Number of Kekulé Structures of Systems with Repeated Units

S. J. Cyvin and B. N. Cyvin
Division of Physical Chemistry, The University of Trondheim, Norway

I. Gutman
Faculty of Science, University of Kragujevac, Yugoslavia.

Z. Naturforsch. 42a, 181–186 (1987); received November 10, 1986

Combinatorial formulas are derived for the number of Kekulé structures of conjugated (benzenoid) systems composed of repeated, mutually condensed units.

Introduction

Continuing our work on the enumeration of Kekulé structures of aromatic (benzenoid) hydrocarbons, especially of pericondensed systems (see [1–3] and the references cited therein), we have examined the system

\[ C_N = A_0: A_1: A_2: \ldots: A_{N-1}: A_N: A_p, \]

whose structure is depicted on the following diagram:

We can describe \( C_N \) as being composed of \( N \) mutually condensed units \( A_1, A_2, \ldots, A_N \), to which two further terminal fragments \( A_0 \) and \( A_p \) are attached. Needless to say that a great number of chemically important conjugated (benzenoid) hydrocarbons and/or systems for which the problem of the enumeration of Kekulé structures has been treated previously, are special cases of \( C_N \). Pertinent examples will be given in the last section of this paper.

Although the main applications of our results are certainly to benzenoid molecules, the entire formalism developed in the present paper applies to non-benzenoids as well.

Definitions

Let \( G \) be a molecular graph [4]. Let further \( a_L, b_L, a_R \) and \( b_R \) be distinct vertices of \( G \), such that \( a_L \) and \( a_R \) are adjacent, and \( a_R \) and \( b_R \) are adjacent. The edges connecting \( a_L \) with \( b_L \) and \( a_R \) with \( b_R \) will be denoted by \( e_L \) and \( e_R \), respectively. Conventionally we say that \( e_L \) is the left edge and \( e_R \) the right edge of \( G \), and we present \( G \) as

\[ a_L \quad e_L \quad b_L \quad e_R \quad a_R \quad b_R \]

In the following we shall always assume that in the molecular graph considered there are a left and a right edge. This particularly applies to the subgraphs \( A_i; i = 1, 2, \ldots, N \), from which \( C_N \) is composed. Exceptionally, in the terminal fragment \( A_0 \) only a right edge, whereas in the fragment \( A_p \) only a left edge need to be distinguished.

Let \( P \) and \( Q \) be two arbitrary molecular graphs. In the following the symbol \( P:Q \) will stand for the graph obtained by identifying the right edge of \( P \) with the left edge of \( Q \), so that the vertices \( a_k \) and \( b_k \) of \( P \) are identified with the vertices \( a_k \) and \( b_k \) of \( Q \):

Reprint requests to Prof. S. J. Cyvin, Division of Physical Chemistry, The University of Trondheim, N-7034 Trondheim-NTH, Norway.

0340-4811 / 87 / 0200-0127 $ 01.30/0. Please order a reprint rather than making your own copy.
Then the structure of \( C_\chi \) is expressed by

\[
A_0: A_1: A_2: \ldots : A_N: A_0.
\]

It is consistent to define \( C_0 = A_0: A_0 \).

The system which coincides with \( C_\chi \) except that it does not possess a right terminal fragment will be denoted \( B_\chi \). Thus, \( B_\chi = A_0: A_1: A_2: \ldots : A_N \), and, in particular, \( B_0 = A_0 \).

In the following we shall be especially interested in the case when all the units \( A_i \) (\( i = 1, 2, \ldots, N \)) are mutually isomorphic; \( A_1 \cong A_2 \cong \ldots \cong A_N \). When we write \( A_i \cong A_j \), we understand not only that \( A_i \) and \( A_j \) are isomorphic, but also that their vertices are labelled in the same manner.

In order to be able to express our results in a compact form we denote the subgraph obtained from \( G \) by deleting the left edge, \( G - e_L \) by \( 'G \). Similarly, the shorthand notation for \( G - e_R - e_L \) is \( 'G' \).

The number of Kekulé structures of the conjugated system whose molecular graph is \( G \) will be denoted by \( K[G] \).

An Auxiliary Result

We demonstrate now that if either \( P \) or \( Q \) have an even number of vertices, then

\[
K[P: Q] = K[P] K[Q] - K[P'] K[Q'] . \tag{1}
\]

If both \( P \) and \( Q \) have an odd number of vertices, then (1) needs not to hold, as shown on the following example:

\[
K[G] = 0 \quad K[G:G] = 9
\]

Of course, if \( K[G] = 0 \), then also \( K'[G] = K'[G'] = 0 \).

If either \( P \) or \( Q \), but not both have an odd number of vertices, then \( P: Q \) has an odd number of vertices, \( K[P: Q] = 0 \), and (1) holds in a trivial manner. Hence the only case of interest is when both \( P \) and \( Q \) have even numbers of vertices.

It has been shown elsewhere [5] that

\[
K[P: Q] = K[P] K[Q - a_L - b_L]
\]

\[
+ K[P - a_R - b_R] K[Q]
\]

\[
- K[P - a_R - b_R] K[Q - a_L - b_L] \tag{2}
\]

\[
+ K[P - a_R] K[Q - b_L] + K[P - b_R] K[Q - a_L],
\]

which for \( P \) and \( Q \) being even systems reduces to

\[
K[P: Q] = K[P] K[Q - a_L - b_L]
\]

\[
+ K[P - a_R - b_R] K[Q]
\]

\[
- K[P - a_R - b_R] K[Q - a_L - b_L]. \tag{3}
\]

According to a well-known formula for the number of Kekulé structures [6],

\[
K[P] = K[P'] + K[P - a_R - b_R], \tag{4a}
\]

\[
K[Q] = K[Q'] + K[Q - a_L - b_L]. \tag{4b}
\]

Substituting (4) back into (3) we obtain (1). All results given in the present paper are consequences of (1).

Here we point out an immediate corollary of (1). Let \( (P: Q)^* \) be the graph obtained by identifying the vertices \( a_R \) and \( b_R \) of \( P \), respectively, with the vertices \( b_L \) and \( a_L \) of \( Q \):

\[
K[(P: Q)^*] = K[P: Q]. \tag{5}
\]

provided either \( P \) or \( Q \) or both have an even number of vertices. One should observe that (3) and (4) remain the same if the symbols \( a_L \) and \( b_L \) exchange places. Therefore

\[
K[(P: Q)^*] = K[P] K[Q] - K[P'] K[Q'], \tag{6}
\]

which together with (1) implies (5).

Number of Kekulé Structures of \( B_N \)

The system \( B_N \) has been defined as

\[
B_N = A_0: A_1: A_2: \ldots : A_{N-1}: A_N.
\]
Let, in addition,

\[ B_N = A_0 : A_1 : A_2 : \ldots : A_{N-1} : A_N. \]

Applying (1) to \( B_N \) and \( B_N' \) we immediately get

\[
\begin{align*}
K[B_N] &= K[A_N] K[B_{N-1}] - K[A'_N] K[B'_{N-1}], \\
K[B_N'] &= K[A'_N] K[B_{N-1}] - K[A_N] K[B'_{N-1}],
\end{align*}
\]

or in matrix form

\[
\begin{bmatrix}
K[B_N] \\
K[B_N']
\end{bmatrix} =
\begin{bmatrix}
K[A_N] & -K[A'_N] \\
K[A'_N] & -K[A_N]
\end{bmatrix}
\begin{bmatrix}
K[B_{N-1}] \\
K[B'_{N-1}]
\end{bmatrix},
\]

where the matrices \( M_i; i = 1, 2, \ldots, N \) are defined as

\[
M_i = \begin{bmatrix}
K[A_i] & -K[A'_i] \\
K[A'_i] & -K[A_i]
\end{bmatrix},
\]

Repeating the same argument additional \( N - 1 \) times, we arrive at

\[
\begin{bmatrix}
K[B_N] \\
K[B_N']
\end{bmatrix} = M_N M_{N-1} \ldots M_2 M_1 \begin{bmatrix}
K[A_0] \\
K[A_0]
\end{bmatrix}.
\]

Formula (9) has the following consequences. If at least one of the units from which \( B_N \) is composed, say \( A_j \), is non-Kekuléan, then irrespective of the structure of \( A_0, A_1, \ldots, A_{j-1}, A_{j+1}, \ldots, A_N \), the entire system \( B_N \) is non-Kekuléan. Indeed, if \( K[A_j] = 0 \) then also \( K[A'_j] = K[A'_N] = K[A'_j] = 0 \) and thus \( M_j = 0 \). Note that \( B_N \) is non-Kekuléan also if \( A_0 \) is non-Kekuléan.

Equations (7) present a system of coupled recurrence relations. They can be easily decoupled into

\[
\begin{align*}
\chi_N K[B_N] + \beta_N K[B_{N-1}] + \gamma_N K[B_{N-2}] &= 0; \quad N \geq 2, \\
\chi_N K[B_N'] + \beta_N K[B'_{N-1}] + \gamma_N K[B'_{N-2}] &= 0; \quad N \geq 2,
\end{align*}
\]

where

\[
\begin{align*}
\chi_N &= K[A_{N-1}], \\
\beta_N &= K[A_N] K[A'_{N-1}] - K[A_N] K[A_{N-1}], \\
\gamma_N &= K[A_N] \det M_{N-1}.
\end{align*}
\]

If all the subgraphs \( A_i \), \( i = 1, 2, \ldots, N \) are mutually isomorphic, then a significant simplification of the above relations occurs. For \( A_i \cong A, i = 1, 2, \ldots, N \), we have

\[
\begin{bmatrix}
K[B_N] \\
K[B_N']
\end{bmatrix} =
\begin{bmatrix}
K[A] & -K[A'] \\
K[A'] & -K[A]
\end{bmatrix}^N
\begin{bmatrix}
K[A_0] \\
K[A_0]
\end{bmatrix}
\]

and

\[
\begin{align*}
K[B_N] &= \beta K[B_{N-1}] + \gamma K[B_{N-2}]; \quad N \geq 2, \\
K[B_N'] &= \beta K[B'_{N-2}] + \gamma K[B'_{N-2}]; \quad N \geq 2,
\end{align*}
\]

where

\[
\begin{align*}
\beta &= K[A] - K[A'], \\
\end{align*}
\]

Using standard techniques of the theory of difference equations [7], the explicit general expressions for \( K[B_N] \) can be calculated from (13a). We shall not pursue this computation any further because the explicit formula for \( K[B_N] \) has a very complicated algebraic form; note that \( K[B_N] \) is a function of \( N, K[A_N], K[A'], K[A'], K[A], K[A], K[A], K[A_A], K[A'] \). It seems quite improbable that such a result ever will be needed in practical applications.

If, however, there is no terminal fragment in \( B_N \), which formally corresponds to the choice \( K[A_0] = 1, K[A_0] = 0 \), then we obtain a somewhat simpler expression:

\[
\begin{align*}
K[B_N] &= (2^{N+1} R^{-1} [(K[A] + K[A']) + R] \\
&\cdot (K[A] - K[A']) + R)^N - (K[A] + K[A'] - R) \\
&\cdot (K[A] - K[A'] - R)^N,
\end{align*}
\]

where

\[
\]

**Number of Kekulé Structures of \( C_N \)**

In the preceding section we pointed out the (not necessarily obvious) fact that both \( K[B_N] \) and \( K[B_N'] \) conform to the same recurrence relation (10). This has the immediate consequence that also \( K[C_N] \) satisfies this relation, viz.

\[
\chi_N K[C_N] + \beta_N K[C_{N-1}] + \gamma_N K[C_{N-2}] = 0; \quad N \geq 2.
\]

In order to see that (10c) holds, observe that \( C_N = B_N : A_\alpha \) and apply (1). Then

\[
\]

Equation (10c) is now obtained by substituting (10a) and (8b) into (17a) and by taking into account that

\[
K[A_\alpha] K[B_{N-1}] - K[A_\alpha] K[B'_{N-1}] = K[C_{N-1}],
\]

\[
\]
If \( A_i \simeq A \) for all \( i = 1, 2, \ldots, N \), then it immediately follows that the recurrence relation for \( K\{C_N\} \) also is of the form (13), namely

\[
K\{C_N\} = \beta K\{C_{N-1}\} + \gamma K\{C_{N-2}\}; \quad N \geq 2. \tag{13c}
\]

From (13c) an explicit formula for \( K\{C_N\} \) could be obtained, but because of already mentioned reasons we avoid this issue.

We have reached the conclusion that the nature of the terminal fragments \( A_0 \) and \( A_w \) does not influence the form of the recurrence relation for \( K\{C_N\} \). In particular, (10a) and (10b) are special cases of (10c), as well as (13a) and (13b) are special cases of (13c). The numerical value of \( K\{C_N\} \) depends, however, on \( A_0 \) and \( A_w \).

It can be shown that a necessary and sufficient condition for \( C_N \) to be Kekuléan (i.e. \( K\{C_N\} > 0 \)) is that all units \( A_0, A_1, A_2, \ldots, A_N, A_w \) are Kekuléan. As we already have pointed out, this is not true if at least one of the \( A_i \)'s has odd number of vertices.

**Examples**

*Two Simple Cases*

Consider first the case when \( C_N \) contains a linear polyacene fragment [5]. This occurs when \( A_i \simeq L \) for all \( i = 1, 2, \ldots, N \). Since \( K\{L\} = 2 \), \( K\{'L\} = K\{L'\} = 1 \), \( K\{'L'\} = 0 \) we have \( \beta = 2 \) and \( \gamma = -1 \), and (13c) reduces to

\[
K\{C_N\} = 2 K\{C_{N-1}\} - K\{C_{N-2}\}. \tag{18}
\]

It is referred to the below illustration.

The recurrence relation (18) for the number of Kekulé structures of systems containing a linear polyacene fragment has not been observed in [2, 5]. From (18) it immediately follows that \( K\{C_N\} \) is a linear function of \( N \), which is a previously known result [5].

The second case is when \( C_N \) contains a zigzag polyacene fragment. This means that \( A_i \simeq Z \) for all \( i \), and then \( K\{Z\} = 2 \), \( K\{'Z\} = K\{Z'\} = K\{'Z'\} = 1 \), \( \beta = 1 \), \( \gamma = 1 \), and we reach the previously known [2] Fibonacci-type recurrence formula

\[
K\{C_N\} = K\{C_{N-1}\} + K\{C_{N-2}\}. \tag{19}
\]

*More Advanced Case*

As a third, somewhat less elementary example we consider “parallelograms on a string”, whose repeated units \( A_i \) are all mutually isomorphic and given as

![Parallelogram](image)

These systems have one terminal fragment \( A_0 \), which is the linear polyacene with \( n \) hexagons, \( L(n) \). The repeated units \( A_i \) can be understood as being fused of a parallelogram-shaped benzenoid system \( L(m, k) \) and a linear polyacene \( L(n) \), i.e. \( A_i = L(m, k): L(n) \):

“Parallelograms on a string” containing \( N \) repeated units will be denoted by \( B_N \). In addition to \( N \), their structure depends on the parameters \( n, m \) and \( k \).

As a special case, the choice \( n = 1, m = 2 \) and \( k = 2 \) corresponds to the class of “pyrenes on a string”. It is a celebrated example of all-benzenoid systems [8, 9]. For \( N = 3 \) we have the following representative of these benzenoid molecules:

![Pyrene](image)

It has been referred specifically to members of this class by several authors [10–13]. One of these papers [10] implies the recurrence relation for the number of Kekulé structures, viz.

\[
K\{B_N\} = 10 K\{B_{N-1}\} - K\{B_{N-2}\}; \quad N \geq 2. \tag{20}
\]

The explicit combinatorial formula for \( K\{B_N\} \) (when \( n = 1, m = 2, k = 2 \)) has also been reported [14]:

\[
K\{B_N\} = \frac{1}{2} 6^{-\frac{1}{2}} \left[ (5 + 2 \cdot 6^{\frac{1}{2}}) N^2 - (5 - 2 \cdot 6^{\frac{1}{2}}) N^2 + 1 \right]. \tag{21}
\]
Let us now return to the general case, namely when the parameters \( n, m \) and \( k \) have arbitrary values. It is well known [15] that
\[
K[L(m,k)] = \binom{m+k}{k}; \quad K[L(n)] = n + 1. \tag{22}
\]
Knowing that \( A_i = L(m,k) \cdot L(n) \) and applying (1) we then easily obtain
\[
K[A_i] = (n+1) \binom{m+k}{k} - n, \tag{23a}
\]
\[
K[A'] = n + 1, \tag{23b}
\]
\[
K[A'_i] = n \binom{m+k}{k} - (n-1), \tag{23c}
\]
\[
K[A'_i'] = n. \tag{23d}
\]
from which \( \beta = (n+1) \binom{m+k}{k} - 2n, \gamma = -1, \) and thus (10a) becomes
\[
K[B_N] = (r+2) K[B_{N-1}] - K[B_{N-2}]; \quad N \equiv 2. \tag{24}
\]
where
\[
r = (n+1) \left[ \binom{k+m}{k} - 2 \right]. \tag{25}
\]
Using the initial conditions
\[
K[B_0] = n + 1, \quad K[B_1] = (n+1)(r+2) \tag{26}
\]
we obtain after a lengthy calculation [7]
\[
K[B_N] = (n+1) r^{-\frac{1}{2}}(r+4)^{-\frac{1}{2}} \left[ \frac{1}{2} (r+2 + r \frac{1}{2}(r+4)) \right]^{N+1} - \left[ \frac{1}{2} (r+2 - r \frac{1}{2}(r+4)) \right]^{N+1}. \tag{27}
\]
In particular, for pyrene \( L(2,2) \) one has \( r = 4(n+1) \) and consequently
\[
K[B_N] = 2(2n+3) K[B_{N-1}] - K[B_{N-2}]; \quad N \equiv 2 \tag{28}
\]
with
\[
K[B_0] = n + 1, \quad K[B_1] = 2(n+1)(2n+3) \tag{29}
\]
and
\[
K[B_N] = \frac{1}{4} (n+1)^{\frac{1}{2}} (n+2)^{-\frac{1}{2}} \cdot \left[ 2n+3 + 2(n+1) \frac{1}{2} (n+2)^{\frac{1}{2}} \right]^{N+1} - \left[ 2n+3 - 2(n+1) \frac{1}{2} (n+2)^{\frac{1}{2}} \right]^{N+1}. \tag{30}
\]
Equations (20) and (21) are special cases for \( n = 1 \) of (28) and (30), respectively.

For anthanthrene \( L(2,3) \) one has \( r = 8(n+1) \) and
\[
K[B_N] = 2(4n+5) K[B_{N-1}] - K[B_{N-2}]; \quad N \equiv 2 \tag{31}\]
with
\[
K[B_0] = n + 1, \quad K[B_1] = 2(n+1)(4n+5) \tag{32}\]
and
\[
K[B_N] = \frac{1}{4} (n+1)^{\frac{1}{2}} (n+2)^{-\frac{1}{2}} \cdot \left[ 4n+5 + 8 \frac{1}{2}(n+1)^{\frac{1}{2}} (2n+3)^{\frac{1}{2}} \right]^{N+1} - \left[ 4n+5 - 8 \frac{1}{2}(n+1)^{\frac{1}{2}} (2n+3)^{\frac{1}{2}} \right]^{N+1}. \tag{33}
\]
Especially, for \( n = 1 \):
\[
K[L(1) \cdot L(2,3)]^{N} \L L(1)\]
\[
= \frac{1}{4} 5^{-\frac{1}{2}} \left[ (9 + 4 \cdot 5^{\frac{1}{2}})^{N+1} - (9 - 4 \cdot 5^{\frac{1}{2}})^{N+1} \right]. \tag{34}
\]
This case of anthanthrene with \( n = 1 \) is illustrated below for \( N = 3 \):

\[
K = 11592
\]

The fact that the above two isomers have equal number of Kekulé structures is a consequence of (5). This feature is well known in connection with the kinks (angular annelation) in single-chain benzenoids [15-18]. Such two variants of a benzenoid chain have been referred to as isoarithmic [16]. The above example shows that isoarithmicity is just a particular manifestation of a more general topological phenomenon. In the above examined case, namely of the parallelograms on a (linear) string, isoarithmic structures occur whenever \( k = m \).

**Conclusion**

The present paper gives a contribution to the enumeration of Kekulé structures, which has
Number of Kekulé Structures of Systems with Repeated Units

attained a considerable interest during the last years. It is sufficient to document this fact by the numerous relevant papers which have appeared only in 1986 [3, 19–33]. The chemical importance of benzenoids with repeated units among other large or very large benzenoid systems seems not necessary to be documented specifically. The polymers with an infinite number of units come into this category, but we have not attempted to extend our results to \( N \to \infty \). The interesting studies of enumeration of Kekulé structures in polymers, which have appeared recently [19, 25], do not interfere with our results.

Acknowledgement

Financial support to BNC from The Norwegian Research Council for Science and the Humanities is gratefully acknowledged.