J. Comput. Biol.
21 (2014), no. 12, 915-935

Zigzag Stacks and m-Regular Linear Stacks

William Y.C. Chen[a,b,*], Qiang-Hui Guo[a], Lisa H. Sun[a] and Jian Wang[a]

[a]Center for Combinatorics, LPMC-TJKLC
Nankai University, Tianjin 300071, P. R. China

[b]Center for Applied Mathematics
Tianjin University, Tianjin 300072, P. R. China

Emails: *chen@nankai.edu.cn, guo@nankai.edu.cn
sunhui@nankai.edu.cn, wj121313@126.com

Abstract. The contact map of a protein fold is a graph that represents the patterns of contacts in the fold. It is known that the contact map can be decomposed into stacks and queues. RNA secondary structures are special stacks in which the degree of each vertex is at most one and each arc has length at least two. Waterman and Smith derived a formula for the number of RNA secondary structures of length n with exactly k arcs. Höner zu Siederdissen et al. developed a folding algorithm for extended RNA secondary structures in which each vertex has maximum degree two. An equation for the generating function of extended RNA secondary structures was obtained by Müller and Nebel by using a context-free grammar approach, which leads to an asymptotic formula. In this paper, we consider m-regular linear stacks, where each arc has length at least m and the degree of each vertex is bounded by two. Extended RNA secondary structures are exactly 2-regular linear stacks. For any $m \geq 2$, we obtain an equation for the generating function of the m-regular linear stacks. For given m, we deduce a recurrence relation and an asymptotic formula for the number of m-regular linear stacks on n vertices. To establish the equation, we use the reduction operation of Chen, Deng and Du to transform an m-regular linear stack to an m-reduced zigzag (or alternating) stack. Then we find an equation for m-reduced zigzag stacks leading to an equation for m-regular linear stacks.

1 Introduction

Proteins are polymer chains consisting of amino acid residues of twenty types. The function of a protein is directly dependent on its three dimensional structure. Due to the complexity of the full-atom protein model, lattice models have been proposed and extensively studied. Lattice models often preserve important features of the protein structure, and enable us to focus on dominant aspects of a protein structure. In such a model, protein folds are represented by self-avoiding walks on the specific lattice.

When two amino acids in a protein fold come very close to each other, say, closer than a predetermined threshold, they presumably form some kind of bond, which is called a contact. Let $S = s_1 s_2 \dots s_n$ represent the amino acid residue sequence of a protein. When we consider the protein fold as a self-avoiding walk on some regular lattice, two residues s_i and s_j are in a contact if they reside on two adjacent points in the lattice, but not consecutive in the sequence. Let the vertex i stand for the residue s_i. The contact map of a folding of S is a diagram with vertices $1, 2, \dots, n$ arranged on a horizontal line and there is an edge between two vertices if they are in contact. See Figure 1.1 for an illustration.

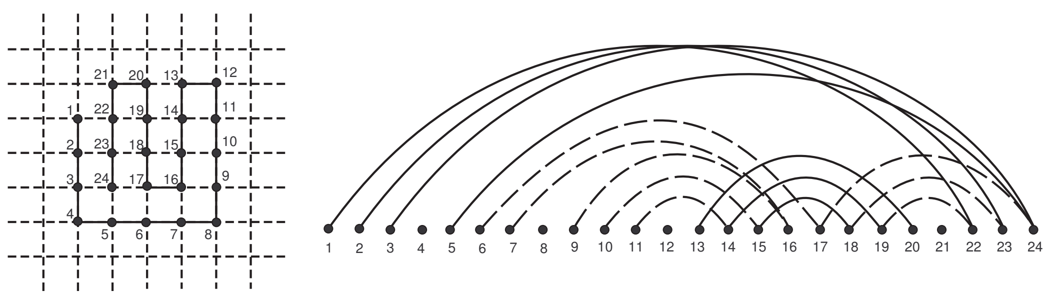

Figure 1.1: A protein fold on the 2D square lattice and its contact map.

Contacts play a fundamental role in the HP-model for protein folding, see [5, 6]. Contact maps of protein folds have been extensively studied from various perspectives, such as protein folding prediction [7,24], structure alignment [1,9,16], protein secondary structure [15, 26], and protein structure data mining [11]. Crippen [4] has studied the enumeration of contact maps. Vendruscolo et al. [23] investigated statistical properties of contact maps. Goldman et al. [9] discovered that contact maps for protein folds in two dimension can be decomposed into "simpler" graphs, called stacks and queues. In combinatorial words, a *stack* is a noncrossing diagram, and a *queue* is a nonnesting diagram.

Theorem 1.1 (Goldman et al. [9]). *For any protein sequence S, the contact map of any two-dimensional fold of S can be decomposed into (at most) two stacks and one queue.*

For example, the contact map in Figure 1.1 can be decomposed into two stacks:

$\{(6,17),\ (7,16),\ (9,16),\ (10,15),\ (11,14),\ (17,24),\ (18,23),\ (19,22)\}$ and $\{(13,20),$ $(14,19),\ (15,18)\}$, and one queue $\{(1,22),\ (2,23),\ (3,24),\ (5,24)\}$.

Recently, Agarwal et al. [1] found a similar decompositions of contact maps of protein folds in the three dimensional cubic lattice.

As pointed out by Istrail and Lam, the enumeration of stacks and of queues is related to an approximation algorithm for computing the partition function of self-avoiding walks in two dimensions. Denote the numbers of stacks on n vertices by $s(n)$. We notice that from the combinatorial intepretation of Schröder numbers a_n in terms of noncrossing graphs, see [22, Exercise 6.39(p)], it is easy to see that $s(n) = 2^{n-1}a_{n-2}$, where a_n is the Schröder number whose generating function is given by

$$\sum_{n\geq 0} a_n x^n = \frac{1 - x - \sqrt{1 - 6x + x^2}}{2x}.$$

When folding a protein on a specific lattice, the lattice model leads to degree and arc length constraints to the corresponding contact map. For instance, in a contact map on 2D square lattice, each internal vertex has maximum degree 2 and each arc has length at least 3. In the case of 2D triangular lattice, each vertex has maximum degree 4 and each arc has length at least 2. In the hexagonal lattice, the degree of each vertex is at most 1, but the length of each arc is at least 5. For other lattice models, see [19]. A stack with arc length at least m is called an m-regular stack and a stack with degree of each vertex bounded by two is called linear.

RNA secondary structures can be viewed as 2-regular stacks with maximum degree 1. By establishing a bijection between RNA secondary structures and linear trees, Schmitt and Waterman [21, 25] provided an explicit formula for the number of RNA secondary structures on n vertices and k arcs. See Figure 1.2 for an example of the bijection.

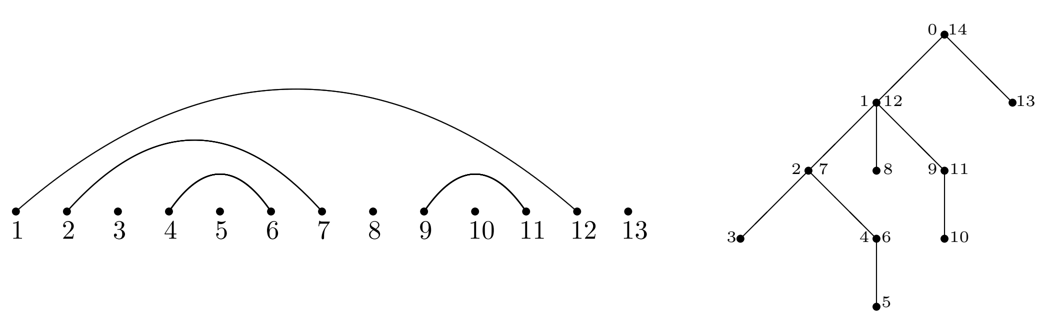

Figure 1.2: An RNA secondary structure and its Schmitt-Waterman contact tree.

Theorem 1.2 (Schmitt and Waterman [21])**.** *The number of RNA secondary structures* *of length n with k arcs is given by*

$$s_2(n,k) = \frac{1}{k}\binom{n-k}{k+1}\binom{n-k-1}{k-1}. \tag{1.1}$$

In a recent survey [12], the authors raised the question concerning generalizations of the Schmitt-Waterman counting formulas to stacks and to queues. In fact, Nebel [18] derived the generating function of m-regular RNA secondary structures by using the binary trees and the Horton-Strahler number. Based on the bijection between matchings and walks inside Weyl-chambers given by [3], Jin, Qin and Reidys [13] derived a formula for the number of k-noncrossing RNA structures with pseudoknots of length n with l isolated vertices.

Theorem 1.3. *Let $f_k(n,l)$ be the number of k-noncrossing digraphs over n vertices with exactly l isolated vertices. Then the number of RNA structures of pseudoknot type $k-2$ with l isolated vertices is given by*

$$s_k(n,l) = \sum_{b=0}^{(n-l)/2} (-1)^b \binom{n-b}{b} f_k(n-2b,l), \tag{1.2}$$

where $n \geq 1$, $1 \leq l \leq n$ and $k \geq 2$.

In particular, when $k=2$, formula (1.2) reduces to Schmitt and Waterman's formula (1.1). When $k=3$, Jin and Reidys [14] obtained the following asymptotic formula for the 3-noncrossing RNA structures

$$S_3(n) \sim \frac{10.4724 \cdot 4!}{n(n-1)\cdots(n-4)} \left(\frac{5+\sqrt{21}}{2}\right)^n.$$

Höner zu Siederdissen et al. [10] presented a model of extended RNA secondary structures in which the degree of each vertex is bounded by two. They provided a folding algorithm which is known to be the first thermodynamics-based algorithm that allows the degrees of vertices to be two. Müller and Nebel [17] studied the enumeration of extended RNA secondary structures by using a context-free grammar approach. They obtained an equation satisfied by the ordinary generating functions of the number of extended RNA secondary structures.

Theorem 1.4. *Let*

$$S(z) = \sum_{n\geq 0} r_2(n) z^n,$$

where $r_2(n)$ is the number of extended RNA secondary structures of length n and $r_2(0) = 0$. Then we have

$$\begin{aligned} S(z) = {} & 4z^5S^5(z) + (4z^3 - 7z^4 + 9z^5)S^4(z) + (-8z^2 + 11z^3 - 14z^4 + 7z^5)S^3(z) \\ & + (5z - 10z^2 + 14z^3 - 9z^4 + 2z^5)S^2(z) + (3z - 7z^2 + 7z^3 - 2z^4)S(z) \\ & + z - 2z^2 + z^3 \end{aligned} \tag{1.3}$$

and

$$r_2(n) \sim 0.250536155 \times 4.1012475^n \cdot n^{-\frac{3}{2}}. \tag{1.4}$$

In this paper, we are mainly concerned with m-regular linear stacks, in which arc length is at least m with $m \geq 2$ and the degree of each vertex is bounded by 2. For example, the two stacks in Figure 1.1 are 3-regular linear stacks. Extended RNA secondary structures are exactly 2-regular linear stacks.

We obtain an equation for the generating function of the m-regular linear stacks. For given m, we can derive an explicit recurrence relation and an asymptotic formula for the number of m-regular linear stacks of length n.

Using the reduction operation on the standard representation of an m-regular partition due to Chen et al. [2], we find a class of zigzag stacks which are in one-to-one correspondence with m-regular linear stacks. More precisely, a zigzag stack, or an alternating stack, is a noncrossing diagram without multiple edges or loops subject to the following conditions:

(1) The degree of each vertex is bounded by 2;

(2) For each vertex v of degree 2, the two arcs are on the same side with respect to the position of v.

Notice that isolated vertices are allowed in a zigzag stack. Figure 1.3 gives a zigzag stack. It is easy to see that each connected component of a zigzag stack forms a zigzag path.

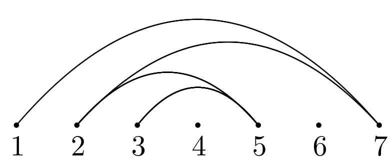

Figure 1.3: A zigzag stack.

Given a vertex v in a stack, we denote the degree of v by $\deg(v)$, and denote the left-degree and right-degree of v by $\mathrm{ldeg}(v)$ and $\mathrm{rdeg}(v)$, respectively. We find that the reduction operation transforms an m-regular linear stack of length $n+m-1$ to a zigzag stack of length n with the following two constraints:

(1) For $1 \leq i \leq n-m+1$, $\mathrm{ldeg}(i) + \mathrm{rdeg}(i+m-1) \leq 2$;

(2) For $1 \leq i < j \leq n$, if $\mathrm{ldeg}(i) > 0$ and $\mathrm{rdeg}(j) > 0$, then $j - i \geq m-1$.

A zigzag stack satisfying the above two constraints is called an m-reduced zigzag stack. The above conditions can be used to characterize the substructures of m-reduced zigzag stacks. For example, the zigzag stack in Figure 1.4 is obtained from a non-linear 3-regular stack by applying the reduction operation twice. Thus none of 3-reduced zigzag

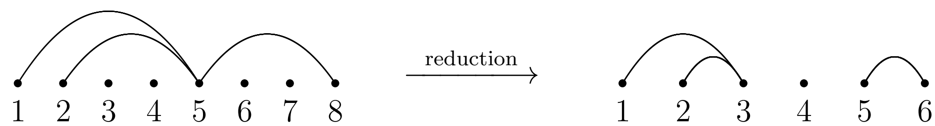

Figure 1.4: The reduction from a non-linear stack to a zigzag stack.

stacks can contain a substructure like the reduced zigzag stack in Figure 1.4. We shall use Conditions (1) and (2) to describe the substructures in the decomposition of m-reduced zigzag stacks. As will be seen, the substructure in Figure 1.4 is not a valid substructure according to the characterizations as given in Theorem 5.1.

Denote the set of m-regular linear stacks and m-reduced zigzag stacks of length n by $\mathcal{R}_m(n)$ and $\mathcal{Z}_m(n)$, respectively. Clearly, when $m = 2$, $\mathcal{R}_2(n)$ reduces to extended RNA secondary structures. Let $r_m(n) = |\mathcal{R}_m(n)|$ and $z_m(n) = |\mathcal{Z}_m(n)|$, and let

$$R_m(x) = \sum_{n=0}^{\infty} r_m(n)x^n,$$

$$Z_m(x) = \sum_{n=0}^{\infty} z_m(n)x^n,$$

where $r_m(0) = z_m(0) = 1$. Then the reduction algorithm implies that $z_m(n) = r_m(n + m - 1)$ and

$$R_m(x) = 1 + x + x^2 + \cdots + x^{m-2} + x^{m-1}Z_m(x). \tag{1.5}$$

By decomposing an m-reduced zigzag stack into a connected component and a list of substructures, we derive an equation satisfied by $Z_m(x)$. Furthermore, by applying relation (1.5), we obtain that the following relation for $R_m(x)$ with polynomial coefficients.

Theorem 1.5. *We have*

$$c_5(x)R_m^5 + c_4(x)R_m^4 + c_3(x)R_m^3 + c_2(x)R_m^2 + c_1(x)R_m + c_0(x) = 0, \tag{1.6}$$

where

$$
\begin{aligned}
c_0(x) &= (x-1)(x^m-1)^3,\\
c_1(x) &= (x^m-1)^2(x^{2m+1}-2x^{m+2}-x^{m+1}+x^m-3x^3+8x^2-3x-1),\\
c_2(x) &= -x(x-1)(x^m-1)(5x^{2m+1}-6x^{m+2}-9x^{m+1}+5x^m-3x^3+12x^2+x-5),\\
c_3(x) &= x^2(x-1)^2(x^m-1)(11x^{m+1}-8x^2-11x+8),\\
c_4(x) &= x^3(x-1)^3(-11x^{m+1}+4x^2+11x-4),\\
c_5(x) &= 4x^5(x-1)^4.
\end{aligned}
$$

From the above equation (1.6), we can deduce a recurrence relation for $r_m(n)$ for given m. Applying Newton-Puiseux Expansion Theorem [8], we are led to the following asymptotic formula

$$r_m(n) \sim \gamma \cdot \omega^n \cdot n^{-\frac{3}{2}},$$

where γ and ω are constants. When $m = 2$, it coincides with Müller and Nebel's formula (1.4) for extended RNA secondary structures, see [17]. For $m = 3, 4, 5, 6$, we have

$$
\begin{aligned}
r_3(n) &\sim 0.19005341 \times 3.5271506^n \cdot n^{-\frac{3}{2}},\\
r_4(n) &\sim 0.145636571 \times 3.2431591^n \cdot n^{-\frac{3}{2}},\\
r_5(n) &\sim 0.112004701 \times 3.0833083^n \cdot n^{-\frac{3}{2}},\\
r_6(n) &\sim 0.086237333 \times 2.9880679^n \cdot n^{-\frac{3}{2}}.
\end{aligned}
$$

2 Zigzag stacks

In this section, we derive an equation satisfied by the generating function of the number of zigzag stacks on n vertices. To enumerate zigzag stacks, we introduce the primary component decomposition.

Denote $[n] = \{1, 2, \ldots, n\}$. Let S be a zigzag stack on $[n]$, and define the primary component of S to be the connected component containing the vertex 1. The primary component will split $[n]$ into disjoint intervals, on which smaller zigzag stacks can be constructed. This enables us to establish a recursive procedure to enumerate zigzag stacks, so that we can derive an equation on the generating function for the number of zigzag stacks. Figure 2.1 illustrates a primary component decomposition of a zigzag stack.

The following lemma shows that the primary component decomposition leads to a primary component along with zigzag stacks on the intervals.

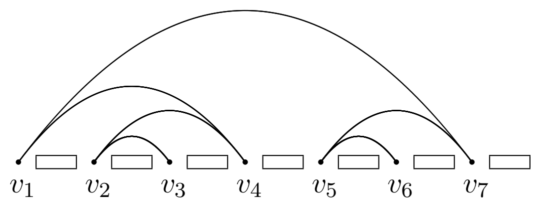

Figure 2.1: Decomposition of a zigzag stack.

Lemma 2.1. *Let S be a zigzag stack, and C be the primary component of S. Then there is no arc of S that connects two vertices in different intervals. In other words, a zigzag stack can be decomposed into a primary component along with a list of zigzag stacks on the intervals.*

Proof. Clearly, any vertex in an interval is not connected to any vertex of C. Assume to the contrary that the lemma does not hold. Then there exists an arc (i, j) with $i < j$ such that $i \in I$ and $j \in J$, where I and J are two different intervals of S. Let k be the vertex of C such that k is next to the last vertex of I. We now have $1 < i < k < j$. Since 1 and k are in the connected component C, the arc (i, j) must intersect with some arc of C, contradicting the assumption that S is a stack. This completes the proof. ∎

Let c_n denote the number of connected zigzag stacks on $[n]$. We have the following formula.

Lemma 2.2. *If $n \geq 2$, we have*

$$c_n = n - 1. \tag{2.1}$$

Proof. The lemma is obvious for $n = 2, 3$. For $n \geq 4$, let C be a connected zigzag stack on $[n]$. Clearly, $(1, n)$ is an arc in C, since C is connected and is zigzag. There are three cases with respect to the degrees of 1 and n, see Figure 2.2.

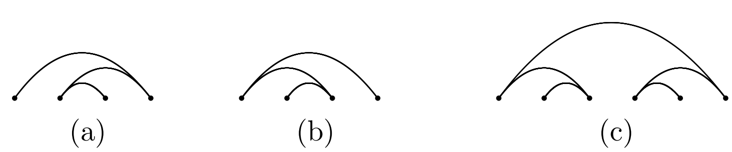

Figure 2.2: Three cases of connected zigzag stacks

If $\deg(1) = 1$, as shown in Figure 2.2 (a), C is the unique diagram consisting of the arcs $(1, n)$, $(2, n)$, $(2, n - 1)$, and so on. Similarly, if $\deg(n) = 1$, C is uniquely determined.

If $\deg(1) = \deg(n) = 2$, assume that $(1, i)$ is an arc of C, where $i < n$. Then $(i + 1, n)$ is an arc of C. Moreover, C is determined once the arcs $(1, i)$ and $(i + 1, n)$

are given. Since i can be any vertex in $\{2, 3, \ldots, n-2\}$, there are $n-3$ choices of C in this case. In summary, there are a total number of $n-1$ connected zigzag stacks on $[n]$. ∎

We set $c_0 = c_1 = 1$. For $n \geq 1$, let $\mathcal{Z}(n)$ denote the set of zigzag stacks on $[n]$, and let $z(n) = |\mathcal{Z}(n)|$. We set $z(0) = 1$. Let

$$Z(x) = \sum_{n \geq 0} z(n) x^n$$

denote the ordinary generating function of $z(n)$. The following theorem gives an equation satisfied by $Z(x)$.

Theorem 2.1. *We have*

$$x^2(x-1)Z^3(x) + 2xZ^2(x) - (x+1)Z(x) + 1 = 0. \tag{2.2}$$

Proof. Let S be a zigzag stack, and let C be the primary component of S. Assume that C has k vertices, then S can be decomposed into a primary component C and k zigzag stacks $S_1, S_2, \ldots, S_k$, where S_i are allowed to be empty.

Let d_i be the number of vertices in S_i. Then we have

$$z(n) = \sum_{k=1}^{n} c_k \sum_{d_1+d_2+\cdots+d_k = n-k} z(d_1)z(d_2)\cdots z(d_k). \tag{2.3}$$

Multiplying both sides of (2.3) by x^n and summing over n, we obtain

$$\begin{aligned}
Z(x) &= \sum_{n=0}^{\infty} \sum_{k=1}^{\infty} c_k \sum_{d_1+d_2+\cdots+d_k = n-k} z(d_1)z(d_2)\cdots z(d_k)x^n \\
&= \sum_{k=1}^{\infty} c_k x^k \sum_{c_1=0}^{\infty} z(d_1)x^{d_1} \sum_{d_2=0}^{\infty} z(d_2)x^{d_2} \cdots \sum_{d_k=0}^{\infty} z(d_k)x^{d_k} \\
&= \sum_{k=1}^{\infty} c_k \big(xZ(x)\big)^k.
\end{aligned}$$

Since $c_0 = c_1 = 1$ and $c_k = k-1$ for $k \geq 2$, we deduce that

$$\begin{aligned}
Z(x) = c_0 + c_1 xZ + \sum_{k=2}^{\infty} c_k (xZ(x))^k \\
&= 1 + xZ(x) + \sum_{k=2}^{\infty} (k-1)(xZ(x))^k \\
&= 1 + xZ(x) + (xZ(x))^2 \sum_{k=0}^{\infty} (k+1)(xZ(x))^k
\end{aligned}$$

$$= 1 + xZ(x) + \frac{(xZ(x))^2}{(1 - xZ(x))^2},$$

which yields (2.2). ∎

Given the equation (2.2), it is straightforward to derive a second order differential equation for $Z(x)$ and a recurrence relation of $z(n)$, see, for example [22, Chapter 6]. The computation can also be carried out by using the Maple package `gfun`, see [20]. Moreover, the asymptotic formula for $z(n)$ can be derived by applying Newton-Puiseux Expansion Theorem [8].

Theorem 2.2. *We have*

$$\begin{aligned}
x^2(23x^3 - 26x^2 + 23x - 4)(4x^2 + x - 1)(x - 1)Z''(x) \\
+ x(368x^6 - 433x^5 + 108x^4 + 260x^3 - 258x^2 + 93x - 10)Z'(x) \\
+ (184x^6 - 87x^5 - 117x^4 + 217x^3 - 129x^2 + 30x - 2)Z(x) \\
- 2(25x^2 - 8x + 1)(x - 1) = 0.
\end{aligned}$$

The number of zigzag stacks on $[n]$ satisfies the following recurrence relation

$$\sum_{i=0}^{6} p_i(n) z(n+i) = 0, \tag{2.4}$$

where

$$\begin{aligned}
p_0(n) &= 184 + 276n + 92n^2, & p_1(n) &= -520 - 606n - 173n^2, \\
p_2(n) &= 347 + 480n + 124n^2, & p_3(n) &= 937 + 210n - 10n^2, \\
p_4(n) &= -1881 - 678n - 60n^2, & p_5(n) &= 1115 + 372n + 31n^2, \\
p_6(n) &= -182 - 54n - 4n^2.
\end{aligned}$$

We also have

$$z(n) \sim 0.4781905 \times 4.6107186^n \cdot n^{-\frac{3}{2}}.$$

The first few values of $z(n)$ are given below

n	1	2	3	4	5	6	7	8	9	10	11	12
$z(n)$	1	2	6	20	70	255	959	3696	14520	57930	234080	955999

3 Reduction of m-Regular linear stacks

In this section, we use the reduction operation in [2] to transform a m-regular linear stack to a zigzag stack. Then we give a characterization of zigzag stacks that are in one-to-one correspondence with m-regular linear stacks. In the next three sections, we derive an equation for the generating function of the number of m-reduce zigzag stacks on $[n]$, which leads to an equation for the generating function of the number of m-regular linear stacks on $[n]$.

An m-reduced zigzag stack in $\mathcal{Z}_m(n)$ is a zigzag stack satisfying the following two conditions:

(1) For $1 \leq i \leq n-m+1$, $\mathrm{ldeg}(i) + \mathrm{rdeg}(i+m-1) \leq 2$;

(2) For $1 \leq i < j \leq n$, if $\mathrm{ldeg}(i) > 0$ and $\mathrm{rdeg}(j) > 0$, then $j - i \geq m-1$.

The reduction operation θ_m is defined by

$$\theta_m \colon \mathcal{R}_m(n+m-1) \longrightarrow \mathcal{Z}_m(n), \tag{3.1}$$

in which for $S \in \mathcal{R}_m(n+m-1)$, $\theta_m(S)$ is obtained from S by replacing each arc (i,j) by an arc $(i, j-m+1)$ and deleting the vertices $n+1, n+2, \ldots, n+m-1$ afterwards. Figure 3.1 is an example for $m = 2$.

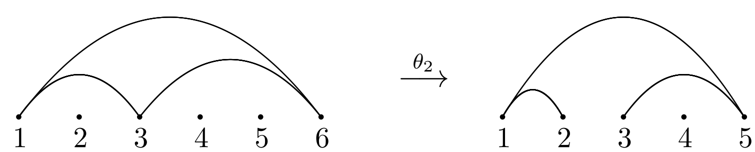

Figure 3.1: An example of the reduction for $m = 2$.

The following theorem shows that θ_m is a bijection.

Theorem 3.1. *The map θ_m is a bijection between m-regular linear stacks on $[n+m-1]$ and m-reduced zigzag stacks on $[n]$.*

Proof. Let S be any m-regular linear stack in $\mathcal{R}_m(n+m-1)$ and $T = \theta_m(S)$. Let (i,j) be an arc of S. Since S is m-regular, that is $j - i \geq m$, we see that the reduced pair $(i, j-m+1)$ is an arc. We claim that T is a stack. Assume to the contrary that T contains two crossing arcs: (i_1, j_1) and (i_2, j_2) with $i_1 < i_2 < j_1 < j_2$. Thus (i_1, j_1+m-1) and (i_2, j_2+m-1) are arcs of S. Moreover, these two arcs form a crossing since $i_1 < i_2 < j_1 + m - 1 < j_2 + m - 1$, a contradiction. This proves that T is a stack.

Next we show that T is a zigzag stack. We claim that for each vertex j of T, either $\mathrm{ldeg}_T(j) = 0$ or $\mathrm{rdeg}_T(j) = 0$. Otherwise, T contains two arcs: (i,j) and (j,k) with $i < j < k$. These two arcs in T correspond to arcs $(i, j+m-1)$ and $(j, k+m-1)$ in S, which form a crossing, since $i < j < j+m-1 < k+m-1$. Thus the claim is proved.

To prove that the degree of each vertex in T is bounded by 2, we notice that for any vertex i in T,

$$\mathrm{ldeg}_T(i) = \mathrm{ldeg}_S(i+m-1) \tag{3.2}$$

and

$$\mathrm{rdeg}_T(i) = \mathrm{rdeg}_S(i). \tag{3.3}$$

Moreover, if $\mathrm{ldeg}_T(i) = 0$, then $\deg_T(i) = \mathrm{rdeg}_S(i)$, and if $\mathrm{rdeg}_T(i) = 0$, then $\deg_T(i) = \mathrm{ldeg}_S(i+m-1)$. In view of the above claim that for any vertex i in T, $\mathrm{rdeg}_T(i) = 0$ or $\mathrm{rdeg}_T(i) = 0$, we deduce that either $\deg_T(i) = \mathrm{ldeg}_S(i+m-1)$ or $\deg_T(i) = \mathrm{rdeg}_S(i)$. Given that S is linear, both $\mathrm{ldeg}_S(i+m-1)$ and $\mathrm{rdeg}_S(i)$ are bounded by 2. It follows that $\deg_T(i) \leq 2$. Hence T is a zigzag stack.

Employing relations (3.2) and (3.3), we deduce that

$$\begin{aligned} \mathrm{ldeg}_T(i) + \mathrm{rdeg}_T(i+m-1) &= \mathrm{ldeg}_S(i+m-1) + \mathrm{rdeg}_S(i+m-1) \\ &= \deg_S(i+m-1) \leq 2, \end{aligned}$$

which is Condition (1) in the definition of an m-reduced zigzag stack.

To verify Condition (2), we assume to the contrary that there exist two vertices j, k in T with $j < k$ such that $\mathrm{ldeg}_T(j) > 0$ and $\mathrm{rdeg}_T(k) > 0$, but $k - j < m - 1$. Since $\mathrm{ldeg}_T(j) > 0$ and $\mathrm{rdeg}_T(k) > 0$, there exist two arcs (i, j), (k, l) in T which are obtained from arcs $(i, j+m-1)$ and $(k, l+m-1)$ in S. Since $k - j < m - 1$, we have $i < k < j + m - 1 < l + m - 1$, so that we get two crossing arcs in S, a contradiction. This implies that T satisfies Condition (2) in the definition of an m-reduced zigzag stack. In summary, we have shown that θ_m is well-defined.

Let $T \in \mathcal{Z}_m(n)$. To prove that θ_m is a bijection, we define the inverse map

$$\varphi_m\colon\ \mathcal{Z}_m(n) \longrightarrow \mathcal{R}_m(n+m-1)$$

by $\varphi_m(T) = S$, where S is a diagram on $[n+m-1]$ whose arcs are obtained by expanding each arc (i, j) of T to an arc $(i, j+m-1)$. Since $j - i \geq 1$ for each arc (i, j) in T, we have $j + m - 1 - i \geq m$, which says that S is m-regular.

Next, to prove that S is linear, we notice that for any vertex i in S,

$$\mathrm{ldeg}_S(i) = \mathrm{ldeg}_T(i-m+1), \tag{3.4}$$

$$\mathrm{rdeg}_S(i) = \mathrm{rdeg}_T(i). \tag{3.5}$$

When $1 \leq i \leq m-1$, it is easy to see that $\mathrm{ldeg}_S(i) = 0$ since S is m-regular. Then $\deg_S(i) = \mathrm{rdeg}_T(i) \leq 2$. When $m \leq i \leq n$, by using Condition (1), we find that

$$\deg_S(i) = \mathrm{ldeg}_S(i) + \mathrm{rdeg}_S(i) = \mathrm{ldeg}_T(i-m+1) + \mathrm{rdeg}_T(i) \leq 2.$$

When $n+1 \leq i \leq n+m-1$, it is clear that $\mathrm{rdeg}_S(i) = 0$, which implies that $\deg_S(i) = \mathrm{ldeg}_T(i-m+1) \leq 2$. Thus the degree of each vertex in S is bounded by 2, namely, S is linear.

To prove S is an m-regular linear stack, we still need to show that S is a stack. Otherwise, suppose that S contains two crossing arcs (i_1, j_1) and (i_2, j_2) with $i_1 < i_2 < j_1 < j_2$. Thus we get two arcs $(i_1, j_1 - m + 1)$ and $(i_2, j_2 - m + 1)$ in T. If $j_1 - m + 1 > i_2$, then $(i_1, j_1 - m + 1)$ and $(i_2, j_2 - m + 1)$ form a crossing in T, contradicting the assumption that T is a stack. If $j_1 - m + 1 = i_2$, then for vertex i_2, we have $\mathrm{ldeg}_T(i_2) > 0$ and $\mathrm{rdeg}_T(i_2) > 0$, contradicting the assumption that T is a zigzag stack. If $j_1 - m + 1 < i_2$, then we have $\mathrm{ldeg}_T(j_1 - m + 1) > 0, \mathrm{rdeg}_T(i_2) > 0$ and $i_2 - (j_1 - m + 1) < j_1 - (j_1 - m + 1) = m - 1$. But this violates Condition (2) for T. So we conclude that by no means does S contain a crossing, namely, S is a stack.

We now have shown that S is an m-regular linear stack, that is, φ_m is well-defined. It is easily seen that $\varphi_m(\theta_m(S)) = S$. Hence θ_m is a bijection and the proof is complete. ∎

4 Decomposition of an m-reduced zigzag stack

Based on the bijection given in Theorem 3.1, we can transform an m-regular linear stack to an m-reduced zigzag stack. The advantage of enumerating m-reduced zigzag stack lies in that there is no restriction on the arc lengths. We apply the primary component decomposition to enumerate m-reduced zigzag stacks.

Let u and v $(u < v)$ be two adjacent vertices in the primary connected component of S. We use $\langle u, v\rangle$ to denote the interval of integers between u and v, that is, $\{u + 1, u + 2, \ldots, v - 1\}$. Note that $\langle u, v\rangle$ is allowed to be empty. We use this notation $\langle u, v\rangle$ to distinguish with the notation (u, v) for an arc. If the primary component C does not include the last vertex n of S, then the vertices after the last vertex of C also form an interval. Since the intervals we are concerned with are determined by the primary component C, we call such intervals C-intervals. Note that the C-intervals are allowed to be empty. However, if C contains the last vertex n, we do not consider the empty set after the vertex n as an interval. In this case, if C contains k vertices, then there are $k - 1$ C-intervals.

Figure 4.1 illustrates a primary component decomposition of a 3-reduced zigzag stack S. The primary component C is a connected zigzag stack on $\{1, 7, 9, 13\}$, it decomposes S into four intervals $I_1 = \{2, 3, 4, 5, 6\}$, $I_2 = \{8\}$, $I_3 = \{10, 11, 12\}$ and the last interval is $I_4 = \{14, 15, 16, 17\}$.

Let S be a zigzag stack on $[n]$, and let C be the primary component of S. Let $I_1, I_2, \ldots, I_k$ be the C-intervals of S. Define the sets $J_1, J_2, \ldots, J_k$ as follows. There are two cases. Case 1: C contains the vertex n. Assume that for $1 \le t \le k$, I_t is the interval between the vertices u_t and v_t in C. Then let $J_t = \{u_t, u_t + 1, \ldots, v_t\}$ for each t. Case 2: C does not contain the vertex n. Assume that for $1 \le t \le k - 1$, I_t is the interval between the vertices u_t and v_t in C and that I_k is the interval after the vertex u_k. Then let $J_t = \{u_t, u_t + 1, \ldots, v_t\}$ for $1 \le t \le k - 1$ and let $J_k = \{u_k, u_k + 1, \ldots, n\}$.

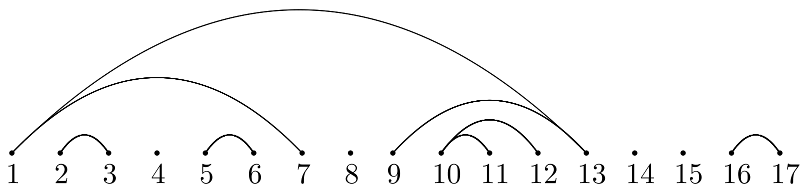

Figure 4.1: The primary component decomposition of a 3-reduced zigzag stack.

For example, in Figure 4.1, we see that $J_1 = \{1, 2, 3, 4, 5, 6, 7\}$, $J_2 = \{7, 8, 9\}$, $J_3 = \{9, 10, 11, 12, 13\}$ and $J_4 = \{13, 14, 15, 16, 17\}$.

Lemma 2.1 shows that a zigzag stack S can be decomposed into a primary component and a list of zigzag stacks. Let $S_1, S_2, \ldots, S_k$ be the zigzag stacks on the C-intervals $I_1, I_2, \ldots, I_k$. If S is an m-reduced zigzag stack, it can be seen that all the substructures S_t are m-reduced zigzag stacks. However, the converse may not be true. For example, in Figure 4.2, the substructure on $\{4, 5, 6\}$ is a 3-reduced zigzag stack, but the stack on $\{1, 2, 3, 4, 5, 6, 7\}$ is not a 3-reduced zigzag stack, because for the vertices 3 and 4, we have $\mathrm{ldeg}(3) > 0$, $\mathrm{rdeg}(4) > 0$, but $4 - 3 = 1 < 2$, contradicting the condition $j - i \geq m - 1$.

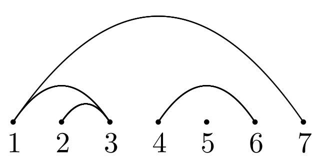

Figure 4.2: A zigzag stack on [7].

The following lemma gives a necessary and sufficient condition for a zigzag stack S to be m-reduced. It shows that the primary component decomposition of S can be used to restrict the verification of the following degree conditions to pairs of vertices in each set J_t, where $1 \leq t \leq k$:

(1) For $1 \leq i \leq n - m + 1$, $\mathrm{ldeg}(i) + \mathrm{rdeg}(i + m - 1) \leq 2$;

(2) For $1 \leq i < j \leq n$, if $\mathrm{ldeg}(i) > 0$ and $\mathrm{rdeg}(j) > 0$, then $j - i \geq m - 1$.

Lemma 4.1. *Let S be a zigzag stack on $[n]$, and let C be the primary component of S. Let $I_1, I_2, \ldots, I_k$ be the C-intervals of S, and $J_1, J_2, \ldots, J_k$ be the subsets of $[n]$ as given before. Then S is an m-reduced zigzag stack if and only if Conditions* (1) *and* (2) *hold for each J_t with respect to the degrees in S.*

Proof. Clearly, we only need to show that if Conditions (1) and (2) hold for each J_t, then they hold for any pair of vertices in $[n]$. The proof consists of two claims. Claim 1. If Condition (1) holds for each J_t, then the violation of Condition (1) on $[n]$ implies the violation of Condition (2) on $[n]$. Claim 2. If Condition (2) holds for each J_t, then Condition (2) holds for any pair of vertices in $[n]$. Combining these two claims, we see that if Conditions (1) and (2) hold for each J_t, then they hold on $[n]$, and hence S is an m-reduced zigzag stack.

To prove Claim 1, we suppose that there exist two vertices i and $i+m-1$ such that $\mathrm{ldeg}(i)+\mathrm{rdeg}(i+m-1) > 2$. Since $\mathrm{ldeg}(i), \mathrm{rdeg}(i+m-1) \leq 2$, we see that $\mathrm{ldeg}(i) > 0$ and $\mathrm{rdeg}(i+m-1) > 0$. On the other hand, i and $i+m-1$ cannot be contained in the same interval J_t. Otherwise, Condition (1) would hold for i and $i+m-1$, a contradiction. Thus there exists at least one vertex u in C such that $i < u < i+m-1$. It follows that either $\mathrm{ldeg}(u) > 0$ or $\mathrm{rdeg}(u) > 0$ since C is a connected zigzag stack. If $\mathrm{ldeg}(u) > 0$, then $(i+m-1)-u < m-1$; If $\mathrm{rdeg}(u) > 0$, then $u-i < m-1$. In either case, Condition (2) is violated.

To verify Claim 2, suppose to the contrary that i and j $(1 \leq i < j \leq n)$ are two vertices that violate Condition (2), namely, $\mathrm{ldeg}(i) > 0$, $\mathrm{rdeg}(j) > 0$ but $j-i < m-1$. We further assume that the vertices i and j are chosen so that $j-i$ is minimum. Since Condition (2) holds for each J_t, i and j cannot be contained in the same interval J_t. Hence there exists at least one vertex u in C such that $i < u < j$. Noting that C is a connected zigzag stack, we have either $\mathrm{ldeg}(u) > 0$ or $\mathrm{rdeg}(u) > 0$. If $\mathrm{ldeg}(u) > 0$, then $j-u < j-i$; If $\mathrm{rdeg}(u) > 0$, then $u-i < j-i$. In either case, this contradicts the choice of i and j. So we conclude that any pair of vertices in S satisfies Condition (2). This completes the proof. ∎

The above lemma shows that an m-reduced zigzag stack can be decomposed into a primary component along with a list of m-reduced zigzag stacks on the intervals. We observe that the primary components C has six patterns as shown in Figure 4.3, where the structure (a) and (b) in Figure 4.4 stand for connected zigzag stacks with at least three vertices.

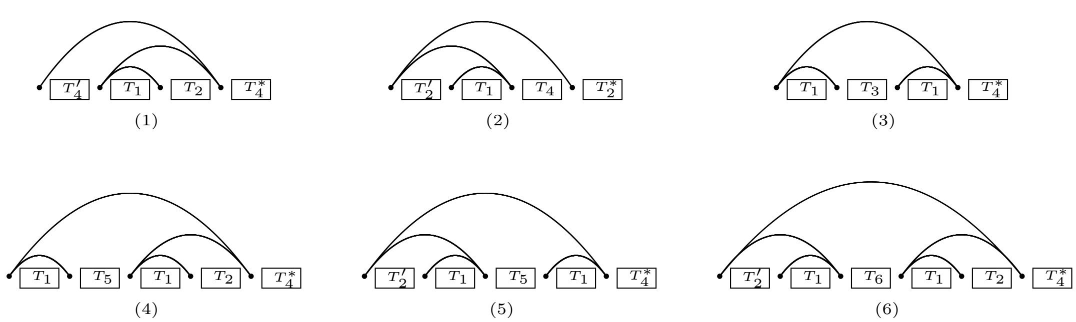

Figure 4.3: The primary component decompositions of m-reduced zigzag stacks.

(a) (b)

Figure 4.4: Substructures with at least three vertices.

We further classify the substructures on the intervals split by C into six classes, see Figure 4.5. The types of intervals lead us to define types of m-reduced zigzag stacks.

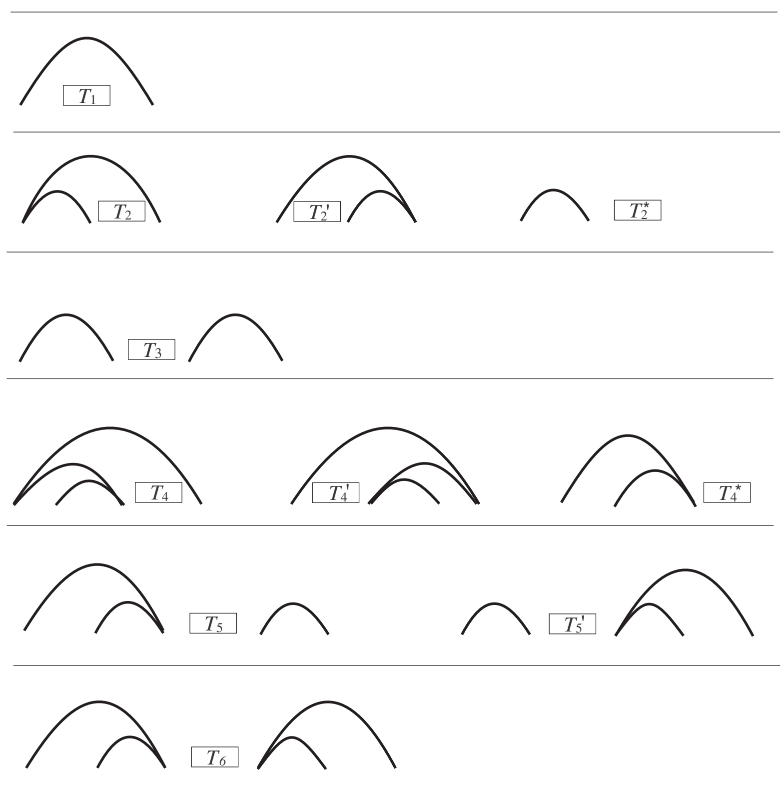

Figure 4.5: Types of intervals.

If S is a zigzag stack on an interval of type T, then we say that S is also of type T.

Let S be an m-reduced zigzag stack on $[n]$, and let C be the primary component of S. For given C, the types of the intervals created by C are determined. Assume that C has k vertices, that create k intervals which are allowed to be empty. In particular, if C containing the vertex n, then the k-th interval becomes empty. Let us discuss the case as shown in Figure 4.6 in detail. The other cases are similar.

When $k = 2$, it is easily seen that the two intervals are of types T_1 and T_2^* respectively. When $k \geq 3$, the types of the substructures are described below.

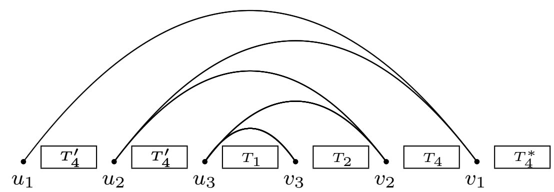

Figure 4.6: Decomposition of m-reduced zigzag stacks with $\deg(1) = 1$.

Lemma 4.2. *Let S be an m-reduced zigzag stack with* $\deg(1) = 1$*, and let C denote the primary component of S. Assume that C contains k vertices with $k \geq 3$. Suppose that the arcs of C are denoted by $e_1 = (u_1, v_1)$, $e_2 = (u_2, v_1)$, $e_3 = (u_2, v_2)$, $\ldots$, $e_{k-1} = \big(u_{[\frac{k+1}{2}]}, v_{[\frac{k}{2}]}\big)$. Then the substructures on the intervals can be described as follows:*

(1) *If k is even, then the substructure on the interval $\langle u_{\frac{k}{2}}, v_{\frac{k}{2}}\rangle$ is of type T_1, the one on the interval $\langle v_{\frac{k}{2}}, v_{\frac{k}{2}-1}\rangle$ is of type T_2, the ones on the intervals $\langle u_i, u_{i+1}\rangle$ for $1 \leq i \leq \frac{k}{2} - 1$ are of type T_4', the ones on the intervals $\langle v_j, v_{j-1}\rangle$ for $2 \leq j \leq \frac{k}{2} - 1$ are of type T_4, and the one on the interval after the vertex v_1 is of type T_4^*.*

(2) *If k is odd, then the substructure on the interval $\langle u_{\frac{k+1}{2}}, v_{\frac{k-1}{2}}\rangle$ is of type T_1, the one on the interval $\langle u_{\frac{k-1}{2}}, u_{\frac{k+1}{2}}\rangle$ is of type T_2', the ones on the intervals $\langle u_i, u_{i+1}\rangle$ for $1 \leq i \leq \frac{k-1}{2} - 1$ are of type T_4', the ones on the intervals $\langle v_j, v_{j-1}\rangle$ for $2 \leq j \leq \frac{k-1}{2}$ are of type T_4, and the one on the interval after the vertex v_1 is of type T_4^*.*

Proof. We only consider the case when k is even. It is clear that $\mathrm{ldeg}_S(u_i) = 0$ for any $1 \leq i \leq k/2$, and $\mathrm{rdeg}_S(u_1) = 1$, $\mathrm{rdeg}_S(u_i) = 2$ for $2 \leq i \leq k/2$. Similarly, we have $\mathrm{ldeg}_S(v_j) = 2$ for $1 \leq j \leq k/2 - 1$, $\deg_S(v_{\frac{k}{2}}) = 1$ and $\mathrm{rdeg}_S(v_j) = 0$ for $1 \leq j \leq k/2$. Thus for the interval $\langle u_{\frac{k}{2}}, v_{\frac{k}{2}}\rangle$, we have $\mathrm{ldeg}_S\big(u_{\frac{k}{2}}\big) = 0$ and $\mathrm{rdeg}_S\big(v_{\frac{k}{2}}\big) = 0$. By Lemma 4.1, we see that the substructure of S on this interval is of type T_1. In other words, the substructure on $\langle u_{\frac{k}{2}}, v_{\frac{k}{2}}\rangle$ can be any m-reduced zigzag stack.

For the interval $\langle v_{\frac{k}{2}}, v_{\frac{k}{2}-1}\rangle$, we have $\mathrm{ldeg}_S\big(v_{\frac{k}{2}}\big) = 1$ and $\mathrm{rdeg}_S\big(v_{\frac{k}{2}-1}\big) = 0$. Thus this interval is of type T_2 and the substructure on this interval is also of type T_2.

Similarly, the other types of the substructures also can be determined from the degrees of endpoints of the corresponding intervals. This completes the proof. ∎

By Lemma 4.1, we see that if S is the substructure in a C-interval $\langle u, v\rangle$, then S is an m-reduced zigzag stack restricted by $\mathrm{ldeg}(u)$ and $\mathrm{rdeg}(v)$; and if S is the substructure in the interval after the last vertex w of the primary component C, then S is an m-reduced zigzag stack restricted by $\mathrm{ldeg}(w)$. For type T_2, since $\mathrm{rdeg}(v) = 0$, then S is determined by $\mathrm{ldeg}(u)$. Noting that $\mathrm{ldeg}(u) = \mathrm{ldeg}(w) = 1$ for types T_2 and T_2^*, this implies that the substructures of types T_2 and T_2^* are of the same class of m-reduced zigzag stacks. Since $\mathrm{ldeg}(u) = \mathrm{ldeg}(w) = 2$ and $\mathrm{rdeg}(v) = 0$ for types T_4 and T_4^*, we see that the substructures of types T_4 and T_4^* are also of the same class of m-reduced zigzag stacks. Moreover, if S is of type T_i, then the reflection or reversal of S is of type T_i' $(i = 2, 4, 5)$.

For $1 \leq i \leq 6$, let $t_i(n)$ denote the number of m-reduced zigzag stacks of type T_i on $[n]$. Similarly, for $i = 2, 4, 5$, let $t_i'(n)$ denote the number of m-reduced zigzag stacks of type T_i' on $[n]$. Moreover, for $i = 2, 4$, let $t_i^*(n)$ denote the number of m-reduced zigzag stacks of type T_i^* on $[n]$. It is clear that $t_i(n) = t_i'(n)$ for $i = 2, 4, 5$. We have shown that $t_i(n) = t_i^*(n)$ for $i = 2, 4$. So we shall identify the types T_i' and T_i^* as T_i and we will restrict our attention to substructures of type T_i $(1 \leq i \leq 6)$. For $1 \leq i \leq 6$, let

$$T_i(x) = \sum_{n=0}^{\infty} t_i(n) x^n.$$

From the primary component decompositions of m-reduced zigzag stacks, we find that $Z_m(x)$ can be expressed in terms of $T_i(x)$.

Theorem 4.1. *We have*

$$(1-x)Z_m = 1 + \frac{x^2 T_1 T_2}{1 - xT_4} + \frac{x^3 T_1 T_2^2}{1 - xT_4} + x^4 T_1^2 T_3 T_4 + \frac{2x^5 T_1^2 T_2 T_4 T_5}{1 - xT_4} + \frac{x^6 T_1^2 T_2^2 T_4 T_6}{1 - xT_4}, \tag{4.1}$$

where Z_m and T_i stand for the generating functions $Z_m(x)$ and $T_i(x)$, respectively.

Proof. Recall that $z_m(n)$ denotes the number of m-reduced zigzag stacks on $[n]$. For $n = 0$, we set $z_m(0) = 1$. For $n \geq 1$, let $u_0(n)$ denote the number of these m-reduced zigzag stacks such that $\deg(1) = 0$, and define

$$U_0(x) = \sum_{n=0}^{\infty} u_0(n) x^n.$$

If 1 is an isolated vertex in an m-reduced zigzag stack S, we get an m-reduced zigzag stack of length $n-1$ by deleting 1. Thus $u_0(n) = z_m(n-1)$ and

$$U_0(x) = xZ_m(x).$$

We next divide m-reduced zigzag stacks with $\deg(1) \geq 1$ into six classes $\mathcal{U}_i(n)$ $(1 \leq i \leq 6)$ according to the patterns of the primary components, see Figure 4.3. For each $1 \leq i \leq 6$, denote the numbers of the m-reduced zigzag stacks in $\mathcal{U}_i(n)$ by $u_i(n)$ and define

$$U_i(x) = \sum_{n=0}^{\infty} u_i(n) x^n.$$

Let $u_i(n,k)$ denote the number of m-reduced zigzag stacks in $\mathcal{U}_i(n)$ such that the primary component C contains k vertices. Assume that C contains k vertices $v_1, v_2, \ldots, v_k$.

Case (1): $\deg(1) = 1$. By Lemma 2.2, C is of the form as shown in Figure 4.3. The primary component C creates k intervals. Notice that if $v_k = n$, we still consider the empty set after the vertex n as an interval. By Lemma 4.2, we see that among these k intervals , there are one interval of type T_1, one interval of type T_2, and $k-2$ intervals of type T_4. Thus

$$u_1(n,k) = \sum_{d_1+\cdots+d_k=n-k} t_1(d_1)t_2(d_2)t_4(d_3)t_4(d_4)\cdots t_4(d_k),$$

where d_i are nonnegative integers. Therefore,

$$\begin{aligned}
U_1(x) &= \sum_{n=2}^{\infty}\sum_{k=2}^{n} u_1(n,k)x^n \\
&= \sum_{n=2}^{\infty}\sum_{k=2}^{n}\sum_{d_1+\cdots+d_k=n-k} t_1(d_1)t_2(d_2)t_4(d_3)\cdots t_4(d_k)x^n \\
&= \sum_{k=2}^{\infty} x^k \sum_{n=0}^{\infty} x^n \sum_{d_1+\cdots+d_k=n} t_1(d_1)t_2(d_2)t_4(d_3)\cdots t_4(d_k) \\
&= \sum_{k=2}^{\infty} x^k \left(\sum_{d_1\geq 0} t_1(d_1)x^{d_1} \sum_{d_2\geq 0} t_2(d_2)x^{d_2} \sum_{d_3\geq 0} t_4(d_3)x^{d_3} \ldots \sum_{d_k\geq 0} t_4(d_k)x^{d_k} \right) \\
&= \sum_{k=2}^{\infty} x^k T_1(x)T_2(x)T_4^{k-2}(x) \\
&= \frac{x^2T_1(x)T_2(x)}{1-xT_4(x)}.
\end{aligned}$$

Case (2): $\deg(1) = 2$ and $\deg(v_k) = 1$. In this case, we have $k \geq 3$ and C creates k intervals, in which there are one interval of type T_1, two of type T_2 and $k-3$ of type T_4. It yields that

$$u_2(n,k) = \sum_{d_1+\cdots+d_k=n-k} t_1(d_1)t_2(d_2)t_2(d_3)t_4(d_4)t_4(d_5)\cdots t_4(d_k),$$

where d_i ranges over nonnegative integers. Hence

$$U_2(x) = \frac{x^3T_1T_2^2}{1-xT_4}.$$

Case (3): $k = 4$, $(1,u)$, $(1,w)$ and (v,w) are arcs of S such that $u < v < w \leq n$ and $\deg(u) = \deg(v) = 1$. These four vertices $1,u,v,w$ create four intervals $\langle 1,u\rangle$, $\langle u,v\rangle$, $\langle v,w\rangle$ and the interval after the vertex w. In these intervals, there are two of type T_1, one of type T_3 and one of type T_4. Thus we get

$$u_3(n) = \sum_{d_1+d_2+d_3+d_4=n-4} t_1(d_1)t_1(d_2)t_2(d_3)t_4(d_4),$$

where d_1, d_2, d_3 and d_4 are nonnegative integers. It follows that

$$U_3(x) = x^4 T_1^2 T_3 T_4.$$

Cases (4) and (5) are symmetric, thus $u_4(n) = u_5(n)$ and $U_4(x) = U_5(x)$. Let us consider Case (4), which means that $(1, v_2)$, $(1, v_k)$ and (v_3, v_k) are arcs of S such that $1 < v_2 < v_3 < v_k \le n$, $\deg(v_2) = 1$ and $\deg(v_3) = 2$. In this case $k \ge 5$ and the vertices of C create k intervals, where the interval after v_k is allowed to be empty. Among these k intervals, there are two of type T_1, one of type T_2, one of type T_5 and $k-4$ of type T_4. So we have

$$u_4(n,k) = \sum_{d_1+\cdots+d_k=n-k} t_1(d_1)t_1(d_2)t_2(d_3)t_5(d_4)t_4(d_5)\cdots t_4(d_k),$$

where $d_1, d_2, \ldots, d_k$ are nonnegative integers. Therefore,

$$U_4(x) = \frac{x^5 T_1^2 T_2 T_4 T_5}{1 - xT_4}.$$

Case (6): $\deg(1) = 2$, $(1, v_i)$, $(1, v_k)$, (v_{i+1}, v_k) are arcs of S such that $1 < v_i < v_{i+1} < v_k$ and $\deg(v_i) = \deg(v_{i+1}) = 2$. In this case $k \ge 6$ and $3 \le i \le k-3$. When i is given, these k vertices create k intervals. It can be seen that among these k intervals, there are two of type T_1, two of type T_2, one of type T_6 and $k-5$ of type T_4. Thus,

$$u_6(n,k) = \sum_{d_1+\cdots+d_k=n-k} (k-5)t_1(d_1)t_1(d_2)t_2(d_3)t_2(d_4)t_6(d_5)t_4(d_6)\cdots t_4(d_k),$$

which implies

$$U_6(x) = \frac{x^6 T_1^2 T_2^2 T_4 T_6}{1 - xT_4}.$$

In summary, we find that

$$\begin{aligned} Z_m(x) &= 1 + U_0(x) + U_1(x) + U_2(x) + U_3(x) + U_4(x) + U_5(x) + U_6(x) \\ &= 1 + xZ_m + \frac{x^2 T_1 T_2}{1 - xT_4} + \frac{x^3 T_1 T_2^2}{1 - xT_4} + x^4 T_1^2 T_3 T_4 + \frac{2x^5 T_1^2 T_2 T_4 T_5}{1 - xT_4} + \frac{x^6 T_1^2 T_2^2 T_4 T_6}{1 - xT_4}, \end{aligned} \tag{4.2}$$

which completes the proof. ∎

5 Generating functions of substructures

In this section, we analyze the substructures of types T_i $(1 \le i \le 6)$ and derive their generating functions in terms of $Z_m(x)$. To do so, we shall consider two classes of

m-reduced zigzag stacks, which will be called m-reduced zigzag stacks of types G and H.

More precisely, an m-reduced zigzag stack on $[n]$ is said to be of type G if $\deg(1) \leq 1$, or of type H if $\deg(1) \leq 1$ and $\deg(n) \leq 1$. For example, given the 3-reduced zigzag stack in Figure 5.1, the zigzag stack on $\{9, 10, 11\}$ is of type G. The zigzag stack on $\{14, 15, 16\}$ is also of type G. Moreover, the substructure on $\{2, 3, 4, 5, 6\}$ is of type H.

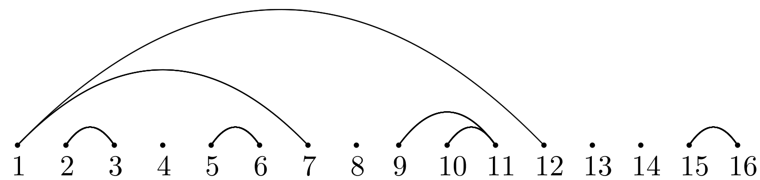

Figure 5.1: A 3-reduced zigzag stack.

We find that the substructure on each interval of type T_i $(1 \leq i \leq 6)$ can be characterized as follows.

Theorem 5.1. *Let S be an m-reduced zigzag stack. Let I be an interval of S of type T_i $(1 \leq i \leq 6)$. Let T be the substructure of S on the interval I. Then T can be described as follows.*

(1) *If T is of type T_1, then T can be any m-reduced zigzag stack.*

(2) *If T is of type T_2, then T may be empty, or consists of not more than $m-3$ isolated vertices, or $m-2$ isolated vertices followed by an m-reduced zigzag stack of type G.*

(3) *If T is of type T_3, then T consists of k isolated vertices with $m-2 \leq k \leq 2m-5$, or T is a substructure beginning with $m-2$ isolated vertices followed by an m-reduced zigzag stack of type H and ending with $m-2$ isolated vertices.*

(4) *If T is of type T_4, then T is empty, or consists of not more than $m-2$ isolated vertices, or $m-1$ isolated vertices followed by an m-reduced zigzag stack.*

(5) *If T is of type T_5, then T can be k isolated vertices with $m-2 \leq k \leq 2m-4$, or $m-1$ isolated vertices followed by an m-reduced zigzag stack of type G and $m-2$ isolated vertices.*

(6) *If T is of type T_6, then T can be k isolated vertices with $m-2 \leq k \leq 2m-3$, or $m-1$ isolated vertices followed by an m-reduced zigzag stack and $m-1$ isolated vertices.*

Proof. We assume that $I = \langle u, v\rangle$, and denote I by $\{w_1, w_2, \ldots, w_\ell\}$ with $w_1 = u+1$, $w_2 = u+2$, …, $w_\ell = v-1$.

If T is of type T_1, then $\mathrm{ldeg}_S(u) = \mathrm{rdeg}_S(v) = 0$. By Lemma 4.1, Conditions (1) and (2) hold on the interval $J = \{u, w_1, \ldots, w_\ell, v\}$. For vertices u and v, since $\mathrm{ldeg}_S(u) = \mathrm{rdeg}_S(v) = 0$, there is no restriction with respect to Condition (2). On the other hand, Condition (1) becomes $\mathrm{rdeg}_S(u+m-1) \leq 2$ and $\mathrm{ldeg}_S(v-m+1) \leq 2$, which are automatically satisfied since S is zigzag. Hence Conditions (1) and (2) hold on $I = \{w_1, \ldots, w_\ell\}$. In other words, T can be any m-reduced zigzag stack.

If T is of type T_2, we have $\mathrm{ldeg}_S(u) = 1$ and $\mathrm{rdeg}_S(v) = 0$. If $\ell < m-2$, we claim that $w_1, w_2, \ldots, w_\ell$ are isolated vertices. Otherwise, by Lemma 2.1, there exists an arc (w_j, w_k) in T. It follows that $\mathrm{rdeg}_S(w_j) > 0$. Now we have $\mathrm{ldeg}_S(u) > 0$ and $\mathrm{rdeg}_S(w_j) > 0$, but $w_j - u < m-2$, which contradicts Condition (2). If $\ell \geq m-2$, by the above claim we see that $\deg_S(w_i) = 0$ for $1 \leq i \leq m-2$. Moreover, by Condition (1), we find that $\mathrm{ldeg}_S(u) + \mathrm{rdeg}_S(w_{m-1}) \leq 2$, so that $\mathrm{rdeg}_S(w_{m-1}) \leq 1$. Consequently, the substructure on $\{w_{m-1}, w_m, \ldots, w_\ell\}$ is an m-reduced zigzag stack of type G. Considering all possible structures of T when T is of type T_2, we write

$$T_2 = \emptyset,\ \bullet,\ \bullet\bullet,\ \cdots,\ \underbrace{\bullet\bullet\cdots\bullet}_{m-3},\ \underbrace{\bullet\bullet\cdots\bullet}_{m-2} G.$$

If T is of type T_3, we have $\mathrm{ldeg}_S(u) = \mathrm{rdeg}_S(v) = 1$. If $\ell < m-2$, then $\mathrm{ldeg}_S(u) > 0, \mathrm{rdeg}_S(v) > 0$ but $v - u = \ell + 1 < m-1$, which violates Condition (2). Thus ℓ is at least $m-2$. If $m-2 \leq \ell < 2m-4$, we claim that $w_1, w_2, \ldots, w_\ell$ are isolated vertices. Otherwise, by Lemma 2.1, there exists an arc (w_j, w_k). It follows that $\mathrm{rdeg}_S(w_j) > 0$ and $\mathrm{ldeg}_S(w_k) > 0$. Since $\mathrm{ldeg}_S(u) > 0$ and $\mathrm{rdeg}_S(w_j) > 0$, Condition (2) implies that $w_j - u \geq m-1$. Furthermore, since $\mathrm{ldeg}_S(w_k) > 0$ and $\mathrm{rdeg}_S(v) > 0$, we see that $v - w_k \geq m-1$. Hence

$$\ell = v - u - 1 > (v - w_k) + (w_j - u) - 1 \geq 2m - 3,$$

contradicting the assumption that $\ell < 2m-4$. This proves the claim.

If $\ell \geq 2m-4$, by the above claim, the first $m-2$ vertices $w_1, w_2, \ldots, w_{m-2}$ and the last $m-2$ vertices $w_{\ell-m+3}$, $w_{\ell-m+4}$, …, $w_{\ell-1}$, w_ℓ are isolated vertices. Moreover, according to Condition (1), we have $\mathrm{ldeg}_S(u) + \mathrm{rdeg}_S(w_{m-1}) \leq 2$ and $\mathrm{ldeg}_S(w_{\ell-m+2}) + \mathrm{rdeg}_S(v) \leq 2$. This yields $\mathrm{rdeg}_S(w_{m-1}) \leq 1$ and $\mathrm{ldeg}_S(w_{\ell-m+2}) \leq 1$, so that the substructure on $\{w_{m-1}, w_m, \ldots, w_{\ell-m+2}\}$ is an m-reduced zigzag stack of type H. Hence we get

$$T_3 = \underbrace{\bullet\bullet\cdots\bullet}_{m-2},\ \underbrace{\bullet\bullet\cdots\bullet}_{m-1}, \cdots,\ \underbrace{\bullet\bullet\cdots\bullet}_{2m-5},\ \underbrace{\bullet\bullet\cdots\bullet}_{m-2} H \underbrace{\bullet\bullet\cdots\bullet}_{m-2}.$$

Similarly, if T is of type T_4, T_5 or T_6, all possible structures of T are given below:

$$T_4 = \emptyset,\ \bullet,\ \bullet\bullet,\ \cdots,\ \underbrace{\bullet\bullet\cdots\bullet}_{m-2},\ \underbrace{\bullet\bullet\cdots\bullet}_{m-1} Z_m,$$

$$T_5 = \underbrace{\bullet\bullet\cdots\bullet}_{m-1},\ \underbrace{\bullet\bullet\cdots\bullet}_{m},\cdots,\ \underbrace{\bullet\bullet\cdots\bullet}_{2m-4},\ \underbrace{\bullet\bullet\cdots\bullet}_{m-1}\,G\,\underbrace{\bullet\bullet\cdots\bullet}_{m-2},$$

$$T_6 = \underbrace{\bullet\bullet\cdots\bullet}_{m-1},\ \underbrace{\bullet\bullet\cdots\bullet}_{m},\cdots,\ \underbrace{\bullet\bullet\cdots\bullet}_{2m-3},\ \underbrace{\bullet\bullet\cdots\bullet}_{m-1}\,Z_m\,\underbrace{\bullet\bullet\cdots\bullet}_{m-1}.$$

This completes the proof. ∎

Let $g(n)$ denote the number of m-reduced zigzag stacks of type G on $[n]$, and let $h(n)$ denote the number of m-reduced zigzag stacks of type H on $[n]$. The generating functions of $g(n)$ and $h(n)$ are defined as follows,

$$G(x) = \sum_{n\geq 0} g(n)x^n, \qquad H(x) = \sum_{n\geq 0} h(n)x^n.$$

By Theorem 5.1, we are led to expressions of $T_i(x)$ in terms of $Z_m(x)$, $G(x)$ and $H(x)$.

Theorem 5.2. *We have*

$$T_1(x) = Z_m(x), \tag{5.1}$$

$$T_2(x) = \frac{1-x^{m-2}}{1-x} + x^{m-2}G(x), \tag{5.2}$$

$$T_3(x) = \frac{x^{m-2}(1-x^{m-2})}{1-x} + x^{2m-4}H(x), \tag{5.3}$$

$$T_4(x) = \frac{1-x^{m-1}}{1-x} + x^{m-1}Z_m(x), \tag{5.4}$$

$$T_5(x) = \frac{x^{m-1}(1-x^{m-2})}{1-x} + x^{2m-3}G(x), \tag{5.5}$$

$$T_6(x) = \frac{x^{m-1}(1-x^{m-1})}{1-x} + x^{2m-2}Z_m(x). \tag{5.6}$$

Next, we consider the primary component decompositions of m-reduced zigzag s-tacks of types G and H. We obtain formulas for the generating functions $G(x)$ and $H(x)$ in terms of $Z_m(x)$ and $T_i(x)$ $(1 \leq i \leq 6)$.

Recall that S is of type G if $\deg_S(1) \leq 1$. If 1 is an isolated vertex, then S becomes an m-reduced zigzag stack of length $n-1$ after deleting the vertex 1. If $\deg_S(1) = 1$, by Lemma 2.2, C is of the form as shown in Figure 4.6. By Lemma 4.2, for $k \geq 2$, S can be decomposed into a connected zigzag stack C and a list of k m-reduced zigzag stacks. Based on this decomposition, we obtain a functional equation satisfied by $G(x)$, $Z_m(x)$, $T_2(x)$ and $T_4(x)$.

Lemma 5.1. *We have*

$$G(x) = 1 + xZ_m(x) + \frac{x^2 Z_m(x) T_2(x)}{1 - xT_4(x)}. \tag{5.7}$$

Proof. Let S be an m-reduced zigzag stack of type G on $[n]$, and let C be the primary component of S. Assume that C has k vertices, and S is decomposed into a connected zigzag stack C and a list of k zigzag stacks $S_1, S_2, \ldots, S_k$, where each S_i is allowed to be empty. Suppose that S_i has d_i vertices for $1 \le i \le k$, so that we have $d_1 + d_2 + \cdots + d_k = n - k$.

Let $g(n,k)$ denote the number of m-reduced zigzag stacks of type G whose primary component contains k vertices. Set $g(n,0) = 1$. When $k = 1$, by deleting the vertex 1 from S, we obtain an m-reduced zigzag stack of length $n-1$. Thus we have $g(n,1) = z_m(n-1)$.

For $k \ge 2$, by Lemma 4.2, C creates k intervals, in which there are one component of type T_1, one of type T_2 and $k-2$ of type T_4. It follows that

$$g(n,k) = \sum_{d_1+\cdots+d_k=n-k} t_1(d_1)t_2(d_2)t_4(d_3)\cdots t_4(d_k). \tag{5.8}$$

Since

$$g(n) = \sum_{k=0}^{n} g(n,k),$$

we find that

$$\begin{aligned}
G(x) &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} g(n,k)x^n \\
&= 1 + \sum_{n=0}^{\infty} g(n,1)x^n + \sum_{n=0}^{\infty}\sum_{k=2}^{n}\sum_{d_1+\cdots+d_k=n-k} t_1(d_1)t_2(d_2)t_4(d_3)\cdots t_4(d_k)x^n \\
&= 1 + \sum_{n=0}^{\infty} z_m(n-1)x^n + \sum_{k=2}^{\infty} x^k \sum_{n=0}^{\infty} x^n \sum_{d_1+\cdots+d_k=n} t_1(d_1)t_2(d_2)t_4(d_3)\cdots t_4(d_k) \\
&= 1 + xZ_m(x) + \frac{x^2 T_1(x) T_2(x)}{1 - xT_4(x)}.
\end{aligned}$$

Since $T_1(x) = Z_m(x)$, we arrive at expression (5.7) for $G(x)$. This completes the proof. ∎

For the m-reduced zigzag stacks of type H, we obtain the following relation for the generating function $H(x)$.

Lemma 5.2. *We have*

$$H(x) = 1 + x + 2xG(x) + x^4 Z_m^2(x)T_3(x) + \frac{2x^5 Z_m^2(x)T_2(x)T_5(x)}{1 - xT_4(x)} + \frac{x^6 Z_m^2(x)T_2^2(x)T_6(x)}{(1 - xT_4(x))^2}. \tag{5.9}$$

Proof. Let $\mathcal{H}(n)$ denote the set of m-reduced zigzag stacks of type H on $[n]$. Recall that $h(n)$ denotes the number of m-reduced zigzag stack in $\mathcal{H}(n)$. We set $h(0) = 1$. It is clear that $h(1) = 1$. For $n \geq 2$, according to the degrees of vertices 1 and n, we divide the set of m-reduced zigzag stacks of type H into four classes $\mathcal{F}_1(n)$, $\mathcal{F}_2(n)$, $\mathcal{F}_3(n)$ and $\mathcal{F}_4(n)$. For each $1 \leq i \leq 4$, denote the cardinality of $\mathcal{F}_i(n)$ by $f_i(n)$ and define the generating function as

$$F_i(x) = \sum_{n=0}^{\infty} f_i(n) x^n.$$

(1) $\mathcal{F}_1(n) = \{S \,|\, S \in \mathcal{H}(n) \text{ and } \deg(1) = 0, \deg(n) = 0\}$.

Deleting the vertices 1 and n from a stack in $\mathcal{F}_1(n)$, we obtain an m-reduced zigzag stack on $n-2$ vertices. Conversely, from any given m-reduced zigzag stack on $n-2$ vertices, we may generate a stack in $\mathcal{F}_1(n)$. Thus we have $f_1(n) = z_m(n-2)$ and the generating function

$$F_1(x) = x^2 Z_m(x).$$

(2) $\mathcal{F}_2(n) = \{S \,|\, S \in \mathcal{H}(n) \text{ and } \deg(1) = 1, \deg(n) = 0\}$.

By deleting the vertex n from a stack in $\mathcal{F}_2(n)$, we obtain an m-reduced zigzag stack of type G of length $n-1$ with $\deg(1) = 1$. This process is clearly reversible. By the definition of type G, we see that the number of m-reduced zigzag stacks of length $n-1$ with $\deg(1) \leq 1$ is $g(n-1)$. On the other hand, given an m-reduced zigzag stack of length $n-1$ with $\deg(1) = 0$, by deleting the vertex 1, we are led to an m-reduced zigzag stack of length $n-2$. Thus the number of m-reduced zigzag stacks of length $n-1$ with $\deg(1) = 0$ is $z_m(n-2)$. Hence,

$$f_2(n) = g(n-1) - z_m(n-2)$$

and

$$F_2(x) = xG(x) - x^2 Z_m(x).$$

(3) $\mathcal{F}_3(n) = \{S \,|\, S \in \mathcal{H}(n) \text{ and } \deg(1) = 0, \deg(n) = 1\}$.

By reversing the order of vertices $1, 2, \ldots, n$, we get a one-to-one correspondence between $\mathcal{F}_2(n)$ and $\mathcal{F}_3(n)$. Therefore, $f_3(n) = f_2(n)$ and

$$F_3(x) = xG(x) - x^2 Z_m(x).$$

(4) $\mathcal{F}_4(n) = \{S \,|\, S \in \mathcal{H}(n) \text{ and } \deg(1) = 1, \deg(n) = 1\}$.

In this case, we further divide $\mathcal{F}_4(n)$ into five classes $\mathcal{F}_{4,j}(n)$, where $1 \leq j \leq 5$, see Figure 5.2. For each $1 \leq j \leq 5$, let $f_{4,j}(n)$ be the number of m-reduced zigzag stacks in $\mathcal{F}_{4,j}(n)$ and define

$$F_{4,j}(x) = \sum_{n=0}^{\infty} f_{4,j}(n) x^n.$$

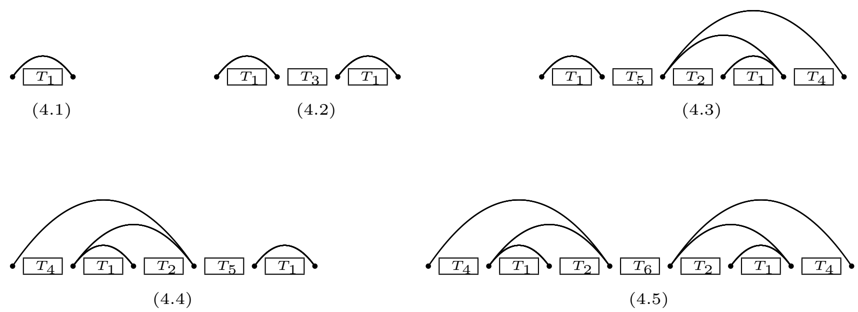

Figure 5.2: Five classes of $\mathcal{F}_4(n)$.

Case (4.1): $(1,n)$ is an arc of S. Then the interval $\langle 1,n\rangle$ is of type T_1. By Theorem 5.1, the substructure on $\langle 1,n\rangle$ is an m-reduced zigzag stack of length $n-2$. Thus $f_{4,1}(n) = z_m(n-2)$ and

$$F_{4,1}(x) = x^2 Z_m(x).$$

Case (4.2): $(1,u)$ and (v,n) are two arcs of S, where $1 < u < v < n$ such that $\deg(u) = \deg(v) = 1$. Consider the types of the three intervals. It can be seen that $\langle 1,u\rangle$ is of type T_1, $\langle u,v\rangle$ is of type T_3, and $\langle v,n\rangle$ is of type T_1. Consequently, for $n \geq 4$,

$$f_{4,2}(n) = \sum_{d_1+d_2+d_3=n-4} t_1(d_1)t_1(d_2)t_3(d_3),$$

where d_1, d_2 and d_3 are nonnegative integers. So we obtain that

$$F_{4,2}(x) = x^4 Z_m^2(x) T_3(x).$$

The Cases (4.3) and (4.4) are symmetric. So we have $f_{4,3}(n) = f_{4,4}(n)$ and $F_{4,3}(x) = F_{4,4}(x)$. Let us consider Case (4.3). In this case, the primary component contains exactly two vertices. Suppose that there are $k-2$ vertices in the connected component containing n, so that there exist k vertices in these two connected components, where $k \geq 5$. Ignoring the empty interval after the vertex n, there are $k-1$ intervals created by these k vertices. Using the previous arguments, we see that the first interval is of type T_1, the second is of type T_5, and among the other $k-3$ intervals, there are one interval of type T_1, one of type T_2, and $k-5$ of type T_4. Therefore, we get

$$f_{4,3}(n) = \sum_{k=5}^{n} \sum_{d_1+\cdots+d_{k-1}=n-k} t_1(d_1)t_1(d_2)t_2(d_3)t_5(d_4)t_4(d_5)t_4(d_6)\cdots t_4(d_{k-1}),$$

where $d_1, d_2, \ldots, d_{k-1}$ are nonnegative integers. It follows that

$$F_{4,3}(x) = F_{4,4}(x) = \frac{x^5 Z_m^2(x) T_2(x) T_5(x)}{1 - xT_4(x)}.$$

Case (4.5): Both the primary component and the connected component containing n have at least three vertices. Suppose that these two connected components have a total number of k vertices, where $k \geq 6$. Given k, the primary connected component may have i vertices, where $3 \leq i \leq k-3$. When i is also given, these k vertices create $k-1$ intervals, not mentioning the empty interval after the vertex n. It can be seen that among these $k-1$ intervals, there are two of type T_1, two of type T_2, $k-6$ of type T_4, and one of type T_6. Hence

$$f_{4,5}(n) = \sum_{k=6}^{n} \sum_{d_1+\cdots+d_{k-1}=n-k} (k-5)t_1(d_1)t_1(d_2)t_2(d_3)t_2(d_4)t_6(d_5)t_4(d_6)t_4(d_7)\cdots t_4(d_{k-1}),$$

which gives

$$F_{4,5}(x) = \frac{x^6 Z_m^2(x) T_2^2(x) T_6(x)}{(1-xT_4(x))^2}.$$

So we find that

$$\begin{aligned} F_4(x) &= F_{4,1}(x) + F_{4,2}(x) + \cdots + F_{4,5}(x) \\ &= x^2 Z_m(x) + x^4 Z_m^2(x) T_3(x) + \frac{2x^5 Z_m^2(x) T_2(x) T_5(x)}{1-xT_4(x)} + \frac{x^6 Z_m^2(x) T_2^2(x) T_6(x)}{(1-xT_4(x))^2}. \end{aligned}$$

Finally, we obtain that

$$H(x) = 1 + x + F_1(x) + F_2(x) + F_3(x) + F_4(x),$$

which leads to relation (5.9), and hence the proof is complete. ∎

Substituting the relations (5.7) and (5.9) into (5.1)–(5.6), we find that the generating functions $T_2(x)$, $T_3(x)$ and $T_5(x)$ can be expressed in terms of $Z_m(x)$.

Theorem 5.3. *We have*

$$T_2(x) = \frac{\big(1 - x^{m-1} + x^{m-1}(1-x)Z_m\big)\big(1 - 2x + x^m - x^m(1-x)Z_m\big)}{(1-x)\big(1-2x+x^m-2x^m(1-x)Z_m\big)}, \tag{5.10}$$

$$\begin{aligned} T_3(x) = x^{m-2}\Big(&x^{4m}(x-1)^3 Z_m^4 + 3x^{3m}(x-1)^2(x^m-2x+1)Z_m^3 \\ &+ x^{2m}(x-1)(3x^{2m} - 13x^{m+1} + 7x^m + 9x^2 - 5x - 1)Z_m^2 \\ &+ x^m(x^m-2x+1)(x^{2m} - 6x^{m+1} + 4x^m + 2x^2 + 2x - 3)Z_m \\ &+ (1-x^m)(x^m-2x+1)^2\Big) \\ &\Big/\Big((1-x)(1-x^m Z_m)\big(1-2x+x^m-2x^m(1-x)Z_m\big)^2\Big), \end{aligned} \tag{5.11}$$

$$T_5(x) = \frac{x^{m-1}\big(1 - x^{m-1} + x^{m-1}(1-x)Z_m\big)\big(1-2x+x^m-x^m(1-x)Z_m\big)}{(1-x)\big(1-2x+x^m-2x^m(1-x)Z_m\big)}. \tag{5.12}$$

We have shown that $T_i(x)$ $(1 \leq i \leq 6)$ can be represented in terms of $Z_m(x)$. These relations will be used in the next section to derive an equation on $Z_m(x)$.

6 The generating function of m-regular linear stacks

In this section, we use the relations between $Z_m(x)$ and $T_i(x)$ $(1 \leq i \leq 6)$ to derive an equation satisfied by $Z_m(x)$. Then we obtain an equation on the generating function $R_m(x)$ of the number $r_m(n)$ of m-regular linear stacks on $[n]$. For given m, we can deduce a recurrence relation and an asymptotic formula for $r_m(n)$. For $m = 3, 4, 5, 6$, we give asymptotic formulas for $r_m(n)$.

In Section 4, we find expression (4.1) of $Z_m(x)$ in terms of $T_i(x)$ $(1 \leq i \leq 6)$. Moreover, in Section 5, we have shown that the generating functions $T_i(x)$ $(1 \leq i \leq 6)$ can be expressed in terms of $Z_m(x)$. Combing these relations, we arrive at the following equation satisfied by $Z_m(x)$.

Theorem 6.1. *We have*

$$a_5(x)Z_m^5(x)+a_4(x)Z_m^4(x)+a_3(x)Z_m^3(x)+a_2(x)Z_m^2(x)+a_1(x)Z_m(x)+a_0(x)=0, \quad (6.1)$$

where

$$\begin{aligned}
a_0(x) &= (x-1)(x^m-2x+1)^2,\\
a_1(x) &= -2x^{3m+1}+x^{3m}+12x^{2m+2}-16x^{2m+1}+7x^{2m}-18x^{m+3}+36x^{m+2}\\
&\quad -28x^{m+1}+7x^m+4x^4-10x^3+12x^2-6x+1,\\
a_2(x) &= x^m(2x^{3m+1}-15x^{2m+2}+14x^{2m+1}-5x^{2m}+33x^{m+3}-60x^{m+2}+47x^{m+1}\\
&\quad -14x^m-16x^4+39x^3-45x^2+25x-5),\\
a_3(x) &= x^{2m}(x-1)(7x^{2m+1}-28x^{m+2}+22x^{m+1}-8x^m+24x^3-36x^2+27x-8),\\
a_4(x) &= x^{3m}(x-1)^2(9x^{m+1}-16x^2+11x-4),\\
a_5(x) &= 4x^{4m+1}(x-1)^3.
\end{aligned}$$

From the bijection between $\mathcal{R}_m(n+m-1)$ and $\mathcal{Z}_m(n)$ as given in Theorem 3.1, we see that

$$R_m(x) = 1+x+x^2+\cdots+x^{m-2}+x^{m-1}Z_m(x),$$

or equivalently,

$$Z_m(x) = \frac{(1-x)R_m(x)-(1-x^{m-1})}{(1-x)x^{m-1}}. \quad (6.2)$$

Substituting the above relation into equation (6.1) on $Z_m(x)$, we arrive at equation (1.6) given in Theorem 1.5.

To conclude this paper, we give some special cases of Theorem 1.5. When $m = 2$, equation (1.6) on $R_m(x)$ reduces to Müller and Nebel's equation (1.3) by replacing

Table 6.1: $r_m(n)$ for $m = 3, 4, 5, 6$.

n	1	2	3	4	5	6	7	8	9	10	11	12
$r_3(n)$	1	1	1	2	6	18	54	162	491	1509	4692	14729
$r_4(n)$	1	1	1	1	2	6	18	52	150	434	1263	3699
$r_5(n)$	1	1	1	1	1	2	6	18	52	148	422	1206
$r_6(n)$	1	1	1	1	1	1	2	6	18	52	148	420

$R_2(x)$ with $S(z) + 1$. For $m = 3, 4, 5, 6$, the asymptotic formulas for $r_m(n)$ have been given in Introduction.

The values of $r_m(n)$ for $m = 3, 4, 5, 6$ and $1 \leq n \leq 12$ are given in Table 6.1.

Acknowledgements. This work was supported by the 973 Project, the PCSIRT Project of the Ministry of Education, and the National Science Foundation of China.

References

[1] P.K. Agarwal, N.H. Mustafa and Y. Wang, Fast molecular shape matching using contact maps, J. Comput. Biol. 14 (2) (2007) 131–143.

[2] W.Y.C. Chen, E.Y.P. Deng and R.R.X. Du, Reduction of m-regular noncrossing partitions, European J. Combin. 26 (2) (2005) 237–243.

[3] W.Y.C. Chen, E.Y.P. Deng, R.R.X. Du, R.P. Stanley and C.H. Yan, Crossings and nestings of matchings and partitions, Trans. Amer. Math. Soc. 359 (2007) 1555–1575.

[4] G.M. Crippen, Enumeration of cubic lattice walks by contact class, J. Chem. Phys. 12 (24) (2000) 11065–11068.

[5] K.A. Dill, Dominant forces in protein folding, Biochemistry 29 (1990) 7133–7155.

[6] K.A. Dill, S. Bromberg, K. Yue, K.M. Fiebig, D.P. Yee, P.D. Thomas and H.S. Chan, Principles of protein folding: A perspective from simple exact models, Prot. Sci. 4 (1995) 561–602.

[7] E. Domany, Protein folding in contact map space, Physica A 288 (2000) 1–9.

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

[9] D. Goldman, S. Istrail and C.H. Papadimitriou, Algorithmic aspects of protein structure similarity, Proc. 40th IEEE Symp. Found. Comput. Sci., 1999, pp. 512–522.

[10] C. Höner zu Siederdissen, S.H. Bernhart, P.F. Stadler and I.L. Hofacker, A folding algorithm for extended RNA secondary structures, Bioinformatics 27 (13) (2011) 129–136.

[11] J. Hu, X. Shen, Y. Shao, C. Bystroff and M.J. Zaki, Mining protein contact maps, Proc. 2nd BIOKDD Workshop Data Mining Bioinformatics, 2002, pp. 3–10.

[12] S. Istrail and F. Lam, Combinatorial algorithms for protein folding in lattice models: a Survey of mathematical results, Commun. Inf. Syst. 9 (4) (2009) 303–346.

[13] E.Y. Jin, J. Qin and C.M. Reidys, Combinatorics of RNA structures with pseudoknots, Bull. Math. Biol. 70 (2008) 45–67.

[14] E.Y. Jin and C.M. Reidys, Asymptotic enumeration of RNA structures with pseudoknots, Bull. Math. Biol. 70 (2008) 951–970.

[15] A.P. Kamat and A.M. Lesk, Contact patterns between helices and strands of sheet define protein folding patterns, Proteins 66 (2007) 869–876.

[16] G. Lancia, R. Carr, B. Walenz and S. Istrail, 101 optimal PDB structure alignments: A branch-and-cut algorithm for the maximum contact map overlap problem, Proc. Fifth Annual Internat. Conf. Comput. Biol., 2001, pp. 193–202.

[17] R. Müller and M.E. Nebel, Combinatorics of RNA secondary structures with base triples, 2013, preprint.

[18] M.E. Nebel, Combinatorial properties of RNA secondary structures, J. Comput. Biol. 9(3) (2002) 541–573.

[19] C.L. Pierri, A. De Grassi and A. Turi, Lattices for ab initio protein structure prediction, Proteins 73(2) (2008) 351–361.

[20] B. Salvy and P. Zimmerman, GFUN: a Maple package for the manipulation of generating and holonomic functions in one variable, ACM T. Math. Software 20(2) (1994) 163–177.

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

[22] R.P. Stanley, Enumerative Combinatorics, Vol. 2, Cambridge University Press, Cambridge, 1999.

[23] M. Vendruscolo, I. Kanter, E. Domany, J. Lebowitz and B. Subramanian, Statistical properties of contact maps, Phys. Rev. E 59 (1) (1999) 977–984.

[24] M. Vendruscolo, E. Kussell and E. Domany, Recovery of protein structure from contact maps, Fold. Des. 2 (5) (1997) 295–306.

[25] M. Waterman and T. Smith, RNA secondary structure: a complete mathematical analysis, Math. Biosci. 42 (1978) 257–266.

[26] M. Zuker and D. Sankoff, RNA secondary structures and their prediction, Bull. Math. Biol. 46(4) (1984) 591–621.