# Riordan Matrices and Lattice Path Enumeration



## Naiomi Cameron and Asamoah Nkwanta

In this friendly survey, we look at some interactions between a certain class of infinite matrices called <u>Riordan</u> <u>matrices</u> and the class of combinatorial objects called <u>lattice paths</u>. There are a variety of interesting algebraic and combinatorial relationships between Riordan

Communicated by Notices Associate Editor Emilie Purvine.

For permission to reprint this article, please contact: reprint-permission@ams.org.

DOI: https://doi.org/10.1090/noti2614

matrices and lattice path enumeration problems. This article attempts to introduce the reader to the concepts and techniques that may be used to explore these relationships.

The lattice paths considered in this article are presented as combinatorial objects which are subject to specific construction rules, and our interest is to enumerate and classify them with counting sequences. Riordan matrices are infinite matrices whose columns can be associated with a certain kind of sequence of generating functions. There are many interesting Riordan matrices that contain column entries or row sums associated with well-known counting numbers, such as the binomial coefficients, the ballot, Catalan, Motzkin, Schröder, Fine, Delannoy, and RNA numbers, and other counting numbers that can be found

Naiomi Cameron is a professor of mathematics at Spelman College. Her email address is naiomi.cameron@spelman.edu.

Asamoah Nkwanta is a professor of mathematics at Morgan State University. His email address is asamoah.nkwanta@morgan.edu.

in Sloane's On-line Encyclopedia of Integer Sequences [S1]. The aim of this article is to introduce the reader to Riordan matrices and to present examples of solutions to particular lattice path counting problems, where the lattice paths are enumerated by the column entries or row sums of Riordan matrices.

Algebraic combinatorics involves the use of techniques from algebra, topology, and geometry in the solution of combinatorial problems. Because of this interplay with many fields of mathematics, algebraic combinatorics is an area in which a wide variety of ideas and methods come together. Riordan arrays appear in algebraic combinatorics and are useful for proving combinatorial sums and identitites. They also appear in various counting problems in enumerative combinatorics. A certain subset of Riordan arrays called proper Riordan arrays, otherwise known as Riordan matrices, form the Riordan group, an infinite noncommutative matrix group, which is the main combinatorial device reported in this article. The Riordan group can be characterized as being algebraic and combinatorial. Thus, as a subfield of enumerative and algebraic combinatorics, the Riordan group brings together a wide variety of combinatorial methods and mathematical ideas. There are also some interesting mathematical contributions outside of combinatorics that touch on geometry, group theory, Lie algebra and groups, functional analysis, representation theory, and queuing theory. For a more thorough survey of contributions of the Riordan group to other areas of mathematics and combinatorics, see the monograph by Shapiro et al. [SSB+].

The algebraic concepts subsequently reported in this article involve finding inverse relations and multiplying, inverting, and manipulating Riordan matrices. The concepts covered herein involve using recurrence relations, generating functions, combinatorial statistics and explicit bijections to solve certain lattice path counting problems. However, the authors do not claim to solve all lattice path counting problems associated with Riordan matrices.

The article proceeds with an overview of the type of lattice paths under consideration, followed by an elementary introduction to generating functions. Riordan arrays are then formally introduced, followed by a section devoted to the algebra of the Riordan group. The last two sections of the paper feature specific demonstrations of how the Riordan group can be applied to lattice path enumeration problems, especially as related to generalized Catalan paths.

**Lattice paths.** The subject of counting paths (walks) on the lattice in Euclidean space is one of the most important areas of combinatorics. The lattice paths described in this article are defined on the integral lattice  $\mathbb{Z}^d$ , where  $d \ge 1$ , under the conditions that each path starts at some fixed origin, moves according to certain steps under specified rules and restrictions, and never crosses or touches certain

coordinate axes or hyperplanes. For excellent introductions to lattice path combinatorics and enumeration, see the surveys by Humphreys [H] and Krattenthaler [K].

A lattice path is a sequence of contiguous (unit) steps which traverses the *d*-dimensional integral lattice  $\mathbb{Z}^d$ . More precisely, a lattice path in  $\mathbb{Z}^d$  with  $k \ge 1$  unit steps is a sequence  $s_1, s_2, ..., s_k \in \mathbb{Z}^d$  such that for each  $i, 1 \le i \le k$ ,  $s_i - s_{i-1} \in S \subset \mathbb{Z}^d$ . In this case, we refer to S as a step set. We will often refer to a lattice path simply as a path, where the path is encoded by a word based on an alphabet containing S. Geometrically, a lattice path is represented by the edges between the consecutive vertices of the path. A path with no steps is a point. The length of a path is the number of unitary steps.

In this article, the paths are considered to be in ddimensional Euclidean space and never pass below a specified hyperplane. Typically, in  $\mathbb{Z}^2$ , the paths never pass below the *x*-axis. The height of a path corresponds to the *y* value of the endpoint (x, y) of the path. In three dimensions the paths are considered to be in three-dimensional Euclidean space and often never pass below the xy-plane. The height of each path corresponds to the z value of the endpoint (x, y, z) of the path. Throughout this article, we will focus mainly on lattice paths in the plane, i.e., d = 2. However, the step sets of particular higher dimensional paths defined in [N2] will be mentioned in the last section. There are many types of path problems as well as methods and models for finding their solutions. Thus, there is a vast amount of literature on this topic from the self-avoiding walk problem to lattice path problems with infinite step sets.

*Examples of lattice paths.* This article will focus mainly on Dyck (Catalan) paths and related paths, such as Motzkin and Schröder paths, with a few other paths discussed briefly.

Example 1 (Dyck Paths). A Dyck (Catalan) path is a lattice path in the first quadrant of  $\mathbb{Z}^2$  which begins at the origin (0,0), ends at (2*n*,0), has the step set  $S = \{(1,1),(1,-1)\}$ , and never goes below the *x*-axis. These paths can be described as paths that consist of unit up steps with slope 1, denoted by *U*, and unit down steps with slope -1, denoted by *D*. We refer to *n* as the semi-length of the path. A Dyck path of semi-length *n* is called a Dyck *n*-path. These paths are counted by the Catalan numbers, which are defined in the next section. The 5 possible paths of length 6 are depicted in Figure 1 and can also be encoded by the words

## UUUDDD, UUDUDD, UUDDUD, UUUUDD, UDUUDD, UDUDUD.

**Example 2** (Motzkin Paths). A Motzkin path of length *n* is a lattice path in  $\mathbb{Z}^2$  that begins at the origin (0,0), ends at (*n*,0), has step set  $\hat{S} = \{(1,1),(1,0),(1,-1)\}$  and never





Figure 2. The 4 Motzkin paths of length 3.

passes below the x-axis. These paths consist of unit up steps with slope 1 denoted by U, unit down steps with slope -1 denoted by D, and unit horizontal (or level) steps with slope 0 denoted by H. Motzkin paths are counted by the Motzkin numbers that are defined in the next section. The 4 possible paths of length 3 are shown in Figure 2 and the 9 possible paths of length 4 can be encoded by the words

#### UUDD, UHHD, UDUD, UDHH, UHDH,

#### HUDH, HUHD, HHUD, HHHH.

**Example 3** (Schröder Paths). A Schröder path is a path in the first quadrant of  $\mathbb{Z}^2$  that begins at the origin (0, 0), ends at (2n, 0), with step set  $\hat{S} = \{(1, 1), (1, -1), (2, 0)\}$  and never passes below the x-axis. These paths consist of U steps, D steps, and double horizontal or level steps denoted by  $\widetilde{H}$ . These paths are counted by the large Schröder numbers that are defined in the next section. The 6 possible paths of length 4 are depicted in Figure 3.



Figure 3. The 6 Schröder paths of length 4.

Motzkin, Dyck, and Schröder paths are interesting combinatorial objects and they appear in a wide variety of combinatorial problems. Riordan arrays have been constructed to study these paths and many other types of lattice path counting problems.

Generating functions. In order to understand Riordan matrices, some working knowledge of generating functions is needed. In this article, Riordan arrays are defined in terms of ordinary generating functions.

Given a sequence  $\langle b_n \rangle_{n>0}$  of elements of a commutative ring, the ordinary generating function (GF) for  $\langle b_n \rangle_{n>0}$  is the formal power series

$$b(z) = \sum_{n \ge 0} b_n z^n = b_0 + b_1 z + b_2 z^2 + \cdots$$

where z is an indeterminate (or auxiliary variable). In some instances "aerated" GFs are of interest. An aerated GF is a GF of the form

$$b(z^2) = \sum_{n \ge 0} b_n z^{2n} = b_0 + b_1 z^2 + b_2 z^4 + \cdots$$

where  $\langle b_0, 0, b_1, 0, b_2, ... \rangle$  is the associated sequence of aerated coefficients. In this article, we will only refer to the concept of aerated GFs once in Example 13, but otherwise they may be encountered frequently in some lattice path enumeration problems.

Note that since GFs are defined algebraically as formal power series, and not as real-valued functions, series convergence is not necessary for the existence of GFs. Nonetheless, convergence of *GFs* is sometimes necessary for finding exact formulae and asymptotic estimates of the *n*th term of a sequence. See Wilf [W2], for more details on algebraic and analytic properties of GFs.

Sequences of coefficients in combinatorics are sometimes called counting sequences and are often computed by recurrence relations. A recurrence relation (or recursion) recursively defines a sequence where the *n*th term of the sequence is expressed in terms of some previous kterms, k < n. In turn, GFs can be derived from their related recurrence relation. The famed Fibonacci numbers are computed by the following recurrence relation

$$F_{n+1} = F_n + F_{n-1}, n \ge 1$$

with initial conditions  $F_0 = F_1 = 1$ . The *GF* for this sequence is perhaps one of the most recognized and studied GFs:

$$F(z) = \sum_{n \ge 0} F_n z^n = 1 + z + 2z^2 + 3z^3 + \dots = \frac{1}{1 - z - z^2}$$

where  $\langle F_0, F_1, F_2, ... \rangle = \langle 1, 1, 2, ... \rangle$  are the popular Fibonacci numbers, which count various combinatorial objects including lattice paths. The *n*th Fibonacci number is given by

$$F_n = \left(\beta^{n+1} - \alpha^{n+1}\right)/\sqrt{5}$$

where  $\alpha = (1 - \sqrt{5})/2$ ,  $\beta = (1 + \sqrt{5})/2$ , and  $\beta$  is known as the golden ratio.

Interestingly, with the same recurrence relation, if the initial conditions are  $\hat{F}_0 = 0, \hat{F}_1 = 1$ , then

$$\hat{F}_n = \left(\beta^n - \alpha^n\right)/\sqrt{5} \text{ and } \hat{F}(z) = z/\left(1 - z - z^2\right).$$

Some suggested exercises for the reader are to use the recursion to derive F(z), and then to use F(z) to derive the exact forumula of  $F_n$  given above.

We now present three examples of GFs that are commonly encountered in lattice path enumeration.

Example 4. The Catalan *GF* is defined by

$$c(z) = \sum_{n \ge 0} c_n z^n = 1 + z + 2z^2 + 5z^3 + \dots$$
$$= \frac{1 - \sqrt{1 - 4z}}{2z}$$

where  $\langle c_n \rangle_{n \ge 0} = \langle 1, 1, 2, 5, ... \rangle$  are the Catalan numbers and  $c_n = 1/(n+1)\binom{2n}{n}$  is the *n*th Catalan number. They are computed recursively by  $c_0 = 1$ ,

$$c_{n+1} = \sum_{k=0}^{n} c_k c_{n-k}, \ n \ge 1.$$

The Catalan numbers appear in many combinatorial problems and have a variety of algebraic applications [S2].

Next, we present the Motzkin and Schröder *GFs* which have lattice path interpretations that are related to the Catalan numbers.

Example 5. The Motzkin GF is defined by

$$m(z) = \sum_{n \ge 0} m_n z^n = 1 + z + 2z^2 + 4z^3 + \cdots$$
$$= \frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z^2}$$

where  $\langle m_n \rangle_{n \ge 0} = \langle 1, 1, 2, 4, ... \rangle$  are the Motzkin numbers. They are computed recursively by  $m_0 = 1$ ,  $m_1 = 1$ ,

$$m_n = m_{n-1} + \sum_{k=0}^{n-2} m_k m_{n-2-k}, \ n \ge 2$$

and are connected to the Catalan numbers by

$$m_n = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} c_k$$

where  $\lfloor x \rfloor$  denotes the floor of *x* (i.e., the greatest integer  $\leq x$ ).

Example 6. The Schröder GF is defined by

$$\tilde{s}(z) = \sum_{n \ge 0} \tilde{s}_n z^n = 1 + 2z + 6z^2 + 22z^3 + \cdots$$
$$= \frac{1 - z - \sqrt{1 - 6z + z^2}}{2z}$$

where  $\langle \tilde{s}_n \rangle_{n \ge 0} = \langle 1, 2, 6, 22, ... \rangle$  are the large Schröder numbers. They are computed recursively by  $\tilde{s}_0 = 1$ ,  $\tilde{s}_2 = 2$ ,

$$\tilde{s}_n = 3\tilde{s}_{n-1} + \sum_{k=1}^{n-2} \tilde{s}_k \tilde{s}_{n-1-k}, \ n \ge 2$$

and are connected to the Catalan numbers by

$$\tilde{s}_n = \sum_{k=0}^n \binom{2n-k}{k} c_{n-k}, n \ge 0.$$

The Catalan, Motzkin, and Schröder *GFs* appear in many combinatorial problems. For more properties and examples of *GFs* and recurrence relations, see [W2].

What are Riordan arrays? The Riordan array concept was introduced in 1991 by Shapiro et al. [SGWW]. A Riordan array is a special infinite lower-triangular matrix where the column entries consist of coefficients of certain formal power series. Representing infinite matrices by coefficients of (formal) power series is not new and goes back to papers by Schur [S] and Jabotinsky [J] on Faber polynomials. Thus, Riordan arrays depend upon certain formal power series and *GFs*.

Suppose  $g(z) = g_0 + g_1z + g_2z^2 + \cdots$  and  $f(z) = f_1z + f_2z^2 + \cdots$ , where  $g_0 \neq 0$  and g(z) and f(z) are formal power series  $\mathbb{C}[[\mathbb{Z}]]$ . Then the infinite array

$$L = \left(l_{n,k}\right)_{n,k>0},$$

with entries in  $\mathbb{C}$  is called a <u>Riordan array</u> if the ordinary generating function for its *k*th column is the Cauchy (convolution) product of g(z) and  $f^k(z)$ . That is, *L* is a Riordan array if the *GF* for the sequence  $\langle l_{n,k} \rangle_{n \ge 0}$  of numbers in the *k*th column of *L* is

$$g(z) \cdot f^{k}(z)$$

for all  $k \ge 0$ . The constant coefficient  $g_0 = 1$  is commonly used for combinatorial convenience. Note that g(z) and f(z) are sometimes abbreviated, respectively, as g and f.

Since the column-generating functions of a Riordan array *L* form a geometric sequence  $g, gf, ..., gf^k, ..., we$  may denote the Riordan array *L* in pair form as

$$L = (g(z), f(z)) = \begin{pmatrix} | & | & | \\ g & gf & gf^2 & \cdots \\ | & | & | & | \end{pmatrix}.$$

The (n, k)-th or generic element of L can be obtained by extracting the *n*th coefficient of  $gf^k$  and obtaining

$$l_{n,k} = [z^n]g(z) \cdot f^k(z)$$

where  $[z^n]$  denotes the operator for extracting the *n*th coefficient of a generating function. For instance,  $[z^4]F(z) = 5 = F_4$ . For rules that govern the actions of  $[z^n]$ , see [SSB<sup>+</sup>]. The rules are sometimes called the method of coefficients and sometimes involve using the binomial theorem and Lagrange inversion formula to extract the coefficients.

Following the definition of *L* and if, in addition,  $f_1 \neq 0$  is satisfied, then *L* is invertible under matrix multiplication. The invertible Riordan arrays are called proper Riordan arrays or simply Riordan matrices. We note that Riordan matrices are also known as "recursive matrices" in the umbral calculus [MRSV]. Riordan matrices (arrays) are also defined by coefficients of exponential generating functions where one encounters the well-known derangement, partition (Bell), Bernoulli, Cauchy, Euler, and Stirling numbers; however, they are not described in this paper. For more information and examples of Riordan arrays and exponential formal power series see [B].

*Examples of Riordan arrays.* Here we provide examples of Riordan arrays related to the Pascal, Catalan, Motzkin, and RNA sequences.

**Example** 7 (Pascal's Triangle). Pascal triangle, the most popular example, denoted by P contains the well-known binomial coefficients. When written in infinite lower-triangular form,

$$P = \left(\frac{1}{1-z}, \frac{z}{1-z}\right) = \left(\begin{array}{ccccc} 1 & & & \\ 1 & 1 & & & \\ 1 & 2 & 1 & & \\ 1 & 3 & 3 & 1 & \\ 1 & 4 & 6 & 4 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right)$$

where

$$p_{n,k} = [z^n] (1/(1-z)) (z/(1-z))^k = \binom{n}{k}$$

is the (n, k)-th element of *P*. More generally, for integers  $t \ge 1$ 

$$P^{t} = (1/(1-tz), z/(1-tz))$$

is a generalized Pascal-Riordan matrix with generic element  $\binom{n}{k}t^{n-k}$ . The entries of *P* count certain lattice paths in  $\mathbb{Z}^2$  with horizontal and diagonal steps that do not go above the line y = x or below the *x*-axis [SSB<sup>+</sup>]. A suggested exercise for the reader is to show that the entries of  $P^t$  count the lattice paths in  $\mathbb{Z}^t$  with different kinds of horizontal steps that come in *t* different colors.

**Example 8** (Aigner-Catalan Array). The Aigner-Catalan array denoted by *C* contains the well-known ballot numbers. *C* is defined as

$$C = (c(z), zc(z)) = \begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ 2 & 2 & 1 & & \\ 5 & 5 & 3 & 1 & \\ 14 & 14 & 9 & 4 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where the leftmost column entries are the Catalan numbers, c(z) is the Catalan *GF*,

$$\tilde{c}_{n,k} = [z^n] c(z) (zc(z))^k = \frac{k+1}{n+1} {2n-k \choose n-k}$$

is the (n, k)-th element of *C*, and  $\tilde{c}_{n,k}$  are the well-known ballot numbers. The entries of *C* count certain lattice paths in  $\mathbb{Z}^2$  with horizontal and vertical steps that are restricted to lie on or below the line y = x in the coordinate plane [SSB<sup>+</sup>].

**Example 9** (Shapiro-Catalan Array). The Shapiro-Catalan array denoted by *B* is defined as

$$B = (c^{2}(z), zc^{2}(z)) = \begin{pmatrix} 1 & & & \\ 2 & 1 & & \\ 5 & 4 & 1 & & \\ 14 & 14 & 6 & 1 & \\ 42 & 48 & 27 & 8 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where the leftmost column entries are the Catalan numbers (minus the leading 1) and

$$b_{n,k} = [z^n] c^2(z) (zc^2(z))^k = \frac{k+1}{n+1} {2n+2 \choose n-k}$$

is the (n, k)-th element of *B*. The entries of *B* count certain pairs of non-intersecting lattice paths in  $\mathbb{Z}^2$  in the first quadrant.

**Example 10** (Radoux-Catalan Array). The Radoux-Catalan array denoted by  $\tilde{A}$  is defined as

$$\tilde{A} = (c(z), zc^{2}(z)) = \begin{pmatrix} 1 & & & \\ 1 & 1 & & & \\ 2 & 3 & 1 & & \\ 5 & 9 & 5 & 1 & \\ 14 & 28 & 20 & 7 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where the leftmost column entries are again the Catalan numbers and

$$a_{n,k} = [z^n] c(z) (zc^2(z))^k$$
$$= \frac{2k+1}{2n+1} {2n+1 \choose n-k}$$

is the (n, k)-th element of  $\tilde{A}$ . The entries of  $\tilde{A}$  count certain Dyck paths [A1].

The Catalan arrays  $\tilde{A}$ , B, and C are of interest and appear in many combinatorial problems.

**Example 11** (RNA Array [N1]). The RNA array, denoted by  $R^*$ , is defined as

$$R^* = (s(z), zs(z)) = \begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ 1 & 2 & 1 & & \\ 2 & 3 & 3 & 1 & \\ 4 & 6 & 6 & 4 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where the leftmost column entries are the RNA numbers (or generalized Catalan numbers) which count RNA secondary structures from molecular biology [S2], [W1]. The GF for the RNA numbers is

$$s(z) = \sum_{n \ge 0} s_n z^n = 1 + z + z^2 + \cdots$$
$$= \frac{1 - z + z^2 - \sqrt{1 - 2z - z^2 - 2z^3 + z^4}}{2z^2}$$

and  $\langle s_n \rangle_{n \ge 0} = \langle 1, 1, 1, 2, 4, 8, ... \rangle$  [W1]. The RNA numbers are computed recursively

$$s_{n+1} = s_n + \sum_{k=1}^n s_{k-1} s_{n-k}, \ n \ge 2$$

with  $s_0 = s_1 = s_2 = 1$ . They are also computed by the following sum

$$\sum_{k \ge 1} \frac{1}{n-k} \binom{n-k}{k} \binom{n-k}{k-1} (k < n).$$

The entries of  $R^*$  count certain Motzkin paths with certain restrictions [N1]. A problem of interest is to find a nice form for the generic element of  $R^*$ 

$$s_{n,k} = [z^n] s(z) (zs(z))^k$$

**Example 12** (An Improper Riordan Array). The array below denoted by  $Q^*$  is an example of an improper Riordan array [MRSV]

$$Q^* = (q(z), z^2 q(z)) = \begin{pmatrix} 1 & & & \\ 1 & 0 & & \\ 2 & 1 & 0 & & \\ 5 & 2 & 0 & 0 & \\ 12 & 5 & 1 & 0 & 0 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where q(z) is the generating function

$$q(z) = \sum_{n \ge 0} q_n z^n = 1 + z + 2z^2 + 5z^3 + \cdots$$
$$= \frac{1 - z - \sqrt{1 - 2z - 3z^2 - 4z^3}}{2z^2 (1 + z)}.$$

Note that  $[z^1]z^2q(z) = 0$ . So,  $Q^*$  is not a proper Riordan array. However, the entries of  $Q^*$  count certain lattice paths in  $\mathbb{Z}^2$  with steep and shallow diagonal steps [MRSV].

We conclude the examples by noting some well-known combinatorial triangles that are not (ordinary) Riordan arrays, as defined in this article. The Narayana array denoted by  $\tilde{T}$  contains the Narayana numbers  $\tilde{T}_{n,k}$ ,  $0 \le k < n$ .  $\tilde{T}$  is typically given as an example of an invertible, infinite lower-triangular array that is not a Riordan array.

$$\tilde{T} = \begin{pmatrix} 1 & & \\ 1 & 1 & & \\ 1 & 3 & 1 & & \\ 1 & 6 & 6 & 1 & \\ 1 & 10 & 20 & 10 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

and the generic element is

$$\tilde{T}_{n,k} = \frac{1}{n} \binom{n}{k} \binom{n}{k+1}.$$

 $\tilde{T}$  is not a Riordan array since the *k*th column is not defined by ordinary generating functions of the form  $gf^k$ . Although  $\tilde{T}$  is not Riordan, there are interesting combinatorial interpretations of  $\tilde{T}$ : the entries count the total number of Dyck *n*-paths with *k* peaks, the row sums are the Catalan numbers  $c_{n+1}$ , and the diagonal sums are the RNA numbers  $s_n$  (minus the leading 1). Suggested exercises for the reader are to verify the row sums and diagonal sums. (The diagonal sums, sometimes called slices, move upward left-to-right within the array in a north-east direction.)

Finally, we make a brief note of the Stirling triangle, which contains the Stirling numbers of the second kind S(n, k), the number of partitions of an *n*-set into *k* blocks (nonempty subsets). The Stirling triangle is not a Riordan array, as we have defined in this article. However, since the exponential generating function for S(n, k) is of the form  $gf^k/k!$  where g and f are exponential generating functions and  $k \ge 0$ , the Stirling triangle is a so-called exponential Riordan array. For more information on exponential generating functions and exponential Riordan arrays, see [B]. Algebra of the Riordan group. As mentioned earlier, the set of Riordan matrices form the Riordan group, which we will define and discuss in more detail in this section. We note that the Riordan group is also known as the Sheffer or 1-umbral group [JLN]. See [SSB<sup>+</sup>] for a connection between the Sheffer and Riordan groups. For an excellent introduction to the Riordan group and Riordan arrays, see Barry [4].

We now introduce Riordan matrix multiplication, starting with multiplication by a column vector. If L = (g(z), f(z)) is a Riordan matrix and h(z) is the *GF* associated with the infinite column vector  $h = (h_k)_{k\geq 0}^T$ , where T denotes the transpose, then one can show that the matrix product *Lh* is an infinite column vector whose associated *GF* is  $g(z) \cdot h(f(z))$ , where  $g \cdot h$  denotes a Cauchy product and h(f) denotes composition of *GFs* [SGWW]. This statement is known as the fundamental theorem of Riordan arrays (FTRA) and it allows us to define multiplication of Riordan arrays in the following way.

Suppose (g(z), f(z)) is a Riordan matrix and h(z) is a *GF*. We define the product

$$(g(z), f(z)) \otimes h(z) \coloneqq g(z) \cdot h(f(z)).$$

Here, we have introduced the  $\otimes$  symbol for use when writing the matrix product of a Riordan matrix, in terms of its associated generating function pair, and a column vector, in terms of its associated generating function. Note that the product represents the usual matrix multiplication. Next, we give some examples of properties of

Riordan matrix-vector multiplication before moving on to multiplication of two Riordan matrices.

For example, the product of Pascal's Triangle *P* and h(z) = 1/(1-z) associated with column vector  $(1, 1, ...)^{T}$  is

$$P \otimes h(z) = (1/(1-z), z/(1-z)) \otimes 1/(1-z)$$
  
= 1/(1-2z).

This example is the well-known result where the row sums of *P* are

$$[z^n] 1/(1-2z) = 2^n$$
 for  $n \ge 0$ .

We now give some important *GFs* of various sums of Riordan matrices that are obtained by multiplying by different column vectors [SSB<sup>+</sup>].

Row sums:  

$$(g(z), f(z)) \otimes 1/(1-z) = g(z)/(1-f(z))$$
Diagonal sums:  

$$(g(z), zf(z)) \otimes 1/(1-z) = g(z)/(1-zf(z))$$
Alternate sums:  

$$(g(z), f(z)) \otimes 1/(1+z) = g(z)/(1+f(z))$$
Weighted sums:  

$$(z(z), f(z)) \otimes 2 - ((1-z)^2) = \frac{g(z)f(z)}{z}$$

$$(g(z), f(z)) \otimes z/(1-z)^2 = \frac{g(z)f(z)}{(1-f(z))^2}$$

Recall that diagonal sums move upward left-to-right within the array in a north-east direction. For instance, in P the fifth diagonal sum is 8 = 1 + 4 + 3 + 0 + 0 + 0. The diagonal sums of all entries of a Riordan array L = (g(z), f(z))are computed by the row sums of the "vertically stretched" Riordan array (g(z), zf(z)). Vertically stretched Riordan arrays are not invertible because in this case the first three coefficients of zf(z) are  $f_0 = 0$ ,  $f_1 = 0$  and  $f_2 \neq 0$ . See Example 12 for an example of a vertically stretched Riordan array, and for a broader introduction to stretched Riordan arrays, see [4]. Some suggested exercises are to show that the diagonal sums of P are the Fibonacci numbers and the weighted row sums starting with leading coefficent 1 are given by the *GF*  $g(z)/(1-f(z))^2$ . Weighted row sums correspond to the first moments or the expected value of each row of the Riordan matrix. For example, the first moments of R\* are the bisected Fibonacci numbers given by *GF*  $1/(1 - 3z + z^2)$  [N1]. A straight forward calculation shows

$$(s(z), zs(z)) \otimes 1/(1-z)^2 = 1/(1-3z+z^2)$$

where s(z) is the RNA *GF*. However, the bisection formula can be applied to the Fibonacci generating function F(z). For more information on the bisection formula, see [SSB<sup>+</sup>]. The result is

$$s(z)/(1-zs(z))^{2} = \left(F\left(\sqrt{z}\right) - F\left(-\sqrt{z}\right)\right)/2\sqrt{z}.$$

This gives a nice relationship between the Fibonacci and RNA *GFs*.

Since FTRA is based on the usual row-by-column matrix multiplication, it can be extended to the matrix product of two Riordan matrices. Let L = (g(z), f(z)) and N = (h(z), l(z)) be Riordan matrices. Then the matrix product *LN* is also a Riordan matrix which can be described as follows:

$$LN = (g(z), f(z)) \otimes (h(z), l(z)) = (g(z) \cdot h(f(z)), l(f(z))).$$

Note that, in the context of this article, a product of Riordan matrices, expressed without the " $\otimes$ " symbol may be interpreted as the usual matrix multiplication.

Example 13. An exercise for the reader is to show that  $PC_0 = M$  where *P* is Pascal's Triangle,  $C_0 = (c(z^2), zc(z^2))$  is the "aerated" version of the Aigner-Catalan Riordan matrix from Example 8, and M = (m(z), zm(z)) is the Motzkin Riordan matrix from Example 5. The entries of  $C_0$  and *M* count certain partial Motzkin paths [4], [N2].

The set of all Riordan matrices forms a group under the operation of matrix multiplication [SGWW]. The identity element is (1, z). This is the usual identity with ones along the main diagonal. The inverse of (g(z), f(z)) is  $(1/g(\bar{f}(z)), \bar{f}(z))$  where  $\bar{f}(z)$  is the compositional inverse of f(z) such that  $f(\bar{f}) = z = \bar{f}(f)$ . For example, the inverse of Pascal's Triangle is

$$P^{-1} = (1/(1+z), z/(1+z))$$

and the (n, k)th element of  $P^{-1}$  is

$$[z^n]\frac{1}{1+z}\left(\frac{z}{1+z}\right)^k = (-1)^{n-k} \binom{n}{k}.$$

A slightly more complicated example, where the compositional inverse is not as obvious to compute, is

$$((1-z), z(1-z))^{-1} = C,$$

where C = (c(z), zc(z)) is the Aigner-Catalan array of Example 8. In many instances, finding  $\overline{f}(z)$  can be laborous and complicated.

**Example 14.** A nice application of P and  $P^{-1}$  involves the FTRA and the notion of binomial inversion (transform) of sequences. If A(z) is the *GF* of the generic sequence  $\langle a_n \rangle_{n \ge 0}$ , then the *GF* of the binomial transform  $\langle b_n \rangle_{n \ge 0}$  is given by

$$B(z) = P \otimes A(z)$$
  
= (1/(1-z))A(z/(1-z)).

Similarly,

$$A(z) = P^{-1} \otimes B(z) = (1/(1+z))B(z/(1+z))$$

As a result of this we have the classical binomial transform of sequences

$$b_n = \sum_{k=0}^n \binom{n}{k} a_k \iff a_n = \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} b_k$$

Example 15. Let *S* and *W* denote, respectively, column vectors associated with the sequences  $\langle s_n \rangle_{n \ge 0}$  and  $\langle w_n \rangle_{n \ge 0}$ . Then  $PW = S \iff P^{-1}S = W$ , where *P* is Pascal's Triangle. Suppose we want to find *W* such that *S* is the RNA numbers described in Example 11. Then, by the FTRA

$$P^{-1} \otimes s(z) = (1/(1+z))s(z/(1+z))$$

where s(z) is the RNA *GF*. Thus,  $P^{-1} \otimes s(z)$  equals

$$\frac{1+z+z^2-\sqrt{1+2z-z^2-6z^3-3z^4}}{2z^2(1+z)} = w(z)$$

where w(z) is the *GF* associated with

$$\langle w_n \rangle_{n>0} = \langle 1, 0, 0, 1, -1, 2, -1, 1, 3, -5, 13, -13, ... \rangle.$$

This curious sequence is not in the OEIS database [S1]. A suggested exercise is to show that  $P \otimes w(z) = s(z)$ . This gives a nice relationship between Pascal's triangle and the RNA numbers. An interesting open problem is to find a combinatorial interpretation of this multiplication that involves lattice paths or RNA secondary structures from molecular biology.

There are many important subgroups of the Riordan group. We list a few special subgroups below:

Appell subgroup:	(g(z), z)
associated subgroup	(1, f(z))
Bell subgroup	(g(z), zg(z))
checkerboard subgroup	$(g_e(z), f_o(z))$
derivative subgroup	(f'(z), f(z))
hitting time subgroup	$\left(zf'(z)/f(z),f(z)\right)$

Note that  $g_e$  and  $f_o$  are even and odd functions, respectively. f'(z) is the first derivative of f(z).  $P, B, C, C_o, M$ , and  $R^*$  are elements of the Bell subgroup. In addition,  $C_o$  is an element of the checkerboard subgroup and P the hitting time subgroup. Interestingly, one can use the fact that

$$(g(z), f(z)) = (g(z), z) \otimes (1, f(z))$$

to show that the Riordan group is a semi-direct product of the associated and Appell subgroups. For more algebraic properties of the Riordan group, see [JLN].

Finally, we mention that Riordan matrices have recursive formulas for their entries, known as "formation rules." A formation rule is a recurrence relation which describes how each entry in a Riordan matrix L can be expressed as a linear combination of entries in the preceding rows. One common formation rule is denoted by [Z;A], which represents two recurrence relations that define the way entries of L are computed. "Z" is a sequence of coefficients for

a recurrence relation that describes the entries of the leftmost or zeroth column of L and "A" is a sequence of coefficients for a recurrence relation that describes the entries of all other columns of L.

More precisely, the *Z*-sequence  $Z = \langle z_0, z_1, ... \rangle$   $(z_0 \neq 0)$  characterizes the zeroth column of  $L = (\ell_{n,k})_{n,k\geq 0}$ . This means every element  $\ell_{n+1,0}$  can be expressed as a linear combination of all the elements in the preceding row by

$$\ell_{n+1,0} = z_0 \ell_{n,0} + z_1 \ell_{n,1} + \dots = \sum_{j \ge 0} z_j \ell_{n,j}.$$

The *A*-sequence  $A = \langle a_0, a_1, ... \rangle$  ( $a_0 \neq 0$ ) characterizes the other columns of *L*. In this case every element  $\ell_{n+1,k+1}$  can be expressed as a linear combination of the elements in the preceding row, starting from the preceding column by

$$\ell_{n+1,k+1} = a_0 \ell_{n,k} + a_1 \ell_{n,k+1} + \dots = \sum_{j \ge 0} a_j \ell_{n,k+j}$$

The GFs of the A- and Z-sequences, respectively, are

$$A(z) = z/\bar{f}(z)$$
 and  $Z(z) = 1/\bar{f}(z)(1 - 1/g(\bar{f}(z)))$ 

where  $\overline{f}(z)$  is the compositional inverse of f(z) [4]. For example, the formation rule of the Shapiro-Catalan Riordan matrix *B* is [2, 1; 1, 2, 1] where Z(z) = 2 + z and  $A(z) = 1 + 2z + z^2$ . In general, the entries are computed by  $\ell_{n+1,0} = 2\ell_{n,0} + \ell_{n,1}$  and  $\ell_{n+1,k+1} = \ell_{n,k} + 2\ell_{n,k+1} + \ell_{n,k+2}$ . See [MRSV], [R] for more information on the *A*- and *Z*sequences.

**Riordan matrix method and lattice path counting.** One can use Riordan matrices to solve many kinds of lattice path enumeration problems. A typical method for using Riordan matrices to count lattice paths, as well as other combinatorial objects, is outlined by the following steps:

- 1. Count a few cases for the objects and observe some counting numbers.
- 2. Set up the counting numbers as a lower-triangular matrix *M*.
- 3. Find a pattern in the way the entries of *M* are computed.
- 4. Use the pattern (formation rules) to conjecture/define *M* as a Riordan matrix *L*.
- 5. Find the generic element of *L*.
- 6. Use the generic element or formation rules to prove combinatorially that *L* counts the objects.
- 7. Compute  $L \otimes (1/(1-z))$  to find the total number of combinatorial objects.

We now apply the Riordan matrix method to a particular lattice path counting problem. The problem was previously solved in [N2]. However, in this survey we provide more details and improve the presentation of the solution. We wish to count paths in  $\mathbb{Z}^2$  that begin at the origin (0,0), remain in the upper *xy* plane, follow the step set

$$\dot{S} = \{N(0,1), S(0,-1), R(1,0), \overline{L}(-1,0)\}$$

with north (*N*), south (*S*), right (*R*), and left  $(\overline{L})$  unit steps, and satisfy the following restrictions: (1) no path passes below the *x*-axis, (2) there are no paths with consecutive *N* and *S* steps, i.e., there are no *NS* steps, (3) no paths begin with an  $\overline{L}$  step, (4) all  $\overline{L}$  steps touch and remain on the *x*-axis. Then, we call these paths  $NSR\overline{L}$  paths. Let  $R_1^* = (\rho_{n,k})_{n,k\geq 0}$  where  $\rho_{n,k}$  denotes the number of  $NSR\overline{L}$ paths of length *n* that end at height *k*. Recall that the height corresponds to the *y* value of the endpoint (*x*, *y*) of the path. A typical path of length 10 and height 1 is denoted by the word

### RĪRRĪNRSĪN.

We want to find the total number of NSRL paths. Start by counting the number of NSRL paths of length *n* that end at height *k* for the first few cases where  $n, k \le 5$  and obtain the following lower-triangular matrix

$$R_1^* = \begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ 2 & 2 & 1 & \\ 5 & 4 & 3 & 1 & \\ 12 & 10 & 7 & 4 & 1 & \\ 29 & 25 & 18 & 11 & 5 & 1 \end{pmatrix}.$$

Then, we find the following patterns for the entries of  $R_1^*$ . We observe that the leftmost (or zeroth) column entry **29** is computed by

$$29 = 2 \cdot 12 + 1 \cdot 4 + 1 \cdot 1 + 1 \cdot 0 + 1 \cdot 0.$$

All other leftmost column entries seem to follow this same pattern (formation rule), that is, the same linear combination of respective entries from the preceding rows. We now observe the other columns entries. Observe that **18** is computed by

$$\mathbf{18} = \underline{1} \cdot \mathbf{10} + \underline{1} \cdot \mathbf{7} + \underline{1} \cdot \mathbf{1} + \underline{1} \cdot \mathbf{0} + \underline{1} \cdot \mathbf{0}.$$

All other **bold** entries in the other columns (not the leftmost) seem to follow this pattern. Thus, we conjecture this pattern (formation rule) and make the assumption that the (n, k)-th entry of  $R_1^*$  is formed by the following recurrence relations. The leftmost or zeroth column is formed for  $\rho_{0,0} = 1$ ,  $\rho_{1,0} = 1$ , and  $n \ge 1$ 

$$\rho_{n+1,0} = 2\rho_{n,0} + \sum_{k \ge 1} \rho_{n-k,k}.$$

The other columns are formed for  $n, k \ge 1$ 

$$\rho_{n+1,k} = \rho_{n,k-1} + \sum_{j \ge 0} \rho_{n-j,k+j}$$
$$= \rho_{n,k-1} + \rho_{n,k} + \cdots.$$

As previously computed, one can easily confirm  $\rho_{5,0} = 29$  and  $\rho_{5,2} = 18$ .

From the recurrence relations, we will derive explicit *GFs* whose coefficients are the column entries of  $R_1^*$ . Following the matrix formation rules, the definition of a Riordan matrix, and making adjustments to properly align the coefficients, the *k*th column *GF* of  $R_1^*$  is defined for  $k \ge 1$  as

$$g \cdot f^{k} = zgf^{k-1} + z(gf^{k} + zgf^{k+1} + z^{2}gf^{k+2} + \cdots).$$

Solving for *f* gives

$$f = zf^2 + (z - z^2)f + z.$$

Now, we solve f in terms of f(z) and apply the quadratic formula to obtain f(z) = zs(z) where s(z) is the RNA *GF*. Similarly, the leftmost column *GF* of  $R_1^*$  is defined as

$$g = 1 + z (2g + zgf + z^2gf^2 + \cdots).$$

Using geometric series and simplifying,

$$g = 1 + 2zg + z^2gf + z^3gf^2 + \cdots$$
  
= 1 + 2zg + z^2gf \cdot (1 + zf + z^2f^2 + \cdot)  
= 1 + 2zg + z^2gf \cdot (1/(1 - zf))  
= (1 - zf)/(1 - 2z - zf + z^2f).

Now, we solve g in terms of g(z) and substitute f(z) = zs(z) to obtain

$$g(z) = (1 - z^2 s(z)) / (1 - 2z - z^2 s(z) + z^3 s(z))$$

Simplifying,

$$g(z) = \frac{1 - z}{2z} \left( \frac{1 - 3z + z^2 - \sqrt{\Delta}}{-1 + 3z - z^2} \right)$$

where  $\triangle = 1 - 2z - z^2 - 2z^3 + z^4$ . This confirms the pattern (formation rule) and we conjecture that  $R_1^*$  is the Riordan matrix defined by the *GFs* derived from the recurrence relations. Thus,

$$R_{1}^{*} = \left(\frac{1-z}{2z} \left(\frac{1-3z+z^{2}-\sqrt{\Delta}}{-1+3z-z^{2}}\right), zs(z)\right).$$

We now give combinatorial arguments that show that the entries of  $R_1^*$  model the lattice path counting problem earlier described in this section. We want to find the total number of *NSRL* paths of length *n* ending at height *k*. We will connect the entries of  $R_1^*$  to the paths by showing the paths satisfy the recurrence relations. To do this we consider the following combinatorial arguments. Consider an *NSRL* path of length *n* ending at height *k*. To form a new path of length n + 1 of height *k*, we consider the following cases. Case 1: Given a path of length *n* ending at height k-1, there is one choice to move the path up to height *k*, a north step *N*. Thus, if the last step is *N*, there are  $1 \cdot \rho_{n,k-1}$ possibilities to construct a new path ending at height *k*. In this case, all paths with last step *N* are counted by  $\rho_{n,k-1}$ . Case 2: Given a path of length *n* ending at height *k*, there is one choice for the path to remain at height  $k_i$  a right step R. Thus, if the last step is R, there are  $1 \cdot \rho_{n,k}$  possibilities to construct a new path ending at height k. In this case, all paths with last step *R* are counted by  $\rho_{n,k}$ . Case 3: For  $j \ge 0$ , given a path of length n - j ending at height k + j, there is one choice of j south steps to move the path down to height *k*. Thus, there are  $1 \cdot \rho_{n-j,k+j}$  possibilities to construct a new path ending at height k. In this case, all paths with *j* south steps are counted by  $\rho_{n-j,k+j}$ . For k > 0there are no paths with a last left step  $\overline{L}$  since  $\overline{L}$  steps touch and remain on the x-axis. Now, summing all cases gives all possible ways to construct a new path of lenth n + 1ending at height k. Therefore, the recurrence relation for the other columns of  $R_1^*$  counts all  $NSR\overline{L}$  paths of length *n* ending at height  $k \ge 1$ . By similar arguments, we can show that the leftmost column of  $R_1^*$  counts all paths of length *n* ending at height k = 0. This proves the formation rule and gives  $R_1^*$  a NSRL lattice path interpretation.

To find the total number of NSRL paths, we apply the FTRA and compute the row sums by

$$R_1^* \otimes (1/(1-z)).$$

Simplifying, the total number of  $NSR\overline{L}$  paths is given by

$$R_1^* \otimes (1/(1-z)) = (1-z)/(1-3z+z^2)$$
  
=  $\sum_{n \ge 0} \rho_n z^n = BF(z)$ 

where the *n*th coefficient is computed by the following sums [A2]

$$[z^{n}]BF(z) = 1 + \sum_{k=0}^{n} F_{2k} = \sum_{k=0}^{n} \binom{2n-k}{k}.$$

For example,

$$[z^{3}]BF(z) = 1 + \sum_{k=0}^{3} F_{2k}$$
$$= 13 = \rho_{3}.$$

Thus, the total number of NSRL paths are counted by the bisected Fibonacci numbers with generating function BF(z). The number of Dyck paths of length 2n and height at most 3 are also counted by these bisected Fibonacci numbers. A suggested exercise is to find an explicit bijection between these two sets of paths. See OEIS [S1] sequence number A001519 for other combinatorial objects enumerated by the counting numbers  $\langle \rho_n \rangle_{n>0}$ .

Generalized Catalan paths and Riordan matrices. The Riordan matrix method can be readily used to study statistics on lattice paths and other combinatorial objects. In this final section, we give some illustrations of how Riordan matrices can be leveraged to study statistical distributions of generalized lattice paths and to observe combinatorial relationships among higher-dimensional lattice paths.

A frequent example of lattice paths are the so-called Catalan paths, which are equivalent to ballot numbers and Dyck paths. A Catalan path is a path starting from the origin (0, 0), using *n* steps of the form (1, 0) and *n* steps of the form (0, 1), that does not go above the line y = x. Catalan paths are equivalent to Dyck *n*-paths, as described in Example 1. The number of Dyck *n*-paths that end at height *k* is given by the (*n*, *k*)th entry of the (aerated) Catalan matrix  $C_0 = (c(z^2), zc(z^2))$ .

Catalan paths can be generalized in a variety of ways to obtain other interesting lattice paths associated with well-known combinatorial sequences. For instance, *t*-Dyck paths are paths of length *tn* from (0,0) to (*tn*,0), where t > 0, that use up steps *U* of the form (1, 1) and down steps *D* of the form (1, -t + 1) and do not go below the *x*-axis. When t = 2, we have Dyck paths while the paths we obtain when t = 3 are called ternary paths which are counted by the ternary numbers  $\frac{1}{2n+1} {3n \choose n}$  with generating function T(z) satisfying  $T(z) = 1 + zT^3(z)$ .

A *t*-Dyck path can also be thought of as a path starting from the origin (0,0) that uses *n* steps of the form (1,0) and (t-1)n steps of the form (0,1) that does not go above the line y = (t-1)x. In general, *t*-Dyck paths are counted by the Fuss-Catalan numbers  $\frac{1}{(t-1)n+1} {tn \choose n}$ .

One may be interested in enumerating certain properties of these generalized lattice paths, such as the number of peaks, valleys, returns, or hills. A return in a Dyck path is a non-origin point on the path that intersects the *x*axis. The Riordan matrix C = (c(z), zc(z)), from Example 8, counts the number of Catalan paths of length 2(n + 1)having exactly k + 1 returns to the *x*-axis. A quick computation of  $(c(z), zc(z)) \otimes \frac{1}{1-z}$  verifies that the row sums of (c(z), zc(z)) indeed count the total number of Catalan paths of length 2(n + 1). Moreover, the computation of

$$\frac{[z^n](c(z), zc(z)) \otimes \frac{1}{(1-z)^2}}{[z^n]c(z)}$$

produces the average number of returns among Dyck *n*-paths. An exercise for the reader is to perform this computation and show that the average number of returns among Dyck *n*-paths approaches 3 as  $n \rightarrow \infty$ .

Similarly, the total number of returns among ternary paths of length 3*n* is given by

$$(1, z(T(z))^2) \otimes \frac{z}{(1-z)^2} = z(T(z))^4.$$

Hence, the expected number of returns among ternary paths is

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

which approaches 2 as  $n \to \infty$ . Furthermore, one can use these methods to show that the probability that a randomly chosen ternary path has exactly k returns approaches  $\frac{4k}{3^{k+1}}$  and that the expected number of returns among nontrivial *t*-Dyck paths approaches  $\frac{t+1}{t-1}$  with variance  $\frac{2t}{(t-1)^2}$ . Ultimately, these results lead to the fact that the number of returns among *t*-Dyck paths approaches a negative binomial distribution with parameters 2 and  $\frac{t-1}{t}$ . [CM]

A hill in a *t*-Dyck path is a subsequence of the form  $U^{t-1}D$  such that the preceding sequence of steps form a *t*-Dyck path. The number of Dyck *n*-paths having exactly  $k \ge 0$  hills is given by the Riordan array  $(\tilde{f}(z), z\tilde{f}(z))$  where

$$\tilde{f}(z) = \frac{1 - \sqrt{1 - 4z}}{z \left(3 - \sqrt{1 - 4z}\right)} = \frac{c(z)}{1 + zc(z)}$$

is the *GF* for the Fine numbers  $\langle \tilde{f}_n \rangle_{n \ge 0} = \langle 1, 0, 1, 2, 6, 18, 57, 186, ... \rangle$ . Dyck paths having no hills, known as Fine paths, are counted by the Fine numbers. The reader may verify as an exercise that the total number of hills among Dyck *n*-paths is given by

$$[z^n]\left(\tilde{f}(z), z\tilde{f}(z)\right) \otimes \frac{z}{(1-z)^2} = c(z) - 1,$$

which implies that the expected number of hills among Dyck *n*-paths is exactly 1. This result suggests that there is a bijection between the set of Dyck *n*-paths and the set of hills among all Dyck *n*-paths. One nice bijection can be described as follows. Given the hill  $\underline{UD}$  in a Dyck *n*-path of the form  $S_1\underline{UDS}_2$ , where *U* is an up step, *D* is a down step, and  $S_1, S_2$  are Dyck paths of semilength at most *n*, map the hill  $\underline{UD}$  to the Dyck *n*-path  $US_1DUS_2$ . It should be clear that the mapping is injective, and since every Dyck *n*-path can be described in the latter form, the mapping is surjective.

More generally speaking, similar methods on the generalized Fine array  $(\tilde{f}_t(z), z\tilde{f}_t(z))$ , where  $\tilde{f}_t(z)$  is the *GF* for the number of *t*-Dyck paths having no hills, can be used to show that the number of hills among *t*-Dyck paths of length *tn* has expected value

$$\frac{2((t-1)n+1)^{t-2}}{t(tn-1)^{t-2}}$$

where  $t^{\underline{j}} \coloneqq t(t-1)\cdots(t-j+1)$ , and approaches a negative binomial distribution with parameters 2 and  $\frac{t^{t-1}}{t^{t-1}+(t-1)^{t-2}}$  [CM].

As mentioned previously, lattice paths from (0,0) to (n,0) that use up steps of the form (1, 1), down steps of the form (1, -1), and level steps of the form (1,0) are known as Motzkin paths, which are counted by the Motzkin numbers. When we allow two possible colors for the level step,

the resulting paths, which we will call 2-colored Motzkin paths, are counted by the Catalan numbers with generating function  $c^2(z)$ . In fact, the (n, k)th entry of the Bell subgroup Riordan matrix  $B = (c^2(z), zc^2(z))$ , from Example 9 is the number of these 2-colored Motzkin paths of length n having terminal height k. Recall that a formation rule for  $(c^2(z), zc^2(z))$  is [2, 1; 1, 2, 1]. Another lattice path interpretation of the array  $(c^2(z), zc^2(z))$  is the number of Dyck paths of length 2n + 1 that go from (0, 0) to (2n + 1, 2k + 1).

On the other hand, ternary paths of length 3n + 1 that go from (0, 0) to (3n+1, 3k+1) are counted by the Bell subgroup matrix  $(T^3(z), zT^3(z))$  where T(z) is the *GF* for the ternary numbers. The formation rule for  $(T^3(z), zT^3(z))$  is [3, 3, 1; 1, 3, 3, 1]. Given this formation rule, it is not hard to see that the (n, k)th entry of  $(T^3(z), zT^3(z))$  also counts paths from (0, 0) to (n, k) that never go below the *x*-axis and use the following 8 types of steps: U(1, 1),  $L_1(1, 0)$ ,  $L_2(1, 0)$ ,  $L_3(1, 0)$ ,  $D_1(1, -1)$ ,  $D_2(1, -1)$ ,  $D_3(1, -1)$ . There is a natural bijection between these paths and the aforementioned ternary paths.

More generally, the Bell subgroup matrix whose formation rule is determined by the *t*th row of Pascal's triangle will count lattice paths of length tn+1 that go from (0,0) to (tn+1,tk+1) and use up steps of the form (1,1) and down steps of the form (1, -t + 1) that never go below the *x*-axis. And furthermore, this same matrix will count paths from (0,0) to (n,k) that never go below the *x*-axis and use  $2^t$  possible steps, each of the form  $L_i(1, -k+1)$  where k = 0, 1, ..., tfor some *i* between 1 and  $\binom{t}{k}$ . Given that there are natural bijections between these two different path interpretations, it would be interesting to explore through the Riordan matrix method how certain statistics may translate between these lattice path interpretations and how the resulting statistical distributions compare.

As a final illustration of the vast connections between lattice path enumeration and Riordan matrices, we present two open problems related to higher-dimensional lattice paths and generalized Riordan arrays.

Consider the generalized Riordan matrix

$$R_t^* = \left( \left( (1-z) / (1-zs(z)) \right)^t s(z), zs(z) \right)$$

where s(z) is the RNA *GF*. Note that  $R^*$  and  $R_1^*$  are special cases where t = 0 and t = 1, and for t = 2, we have

$$R_2^* = \left( \left( 1 - 2z + z^2 \right) / \left( 1 - 3z + z^2 \right), zs(z) \right)$$

Interesting open problems are to find lattice path interpretations of the generalized Riordan matrix  $R_t^*$  for  $t \ge 2$ . In Figure 4, there is a nice example of an infinite twodimensional array made up of products of Riordan matrices that involve Pascal's triangle *P* and the "aerated" Aigner-Catalan matrix  $C_0 = (c(z^2), zc(z^2))$ . The infinite array denoted by  $\pounds_{t,\nu} = P^t C_0 E^{\nu}$  constructed in [N2] is a triple product of Riordan matrices where  $E^{\nu} = (1/(1-z)^{\nu}, z)$ . Some

**Figure 4.** Each entry is a product of Riordan matrices  $\pounds_{t,\nu} = P^t C_0 E^{\nu}$ , where *P* is Pascal's triangle,  $C_0$  is a special Shapiro-Catalan array and *E* is the lower-triangular array whose entries equal 1.

interesting matrices that appear in  $\pounds_{t,\nu}$  are the Pascal, Catalan, Motzkin, hexagonal (Hex), and directed animal arrays. Solutions of certain higher-dimensional lattice path counting problems are known for the first two columns of this infinite two-dimensional array, but less is known about the remaining columns. Thus,  $\pounds_{t,\nu}$  is of combinatorial interest. See [N2] for more information on Riordan matrices and generalized lattice paths.

 $f_{t,\nu}$  is the Riordan matrix

$$\pounds_{t,\nu} = \left(k_t(z) / (1 - zk_t(z))^{\nu}, zk_t(z)\right)$$

where

$$k_t(z) = \left(1 - tz - \sqrt{1 - 2tz + (t^2 - 4)z^2}\right)/2z^2$$
$$= \left(1/(1 - tz)\right)c\left(\left(z/(1 - tz)\right)^2\right)$$

and c(z) is the Catalan GF. Thus,  $\pounds_{t,\nu}$  is a generalized Catalan Riordan matrix. See Figure 4 for the first few entries of  $f_{t,\nu}$ . Surprisingly, by moving down the leftmost column of  $\mathfrak{L}_{t,\nu}$  a solution of Sands's problem is found in [N2]. In addition, Sands's problem was extended to higher dimensions in  $\mathbb{Z}^d$  for d > 3 where there are countably many step directions and the paths never pass below the (d-1)th hyperplane  $x_1 + \cdots + x_{d-1} = 0$ . The height of each higherdimensional path corresponds to the value  $x_d$  of the endpoint  $(x_1, \dots, x_d)$  of the path. The step sets of the higherdimensional paths are defined in [N2]. Another subclass of paths called generalized (or partial-t) Motzkin walks are also counted by the entries of the leftmost column. In the same paper [N2], higher-dimensional path results were also obtained by moving down the first column of  $\pounds_{t,\nu}$ . By moving down the first column, surprisingly, a more restrictive subset of Sands type paths called *power walks* are obtained. These paths are generalized to higher dimensions and a Motzkin analog is given [N2]. An interesting open problem is to find a higher-dimensional lattice path interpretation of the matrix entries  $P^t C_0 E^2$  of the second column of  $\pounds_{t,\nu}$ . There is not much known about  $\pounds_{t,\nu}$  for column entries for  $\nu \geq 2$ .

ACKNOWLEDGMENT. The authors would like to thank the associate editor and anonymous referees for useful comments that improved this article.

#### References

- [A1] Martin Aigner, Catalan-like numbers and determinants, J. Combin. Theory Ser. A 87 (1999), no. 1, 33–51, DOI 10.1006/jcta.1998.2945. MR1698277
- [A2] Mohammad K. Azarian, Fibonacci identities as binomial sums, Int. J. Contemp. Math. Sci. 7 (2012), no. 37-40, 1871–1875. MR2959001
- [B] Paul Barry, Exponential Riordan arrays and permutation enumeration, J. Integer Seq. 13 (2010), no. 9, Article 10.9.1, 16. MR2746249
- [4] Paul Barry, *Riordan arrays: a primer*, Kildare, Ireland: Logic Press, 2016.
- [CM] Naiomi T. Cameron and Jillian E. McLeod, Returns and hills on generalized Dyck paths, J. Integer Seq. 19 (2016), no. 6, Article 16.6.1, 28, DOI 10.9734/bjmcs/2016/30398. MR3546615
- [H] Katherine Humphreys, A history and a survey of lattice path enumeration, J. Statist. Plann. Inference 140 (2010), no. 8, 2237–2254, DOI 10.1016/j.jspi.2010.01.020. MR2609483
- [J] Eri Jabotinsky, Representation of functions by matrices. Application to Faber polynomials, Proc. Amer. Math. Soc. 4 (1953), 546–553, DOI 10.2307/2032522. MR59359
- [JLN] Candice Jean-Louis and Asamoah Nkwanta, Some algebraic structure of the Riordan group, Linear Algebra Appl. 438 (2013), no. 5, 2018–2035, DOI 10.1016/j.laa.2012.10.027. MR3005272
- [K] Christian Krattenthaler, Lattice path enumeration, Handbook of enumerative combinatorics, Discrete Math. Appl. (Boca Raton), CRC Press, Boca Raton, FL, 2015, pp. 589– 678. MR3409351
- [MRSV] Donatella Merlini, Douglas G. Rogers, Renzo Sprugnoli, and M. Cecilia Verri, On some alternative characterizations of Riordan arrays, Canad. J. Math. 49 (1997), no. 2, 301–320, DOI 10.4153/CJM-1997-015-x. MR1447493
- [N1] Asamoah Nkwanta, Lattice paths and RNA secondary structures, African Americans in mathematics (Piscataway, NJ, 1996), DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 34, Amer. Math. Soc., Providence, RI, 1997, pp. 137–147. MR1482263
- [N2] Asamoah Nkwanta, Riordan matrices and higherdimensional lattice walks, J. Statist. Plann. Inference 140 (2010), no. 8, 2321–2334, DOI 10.1016/j.jspi.2010.01.027. MR2609490
- [R] D. G. Rogers, Pascal triangles, Catalan numbers and renewal arrays, Discrete Math. 22 (1978), no. 3, 301–310, DOI 10.1016/0012-365X(78)90063-8. MR522725
- [S] Issai Schur, On Faber polynomials, Amer. J. Math. 67 (1945), 33–41, DOI 10.2307/2371913. MR11740

- [SSB<sup>+</sup>] Louis Shapiro, Renzo Sprugnoli, Paul Barry, Gi-Sang Cheon, Tian-Xiao He, Donatella Merlini, and Weiping Wang, *The Riordan group and applications*, Springer Monographs in Mathematics, Springer, Cham, [2022] @2022. With a foreword by Richard Stanley, DOI 10.1007/978-3-030-94151-2. MR4424807
- [SGWW] Louis W. Shapiro, Seyoum Getu, Wen Jin Woan, and Leon C. Woodson, *The Riordan group*, Discrete Appl. Math. 34 (1991), no. 1-3, 229–239, DOI 10.1016/0166-218X(91)90088-E. Combinatorics and theoretical computer science (Washington, DC, 1989). MR1137996
- [S1] Neil J. A. Sloane, The on-line encyclopedia of integer sequences, Notices Amer. Math. Soc. 65 (2018), no. 9, 1062– 1074. MR3822822
- [S2] Richard P. Stanley, Enumerative combinatorics. Vol. 2, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin, DOI 10.1017/CBO9780511609589. MR1676282
- [W1] Michael S. Waterman, Secondary structure of singlestranded nucleic acids, Studies in foundations and combinatorics, Adv. in Math. Suppl. Stud., vol. 1, Academic Press, New York-London, 1978, pp. 167–212. MR520559
- [W2] Herbert S. Wilf, generatingfunctionology, 2nd ed., Academic Press, Inc., Boston, MA, 1994. MR1277813





Naiomi Cameron

Credits

## ron Asamoah Nkwanta

Opening image is courtesy of shaunl via Getty. Figures 1, 2, 3, and 4 are courtesy of Naiomi Cameron. Photo of Naiomi Cameron is courtesy of Aaron Fagerstrom. Photo of Asamoah Nkwanta is courtesy of Charlita Woodruff-White.

# ew from the AMS

## **Linear Algebra**

## **Vector Spaces and Linear Transformations**

## Meighan I. Dillon, Kennesaw State University, Marietta, GA

This textbook is directed towards students who are familiar with matrices and their use in solving systems of linear equations. The emphasis is on the algebra supporting the ideas that make linear algebra so important, both in theoretical and practical applications. The narrative is written to bring along students who may be new to the level of abstraction essential to a working understanding of linear algebra. The determinant is used throughout, placed in some historical perspective, and defined several different ways, including in the context of exterior algebras. The text details proof of the existence of a basis for an arbitrary vector space and addresses vector spaces over arbitrary fields. It develops LU-factorization, Jordan canonical form, and real and complex inner product spaces. It includes examples of inner product spaces of continuous complex functions on a real interval, as well



as the background material that students may need in order to follow those discussions. Special classes of matrices make an entrance early in the text and subsequently appear throughout. The last chapter of the book introduces the classical groups.

Pure and Applied Undergraduate Texts, Volume 57; 2023; 367 pages; Softcover; ISBN: 978-1-4704-6986-3; List US\$85; AMS members US\$68; MAA members US\$76.50; Order code AMSTEXT/57



## Learn more at bookstore.ams.org/amstext-57

Background image credit: Liudmila Chernetska / iStock / Getty Images Plus via Getty Images