Dimacs Services (1997) pp. 137–

Lattice Paths and RNA Secondary Structures

Asamoah Nkwanta

January 13, 1997

Abstract. Four infinite lower-triangular matrices, each of whose entries count lattice paths or random walks, are presented and denoted as P, C_0, C and M where M is a Motzkin triangle, C and C_0 are Catalan triangles, and P is Pascal's triangle. By matrix multiplication, another infinite lower-triangular matrix denoted as R is defined by $C_0 \cdot R = M$. Then, the following results are proved about R:

1) $M \cdot R = C$.
2) The entries in the left most column of R count the number of RNA secondary structures of length n (from molecular biology).
3) A combinatorial interpretation of R is given in terms of lattice paths.
4) There is a one-to-one correspondence between RNA secondary structures and these lattice paths.
5) The first moments of R are every other Fibonacci number.

Results related to the Narayana numbers and noncrossing partitions are also discussed.

1. Introduction

We consider four infinite lower-triangular matrices denoted as C_0, M, C and P, where the first few terms of each triangle are listed below:

$$C_0 = \begin{bmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ 1 & 0 & 1 & & & \\ 0 & 2 & 0 & 1 & & \\ 2 & 0 & 3 & 0 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \qquad M = \begin{bmatrix} 1 & & & & & \\ 1 & 1 & & & & \\ 2 & 2 & 1 & & & \\ 4 & 5 & 3 & 1 & & \\ 9 & 12 & 9 & 4 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & & & & & \\ 2 & 1 & & & & \\ 5 & 4 & 1 & & & \\ 14 & 14 & 6 & 1 & & \\ 42 & 48 & 27 & 8 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \text{and} \quad P = \begin{bmatrix} 1 & & & & & \\ 1 & 1 & & & & \\ 1 & 2 & 1 & & & \\ 1 & 3 & 3 & 1 & & \\ 1 & 4 & 6 & 4 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

1991 *Mathematics Subject Classification.* Primary 05A15; Secondary 92D20.

Partially supported by an HBCU fellowship from the Jet Propulsion Laboratory, Pasadena, CA.

P is the well known Pascal triangle. M is called a Motzkin triangle since its left most column contains the Motzkin numbers, $m_n = \sum_{k\geq 0} \frac{1}{k+1}\binom{2k}{k}\binom{n}{2k}$. Similarly, C_0 and C are called Catalan triangles since their left most columns contain the Catalan numbers, $c_n = \frac{1}{n+1}\binom{2n}{n}$. The entries of C_0 are sometimes referred to as the aerated Catalan numbers since zeros are between each number. These triangles often arise in combinatorial applications. For instance, the Motzkin triangle has interpretations as random walks [4], and as interval graphs [8]. The Catalan triangles have interpretations as ballot sequences [17], and as lattice paths (or walks) with various restrictions [6], [7]. Most of the interpretations mentioned in this paper are related to combinatorial objects called lattice paths. What we mean by a *lattice path* is a sequence of contiguous unit steps of length n which traverse an integer lattice. The lattice paths are in the (x, y) plane such that all paths begin at the origin, $(0, 0)$, and never go below the x-*axis*. The length of each path is the number of unit-steps and the height corresponds to the y value of the point (x, y) at the end of the path. The symbols N, S, E and W denote unit-steps in the north, south, east and west directions, respectively. Thus, from the set of lattice paths, we have the following interpretation. The (n, k)th entry of C_0 (M and C, respectively) is the number of unit-step NS (NSE and NSEW, respectively) lattice paths of length n and height k.

Since C_0 is invertible, another infinite lower-triangular matrix R can be defined by $C_0 \cdot R = M$. The first few terms of R and $(C_0)^{-1}$ are

$$R = \begin{bmatrix} 1 \\ 1 & 1 \\ 1 & 2 & 1 \\ 2 & 3 & 3 & 1 \\ 4 & 6 & 6 & 4 & 1 \\ 8 & 13 & 13 & 10 & 5 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \text{ and } (C_0)^{-1} = \begin{bmatrix} 1 \\ 0 & 1 \\ -1 & 0 & 1 \\ 0 & -2 & 0 & 1 \\ 1 & 0 & -3 & 0 & 1 \\ 0 & 3 & 0 & -4 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

R can also be defined by $R = (C_0)^{-1} \cdot P \cdot C_0$ since $M = P \cdot C_0$. These relations and other matrix relations involving R are defined in section 2. A surprising fact about R is that the entries in the left most column, sequence $\{1, 1, 1, 2, 4, 8, 17, \ldots\}$, count ribonucleic acid (RNA) secondary structures of length n. As a result of the sequence, we summarize some of the combinatorial aspects of RNA secondary structures in section 3. Then, a lattice path interpretation of R is defined in section 4. Given the interpretation, a one-to-one correspondence between RNA secondary structures and lattice paths is constructed in section 5. Then, the first moments of R are computed and shown to be equal to the alternating Fibonacci numbers in section 6. We then conclude with applications of the Narayana numbers and noncrossing partitions in section 7.

2. Matrix Relations

Relations involving all of the above triangles are defined in this section. An outline of a proof of the relations is mentioned. The invertibility of C_0 is also mentioned.

Multiplying by R, the matrix relations $M \cdot R = C$ and $C \cdot R = H$ are defined. In the latter relation, H is another infinite lower-triangular matrix where the first few terms are

$$H = \begin{bmatrix} 1 & & & & & \\ 3 & 1 & & & & \\ 10 & 6 & 1 & & & \\ 36 & 29 & 9 & 1 & & \\ 137 & 132 & 57 & 12 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The entries in the left most column, sequence $\{1, 3, 10, 36, 137, 543, \ldots\}$, count edge rooted polyhexes with n hexagons [9], [14]. These are graphs which are constructed by connecting n hexagons with certain restrictions. However, to remain within the context of lattice paths, we define the (n, k)th entry of H as the number of unit-step NSEWF lattice paths of length n and height k. The F denotes a forward unit-step. These lattice paths are 3-dimensional and they correspond to integer points (x, y, z) such that all paths begin at the origin, $(0, 0, 0)$, and never go below the (x, y) plane. The F steps are along the *y-axis*, and the height corresponds to the z value of (x, y, z) at the end of the path.

From all lattice path interpretations mentioned above and the associated matrix relations, we observe that right multiplication by R takes NS paths to NSE paths, NSE paths to NSEW paths, and NSEW paths to NSEWF paths. These path relations are illustrated as $C_0 \to M \to C \to H$ where the arrow means "goes to." In a combinatorial sense, R acts as a matrix transformation which transforms a selected set of unit-step lattice paths of length n and height k from 1-dimension to 2-dimensions to 3-dimensions. Likewise by left multiplication by P, the same transformation emerges since $P \cdot C_0 = M$, $P \cdot M = C$, and $P \cdot C = H$. The following proposition arises as a result of the matrix relations.

2.1. Proposition. *Given infinite lower-triangular matrices C_0, C, M, P and H, the following matrix relations are satisfied:*

$$\begin{array}{llcl} (a) & (C_0)^{-1} \cdot P \cdot C_0 & = & R \\ (b) & C_0 \cdot R = M & = & P \cdot C_0 \\ (c) & M \cdot R = C & = & P \cdot M \\ (d) & C \cdot R = H & = & P \cdot C \end{array}$$

Since the matrices are infinite it may not be obvious that the proposition is true. One way to prove the proposition is to consider that all of the matrices mentioned above are Riordan. *Riordan matrices* are infinite lower-triangular matrices made up of columns of the form $g(x) \cdot [f(x)]^i$ where $g(x) = 1 + g_1 x + g_2 x^2 + g_3 x^3 + \cdots$ corresponds to the left most column. For $i > 0$, the expression $g(x) \cdot [f(x)]^i$ corresponds to the *ith* column where $f(x) = 1x + f_2 x^2 + f_3 x^3 + f_4 x^4 + \cdots$. These columns are characterized as column generating functions (GFs) with integer coefficients f_i and g_i, and 1's along the main diagonal. A useful property of Riordan matrices is that they are invertible. Thus, the existence of $(C_0)^{-1}$ follows since C_0 is Riordan. Applying the Riordan matrix enumeration we can prove that the column GF associated with the left most column of H equals the GF associated with edge rooted polyhexes with n hexagons.

In the next two sections, the notions that motivate the correspondence between lattice paths and RNA secondary structures are mentioned. The relevant combinatorial aspects of RNA secondary structures are summarized. and then a combinatorial interpretation of R is given in terms of lattice paths.

3. RNA Secondary Structure

The single-stranded RNA molecule consists of a chain of base pairs derived from one of four bases (nucleotides): A (adenine), C (cytosine), G (guanine), and U (uracil) where A bonds with U, and G bonds with C. The linear sequence of such bases along the chain is defined as the *primary structure*. When an RNA molecule folds back on itself and forms new hydrogen bonds which form helical regions, the sequence is referred to as the secondary structure. As an example of secondary structure, we give the following RNA sequence s denoted as

$$s = CAGCAUCACAUCCGCGGGGUAAACGCU.$$

This sequence is referred to as a cloverleaf, in the biological literature, and is the secondary structure assumed by transfer RNA molecules. Two representations of s appear below in Fig. 1. In the figure we ignore the C, A, G and U and focus on the secondary structure.

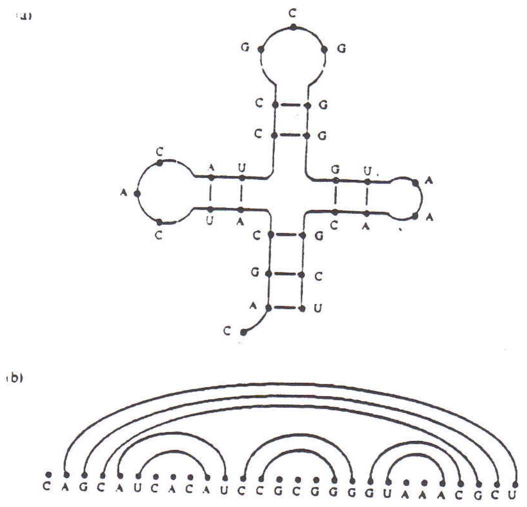


Fig. 1. Two representations of secondary structure.

In Fig. 1(a) the base pairs are indicated by dashes. In Fig. 1(b), the primary structure is given along the horizontal axis and the base pairs are shown as arcs. The above example, description and figure comes from Schmitt and Waterman [16]. The enumeration of secondary structures was studied, from a graph theoretic point of view, by Waterman [25]. He gives a graph theoretic definition of *secondary structure* as a planar graph defined on a set of n labeled points $\{1, 2, \ldots, n\}$. Thus, if $s(n)$ denotes the total number of secondary structures defined on n labeled points, then the associated recurrence relation is

$$(1) \qquad s(n+1) = s(n) + \sum_{j=1}^{n-1} s(j-1)s(n-j)$$

for $n \geq 2$, and $s(0) = s(1) = s(2) = 1$. The first few values of $s(n)$ for $n = 0, 1, \ldots, 6$ are $1, 1, 1, 2, 4, 8$, and 17. Donaghey [3] notes that these numbers can also be computed by the following sum

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

We call these numbers the RNA numbers. It turns out that these numbers are the same numbers as the entries in the left most column of R. The generating function derived from the recurrence relation is

$$(2) \qquad s(x) = \frac{(1-x+x^2) - \sqrt{(1+x+x^2)(1-3x+x^2)}}{2x^2}$$

Proofs of the recurrence relation, and generating function can be found in [10], [14] and [25].

4. Lattice Path Interpretation

Recursions for R are defined in this section by using the rule of formation of R. What we mean by rule of formation is a recursion which defines the way the entries of R are formed or computed. From the recursions, a combinatorial argument is proved showing that the entries of R denote the number of unit-step NSE* lattice paths of length n and height k. These NSE* lattice paths are explicitly defined later in this section. Then, the GF associated with the left most column of R is derived and shown to be equal to $s(x)$.

As examples of the way the elements of R are formed, we observe that the second column entry **6** is computed from $2+3+1$, and the left most column entry **8** is computed from $4+3+1$. The following illustrations:

$$(a)\quad \begin{array}{ccc} & & 1 \\ & & \swarrow \\ 2 & 3 & \\ \searrow & \downarrow & \\ & [\mathbf{6}] & \end{array} \qquad\qquad (b)\quad \begin{array}{ccc} & & 1 \\ & & \swarrow \\ & 3 & \\ & \swarrow & \\ 4 & & \\ \downarrow & & \\ [\mathbf{8}] & & \end{array}$$

depict the rule of formation of the entries shown in the above examples. These patterns continue to form all of R. In general, the (n, k)th entry of R is formed or computed recursively by the following recursions. For $n \geq 0$ and $k \geq 1$

$$(3) \qquad r(n+1, k) = r(n, k-1) + r(n, k) + r(n-1, k+1) + \cdots, \text{ and}$$

(4) $$r(n+1,0) = r(n,0) + r(n-1,1) + r(n-2,2) + \cdots$$

where $r(0,0) = 1$, and $r(n+1,k) = 0$ if $k > n+1$. Recursion 4 is defined for the left most column of R, and the associated column GF is denoted as $r(x)$. Recursion 3 is defined for the rest of the columns of R, and the associated kth column GF is denoted as $r(x) \cdot [f(x)]^k$ for $k > 0$. These associated GFs follow since R is a Riordan matrix. The GFs $r(x)$ and $f(x)$ are derived later in this section.

The NSE* lattice paths are now defined. These paths are also 2-dimensional and they have the same path restrictions as the NSE paths, mentioned in section 1, except for the additional restriction that consecutive N and S steps are **not** allowed. That is, a NSE* lattice path is a unit-step NSE lattice path which does not have any consecutive pair of NS steps. We combinatorially interpret the recursions in terms of these paths by letting $r(n,k)$ denote the number of unit-step NSE* lattice paths of length n and height k. Given the recursions and the interpretation, we can prove the following proposition.

4.1. Proposition. *For $r(0,0) = 1$, $n \geq 0$ and $k \geq 1$, $r(n+1,k)$ satisfy the following equations:*

$$\begin{aligned}
(a)\ r(n+1,k) &= \begin{cases} r(n,k-1) + \sum_{j\geq 0} r(n-j,k+j) \\ 0, \text{ if } k > n+1 \end{cases} \\
(b)\ r(n+1,0) &= \sum_{j\geq 0} r(n-j,j).
\end{aligned}$$

Proof. Suppose we have a unit-step NSE* lattice path of length n and height k. Then, to form a new path of length $(n+1)$ and height k consider the following cases. Case (i): if we have a path of length n and height $k-1$, then on the last step there is 1 choice for height $k-1$ (the N step). In this case, all paths whose last step is N are counted by $r(n,k-1)$. Case (ii): if we have a path of length n and height k, then on the last step there is also 1 choice for height k (the E step). In this case, all paths whose last step is E are counted by $r(n,k)$. Case (iii): if we have a path of length $(n-1)$ and height $k+1$, then the last possible sequence of steps for height $k+1$ is ES (east, south). In this particular case, all paths whose last sequence of steps is ES are counted by $r(n-1,k+1)$. Case (iv): if we continue and have a path of length $(n-j)$ and height $k+j$, then the last possible sequence of steps for height $k+j$ is ES^j (east, south j-times). These sequences occur since there are no NS steps. Here, all paths whose last sequence of steps is ES^j are counted by $r(n-j,k+1)$. Combining all of the cases give all possible ways of forming a new $(n+1)$st path of height k. Applying the addition principle, recursion (a) is proved. Recursion (b) is proved by similar reasoning.□

Thus, the combinatorial interpretation of R is proved. Also by similar reasoning, we can prove the lattice path interpretations defined above for H, C, M and C_0 [12].

Explicit GFs for $r(x)$ and $f(x)$ are now derived. Recall that R is Riordan, so each column is of the form $r(x) \cdot [f(x)]^k$. By the rule of formation of R, the kth column GF is defined as

$$r \cdot f^k = x\left(rf^{k-1} + rf^k + xrf^{k+1} + x^2rf^{k+2} + \cdots\right).$$

Solving for f, we find $f = x + xf + x^2f^2 + x^3f^3 + \cdots$. Since the sum is a geometric series, we obtain $f = xf^2 + x(1-x)f + x$. Now, solving f in terms of $f(x)$ and simplifying we obtain $f(x) = x \cdot s(x)$. Similarly, the left most column GF is defined as

$$r = 1 + x\left(r + xrf + x^2rf^2 + x^3rf^3 + \cdots\right).$$

Simplifying this equation, we find that r expressed in terms of $r(x)$ is the same GF as equation 2, i.e., $r(x) = s(x)$. Thus, we have explicit GF representations for $r(x)$ and $f(x)$.

The GF $r(x)$ is associated with the unit-step NSE* lattice paths of length n and height $k = 0$. As a result of $k = 0$, $r(x)$ corresponds to the paths that return to the *x-axis*. In general, as a consequence of deriving $r(x)$ and $f(x) = x \cdot r(x)$, we can define for $k \geq 0$ the GF $\frac{r(x)}{1-x \cdot r(x)}$ whose associated sequence is $\{1, 2, 4, 9, 21, 50, \ldots\}$. These numbers are the values of the row sums of R and they count the total number of unit-step NSE* lattice paths of length n and height k [12]. However, of all NSE* paths, we are only concerned with those paths that return to the *x-axis*.

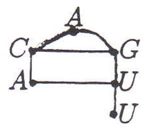
5. RNA Correspondence with Lattice Paths

A one-to-one correspondence between NSE* lattice paths and RNA secondary structures is constructed in this section. Since $r(x)$ and $s(x)$ denote the same GF, the order of the set of unit-step NSE* lattice paths of length n and height $k = 0$ and the order of the set of RNA secondary structures of length n are the same. So, the *nth* terms are such that $|s(n)| = |r(n, 0)|$. We can now state the correspondence theorem and the corresponding constructive proof.

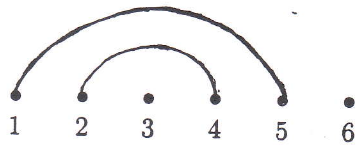
5.1. Theorem. *There is a one-to-one correspondence between the set of RNA secondary structures of length n and the set of unit-step NSE* lattice paths of length n and height $k = 0$.*

Proof. First, we show how the unit steps of a NSE* lattice path of length n and height $k = 0$ are assigned. To establish the required correspondence, let s be a secondary structure of length n. Then, for a given s list the RNA sequence as a sequence of integers increasing in order from left to right as a primary structure along a horizontal axis and denote base pairs (or bondings) as arcs. Now, consider whether an integer is paired or unpaired. If an integer (or base) k is unpaired label the *kth* integer as an E step. Otherwise, if a base pairing of integers occurs where an integer i represents a pairing with a larger integer j, then label the *ith* integer as an N step and the *jth* integer as an S step. Therefore, the correspondence is setup according to the rules $k \to E$, $i \to N$, and $j \to S$ where the arrow means "corresponds to." The definition of secondary structure ensures that the NSE* paths do not have any consecutive pair of NS steps since no two adjacent points (i.e., with labels i and $i + 1$) can be connected by an arc, and no two arcs may intersect. Also, the paths are of height zero since each N step is paired with an S step.

As an example of the correspondence consider the sequence $\varsigma = ACAGUU$ where A bonds with U and C bonds with G. The bondings are indicated by dashes in the following graph:



Now list ς as a sequence of integers along a horizontal axis as described above to obtain the following linear representation:



Applying the rules of the correspondence, the integers are assigned to the path steps as follows: 1 $\longrightarrow$N, 2 $\longrightarrow$N, 3 $\longrightarrow$E, 4 $\longrightarrow$S, 5 $\longrightarrow$S, and 6 $\longrightarrow$E. From the correspondence, we obtain the unit-step NSE* path NNESSE which has length 6 and height 0.

The correspondence is reversible. Thus, the correspondence is one-to-one and the theorem is proved.□

6. **First Moments**

The first moments (weighted row sums) of R are computed in this section. These moments can be used, in a combinatorial sense, to compute the average distance from the origin of all unit-step NSE* lattice paths of length n.

Multiplying R on the right side by the column vector $V^T = \{1, 2, 3, 4, 5, 6, 7, \ldots\}^T$, we make the following observation:

$$R \cdot V^T = \begin{array}{llllll} 1 \cdot 1 & & & & & = 1 \\ 1 \cdot 1 & +1 \cdot 2 & & & & = 3 \\ 1 \cdot 1 & +2 \cdot 2 & +1 \cdot 3 & & & = 8 \\ 2 \cdot 1 & +3 \cdot 2 & +3 \cdot 3 & +1 \cdot 4 & & = 21 \\ 4 \cdot 1 & +6 \cdot 2 & +6 \cdot 3 & +4 \cdot 4 & +1 \cdot 5 & = 55 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

Then, we conjecture that the first moments of R are defined by

$$R_n = F_{2n-1} = \sum_{k=0}^{n} k \cdot r(n-1, k-1) \text{ for } n \geq 1,$$

where F_{2n-1} denotes the alternating Fibonacci numbers 1, 3, 8, 21, 55, ... The conjecture can be proved using the Riordan matrix technique mentioned in section 2. The proof is outlined as follows. A Riordan matrix can be represented as a pair $[g(x), f(x)]$ where $g(x)$ and $f(x)$ are defined in section 2. A compositional functional $B(x)$ is obtained when a Riordan pair is multiplied on the right side by a GF denoted as $A(x)$. The GF $A(x)$ is associated with an appropriate column vector. Thus, $B(x)$ is defined as

$$\begin{aligned} B(x) &= [g(x), f(x)] * A(x) \\ &= g(x) \cdot A(f(x)) \end{aligned}$$

where the symbol ' * ' denotes Riordan matrix multiplication. The Riordan pair associated with R is the pair $[r(x), x \cdot r(x)]$ where $r(x)$ is the GF defined by equation 2 of section

3. The GF associated with the column vector V^T is defined by $v(x) = (1-x)^{-2}$. Then by Riordan multiplication, a compositional functional $F(x)$ is obtained and defined as

$$\begin{aligned} F(x) &= [r(x), x \cdot r(x)] * v(x) \\ &= r(x) \cdot v(x \cdot r(x)) \\ &= \tfrac{1}{1-3x+x^2}. \end{aligned}$$

Therefore, $F(x) = \frac{1}{1-3x+x^2}$ which is the GF for the alternating Fibonacci numbers [1], [5]. This proves the conjecture.

7. Other Applications

In the previous sections, we showed that the R triangle has matrix properties that are of combinatorial interest, and an application related to the RNA numbers. In this section, we mention other appearances of the RNA numbers related to noncrossing partitions, and the Narayana numbers.

The Narayana numbers are also of combinatorial interest and are defined as $N(n,k) = \frac{1}{n}\binom{n}{k}\binom{n}{k-1}$ for $n \geq 1$ and $k \geq 1$. These numbers can be put into infinite lower-triangular matrix form, denoted as N. The triangle N is not Riordan, and the first few terms are

$$N = \begin{bmatrix} 1 & & & & & \\ 1 & 1 & & & & \\ 1 & 3 & 1 & & & \\ 1 & 6 & 6 & 1 & & \\ 1 & 10 & 20 & 10 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

It is known that the row sums of N are the nth Catalan numbers C_n [11]. A combinatorial interpretation of N is that $N(n,k)$ counts the number of Dyck paths of length $2n$ with k peaks. A *Dyck* path is a path in the first quadrant, which begins at the origin, ends at $(2n, 0)$, and consists of north-east and south-east steps. We note that the diagonal slices $1, 1, 1+1, 1+3, 1+6+1, \ldots$ of the N triangle give the first few RNA numbers. This can be proved by using generating functions, where the bivariate GF associated with the Narayana numbers is noted by Stanley [23].

To find a combinatorial relation between the RNA numbers and noncrossing partitions consider the set $[n] := \{1, 2, \ldots, n\}$. A partition π of $[n]$ is said to be *noncrossing* if $1 \leq a < b < c < d \leq n$ and if B_1 and B_2 are blocks of π such that $a, c \in B_1$ and $b, d \in B_2$, then $B_1 = B_2$. That is, given that the conditions are satisfied, a, b, c and d are all in the same block. As an example of a noncrossing partition of $[6] = \{1, 2, 3, 4, 5, 6\}$ consider $\pi = 1\ 5\ /\ 2\ 4\ /\ 3\ /\ 6$ where the slashes separate the blocks. The linear representation of π is illustrated above in section 5, where successive elements in the same block are joined by arcs.

Following Simion and Ullman [20], a word w of length $n-1$ over the alphabet $\{b, e, l, r\}$ can be associated with a noncrossing partition π. See the reference for detailed definitions of each letter in the alphabet. By eliminating the letter r and the consecutive pair of letters b and e from any potential word, another word w^* can be defined over the alphabet $\{b, e, l, \}$. The word w^* can also be associated with a noncrossing partition π.

For example, noncrossing partition $\pi = 1\ 5\ /\ 2\ 4\ /\ 3\ /\ 6$ is associated with the word $w^*(\pi) = bbleel$. A one-to-one correspondence can be constructed between the subset of noncrossing partitions associated with words w^* and the set of NSE* lattice paths. The correspondence is setup according to the following rules where $b \to N$, $e \to S$, and $l \to E$. Thus, via NSE* lattice paths, a relation is established between noncrossing partitions and the RNA numbers.

Several appearances of the RNA numbers related to planar trees are in [3]. Other combinatorial applications and interpretations of the RNA numbers are discussed in [2], [12] and [24].

Acknowledgment

This paper was developed during research on the author's dissertation. The author wishes to thank his advisor, Professor Louis W. Shapiro, and the combinatorics research group of Howard University for the many useful discussions pertaining to the material presented here.

References

[1] E.F. Beckenbach, *Applied Combinatorial Mathematics*, Wiley, New York, 1964.

[2] M. Bernstein and N.J.A. Sloane, *Some canonical sequences of integers*, Linear Algebra and its Applications **226-228** (1995) 57-72.

[3] R. Donaghey, *Automorphisms on Catalan trees and bracketings*, J. Combinatorial Theory, Series B **29** (1980) 75-90.

[4] R. Donaghey and L. W. Shapiro, *Motzkin numbers*, J. Combinatorial Theory, Series A **23** (1977) 291-301.

[5] R.L. Graham, D.E. Knuth and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, Reading, MA, 1989.

[6] R.K. Guy, *Catwalks, sandsteps & Pascal pyramids*, preprint.

[7] R.K. Guy, C. Krattenthaler and B.E. Sagan, *Lattice paths, reflections, & dimension-changing bijections*, ARS Combinatoria **34** (1992) 3-15.

[8] P. Hanlon, *Counting interval graphs*, Trans. Am. Math. Soc. **272** (1982) 383-426.

[9] F. Harary and R. Read, *The enumeration of tree-like polyhexes*, Proc. Edinburgh Math. Soc. **17** (1972) 1-13.

[10] J.A. Howell, T.F. Smith and M.S. Waterman, *Computation of generating functions for biological molecules*, SIAM J. Appl. Math. **39** (1980) 119-133.

[11] T.V. Narayana, *Lattice Path Combinatorics with Statistical Applications*, University of Toronto Press, 1979.

[12] A. Nkwanta, *Lattice paths, generating functions, and the Riordan group*, Ph.D. Dissertation, Howard University, in progress, 1996.

[13] P. Peart and L. Woodson, *Triple factorization of some Riordan matrices*, Fibonacci Quart. **31** (1993) 121-128.

[14] F.S. Roberts, *Applied Combinatorics*, Prentice Hall, New Jersey, 1984.

[15] B. Sands, *Problem 1517**, Crux Mathematicorum **17** (1991) 119-122.

[16] W.R. Schmitt and M.S. Waterman, *Linear trees and RNA secondary structures*, Discrete Appl. Math. **51** (1994) 317-323.

[17] L.W. Shapiro, *A Catalan triangle*, Discrete Math. **14** (1976) 83-90.

[18] L.W. Shapiro, S. Getu, W.J. Woan and L. Woodson, *The Riordan Group*, Discrete Appl. Math. **34** (1991) 229-239.

[19] L.W. Shapiro, W.J. Woan and S. Getu, *Runs, slides and moments*, SIAM J. Alg. Disc. Meth. **4** (1983) 459-466.

[20] R. Simion and D. Ullman, *On the structure of the lattice of noncrossing partitions*, Discrete Math. **98** (1991) 193-206.

[21] N.J.A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press, San Diego, 1995.

[22] R. Sprugnoli, *Riordan arrays and combinatorial sums*, Discrete Math. **132** (1994) 267-290.

[23] R. Stanley, *Enumerative combinatorics*, Vol. II, Cambridge University Press, preliminary version, 1996.

[24] P.R. Stein and M.S. Waterman, *On some new sequences generalizing the Catalan and Motzkin numbers*, Discrete Math. **26** (1979) 261-272.

[25] M.S. Waterman, *Secondary structure of single-stranded nucleic acids*, Adv. Math. **I** (suppl.) (1978) 167-212.

Department of Mathematics, Howard University, Washington, D.C. 20059
E-mail address: nkwanta@scs.howard.edu