

Algorithms and Extremal Problem on Wiener Polarity Index *

Wenxue Du, Xueliang Li, Yongtang Shi

Center for Combinatorics and LPMC-TJKLC, Nankai University

Tianjin 300071, P.R. China. E-mail: lxl@nankai.edu.cn

(Received October 6, 2008)

Abstract

The Wiener polarity index $W_P(G)$ of a graph $G = (V, E)$ is the number of unordered pairs of vertices $\{u, v\}$ of G such that $d_G(u, v) = 3$. In this paper, we consider the index for connected graphs. In the first part, we describe a linear time algorithm **APT** for computing the index of trees, and then characterize the trees maximizing the index among all trees of given order. In the second part, we present an algorithm which computes the index $W_P(G)$ for any given connected graph G on n vertices in time $O(M(n))$, where $M(n)$ denotes the time necessary to multiply two $n \times n$ matrices of small integers (which is currently known to be $O(n^{2.376})$).

1 Introduction

We use Trinajstić [14] for terminology and notations. Let G be a connected (molecular) graph. The *distance* between two vertices u and v in G , denoted by $d_G(u, v)$, is the length of a shortest path between u and v in G . A tree is a connected acyclic graph. It is well known that for any two vertices u and v in a tree T , there exists exactly one path between u and v in T . Thus, the distance between two vertices u and v in T is the length of the path between u and v in T . The *Wiener polarity index* of a graph $G = (V, E)$, denoted by $W_P(G)$, is defined by

$$W_P(G) := \#\{\{u, v\} \mid d_G(u, v) = 3, u, v \in V\}, \quad (1)$$

*Supported by PCSIRT, NSFC and the "973" program.

which is the number of unordered pairs of vertices $\{u, v\}$ of G such that $d_G(u, v) = 3$. In organic compounds, say paraffin, this number is the number of pairs of carbon atoms which are separated by three carbon-carbon bonds. The name “Wiener polarity index” for the quantity defined in Equation (1) is introduced by Harold Wiener [15] in 1947. Wiener himself conceived the index only for acyclic molecules and defined it in a slightly different – yet equivalent – manner. In the same paper, Wiener also introduced another index for acyclic molecules, called *Wiener index* or *Wiener distance index* and defined by

$$W(G) := \sum_{\{u,v\} \subseteq V} d_G(u, v).$$

Wiener [15] used a linear formula of W and W_P to calculate the boiling points t_B of the paraffins, *i.e.*,

$$t_B = aW + bW_P + c,$$

where a, b and c are constants for a given isomeric group.

The Wiener index $W(G)$ is popular in chemical literatures. In the mathematical literature, it seems to be studied firstly by Entringer *et al.* [8] in 1976. From then on, many researchers studied the Wiener index in different ways. For instance, one can see [1], [2], [3] [6], [8], [9], [11], [12] and [15] for theoretical aspects, and [4], [10] and [13] for algorithmic and computational aspects. Recently, Dobrynin *et al.* wrote a comprehensive survey [7] for the Wiener index. The reader is referred to the paper for further details.

However, it seems that less attention has been paid for the Wiener polarity index $W_P(G)$ up to now. In the present paper, we consider the index for connected graphs. By the definition of Wiener polarity index, one can readily check that $W_P(K_{1,n-1}) = 0$. Moreover, $W_P(T) > 0$ for any tree T of order $n \geq 4$. Thus, a star $K_{1,n-1}$ minimize the Wiener polarity index among all trees of given order. In Section 2, we first give a linear time algorithm **APT** for computing the index of trees, and then characterize the trees maximizing the index among all trees of given order. In Section 3, we present an algorithm which computes the index $W_P(G)$ for any given connected graph G on n vertices in time $O(M(n))$, where $M(n)$ denotes the time necessary to multiply two $n \times n$ matrices of small integers (which is currently known to be $O(n^{2.376})$ [5]).

2 Wiener Polarity Index for Trees

In this section, we consider the Wiener polarity index W_P for trees. We first introduce a linear time algorithm **APT** for computing the index of trees, and then consider the problem of determining which trees maximize the index among all trees of given order.

2.1 A Linear Time Algorithm

According to the definition of Wiener polarity index, Equation (1), one can readily design an algorithm in $O(|V(T)|(\Delta(T) - 1)^2)$ time for computing the index $W_P(T)$ of a tree T by exhausted searching. The algorithm, however, might be not linear time if the maximum degree $\Delta(T)$ of T is large. In fact, we can get a linear time algorithm **APT** for computing the index of a tree T due to a good property that for any two vertices u and v in a tree T , there exists exactly one path between u and v in T . Furthermore, we have the following result.

Lemma 1. *Let $T = (V, E)$ be a tree. Then*

$$W_P(T) = \sum_{uv \in E} (d_T(u) - 1)(d_T(v) - 1). \quad (1)$$

Proof. We first define a set $D_3(T)$ as follows:

$$D_3(T) := \{\{u, v\} \mid d_T(u, v) = 3, u, v \in V\}.$$

Clearly, $W_P(T) = |D_3(T)|$ by the definition of Wiener polarity index. Next, we introduce another set $S_E(T)$ as follows:

$$S_E(T) := \{\{u, v\} \mid \exists xy \in E \text{ such that } ux \text{ and } vy \in E\}.$$

One can readily see that

$$|S_E(T)| = \sum_{xy \in E} (d_T(x) - 1)(d_T(y) - 1).$$

Let $\varphi : D_3(T) \rightarrow S_E(T)$ be a mapping such that $\varphi(\{u, v\}) = \{u, v\}$ for any $\{u, v\} \in D_3(T)$. One can easily check that the mapping φ is a bijection. Thus, $|D_3(T)| = |S_E(T)|$, and then Equation (1) follows. \square

Let T be a tree of order n . In the sequel, we use a list $li(T)$ concerning edges of T and degrees of $V(T)$ to represent T . Formally, we define

$$li(T) := \{e_1 = x_1y_1, e_2 = x_2y_2, \dots, e_{n-1} = x_{n-1}y_{n-1}, d_T(v_1), \dots, d_T(v_n)\}.$$

The following is a linear time algorithm for computing the Wiener polarity index $W_P(T)$ of a tree T represented by a list $li(T)$ of T .

APT

Input: A tree T of order n represented by a list $li(T)$ of T .

Output: Wiener polarity index $WP(T)$ of T .

```

begin
  WP(T) := 0
  for all edges u[i]v[i] of T, i:=1 to (n-1) do
    WP(T) = WP(T) + (d(u[i])-1)(d(v[i])-1)
end
```

According to Lemma 1, the algorithm **APT** correctly computes the Wiener polarity index $W_P(T)$ of T . Obviously, the algorithm **APT** can be done in $O(n)$ time. Hence, we have the following result.

Theorem 1. *Let T be a tree of order n . Then the algorithm **APT** correctly computes the Wiener polarity index $W_P(T)$ of T in $O(n)$ time.*

2.2 Extremal Trees

As we mentioned in the introduction, a star $K_{1,n-1}$ minimizes the Wiener polarity index among all trees of order n . The goal of this part is to characterize the trees maximizing the index among all trees of given order. For this purpose, we first consider a simple case.

The *diameter* of a connected graph G , denoted by $diam(G)$, is the maximum distance between two vertices of G . Since there exists exactly one path between any two vertices of a tree, the diameter of a tree is the length of a longest path in the tree. In what follows, we use $\mathcal{T}(n)$ to denote the set of trees on n vertices. Let $T \in \mathcal{T}(n)$ with $diam(T) = 3$, and let $P_L(T) = v_0v_1v_2v_3$ be a longest path in T . Then, every longest path in T goes through v_1v_2 . It follows from Equation (1) that

$$W_P(T) = (d_T(v_1) - 1)(d_T(v_2) - 1).$$

Thus,

$$W_P(T) \leq \left(\left\lceil \frac{n}{2} \right\rceil - 1 \right) \left(\left\lfloor \frac{n}{2} \right\rfloor - 1 \right) = \left\lceil \frac{n-2}{2} \right\rceil \left\lfloor \frac{n-2}{2} \right\rfloor. \quad (2)$$

On the other hand, if $T \in \mathcal{T}(n)$, $n \geq 4$, and $P_L(T) = v_0v_1v_2v_3$ is a longest path of T with $d_T(v_1) = \lceil \frac{n}{2} \rceil$ and $d_T(v_2) = \lfloor \frac{n}{2} \rfloor$ then $W_P(T) = \lceil \frac{n-2}{2} \rceil \lfloor \frac{n-2}{2} \rfloor$. Hence, the following set

$$\mathcal{T}_3(n) := \{T \in \mathcal{T}(n) \mid diam(T) = 3 \text{ and } W_P(T) = \lceil \frac{n-2}{2} \rceil \lfloor \frac{n-2}{2} \rfloor\}$$

is not empty. In the following, we will see that the above value $\lceil \frac{n-2}{2} \rceil \lfloor \frac{n-2}{2} \rfloor$ is the maximum value of W_P for trees of order n . Hence, $\mathcal{T}_3(n)$ is one of the classes of extremal trees maximizing the Wiener polarity index W_P .

To characterize extremal trees with diameter larger than 3, we introduce an operation on trees. Let $T \in \mathcal{T}(n)$ be a tree with $diam(T) = k$ where $k \geq 4$ is an integer. We suppose that $P_L(T) = v_0v_1v_2v_3v_4 \dots v_k$ is a longest path of T . Let $T \circledast v_0$ denote the tree obtained from T by deleting the edge v_0v_1 and adding a new edge v_0v_3 as shown in *Figure 1*.

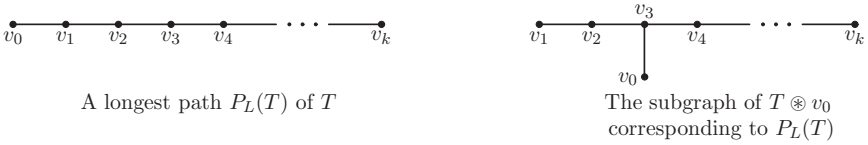


Figure 1. Maximization operation of a tree T with $\text{diam}(T) \geq 4$.

The above operation is called *maximization operation*. One can easily see that $\text{diam}(T \otimes v_0) \leq \text{diam}(T)$. To establish the main theorem of this subsection, we first prove the following lemma concerning the relation between $W_P(T)$ and $W_P(T \otimes v_0)$.

Lemma 2. *Let $T = (V, E)$ be a tree, and let $P_L(T) = v_0v_1 \dots v_k$ be a longest path of T , where $k \geq 4$ is an integer. Then*

- i) $W_P(T) < W_P(T \otimes v_0)$ if $\text{diam}(T) \geq 5$,
- ii) $W_P(T) \leq W_P(T \otimes v_0)$ if $\text{diam}(T) = 4$.

Proof. We only prove the first assertion, and suppose that T is a tree with $\text{diam}(T) \geq 5$. Let

$$D_3(T) := \{\{u, v\} \mid d_T(u, v) = 3, u, v \in V\}.$$

Obviously, $W_P(T) = |D_3(T)|$. Thus, to show the first assertion, it is sufficient to show that

$$|D_3(T)| < |D_3(T \otimes v_0)|.$$

In order to establish the above inequality, we introduce a notion concerning a subset of $D_3(T)$. Let

$$D_3(T, u) := \{\{u, v\} \mid \{u, v\} \in D_3(T)\}.$$

Obviously, $D_3(T, u) \subseteq D_3(T)$ for any vertex $u \in V$, and $|D_3(T)| = \frac{1}{2} \sum_{u \in V} |D_3(T, u)|$. One can verify that

$$|D_3(T, v_0)| < |D_3(T \otimes v_0, v_0)|, \quad |D_3(T, v_1)| = |D_3(T \otimes v_0, v_1)| - 1,$$

and

$$|D_3(T, v_3)| = |D_3(T \otimes v_0, v_3)| + 1.$$

We only give the explanation of the last equality. In fact, for the vertex v_3 , the pair $\{v_0, v_3\}$ is in $D_3(T, v_3)$ but not in $D_3(T \otimes v_0, v_3)$ and all other pairs in $D_3(T, v_3)$ are also in $D_3(T \otimes v_0, v_3)$. Hence we have $|D_3(T, v_3)| = |D_3(T \otimes v_0, v_3)| + 1$.

Furthermore, one can also verify that if $u \in V \setminus \{v_0, v_1, v_3\}$ then

$$|D_3(T, u)| \leq |D_3(T \otimes v_0, u)|.$$

Thus, $|D_3(T)| < |D_3(T \otimes v_0)|$ and then the first assertion holds.

Using a similar method, one can readily prove the second assertion. □

By the second assertion of Lemma 2, if T is a tree with $diam(T) \geq 4$ then we can construct another tree $T \otimes v_0$ by the maximization operation such that

$$W_P(T) \leq W_P(T \otimes v_0),$$

where v_0 is one end of a longest path in T . Since $W_P(T) \leq \lceil \frac{n-2}{2} \rceil \lfloor \frac{n-2}{2} \rfloor$ for any tree $T \in \mathcal{T}(n)$ with $diam(T) = 3$, we have the following result.

Theorem 2. *For any tree T of order n we have that $0 \leq W_P(T) \leq \lceil \frac{n-2}{2} \rceil \lfloor \frac{n-2}{2} \rfloor$.*

By the first assertion of Lemma 2, if T is a tree with $diam(T) \geq 5$ then we can construct another tree $T \otimes v_0$ by the maximization operation such that

$$W_P(T) < W_P(T \otimes v_0),$$

where v_0 is one end of a longest path in T . Moreover, by Theorem 2, one can readily see that if T is a tree with $|V(T)| = n$ and $W_P(T) = \lceil \frac{n-2}{2} \rceil \lfloor \frac{n-2}{2} \rfloor$ then $diam(T) \leq 4$. In the following, we characterize trees T of order n with $diam(T) = 4$ and $W_P(T) = \lceil \frac{n-2}{2} \rceil \lfloor \frac{n-2}{2} \rfloor$.

Lemma 3. *There exists a tree T of order n such that $diam(T) = 4$ and $W_P(T) = \lceil \frac{n-2}{2} \rceil \lfloor \frac{n-2}{2} \rfloor$.*

Proof. Let T be a tree with $|V(T)| = n$ and $diam(T) = 4$. It is not difficult to see that T can be represented by $m + 3$ integers (see Figure 2) $k_1, k_2, k_3, l_1, \dots, l_m$ satisfying that $k_i \geq 0$ ($i = 1, 2, 3$), $m \geq 0$, $l_j \geq 1$ when $m \geq 1$ and $1 \leq j \leq m$, and

$$k_1 + k_2 + k_3 + l_1 + \dots + l_m = n - 5 - m. \tag{3}$$

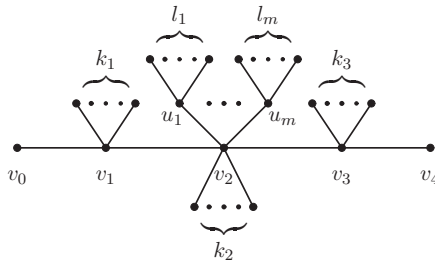


Figure 2. The structure of $T(k_1, k_2, k_3, l_1, \dots, l_m)$.

Clearly, the above representation is unique for a given tree $T \in \mathcal{T}(n)$ with diameter 4. In what follows, we use $T(k_1, k_2, k_3, l_1, \dots, l_m)$ to denote a tree which can be represented by integers $k_1, k_2, k_3, l_1, \dots, l_m$. By Equation (1), we have

$$\begin{aligned} W_P(T(k_1, k_2, k_3, l_1, \dots, l_m)) &= (m + k_2 + 1)(k_1 + 1) + (m + k_2 + 1)(k_3 + 1) \\ &\quad + (m + k_2 + 1)l_1 + \dots + (m + k_2 + 1)l_m \\ &= (m + k_2 + 1)(k_1 + k_3 + l_1 + \dots + l_m + 2). \end{aligned}$$

Using Equation (3), we have

$$W_P(T(k_1, k_2, k_3, l_1, \dots, l_m)) = (m + k_2 + 1)(n - 2 - (m + k_2 + 1)).$$

One can readily check that $n - 2 - (m + k_2 + 1) = \lfloor \frac{n-2}{2} \rfloor$ if $m + k_2 + 1 = \lceil \frac{n-2}{2} \rceil$, and $n - 2 - (m + k_2 + 1) = \lceil \frac{n-2}{2} \rceil$ if $m + k_2 + 1 = \lfloor \frac{n-2}{2} \rfloor$. Thus, for a tree $T(k_1, k_2, k_3, l_1, \dots, l_m)$ with $m + k_2 + 1 = \lceil \frac{n-2}{2} \rceil$ or $\lfloor \frac{n-2}{2} \rfloor$, we have

$$W_P(T(k_1, k_2, k_3, l_1, \dots, l_m)) = \left\lceil \frac{n-2}{2} \right\rceil \left\lfloor \frac{n-2}{2} \right\rfloor,$$

and then the lemma follows. □

Let

$$\mathcal{T}_4(n) := \{T(k_1, k_2, k_3, l_1, \dots, l_m) \in \mathcal{T}(n) \mid m + k_2 + 1 = \lceil \frac{n-2}{2} \rceil \text{ or } \lfloor \frac{n-2}{2} \rfloor\}.$$

Recall that all the trees in $\mathcal{T}_4(n)$ have a diameter equal to 4. By the proof of the above lemma, one can easily see that $W_P(T) = \lceil \frac{n-2}{2} \rceil \lfloor \frac{n-2}{2} \rfloor$ if $T \in \mathcal{T}_4(n)$. According to our analysis above, we can obtain the main result of this subsection.

Theorem 3. *Among all trees of order n , a tree T has the maximal Wiener polarity index $W_P(T) = \lceil \frac{n-2}{2} \rceil \lfloor \frac{n-2}{2} \rfloor$ if and only if T belongs to $\mathcal{T}_3(n) \cup \mathcal{T}_4(n)$.*

3 An Algorithm for Connected Graphs

Let G be a graph with ω components C_1, \dots, C_ω . Obviously,

$$W_P(G) = \sum_{i=1}^{\omega} W_P(C_i).$$

Thus, to calculate the Wiener polarity index for general graphs, it is sufficient to study how to calculate the index for connected graphs. In this section, we present an algorithm which computes the index $W_P(G)$ for any given connected graph G on n vertices in time $O(M(n))$, where $M(n)$ denotes the time necessary to multiply two $n \times n$ matrices of small integers (which is currently known to be $O(n^{2.376})$ [5]).

To any graph $G = (V, E)$ with the vertex set $V = \{v_1, \dots, v_n\}$ there corresponds an $n \times n$ matrix, called the *adjacency matrix* of G and denoted by $A(G)$ or A , in which $a_{ij} = 1$ if and only if $v_i v_j \in E$. We use $A^k = (a_{ij}^{(k)})_{n \times n}$ to denote the k -th repeated product of A where k is a positive integer. To establish our main result of this section, we first introduce some lemmas.

Lemma 4. *Let G be a connected graph, and let $A = (a_{ij})_{n \times n}$ be the adjacency matrix of G . If G has a path of length k between two vertices v_i and v_j , then $a_{ij}^{(k)} > 0$ where $(a_{ij}^{(k)})_{n \times n} = A^k$ and k is a positive integer.*

Lemma 5. *Let G be a connected graph, and let $A = (a_{ij})_{n \times n}$ be the adjacency matrix of G . If v_i and v_j are two vertices of G , and $a_{ij}^{(k)} > 0$ then $d_G(v_i, v_j) \leq k$ where $(a_{ij}^{(k)})_{n \times n} = A^k$ and k is a positive integer.*

The above two lemmas are well-known results. In fact, one can readily prove them by induction on k .

We use B to denote an $n \times n$ $(0, 1)$ -matrix, called the distance-2 matrix of G , in which $b_{ij} = 1$ ($i \neq j$) if and only if $a_{ij} = 1$ or $a_{ij}^{(2)} > 0$, and $b_{ii} = 0$. Furthermore, we use C to denote another $n \times n$ $(0, 1)$ -matrix, called the distance-3 matrix of G , in which $c_{ij} = 1$ ($i \neq j$) if and only if $b_{ij} = 1$ or $a_{ij}^{(3)} > 0$, and $c_{ii} = 0$. Using the above notations, we can characterize the distance between two vertices of a connected graph G by the distance-2 matrix and distance-3 matrix of G as follows.

Lemma 6. *Let G be a connected graph of order n , and let $B = (b_{ij})_{n \times n}$ and $C = (c_{ij})_{n \times n}$ be the distance-2 matrix and distance-3 matrix of G , respectively. If v_i and v_j are two distinct vertices of G , then*

- i) $b_{ij} = 1$ if and only if $d_G(v_i, v_j) \leq 2$,
- ii) $c_{ij} = 1$ if and only if $d_G(v_i, v_j) \leq 3$.

Proof. We only show the first assertion. If $d_G(v_i, v_j) = 2$ then $a_{ij}^{(2)} > 0$ by Lemma 4. Clearly, $a_{ij} = 1$ if $v_i v_j$ is an edge of G . Thus $b_{ij} = 1$ by the definition the distance-2 matrix. Conversely, if $b_{ij} = 1$ then $a_{ij} = 1$ or $a_{ij}^{(2)} > 0$ by the definition the distance-2 matrix. Thus $d_G(v_i, v_j) \leq 2$ by Lemma 5.

One can easily prove the second assertion by a similar manner. □

Using above lemmas, we can prove the main theorem in this section.

Theorem 4. *Let G be a connected graph of order n , and let B and C be the distance-2 matrix and distance-3 matrix of G , respectively. If $Z := C - B$ then*

$$W_P(G) = \sum_{i=1}^n \sum_{j>i} z_{ij},$$

where $(z_{ij})_{n \times n} = Z$.

Proof. Let v_i and v_j be two distinct vertices of G . By Lemma 6, $z_{ij} = 1$ if and only if $d_G(v_i, v_j) = 3$. Thus,

$$\sum_{i=1}^n \sum_{j>i} z_{ij} = \#\{\{u, v\} \mid d_G(u, v) = 3, u, v \in V\}.$$

Therefore, $W_P(G) = \sum_{i=1}^n \sum_{j>i} z_{ij}$ due to the definition of the Wiener polarity index. \square

According to the above theorem, we can design an algorithm to compute the Wiener polarity index $W_P(G)$ for any connected graph G represented by the adjacency matrix A of G . The algorithm is simply to compute A^2 and A^3 first and then to construct B and C as described before Lemma 6, and finally to compute $Z = C - B$. Then the Wiener polarity index of G is the sum of all the elements in the upper triangular part of Z . It is not difficult to see that the algorithm can be done in $O(M(n))$ time, where $M(n)$ denotes the time necessary to multiply two $n \times n$ matrices of small integers. Up to now, the complexity of the known fast matrix multiplication algorithm $M(n)$ by Coppersmith and Winograd [5] is $O(n^{2.376})$.

Acknowledgement: The authors would like to thank the referee for helpful comments and suggestions.

References

- [1] C.A. Barefoot, R.C. Entringer and L.A. Székely, Extremal values for ratios of distances in trees, *Discrete Math.* **80** (1997), 37–56.
- [2] D. Bonchev and O.E. Polansky, The Wiener number of graphs, I. General theory and change due to graph operations, *MATCH Commun. Math. Comput. Chem.* **21** (1986), 133–186.
- [3] D. Bonchev and N. Trinajstić, Information theory, distance matrix and molecular branching, *J. Chem. Phys.* **67**(1977), 4517–4533.
- [4] E.R. Canfield, R.W. Robinson and D.H. Rouvray, Determination of the Wiener molecular branching index for the general tree, *J. Comput. Chem.* **6** (1985), 598–609.
- [5] D. Coppersmith and S. Winograd, Matrix multiplication via arithmetic progression, *J. Symb. Comput.* **9**(3)(1990), 251–280.
- [6] P. Dankelmann, R.C. Entringer, Average distance, minimum degree, and spanning trees, *J. Graph Theory* **33**(1) (2000), 1–13.

- [7] A.A. Dobrynin, R.C. Entringer and I. Gutman, Wiener index of trees: theory and applications, *Acta Appl. Math.* **66** (2001), 211–249.
- [8] R.C. Entringer, D.E. Jackson and D.A. Snyder, Distance in graphs, *Czechoslovak Math. J.* **26** (1976), 283–296.
- [9] R.C. Entringer, A. Meir, J.W. Moon and L.A. Székely, On the Wiener index of trees from certain families, *Australas. J. Combin.* **10**(1994), 211–224.
- [10] I. Gutman, A new method for the calculation of the Wiener number of acyclic molecules, *J. Mol. Struct. (Theochem)* **285** (1993), 137–142.
- [11] I. Gutman, S.L. Lee, Y.L. Luo and Y.N. Yeh, Some recent results in the theory of Wiener index, *Indian J. Chem.* **32A** (1993), 137–142.
- [12] B. Mohar, Eigenvalues, diameter and mean distance in graphs, *Graph Combin.* **7** (1991), 53–64.
- [13] B. Mohar and T. Pisanski, How to compute the Wiener index of a graph, *J. Math. Chem.* **2** (1988), 267–277.
- [14] N. Trinajstić, Chemical graph theory, CRC Press: Boca Raton, FL, 1993.
- [15] H. Wiener, Structural determination of paraffin boiling points, *J. Amer. Chem. Soc.* **69** (1947), 17–20.