Journal of Integer Sequences, Vol. 10 (2007),
Article 07.6.1

Parametric Production Matrices and Weighted Succession Rules: a Dyck Path Example

Robert Parviainen
ARC Centre of Excellence for Mathematics and
Statistics of Complex Systems
The University of Melbourne
139 Barry Street
3010 Victoria
Australia
robertp@ms.unimelb.edu.au

Abstract

We introduce weighted succession rules and parametric production matrices — simple extensions of the standard ECO method succession rules and production matrices. The purpose is to enumerate combinatorial objects with respect to several variables. We consider one main example, from the theory of Dyck paths. The path statistics primarily considered are peak height, rise height, and certain subwalks. Many classical sequences (such as the Catalan, Motzkin, Narayana and Schroeder numbers) are incorporated in this example.

1 Introduction

Recently Deutsch, Ferrari and Rinaldi [2] introduced production matrices as a device for ECO method calculations. Building on the concept of generating trees, the *Enumeration of Combinatorial Objects* method was developed and systematised by Barcucci and collaborators [1]. It is a general method for enumerating combinatorial objects by studying how the objects grow according to a parameter. At the centre of method are the succession rules that describe how the objects grow — production matrices translate these to matrix notation.

The production matrices considered by Deutsch et al. are integer valued. The step to include matrices with parameters is natural and straightforward, and corresponds to weighted succession rules. In this paper we study a rich example of a parametric production matrix. Although only one matrix, it embodies many well known combinatorial number sequences, such as the Catalan, Motzkin, and Narayana numbers. The underlying objects for our matrix are Dyck paths, and the main statistics considered are heights of peaks and rises, and occurrences of certain subwalks. We recover known results about the height distribution of peaks, and use the matrix to show that it is equivalent to a rise height distribution.

In Section 2 we give a brief description of the ECO method, define succession rules and production matrices, and introduce weighted successions. We recall results from Deutsch et. al [2] for operations on product matrices (which continue to hold in the parametric case). In Section 3 we define Dyck paths and some of their statistics. We then define a parametric production matrix and show how it is related to those statistics.

2 The ECO method

The core of the ECO method is a recursive description of a class of combinatorial objects. This should be done in such a way that, if $\mathcal{O}_n$ denotes the set of objects of size n, each object $O' \in \mathcal{O}_{n+1}$ is achieved from one and only one object $O \in \mathcal{O}_n$. We say that O' is a *successor* of O.

We assign a label (k), $k \in \mathbb{N}^+$ (the positive integers), to each object. An object's label gives the number of successors of that object. The succession rule dictates the labels of these successors. The rule also includes an axiom (a), $a \in \mathbb{N}^+$, which gives the label of the smallest object. In the basic case a succession rule Ω is written as

$$\Omega = \begin{cases} (a) \\ (k) \rightsquigarrow (e_1(k))\,(e_2(k))\cdots(e_k(k)) \end{cases},$$

where $e_i(k) : \mathbb{N}^+ \to \mathbb{N}^+$ gives the labels of the k successors of an object with label k.

It is natural to think of a succession rule as a generating tree: the root of the tree is the axiom, (a), and if a node has label (k), it has k children labelled $e_1(k), e_2(k), \ldots, e_k(k)$.

A succession rule Ω defines a sequence of positive integers a_n = the number of nodes at level n, $n \geq 0$. The generating function for a_n is defined by $f_\Omega(z) = \sum_n a_n z^n$.

The production matrix $P(\Omega) = P = (p_{i,j})_{i,j\geq 1}$ of a succession rule is defined as follows. List all labels of a rule. Let this list be $l_1, l_2, \ldots$. Then $p_{i,j}$ equals the number of successors with label l_j produced by an object with label l_i.

Example 1. The succession rule

$$\Omega = \begin{cases} (1) \\ (1) \rightsquigarrow (2) \\ (2) \rightsquigarrow (1)(2) \end{cases},$$

has production matrix

$$P(\Omega) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

since label (1) has one successor with label (2) and (2) has two successors with labels (1) and (2). The sequence defined by Ω is the Fibonacci sequence (A000045[1]), $a_n = 1, 1, 2, 3, 5, 8, \ldots$.

2.1 Parametric production matrices and weighted succession rules

Naturally, one would like to use the ECO method to obtain not only the size distribution of the objects, but also multi–dimensional distributions.

One strategy is to choose the labels wisely, since the distribution of labels on objects of a given size might be acquired. Indeed, using production matrices this is essentially as easy (or hard) as getting the size distribution. Another strategy is to use the succession rule and other data to find functional relations for the multivariate generating function.

One of the aims of this paper is to demonstrate an alternative way of using production matrices to find multivariate distributions, by introducing weighted succession rules and parametric production matrices. The idea is simple:

Assume we are interested in the distribution of *property*, jointly with size. If a succession $(k) \rightsquigarrow (l)$ adds one *property*, we give that succession weight x. The sequence generated by the succession rule now is the weighted sum over objects of given size, and the generating function for the sequence is the bivariate generating function for *property* and size.

Let $\mathcal{W}$ be a set of weights, and consider a matrix $P = (p_{i,j})_{i,j\geq 1}$ where $p_{i,j} \in \mathcal{W}$. This *parametric* production matrix corresponds to a *weighted* succession rule. A weighted succession rule is written as

$$\Omega = \begin{cases} (a; w\,(a)) \\ (k) \rightsquigarrow (e_1(k); w\,(e_1(k)))\,(e_2(k); w\,(e_2(k))) \cdots (e_k(k); w\,(e_k(k))) \end{cases},$$

where again $e_i(k) : \mathbb{N}^+ \to \mathbb{N}^+$ gives the labels of the k successors of an object with label (k), and $w\,(e_i(k)) \in \mathcal{W}$ gives the weight of that particular succession. We also have the possibility to give the axiom a a weight $w(a)$. Of course, if $\mathcal{W} = \{1\}$ we recover the unweighted succession rules.

The generating tree interpretation is as follows. Give the root weight $w(a)$ and label (a). A node in the tree with label (k) has k children, labelled $e_1(k)$ to $e_k(k)$. Give the edge between the parent and child $e_i(k)$ weight $w\,(e_i(k))$. Using the tree interpretation, the weight of an object is easily defined: it is the product of the weight of the root and the weights of the edges in the (unique) path in the generating tree from the root to the node corresponding to the object.

From a weighted succession rule Ω, a parametric production matrix $P(\Omega) = P = (p_{i,j})_{i,j\geq 1}$ is defined as follows. List all labels of a rule. Let this list be $l_1, l_2, \ldots$. Then $p_{i,j}$ equals the sum of weights of successors with label l_j produced by an object with label l_i. The sequence $\{a_n\}_{n\geq 0}$ produced by Ω is now a sequence of polynomials in $\mathcal{W}$, and we write

$$f_\Omega(\mathcal{W}, z) = \sum_{n=0}^{\infty} a_n z^n.$$

[1]Six-digit numbers prefixed by 'A' indicate the corresponding entry in *The On-Line Encyclopedia of Integer Sequences* [4].

Example 2. The succession rule

$$\Omega = \begin{cases} (1;1) \\ (1) \rightsquigarrow (2;1) \\ (2) \rightsquigarrow (1;x)(2;1) \end{cases},$$

has production matrix

$$P(\Omega) = \begin{pmatrix} 0 & 1 \\ x & 1 \end{pmatrix}.$$

The sequence defined by Ω gives the Fibonacci polynomials (A011973) $a_n = 1, 1, 1+x, 1+2x, 1+3x+x^2, 1+4x+3x^2, \ldots$.

Let us fix some notation. Let $u^\top = \{1\ 0\ 0\ \ldots\}$, $e = \{1\ 1\ 1\ \ldots\}^\top$ and I define a row vector, column vector, and identity matrix, respectively, of appropriate size. If a succession rule Ω defines a sequence a_n with generating function $f_\Omega(\mathcal{W}, z)$, we say that $P = P(\Omega)$ defines a_n and write $P \to a_n$ and $P \to f_\Omega(\mathcal{W}, z)$.

We now cite some useful results for production matrices. These are from Deutsch et. al [2], where they were stated in terms of integer valued production matrices. To verify that these still hold in the parametric case is trivial.

Fact 1 (Fact (iii), [2])**.** *If $P \to a_n$ then $a_n = u^\top P^n e$.*

Fact 2 (Fact (v), [2])**.** *If $P \to f(\mathcal{W}, z)$ then $f(\mathcal{W}, z) = u^\top (I - zP)^{-1} e$.*

Theorem 3 (Theorem 3.2, [2])**.** *If b, c and r are nonnegative integers or weights and $P \to f(\mathcal{W}, z)$, then*

$$M = \begin{pmatrix} b & ru^\top \\ ce & P \end{pmatrix} \to \frac{1 + rzf(\mathcal{W}, z)}{1 - bz - rcz^2 f(\mathcal{W}, z)}.$$

Corollary 4 (Corollary 3.1, [2])**.** *If b, c and r are nonnegative integers or weights and*

$$P = \begin{pmatrix} b & ru^\top \\ ce & P \end{pmatrix} \to f(\mathcal{W}, z),$$

then $f(\mathcal{W}, z)$ satisfies

$$rcz^2 f^2(\mathcal{W}, z) - (1 - (b+r)z) f(\mathcal{W}, z) + 1 = 0.$$

3 The Main Example

3.1 Dyck paths

A Dyck path is a lattice path in the first quadrant from $(0,0)$ to $(2n, 0)$ that uses steps $(1,1)$, we call these up, or u, steps, and $(1,-1)$ (down, or d, steps). We will call n the *semilength* of the path. Let $\mathbb{D}_n$ denote the set of Dyck paths of semilength n. The number of Dyck paths of semilength n are given by the nth Catalan number c_n (A000108),

$$c_n = \frac{1}{n+1}\binom{2n}{n}.$$

The generating function for the Catalan numbers satisfies

$$C(z) = 1 + zC^2(z) = \sum_{n=0}^{\infty} c_n z^n = \frac{1-\sqrt{1-4z}}{2z}.$$

A *peak* in a Dyck path p is an occurrence of an up step immediately followed by a down step. The height of the peak is the height of the start point of the u step. A *high peak* is a peak at height 1 or higher. A *rise* is a maximal sequence of consecutive u steps. The height of the rise is the number of u steps minus 1. A *high rise* is a rise with positive height. A *contact* is any place the path touches the x-axis. A *segment* is any subpath consisting of consecutive steps. The segment is a *level* segment if it starts and ends at the same height. An *excursion* is the segment of the walk between two contacts with no other contact in between. Note that a Dyck path p with $k+1$ contacts can be decomposed uniquely into its k excursions $p = e_1 e_2 \cdots e_k$.

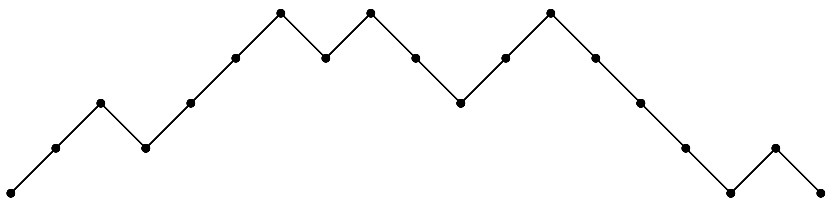

Figure 1: A Dyck path of semilength 9. The height of the rises are, in order, 1,2,0,1,0. The path has 4 high peaks, of heights 1,3,3,3. Using succession rule Ω of Theorem 5 the path will have weight $x_0^5 x_1 x_2^3 y_1^1 y_2^2 y_4^1$. Using succession rule Ω' of Theorem 8 the path will have weight $y_1^1 y_3^3$.

3.2 The matrix

We are interested in the following matrix.

$$P = \begin{pmatrix} x_1y_1 & x_0 & 0 & 0 & 0 & 0 & \cdots \\ x_2y_1 & x_1y_2 & x_0 & 0 & 0 & 0 & \cdots \\ x_3y_1 & x_2y_2 & x_1y_3 & x_0 & 0 & 0 & \cdots \\ x_4y_1 & x_3y_2 & x_2y_3 & x_1y_4 & x_0 & 0 & \cdots \\ x_5y_1 & x_4y_2 & x_3y_3 & x_2y_4 & x_1y_5 & x_0 & \cdots \\ x_6y_1 & x_5y_2 & x_4y_3 & x_3y_4 & x_2y_5 & x_1y_6 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

It is, as is shown below, the production matrix for Dyck paths, where the y_k marks the height of rise number k and x_k marks level segments of the form $u(uq_kd)^kd$, where q_k are Dyck paths. We note that many classical number sequences arise from specialisations of the sequence associated with P. See Section 3.3.

Theorem 5. *For a Dyck path p, let $h_k(p)$ denote the height of rise number k in p and let $s_k(p)$ denote the number of level segments of type $u(uq_kd)^kd$, where q_k are Dyck paths. Then the weighted sum a_n over Dyck paths of semilength n is given by*

$$a_n := \sum_{p\in\mathbb{D}_n} x_0^{h_0(p)} y_1^{s_1(p)} x_1^{h_1(p)} y_2^{s_2(p)} \cdots = u^\top P^n e,$$

and the generating function is given by

$$\begin{aligned} F(x_0, x_1, \ldots; y_1, y_2, \ldots; z) := &\sum_n z^n \sum_{p\in\mathbb{D}_n} x_0^{h_0(p)} y_1^{s_1(p)} x_1^{h_1(p)} y_2^{s_2(p)} \cdots \\ = &\; u^\top (I - zP)^{-1} e. \end{aligned}$$

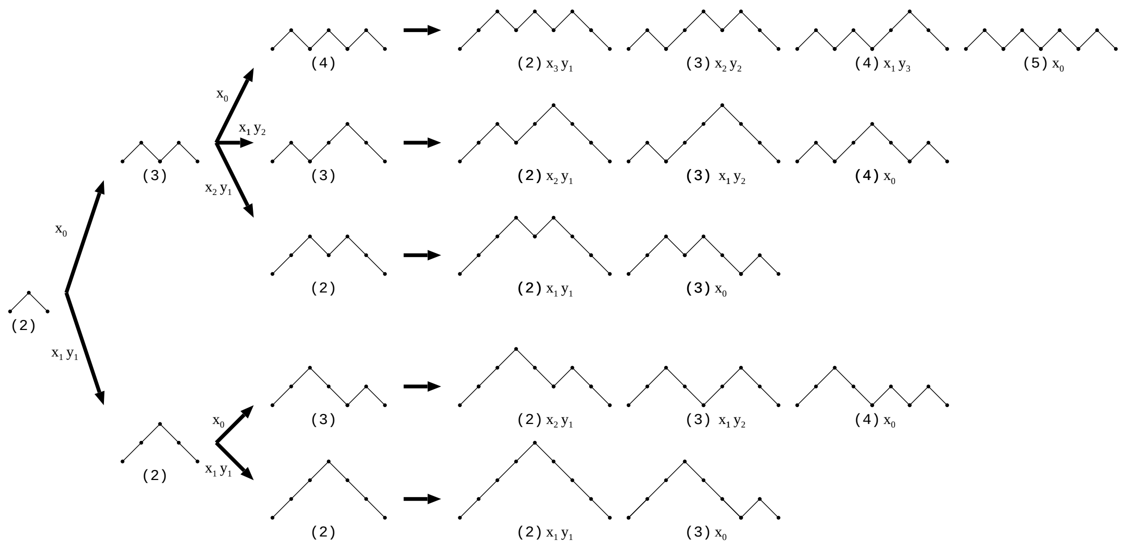

Figure 2: The generating tree associated with the succession rule of Theorem 5.

Proof. We will describe a weighted succession rule Ω for Dyck paths whose parametric production matrix is P. The labels are the number of contacts (= the number of excursions + 1). The axiom is the path ud with label $(2; x_0)$. The generating tree for this rule is sketched in Figure 2.

Given a Dyck path p of length n with k excursions, $p = e_1e_2\cdots e_k$ we now explain how the successors are derived. There are $k+1$ successors to p, with $1, 2, \cdots, k+1$ excursions. They are, respectively,

1 $ue_1e_2\cdots e_kd$

2 $e_1ue_2\cdots e_kd$

$\vdots$

k $e_1e_2\cdots e_{k-1}ue_kd$

k+1 $e_1e_2\cdots e_kud$.

Now, let a succession from a path with k excursions to one with l, $1\le l\le k+1$, have weight y_lx_{k-l+1}. The succession increases the height of rise number l by one and adds one excursion, by default a level segment, of type $u(uqd)^{k-l+1}d$ to the path as desired. Furthermore, each path has exactly one parent: the path obtained by lowering the last excursion uqd to q, a Dyck path. $\square$

In the following, we will use the notation $x=(x_0,x_1,x_2,\ldots)$ and $y=(y_1,y_2,y_3,\ldots)$. If $x_k=1$ for $k>n$ we will write $F(x_1,\ldots,x_n;y;z)$ instead of $F(x;y;z)$, and similarly for y.

Corollary 6.

$$F_0 := F(x_0,x_1;;z) = \frac{1-(x_0+x_1)z-\sqrt{(1-(x_0+x_1)z)^2-4x_0z^2}}{2x_0z^2}$$

Proof. It follows from Corollary 4 that

$$x_0z^2F_0^2-(1-(x_0+x_1)z)\,F_0+1=0.$$

$\square$

Applying Theorem 3 we get

Corollary 7.

$$F_1 := F(x_0,x_1;y_1;z) = \frac{1+x_0zF_0(x_0,x_1;;z)}{1-x_1y_1z-x_0y_1z^2F_0(x_0,x_1;;z)},$$

and in general

$$F_k := F(x_0,x_1;y_1,y_2,\ldots,y_k;z) = \frac{1+x_0zF_{k-1}(x_0,x_1;y_2,\ldots,y_k;z)}{1-x_1y_1z-x_0y_1z^2F_{k-1}(x_0,x_1;y_2,\ldots,y_k;z)}.$$

Note the use of $F_{k-1}(x_0,x_1;y_2,\ldots,y_k;z)$ *and not of* $F_1(x_0,x_1;y_1,\ldots,y_{k-1};z)$.

Below we find bivariate generating functions $F(1,\ldots,1,x_k;1;z)$ for length and height of rise number k in a Dyck path. First we show that the heights of high rises have the same distribution as the number of high peaks. The reason for not using the perhaps more familiar high peak statistics in the definition of the main production matrix is that the *joint* distribution of number of high peaks and level segments is not the same as the joint distribution of heights of rises and level segments.

Theorem 8. *The number of Dyck paths with whose high rises have heights (in order)* $k_1,\ldots,k_l$ *equals the number of Dyck paths with* $k_1,\ldots,k_l$ *high peaks of heights* $1,\ldots,l$, *respectively.*

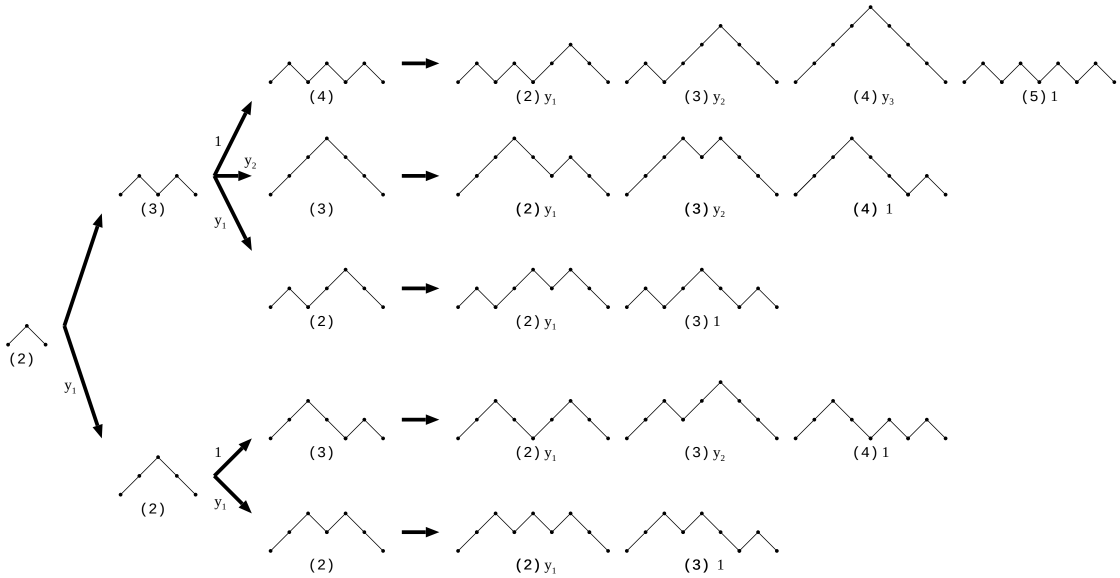

Figure 3: The generating tree associated with the succession rule of Theorem 8.

Proof. The result can be proved by describing an appropriate succession rule Ω'. This will be such that $P(\Omega') = P(\Omega)|_{x=0}$. The generating tree for this rule is sketched in Figure 3.

Paths are labeled by the number of steps in the last descent minus 1, except if that number is 0, in which we have two subcases:

i. If the path is a sequence of uds, the label is the number of those uds plus 1.

ii. Otherwise the label is the number of successive uds at the end plus the number of steps in the last descent.

In both cases the label equals the parent's label plus 1. The axiom is the path ud, and is given weight 1. The successors of a path p with label (k) are paths $p_1, \ldots, p_{k+1}$ with labels (1) to $(k+1)$. The path p_{k+1} is pud. The path p_l, $l \leq k$, is achieved from p as follows. It is a combination of the paths p and $q = u^{l+1}d^{l+1}$; the latter is placed such that it ends at $2n+2$ and in p_l we follow p up to the first intersection of p and q, after which we follow q. See Figure 4 for illustration.

Now, note that the high peaks (if any) of p are preserved in p_l, and if, $l \leq k$, p_l has an extra peak at height l. It is thus clear that if we give the updates of (k) weights $y_1, \ldots, y_k, 1$ the succession rule corresponds to heights of high peaks as desired. It is also easy to see that each path has a unique parent. $\square$

The next result is essentially equivalent to one of Mansour, [3, Theorem 1.1(ii)]. Mansour gives generating functions for k peaks of height h in terms of Chebyshev polynomials. These can be summed to give an equivalent function for peaks of height h. As a function of t, the function G_n below is of the form $\frac{a+bt}{c+dt}$, and so can easily be expanded in t to give one–dimensional generating functions if desired.

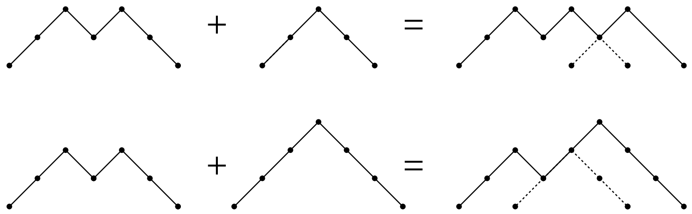

Figure 4: Example of the use of the concatenation rule used in the proof of Theorem 8.

Corollary 9. *Let $G_n(t,z)$ be the generating function for high peaks of height $n \geq 1$,*

$$G_n(t,z) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} g_{i,j} z^i t^j,$$

where $g_{i,j}$ is the number of Dyck paths of semilength $i+1$ with j high peaks at height n. Then

$$G_1(t,z) = \frac{1 + zC(z)^2}{1 - tz - tz^2C(z)^2},$$

and for $n \geq 2$

$$G_n(t,z) = \frac{G_1(t,z)zU(n-4, \frac{1}{2\sqrt{z}}) - U(n-2, \frac{1}{2\sqrt{z}})}{G_1(t,z)z^2U(n-2, \frac{1}{2\sqrt{z}}) - zU(n, \frac{1}{2\sqrt{z}})}$$

Proof. The first statement follows from Corollaries 6 and 7.

Let P_n be the matrix P with all parameters except y_n set to 1. Now

$$P_n = \begin{pmatrix} 1 & u^\top \\ e & P_{n-1} \end{pmatrix}.$$

By Theorem 3 the generating function for $P_n, n \geq 2$ is

$$G_n = \frac{1 + zG_{n-1}}{1 - z - z^2G_{n-1}}$$

Now, straightforward but cumbersome calculations show that $\frac{G_1 zU(n-4,\frac{1}{2\sqrt{z}}) - U(n-2,\frac{1}{2\sqrt{z}})}{G_1 z^2U(n-2,\frac{1}{2\sqrt{z}}) - U(n,\frac{1}{2\sqrt{z}})}$ satisfies the recursion. □

Remark 1. The stated formula for G_n in fact holds for $n = 0$ and $n = 1$ as well.

3.3 Examples

As mentioned earlier, many well–known combinatorial sequences are contained in the sequence produced by the main matrix P. Here we list a selection of the most interesting of these.

- $F(;;z) \to 1, 2, 5, 14, \ldots =$ the Catalan numbers $=$ the number of Dyck paths. (A000108).

- $F(1,0;;z) \to 1,1,2,4,9,\ldots$ = the Motzkin numbers = the number of Motzkin paths. (A001006).
- $F(2;;z) \to 1,3,11,45,\ldots$ = the little Schroeder numbers = the number of Schroeder paths with no peaks at height 1 (A001003).
- $F(s;;z) \to 1, 1+t, 1+3t+t^2, 1+6t+6t^2+t^3,\ldots$ = the Narayana numbers = the number of Dyck paths with k peaks. (A001263).
- $F(2s;;z) \to 1, 1+2t, 1+6t+4t^2, 1+12t+24t^2+8t^3,\ldots$ = the number of double rise bi-colored Dyck paths with k double rises. (A114687).
- $F(1+s;;z) \to 1, 2+t, 5+5t+t^2, 14+21t+9t^2+t^3,\ldots$ = the number of Schroeder paths with k peaks, but none at height 1 = the number of double rise bi-colored Dyck paths with k double rises of a given color. (A126216).
- $F(1,2;;z) \to 1,3,10,36,137,\ldots$ = the number of Schroeder paths with 0 peaks at odd level. (A002212).
- $F(1,s;;z) \to 1, 1+t, 2+2t+t^3, 4+6t+3t^2+t^3,\ldots$ = the number of Dyck paths with k peaks at even height. (A091869).
- $F(t,2;;z) \to 1, 2+t, 4+5t+t^2, 8+18t+9t^2+t^3,\ldots$ = the number of hex tree by edges and left edges (A126182).
- $F(1,1+t;;z) \to 1, 2+t, 5+4t+t^2, 14+15t+6t^2+t^3,\ldots$ = the number of hex tree by edges and median children (A126181).
- $F(;2;z) \to 1,3,10,35,\ldots$ = the number of times a fixed Dyck word of length 2k occurs in all Dyck words of length $2n+2k = \binom{2n+1}{n+1}$, (A001700).
- $F(;1+t;z) \to 1, 2+t, 5+4t+t^2, 14+14t+6t^2+t^3,\ldots$ = the number of leaves at level k+1 in all ordered trees with n+1 edges (A039598).
- $F(0,1;1+t;z) \to 1, 1+t, 1+2t+t^3, 1+3t+3t^2+1,\ldots$ = Pascal's triangle (A007318).
- $F(0,1+t;1;z) \to 1, 1+t, 1+2t+t^3, 1+3t+3t^2+1,\ldots$ = Pascal's triangle (A007318).
- $F(1,0;t;z) \to 1, 1, 1+t, 2+2t, 4+4t+t^2,\ldots$ = the number of Motzkin paths by height of final descent (A098979).

Note that in all these cases Corollary 4 can be use to find a quadratic equation for the generating function. Conversely, we have the following result.

Corollary 10 (Corollary 3.2, [2]). *If a generating function f satisfies an equation of the form*

$$x_0 y_1 z^2 f^2 - (1-(x_1 y_1 + x_0)z)\, f + 1 = 0,$$

then $f = F(x_0, x_1; y_1; z)$.

References

[1] E. Barcucci, A. Del Lungo, E. Pergola, and R. Pinzani. Eco: A methodology for the Enumeration of Combinatorial Objects. *J. Difference Eq. Appl.*, **5** (1999), 435–490.

[2] E. Deutsch, L. Ferrari, and S. Rinaldi. Production matrices. *Advances Appl. Math.*, **34** (2005), 101–122.

[3] T. Mansour. Counting peaks at height k in a Dyck path. *J. Integer Sequences*, **5** (2002), Article 02.1.1.

[4] N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences. Published electronically at `http://www.research.att.com/~njas/sequences/`.

2000 *Mathematics Subject Classification*: Primary 05A05; Secondary 05A15.

Keywords: ECO method, succession rule, generating tree, production Matrix, Dyck path.

(Concerned with sequences A000045, A000108, A001003, A001006, A001263, A001700, A002212, A007318, A011973, A039598, A091869, A098979, A114687, A126181, A126182, and A126216.)

Received January 29 2007; revised version received June 6 2007. Published in *Journal of Integer Sequences*, June 6 2007.

Return to Journal of Integer Sequences home page.