UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2011:29

Brownian Motions and Scaling Limits of Random Trees

Hady Ahmady Phoulady

Examensarbete i matematik, 30 hp
Handledare och examinator: Takis Konstantopoulos
November 2011

Department of Mathematics
Uppsala University

I would like to dedicate this thesis to my brother, Rasta, who is more than a brother to me. Guiding me through my studies and supporting me by any means is the least he does!

Abstract

Brownian motions as continuous time stochastic processes are scaling limit of different mathematical objects in distribution. When a polygon triangulation tends to a triangulation of the circle, a simple random walk tends to normalized Brownian motion in distribution. A similar thing can be argued for plane, real and labeled trees. The aim of this thesis is to define and check the random triangulation of the circle in more details and find the scaling limit and distribution of random trees above.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis is mainly based on an article[2] by *David Aldous* and some notes[12] by *Jean-François Le Gall*.

We will start by discussing Catalan numbers, one of the most famous and frequently occurring sequences in Combinatorics. There exist many different examples where Catalan numbers appear when we try to count a set. Among these sets, we check three sets more in depth and use its results later.

In Chapter 3, we will state several basis definitions for the rest of the thesis such as Hausdorff dimension and Brownian motions. We show that how the random triangulation of the circle can be encoded by normalized Brownian excursions.

In the other two Chapters we will discuss about different classes of discrete trees, like plane and labeled trees and also real trees and their scaling limits will be driven.

In this thesis it has been tried to provide the proof of every statement we claim. In many cases we provided proofs if there were no proofs available in two text mentioned above and also in several cases, although a proof was provided, it has been tried to present a new proof.

Chapter 2

Catalan Numbers

Amongst the set of infinite sequences of positive integers we can mention some simple and obvious series like doubling series $(1, 2, 4, 8, 16, ...)$ or the squares $(1, 4, 9, 16, 25, ...)$ or some others that a few mathematicians would fail to recognize like the Fibonacci numbers $(1, 1, 2, 3, 5, 8, ...)$ or the triangular numbers $(1, 3, 6, 10, 15, 21, ...)$. In case of an unfamiliar sequence, however, we may have to spend an enormous amount of time to find a recursive or non-recursive formula that generates the sequence[13].

When someone encounters an infinite sequence of positive integers, a possible way is to look it up in *A Handbook of Integer Sequences*[20]. In that handbook at the top of the page 71, the 557-th sequence is the sequence that we will talk about it in this chapter, the sequence of the *Catalan numbers*: $1, 2, 5, 14, 42, 132, 429, 1430, ...$

2.1 A Brief History

Around 1751, *Leonard Euler* found the Catalan numbers[15] after asking himself: In how many ways can we divide a fixed convex polygon into triangles by drawing diagonals that do not intersect[13]? Denote the set of different polygon triangulations by S_1. The first few examples can be seen in Figure 2.1.

Let us suggest a question whose solution generates the sequence of Catalan numbers too. What are the number of well-formed sequences of parentheses?

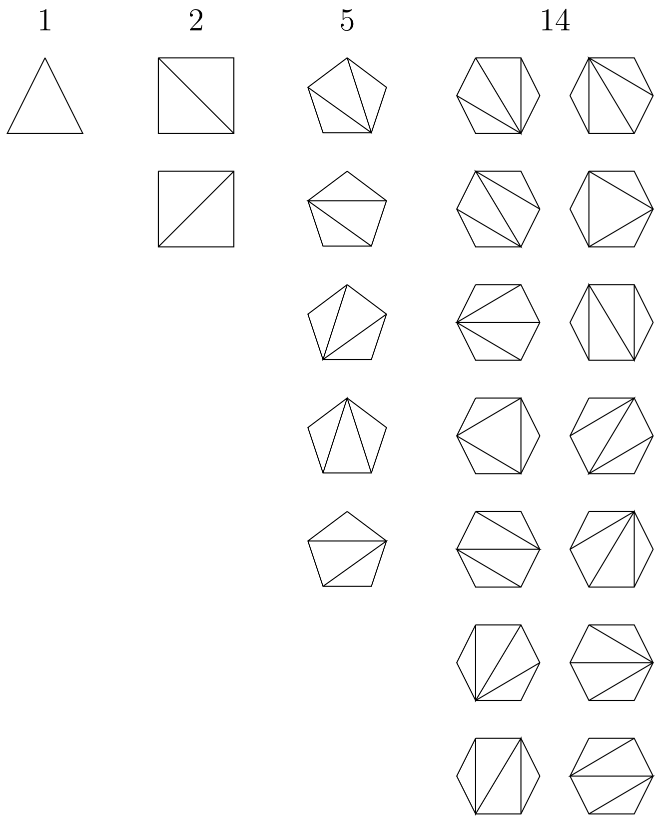

Figure 2.1: Euler's polygon triangulation problem.

"Well-formed" means that each open parenthesis has a matching closed parenthesis. For example amongst different sequences of $n = 3$ pair of open and closed parentheses, "()(())" is well-formed while "(()))(" is not. Denote the set of well-formed sequences by S_2. Look at Table 2.1 for the smallest examples of such sequences.

Now let us define the "valid" sequences of parentheses with l letters as below.

Suppose that we have a chain of l letters in a fixed order. We want to add $l-1$ pairs of parentheses so that each pair of matched parentheses contain exactly two "parts". These parts can be two adjacent letters, a letter and an adjacent parenthetical grouping or two adjacent parenthetical groupings. We call these sequences valid sequences of parentheses with l letters and denote its set by S_3. The examples of such sequences of parentheses with 2, 3, 4 and 5 letters are shown in Table 2.2.

Table 2.1: The first smallest well-formed sequences of parentheses.

Number	1	2	5	14	
	()	(())	((()))	(((())))	()()(())
		()()	(()())	((()()))	()()()()
			()(())	(()(()))	()(())()
Sequences			()()()	(()()())	(())(())
			(())()	((())())	(())()()
				()((()))	((()))()
				()(()())	(()())()
n	1	2	3	4	

Table 2.2: Examples of valid sequences of parentheses with 2, 3, 4 and 5 letters.

Number	1	2	5	14	
	(ab)	$(a(bc))$	$(a(b(cd)))$	$(a(b(c(de))))$	$(((ab)c)(de))$
		$((ab)c)$	$(a((bc)d))$	$(a(b((cd)e)))$	$((((ab)c)d)e)$
			$((ab)(cd))$	$(a((bc)(de)))$	$(((ab)(cd))e)$
Sequences			$(((ab)c)d)$	$(a(((bc)d)e))$	$((a(bc))(de))$
			$((a(bc))d)$	$(a((b(cd))e))$	$(((a(bc))d)e)$
				$((ab)(c(de)))$	$((a(b(cd)))e)$
				$((ab)((cd)e))$	$((a((bc)d))e)$
l	2	3	4	5	

In 1838, Belgian mathematician *Eugene C. Catalan* discovered Catalan numbers while studying valid sequences of parentheses with l letters.

2.2 Isomorphism Between S_1, S_2 and S_3.

Proposition 2.1. *There is an injection from* S_3 *to* S_2*.*

Proof. [6] Let us define an injection from the set of well-formed sequences of n pair of parentheses to the set of valid sequences of parentheses with $l = n + 1$ letters with a simple example. Suppose that we have a sequence of parentheses with $l = 5$ letters, like $((ab)(c(de)))$. First put a dot between 2 parts of each pair of matched parentheses and get the sequence $((a.b).(c.(d.e)))$. Now delete all open parentheses and letters and get .)...))). Finally put open parentheses instead of each of the dots and get ()(((()))). It is easy to see that we get a well-formed sequence.

The valid sequences in Table 2.2 are arranged in the same order as their corresponded well-formed sequences in Table 2.1. □

Proposition 2.2. *There is an injection from* S_2 *to* S_1*.*

Proof. We define an injection from well-formed sequence of parentheses with n pair of parentheses to the triangulation of polygons with $n+2$ sides.

Suppose that we have ()(((()))) (a well-formed sequence with $n = 4$ pair of parentheses). Start from a corresponded regular polygon which is a hexagon in our example. Set the base node the lower left node. Starting from 1, assign numbers to other nodes clockwise from the second node after the base node. Also assign numbers to the closed parentheses in the sequence from left to right. Consider the assigned number of the matched close parenthesis of the first open parenthesis.

If it is between 1 and $n-1$, then draw a diagonal between the base node and the node with the same assigned number. By this diagonal the original polygon will be divided into 2 polygons. For the first polygon, base node remains the same. For the second polygon make the other node (the other end of the diagonal) the base node. Also divide the sequence to 2 subsequences, the first subsequence starts from the first parenthesis to the matched closed parenthesis (with the assigned number) and the other subsequence the rest of the sequence. Assign them respectively to the new and old polygons and do the same procedure for them by induction.

If the matched close parenthesis is the n-th closed parenthesis, then draw a diagonal between the nodes to the left and the right of the base node, make the next node of base node (clockwise) the base node of the new polygon. Remove the first and last parentheses from the original sequence and do the same procedure by induction for the new polygon and with the new sequence.

The basis for the induction step is of course the single pair of parenthesis which is equivalent to a triangle itself. For clarification, look at the Figure 2.2. The sequences in Table 2.1 are arranged in the same order as their corresponded polygon triangulations in Figure 2.1. □

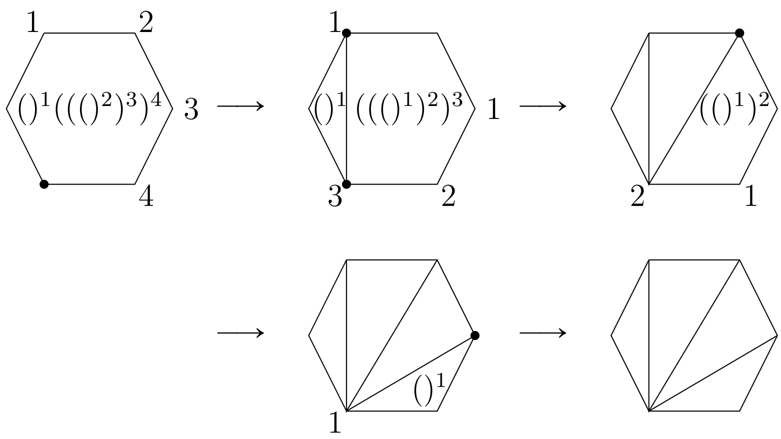

Figure 2.2: The process of getting a polygon's triangulation from a well-formed sequence of parentheses.

Proposition 2.3. *There is an injection from* S_1 *to* S_3.

Proof. [13] In 1961, *H. G. Forder* showed a simple way to prove a one-on-one correspondence between the triangulated polygons with n sides and the valid sequences of parentheses with $l = n - 1$ letters. Let us describe the injection part of it with a simple example on a hexagon.

Except the base side, label the other sides by letters a, b, c, d, e. Each diagonal spanning the the adjacent sides is labeled with the letters of those side in parentheses. The other diagonals are then labeled in similar fashion by combining the labels on the other two sides of the triangle. The base is labeled last. Look at Figure 2.3. □

Propositions 2.1, 2.2 and 2.3 show that these sets are isomorphic. In fact a lot of other seemingly unrelated sets are isomorphic to S_i's which we will give a few examples in Section 2.4.

2.3 The Values of C_n

But what can we say about the value of C_n, the n-th Catalan number?

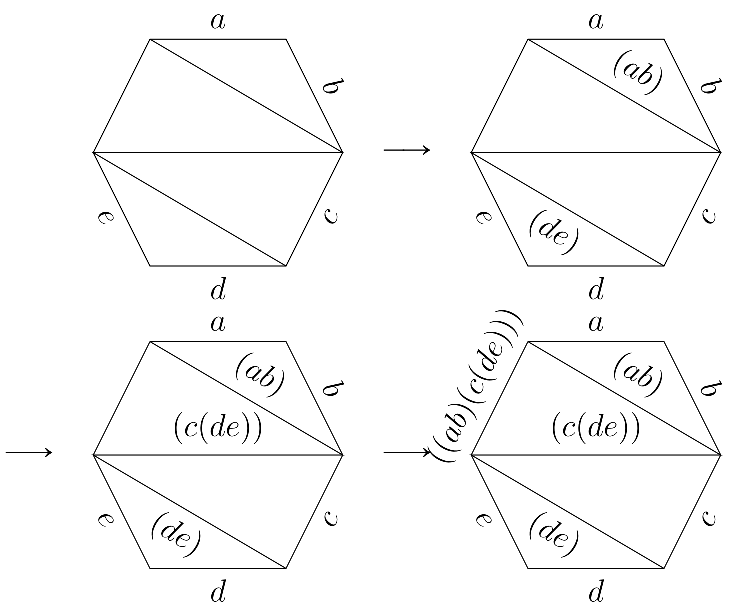

Figure 2.3: A valid sequence of parenthesis with 5 letters driven from a hexagon triangulation.

For getting more simpler recursive formulas, usually they add a "1" to the first of the sequences. If we do that with the sequence of Catalan numbers, we get the sequence $(1, 1, 2, 5, 14, 42, 132, 429, ...)$. Put $\mathrm{C}_0 = 1, \mathrm{C}_1 = 1, \mathrm{C}_2 = 2, \mathrm{C}_3 = 5, \mathrm{C}_4 = 14, ...$

Consider the polygon triangulation problem. Let us try to count the number of different triangulations of a polygon with $n > 3$ sides recursively. The first diagonal can be drawn between any two nodes with at least one node between them. This diagonal will divide the polygon into two smaller polygons one with $k+1$ sides and the other with $n-k+1$ sides where $1 < k < n-1$. Each of these new polygons can be triangulated independently. Also easily we can guess than the number of different triangulations of a polygon with n sides is equal to C_{n-2}, then get the following recursive formula.

$$\mathrm{C}_n = \sum_{i=0}^{n-1} \mathrm{C}_i \mathrm{C}_{n-i-1}. \tag{2.1}$$

On the other hand, if we denote the number of ways of triangulating a polygon with n sides by T_n, *Euler*, using an inductive argument that he described as "quite laborious" established that[15]

$$T_n = \frac{2 \times 6 \times 10 \times \cdots \times (4n-10)}{(n-1)!},$$

where $n \geqslant 3$. Also $T_n = \mathrm{C}_{n-2}$, so

$$\begin{aligned}
T_{n+2} &= \mathrm{C}_n \\
&= \frac{2 \times 6 \times 10 \times \cdots \times (4n-2)}{(n+1)!} \\
&= \frac{4n-2}{n+1} \times \frac{2 \times 6 \times 10 \times \cdots \times (4n-6)}{n!} \\
&= \frac{4n-2}{n+1} T_{n+1};
\end{aligned}$$

and now we conclude that

$$\mathrm{C}_n = \frac{4n-2}{n+1}\mathrm{C}_{n-1}.. \tag{2.2}$$

From recursive equation 2.2, we can derive an explicit formula for Catalan numbers:

$$\begin{aligned}
\mathrm{C}_n &= \frac{4n-2}{n+1}\mathrm{C}_{n-1} \\
&= \frac{(4n-2)(4n-6)}{(n+1)n}\mathrm{C}_{n-2} \\
&\vdots \\
&= \frac{(4n-2)(4n-6) \times \cdots \times 6 \times 2}{(n+1)n \times \cdots \times 3 \times 2}\mathrm{C}_0 \\
&= \frac{(2n-1)(2n-3) \times \cdots \times 3 \times 1}{(n+1)!} \times 2^n \\
&= \frac{(2n)! \times 2^n}{(n+1)! \times 2^n \times n!} \\
&= \frac{(2n)!}{(n+1)!n!},
\end{aligned}$$

and it means that

$$C_n = \frac{1}{n+1}\binom{2n}{n}. \tag{2.3}$$

So Catalan numbers can be defined by any recursive formulas in (2.1), (2.2) or the explicit formula in (2.3).

2.4 Some Examples

Here let us present a few problems[6] which introduces sets which are isomorphic to each other and also to sets in Section 2.2.

Example 2.1 (Simple Random Walks)**.** Consider the random walks which consists of n up-walks and n down-walks in such a way that we never go below the horizontal line (see Figure 2.4).

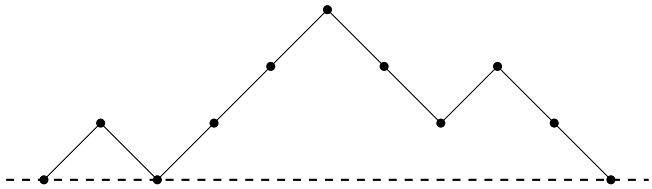

Figure 2.4: A Simple Random Walk with $2n = 10$ Walks.

The number of these random walks are equal to C_n.

Example 2.2 (Hands Across a Table[6])**.** If $2n$ people are seated around a circular table, in how many ways can all of them simultaneously shake hands with another person such that none of their arms cross each other? To see the all possible ways when $n = 3$ look at the Figure 2.5.

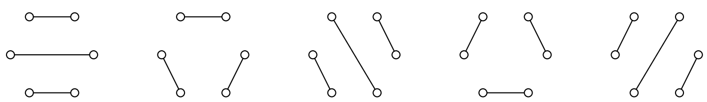

Figure 2.5: Different ways of how $2n = 6$ people around a table can handshake without crossing each other hands.

The number of different handshaking with $2n$ people around table is equal to C_n.

Example 2.3 (Plane Rooted Trees)**.** The set of rooted trees with n edge in which a specific node is root is also isomorphic to the sets defined above.

The number of rooted trees with n edges is equal to C_n.

Example 2.4 (Rooted Binary Trees)**.** The number of rooted binary trees with n internal node (none leaf) which each nod is either a leaf or an internal node with exactly two children also generate the sequence of Catalan numbers.

The number of rooted binary trees with n internal node is equal to C_n.

There are a lot of other examples which produce the sequence of Catalan numbers. In 1971 *Henry W. Gould*, a mathematician at West Virginia University, privately issued a bibliography of 243 references on Catalan numbers. In 1976 he increased the number of references to 450 [13]. In many cases people were not aware that they were dealing with the sequence of Catalan numbers!

In next two chapters, we will talk about the polygon triangulation and plane trees and their limits more in depth.

Chapter 3

Limit of Polygon Triangulation

In Chapter 2 we saw that the number of different triangulation of a polygon with n sides is equal to the $(n-2)$-th Catalan number,

$$\mathrm{C}_{n-2} = \frac{1}{n-1}\binom{2(n-2)}{n-2}.$$

In this chapter we will discuss in more depth about the set of different triangulations of a polygon and its limit when $n \to \infty$.

3.1 Random Triangulation of the Circle

When $n \to \infty$, a regular polygon with n sides will converge to a circle; so we can consider that it is a triangulation of a circle! Based on this convergence, a random triangulation of a regular polygon with n sides when $n \to \infty$, can be considered as a *random triangulation of the circle*. We consider the definition below for a triangulation of the circle.

Definition 3.1. A triangulation of the circle is a closed subset of the closed disc whose complements is a disjoint union of open triangles with vertices on the circumference of the circle.[2]

The triangulations defined as above are exactly the possible limits of triangulations of n-gons. Before talking more about different triangulations, let us define the *Hausdorff content* and *dimension*.

3.1.1 Hausdorff content and dimension

Definition 3.2. *d-dimensional Hausdorff content* of S is defined by

$$C_H^d(S) := \inf \left\{ \sum_i r_i^d : \text{there is a cover of } S \text{ by balls with radii } r_i > 0 \right\},$$

where S is a subset of a metric space X and $d \in [0, \infty)$.

Suppose that $X \subset \mathbb{R}^n$ and and $\lambda > 0$ then it is easy to see that[16]

$$C_H^d(\lambda X) = \lambda^d C_H^d(X). \tag{3.1}$$

For proving that, suppose that a covering C of X gives the infimum of the set above, $\sum_i r_i^d$ which is equal to C_H^d, then if we replace each ball in C with a ball λ times bigger, the new sum will be $\sum_i (\lambda r_i)^d = \sum_i \lambda^d r_i^d = \lambda^d \sum_i r_i^d = \lambda^d C_H^d$. It proves that $C_H^d(\lambda X) \leqslant \lambda^d C_H^d(X)$. With a similar argument we can prove that $C_H^d(X) \leqslant \frac{1}{\lambda^d} C_H^d(\lambda X)$; and together we conclude (3.1).

Definition 3.3. *Hausdorff dimension* of X is defined by

$$\dim_H(X) := \inf \left\{ d \geqslant 0 : C_H^d(X) = 0 \right\}.$$

With the definition of Hausdorff dimension, we can easily observe that[10]

$$C_H^d(X) = \begin{cases} \infty & \text{if } 0 \leqslant d < \dim_H(X) \\ 0 & \text{if } d > \dim_H(X) \end{cases}. \tag{3.2}$$

But when $d = \dim_H(X)$, $C_H^d(X)$ is not determined. In fact it can be either 0 or infinity or may take any value between 0 and infinity.

In general finding the Hausdorff dimension of a space directly is rather a hard work! The calculating is usually done by using some basic techniques that are available for dimension calculations. For example the equation 3.1 can be used sometimes to establish an upper or a lower bound for the Hausdorff dimension.

Let us explain a simple example in which we calculate the Hausdorff dimension directly. We prove that the Hausdorff dimension of the interval $[0, 1]$ is 1.

We define different coverings of the interval $[0, 1]$ for each $k = 1, 2, ...,$ and denote them by C_k as follows. C_k is a covering consists of k segments $[\frac{i}{k}, \frac{i+1}{k}]$ for each $i = 0, 1, ..., k-1$. Now let us calculate $\sum_i r_i^d$ for a specific k and an arbitrary d:

$$\sum_i r_i^d = \sum_{i=0}^{k} (\frac{1}{k})^d = \frac{k}{k^d} = k^{1-d},$$

so when $k \to \infty$, the sum goes to 0 if $d > 1$ and goes to infinity if $d < 1$ and according to (3.2) we conclude that the Hausdorff dimension of the interval $[0, 1]$ is 1.

Other examples can be a countable set, circle and $\mathbb{R}^n$ which have Hausdorff dimension respectively 0, 1 and n.

3.1.2 Two Examples of Different Triangulations

Let us now give two examples of different triangulations of an n-gon.

Assign numbers $1, 2, ..., n$ respectively to nodes of the n-gon. We show a triangulation of this n-gon by a set of pairs of nodes. An arbitrary triangulation can be the set

$$T_1 = \{(1, 2), (1, 3), (1, 4), \cdots, (1, n)\}.$$

Clearly when $n \to \infty$, T_1 will be the whole interior of the circle and will have the dimension 2.

Another example can be

$$T_2 = \left\{ \left(\frac{n}{2}, n\right), \left(n, \frac{n}{4}\right), \left(\frac{n}{4}, \frac{n}{2}\right), \left(\frac{n}{2}, \frac{3n}{4}\right), \left(\frac{3n}{4}, n\right), \left(n, \frac{n}{8}\right), \left(\frac{n}{8}, \frac{n}{4}\right), \cdots \right\},$$

which can be considered as a portion of a straight line and has Hausdorff dimension 1.

Although the triangulations above have Hausdorff dimension 1 and 2, but it turns out that the limit of random triangulation of the circle has Hausdorff dimension $\frac{3}{2}$ almost surely[2]. We will show it in Section 3.5.2.

Also in any random triangulation of the circle, the length of the longest chord is at most the diameter l_0 of the circle and at least the length l_1 of the side of an inscribed equilateral triangle. Noting this, we may ask ourselves the question below[2].

Question 3.1. *In a random triangulation of the circle, what is the chance that a longest chord has length greater than* $(l_0 + l_1)/2$*?*

This question is phrased to resemble the well known *Bertrand's paradox*.

Question 3.2. *What is the chance that a random chord in the circle has length greater than* l_1*?*

This is a paradox because as *Martin Gardner* explained[14] (Chapter 19), we can get at least three different answers by three equally plausible calculations. The point is that here *randomness* has no canonical meaning. There are several different mechanisms for physically drawing a chord in some ways influenced by chance and these different mechanisms, mathematically, lead to different *probability measures* on the set of chords. The same is about the notation of *a random* triangulation of the circle. For solving this problem of ours, we use the measure which is the limit of uniform random triangulations of n-gons and so then we will need to prove the *existence* of such a limit.

3.2 Continuous Functions and Triangulations of the Circle

Consider a continuous function $f : [0, 1] \to [0, \infty)$ which satisfies

$$\begin{aligned} &f(0) = f(1) = 0, \\ &f(t) > 0 \text{ for } 0 < t < 1. \end{aligned} \tag{3.3}$$

We explain a simple way to establish a mapping from these kind of functions to triangulations of the circle.

Suppose t_2 is a strict local minimum of f, that is $f(t_2) < f(t)$ for all $t \neq t_2$ in some neighborhoods of t_2. Amongst these neighborhoods, suppose that (t_1, t_3) is the largest one. By continuity we observe that $f(t_1) = f(t_2) = f(t_3)$. Now regard the interval $[0, 1]$ as the circumference of the circle and draw a triangle with vertices t_1, t_2 and t_3. Do that for each strict local minimum t'_2. If $f(t'_2) > f(t_2)$

(the same goes for the case $f(t_2') < f(t_2)$) then no matter where t_2' lies (in which arc of arcs (t_1, t_2), (t_2, t_3) or (t_3, t_1)), also t_1' and t_3' lie in the same arc and thus triangles are disjoint. If $f(t_2') = f(t_2)$ then 0, 1 or 2 of the points t_i''s are equal to one of t_i's. If none of them or only one of them are equal to each other, then it would be almost like before. In the case when 2 of them are equal to t_i's, then we would have some chords crossing each other and so we assume that this case never happens with our specific function f. So we can define our triangulation to be the complement of the union of all the open triangles associated with the local minimums.

By some functions we may get a finite number of triangles and thus our triangulation will have a non-zero area, but there exist some functions f with the property that the set of strict local minimums of f is dense in $[0, 1]$. By these kind of functions our triangulation will have no non-zero area.

This mapping from continuous functions in $[0, 1]$ to triangulations of the circle is useful because with using it we can define the *random* triangulation of the circle indirectly by first defining random functions and then use the mapping. Random functions, on the other hand, are actually *stochastic processes* which are well known and thus with the mapping we related a well studied subject to our new subject! Also this mapping is in fact the continuous analog of the mapping from discrete walks to triangulations of the n-gons which we will talk about in next section.

3.3 Walks, Trees and Triangulations of n-gons

Let us first define four sets:

- S_1 := Set of positive (except the two ends of the walk) walks with steps +1 or -1 and length $2n$. For an example see Figure 3.1.
- S_2 := Set of rooted (and ordered) plane trees with $n - 1$ edges. For an example see Figure 3.2.
- S_3 := Set of binary trees with $n - 1$ nodes. For an example see Figure 3.3.
- S_4 := Set of triangulations of $(n + 1)$-gons. For an example see Figure 3.4.

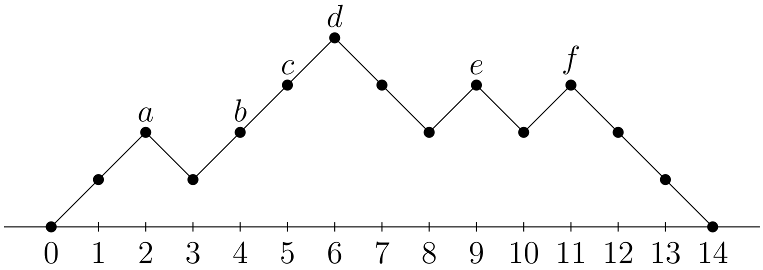

Figure 3.1: A Positive Walk with $2n = 14$ Walks.

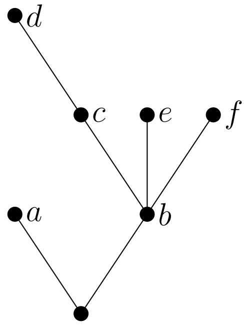

Figure 3.2: A Plane Rooted (and Ordered) Tree with $n - 1 = 6$ Edges.

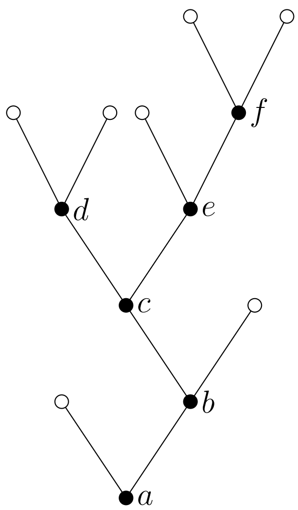

Figure 3.3: A Binary Tree with $n - 1 = 6$ Nodes.

For each $i = 1, 2, 3$ we will present a one-to-one mapping from $\mathrm{S_i}$ to $\mathrm{S_{i+1}}$. In fact with these mappings we can get the ordered tree in Figure 3.2 from the positive walk in Figure 3.1, the binary tree in Figure 3.3 from the ordered tree in

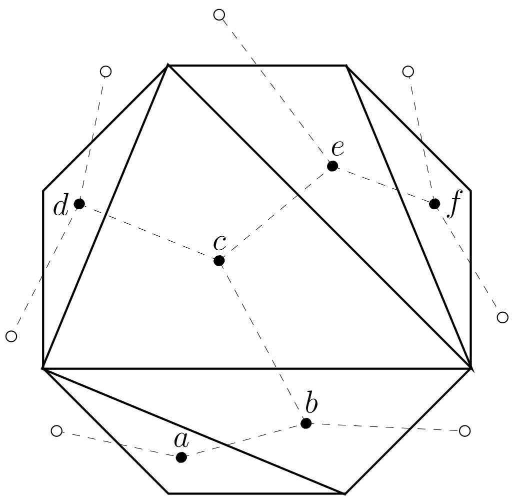

Figure 3.4: A Triangulation of the Octagon.

Figure 3.2 and the triangulation of octagon in Figure 3.4 from the binary tree in Figure 3.3, and vice versa!

Mapping 1. This mapping is from S_1 to S_2 and vice versa. Consider a positive walk. For the first +1 walk (the first walk is surely +1), we draw the root. After that for each +1 walk we draw a node from current node and for each -1 walk we go back to previous node that we came from (the parent of current node). Each new node from a node is drawn to the right of the other nodes.

Look at the walk in Figure 3.1. This mapping takes the points a, b, c, d, e and f in the walk to corresponding nodes with the same labels in the ordered tree in Figure 3.2.

For producing a walk from an ordered tree, we start from root and then visit children of each node from left to right (for each new node we first visits its children and then go to visit its siblings on its right). This will be actually a *depth-first search* algorithm which traverses the tree. In this process, for each up movement, we draw a +1 walk and vice versa.

So this mapping is a one-to-one correspondence between S_1 and S_2.

Mapping 2. This mapping is from S_2 to S_3 and vice versa. Consider an ordered plane tree with a root. Start from the root. Use the depth-first search algorithm above to traverse the ordered tree. For the first move (from root to its leftmost child), draw a node. After that if you are visiting a new node that has not a left sibling in ordered tree, then draw a left edge from its parent's mapped node in binary tree which produces a new node who will be the mapped node for current node in ordered tree. On the other hand, if you are visiting a new node which has a left sibling, then draw a right edge from its left sibling's mapped node in binary tree and map it to this new node. In the end until there exists a node in binary tree that has not two children, draw a new edge from it and add it to a leaf. So the left/right child of a node in binary tree will be a leaf if its corresponded node in ordered tree has no child/sibling. In Figure 3.3, these new nodes are specified with white circles.

With this mapping, each node in an ordered tree will be mapped to an internal node in binary tree. Take a look at the ordered tree in Figure 3.2 and its mapped binary tree in Figure 3.3 and note the labels of the nodes.

On the other hand, from a binary tree we can get an ordered tree briefly as follows. Use depth-first search algorithm again and when you arrive to a new node which is not a leaf, then if it is the left child of its parent (in binary tree), draw a new edge from the corresponded node of its parent and if it is the right child, draw a new edge from the parent (in ordered tree) of the corresponded nod of its parent (in binary tree). New edges from each node should be drawn in a way that the most new edge be on the right of other edges on the node.

In this way we established a mapping which is a one-to-one correspondence between S_2 and S_3.

Mapping 3. This mapping is from S_3 to S_4 and vice versa. Consider a binary tree with $n-1$ nodes and a regular $(n+1)$-gon. Choose a base side s in the $(n+1)$-gon. This side should be side of a triangle in the final triangulation. This triangle will be specified by the number of internal nodes to the left and right of the root of the binary tree. In our example of triangulation in Figure 3.4 and binary tree in Figure 3.3, the base side s is the bottom horizontal side and these numbers are 0 and 5. So the triangle should be drawn so that the number of polygon's node on the right side of the triangle be 5 and this number of left be 0. The edges from the root are drown so that they cross the two other sides of the triangle (other than the side s). When an edges is drawn in a way that crosses a side of the polygon, it denotes a leaf. After this, the part of the binary tree on the right/left of the root (except any leaves) is drawn with the same process on the right/left of the first triangle.

Finally in fact each chord in triangulation represents an edge in binary tree which is between two internal nodes and each side, except the base s, represents an edge to a leaf. Also each internal node will be inside a specific triangle in final triangulation.

The reverse mapping for getting the binary tree from a triangulation is much simpler. First we put $n-1$ nodes inside the $n-1$ triangles and connect those two nodes to each other whose triangles are adjacent. The root is the node inside the triangle containing the base side s. Also its the base side s that determines which child is a right child and which is a left child. Also the leaves are clearly connected to each node whose triangle has a side in the set of polygon's sides!

So this is in fact a one-to-one correspondence between S_3 and S_4.

3.4 Brownian Motions

Consider an arbitrary (not necessarily positive or starting from zero) random walk with length m (consisted of m walks with steps +1 or -1). Suppose that we want to scale it and draw it on a paper with width w and an unlimited height. Then if we want to fit the walk to the paper, the width of each walk should be $\frac{w}{m}$ and thus the steps should be $+\frac{w}{m}$ or $-\frac{w}{m}$. To be sure that we can draw the scaled walk on paper, we are in need of a height of paper equal to $2w$, (from $+w$ to $-w$ in case if all steps in original walk are +1 or all are -1) but it turns out that in general the height of the original walk is of order $\pm\sqrt{m}$ and thus for the scaled walk, generally we are in need of a height of order $2\frac{w}{\sqrt{m}}$.

In the limit, when $m \to \infty$ we will have a path that we call it *Brownian motion*. If the original random walk is constrained to be positive at all times except at the two ends which is zero and also we put $w = 1$, then the path is constrained to satisfy (3.3) and it is called *normalized Brownian excursion*. In the case that w is any arbitrary value the path is simply called *Brownian excursion*.

Now we can prove the following Proposition (with of course skipping a lot of technicalities).

Proposition 3.1. *Random triangulation of the circle is the limit of uniform* *random triangulation of polygon.*

Proof. By mappings defined in Section 3.3, we indeed showed that there exists a one-to-one correspondence between constrained random walks and triangulations of polygons. On the other hand, by applying the mapping (from continuous functions to triangulation of the circle) in Section 3.2 to normalized Brownian motions we get a random triangulation of the circle. So the proof is complete, because normalized Brownian motions are the limits of constrained random walks. □

Also we can say that there exists a mapping from normalized Brownian excursions to triangulations of the circle and vice versa. Because they are respectively limits of constrained random walks and triangulation of polygons and the last two sets have a one-to-one correspondence between each other.

3.5 Random Triangulation of the Circle, Revisited

Let us first talk more about Brownian motions.

3.5.1 Zero Set of Brownian Motion

Let us first present and prove a fundamental principle.

Theorem 3.1 (Mass Distribution Principle)**.** *If $A \subset X$ supports a positive Borel measure μ such that $\mu(D) \leqslant C|D|^d$ for any Borel set D, then $C_H^d(A) \geqslant \frac{\mu(A)}{C}$ and hence* $\dim_H(A) \geqslant d$.

Proof. Consider a covering of A, like $\bigcup_j A_j$, then

$$\sum_j |A_j|^d \geqslant C^{-1} \sum_j \mu(A_j) \geqslant C^{-1}\mu(A),$$

and thus $C_H^d(A) \geqslant \frac{\mu(A)}{C}$ which means that $C_H^d(A)$ is positive and therefore $\dim_H(A) \geqslant d$. □

It is known that Brownian motion's transition kernel $p(t, x, \cdot)$ has $N(x, t)$ distribution. In general, if Z has $N(x, t)$ distribution we define $|N(x, t)|$ to be the distribution of $|Z|$.

Next Theorem is known as *Lévy's identity*.

Theorem 3.2 (Lévy, 1984)**.** *Let M_t be the maximum process of a one dimensional Brownian motion B_t, i.e. $M_t = max_{0 \leqslant s \leqslant t} B_s$. Then, the process $Y_t = M_t - B_t$ is Markov and its transition kernel $p(t, x, \cdot)$ has $|N(x, t)|$ distribution*[17].

We will not provide a proof for this theorem, but note that it actually states that $Y_t = M_t - B_t$ and $|B_t|$ has the same distribution.

Now let us define *zero set* and also *record time* which is in fact zero set of Y_t defined above.

Definition 3.4. We denote *Zero set* of a Brownian motion B_t by $\mathcal{Z}_B$ and define it as

$$\mathcal{Z}_B = \{t \geqslant 0 : B_t = 0\}.$$

Definition 3.5. We call a time t a *record time* for Brownian motion B_t if it is a zero of Y_t, i.e. $Y_t = M_t - B_t = 0$. In other words, t is a record time if it is a global maximum from left.

Although almost surely Brownian motions have isolated zeros from left (first zero after a specific time) or from right, but zero set of a Brownian motion is an uncountable closed set with no isolated point with probability one[17]!

Before presenting and proving the next Lemma, let us define *Hölder continuity.*

Definition 3.6. A function f defined on $\mathbb{R}$ is Hölder continuous with exponent α if there exists a constant, C_α, such that

$$\forall x, y : |f(x) - f(y)| \leqslant C_\alpha |x - y|^\alpha.$$

Lemma 3.1. $\dim_H(\mathcal{Z}_B) \geqslant \frac{1}{2}$ *with probability one.*

Proof. Instead of showing directly the lemma above, we show that with probability one, the set of record times for a Brownian motion B_t

$$\{t \geqslant 0 : Y_t = M_t - B_t = 0\},$$

has Hausdorff dimension $\frac{1}{2}$[17].

M_t is an increasing function, so we can regard it as a distribution function of a measure μ, with $\mu(a, b] = M_b - M_a$. Then set of record times is a support on this measure. Also we know that with probability one, the Brownian motion is Hölder continuous with any exponent $\alpha < \frac{1}{2}$. Therefore

$$M_b - M_a \leqslant \max_{0 \leqslant h \leqslant b-a} B_{a+h} - B_a \leqslant C_\alpha (b-a)^\alpha,$$

where C_α is a constant that does not depend on a or b[17]. Now according to Mass Distribution Principle, we get that almost surely, $\dim_H(\{t \geqslant 0 : Y_t = M_t - B_t = 0\}) \geqslant \alpha$. $\square$

There is a rather longer proof for the reverse Lemma which gives an upper bound for the Hausdorff dimension of zero set of a Brownian motion that is the same as the lower bound above, $\frac{1}{2}$. With combining that Lemma with Lemma 3.1 we get the proof for the following Theorem.

Theorem 3.3. *Zero set of a Brownian motion has Hausdorff dimension $\frac{1}{2}$ with probability one.*

In fact the Theorem above implies that the set of record times has Hausdorff dimension $\frac{1}{2}$ too, because Y_t has the same distribution as $|B_t|$ (Theorem 3.2) and zero set of $|B_t|$ is the same as zero set of B_t.

3.5.2 Hausdorff Dimension of Random Triangulation of the Circle

In this Section we will show briefly that the random triangulation of the circle has Hausdorff dimension $\frac{3}{2}$ with probability one.

In fact we will show that for probability one for any given $\varepsilon > 0$, S_ε has dimension $\frac{1}{2}$, in which S_ε is the set of endpoints of chords with length at least ε. After showing that we actually proved what we wanted to prove, because each point in S_ε corresponds to a chord in circle (which clearly has a Hausdorff dimension 1) and thus the set of all of those chords has Hausdorff dimension $\frac{3}{2}$ and when $\varepsilon \to 0$, S_ε converges to triangulation of the circle!

In mapping from normalized Brownian excursion to triangulation of the circle, chords correspond to intervals $[s, s']$ for which $f(s) = f(s')$ and $f(t) > f(s)$ for all $t \in (s, s')$ (in fact such an interval may not be part of a local minimum interval-pair, but it will be a limit of intervals which are[2]). Consider such intervals straddling time 0.5. These are intervals $[s_y, s'_y]$ where $0 < y < f(0.5)$ and

$$\begin{cases} s_y = \sup\{t < 0.5 : f(t) = y\}, \\ s'_y = \inf\{t > 0.5 : f(t) = y\}. \end{cases}$$

So now we need to show that

$$\text{the set } \{s_y : 0 < y < f(0.5)\} \text{ has dimension } \frac{1}{2}, \tag{3.4}$$

with probability one and then replacing 0.5 by any rational shows that S_ε has dimension $\frac{1}{2}$.

For proving 3.4, note that the set of record times of a normalized Brownian excursion has Hausdorff dimension $\frac{1}{2}$ with probability one[1]. It is essentially the same as saying that with probability one

$$\text{the set } \{t_y : 0 < y\} \text{ has Hausdorff dimension } \frac{1}{2}, \tag{3.5}$$

where

$$t_y = \inf\{t > 0 : g(t) = y\}.$$

$g(t)$ is a normalized Brownian excursion here, but in general it can be any Brownian motion.

Also note that Brownian motions has *time-reversal property* which means that if B_t is a Brownian motion, then $\tilde{B}_t = B_{u-t}$ is also a Brownian motion. For normalized Brownian excursion $f(t)$ if we choose $u = 0.5$ we observe that $\tilde{f}(t) = f(0.5 - t)$ for $t \in [0, 0.5]$ is also a Brownian motion. This and (3.5) give us a proof of (3.4).

Finally because of (3.4) we conclude that S_ε has Hausdorff dimension $\frac{1}{2}$ and

[1] We proved this for Brownian *motions*, but it is also true for normalized Brownian excursion, because the conditioning involved in producing normalized Brownian excursion from Brownian motion does not effect *local* properties of the random functions[2].

thus the Hausdorff dimension of random triangulation is $\frac{3}{2}$.

Chapter 4

Discrete Trees

In previous Chapter we showed that simple (positive) walks have a one-to-one correspondence with ordered trees and then binary trees and finally triangulations of polygons. In the end we showed that the limit of simple random walks tends to normalized Brownian excursion. Due to one-to-one correspondence of simple walks and ordered trees, it is intuitive to guess that the limit of ordered trees will tend to Brownian excursions too. In this Chapter we will check this in more details. In fact we first map ordered trees to *contour functions* and then we show that the limit of contour functions tend to (normalized) Brownian excursions.

4.1 Dyck Path and Contour Function

One of the fundamental tools in enumerative combinatorics is *bijections*. Two sets A and B have the same cardinality if and only if there exists a bijection from A to B[21]. With such a bijection we can count the elements of A by counting the elements of B. We do not need any example: We used this tool several times in previous two Chapters! But let us give another interesting example which is also useful in this Chapter: The enumeration of *Dyck words*.

Dyck words are words in letters X and Y with as many X's as Y's such that in any initial segment of the word we have at least as many X's as Y's[21]. For example $XYXXYXXYYY$ is a Dyck word, but $XXYXYYYXXY$ is not a Dyck word. If we replace each X with a left parenthesis and each Y with a right parenthesis and vice versa, we clearly get a bijection from Dyck words to well-formed sequence of parentheses and thus we observe that the number of different Dyck words with n letter X's and n letter Y's is equal to $\mathrm{C}_n = \frac{1}{n+1}\binom{2n}{n}$. But let us count the number of Dyck words in another way.

We can count the number of Dyck words of length $2n$ by starting to count all words with n X's and n Y's which is $\binom{2n}{n}$ and then subtract the wrong words. A bijection due to *D. André*[3] shows that the number of wrong words[1] is $\binom{2n}{n-1}$: Given a word with n X's and n Y's that is not a Dyck word, locate the first Y that violates the restriction of Dyck words and interchange all X's and Y's that come after it. This will be a bijection from the set of wrong words to the set of words with $n-1$ X's and $n+1$ Y's. Number of the elements of the second set is clearly $\binom{2n}{n-1}$ and so is the number of wrong words! Thus the number of Dyck words will be equal to

$$\binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{n+1}\binom{2n}{n} = \mathrm{C}_n.$$

If we write the number of X minus the number of Y for each initial segment[2] of a Dyck word, we get a sequence of nonnegative numbers that we call it *Dyck path*. For example from the Dyck word $XYXXYXXYYY$ we get the Dyck path $0, 1, 0, 1, 2, 1, 2, 3, 2, 1, 0$. Let us define it *mathematically* rather than *combinatorially*!

Definition 4.1. Let $n \geqslant 0$ be an integer. A *Dyck path* of length $2n$ is a sequence $(x_0, x_1, ..., x_{2n})$ of nonnegative integers such that $x_0 = x_{2n} = 0$ and for each $i = 1, 2, ..., 2n$, $|x_i - x_{i-1}| = 1$[12].

If we plot Dyck path in a Cartesian coordinate plane we get some isolated points and if we use linear interpolation between each of these points, we get the plot of a function that we call it *contour function*. Obviously the plot of contour functions will remind us of (nonnegative) simple random walks.

[1] Which are not Dyck words!

[2] Can be also an empty segment or the whole word.

4.2 Discrete Trees

4.2.1 Plane Trees

For defining plane trees we introduce the set [12]

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

where $\mathbb{N} = \{1, 2, ...\}$ and $\mathbb{N}^0 = \{\varnothing\}$.

Thus $\mathcal{U}$ is a set of elements like $u = (u^1, u^2, ..., u^n)$ and we set $|u| = n$. If $u = (u^1, u^2, ..., u^m)$ and $v = (v^1, v^2, ..., v^n)$ belong to $\mathcal{U}$, we define the concatenation of u and v by $uv = (u^1, ..., u^m, v^1, ..., v^n)$. Also $u\varnothing = \varnothing u = u$. In fact $|\varnothing| = 0$ and in general $|uv| = |u| + |v|$.

We define mapping $\pi = \mathcal{U}\backslash\varnothing \to \mathcal{U}$ by $\pi((u^1, u^2, ..., u^n)) = (u^1, u^2, ..., u^{n-1})$. Along with the definition below we see that $\pi(u)$ is the parent of u.

Definition 4.2. A *plane tree* τ is a finite subset of $\mathcal{U}$ such that:

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

(ii) for every $u \in \tau\backslash\{\varnothing\}$, $\pi(u) \in \tau$;

(iii) for every $u \in \tau$ there exists an integer $n_u(\tau) \geqslant 0$ such that for every $j \in \mathbb{N}$, $uj \in \tau$ if and only if $1 \leqslant j \leqslant n_u(\tau)$.

So node u in plane tree τ has $n_u(\tau)$ children.

We denote the set of all trees by $\mathbf{A}$ and define $|\tau|$ to be the number of edges of tree τ: $|\tau| = \#\tau - 1$. Also for every integer $k \geqslant 0$, we let $\mathbf{A}_n$ be the set of trees with n edges:

$$\mathbf{A}_n = \{\tau \in \mathbf{A} : |\tau| = n\}.$$

Proposition 4.1. *Cardinality of $\mathbf{A}_n$ is the n-th Catalan number*

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

Proof. It was proved in previous Chapter. □

Let us explain briefly how to get the contour function of tree τ. Suppose that τ is the tree shown in Figure 3.2. If we suppose that each edge of τ is drawn such that all of them have unit length, then contour function of τ is the distance (in tree) of a parcel which starts to move from the root and traverse the tree like as shown in Figure 4.1.

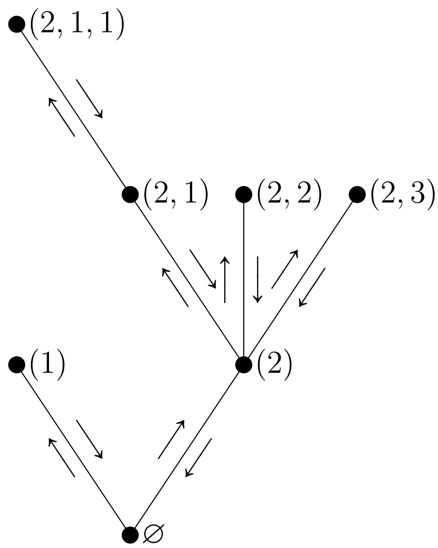

Figure 4.1: Traversing a plane tree and its nodes' sequences.

In this traverse, we visit the children from left to right and create their sequences upon that ordering. Also each edge is traversed two times, so in general, contour function C_s of tree τ is the function

$$C_s : s \in [0, 2|\tau|] \to [0, |\tau|].$$

By convention $C_s = 0$ for $s > 2|\tau|$. Note that C_s above might not be surjective[1]. In this way it is easy to see that the contour function will look alike the equivalent simple walk of the tree τ which is shown in Figure 4.2.

Proposition 4.2. *The mapping* $\tau \mapsto (C_0, C_1, ..., C_{2n})$ *is a bijection from* A_n *onto the set of all Dyck paths of length* $2n$*.*

[1]With probability one in its limit!

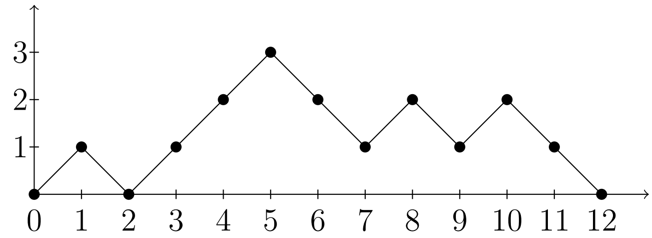

Figure 4.2: Contour function of the plane tree in Figure 4.1.

Proof. Mapping 1 in Section 3.3 which we showed that it is a bijection is indeed the mapping in this Proposition. □

4.2.2 Galton-Watson Trees

An offspring distribution $\{p_k\}_{k\geqslant 0}$ is simply a probability measure on $\mathbb{N}_0 = \{0, 1, 2, ...\}$. Let us define Galton-Watson process.

Definition 4.3. A *Galton-Watson process* $(Z_n)_{n\geqslant 0}$ is a discrete Markov chain with values in $\mathbb{N}_0$ with transition probabilities

$$P(Z_{n+1} = k | Z_n = m) = p_k^{*m},$$

where p_k^{*m} denotes the m-th convolution power of offspring distribution $\{p_k\}_{k\geqslant 0}$. In other words the conditional distribution of Z_{n+1} given Z_n is the distribution of the sum of Z_n i.i.d. random variables with distribution $\{p_k\}_{k\geqslant 0}$. Initial value is $Z_0 = 1$.

If the expected value of a random variable with law $\{p_k\}_{k\geqslant 0}$ is equal to 1, we will have some interesting distributions defined below.

Definition 4.4. Probability measure μ on $\mathbb{N}_0$ is called *critical* or *subcritical offspring* distribution if

$$\sum_{n=0}^{\infty} n\mu(n) \leqslant 1.$$ [12]

We suppose that $\mu(1) \neq 1$.

Now let $(N_u, u \in \mathcal{U}$[1]$)$ be a collection of i.i.d. random variables with distribution

[1]Note that this is the set $\mathcal{U}$ which we defined in previous Section.

μ defined above and indexed by set $\mathcal{U}$. Denote by θ the random subset of $\mathcal{U}$ defined by

$$\theta = \left\{u = (u^1, u^2, ..., u^n) \in \mathcal{U} : \forall j \in \{1, 2, ..., n\}, u^j \leqslant N_{(u^1,u^2,...,u^{j-1})}\right\}. \tag{4.1}$$

Proposition 4.3. *θ is a.s. a tree. Also if*

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

$(Z_n)_{n\geqslant 0}$ is a Galton-Watson process with offspring distribution μ and initial value $Z_0 = 0$.

Proof. If θ is finite then it is a tree, because for every $u \in \theta$, due to the definition of θ, we have all of its left siblings and their parents. Also $\varnothing \in \theta$. In fact N_u is the number of children of node u.

If θ is infinite, then there exists $u = (u^1, u^2, ...) \in \theta$ for which $|u| = \infty$. It means that for each n, $N_{u_n} > 0$, where $u_n = (u^1, u^2, ..., u^n)$. The probability of this is at most $\Pi_{n=0}^{\infty} 1 - \mu(0)$ which converges to 0 if $\mu(0) > 0$ which is clearly the case because $\mu(1) \neq 1$ and $\sum_{n=0}^{\infty} n\mu(n) \leqslant 1$.

The proof that $(Z_n)_{n\geqslant 0}$ is indeed a Galton-Watson process can be done easily by induction. □

The finiteness of θ can also be concluded from the fact that the Galton-Watson process with offspring distribution μ becomes extinct a.s.: $Z_n = 0$ for n large.

Definition 4.5. The tree θ defined by (4.1), or any random tree with the same distribution is called *Galton-Watson tree* with offspring distribution μ, or in short *μ-Galton-Watson tree*[12].

Suppose that τ is a tree and $1 \leqslant j \leqslant n_\varnothing(\tau)$, then we denote by $T_j\tau$ the branch that starts from the j-th child of the root:

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

We write Π_μ for the distribution of θ on the space $\mathbf{A}$. Π_μ can be characterized by the following two properties[12]:

(i) $\Pi_\mu(n_\varnothing = j) = \mu(j)$ for every $j \in \mathbb{N}_0$;

(ii) for every $j \geqslant 1$ with $\mu(j) > 0$, the branches $T_1\tau, T_2\tau, ..., T_j\tau$ are independent under the conditional probability $\Pi_\mu(d\tau | n_\varnothing = j)$ and their conditional distribution is Π_μ.

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

Proposition 4.4. *For every $\tau \in \mathbf{A}$,*

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

Proof. It is easily understood that knowing that a randomly generated tree θ with offspring distribution μ is the same as τ is equivalent to knowing that for each $u \in \tau$, $N_u = n_u(\tau)$[1]! So

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

□

In particular the case when $\mu = \mu_0$ for which μ_0 is the (critical) geometric offspring distribution, $\mu_0(n) = 2^{-n-1}$ for every $n \in \mathbb{N}_0$, is interesting and we check it more in what follows. In that case, the Proposition above tells us that

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

because for every $\tau \in \mathbf{A}$, we have $\sum_{u \in \tau} n_u = \#(\tau) - 1 = |\tau|$.

It means that $\Pi_{\mu_0}(\tau)$ depends only on $|\tau|$. So the conditional distribution when given $|\tau| = n$ will be a uniform distribution on $\mathbf{A}_n$.

Let us check the contour function when $\mu = \mu_0$.

[1] The first clearly implies the second. Knowing that for each $u \in \tau$, $N_u = n_u(\tau)$, we conclude that all the nodes $u \in \tau$ are also in θ and also no other node can be added to θ because for each leaf in $v \in \tau$, surely $n_v(\tau) = 0$.

4.3 The Contour Function in the Geometric Case

In general, the contour function does not have a "nice" probabilistic structure[12]. But when the distribution is the geometric offspring distribution, μ_0, there exists a bijection between Dyck paths and random walks.

Recall that if $(S_n)_{n\geqslant 0}$ is a simple random walk on $\mathbb{Z}$ starting from 0, then it can be written as

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

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

We are interested in nonnegative random walks, so put

$$T = \min\{n \geqslant 0 : S_n = -1\}, \tag{4.2}$$

and consider the walk from the start until the $(n-1)$-th walk:

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

This path in finite a.s. and we call it an excursion of simple random walk. Note that each excursion of simple random walk of length $T-1$ is also a contour function of a tree with $\frac{T-1}{2}$ edges.

Before stating the next Proposition let us introduce the upcrossing times of random walk S from 0 to 1:

$$U_1 = \min\{n \geqslant 0 : S_n = 1\} \text{ and } V_1 = \min\{n \geqslant U_1 : S_n = 0\}$$

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

$$U_{j+1} = \min\{n \geqslant V_j : S_n = 1\} \text{ and } V_{j+1} = \min\{n \geqslant U_{j+1} : S_n = 0\}.$$

If S is an excursion of simple random walk of length $T-1$, then if we put

$$N = \max\{j : U_j \leqslant T-1\}, \tag{4.3}$$

it means that S is consisting of K parts that are each positive random walks and due to Markov property they are independent from each other.

If S is a simple random walk (not necessarily an excursion of it) which starts at 0 then for each $j = 1, 2, \ldots$, the part starting at V_j and ending at $U_{j+1} - 1$ can be either empty or nonempty. Thus S can be partitioned to some *i.i.d.* simple random walks which endpoints are 0. Denote these parts by ξ_i ($i = 1, 2, \ldots$) and also let $T_0 = 0$ and T_i be the i-th time when S comes back to 0, for $i = 1, 2, \ldots$ (it means that ξ_i starts at time T_{i-1} and ends at time T_i).

The fact that ξ_i's are i.i.d., is the essential of excursion theory[19]. To show that how it can be used for calculations, let us find the distribution of the number of returns to 0 before the time $\tau = \inf\{n : S_n = -2\}$. Obviously,

$$\begin{aligned} P(\xi_1 \text{ visits } -2) &\equiv P(S_n = -2 \text{ for some } k \text{ with } 0 < k < T_1) \\ &= P(S_1 = -1, S_2 = -2) \\ &= \frac{1}{4}. \end{aligned}$$

So

$$\begin{aligned} &P(\text{number of returns to 0 before } \tau, \text{ is at least } k \\ &\quad = P(\text{excursions } \xi_1, \xi_2, \ldots, \xi_k \text{ do not visit } -2) \\ &\quad = P(\xi_1 \text{ does not visit } -2)^k \\ &\quad = \left(\frac{3}{4}\right)^k \end{aligned}$$

Proposition 4.5. *Contour function of μ_0-Galton-Watson tree θ is an excursion of simple random walk.*

Proof. According to Proposition 4.2, plane trees are in one-to-one correspondence with Dyck paths. Also Dyck paths are clearly in one-to-one correspondence with nonnegative random walks. Thus the statement of this Proposition is equivalent to saying that the random plane tree θ coded by an excursion of simple random walk is a μ_0-Galton-Watson tree. To prove this, suppose that we coded tree θ by an excursion of simple random walk, S. Now if we consider N, as defined in (4.3), it is easily understood that $n_\varnothing(\theta) = N$ and for every $i \in 1, 2, \ldots, N$, the branch $T_i\theta$

is coded with the path $(\omega_i(n))_{0\leqslant n\leqslant V_i-U_i-1}$, where for each $n \in \{0,1,...,V_i-U_i-1\}$

$$\omega_i(n) = S_{U_i+n} - 1.$$

Also N is distributed according to geometric offspring distribution μ_0 and conditioned on $N = m$, paths $\omega_1, \omega_2, ..., \omega_m$ are independent excursions of simple random walks. Now according to characterization of Π_{μ_0} these who that θ is a μ_0-Galton-Watson tree. □

4.4 Brownian Excursions

In Section 3.4 we talked briefly about Brownian motions and also normalized Brownian excursions. We defined Brownian motions by limit of simple random walks when rescaled properly. In this section we talk about Brownian excursions in more depth and we show that the contour function of a tree uniformly distributed over $\mathbf{A}_n$ converges in distribution as $n \to \infty$ towards a normalized Brownian motion.

4.4.1 Local Time Process and Excursion Space

Consider a standard linear Brownian motion $B = (B_t)_{t\geqslant 0}$ starting from 0. We define *local time process* of Brownian motion B as follows.

Definition 4.6. The *local time process* $(L_t^x)_{t\geqslant 0}$ of standard linear Brownian motion B at level x is mathematically defined by[1]

$$L_t^x = \int_0^t \delta(x - B_s)ds,$$

where δ is the Dirac delta function[2]. It can be approximated a.s. for every $t \geqslant 0$ by

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

[1]Some authors denote the Local time process by $l(t,x)$

[2]It is zero for all values except at zero and its integral over any interval containing zero is equal to one.

We are particularly interested in local time process at level 0. Now if we define the reflected Brownian motion by $\beta_t = |B_t|$, the local time process at level 0 of B_t or of β_t is approximated a.s. for every $t \geqslant 0$ by

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

Thus local time process is a continuous increasing process. The increasing points of this process at level 0 is the zero set of B_t,

$$\mathcal{Z}_B = \{t \geqslant 0 : B_t = 0\},$$

which is the same as $\mathcal{Z}_\beta$. If we define the right-continuous inverse of the local time process as

$$\sigma_l := \inf\{t \geqslant 0 : L_t^0 > l\}$$

for every $l > 0$, we will have

$$\mathcal{Z}_\beta = \{\sigma_l : l \geqslant 0\} \cup \{\sigma_{l-} : l \in D\}$$

where D denotes the countable set of all discontinuity times of mapping $l \to \sigma_l$.

excursion intervals (away from 0) of β are any connected component of the open set $\mathbb{R}_+ \backslash \mathcal{Z}_\beta$. Then excursion intervals away from 0 of β are intervals of the form (σ_{l-}, σ_l) where $l \in D$. We define the excursion $e_l = (e_l(t))_{t \geqslant 0}$ associated to the interval (σ_{l-}, σ_l) for every $l \in D$ by

$$e_l(t) = \begin{cases} \beta_{\sigma_{l-}+t} & \text{if } 0 \leqslant t \leqslant \sigma_l - \sigma_{l-}, \\ 0 & \text{if } t > \sigma_l - \sigma_{l-}. \end{cases}$$

In fact different excursions e_l are defined somewhat like ω_i's in Proposition 4.5. We view these excursions as elements of the excursions space E that is defined as follows[12].

Definition 4.7. The *excursion space* E is a metric space with elements

$$e \in C(\mathbb{R}_+, \mathbb{R}_+) \text{ such that } e(0) = 0 \text{ and } \zeta(e) > 0,$$

and metric d,

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

and with the associated Borel σ-field. $\zeta(e)$ above is defined by

$$\zeta(e) := \sup\{s > 0 : e(s) > 0\}$$

where $\sup \varnothing = 0$.

Note that zero function does not belong to the excursion space because we require $\zeta(e) > 0$ and $\zeta(e)$ can be seen as the length of excursion e. Also for every $l \in D$, $\zeta(e_l) = \sigma_l - \sigma_{l-}$.

4.4.2 The Itô Excursion Measure

Put

$$q_t(x) = \frac{x}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right). \tag{4.4}$$

The function $t \to q_t(x)$ is the density of first hitting time of x by B[12] (starting at 0) or of first hitting time of 0 by a linear Brownian motion which starts at x[11].

Itô measure $\mathbf{n}(de)$ of positive excursions is an infinite measure on the set of elements of excursion space E and has the following two (characteristic) properties[11]:

(i) For every $t > 0$ and every measurable function $f : R_+ \to R_+$ such that $f(0) = 0$,

$$\int \mathbf{n}(de) f(e(t)) = \int_0^\infty dx q_t(x) f(x);$$

(ii) if $t > 0$ and Φ and Ψ are two nonnegative measurable functions defined respectively on $C([0,t], R_+)$ and $C(R_+, R_+)$, then

$$\begin{aligned} &\int \mathbf{n}(de)\Phi(e(r), 0 \leqslant r \leqslant t)\Psi(e(t+r), r \geqslant 0) \\ = &\int \mathbf{n}(de)\Phi(e(r), 0 \leqslant r \leqslant t) E_{e(t)}(\Psi(B_{r\wedge T_0}, r \geqslant 0)), \end{aligned}$$

where E_x is the set of excursions e for which $e(s) > x$ if and only if $s \in (0, \sigma)$ for some positive $\sigma(e)$[1], $(B_t)_{t \geqslant 0}$ is a linear Brownian motion which starts at x and $T_0 = \inf \mathcal{Z}_B$

The following theorem is the basic result of excursion theory in our particular setting.

Theorem 4.1 (Itô). *The point measure*

$$\sum_{l \in D} \delta_{(l, e_l)}(dsde)$$

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

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

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

A proof of this Theorem can be found in an article by *L. C. G. Rogers*[19].

The measure $\mathbf{n}(de)$ is called the Itô excursion measure. From standard properties of Poisson measures we can conclude the next Corollary.

Corollary 4.1. *Suppose A be a measurable set of E with finite positive measure. Put $T_A = \inf\{l \in D : e_l \in A\}$. Then T_A is exponentially distributed with parameter of the measure of A, $\mathbf{n}(A)$, and the distribution of e_{T_A} is the conditional measure*

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

Moreover, T_A and e_{T_A} are independent.

This corollary can be used for calculating various distributions like height and length of excursions, under the Itô excursion measure.

The distribution of height of excursion $e(t)$ is

$$\mathbf{n}\left(\sup_{t \geqslant 0} e(t) > \varepsilon\right) = \frac{1}{2\varepsilon}$$

[1] $\sigma(e)$ is called the length of duration of the excursion e, particularly when $x = 0$.

and its length distribution is

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

The Itô excursion measure have scaling property: For every $\lambda > 0$, define mapping $\Phi_\lambda : E \to E$ by putting $\Phi_\lambda(e)(t) = \sqrt{\lambda}e(t/\lambda)$, for every $e \in E$ and $t \geqslant 0$. Then we have $\Phi_\lambda(\mathbf{n}) = \sqrt{\lambda}\mathbf{n}$.

The scaling property is especially useful when defining conditional versions of Itô excursion measure[12]. Let us discuss $\mathbf{n}(de)$ when conditioning with respect to length $\zeta(e)$.

There exists a unique collection of probability measures $(\mathbf{n}_{(s)}, s > 0)$ on E with the following properties[12]:

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

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

(iii) for every measurable subset A of E,

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

Notice that $\frac{ds}{2\sqrt{2\pi s^3}}$ can be seen as the measure of the set of excursions like e with length $\zeta(e) \in ds$. We may and will write $\mathbf{n}_{(s)} = \mathbf{n}(\cdot|\zeta = s)$, and the measure $\mathbf{n}_{(1)}$ is called the *law of the normalized Brownian excursions.*

Before continuing, let us first state the famous *Radon-Nikodym theorem.*

Theorem 4.2 (Radon-Nikodym). *If μ and λ are two σ-finite measures on measurable space (X, Σ) and μ is absolutely continuous*[1] *with respect to λ, then there is a measurable function f on X taking values in $[0, \infty)$ such that for any measurable set A*

$$\mu(A)=\int_A f d\lambda.$$

The following Proposition emphasizes the Markovian properties of $\mathbf{n}$[12].

[1]Measure μ is absolutely continuous with respect to measure λ if $\mu(A) = 0$ for every set A for which $\lambda(A) = 0$ and we write it as $\mu \ll \lambda$.

Proposition 4.6. *The Itô excursion measure* $\mathbf{n}$ *is the only* σ*-finite measure on excursion space* E *that specifies the following two properties:*

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

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

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

We can use this Proposition to establish the absolute continuity properties of the conditional measures $\mathbf{n}_{(s)}$ with respect to $\mathbf{n}$. By Radon-Nikodym theorem this is equivalent to saying that for any measurable set A in excursion space and some measurable function f

$$\mathbf{n}_{(s)}(A) = \int_A f d\mathbf{n}.$$

Here f is called the *Radon-Nikodym derivative* of $\mathbf{n}_{(s)}$. Now let us denote the σ-field on E generated by the mappings $r \to e(r)$, for every $t \geqslant 0$ and $0 \leqslant r \leqslant t$, by $\mathcal{F}_t$. If $0 < t < 1$, then the measure $\mathbf{n}_{(1)}$ is absolutely continuous with respect to $\mathbf{n}$ on the σ-field $\mathcal{F}_t$ and the Radon-Nikodym derivative, f, will be equal to

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

Using the derivative above we can derive the density of the distribution of $(e(t_1), e(t_2), ..., e(t_p)$ under $\mathbf{n}_{(1)}(de)$ for every integer $p \geqslant 1$ and every choice of $0 < t_1 < t_2 < \cdots < t_p < 1$[12]:

$$2\sqrt{2\pi}\, q_{t_1}(x_1) p^*_{t_2-t_1}(x_1, x_2) p^*_{t_3-t_2}(x_2, x_3) \cdots p^*_{t_p-t_{p-1}}(x_{p-1}, x_p) q_{1-t_p}(x_p), \tag{4.5}$$

where

$$p^*_t(x, y) = p_t(x, y) - p_t(x, -y), \quad t > 0, \quad x, y > 0,$$

is the transition density of Brownian motion killed when it hits 0. This density shows that law of $(e(t))_{0\leqslant t\leqslant 1}$ under $\mathbf{n}_{(1)}$ is invariant under time reversal.

4.5 Convergence of Contour Functions Towards Brownian Excursions

The convergence of contour functions to Brownian excursions can be seen as a special case of results provided in article *The continuum random tree III* by *Aldous*[1]. Before proving this convergence let us first present two lemmas.

Lemma 4.1. *For every $\varepsilon > 0$,*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \sup_{s\geqslant\varepsilon} \left|\sqrt{n}P\left(S_{\lfloor ns\rfloor} = \lfloor x\sqrt{n}\rfloor \text{ or } \lfloor x\sqrt{n}\rfloor + 1\right) - p_s(0,x)\right| = 0.$$

This lemma is a very special case of classical local limit theorems and can be easily obtained by direct calculations, using the explicit form of the law of S_n and Stirling's formula[12].

On the other hand the next lemma is a special case of famous *Kemperman's formula*[18].

For every integer $\ell \in \mathbb{Z}$, denote a probability measure under which the simple random walk S starts from ℓ by P_ℓ.

Lemma 4.2. *For every $\ell \in \mathbb{N}_0$ and every integer $n \geqslant 1$,*

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

Proof. There are several different proofs to this lemma[12;18], but we will provide a more enumerative combinatorial proof.

If both sides of the equation above are 0, then there is nothing to prove. Otherwise let us propose a simple question and solve it first.

Question 4.1. *Suppose that we have k "X" and $k + l$ "Y". In how many ways we can put them in a line that for no initial segment the number of Y's be more than l more than X's.*

For solving this question we can make a bijection to count the number of wrong sequences as we did for Dyck words, and then subtract it from the number of all sequences. Suppose that we have a *wrong* sequence of X's and Y's. Consider the shortest initial segment in which we have exactly k' X and $(k'+l+1)$ Y. Interchange all the other X's and Y's that come after this segment. Now we will have a sequence of $(k-1)$ X and $(k+l+1)$ Y. It is easy to see that this is a bijection. So the solution will be equal to

$$\begin{aligned}
\binom{2k+l}{k} - \binom{2k+l}{k-1} &= \frac{(2k+l)!}{k!(k+l)!} - \frac{(2k+1)!}{(k-1)!(k+l+1)!} \\
&= \frac{(k+l+1)(2k+l)! - k(2k+l)!}{k!(k+l+1)!} \\
&= \frac{(l+1)(2k+l)!}{k!(k+l+1)!} \\
&= \frac{l+1}{k+l+1}\binom{2k+l}{k+l}.
\end{aligned}$$

Note that if we put $l=0$ in question above, we get the k-th Catalan number as the solution and if we put $\frac{n-\ell-1}{2}$ instead of k and ℓ instead of l in above question[1], we will get

$$2\frac{\ell+1}{n+\ell+1}\binom{n-1}{\frac{n+\ell-1}{2}},$$

which is equivalent to the number of random walks like S that start from ℓ, for which we have $S_{n-1}=0$ and for no $i=0,1,...,n-1$, $S_i=-1$. So it is equal to the number of random walks starting from ℓ and for which $T=n$.

Coming back to proof of the lemma, note that it is enough to prove

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

It is equal to the number of random walks reaching -1 from ℓ for the first time in n-th step divided by number of all random walks starting from ℓ and reaching to

[1]Note that $\frac{n-\ell-1}{2}$ and thus $\frac{n+\ell-1}{2}$ are integers if and only if the probabilities defined in the statement of lemma are not equal to 0.

0 or -2 at $(n-1)$-th step. So

$$\begin{aligned} P_\ell(T = n \mid S_n = -1) &= \frac{2\frac{\ell+1}{n+\ell+1}\binom{n-1}{\frac{n+\ell-1}{2}}}{\binom{n-1}{\frac{n+\ell-1}{2}} + \binom{n-1}{\frac{n+\ell+1}{2}}} \\ &= \frac{2\frac{\ell+1}{n+\ell+1}}{1 + \frac{n-\ell-1}{n+\ell+1}} \\ &= \frac{2\frac{\ell+1}{n+\ell+1}}{\frac{2n}{n+\ell+1}} \\ &= \frac{l+1}{n}. \end{aligned}$$

$\square$

Using two lemmas above we can prove the following theorem which says that contour functions of random trees in $\mathbf{A}_n$ converge to Brownian excursions as $n \to \infty$.

Theorem 4.3. *For every $n \in \mathbb{N}$, let θ_n be a random tree uniformly distributed over $\mathbf{A}_n$, and let $(C_n(t))_{t\geqslant 0}$ be its contour function. Then*

$$\left(\frac{1}{\sqrt{2n}} C_n(2nt)\right)_{0\leqslant t\leqslant 1} \xrightarrow[n\to\infty]{(d)} (\mathbf{e}_t)_{0\leqslant t\leqslant 1}$$

where $\mathbf{e}$ is a normalized Brownian excursion distributed according to $\mathbf{n}_{(1)}$ and the space $C([0,1],\mathbb{R}_+)$ is equipped with the topology of uniform convergence.

Proof. Using Proposition 4.5 and that $\Pi_{\mu_0}(\cdot \mid |\tau| = n)$ coincides with the uniform distribution over $\mathbf{A}_n$, we get that $(C_n(0), C_n(1), ..., C_n(2n))$ is distributed as an excursion of simple random walk conditioned to have length $2n$. Thus we need to verify that the law of

$$\left(\frac{1}{\sqrt{2n}} S_{[2nt]}\right)_{0\leqslant t\leqslant 1}$$

given that $T = 2n+1$ converges to $\mathbf{n}_{(1)}$ as $n \to \infty$. This can be seen as a conditional version of *Donsker's theorem*. We will divide the proof into two parts: Proving the convergence of finite-dimensional marginals and then establishing the tightness of the sequence of laws[12].

Finite-dimensional marginals. Let us first consider one-dimensional marginals and then base the proof of higher dimensional marginals on it. Fix $t \in (0,1)$ and we will show that

$$\lim_{n\to 0} \sqrt{2n} P\left(S_{\lfloor 2nt \rfloor} = \left\lfloor x\sqrt{2n} \right\rfloor \text{ or } \left\lfloor x\sqrt{2n} \right\rfloor + 1 \mid T = 2n+1\right) = 2\sqrt{2\pi} q_t(x) q_{1-t}(x), \tag{4.6}$$

uniformly when x varies over a compact subset of $(0, \infty)$. Note that right hand side of the above equation is the same as (4.5) for $p = 1$. It means that the law of $\frac{S_{\lfloor 2nt \rfloor}}{\sqrt{2n}}$ under $P(\cdot \mid T = 2n+1)$ converges to the law of $e(t)$ under $\mathbf{n}_{(1)}(de)$.

For every $i \in \{1, 2, ..., 2n\}$ and $\ell \in \mathbb{N}_0$,

$$P(S_i = \ell \mid T = 2n+1) = \frac{P(\{S_i = \ell\} \cap \{T = 2n+1\})}{P(T = 2n+1)}.$$

But

$$\begin{aligned} P(\{S_i = \ell\} \cap \{T = 2n+1\}) &= P(\{S_i = \ell, T > i\} \cap \{T = 2n+1\}) \\ &= P(\{S_i = \ell, T > i\}) P_\ell(T = 2n+1-i), \end{aligned}$$

also

$$\begin{aligned} P_\ell(T = i+1) &= P_\ell(\{S_{i+1} = -1\} \cap \{S_i = 0, T > i\}) \\ &= P_\ell(S_{i+1} = -1 \mid S_i = 0) P_\ell(S_i = 0, T > i) \\ &= \frac{1}{2} P_\ell(S_i = 0, T > i), \end{aligned}$$

both because of markovian property of S. Also

$$\begin{aligned} P_\ell(S_i = 0, T > i) &= P(S_i = \ell, T > i) \\ \Rightarrow P(S_i = \ell, T > i) &= 2P_\ell(T = i+1). \end{aligned}$$

Thus

$$\begin{aligned}P(S_i = \ell \mid T = 2n+1) &= \frac{2P_\ell(T = i+1)P_\ell(T = 2n+1-i)}{P(T = 2n+1)} \\ &= \frac{2(2n+1)(n+1)^2}{(i+1)(2n+1-i)} \cdot \frac{2P_\ell(S_{i+1} = -1)P_\ell(S_{2n+1-i} = -1)}{P(S_{2n+1} = -1)},\end{aligned} \tag{4.7}$$

where we used Lemma 4.2 for deriving the second equality[1].

Recall that $p_t(0,x) = (t/x)q_t(x)$ where $q_t(x)$ is defined as (4.4), so

$$\begin{aligned}p_t(0,x) &= \frac{t}{x} \cdot \frac{x}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right) \\ &= \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right).\end{aligned} \tag{4.8}$$

As an important special case, when $x := 0$, we get $p_t(0,0) = (2\pi t)^{-\frac{1}{2}}$. For large n,

$$\begin{aligned}P(S_{2n+1} = -1) &\approx P(S_{2n} = 0) \\ &\approx \frac{p_1(0,0)}{\sqrt{2n}} \\ &= \frac{1}{\sqrt{2n}} \cdot \frac{1}{\sqrt{2\pi}},\end{aligned}$$

using Lemma 4.1 if we set $x := 0$, $s := 1$ and $n := 2n$ in the second approximation. Also we have the approximations

$$\frac{2n+1}{\lfloor 2nt \rfloor + 1} \approx \frac{1}{t} \text{ and } \frac{(\lfloor x\sqrt{2n} \rfloor + 1)^2}{2n+1-\lfloor 2nt \rfloor} \approx \frac{x^2 \cdot 2n}{2n - 2nt} \approx \frac{x^2}{1-t}.$$

Using all of the approximations above we get

$$\frac{2(2n+1)(\lfloor x\sqrt{2n} \rfloor + 1)^2}{(\lfloor 2nt \rfloor + 1)(2n+1-\lfloor 2nt \rfloor)} \cdot \frac{1}{P(S_{2n+1} = -1)} \approx 2\sqrt{2\pi}\sqrt{2n}\frac{x^2}{t(1-t)}. \tag{4.9}$$

[1] Note that $P_\ell = P$ when $\ell = 0$.

Again using Lemma 4.1, we have the approximation

$$\begin{aligned} &P_{\lfloor x\sqrt{2n}\rfloor \text{ or } \lfloor x\sqrt{2n}\rfloor+1}(S_{\lfloor 2nt\rfloor+1}=-1)P_{\lfloor x\sqrt{2n}\rfloor \text{ or } \lfloor x\sqrt{2n}\rfloor+1}(S_{2n+1-\lfloor 2nt\rfloor}=-1)\\ &\quad\approx \frac{p_t(0,x)p_{1-t}(0,x)}{2n}=\frac{t(1-t)}{x^2\cdot 2n}q_t(x)q_{1-t}(x), \end{aligned} \tag{4.10}$$

where in general with $P_{\ell \text{ or } \ell'}$ we mean $P_\ell+P_{\ell'}$. Now by multiplying approxiamtions (4.9) and (4.10) to each other and putting $i=\lfloor 2nt\rfloor$ and $\ell=\lfloor x\sqrt{2n}\rfloor$ or $\ell=\lfloor x\sqrt{2n}\rfloor+1$ in right hand side of (4.7) we get

$$P(S_{\lfloor 2nt\rfloor}=\left\lfloor x\sqrt{2n}\right\rfloor \text{ or } \left\lfloor x\sqrt{2n}\right\rfloor+1 \mid T=2n+1)\approx \frac{2\sqrt{2\pi}}{\sqrt{2n}}q_t(x)q_{1-t}(x),$$

and the proof of (4.6) is complete.

For higher dimensional marginals we can use a similar way. For example for two-dimensional marginals, we can observe that if $0<i<j<2n$ and if $\ell\in\mathbb{N}_0$,

$$\begin{aligned} &P(S_i=\ell,S_j=m,T=2n+1)\\ &\quad=2P_\ell(T=i+1)P_\ell(S_{j-i}=m,T>j-i)P_m(T=k+1-j). \end{aligned}$$

Here, only the middle term, $P_\ell(S_{j-i}=m,T>j-i)$, needs a treatment that we didn't discuss before. However we can see that

$$P_\ell(S_{j-i}=m,T>j-i)=P_\ell(S_{j-i}=m)-P_\ell(S_{j-i}=-m), \tag{4.11}$$

because if for a random walk that passed through -1 and yet arrived to m at $(j-i)$-th step, we reflect the part from the first time that random walk hit -1 to the end, we get a random walk that arrives to $-m$ at $(j-i)$-th step. It is easy to see that it is also a bijection. On the other hand, by putting $i=\lfloor 2ns\rfloor$, $j=\lfloor 2nt\rfloor$ and $\ell=\lfloor x\sqrt{2n}\rfloor$ or $\ell=\lfloor x\sqrt{2n}\rfloor+1$ in (4.11) and then using Lemma 4.1, we get

$$\begin{aligned} P_{\lfloor x\sqrt{2n}\rfloor \text{ or } \lfloor x\sqrt{2n}\rfloor+1}\left(S_{\lfloor 2nt\rfloor-\lfloor 2ns\rfloor}=\lfloor y\sqrt{2n}\rfloor\right) &\approx \frac{p_{t-s}(0,y-x)}{\sqrt{2n}}\\ &=\frac{p_{t-s}(x,y)}{\sqrt{2n}}. \end{aligned}$$

Similarly

$$P_{\lfloor x\sqrt{2n}\rfloor \text{ or } \lfloor x\sqrt{2n}\rfloor+1}\left(S_{\lfloor 2nt\rfloor-\lfloor 2ns\rfloor}=\lfloor -y\sqrt{2n}\rfloor\right)\approx \frac{p_{t-s}(0,-y-x)}{\sqrt{2n}}$$
$$=\frac{p_{t-s}(x,-y)}{\sqrt{2n}}.$$

Subtracting the second approximation from the first, we get

$$P_{\lfloor x\sqrt{2n}\rfloor \text{ or } \lfloor x\sqrt{2n}\rfloor+1}\left(\left\{S_{\lfloor 2nt\rfloor-\lfloor 2ns\rfloor}=\lfloor y\sqrt{2n}\rfloor\right\}\cap\{T>\lfloor 2nt\rfloor-\lfloor 2ns\rfloor\}\right)$$
$$\approx \frac{p^{*}_{t-s}(x,y)}{\sqrt{2n}},$$

and the result follows in a straightforward way (approximating the other two terms as previous way and putting this approximation for the middle term, we get want we want).

Tightness. Let $(x_0, x_1, ..., x_{2n})$ be a Dyck path with length $2n$, and for each $i \in \{0, 1, ..., 2n-1\}$ and $j \in \{0, 1, ..., 2k\}$, set

$$x_j^{(i)} = x_i + x_{i\oplus j} - 2\min_{i\wedge(i\oplus j)\leqslant m\leqslant i\vee(i\oplus j)} x_m$$

with the notation $i \oplus j = i + j$ if $i + j \leqslant 2n$ and $i \oplus j = i + j - 2n$ if $i + j > 2n$.

Proposition 4.7. *For each $i \in \{0, 1, ..., 2n-1\}$, $(x_0^{(i)}, x_1^{(i)}, ..., x_{2n}^{(i)})$ is also a Dyck path, where $x_j^{(i)}$ is defined as above.*

Proof. We should prove that $x_0^{(i)} = 0$, $x_{2n}^{(i)} = 0$ and for each $k \in \{1, ..., 2n\}$, $\left|x_k^{(i)} - x_{k-1}^{(i)}\right| = 1$. By definition we have

$$\begin{aligned}
x_0^{(i)} &= x_i + x_{i\oplus 0} - 2\min_{i\wedge(i\oplus 0)\leqslant m\leqslant i\vee(i\oplus 0)} x_m\\
&= x_i + x_i - 2\min_{i\wedge i\leqslant m\leqslant i\vee i} x_m\\
&= 2x_i + -2x_i\\
&= 0.
\end{aligned}$$

Similarly $x_{2n}^{(i)} = 0$. Now fix $k \in \{1, ..., 2n\}$. So

$$\begin{aligned}
\left|x_k^{(i)} - x_{k-1}^{(i)}\right| &= \Bigg|\left(x_i + x_{i\oplus k} - 2 \min_{i\wedge(i\oplus k)\leqslant m\leqslant i\vee(i\oplus k)} x_m\right) \\
&\qquad - \left(x_i + x_{i\oplus(k-1)} - 2 \min_{i\wedge(i\oplus(k-1))\leqslant m\leqslant i\vee(i\oplus(k-1))} x_m\right)\Bigg| \\
&= \Bigg|x_{i\oplus k} - x_{i\oplus(k-1)} - 2 \min_{i\wedge(i\oplus k)\leqslant m\leqslant i\vee(i\oplus k)} x_m \\
&\qquad - 2 \min_{i\wedge(i\oplus(k-1))\leqslant m\leqslant i\vee(i\oplus(k-1))} x_m\Bigg|.
\end{aligned} \tag{4.12}$$

If $k = 2n - i + 1$ then both " min " parts of right side of above equation will be equal to 0 and leaves $|x_{i\oplus k} - x_{i\oplus(k-1)}| = |x_1 - x_{2n}| = 1$. If $k \neq 2n - i + 1$ then $i \oplus k = i \oplus (k-1) + 1$. Put $i \oplus k = K$. Now if we check three different cases $K < i$, $K = i$ and $K > i$ we observe that the difference between "min" parts are either 0 or $x_K - x_{K-1}$ and the right side of (4.12) will be

$$\left|x_K - x_{K-1} - 2\begin{cases} 0 \\ x_K - x_{K-1} \end{cases}\right| \overset{\text{in both cases}}{=} |x_K - x_{K-1}| = 1,$$

and proof will be complete. □

Define the mapping $\Phi_i : (x_0, x_1, ..., x_{2n}) \to \left(x_0^{(i)}, x_1^{(i)}, ..., x_{2n}^{(i)}\right)$. It is possible to show that this mapping is a bijection from the set of all Dyck paths with length $2n$ onto itself[12]. In fact $\Phi_{2n-i} \circ \Phi_i$ is the identity mapping. It can be verified if we check that Φ_i is indeed the Dyck path of a tree which is obtained from the original tree rooted at i-th node that we encounter while exploring it with its Dyck path. Checking that, it will be obvious that $\Phi_{2n-i} \circ \Phi_i$ leads us back to the original tree.

For every $i, j \in \{0, 1, ..., 2n\}$ set

$$\check{C}_n^{i,j} = \min_{i\wedge j\leqslant m\leqslant i\vee j} C_n(m).$$

The discussion above then shows that

$$\left(C_n(i) + C_n(i\oplus j) - 2\check{C}_n^{i,i\oplus j}\right)_{0\leqslant j\leqslant 2n} \overset{(d)}{=} \left(C_n(j)\right)_{0\leqslant j\leqslant 2n} \tag{4.13}$$

Lemma 4.3. *[12] For every integer $p \geqslant 1$, there exists a constant C_p such that for every $n \geqslant 1$ and every $i \in \{0, 1, ..., 2n\}$,*

$$\mathrm{E}\left[C_n(i)^{2p}\right] \leqslant C_p i^p.$$

If we prove this Lemma, the proof of tightness is complete, because after proving the Lemma, considering (4.13), for any i and j such that $0 \leqslant i < j \leqslant 2n$ we will have

$$\begin{aligned}\mathrm{E}\left[(C_n(j) - C_n(i))^{2p}\right] &\leqslant \mathrm{E}\left[\left(C_n(i) + C_n(j) - 2\check{C}_n^{i,j}\right)^{2p}\right] \\ &= \mathrm{E}\left[C_n(j-i)^{2p}\right] \\ &\leqslant C_p(j-i)^p.\end{aligned}$$

It means that we have

$$\mathrm{E}\left[\left(\frac{C_n(2nt) - C_n(2ns)}{\sqrt{2n}}\right)^{2p}\right] \leqslant C_p(t-s)^p, \tag{4.14}$$

at least for all s and t of the forms $s = \frac{i}{2n}$ and $t = \frac{j}{2n}$ where $0 \leqslant i < j \leqslant 2n$. But C_n is 1-Lipschitz[1] and with a simple argument we can see that (4.14) holds for every s and t such that $0 \leqslant s < t \leqslant 1$ (possibly with different C_p)[12].

Now let us give the proof of the Lemma.

Proof of Lemma 4.3. Obviously $(C_n(2n-i))_{0 \leqslant i \leqslant 2n}$ has the same distribution as $(C_n(i))_{0 \leqslant i \leqslant 2n}$ and thus we can restrict our attention to the case $1 \leqslant i \leqslant n$. Also $C_n(i)$ has the same distribution as Si conditioned on $T = 2n+1$. Thus according to (4.7), for every $\ell \in \mathbb{N}_0$ we have

$$P(C_n(i) = \ell) = \frac{2(2n+1)(\ell+1)^2}{(i+1)(2n+1-i)} \cdot \frac{P_\ell(S_{i+1} = -1)P_\ell(S_{2n+1-i} = -1)}{P(S_{2n+1} = -1)}.$$

[1] $\forall x, y \in [0,1] : |C_n(x) - C_n(y)| \leqslant |x - y|$

In Lemma 4.1, if we set $x := \frac{-1}{\sqrt{2n}}$, $s := \frac{2n+1}{2n}$ and $n := 2n$, we have

$$\begin{aligned} P(S_{2n+1} = -1) &\approx \frac{1}{\sqrt{2n}} p_{\frac{2n+1}{2n}}\left(0, \frac{-1}{\sqrt{2n}}\right) \\ &= \frac{1}{\sqrt{2n}}\left(\frac{1}{\sqrt{2\pi\frac{2n}{2n+1}}}\exp\left(-\frac{\frac{1}{2n}}{2\frac{2n+1}{2n}}\right)\right), \end{aligned}$$

where for observing the equality we used (4.8). Thus

$$P(S_{2n+1} = -1) \lesssim \frac{1}{\sqrt{2n}}\frac{1}{\sqrt{2\pi}},$$

and

$$P(S_{2n+1} = -1) \gtrsim \frac{1}{\sqrt{2n}}\frac{1}{\sqrt{\frac{4}{3}\pi}}\exp(-\frac{1}{6}).$$

Similarly we can get the bounds for $P_\ell(S_{2n+1-i} = -1)$ and we can find two constants c_0 and c_1 such that

$$P(S_{2n+1} = -1) \geqslant \frac{c_0}{\sqrt{2n}}, \quad P_\ell(S_{2n+1-i} = -1) \leqslant \frac{c_1}{\sqrt{2n}}.$$

We assumed that $i \leqslant n$ and therefore

$$P(C_n(i) = \ell) \leqslant 4\frac{c_1(\ell+1)^2}{c_0(i+1)}P_\ell(S_{i+1} = -1) = 4\frac{c_1(\ell+1)^2}{c_0(i+1)}P(S_{i+1} = \ell+1).$$

Consequently

$$\begin{aligned} \mathrm{E}\left[C_n(i)^{2p}\right] = \sum_{\ell=0}^{\infty}\ell^{2p}P(C_n(i) = \ell) &\leqslant 4\frac{c_1}{c_0(i+1)}\sum_{\ell=0}^{\infty}\ell^{2p}(\ell+1)^2P(S_{i+1} = \ell+1) \\ &\leqslant 4\frac{c_1}{c_0(i+1)}\mathrm{E}\left[(S_{i+1})^{2p+2}\right], \end{aligned}$$

where we have the second inequality because

$$\begin{aligned} \mathrm{E}\left[(S_{i+1})^{2p+2}\right] &= \sum_{\ell=0}^{\infty} \ell^{2p+2} P(S_{i+1}=\ell) \\ &= \sum_{\ell=0}^{\infty} (\ell+1)^{2p+2} P(S_{i+1}=\ell+1) \\ &\geqslant \sum_{\ell=0}^{\infty} \ell^{2p}(\ell+1)^{2} P(S_{i+1}=\ell+1). \end{aligned}$$

It is well known that $\mathrm{E}\left[(S_{i+1})^{2p+2}\right] \leqslant C'_p(i+1)^{p+1}$, with some constant C'_p independent of i. Thus

$$\begin{aligned} \mathrm{E}\left[C_n(i)^{2p}\right] &\leqslant 4\frac{c_1}{c_0(i+1)} C'_p(i+1)^{p+1} \\ &= 4\frac{c_1}{c_0} C'_p(i+1)^{p} \\ &\leqslant C_p i^p, \end{aligned}$$

where $C_p = 2^{p+2}c_0^{-1}c_1C'_p$, and the proof is complete. □

With proving the Lemma, the proof of Theorem is complete too. □

The Theorem above is powerful and useful. We state the following Corollary as a typical application of this Theorem. Note that the height $H(\tau)$ of a plane tree τ is the maximal generation of a vertex of τ[12].

Corollary 4.2. *Let θ_n be uniformly distributed over $\mathbf{A}_n$. Then*

$$\frac{1}{\sqrt{2n}} H(\theta_n) \xrightarrow[n\to\infty]{(d)} \max_{0\leqslant t\leqslant 1} \mathbf{e}_t.$$

Proof. We have

$$\frac{1}{\sqrt{2n}} H(\theta_n) = \max_{0\leqslant t\leqslant 1}\left(\frac{1}{\sqrt{2n}} C_n(2nt)\right),$$

and so the result is immediate from Theorem 4.3. □

Chapter 5

Real and Labeled Trees

In this Chapter we will define and discuss real and labeled trees and deriving similar results we did for discrete trees.

5.1 Real Trees

5.1.1 Definition

Definition 5.1. A metric space (X, d) is a real tree (or $\mathbb{R}$-tree) if it has the following three properties:

(i) *Completeness.* It is complete.

(ii) *Unique geodesics.* For all $x, y \in X$, there is a unique isometric map $f_{x,y} : [0, d(x, y)] \to X$ such that $f_{x,y}(0) = x$ and $f_{x,y}(d(x, y)) = y$.

(iii) *Loop-free.* For any injective continuous map $q : [0, 1] \to X$ we have

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

A useful fact is that a metric space (X, d) is an $\mathbb{R}$-tree if and only if it is complete, path-connected and satisfies the so-called *four point condition*,

$$d(x_1, x_2) + d(x_3, x_4) \leqslant \max\{d(x_1, x_3) + d(x_2, x_4), d(x_1, x_4) + d(x_2, x_3)\}, \quad (5.1)$$

for all $x_i \in X$ $(i \in \{1, 2, 3, 4\})$[9].

In this Chapter, however, we restrict our attention only to *compact* real trees which has the same definition above as real trees except that (X, d) needs to be a compact metric space. If we distinguish a vertex ρ of X, we will have a rooted real tree. $\rho = \rho(X)$ is called the root. From now on, when we say real tree, we mean rooted compact real tree like $(\mathcal{T}, d)$ with the root $\rho(\mathcal{T})$.

Informally, a connected union of line segments (property (ii)) in the plane without any loops (property (iii)) is a (compact) real tree. We assume that there are finitely many line segments and therefore for any two points x and y in the tree, there is a unique path going from x to y in the tree which is consisted of finitely many line segments and the distance between x and y is the sum of the lengths of these line segments[12].

Let $[\![x, y]\!]$ denote the whole range of the mapping $f_{x,y}$ in (ii). Particularly, $[\![\rho, x]\!]$ is the path going from the root ρ to x which we will interpret as the ancestral line of vertex x. By $[\![x, y[\![$, $]\!]x, y]\!]$ and $]\!]x, y[\![$ we mean the images of respectively $[0, d(x, y))$, $(0, d(x, y)]$ and $(0, d(x, y))$ of the mapping $f_{x,y}$ in (ii)[7].

We define a partial order on the tree by setting $x \preccurlyeq y$ (x is ancestor of y) if and only if $x \in [\![\rho, y]\!]$. For every $x, y \in \mathcal{T}$, there is a unique z, we call the most recent common ancestor to x and y for which we have $[\![\rho, x]\!] \cap [\![\rho, y]\!] = [\![\rho, z]\!]$[12]. We denote it by $z = x \wedge y$.

By definition, the multiplicity of a vertex $x \in \mathcal{T}$ is the number of connected components of $\mathcal{T} \backslash x$ and we denote it by $\mathrm{n}(x, \mathcal{T})$ and if we are sure of which tree we are talking about we simply denote it by $\mathrm{n}(x)$. Also any non-root vertex of a tree which has multiplicity one is called *leaf*. The set of all leaves is denoted by

$$\mathrm{Lf}(\mathcal{T}) = \{x \in \mathcal{T} \backslash \{\rho\} : \mathrm{n}(x, \mathcal{T}) = 1\}.$$

Also we denote the *branching points* of $\mathcal{T}$ by

$$\mathrm{Br}(\mathcal{T}) = \{x \in \mathcal{T} \backslash \{\rho\} : \mathrm{n}(x, \mathcal{T}) \geqslant 3\}.$$

By convention, the root ρ is neither a leaf nor a branching point[7]. We also denote the *internal skeleton* of $\mathcal{T}$ by $\mathrm{Sk}(\mathcal{T})$ and is defined as $\mathrm{Sk}(\mathcal{T}) = \mathcal{T} \backslash \mathrm{Lf}(\mathcal{T})$.

Proposition 5.1. *For any dense sequence* $\{x_n\}_{n\geqslant 1}$ *in* $\mathcal{T}$*, we have*

$$Sk(\mathcal{T}) = \bigcup_{n\geqslant 1} [\![\rho, x_n[\![.$$

Proof. It is obvious that the union does not contain any leaves.

On the other hand, suppose $x \in \mathcal{T}$ is not a leaf. If it does not belong to the union, then no point in set $Y = \{y : y \neq x, x \leqslant y\}$ exists in the sequence. Clearly there exists a leaf $z \in \mathcal{T}$. The sequence does not contain any points of $]\!]x, z]\!]$, which is a contradiction to the assumption that it is dense. □

It is also easy to show that $\mathrm{Br}(\mathcal{T})$ is at most countable[8].

5.1.2 Coding

As in previous Chapter for coding discrete trees by contour functions, we can describe a method for coding real trees by continuous functions.

Set

$$U^\ell = \{f \in C([0,\ell], \mathbb{R}_+) \mid f(0) = f(\ell) = 0\}.$$

Let $g \in U^1$ and for every $u_1, u_2 \in [0,1]$, set

$$m_g(u_1, u_2) = \inf_{u\in[u_1\wedge u_2, u_1\vee u_2]} g(u),$$

and

$$d_g(u_1, u_2) = g(u_1) - 2m_g(u_1, u_2) + g(u_2).$$

Clearly $d_g(u_1, u_2) = d_g(u_2, u_1)$ and it is easy to verify the triangle inequality for every $u_1, u_2, u_3 \in [0,1]$:

$$d_g(u_1, u_3) \leqslant d_g(u_1, u_2) + d_g(u_2, u_3).$$

Define an equivalence relation $\sim_g$ on $[0,1]$ by

$$u_1 \sim_g u_2 \text{ if and only if } g(u_1) = m_g(u_1, u_2) = g(u_2),$$

and let $\mathcal{T}_g$ be the quotient space

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

Theorem 5.1. *For every $g \in U^1$, $(\mathcal{T}_g, d_g)$ is a metric space and is a (compact) real tree.*

Proof. For every $g \in U^1$, it is obvious from the definition of d_g that it is a metric on $\mathcal{T}_g$. Also if we denote the canonical projection by $p_g : [0,1] \to \mathcal{T}_g$, clearly it is continuous[1] and thus the metric space $(\mathcal{T}_g, d_g)$ is path-connected and compact.

After observing that $(\mathcal{T}_g, d_g)$ is a path-connected compact metric space, now it suffices to verify the four point condition (5.1).

Consider $u_1, u_2, u_3, u_4 \in \mathcal{T}_g$. We should show that

$$d_g(u_1,u_2) + d_g(u_3,u_4) \leqslant \max\{d_g(u_1,u_3) + d_g(u_2,u_4), d_g(u_1,u_4) + d_g(u_2,u_3)\}.$$

But

$$d_g(u_1,u_2) + d_g(u_3,u_4) = g(u_1) + g(u_2) + g(u_3) + g(u_4) - 2(m_g(u_1,u_2) + m_g(u_3,u_4)),$$

and similarly for

$$d_g(u_1,u_3) + d_g(u_2,u_4) \text{ and } d_g(u_1,u_4) + d_g(u_2,u_3).$$

Thus we should show that

$$m_g(u_1,u_2) + m_g(u_3,u_4) \geqslant \min\{m_g(u_1,u_3) + m_g(u_2,u_4), m_g(u_1,u_4) + m_g(u_2,u_3)\}.$$

It is easy to see that

$$m_g(u_{(1)},u_{(2)}) + m_g(u_{(3)},u_{(4)}) \geqslant m_g(u_{(1)},u_{(3)}) + m_g(u_{(2)},u_{(4)}),$$

and

$$m_g(u_{(1)},u_{(4)}) + m_g(u_{(2)},u_{(3)}) \geqslant m_g(u_{(1)},u_{(3)}) + m_g(u_{(2)},u_{(4)}),$$

[1] [0,1] and $\mathcal{T}_g$ are equipped respectively by the Euclidean metric and metric d_g.

where $u_{(i)}$ is the i-th smallest number in $\{u_1, u_2, u_3, u_4\}$ for each $i \in \{1, 2, 3, 4\}$. On the other hand, with checking different cases[1], we can see that at least one of those inequalities will be equality and thus the proof is complete. $\square$

According to Theorem above, therefore we can view $(\mathcal{T}_g, d_g)$ as a rooted tree with root $\rho = p_g(0) = p_g(1)$.

Remark 5.1. Any real tree $\mathcal{T}$ is isometric to $\mathcal{T}_g$ for some $g \in U^1$ [9].

To get an insight on how this coding works, look at Figure 5.1, where a tree obtained from three given points, u_1, u_2, u_3, of a continuous function.

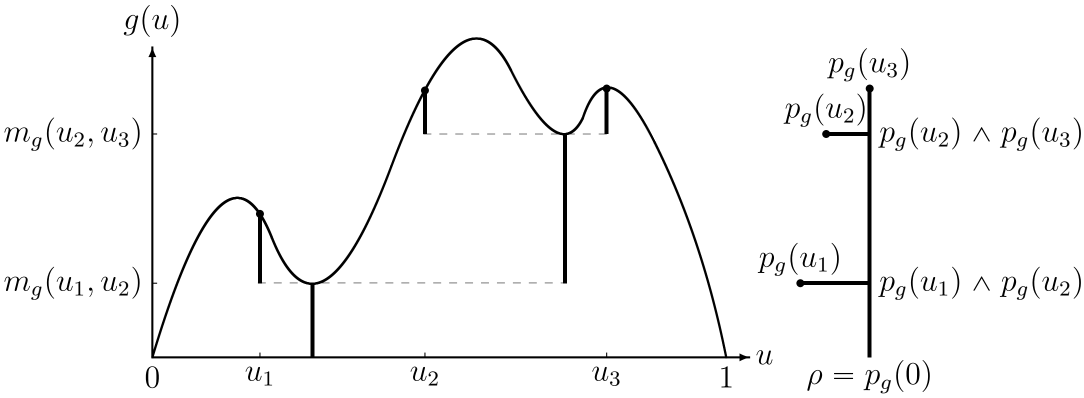

Figure 5.1: Real Tree Coded by Three Points of a Continuous Function.

In the Figure, the real tree on the right obtained by bold segments shown on the left. The heights of endpoints of each bold segment on the left is either 0, $m_g(u_i, u_j))$ or $g(u_i)$ for each $i, j \in \{u_1, u_2, u_3\}$. Clearly the path from the root to $p_g(u_i)$ has length $g(u_i)$ and its common part with $p_g(u_j)$ has length $m_g(u_i, u_j)$. With knowing these two properties about different paths between root and $p_g(u_i)$ we can build the real tree on the right. One approach can be as follows.

We start from any arbitrary point, for example the most far node from root, here u_3. We draw a straight line between root and $p_g(u_3)$ with length $g(u_3)$. Now we can choose another node, for example whose path has the longest common part with $p_g(u_3)$, i.e. u_2. We can see that the path from root to $p_g(u_2) \wedge p_g(u_3)$ will have length $m_g(u_2, u_3)$. From there, we can draw a new line segment with

[1]Whether $\inf_{[u_{(1)}, u_{(4)}]} g(u)$ happens in interval $[u_{(1)}, u_{(2)}]$, $[u_{(2)}, u_{(3)}]$ or $[u_{(3)}, u_{(4)}]$.

length $g(u_2) - m_g(u_2, u_3)$ and its endpoint will be $p_g(u_2)$. The same goes for any other nodes.

Note that $p_g(0) = p_g(1)$, $m_g(u_1, u_2) = m_g(u_1, u_3)$ and also $p_g(u_1) \wedge p_g(u_2) = p_g(u_1) \wedge p_g(u_3)$.

For the specific continuous function in Figure 5.1, if the chosen three points were the only three local maximums of the function, we could obtain the complete structure of its equivalent real tree. It means that after building the real tree according to those three points, adding any new points of function won't change the way the real tree looks like and the new points will lie on the previous drawn real tree. Look at the Figure 5.2 for clarification.

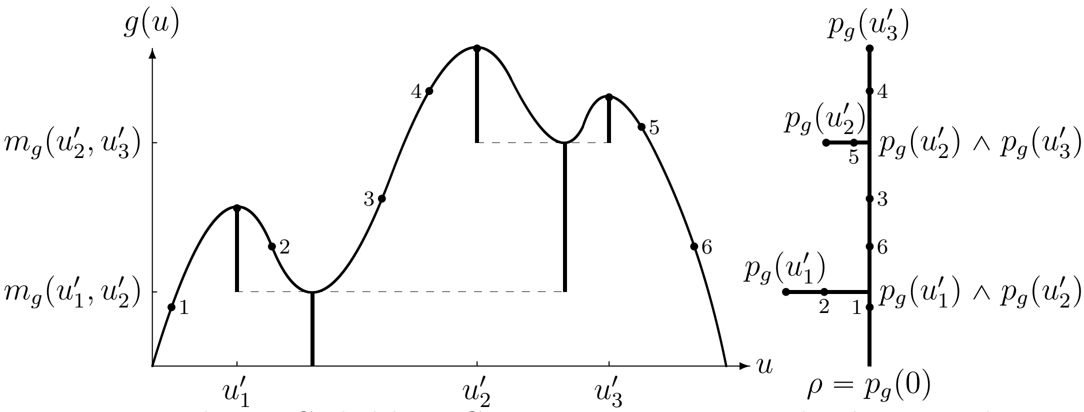

Figure 5.2: Real Tree Coded by a Continuous Function with Three Local Maximum Points.

It is easily understood that the number of nodes in real tree coded by continuous function f, is equal to the number of local maximums of f.

5.1.3 Convergence of Real Tree Towards CRT

We end this Section by presenting a Lemma which is a kind of continuous version of Proposition 4.7 and a Theorem which can be proved easily using the statement of Theorem 4.3. Both of them have been proved by *Le Gall*[12] and we will not provide their proofs here.

Lemma 5.1. *Let* $u_0 \in [0, 1)$. *For any* $r \geqslant 0$, *denote the fractional part of* r *by*

$\overline{r} = r - [r]$. *Set*

$$g'(u) = g(u_0) + g(\overline{u_0 + u}) - 2m_g(u_0, \overline{u_0 + u}),$$

for every $u \in [0, 1]$. *Then, the function* g' *is continuous and satisfies* $g'(0) = g'(1) = 0$ *and thus we can define* $\mathcal{T}_{g'}$. *Moreover, for every* $u_1, u_2 \in [0, 1]$, *we have*

$$d_{g'}(u_1, u_2) = d_g(\overline{u_0 + u_1}, \overline{u_0 + u_2}),$$

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

$$R(p_{g'}(u)) = p_g(\overline{u_0 + u}).$$

$\mathcal{T}_{g'}$ *can be seen as* $\mathcal{T}_g$ *re-rotted at* $p_g(u_0)$.

Before stating the Theorem, we need to introduce a couple of notations and definition of CRT.

For a metric space (E, δ), the notation $\delta_H(K, K')$ stands for the usual Hausdorff metric between compact subsets of E:

$$\delta_H(K, K') = \inf\{\varepsilon > 0 \mid K \subset U_\varepsilon(K') \text{ and } K' \subset U_\varepsilon(K)\},$$

where $U_\varepsilon(K) := \{x \in E \mid \delta(x, K) \leqslant \varepsilon\}$. Also we define the distance $d_{GH}(E_1, E_2)$ by

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

where E_1 and E_2 are two rooted compact metric spaces respectively with roots ρ_1 and ρ_2 and the infimum is over all possible choices of the metric space (E, δ) and the isometric embeddings $\varphi_1 : E_1 \to E$ and $\varphi_2 : E_2 \to E$ of E_1 and E_2 into E.

If we call E_1 and E_2 equivalent when there exists a root preserving isometry mapping E_1 onto E_2, then $d_{GH}(E_1, E_2)$ clearly only depends on the equivalence classes of E_1 and E_2. We denote the space of all equivalence classes of rooted compact metric spaces by $\mathbb{K}$[12]. We can prove that the metric space $(\mathbb{K}, d_{GH})$ is separable and complete[4;9].

Definition 5.2. The real tree $\mathcal{T}_{\mathbf{e}}$ which is coded by normalized Brownian motion excursion is called *continuum random tree* or briefly, CRT.

Theorem 5.2. *For every integer $n \geqslant 1$, let θ_n be uniformly distributed over $\mathbf{A}_n$ equipped with the usual graph distance, d_{gr}. Then*

$$\left(\theta_n, \frac{1}{\sqrt{2n}} d_{gr}\right) \xrightarrow[n\to\infty]{(d)} (\mathcal{T}_{\mathbf{e}}, d_{\mathbf{e}}),$$

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

5.2 Labeled Trees

Also for labeled trees we can derive the similar results as for discrete and real trees. We will define labeled trees and Brownian snakes, will prove a couple of different statements and in the end will present the Theorem without its proof which is somewhat similar to the proof of Theorem 4.3 and is provided by *Le Gall* in full details[12].

Definition 5.3. A *labeled tree* is a pair $(\tau, (\ell(v))_{v\in\tau})$ where τ is a plane tree and $(\ell(v))_{v\in\tau}$ is a collection of integer labels assigned to the vertices of τ, such that

(i) $\ell(\varnothing) = 0$;

(ii) for every $v \in \tau$, $\ell(v) \in \mathbb{Z}$;

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

For every integer $n \geqslant 0$, denote the set of all labeled trees with n edges by $\mathbf{T}_n$.

Proposition 5.2. *We have*

$$\#(\mathbf{T}_n) = 3^n \#(\mathbf{A}_n) = \frac{3^n}{n+1}\binom{2n}{n}.$$

Proof. The second equality comes from Proposition 4.1. The first equality is obvious, because moving from the root, we have three choices to choose a label for each new visited node. □

Consider a labeled tree $(\tau, (\ell(v))_{v\in\tau})$ with n edges. We know that τ can be coded by its contour function, $(C_t)_{t\geqslant 0}$. Let us define a function to code $(\ell(v))_{v\in\tau}$. One intuitive way to code it would be as follows. Suppose we are traversing τ according to its contour function. Then we visit different nodes in an order like $v_0 = \varnothing, v_1, v_2, ..., v_{2n} = \varnothing$. Note that a node will appear in this sequence exactly once if and only if it is a leaf. Now put

$$V_i = \ell(v_i),$$

for every $i = 0, 1, ..., 2n$. Also we can set $V_t = 0$ for any $t \geqslant 2n$ and by using linear interpolation we define V_t for every $t \geqslant 0$. We call $(V_t)_{t\geqslant 0}$ the *spatial contour function* of the labeled tree $(\tau, (\ell(v))_{v\in\tau})$.

We code the labeled tree $(\tau, (\ell(v))_{v\in\tau})$ by the pair $(C_t, V_t)_{t\geqslant 0}$ and our goal in this Section is to describe the scaling limit of this pair when chosen uniformly random in $\mathbf{T}_n$ as $n \to \infty$. Theorem 4.3 says that the scaling limit of $(C_t)_{t\geqslant 0}$ is indeed the normalized Brownian motion and thus is remains to find the scaling limit of $(V_t)_{t\geqslant 0}$. For this purpose we introduce the Brownian snakes.

5.2.1 Brownian Snakes

Let g be a continuous function as previous Section such that $g(0) = g(1) = 0$, and assume that it is also Hölder continuous (Definition 3.6). It means that

$$|g(u_1) - g(u_2)| \leqslant C_\alpha |u_1 - u_2|^\alpha,$$

for some exponents α and constant C_α.

Lemma 5.2. *The function $(m_g(u, v))_{u,v\in[0,1]}$ is nonnegative definite: for every integer $n \geqslant 1$, every $u_1, u_2, ..., u_n \in [0, 1]$ and every $\lambda_1, \lambda_2, ..., \lambda_n \in \mathbb{R}$, we have*

$$sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j m_g(u_i, u_j) \geqslant 0.$$

Proof. Proof can be done using complete induction. Denote the sum by A:

$$A = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j m_g(u_i, u_j).$$

The basis is $n = 1$. For $n = 1$, clearly

$$A = \sum_{i=1}^{1} \sum_{j=1}^{1} \lambda_i \lambda_j m_g(u_i, u_j) = \lambda_1^2 m_g(u_1, u_1) \geqslant 0.$$

Now suppose that $n = k > 1$. Define

$$m = \inf_{u_1 \leqslant u \leqslant u_n} g(u)$$

.

If $m = g(k') = 0$, let $i < n$ be such that $u_i < k' < u_{i+1}$ (if for an i, $u_i = k'$, the argument would be almost the same). Then we would have

$$A = \sum_{i=1}^{k'} \sum_{j=1}^{k'} \lambda_i \lambda_j m_g(u_i, u_j) + \sum_{i=k'+1}^{n} \sum_{j=k'+1}^{n} \lambda_i \lambda_j m_g(u_i, u_j).$$

By induction, left and right expressions of plus sign are both non-negative (number of points in both expressions are less than k) so A would be non-negative.

If $m > 0$, put

$$g'(u) = g(u) - m,$$

for $u_1 \leqslant u \leqslant u_k$. Define

$$A' = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j m_{g'}(u_i, u_j).$$

Since function g' is nonnegative, we can use the previous argument and prove that A' is non-negative. Also

$$A = A' + \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j m,$$

because from the definition of g', for any u and v, $u_1 \leqslant u, v \leqslant s_k$, we have $m_g(u,v) = m_{g'}(u,v) + m$. But

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j m = m\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j = m(\sum_{i=1}^{n}\lambda_i)^2 \geqslant 0.$$

So A is nonnegative. □

By Lemma 5.2 and a standard application of the Kolmogorov extension theorem, there exists a centered Gaussian process $(Z_u^g)_{u\in[0,1]}$ with covariance

$$\mathrm{E}[Z_{u_1}^g Z_{u_2}^g] = m_g(u_1,u_2),$$

for every $u_1, u_2 \in [0,1]$ [12]. Therefore we have

$$\begin{aligned}
\mathrm{E}[(Z_{u_1}^g - Z_{u_2}^g)^2] &= \mathrm{E}[(Z_{u_1}^g)^2] + \mathrm{E}[(Z_{u_2}^g)^2] - 2\mathrm{E}[Z_{u_1}^g Z_{u_2}^g] \\
&= g(u_1) + g(u_2) - 2m_g(u_1,u_2) \quad (= d_g(u_1,u_2)) \\
&= (g(u_1) - m_g(u_1,u_2)) + (g(u_2) - m_g(u_1,u_2)) \\
&\leqslant 2C_\alpha[u_1 - u_2]^\alpha.
\end{aligned}$$

From this bound and an application of the Kolmogorov continuity criterion, we observe that the process $(Z_u^g)_{u\in[0,1]}$ has a modification with continuous sample paths.

Definition 5.4. The *snake driven by the function* g is the centered Gaussian process $(Z_u^g)_{u\in[0,1]}$ with continuous sample paths and covariance

$$\mathrm{E}[Z_{u_1}^g Z_{u_2}^g] = m_g(u_1,u_2), \quad u_1, u_2 \in [0,1].$$

In particular, we have $Z_0^g = Z_1^g = 0$ and for every $u \in [0,1]$, Z_u^g is normal with mean 0 and variance $g(t)$.

As previously stated, Brownian motions are Hölder continuous with any exponent $\alpha < \frac{1}{2}$ almost surely and thus given a normalized Brownian excursion $(\mathbf{e}_t)_{t\in[0,1]}$, we can construct a snake $(Z_t)_{t\in[0,1]}$ from it.

Definition 5.5. We construct a pair $(\mathbf{e_t}, Z_t)_{t\in[0,1]}$ of continuous random processes whose distribution has the following two properties[12]:

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

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

The snake Z driven from normalized Brownian excursion $\mathbf{e}$ is called the *Brownian snake.*

5.2.2 Convergence of Labeled Tree Towards Brownian Snake

The following Theorem is due to *Chassaing* and *Schaeffer*[5;12].

Theorem 5.3. *For every integer $n \geqslant 1$, let $(\theta_n, (\ell^k(v))_{v\in\theta_n})$ be uniformly distributed over $\mathbf{T}_n$ and let $(C_n(t))_{t\geqslant 0}$ and $(V_n(t))_{t\geqslant 0}$ be respectively contour function and the spatial contour function of the labeled tree $(\theta_n, (\ell^k(v))_{v\in\theta_n})$. Then*

$$\left(\frac{1}{\sqrt{2n}}C_n(2nt), \left(\frac{9}{8k}\right)^{\frac{1}{4}} V_n(2nt)\right)_{t\in[0,1]} \xrightarrow[n\to\infty]{(d)} (\mathbf{e}_t, Z_t)_{t\in[0,1]},$$

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

References

[1] David Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993. 40

[2] David Aldous. Triangulating the circle, at random. *American Mathematical Monthly*, 101:223–233, 1994. 1, 11, 13, 14, 23

[3] D. André. Solution directe du problème résolu par m. bertrand. *C. R. Acad. Sci. Paris Sér.*, 105:436–437, 1887. 26

[4] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, 2001. ISBN 0-8218-2129-6. 57

[5] P. Chassaing and G. Schaeffer. Random planar lattices and integrated superbrownian excursion. *ArXiv Mathematics e-prints*, May 2002. URL http://arxiv.org/abs/math/0205226v1. 62

[6] Tom Davis. Catalan numbers. November 26, 2006. URL http://mathcircle.berkeley.edu/BMC6/pdf0607/catalan.pdf. 4, 9

[7] T. Duquesne. The coding of compact real trees by real valued functions. *ArXiv Mathematics e-prints*, April 2006. URL http://arxiv.org/abs/math/0604106v1. 52

[8] T. Duquesne and M. Winkel. Growth of levý trees. *ArXiv Mathematics e-prints*, September 2005. URL http://arxiv.org/abs/math/0509518. 53

[9] Steven N. Evans and Anita Winter. Subtree prune and regraft: A reversible real tree-valued markov process. *The Annals of Probability*, 34(3), 2006. URL http://arxiv.org/PS_cache/math/pdf/0502/0502226v3.pdf. 52, 55, 57

[10] Kenneth Falconer. *Fractal geometry: mathematical foundations and applications*. Wiley, 2nd edition, 2003. ISBN 0470848618. 12

[11] Jean-François Le Gall. *Spatial Branching Processes, Random Snakes and Partial Differential Equations*. Birkhäuser Basel, 1999. ISBN 9783764361266. 36

[12] Jean-François Le Gall and Grégory Miermont. Scaling limits of random trees and graphs. Notes pour cours d'école doctorale au CIRM, September, 2010. URL http://www.math.ens.fr/~legall/Cours-CIRM.pdf. 1, 26, 27, 29, 30, 32, 35, 36, 38, 39, 40, 42, 47, 48, 50, 52, 56, 57, 58, 61, 62

[13] Martin Gardner. *Time Travel and Other Mathematical Bewilderments*. W. H. Freeman, 1987. ISBN 0716719258. 2, 6, 10

[14] Martin Gardner. *Mathematical Puzzles and Diversions*. Penguin Books Ltd., 1991. ISBN 0140136355. 14

[15] Thomas Koshy. *Catalan Numbers with Applications*. Oxford University Press, 2008. ISBN 019533454X. 2, 8

[16] G. Morais. The study of hausdorff dimension. November, 2002. URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.3.2457&rep=rep1&type=pdf. 12

[17] Yuval Peres. An invitation to sample paths of brownian motion. Lecture notes edited by Bálint Virág, Elchanan Mossel and Yimin Xiao, November 1, 2001. URL http://stat-www.berkeley.edu/~peres/bmall.pdf. 21, 22

[18] Jim Pitman. Enumerations of trees and forests related to branching processes and random walks. In D. Aldous and J. Propp, editors, *Microsurveys in Discrete Probability*, number 41 in DIMACS Ser. Discrete Math. Theoret. Comp. Sci, pages 163–180. Amer. Math. Soc., 1998. 40

[19] L. C. G. Rogers. A guided tour through excursions. *Bull. London Math. Soc.*, 21:305–341, 1989. 33, 37

[20] N. J. A. Sloane. *A Handbook of Integer Sequences*. Academic Pr, 1973. ISBN 012648550X. 2

[21] Richard P. Stanley. Enumerative combinatorics, volume 2. *Bulletin (New Series) of the American Mathematical Society*, 39(1):129–135, 1999. URL http://www.ams.org/journals/bull/2002-39-01/S0273-0979-01-00928-4/S0273-0979-01-00928-4.pdf. 25