Monatshefte für Chemie 124, 117–125 (1993)

Monatshefte für Chemie
Chemical Monthly

Number of Kekulé Structures of Multiple Zigzag Chain Aromatics

Ivan Gutman[1], Sven J. Cyvin[2], Rong-si Chen[3], and Ke-rong Lin[4]

[1] Faculty of Science, University of Kragujevac, YU-34000 Kragujevac, Yugoslavia
[2] Division of Physical Chemistry, The University of Trondheim, N-7034 Trondheim – NTH, Norway
[3] College of Finance and Economics, Fuzhou University, Fuzhou, Fujian 350002, P.R. China
[4] Department of Mathematics, Fuzhou University, Fuzhou, Fujian 350002, P.R. China

Summary. The Kekulé structures of benzenoid hydrocarbons of type A(n, m) are enumerated. Here A(n, m) stands for a system composed of n condensed zigzag aromatic chains, each containing m hexagons. We study the hitherto unresolved case when the chain length m is variable and the number of chains n is constant.

Keywords. Aromatics; Benzenoids; Kekulé structures.

Die Anzahl von Kekulé-Strukturen von Zick-zack-Mehrfachketten-Aromaten

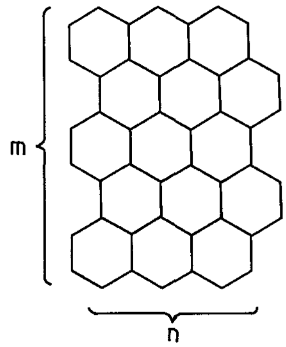
Zusammenfassung. Es wird die Anzahl der Kekulé-Strukturen von bezenoiden Kohlenwasserstoffen des Typs A(n, m) ermittelt. A(n, m) bezeichnet dabei Systeme, die aus n kondensierten aromatischen Zick-zack-Ketten bestehen, wobei jede m Sechsecke enthält. Es wird der bis jetzt ungelöste Fall behandelt, wo die Kettenlänge m variabel ist, die Anzahl der Ketten n jedoch konstant bleibt.

Introduction

In this paper we are concerned with the Kekulé structure counts of benzenoid systems of the type A(n, m) which we call multiple zigzag chains. Such systems can be viewed as being composed of n condensed zigzag-chains, each containing m hexagons. As an example compare the formula for A(n, m) with $m=5$ and $n=3$.

Evidently, the system A(n, m) possesses $n \times m$ hexagons. The respective benzenoid hydrocarbon has the formula $C_{2(m+n+mn)}H_{2(m+n+1)}$. The single-chain zigzag benzenoid hydrocarbons (benzene, naphthalene, phenanthrene, chrysene, picene, ...) correspond to the special cases of A(n, m) when $n=1$ (and $m=1, 2, 3, 4, 5, \ldots$, respectively). The case $n=0$ formally corresponds to the linear polyene series [ethylene ($m=1$), butadiene ($m=2$), hexatriene ($m=3$), ...]; the respective Kekulé structure counts are always equal to unity.

The enumeration of the Kekulé structures of A(n, m) attracted a considerable attention of theoretical chemists. The fact that the Kekulé structure count of A(1, m) is the ($m+1$)-th Fibonacci number was known already to Gordon and Davison in the early fifties [1] and was repeatedly mentioned in several subsequent publications



A(n, m)

[2–5]. Randić [3] seems to be the first to count the Kekulé structures of A(2, m). Results pertaining to the Kekulé structures of A(3, m) were communicated by Ohkami and Hosoya [6]. Nevertheless, the first systematic enumeration of the Kekulé structures of multiple zigzag chains was undertaken by two of the present authors [7] and was eventually further extended [8, 9]. A detailed survey of the numerous results obtained in this area is given in the monograph [10].

In line with the notation used in [9, 10] the number of Kekulé structures of A(n, m) will be denoted by $Z_n(m)$. The problem of determining $Z_n(m)$, preferably by means of explicit combinatorial expressions, turned out to be a rather difficult one. The knowledge accumulated so far makes it highly improbable that a single algebraic formula can be found, which would reproduce the numbers of Kekulé structures of A(n, m) for all values of m and n. The way out of such a difficulty (which is not at all unusual in the theory of Kekulé structure enumeration [10]), is to search for particular solutions. The natural procedure is to separately examine the below two cases:

(a) the chain length m has a fixed value whereas the number n of chains varies;
(b) the chain length m varies whereas the number n of chains has a fixed value.

The study of the $Z_n(m)$-problem [7–10] revealed that the approach (a) is much easier. Explicit combinatorial expressions for $Z_n(m)$ have been deduced for $m = 1, 2, \ldots, 10$ [8, 10]. For a fixed value of m, $Z_n(m)$ is a polynomial in the variable n. Its degree is equal to m. The first ten polynomials $Z_n(m)$ are collected on p. 143 of [10]. There seems to be no obstacle to determine such expressions also for higher values of m, except that these tasks would require very lengthy calculations and perplexed algebraic manipulations.

The treatment of the chemically more interesting case (b) encounters serious difficulties. If m is the variable parameter, then for small values of $n (n \leqslant 5)$ it was possible to deduce [8] recurrence relations for $Z_n(m)$; these are collected on p. 138 of [10]. The original method for obtaining such recurrence relations – described in detail in [8] – is very laborious and has to be employed separately for each particular value of n.

In this paper we report some progress along these lines by establishing the general form of the recurrence relation for $Z_n(m)$. We also show how this recurrence can be solved and, in particular, that $Z_n(m)$ is an exponential function of m.

Furthermore, we determine the approximate behaviour of $Z_n(m)$ for large values of n and m. By this we come somewhat closer to the complete enumeration of the Kekulé structures of multiple zigzag chain aromatics.

Results and Discussion

Recurrence Relations for $Z_n(m)$ for Fixed n

As already mentioned, the first few recurrence relations for $Z_n(m)$ with fixed values of $n (n=1, 2, \ldots, 5)$ have been obtained in [8]. In this section we show how from these formulas the recurrence relations for higher values of n can be determined in a systematic way. First of all, however, we list the respective recurrences for $n \leqslant 10$.

$$Z_0(m) = Z_0(m-1) \tag{1a}$$

$$Z_1(m) = Z_1(m-1) + Z_1(m-2) \tag{1b}$$

$$Z_2(m) = 2\,Z_2(m-1) + Z_2(m-2) - Z_2(m-3) \tag{1c}$$

$$Z_3(m) = 2\,Z_3(m-1) + 3\,Z_3(m-2) - Z_3(m-3) - Z_3(m-4) \tag{1d}$$

$$Z_4(m) = 3\,Z_4(m-1) + 3\,Z_4(m-2) - 4\,Z_4(m-3) - Z_4(m-4) + Z_4(m-5) \tag{1e}$$

$$\begin{aligned} Z_5(m) = {} & 3\,Z_5(m-1) + 6\,Z_5(m-2) - 4\,Z_5(m-3) - 5\,Z_5(m-4) \\ & + Z_5(m-5) + Z_5(m-6) \end{aligned} \tag{1f}$$

$$\begin{aligned} Z_6(m) = {} & 4\,Z_6(m-1) + 6\,Z_6(m-2) - 10\,Z_6(m-3) - 5\,Z_6(m-4) \\ & + 6\,Z_6(m-5) + Z_6(m-6) - Z_6(m-7) \end{aligned} \tag{1g}$$

$$\begin{aligned} Z_7(m) = {} & 4\,Z_7(m-1) + 10\,Z_7(m-2) - 10\,Z_7(m-3) - 15\,Z_7(m-4) \\ & + 6\,Z_7(m-5) + 7\,Z_7(m-6) - Z_7(m-7) - Z_8(m-8) \end{aligned} \tag{1h}$$

$$\begin{aligned} Z_8(m) = {} & 5\,Z_8(m-1) + 10\,Z_8(m-2) - 20\,Z_8(m-3) - 15\,Z_8(m-4) + 21\,Z_8(m-5) \\ & + 7\,Z_8(m-6) - 8\,Z_8(m-7) - Z_8(m-8) + Z_8(m-9) \end{aligned} \tag{1i}$$

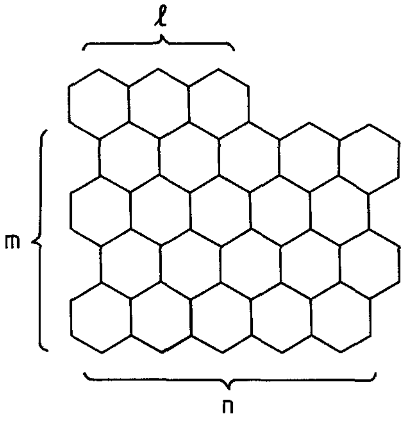
$$\begin{aligned} Z_9(m) = {} & 5\,Z_9(m-1) + 15\,Z_9(m-2) - 20\,Z_9(m-3) - 35\,Z_9(m-4) \\ & + 21\,Z_9(m-5) + 28\,Z_9(m-6) - 8\,Z_9(m-7) - 9\,Z_9(m-8) \\ & + Z_9(m-9) + Z_9(m-10) \end{aligned} \tag{1j}$$

$$\begin{aligned} Z_{10}(m) = {} & 6\,Z_{10}(m-1) + 15\,Z_{10}(m-2) - 35\,Z_{10}(m-3) - 35\,Z_{10}(m-4) \\ & + 56\,Z_{10}(m-5) + 28\,Z_{10}(m-6) - 36\,Z_{10}(m-7) - 9\,Z_{10}(m-8) \\ & + 10\,Z_{10}(m-9) + Z_{10}(m-10) - Z_{10}(m-11). \end{aligned} \tag{1k}$$

Define an auxiliary benzenoid system $A(n, m, l)$, obtained by adding l new hexagons, $0 \leqslant l \leqslant n$, to the top of $A(n, m)$ (see [10], p. 136); as an example compare the formula $A(n, m, l)$ with $m=4$, $n=5$, $l=3$. Note that $A(n, m, 0) \equiv A(n, m)$ whereas $A(n, m, n) \equiv A(n, m+1)$. Let $Z_n(m, l)$ stand for the number of Kekulé structures of $A(n, m, l)$.

Now, by means of the fragmentation method [8, 10] one immediately attains at

$$Z_n(m) = Z_n(m-1) + Z_n(m-2) + \sum_{l=1}^{n-1} Z_n(m-2, l)$$



A(n,m,ℓ)

and

$$Z_n(m-2,l)=Z_n(m-2)+\sum_{l_1=n-l}^{n-1} Z_n(m-3,l_1).$$

The combination of the above two relations results in

$$Z_n(m)=Z_n(m-1)+Z_n(m-2)+\sum_{i\geqslant 1} K_i Z_n(m-i-1) \tag{2}$$

where

$$K_i=\sum_{l_1=1}^{n-1}\sum_{l_2=n-l_1}^{n-1}\sum_{l_3=n-l_2}^{n-1}\cdots\sum_{l_i=n-l_{i-1}}^{n-1} 1. \tag{3}$$

Comparing formula (3) with the known algorithmic procedure for the calculation of $Z_n(m)$ (see [10], pp.144–145), we conclude that

$$K_i=Z_{n-2}(i). \tag{4}$$

Therefore, the coefficients $K_1, K_2, \ldots, K_i, \ldots$ in Eq. (2) satisfy the same recurrence relation as the number of Kekulé structures of $A(n-2,m)$. If this latter recurrence is known, then we can easily eliminate the coefficients K_i from Eq. (2) and arrive at a recurrence relation for the Kekulé structure count of $A(n,m)$.

In the below example we show how formula (1e) is deduced from Eqs. (2) and (1c). Bearing in mind the form of (1c), define an auxiliary quantity $R_n(m)$:

$$R_n(m)=Z_n(m)-2Z_n(m-1)-Z_n(m-2)+Z_n(m-3)$$

and observe that Eq. (1c) is tantamount to the condition $R_2(m)=0$. The application of formula (2), $n=4$, to the right-hand side of $R_4(m)$ results in

$$R_4(m)=\left[Z_4(m-1)+Z_4(m-2)+\sum_{i\geqslant 1} K_i Z_4(m-i-1)\right]-2\left[Z_4(m-2)+Z_4(m-3)\right.$$
$$\left.+\sum_{i\geqslant 1} K_i Z_4(m-i-2)\right]-\left[Z_4(m-3)+Z_4(m-4)+\sum_{i\geqslant 1} K_i Z_4(m-i-3)\right]$$

$$+\left[Z_4(m-4)+Z_4(m-5)+\sum_{i\geqslant 1} K_i Z_4(m-i-4)\right]$$

$$=Z_4(m-1)+(1-2+K_1)Z_4(m-2)+(-2-1+K_2-2K_1)Z_4(m-3)$$

$$+(-1+1+K_3-2K_2-K_1)Z_4(m-4)+Z_4(m-5)$$

$$+\sum_{i\geqslant 4}(K_i-2K_{i-1}-K_{i-2}+K_{i-3})Z_4(m-i-1).$$

Because of (4), $K_i-2K_{i-1}-K_{i-2}+K_{i-3}=0$ for all $i\geqslant 4$ and consequently,

$$Z_4(m)-2Z_4(m-1)-Z_4(m-2)+Z_4(m-3)=Z_4(m-1)+(K_1-1)Z_4(m-2)$$

$$+(K_2-2K_1-3)Z_4(m-3)+(K_3-2K_2-K_1)Z_4(m-4)+Z_4(m-5).$$

Equation (1e) follows from this latter identity by taking into account the conditions $K_1=Z_2(1)=3$, $K_2=Z_2(2)=6$ and $K_3=Z_2(3)=14$.

By means of the above illustrated procedure the recurrence relations for $Z_n(m)$ can be obtained one-by-one, leading e.g. to Eqs. (1a)–(1k). However, with the increasing value of the parameter n the respective calculations, although routine and straightforward, become more and more involved.

In the subsequent section this problem is overcome by finding a general expression for the recurrence relation for $Z_n(m)$.

A General Formulae for the Recurrence Relation for $Z_n(m)$ for Fixed n

By inspecting the recurrence relations for $Z_n(m)$ for $n=0,1,\ldots,10$, given by Eqs. (1a)–(1k), we observe the following regularities.

(i) $Z_n(m)$ obeys a linear recurrence relation of the order $n+1$:

$$Z_n(m)=\sum_{k=1}^{n+1} a_{kn} Z_n(m-k);\ m\geqslant n+1,\ n\geqslant 0. \tag{5a}$$

Here a_{kn}, $k=1,2,\ldots,n+1$, denote the respective coefficients which are independent of m.

(ii) For $k=1,2,\ldots,n+1$ the signs of the coefficients a_{kn} follow the pattern $++--++--\ldots$.

(iii) For $n\geqslant 0$ the first coefficient in Eq. (5a) is given by $a_{1n}=\lceil (n+1)/2\rceil$.
Here and later $\lceil x\rceil$ denotes the smallest integer which is not smaller than x. (For instance, $\lceil 3.9\rceil=4$, $\lceil 4\rceil=4$, $\lceil 4.1\rceil=5$.)

(iv) Each coefficient, except the first, is equal to the sum of two previous coefficients, namely the identity

$$|a_{kn}|=|a_{k,n-2}|+|a_{k-1,n-1}|$$

is obeyed for all $n\geqslant 2$, $k\geqslant 2$.

The properties (i)–(iv) together with the condition $a_{21}=1$, fully determine all the coefficients a_{kn}. Then an elementary combinatorial reasoning leads to the formula

$$a_{kn}=(-1)^{\lceil k/2\rceil+1}\binom{\lceil (n+k)/2\rceil}{k}$$

i.e.

$$Z_n(m) = \sum_{k=1}^{n+1} (-1)^{\lceil k/2 \rceil + 1} \binom{\lceil (n+k)/2 \rceil}{k} Z_n(m-k). \tag{5b}$$

The above general combinatorial expression for the recurrence relations of $Z_n(m)$ with fixed n applies to all $n \geqslant 0$ and all $m \geqslant n+1$. Formulas (1a)–(1k) are just special cases of (5b) for $n = 0, 1, \ldots, 10$.

A somewhat more compact way of writing Eq. (5b) is

$$\sum_{k=0}^{n+1} (-1)^{\lceil k/2 \rceil} \binom{\lceil (n+k)/2 \rceil}{k} Z_n(m-k) = 0. \tag{5c}$$

In the work [9] another recurrence relation for $Z_n(m)$ was reported, namely

$$Z_n(m) = \sum_{k=1}^{n+1} b_{kn} Z_n(m-2k);\ m \geqslant 2n+2,\ n \geqslant 0 \tag{6}$$

where

$$b_{kn} = (-1)^{k+1} \binom{n+k+1}{2k}. \tag{7}$$

Although the forms of (5) and (6) look quite similar, they are far from being equivalent. Whereas Eqs. (5) express the Kekulé structure count of $A(n, m)$ by means of a recurrence relation of order $n+1$, the respective recurrence relation given by Eq. (6) is of order $2n+2$. In particular, for $n=1$ and $n=2$ the special cases of (6) read:

$$Z_1(m) = 3\,Z_1(m-2) - Z_1(m-4) \tag{8a}$$

$$Z_2(m) = 6\,Z_2(m-2) - 5\,Z_2(m-4) + Z_2(m-6) \tag{8b}$$

which should be compared with Eqs. (1b) and (1c). Using Eq. (1c) and the simple initial conditions $Z_2(0) = 1$, $Z_2(1) = 3$, $Z_2(2) = 6$ we can compute $Z_2(m)$ for all $m \geqslant 3$. On the other hand, Eq. (8b) requires the significantly more complicated initial conditions $Z_2(0) = 1$, $Z_2(2) = 6$, $Z_2(4) = 31$ (if n is even) and $Z_2(1) = 3$, $Z_2(3) = 13$, $Z_2(5) = 70$ (if m is odd) and enables the calculation of $Z_2(m)$ only for $m \geqslant 6$. Such differences between (5) and (6) are even more pronounced for larger values of n.

Evidently, it is much more expedient to use (5) than (6). The coefficients of Eqs. (5) and (6) are mutually related in a fairly complicated manner:

$$b_{kn} = 2\,a_{2k,n} - \sum_{j=1}^{2k-1} (-1)^j a_{jn} a_{2k-j,n}.$$

Whence, for any particular value of n, Eq. (6) can always be deduced from the respective Eq. (5).

We wish to point out another intriguing relation between Eqs. (5) and (6). For even values of the parameter n, Eq. (5b) can be transformed into

$$Z_n(m) = \sum_{k=1}^{n/2+1} (-1)^{k+1} \binom{n/2+k}{2k-1} Z_n(m-2k+1) + \sum_{k=1}^{n/2} (-1)^{k+1} \binom{n/2+k}{2k} Z_n(m-2k). \tag{9}$$

On the other hand, from (6) and (7) one immediately obtains

$$Z_{n/2-1}(m) = \sum_{k=1}^{n/2} (-1)^{k+1} \binom{n/2+k}{2k} Z_{n/2-1}(m-2k). \tag{10}$$

The coefficients in the last summation of (9) are precisely the same as the coefficients in (10).

Solving the Recurrence Relation for $Z_n(m)$ for Fixed n

The standard method [11] for solving recurrence relations of the type (5) is to find the roots $t_1, t_2, \ldots, t_{n+1}$ of the auxiliary equation

$$t^{n+1} = \sum_{k=1}^{n+1} a_{kn} t^{n+1-k}. \tag{11}$$

If all the roots $t_1, t_2, \ldots, t_{n+1}$ are mutually different and all are real-valued (i.e. no one of them is complex-valued), then the general solution of (5) is of the form

$$Z_n(m) = C_1(t_1)^m + C_2(t_2)^m + \cdots + C_{n+1}(t_{n+1})^m \tag{12}$$

where $C_1, C_2, \ldots, C_{n+1}$ are multipliers which can be determined from the initial conditions i.e. from the known Kekulé structure counts of $A(n, m)$, $m = 0, 1, \ldots, n$.

In Table 1 are collected the roots of the equation (11) for $n = 1, 2, \ldots, 10$.

Table 1. The roots of the auxiliary equation of the recurrence relation (5) for the first few values of the parameter n

	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$
t_1	1.618034	2.246980	2.879385	3.513337	4.148115
t_2	−0.618034	0.554958	0.652704	0.763521	0.880181
t_3		−0.801938	−0.532089	0.521109	0.564681
t_4			−1.000000	−0.594351	−0.514964
t_5				−1.203616	−0.667993
t_6					−1.410020

	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
t_1	4.783386	5.418976	6.054783	6.690745	7.326822
t_2	1.000000	1.121734	1.244724	1.368584	1.493074
t_3	0.618034	0.676582	0.738245	0.801938	0.867030
t_4	0.511170	0.536209	0.568521	0.605152	0.644570
t_5	−0.547318	−0.508661	0.506914	0.523246	0.545129
t_6	−0.747238	−0.588085	−0.528643	−0.505648	0.504702
t_7	−1.618034	−0.829690	−0.633601	−0.554957	−0.519255
t_8		−1.827065	−0.914164	−0.682079	−0.585193
t_9			−2.036780	−1.000000	−0.732544
t_{10}				−2.246980	−1.086803
t_{11}					−2.457534

Based on the data from Table 1 we can formulate the following rules. Although the validity of these rules was verified only for the first few values of n, there is little doubt that they apply to all recurrence relations of the type (5).

Rule 1. For all $n \geqslant 0$, all the roots of the Eq. (11) are real-valued numbers.
Rule 2. For a given value of n no two of the roots of (11) coincide.
Rule 3. For two consecutive values of n the roots of (11) interlace each other. In other words, if $t_1(p), t_2(p), \ldots, t_{p+1}(p)$ denote the roots of Eq. (11) for $n=p$, then

$$t_1(p) > t_1(p-1) > t_2(p) > t_2(p-1) > t_3(p) > \cdots > t_p(p) > t_p(p-1) > t_{p+1}(p).$$

When the roots of Eq. (11) are considered as functions of the parameter n then this functional dependence is almost perfectly linear. This remarkable regularity is illustrated in Fig. 1.

For $n \leqslant 10$ the following approximate formulas were established using the method of least squares:

$$t_1 = 0.6336\,n + 0.9847 \tag{13a}$$

$$t_2 = 0.1184\,n + 0.2984 \tag{13b}$$

$$t_n = -0.0801\,n - 0.2756 \tag{13c}$$

$$t_{n+1} = -0.2056\,n - 0.3909. \tag{13d}$$

The correlation coefficients of Eqs. (13a)–(13d) are 0.999995, 0.9995, -0.9989 and -0.99987, respectively.

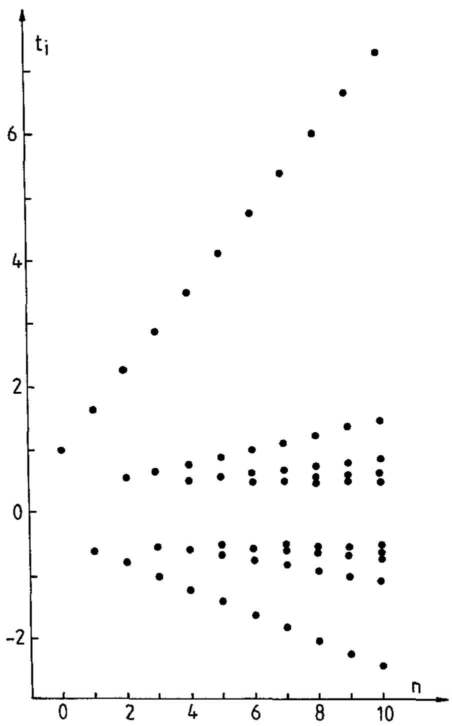


Fig. 1. The roots of the auxiliary equation of the recurrence relation (5) as functions of the parameter n

The linear dependence of the roots of Eq. (11) on the parameter n is in harmony with the previously mentioned fact that $Z_n(m)$ is an m-degree polynomial in the variable n.

From Fig. 1 we see that one root of (11), namely t_1, is much larger than the other roots. In addition to this, with increasing n, t_1 increases much faster than the other roots (see, for instance, Eqs. (13)). This means that for large values of m and n the first term on the right-hand side of (12) will have the dominant contribution to $Z_n(m)$. In other words, $Z_n(m)$ behaves asymptotically as $C_1(t_1)^m$ i.e. as $C_1(an+b)^m$ where C_1, a and b are constants. This, in turn, means that

$$Z_n(m) \approx C_1(a\,n+b)^m \tag{14a}$$

is a reasonable approximation for $Z_n(m)$, provided n and m are sufficiently large. Employing the relation (13a) and using the value $Z_{10}(10) = 565424068$ [10], we arrive at our final result

$$Z_n(m) \approx 1.279\,(0.6336\,n + 0.9847)^m. \tag{14b}$$

Since the $Z_n(m)$-values have been calculated and tabulated for $m \leqslant 10$, $n \leqslant 10$ (see [10], p. 144), formula (14) covers just the case for which the Kekulé structures of $A(n, m)$ have not been enumerated.

References

[1] Gordon M, Davison W. H. T. (1952) J. Chem. Phys. **20**: 428

[2] Yen T. F. (1971) Theor. Chim. Acta **20**: 399

[3] Randić M. (1980) Int. J. Quantum Chem. **17**: 549

[4] Cyvin S. J. (1983) Monatsh. Chem. **114**: 13

[5] Balaban A. T., Tomescu I. (1985) Match **17**: 91

[6] Ohkami N., Hosoya H. (1983) Theor. Chim. Acta **64**: 153

[7] Cyvin S. J., Gutman I. (1986) Comput. and Math. with Appl. **12B**: 859

[8] Gutman I., Cyvin S. J. (1987) Monatsh. Chem. **118**: 541

[9] Cyvin S. J., Cyvin B. N., Brunvoll J., Gutman I. (1987) Z. Naturforsch. **42a**: 722

[10] Cyvin S. J., Gutman I. (1988) Kekulé Structures in Benzenoid Hydrocarbons. Springer, Berlin Heidelberg New York Tokyo

[11] Spiegel M. R. (1971) Finite Differences and Difference Equations. McGraw-Hill, New York

Received December 17, 1991. Accepted June 16, 1992