The Riordan group, additional algebraic structure and the uplift principle.

by

Thomas Yaw Osei

(10283762)

THIS THESIS IS SUBMITTED TO THE UNIVERSITY OF GHANA, LEGON, IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE AWARD OF MPHIL MATHEMATICS DEGREE.

July, 2015

Declaration

This thesis was written in the Department of Mathematics, University of Ghana, Legon from September 2014 to July 2015 in partial fulfilment of the requirement for the award of Master of Philosophy degree in Mathematics under the supervision of Dr. Margaret McIntyre and Dr. Douglass Adu-Gyamfi of the University of Ghana.

I hereby declare that except where due acknowledgement is made, this work has never been presented wholly or in part for the award of a degree at this University or any other University.

....

Thomas Y. Osei Date

....

Dr. Margaret McIntyre Date

....

Dr. Douglass Adu-Gyamfi Date

Abstract

We show that the normal Appell subgroup of the Riordan group is a pseudo ring under a multiplication given by the componentwise composition.

We develop formulae for calculating the degree of the root in generating trees and we establish isomorphisms between the four groups : the hitting time, Bell, associated and the derivative which are all subgroups of the Riordan group.

We have found the average number of trees with left branch length in the class of ordered trees and the Motzkin trees.

In the last chapter we examine the uplift principle and some known examples. We generalise some of the examples and we show that the average portion of protected points in the hex trees approaches 76/125 as $n \to \infty$.

Acknowledgements

I thank the Lord Jesus Christ who is the embodiment of all wisdom, knowledge and understanding for helping me complete this thesis successfully.
My sincerest gratitude also go to my supervisors, Dr. Margaret McIntyre and Dr. Douglass Adu-Gyamfi for spending time overseeing my work and for their constructive critiques.

I also highly appreciate the Carnegie Co-operation and Prof. F. K. Allotey of the institute for Mathematical Sciences for financially assisting me during the periods of my course work and thesis.

Finally I thank my God sent sister Maryjane Akua Mensah and brother Abraham Osei Boateng and last but not the least my friend and room mate Richard N. Y. Ofori and all my loved ones for their support throughout this thesis, without them I wouldn't have made it this far. I say may God richly bless you all.

Dedication

I dedicate this work to my biological father Mr. Osei Adu-Gyamfi, biological brothers George and Desmond and my biological mother Elizabeth Osei who is struggling for her life as she battles with stroke. May God almighty bless them and see them through all their endeavours.

I also dedicate this work to Professor Daniel Afedzi Akyeampong (a.k.a Akye) who recently passed away. Prof, you have been an inspiration to most of us. May your soul rest in perfect peace.

Contents

Chapter 1

Introduction

In this work, we begin by describing the concept of the group of infinite lower triangular matrices*(Riordan arrays)* called the *Riordan group*, introduced in 1991 by Shapiro, Getu, Woan, and Woodson.
In 1999, Donatella Merlini, M. Cecilia Verri used an algebraic approach to study the connection between some Riordan arrays and generating trees[7]. A formula for calculating the degree of the root and an arbitrary vertex of the generating tree are developed.

We study some isomorphism properties between some subgroups of the Riordan group established by Candice Jean-Louis and Asamoah Nkwanta in [3] (2012) and establish other isomorphisms based on their results.
A *Pseudo-ring(rng)* is an algebraic structure satisfying the same properties as a ring, without assuming the existence of a multiplicative(second operation) identity. The fact that there is a normal subgroup of the Riordan group presents an opportunity to establish a *pseudo-ring* structure of the Riordan group in this thesis.

This thesis presents the uplift principle for ordered trees which let us solve a variety of combinatorial problems in two simple steps. The first is to find the appropriate generating function at the root of the tree, the second is to lift the result to an arbitrary vertex in the tree by the leaf generating function.

1.1 Basic Definitions And Outline

1.1.1 Definitions

We set out most of the terminologies we intend to use throughout this thesis here. Vertices and trees are things we will often refer to in this thesis.

Definition 1.1.2. [2] A *graph* is a triple, $G = (V, E, \phi)$, where V is a finite set of vertices(nodes or points), E is a finite set of edges and ϕ is a function which assigns a unique 2-element set of vertices $\{u, v\}, u \neq v$ to each edge e. A graph G is *connected* if for any two distinct vertices u and v, there exists a sequence $v_0v_1 \cdots v_n$ of vertices such that $v_0 = u, v_n = v$ and any two consecutive terms v_i and v_{i+1} are adjacent. A *cycle* is a sequence $v_0v_1 \cdots v_n, n \geq 1$ of vertices such that $v_0 = v_n$.

Definition 1.1.3. A *tree* is a connected graph with no cycles, where one vertex is designated as a root. There are different classes of trees. Unless otherwise stated, by a tree, we will mean an ordered tree with a root and ordered list of subtrees at the root[4].
They will be drawn going down.

Definition 1.1.4. [2] Let S be a finite sequence of points $(x_0, y_0), (x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n) = (x, y)$ in the plane, such that all $(x_{i+1} - x_i, y_{i+1} - y_i) \in S$. These points are called *vertices* and a line connecting adjacent vertices is an *edge*
We define the *level* of a vertex, v, to be the ordinate of that point.

Definition 1.1.5. The *degree(updegree)* of a vertex at level k is the updegree of the vertex, i.e the number of edges emanating from that vertex to level $k + 1$. If every vertex has the same possibilities of updegrees we say it satisfies the *uniform updegree requirement(UUR)*.

Definition 1.1.6. A *leaf* is a vertex with updegree zero.

Definition 1.1.7. A marked vertex is called a *mutator*[5].

Definition 1.1.8. The *Ordinary Generating function*(OGF) of an infinite sequence $\{a_n\}_{n\geq 0}$ is the formal power series

$$a_0 + a_1z^1 + a_2z^2 + a_3z^3 + \cdots = \sum_{n\geq 0} a_nz^n.$$

One can also define an exponential generating function but in this thesis by generating function we mean an *OGF*.

Definition 1.1.9. A *generating tree*[7] is a rooted labelled tree with the property that if v_1 and v_2 are any two nodes with the same label then, for each label l, v_1 and v_2 have exactly the same number of children with label l. To specify a generating tree it therefore suffices to specify:
(1) the label of the root;
(2) a set of rules explaining how to derive from the label of a parent the label of all its children.

1.1.10 Outline of Thesis

Chapter 1 of this thesis contains the mathematical material which lays the foundation and the background for the remainder of the document. We set out to give the reader the notation and the language we need.
In section 1.2, we discuss the combinatorial interpretations of the Catalan numbers in terms of ordered trees. We discuss some generating function of the Central Binomial function, $B(z)$, the Fine function, $F(z)$, and the Motzkin function which form interesting relationships with $C(z)$, the generating function of the Catalan numbers. We will discuss the Hex function. The hex trees have interesting relationship with the Motzkin trees.

In chapter 2, we define the Riordan group and give a practical explanation of its usefulness in solving sums involving combinatorial inversions.
We briefly discuss the connection between elements of the Riordan group (aka Riordan arrays) and generating trees which was established by Merlini et al. in [7] and develop a formula for calculating the degree of the root of the generating tree.

In chapter 3, we look at some algebraic properties of the Riordan group.
In sub-section 3.1.7, we ask the question: Can we extend the Appell subgroup since it is a commutative subgroup to a ring?.
In section 3.2, we present the proofs of certain Riordan group isomorphisms and illustrate this in a form of a commutative diagram.
In the last section of chapter 3, we shift our attention back to the fundamental theorem of Riordan arrays where we obtain the average number of trees with left branch length for some classes of ordered trees.

In chapter 4, we discuss the uplift principle and its relationship with the Riordan group and demonstrate its utility combinatorially.

1.2 Catalan Numbers

The Catalan numbers can be defined in many ways. Recursively they are given by

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k} \text{ with } C_0 = 1.$$

In terms of generating functions this becomes $C(z) = 1 + z(C(z))^2$ or more compactly
$C = 1 + zC^2$ or $C = \frac{1-\sqrt{1-4z}}{2z}$.

1.2.1 Combinatorial Interpretation

The Catalan numbers have several combinatorial interpretations including the number of triangulations of a regular $(n+2)-$ gon (Euler), the number of parenthesizations of $(n+1)$ letters (Catalan), the number of paths from $(0,0)$ to $(2n,0)$ using the steps $Up = U = (1,1)$ and $Down = D = (1,-1)$ that never go below the x-axis (Dyck)[2].
Our interest will mostly be in the interpretation of Catalan numbers as the number of ordered trees with n edges. We demonstrate this in the next theorem.

Theorem 1.2.2. *$C_n = \frac{1}{n+1}\binom{2n}{n}$ is the number of ordered trees with n edges.*

Proof. Figure 1 shows the number of ordered trees with $n = 0, 1, 2, 3, 4, 5.$ edges. So if we let T represent the generating function for the trees then we have $T = C$. Now we let $V_{n,k}$ be the number of vertices at height k over all the trees with n edges and find the matrix $(V_{n,k})$ and take the row sum. Thus

$$V = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 3 & 1 & 0 & 0 \\ 5 & 9 & 5 & 1 & 0 \\ 14 & 28 & 20 & 7 & 1 \\ 42 & 90 & 75 & 35 & 9 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 6 \\ 20 \\ 70 \\ 252 \\ \vdots \end{bmatrix}$$

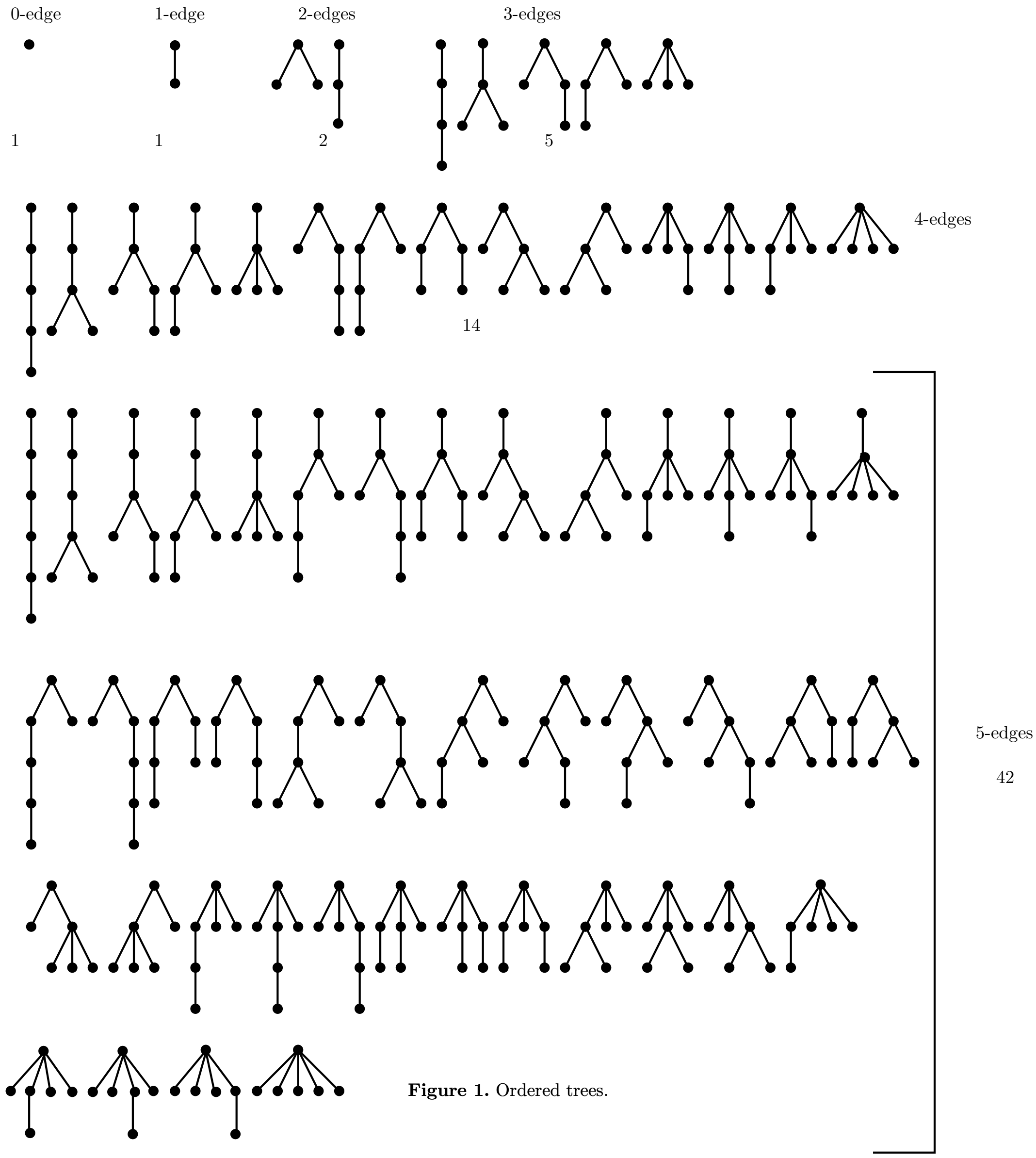


Figure 1. Ordered trees.

□

Definition 1.2.3. *The Central Binomial function* $B(z) = B$ is the generating function for the central binomial coefficients, $\binom{2n}{n}$. That is

$$B(z) = \sum_{n=0}^{\infty} \binom{2n}{n} z^n = \frac{1}{\sqrt{1-4z}}.$$

We notice that the row sum of $(V_{n,k})$ is $1, 2, 6, 20, 70, 252, \cdots [A000984] = B$ consequently we say B is the companion generating function for C meaning as $C(z)$ counts the number of ordered trees, $B(z)$ counts the number of vertices in the ordered trees.

We have the identity $B(z) = 1 + 2zC(z)B(z)$.[2]

Definition 1.2.4. [2] The *Fine function, F(z)*, is the generating function for the Fine numbers, 1,0,1,2,6,18,57,186, (named after information theorist Terrence Fine of Cornell University.)

$$F(z) = \sum_{n=0}^{\infty} f_n z^n = \frac{1-\sqrt{1-4z}}{z(3-\sqrt{1-4z})} \cdots [A000957]$$

The fine numbers count trees with even number of updegree at the root and they also count ordered trees with no leaf at level 1.
The *Fine function* satisfies the identity $F(z) = \frac{C}{1+zC}$. [2]

Definition 1.2.5. [12] *The Motzkin function,* $M(z)$, is the generating function for the Motzkin numbers $1, 1, 2, 4, 9, 21, 51, 127, 323, \cdots$ given by

$$M(z) = \sum_{n=0}^{\infty} m_n z^n = \frac{1-z-\sqrt{1-2z-3z^2}}{2z^2} \cdots [A001006].$$

The Motzkin numbers have several combinatorial interpretations. The one of interest to us is that they count the number of trees with n edges, where every vertex has degree $0, 1$, or 2. Such trees are called $\{0.1.2\}-$trees. Figure 2 shows the Motzkin trees with $n = 0, 1, 2, 3, 4, 5$ edges.
The Motzkin function satisfies the identity $M(z) = \frac{1}{1-z} C\left(\frac{z^2}{(1-z)^2}\right)$. [2]

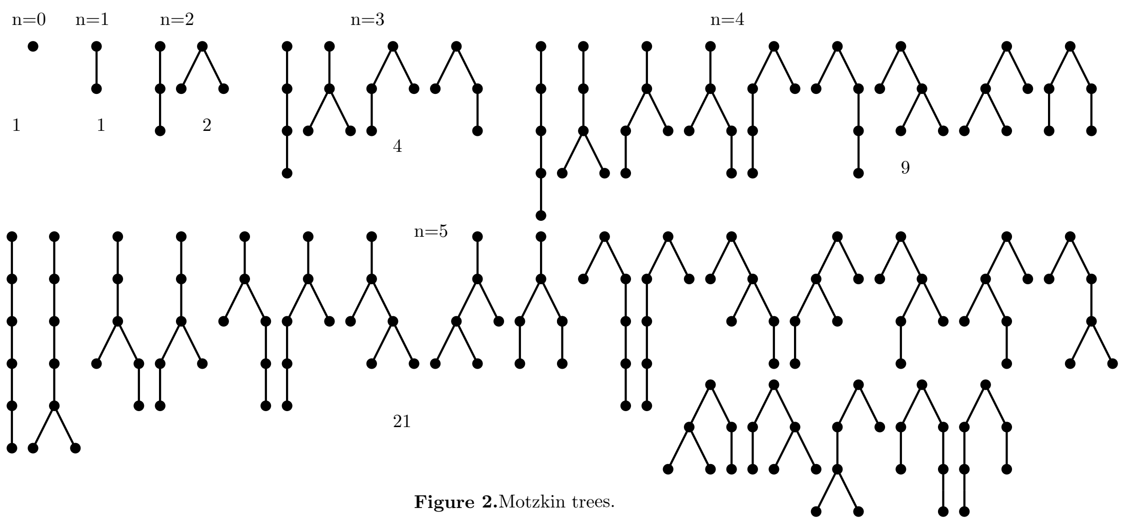
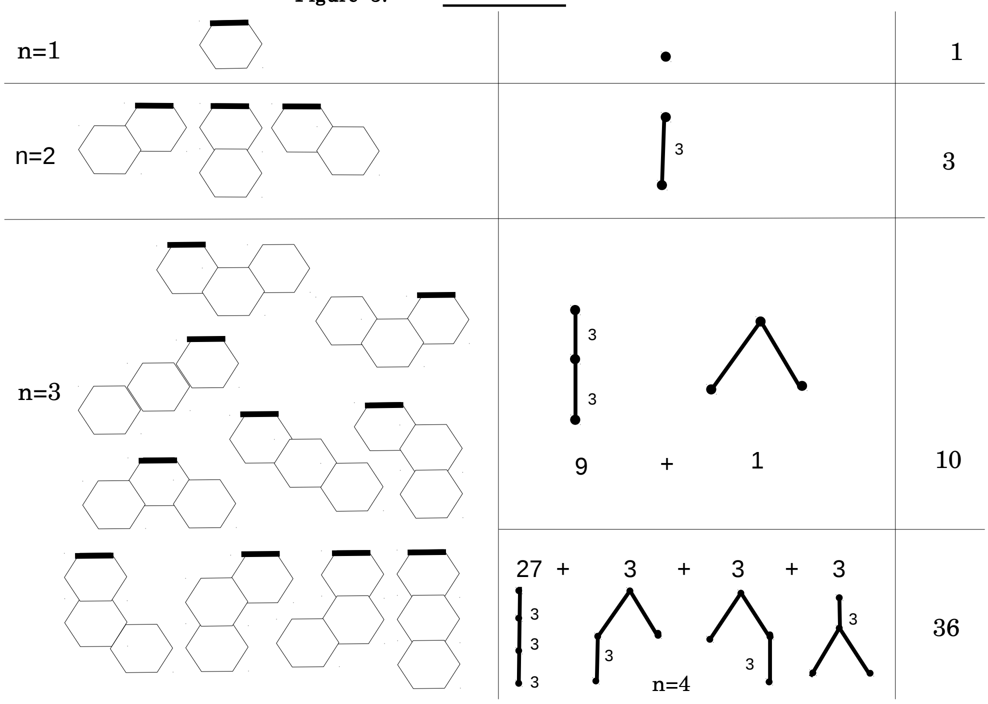


Figure 2. Motzkin trees.

Definition 1.2.6. [5] *The Hex function, $H(z)$,* is the generating function for the Hex numbers $1, 3, 10, 36, 137, 543, \cdots$ [$A002212$] which arise as a result of putting benzene molecules together with no three meeting at a point, see Figure 2 below. As trees, these are trees with every vertex having updegree $0, 1,$ or 2 but the one case can be left, vertical or right but if two they must be left and right edges.[Name due to an obvious tree-like poly-hexes[14]]. The Hex function is given by

$$H(z) = \sum_{n=0}^{\infty} h_n z^n = \frac{1 - 3z - \sqrt{1 - 6z + 5z^2}}{2z^2}.$$

Figure 3. Hex Trees



The thick edges represent the roots of the hex trees.

Chapter 2

Riordan Group

2.1 Riordan Array

Definition 2.1.1. Given two power series $g(z)$ and $f(z)$ from the ring of formal power series $\mathbb{C}[[z]]$ where

$$g(z) = g = 1 + g_1 z + g_2 z^2 + g_3 z^3 + \cdots \quad = \quad 1 + \sum_{n\geq 1} g_n z^n \text{ and} \tag{2.1}$$

$$f(z) = f = f_1 z + f_2 z^2 + f_3 z^3 + \cdots \quad = \quad \sum_{n\geq 1} f_n z^n. \tag{2.2}$$

a Riordan array denoted by $(g(z), f(z))$ is an infinite lower triangular matrix $L = [l_{n,k}]_{n,k\geq 0}$ where $l_{n,k}$ is the coefficient of z^n of the generating function of the k^{th} column $g(z)(f(z))^k$ of $(g(z), f(z))$, for $k = 0, 1, 2, ...$

We can define the linear operator for extracting coefficients as

$$[z^n]A(z) = a_n \quad \text{where} \quad A(z) = \sum_{n\geq 0} a_n z^n.[12] \tag{2.3}$$

Facts that we will use without proof are that

$$\textbf{a.)} \quad [z^n]\, C^k = \frac{k}{2n+k}\binom{2n+k}{n}, \quad \textbf{b.)} \quad [z^n]\, BC^k = \binom{2n+k}{n}.[12] \tag{2.4}$$

With the notation we have $l_{n,k} = [z^n]\, g f^k$. The standard notation for a Riordan array is the rather unimaginative (g, f) . It can be thought of as a geometric sequence of generating functions with lead term g and multiplier term f.

Next if we multiply the Riordan array by the column vector $[a_0, a_1, a_2, \cdots]^T$ where $A(z) = \sum_{n\geq 0} a_n z^n$ we see that a_0 multiplies the column g, a_1 multiplies the column gf, and in general a_k multiplies the column gf^k. If the resulting column is $[b_0, b_1, b_2, \cdots]^T$ with
$B(z) = B = \sum_{n\geq 0} b_n z^n$ then we obtain the key result

$$\begin{aligned} a_0 g + a_1 g f + a_2 g f^2 + \cdots &= \sum_{k\geq 0} a_k g f^k \\ &= g \sum_{k\geq 0} a_k f^k \\ &= g(z) A(f(z)) = B(z), \end{aligned}$$

we could write this operation as

$$(g(z), f(z)) * A(z) = g(z) A(f(z)) = B(z). \qquad [10] \tag{2.5}$$

So to multiply two Riordan arrays (g, f) and (G, F) we take a typical column $G(z)[F(z)]^k$ of (G, F) and use this as $A(z)$ in (2.3). This quickly yields the matrix multiplication

$$(g(z), f(z)) * (G(z), F(z)) = (g(z)G(f(z)), F(f(z))),$$

which is the fundamental Theorem of Riordan Arrays and is abbreviated as the **FTRA.** It is our key building block.

Definition 2.1.2. A Riordan array is an element of a non commutative group under the above operation $*$ called the *Riordan group* denoted $(\mathcal{R}, *)$ if $f_1 \neq 0$. But for convenience we assume $f_1 = 1$.

This group $(\mathcal{R}, *)$ was introduced by Shapiro et al.[13]

2.2 Some Structural properties of $(\mathcal{R}, *)$

In this section, we discuss some properties of $\mathcal{R}$ to show that, it is indeed a group under the matrix multiplication $*$. We start with the identity and inverse.

Theorem 2.2.1. *(a). The identity element of $(\mathcal{R}, *)$ is $(1, z)$.*

(b). For any element $(g, f) \in \mathcal{R}$ the inverse $(g, f)^{-1} = \left(\frac{1}{g(\bar{f}(z))}, \bar{f}(z)\right)$ where $\bar{f}(f(z)) = f(\bar{f}(z)) = z$, so $\bar{f}$ is the compositional inverse of f.

Proof. (a). Let $(g, f) \in (\mathcal{R}, *)$, then by the **FTRA** $(g(z), f(z)) * (1, z) = (g(z) \cdot 1, f(z)) = (g(z), f(z))$ and $(1, z) * (g(z), f(z)) = (g(z), f(z))$. Thus the identity of $(\mathcal{R}, *)$ is $(1, z)$.

(b). For the inverse, we use the fact that any element in a group must have an inverse so if $(g, f) \in \mathcal{R}$ then it has an inverse and multiplying an element by its inverse in the group should give us the identity.
So given $(g, f) \in (\mathcal{R}, *)$, consider $\left(\frac{1}{g(\overline{f}(z))}, \overline{f}(z)\right) = (g(z), f(z))^{-1}$, then by **FTRA**, we have

$$\begin{aligned}
(g(z), f(z)) * (g(z), f(z))^{-1} &= (g(z), f(z)) * \left(\frac{1}{g(\overline{f}(z))}, \overline{f}(z)\right) \\
&= \left(g(z) \cdot \frac{1}{g(\overline{f}(f(z)))}, \overline{f}(f(z))\right) \\
&= \left(g(z) \cdot \frac{1}{g(z)}, \overline{f}(f(z))\right) \\
&= (1, z) \\
\text{and} \quad (g(z), f(z))^{-1} * (g(z), f(z)) &= \left(\frac{1}{g(\overline{f}(z))}, \overline{f}(z)\right) * (g(z), f(z)) \\
&= \left(\frac{1}{g(\overline{f}(z))} \cdot g(\overline{f}(z)), f(\overline{f}(z))\right) \\
&= (1, z)
\end{aligned}$$

$$\text{Hence} \quad (g, f)^{-1} = \left(\frac{1}{g(\overline{f}(z))}, \overline{f}(z)\right) \text{ is the inverse of } \quad (g(z), f(z)).$$

Since matrix multiplication is associative we have shown that indeed $(\mathcal{R}, *)$ is a group. □

The example below shows how the inverse of a given function is found.

Example 2.2.2. We find the compositional inverse $\overline{f}(z)$ when $f(z) = \frac{z}{1-z}$, also when $f(z) = zC$.
Given $f(z) = \frac{z}{1-z}$ we replace z by $\overline{f}(z)$ so we have $z = \frac{\overline{f}}{1-\overline{f}}$ and then solve for $\overline{f}$.
$z(1 - \overline{f}) = \overline{f}$, $\quad z - z\overline{f} = \overline{f}$, $\quad \overline{f} + z\overline{f} = z$, $\quad$ therefore $\quad \overline{f}(z) = \frac{z}{1+z}$.

Similarly If $f(z) = zC$, we know $C = 1 + zC^2 = \frac{1}{1-zC}$, so $\quad f(z) = zC = \frac{z}{1-zC}$ and replacing zC in the denominator by $f(z)$ we obtain $f(z) = \frac{z}{1-f(z)}$, now replace z

by $\overline{f}(z)$ to obtain $z = \frac{\overline{f}(z)}{1-z}$, hence our $\overline{f}(z) = z(1-z)$.

Using the technique in the above example and Theorem 2.2.1 (b), the inverse of a Riordan group element can be found.

Example 2.2.3. [3] The matrix $F = (1, z(1+z))$ is called the Fibonacci matrix since its row sum is the Fibonacci number sequence $1, 1, 3, 5, 8, \cdots [A800045]$. To find F^{-1} all we need to do is to find $\overline{f}(z)$, for $f(z) = z(1+z)$ We replace z by $\overline{f}$ giving us $\overline{f}^2 + \overline{f} - z = 0$ and now we solve for $\overline{f}$ using the quadratic formula $i.e \quad \overline{f} = \frac{-1-\sqrt{1+4z}}{2} = zC(-z)$. Thus

$$F^{-1} = (1, z(1+z))^{-1} = (1, zC(-z)) = \begin{bmatrix} 1 & & & & \\ 0 & 1 & & & \\ 0 & -1 & 1 & & \\ 0 & 2 & -2 & 1 & \\ 0 & -5 & 5 & -3 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

From (2.5) it is convenient to write $(g, f) * A = B$ to indicate that we have shifted from matrix multiplication to composition of functions via the **FTRA**. We are now in the position to invert some binomial identities by finding the inverses in the Riordan group $(\mathcal{R}, *)$, we have that,

$$(g, f) * A = B \text{ if and only if } (g, f)^{-1} B = A.$$

In the example below we evaluate and invert an identity using both the usual primitive way and the **FTRA**.

Example 2.2.4. We evaluate the sum below and invert the identity

$$\sum_{k\geq 0}^{n} \binom{n+k}{2k} (-4)^k.$$

To get a feel of this, we begin by letting

$$S(n) = \sum_{k\geq 0}^{n} \binom{n+k}{2k} (-4)^k$$

and trying a few cases when $n = 0, 1, 2, 3, \cdots$

$$\text{for} \quad n = 0; \quad S(0) = \sum_{k\geq 0}^{0} \binom{0+k}{2k} (-4)^k = 1$$

$$\text{for} \quad n = 1; \quad S(1) = \sum_{k\geq 0}^{1} \binom{1+k}{2k} (-4)^k = \binom{1}{0}(-4)^0 + \binom{2}{2}(-4)^1 = 1.1 + 1.-4 = -3$$

$$\text{for} \quad n = 2; \quad S(2) = \sum_{k\geq 0}^{2} \binom{2+k}{2k} (-4)^k = \binom{2}{0}(-4)^0 + \binom{3}{2}(-4)^1 + \binom{4}{4}(-4)^2$$

$$= 1 \cdot 1 + 3 \cdot (-4) + \cdot 16 = 5$$

$$\text{for} \quad n = 3; \quad S(3) = \sum_{k\geq 0}^{3} \binom{3+k}{2k} (-4)^k = \binom{3}{0}(-4)^0 + \binom{4}{2}(-4)^1 + \binom{5}{4}(-4)^2$$

$$+ \binom{6}{6}(-4)^3$$

$$= 1 \cdot 1 - 6 \cdot (-)4 + 5 \cdot 16 - 64 = -7$$

for $n = 4$ and $n = 5$ we get 9 and -11 respectively, so we can guess $S(n)$ is the sequence of coefficients of the generating function $1-3z+5z^2-7z^3+9z^4-11z^5+\cdots$ i.e the alternating odd numbers.
Now we discover the pattern and put this into matrix form as

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 0 \\ 1 & 6 & 5 & 1 & 0 \\ 1 & 10 & 15 & 7 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} 1 \\ -4 \\ 16 \\ -64 \\ 256 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ 5 \\ -7 \\ 9 \\ \vdots \end{bmatrix}.$$

The generating function for the left most column is $1/(1-z) = 1 + z + z^2 + z^3 + \cdots$ and the next column $[0, 1, 3, 6, 10, \cdots]^T$ has as its generating function $z/(1-z)^3$ followed by $[0, 0, 1, 5, 15, 35,]^T$ with generating function $z^2(1-z)^5$. Each move to the right by one column corresponds to multiplying the generating function by $z/(1-z)^2$. Meanwhile the generating function (briefly GF) for $[1, -4, 16, -64, \cdots]^T$ is $1/(1+4z)$ and the GF for $[1, -3, 5, -7, 9, \cdots]^T$ is $(1-z)/(1+z)^2$. Expanding the LHS of the matrix equation using generating functions gives us

$$1\left(\frac{1}{1-z}\right)-4\left(\frac{z}{(1-z)^3}\right)+4^2\left(\frac{z^2}{(1-z)^5}\right)-\cdots$$
$$=\left(\frac{1}{1-z}\right)\cdot\left(1-4\left(\frac{z}{(1-z)^2}\right)+4^2\left(\frac{z^2}{(1-z)^4}\right)-\cdots\right)$$
$$=\left(\frac{1}{1-z}\right)\left(\frac{1}{1+4\left(\frac{z}{(1-z)^2}\right)}\right)=\left(\frac{1}{1-z}\right)\left(\frac{1}{\frac{1-2z+z^2+4z}{(1-z)^2}}\right)$$
$$=\left(\frac{1}{1-z}\right)\left(\frac{(1-z)^2}{(1+z)^2}\right)=\frac{1-z}{(1+z)^2}=1-3z+5z^2-7z^3+\cdots$$

Thus we have shown using generating functions, that indeed $S(n)=(-1)^n(2n+1)$, is the alternating odd numbers.

Using the **FTRA** we have $g(z)=\frac{1}{1-z}$ and $f(z)=\frac{z}{(1-z)^2}$ and $A=\frac{1}{1+4z}$ so we have

$$\begin{aligned}\left(\frac{1}{1-z},\frac{z}{(1-z)^2}\right)*\left(\frac{1}{1+4z}\right) &= \left(\frac{1}{1-z}\right)\cdot\left(\frac{1}{1+4\left(\frac{z}{(1-z)^2}\right)}\right)\\ &= \left(\frac{1}{1-z}\right)\cdot\left(\frac{(1-z)^2}{1-2z+z^2+4z}\right)\\ &= \frac{1-z}{1+2z+z^2}\\ &= \frac{1-z}{(1+z)^2}.\end{aligned}$$

From above which now can be written more compactly as $\left(\frac{1}{1-z},\frac{z}{(1-z)^2}\right)*\frac{1}{1+4z}=\frac{1-z}{(1+z)^2}$, we can invert it to say $\left(\frac{1}{1-z},\frac{z}{(1-z)^2}\right)^{-1}*\frac{1-z}{(1+z)^2}=\frac{1}{1+4z}$. Now to find the inverse $\left(\frac{1}{1-z},\frac{z}{(1-z)^2}\right)^{-1}$, we can take the primitive approach of finding

$$\begin{bmatrix}1&0&0&0&0\\1&1&0&0&0\\1&3&1&0&0\\1&6&5&1&0\\1&10&15&7&1\\\vdots&\vdots&\vdots&\vdots&\vdots\end{bmatrix}^{-1}$$

using the row reduction formula one subdiagonal at a time. This yields

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ \cdots & -3 & 1 & 0 & 0 \\ \cdots & \cdots & -5 & 1 & 0 \\ \cdots & \cdots & \cdots & -7 & 1 \\ & & \cdots & & \end{bmatrix}, \quad \text{then} \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 \\ \cdots & 9 & -5 & 1 & 0 \\ \cdots & \cdots & 20 & -7 & 1 \\ & & \cdots & & \end{bmatrix}$$

then after another couple of steps

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 \\ -5 & 9 & -5 & 1 & 0 \\ 14 & -28 & 20 & -7 & 1 \\ -42 & 90 & -75 & 35 & -9 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

At this point we need to recognize the sequence $1, 1, 2, 5, 14, 42, \cdots [A000108]$ as the first few terms of $C(z)$ (Catalan numbers). Even if we do not recognize the sequence, the first few terms of $\overline{f} = \overline{f}\,(z)$ can be determined from the equations

$$\begin{aligned} \frac{1}{g(\overline{f}(z))} &= 1 - z + 2z^2 - 5z^3 + 14z^4 - 42z^5 + \cdots \\ \text{and} \quad \frac{1}{g(\overline{f}(z))} \cdot \overline{f}(z) &= z - 3z^2 + 9z^3 - 28z^4 + 90z^5 - \cdots \end{aligned}$$

and this yields $\overline{f}(z) = z - 2z^2 + 5z^3 - 14z^4 + 42z^5 - \cdots$.
The second approach is to compute $(g, f)^{-1} = \left(\frac{1}{g(\overline{f})}, \overline{f}\right)$. The main step is finding $\overline{f}$. Here we have $f = \frac{z}{(1-z)^2}$ so that $z = \frac{\overline{f}}{(1-\overline{f})^2}$ and by the quadratic formula, $\overline{f} = \frac{1+2z-\sqrt{1+4z}}{2z} = z - 2z^2 + 5z^3 - 14z^4 + 42z^5 - 132z^6 + 429z^7 - 1430z^8 + O\left(z^9\right)$. Thus from (2.4), the (n, k) entry of the inverse matrix is

$$\begin{aligned} [z^n]\, z^k C^{2k+1}(-z) &= \left[z^{n-k}\right] C^{2k+1}(-z) \\ &= (-1)^{n-k} \frac{2k+1}{2(n-k)+2k+1} \binom{2(n-k)+2k+1}{n-k} \\ &= (-1)^{n-k} \frac{2k+1}{2n+1} \binom{2n+1}{n-k}. \end{aligned}$$

The inverse identity is thus

$$\sum_{k\geq 0}(-1)^{n-k}\frac{2k+1}{2n+1}\binom{2n+1}{n-k}(-1)^k(2k+1)=(-4)^n.$$

or more elegantly

$$\sum_{k\geq 0}\frac{2k+1}{2n+1}\binom{2n+1}{n-k}(2k+1)=4^n.$$

2.2.5 Riordan Arrays and Generating Trees

Riordan arrays can also be characterized using two sequences called the A- and Z- sequences. Here we look at the connection between Riordan arrays and the generating trees and investigate what $g(z)$, $f(z)$ and the $A-$ sequence and $Z-$ sequence tell us about the degree of the root and the degree of an arbitrary vertex in a generating tree.

Definition 2.2.6. [6] A Riordan array $L=(g(z),f(z))=[l_{n,k}]_{n,k\geq 0}$ can be characterized by two sequences $A=(a_0,a_1,\cdots)$ and $Z=(z_0,z_1,\cdots)$ satisfying

$$l_{n+1,0}=\sum_{j\geq 0}z_jr_{n,j}\quad\text{and}\quad l_{n+1,k+1}=\sum_{j\geq 0}a_jr_{n,k+j}\quad\text{for}\quad n,k\geq 0.$$

These are referred to respectively as the A-*sequence* and the Z-*sequence*.
The Z-sequence characterizes column 0 while the A-sequence characterizes all the other columns. If $A(z)=a_0+a_1z^1+a_2z^2+\cdots$ and $Z(z)=z_0+z_1z+z_2z^2+\cdots$ are the generating functions for the $A-$ and $Z-$ sequences respectively, then it follows that

$$\begin{aligned} g(z)=\frac{1}{1-zZ(f(z))} &= 1+zZ(f(z))+(zZ(f(z)))^2+(zZ(f(z)))^3+\cdots. \\ &= 1+[Z(f(z))]z+[Z(f(z))]^2z^2+[Z(f(z))]^3z^3+\cdots \end{aligned} \tag{2.6}$$

and

$$\begin{aligned} f(z) &= zA(f(z))=a_0z+(a_1f(z))z+(a_2f^2(z))z+\cdots. \\ &= a_0z+\left[\frac{a_1f(z)}{z}\right]z^2+\left[\frac{a_2f^2(z)}{z^2}\right]z^3+\cdots. \end{aligned} \tag{2.7}$$

Definition 2.2.7. [7] An infinite matrix $[l_{n,k}]_{n,k\geq 0}$ is said to be *'associated' to a generating tree* with root (c) *(AGT matrix for short)* if $l_{n,k}$ is the number of nodes at level n with label $k+c$. By convention, the level of the root is 0.

The following examples demonstrate how to obtain a generating tree given matrix $[l_{n,k}]_{n,k\geq 0}$ and conversely how to obtain a matrix $[l_{n,k}]_{n,k\geq 0}$ given the generating tree.

Example 2.2.8. [7] The Catalan triangle is

$$(C(z), zC(z)) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 2 & 1 & 0 & 0 \\ 5 & 5 & 3 & 1 & 0 \\ 14 & 14 & 9 & 4 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

In order to specify a generating tree, we have to specify a label for the root and a set of rules explaining how to derive the children from the labelled parent.

$$\text{specification}(1) = \begin{cases} root : (2) \\ rule : (k) \rightarrow (2)(3)\cdots(k+1). \end{cases}$$

Figure 4a below is the generating tree for specification (1) which is referred to as the *Catalan generating tree.*

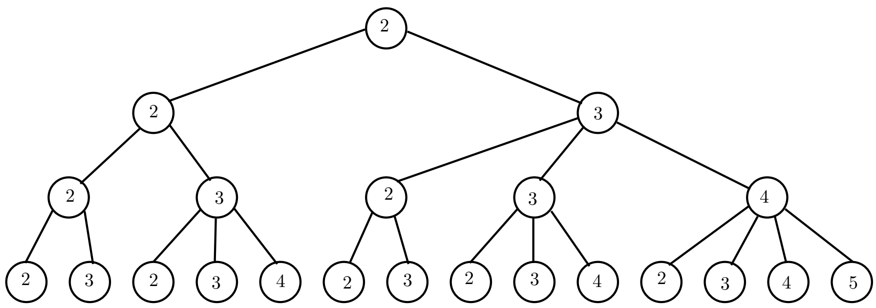


Figure 4a. Catalan generating tree.

Conversely given a generating tree and a specified label for the root and a set of rules explaining how to derive the children from a labelled parent, the AGT matrix can be derived.

Example 2.2.9. [7] Given

$$\text{specification}(2) = \begin{cases} root : (2) \\ rule : (k) \rightarrow (k)(k+1) \end{cases}$$

which gives the *Pascal generating tree* in Figure 4b below,

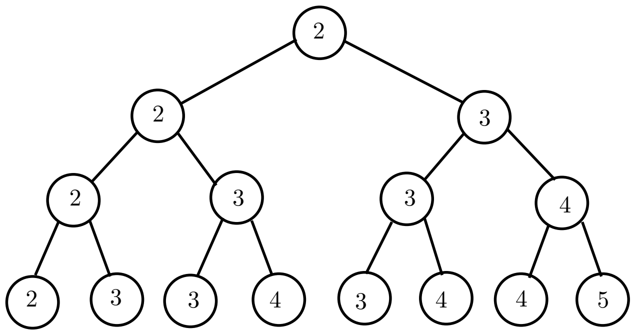


Figure 4b. Pascal generating tree.

then the AGT matrix is given as

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 1 & 3 & 3 & 1 & 0 \\ 1 & 4 & 6 & 4 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}. \quad \text{This is the Pascal triangle.}$$

Definition 2.2.10. [8] Given a generating tree specification t_1 and the corresponding AGT matrix T_1, we define the *generating tree specification inverse of* t_1 as the specification t_2 having $T_2 = T_1^{-1}$ as AGT matrix.

Example 2.2.11. We look at the inverse of the Pascal triangle (P^{-1}) and its generating tree.

$$P^{-1} = \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ 1 & -2 & 1 & & \\ -1 & 3 & -3 & 1 & \\ 1 & -4 & 6 & -4 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

$$\text{specification}(3) = \begin{cases} root : (2) \\ rule : (k) \rightarrow \overline{(k)}(k+1) \\ \qquad\quad \overline{(k)} \rightarrow (k)\overline{(k+1)} \end{cases}$$

where $(\overline{k})$ means we assign a negative number to its corresponding $l_{n,k}$ in the AGT matrix.
Figure 4c below is the generating tree for specification (3).

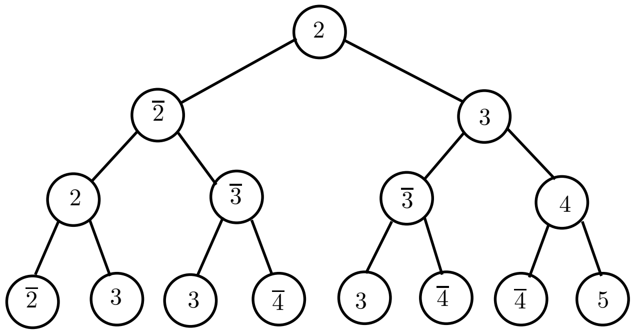


Figure 4c. Inverse pascal generating tree.

In [7], they discussed two interesting quantities related to generating trees, namely s_n, the total number of vertices at level n and the average label v_n of a vertex at the same level n. The obvious combinatorial meanings of these quantities are that, going to the AGT matrix $[l_{n,k}]_{n,k\geq 0}$ associated to the generating tree, we see s_n is equal to the row sums of $[l_{n,k}]_{n,k\geq 0}$ and v_n is the ratio between the weighted row sums W_n of $[l_{n,k}]_{n,k\geq 0}$ and s_n, plus c :

$$s_n = \sum_{k=0}^{n} l_{n,k}, \quad v_n = \frac{\sum_{k=0}^{n}(k+c)l_{n,k}}{s_n} = \frac{\sum_{k=0}^{n} k l_{n,k}}{s_n} + c = \frac{W_n}{s_n} + c.$$

The quantities s_n and W_n are easily computed if the AGT matrix is a Riordan array.

Degree of the root

Here we develop the formula for calculating the degree of the root for generating trees. The AGT matrix from a generating tree can be written as

$$(g(z), f(z)) = \begin{bmatrix} \vdots & & & & \\ \vdots & \vdots & & & \\ \vdots & \vdots & \vdots & . & \\ g & gf & gf^2 & gf^3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The g column is just (2.1) and the gf column is

$$= f_1 z + (g_1 f_1 + f_2)z^2 + (g_2 f_1 + g_1 f_2 + f_3)z^3 + \cdots . \tag{2.8}$$

So from generating trees, it implies that degree of the root (deg(root)) is the sum of g_1 and f_1. In finding degrees we ignore the minus signs that may appear in inverse AGT matrices.

$$\begin{aligned} \deg(\text{root}) &= |g_1| + |f_1| \geq 0 \\ &= 1 + |g_1| \geq 0 \quad \text{since} \quad f_1 = 1 \quad \text{from definition} \quad (2.1.2). \end{aligned}$$

Comparing (2.1) and (2.6) we have

$$g_1 = Z(f(z)).$$

And also comparing (2.2) and (2.7) we have

$$f_1 = a_0 = 1.$$

So the degree of the root from the A- and Z- sequences is given by

$$\begin{aligned} \deg(\text{root}) &= |Z(f(z))| + |a_0| \geq 0 \\ &= 1 + |Z(f(z))| \geq 0. \end{aligned}$$

Degree of an arbitrary vertex

Similarly, we develop a formula for calculate the degree of an arbitrary vertex in a generating tree in terms of $g(z)$ and $f(z)$ and the A- and Z-sequence.

We recall that,

$$g(z) = \frac{1}{1 - zZ(f(z))} \quad \text{and} \quad f(z) = zA(f(z))$$

and $l_{n,k}$ is the coefficient of z^n of the generating function of the k^{th} column $g(z)(f(z))^k$ i.e

$$\begin{aligned} l_{n,k} = [z^n]gf^k &= [z^n]\frac{(zA(f(z)))^k}{1 - zZ(f(z))} \\ &= [z^{n-k}]\frac{(A(f(z)))^k}{1 - zZ(f(z))} \end{aligned}$$

To find the degree of an arbitrary vertex with label $(k+c)$ at level n in a generating tree involving $g(z)$ and $f(z)$, we note that, first of all

$$deg((k + c)_n) = \#(k + c)\#(k + c)^r$$

where $(k+c)_n$ is the vertex with label $(k+c)$ at level n and $\#(k+c)$ is the number of vertices with label $(k + c)$ at level n and $\#(k + c)^r$ is the number of children of the vertex with label $(k + c)$ in the rule. We also note that, every label has a corresponding $l_{n,k} \in L$ which is the same as $\#(k + c)$, so we obtain

$$deg((k + c)_n) = (l_{n,k})\#(k + c)^r.$$

Now we can write our formula for estimating the degree of an arbitrary vertex with label $(k + c)$ at level n in a generating tree as

$$deg((k + c)_n) = [z^n]gf^k\#(k + c)^r$$

Now we write the formula for estimating the degree of an arbitrary vertex with label $(k + c)$ at level n in a generating tree involving the $A-$ and $Z-$ sequence as

$$deg((k + c)_n) = [z^{n-k}]\frac{(A(f(z)))^k}{1 - zZ(f(z))}\#(k + c)^r$$

Chapter 3

Subgroups, Isomorphisms and Left Branch Length

In this section we present some subgroups of the Riordan group $(\mathcal{R}, *)$. We define a pseudo-ring and demonstrate that a normal subgroup of $(\mathcal{R}, *)$ forms a pseudo-ring. We also discuss isomorphisms between the subgroups and from the isomorphisms we construct a commutative diagram.

3.1 Some special subgroups

By special, we mean the lead sequence g of the Riordan array depends on the multiplier sequence f.

3.1.1 The Hitting time subgroup($\mathcal{H}$)

[9] This consists of elements of the form $\left(\frac{zf'(z)}{f(z)}, f(z)\right)$. To see that the members of $\mathcal{R}$ with this form constitute a subgroup, we note the following:

- The identity of $(\mathcal{R}, *)$, $(1, z)$ corresponds to $\left(\frac{z(z)'}{z}, z\right) \in \mathcal{H}$ where we have taken $f(z) = z$, so $f'(z) = 1 := (z)'$

- Closure under the matrix multiplication $*$

$$\left(\frac{zf'(z)}{f(z)}, f(z)\right) * \left(\frac{zF'(z)}{F(z)}, F(z)\right) = \left(\frac{zf'(z)}{f(z)} \cdot \frac{f(z)F'(f(z))}{F(f(z))}, F(f(z))\right)$$
$$= \left(\frac{z(F(f(z)))'}{F(f(z))}, F(f(z))\right) \in \mathcal{H}$$

- The inverse of an element $\left(\frac{zf'(z)}{f(z)}, f(z)\right)$ is $\left(\frac{z}{\overline{f}(z)f'(\overline{f}(z))}, \overline{f}(z)\right)$ since

$$\left(\frac{Id \cdot f'}{f}, f\right) * \left(\frac{Id}{\overline{f} \cdot (f' \circ \overline{f})}, \overline{f}\right) = \left(\frac{Id \cdot f'}{f} \cdot \frac{f}{Id \cdot f'}, \overline{f} \circ f\right)$$
$$= (1, Id)$$

with $\overline{f}$ as the compositional inverse of f.

3.1.2 The Bell Subgroup ($\mathcal{B}$)

Elements in this subgroup are of the form $(g(z), zg(z))$ or $(f(z)/z, f(z))$. To see that this subset of elements of $(\mathcal{R}, *)$ forms a subgroup we have

- Identity

  $\left(\frac{z}{z}, z\right) = (1, z) \in \mathcal{B}$ i.e $f(z) = z$.

- Closure under the matrix multiplication $*$

$$\left(\frac{f(z)}{z}, f(z)\right) * \left(\frac{F(z)}{z}, F(z)\right) = \left(\frac{f(z)}{z} \cdot \frac{F(f(z))}{f(z)}, F(f(z))\right)$$
$$= \left(\frac{F(f(z))}{z}, F(f(z))\right) \in \mathcal{B}$$

- The inverse of an element $\left(\frac{f(z)}{z}, f(z)\right)$ is

$$\left(\frac{1}{\frac{f(\overline{f}(z))}{\overline{f}(z)}}, \overline{f}(z)\right) = \left(\frac{\overline{f}(z)}{z}, \overline{f}(z)\right) \in \mathcal{B}$$

3.1.3 The Derivative Subgroup ($\mathcal{D}$)

Elements here have the lead sequence being the first derivative of the multiplier sequence *i.e* $(f'(z), f(z))$ and the fact that these elements of $\mathcal{R}$ form a subgroup follows from the observations below

- Identity

  $(1, z) \in \mathcal{D}$ since $f(z) = z$ and $f'(z) = 1$

- Closure under the multiplication.

$$
\begin{aligned}
(f'(z), f(z)) * (F'(z), F(z)) &= (f'(z) \cdot F'(f(z), F(f(z))) \\
&= ((F \circ f)'(z), (F \circ f)(z))) \in \mathcal{D}
\end{aligned}
$$

- The inverse of an element $(f'(z), f(z))$ is $\left(\frac{1}{f'(\bar{f}(z))}, \bar{f}(z)\right) \in \mathcal{D}$ since

$$
\begin{aligned}
(f'(z), f(z)) * \left(\frac{1}{f'(\overline{f}(z))}, \overline{f}(z)\right) &= \left(f'(z) \cdot \frac{1}{f'(z)}, (\overline{f} \circ f)(z)\right) \\
&= (1, (\overline{f} \circ f)(z)) \\
&= (1, z).
\end{aligned}
$$

Example 3.1.4. Given $f(z) = \frac{z}{1-z}$ then $f'(z) = \frac{(1-z)(z)'-z(1-z)'}{(1-z)^2} = \frac{1}{(1-z)^2}$, $\overline{f}(z) = \frac{z}{1+z}$ and an element in $\mathcal{D}$ will be $\left(\frac{1}{(1-z)^2}, \frac{z}{1-z}\right)$ and so the inverse $\left(\frac{1}{f'(\overline{f}(z))}, \overline{f}(z)\right) = \left(\left(1 - \frac{z}{1+z}\right)^2, \frac{z}{1+z}\right) = \left(\frac{1}{(1+z)^2}, \frac{z}{1+z}\right) \in \mathcal{D}$ since $\overline{f}'(z) = \frac{(1+z)(z)'-z(1+z)'}{(1+z)^2} = \frac{1}{(1+z)^2} = \frac{1}{f'(\overline{f}(z))}$.

3.2 Generalised subgroups

Here the subgroups we discuss may or may not have a direct connection between the lead sequence and the multiplier sequence.

3.2.1 Associated subgroup($\mathcal{S}$)

[3] Elements of the associated subgroup are of the form $(1, f(z))$. So from the definition of our lead sequence g and multiplier sequence f for the Riordan arrays

we can think of the lead sequence for the associated subgroup to be $g(0)$. An arbitrary element $(g, f) \in (\mathcal{R}, *)$ can be converted into an element in the associated subgroup by evaluating $g(z)$ at zero(0).

3.2.2 Appell Subgroup ($\mathcal{A}$)

Elements of this subgroup are of the form $(g(z), z)$. The Appell subgroup is a normal subgroup and we provide a proof in the theorem below.

Theorem 3.2.3. *[3] The Appell subgroup is a normal subgroup of* $(\mathcal{R}, *)$

Proof. Let $r = (g(z), f(z))$ be an arbitrary element of $(\mathcal{R}, *)$ and $T = (G(z), z)$ be an element of the Appell subgroup. Then

$$\begin{aligned} r^{-1}Tr &= (g(z), f(z))^{-1} * (G(z), z) * (g(z), f(z)) \\ &= \left(\frac{1}{g(\bar{f}(z))}, \bar{f}(z)\right) * (G(z) \cdot g(z), f(z)) \\ &= \left(\frac{G(\bar{f}(z)) \cdot g(\bar{f}(z))}{g(\bar{f}(z))}, f(\bar{f}(z))\right) \\ &= (G(\bar{f}(z)), z) \end{aligned}$$

which is an element of the Appell subgroup hence the Appell subgroup is normal. □

The Riordan group is a semi-direct product of (i) the $\mathcal{A}$ and the $\mathcal{S}$ subgroups and(ii) the $\mathcal{A}$ and the $\mathcal{B}$.[3]. The products are as follows

(i) $(g(z), z) * (1, f(z)) = (g(z), f(z))$

(ii) $(zg(z)/f(z), z) * (f(z)/z, f(z)) = (g(z), f(z))$.

Given normality in $(\mathcal{R}, *)$, we now ask ourself the question: Can we extend the $\mathcal{A}$ Subgroup to a ring under a second defined operation? $\underline{ANSWER:}$ Lets see

Definition 3.2.4. Let $\circ$ be the componentwise composition operation, then $\forall (g(z), z), (G(z), z) \in \mathcal{A}$, $(g(z), z) \circ (G(z), z) = (g \circ G(z), z \circ z) = (g(G(z)), z)$ $\in \mathcal{A}$ $\quad i.e$ componentwise composition.

Definition 3.2.5. A *Pseudo-ring* is an algebraic structure which satisfies all the ring axioms except the existence of a multiplicative (second operation) identity.

Proposition 3.2.6. *The Appell subgroup is a pseudo-ring under composition*$(\circ)$.

Proof. $(\mathcal{A}, *)$ is an abelian group since it is a subgroup of $(\mathcal{R}, *)$ and $\forall (g(z), z)$, $(G(z), z) \in (\mathcal{R}, *)$ we have

$$(g(z), z) * (G(z), z) = (g(z)G(z), z) = (G(z)g(z), z) = (G(z), z) * (g(z), z).$$

In addition $\circ$ is associative and distributive over $*$. We have

$$\{(g(z), z) \circ (G(z), z)\} \circ (h(z), z) = (((g \circ G) \circ h)(z), z) = ((g \circ (G \circ h))(z), z)$$

by associativity of composition, and

$$(g(z), z) * \{(G(z), z) \circ (h(z), z)\} = (g(z), z) * ((G \circ h)(z), z) = (g(z) \cdot (G \circ h)(z), z)$$

$$= ([g \cdot (G \circ h)](z), z) = (g(z)G(z), z) \circ (g(z)h(z), z)$$

$$= \{(g(z), z) * (G(z), z)\} \circ \{(g(z), z) * (h(z), z)\}$$

$$\{(G(z), z) \circ (h(z), z)\} * (g(z), z) = ((G \circ h)(z), z) * (g(z), z) = ((G \circ h) \cdot g)(z), z)$$

$$= (G(z) \cdot g(z), z) \circ (h(z) \cdot g(z), z)$$

$$= \{(G(z), z) * (g(z), z)\} \circ \{(h(z), z) * (g(z), z)\}$$

so * is distributive over $\circ$ from the right and left.
Similarly,

$$(g(z), z) \circ \{(G(z), z) * (h(z), z)\} = \{(g(z), z) \circ (G(z), z)\} * \{(g(z), z) \circ (h(z), z)\}$$

and

$$\{(G(z), z) * (h(z), z)\} \circ (g(z), z) = \{(G(z), z) \circ (g(z), z)\} * \{(h(z), z) \circ (g(z), z)\}.$$

so $\circ$ is distributive over * from both left and right .

This proves the result. □

3.3 Isomorphism Of Riordan Subgroups

In [3], an isomorphism between the associated subgroup and the Bell subgroup was established through a mapping $(1, z(1+z)) \to (1+z, z(1+z))$ that removes the left most column and the topmost row of the Fibonacci matrix $F = (1, z(1+z))$ thus,

$$\begin{bmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ 0 & 1 & 1 & & & \\ 0 & 0 & 2 & 1 & & \\ 0 & 0 & 1 & 3 & 1 & \\ 0 & 0 & 0 & 3 & 4 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \mapsto \begin{bmatrix} & & & & \\ 1 & & & & \\ 1 & 1 & & & \\ 0 & 2 & 1 & & \\ 0 & 1 & 3 & 1 & \\ 0 & 0 & 3 & 4 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

gives the following propositions.

Proposition 3.3.1. *[3] The mapping $\alpha : (\underline{1}, f(z)) \mapsto (f'(z), f(z))$ is an isomorphism between the associated subgroup and the derivative subgroup.*

Proof. Consider the mapping that takes $(\underline{1}, f)$ to (f', f), where $\underline{1}$ is the constant mapping. Clearly, α is injective and onto. We now show using the **FTRA** and the definition of the mapping that α is a homomorphism.

$$\begin{aligned} \alpha[(\underline{1}, f_1) * (\underline{1}, f_2)] &= \alpha[(\underline{1}, f_2(f_1))] \\ &= \alpha[(\underline{1}, (f_2 \circ f_1)] \\ &= ((f_2 \circ f_1)', (f_2 \circ f_1)) \\ &= (f' \cdot f_2' \circ f_1, f_2 \circ f_1) \\ &= (f_1', f_1) * (f_2', f_2) \\ &= \alpha[(\underline{1}, f_1)] * \alpha[(\underline{1}, f_2)]. \end{aligned}$$

This proves the result. □

Proposition 3.3.2. *[3] The mapping $\beta : \left(\frac{f(z)}{z}, f(z)\right) \mapsto (\underline{1}, f(z))$ is an isomorphism between the Bell subgroup and the associated subgroup.*

Proof. Consider the mapping β which takes $(f(z)/z, f(z))$ to $(1, f(z))$. Clearly β is injective and onto. By the **FTRA** and the definition of the mapping we show that β is a homomorphism.

$$\begin{aligned}
\beta\left[\left(\frac{f_1}{Id}, f_1\right) * \left(\frac{f_2}{Id}, f_2\right)\right] &= \beta\left[\left(\frac{f_1}{Id}\cdot\frac{(f_2\circ f_1)}{f_1}, (f_2\circ f_1)\right)\right] \\
&= \beta\left[\left(\frac{(f_2\circ f_1)}{Id}, (f_2\circ f_1)\right)\right] \\
&= (1, (f_2\circ f_1)) \\
&= ((1, f_1) * (1, f_2)) \\
&= \beta\left[\left(\frac{f_1}{Id}, f_1\right)\right] * \beta\left[\left(\frac{f_2}{Id}, f_2\right)\right].
\end{aligned}$$

This proves the result. □

Using the Fundamental Theorem of Calculus and the fact that isomorphism is an equivalence relation, the isomorphism between the Bell and derivative and subgroups is given in the proposition below.

Proposition 3.3.3. *[3] The mapping $\gamma : \left(\frac{f(z)}{z}, f(z)\right) \mapsto (f'(z), f(z))$ is an isomorphism between the Bell subgroup and the derivative subgroup.*

Proof. Consider the mapping γ defined by

$$\gamma\left(\frac{f(z)}{z}, f(z)\right) = \left[\left(\int_0^z f'(t)dt\right)', f(z)\right].$$

Clearly γ is injective and onto. We now prove γ is a homomorphism using the **FTRA**, the fundamental theorem of calculus, and the definition of the mapping.

$$\begin{aligned}
\gamma\left[\left(\frac{f_1(z)}{z}, f_1(z)\right) * \left(\frac{f_2(z)}{z}, f_2(z)\right)\right] &= \gamma\left[\left(\frac{f_1(z)}{z}\cdot\frac{f_2(f_1(z))}{f_1(z)}, f_2(f_1(z))\right)\right] \\
&= \gamma\left[\left(\frac{f_2(f_1(z))}{z}, f_2(f_1(z))\right)\right] \\
&= \left[\left(\int_0^z (f_2(f_1(t)))'dt\right)', f_2(f_1(z))\right] \\
&= [(f_2(f_1(z)))', (f_2(f_1(z))] \\
&= [f_1'(z)(f_2'(f_1(z)), f_2(f_1(z))] \\
&= \left[\left(\int_0^z f_1'(t)dt\right)'\cdot\left(\int_0^{f_1(z)} f_2'(t)dt\right)', f_2(f_1(z))\right] \\
&= \left[\left(\int_0^z f_1'(t)dt\right)', f_1(z)\right] * \left[\left(\int_0^z f_2'(t)dt\right)', f_2(z)\right] \\
&= \gamma\left[\left(\frac{f_1(z)}{z}, f_1(z)\right)\right] * \gamma\left[\left(\frac{f_2(z)}{z}, f_2(z)\right)\right].
\end{aligned}$$

This proves the result. $\square$

Proposition 3.3.4. *The mapping* $\psi : \left(\frac{zf'(z)}{f(z)}, f(z)\right) \mapsto \left(\frac{f(z)}{z}, f(z)\right)$ *is an isomorphism between the hitting time subgroup and the Bell subgroup.*

Proof. Consider the mapping ψ that takes $\left(\frac{zf'}{f}, f\right)$ to $\left(\frac{f}{z}, f\right)$. Clearly ψ is injective and onto. We now show that ψ is a homomorphism. By the Riordan matrix multiplication and the definition of the mapping we obtain

$$\begin{aligned}
\psi\left[\left(\frac{zf'(z)}{f(z)}, f(z)\right) * \left(\frac{zF'(z)}{F(z)}, F(z)\right)\right] &= \psi\left[\left(\frac{zf'(z)f(z)F'(f(z))}{f(z)F(f(z))}, F(f(z))\right)\right] \\
&= \psi\left[\left(\frac{z(F\circ f)'(z)}{(F\circ f)(z)}, (F\circ f)(z)\right)\right] \\
&= \left(\frac{F(f(z))}{z}, F(f(z))\right) \\
&= \left(\frac{f(z)}{z}, f(z)\right) * \left(\frac{F(z)}{z}, F(z)\right) \\
&= \psi\left(\frac{zf'(z)}{f(z)}, f(z)\right) * \psi\left(\frac{zF'(z)}{F(z)}, F(z)\right)
\end{aligned}$$

This proves the result. $\square$

Corollary 3.3.5. *The hitting time group is isomorphic to both the derivative and associated subgroups.*

Proof. Combine propositions 3.2.1, 3.2.2, 3.2.3 and 3.2.4 .

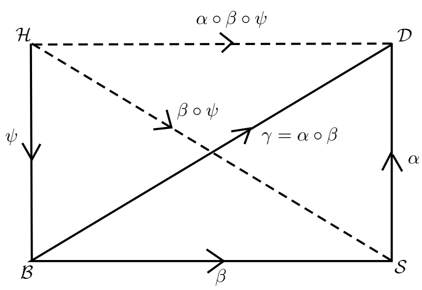


Figure 5. Commutative diagram.

We have

$$\begin{aligned}
&\beta\circ\psi : \mathcal{H} \to \mathcal{S} \quad \text{as an isomorphism} \\
\text{and} \quad &\alpha\circ\beta\circ\psi : \mathcal{H} \to \mathcal{D} \quad \text{as an isomorphism.}
\end{aligned}$$

$\square$

3.4 Left Branch Length

The *leftmost* edge of a tree is an edge on the left side of a tree seen from level zero (root level) to higher levels. At each level of a tree if there is only one edge it is considered as a leftmost edge. We define the *left branch length* as the total number of leftmost edges of a tree with a leaf at each level. In this section, we find the average number of left branch length in some classes of trees.

3.4.1 Left Branch Length for Motzkin Trees.

In the Motzkin trees every vertex has an updegree requirement of 0,1, or 2.
Step1: Draw the trees

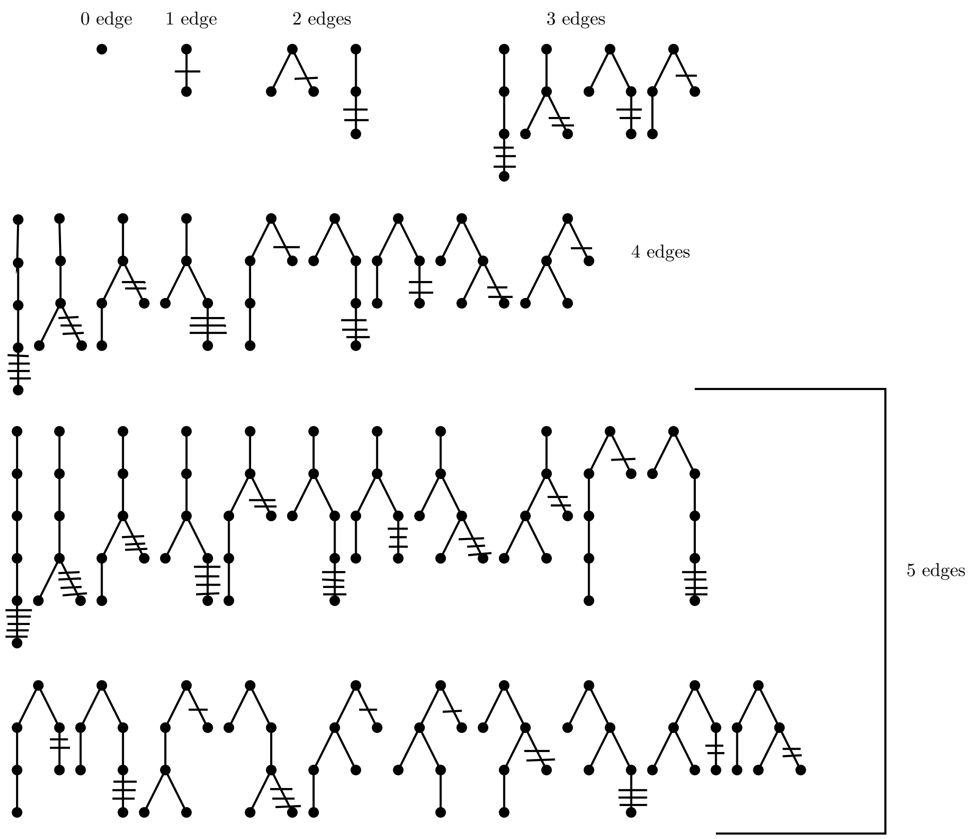


Figure 6. Motzkin trees with leftmost edges crossed.

The edges under consideration are the ones with crossings and the number of crossings on an edge indicates the level location of the leftmost edge.

Step2: We construct a matrix $(a_{n,h})$ where n is the number of edges in the tree and h is the level of the leftmost edge and we take row sums. So $a_{n,h}$ is the number of leftmost edges at level h for a tree with n edges.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 3 & 3 & 1 & 0 \\ 0 & 4 & 6 & 6 & 4 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 4 \\ 9 \\ 21 \\ \vdots \end{bmatrix} = M.$$

Step 3: What is f?

$$\begin{aligned} \text{we have} \quad (1, f) \cdot \frac{1}{1-z} &= M \\ \text{so} \quad \frac{1}{1-f} &= M \\ \text{thus} \quad fM &= M - 1. \\ \text{Recall} \quad M &= 1 + z + 2z^2 + 4z^3 + 9z^3 + 21z^4 + \cdots \\ &= 1 + zM + z^2M^2 \\ \text{so} \quad M - 1 &= zM + z^2M^2 \\ \text{thus} \quad fM &= zM + z^2M^2 \\ \text{i.e} \quad f &= z + z^2M. \end{aligned}$$

Thus f counts the number of Motzkin trees with left hand branch at level 1

Step 4: Find T_L = the generating function Motzkin trees for the total levels of the left branch.

NB: The $A(z)$ function is $\begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ \vdots \end{bmatrix} \leftrightarrow \frac{z}{(1-z)^2}$.

By the **FTRA** we have

$$\begin{aligned} T_L &= (1, f) \cdot \frac{z}{(1-z)^2} = \frac{f}{(1-f)^2} = fM^2 \\ \text{so} \quad T_L &= (z + z^2M)M^2 \\ &= zM^2 + z^2M^3 \\ \text{thus} \quad T_L &= M^2 - M. \end{aligned}$$

Step 5: We find the average number of trees in the Motzkin trees with left branches.

$$\frac{[z^n]T_L}{[z^n]M} = \frac{[z^n](M^2 - M)}{[z^n]M} = \frac{[z^n]M^2 - [z^n]M}{[z^n]M}$$
$$\text{but} \quad z^2M^2 = M - zM - 1$$
$$M^2 = \frac{M}{z^2} - \frac{M}{z} - \frac{1}{z^2}$$
$$\text{so} \quad [z^n]M^2 = [z^{n-2}]M - [z^{n-1}]M = m_{n+2} - m_{n+1}.$$

Thus

$$\begin{aligned}\frac{[z^n]T_L}{[z^n]M} &= \frac{m_{n+2} - m_{n+1} - m_n}{m_n} \\ &= \frac{3^2 - 3 - 1}{1} \to 5 \\ &\text{as} \quad n \to \infty.\end{aligned}$$

3.4.2 Left Branch Length for Ordered Trees

For an ordered tree with n edges, every vertex has the possibility of having an updegree of $0, 1, 2, 3, 4, \cdots (n+1)$.

Step 1: We draw the trees as in figure 6.

Step2: We find the $(a_{n,h})$ matrix and row sum as

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 2 & 1 & 0 & 0 \\ 0 & 5 & 5 & 3 & 1 & 0 \\ 0 & 14 & 14 & 9 & 4 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 5 \\ 14 \\ 42 \\ \vdots \end{bmatrix} = C.$$

Step 3: What is f?

$$\begin{aligned} \text{We have} \quad (1,f)\cdot\frac{1}{1-z} &= C \\ \text{so} \quad \frac{1}{1-f} &= C \\ \text{thus} \quad fC &= C-1. \\ \text{Recall} \quad C &= 1+zC^2 \\ C-1 &= zC^2 \\ \text{thus} \quad fC &= zC^2 \\ \text{i.e} \quad f &= zC. \end{aligned}$$

So f counts the number of ordered trees with left branch at level 1.

Step 4: Find T_L = the generating function for ordered trees for the total levels of the left branch.

NB: The $A(z)$ function is $\begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ \vdots \end{bmatrix} \leftrightarrow \frac{z}{(1-z)^2}$.

By the **FTRA** we have

$$\begin{aligned} T_L &= (1,f)\cdot\frac{1}{(1-z)^2} = \frac{f}{(1-f)^2} = fC^2. \\ \text{But} \quad f &= zC \\ \text{so} \quad T_L &= zC^3. \end{aligned}$$

Step 5: Find the average number of ordered trees with left branch length.

$$\begin{aligned} \frac{[z^n]zC^3}{[z^n]C} &= \frac{[z^{n-1}]C^3}{[z^n]C} \\ &= \frac{\frac{3n}{n+2}c_n}{c_n} \\ &= \frac{3n}{n+2} \to 3 \\ &\quad \text{as} \quad n\to\infty. \end{aligned}$$

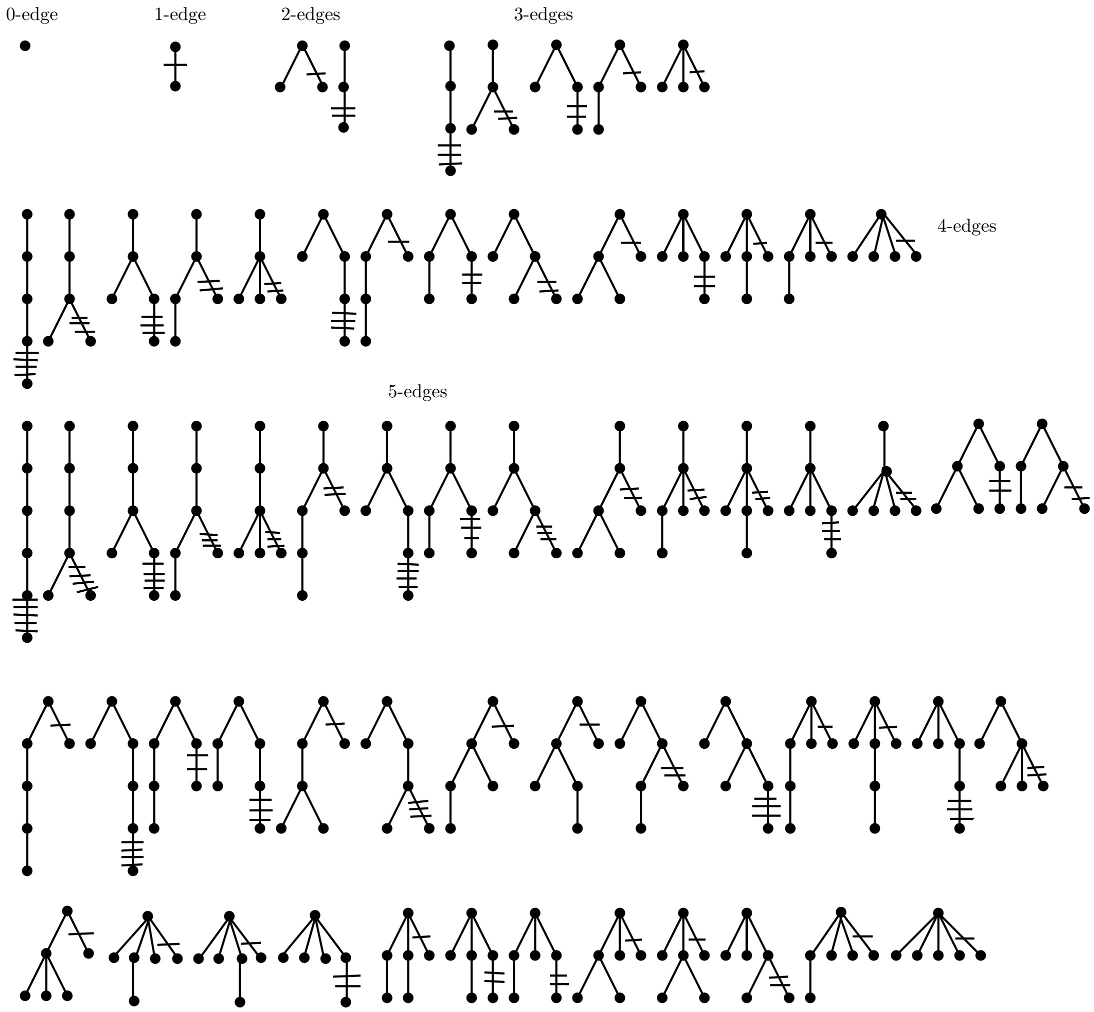


Figure 7. Ordered trees with leftmost edges crossed.

Chapter 4

The Uplift Principle

This section is studied to enable us solve a variety of combinatorial problems in two simple steps. The first is to find the appropriate generating function at the root of the tree, the second is to lift the result to an arbitrary vertex in the tree by the leaf generating function.

4.1 Definition

The *uplift principle* has two steps.
First, we find the generating function at the root, for whatever is being counted.
Then uplift the result to an arbitrary vertex by multiplying by the leaf generating function L.
It is known that the generating function for $B(z)$ counts the number of vertices in ordered trees and the number of vertices in any tree with n edges is $n+1$ so the generating function for the number of ordered trees with a **mutator** [5] is

$$B(z) = B = \sum_{n\geq 0}(n+1)\frac{1}{n+1}\binom{2n}{n}z^n = \sum_{n\geq 0}\binom{2n}{n}z^n = \sum_{n\geq 0} b_n z^n = \frac{1}{\sqrt{1-4z}}.$$

If we count **leaves** which are vertices of up degree 0 then we find the numbers $1, 1, 3, 10, 35, 126, \cdots$ which suggests the generating function $(B+1)/2$.
For classes of trees that satisfy the **UUR** and similar classes of trees e.g. ordered trees, even trees, 0.1.2 or Motzkin trees, complete or incomplete binary trees, complete or incomplete ternary trees, and Hex trees, we get the basic equation

$$V = TL, \tag{4.1}$$

where V is the generating function for trees with the mutator m, L is the generating function for trees with marked leaf and T is the generating function for the number of trees in the class.
Here is the idea of proof of (4.1). Consider a tree with mutator marked m. Snip the tree into two parts at the mutator as in figure 8. Then we can easily get $V = TL$ since the mutator just became a distinguished leaf for one tree and a distinguished root for the subtree on top.

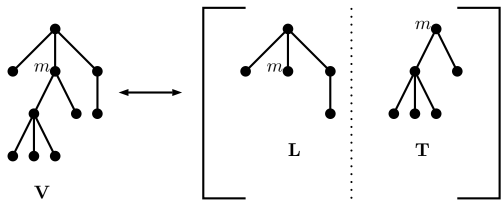


Figure 8. A tree snipped at the mutator m.

For the class of all ordered trees, $V = B$ and $T = C$ so that from (4.1) , we have $L = \frac{B}{C} = 1 + z + 3z^2 + 10z^3 + 35z^4 + \cdots$.
Now if we let $l_{n,k}$ be the number of leaves at height k in an ordered tree with n edges, we find the matrix $(l_{n,k})$ and also find the row sum to obtain

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 5 & 4 & 1 & 0 \\ 0 & 14 & 14 & 6 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 10 \\ 35 \\ \vdots \end{bmatrix} = \frac{B+1}{2}$$

This yields

$$\frac{B}{C} = \frac{B+1}{2} \tag{4.2}$$

We look at ordered trees again and with $L = B/C$ in hand, we can examine some questions at the root and then uplift to an arbitrary vertex. We will examine some examples to illustrate the uplift principle.

Example 4.1.1. *(Mutator).*[5] Let us consider trees with a mutator. For instance the figure below shows the mutators in some ordered trees.

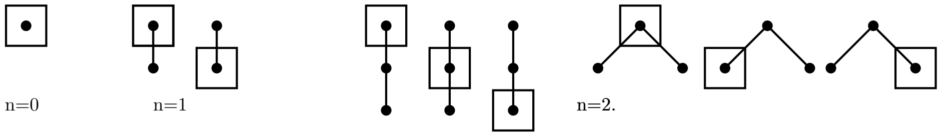


Figure 9. Ordered trees with the mutator marked with a square.

Suppose all points above the mutator including the mutator are infected. How many infected points are there? We could say these infected points are new type or polluted or unpolluted points depending on the context.
Step 1. There are $B = \sum_{n\geq 0}\binom{2n}{n}z^n = \frac{1}{\sqrt{1-4z}}$ vertices counting the root and all the vertices above the root.
Step 2. Pick a point anywhere in the tree to be the mutator and multiply by $L = B/C$. By the uplift principle the generating function is
$\frac{B}{C}\cdot B = \frac{B^2}{C} = 1 + 3z + 11z^2 + 42z^3 + \ldots$. Since $\frac{B}{C} = \frac{B+1}{2}$, we can write this as

$$B \cdot \frac{B+1}{2} = \frac{B^2+B}{2} = \frac{1}{2}\sum_{n\geq 0}\left\{4^n + \binom{2n}{n}\right\}z^n.$$

Hence the expected number of infected points is

$$\frac{[z^n](B^2+B)/2}{[z^n]B} = \frac{1/2\{4^n + \binom{2n}{n}\}}{\binom{2n}{n}}.$$

Stirling's formula gives us $4^n \sim \sqrt{\pi n}\binom{2n}{n}$ so that as n increases, the expected number of points approaches $\frac{1}{2}(\sqrt{\pi n}+1) \sim \frac{\sqrt{\pi n}}{2}$.

Example 4.1.2. *(Vertices by Updegree).*[5] Let us consider all ordered trees with n edges. What is the generating function counting all vertices of updegree k?
If $k = 0$, clearly we are counting all leaves and the leaf function is $L = B/C$. Let $k \geq 1$. Since the generating function is $(zC)^k$ when k is the updegree at the root, by the uplift principle, the generating function for all vertices with updegree k is

$$\frac{B}{C}\cdot(zC)^k = z^k BC^{k-1}.$$

Then we have

$$[z^n]z^kBC^{k-1} = [z^{n-k}]BC^{k-1} = \binom{2(n-k)+k-1}{n-k} = \binom{2n-k-1}{n-k}.$$

Example 4.1.3. *(Twigs)* [5] . We define a twig to be a vertex with 2 children

and no grandchildren. What is the generating function counting all twigs in all ordered trees with n edges ?
For step 1, we figure out the generating function at the root. The only possibility is a twig with the generating function z^2. We uplift to the general case by multiplying by the leaf function $L = \frac{B}{C}$. Thus by the uplift principle the generating function is

$$z^2 \cdot \frac{B}{C} = z^2 \cdot \frac{B+1}{2} = z^2 + z^3 + 3z^4 + 10z^5 + 35z^6 + \ldots \quad [A088218]$$

If t_n is the number of twigs from all trees with n edges, then $t_n = \frac{1}{2}\binom{2(n-2)}{n-2}$ for $n \geq 2$. We note that the average number of twigs is

$$\begin{aligned}\frac{t_n}{c_n} &= \frac{\frac{1}{2}\binom{2n-4}{n-2}}{\frac{1}{n+1}\binom{2n}{n}} = \frac{n+1}{2} \cdot \frac{n(n-1)n(n-1)}{(2n)(2n-1)(2n-2)(2n-3)} \\ &= \frac{(n+1)n(n-1)}{8(2n-1)(2n-3)} \to \frac{n}{32} \quad \text{as} \quad n \to \infty.\end{aligned}$$

Thus it takes 32 new edges to get a new twig as n gets larger.

Example 4.1.4. *(Descendants).*
We first look at *children vs grandchildren*[5]
A child at the root produces the generating function zC^3 with the z for the edge connecting the root to the child, and the three possible subtrees are to the left of this edge, to the right of this edge and, and on top of this edge. Then the uplift principle gives

$$\frac{B}{C} \cdot zC^3 = zBC^2 = \sum_{n \geq 0} \frac{n}{n+1}\binom{2n}{n} = z + 4z^2 + 15z^3 + 65z^4 + \ldots \quad [A001791]$$

By the coefficient extraction formula $[z^n]zC^3 = \frac{3n}{n+2}c_n$, so the average number of children of the root is $\frac{3n}{n+2} \to 3$ as $n \to \infty$.
But the average number of children is $\frac{n}{n+1}\binom{2n}{n} / \binom{2n}{n} = \frac{n}{n+1} \to 1$ as $n \to \infty$.
An alternate approach is to note that a tree with n edges has $n+1$ vertices and all but the root are someone's child. Thus there are $\frac{n}{n+1}$ children per vertex on average.
Next, counting grandchildren at the root leads to the generating function z^2C^5. Since

$$[z^n]z^2C^5 = \frac{5n(n-1)}{(n+2)(n+3)}C_n,$$

the average number of grandchildren of the root is $\frac{5n(n-1)}{(n+2)(n+3)} \to 5$ as $n \to \infty$. The situation changes after the uplift to an arbitrary vertex.The uplift gives us the generating function

$$\frac{B}{C} \cdot z^2C^5 = z^2BC^4 = \sum_{n\geq 2} \binom{2n}{n-2} z^n = z^2 + 6z^3 + 28z^4 + 120z^5 + \dots \quad [A002694]$$

and

$$[z^n]z^2BC^4 = \binom{2n}{n-2} = \frac{n(n-1)}{(n+1)(n+2)} \binom{2n}{n},$$

the average number of grandchildren is $\frac{n}{n+1} \cdot \frac{n-1}{n+2} \to 1$ as $n \to \infty$
Now if we consider three descendants or siblings at the root where one descendant is a child, the second a grandchild and the other a great-grandchild. Then counting great-grandchildren at the root leads to the generating function z^3C^7. Since

$$[z^n]z^3C^7 = \frac{7n(n-1)(n-2)}{(n+2)(n+3)(n+4)} C_n,$$

the average number of great-grandchildren of the root is $\frac{7n(n-1)(n-2)}{(n+2)(n+3)(n+4)} \to 7$ as $n \to \infty$.

The situation also changes after the uplift to an arbitrary vertex. The uplift gives us the generating function

$$\frac{B}{C} \cdot z^3C^7 = z^3BC^6 = \sum_{n\geq 3} \binom{2n}{n-3} z^n = z^3 + 8z^4 + 45z^5 + 220z^6 + 1001z^7 + \dots \quad [A002696]$$

and

$$[z^n]z^3BC^6 = \binom{2n}{n-3} = \frac{n(n-1)(n-2)}{(n+1)(n+2)(n+3)} \binom{2n}{n},$$

the average number of great-grandchildren is $\frac{n}{n+1} \cdot \frac{n-1}{n+2} \cdot \frac{n-2}{n+3} \to 1$ as $n \to \infty$.

Thus the root definitely has more great-grandchildren than grandchildren than children on average.
There are three surprising conclusions here, one is that, at the root there are fewer children than grandchildren than great-grandchildren, two, all uplifted limits approach 1 and thirdly, all the limits are integers.

Proposition 4.1.5. *If there are k descendants at the root then the average number of descendants at the root approaches $2k+1$ as $n \to \infty$.*

Based on the results above we can make a deduction that if there are k edges at the root then the generating function is $z^k C^{2k+1}$. Since

$$[z^n]z^k C^{2k+1} = \frac{(2k+1)n(n-1)(n-2)...(n-k+1)}{(n+k)(n+k-1)(n+k-2)...(n+2)} C_n,$$

the average number at the root is $\frac{(2k+1)n(n-1)(n-2)...(n-k+1)}{(n+k)(n+k-1)(n+k-2)...(n+2)} \to 2k+1$ as $n \to \infty$. The situation after the uplift to an arbitrary vertex gives us the generating function

$$\begin{aligned} \frac{B}{C} \cdot z^k C^{2k+1} &= z^k B C^{2k} = \sum_{n \geq k} \binom{2n}{n-k} z^n \\ &= \binom{2k}{0} z^k + \binom{2k+2}{1} z^{k+1} + \binom{2k+4}{2} z^{k+2} \\ &\quad + \binom{2k+6}{3} z^{k+3} + \cdots . \end{aligned}$$

and since

$$[z^n]z^k B C^{2k} = \binom{2n}{n-k} = \frac{n(n-1)(n-2)...(n-k+1)}{(n+1)(n+2)(n+3)...(n+k-1)} \binom{2n}{n},$$

the average number away from the root is $\frac{n}{n+1} \cdot \frac{n-1}{n+2} \cdot \frac{n-2}{n+3} \cdots \frac{n-k+1}{n+k-1} \to 1$ as $n \to \infty$.

4.2 Protected Points.

A protected point is a vertex which is neither a leaf nor distance 1 from a leaf. The root is not considered to be a leaf except in case $n = 0$ i.e the empty tree with just the root.
[4] shows that as the number of edges gets large the average proportion of protected points in all ordered trees approaches $1/6$. The tool used is generating functions. A reasonable variation occurs if we have an organisational tree such that the maximum number of employees directly under any one manager is at most two. If the updegree of any point is at most two then we are looking at *Motzkin* trees. The same tools can be used to show that the portion of protected points in Motzkin trees approaches $10/27$. To see that the average proportion of protected points in all ordered trees approach $1/6$, the generating function for the number of trees where the root is a protected point or is the empty tree was used. Trying small cases gave the numbers $1, 0, 1, 2, 6, 18, 57, \cdots$. The figure below shows the number of trees in the ordered trees with 3 edges with protected points in squares.

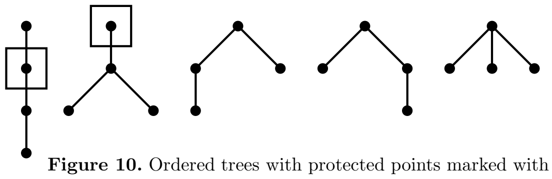


Figure 10. Ordered trees with protected points marked with squares.

Since each *subtree* out of the root must have one edge connecting to the root (the generating function for a single edge is z) and a nontrivial tree attached to this edge (with generating function $(C-1)$), each subtree at the root contributes $z(C-1) = z^2C^2$ and the total generating function is

$$1 + z^2C^2 + (z^2C^2)^2 + (z^2C^2)^3 + \cdots = \frac{1}{1-z^2C^2}.$$

$$\text{now} \quad F(z) = F = \frac{1}{1-z^2C^2} = \frac{1}{(1-zC)(1+zC)} = \frac{C}{1+zC}. \tag{4.3}$$

For asymptotic estimates the following lemma of Bender is easy to apply and very useful.

Theorem 4.2.1. ***(Bender's Lemma)*** *Suppose that $A(z) = \sum_{n\geq 0} a_n z^n$ and $B(z) = \sum_{n\geq 0} b_n z^n$ are two generating functions, and the radii of convergence of $A(z)$ and $B(z)$ are α and β respectively with $\alpha > \beta$.*
Let $C(z) = \sum_{n\geq 0} c_n z^n = A(z)B(z)$. Suppose further that b_{n-1}/b_n approaches a limit β as $n \to \infty$. If $A(\beta) \neq 0$, then $c_n \sim A(\beta)b_n$.

We have seen that ordered trees and similar classes of trees satisfy (4.1). We use the same idea of proof of (4.1) , i.e as in figure 8 in this section.

Theorem 4.2.2. *[4] The average portion of protected points in all ordered trees with n edges approaches 1/6 as $n \to \infty$.*

Does a system where each staff member can hire at most two underlings afford a higher percentage of protected points? To determine this a look at {0.1.2}-trees where the outdegree of every vertex is 0,1, or 2 is examined in [4]. The numbers of these trees are counted by the Motzkin numbers m_n with the generating function given in definition (1.2.5). It is known that the number of vertices in {0.1.2}-trees with distinguished vertex, has the generating function V given by

$$V = \sum_{n\geq 0}(n+1)M_n z^n = \frac{d}{dz}(zM)$$

since any tree with n edges has $n+1$ vertices. But then, after manipulation,

$$V = \frac{d}{dz}(zM) = \frac{1-z-\sqrt{1-2z-3z^2}}{2z^2}.\frac{1}{\sqrt{1-2z-3z^2}}$$

Since $V = TL$ we see that the number of 0.1.2-trees with the distinguished leaf has the generating function

$$L = \frac{V}{T} = \frac{1}{\sqrt{1-2z-3z^2}}.$$

We note that L has a singularity at $z = 1/3$ so the radius of convergence about $z = 0$ is also $1/3$ and the ratio test then tells us that

$$\lim_{n\to\infty} \frac{l_{n+1}}{l_n} = 3, \quad \text{where} \quad l_n = [z^n]L.$$

To get our asymptotic result we now express everything in terms of l_n. Since V may be rewritten as

$$V = \frac{1-z}{2z^2\sqrt{1-2z-3z^2}} = \frac{1}{2}.(l_{n+2} - l_{n+1})$$

$$\sim \frac{1}{2}.(9l_n - 3l_n) = 3l_n \quad \text{as} \quad n \to \infty \tag{4.4}$$

This is a result of some independent interest since it tells us that for $\{0.1.2\}$-trees as n gets large about $1/3$ of the vertices are leaves.

Theorem 4.2.3. *[4] The average portion of protected points in $\{0.1.2\}$-trees with n edges approaches 10/27 as $n \to \infty$.*

Proof. By similar argument used in [4] to proof the average proportion of protected points in all ordered trees, the number of $\{0.1.2\}$-trees where the root is the protected point or the empty tree has the generating function

$$K := 1 + z(M-1) + z^2(M-1)^2,$$

where M is the generating function for the Motzkin numbers. Thus, the generating function for the number of protected points on {0.1.2}-trees is given by

$$\begin{aligned} L(K-1) &= L(z(M-1)+z^2(M-1)^2) \\ &= L\left(zM - z + z^2M^2 - 2z^2M + z^2\right) \\ &= L\left(zM - z + (M-1-ZM) - 2z^2M + z^2\right) \\ &= L\left((1-2z^2)M + z^2 + -z - 1\right) \\ &= \frac{(1-2z^2)M}{\sqrt{1-2z-3z^2}} + \frac{z^2-z-1}{\sqrt{1-2z-3z^2}} \end{aligned} \tag{4.5}$$

But $M = \frac{1-z-\sqrt{1-2z-3z^2}}{2z^2}$ so we have

$$\begin{aligned} L(K-1) &= \frac{(1-2z^2)(1-z-\sqrt{1-2z-3z^2})}{2z^2\sqrt{1-2z-3z^2}} + \frac{z^2-z-1}{\sqrt{1-2z-3z^2}} \\ &= \frac{2z^4-4z^2-z+1}{2z^2\sqrt{1-2z-3z^2}} + 1 - \frac{1}{2z^2}. \end{aligned} \tag{4.6}$$

The first few terms of $L(K-1)$ are

$$z^2 + 3z^3 + 10z^4 + 31z^5 + 94z^6 + 281z^7 + 834z^8 + 2465z^9 + O(z^{10}).$$

Asymptotically the last two terms of the right hand side in (4.6) are irrelevant and

$$\begin{aligned} [z^n]L(K-1) &= \frac{2z^4-4z^2-z+1}{2z^2\sqrt{1-2z-3z^2}} \\ &= \frac{1}{2}(2l_{n-2} - 4l_n - l_{n+1} + l_{n+2}) \\ &\sim \frac{1}{2}(2.\frac{1}{9}l_n - 4l_n - 3l_n + 9l_n) = \frac{10}{9}l_n \quad \text{as} \quad n \to \infty. \end{aligned} \tag{4.7}$$

From (4.4) and (4.7) , we can now estimate the average number of protected points in {0.1.2}-trees as $n \to \infty$:

$$\frac{[z^n]L(K-1)}{[z^n]V} \sim \frac{\frac{10}{9}l_n}{3l_n} = \frac{10}{27} = 0.37037.$$

We note that numerically

$$\begin{aligned} \frac{[z^{100}]L(K-1)}{[z^{100}]V} &= \frac{270313833066464875926159094652788199393380l8482}{7447897271050759943050224248101637348052367 0569} \\ &= 0.36292 \end{aligned}$$

□

We use the same technique to provide a result for the hex trees.

Proposition 4.2.4. *The average portion of protected points in Hex-trees(Polyhexes rooted at an edge) approaches 76/125 as $n \to \infty$.*

Proof. By a similar argument, the number of Hex-trees where the root is the protected point or the empty tree has the generating function

$$P := 1 + 3z(H-1) + z^2(H-1)^2,$$

where H is the generating function for the Hex numbers given as

$$H(z) = \frac{1 - 3z - \sqrt{1 - 6z + 5z^2}}{2z^2}. \tag{4.8}$$

Thus, the generating function for the number of protected points on Hex-trees is given by

$$\begin{aligned} L(P-1) &= L(3z(H-1) + z^2(H-1)^2) \\ &= \frac{(1-2z^2)H}{\sqrt{1-6z+5z^2}} + \frac{z^2 - 3z - 1}{\sqrt{1-6z+5z^2}} \\ &= \frac{2z^4 - 4z^2 - 3z + 1}{2z^2\sqrt{1-6z+5z^2}} + 1 - \frac{1}{2z^2}. \end{aligned} \tag{4.9}$$

The first few terms of $L(P-1)$ are
$9z^2 + 57z^3 + 306z^4 + 1557z^5 + 7750z^6 + 31055z^7 + 162467z^8 + O(z^9)$.
Asymptotically the last two terms of the right hand side in (4.9) are irrelevant and

$$\begin{aligned} [z^n]L(P-1) &= \frac{2z^4 - 4z^2 - 3z + 1}{2z^2\sqrt{1-6z+5z^2}} \\ &= \frac{1}{2}(2l_{n-2} - 4l_n - 3l_{n+1} + l_{n+2}) \\ \sim \frac{1}{2}(2.\frac{1}{25}l_n - 4l_n - 15l_n + 25l_n) &= \frac{76}{25}l_n \quad \text{as} \quad n \to \infty. \end{aligned} \tag{4.10}$$

For the total number of points in the Hex-trees we consider (4.8). It is also known that the number of vertices in Hex-trees with distinguished vertex, has the generating function V given by

$$V = \sum_{n \geq 0} (n+1) H_n z^n = \frac{d}{dz}(zH)$$

since any tree with n edges has $n+1$ vertices. But then, after manipulation,

$$\begin{aligned}
V = \frac{d}{dz}(zH) &= \frac{d}{dz}\left(\frac{1-3z-\sqrt{1-6z+5z^2}}{2z}\right) \\
&= \frac{(2z)(1-3z-\sqrt{1-6z+5z^2})' - (1-3z-\sqrt{1-6z+5z^2})(2z)'}{(2z)^2}
\end{aligned}$$

$$\begin{aligned}
&= \frac{(2z)[-3-(\frac{1}{2})(-6+10z)](1-6z+5z^2)^{-\frac{1}{2}} - 2(1-3z-\sqrt{1-6z+5z^2})}{(2z)^2} \\
&= \frac{(2z)(-3\sqrt{1-6z+5z^2}+3-5z)}{(2z)^2\sqrt{1-6z+5z^2}} - \frac{(2-6z-2\sqrt{1-6z+5z^2})}{(2z)^2} \\
&= \frac{1-3z-\sqrt{1-6z+5z^2}}{2z^2} \cdot \frac{1}{\sqrt{1-6z+5z^2}} \qquad (4.11)
\end{aligned}$$

And from (4.1) we see that the number of Hex-trees with the distinguished leaf has the generating function

$$L = \frac{V}{T} = \frac{1}{\sqrt{1-6z+5z^2}}.$$

We note that L has a singularity at $z = 1/5$ so the radius of convergence about $z = 0$ is $1/5$ and the ratio test then tells us that

$$\lim_{n \to \infty} \frac{l_{n+1}}{l_n} = 5, \quad \text{where} \quad l_n = [z^n]L.$$

To get our asymptotic result we now express everything in terms of l_n. Since V may be rewritten as

$$\begin{aligned}
V &= \frac{1-3z}{2z^2\sqrt{1-6z+5z^2}} = \frac{1}{2}.(l_{n+2} - 3l_{n+1}) \\
&\sim \frac{1}{2}.(25l_n - 15l_n) = 5l_n \quad \text{as} \quad n \to \infty. \qquad (4.12)
\end{aligned}$$

So for Hex-trees, as n gets large about $1/5$ of the vertices are leaves. From (4.10) and (4.12), we can now estimate the average number of protected points in Hex-trees as $n \to \infty$:

$$\frac{[z^n]L(P-1)}{[z^n]V} \sim \frac{\frac{76}{25}l_n}{5l_n} = \frac{76}{125} = 0.608.$$

□

Comparing the Protected points in Ordered, Motzkin and Hex Trees

Here, we try and compare the protected points in ordered, Motzkin and the Hex trees when; leaves represent infected people, leaves represent computer hackers, and when leaves represent customers.
To do this comparison, we first look at these three cases.

Case 1: In a tree structure if leaves represent infected people, we need more of the vertices to be protected. If we focus on trying to contain the spread of a disease it is advisable to begin the process on the protected points.

Case 2: If leaves represent computer hackers , then we need more vertices to be protected as well.

Case 3: Many vertices are required to be unprotected if the leaves in an organizational tree represent customers.

For case 1 and case 2 where more vertices need to be protected, the hex tree is the best model. For case 3 where fewer vertices are required to be unprotected then the ordered tree is the best model. The motzkin tree model is the best compared to ordered and hex tree models when leaves represent customers and infected people.

4.2.5 Riordan group elements and the uplift principle.

In several combinatorial counting problems, we have seen the appearance of an element of the Riordan group. The uplift principle provides interesting Riordan matrices with combinatorial implications. This section discusses some Riordan matrices with some combinatorial implications as a result of the uplift principle.

Example (4.1.2) provides the Riordan matrix

$$\left(\frac{B}{C}, zC\right) = \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 3 & 2 & 1 & & \\ 10 & 6 & 3 & 1 & \\ 35 & 20 & 10 & 4 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

where the (n, k)-entry is the number of vertices of the ordered trees with n edges and updegree k [5]. We can use this matrix to compute many other statistics concerning ordered trees. If, for instance, we wanted to know how many vertices have at least 2 children, we could start with

$$\left(\frac{B}{C}, zC\right)(0, 0, 1, 1, \cdots)^T, \quad \text{where} \quad (0, 0, 1, 1, \cdots)^T = \frac{z^2}{1-z}.$$

Applying the **FTRA** yields

$$\begin{aligned} \left(\frac{B}{C}, zC\right) * \frac{z^2}{1-z} &= \frac{B}{C} \cdot \frac{(zC)^2}{1-zC} = z^2 B \frac{C}{1-zC} = z^2 BC^2 \\ &= z^2 + 4z^3 + 15z^4 + \cdots = \sum_{n \geq 0} \binom{2n-2}{n-2} z^n. \end{aligned}$$

The probability that a randomly chosen vertex from a random chosen ordered tree with n edges has updegree at least 2 is

$$\frac{\binom{2n-2}{n-2}}{\binom{2n}{n}} = \frac{n-1}{2(2n-1)} \to \frac{1}{4}, \quad [5]$$

as $n \to \infty$.

Also if we want to know how many vertices have at least 3 children, we go through the same steps above using $\frac{z^3}{1-z}$ instead of $\frac{z^2}{1-z}$.

The probability that a randomly chosen vertex from a randomly chosen ordered tree with n edges, has updegree at least 3, approaches $\frac{1}{8}$ as $n \to \infty$.

Proposition 4.2.6. *The probability that a randomly chosen vertex from a randomly chosen ordered tree with n edges, has updegree at least $k \geq 1$, approaches $\frac{1}{2^k}$ as $n \to \infty$.*

Here we start with

$$\left(\frac{B}{C}, zC\right)(0,0,0,\cdots,1^{kth}\cdots)^T. \quad \text{where} \quad (0,0,\cdots,1^{kth},\cdots)^T = \frac{z^k}{1-z}$$

Applying the **FTRA** yields

$$\begin{aligned}\left(\frac{B}{C}, zC\right) * \frac{z^k}{1-z} &= \frac{B}{C}\cdot\frac{(zC)^k}{1-zC} = z^k BC^{k-2}\frac{C}{1-zC} = z^k BC^k \\ &= \sum_{n\geq 0}\binom{2n-k}{n-k} z^n.\end{aligned}$$

Now the probability that a randomly chosen vertex from a random chosen ordered tree with n edges has updegree at least k is

$$\frac{\binom{2n-k}{n-k}}{\binom{2n}{n}} = \frac{n(n-1)(n-2)\cdots(n-k+1}{2n(2n-1)(2n-2)\cdots(2n-k+1)} \to \frac{1}{2^k} \quad \text{as} \quad n \to \infty\cdot$$

Chapter 5

Conclusion

5.0.1 Summary

In chapter 2, we established that the Appell subgroup is a pseudo ring and we also established advantages of the **FTRA** over the more primitive row reduction formula when we are estimating and inverting some combinatorial identities which involve infinite matrices.
We also looked at the relationship between the Riordan arrays and generating trees and developed a formula for calculating the degree of the root and degree of an arbitrary vertex in a generating tree involving $g(z)$, $f(z)$ and the A- and Z-sequences.
In chapter 3, we established that, the hitting time subgroup is isomorphic to the Bell, derivative and associated subgroups. We also estimated the average number of trees with left branch length in the motzkin and ordered trees.
In chapter 4, we presented the uplift principle and we established that, if there are k number of descendants at the root then the average number of descendants at the root approaches $2k + 1$ as $n \to \infty$. We also established that the average portion of protected points in the hex trees approaches $76/125$ as $n \to \infty$. Lastly, we established that, the probability that a randomly chosen vertex from a randomly chosen ordered tree with n edges has updegree at least k approaches $\frac{1}{2^k}$.

References

[1] Aigner, M. (1999). *Enumerative Combinatorics.* Cambridge.

[2] Cameron, N. T. (2002). *Random walks, Trees and Extensions of Riordan group tchniques.* PhD thesis, Howard University.

[3] Candice, J.-L. and Nkwanta, A. (2013). *Some Algebraic Structure of the Riordan group.* Linear Algebra and Applications 438, pages 229-239.

[4] Cheon, G.-S. and Shapiro, L. W. (2008). *Protected points in ordered trees.* Applied Mathematics Letters 21, pages 516-520.

[5] Cheon, G.-S. and Shapiro, L. W. (2012a). *The Uplift principle for ordered trees.* Applied Mathematics Letters 25, pages 1010-1015.

[6] Cheon, G-S, K. H. and Shapiro, L. W. (2012b). *Combinatorics of Riordan arrays with identical A- and Z- sequences.* Discrete Mathematics 312, pages 2040-2049.

[7] Merlini, D. and Verri, M. C. (2000a). *Generating trees and proper Riordan Arrays.* Discrete Mathematics 218 , pages 167-183.

[8] Merlini, D, S. R. and Verri, M. C. (2000b). *An algebra for proper generating trees.* Mathematics and Computer science, Birkause, pages 127-139.

[9] Peart, P. and Woan, W.-J. (2000). *A divisibility property for a subgroup of Riordan matrices.* Discrete Applied Math 98, pages 255-263.

[10] Shapiro, L. W. (2003). *Bijections and the Riordan group.* Theoretical Computer Science 307, pages 403-413.

[11] Shapiro, L. W. (2005). *Lecture notes on : A Survey of the Riordan group.* Nankai University.

[12] Shapiro, L. W. (2009). *Trees and Logs.* Congressus Numerantium. 199, pages 217-222.

[13] Shapiro L. W, Getu, S. W. W.-J. and Woodson, L. C. (1991). *The Riordan group.* Discrete Applied Mathematics 34, pages 229-239.

[14] Sloan, N. (1964). *The On-line encyclopedia of integer sequences.* Available from: http://www.oeis.org.

[15] Wilf, H. (1990). *Generating functionology.* Academic Press.