Proper generating trees
and their internal path length

Donatella Merlini
Dipartimento di Sistemi e Informatica
viale Morgagni 65, 50134, Firenze, Italia
merlini@dsi.unifi.it

Abstract

We find the generating function counting the total internal path length of any proper generating tree. This function is expressed in terms of the functions $(d(t), h(t))$ defining the associated proper Riordan array. This result is important in the theory of Riordan arrays and has several combinatorial interpretations.

Keywords: Riordan arrays; generating trees; internal path length; generating functions.

1 Introduction

The concept of a *proper Riordan array* (pRa, for short) is very useful in combinatorics. The infinite triangles of Pascal, Catalan, Motzkin and Schröder are important and meaningful examples of pRa's, and many others have been proposed and developed (see, e.g. [10, 18, 20]). In the recent literature, Riordan arrays have attracted the attention of various authors from many point of view and many examples and applications can be found (see, e.g., [1, 6, 11, 13, 14, 15, 17, 23, 24] but the literature is still developing now). In particular, Merlini and Verri [15] pointed out an important connection between pRa's and *generating trees* and they call *proper generating trees* the corresponding trees. There exists a vast literature about the concept of generating trees: it was used for the first time, without any specific name, by Chung, Graham, Hoggat and Kleiman [5] and successively by West [21, 22]. Generating trees are a device to represent the development of many classes of combinatorial objects (see, e.g., [4]) which can then be enumerated by counting the different labels in the various levels of the tree (see, e.g., [2]). The proved relation between pRa's and generating trees allows one to combine the counting capabilities of both approaches and thus improve our understanding of the problem under consideration.

In this paper we study the total internal path length of a proper generating tree and find, in particular, an explicit formula for the corresponding generating function expressed in terms of the functions $(d(t), h(t))$ defining the pRa, thus deepening and generalizing a preliminary result presented in [9] in the context of lattice paths (see, below). In this sense, Theorems 3.8 and 3.10 are the main results of the present paper: the first one solves an proves the problem in the *renewal case*, that is when $d(t) = h(t)$, while the second solves and proves the general case.

The internal path length of a proper generating tree is a quantity interesting on its own, due to its relation with the theory of Riordan arrays. In fact, the present paper can be seen as a natural continuation of the paper [15], where the connections between generating trees and Riordan arrays were originally studied.

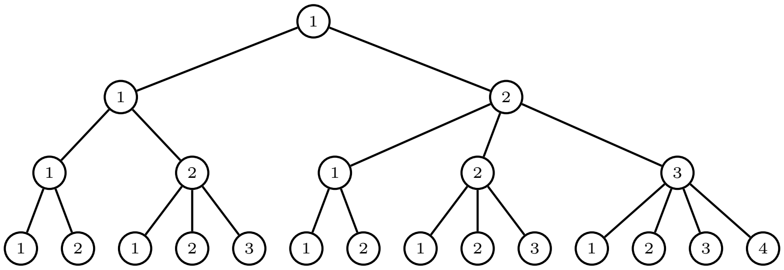

Figure 1.1: The Catalan generating tree

Moreover, the internal path length of a proper generating tree has also some nice combinatorial interpretations, some of which will be discussed in Sections 3 and 4.

In particular, the total internal path length corresponds to the total area of the histograms obtained by juxtaposing n columns having height equal to the labels found in all the paths of length n in the generating tree. For example, Figure 1.2 illustrates the histograms with $n \leq 3$ corresponding to the generating tree in Figure (1.1). This example will be examined in Section 4.1 where we will find a relation with the Catalan numbers.

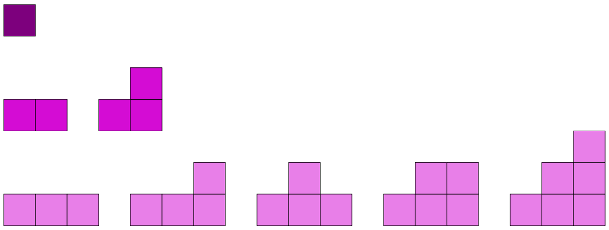

Figure 1.2: The histograms, with ($n \leq 3$), corresponding to the tree in Figure 1.1

The same quantity corresponds to the area below the lattice paths under the model of [3]. This relation has been examined in [9] and the resulting paths are called *proper paths on* $\mathbb{N}$. In this model, a path (of length n) links the points $((0,p_0),(1,p_1),\ldots,(n,p_n))$ of the lattice $\mathbb{N}\times\mathbb{N}$: if at time j the path is at altitude k, that is $p_j = k$, then the new position p_{j+1} is obtained by making a jump which belongs to a fixed set $\mathcal{P}_k$ containing the only positive jump $+1$ and some negative jumps belonging to a specified set which depends on the position k. The study of this model in terms of proper generating trees allows to find the generating function for the total area below the corresponding paths. This generating function is presented in [9, Theorem 4] without a complete proof and corresponds to Corollary 3.13 in the present paper. We refer to [9] for more details and examples concerning the interpretation of the internal path length in terms of the area below proper paths on $\mathbb{N}$.

Finally, the internal path length of proper generating trees has been used in [12] in connection with the "tennis ball problem" and in [13] to study the behavior of devices like printers under a particular combinatorial model (see Examples 4.1 and 4.3). These are examples of situations in which the generating function for the internal path length needs to be computed with the general Theorem 3.10, or, equivalently, Theorem 3.8.

Besides these combinatorial interpretations, in this paper we show how Theorems 3.10 and 3.8 can be used to find some interesting new combinatorial identities (see formulas (4.22), (4.23) and (4.24)).

The paper is organized as follows: in Section 2 we give the main properties of pRa's and their relations with generating trees. In Section 3 the formulas for the total internal path length are found and proved and Theorem 3.10 is applied to some particular situations corresponding to Riordan arrays in the Bell, associated and Appel subgroups (Theorem 3.8 and Corollaries 3.11 and 3.12). Another particular application leads to Corollary 3.13. Finally, in Section 4, we present a series of examples of generating trees already appeared in the literature and find their total internal path length by applying the results of Section 3. Moreover, we present some applications of these results in the context of combinatorial identities.

2 Background knowledge

A generating tree is a rooted labeled 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 define a generating tree it therefore suffices to specify: the label of the root and a set of rules explaining how to derive from the label of a parent node the labels of all of its children. For example, Figure 2.3 illustrates the upper part of the generating tree which corresponds to the following specification:

$$\begin{cases} root: & (1) \\ rule: & (k) \quad \rightarrow (k)(k+1) \end{cases} \tag{2.1}$$

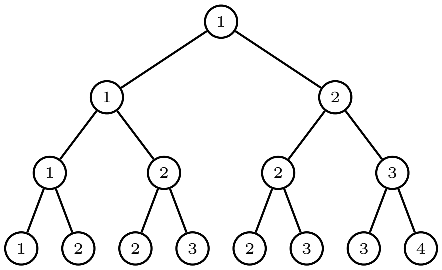

Figure 2.3: The Pascal generating tree: specification (2.1)

We can associate a matrix to any generating tree: a *matrix associated to a generating tree* (AGT matrix, for short) is an infinite matrix $(d_{n,k})_{n,k\in\mathbb{N}}$ where $d_{n,k}$ is the number of nodes at level n with label $k+c$, c being the label of the root. For example, for rule (2.1) we have the following AGT matrix:

n/k	0	1	2	3	4
0	1				
1	1	1			
2	1	2	1		
3	1	3	3	1	
4	1	4	6	4	1

where we recognize the Pascal triangle. Many AGT matrices can be studied from a Riordan array point of view. The concept of a *Riordan array* was introduced in 1991 by Shapiro, Getu, Woan and Woodson [18] (they chose this name in honour of John Riordan), with the aim of generalizing the concept of *Renewal array* defined by Rogers [16] in 1978. Their basic idea was to define a class of infinite lower triangular arrays with properties analogous to those of the Pascal triangle whose elements, as is well-known, are the binomial coefficients $\binom{n}{k}$. This concept has also been studied by Sprugnoli [20], who pointed out the relevance of these matrices in the computation of combinatorial sums. Successively, some other aspects of the theory have been studied in [10] and, as already observed in the introduction, the literature about Riordan arrays is vast and still growing.

A Riordan array is an infinite lower triangular array $(d_{n,k})_{n,k\in\mathbb{N}}$, defined by a pair of formal power series $D = (d(t), h(t))$, such that the generic element $d_{n,k}$ is the n-th coefficient in the series $d(t)(th(t))^k$, i.e.:

$$d_{n,k} = [t^n]d(t)(th(t))^k, \qquad n, k \geq 0. \tag{2.2}$$

From this definition[1] we have $d_{n,k} = 0$ for $k > n$. The bivariate generating function of a Riordan array is given by:

$$d(t, w) = \sum_{n,k\geq 0} d_{n,k} t^n w^k = \frac{d(t)}{1 - wth(t)}. \tag{2.3}$$

In the sequel we always assume that $d(0) \neq 0$; if we also have $h(0) \neq 0$ then the Riordan array is said to be *proper*; in the proper case the diagonal elements $d_{n,n}$ are different from zero for all $n \in \mathbb{N}$. The most simple example is the Pascal triangle for which we have

$$\binom{n}{k} = [t^n]\frac{1}{1-t}\left(\frac{t}{1-t}\right)^k,$$

where we recognize the pRa $d(t) = h(t) = 1/(1 - t)$ or $d(t, w) = 1/(1 - t(1 + w))$, as can be easily proved from (2.2) and (2.3).

Proper Riordan arrays are characterized by the following fundamental property found by Rogers [16] in 1978 and then examined closely by Sprugnoli [20].

Theorem 2.1 *A matrix $(d_{n,k})_{n,k\in\mathbb{N}}$ is a proper Riordan array iff there exists a sequence $A = (a_i)_{i\in\mathbb{N}}$ with $a_0 \neq 0$ s.t. every element $d_{n+1,k+1}$ can be expressed as a linear combination, with coefficients in A, of the elements in the preceding row, starting from the preceding column:*

$$d_{n+1,k+1} = a_0 d_{n,k} + a_1 d_{n,k+1} + a_2 d_{n,k+2} + \cdots$$

The previous sum is finite since $d_{n,k} = 0$ for $k > n$. The sequence A is called the *A-sequence* of the pRa and is characteristic of the matrix since it determines the function $h(t)$ and vice versa. In fact we have the following:

Corollary 2.2 *Let $D = (d(t), h(t))$ be a pRa, and let $A = (a_j)_{j\in\mathbb{N}}$ be its A-sequence. Then, if $A(t)$ is the generating function of the sequence A, we have:*

$$h(t) = A(th(t)).$$

For example, for the Pascal triangle we have: $A(t) = 1 + t$, and thus the A-sequence for this triangle is $(1, 1, 0, 0, \ldots)$; the relation of Theorem 2.1 reduces to the well-known recurrence relation for binomial coefficients:

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}.$$

More recently some new aspects of the Riordan array theory have been studied (see Merlini, Rogers, Sprugnoli and Verri [10]). The A-sequence doesn't characterize completely $(d(t), h(t))$ because $d(t)$ is independent of $A(t)$. But we have the following:

Theorem 2.3 *Let $(d_{n,k})_{n,k\in\mathbb{N}}$ be an infinite lower triangular array with $d_{n,n} \neq 0, \forall n \in \mathbb{N}$ (in particular, let it be a pRa); then there exists a unique sequence $Z = (z_0, z_1, z_2, \ldots)$ such that every element in column 0 can be expressed as a linear combination of all the elements of the preceding row:*

$$d_{n+1,0} = z_0 d_{n,0} + z_1 d_{n,1} + z_2 d_{n,2} + \cdots$$

---

[1] An equivalent definition includes the t within the function $h(t)$ and requires $h(0) = 0$.

The Z-sequence characterizes column 0 while the A-sequence characterizes all the other columns. We can conclude that the triple $(d_0, Z(t), A(t))$, with $d_0 = d(0)$, characterizes any pRa:

Theorem 2.4 *Let $(d(t), h(t))$ be a pRa and let $Z(t)$ be the generating function for the Z-sequence of the matrix. Then we have:*

$$d(t) = \frac{d_0}{1 - tZ(th(t))}.$$

The complete theory of Riordan arrays and the proofs of their properties can be found in [8, 10]. Another interesting result concerns the computation of combinatorial sums involving Riordan arrays:

Theorem 2.5 *Let $D = (d(t), h(t))$ be a Riordan array and $f(t)$ the generating function for the sequence $(f_k)_{k\in\mathbb{N}}$. Then:*

$$\sum_{k=0}^{n} d_{n,k} f_k = [t^n]d(t)f(th(t)).$$

The connection between pRa and generating trees has been studied in [15] where it is proved the following:

Theorem 2.6 *Let $c \in \mathbb{N}$, $a_j, z_j \in \mathbb{N}$, $\forall j \geq 0$, $a_0 \neq 0$ and $k \geq c$ and let*

$$\begin{cases} \textit{root}: & (c) \\ \textit{rule}: & (k) \quad \to (c)^{z_{k-c}}(c+1)^{a_{k-c}}(c+2)^{a_{k-c-1}} \cdots (k+1)^{a_0} \end{cases} \tag{2.4}$$

be a generating tree specification. Then, the AGT matrix associated to (2.4) is a proper Riordan array D defined by the triple (d_0, A, Z), such that

$$d_0 = 1, \quad A = (a_0, a_1, a_2, \ldots), \quad Z = (z_0, z_1, z_2, \ldots).$$

On the contrary, if D is a proper Riordan array defined by the triple (d_0, A, Z) with $d_0 = 1$ and $a_j, z_j \in \mathbb{N}$, $\forall j \geq 0$, then D is the AGT matrix associated to the generating tree specification (2.4).

We call *proper generating trees* the generating trees corresponding to Theorem 2.6.

3 The internal path length

In this paper we are interested in studying the total internal path length of pRa weighted with the value of each node label. More precisely, if we fix level n in the tree (being 0 the level of the root), by *total internal path length up to level n* we mean the total sum of the labels in all the paths from the root to level n (in the sequel, the term "weighted" will be understood). Referring to Figure 2.3, we have a total path length equal to 1 for paths of length 1, equal to 5 for paths of length 2, equal to 18 for paths of length 3 and equal to 56 for paths of length 4. In fact, we will prove that the generating function counting the total path length in the Pascal case is given by:

$$P(t) = \frac{1-t}{(1-2t)^3} = 1 + 5\,t + 18\,t^2 + 56\,t^3 + 160\,t^4 + 432\,t^5 + O\left(t^6\right).$$

Let us consider a proper generating tree with specification (2.4) and the corresponding pRa. The question is: how can one compute the generating function $P(t) = \sum_{n\geq 0} P_n t^n$ counting the

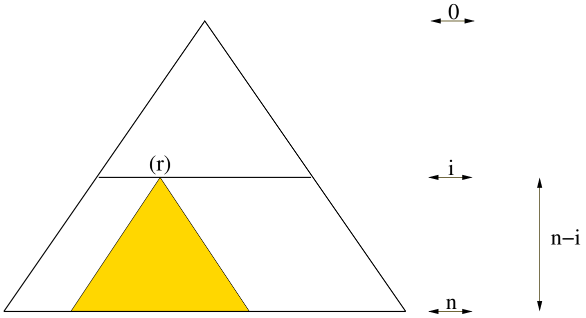

Figure 3.4: The computation of the internal path length.

internal path length up to level n in the tree? If P_n is the total sum of the labels in all the paths from level 0 to level n in a generating tree, then P_n can be seen as the sum

$$P_n = \sum_{i=0}^{n} P_{i,n} \tag{3.5}$$

where $P_{i,n}$ is the sum of the labels at level i counted with their multiplicity. Figure 3.4 illustrates how $P_{i,n}$ can be computed: if we fix level n and consider a label (r) at level i, $0 \leq i \leq n$, the multiplicity of this label is given by the number of nodes at level $n-i$ in the marked sub-tree, that is, in the generating tree having the same specification (2.4) but root labeled (r). This quantity must be multiplied by the number of nodes at level i having label (r) and obviously by the value r of the label. On the other hand, we know that the element $d_{i,r}$ of the associated pRa counts the number of nodes at level i with label $(r+c)$, being (c) the label of the root. So, if we let $f_j(t)$ be the generating function counting the number of nodes at a given level in the generating tree having root labeled (j) we have this important result:

Theorem 3.1 *The sum of the labels at level i in all the paths from levels 0 through n in a proper generating tree is:*

$$P_{i,n} = \sum_{r\geq 0} d_{i,r}(r+c)[t^{n-i}]f_{r+c}(t). \tag{3.6}$$

This theorem has been already used in previous works to study some particular situations. In [9], as already observed, it has been used for studying the area of lattice paths under a particular model. In [12], a combinatorial problem is reduced to the computation of the internal path length of a particular generating tree. However, in this case, the involved generating tree depends on a parameter s and the associated matrix is a pRa only for $s = 2$. For this value of s we find the example of Section 4.1. In [13], formulas (3.5) and (3.6) have been used in correspondence with the specification rule which will be examined in Section 4.3: in that case the internal path length was used to determine the behavior of devices like printers, under a particular combinatorial model.

In this paper, we use Theorem 3.1 for finding the generating function for the internal path length of *any* proper generating tree.

The first step in the computation of the sum (3.6) consists in the computation of the generating function $f_j(t)$.

Theorem 3.2 *Let $f_j(t)$ be the generating function counting the number of nodes at a given level in the generating tree (2.4) having root labeled (j). We have:*

$$f_c(t) = \frac{d(t)}{1-th(t)}, \quad f_{r+c}(t) = \frac{p_r(t)f_c(t) - q_r(t)}{a_0^r t^r}$$

where

$$p_0(t)=1, \quad p_r(t)=\sum_{k=0}^{r} p_{r,r-k}t^k$$

$$q_0(t)=0, \quad q_r(t)=\sum_{k=0}^{r-1} q_{r-1,r-1-k}t^k,$$

that is, $p_r(t)$ *and* $q_r(t)$ *are polynomials of degree* r *and* $r-1$*, respectively.*

Proof: The number of nodes at a given level in the generating tree (2.4) is given by $\sum_{k=0}^{n} d_{n,k}$, if $(d_{n,k})_{n,k\in\mathbb{N}}$ is the associated Riordan array. Hence, from Theorem 2.5 we have:

$$f_c(t)=\frac{d(t)}{1-th(t)}.$$

Now, the first application of rule (2.4) gives

$$(c)\to(c)^{z_0}(c+1)^{a_0}$$

from which we deduce the following equation:

$$f_c(t)=1+t(z_0f_c(t)+a_0f_{c+1}),$$

hence:

$$f_{c+1}(t)=\frac{(1-z_0t)f_c(t)-1}{a_0t}.$$

In a similar way we find:

$$f_{c+1}(t)=1+t(z_1f_c(t)+a_1f_{c+1}+a_0f_{c+2}(t))$$

which corresponds to the rule

$$(c+1)\to(c)^{z_1}(c+1)^{a_1}(c+2)^{a_2}$$

and gives:

$$f_{c+2}(t)=\frac{(1-(z_0+a_1)t+(z_0a_1-z_1a_0)t^2)f_c(t)-1+(a_0-a_1)t}{a_0^2t^2}$$

In general we have:

$$f_{c+r}(t)=1+t(z_rf_c(t)+a_rf_{c+1}+a_{r-1}f_{c+2}(t)+\ldots+a_0f_{c+r+1}(t)) \tag{3.7}$$

and

$$f_{c+r}(t)=\frac{p_r(t)f_c(t)-q_r(t)}{a_0^rt^r} \tag{3.8}$$

where $p_r(t)$ is a polynomial of degree r and $q_r(t)$ is a polynomial of degree $r-1$. The problem is determining the nature of polynomials p and q. ∎

Theorem 3.3 *The matrices* $P=(p_{r,k})_{r,k\in\mathbb{N}}$ *and* $Q=(q_{r,k})_{r,k\in\mathbb{N}}$ *defining the coefficients of polynomials* $p_r(t)$ *and* $q_r(t)$ *in Theorem 3.2 correspond to the following pRa:*

$$P=(d_P(t),h_P(t))=\left(1-\frac{a_0tZ(a_0t)}{A(a_0t)},\frac{a_0}{A(a_0t)}\right),$$

$$Q=(d_Q(t),h_Q(t))=\left(\frac{a_0}{(1-a_0t)A(a_0t)},\frac{a_0}{A(a_0t)}\right).$$

Proof: By solving equation (3.7) in terms of $f_{c+r+1}(t)$ and by using equation (3.8) we find:

$$p_{r+1}(t) = p_r(t) - z_r a_0^r t^{r+1} - a_0^r t^{r+1} \sum_{j=1}^{r} a_{r-j+1} \frac{p_j(t)}{(a_0 t)^j}$$

and

$$q_{r+1}(t) = q_r(t) + a_0^r t^{r+1} - a_0^r t^{r+1} \sum_{j=1}^{r} a_{r-j+1} \frac{q_j(t)}{(a_0 t)^j}.$$

If we multiply both equations by w^r and sum over r we find:

$$\frac{\bar{P}(t,w) - 1}{w} = \bar{P}(t,w) - tZ(a_0tw) - \frac{(A(a_0tw) - a_0)\bar{P}(t,w)}{a_0 w} + \frac{A(a_0tw) - a_0}{a_0 w}$$

and

$$\frac{\bar{Q}(t,w)}{w} = \bar{Q}(t,w) + \frac{1}{1 - a_0tw} - \frac{(A(a_0tw) - a_0)\bar{Q}(t,w)}{a_0 w}$$

with $\bar{P}(t,w) = \sum_{r\geq 0} p_r(t)w^r = \sum_{r,k\geq 0} p_{r,k}t^{r-k}w^r$ and $\bar{Q}(t,w) = \sum_{r\geq 0} q_r(t)w^r = w\sum_{r,k\geq 0} q_{r,k}t^{r-k}w^r$. By solving the two equations, we have:

$$\bar{P}(t,w) = \frac{A(a_0tw) - a_0twZ(a_0tw)}{A(a_0tw) - a_0w} = \frac{1 - a_0twZ(a_0tw)/A(a_0tw)}{1 - a_0w/A(a_0tw)}$$

and

$$\bar{Q}(t,w) = \frac{a_0 w}{(a_0w - A(a_0tw))(a_0tw - 1)} = \frac{a_0 w}{(1 - a_0tw)A(a_0tw)\,(1 - a_0w/A(a_0tw))}.$$

Now, if $P(t,w) = \sum_{r,k\geq 0} p_{r,k}t^r w^k$ and $Q(t,w) = \sum_{r,k\geq 0} q_{r,k}t^r w^k$ we have $\bar{P}(t,w) = P(tw, 1/t)$ and $\bar{Q}(t,w) = wQ(tw, 1/t)$, from which we deduce:

$$P(t,w) = \frac{d_P(t)}{1 - twh_P(t)}, \quad Q(t,w) = \frac{d_Q(t)}{1 - twh_Q(t)}.$$

From (2.3) we have immediately the statement of the theorem. ∎

Another result involving functions $f_j(t)$ is the following:

Theorem 3.4 *Let* $F(t,w) = \sum_{r\geq 0} f_{r+c}(t)w^r$; *then we have:*

$$F(t,w) = \frac{td(t)(1-w)(A(w) - wZ(w)) + twh(t) - w}{(tA(w) - w)(1 - th(t))(1-w)}.$$

Proof: We have:

$$F(t,w) = f_c(t)\sum_{r\geq 0} p_r(t)\left(\frac{w}{a_0 t}\right)^r - \sum_{r\geq 0} q_r(t)\left(\frac{w}{a_0 t}\right)^r =$$

$$= f_c(t)\bar{P}\left(t, \frac{w}{a_0 t}\right) - \bar{Q}\left(t, \frac{w}{a_0 t}\right),$$

where $\bar{P}(t,w)$, and $\bar{Q}(t,w)$ are the functions defined in Theorem 3.3; after some simplifying we get the expression in the statement of the theorem. ∎

Since the computation of $P_{i,n}$ and P_n is quite complex we proceed by first solving a particular case, which however covers a lot of significant combinatorial cases, and then go on with the general one.

3.1 The renewal case

When $d(t) = h(t)$ in a Riordan array $D = (d(t), h(t))$, the corresponding matrix is known as a *renewal array* and things simplify a bit. Many such matrices are well known and arise in various combinatorial problems as we will see in Section 4.

Theorem 3.5 *Let $d_0 = h_0 \neq 0$. Then $d(t) = h(t)$ iff $A(t) = d_0 + tZ(t)$.*

In particular, when D is associated to the generating tree (2.4) we have $d_0 = 1$, hence $a_0 = 1$; moreover we have $z_{k-c} = a_{k-c+1}$. This means that in the renewal case we are dealing with the following specification rule:

$$\begin{cases} root: & (c) \\ rule: & (k) \quad \rightarrow (c)^{a_{k-c+1}}(c+1)^{a_{k-c}}(c+2)^{a_{k-c-1}} \cdots (k)^{a_1}(k+1) \end{cases} \tag{3.9}$$

and with an associated matrix defined by the renewal array $D = (d(t), d(t))$, with $d_0 = 1$ and $d(t) = A(td(t))$.

In this case Theorems 3.2 and 3.3 reduce to the following:

Theorem 3.6 *Let $f_j(t)$ be the generating function counting the number of nodes at a given level in the generating tree (3.9) having root labeled (j). We have:*

$$f_c(t) = \frac{d(t)}{1 - td(t)}, \qquad f_{r+c}(t) = \frac{1}{t^r}(p_r(t)f_c(t) - q_r(t))$$

where

$$p_0(t) = 1, \quad p_r(t) = \sum_{k=0}^{r} p_{r,r-k}t^k$$

$$q_0(t) = 0, \quad q_r(t) = \sum_{k=0}^{r-1} p_{r-1,r-1-k}t^k;$$

the matrices $P = (p_{r,k})_{r,k\in\mathbb{N}}$ and $Q = (q_{r,k})_{r,k\in\mathbb{N}}$ correspond to the following pRa:

$$P = \left(\frac{1}{A(t)}, \frac{1}{A(t)}\right), \qquad Q = \left(\frac{1}{(1-t)A(t)}, \frac{1}{A(t)}\right).$$

In the renewal case $A(w) = 1 + wZ(w)$, hence the function $F(t, w)$ defined in Theorem 3.4 reduces to:

$$F(t, w) = \sum_{r\geq 0} f_{r+c}(t)w^r = \frac{td(t) - w}{(tA(w) - w)(1 - td(t))(1 - w)}. \tag{3.10}$$

By differentiating $F(t, w)$ with respect to w we obtain:

$$G(t, w) = \sum_{r\geq 0} rf_{r+c}(t)w^r = \frac{twA'(w)(td(t) - w) + w(tA(w)(td(t) - 1) + td(t) - 2twd(t) + w^2)}{(tA(w) - w)^2(1 - td(t))(1 - w)^2}. \tag{3.11}$$

The previous functions can be used in the following:

Theorem 3.7 *Let $F(t, w)$ and $G(t, w)$ be the functions defined in formulas 3.10 and 3.11. Then:*

$$\sum_{r\geq 0} d_{i,r}(r + c)f_{r+c}(t) = [w^i]\left(d(w)G(t, wd(w)) + cd(w)F(t, wd(w))\right).$$

Proof: We have:

$$\sum_{r\geq 0} d_{i,r}(r+c)f_{r+c}(t) = \sum_{r\geq 0} d_{i,r} r f_{r+c}(t) + c\sum_{r\geq 0} d_{i,r} f_{r+c}(t) =$$

$$= \sum_{r\geq 0} d_{i,r}[w^r]G(t,w) + c\sum_{r\geq 0} d_{i,r}[w^r]F(t,w).$$

We have two combinatorial sums involving Riordan arrays and Theorem 2.5 yields the proof. ■

We can finally state the following fundamental theorem:

Theorem 3.8 *Let* $P(t)=\sum_{n\geq 0} P_n t^n$ *be the generating function counting the total internal path length of any generating tree corresponding to the specification rule (3.9). Then*

$$P(t) = \frac{t^2 d(t)(1-td(t))d''(t) - 2t^2(1-c+(c-2)td(t))d'(t)^2}{2(d(t)+td'(t))(1-td(t))^3} +$$

$$+\frac{4td(t)(c+(1-c)td(t))d'(t) + 2(1-c)td(t)^3 + 2cd(t)^2}{2(d(t)+td'(t))(1-td(t))^3}$$

Proof: From Theorem 3.7 we have

$$P_n = \sum_{i=0}^{n} P_{i,n} = \sum_{i=0}^{n} [t^{n-i}][w^i]\left(d(w)G(t,wd(w)) + cd(w)F(t,wd(w))\right). \tag{3.12}$$

From the relation $d(w)=A(wd(w))$ we obtain:

$$A'(wd(w)) = \frac{d'(w)}{d(w)+wd'(w)}$$

and after some computation we have:

$$F(t,wd(w)) = \frac{td(t)-wd(w)}{d(w)(t-w)(1-td(t))(1-wd(w))},$$

and

$$G(t,wd(w)) = \frac{w(t-w)(td(t)-2d(t)twd(w)+w^2d(w)^2)d'(w)}{d(w)(d(w)+wd'(w))(t-w)^2(1-td(t))(1-wd(w))^2} +$$

$$-\frac{wd(w)(td(t)-2twd(t)d(w)+w^2d(w)^2-td(w)+t^2d(w)d(t))}{d(w)(d(w)+wd'(w))(t-w)^2(1-td(t))(1-wd(w))^2}.$$

The second sum in formula (3.12) corresponds to a convolution, so, in order to find $P(t)$, we have to put $t=w$ in $H(t,w)=d(w)G(t,wd(w))+cd(w)F(t,wd(w))$. Unfortunately there is a $(t-w)^2$ factor at the denominator but this factor can be eliminated by taking the numerator $N_H(t,w)$ of $H(t,w)$ and developing it into a series at $w=t$:

$$N_H(t,w) = N_H(t,t) + \frac{\partial}{\partial w}N_H(t,w)\bigg|_{w=t}(w-t) + \frac{1}{2}\frac{\partial^2}{\partial w^2}N_H(t,w)\bigg|_{w=t}(w-t)^2 + O\left((w-t)^3\right).$$

By using Maple one can easily find that both $N_H(t,t)$ and $\frac{\partial}{\partial w}N_H(t,w)\big|_{w=t}$ are equal to zero and the other term yields to the desired expression for $P(t)$. ■

In the Pascal case $d(t) = 1/(1-t)$, Theorem 3.8 gives, as expected:

$$P_n = [t^n]\frac{1-t}{(1-2t)^3} = (n+1)(n+4)2^{n-2}$$

if the root has label (1) (see Figure 2.3). This sequence corresponds to the coefficients sequence of Chebyshev polynomials and also counts the number of 132-avoiding permutations of $[n+4]$ containing exactly the 123 patterns (see sequence A001793 in [19]). If the root has label (c), as in Figure 3.5, we have:

$$P_n = [t^n]\frac{c-(2c-1)t}{(1-2t)^3} = c + (4\,c+1)\,t + (12\,c+6)\,t^2 + (32\,c+24)\,t^3 + O\left(t^4\right).$$

Some other examples corresponding to the renewal case will be examined in Section 4. In all of them, the use of a symbolic computation system like Maple allows to obtain a simplified, nice, expression for $P(t)$.

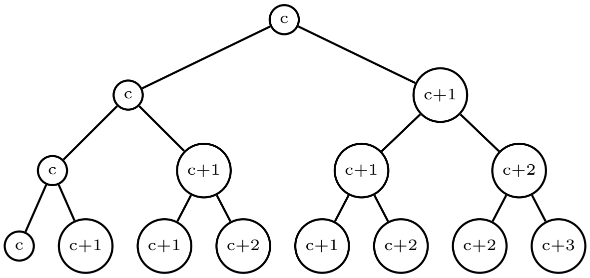

Figure 3.5: The Pascal generating tree with root (c)

3.2 The general case

When a generating tree corresponds to specification (2.4) then the computation of its internal path length leads to complicated expressions. However, the various steps of computation are analogous to those done in the renewal case and we only give the final results. In particular, the sum $\sum_{r\geq 0} d_{i,r}(r+c)f_{r+c}(t)$ can be computed again by extracting the $[w^i]$ coefficient from $H(t,w) = d(w)G(t,wh(w)) + cd(w)F(t,wh(w))$. We don't write the expressions for $F(t,w)$ and $G(t,w)$ but rather those for $F(t,wh(w))$ and $G(t,wh(w))$ which take quite simple forms by using the relations $h(w) = A(wh(w))$, $d(w) = 1/(1-wZ(wh(w)))$ and

$$A'(wh(w)) = \frac{h'(w)}{h(w)+wh'(w)}, \quad Z'(wh(w)) = \frac{wd'(w)+d(w)-d(w)^2}{w^2d(w)^2(h(w)+wh'(w))}.$$

We give this result as a corollary:

Corollary 3.9 *Let* $F(t,w) = \sum_r f_{r+c}(t)w^r$ *and* $G(t,w) = w\frac{\partial}{\partial w}F(t,w)$; *then we have:*

$$F(t,wh(w)) = \frac{td(t) - wd(w) - tw(d(t)h(w) - d(w)h(t))}{d(w)(t-w)(1-th(t))(1-wh(w))}$$

and

$$G(t,wh(w)) = \frac{N_G(t,w)}{D_G(t,w)}$$

where

$$N_G(t,w) = wh(w)\left((t-w)(td(t)(1-wh(w))^2 d'(w) - w^2 d(w)^2(1-th(t))h'(w)) + \right.$$

$$\left. + d(w)^2(t^2h(t) + w^2h(w) - tw^2h(t) - t) + d(t)d(w)(tw^2h(w)^2 - 2twh(w) + t)\right),$$

$$D_G(t,w) = d(w)^2(h(w) + wh'(w))(t-w)^2(1-th(t))(1-wh(w))^2$$

The final result is obtained by putting $t = w$ in $H(t,w)$ and proceeding as in Theorem 3.10.

Theorem 3.10 *Let $P(t) = \sum_{n\geq 0} P_n t^n$ be the generating function counting the total internal path length of any generating tree corresponding to the specification rule (2.4). Then we have:*

$$P(t) = \frac{N_P(t)}{D_P(t)}$$

where

$$N_P(t) = \frac{1}{2}t^2 d(t)h(t)d''(t)(1-th(t))^2 + \frac{1}{2}t^3 d(t)^2 h(t)h''(t)(1-th(t)) +$$

$$+ t^3 d(t)^2 h'(t)^2 (c + (1-c)th(t)) - t^2 h(t)d'(t)^2(1-th(t))^2 +$$

$$+ ctd(t)h(t)d'(t)(1-th(t))^2 + d(t)^2 h(t)(c + (1-c)th(t)) +$$

$$+ td(t)h'(t)((ct - 2ct^2h(t) + ct^3h(t)^2)d'(t) + (c + 2th(t) - ct^2h(t)^2)d(t)),$$

$$D_P(t) = d(t)(h(t) + th'(t))(1-th(t))^3.$$

3.3 Some particular cases

In Section 3.1 we have examined the case of renewal arrays, that is, Riordan arrays having $d(t) = h(t)$, and we have observed that many such matrices are well known and arise in various combinatorial contexts. It is well known that Riordan arrays constitute a group with respect to the usual row by column product between matrices and that renewal arrays constitute a subgroup named the *Bell subgroup*. There are other two particular cases which occur frequently in practice, $d(t) = 1$ and $h(t) = 1$: the case $d(t) = 1$ corresponds to the *associated subgroup* while the case $h(t) = 1$ to the *Appel subgroup* (see, e.g., [18, 17]).

Since the generating function in Theorem 3.10 is quite complex and due to the importance of the above particular cases, in the following corollaries we give the generating functions for the internal path length of generating trees corresponding to matrices in the associated and Appel subgroups.

In the first case, from $d(t) = 1$ we obtain $Z(t) = 0$ and the following specification rule:

$$\begin{cases} root: & (c) \\ rule: & (k) \quad \rightarrow (c+1)^{a_{k-c}}(c+2)^{a_{k-c-1}} \cdots (k)^{a_1}(k+1)^{a_0} \end{cases} \tag{3.13}$$

Therefore, we have the following:

Corollary 3.11 *Let $P(t) = \sum_{n\geq 0} P_n t^n$ be the generating function counting the total internal path length of any generating tree corresponding to the specification rule (3.13). Then we have:*

$$P(t) = \frac{t^2h'(t) + 1}{(1-th(t))^2}c + \frac{th(t)\left(t^2h''(t)(1-th(t)) + 2th'(t)(t^2h'(t) + 2) + 2h(t)\right)}{2(1-th(t))^3(h(t) + th'(t))}$$

When $h(t) = 1$, we have $A(t) = 1$ and the following specification rule:

$$\begin{cases} \mathit{root}: & (c) \\ \mathit{rule}: & (k) \quad \rightarrow (c+1)^{z_{k-c}}(k+1) \end{cases} \tag{3.14}$$

In this case, Theorem 3.10 becomes:

Corollary 3.12 *Let $P(t) = \sum_{n\geq 0} P_n t^n$ be the generating function counting the total internal path length of any generating tree corresponding to the specification rule (3.14). Then we have:*

$$P(t) = \frac{d(t) + t(1-t)d'(t)}{(1-t)^2}c + \frac{t(td(t)d''(t)(1-t)^2 - 2td'(t)(1-t)^2 + 2d(t)^2)}{2(1-t)^3 d(t)}$$

Another case which has an interesting combinatorial interpretation is the one examined in [9]. In the model of lattice paths studied in that paper, the label of a node in the corresponding generating trees correspond to the altitude of a path after a series of steps. Since these paths start from the origin of the lattice the label of the root needs to be zero, that is, $c = 0$ in rule (2.4). Therefore, we have the following specification rule:

$$\begin{cases} \mathit{root}: & (0) \\ \mathit{rule}: & (k) \quad \rightarrow (0)^{z_k}(1)^{a_k}(2)^{a_{k-1}}\cdots(k+1)^{a_0} \end{cases} \tag{3.15}$$

and the following corollary, which corresponds to [9, Theorem 4]:

Corollary 3.13 *Let $P(t) = \sum_{n\geq 0} P_n t^n$ be the generating function counting the total internal path length of any generating tree corresponding to the specification rule (3.15). Then we have:*

$$P(t) = \frac{N_P(t)}{D_P(t)}$$

where

$$N_P(t) = \frac{1}{2}t^2 d(t)h(t)d''(t)(1-th(t))^2 + \frac{1}{2}t^3 d(t)^2 h(t)h''(t)(1-th(t)) +$$
$$+t^4 d(t)^2 h(t)h'(t)^2 - t^2 h(t)d'(t)^2(1-th(t))^2 + td(t)^2 h(t)^2 + 2t^2 d(t)^2 h(t)h'(t),$$
$$D_P(t) = d(t)(h(t) + th'(t))(1-th(t))^3.$$

We wish to point out that when proper generating trees are used as a device to represent the development of classes of combinatorial objects, as in the ECO method [4], the label of a node represents the number of sons of the node, or equivalently, the number of combinatorial objects generated from that node. Therefore, in these cases we have always $c \neq 0$.

In the next section we will illustrate some applications of Theorem 3.8 and Corollaries 3.11 and 3.12. The use of Maple in this cases becomes essential. Other examples concerning Corollary 3.13 can be found in [9].

Finally, a nice consequence of Theorems 3.8 and 3.10 which we wish to point out is given by the following corollary:

Corollary 3.14 *Let $P(t) = \sum_{n\geq 0} P_n t^n$ be the generating function counting the total internal path length of any generating tree corresponding to the specification rule (2.4). Then $P(t)$ is algebraic as soon as $d(t)$ and $h(t)$ are algebraic.*

4 Examples

We now take into consideration some well known generating trees and Riordan arrays (see, e.g., [2, 4, 5, 15, 21, 22]) and find their internal path length by using the results of the previous sections. Moreover, in Section 4.6, we present an application of the same results for determining some interesting combinatorial identities.

4.1 The Catalan case

The first example is related to the Catalan numbers $C_n = \frac{1}{n+1}\binom{2n}{n}$ which count the number of nodes at each level in the generating tree of Figure 1.1. The specification rule is:

$$\begin{cases} root: & (1) \\ rule: & (k) \quad \rightarrow (1)\cdots(k)(k+1) \end{cases} \tag{4.16}$$

and the associated matrix begins:

n/k	0	1	2	3	4
0	1				
1	1	1			
2	2	2	1		
3	5	5	3	1	
4	14	14	9	4	1

By using the results of Section 2, this matrix corresponds to the renewal array defined by

$$A(t) = Z(t) = \frac{1}{1-t},$$

or, equivalently, by

$$d(t) = h(t) = \frac{1-\sqrt{1-4t}}{2t}.$$

The formula in Theorem 3.8 simplifies a lot and gives:

$$P(t) = \frac{1-\sqrt{1-4t}}{2t(1-4t)} = 1 + 5\,t + 22\,t^2 + 93\,t^3 + 386\,t^4 + O\left(t^5\right).$$

By extracting the n^{th} coefficient from $P(t)$ we have:

$$P_n = [t^n]P(t) = \frac{1}{2}[t^{n+1}]\frac{1}{1-4t} - \frac{1}{2}[t^{n+1}]\frac{1}{\sqrt{1-4t}} =$$

$$= \frac{1}{2}4^{n+1} - \frac{1}{2}\binom{2(n+1)}{n+1} \approx \frac{1}{2}4^{n+1}\left(1 - \frac{1}{\sqrt{\pi(n+1)}}\right),$$

where we used the approximation $\binom{2n}{n} \approx 4^n/\sqrt{\pi n}$. By Theorem 3.2, the number S_n of nodes at level n is given by:

$$S_n = [t^n]\frac{d(t)}{1-td(t)} = [t^n]\frac{d(t)-1}{t} = C_{n+1},$$

hence the average internal path length up to level n behaves as follows:

$$\frac{P_n}{S_n} \approx \frac{1}{2}\sqrt{\pi}n^{3/2}.$$

Sequence P_n corresponds to sequence A000346 in [19] and can also be found in [12] in relation with "the tennis ball problem".

4.2 The Motzkin case

This example is related to Motzkin numbers $M_n = [t^n](1-t-\sqrt{1-2t-3t^2})/(2t^2)$ which count the number of nodes at each level in the following generating tree:

$$\begin{cases} root: & (1) \\ rule: & (k) \quad \rightarrow (1)\dots(k-1)(k+1) \end{cases} \tag{4.17}$$

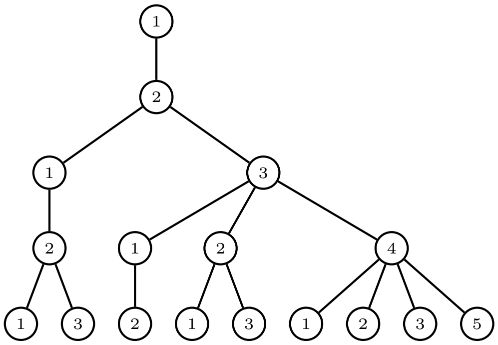

Figure 4.6: The Motzkin generating tree: specification (4.17)

The associated matrix is:

n/k	0	1	2	3	4
0	1				
1	0	1			
2	1	0	1		
3	1	2	0	1	
4	3	2	3	0	1

and corresponds to the renewal array:

$$d(t) = h(t) = \frac{1+t-\sqrt{1-2t-3t^2}}{2t(1+t)}, \qquad A(t) = \frac{1-t+t^2}{1-t}, \qquad Z(t) = \frac{t}{1-t}$$

The generating function for the total internal path length can be determined by applying Theorem 3.8:

$$P(t) = 2\frac{(1+t-\sqrt{1-2t-3t^2})(2t^2+t-1-\sqrt{1-2t-3t^2})}{t(3t-1)(1+t+\sqrt{1-2t-3t^2})^3} =$$

$$= 1 + 3\,t + 10\,t^2 + 31\,t^3 + 96\,t^4 + O\left(t^5\right).$$

The asymptotic approximation of P_n can be found by performing a series development of $P(t)$ around its dominating singularity $t = 1/3$. If we put $t = (1-w)/3$ in $P(t)$, so that $w = 1-3t$, and then compute the series development around $w = 0$ we get

$$P(t) = \frac{3}{2}\frac{1}{w} - \frac{3\sqrt{3}}{4}\frac{1}{w^{1/2}} + O(1),$$

hence

$$P_n \approx \frac{1}{2}3^{n+1}\left(1 - \frac{\sqrt{3}}{2}\frac{1}{\sqrt{\pi n}}\right).$$

The number S_n of nodes at level n, by Theorem 3.2, is equal to the n^{th} Motzkin number M_n and, by the same arguments used to find the approximation for P_n, we have $M_n \approx \frac{3^{n+1}}{(2\,n+3)}\sqrt{\frac{3}{\pi\,(n+2)}}$. Therefore, the average internal path length up to level n satisfies:

$$\frac{P_n}{S_n} \approx \frac{1}{3}\sqrt{3}\sqrt{\pi}n^{3/2}.$$

Sequence P_n corresponds to sequence A055217 in [19] and counts the maximal number of different sequences that can be obtained from a ternary sequence of length $2n+1$ by deleting n symbols. For example, we have 27 ternary sequences of length 3 and the maximal number of different sequences one can obtain by deleting 1 symbol is 3, similarly, we have 243 ternary sequences of length 5 and by deleting any two symbols in all of them we obtain at most 10 different sequences, so 10 is the required number. Note, the case $n = 3$ gives 31 which is different from $\binom{7}{3}$.

The next two examples are related to Schröder numbers $[t^n](1 + t - \sqrt{1 - 6t + t^2})/(4t)$. Again, the relation with these special numbers regards the enumeration of nodes at each level in the generating trees (4.18) and (4.19). In particular, the number of nodes at level n in the generating tree (4.19) is twice the same quantity in (4.18).

4.3 The "printer" case

The following generating tree specification has been introduced in [13] in connection with the study of the behaviour of devices like printers, under a particular combinatorial model introduced in the same paper. The rule is:

$$\begin{cases} \mathit{root}: & (1) \\ \mathit{rule}: & (k) \quad \rightarrow (1)^2(2)^2 \ldots (k)^2(k+1) \end{cases} \tag{4.18}$$

and the first rows of the associated matrix are in this case:

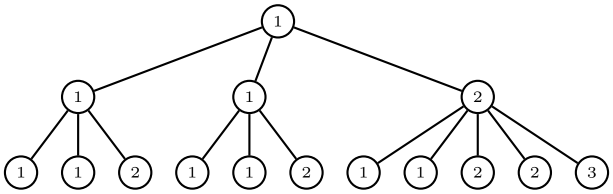

Figure 4.7: The generating tree (4.18)

n/k	0	1	2	3	4
0	1				
1	2	1			
2	6	4	1		
3	22	16	6	1	
4	90	68	30	8	1

This matrix corresponds to the renewal array:

$$d(t) = h(t) = \frac{1 - t - \sqrt{1 - 6t + t^2}}{2t}, \quad A(t) = \frac{1+t}{1-t}, \quad Z(t) = \frac{2}{1-t}$$

and the generating function for the total internal path length is given by:

$$P(t)=\frac{1+t-\sqrt{1-6t+t^2}}{4t(1-6t+t^2)}=1+7\,t+44\,t^2+268\,t^3+1609\,t^4+9583\,t^5+O\left(t^6\right).$$

The series devolopment of $P(t)$ around its dominating singularity $t=3-2\sqrt{2}$ gives:

$$P_n\approx\frac{1}{8}\left(7+5\sqrt{2}\right)\left(3+2\sqrt{2}\right)^n\left(1-\frac{\left(3+2\sqrt{2}\right)^2}{\sqrt{4+3\sqrt{2}}\left(7+5\sqrt{2}\right)\sqrt{\pi\,n}}\right).$$

For the number S_n of nodes at level n we have $S_n\approx\frac{\sqrt{4+3\sqrt{2}}\left(3+2\sqrt{2}\right)^n}{2(2\,n+1)\sqrt{\pi\,(n+1)}}$ hence:

$$\frac{P_n}{S_n}\approx\frac{\sqrt{\pi}\left(7+5\sqrt{2}\right)n^{3/2}}{2\sqrt{4+3\sqrt{2}}}$$

4.4 The Schröder case

In this section we examine the generating tree having the following specification:

$$\begin{cases} \mathit{root}: & (2) \\ \mathit{rule}: & (k) \quad\rightarrow (3)(4)\dots(k)(k+1)^2 \end{cases} \tag{4.19}$$

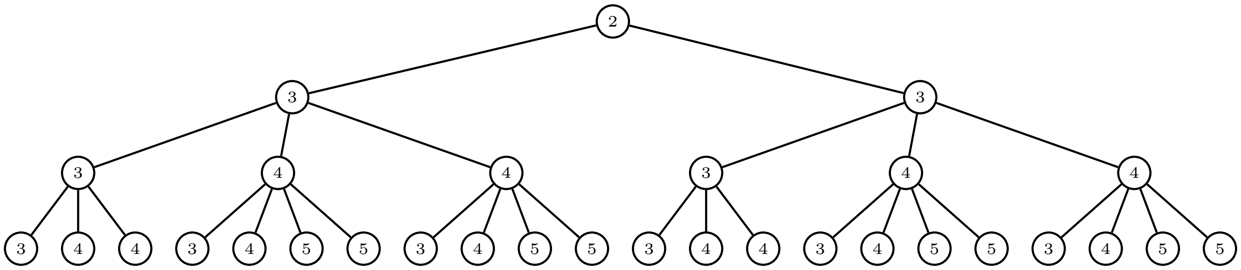

Figure 4.8: The Schröder generating tree: specification (4.19).

The associated matrix in this case is a pRa with $d(t)\neq h(t)$:

n/k	0	1	2	3	4
0	1				
1	0	2			
2	0	2	4		
3	0	6	8	8	
4	0	22	28	24	16

$$d(t)=1,\quad h(t)=\frac{1+t-\sqrt{1-6t+t^2}}{2t},\quad A(t)=\frac{2-t}{1-t},\quad Z(t)=0$$

and the function $P(t)$ can be found by applying Corollary 3.11 and simplifying:

$$P(t)=8\frac{(1-4t+t^2)\sqrt{1-6t+t^2}+1-7t+9t^2-t^3}{(1-6t+t^2)(1-t+\sqrt{1-6t+t^2})^3}=$$

$$=2+10\,t+52\,t^2+282\,t^3+1564\,t^4+8786\,t^5+O\left(t^6\right).$$

Finally, the n^{th} coefficient can be approximated as follows:

$$P_n \approx \frac{1}{2}\frac{\left(3+2\sqrt{2}\right)^3\left(3+2\sqrt{2}\right)^n}{\left(\sqrt{2}+1\right)^3\left(4+3\sqrt{2}\right)}\left(1-\frac{\sqrt{4+3\sqrt{2}}\,\left(1-2\sqrt{2}\right)}{\left(\sqrt{2}+1\right)\sqrt{\pi\,n}}\right).$$

Note, in this example and in the following one, the label of the root is $c = 2$.

4.5 The Fibonacci case

The odd Fibonacci numbers $F_{2n+1} = [t^{2n+1}]t/(1-t-t^2)$ count the number of nodes in the generating tree having specification:

$$\begin{cases} root: & (2) \\ rule: & (k) \quad \rightarrow (2)^{k-1}(k+1) \end{cases} \tag{4.20}$$

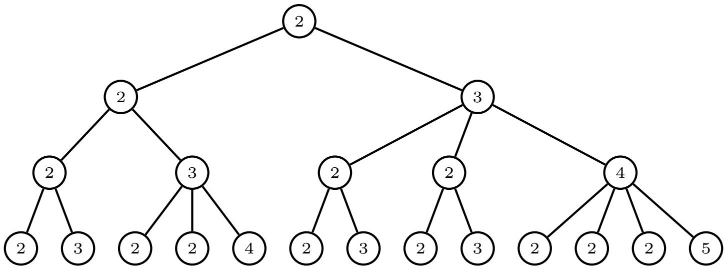

Figure 4.9: The odd Fibonacci generating tree: specification (4.20)

The associated matrix and function $P(t)$ are defined as follows:

n/k	0	1	2	3	4
0	1				
1	1	1			
2	3	1	1		
3	8	3	1	1	
4	21	8	3	1	1

$$d(t) = \frac{(1-t)^2}{1-3t+t^2}, \quad h(t) = 1, \quad A(t) = 1, \quad Z(t) = \frac{1}{(1-t)^2}$$

$$P(t) = \frac{2t^3-7t^2+5t-2}{(-1+t)(1-3t+t^2)^2} = 2+9\,t+36\,t^2+131\,t^3+444\,t^4+1432\,t^5+O\left(t^6\right)$$

For what concerns the n^{th} coefficient we have:

$$P_n \approx \frac{2\,\left(3+\sqrt{5}\right)^2\left(9+4\sqrt{5}\right)(n+1)\left(\frac{3+\sqrt{5}}{2}\right)^n}{\left(5+3\sqrt{5}\right)^2\left(\sqrt{5}+1\right)}\left(1-\frac{33\sqrt{5}-70}{5(n+1)}\right).$$

4.6 Some combinatorial identities

As we observed in Section 2, one of the main properties of Riordan arrays is their ability in dealing with combinatorial sums (see Theorem 2.5 in the present paper and Sprugnoli [20]). In this paper we have studied the internal path length of proper generating trees by finding the generating function $P(t)$ of the sequence P_n defined as:

$$P_n = \sum_{i=0}^{n} \sum_{r=0}^{i} d_{i,r}(r+c)[t^{n-i}]f_{r+c}(t), \tag{4.21}$$

where $d_{i,r}$ is the generic element of the Riordan array associated with the generating tree and $f_j(t)$ is given by Theorem 3.2. Besides its combinatorial interpretation, formula (4.21) can be used to find some interesting new combinatorial identities.

For example, for the Pascal generating tree (2.1), Theorem 3.10 gives

$$f_n^{(r+1)} = [t^n w^r] \frac{1}{(1-w)(1-2t)} = 2^n,$$

therefore we find the identity:

$$\sum_{i=0}^{n} \sum_{r=0}^{i} \binom{i}{r} (r+1) 2^{n-i} = (n+1)(n+4)2^{n-2}. \tag{4.22}$$

This is a rather simple identity but with the same approach we can find more complex examples. For the Catalan generating tree (4.16), applying Theorem 3.10 and using the relations $d(t) = 1 + td(t)^2$ and $[t^n]d(t)^p = \frac{p}{p+2n}\binom{p+2n}{n}$ (see, e.g., [7]), we have:

$$f_n^{(r+1)} = [t^n w^r] \frac{d(t)^2}{1 - wd(t)} = [t^n]d(t)^{2+r} = \frac{2+r}{2+r+2n} \binom{2+r+2n}{n},$$

with $d(t) = (1 - \sqrt{1-4t})/(2t)$. Moreover,

$$d_{i,r} = [t^{i-r}]d(t)^{r+1} = \frac{r+1}{2i-r+1} \binom{2i-r+1}{i-r}$$

and therefore we have the following non trivial identity:

$$\sum_{i=0}^{n} \sum_{r=0}^{i} \frac{(r+1)^2}{2i-r+1} \binom{2i-r+1}{i-r} \frac{2+r}{2+r+2(n-i)} \binom{2+r+2(n-i)}{n-i} =$$

$$= \frac{1}{2} 4^{n+1} - \frac{1}{2} \binom{2(n+1)}{n+1}. \tag{4.23}$$

For the generating tree (4.20), examined in the previous section, Theorem 3.4 gives:

$$F(t,w) = \frac{1 - t - w(1-2t)}{(1-3t+t^2)(1-w)^2}.$$

Hence we have:

$$f_n^{(r+2)} = [t^n w^r] \frac{1 - t - w(1-2t)}{(1-3t+t^2)(1-w)^2} = [t^n w^r] \frac{1 - t - w(1-2t)}{(1-3t+t^2)} \sum_{j \geq 0} (j+1) w^j =$$

$$= (r+1)[t^n] \frac{1-t}{(1-3t+t^2)} - r[t^n] \frac{1-2t}{(1-3t+t^2)} =$$

$$= [t^n]\frac{1+(r-1)t}{(1-3t+t^2)} = F_{2(n+1)} + (r-1)F_{2n},$$

where F_n is the nth Fibonacci number. On the other hand, we have:

$$d_{i,r} = [t^{i-r}]\frac{1-2t+t^2}{1-3t+t^2} = F_{2(i-r+1)} - 2F_{2(i-r)} + F_{2(i-r-1)}$$

and

$$P_n = [t^n]\frac{2t^3-7t^2+5t-2}{(-1+t)(1-3t+t^2)^2} = [t^n]\left(\frac{2}{1-t} - \frac{2-2t}{1-3t+t^2} + \frac{2-t}{(1-3t+t^2)}\right).$$

Therefore, we have the following interesting identity,

$$\sum_{i=0}^{n}\sum_{r=0}^{i}\left(F_{2(i-r+1)} - 2F_{2(i-r)} + F_{2(i-r-1)}\right)(r+2)\left(F_{2(n-i+1)} + (r-1)F_{2(n-i)}\right) =$$

$$= 2 - 2F_{2(n+1)} + 2F_{2n} + 2a_n - a_{n-1}, \tag{4.24}$$

where (see, sequence A001793 in [19]),

$$a_n = \frac{1}{5}\left(2(2n+1)F_{2(n+1)} + 3(n+1)F_{2n+1}\right).$$

Acknowledgments

We wish to thank Renzo Sprugnoli for very useful suggestions and comments on this paper.

References

[1] D. Baccherini, D. Merlini, and R. Sprugnoli. Binary words excluding a pattern and proper Riordan arrays. *Discrete Mathematics*, 307:1021–1037, 2007.

[2] C. Banderier, M. Bousquet-Mélou, A. Denise, P. Flajolet, D. Gardy, and D. Gouyou-Beauchamps. Generating functions of generating trees. *Discrete Mathematics*, 246:29–55, 2002.

[3] C. Banderier and D. Merlini. Lattice paths with an infinite set of jumps. *Proceedings of the 14-th International Conference on Formal Power Series and Algebraic Combinatorics, Melbourne, Australia*, 2002.

[4] E. Barcucci, A. Del Lungo, E. Pergola, and R. Pinzani. Eco: a methodology for the Enumeration of Combinatorial Objects. *Journal of Difference Equations and Applications*, 5:435–490, 1999.

[5] F. R. K. Chung, R. L. Graham, V. E. Hoggat, and M. Kleiman. The number of Baxter permutations. *Journal of Combinatorial Theory, Series A*, 24:382–394, 1978.

[6] G. P. Egorychev and E. V. Zima. Decomposition and group theoretic characterization of pairs of inverse relations of the Riordan type. *Acta Applicandae Mathematicae*, 85:93–109, 2005.

[7] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison-Wesley, New York, 1989.

[8] D. Merlini. I Riordan Array nell'Analisi degli Algoritmi. Tesi di Dottorato, Università degli Studi di Firenze, 1996.

[9] D. Merlini. Generating functions for the area below some lattice paths. In *Proceedings of Discrete Random Walks'2003*, volume AC of *Discrete Mathematics and Theoretical Computer Science*, pages 217–228, 2003.

[10] D. Merlini, D. G. Rogers, R. Sprugnoli, and M. C. Verri. On some alternative characterizations of Riordan arrays. *Canadian Journal of Mathematics*, 49(2):301–320, 1997.

[11] D. Merlini, D. G. Rogers, R. Sprugnoli, and M. C. Verri. Underdiagonal lattice paths with unrestricted steps. *Discrete Applied Mathematics*, 91:197–213, 1999.

[12] D. Merlini, R. Sprugnoli, and M. C. Verri. The tennis ball problem. *Journal of Combinatorial Theory, A*, pages 307–344, 2002.

[13] D. Merlini, R. Sprugnoli, and M. C. Verri. Waiting patterns for a printer. *Discrete Applied Mathematics*, pages 359–373, 2004.

[14] D. Merlini, R. Sprugnoli, and M. C. Verri. The Cauchy numbers. *Discrete Mathematics*, 306:1906–1920, 2006.

[15] D. Merlini and M. C. Verri. Generating trees and proper Riordan arrays. *Discrete Mathematics*, 218:167–183, 2000.

[16] D. G. Rogers. Pascal triangles, Catalan numbers and renewal arrays. *Discrete Mathematics*, 22:301–310, 1978.

[17] L. W. Shapiro. Bijections and the Riordan group. *Theoretical Computer Science*, 307:403–413, 2003.

[18] L. W. Shapiro, S. Getu, W.-J. Woan, and L. Woodson. The Riordan group. *Discrete Applied Mathematics*, 34:229–239, 1991.

[19] N. Sloane. On-line Encyclopedia of Integer Sequences. http://www.research.att.com/∼njas/sequences.

[20] R. Sprugnoli. Riordan arrays and combinatorial sums. *Discrete Mathematics*, 132:267–290, 1994.

[21] J. West. Generating trees and the Catalan and Schröder numbers. *Discrete Mathematics*, 146:247–262, 1995.

[22] J. West. Generating trees and forbidden subsequences. *Discrete Mathematics*, 157:363–374, 1996.

[23] M. Wilson. Asymptotics for generalized Riordan arrays. In *2005 International Conference on Analysis of Algorithms*, volume AD of *Discrete Mathematics and Theoretical Computer Science*, pages 323–334, 2005.

[24] T. Wang X. Zhao, S. Ding. Some summation rules related to Riordan arrays. *Discrete Mathematics*, 281:295–307, 2004.