MATCH Communications in Mathematical and in Computer Chemistry

ISSN 0340 - 6253

## An Algorithm and FORTRAN Program for Automatic Computation of the Zhang–Zhang Polynomial of Benzenoids

## Chien-Pin Chou, Henryk A. Witek\*

Department of Applied Chemistry and Institute of Molecular Science, National Chiao Tung University, Hsinchu, Taiwan \*e-mail: hwitek@mail.nctu.edu.tw

(Received November 30, 2011)

#### Abstract

An automatic computer code is developed to calculate the Zhang–Zhang (ZZ) polynomial (*aka* Clar covering polynomial) for benzenoid systems. The code can be routinely applied to dense pericondensed benzenoids containing up to 500 carbon atoms. For catacondensed and quasi-linear pericondensed benzenoid systems, the limiting number of atoms is much larger and may exceed 10000. A parallel implementation of the code is also presented, which allows one to surpass these limits with a large number of CPUs. The developed program is applied for finding the ZZ polynomials of various classes of benzenoid systems; general techniques and algorithms applicable in this context are reviewed and discussed. A survey of new results obtained in that way is presented in a sequel to this paper [C.-P. Chou, Y. Li, and H.A. Witek, MATCH Commun. Math. Comput. Chem., submitted].

## 1. Introduction

Determination of topological invariants of chemical structures in many cases provides a basis for understanding difficult theoretical problems, particularly for system with a large number of atoms and with a dense network of bonds. The knowledge of topological descriptors is important for qualitative understanding of various molecular aspects in their chemistry and physics. In many cases, simple topological analysis can replace expensive quantum chemical calculations, yielding approximately equivalent information[1-21]. The analysis of molecular graphs for large aromatic molecules is sometimes the only source of chemical knowledge for systems exceedingly large for quantum chemical calculations. The knowledge of graph-theoretical invariants is particularly important in chemistry and physics of benzenoid molecules[6, 12, 22-29]. The number of useful topological descriptors characterizing such structures is abundant. Numerous theoretical studies employ these invariants for explication of benzenoid reactivity, geometric and electronic structure, and their various physicochemical properties [15, 17, 18, 30-34]. It is regrettable that in many cases the determination of many of such invariants (e.g., Kekulé number corresponding to a number of distributions of single and double bonds in an aromatic system or a Clar number corresponding to the maximal number of aromatic sextets in an aromatic system) still often provides such a painstaking problem.

The present manuscript gives a partial solution to this problem by reporting a computer code for automatic generation of the Zhang–Zhang (ZZ) polynomial[35-40], a combinatorial polynomial that can be defined for an aromatic structure on the base of its molecular graph. A wide class of topological invariants can be readily extracted from the ZZ polynomial, usually as its various coefficients; for details see Section 2. The only input necessary for the determination of the ZZ polynomial is the molecular graph, usually provided in the form of Cartesian coordinates of a given structure. The reported here program can be routinely applied to structures containing up to 200 carbon atoms. For larger structures, it is necessary to provide an additional input file containing straightforward, graph-theoretical information about minimum cuts of a given molecular graph; details are given in Section 4. To illustrate the computational complexity associated with the determination of the Zhang–Zhang polynomial for larger structures, it is insightful to quote the number of Kekulé structures that can be written for the hexagonal  $D_{6h}$  graphene flake containing 384 carbon atoms. This

number—directly corresponding to the zero-degree coefficient of the ZZ polynomial—is larger than  $10^{22}$ . Note that the coefficients in the ZZ polynomials tend to grow very fast with size for dense, two-dimensional benzenoid structures; it is reasonable to say that the computational time limitations set an upper limit of carbon atoms in our program for such structures to approximately 500. For linear and quasi-linear benzenoid systems, this limit is much larger and may exceed 10000 carbon atoms.

The reported here automatic program for determination of the Zhang–Zhang polynomial can be used in a twofold way. The computer code provided in Auxiliary Materials can be downloaded and compiled (or directly installed) on any UNIX platform. Examples, given in Section 5, show how to use it. An alternative—and much simpler—way of computing the ZZ polynomial can be achieved via the homepage (http://qcl.ac.nctu.edu.tw/zzpolynomial) we have established and will maintain for this purpose. Everything one needs for computing the ZZ polynomial via the homepage is to copy-and-paste the molecular coordinates of a given aromatic system into the special form available in the homepage. For the convenience of the user, simple graphical builder can be also used for generating the required benzenoid structure. For smaller systems, the corresponding ZZ polynomial will be printed instantaneously, for larger ones, within seconds or minutes. We hope that the automatic program presented here will be a useful tool for all scientists working in this field.

The structure of the current manuscript is as follows. Section 2 introduces the concept of the ZZ polynomial and illustrates it on two simple examples enabling a novice in the field understanding its structure. Section 3 presents a compilation of recursive properties of the ZZ polynomial that can be used for its determination. These two sections present well-known facts about the ZZ polynomial, which are included here to make the current manuscript self-contained. An additional motivation was the desire of further popularization of the ZZ polynomial as, in our opinion, its knowledge is not sufficiently spread among non-specialists. The algorithm used in our program is presented in Section 4 together with various important technical details concerning the presented program; their understanding will help in robust and skillful application of our program for solving various kind of combinatorial problems in the chemistry of benzenoids. This section is primarily meant to introduce various techniques that can be used for obtaining closed-form solutions for various classes of

benzenoid structures. Application of these techniques to general problems is presented in a sequel to the current manuscript[41] (hereafter referred to as II), which presents a derivation of Zhang–Zhang polynomials for various subclasses of benzenoid systems. The results presented in the current work are in general well-known, while the results presented in II are mostly new.

## 2. Zhang-Zhang polynomial

The combinatorial polynomial, known as the Zhang–Zhang polynomial or the ZZ polynomial, is a powerful and convenient tool for classifying resonance structures of a given aromatic system. The ZZ polynomial was introduced in 1996 in a series of papers [35-37, 39] by two Chinese mathematicians, Heping Zhang and Fuji Zhang, who were able to show that it possesses a number of very inviting recursive properties that can be used for its fast determination [35-37, 39]. As an example, they used these properties to calculate the ZZ polynomial for some small benzenoid systems and several series of catacondensed hexagonal systems. Zhang and Zhang proposed also a simple recursive algorithm for determination of the ZZ polynomial of any finite hexagonal system. Relatively high complexity and the graphtheoretical language, in which the results of Zhang and Zhang were announced, made their results available only to experts in this field. An important step toward the popularization of the ZZ polynomial in chemical community was the publication of a review article by Gutman, Furtula, and Balaban[42], who in brilliantly clear and ingenious exposition explained the theoretical issues leading to the definition of the ZZ polynomial and presented an algorithmically refined version of its recursive determination given originally by Zhang and Zhang. The computer program reported in this manuscript was largely stimulated by this work.

The basic idea necessary for understanding the importance of the ZZ polynomial is the concept of a Clar cover introduced by Zhang and Zhang. In general, a Clar cover of order k of a benzenoid system containing n carbon atoms is a permissible resonance structure containing exactly k aromatic Clar sextets and n/2-3k double bonds. The concept of Clar covers can be best understood on a simple example. In Figure 1 we show all the 13 conceivable Clar covers that can be written for a simple benzenoid molecule, pyrene; these resonance structures are constructed using either delocalized aromatic sextets of Clar or the localized double bonds of Kekulé. Structures **1–6** represent the well-known Kekulé formulas (Clar covers of order 0) of

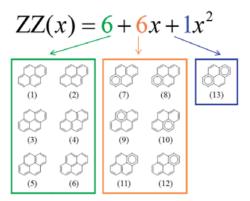


Figure 1. All possible Clar covers of pyrene with zero (1–6), one (7–12), and two (13) aconatic rings. The number of Clar covers can be conveniently summarized in the form of a combinatorial Zhang–Zhang polynomial  $ZZ(x) = 6 + 6x + x^2$ .

pyrene. Structure **13** represents the only Clar formula (Clar cover of order 2) of pyrene. Structures **7–12** are non-standard and correspond to Clar covers of order 1. Zhang and Zhang suggested that the number of Clar covers of each type can be conveniently summarized in a form of a combinatorial polynomial ZZ(x) in a dummy variable x, which only function is to represent the order of Clar covers of a given type. The ZZ polynomial of pyrene is thus  $ZZ(x) = 6 + 6x + x^2$ , which should be interpreted as follows: pyrene possesses six Clar covers of order 0, six Clar covers of order 1, and a single Clar cover of order 2. The advantages of such a representation will become apparent later, when we discuss the recursive properties of the ZZ polynomial. Before proceeding to this point, let us discuss another, slightly more complicated example. Figure 2 shows all the 69 conceivable Clar covers that can be written for coronene. We see from Figure 2 that all these Clar covers can be grouped into four families: 20 Clar covers of order 0, 32 Clar covers of order 1, 15 Clar covers of order 2, and two Clar covers of order 3. Consequently, the ZZ polynomial of coronene is  $ZZ(x) = 20 + 32x + 15x^2 + 2x^3$ .

The knowledge of the ZZ polynomial yields immediately the most important topological descriptors. The free coefficient of the ZZ polynomial, equal to 6 for pyrene and 20 for coronene, determines the number of Kekulé formulas *K*. The highest degree of the ZZ polynomial, equal to 2 for pyrene and 3 for coronene, corresponds to the Clar number *Cl*. The coefficient accompanying the  $x^{Cl}$  term determines the number of Clar formulas *C*, which is

equal to 1 for pyrene and 2 for coronene. Finally, the sum of all coefficients in the ZZ polynomial gives the total number of conceivable resonance structures that can be written for a given aromatic system.

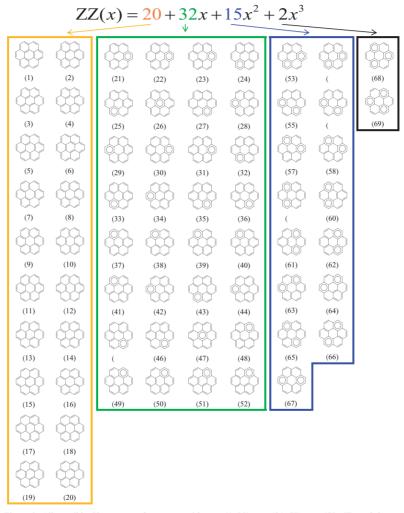


Figure 2. All possible Clar covers of coronene with zero (1–20), one (21–52), two (53–67), and three (68 and 69) aromatic rings. The number of Clar covers can be conveniently summarized in the form of a combinatorial Zhang–Zhang polynomial  $ZZ(x) = 20 + 32x + 15x^2 + 2x^3$ .

## 3. Recursive properties of Zhang-Zhang polynomial

The ZZ polynomial described in the previous Section possesses a number of properties, which enable its fast and convenient recursive determination. Proofs of these properties are elementary and can be found elsewhere[35-39, 42]. The properties are a simple consequence of the fact that in any  $\pi$ -aromatic system each carbon atom must participate either in a localized double bond or in a delocalized aromatic sextet.

Let us choose arbitrarily a bond connecting two neighboring carbon atoms A and B in some  $\pi$ -aromatic system S. If both of these atoms have more than one carbon neighbor, we can distinguish three distinct situations: i) the bond AB belongs simultaneously to two distinct benzene rings a and b, ii) the bond AB belongs to only one benzene ring a, and iii) the bond AB does not belong to any benzene rings. Each of these cases leads to a specific recursive property of the ZZ polynomial.

**Property 1** The bond between the atoms *A* and *B* is either a localized single or double bond or belongs to a delocalized aromatic sextet centered on the ring *a* or on the ring *b*. In this case, the ZZ polynomial of *S* can be expressed as a sum of four simpler ZZ polynomials, corresponding to four subsystems of *S*: i)  $S_1$  with the *AB* bond removed, ii)  $S_2$  with the atoms *A* and *B* removed, iii)  $S_3$  with the benzene ring *a* removed, and iv)  $S_4$  with the benzene ring *b* removed.

$$ZZ(S,x) = ZZ(S_1,x) + ZZ(S_2,x) + x \cdot ZZ(S_3,x) + x \cdot ZZ(S_4,x)$$
(1)

**Property 2** The bond between the atoms *A* and *B* is either a localized single or double bond or belongs to a delocalized aromatic sextet centered on the ring *a*. In this case, the *ZZ* polynomial of *S* can be expressed as a sum of three simpler *ZZ* polynomials, corresponding to three subsystems of *S*: i)  $S_1$  with the *AB* bond removed, ii)  $S_2$  with the atoms *A* and *B* removed, and iii)  $S_3$  with the benzene ring *a* removed.

$$ZZ(S,x) = ZZ(S_1,x) + ZZ(S_2,x) + x \cdot ZZ(S_3,x)$$
<sup>(2)</sup>

**Property 3** The bond between the atoms *A* and *B* is either a localized single bond or a localized double bond. In this case, the ZZ polynomial of *S* can be expressed as a sum of

two simpler ZZ polynomials, corresponding to two subsystems of S: i)  $S_1$  with the AB bond removed and ii)  $S_2$  with the atoms A and B removed.

$$ZZ(S,x) = ZZ(S_1,x) + ZZ(S_2,x)$$
(3)

Note that the ZZ polynomials of the subgraphs  $S_3$  and  $S_4$  in Eqs. (1) and (2) must be multiplied by *x* to compensate for removing the aromatic Clar sextets *a* and *b*.

If one of the atoms has only one neighbor (we call such an atom terminal), we obtain next recursive property of the ZZ polynomial.

**Property 4** If one of the atoms *A* and *B* has terminal character, then clearly the bond *AB* must have double character. In this case, the ZZ polynomial of *S* is equal to the ZZ polynomial of the subsystem  $S_1$  with the atoms *A* and *B* removed.

$$ZZ(S,x) = ZZ(S_1,x)$$
(4)

It may happen that removing a bond, ring, or atoms from the system S generates a subsystem that is disconnected, i.e., composed of *m* chemical fragments  $S_1, S_2, ..., S_m$  not connected by chemical bonds. This situation generates next recursive property of the ZZ polynomial.

**Property 5** The ZZ polynomial of a disconnected system *S* is equal to the product of the ZZ polynomials of its all fragments  $S_1, S_2, ..., S_m$ .

$$ZZ(S, x) = ZZ(S_1, x) \cdot ZZ(S_2, x) \cdot \ldots \cdot ZZ(S_m, x)$$
(5)

The listed above properties of the ZZ polynomial can be referred to as recursive in the following sense. For each of the listed properties, we express the ZZ polynomial of the molecular graph S as a sum or a product of the ZZ polynomials of its subgraph(s) with a smaller number of vertices or edges. Clearly, if one repeats such a procedure an appropriate number of times, the ZZ polynomial of S can be computed recursively in terms of the ZZ polynomials of some minimal subgraphs of S. This can be accomplished owing to the next two properties of the ZZ polynomial giving explicitly its values in two important cases.

**Property 6** The ZZ polynomial of a system *S* consisting of an odd number of carbon atoms vanishes.

$$ZZ(S,x) = 0 \tag{6}$$

**Property 7** The ZZ polynomial of a system *S* consisting of zero atoms is equal to 1. (This fact signifies that there is only one way of writing "nothing" similarly like in mathematics only one empty set exists.)

$$ZZ(S,x) = 1 \tag{7}$$

The presented here properties are sufficient to determine the ZZ polynomial of any  $\pi$ -aromatic structure consisting of carbon atoms. In next Section we discuss how these properties can be used for creating a general algorithm capable of determination of the ZZ polynomial for an arbitrary aromatic molecule.

# 4. Automatic program for determination of the Zhang–Zhang polynomial

The algorithm for determination of the ZZ polynomial was first proposed by Zhang and Zhang as a series of theorems that were subsequently used for computing the ZZ polynomial of a series of simple systems. The algorithm was given in more obvious form by Gutman, Furtula, and Balaban[42], who demonstrated an explicit sequence of operations that can be used for the determination of the ZZ polynomial for tetracene and benzo[a]pyrene. They noticed that every planar benzenoid system contains an edge bond belonging maximally to one benzene ring, which makes **Property 1** obsolete. This observation is important from the computational point of view since using **Property 1** results in higher-order branching of the algorithm and larger number of operations needed for determination of the ZZ polynomial. We note in passing that **Property 1** is important for non-planar aromatic systems such as nanotubes, where bonds belonging only to a single hexagon may not exist.

In the present manuscript we employ a slightly modified version of this algorithm to develop a computer program capable of computing the Zhang–Zhang polynomial in an automatic fashion. The algorithm has a recursive character and therefore it is useful to define first a recursive function ZZ(S), which constitutes the heart of our program. The flow chart for this function is shown in Figure 3. The value of the function ZZ for the structure S is computed explicitly only in two cases, when S has no atoms (using **Property** 7) or when S has an odd number of atoms (using **Property** 6). In all other cases, the ZZ polynomial of S is computed recursively as a product or a sum of the ZZ polynomials of simpler substructures of S using **Properties 2, 3,** and **5**. Before dividing S into substructures, an extensive use of **Property 4** eliminates all terminal atoms from the structure S, seriously reducing the computational cost in the subsequent partition steps. The inset in Figure 3 presents the structure of the main program in form of a few lines of meta-code; it is clear that the main program acts only as a wrapper for the recursive function ZZ.

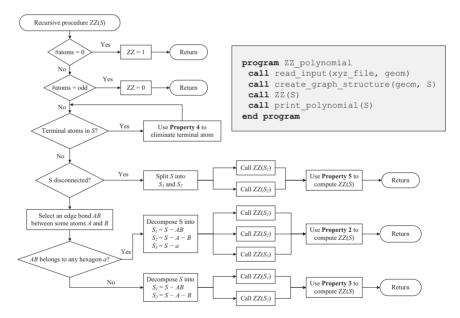


Figure 3. The flowchart of the Zhang-Zhang polynomial calculator program.

Our program is written in Fortran 95 and uses extensively many features of this programming language. Since the coefficients of the ZZ polynomial can attain very large values, it was necessary to implement in our code a new multi-precision integer type together with an appropriate interface allowing robust multiplication and addition of such large

numbers. At present the maximum number of digits for the multi-precision integer type is set to 1000, which allows for ZZ polynomials with coefficients up to  $10^{1000}$ ; it is highly unlikely that it will ever be achieved in practical calculations due to computational cost limitations. The computational cost grows very fast with the size of the system S; the scaling is highly non-linear. Therefore, in order to facilitate calculations of the ZZ polynomial for large structures, a parallel version of the program has been developed. We have used the OpenMP extension of the Fortran 95 compiler to achieve parallel execution. Since the formula for computing the ZZ polynomial has a recursive character, enabling the capability for parallel execution needs an automatic pretreatment performed by decomposing the initial structure S into a certain amount of its substructures (of the order of 2000), which are stored inside a temporary array T. Subsequently, the OpenMP library distributes the substructures between all the available processors, which compute the appropriate ZZ polynomial and return it to the array. After the ZZ polynomials are computed for all the generated substructures, the ZZ polynomial of the structure S is expressed in terms of the ZZ polynomials of the substructures. The initial pretreatment allows us also for reducing the computational cost associated with the computation of ZZ(S), because the initial steps of recursive decomposition of S often generate identical substructures, while the array T contains only unique substructures. The speed-up generated by this operation varies from system to system, ranging from 10% to 50%.

Despite of the parallel execution feature and the speed-up recursive reduction pretreatment discussed above, the computational cost associated with determination of the ZZ polynomial for dense 2D benzenoid molecules can be very large. To overcome this obstacle, we have implemented another very important modification to the original algorithm. The modification is based on the observation that from the computational point of view, using **Property 5** is much more robust in computing the ZZ polynomial than using the remaining properties. Therefore, it is always desirable to find a sequence of bond selections that leads to a decomposition of the structure *S* into disconnected fragments. Clearly, many such sequences exist, leading to various time savings during the ZZ polynomial calculations. The optimal partitioning of the structure *S* is achieved when it is decomposed into two segments having the same or similar size. The sequence of bond selections defining the optimal decomposition path can be specified by the user in an auxiliary input file *bondlist*. This particular feature is probably best explained on an example. In Figure 4 we show a benzenoid structure that can be conceived as a fusion of two hexagonal graphene flakes (supercoronenes). Computing the ZZ

polynomial in a traditional way with a random choice of the decomposition path requires rather long calculations taking as much as 126 seconds. However, if the user provides an auxiliary file called *bondlist* containing a list of two bonds, AB and CD (for details, see Figure 4), then the computational time is reduced to 0.02 second. For larger structures, the speed-up ratio may be considerably larger. It is desirable that the auxiliary file *bondlist* contains not only the optimal cut of the structure *S*, but also optimal cuts for the resulting substructures. In the best case, the user can provide the full decomposition path containing as many as needed bipartitions of the molecular graph *S* and its substructures. If the file *bondlist* is not provided by the user, the program attempts to define the optimal bipartitioning in an automatic fashion using a spectral partitioning algorithm, proposed originally by Fiedler[43-45] and popularized by Pothen, Simon, and Liou[46].

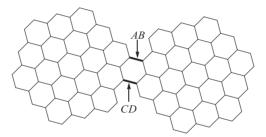


Figure 4. Starting the determination of the ZZ polynomial of two fused supercoronenes by selecting initially the bonds AB and CD speeds-up the computation process by a factor of 1000. For further discussion, see text.

The input file for our program is a standard XYZ file containing the Cartesian coordinates of a structure *S*. The parser ignores all other atoms than carbons. The user should make sure that the structure given as the input consists only of the  $sp^2$  carbons by removing all the side alkyl chains containing the  $sp^3$  carbons from the structure. In addition to the main input file, the user can specify an auxiliary input file defining the optimal decomposition path; details have been discussed in the previous paragraph. The *bondlist* file contains *m* lines, each line specifying two atom indices corresponding to the numbering of atoms in the original XYZ input file. The program chooses an edge *AB* following the order given in the auxiliary input file; if the bond list is empty or all its entries has been exhausted, the selection of the bond *AB* is performed by the spectral partitioning algorithm.

The program source code is available as Auxiliary Materials accompanying this manuscript available from XXX. Alternatively, the program can be accessed via the ZZ polynomial calculator homepage http://qcl.ac.nctu.edu.tw/zzpolynomial, which provides a possibility of computing the ZZ polynomial for small and moderate size systems (the calculation time has been restricted to 1 minute). The homepage contains also a convenient graphical tool for constructing benzenoid structures, which may serve as the input for the ZZ polynomial calculator. The graphical interface can be also used for defining the optimal decomposition path of a given benzenoid system. Graphical representation of the ZZ polynomial calculator is given in Figure 5.

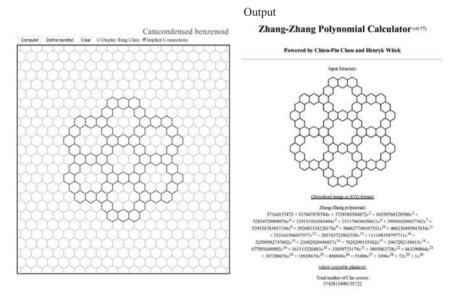


Figure 5. Snapshot of the web-based Zhang-Zhang polynomial calculator application.

## 5. Exemplary applications

To demonstrate the capabilities of the program presented in this manuscript we have applied it for determination of the ZZ polynomials of a series of benzenoid systems. Most of the results presented in this Section are well-known, some are new. We treat the results shown in this Section mainly as an introduction of techniques that can be used for determination of general closed-form of the ZZ polynomial of more complex systems. An application of these techniques to more complex system is given in the subsequent paper referred to as **II**.

#### a. Polyacenes

$$ZZ(L(1), x) = 2 + x$$

$$ZZ(L(2), x) = 3 + 2x$$

$$ZZ(L(3), x) = 4 + 3x$$

$$ZZ(L(4), x) = 5 + 4x$$

$$ZZ(L(4), x) = 5 + 4x$$

$$ZZ(L(5), x) = 6 + 5x$$

Figure 6. Linear polyacene molecule containing m rings

We start our exposition by discussing the ZZ polynomials of polyacenes, probably the simplest and best-studied benzenoid structures. Their Clar number is equal to 1 and their Kekulé number is equal to m + 1, where m denotes the number of benzene rings in a given polyacene L(m). The ZZ polynomials for this series of molecules have a remarkably simple structure (for details see Figure 6), which can be immediately generalized to any number of rings, giving

$$ZZ(L(m), x) = 1 + m(1 + x),$$
 (8)

which agrees with the analogous result given by Zhang and Zhang[37]. Note that the number of Kekulé structure determined in this way, 1 + m, is consistent with the previous studies[47].

#### b. Single armchair chain

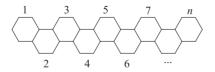


Figure 7. Single armchair chain N(n) of length *n*. Note that N(0) corresponds to ethylene, N(1), to benzene, N(2), to naphthalene, and N(3), to phenanthrene.

The example given in the previous paragraph can be considered as a one-dimensional sequence of benzene rings selected from a graphene sheet along the zigzag direction. In this paragraph we consider a related system, in which the hexagons are selected from the graphene sheet along the armchair direction. The single armchair chain N(n) of length n obtained in that way is shown schematically in Figure 7.

The ZZ polynomials for the shortest few structures, N(0)-N(11), are given by

$$ZZ(N(0),x) = 1$$

$$ZZ(N(1),x) = 2 + x$$

$$ZZ(N(2),x) = 3 + 2x$$

$$ZZ(N(3),x) = 5 + 5x + x^{2}$$

$$ZZ(N(4),x) = 8 + 10x + 3x^{2}$$

$$ZZ(N(5),x) = 13 + 20x + 9x^{2} + x^{3}$$

$$ZZ(N(6),x) = 21 + 38x + 22x^{2} + 4x^{3}$$

$$ZZ(N(7),x) = 34 + 71x + 51x^{2} + 14x^{3} + x^{4}$$

$$ZZ(N(8),x) = 55 + 130x + 111x^{2} + 40x^{3} + 5x^{4}$$

$$ZZ(N(9),x) = 89 + 235x + 233x^{2} + 105x^{3} + 20x^{4} + x^{5}$$

$$ZZ(N(10),x) = 144 + 420x + 474x^{2} + 256x^{3} + 65x^{4} + 6x^{5}$$

$$ZZ(N(11),x) = 233 + 744x + 942x^{2} + 594x^{3} + 190x^{4} + 27x^{5} + x^{6}.$$

An analysis of these formulas suggests that they are related by some kind of recurrence pattern, as, for example, the free coefficient of the ZZ polynomial of N(n) is equal to the sum of analogous coefficients for N(n-1) and N(n-2). To find this recurrence formula, we denote the unknown recurrence factors by *a* for N(n-1) and b + cx for N(n-2); their form is obvious from (9). (Note that multiplying the ZZ polynomial of N(2) by *x* would produce  $x^3$  for N(3), a term that is obviously obsolete.) It is easy to find the unknown factors (a = b = c = 1) by a simple analysis of the first two equations. Thus, we find that the ZZ polynomial of N(n) obeys a recursion relation

$$ZZ(N(n), x) = ZZ(N(n-1), x) + (x+1) \cdot ZZ(N(n-2), x).$$
(10)

Its correctness has been verified for a large number of cases. This recursion relation is a second-order linear homogeneous recurrence relation with constant coefficients. Its solution, found using standard techniques for solving linear recurrence equations as implemented in the symbolic algebra package MAPLE[48], can be expressed in a closed form, giving the ZZ polynomial of N(n) as

$$ZZ(N(n), x) = \frac{1}{2} \left( 1 + \frac{2x+3}{\sqrt{4x+5}} \right) \left( \frac{1+\sqrt{4x+5}}{2} \right)^n + \frac{1}{2} \left( 1 - \frac{2x+3}{\sqrt{4x+5}} \right) \left( \frac{1-\sqrt{4x+5}}{2} \right)^n.$$
(11)

The recurrence equation (10) was derived before by Zhang and Zhang directly from a recursive decomposition of N(n)[37]. An expression equivalent to Eq. (11) was also given in [37]; however, the formula given here by Eq. (11) is more explicit and simpler than the expression given by Zhang and Zhang. Note that setting x = 0 reduces Eq. (11) to Eq. (4.12) in [49], which gives the number of Kekulé structures of a single armchair chain. It may be useful to notice that Eq. (11) can be also expressed in a simple additive form

$$ZZ(N(n), x) = \sum_{k=0}^{n} {\binom{n+1-k}{k}} (1+x)^{k},$$
(12)

which can be more useful for practical calculations.

#### c. Parallelograms

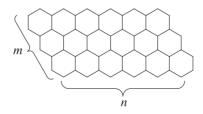


Figure 8. Parallelogram M(m,n)

Let us consider next a two-dimensional paralleloidal graphene flake as shown in Figure 8, which we will denote by M(m,n). The derivation of a general form of the ZZ polynomial for this type of structures may proceed as follows. First, we consider a narrower family of structures by fixing the value of *m* to 2. The first eight ZZ polynomials for the M(2,n) family are given by

$$ZZ(M(2,1),x) = 3 + 2x$$

$$ZZ(M(2,2),x) = 6 + 6x + x^{2}$$

$$ZZ(M(2,3),x) = 10 + 12x + 3x^{2}$$

$$ZZ(M(2,4),x) = 15 + 20x + 6x^{2}$$

$$ZZ(M(2,5),x) = 21 + 30x + 10x^{2}$$

$$ZZ(M(2,6),x) = 28 + 42x + 15x^{2}$$

$$ZZ(M(2,7),x) = 36 + 56x + 21x^{2}$$

$$ZZ(M(2,8),x) = 45 + 72x + 28x^{2}.$$
(13)

First thing that could be noticed is that the Clar number Cl for M(2,n) is never larger than 2, suggesting that the general form of the ZZ polynomial can be given by  $A + Bx + Cx^2$ , where A, B, and C are some functions of n. These coefficients are easily found by inspection of the integer sequences appearing in the subsequent ZZ polynomials. Before giving their closed form, let us first discuss general techniques used for their identification. We use two principal tools for finding the closed form of integer sequences appearing in our work. The first tool is the On-Line Encyclopedia of Integer Sequences (OEIS)[50], which requires specifying the first few terms of a given sequence in order to recognize it. In general, the recognition is not unique as many sequences share the same initial terms. A specification of larger number of entries obviously helps, in many cases yielding the unique solution. For each of the recognized sequences, a discussion of its properties is given along with methods of its reproducing, either by a direct equation, or by recurrence, or by means of a generating function. The second tool used here is the Wolfram  $Alpha (W\alpha)$  [51] environment, which often is able to provide us with a simple closed form of a function generating given sequence. The sequence (1, 3, 6, 10, 15, 21, 28, ...) corresponding in (13) to the function A is readily identified by both tools as  $\frac{1}{2}(n+2)(n+1)$  which can be written more compactly as the

binomial coefficient  $\binom{n+2}{2}$ . Similarly, the functions *B* and *C* are identified from the corresponding sequences as  $n(n+1) = 2\binom{n+1}{2}$  and  $\frac{1}{2}n(n-1) = \binom{n}{2}$ , respectively. Therefore, the general form of the ZZ polynomial formula for M(2,n) is given by

$$ZZ(M(2,n),x) = \binom{n+2}{2} + 2\binom{n+1}{2}x + \binom{n}{2}x^2.$$
 (14)

Similar analysis applied to the M(3,n), M(4,n) and M(5,n) families of structures gives the following series of closed-form ZZ polynomials

$$ZZ(M(3,n),x) = \binom{n+3}{3} + 3\binom{n+2}{3}x + 3\binom{n+1}{3}x^2 + \binom{n}{3}x^3,$$
 (15)

$$ZZ(M(4,n),x) = \binom{n+4}{4} + 4\binom{n+3}{4}x + 6\binom{n+2}{4}x^2 + 4\binom{n+1}{4}x^3 + \binom{n}{4}x^4, \quad (16)$$

$$ZZ(M(5,n),x) = \binom{n+5}{5} + 5\binom{n+4}{5}x + 10\binom{n+3}{5}x^2 + 10\binom{n+2}{5}x^3 + 5\binom{n+1}{5}x^4 + \binom{n}{5}x^5.$$
(17)

The numerical factors appearing in (14)–(17), i.e., the sequences (1, 2, 1), (1, 3, 3, 1), (1, 4, 6, 4, 1), and (1, 5, 10, 10, 5, 1) are easily identified by OEIS as another set of binomial coefficients  $\binom{m}{i}$  with i = 0, ..., m. Combining (14)–(17), the general formula of the ZZ polynomial for M(m,n) is

$$ZZ(M(m,n),x) = \sum_{i=0}^{m} \left( \binom{m}{i} \binom{n+m-i}{m} \cdot x^{i} \right).$$
(18)

This formula has been given previously by Gutman and Borovićanin, who derived it directly from the recursive properties of the ZZ polynomials of M(m,n)[52]. Note that the number of Kekulé structure of M(m,n) obtained by setting x = 0 also agrees with the previously reported formula[47]. Eq. (18) is certainly correct but it does not emphasize the permutational

invariance  $n \leftrightarrow m$  of the ZZ polynomial of the parallelogram M(m,n). From Figure 8 it is clear that a mirror inversion of M(m,n) produces M(n,m) with all Clar covers of one structure being mirror images of the second one. Obviously, the ZZ polynomials are equal, which should be also obvious from the form of the ZZ polynomial. Therefore, we prefer to represent Eq. (18) in an equivalent permutationally invariant form given by

$$ZZ(M(m,n),x) = \sum_{i=0}^{\min(m,n)} \frac{(m+n-i)!}{(m-i)!(n-i)!i!} x^{i} = \sum_{i=0}^{\min(m,n)} \frac{\binom{m+n}{m}\binom{n}{i}}{\binom{m+n}{i}} x^{i}.$$
 (19)

#### d. Polyphenylenes

$$ZZ(P(1),x) = (2+x)$$

$$ZZ(P(2),x) = (2+x)^{2}$$

$$ZZ(P(3),x) = (2+x)^{3}$$

$$ZZ(P(4),x) = (2+x)^{4}$$

$$ZZ(P(5),x) = (2+x)^{5}$$

Figure 9. Linear polyphenylene P(n) of length n.

The linear chain of *n* phenyl rings is usually referred to as *n*-polyphenylene; we abbreviate it here as P(n). In close analogy to linear polyacenes, the ZZ polynomials for this series of molecules have a remarkably simple structure (for details see Figure 9), which can be immediately generalized to any number of rings, giving

$$ZZ(P(n), x) = (2+x)^n.$$
 (20)

The number of Kekulé structures can be easily obtained by setting x = 0; the result  $2^n$  agrees with the previously reported formulae[24, 53].

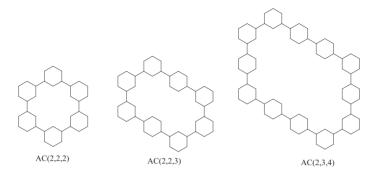


Figure 10. Armchair-type coronoids for different n, m, and l,

The next family of systems considered by us here are cyclic polyphenylenes consisting of six consecutive linear polyphenylene segments of length n, m, l, n, m, and l, respectively. We refer to these structures as armchair coronoids and abbreviate them as AC(n,m,l). The ZZ polynomials computed for the smallest few structures are given by

$$\begin{cases} ZZ(AC(2,2,2),x) = (x+2)^{6} \\ ZZ(AC(2,2,3),x) = (x+2)^{8} \\ ZZ(AC(2,2,4),x) = (x+2)^{10} \\ ZZ(AC(2,3,3),x) = (x+2)^{10} \\ ZZ(AC(2,4,3),x) = (x+2)^{12} \\ ZZ(AC(3,3,3),x) = (x+2)^{12} \\ ZZ(AC(4,4,4),x) = (x+2)^{18} \end{cases}$$
(21)

It is immediately clear that the ZZ polynomials of AC(n,m,l) do not depend on the particular values of the indices *n*, *m*, *l* but only on their sum n + m + l. The closed form of the Zhang–Zhang polynomial of armchair–type coronoids is then given by

$$ZZ(AC(n,m,l),x) = (x+2)^{2(n+m+l-3)}.$$
(22)

This formula is in fact identical to Eq. (20) obtained for linear polyphenylenes, as 2n + 2m + 2l - 6 simply denotes the number of phenyl rings in a given structure.

Note that in the case of linear polyacenes adding an additional benzene ring to the structure was introducing an extra *additive* factor 1 + x; here, adding an additional benzene ring introduces an extra *multiplicative* factor 2 + x.

#### e. The S(n) series of structures

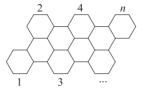


Figure 11. The S(n) structure

The next system studied here is shown in Figure 11. It was first studied by Randić, who derived a recurrence formula for the number of Kekulé structure of the S(n). The ZZ polynomial of the shortest few S(n) are given by

$$ZZ(S(1),x) = 2 + x$$

$$ZZ(S(2),x) = 4 + 4x + x^{2}$$

$$ZZ(S(3),x) = 9 + 13x + 6x^{2} + x^{3}$$

$$ZZ(S(4),x) = 20 + 38x + 26x^{2} + 8x^{3} + x^{4}$$

$$ZZ(S(5),x) = 45 + 106x + 96x^{2} + 43x^{3} + 10x^{4} + x^{5}$$

$$ZZ(S(6),x) = 101 + 284x + 321x^{2} + 190x^{3} + 64x^{4} + 12x^{5} + x^{6}$$

$$ZZ(S(7),x) = 227 + 742x + 1006x^{2} + 742x^{3} + 328x^{4} + 89x^{5} + 14x^{6} + x^{7}$$

$$ZZ(S(8),x) = 510 + 1900x + 3006x^{2} + 2660x^{3} + 1460x^{4} + 518x^{5} + 118x^{6} + 16x^{7} + x^{8}.$$
(23)

The recurrence relation for this series can be found easily as a generalization of the Randić, Ohkami, and Hosoya[8, 29] results. It has the following form

$$ZZ(S(n), x) = (2+x)ZZ(S(n-1), x) + (1+x)ZZ(S(n-2), x) -(1+x)^2 ZZ(S(n-3), x).$$
(24)

One can use this formula to obtain the ZZ polynomial for any number of this series. A closed-form formula can be found by solving recurrence given by Eq. (24); however, due to its third-order characteristic polynomial with complex roots, the recursion formula is too

complicated for practical use. It is much more convenient to find a generating function F(z) instead, by

$$F(z) = \frac{1 - (x+1)z^2}{(x+1)^2 z^3 - (x+1)z^2 - (x+2)z + 1}.$$
(25)

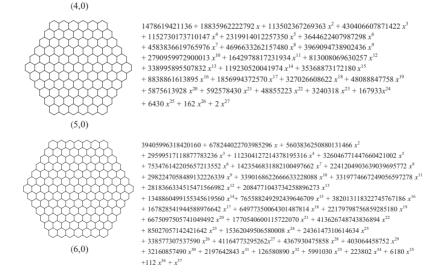
For those not familiar with this technique, we briefly explain that the ZZ polynomial for the S(n) structures can be obtained as the n<sup>th</sup> coefficient in the Taylor expansion of this function at z = 0, which can be written explicitly as

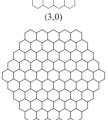
$$ZZ(S(n), x) = \frac{1}{n!} \frac{d}{dz^n} F(z) \Big|_{z=0}.$$
 (26)

#### f. Hexagonal graphene flakes

The structures considered until now were relatively small. We want to show now that the ZZ polynomials of simple benzenoid structures can grow very fast and their coefficients can attain very large numerical values, but nevertheless our program is still capable of finding the results. In Figure 12, we present the ZZ polynomials of a series of hexagonal graphene flakes (HGF), which can be considered as a continuation of the series: benzene, coronene, supercoronene, etc. We refer to these structures as (n, 0), which signifies two important properties of these structures. First, the studied here graphene flakes have a zigzag-like edges, which resembles the structure of (n, 0) carbon nanotubes. Second, the index n denotes the number of benzene ring layers encircling the central benzene ring. For benzene, it is equal to 0 resulting in the (0, 0) notation and for coronene, it is equal to 1 giving the notation of (1, 0). It is immediately clear that the ZZ polynomial grows very fast with n in quite irregular way. It is not obvious if a closed formula for the ZZ polynomial exists; if it does, the form must be quite complex. Usually it is easier to find a recurrence formula relating the structures in a given families of molecules. The complexity of the presented formulas suggests that even this task cannot be easily achieved for the (n, 0) hexagonal graphene flakes.

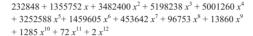
Figure 12. The ZZ polynomials of HGF(n,0), for n = 0 - 6





 $\begin{array}{l} 267227532+2389804604\,x+9865265424\,x^2+24946185302\,x^3\\ +\,43244524658\,x^4+54500920980\,x^5+51684194860\,x^6+37663193263\,x^7\\ +\,21358371744\,x^8+9490314264\,x^9+3312396458\,x^{10}+907310265\,x^{11}\\ +\,194327374\,x^{12}+32357890\,x^{13}+4158672\,x^{14}+408889\,x^{15}+30314\,x^{16}\\ +\,1635\,x^{17}+58\,x^{18}+x^{19} \end{array}$ 







(0.0)

2 + x

 $980 + 3308 x + 4458 x^{2} + 3065 x^{3} + 1140 x^{4} + 225x^{5} + 22 x^{6} + x^{7}$ 



## 6. Conclusion

We have developed a computer program for automatic determination of the Zhang-Zhang (ZZ) polynomial for an arbitrary benzenoid system. The program can be downloaded from the Auxiliary Materials accompanying this manuscript or it can be directly accessed online at http://qcl.ac.nctu.edu.tw/zzpolynomial together with a handy graphical interface allowing for fast and convenient definition of the studied system. The program can be routinely used for determination of the ZZ polynomial for systems containing up to 200 carbon atoms. For larger benzenoids, in order to speed up the computational process, the user may choose to provide an additional input file containing straightforward information about optimal bi-partitioning of the molecular graph. Optimal bipartitioning allows one to reduce seriously the computational time and to extend an upper limit of carbon atoms in the studied structures to approximately 500 for dense pericondensed hexagonal systems. For catacondensed and quasi-linear pericondensed systems, this limit is much larger and may exceed 10000 carbon atoms. Further extension to even larger structures can be achieved in parallel mode execution. Numerous examples show how to apply the presented program to combinatorial problems in the theory of benzenoid systems. A survey of new results obtained in that way is presented in the subsequent manuscript[41]. We truly hope that this code will prove helpful for anyone who is interested in fast and straightforward determination of the Zhang-Zhang polynomial for any class of benzenoid systems.

## 7. Acknowledgment

National Science Council of Taiwan (grant NSC 99-2113-M-009-011-MY3) and Ministry of Education (MOE-ATU project) are acknowledged for financial support.

## References

- G. A. Koppel, R. E. Koehler, Resonance energies of aromatic hydrocarbons. A quantitative test of resonance theory, J. Am. Chem. Soc. 95 (1973) 2404-2406.
- [2] W. C. Herndon, M. L. Ellzey, Resonance theory. V. Resonance energies of benzenoid and nonbenzenoid p systems, J. Am. Chem. Soc. 96 (1974) 6631-6642.
- [3] I. Gutman, W. C. Herndon, On resonance theory, Chem. Phys. Lett. 34 (1975) 387-391.

- [4] M. Randić, Conjugated circuits and resonance energies of benzenoid hydrocarbons, *Chem. Phys. Lett.* 38 (1976) 68-70.
- [5] D. Biermann, W. Schmidt, Diels-Alder reactivity of polycyclic aromatic-hydrocarbons.
   3. New experimental and theoretical results, *Isr. J. Chem.* 20 (1980) 312-318.
- [6] A. Graovac, N. Trinajstić, M. Randić, Graph-theoretical search for benzenoid polymers with zero energy-gap, *Croat. Chem. Acta.* 53 (1980) 571-579.
- [7] Y. S. Jiang, Graph theory of molecular orbitals Calculation of aN the determinant of adjacency matrix and the stability of molecules, *Sci. Sin.* 23 (1980) 847-861.
- [8] M. Randić, Resonance energy of very large benzenoid hydrocarbons, *Int. J. Quantum Chem.* 17 (1980) 549-586.
- [9] I. Gutman, Topological studies on heteroconjugated molecules. 3. On the law of alternating polarity, Z. Naturforsch. 36a (1981) 1112-1114.
- [10] I. Gutman, Topological resonance energy of very large benzenoid hydrocarbons, Z. Naturforsch. 36a (1981) 128-131.
- [11] I. Gutman, On the topological resonance energy of hetero-conjugated molecules, *Croat. Chem. Acta.* 54 (1981) 75-80.
- [12] I. Gutman, B. Mohar, More difficulties with topological resonance energy, *Chem. Phys. Lett.* 77 (1981) 567-570.
- [13] I. Gutman, O. E. Polansky, Cyclic conjugation and the Hückel molecular-orbital model, *Theor. Chim. Acta*. 60 (1981) 203-226.
- [14] J. Cioslowski, I. Gutman, Upper-bounds for the total  $\pi$ -electron energy of benzenoid hydrocarbons and their relations, *Z. Naturforsch.* **41a** (1986) 861-865.
- [15] A. R. Ashrafi, K. Amini, Clar covers and Kekulé index of dendrimer nanostars, Optoelectron. Adv. Mat. 4 (2010) 877-880.
- [16] H. Hosoya, From how to why. Graph-theoretical verification of quantum-mechanical aspects of  $\pi$ -electron behaviors in conjugated systems, *Bull. Chem. Soc. Jpn.* **76** (2003) 2233-2252.
- [17] I. Gutman, S. Gojak, B. Furtula, S. Radenković, A. Vodopivec, Relating total  $\pi$ electron energy and resonance energy of benzenoid molecules with Kekulé- and Clarstructure-based parameters, *Monatsh. Chem.* **137** (2006) 1127-1138.
- [18] M. Randić, A. T. Balaban, Partitioning of π-electrons in rings for Clar structures of benzenoid hydrocarbons, J. Chem. Inf. Model. 46 (2006) 57-64.

- [20] H. Hosoya, Aromaticity index can predict and explain the stability of polycyclic conjugated hydrocarbons, *Monatsh. Chem.* 136 (2005) 1037-1054.
- [21] H. Hosoya, Mathematical foundation of the organic electron theory How do πelectrons flow in conjugated systems? J. Mol. Struct. (THEOCHEM). 461-462 (1999) 473-482.
- [22] G. G. Hall, A graphical model of a class of molecules, Int. J. Math. Educ. Sci. Tech. 4 (1973) 233-240.
- [23] G. G. Hall, On the eigenvalues of molecular graphs, Mol. Phys. 33 (1977) 551-557.
- [24] N. Trinajstić, Chemical Graph Theory, CRC Press, Boca Raton, 1983.
- [25] S. E. Stein, R. L. Brown, Chemical theory of graphite-like molecules, *Carbon* 23 (1985) 105-109.
- [26] N. Ohkami, Graph-theoretical analysis of the sextet polynomial. Proof of the correspondence between the sextet patterns and Kekulé patterns, J. Math. Chem. 5 (1990) 23-42.
- [27] M. Randić, S. El-Basil, Graph theoretical analysis of large benzenoid hydrocarbons, J. Mol. Struct. (Theochem) 304 (1994) 233-245.
- [28] I. Gutman, Topological properties of benzenoid systems. 21. Theorems, conjectures, unsolved problems, *Croat. Chem. Acta.* 56 (1983) 365-374.
- [29] N. Ohkami, H. Hosoya, Topological dependency of the aromatic sextets in polycyclic benzenoid hydrocarbons - Recursive relations of the sextet polynomial, *Theor. Chim. Acta.* 64 (1983) 153-170.
- [30] S. Zhou, H. Zhang, I. Gutman, Relations between Clar structures, Clar covers, and the sextet-rotation tree of a hexagonal system, *Discr. Appl. Math.* 156 (2008) 1809-1821.
- [31] K. Amini, A. R. Ashrafi, Relations between Clar structures, Clar covers and sextetrotations of dendrimer nanostars, *Optoelectron. Adv. Mat.* 3 (2009) 1076-1079.
- [32] S. Gojak, I. Gutman, S. Radenković, A. Vodopivec, Relating resonance energy with the Zhang-Zhang polynomial, J. Serb. Chem. Soc. 72 (2007) 673-679.
- [33] I. Gutman, S. Gojak, B. Furtula, Clar theory and resonance energy, *Chem. Phys. Lett.* 413 (2005) 396-399.

- [34] S. Gojak, S. Radenković, R. Kovačević, S. Stanković, J. Đurđević, I. Gutman, A difference between the π-electron properties of catafusenes and perifusenes, *Polyc. Aromat. Comp.* 26 (2006) 197-206.
- [35] F. Zhang, H. Zhang, A new enumeration method for Kekulé structures of hexagonal systems with forcing edges, J. Mol. Struct. (Theochem) 331 (1995) 255-260.
- [36] F. J. Zhang, H. P. Zhang, Y. T. Liu, The Clar covering polynomial of hexagonal systems. 2. An application to resonance energy of condensed aromatic hydrocarbons, *Chin. J. Chem.* 14 (1996) 321-325.
- [37] H. P. Zhang, F. J. Zhang, The Clar covering polynomial of hexagonal systems. 1. Discr. Appl. Math. 69 (1996) 147-167.
- [38] H. P. Zhang, The Clar covering polynomial of hexagonal systems with an application to chromatic polynomials, *Discr. Math.* **172** (1997) 163-173.
- [39] H. P. Zhang, F. J. Zhang, The Clar covering polynomial of hexagonal systems III. Discr. Math. 212 (2000) 261-269.
- [40] S. Gojak, S. Stanković, I. Gutman, B. Furtula, Zhang-Zhang polynomial and some of its applications, in: I. Gutman (Ed.), *Mathematical Methods in Chemistry*, Prijepolje Museum, Prijepolje, 2006, pp. 141-158.
- [41] C. P. Chou, Y. Li, H. A. Witek, Zhang-Zhang polynomials of various classes of benzenoid systems, *MATCH Commun. Math. Comput. Chem.* 68 (2012) 31-64.
- [42] I. Gutman, B. Furtula, A. Balaban, Algorithm for simultaneous calculation of Kekulé and Clar structure counts, and Clar number of benzenoid molecules, *Polyc. Aromat. Comp.* 26 (2006) 17-35.
- [43] M. Fiedler, Algebraic connectivity of graphs, Czech. Math. J. 23 (1973) 298-305.
- [44] M. Fiedler, Eigenvectors of acyclic matrices, Czech. Math. J. 25 (1975) 607-618.
- [45] M. Fiedler, Property of eigenvectors of nonnegative symmetric matrices and its application to graph theory, *Czech. Math. J.* 25 (1975) 619-633.
- [46] A. Pothen, H. D. Simon, K. P. Liou, Partitioning sparse matrices with eigenvectors of graphs, SIAM J. Matrix Anal. Appl. 11 (1990) 430-452.
- [47] M. Gordon, W. H. T. Davison, Theory of resonance topology of fully aromatic hydrocarbons. I, J. Chem. Phys. 20 (1952) 428-435.
- [48] M. B. Monagan, W. M. Software, *Maple V programming guide*, Springer, Berlin, 1998.

- [49] S. J. Cyvin, I. Gutman, Kekulé Structures in Benzenoid Hydrocarbons, Springer, Berlin, 1988.
- [50] The On-Line Encyclopedia of Integer Sequences, http://oeis.org, OEIS Foundation, 2011.
- [51] Wolfram|Alpha, http://www.wolframalpha.com, Wolfram Alpha LLC
- [52] I. Gutman, B. Borovićanin, Zhang-Zhang polynomial of multiple linear hexagonal chains, Z. Naturforsch. 61a (2006) 73-77.
- [53] T. F. Yen, Resonance topology of polynuclear aromatic hydrocarbons, *Theor. Chim. Acta.* 20 (1971) 399-404.