Poisson line-breaking construction of the Brownian continuum random tree in Section 2.5 in the same manner that the chain of Algorithm 5.1 is related to Algorithm 2.4 for generating uniform rooted labeled trees. This is, of course, what we should expect, because the Poisson line-breaking construction arises as a limit of Algorithm 2.4 when the number of vertices goes to infinity.

5.2 Construction of the Root Growth with Re-Grafting Process

5.2.1 Outline of the Construction

- We want to construct a $\mathbf{T}^{\text{root}}$-valued process X with the root growth and re-grafting dynamics.
- Fix $(T, d, \rho) \in \mathbf{T}^{\text{root}}$. This will be X_0.
- We will construct simultaneously for each finite rooted subtree $T^* \leq^{\text{root}} T$ a process X^{T^*} with $X_0^{T^*} = T^*$ that evolves according to the root growth with re-grafting dynamics.

- We will carry out this construction in such a way that if T^* and T^{**} are two finite subtrees with $T^* \preceq^{\text{root}} T^{**}$, then $X_t^{T^*} \preceq^{\text{root}} X_t^{T^{**}}$ and the cut points for X^{T^*} are those for $X^{T^{**}}$ that happen to fall on $X_{\tau-}^{T^*}$ for a corresponding cut time τ of $X^{T^{**}}$. Cut times τ for $X^{T^{**}}$ for which the corresponding cut point does not fall on $X_{\tau-}^{T^*}$ are not cut times for X^{T^*}.
- The tree (T, ρ) is a rooted Gromov–Hausdorff limit of finite $\mathbb{R}$-trees with root ρ (indeed, any subtree of (T, ρ) that is spanned by the union of a finite ε-net and $\{\rho\}$ is a finite $\mathbb{R}$-tree that has rooted Gromov–Hausdorff distance less than ε from (T, ρ)).
  In particular, (T, ρ) is the "smallest" rooted compact $\mathbb{R}$-tree that contains all of the finite rooted subtrees of (T, ρ).
- Because of the consistent projective nature of the construction, we can define $X_t := X_t^T$ for $t \geqslant 0$ as the "smallest" element of $\mathbf{T}^{\text{root}}$ that contains $X_t^{T^*}$, for all finite trees $T^* \preceq^{\text{root}} T$.

5.2.2 A Deterministic Construction

It will be convenient to work initially in a setting where the cut times and cut points are fixed.

There are two types of cut points: those that occur at points that were present in the initial tree T and those that occur at points that were added due to subsequent root growth.

Accordingly, we consider two countable subsets $\pi_0 \subset \mathbb{R}^{++} \times T^o$ and $\pi \subset \{(t, x) \in \mathbb{R}^{++} \times \mathbb{R}^{++} : x \leqslant t\}$. See Figure 5.2.

Assumption 5.2. *Suppose that the sets π_0 and π have the following properties.*

(a) For all $t_0 > 0$, each of the sets $\pi_0 \cap (\{t_0\} \times T^o)$ and $\pi \cap (\{t_0\} \times]0, t_0])$ has at most one point and at least one of these sets is empty.
(b) For all $t_0 > 0$ and all finite subtrees $T' \subseteq T$, the set $\pi_0 \cap (]0, t_0] \times T')$ is finite.
(c) For all $t_0 > 0$, the set $\pi \cap \{(t, x) \in \mathbb{R}^{++} \times \mathbb{R}^{++} : x \leqslant t \leqslant t_0\}$ is finite.

Remark 5.3. Conditions (a)–(c) of Assumption 5.2 will hold almost surely if π_0 and π are realizations of Poisson point processes with respective intensities $\lambda \otimes \mu$ and $\lambda \otimes \lambda$ (where λ is Lebesgue measure), and it is this random mechanism that we will introduce later to produce a stochastic process having the root growth with re-grafting dynamics.

Consider a finite rooted subtree $T^* \preceq^{\text{root}} T$. It will avoid annoying circumlocutions about equivalence via root-invariant isometries if we work with particular class representatives for T^* and T, and, moreover, suppose that T^* is embedded in T.

Put $\tau_0^* := 0$, and let $0 < \tau_1^* < \tau_2^* < \ldots$ (the cut times for X^{T^*}) be the points of $\{t > 0 : \pi_0(\{t\} \times T^*) > 0\} \cup \{t > 0 : \pi(\{t\} \times \mathbb{R}^{++}) > 0\}$.

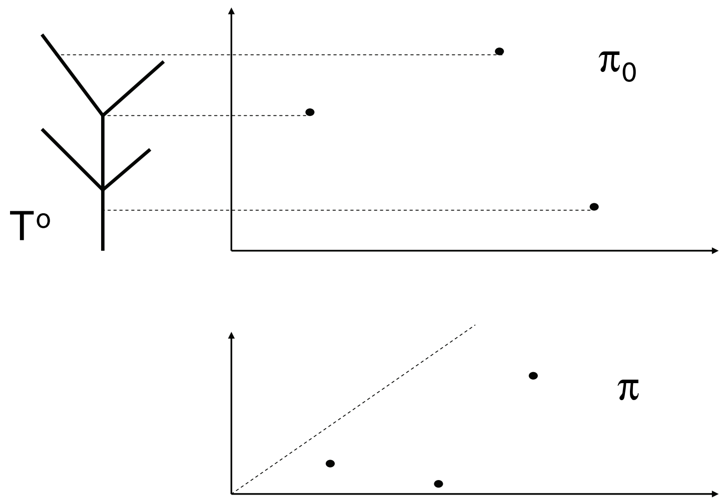

Fig. 5.2. The sets of points π_0 and π

Step 1 (Root growth). At any time $t \geqslant 0$, $X_t^{T^*}$ as a set is given by the disjoint union $T^* \amalg]0,t]$. For $t > 0$, the root of $X_t^{T^*}$ is the point $\rho_t := t \in]0,t]$. The metric $d_t^{T^*}$ on $X_t^{T^*}$ is defined inductively as follows.

Set $d_0^{T^*}$ to be the metric on $X_0^{T^*} = T^*$; that is, $d_0^{T^*}$ is the restriction of d to T^*. Suppose that $d_t^{T^*}$ has been defined for $0 \leqslant t \leqslant \tau_n^*$. Define $d_t^{T^*}$ for $\tau_n^* < t < \tau_{n+1}^*$ by

$$d_t^{T^*}(a,b) := \begin{cases} d_{\tau_n^*}(a,b), & \text{if } a,b \in X_{\tau_n^*}^{T^*}, \\ |b-a|, & \text{if } a,b \in]\tau_n^*, t], \\ |a-\tau_n^*| + d_{\tau_n^*}(\rho_{\tau_n^*}, b), & \text{if } a \in]\tau_n^*, t],\ b \in X_{\tau_n^*}^{T^*}. \end{cases} \tag{5.1}$$

Step 2 (Re-Grafting). Note that the left-limit $X_{\tau_{n+1}^*-}^{T^*}$ exists in the rooted Gromov–Hausdorff metric. As a set this left-limit is the disjoint union

$$X_{\tau_n^*}^{T^*} \amalg]\tau_n^*, \tau_{n+1}^*] = T^* \amalg]0, \tau_{n+1}^*],$$

and the corresponding metric $d_{\tau_{n+1}^*-}$ is given by a prescription similar to (5.1).

Define the $(n+1)^{\text{st}}$ cut point for X^{T^*} by

$$p^*_{n+1} := \begin{cases} a \in T^*, & \text{if } \pi_0(\{(\tau^*_{n+1}, a)\}) > 0, \\ x \in]0, \tau^*_{n+1}], & \text{if } \pi(\{(\tau^*_{n+1}, x)\}) > 0. \end{cases}$$

Let S^*_{n+1} be the *subtree above* p^*_{n+1} in $X^{T^*}_{\tau^*_{n+1}-}$, that is,

$$S^*_{n+1} := \{b \in X^{T^*}_{\tau^*_{n+1}-} : p^*_{n+1} \in [\rho_{\tau^*_{n+1}-}, b[\,\}. \tag{5.2}$$

Define the metric $d_{\tau^*_{n+1}}$ by

$$\begin{aligned} &d_{\tau^*_{n+1}}(a,b) \\ &:= \begin{cases} d_{\tau^*_{n+1}-}(a,b), & \text{if } a, b \in S^*_{n+1}, \\ d_{\tau^*_{n+1}-}(a,b), & \text{if } a, b \in X^{T^*}_{\tau^*_{n+1}} \backslash S^*_{n+1}, \\ d_{\tau^*_{n+1}-}(a, \rho_{\tau^*_{n+1}}) + d_{\tau^*_{n+1}-}(p^*_{n+1}, b), & \text{if } a \in X^{T^*}_{\tau^*_{n+1}} \backslash S^*_{n+1},\ b \in S^*_{n+1}. \end{cases} \end{aligned}$$

In other words $X^{T^*}_{\tau^*_{n+1}}$ is obtained from $X^{T^*}_{\tau^*_{n+1}-}$ by pruning off the subtree S^*_{n+1} and re-attaching it to the root. See Figure 5.3.

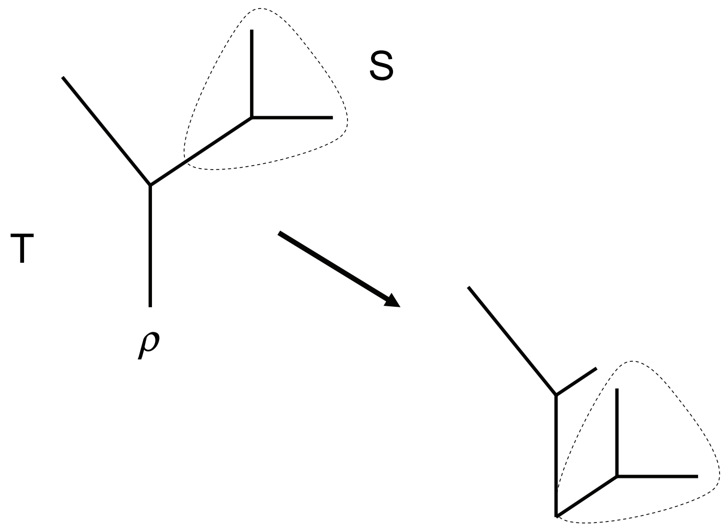

Fig. 5.3. Pruning off the subtree S and regrafting it at the root ρ

Now consider two other finite, rooted subtrees (T^{**}, ρ) and (T^{***}, ρ) of T such that $T^* \cup T^{**} \subseteq T^{***}$ (with induced metrics).

Build $X^{T^{**}}$ and $X^{T^{***}}$ from π_0 and π in the same manner as X^{T^*} (but starting at T^{**} and T^{***}). It is clear from the construction that:

- $X_t^{T^*}$ and $X_t^{T^{**}}$ are rooted subtrees of $X_t^{T^{***}}$ for all $t \geqslant 0$,
- the Hausdorff distance between $X_t^{T^*}$ and $X_t^{T^{**}}$ as subsets of $X_t^{T^{***}}$ does not depend on T^{***},
- the Hausdorff distance is constant between jumps of X^{T^*} and $X^{T^{**}}$ (when only root growth is occurring in both processes).

The following lemma shows that the Hausdorff distance between $X_t^{T^*}$ and $X_t^{T^{**}}$ as subsets of $X_t^{T^{***}}$ does not increase at jump times.

Lemma 5.4. *Let T be a finite rooted tree with root ρ and metric d, and let T' and T'' be two rooted subtrees of T (both with the induced metrics and root ρ). Fix $p \in T$, and let S be the subtree in T above p (recall (5.2)). Define a new metric $\hat{d}$ on T by putting*

$$\hat{d}(a,b) := \begin{cases} d(a,b), & \text{if } a, b \in S, \\ d(a,b), & \text{if } a, b \in T \backslash S, \\ d(a,p) + d(\rho, b), & \text{if } a \in S,\ b \in T \backslash S. \end{cases}$$

Then the sets T' and T'' are also subtrees of T equipped with the induced metric $\hat{d}$, and the Hausdorff distance between T' and T'' with respect to $\hat{d}$ is not greater than that with respect to d.

Proof. Suppose that the Hausdorff distance between T' and T'' under d is less than some given $\varepsilon > 0$. Given $a \in T'$, there then exists $b \in T''$ such that $d(a,b) < \varepsilon$. Because $d(a, a \wedge b) \leqslant d(a,b)$ and $a \wedge b \in T''$, we may suppose (by replacing b by $a \wedge b$ if necessary) that $b \leqslant a$.

We claim that $\hat{d}(a,c) < \varepsilon$ for some $c \in T''$. This and the analogous result with the roles of T' and T'' interchanged will establish the result.

If $a, b \in S$ or $a, b \in T\backslash S$, then $\hat{d}(a,b) = d(a,b) < \varepsilon$. The only other possibility is that $a \in S$ and $b \in T\backslash S$, in which case $p \in [b,a]$ (for T equipped with d). Then $\hat{d}(a,\rho) = d(a,p) \leqslant d(a,b) < \varepsilon$, as required (because $\rho \in T''$). □

Now let $T_1 \subseteq T_2 \subseteq \cdots$ be an increasing sequence of finite subtrees of T such that $\bigcup_{n \in \mathbb{N}} T_n$ is dense in T. Thus, $\lim_{n\to\infty} d_{\mathrm{H}}(T_n, T) = 0$.

Let $X^1, X^2, \ldots$ be constructed from π_0 and π starting with $T_1, T_2, \ldots$. Applying Lemma 5.4 yields

$$\lim_{m,n\to\infty} \sup_{t \geqslant 0} d_{\mathrm{GH^{root}}}(X_t^m, X_t^n) = 0.$$

Hence, by completeness of $\mathbf{T}^{\mathrm{root}}$, there exists a càdlàg $\mathbf{T}^{\mathrm{root}}$-valued process X such that $X_0 = T$ and

$$\lim_{m\to\infty}\sup_{t\geqslant 0} d_{\mathrm{GH}^{\mathrm{root}}}(X_t^m, X_t) = 0.$$

A priori, the process X could depend on the choice of the approximating sequence of trees $\{T_n\}_{n\in\mathbb{N}}$. To see that this is not so, consider two approximating sequences $T_1^1 \subseteq T_2^1 \subseteq \cdots$ and $T_1^2 \subseteq T_2^2 \subseteq \cdots$.

For $k \in \mathbb{N}$, write T_n^3 for the smallest rooted subtree of T that contains both T_n^1 and T_n^2. As a set, $T_n^3 = T_n^1 \cup T_n^2$. Now let $\{(X_t^{n,i}\}_{t\geqslant 0})_{n\in\mathbb{N}}$ for $i = 1, 2, 3$ be the corresponding sequences of finite tree-value processes and let $(X_t^{\infty,i})_{t\geqslant 0}$ for $i = 1, 2, 3$ be the corresponding limit processes. By Lemma 5.4,

$$\begin{aligned}
d_{\mathrm{GH}^{\mathrm{root}}}(X_t^{n,1}, X_t^{n,2}) &\leqslant d_{\mathrm{GH}^{\mathrm{root}}}(X_t^{n,1}, X_t^{n,3}) + d_{\mathrm{GH}^{\mathrm{root}}}(X_t^{n,2}, X_t^{n,3}) \\
&\leqslant d_{\mathrm{H}}(X_t^{n,1}, X_t^{n,3}) + d_{\mathrm{H}}(X_t^{n,2}, X_t^{n,3}) \\
&\leqslant d_{\mathrm{H}}(T_n^1, T_n^3) + d_{\mathrm{H}}(T_n^2, T_n^3) \\
&\leqslant d_{\mathrm{H}}(T_n^1, T) + d_{\mathrm{H}}(T_n^2, T) \to 0
\end{aligned} \tag{5.3}$$

as $n \to \infty$.

Thus, for each $t \geqslant 0$ the sequences $\{X_t^{n,1}\}_{n\in\mathbb{N}}$ and $\{X_t^{n,2}\}_{n\in\mathbb{N}}$ do indeed have the same rooted Gromov–Hausdorff limit and the process X does not depend on the choice of approximating sequence for the initial tree T.

5.2.3 Putting Randomness into the Construction

We constructed a $\mathbf{T}^{\mathrm{root}}$-valued function $t \mapsto X_t$ starting with a fixed triple (T, π_0, π), where $T \in \mathbf{T}^{\mathrm{root}}$ and π_0, π satisfy the conditions of Assumption 5.2. We now want to think of X as a function of time and such triples.

Let Ω^* be the set of triples (T, π_0, π), where T is a rooted compact $\mathbb{R}$-tree (that is, a class representative of an element of $\mathbf{T}^{\mathrm{root}}$) and π_0, π satisfy Assumption 5.2.

The root invariant isometry equivalence relation on rooted compact $\mathbb{R}$-trees extends naturally to an equivalence relation on Ω^* by declaring that two triples (T', π_0', π') and (T'', π_0'', π''), where $\pi_0' = \{(\sigma_i', x_i') : i \in \mathbb{N}\}$ and $\pi_0'' = \{(\sigma_i'', x_i'') : i \in \mathbb{N}\}$, are equivalent if there is a root invariant isometry f mapping T' to T'' and a permutation γ of $\mathbb{N}$ such that $\sigma_i'' = \sigma_{\gamma(i)}'$ and $x_i'' = f(x_{\gamma(i)}')$ for all $i \in \mathbb{N}$. Write Ω for the resulting quotient space of equivalence classes. There is a natural measurable structure on Ω: we refer to [63] for the details.

Given $T \in \mathbf{T}^{\mathrm{root}}$, let $\mathbf{P}^T$ be the probability measure on Ω defined by the following requirements.

- The measure $\mathbf{P}^T$ assigns all of its mass to the set $\{(T', \pi_0', \pi') \in \Omega : T' = T\}$.
- Under $\mathbf{P}^T$, the random variable $(T', \pi_0', \pi') \mapsto \pi_0'$ is a Poisson point process on the set $\mathbb{R}^{++} \times T^o$ with intensity $\lambda \otimes \mu$, where μ is the length measure on T.

- Under $\mathbf{P}^T$, the random variable $(T', \pi_0', \pi') \mapsto \pi'$ is a Poisson point process on the set $\{(t,x) \in \mathbb{R}^{++} \times \mathbb{R}^{++} : x \leqslant t\}$ with intensity $\lambda \otimes \lambda$ restricted to this set.
- The random variables $(T', \pi_0', \pi') \mapsto \pi_0'$ and $(T', \pi_0', \pi') \mapsto \pi'$ are independent under $\mathbf{P}^T$.

Of course, the random variable $(T', \pi_0', \pi') \mapsto \pi_0'$ takes values in a space of equivalence classes of countable sets rather than a space of sets *per se*, so, more formally, this random variable has the law of the image of a Poisson process on an arbitrary class representative under the appropriate quotient map.

For $t \geqslant 0$, g a bounded Borel function on $\mathbf{T}^{\text{root}}$, and $T \in \mathbf{T}^{\text{root}}$, set

$$P_t g(T) := \mathbf{P}^T[g(X_t)]. \tag{5.4}$$

With a slight abuse of notation, let $\tilde{R}_\eta$ for $\eta > 0$ also denote the map from Ω into Ω that sends (T, π_0, π) to $(R_\eta(T), \pi_0 \cap (\mathbb{R}^{++} \times (R_\eta(T))^o), \pi)$.

Theorem 5.5. *(i) If $T \in \mathbf{T}^{\text{root}}$ is finite, then $(X_t)_{t\geqslant 0}$ under $\mathbf{P}^T$ is a Markov process that evolves via the root growth with re-grafting dynamics on finite trees.*

(ii) For all $\eta > 0$ and $T \in \mathbf{T}^{\text{root}}$, the law of $(X_t \circ \tilde{R}_\eta)_{t\geqslant 0}$ under $\mathbf{P}^T$ coincides with the law of $(X_t)_{t\geqslant 0}$ under $\mathbf{P}^{R_\eta(T)}$.

(iii) For all $T \in \mathbf{T}^{\text{root}}$, the law of $(X_t)_{t\geqslant 0}$ under $\mathbf{P}^{R_\eta(T)}$ converges as $\eta \downarrow 0$ to that of $(X_t)_{t\geqslant 0}$ under $\mathbf{P}^T$ (in the sense of convergence of laws on the space of càdlàg $\mathbf{T}^{\text{root}}$-valued paths equipped with the Skorohod topology).

(iv) For $g \in \mathrm{b}\mathcal{B}(\mathbf{T}^{\text{root}})$, the map $(t,T) \mapsto P_t g(T)$ is $\mathcal{B}(\mathbb{R}^+) \times \mathcal{B}(\mathbf{T}^{\text{root}})$-measurable.

(v) The process $(X_t, \mathbf{P}^T)$ is strong Markov and has transition semigroup $(P_t)_{t\geqslant 0}$.

Proof. (i) This is clear from the definition of the root growth and re-grafting dynamics.

(ii) It is enough to check that the push-forward of the probability measure $\mathbf{P}^T$ under the map $R_\eta : \Omega \to \Omega$ is the measure $\mathbf{P}^{R_\eta(T)}$.

This, however, follows from the observation that the restriction of length measure on a tree to a subtree is just length measure on the subtree.

(iii) This is immediate from part (ii) and part (iv) of Lemma 4.32. Indeed, we have that

$$\sup_{t\geqslant 0} d_{\mathrm{GH}^{\text{root}}}(X_t, X_t \circ \tilde{R}_\eta) \leqslant d_{\mathrm{H}}(T, R_\eta(T)) \leqslant \eta.$$

(iv) By a monotone class argument, it is enough to consider the case where the test function g is continuous. It follows from part (iii) that $P_t g(R_\eta(T))$ converges pointwise to $P_t g(T)$ as $\eta \downarrow 0$, and it is not difficult to show using Lemma 4.32 and part (i) that $(t,T) \mapsto P_t g(R_\eta(T))$ is $\mathcal{B}(\mathbb{R}^+) \times \mathcal{B}(\mathbf{T}^{\text{root}})$-measurable, but we omit the details.

(v) By construction and Lemma 4.33, we have for $t \geqslant 0$ and $(T, \pi_0, \pi) \in \Omega$ that, as a set, $X_t^o(T, \pi_0, \pi)$ is the disjoint union $T^o \amalg]0, t]$.

Put

$$\begin{aligned}
&\theta_t(T, \pi_0, \pi) \\
&\quad := \Big(X_t(T, \pi_0, \pi), \{(s, x) \in \mathbb{R}^{++} \times T^o : (t + s, x) \in \pi_0\}, \\
&\qquad \{(s, x) \in \mathbb{R}^{++} \times \mathbb{R}^{++} : (t + s, t + x) \in \pi\}\Big) \\
&\quad = \Big(X_t(T, \pi_0, \pi), \{(s, x) \in \mathbb{R}^{++} \times X_t^o(T, \pi_0, \pi) : (t + s, x) \in \pi_0\}, \\
&\qquad \{(s, x) \in \mathbb{R}^{++} \times \mathbb{R}^{++} : (t + s, t + x) \in \pi\}\Big).
\end{aligned}$$

Thus, θ_t maps Ω into Ω. Note that $X_s \circ \theta_t = X_{s+t}$ and that $\theta_s \circ \theta_t = \theta_{s+t}$, that is, the family $(\theta_t)_{t \geqslant 0}$ is a semigroup.

Fix $t \geqslant 0$ and $(T, \pi_0, \pi) \in \Omega$. Write μ' for the measure on $T^o \amalg]0, t]$ that restricts to length measure on T^o and to Lebesgue measure on $]0, t]$. Write μ'' for the length measure on $X_t^o(T, \pi_0, \pi)$.

The strong Markov property will follow from a standard strong Markov property for Poisson processes if we can show that $\mu' = \mu''$.

This equality is clear from the construction if T is finite: the tree $X_t(T, \pi_0, \pi)$ is produced from the tree T and the set $]0, t]$ by a finite number of dissections and rearrangements.

The equality for general T follows from the construction and Lemma 4.33. □

5.2.4 Feller Property

The proof of Theorem 5.5 depended on an argument that showed that if we have two finite subtrees of a given tree that are close in the Gromov–Hausdorff distance, then the resulting root growth with re-grafting processes can be coupled together on the same probability space so that they stay close together. It is believable that if we start the root growth with re-grafting process with any two trees that are close together (whether or not they are finite or subtrees of of a common tree), then the resulting processes will be close in some sense. The following result, which implies that the measure induced by the root growth with re-grafting process on path space is weakly continuous in the starting state with respect to the Skorohod topology on path space can be established by a considerably more intricate coupling argument: we refer to [63] for the details.

Proposition 5.6. *If the function $f : \mathbf{T}^{\mathrm{root}} \to \mathbb{R}$ is continuous and bounded, then the function $P_t f$ is also continuous and bounded for each $t \geqslant 0$.*

5.3 Ergodicity, Recurrence, and Uniqueness

5.3.1 Brownian CRT and Root Growth with Re-Grafting

Recall that Algorithm 2.4 for generating uniform rooted tree on n labeled vertices was derived from Algorithm 5.1, the tree-valued Markov chain appearing in the proof of the Markov chain tree theorem that has the uniform rooted tree on n labeled vertices as its stationary distribution. Recall also that the Poisson line-breaking construction of the Brownian continuum random tree in Section 2.5 is an asymptotic version of Algorithm 2.4, whilst the root growth with re-grafting process was motivated as an asymptotic version of Algorithm 5.1. Therefore, it seems reasonable that there should be a connection between the Poisson line-breaking construction and the root growth with re-grafting process. We establish the connection in this subsection.

Let us first present the Poisson line-breaking construction in a more "dynamic" way that will make the comparison with the root growth with regrafting process a little more transparent.

- Write $\tau_1, \tau_2, \ldots$ for the successive arrival times of an inhomogeneous Poisson process with arrival rate t at time $t \geqslant 0$. Call τ_n the n^{th} *cut time*.
- Start at time 0 with the 1-tree (that is a line segment with two ends), $\mathcal{R}_0$, of length zero ($\mathcal{R}_0$ is "really" the *trivial tree* that consists of one point only, but thinking this way helps visualize the dynamics more clearly for this semi-formal description). Identify one end of $\mathcal{R}_0$ as the root.
- Let this line segment grow at unit speed until the first cut time τ_1.
- At time τ_1 pick a point uniformly on the segment that has been grown so far. Call this point the *first cut point*.
- Between time τ_1 and time τ_2, evolve a tree with 3 ends by letting a new branch growing away from the first cut point at unit speed.
- Proceed inductively: Given the n-tree (that is, a tree with $n + 1$ ends), $\mathcal{R}_{\tau_n-}$, pick the *n-th cut point* uniformly on $\mathcal{R}_{\tau_n-}$ to give an $n + 1$-tree, $\mathcal{R}_{\tau_n}$, with one edge of length zero, and for $t \in [\tau_n, \tau_{n+1}[$, let $\mathcal{R}_t$ be the tree obtained from $\mathcal{R}_{\tau_n}$ by letting a branch grow away from the n^{th} cut point with unit speed.

The tree $\mathcal{R}_{\tau_n-}$ is n^{th} step of the Poisson line-breaking construction, and the Brownian CRT is the limit of the increasing family of rooted finite trees $(\mathcal{R}_t)_{t\geqslant 0}$.

We will now use the ingredients appearing in the construction of $\mathcal{R}$ to construct a version of the root growth with re-grafting process started at the trivial tree.

- Let $\tau_1, \tau_2, \ldots$ be as in the construction of the $\mathcal{R}$.
- Start with the 1-tree (with one end identified as the *root* and the other as a *leaf*), $\mathcal{T}_0$, of length zero.

- Let this segment grow at unit speed on the time interval $[0, \tau_1[$, and for $t \in [0, \tau_1[$ let $\mathcal{T}_t$ be the rooted 1-tree that has its points labeled by the interval $[0, t]$ in such a way that the root is t and the leaf is 0.
- At time τ_1 sample the *first cut point* uniformly along the tree $\mathcal{T}_{\tau_1-}$, prune off the piece of $\mathcal{T}_{\tau_1-}$ that is above the cut point (that is, prune off the interval of points that are further away from the root t than the first cut point).
- Re-graft the pruned segment such that its cut end and the root are glued together. Just as we thought of $\mathcal{T}_0$ as a tree with two points, (a leaf and a root) connected by an edge of length zero, we take $\mathcal{T}_{\tau_1}$ to be the the rooted 2-tree obtained by "ramifying" the root $\mathcal{T}_{\tau_1-}$ into two points (one of which we keep as the root) that are joined by an edge of length zero.
- Proceed inductively: Given the labeled and rooted n-tree, $\mathcal{T}_{\tau_{n-1}}$, for $t \in [\tau_{n-1}, \tau_n[$, let $\mathcal{T}_t$ be obtained by letting the edge containing the root grow at unit speed so that the points in $\mathcal{T}_t$ correspond to the points in the interval $[0, t]$ with t as the root. At time τ_n, the n^{th} cut point is sampled randomly along the edges of the n-tree, $\mathcal{T}_{\tau_n-}$, and the subtree above the cut point (that is the subtree of points further away from the root than the cut point) is pruned off and re-grafted so that its cut end and the root are glued together. The root is then "ramified" as above to give an edge of length zero leading from the root to the rest of the tree.

Let $(\mathcal{R}_t)_{t\geqslant 0}$, $(\mathcal{T}_t)_{t\geqslant 0}$, and $\{\tau_n\}_{n\in\mathbb{N}}$ be as above. Note that $(\mathcal{T}_t)_{t\geqslant 0}$ has the same law as $(X_t)_{t\geqslant 0}$ under $\mathbf{P}^{T_0}$, where T_0 is the trivial tree.

Proposition 5.7. *The two random finite rooted trees $\mathcal{R}_{\tau_n-}$ and $\mathcal{T}_{\tau_n-}$ have the same distribution for all $n \in \mathbb{N}$.*

Proof. Let R_n denote the object obtained by taking the rooted finite tree with edge lengths $\mathcal{R}_{\tau_n-}$ and labeling the leaves with $1, \ldots, n$, in the order that they are added in Aldous's construction. Let T_n be derived similarly from the rooted finite tree with edge lengths $\mathcal{T}_{\tau_n-}$, by labeling the leaves with $1, \ldots, n$ in the order that they appear in the root growth with re-grafting construction. It will suffice to show that R_n and T_n have the same distribution. Note that both R_n and T_n are rooted bifurcating trees with n labeled leaves and edge lengths. Such a tree S_n is uniquely specified by its *shape*, denoted $\text{shape}(S_n)$, that is a rooted, bifurcating, leaf-labeled combinatorial tree, and by the list of its $(2n-1)$ edge lengths in a canonical order determined by its shape, say

$$\text{lengths}(S_n) := (\text{length}(S_n, 1), \ldots, \text{length}(S_n, 2n-1)),$$

where the edge lengths are listed in order of traversal of edges by first working along the path from the root to leaf 1, then along the path joining that path to leaf 2, and so on.

Recall that τ_n is the nth point of a Poisson process on $\mathbb{R}^{++}$ with rate $t\,dt$. We construct R_n and T_n on the same probability space using cuts at points $U_i\tau_i$, $1 \leqslant i \leqslant n-1$, where $U_1, U_2, \ldots$ is a sequence of independent random variables uniformly distributed on the interval $]0,1]$ and independent of the sequence $\{\tau_n\}_{n\in\mathbb{N}}$. Then, by construction, the common collection of edge lengths of R_n and of T_n is the collection of lengths of the $2n-1$ subintervals of $]0,\tau_n]$ obtained by cutting this interval at the $2n-2$ points

$$\{X_i^{(n)} : 1 \leqslant i \leqslant 2n-2\} := \bigcup_{i=1}^{n-1} \{U_i\tau_i, \tau_i\}$$

where the $X_i^{(n)}$ are indexed to increase in i for each fixed n. Let $X_0^{(n)} := 0$ and $X_{2n-1}^{(n)} := \tau_n$. Then

$$\text{length}(R_n, i) = X_i^{(n)} - X_{i-1}^{(n)}, \quad 1 \leqslant i \leqslant 2n-1, \tag{5.5}$$

$$\text{length}(T_n, i) = \text{length}(R_n, \sigma_{n,i}), \quad 1 \leqslant i \leqslant 2n-1, \tag{5.6}$$

for some almost surely unique random indices $\sigma_{n,i} \in \{1, \ldots 2n-1\}$ such that $i \mapsto \sigma_{n,i}$ is almost surely a permutation of $\{1, \ldots 2n-1\}$. According to [10, Lemma 21], the distribution of R_n may be characterized as follows:

(i) the sequence lengths(R_n) is exchangeable, with the same distribution as the sequence of lengths of subintervals obtained by cutting $]0,\tau_n]$ at $2n-2$ uniformly chosen points $\{U_i\tau_n : 1 \leqslant i \leqslant 2n-2\}$;
(ii) shape(R_n) is uniformly distributed on the set of all $1\times 3\times 5\times\cdots\times(2n-3)$ possible shapes;
(iii) lengths(R_n) and shape(R_n) are independent.

In view of this characterization and (5.6), to show that T_n has the same distribution as R_n it is enough to show that

(a) the random permutation $\{i \mapsto \sigma_{n,i} : 1 \leqslant i \leqslant 2n-1\}$ is a function of shape(T_n);
(b) shape$(T_n) = \Psi_n(\text{shape}(R_n))$ for some bijective map Ψ_n from the set of all possible shapes to itself.

This is trivial for $n=1$, so we assume below that $n \geqslant 2$. Before proving (a) and (b), we recall that (ii) above involves a natural bijection

$$(I_1, \ldots, I_{n-1}) \leftrightarrow \text{shape}(R_n) \tag{5.7}$$

where $I_{n-1} \in \{1, \ldots, 2n-3\}$ is the unique i such that

$$U_{n-1}\tau_{n-1} \in (X_{i-1}^{(n-1)}, X_i^{(n-1)}).$$

Hence, I_{n-1} is the index in the canonical ordering of edges of R_{n-1} of the edge that is cut in the transformation from R_{n-1} to R_n by attachment of an additional edge, of length $\tau_n - \tau_{n-1}$, connecting the cut-point to leaf n. Thus, (ii) and (iii) above correspond via (5.7) to the facts that $I_1, \ldots, I_{n-1}$

are independent and uniformly distributed over their ranges, and independent of lengths(R_n). These facts can be checked directly from the construction of $\{R_n\}_{n\in\mathbb{N}}$ from $\{\tau_n\}_{n\in\mathbb{N}}$ and $\{U_n\}_{n\in\mathbb{N}}$ using standard facts about uniform order statistics.

Now (a) and (b) follow from (5.7) and another bijection

$$(I_1, \ldots, I_{n-1}) \leftrightarrow \text{shape}(T_n) \tag{5.8}$$

where each possible value i of I_m is identified with edge $\sigma_{m,i}$ in the canonical ordering of edges of T_m. This is the edge of T_m whose length equals length(R_m, i). The bijection (5.8), and the fact that $\sigma_{n,i}$ depends only on shape(T_n), will now be established by induction on $n \geqslant 2$. For $n = 2$ the claim is obvious. Suppose for some $n \geqslant 3$ that the correspondence between $(I_1, \ldots, I_{n-2})$ and shape(T_{n-1}) has been established, and that the length of edge $\sigma_{n-1,i}$ in the canonical ordering of edges of T_{n-1} is equals the length of the ith edge in the canonical ordering of edges of R_{n-1}, for some $\sigma_{n-1,i}$ that is a function of i and shape(T_{n-1}). According to the construction of T_n, if $I_{n-1} = i$ then T_n is derived from T_{n-1} by splitting T_{n-1} into two branches at some point along edge $\sigma_{n-1,i}$ in the canonical ordering of the edges of T_{n-1}, and forming a new tree from the two branches and an extra segment of length $\tau_n - \tau_{n-1}$. Clearly, shape(T_n) is determined by shape(T_{n-1}) and I_{n-1}, and in the canonical ordering of the edge lengths of T_n the length of the ith edge equals the length of the edge $\sigma_{n,i}$ of R_n, for some $\sigma_{n,i}$ that is a function of shape(T_{n-1}) and I_{n-1}, and, therefore, a function of shape(T_n). To complete the proof, it is enough by the inductive hypothesis to show that the map

$$(\text{shape}(T_{n-1}), I_{n-1}) \to \text{shape}(T_n)$$

just described is invertible. But shape(T_{n-1}) and I_{n-1} can be recovered from shape(T_n) by the following sequence of moves:

- delete the edge attached to the root of shape(T_n)
- split the remaining tree into its two branches leading away from the internal node to which the deleted edge was attached;
- re-attach the bottom end of the branch not containing leaf n to leaf n on the other branch, joining the two incident edges to form a single edge;
- the resulting shape is shape(T_{n-1}), and I_{n-1} is the index such that the joined edge in shape(T_{n-1}) is the edge $\sigma_{n-1,I_{n-1}}$ in the canonical ordering of edges on shape(T_{n-1}).

□

5.3.2 Coupling

Lemma 5.8. *For any $(T, d, \rho) \in \mathbf{T}^{\text{root}}$ we can build on the same probability space two $\mathbf{T}^{\text{root}}$-valued processes X' and X'' such that:*

- *X' has the law of X under $\mathbf{P}^{T_0}$, where T_0 is the trivial tree consisting of just the root,*
- *X'' has the law of X under $\mathbf{P}^{T}$,*
- *for all $t \geqslant 0$,*

$$d_{\mathrm{GH}^{\mathrm{root}}}(X'_t, X''_t) \leqslant d_{\mathrm{GH}^{\mathrm{root}}}(T_0, T) = \sup\{d(\rho, x) : x \in T\} \tag{5.9}$$

- $$\lim_{t \to \infty} d_{\mathrm{GH}^{\mathrm{root}}}(X'_t, X''_t) = 0, \quad \textit{almost surely.} \tag{5.10}$$

Proof. The proof follows almost immediately from construction of X and Lemma 5.4. The only point requiring some comment is (5.10).

For that it will be enough to show for any $\varepsilon > 0$ that for $\mathbf{P}^T$-a.e. $(T, \pi_0, \pi) \in \Omega$ there exists $t > 0$ such that the projection of $\pi_0 \cap (]0, t] \times T^o)$ onto T is an ε-net for T.

Note that the projection of $\pi_0 \cap (]0, t] \times T^o)$ onto T is a Poisson process under $\mathbf{P}^T$ with intensity $t\mu$, where μ is the length measure on T. Moreover, T can be covered by a finite collection of ε-balls, each with positive μ-measure.

Therefore, the $\mathbf{P}^T$-probability of the set of $(T, \pi_0, \pi) \in \Omega$ such that the projection of $\pi_0 \cap (]0, t] \times T^o)$ onto T is an ε-net for T increases as $t \to \infty$ to 1. □

5.3.3 Convergence to Equilibrium

Proposition 5.9. *For any $T \in \mathbf{T}^{\mathrm{root}}$, the law of X_t under $\mathbf{P}^T$ converges weakly to that of the Brownian CRT as $t \to \infty$.*

Proof. It suffices by Lemma 5.8 to consider the case where T is the trivial tree.

We saw in the Proposition 5.7 that, in the notation of that result, $\mathcal{T}_{\tau_n -}$ has the same distribution as $\mathcal{R}_{\tau_n -}$.

Moreover, $\mathcal{R}_t$ converges in distribution to the continuum random tree as $t \to \infty$ if we use Aldous's metric on trees that comes from thinking of them as closed subsets of ℓ^1 with the root at the origin and equipped with the Hausdorff distance.

By construction, $(\mathcal{T}_t)_{t \geqslant 0}$ has the root growth with re-grafting dynamics started at the trivial tree. Clearly, the rooted Gromov–Hausdorff distance between $\mathcal{T}_t$ and $\mathcal{T}_{\tau_{n+1}-}$ is at most $\tau_{n+1} - \tau_n$ for $\tau_n \leqslant t < \tau_{n+1}$.

It remains to observe that $\tau_{n+1} - \tau_n \to 0$ in probability as $n \to \infty$. □

5.3.4 Recurrence

Proposition 5.10. *Consider a non-empty open set $U \subseteq \mathbf{T}^{\mathrm{root}}$. For each $T \in \mathbf{T}^{\mathrm{root}}$,*

$$\mathbf{P}^T\{\textit{for all } s \geqslant 0, \textit{ there exists } t > s \textit{ such that } X_t \in U\} = 1. \tag{5.11}$$

Proof. It is straightforward, but notationally rather tedious, to show that if $B' \subseteq \mathbf{T}^{\mathrm{root}}$ is any ball and T_0 is the trivial tree, then

$$\mathbf{P}^{T_0}\{X_t \in B'\} > 0 \tag{5.12}$$

for all t sufficiently large.

Thus, for any ball $B' \subseteq \mathbf{T}^{\mathrm{root}}$ there is, by Lemma 5.8, a ball $B'' \subseteq \mathbf{T}^{\mathrm{root}}$ containing the trivial tree such that

$$\inf_{T \in B''} \mathbf{P}^{T}\{X_t \in B'\} > 0 \tag{5.13}$$

for each t sufficiently large.

By a standard application of the Markov property, it therefore suffices to show for each $T \in \mathbf{T}^{\mathrm{root}}$ and each ball B'' around the trivial tree that

$$\mathbf{P}^{T}\{\text{there exists } t > 0 \text{ such that } X_t \in B''\} = 1. \tag{5.14}$$

By another standard application of the Markov property, equation (5.14) will follow if we can show that there is a constant $p > 0$ depending on B'' such that for any $T \in \mathbf{T}^{\mathrm{root}}$

$$\liminf_{t \to \infty} \mathbf{P}^{T}\{X_t \in B''\} > p.$$

This, however, follows from Proposition 5.9 and the observation that for any $\varepsilon > 0$ the law of the Brownian CRT assigns positive mass to the set of trees with height less than ε: this is just the observation that the law of the Brownian excursion assigns positive mass to the set of excursion paths with maximum less that $\varepsilon/2$. □

5.3.5 Uniqueness of the Stationary Distribution

Proposition 5.11. *The law of the Brownian CRT is the unique stationary distribution for X. That is, if ξ is the law of the CRT, then*

$$\int \xi(dT) P_t f(T) = \int \xi(dT) f(T)$$

for all $t \geqslant 0$ and $f \in \mathrm{b}\mathcal{B}(\mathbf{T}^{\mathrm{root}})$, and ξ is the unique probability measure on $\mathbf{T}^{\mathrm{root}}$ with this property.

Proof. This is a standard argument given Proposition 5.9 and the Feller property for the semigroup $(P_t)_{t \geqslant 0}$ established in Proposition 5.6, but we include the details for completeness.

Consider a test function $f : \mathbf{T}^{\mathrm{root}} \to \mathbb{R}$ that is continuous and bounded. By Proposition 5.6, the function $P_t f$ is also continuous and bounded for each $t \geqslant 0$.

Therefore, by Proposition 5.9,

$$\begin{aligned}\int \xi(dT)f(T) &= \lim_{s\to\infty}\int \xi(dT)P_s f(T) = \lim_{s\to\infty}\int \xi(dT)P_{s+t}f(T) \\ &= \lim_{s\to\infty}\int \xi(dT)P_s(P_t f)(T) = \int \xi(dT)P_t f(T)\end{aligned} \tag{5.15}$$

for each $t \geqslant 0$. Hence, ξ is stationary.

Moreover, if ζ is a stationary measure, then

$$\begin{aligned}\int \zeta(dT)f(T) &= \int \zeta(dT)P_t f(T) \\ &\to \int \zeta(dT)\left(\int \xi(dT)f(T)\right) = \int \xi(dT)f(T),\end{aligned} \tag{5.16}$$

and $\zeta = \xi$, as claimed. □

5.4 Convergence of the Markov Chain Tree Algorithm

We would like to show that Algorithm 5.1 converges to a process having the root growth with re-grafting dynamics after suitable re-scaling of time and edge lengths of the evolving tree. It will be more convenient for us to work with the continuous time version of the algorithm in which the transitions are made at the arrival times of an independent Poisson process with rate 1.

The continuous time version of Algorithm 5.1 involves a labeled combinatorial tree, but, by symmetry, if we don't record the labeling and associate rooted labeled combinatorial trees with rooted compact real trees having edges that are line segments with length 1, then the resulting process will still be Markovian.

It will be convenient to use the following notation for re-scaling the distances in a $\mathbb{R}$-tree: $T = (T, d, \rho)$ is a rooted compact real tree and $c > 0$, we write cT for the tree $(T, c\,d, \rho)$ (that is, $cT = T$ as sets and the roots are the same, but the metric is re-scaled by c).

Proposition 5.12. *Let $Y^n = (Y^n_t)_{t\geqslant 0}$ be a sequence of Markov processes that take values in the space of rooted compact real trees with integer edge lengths and evolve according to the dynamics associated with the continuous-time version of Algorithm 5.1. Suppose that each tree Y^n_0 is non-random with total branch length N_n, that N_n converges to infinity as $n \to \infty$, and that $N_n^{-1/2}Y^n_0$ converges in the rooted Gromov–Hausdorff metric to some rooted compact real tree T as $n \to \infty$. Then, in the sense of weak convergence of processes on the space of càdlàg paths equipped with the Skorohod topology, $(N_n^{-1/2}Y^n(N_n^{1/2}t))_{t\geqslant 0}$ converges as $n \to \infty$ to the root growth with re-grafting process X under $\mathbf{P}^T$.*

Proof. Define $Z^n = (Z^n_t)_{t\geqslant 0}$ by

$$Z^n_t := N_n^{-1/2}Y^n(N_n^{1/2}t).$$

For $\eta > 0$, let $Z^{\eta,n}$ be the $\mathbf{T}^{\text{root}}$-valued process constructed as follows.

- Set $Z_0^{\eta,n} = R_{\eta_n}(Z_0^n)$, where $\eta_n := N_n^{-1/2}\lfloor N_n^{1/2}\eta\rfloor$.
- The value of $Z^{\eta,n}$ is unchanged between jump times of $(Z_t^n)_{t\geqslant 0}$.
- At a jump time τ for $(Z_t^n)_{t\geqslant 0}$, the tree $Z_\tau^{\eta,n}$ is the subtree of Z_τ^n spanned by $Z_{\tau-}^{\eta,n}$ and the root of Z_τ^n.

An argument similar to that in the proof of Lemma 5.4 shows that

$$\sup_{t\geqslant 0} d_{\mathrm{H}}(Z_t^n, Z_t^{\eta,n}) \leqslant \eta_n,$$

and so it suffices to show that $Z^{\eta,n}$ converges weakly as $n \to \infty$ to X under $\mathbf{P}^{R_\eta(T)}$.

Note that $Z_0^{\eta,n}$ converges to $R_\eta(T)$ as $n \to \infty$. Moreover, if Λ is the map that sends a tree to its total length (that is, the total mass of its length measure), then $\lim_{n\to\infty} \Lambda(Z_0^{\eta,n}) = \Lambda \circ R_\eta(T) < \infty$ by Lemma 4.36 below.

The pure jump process $Z^{\eta,n}$ is clearly Markovian. If it is in a state (T', ρ'), then it jumps with the following rates.

- With rate $N_n^{1/2}(N_n^{1/2}\Lambda(T'))/N_n = \Lambda(T')$, one of the $N_n^{1/2}\Lambda(T')$ points in T' that are at distance a positive integer multiple of $N_n^{-1/2}$ from the root ρ' is chosen uniformly at random and the subtree above this point is joined to ρ' by an edge of length $N_n^{-1/2}$. The chosen point becomes the new root and a segment of length $N_n^{-1/2}$ that previously led from the new root toward ρ' is erased. Such a transition results in a tree with the same total length as T'.
- With rate $N_n^{1/2} - \Lambda(T')$, a new root not present in T' is attached to ρ' by an edge of length $N_n^{-1/2}$. This results in a tree with total length $\Lambda(T') + N_n^{-1/2}$.

It is clear that these dynamics converge to those of the root growth with re-grafting process, with the first class of transitions leading to re-graftings in the limit and the second class leading to root growth. □