PROPERTIES OF LEVEL SET TREES OF GEOMETRIC RANDOM WALKS

CLAIRE CAMPREGHER AND JESSE JOHNSON

ADVISOR: YEVGENIY KOVCHEGOV
OREGON STATE UNIVERSITY

ABSTRACT. In this report we outline the results that we found regarding the parameters of a pruned walk, as well as the behavior after repeated pruning. Specifically, looking at parameters after repeated pruning of mean zero walks, as well as representation as a Galton Watson process. Horton and Tokunaga self similarity were also considered.

1. INTRODUCTION

In our results, we are primarily concerned with establishing the ways in which the pruned walks of a geometric random walk behave. In order to do this, we examine the structure and parameters of a random walk pruned once, and then extend this to walks pruned many times. We were able to uncover unique properties about these walks, including accordance with Horton Law, and classification as a Galton Watson process if the walk's mean is less than or equal to zero. Moreover, we have preliminary results regarding Tokunaga self similarity. Throughout our paper, we have built upon the results from the 2016 OSU REU group (Anna-Sophia Hirst, Rachel Linder, and Daniel Malmuth), both in checking and extending their ideas.

2. TREES

In order to understand the relationship between our random walk and the associated level set tree, we first need to examine some of the basic properties of nonbinary rooted trees.

2.1. Structure of Trees.

Definition 2.1. *A* tree*, T, is a connected simple non-cyclic graph. It consists of a sequence of points called* vertices*, V, and lines connecting them called* edges*, E. No pair of vertices are connected by more than one edge, and no sequence of vertices are connected cyclically.* $T = (V,E)$.

The second tree in Figure 1 may also illustrate the parent child relationship present in rooted trees. Each edge connected to the root is the roots *child*, and the root is the *parent*. After this, for any edge e, every edge connected to e that is not a parent of e are called the *children* of e, and e is their *parent*. [1]

Date: 18 August 2017.

This work was done during the Summer 2017 REU program in Mathematics at Oregon State University, with support by National Science Foundation Grant DMS-1359173.

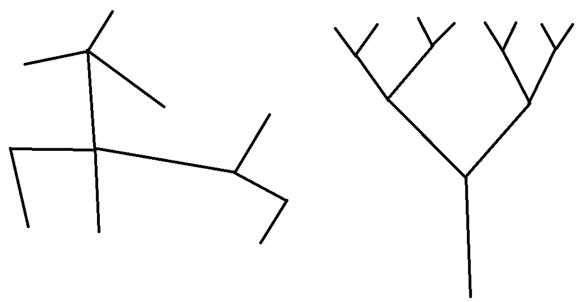

FIGURE 1. Two trees. The tree on the right is rooted.

2.2. **Pruning.**

Definition 2.2. Pruning *is the operation of removing all of the leaves from a tree, followed by a series reduction.*

Definition 2.3. *A* series reduction *consists of merging all edges with one child with their children.*

For the purposes of this paper, consider series reduction to happen automatically after any pruning.

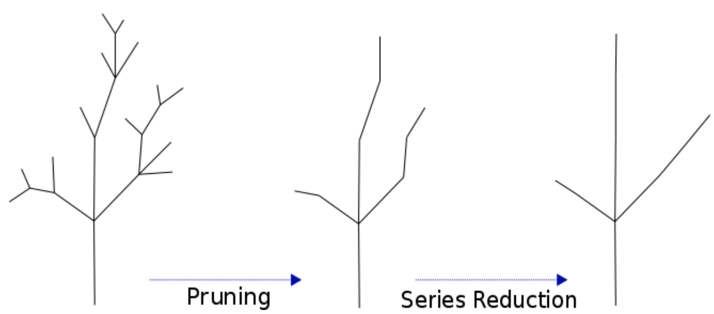

FIGURE 2. Pruning a tree by cutting off leaves and performing a series reduction.

2.3. **Horton Strahler Ordering and Tokunaga Indexing.**

Definition 2.4. *The* Horton-Strahler Order *of an edge is the minimum number of prunings required to completely eliminate it from the tree.*

Similarly, the Horton-Strahler Order of a tree, denoted Ω, is the minimum number of prunings required to completely eliminate the tree. The Horton-Strahler order of any vertex is the Horton-Strahler order of that vertex's parent edge. For the purposes of this paper, a *branch* is the full sequence of an edge and all it's parent or child edges of the same order.

Definition 2.5. Tokunaga indexing *is a valuable way of describing how edges of differing orders merge together. Let* $T_{ij} = \frac{N_{ij}}{N_j}$ *denote the average number of branches of order* $i < j$ *per branch of order* j *in a finite tree order* $\Omega \leq j$.

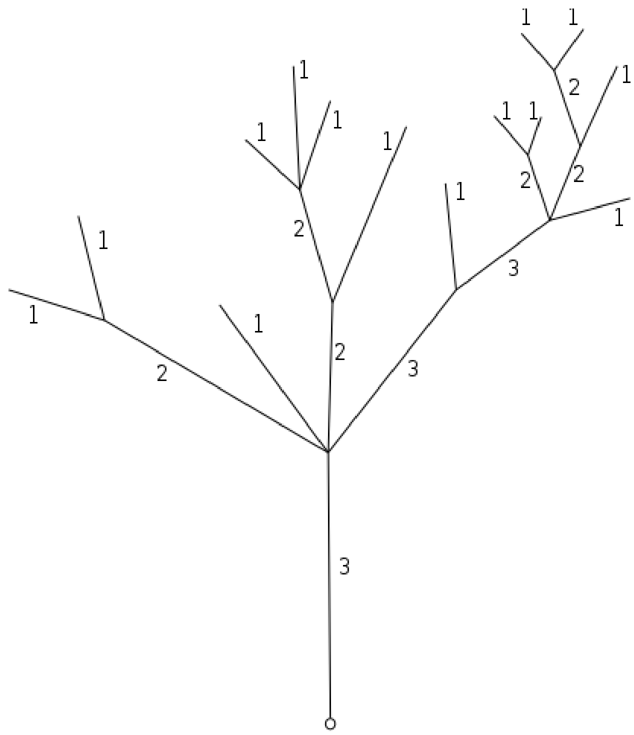

FIGURE 3. Example of how to order the edges of a tree

2.4. **Self-Similarity.**

Definition 2.6. Self Similarity *occurs when statistical structures in trees are preserved under pruning.*

We can think of this in relation to Horton orders, as well as how branches of different orders merge into each other.

Definition 2.7. *We can define* Horton Self Similarity *by the following:*

$$\frac{N_r}{N_{r+1}} = R_b$$

A tree will be Horton Self Similar, assuming it satisfies at least one of the following (and is finite and positive):

$$Root\,Law: \quad \lim_{k\to\infty} (\mathcal{N}_k)^{-1/k} = R \quad R > 0$$
$$Ratio\,Law: \quad \lim_{k\to\infty} \frac{\mathcal{N}_k}{\mathcal{N}_{k+1}} = R \quad R > 0$$
$$Geometric\,Law: \quad \lim_{k\to\infty} \mathcal{N}_k * R^k = N_0 \quad N_0 > 0$$

Where we define:

$$\mathcal{N}_k = \lim_{N\to\infty} \frac{N_k^{(P_k)}}{N}$$

Tokunaga Self Similarity is used to look at branching statistics.

Definition 2.8. *A tree is self similar in regard to branching statistics if* $E\left(\tau^j_{i(i+k)}\right) =: T_k$.

This means that for a particular branch order $i+k$ we can expect a certain number of mergings with branches order k smaller (order i) .

Definition 2.9. *We acheive* Tokunaga Self Similarity *if:*

$$\frac{T_{k+1}}{T_k} = c \iff T_k = ac^{k-1}$$

2.5. The Geometric Random Walk.

Definition 2.10. *A* Geometric random walk *is a stochastic process with a transition kernel* $K(x)$ *given by*

$$K(x) = p_1 g_1(x) + \alpha\delta_0 + p_2 g_2(-x)$$

where

$$g_i(x) = \sum_{a=1}^{\infty} r_i(1-r_i)^{a-1}\delta_a(x) \qquad i = 1,2$$

Note that p_1 is the probability of going up, p_2 is the probability of going down, $\alpha = (1-p_1-p_2)$ is the probability you stay the same height, $g_1(x)$ gives the distance up with a geometric random variable r_1, and $g_2(x)$ gives the distance down with a geometric random variable r_2.

3. Results

In this section we will outline our main theorems and results regarding the parameters of our random walk and tree branching statistics.

Theorem 3.1. *The distribution of the pruned geometric random walk on* $\mathbb{Z}$ *will also be geometric, with the following parameters:*

$$r_1^{(1)} = \frac{p_2 r_1}{p_1+p_2}$$
$$r_2^{(1)} = \frac{p_1 r_2}{p_1+p_2}$$
$$p_1^{(1)} = \frac{r_2^{(1)}(1-r_1^{(1)})}{1-(1-r_1^{(1)})(1-r_2^{(1)})}$$
$$p_2^{(1)} = \frac{r_1^{(1)}(1-r_2^{(1)})}{1-(1-r_1^{(1)})(1-r_2^{(1)})}$$

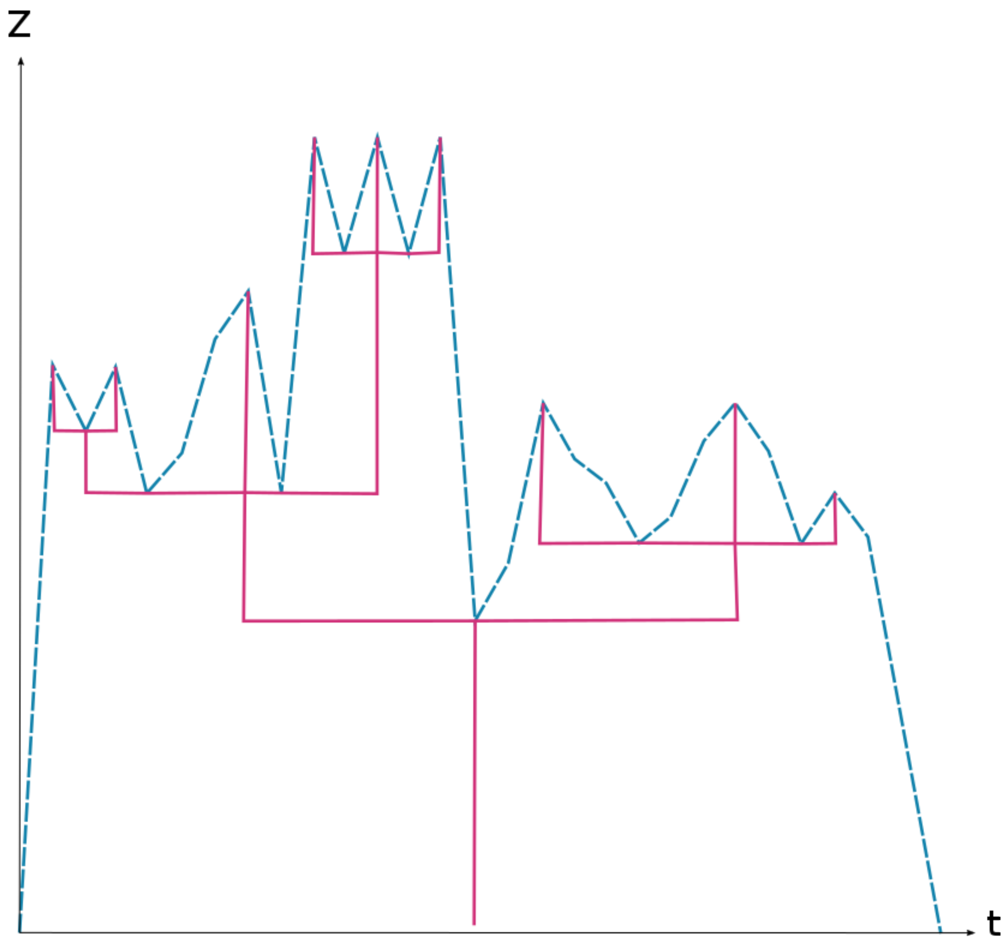

FIGURE 4. The teal line represents the random walk, and the pink line is our level set tree.

Proof. Recall that the kernel of our walk is given by

$$K(x) = p_1 g_1(x) + \alpha\delta_0 + p_2 g_2(-x)$$

where

$$g_i(x) = \sum_{a=1}^{\infty} r_i(1-r_i)^{a-1}\delta_a(x) \qquad i = 1,2$$

The characteristic equation of the kernel, found by taking the Fourier transform, is:

$$\hat{K}(x) = p_1\left(\frac{r_1 e^{is}}{1-(1-r_1)e^{is}}\right) + \alpha + p_2\left(\frac{r_2 e^{is}}{1-(1-r_2)e^{is}}\right)$$

Let our difference in heights of local minimums be:

$$d_j = \sum_{j=1}^{\xi_+} Y_j - \sum_{j=1}^{\xi_-} Z_j$$

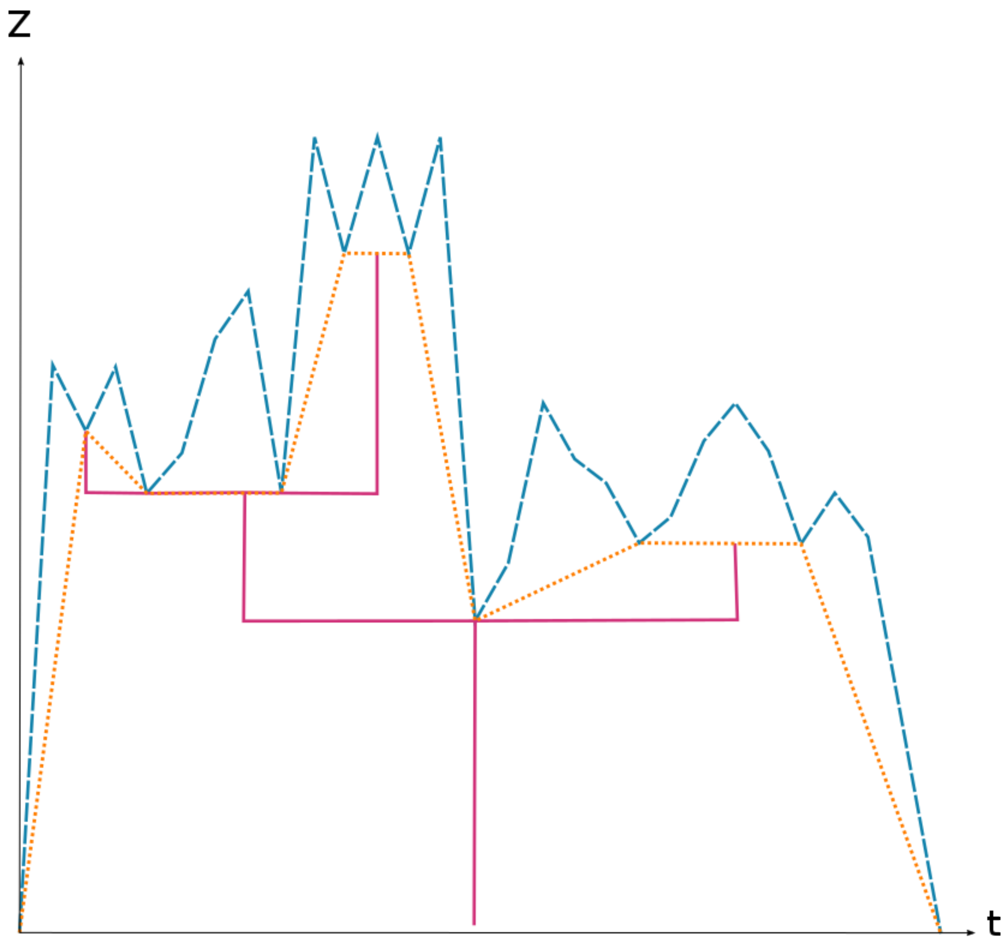

FIGURE 5. The pruned tree of our original random walk is the level set tree of the path connecting the minima of our random walk.

The Y_j sum is the total distance up before going down, and the Z_j is the total distance down before reaching the next minimum. We also must acknowledge that $\xi_+ = geom\left(\frac{p_2}{p_1+p_2}\right)$ and $\xi_- = geom\left(\frac{p_1}{p_1+p_2}\right)$.
Once we calculate these two expectations:

$$E\left(e^{is\left(\sum_{j=1}^{\infty} Y_j - \sum_{j=1}^{\infty} Z_j\right)}\right) = E\left(e^{is\sum_{j=1}^{\infty} Y_j}\right) E\left(e^{-is\sum_{j=1}^{\infty} Z_j}\right)$$

$$E\left(e^{is\sum_{j=1}^{\infty} Y_j}\right) = \frac{\left(\frac{p_2 r_1}{p_1+p_2}\right) e^{is}}{1-\left(1-\frac{p_2 r_1}{p_1+p_2}\right) e^{is}} \implies geo\left(\frac{p_2 r_1}{p_1+p_2}\right)$$

$$E\left(e^{-is\sum_{j=1}^{\infty} Z_j}\right) = \frac{\left(\frac{p_1 r_2}{p_1+p_2}\right) e^{-is}}{1-\left(1-\frac{p_1 r_2}{p_1+p_2}\right) e^{-is}} \implies geo\left(\frac{p_1 r_2}{p_1+p_2}\right)$$

Then we know the characteristic equation of d_j is:

$$E\left(e^{is\left(\sum_{j=1}^{\xi_+} Y_j - \sum_{j=1}^{\xi_-} Z_j\right)}\right) = \frac{\left(\frac{p_2 r_1}{p_1+p_2}\right)e^{is}}{1-\left(1-\frac{p_2 r_1}{p_1+p_2}\right)e^{is}} \cdot \frac{\left(\frac{p_1 r_2}{p_1+p_2}\right)e^{-is}}{1-\left(1-\frac{p_1 r_2}{p_1+p_2}\right)e^{-is}}$$

This implies that:

$$r_1^{(1)} = \frac{p_2 r_1}{p_1+p_2}, \quad r_2^{(1)} = \frac{p_1 r_2}{p_1+p_2}$$

Take $k \geq 1$.

$$\begin{aligned}
& p(x_u - x_d = k) \\
&\qquad = \sum_{j=k+1}^{\infty} p(x_u = j) p(x_d = j-k) \\
&\qquad = \sum_{j=k+1}^{\infty} (1-r_1^{(1)})^{j-1} r_1^{(1)} \cdot (1-r_2^{(1)})^{j-k-1} r_2^{(1)} \\
&\qquad = \frac{r_2^{(1)}(1-r_1^{(1)})}{1-(1-r_1^{(1)})(1-r_2^{(1)})} \cdot r_1^{(1)}(1-r_1^{(1)})^{k-1} \\
\implies & p_1^{(1)} = \frac{r_2^{(1)}(1-r_1^{(1)})}{1-(1-r_1^{(1)})(1-r_2^{(1)})}
\end{aligned}$$

$$\begin{aligned}
p(x_u - x_d = -k) &= \sum_{j=k+1}^{\infty} p(x_u = j-k) p(x_d = j) \\
&= \sum_{j=k+1}^{\infty} (1-r_1^{(1)})^{j-k-1} r_1^{(1)} (1-r_2^{(1)})^{j-1} r_2^{(1)} \\
&= \frac{r_1^{(1)}(1-r_2^{(1)})}{1-(1-r_1^{(1)})(1-r_2^{(1)})} \cdot r_2^{(1)}(1-r_2^{(1)})^{k-1} \\
\implies p_2^{(1)} &= \frac{r_1^{(1)}(1-r_2^{(1)})}{1-(1-r_1^{(1)})(1-r_2^{(1)})}
\end{aligned}$$

□

Corollary 3.2. *The parameters of a pruned geometric random walk, where we guarantee that each step will be nonzero in height, will be geometric with the following parameters. Note that* $p_1 = p$

and $p_2 = 1 - p$.

$$\begin{aligned}
r_1^{(1)} &= pr_2 \\
r_2^{(1)} &= pr_2 \\
p_1^{(1)} &= \frac{p^2 r_1 r_2 - pr_1 r_2 + pr_2}{p^2 r_1 r_2 - pr_1 r_2 + pr_2 + r_1} \\
p_2^{(1)} &= \frac{p^2 r_1 r_2 - pr_1 r_2 - pr_1 + r_1}{p^2 r_1 r_2 - pr_1 r_2 + pr_2 - pr_1}
\end{aligned}$$

Proof. Recall that from the above theorem, and substituting in $p_1 = p$ and $p_2 = 1 - p$, we have:

$$\begin{aligned}
r_1^{(1)} &= \frac{p_2 r_1}{p_1 + p_2} \\
&= \frac{(1-p)r_1}{1-p+p} \\
&= (1-p)r_1 \\
r_2^{(1)} &= \frac{p_1 r_2}{p_1 + p_2} \\
&= \frac{pr_2}{1-p+p} \\
&= pr_2 \\
p_1^{(1)} &= \frac{r_2^{(1)}(1 - r_1^{(1)})}{1 - (1 - r_1^{(1)})(1 - r_2^{(1)})} \\
&= \frac{pr_2(1 - (1-p)r_1)}{1 - (1 - (1-p)r_1)(1 - pr_2)} \\
&= \frac{p^2 r_1 r_2 - pr_1 r_2 + pr_2}{p^2 r_1 r_2 - pr_1 r_2 + pr_2 + r_1} \\
p_2^{(1)} &= \frac{r_1^*(1 - r_2^*)}{1 - (1 - r_2^*)(1 - r_1^*)} \\
&= \frac{(1-p)r_1(1 - pr_2)}{1 - (1 - (1-p)r_1)(1 - pr_2)} \\
&= \frac{p^2 r_1 r_2 - pr_1 r_2 - pr_1 + r_1}{p^2 r_1 r_2 - pr_1 r_2 + pr_2 - pr_1}
\end{aligned}$$

□

Corollary 3.3. *The parameters of a the pruned symmetric geometric random walk will be geometric and symmetric with the following parameters. Note that $r = r_1 = r_2$, $p = p_1 = p_2$.*

$$\begin{aligned}
r_1^{(1)} &= r_2^{(1)} = \frac{r}{2} \\
p_1^{(1)} &= p_2^{(1)} = \frac{\left(1-\frac{r}{2}\right)}{\left(2-\frac{r}{2}\right)} \\
\alpha^{(1)} &= \frac{r}{4-r}
\end{aligned}$$

Proof.

$$\begin{aligned}
r_1^{(1)} &= \frac{p_2 r_1}{p_1+p_2} = \frac{pr}{2p} = \frac{r}{2} \\
r_2^{(1)} &= \frac{p_1 r_2}{p_1+p_2} = \frac{pr}{2p} = \frac{r}{2} \\
p_1^{(1)} &= \frac{r_2^{(1)}\left(1-r_1^{(1)}\right)}{1-\left(1-r_1^{(1)}\right)\left(1-r_2^{(1)}\right)} \\
&= \frac{\frac{r}{2}\left(1-\frac{r}{2}\right)}{1-\left(1-\frac{r}{2}\right)\left(1-\frac{r}{2}\right)} \\
&= \frac{\left(1-\frac{r}{2}\right)}{\left(2-\frac{r}{2}\right)} \\
p_2^{(1)} &= \frac{r_1^{(1)}\left(1-r_2^{(1)}\right)}{1-\left(1-r_1^{(1)}\right)\left(1-r_2^{(1)}\right)} \\
&= \frac{\frac{r}{2}\left(1-\frac{r}{2}\right)}{\left(1-\frac{r}{2}\right)\left(1-\frac{r}{2}\right)} \\
&= \frac{\left(1-\frac{r}{2}\right)}{\left(2-\frac{r}{2}\right)} \\
\alpha^{(1)} &= 1-p_1^{(1)}-p_2^{(1)} \\
&= \frac{r}{4-r}
\end{aligned}$$

□

Lemma 3.4. *The pruned walk of any geometric random walk with mean zero will be symmetric geometric.*

Proof. Recall that for a mean zero random walk we have $\frac{p_1}{r_1} = \frac{p_2}{r_2} \implies p_1 r_2 = p_2 r_1$.
If we divide both sides of this equality by $p_1 + p_2$ we get that $\frac{p_2 r_1}{p_1+p_2} = \frac{p_1 r_2}{p_1+p_2}$, which we can recall

is the same as saying that $r_1^{(1)} = r_2^{(1)}$ by definition of $r_1^{(1)}$ and $r_2^{(2)}$. This implies that:

$$\begin{aligned}
r_1^{(1)} &= r_2^{(1)} \\
r_1^{(1)} - r_1^{(1)} r_2^{(1)} &= r_2^{(1)} - r_1^{(1)} r_2^{(1)} \\
\frac{r_1^{(1)} - r_1^{(1)} r_2^{(1)}}{1-(1-r_1^{(1)})(1-r_2^{(1)})} &= \frac{r_2^{(1)} - r_1^{(1)} r_2^{(1)}}{1-(1-r_1^{(1)})(1-r_2^{(1)})} \\
\frac{r_1^{(1)}(1-r_2^{(1)})}{1-(1-r_1^{(1)})(1-r_2^{(1)})} &= \frac{r_2^{(1)}(1-r_1^{(1)})}{1-(1-r_1^{(1)})(1-r_2^{(1)})}
\end{aligned}$$

Of course, we already know that this is the same as saying that $p_1^{(1)} = p_2^{(1)}$.
Since we have both $r_1^{(1)} = r_2^{(1)}$ and $p_1^{(1)} = p_2^{(1)}$, we fulfill the definition of a symmetric walk. □

Lemma 3.5. *For any number of prunings $n \geq 1$ of a geometric random walk with mean zero, the n^{th} pruned walk will be geometric symmetric with the following parameters:*

$$\begin{aligned}
r^{(n)} &= 2^{-n} r \\
p^{(n)} &= \frac{1-2^{-n} r}{2-2^{-n} r} \\
\alpha^{(n)} &= \frac{r}{2^{n+1}-r}
\end{aligned}$$

Proof. By Lemma 3.4, we know that the pruned walk of any geometric random walk with mean zero will be symmetric geometric. After this, by further prunings will also be symmetric geometric by Corollary 3.3. □

Lemma 3.6. *For any symmetric geometric random walk, the both $\alpha^{(n)}$ and $r^{(n)}$ approach towards zero as n approaches infinity.*

Proof. Recall that

$$\begin{aligned}
\alpha^{(n)} &= \frac{r}{2^{n+1}-r} \\
r^{(n)} &= 2^{-n} r
\end{aligned}$$

Now observe that for a fixed geometric parameter r, and number of prunings n:

$$\begin{aligned}
\lim_{n\to\infty} \alpha^{(n)} &= \lim_{n\to\infty} \frac{r}{2^{n+1}-r} \\
&= \lim_{n\to\infty} \frac{\frac{r}{2^{n+1}}}{1-\frac{r}{2^{n+1}}} \\
&= 0
\end{aligned}$$

Similarly, we have:

$$\begin{aligned}
\lim_{n\to\infty} r^{(n)} &= \lim_{n\to\infty} 2^{-n} r \\
&= \lim_{n\to\infty} \frac{r}{2^n} \\
&= 0
\end{aligned}$$

$\square$

Lemma 3.7. *The ratio of maximums to minimums in a geometric random walk is* $\frac{p_1 p_2}{p_1+p_2}$*. For a symmetric geometric random walk, this simplifies to* $\frac{1-\alpha}{4}$*.*

Proof. For a particular point to be a maximum, we need the following to occur.

(1) To have previously gone up with probability p_1
(2) To then plateau for n steps total, for a probability of $(1-p_1-p_2)$
(3) To finally go down with a probability of p_2

For all of these to happen, we need to multiply the probability of each of these conditions together, and sum over all possible n.

$$\begin{aligned}
\sum_{n=0}^{\infty} p_1(1-p_1-p_2)^n p_2 &= p_1 p_2 \sum_{n=0}^{\infty} (1-p_1-p_2)^n \\
&= p_1 p_2 \frac{1}{1-(1-p_1-p_2)} \\
&= \frac{p_1 p_2}{p_1+p_2}
\end{aligned}$$

For symmetric, recognize that $p_1 = p_2 = \frac{1-\alpha}{2}$. Substitute this value into the probability of a maximum as stated above to obtain:

$$\begin{aligned}
\frac{p_1 p_2}{p_1+p_2} &= \frac{\frac{1-\alpha}{2}\frac{1-\alpha}{2}}{\frac{1-\alpha}{2}+\frac{1-\alpha}{2}} \\
&= \frac{\frac{(1-\alpha)^2}{4}}{(1-\alpha)} \\
&= \frac{1-\alpha}{4}
\end{aligned}$$

$\square$

Theorem 3.8. *A geometric random walk with mean zero satisfies the Horton Ratio Law, with R = 4.*

Proof. Since our original random walk has mean zero, we know that all following pruned walks will be symmetric. Because of this, we have that:

$$\begin{aligned}
\lim_{k\to\infty}\frac{\mathcal{N}_k}{\mathcal{N}_{k+1}} &= \lim_{k\to\infty}\left(\lim_{N\to\infty}\frac{N_k^{(P_N)}}{N}\frac{N}{N_{k+1}^{(P_N)}}\right)\\
&= \lim_{k\to\infty}\frac{N_k^{(P_N)}}{N_{k+1}^{(P_N)}}\\
&= \lim_{k\to\infty}\frac{Number\,of\,order\,k\,branches}{Number\,of\,order\,k+1\,branches}\\
&= \lim_{k\to\infty}\frac{Number\,of\,local\,minimums\,after\,k-1\,prunings}{Number\,of\,local\,maximums\,after\,k-1\,prunings}\\
&= \lim_{k\to\infty}\frac{1}{\frac{1-\alpha^{k-1}}{4}}\\
&= \lim_{k\to\infty}\frac{4}{1-\alpha^{k-1}}\\
&= 4
\end{aligned}$$

$\square$

Definition 3.9. *A* Galton-Watson Process *is a tree branching process in which at each edge, offspring are created with probability q_k of having k offspring at each stage.*

Definition 3.10. *By looking at the expected number of children, we can classify a Galton-Watson Process to be either* subcritical, critical, *or* supercritical.

$$\sum_{k=0}^{\infty} kp_k \begin{cases} <1 & Subcritical\\ =1 & Critical\\ >1 & Supercritical \end{cases}$$

Theorem 3.11. *The Galton Watson Process associated with the geometric random walk is subcritical if the mean is less than zero, meaning $\frac{p_1^{(1)}}{r_1^{(1)}} < \frac{p_2^{(1)}}{r_2^{(1)}}$. The Galton Watson Process will be critical if the mean is equal to zero, meaning $\frac{p_1^{(1)}}{r_1^{(1)}} = \frac{p_2^{(1)}}{r_2^{(1)}}$.*

Proof. Our tree will converge to critical Galton-Watson if we are able to satisfy the following:

$$\sum_{k=0}^{\infty} kp_k = 1$$

Since we know $p_1^{(1)} = 1 - p_0$ we can temporarily rewrite p_k as

$$p_k = p_1^{(1)}(1 + geom(1 - r_2^{(1)}))$$

When we sum over all k, we obtain:

$$\begin{aligned}
\sum_{k=0}^{\infty} kp_k &= 0p_0 + p_1^{(1)}\left(1 + \frac{1}{1-r_2^{(1)}}\right) \\
&= p_1^{(1)}\left(\frac{2-r_2^{(1)}}{1-r_2^{(1)}}\right) \\
&= \frac{(1-r_1^{(1)})}{(1-r_2^{(1)})}\frac{(2r_2^{(1)}-(r_2^{(1)})^2)}{(r_1^{(1)}+r_2^{(1)}-r_1^{(1)}r_2^{(1)})} \\
&= \frac{1-r_1^{(1)}}{(1-r_2^{(1)})}\frac{(r_2^{(1)}+r_2^{(1)}(1-r_2^{(1)})}{(r_2^{(1)}+r_1^{(1)}(1-r_2^{(1)}))}
\end{aligned}$$

From this, we can see that this sum is 1 iff $\frac{p_1}{r_1} = \frac{p_2}{r_2} \iff r_1^{(1)} = r_2^{(1)}$ This means that we achieve Critical Galton Watson only if our geometric random walk is mean zero.

For the subcritical case, consider the case of having $\frac{p_1}{r_1} < \frac{p_2}{r_2}$, which is equivalent to stating that $p_1r_2 < p_2r_1$. Start by analyzing $r_1^{(1)}$ and $r_2^{(1)}$.
Recall that $r_1^{(1)} = \frac{p_2r_1}{p_1+p_2}$ and $r_2^{(1)} = \frac{p_1r_2}{p_1+p_2}$. Since we know $p_1r_2 < p_2r_1$, we can say that:

$$r_1^{(1)} = \frac{p_2r_1}{p_1+p_2} < \frac{p_1r_2}{p_1+p_2} = r_2^{(1)}$$

From this, we can prove that $\frac{p_1^{(1)}}{r_1^{*}(1)} < \frac{p_2^{(1)}}{r_2^{(1)}}$.

$$\begin{aligned}
\frac{p_1^{(1)}}{r_1^{(1)}} &= \frac{r_2^{(}1-r_1^{(1)})}{r_1(1)\left(1-(1-r_1(1))(1-r_2^{(1)})\right)} \\
&< \frac{r_1^{(1)}(1-r_1^{(1)})}{r_2^{(1)}\left(1-(1-r_1^{(1)})(1-r_2^{(1)})\right)} \\
&< \frac{r_1^{(1)}(1-r_2^{(1)})}{r_2^{(1)}\left(1-(1-r_1^{(1)})(1-r_2^{(1)})\right)} = \frac{p_2^{(1)}}{r_2^{(1)}}
\end{aligned}$$

Finally, we can prove that a distribution of this variety will be Subcritical Galton Watson. Meaning,

$$\sum_{k=0}^{\infty} kp_k < 1$$

From our results above when we found the value of this sum, we know that:

$$\sum_{k=0}^{\infty} kp_k = \frac{1-r_1^{(1)}}{1-r_2^{(1)}}\frac{r_2^{(1)}+r_2^{(1)}(1-r_2^{(1)})}{r_2^{(1)}+r_1^{(1)}(1-r_2^{(1)})}$$

Since we know $r_2^{(1)} < r_1^{(1)}$:

$$(1 - r_1^{(1)}(r_2^{(1)} + r_2^{(1)}(1 - r_2^{(1)})) < (1 - r_2^{(1)})(r_2^{(1)} + r_1^{(1)}(1 - r_2^{(1)}))$$

This implies that

$$\sum_{k=0}^{\infty} k p_k = \frac{1 - r_1^{(1)}}{1 - r_2^{(1)}} \frac{r_2^{(1)} + r_2^{(1)}(1 - r_2^{(1)})}{r_2^{(1)} + r_1^{=}(1)(1 - r_2^{(1)})} < 1$$

which satisfies the conditions for subcritical Galton-Watson. □

Theorem 3.12. *For a random walk mean zero, after a large amount of prunings, we eventually converge to a binary Galton-Watson process.*

Proof. For a symmetric walk, show that as we do $n \to \infty$ prunings, we achieve (assuming symmetric):

$$p_0^{(n)} \to 1/2$$

$$p_2^{(n)} \to 1/2$$

$$p_k^{(n)} \to 0$$

To establish the first:

$$\begin{aligned}
\lim_{n\to\infty} p_0^{(n)} &= \lim_{n\to\infty} \frac{1}{2 - r^{(n)}} \\
&= \lim_{n\to\infty} \frac{1}{2 - \frac{r}{2^n}} \\
&= \frac{1}{2-0} \\
&= \frac{1}{2}
\end{aligned}$$

To establish the second:

$$\begin{aligned}
\lim_{n\to\infty} p_2^{(n)} &= \lim_{n\to\infty} p_1^*(r^{(n)})^{2-2}(1 - r^{(n)}) \\
&= \lim_{n\to\infty} \frac{r^{(n)}(1 - r^{(n)})}{(-r^{(n)})^2 + 2r^{(n)}} \\
&= \lim_{n\to\infty} \frac{1 - r^{(n)}}{2 - r^{(n)}} \\
&= \frac{1}{2}
\end{aligned}$$

To establish the third:

$$\begin{aligned}
\lim_{n\to\infty} p_k^{(n)} &= \lim_{n\to\infty} p_1^*(r^{(n)})^{k-2}(1-r^{(n)}) \\
&= \lim_{n\to\infty} \frac{r^{(n)}(1-r^{(n)})}{2r^{(n)}-(r^{(n)2})}(r^{(n)})^{k-2}(1-r^{(n)}) \\
&= \lim_{n\to\infty} \frac{(r^{(n)})^{k-1}(1-r^{(n)})^2}{r^{(n)}(2-r^{(n)})} \\
&= \lim_{n\to\infty} \frac{(r^{(n)})^{k-2}(1-r^{(n)})^2}{2-r^{(n)}} \\
&= \frac{0(1-0)}{2-0} \\
&= 0
\end{aligned}$$

□

Theorem 3.13. *The lengths of edges of any order is distributed geometrically with parameter* $r_1^{(1)} + r_2^{(1)} - r_1^{(1)} r_2^{(1)}$.

Proof. Let ξ be the point that starts an excursion, with height 0. Let k be any value, and let ξ_k be the first minimum with height less than k. Let Z be the first maximum after ξ, and let Z_k be the last maximum before ξ_k. Let Y be the horizontal distance between Z_k and ξ_k. Finally, let L be the height of the lowest extreme point in the excursion.

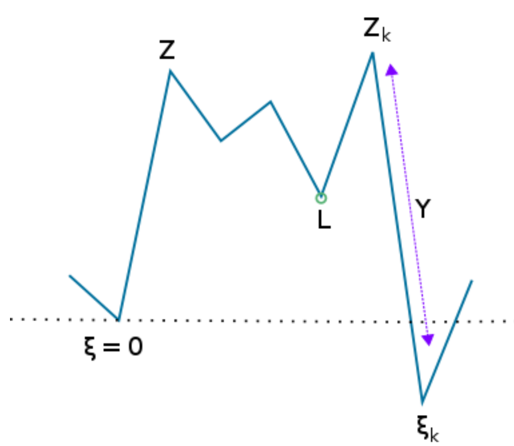

FIGURE 6. Visual representation of ξ, ξ_k, Z, Z_k, L, and Y

Then we have

$$P(L > k) = P(Z > k, \xi_k \leq 0) = P(\xi_k \leq 0 | Z > k) P(Z > k)$$

Because the height of Z is geometrically distributed with parameter $r_1^{(1)}$, we have

$$P(Z > k) = (1 - r_1^{(1)})^k$$

Let a be the height of Z_k, so that $Y = a - \xi_k$ and

$$Y \sim Geom(r_2^{(1)})$$

where it is given that $Y \geq (a-k)$. We then calculate that

$$\begin{aligned}
P(\xi_k \leq 0|Z_k = a) &= P(Y \geq a|Z_k = a, Y \geq a-k)\\
&= P(Y \geq a|Y \geq a-k)\\
&= (1-r_2^{(1)})^k
\end{aligned}$$

Knowing this, we then calculate

$$\begin{aligned}
P = (\xi_k \leq 0|Z_k > k) &= \sum_{a=k+1}^{\infty} P(\xi_k \leq 0|Z_k = a)\\
&= \left[\sum_{a=k+1}^{\infty} P(Z_k = a|Z_k > k)\right](1-r_2^{(1)})^k\\
&= P(Z_k > k|Z > k)(1-r_2^*)^k\\
&= (1-r_2^{(1)})^k
\end{aligned}$$

Referring to what we determined previously, we can rewrite

$$\begin{aligned}
P(L > k) &= P(Z > k, \xi_k \leq 0)\\
&= P(\xi_k \leq 0|Z > k)P(Z > k)\\
&= (1-r_2^{(1)})^k(1-r_1^{(1)})^k\\
&= ((1-r_2^{(1)})(1-r_1^{(1)}))^k\\
&= (1-(r_1^{(1)}+r_2^{(1)}-r_1^{(1)}r_2^{(1)}))^k
\end{aligned}$$

From which we determine that

$$L \sim Geom(r_1^{(1)}+r_2^{(1)}-r_1^{(1)}r_2^{(1)})$$

Therefore, the distribution of lengths of edges of any order is geometric with parameter $r_1^{(1)} + r_2^{(1)} - r_1^{(1)}r_2^{(1)}$. For mean-zero walks, this simplifies to

$$L \sim Geom(2r^{(1)} - (r^{(1)})^2)$$

□

Corollary 3.14. *The lengths of branches of order k from a mean-zero walk is distributed geometrically with parameter* $\frac{r^{(1)}}{2^{k-1}} - \frac{(r^{(1)})^2}{2^{2k}}$.

Proof. a branch of order k corresponds exactly to a leaf of the walk pruned $k-1$ times. as edges, the distribution of lengths of leaves of walks pruned $k-1$ times is

$$L^(k) \sim Geom(\frac{r^*}{2^{k-1}} - \frac{(r^{(1)})^2}{2^{2k}})$$

□

Theorem 3.15. *The Toknaga indices for the level set tree of a mean-zero random walk is given by* $T_{jk} = \frac{2^{k-j}}{\left(2-\frac{r^{(1)}}{2^k}\right)}$

Proof. Let $\{e_n^k\}^{(j)}$ be a complete sequence of n edges with order k, such that every edge has a parent/child relationship to the one before and after it, after j prunings. In other words, $\{e_n^k\}^{(j)}$ is an order k branch after j prunings. Let N be the distribution of number of edges in $\{e_n^k\}^{(j)}$, let S be the distribution of number of side branches of any order between each edge, and let T be the distribution of side branches emerging from the top edge. If E_k^j is the distribution of number of side branches in $\{e_n^k\}^{(j)}$, then

$$E(E_k^j) = (E(N)-1)(E(S)) + E(T)$$

The expected total length of $\{e_n^k\}^{(j)}$ is given by

$$E(L^{(k)}) = \frac{1}{\frac{r^{(1)}}{2^{k-1}} - \frac{(r^{(1)})^2}{2^{2k}}}$$

The expected length of $\{e_n^k\}^{(j)}$'s edges is given by

$$E(L^{(j)}) = \frac{1}{\frac{r^{(1)}}{2^{j-1}} - \frac{(r^{(1)})^2}{2^{2j}}}$$

Then is is simple to calculate

$$E(N) = \frac{E(L^{(k)})}{E(L^{(j)})} = \frac{\frac{1}{\frac{r^{(1)}}{2^{k-1}} - \frac{(r^{(1)})^2}{2^{2k}}}}{\frac{1}{\frac{r^{(1)}}{2^{j-1}} - \frac{(r^{(1)})^2}{2^{2j}}}} = \frac{\frac{r^{(1)}}{2^{j-1}} - \frac{(r^{(1)})^2}{2^{2j}}}{\frac{r^{(1)}}{2^{k-1}} - \frac{(r^{(1)})^2}{2^{2k}}} = \frac{\frac{1}{2^{j-1}} - \frac{r^{(1)}}{2^{2j}}}{\frac{1}{2^{k-1}} - \frac{r^{(1)}}{2^{2k}}}$$

let $s_i^{k(j)}$ be the likelihood that a vertex has exactly one child of order k and i children of order less than k and greater than j. Immediately, we have that

$$s_i^{k(j)} = q_{i+1}^{(j)} H_k (d_j^k)^i$$

Where d_j^k is the likelihood that any given child of an order k edge is of order less than k and greater than j. Consider any edge in $\{e_n^k\}^{(j)}$ before the last. It is given that there is at least two children, one of order k, as $\{e_n^k\}^{(j)}$ continues. it is also given that all other children will be of orders below k, and as there have been j prunings, all children will be of order greater than j. Therefore, the likelihood that any given non-terminal edge of $\{e_n^k\}^{(j)}$ has i children is given by

$$P(S=i) = \frac{s_i^{k(j)}}{(1-q_0^{(j)})H_k(d_j^k)^i} = \frac{q_{i+1}^{(j)}}{(1-q_0^{(j)})}$$

So the expected number of side children from an edge is given by

$$
\begin{aligned}
E(S) &= \sum_{i=2}^{\infty} \frac{(i)q_{i+1}^{(j)}}{(1-q_0^{(j)})} \\
&= \sum_{i=1}^{\infty} \frac{(i-1)q_i^{(j)}}{(1-q_0^{(j)})} \\
&= \frac{1}{1-q_0^{(j)}} \sum_{i=1}^{\infty} iq_i^{(j)} - \sum_{i=1}^{\infty} r^{(1)}(1-r^{(1)})^{i-1} \\
&= \frac{q_0^{(j)}}{1-q_0^{(j)}}
\end{aligned}
$$

It has been shown that

$$q_0 = \frac{1}{2-r^{(1)}}$$

so

$$q_0^{(j)} = \frac{1}{2-r^{(j+1)}} = \frac{1}{2-\frac{r^{(1)}}{2^j}}$$

giving

$$
\begin{aligned}
E(S) = \frac{q_0^{(j)}}{1-q_0^{(j)}} &= \frac{\frac{1}{2-\frac{r^{(1)}}{2^j}}}{1-\frac{1}{2-\frac{r^{(1)}}{2^j}}} \\
&= \frac{1}{1-\frac{r_1^{(1)}}{2^j}}
\end{aligned}
$$

finally, let $t_i^{k(j)}$ be the likelihood that a vertex of a tree pruned j times has 2 children of order k, as well as i children with order greater than j and less than or equal to k. Immediately, we calculate

$$t_i^{k(j)} = q_{i+2}H_k^2(d_j^{k+1})^i$$

One knows that, for the top edge, it is given that there will be at least 2 branches of order k, and that all other branches will be of orders between j and k, with k inclusive, therefore, the likelihood that any given non-terminal edge of $\{e_n^k\}^{(j)}$ has i children is given by

$$P(T=i) = \frac{t_i^{k(j)}}{(1-q_0^{(j)})H_k^2(d_j^{k+1})^i} = \frac{q_{i+2}^{(j)}}{(1-q_0^{(j)})}$$

We then find

$$
\begin{aligned}
E(T) &= \sum_{i=2}^{\infty} \frac{(i)q_{i+2}^{(j)}}{(1-q_0^{(j)})} \\
&= \sum_{i=1}^{\infty} \frac{((i-1)-1)q_i^{(j)}}{(1-q_0^{(j)})} \\
&= \sum_{i=1}^{\infty} \frac{((i-1))q_i^{(j)}}{(1-q_0^{(j)})} - \sum_{1}^{\infty} \frac{q_i^{(j)}}{(1-q_0^{(j)}))} \\
&= E(S) - 1
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
E(E_k^j) &= (E(N)-1)(E(S)) + E(T) \\
&= \frac{\frac{1}{2^{j-1}} - \frac{r^{(1)}}{2^{2j}}}{\frac{1}{2^{k-1}} - \frac{r^{(1)}}{2^{2k}}} \frac{1}{1-\frac{r^{(1)}}{2^j}} - 1
\end{aligned}
$$

To find the number of side branches of order j merging into a branch of order k, calculate

$$
T_{jk} = E(E_k^{j-1}) - E(E_k^j)
$$

This works because we are calculating the number of side branches with order from j to $k-1$, and subtracting away the number of side branches with orders from $j+1$ to $k-1$. this is equal to

$$
\begin{aligned}
T_{jk} &= \frac{2^k}{2^j} \frac{\frac{1}{2^{-2}} - \frac{r^{(1)}}{2^{j-2}}}{\frac{1}{2^{-1}} - \frac{r^{(1)}}{2^k}} \frac{1-\frac{r^{(1)}}{2^j}}{(1-\frac{r^{(1)}}{2^j})(1-\frac{r^{(1)}}{2^{j-1}})} - \frac{2^k}{2^j} \frac{\frac{1}{2^{-1}} - \frac{r^{(1)}}{2^j}}{\frac{1}{2^{-1}} - \frac{r^{(1)}}{2^k}} \frac{1-\frac{r^{(1)}}{2^{j-1}}}{(1-\frac{r^{(1)}}{2^j})(1-\frac{r^{(1)}}{2^{j-1}})} \\
&= 2^{k-j} \left(\frac{\left(2^2 - \frac{2^2 r^{(1)}}{2^j}\right)\left(1-\frac{r^{(1)}}{2^j}\right) - \left(2-\frac{r^{(1)}}{2^j}\right)\left(1-\frac{2r^{(1)}}{2^j}\right)}{\left(2-\frac{r^{(1)}}{2^k}\right)(1-\frac{r^{(1)}}{2^j})(1-\frac{2r^{(1)}}{2^j})} \right) \\
&= 2^{k-j} \left(\frac{2 - \frac{3r^{(1)}}{2^j} + \frac{2(r^{(1)})^2}{2^{2j}}}{\left(2-\frac{r^{(1)}}{2^k}\right)(1-\frac{r^{(1)}}{2^j})(1-\frac{2r^{(1)}}{2^j})} \right) \\
&= 2^{k-j} \left(\frac{(1-\frac{r^{(1)}}{2^j})(1-\frac{2r^{(1)}}{2^j})}{\left(2-\frac{r^{(1)}}{2^k}\right)(1-\frac{r^{(1)}}{2^j})(1-\frac{2r^{(1)}}{2^j})} \right) \\
&= \frac{2^{k-j}}{\left(2-\frac{r^{(1)}}{2^k}\right)}
\end{aligned}
$$

$\square$

4. Conclusion

In the future, we wish to refine the contents of this paper and publish in a scholarly journal. before this is done, we will have to check our results for Tokunaga and Horton self similarity. We also wish to explore the ties between these results and abstract algebra. Finally, we are curious as to how similar results may be found for other symmetric discrete random walks.

References

[1] Yevgeniy Kovchegov, Ilya Zaliapin *Random Self-Similar Trees and a Hierarchical Branching Process*. preprint, 2017.

[2] Neveu, J *Renewal property of the extrema and tree property of the excursion of a one-dimensional brownian motion*. Séminaire de Probabilités XXIII, 1989.

[3] Le Gall, Jean-François *The uniform random tree in a Brownian excursion*. Probability Theory and Related Fields, 1993.

[4] Jim Pitman *Combinatorial Stochastic Processes*. Lecture notes for St. Flour Course, 2002.

Appendix

California State University Long Beach
E-mail address: claire_campregher@aol.com

Oregon State University
E-mail address: johnsoj3@oregonstate.edu