AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **84(2)** (2022), Pages 325–340

Set partitions and non-crossing partitions with ℓ-neighbors and ℓ-isolated elements

Beáta Bényi

Department of Hydraulic Engineering
University of Public Service, Baja
Hungary
benyi.beata@uni-nke.hu

Toufik Mansour

Department of Mathematics
University of Haifa, 3498838 Haifa
Israel
tmansour@univ.haifa.ac.il

José L. Ramírez

Departamento de Matemáticas
Universidad Nacional de Colombia, Bogotá
Colombia
jlramirezr@unal.edu.co

Abstract

In this paper we introduce the notion of an ℓ-neighbor element of set partitions, that is, an element a in a block that contains $\ell+1$ consecutive elements, among which is a. Elements that are not ℓ-neighbors are called ℓ-isolated elements. We explore combinatorial results to study these new statistics over the set partitions. In particular, we use combinatorial arguments, recurrence relations, and generating functions to describe our results. We also discuss possible relations with ribonucleic acids (RNA) structures.

1 Introduction

A *set partition* of a set $[n] := \{1, 2, \ldots, n\}$ is a collection of non-empty disjoint subsets, called *blocks*, whose union is $[n]$. Let $\left\{ {n \atop k} \right\}$ denote the number of set partitions of $[n]$ into k non-empty blocks. This sequence is called the *Stirling numbers of the*

ISSN: 2202-3518

second kind. Let $\Pi(n,k)$ denote the set of partitions of $[n]$ having k blocks. Suppose $\pi \in \Pi(n,k)$ is represented as $\pi = B_1/B_2/\cdots/B_k$, where B_i denotes the i-th block, with $\min(B_1) < \min(B_2) < \cdots < \min(B_k)$. The graph on the vertex set $[n]$ whose edge set consists of arcs connecting the elements of each block in numerical order is called the *graph representation* of π. For example, in Figure 1 we depict the graph representation of the set partition

$$\pi = \{1,2,3\} / \{4\} / \{5\} / \{6,7,8,9\} / \{10,11,13\} / \{12,14\}. \tag{1}$$

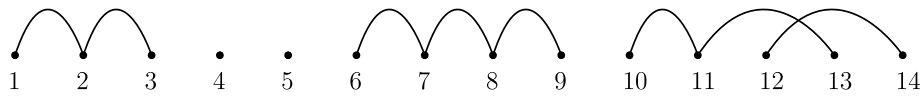

Figure 1: Graph representation of π.

Let $\pi = B_1/B_2/\cdots/B_k$ be a partition of $[n]$ and $a \in B_i$, $1 \le i \le k$. If ℓ is a positive integer then we say that a is an *ℓ-neighbor* if there is a subset of B_i, with $\ell+1$ consecutive elements that contain the element a. For example, the partition π defined in (1) contains nine 1-neighbors, namely, 1, 2, 3, 6, 7, 8, 9, 10, 11, seven 2-neighbors, 1, 2, 3, 6, 7, 8, 9, four 3-neighbors, 6, 7, 8, 9, and does not contain ℓ-neighbors for $\ell \ge 4$.

Note that if an element is an ℓ-neighbor then it is an ℓ'-neighbor for all $\ell' \le \ell$. The (maximal) sequence of consecutive elements contained in the same block is called *chain*. The *length* of the chain is the number of its elements. Let ℓ^* denote the length of the longest chain in a partition. For example, for the partition π defined in (1) $\ell^* = 4$?. Another easy observation is that a partition contains at least ℓ^* ℓ-neighbors for all $\ell < \ell^*$ and no ℓ-neighbors for $\ell \ge \ell^*$.

If an element a is not a 1-neighbor, then we say that a is an *isolated singleton*. This definition was recently studied by Munagi [10]. For example, the elements $4, 5, 12, 13, 14$ are isolated singletons. In general an element a is *ℓ-isolated* if it is not an ℓ-neighbor. Elements that are contained in a block of size less than $\ell+1$ are ℓ-isolated in an obvious sense. The partitions of $[n]$ into k non-empty blocks, each of size less than or equal to ℓ, are counted by the $\left\{ {n \atop k} \right\}_{\le \ell}$ numbers (cf. [1]), this is the number of partitions having only trivial ℓ-isolated elements. This definition was also considered by Mansour and Munagi in [9]. In this paper we find additional results, in particular we consider the notion of ℓ-neighbors for non-crossing partitions, and we also give different proofs and interpretations to many of the results.

Let $N_{\ell,r}(n,k)$ denote the number of partitions of $[n]$ into k blocks containing r ℓ-neighbors. Analogously, let $L_{\ell,r}(n,k)$ denote the number of partitions of $[n]$ into k blocks containing r ℓ-isolated elements. The sequence $L_{1,r}(n,k)$ coincides with the sequence $g_r(n,k)$ studied by Munagi in [10]. Since for a fixed ℓ an element a is either an ℓ-neighbor or ℓ-isolated, the relation $N_{\ell,r}(n,k) = L_{\ell,n-r}(n,k)$ holds. By definition we take $L_{0,0}(n,k) = \left\{ {n \atop k} \right\}$. If there is an ℓ-neighbor a in the partition then there are at least ℓ other ℓ-neighbors in the partitions since an ℓ-neighbor is contained in a chain of length $\ell+1$. Hence, $N_{\ell,r}(n,k)=0$, if $r\leq\ell$.

In the present article, we use combinatorial arguments, generating functions, and recurrence relations to calculate the sequence $L_{\ell,r}(n,k)$. Afterward, we find explicit combinatorial formulas in terms of this new sequence. Finally, we give analogous results in the context of non-crossing partitions, that is, set partitions such that none of the edges in the graph representation cross. We also give a connection between the ℓ-neighbor elements and the ribonucleic acids (RNA) structures.

2 Set partitions with no ℓ-isolated singletons

This section discusses the sequence $L_{\ell,0}(n,k)$, which enumerates set partitions with no ℓ-isolated singletons. In particular, Theorem 2.1 and Theorem 2.2 give recurrence relations to calculate this sequence, and Theorem 2.3 gives an explicit expression in terms of Stirling numbers of the second kind. We provided three different proofs for this expression.

Theorem 2.1. *If $n\geq\ell+3$ and $2\leq k<n$, then*

$$L_{\ell,0}(n,k)=L_{\ell,0}(n-\ell-1,k-1)+L_{\ell,0}(n-1,k)+(k-1)L_{\ell,0}(n-\ell-1,k),$$

with the initial values $L_{\ell,0}(n,k)=0$ for $1\leq k,n\leq\ell$, $L_{\ell,0}(n,k)=\delta_{k,1}$ for $\ell+1\leq n\leq 2\ell+1$, $L_{\ell,0}(n,1)=1$ for $n>\ell$.

Proof. The left-hand side counts the number of set partitions of $[n]$ into k blocks with no ℓ-isolated singletons. Now we consider the right-hand side. Let π be any set partition in $\Pi_{n,k}$ with no ℓ-isolated singletons. There are three options. The first case is that $\{n-\ell,\dots,n-1,n\}$ is a block of π. The remaining $n-\ell-1$ elements create a partition with $k-1$ blocks, hence, there are $L_{\ell,0}(n-\ell-1,k-1)$ such partitions. The second case is when n is contained in a block in that all the elements $n-(\ell+1),n-\ell,\dots,n-1$ are also contained. In this case, if we delete the element n, the remaining partition of the $n-1$ elements have any ℓ-isolated elements, hence, there are $L_{\ell,0}(n-1,k)$ such partitions. The third case is when the block that contains n, does not contain the element $n-(\ell+1)$. We count these partition as follows. The number of partitions of the elements $[n-(\ell+1)]$ is $L_{\ell,0}(n-\ell-1,k)$. We can add now the chain $n-\ell,\dots,n$ to any block that does not contain the element $n-(\ell+1)$ in order to obtain a partition with the required property. This gives $(k-1)L_{\ell,0}(n-\ell-1,k)$ possibilities. □

Theorem 2.2. *If $n\geq\ell+3$ and $2\leq k<n$, then*

$$L_{\ell,0}(n,k)=\sum_{i\geq 0}L_{\ell,0}(n-\ell-i-1,k-1)+(k-1)L_{\ell,0}(n-\ell-i-1,k),$$

with the initial values $L_{\ell,0}(n,k)=0$ for $1\leq k,n\leq\ell$, $L_{\ell,0}(n,k)=\delta_{k,1}$ for $\ell+1\leq n\leq 2\ell+1$, $L_{\ell,0}(n,1)=1$ for $n>\ell$.

Proof. Assume that the length of the chain having the last element n is of length $\ell+1+i$ $(i \geq 0)$, i.e, the consecutive elements $n-(\ell+i), \ldots, n$ are contained in the same block. This chain is itself a block, or there are other elements in its block. In the first case, the remaining $n-\ell-i-1$ elements create a partition with $k-1$ block (in $L_{\ell,0}(n-\ell-i-1,k-1)$ ways), while in the second case we can add this chain to any of the blocks in a partition of $\Pi_{n-\ell-i-1,k}$ without ℓ-isolated elements that does not contain $n-\ell-i-1$ itself. There are $(k-1)L_{\ell,0}(n-i-\ell-1,k)$ possibilities in this case. □

Next we generalize the explicit expression Munagi [10] provided for the number of partitions without isolated elements.

$$L_{1,0}(n,k) = \sum_{j\geq 1} \binom{n-j-1}{j-1} \left\{ {j-1 \atop k-1} \right\}.$$

We present three different proofs for Theorem 2.3.

Theorem 2.3. *We have the combinatorial identity*

$$L_{\ell,0}(n,k) = \sum_{j=k}^{\lfloor n/(\ell+1) \rfloor} \binom{n-1-j\ell}{j-1} \left\{ {j-1 \atop k-1} \right\}.$$

First proof of Theorem 2.3. We will prove that the combinatorial sum satisfies the recurrence of Theorem 2.1 and the same initial values. In fact, from the recurrence relation for the Stirling numbers $\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k\left\{ {n-1 \atop k} \right\}$, with the initial conditions $\left\{ {0 \atop 0} \right\} = 1$ and $\left\{ {n \atop 0} \right\} = 0$ for $n > 0$, we obtain that

$$\begin{aligned}
&L_{\ell,0}(n-1,k) + L_{\ell,0}(n-\ell-1,k-1) + (k-1)L_{\ell,0}(n-\ell-1,k) \\
&\qquad = \sum_{j\geq 1} \binom{n-2-j\ell}{j-1} \left\{ {j-1 \atop k-1} \right\} + \sum_{j\geq 1} \binom{n-2-(j+1)\ell}{j-1} \left\{ {j-1 \atop k-2} \right\} \\
&\qquad \quad + (k-1) \sum_{j\geq 1} \binom{n-2-(j+1)\ell}{j-1} \left\{ {j-1 \atop k-1} \right\} \\
&\qquad = \sum_{j\geq 1} \binom{n-2-j\ell}{j-1} \left\{ {j-1 \atop k-1} \right\} + \sum_{j\geq 1} \binom{n-2-(j+1)\ell}{j-1} \left\{ {j \atop k-1} \right\} \\
&\qquad = \sum_{j\geq 0} \binom{n-2-(j+1)\ell}{j} \left\{ {j \atop k-1} \right\} + \sum_{j\geq 1} \binom{n-2-(j+1)\ell}{j-1} \left\{ {j \atop k-1} \right\} \\
&\qquad = \sum_{j\geq 0} \binom{n-1-(j+1)\ell}{j} \left\{ {j \atop k-1} \right\} \\
&\qquad = L_{\ell,0}(n,k).
\end{aligned}$$

□

Second proof of Theorem 2.3. Define $L_\ell^k(x) = \sum_{n\geq 0} L_{\ell,0}(n,k)x^n$. Then Theorem 2.1 gives that $L_\ell^1(x) = \frac{x^{\ell+1}}{1-x}$, $L_\ell^2(x) = \frac{x^{2\ell+2}}{(1-x)(1-x-x^{\ell+1})}$, and

$$L_\ell^k(x) = \frac{x^{\ell+1}}{1-x-(k-1)x^{\ell+1}} L_\ell^{k-1}(x),$$

for all $k \geq 3$. Hence, by induction on k, we have

$$L_\ell^k(x) = \frac{x^{k(\ell+1)}}{\prod_{j=1}^k (1-x-(j-1)x^{\ell+1})} = \frac{(x^{\ell+1}/(1-x))^k}{\prod_{j=1}^{k-1}(1-jx^{\ell+1}/(1-x))}.$$

Using the fact $\frac{x^k}{\prod_{j=1}^k(1-jx)} = \sum_{n\geq k} {n \brace k} x^n$, we obtain

$$L_\ell^k(x) = \sum_{j\geq k} {j-1 \brace k-1} \frac{x^{j(\ell+1)}}{(1-x)^j},$$

which, by $\frac{1}{(1-x)^j} = \sum_{i\geq 0}\binom{j-1+i}{j-1}x^i$, implies

$$L_\ell^k(x) = \sum_{j\geq k}\sum_{i\geq 0}\binom{j-1+i}{j-1}{j-1 \brace k-1} x^{j(\ell+1)+i}.$$

Hence, by comparing the coefficient of x^n in both sides, we obtain

$$L_{\ell,0}(n,k) = \sum_{j=k}^{n/(\ell+1)} \binom{n-1-j\ell}{j-1}{j-1 \brace k-1}.$$

□

Third proof of Theorem 2.3. Let us consider first an example. Within the summand, for $j = k$, we have $\binom{n-1-k\ell}{k-1}$. This is the number of ℓ-isolated partitions into k blocks such that each block is a chain.

How can we get the formula $\binom{n-1-k\ell}{k-1}$? First, consider the subsets of consecutive ℓ elements:

$$H_1 = \{1, 2, \ldots, \ell\}, \quad H_2 = \{\ell+1, \ldots, 2\ell\}, \ldots, H_k = \{(k-1)\ell+1, \ldots, k\ell\}.$$

Now we modify these blocks by sliding the limits in the following way: consider the sequence of the remaining $n - k\ell$ elements and separate the sequence into k sections by inserting $k-1$ bars between the elements. In other words, take a composition of the $n-k\ell$ elements into k parts, (the order of the parts matters) $C = c_1+c_2+\cdots+c_k$. For example,

$$\cdot\cdot \mid \cdot\cdot\cdot \mid \cdot \mid \cdot\cdot = 2+3+1+2$$

Then define the new blocks as

$$\begin{aligned}
H'_1 &= H_1 \cup \{\ell+1,\dots,\ell+c_1\} = \{1,2,\dots,\ell,\ell+1,\dots\ell+c_1\}\\
H'_2 &= H_2^{\to c_1} \cup \{2\ell+c_1+1,\dots,2\ell+c_1+c_2\} = \{\ell+c_1+1,\dots,2\ell+c_1+c_2\}\\
\vdots &= \vdots\\
H'_i &= H_i^{\to c_1+\cdots+c_{i-1}} \cup \{i\ell+c_1+\cdots+c_{i-1}+1,\dots,i\ell+c_1+\cdots+c_i\}\\
&= \{(i-1)\ell+c_1+\cdots+c_{i-1}+1,\dots,i\ell+c_1+\cdots+c_i\},
\end{aligned}$$

where $H^{\to p}$ denotes the set where we add to each element of the set H the value p.

Note that since there are k parts in the composition, (there are dots between the bars and before and after), each set H_i contains at least $\ell+1$ consecutive elements, hence, in the partition $H'_1,\dots,H'_k$ there are no ℓ-isolated elements.

Next, let j be the number of chains in the partition. Since the partition is ℓ-isolated each chain has to be of length at least $\ell+1$. Consider again first the blocks $H_i = \{(i-1)\ell,\dots,i\ell\}$ (as they would be distinct elements) and partition these into j blocks such that two consecutive H_i do not come into the same block ($a,b\in B_i$ implies that $|a-b|>1$). It is known [12] that the number of such partitions is $\left\{ {n-1 \atop j-1} \right\}$. □

Denote by $L_{\ell,0}(n)$ the total number of partitions of $[n]$ that contain no ℓ-isolated elements, that is, $L_{\ell,0}(n) = \sum_{k\ge1} L_{\ell,0}(n,k)$. The total number of set partitions of $[n]$ is counted by the Bell numbers B_n, that is, $B_n = \sum_{k=0}^{n} \left\{ {n \atop k} \right\}$.

Corollary 2.4. *For $\ell \ge 0$*

$$L_{\ell,0}(n) = \sum_{j\ge1} \binom{n-1-j\ell}{j-1} B_{j-1},$$

where B_n are the Bell numbers.

In Table 1 we show the first few values of the sequence $L_{\ell,0}(n)$.

$\ell \backslash n-\ell$	1	2	3	4	5	6	7	8	9	10	11	12
$\ell=0$	1	2	5	15	52	203	877	4140	21147	115975	678570	4213597
$\ell=1$	1	1	2	3	6	11	23	47	103	226	518	1200
$\ell=2$	1	1	1	2	3	4	7	12	19	33	59	102
$\ell=3$	1	1	1	1	2	3	4	5	8	13	20	29
$\ell=4$	1	1	1	1	1	2	3	4	5	6	9	14

Table 1: Values of $L_{\ell,0}(n)$ for $\ell+1 \le n \le 12+\ell$, $\ell = 0,1,\dots,4$.

Notice that the array $[L_{\ell,0}(n)]_{0\le\ell,\ell+1\le n}$ coincides with the array A211700 in [15]. The interpretation given there is the corresponding set of words, that we obtain by the trivial encoding of our partitions by words: w_i is the index of the block that contains i (taking the canonical order of the blocks). Example: $\{1,4,5\} / \{2,3,9\} / \{6,8\} / \{7\} = 01100232$.

3 Set partitions with only ℓ-isolated singletons

We consider now the other extreme case, partitions in which all the elements are ℓ-isolated. We also give the recurrence relation and the explicit formula for $L_{\ell,n}(n,k)$.

Proposition 3.1. *Let n,k,ℓ be integers with $n \geq \ell+1$, $2 \leq k < n$. Then*

$$L_{\ell,n}(n,k) = \sum_{i=1}^{\ell} L_{\ell,n-i}(n-i,k-1) + (k-1)\sum_{i=1}^{\ell} L_{\ell,n-i}(n-i,k),$$

where $L_{\ell,n}(n,k) = \left\{ {n \atop k} \right\}$ for $1 \leq n,k \leq \ell$ and $L_{\ell,n}(n,1) = 0$ for $n > \ell$.

Proof. The left-hand side counts the number of set partitions of $[n]$ into k block with exactly n ℓ-isolated elements. Now we consider the right-hand side. Let π be any set partition in $\Pi_{n,k}$ with n ℓ-isolated elements. So, we can do the following construction: either n (the last element) forms a block (possibly a singleton) with the elements $n-i,\dots,n-2,n-1$, for $i=1,\dots,\ell-1$ or n is in a block of size at least 2 with the elements $n-i+1,\dots,n-2,n-1$ but not with the element $n-i$, for $i=1,\dots,\ell-1$. In the first case, it is clear that there are $\sum_{i=1}^{\ell} L_{\ell,n-i}(n-i,k-1)$, while in the second case, have $k-1$ blocks that do not contain the element $n-i$. Therefore, there are $(k-1)\sum_{i=1}^{\ell} L_{\ell,n-i}(n-i,k)$ options. $\square$

For example, for $\ell = 5$ we obtain the following values

$$[L_{5,n}(n,k)]_{n,k\geq 1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 7 & 6 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 15 & 25 & 10 & 1 & 0 & 0 & 0 & 0 \\ 0 & 31 & 90 & 65 & 15 & 1 & 0 & 0 & 0 \\ 0 & 61 & 301 & 350 & 140 & 21 & 1 & 0 & 0 \\ 0 & 120 & 963 & 1701 & 1050 & 266 & 28 & 1 & 0 \\ 0 & 236 & 3004 & 7766 & 6951 & 2646 & 462 & 36 & 1 \end{pmatrix}.$$

Munagi [10] noted that the case $\ell = 1$ gives $L_{1,n}(n,k) = \left\{ {n-1 \atop k-1} \right\}$. In Theorem 3.2 we generalize this result.

Theorem 3.2. *For $\ell \geq 1$, we have*

$$L_{\ell,n}(n,k) = \sum_{j\geq 1} f_\ell(n,j) \left\{ {j-1 \atop k-1} \right\},$$

where $f_\ell(n,j)$ counts the number of compositions of n with j parts in $\{1,2,\dots,\ell\}$.

Proof. The sequence $L_{\ell,n}(n,k)$ counts the number of partitions into k blocks where each element is ℓ-isolated, i.e., each chain is at most of length ℓ. To obtain such a partition, take a composition of n with j parts in $\{1, 2, \ldots, \ell\}$, denoted it by $\sigma_1 \cdots \sigma_j$. By definition these compositions are counted by $f_\ell(n, j)$. Then, take a set partition of $[j]$ into k blocks such that no two consecutive elements are in the same block. This can be done in $\left\{ {j-1 \atop k-1} \right\}$ ways. Summing over $j \geq 1$, we obtain the desired result. $\square$

It is well-known that the generating function of the sequence $f_\ell(n, j)$ is given by (cf. [8])

$$\sum_{n \geq 0} f_\ell(n, j) x^n = \left(\sum_{i=1}^{\ell} x^i \right)^j .$$

In particular, $f_2(n, j)$, the number of compositions with parts 1 or 2 is $\binom{j}{n-j}$. Hence, the number of partitions where there are at most 2 consecutive elements in each block is $L_{2,n}(n, k) = \sum_{j \geq 1} \binom{j}{n-j} \left\{ {j-1 \atop k-1} \right\}$.

4 The general case

In this section we give a recurrence relation for the sequence $L_{\ell,r}(n, k)$, i.e., for the number of partitions of $[n]$ into k non-empty blocks with exactly r ℓ-isolated elements.

Theorem 4.1. *Let n, k, r, ℓ be integers with $n \geq \ell + 2$, $2 \leq k < n$, $1 \leq r < n$. Then*

$$\begin{aligned} &L_{\ell,r}(n, k) \\ &= L_{\ell,r}(n-1, k) - \left(\sum_{j=1}^{\ell} L_{\ell,r-j}(n-1-j, k-1) + (k-1) \sum_{j=1}^{\ell} L_{\ell,r-j}(n-1-j, k) \right) \\ &\qquad + L_{\ell,r}(n-\ell-1, k-1) + (k-1) L_{\ell,r}(n-\ell-1, k) \\ &\qquad\qquad + \sum_{j=1}^{\ell} L_{\ell,r-j}(n-j, k-1) + (k-1) \sum_{j=1}^{\ell} L_{\ell,r-j}(n-j, k), \end{aligned}$$

where $L_{\ell,r}(n, 1) = 0$ for $\ell < r$ and $L_{\ell,r}(n, 1) = \delta_{n,r}$ for $r \leq \ell$, $L_{\ell,r}(n, k) = 0$ if $1 \leq n \leq \ell + r$ and $n \neq r$, $L_{\ell,r}(n, n) = \delta_{n,r}$, $L_{\ell,n}(n, k)$ is as in Proposition 3.1, $L_{\ell,0}(n, k)$ is as in Theorem 2.1, and $L_{\ell,r}(n, k) = 0$ for $r, \ell < 0$.

Proof. The left-hand side counts the number of set partitions of $[n]$ into k block with r ℓ-isolated elements. Now we consider the right hand side. Let π be any set partition in $\Pi_{n,k}$ with r ℓ-isolated singletons. We have two options: either n (the last element) is not an ℓ-isolated element or n is one of the r ℓ-isolated elements. In the first case we can do the following construction: if the elements $n-1-\ell, \ldots, n-2, n-1$ are in the same block, then we put the element n into this block, so by counting the complement we obtain the the following enumeration

$$L_{\ell,r}(n-1, k) - \left(\sum_{j=1}^{\ell} L_{\ell,r-j}(n-1-j, k-1) + (k-1) \sum_{j=1}^{\ell} L_{\ell,r-j}(n-1-j, k) \right).$$

We can also have $\{n-\ell, \ldots, n-1, n\}$ as a block of π, then we have $L_{\ell,r}(n-\ell-1, k-1)$ options, or a block that contains the elements $n-\ell, \ldots, n-1, n$ but not the element $n-\ell-1$. In this case we have $(k-1)L_{\ell,r}(n-\ell-1, k)$ ways.

On the other hand, if n is one of the r ℓ-isolated elements, then by a similar argument as in Proposition 3.1 we obtain that there are

$$\sum_{j=1}^{\ell} L_{\ell,r-j}(n-j, k-1) + (k-1)\sum_{j=1}^{\ell} L_{\ell,r-j}(n-j, k)$$

options. Summing the different options we obtain the desired result. □

Munagi gave the following closed formula for the case $\ell = 1$

$$L_{1,r}(n,k) = \sum_{j\geq 1} \binom{j}{r}\binom{n-j-1}{j-r-1}\left\{ {j-1 \atop k-1} \right\}.$$

In Theorem 4.2 we generalize this result.

Theorem 4.2. *We have*

$$L_{\ell,r}(n,k) = \sum_{j\geq 1} g_\ell(n,j;r)\left\{ {j-1 \atop k-1} \right\},$$

where $g_\ell(n,j;r)$ counts the number of compositions of n with j parts, where the sum of parts that are greater than ℓ is equal to $n-r$.

Proof. The sequence $L_{\ell,r}(n,k)$ is the number of partitions of $[n]$ into k non-empty blocks such that there are r ℓ-isolated elements, i.e., $n-r$ ℓ-neighbors. Take a composition $n = c_1 + \cdots + c_j$, and form from the first c_1 consecutive elements a set P_1, from the second c_2 consecutive elements a set P_2 etc. Take now a set partition π of $[j]$ into k non-empty blocks such that no two consecutive elements are in the same block. Finally, exchange each element i in the partition by the sets of consecutive elements P_i, obtaining the partition π' of $[n]$. The number of ℓ-neighbors in the so obtained partition π' is equal to the number of the elements in all P_i's that have size greater than ℓ. From definition of $g_\ell(n,j;r)$, the statement follows. □

In Lemma 5.1 we give a generating function $C_\ell(x,t,w)$, such that

$$g_\ell(n,j;r) = [x^n t^{n-r} w^{j-1}]C_\ell(x,t,w).$$

5 Non-crossing partitions with a given number of ℓ-neighbors

A set partition is called *non-crossing* if none of the edges on the graph representation cross. In other words if i,j with $i<j$ is contained in a block and s,t with $s<t$ is contained in another block and $i<s<j$ then $i<t<j$.

Let $\mathcal{NC}(n)$ denote the set of non-crossing set partitions of $[n]$. It is well-known that $|\mathcal{NC}(n)| = C_n$, where $C_n = \frac{1}{n+1}\binom{2n}{n}$ is the n-th Catalan number. In this section we study the distribution of the neighbors in the non-crossing partitions. Let $H_{\ell,r}(n,k)$ denote the number of non-crossing partitions of $[n]$ into k blocks containing r ℓ-neighbors.

We define the multivariate generating function

$$H_\ell(x,y,z) := \sum_{\pi\in\mathcal{NC}(n,k)} x^n y^k z^{\eta_\ell(\pi)} = \sum_{n,k,r\geq 0} H_{\ell,r}(n,k)x^n y^k z^r,$$

where $\eta_\ell(P)$ is the number of ℓ-neighbors in π and $\mathcal{NC}(n,k)$ is the set of non-crossing set partitions of $[n]$ into k blocks.

In order to obtain a functional equation for this generating function, we translate our problem into the language of Dyck paths. Recall that a *Dyck path* of semi-length n is a lattice path in the first quadrant of the xy-plane that starts at the point $(0,0)$, ends on the point $(2n,0)$, and consists of the same number of up-steps $X=(1,1)$ and down-steps $Y=(1,-1)$.

There is a well-known bijection between $\mathcal{NC}(n)$ and the Dyck paths of semi-length n [14]. The bijection works as follows. First, label with positive integers, from 1 to n, the up-steps left to right. Second, label each down-step with the number on its matching up-step. Finally, define the partition of $[n]$, whose blocks are the labels on the descents (maximal number of Y's). For example, the Dyck path $X^4Y^2X^4Y^3XY^4X^2Y^2$ corresponds to the non-crossing partition $\{1,2,5,9\},\{3,4\},\{6,7,8\},\{10,11\}$, see Figure 2. A *peak* is a subpath of the form XY and a *valley* is a subpath of the form YX. A *pyramid* of height h in a Dyck path is a subpath of the form X^hY^h; it is called *maximal* if it can not be extended to a pyramid $X^{h+1}Y^{h+1}$. We use Δ_h to denote a maximal pyramid of the form X^hY^h.

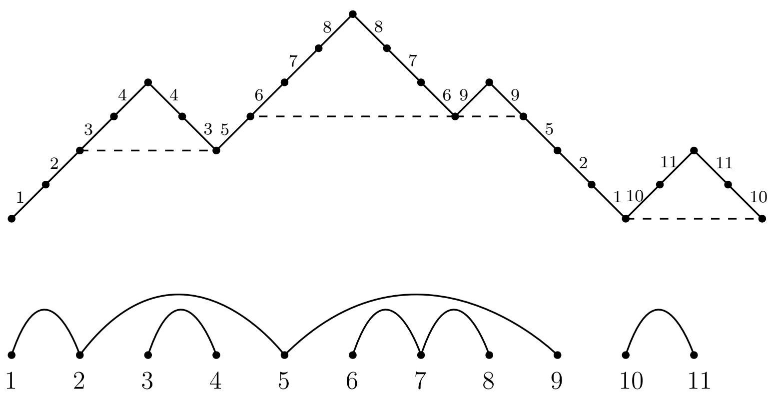

Figure 2: Bijection between non-crossing partitions and Dyck paths.

Let $\mathcal{D}$ denote the family of Dyck paths. We have

$$H_\ell(x,y,z) := \sum_{P\in\mathcal{D}} x^{|P|}y^{\rho(P)}z^{\eta_\ell(P)},$$

where $|P|$ is the semi-length of P, $\rho(P)$ is the number of peaks contained in P. We explain the definition of $\eta_\ell(P)$ later since the correspondence of the ℓ-neighbors in the case of the Dyck paths is a little more complicated. From the symbolic method [6], in Theorem 5.2 we give a functional equation satisfied by $H_\ell(x, y, z)$. Before, we need an auxiliary generating function.

Let $\mathcal{C}(n)$ denote the set of compositions of n. We introduce the generating function

$$C_\ell(x, t, w) = 1 + \sum_{n\geq 1} x^n \sum_{\sigma\in\mathcal{C}(n)} t^{\mathtt{Part}_\ell(\sigma)} w^{\mathtt{Part}(\sigma)-1},$$

where $\mathtt{Part}(\sigma)$ denotes the number of parts of σ and $\mathtt{Part}_\ell(\sigma)$ denotes the sum of the parts of σ strictly greater than ℓ.

Lemma 5.1. *We have*

$$C_\ell(x, t, w) = 1 + \frac{P_\ell(x, t)}{1 - wP_\ell(x, t)},$$

where $P_\ell(x, t) = \sum_{k=1}^{\ell} x^k + \sum_{k\geq \ell+1}(xt)^k$.

Proof. Let $\sigma = \sigma_1 \cdots \sigma_m$ be a non-empty composition having m parts. If $\sigma_j = k$ with $k > \ell$, then σ_j contributes to the generating function with the term $x^k t^k$, while if $k \leq \ell$, then it contributes with the term x^k. Therefore,

$$C_\ell(x, t, w) = 1 + \sum_{m\geq 1} w^{m-1} \left(\sum_{k=1}^{\ell} x^k + \sum_{k\geq \ell+1} (xt)^k \right) = 1 + \frac{P_\ell(x, t)}{1 - wP_\ell(x, t)}.$$

□

Theorem 5.2. *The generating function $H_\ell(x, y, z)$ satisfies the functional equation*

$$H_\ell(x, y, z) = 1 + y(C_\ell(x, z, H_\ell(x, y, z) - 1) - 1)H_\ell(x, y, z).$$

Proof. A Dyck path P is either empty or has the form $P = X^n Y P_1 Y P_2 \cdots Y P_n$, where P_i are Dyck paths (possible empty) for $1 \leq i \leq n$. This decomposition is equivalent to $P = X^n Y^{i_1} P_1^+ Y^{i_2} P_2^+ \cdots Y^{i_{r-1}} P_{r-1}^+ Y^{i_r} P_r$, where P_k^+ are non-empty Dyck paths, for $1 \leq k < r \leq n$, and $i_1 + i_2 + \cdots + i_r = n$, with $i_k \geq 1$. Figure 3 illustrates this decomposition.

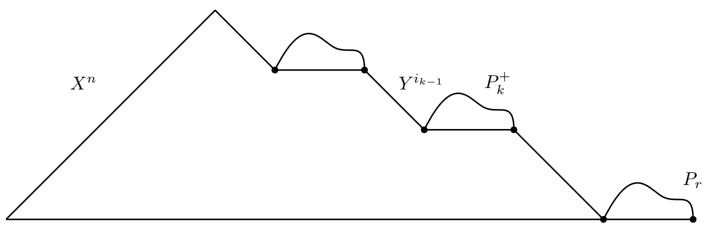

Figure 3: Decomposition of a Dyck path.

Notice that the number of ℓ-neighbors in a non-crossing partition corresponds to the sum $\sum_{\substack{i_k>\ell \\ 1\le k\le r}} i_k$ in the decomposition of the path plus the total number of ℓ-neighbors associated with the Dyck path P_r in the decomposition. For example, for $\ell = 1$ and $n = 4$ we have the decompositions given in Figure 4 with the corresponding compositions of $n = 4$. Therefore, from Lemma 5.1 we obtain the desired result. $\square$

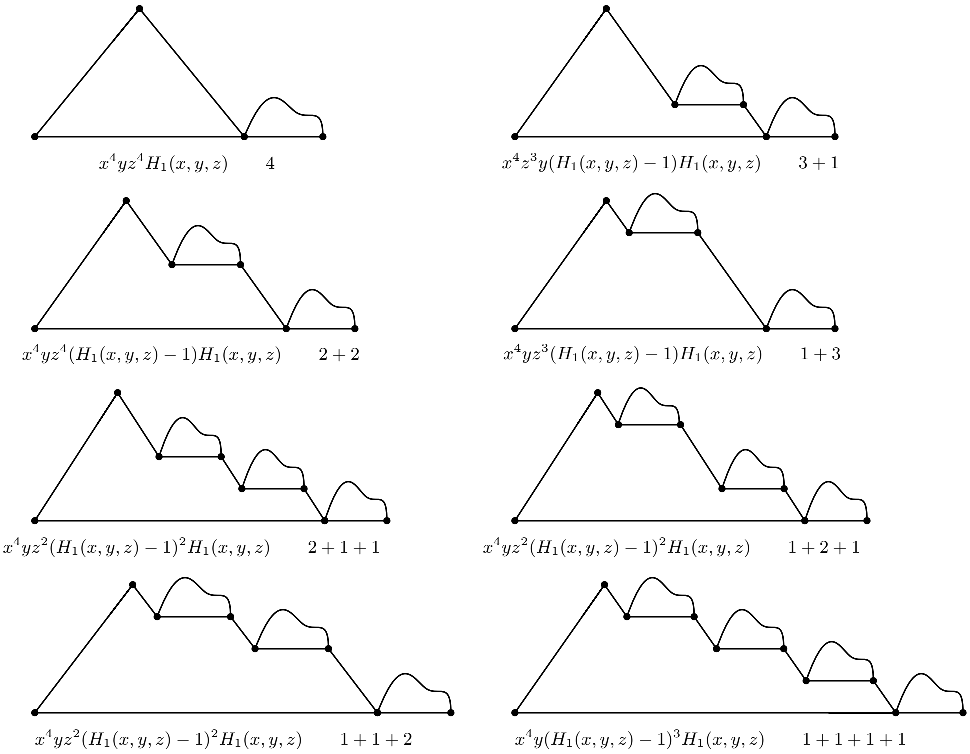

Figure 4: Decomposition of the Dyck paths for $n = 4$.

The solution of the functional equation in Theorem 5.2 produces an algebraic generating function of degree 2. For example, for $\ell = 1$ we have the explicit expression

$$H_1(x, y, z) = \frac{1 - x^2(2 - y)(1 - z)z + x(2 - y - z) - \sqrt{p(x, y, z)}}{2x(1 - x(1 - z)z)},$$

where

$$\begin{aligned} &p(x, y, z) \\ &= 1-xz(2-xz)+(xy-x^2y(1-z)z)^2-2yx(1-x(1-z)z)(1+x(2-z)-2x^2(1-z)z). \end{aligned}$$

As a series expansion, the generating function $H_1(x, y, z)$ begins with

$$\begin{aligned} H_1(x, y, z) = 1 + yx + (y^2 + yz^2)x^2 + \boldsymbol{(y^2 + y^3 + 2y^2z^2 + yz^3)x^3} \\ + (3y^3 + y^4 + 3y^2z^2 + 3y^3z^2 + 2y^2z^3 + yz^4 + y^2z^4)x^4 + \cdots \end{aligned}$$

Example 5.3. Figure 5 shows the non-crossing partitions of [3] corresponding to the bold coefficient in the above series. The red vertices represent the 1-neighbors.

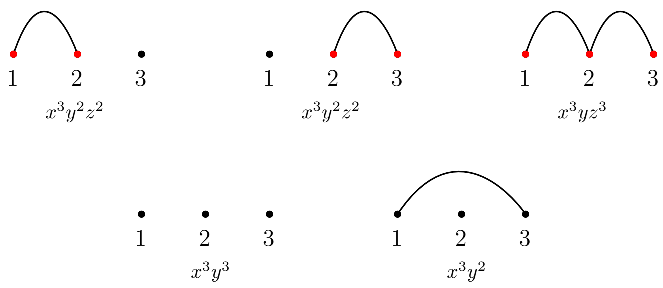

Figure 5: Non-crossing partitions of [3] and the 1-neighbors.

6 An application: RNA shapes

In this section we investigate the non-crossing partitions with only ℓ-singletons. In particular, we point out an interesting connection to the RNA shapes, an abstract object that are defined in molecular biology, and we provide an explicit formula for $H_{\ell,0}(n)$ involving the Motzkin numbers.

Denote by $H_{\ell,0}(n)$ the total number of non-crossing partitions of $[n]$ that contain no ℓ-neighbors. Let $H_\ell(x,y)$ be the bivariate generating function $H_\ell(x,y) = \sum_{n,k\geq 0} H_{\ell,0}(n,k)x^n y^k$. By setting $z=0$ in Theorem 5.2 we obtain Corollary 6.1.

Corollary 6.1. *The generating function $H_\ell(x,y)$ is given by*

$$\frac{1+x(1-y)-x^{\ell+1}(2-y)-\sqrt{(1+x(1-y)-x^{\ell+1}(2-y))^2-4x(1-x^\ell)(1-x^{\ell+1})}}{2x(1-x^\ell)}.$$

In particular, the generating function of the sequence $H_{\ell,0}(n) = \sum_{k\geq 1} H_{\ell,0}(n,k)$ is given by

$$\begin{aligned} H_\ell(x) := \sum_{n\geq 0} H_{\ell,0}(n)x^n &= H_\ell(x,1) \\ &= \frac{1+x-x^{\ell+1}-\sqrt{(1-x^{\ell+1})^2-4x(1-x^\ell)(1-x^{\ell+1})}}{2x(1-x^\ell)}. \end{aligned}$$

In Table 2 we show the first few values of the sequence $H_{\ell,0}(n)$.

$\ell\backslash n$	1	2	3	4	5	6	7	8	9	10
$\ell=1$	1	1	2	4	9	21	51	127	323	835
$\ell=2$	1	2	4	11	31	92	283	893	2875	9407
$\ell=3$	1	2	5	13	39	121	388	1277	4288	14630
$\ell=4$	1	2	5	14	41	129	418	1389	4708	16215
$\ell=5$	1	2	5	14	42	131	426	1419	4821	16642

Table 2: Values of the sequence $H_{\ell,0}(n)$.

Note that $H_{1,0}(n)$, the number of non-crossing partitions such that two consecutive elements cannot be included in the same block, are the Motzkin numbers A001006. Using the correspondence between non-crossing partitions and Dyck paths as above and the classical translation of Dyck paths into bracketings, we have that $H_{1,0}(n)$ is the number of bracketings with no occurrence of $((*))$, where $*$ is itself any (possibly vacuous) bracketings.

Example 6.2. For $n = 5$ the $H_{1,0}(5) = 9$ bracketings are

$$\begin{array}{lllll} ()()()()(), & (()())()(), & ()(()())(), & ()()(()()), & (()()())(), \\ ()(()()()), & (()()()()), & ((()())()), & (()(()())). \end{array}$$

It is known that these objects are counted by the Motzkin numbers [4]. Here we point out an application in molecular biology where these sequences occur.

RiboNucleic Acids (RNAs) are a category of bio-molecules. RNA may be described by its sequence of bases, nucleotides Adenine (A), Cytosine (C), Guanine (G), and Uracil (U). The sequence of bases is known to be the primary structure of the molecule. The RNA is folded then into a helical 3D structure. In the folding process, the secondary structure plays a crucial role, which is completed by complex topological motifs, including pseudoknots and non-canonical pairs and motifs, in reaching the final 3D form of the molecule. Hence, the secondary structure is used to predict the 3D structure [5, 13, 16].

The experimental determination of RNA structure is time-consuming and expensive. Therefore, computational prediction is of great interest; however, the users are often only interested in structures with fundamental differences. For this reason, the consideration of classes comprising similar structures in a reasonable way is useful. *RNA shapes* [7] represent a hierarchy of abstract representations for the secondary structures of RNA. The coarsest shape [11, p. 119] is defined so that the unpaired nucleotides are ignored and consecutive helices separated only by an internal loop are contracted. One of the visualizations of such a class (and its representative) is by bracketing as described above.

In terms of bracketings, $H_{\ell,0}(n)$ count those bracketings, where $((\cdots(*)\cdots))$ with direct nestings of length more than $\ell+1$ collapse to ℓ, or in other words the depth of direct nestings is at most ℓ. Such bracketings are in one-to-one correspondence with ordered trees with branch lengths at most ℓ. (In the special case, when $\ell = 1$, these trees are called *branch-reduced plane trees* [3] or *bushes* [4, (M8)], and are in bijection with the bracketings without $((*))$, by the well-known "walking around the tree"–transformation.)

We recall briefly some facts from Deutsch [2]. If T denotes the generating function of all ordered trees according to size and P the generating function of all paths, then we have

$$T = 1 + P + PT(T-1) = 1 + P(1 - T + T^2),$$

which in turn implies that the generating function of trees with branches of length at most ℓ satisfies

$$T = 1 + (z + z^2 + \cdots + z^\ell)(1 - T + T^2).$$

By extracting the coefficient of z^n we obtain

$$1 + M_0(z + z^2 + \cdots + z^\ell) + M_1(z + z^2 + \cdots + z^\ell)^2 + M_2(z + z^2 + \cdots + z^\ell)^3 + \cdots,$$

where M_i is the ith Motzkin number. Hence,

$$H_{\ell,0}(n) = \sum_{j\geq 1} f_\ell(n, j) M_{j-1},$$

where $f_\ell(n, j)$ counts the number of compositions of n with j parts in $\{1, 2, \ldots, \ell\}$.

The combinatorial explanation is as follows: take a composition $n = c_1 + \cdots + c_j$ of n into j parts from the set $\{1, 2, \ldots, \ell\}$, in other words with parts at most ℓ and a branch-reduced plane tree with j edges. Cut the ith edge of the tree into a path of c_i edges by inserting $c_i - 1$ new vertices. Since the number of compositions is given by $f_\ell(n, j)$ and the number of branch-reduced trees is given by the Motzkin numbers we have the above formula. For $\ell = 2$ and $\ell = 3$ the sequences are A247333 and A239106 in [15].

$$H_{2,0}(n) = \sum_{k=1}^{n} \binom{k}{n-k} M_{k-1} \quad \text{and} \quad H_{3,0}(n) = \sum_{k=n/3}^{n} \binom{k}{n-k}_3 M_{k-1},$$

where $\binom{n}{k}_3$ denotes the trinomial coefficient.

Acknowledgements

The authors thank the referees for their comments which helped to improve the paper.

References

[1] J. B. Caicedo, V. H. Moll, J. L. Ramírez and D. Villamizar, Extensions of set partitions and permutations, *Electron. J. Comb.* **26**(2) (2019), #P2.20.

[2] E. Deutsch, Ordered trees with described root degrees, node degrees, and branch lengths, *Discrete Math.* **282** (2004), 89–94.

[3] R. Donaghey, Restricted plane tree representations of four Motzkin-Catalan equations, *J. Combin. Theory Ser. B* **22** (1977), 114–121.

[4] R. Donaghey and L. W. Shapiro, Motzkin numbers, *J. Combin. Theory Ser. A* **23** (1977), 291–301.

[5] T. Došlić and D. Veljan, Secondary structures, plane trees and Motzkin numbers, *Math. Commun.* **12** (2007), 163–169.

[6] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2009.

[7] R. Giegerich, B. Voss and M. Rehmeister, Abstract shapes of RNA, *Nucleic Acids Res.* **32**(16) (2004), 4843–4851.

[8] S. Heubach and T. Mansour, *Combinatorics of Compositions and Words*, CRC Press, 2010.

[9] T. Mansour and A. Munagi, Set partitions with isolated successions, *Notes Number Theory Discrete Math.* **28**(1) (2022), 81–91.

[10] A. Munagi, Set partitions with isolated singletons, *Amer. Math. Monthly* **125**(5) (2018), 447–452.

[11] Y. Ponty, Ensemble algorithms and analytic combinatorics in RNA bioinformatics and beyond, Habilitation Université Paris-Saclay, (2020).

[12] H. Prodinger, On the number of Fibonacci partitions of a set, *Fibonacci Quart.* **19** (1981), 463–465.

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

[14] R. Simion, Noncrossing partitions, *Discrete Math.* **217** (2000), 367–409.

[15] N. J. A. Sloane, *The On-line Encyclopedia of Integer Sequences*, https://oeis.org.

[16] P. R. Stein and M. S. Waterman, On some sequences generalizing the Catalan and Motzkin numbers, *Discrete Math.* **26** (1978), 261–272.

(Received 10 Feb 2022; revised 4 July 2022)