J. Math. Anal. Appl. 383 (2011) 200–207

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Asymptotic enumeration of some RNA secondary structures

Hongmei Liu [a,*], Wensong Mou [b], You Zhang [a], Tianming Wang [c]

[a] School of Science, Dalian Nationalities University, Dalian 116600, China
[b] Dalian Economic & Technological Development Area Hospital, Dalian 116600, China
[c] School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China

ARTICLE INFO

Article history:
Received 17 May 2009
Available online 11 May 2011
Submitted by B.C. Berndt

Keywords:
RNA secondary structures
Asymptotic enumeration
Recursion formulae
Generating functions

ABSTRACT

In this paper, we derive recursions of some RNA secondary structures with certain properties under two new representations. Furthermore, by making use of methods of asymptotic analysis and generating functions we present asymptotic enumeration of these RNA secondary structures.

1. Introduction

RNA molecules are single stranded nucleic acids composed of four nitrogen bases: adenine (A), guanine (G), cytosine (C) and uracil (U). As Watson–Crick pairing rule shows, in RNA, A can pair with U, and C can pair with G. The *primary structure* of a RNA is a linear sequence of bases. For the *secondary structure*, we have the following definition.

Definition 1.1. (See [2,3,8].) A *secondary structure* is a vertex-labeled graph on n vertices with an adjacency matrix A fulfilling:

(i) $a_{i,i+1} = 1$ for $1 \leqslant i \leqslant n-1$;
(ii) for each i there is at most a single $k \neq i-1, i+1$ such that $a_{i,k} = 1$;
(iii) if $a_{i,j} = a_{k,l} = 1$ and $i < k < j$ then $i < l < j$.

We call an edge (i,k), $|i-k| \neq 1$ a *bond* or *base pair*. A vertex i connected only to $i-1$ and $i+1$ will be called *unpaired*. A vertex i is said to be *interior* to the base pair (k,l) if $k < i < l$. If, in addition, there is no base pair (p,q) such that $k < p < i < q$, we will say that i is *immediately interior* to the base pair (k,l).

A *stack* consists of subsequent base pairs $(p-k, q+k), (p-k+1, q+k-1), \ldots, (p,q)$ such that neither $(p-k-1, q+k+1)$ nor $(p+1, q-1)$ is a base pair. $k+1$ is the length of the stack and $(p-k, q+k)$ is the terminal base pair of the stack. The sequence $i+1, i+2, \ldots, j-1$ is a *loop*, if $i+1, i+2, \ldots, j-1$ are all unpaired and $a_{i,j} = 1$. The pair (i,j) is said to be the foundation of the loop. A *hairpin* is the longest sequence $i+1, i+2, \ldots, j-1$ containing exactly one loop such that $a_{i+1,j-1} = 1$ and $a_{i,j} = 0$. The paired points $i+1$ and $j-1$ will be called the foundation of the hairpin. See Fig. 1.

* Corresponding author.
E-mail address: hmliu99@yahoo.com.cn (H. Liu).

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2011.05.015

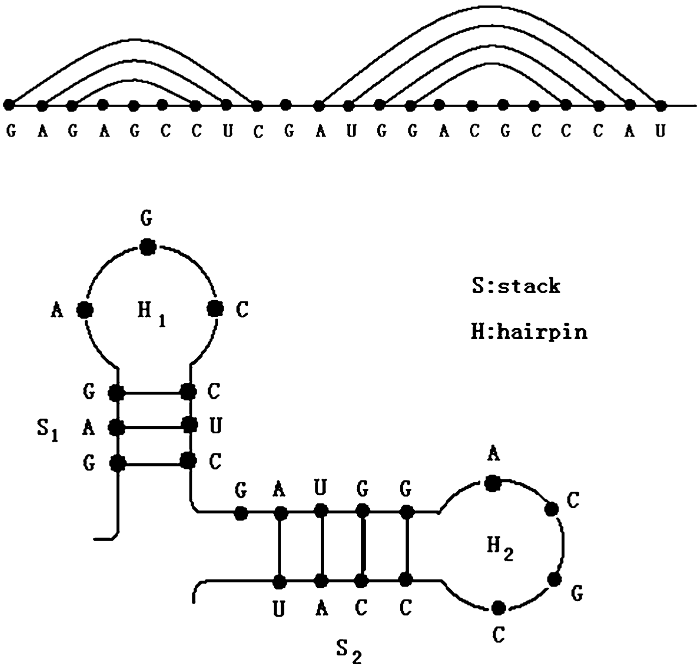

Fig. 1. RNA secondary structures.

Enumeration problems of RNA secondary structures of single-stranded nucleic acids have a long history starting with the investigations of Waterman [7,8] who gave the first formal framework for the topic [4]. Recently, there are many published work on the asymptotic enumeration of RNA secondary structures by making use of generating functions, for example in [3,2,6]. In the previous representation of RNA secondary structures, all bases A, U, G, C are regarded same. We call this representation *the first representation*. Under this representation, some asymptotic enumeration formulae on a variety subclasses and structural elements are derived in [2,3]. In [6], for the first time Wang et al. consider the difference between $A(U)$ and $G(C)$ bases. We call this representation *the second representation*. Based on the second representation, Wang et al. obtain some asymptotic enumeration formulae on RNA secondary structures with prescribed size m for hairpin loops and minimum stack length l.

In this paper, based on the second representation we present some recursions about some other secondary structures with a prescribed size m (for the biological reasons, throughout this paper, m denotes the minimum number of unpaired digits in a hairpin), for example, the structures with exactly b bases, the structures with exactly d hairpins and the structures with exactly r stacks. By means of generating functions, we derive their asymptotic enumeration. Furthermore, we define *the third representation* of RNA secondary structures by considering the bases A, U, G, C are all different. Under this representation, the number of structures with no base pairs on n vertices is 4^n. The asymptotic enumeration on the above mentioned structures are also obtained.

2. Asymptotic enumeration under the second representation

2.1. Basic recursions of secondary structure with certain properties

Let $H_n(b)$ denote the number of structures with exactly b base pairs on n vertices. The number $H_n(0)$ (i.e., the number of structures on n vertices without base pair) is 2^n, because each vertex is either the base $A(U)$ or the base $G(C)$. The number $H_{n+1}(b)$ may be obtained from a structure on n vertices either by adding a free end $A(U)$ or $G(C)$ at the right-hand end or by inserting a base pair $(1, k+2)$ which is AU or GC. In the first case, we obtain $2H_n(b)$ structures. In the second case, the substructure enclosed by the base pair $(1, k+2)$ is an arbitrary structure on k vertices with l base pairs, and the remainder is an arbitrary structure on $n-k-1$ vertices with $b-l-1$ base pairs. Therefore, we obtain the following recursion:

$$
\begin{aligned}
&H_{n+1}(b) = 2\left\{H_n(b) + \sum_{k=m}^{n-1}\sum_{l=0}^{b-1} H_k(l)H_{n-k-1}(b-l-1)\right\} \quad \text{for } b>0,\ n \geqslant m+1;\\
&H_n(b) = 0 \quad \text{for } b>0,\ n \leqslant m+1;\\
&H_n(0) = 2^n \quad \text{for } n \geqslant 0.
\end{aligned}
\tag{2.1}
$$

Let $H_n(b,c)$ denote the number of structures with exactly b base pairs in which there are c AU base pairs. Similarly to the above, by observing the inserting base pair $(1,k+2)$ is either AU or GC, we obtain the recursion

$$\begin{aligned} H_{n+1}(b,c) &= 2H_n(b,c) + \sum_{k=m}^{n-1}\sum_{i=0}^{b-1}\sum_{j=0}^{c} H_k(i,j)H_{n-k-1}(b-1-i,c-j) \\ &\quad + \sum_{k=m}^{n-1}\sum_{i=0}^{b-1}\sum_{j=0}^{c-1} H_k(i,j)H_{n-k-1}(b-1-i,c-1-j) \quad \text{for } b>0,\ n\geqslant m+1; \\ H_n(b,c) &= 0 \quad \text{for } b>0,\ n\leqslant m+1; \\ H_n(0,0) &= 2^n \quad \text{for } n\geqslant 0. \end{aligned} \tag{2.2}$$

Let $A_n(d)$ denote the number of structures with exactly d hairpins. The number $A_{n+1}(d)$ may be computed as follows. By adding an unpaired base $A(U)$ or $G(C)$ to a structure on n vertices, we obtain $2A_n(d)$ structures. By adding a base pair $(1,k+2)$, we have two cases to consider: (i) the number of hairpins is unchanged by the additional base pair; (ii) the additional base pair produces a new hairpin. So we obtain the following recursion for the number $A_n(d)$:

$$\begin{aligned} A_{n+1}(d) &= 2A_n(d) + 2\sum_{k=m}^{n-1}\left\{\sum_{l=1}^{d} A_k(l)A_{n-k-1}(d-l) + 2^k A_{n-k-1}(d-1)\right\} \quad \text{for } n\geqslant m+1; \\ A_n(d) &= 0 \quad \text{for } d>0,\ n\leqslant m+1; \\ A_n(0) &= 2^n \quad \text{for } n\geqslant 0. \end{aligned} \tag{2.3}$$

Let $N_n(r)$ denote the number of structures with exactly r stacks. In order to compute it, we need the auxiliary variable $Z_n(r)$ counting the number of structures with exactly r stacks given that $(1,n)$ is a base pair. The corresponding recursion is

$$\begin{aligned} N_{n+1}(r) &= 2N_n(r) + 2\sum_{k=m}^{n-1}\sum_{l=0}^{r} Z_{k+2}(l)N_{n-k-1}(r-l) \quad \text{for } r>0,\ n\geqslant m+1; \\ N_n(r) &= 0 \quad \text{for } r>0,\ n\leqslant m+1; \\ N_n(0) &= 2^n \quad \text{for } n\geqslant 0. \end{aligned} \tag{2.4}$$

At the same time, for the auxiliary variable $Z_n(r)$, we have

$$\begin{aligned} Z_n(r) &= 2\big(Z_{n-2}(r) + N_{n-2}(r-1) - Z_{n-2}(r-1)\big) \quad \text{for } r>0,\ n\geqslant 2; \\ Z_0(r) &= Z_1(r) = 0 \quad \text{for } r>0; \\ Z_0(0) &= 1. \end{aligned} \tag{2.5}$$

2.2. Asymptotic enumeration

Based on the generating functions of the second representation, we use the simplified version of Darboux's theorem [1] (see also [2,3,5,6]) to give the asymptotic enumeration of the above RNA secondary structures.

Lemma 2.1. *Suppose $y_n \geqslant 0$ and $y(x) = \sum_{n=0}^{\infty} y_n x^n$ is of the form*

$$y(x) = \beta(x) + g(x)\left(1-\frac{x}{\alpha}\right)^{w},$$

where $\alpha>0$ is real, $\beta(x)$ and $g(x)$ are analytic near α, and w is real but not a nonnegative integer. If $y(x)$ is analytic for $|x|<\alpha$ and $x=\alpha$ is the only singularity of y on its circle of convergence, then

$$y_n \sim \frac{g(\alpha)}{\Gamma(-w)} n^{-1-w}\left(\frac{1}{\alpha}\right)^{n}.$$

Theorem 2.2. *The number of structures with exactly b base pairs on n vertices, $H_n(b)$, satisfies*

$$H_n(b) \sim \frac{C_b n^{2b} 2^{n-b}}{(2b)!},$$

where $C_b = \frac{1}{b+1}\binom{2b}{b}$, $b\geqslant 0$, are the famous Catalan numbers.

Proof. Let $h_b(x) = \sum_{n=0}^{\infty} H_n(b)x^n$ be the generating function of $H_n(b)$. From recursion (2.1) we obtain

$$(1-2x)h_b(x) = 2x^2 h_{b-1}(x)\frac{(2x)^m}{1-2x} + 2x^2 \sum_{l=1}^{b-1} h_l(x)h_{b-l-1}(x)$$

and

$$h_0(x) = \sum_{n=0}^{\infty} H_n(0)x^n = \sum_{n=0}^{\infty} 2^n x^n = \frac{1}{1-2x}.$$

Suppose

$$h_b(x) = p_b(x)(2x)^{2b}(1-2x)^{-2b-1},$$

then we find that $p_b(x)$ must be polynomials satisfying

$$p_b(x) = \frac{1}{2}(2x)^m p_{b-1}(x) + \frac{1}{2}\sum_{l=1}^{b-1} p_l(x)p_{b-l-1}(x) \quad \text{with } p_0(x) = 1. \tag{2.6}$$

Then Lemma 2.1 indicates

$$H_n(b) \sim \frac{p_b(\frac{1}{2})}{\Gamma(2b+1)} n^{2b} 2^n.$$

By means of Eq. (2.6), we have

$$p_b\left(\frac{1}{2}\right) = \frac{(2b)!}{(b+1)!b!2^b} = 2^{-b} C_b,$$

which gives the result of the theorem. □

Theorem 2.3. *The number of structures with exactly b base pairs in which there are c AU base pairs, $H_n(b,c)$, satisfies*

$$H_n(b,c) \sim \frac{n^{2b} 2^{n-2b}}{(b+1)!c!(b-c)!}.$$

Proof. Let $h_{b,c}(x) = \sum_{n=0}^{\infty} H_n(b,c)x^n$ be the generating function of $H_n(b,c)$. From recursion (2.2), we have

$$(1-2x)h_{b,c}(x) = -x^2 \frac{1-(2x)^m}{1-2x}\big(h_{b-1,c}(x) + h_{b-1,c-1}(x)\big) + x^2 \sum_{i=0}^{b-1}\sum_{j=0}^{c} h_{i,j}(x)h_{b-1-i,c-j}(x) + x^2 \sum_{i=0}^{b-1}\sum_{j=0}^{c-1} h_{i,j}(x)h_{b-1-i,c-1-j}(x)$$

and

$$h_{0,0}(x) = \frac{1}{1-2x}.$$

With the assumption

$$h_{b,c}(x) = f_{b,c}(x)(2x)^{2b}(1-2x)^{-2b-1} \quad \text{and} \quad f_{0,0}(x) = 1,$$

it can be found that $f_{b,c}(x)$ are polynomials fulfilling

$$f_{b,c}(x) = -\frac{1}{4}\big(1-(2x)^m\big)\big(f_{b-1,c}(x) + f_{b-1,c-1}(x)\big) + \frac{1}{4}\sum_{i=0}^{b-1}\sum_{j=0}^{c} f_{i,j}(x)f_{b-1-i,c-j}(x) + \frac{1}{4}\sum_{i=0}^{b-1}\sum_{j=0}^{c-1} f_{i,j}(x)f_{b-1-i,c-1-j}(x). \tag{2.7}$$

Then Lemma 2.1 gives

$$H_n(b,c) \sim \frac{f_{b,c}(\frac{1}{2})}{\Gamma(2b+1)} n^{2b} 2^n. \tag{2.8}$$

Let $\hat{f}_{n,m} = f_{n,m}(\frac{1}{2})$. From Eq. (2.7), we obtain the following recursion:

$$4\hat{f}_{n+1,m} = \sum_{i=0}^{n}\sum_{j=0}^{m}\hat{f}_{i,j}\hat{f}_{n-i,m-j} + \sum_{i=0}^{n}\sum_{j=0}^{m-1}\hat{f}_{i,j}\hat{f}_{n-i,m-1-j}. \tag{2.9}$$

Define

$$\hat{f}(x,y) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\hat{f}_{n,m}x^n y^m;$$

then Eq. (2.9) reduces to

$$\frac{4}{x}\big(\hat{f}(x,y) - 1\big) = \hat{f}^2(x,y) + y\hat{f}^2(x,y),$$

which yields

$$\hat{f}(x,y) = 2\sum_{n=0}^{\infty}\sum_{m=0}^{n}(-1)^n\binom{\frac{1}{2}}{n+1}\binom{n}{m}x^n y^m.$$

Therefore $\hat{f}_{n,m} = 2(-1)^n\binom{\frac{1}{2}}{n+1}\binom{n}{m}$ and the final result of the theorem can be derived from (2.8). □

Theorem 2.4. *The number of structures with exactly d hairpins, $A_n(d)$, satisfies*

$$A_n(d) \sim \frac{(2-\sqrt{2})^{dm}(3-2\sqrt{2})^{d-1}}{2d!(d-1)!}(n/2)^{2d-2}(2+\sqrt{2})^n.$$

Proof. Let $a_d(x) = \sum_{n=0}^{\infty} A_n(d)x^n$ be the generating function of $A_n(d)$. From recursion (2.3) we obtain the functional equation

$$\big(1-4x+2x^2\big)a_d(x) = \frac{1}{2}(2x)^{2+m}a_{d-1}(x) + 2x^2(1-2x)\sum_{i=1}^{d-1}a_i(x)a_{d-i}(x)$$

and

$$a_0(x) = \frac{1}{1-2x}.$$

Suppose

$$\begin{aligned} a_d(x) &= \left(\frac{(2x)^{m+2}}{1-2x}\right)^d \frac{1}{(1-4x+2x^2)^{2d-1}}q_d(x) \\ &= (2x)^{md+2d}(1-2x)^{-d}\big(1-(2-\sqrt{2})x\big)^{-2d+1}q_d(x)\left(1-\frac{x}{\frac{2-\sqrt{2}}{2}}\right)^{-2d+1}. \end{aligned}$$

It can be verified that $q_d(x)$ are polynomials satisfying the recursion

$$\begin{aligned} q_d(x) &= 2x^2(1-2x)\sum_{i=1}^{d-1}q_i(x)q_{d-i}(x) + \frac{1}{2}(1-2x)\big(1-4x+2x^2\big)q_{d-1}(x), \\ q_1(x) &= \frac{1}{2}. \end{aligned} \tag{2.10}$$

According to Lemma 2.1, we have

$$A_n(d) \sim \frac{(2\alpha)^{dm+2d}(1-2\alpha)^{-d}(1-(2-\sqrt{2})\alpha)^{-2d+1}q_d(\alpha)}{\Gamma(2d-1)}n^{2d-2}\alpha^{-n}, \tag{2.11}$$

where $\alpha = \frac{2-\sqrt{2}}{2}$. Next, from Eq. (2.10), we find

$$q_d(\alpha) = \binom{\frac{1}{2}}{d}(-2)^{d-1}(5\sqrt{2}-7)^{d-1},$$

which, combined with (2.11), gives the desired result. □

Theorem 2.5. *The number of structures with exactly r stacks, $N_n(r)$, satisfies*

$$N_n(r) \sim \frac{C_r n^{r-1} 2^{r+n}}{(r-1)!}.$$

Proof. Let the generating functions of $N_n(r)$ and $Z_n(r)$ be $v_r(x) = \sum_{n=0}^{\infty} N_n(r)x^n$ and $u_r(x) = \sum_{n=0}^{\infty} Z_n(r)x^n$, respectively. From recursions (2.4) and (2.5) we obtain

$$v_r(x) = \frac{2}{1-2x}\sum_{l=1}^{r} u_l(x)v_{r-l}(x),$$
$$u_r(x) = \frac{2x^2}{1-2x^2}\big(v_{r-1}(x) - u_{r-1}(x)\big).$$

Suppose

$$v_r(x) = \mu_r(x)\big(1-2x^2\big)^{-r}(1-2x)^{-r},$$
$$u_r(x) = \xi_r(x)\big(1-2x^2\big)^{-r}(1-2x)^{-r+1},$$

then $\mu_r(x)$ and $\xi_r(x)$ are polynomials satisfying

$$\mu_r(x) = 2\sum_{l=1}^{r} \xi_l(x)\mu_{r-l}(x), \tag{2.12}$$
$$\xi_r(x) = 2x^2\big(\mu_{r-1}(x) - (1-2x)\xi_{r-1}(x)\big). \tag{2.13}$$

By Eq. (2.13), we have

$$\xi_r\left(\frac{1}{2}\right) = \frac{1}{2}\mu_{r-1}\left(\frac{1}{2}\right).$$

Inserting the above formula into Eq. (2.12) yields

$$\mu_r\left(\frac{1}{2}\right) = \sum_{l=0}^{r-1} \mu_l\left(\frac{1}{2}\right)\mu_{r-1-l}\left(\frac{1}{2}\right),$$

which indicates that $\mu_r(\frac{1}{2})$ are just the Catalan numbers C_r. Now Lemma 2.1 shows that

$$N_n(r) \sim \frac{\mu_r(\frac{1}{2})(\frac{1}{2})^{-r} n^{r-1}}{\Gamma(r)} 2^n = \frac{C_r n^{r-1} 2^{r+n}}{(r-1)!}.$$

This completes the proof. □

3. Asymptotic enumeration under the third representation

In this section, the asymptotic enumeration of all structures is under the third representation.

Let $\tilde{H}_n(b)$ denote the number of structures with exactly b base pairs on n vertices based on the third representation. Then the following recursion holds:

$$\tilde{H}_{n+1}(b) = 4\left\{\tilde{H}_n(b) + \sum_{k=m}^{n-1}\sum_{l=0}^{b-1} \tilde{H}_k(l)\tilde{H}_{n-k-1}(b-l-1)\right\} \quad \text{for } b > 0,\ n \geqslant m+1;$$
$$\tilde{H}_n(b) = 0 \quad \text{for } b > 0,\ n \leqslant m+1;$$
$$\tilde{H}_n(0) = 4^n \quad \text{for } n \geqslant 0.$$

Similarly to Theorem 2.2, we obtain the asymptotic enumeration

$$\tilde{H}_n(b) \sim \frac{C_b n^{2b} 4^{n-b}}{(2b)!}.$$

Let $\tilde{H}_n(b,c)$ denote the number of structures with exactly b base pairs in which there exist c AU base pairs. The corresponding recursion is

$$\tilde{H}_{n+1}(b,c) = 4\tilde{H}_n(b,c) + 2\sum_{k=m}^{n-1}\sum_{i=0}^{b-1}\sum_{j=0}^{c}\tilde{H}_k(i,j)\tilde{H}_{n-k-1}(b-1-i,c-j)$$
$$+2\sum_{k=m}^{n-1}\sum_{i=0}^{b-1}\sum_{j=0}^{c-1}\tilde{H}_k(i,j)\tilde{H}_{n-k-1}(b-1-i,c-1-j) \quad \text{for } b>0,\ n \geqslant m+1;$$
$$\tilde{H}_n(b,c) = 0 \quad \text{for } b>0,\ n \leqslant m+1;$$
$$\tilde{H}_n(0,0) = 4^n \quad \text{for } n \geqslant 0.$$

Similarly to Theorem 2.3, we have

$$\tilde{H}_n(b,c) \sim \frac{n^{2b}2^{2n-3b}}{(b+1)!c!(b-c)!}.$$

Let $\tilde{A}_n(d)$ denote the number of structures with exactly d hairpins. The corresponding recursion is

$$\tilde{A}_{n+1}(d) = 4\tilde{A}_n(d) + 4\sum_{k=m}^{n-1}\left\{\sum_{l=1}^{d}\tilde{A}_k(l)\tilde{A}_{n-k-1}(d-l) + 4^k\tilde{A}_{n-k-1}(d-1)\right\} \quad \text{for } n \geqslant m+1;$$
$$\tilde{A}_n(d) = 0 \quad \text{for } d>0,\ n \leqslant m+1;$$
$$\tilde{A}_n(0) = 4^n \quad \text{for } n \geqslant 0.$$

Similarly to Theorem 2.4, we can obtain the following asymptotic enumeration:

$$\tilde{A}_n(d) \sim \frac{n^{2d-2}2^{dm}6^{n-2d+1}}{3^{dm-1}d!(d-1)!}.$$

Finally, based on the third representation let $\tilde{N}_n(r)$ denote the number of structures with exactly r stacks and let the auxiliary variable $\tilde{Z}_n(r)$ denote the number of structures with exactly r stacks given that $(1,n)$ is a base pair. The recursion for $\tilde{N}_n(r)$ is

$$\tilde{N}_{n+1}(r) = 4\tilde{N}_n(r) + 4\sum_{k=m}^{n-1}\sum_{l=0}^{r}\tilde{Z}_{k+2}(l)\tilde{N}_{n-k-1}(r-l) \quad \text{for } r>0,\ n \geqslant m+1;$$
$$\tilde{N}_n(r) = 0 \quad \text{for } r>0,\ n \leqslant m+1;$$
$$\tilde{N}_n(0) = 4^n \quad \text{for } n \geqslant 0.$$

The recursion for $\tilde{Z}_n(r)$ is

$$\tilde{Z}_n(r) = 4\big(\tilde{Z}_{n-2}(r) + \tilde{N}_{n-2}(r-1) - \tilde{Z}_{n-2}(r-1)\big) \quad \text{for } r>0,\ n \geqslant 2;$$
$$\tilde{Z}_0(r) = \tilde{Z}_1(r) = 0 \quad \text{for } r>0;$$
$$\tilde{Z}_0(0) = 1.$$

Similarly to Theorem 2.5, we can find

$$\tilde{N}_n(r) \sim \frac{C_r n^{r-1}4^{n+r}}{3^r(r-1)!}.$$

4. Conclusions

In this paper, inspired by the representation in Wang et al. [6], we give recursions of some other RNA secondary structures, and derive the asymptotic formulae of these structures. In addition, we present a new representation about RNA secondary structures, i.e., the bases A, U, G, C are all considered different. From the biological point of view, this representation is more attractive. Under this representation, the asymptotic formulae of the above-mentioned structures are also derived.

Based on the new representation, the structures involved in Wang et al. [6] and some structures involved in Hofacker et al. [2] are not considered in this paper. So in the near future, we can investigate them under this representation.

Acknowledgments

The authors would like to thank Dr. Weiping Wang for his kind help. The authors are also very grateful to the anonymous referee for his valuable comments. This work was supported financially by the Natural Science Foundation of China (10871219) and Dalian Nationalities University foundation (20076210).

References

[1] G. Darboux, Mémoir sur l'approximation des fonctions de très grande nombres, et sur une classe étendu de développements en série, J. Math. Pures Appl. 4 (1878) 5–56.
[2] I.L. Hofacker, P. Schuster, P.F. Standler, Combinatorics of RNA secondary structures, Discrete Appl. Math. 88 (1998) 207–237.
[3] B. Liao, T. Wang, General combinatorics of RNA secondary structures, Math. Biosci. 191 (2004) 69–81.
[4] M.E. Nebel, Combinatorial properties of RNA secondary structures, J. Comput. Biol. 9 (2002) 541–573.
[5] A.M. Odlyzko, Asymptotic enumeration methods, in: R.L. Graham, M. Grötschel, L. Lovász (Eds.), Handbook of Combinatorics, vol. II, Elsevier, Amsterdam, 1995, pp. 1021–1229.
[6] W. Wang, M. Zhang, T. Wang, Asymptotic enumeration of RNA secondary structure, J. Math. Anal. Appl. 342 (2008) 514–523.
[7] M.S. Waterman, Combinatorics of RNA hairpins and cloverleaves, Stud. Appl. Math. 60 (1978) 91–96.
[8] M.S. Waterman, Secondary structure of single-stranded nucleic acids, Adv. Math. Suppl. Stud. 1 (1978) 167–212.