Wiener–Type Invariants of Trees and Their Relation

Ivan Gutman
Faculty of Science, University of Kragujevac, P. O. Box 60,
34000 Kragujevac, Yugoslavia;
e-mail: gutman@knez.uis.kg.ac.yu

Andrey A. Dobrynin
Sobolev Institute of Math., Russian Academy of Sciences,
Siberian Branch, Novosibirsk 630090, Russia
e-mail: dobr@math.nsc.ru

Sandi Klavžar
Department of Mathematics, PeF, University of Maribor
Koroška cesta 160, 2000 Maribor, Slovenia
e-mail: sandi.klavzar@uni-mb.si

Ljiljana Pavlović
Faculty of Science, University of Kragujevac, P. O. Box 60,
34000 Kragujevac, Yugoslavia;

July 15, 2002

Abstract

The distance $d(u, v|G)$ between the vertices u and v of a (connected) graph G is the length (= number of edges) of a shortest path connecting u and v. The Wiener number $W(G)$ of G is the sum of distances between all pairs of vertices of G. We consider a class of Wiener–type invariants $W_\lambda(G)$, defined as the sum of the terms $d(u, v|G)^\lambda$ over all pairs of vertices of G. Several special cases of $W_\lambda(G)$, namely the invariants for $\lambda = +1$ (the original Wiener number) as well as for $\lambda = -2, -1, +1/2, +2$ and $+3$, were previously studied in the chemical literature, and found applications as molecular structure descriptors. We modify the definition of $W_\lambda(G)$ so that it extends also to non-connected graphs and then deduce the identity $W_{\lambda+1}(T) = (n-1)\, W_\lambda(T) - \sum W_\lambda(T-e)$, valid for any n-vertex tree T, with the summation embracing all edges e of T.

1 Introduction

In this paper we are concerned with finite undirected graphs. The metric on these graphs is defined in the usual manner [1]: Let u and v be two vertices belonging to the same component of the graph G. The *distance* $d(u, v|G)$ between the vertices u and v is the length (= number of edges) of a shortest path connecting u and v. If $u = v$, then $d(u, v|G) = 0$. If u and v belong to different components of G, then the distance between them is not determined.

Let G be a graph with vertex set $V(G)$ and edge set $E(G)$, and let $|V(G)| = n$ and $|E(G)| = m$.

The *Wiener number* (or *Wiener index*) of a connected graph G is defined as [15]

$$W = W(G) = \sum_{\{u,v\} \subseteq V(G) \times V(G)} d(u, v|G) \ . \tag{1}$$

In words: the Wiener number is the sum of distances between all pairs of vertices of the respective graph. Therefore, $\binom{n}{2}^{-1} W(G)$ is just the average distance between the vertices of the graph G.

The graph invariant W was introduced in 1947 by Wiener [15], who used it for modeling the shape of organic molecules and for calculating several of their physico–chemical properties. Mathematical research on W started in 1976 [6] and since then this distance–based quantity was much studied; for details of the theory of the Wiener number and for an exhaustive list of references see the recent reviews [4, 5].

The definition (1) of the Wiener number requires that the graph G be connected. As a consequence, practically the entire research on W, done so far [4, 5], was restricted to connected graphs. Yet, this restriction can easily be overcome.

Denote by $d(G, k)$ the number of pairs of vertices of the graph G that are at distance k, and note that this quantity is well defined for both connected and disconnected graphs. In particular, $d(G, 0) = n$ and $d(G, 1) = m$. Now, evidently, the right–hand side of Eq. (1) can be rewritten as $\sum_{k \geq 1} k\, d(G, k)$, which hints towards the possibility to *define* the Wiener number of a graph G as

$$W = W(G) = \sum_{k \geq 1} k\, d(G, k) \ . \tag{2}$$

If G is a connected graph, then Eq. (2) reduces to Eq. (1). If G is disconnected, then the right–hand side of (1) is ill-determined, which is not the case with the right–hand side of Eq. (2).

From (2) follows that if G is a graph consisting of components G_1, G_2, $\ldots$, G_p , then

$$W(G) = W(G_1) + W(G_2) + \cdots + W(G_p) \ . \tag{3}$$

An immediate generalization of the Wiener number is

$$W_\lambda = W_\lambda(G) = \sum_{k \geq 1} d(G,k)\, k^\lambda \tag{4}$$

where λ is some real (or complex) number. For connected graphs formula (4) is tantamount to

$$W_\lambda = W_\lambda(G) = \sum_{\{u,v\} \subseteq V(G) \times V(G)} d(u,v|G)^\lambda \ .$$

In an explicit form the Wiener–type graph invariant W_λ was first put forward in the works [7] and [8]. However, various of its special cases have independently been considered in the chemical literature, where they found considerable applications. Thus W_{-2} and W_{-1} , named Harary index and reciprocal Wiener index, were introduced in the papers [11] and [3], respectively, and eventually studied in numerous subsequent publications. The case $\lambda = \frac{1}{2}$ was analyzed in the article [16]. The so-called "hyper–Wiener index" [12] was shown [10] to be equal to $\frac{1}{2}\, W_2 + \frac{1}{2}\, W_1$. The so-called "Tratch–Stankevich–Zefirov index" [13] was shown [9] to be equal to $\frac{1}{6}\, W_3 + \frac{1}{2}\, W_2 + \frac{1}{3}\, W_1$. (Recall that the hyper–Wiener and Tratch–Stankevich–Zefirov indices were originally defined in terms completely different from the presently considered Wiener–type invariants; for details see [12, 13].) More details on the chemical applications and interconnections of various distance–based graph invariants are found in the review [2] and the book [14].

2 Two identities for distances in trees

A tree is a connected acyclic graph. Any two vertices of a tree are connected by a unique path; the number of edges of this unique path is the distance between the respective two vertices.

Let T be a tree on n vertices and let e be one of its edges. The subgraph $T - e$ is obtained by deleting from T the edge e . Thus, $V(T-e) = V(T)$.

The subgraph $T-e$ is disconnected, possessing two components. Denote them by $T_1(e)$ and $T_2(e)$, and let the number of their vertices be $n_1(e)$ and $n_2(e)$, respectively, $n_1(e) + n_2(e) = |V(T-e)| = n$.

Lemma 1. Let T be a tree on n vertices. Then

$$(n-1-k)\, d(T,k) = \sum_{e \in E(T)} d(T-e,k) \tag{5}$$

holds for all $k = 0, 1, 2, \ldots$.

Proof. Consider the difference $d(T,k) - d(T-e,k)$. In view of the uniqueness of the path connecting any given pair of vertices of a tree, any two vertices of T , connected by a path that contains the edge e , belong to different components of $T-e$. Consequently, the difference $d(T,k) - d(T-e,k)$ counts the pairs of vertices of T that are at distance k and whose connecting path contains the edge e . By summing this difference over all edges of T we will count any pair of vertices of T at distance k . Furthermore, every such pair will be counted exactly k times, because there are exactly k edges in the path connecting them. Hence,

$$\sum_{e \in E(T)} [d(T,k) - d(T-e,k)] = k\, d(T,k) \ .$$

Formula (5) follows now by taking into account that T has $n-1$ edges. □

Lemma 2 is deduced in a fully analogous manner. Here u stands for a vertex of the tree T and $T-u$ is the subgraph obtained by deleting u (together with its incident edges) from T .

Lemma 2. Let T be a tree on n vertices. Then

$$(n-1-k)\, d(T,k) = \sum_{u \in V(T)} d(T-u,k)$$

holds for all $k = 0, 1, 2, \ldots$.

Theorem 3. Let T be a tree on n vertices. Let λ be a real (or complex) number. Then

$$W_{\lambda+1}(T) = (n-1)\, W_\lambda(T) - \sum_{e \in E(T)} W_\lambda(T-e) \ . \tag{6}$$

Proof. By multiplying Eq. (5) by k^λ one obtains

$$d(T,k)\, k^{\lambda+1} = (n-1)\, d(T,k)\, k^\lambda - \sum_{e \in E(T)} d(T-e,k)\, k^\lambda$$

which summed over all $k \geq 1$ and in view of Eq. (4) yields (6). □

In an analogous manner, from Lemma 2 follows:

Theorem 4. Let T be a tree on n vertices. Let λ be a real (or complex) number. Then

$$W_{\lambda+1}(T) = (n-1)\,W_\lambda(T) - \sum_{u \in V(T)} W_\lambda(T-u) \ .$$

Remark. The identity (5) can be rewritten as

$$(m-k)\,d(F,k) = \sum_{e \in E(F)} d(F-e,k)\,,$$

in which case it holds for any forest F (= acyclic graph, not necessarily connected), with $m \leq n-1$ edges. Analogously, relation (6) then becomes $W_{\lambda+1}(F) = m\,W_\lambda(F) - \sum\limits_{e \in E(F)} W_\lambda(F-e)\,.$

3 Applications of relation (6)

First of all, using Eq. (3) and the notation defined above, Eq. (6) can be rewritten as

$$W_{\lambda+1}(T) = (n-1)\,W_\lambda(T) - \sum_{e \in E(T)} \left[W_\lambda(T_1(e)) + W_\lambda(T_2(e))\right] \ . \tag{7}$$

Note that all graphs occurring in formula (7) are connected.

For any connected n-vertex graph G , $W_0(G) = \binom{n}{2}$.

Formulas (6) holds for any value of λ . By setting $\lambda = 0$ and by taking into account that $n_1(e) + n_2(e) = n$, we obtain:

$$\begin{aligned}
W_1(T) &= (n-1)\,W_0(T) - \sum_e \left[W_0(T_1(e)) + W_0(T_2(e))\right] \\
&= (n-1)\binom{n}{2} - \sum_e \left[\binom{n_1(e)}{2} + \binom{n_2(e)}{2}\right] \\
&= \frac{1}{2}\,n\,(n-1)^2 - \frac{1}{2}\sum_e \left[n_1(e)^2 + n_2(e)^2 - (n_1(e) + n_2(e))\right] \\
&= \frac{1}{2}\,n\,(n-1)^2 - \frac{1}{2}\sum_e \left[n^2 - n - 2\,n_1(e)\,n_2(e)\right]
\end{aligned}$$

$$= \quad \frac{1}{2}\,n\,(n-1)^2 - \frac{1}{2}\,(n-1)(n^2-n) + \sum_e n_1(e)\,n_2(e)$$

which finally yields

$$W(T) = \sum_e n_1(e)\,n_2(e) \tag{8}$$

a result first reported by Wiener himself [15]. Thus, the relation (6) may be viewed as a generalization of the Wiener formula (8).

The n-vertex tree possessing a maximum number ($= n-1$) vertices of degree 1 is called the star (S_n). The n-vertex tree possessing a minimum number ($= 2$) vertices of degree 1 is the path graph (P_n). In the set of all n-vertex trees, S_n and P_n usually have extremal properties. It has been shown elsewhere [7] that for T_n being any n-vertex tree different from S_n and P_n, and for any $\lambda > 0$,

$$W_\lambda(S_n) \;<\; W_\lambda(T_n) \;<\; W_\lambda(P_n)\ .$$

If $\lambda < 0$, then in the above inequalities "less than" has to be exchanged into "greater than".

Because $d(S_n, k) = 0$ for $k \geq 3$, one directly gets

$$W_\lambda(S_n) = n - 1 + \binom{n-1}{2}\,2^\lambda\ .$$

The calculation of the Wiener–type invariants of P_n is less easy.

By means of formulas (6) or (7) the Wiener–type invariants of a tree can be computed recursively. This is especially efficient if the respective tree possesses some structural regularity. For instance, for P_n, formula (7) reduces to

$$W_{\lambda+1}(P_n) = (n-1)\,W_\lambda(P_n) - 2\sum_{i=1}^{n-1} W_\lambda(P_i)\ . \tag{9}$$

We start with $\lambda = 0$ and the obvious relation $W_0(P_n) = \binom{n}{2}$. Then, by applying (9),

$$W_1(P_n) = \binom{n}{2} - 2\sum_{i=1}^{n-1}\binom{i}{2} = \binom{n+1}{3}\ . \tag{10}$$

For $\lambda = 1, 2, \ldots, 5$ analogous calculations yield

$$W_2(P_n) = \frac{n}{2}\binom{n+1}{3} \qquad\qquad W_3(P_n) = \frac{3n^2-2}{10}\binom{n+1}{3} \tag{11}$$

$$W_4(P_n) = \frac{n(2n^2-3)}{10}\binom{n+1}{3} \qquad W_5(P_n) = \frac{(n^2-2)(2n^2-1)}{14}\binom{n+1}{3}$$

$$W_6(P_n) = \frac{n(n^2-2)(3n^2-5)}{28}\binom{n+1}{3} \, .$$

By induction it can be shown that for λ being a positive integer, $W_\lambda(P_n)$ has the following properties:
— $W_\lambda(P_n)$ is a polynomial in the variable λ, of degree $n+2$;
— if n is even/odd, the coeficcients at odd/even terms are 0;
— the nonzero coefficients alternate in sign.

Using expressions (10) and (11) one can immediately check that

$$\frac{1}{2}\,W_2(P_n) + \frac{1}{2}\,W_1(P_n) = \binom{n+2}{4}$$

and

$$\frac{1}{6}\,W_3(P_n) + \frac{1}{2}\,W_2(P_n) + \frac{1}{3}\,W_1(P_n) = \binom{n+3}{5}$$

Thus we arrive at the remarkable result that the Wiener number [15], the hyper–Wiener index [12] and the Tratch–Stankevich–Zefirov index [13] of the n-vertex path graph are given by

$$\binom{n+1}{3} \quad , \quad \binom{n+2}{4} \quad , \quad \binom{n+3}{5}$$

respectively.

Acknowledgements

Parts of this research were supported by the Ministry of Science and Technology of Serbia, under the grant 1389, by the Russian Foundation for Basic Research (project code 02-01-00039), and by the Ministry of Education, Science and Sport of Slovenia under the grant 0101–504.

References

[1] F. Buckley and F. Harary, *Distance in Graphs*, Addison–Wesley, Redwood, 1990.

[2] M. V. Diudea and I. Gutman, Wiener–type topological indices, *Croat. Chem. Acta* **71** (1998) 21–51.

[3] M. V. Diudea, O. Ivanciuc, S. Nikolić and N. Trinajstić, Matrices of reciprocal distance, polynomials and derived numbers, *MATCH - Commun. Math. Comput. Chem.* **35** (1997) 41–64.

[4] A. A. Dobrynin, R. Entringer and I. Gutman, Wiener index of trees: theory and applications, *Acta Appl. Math.* **66** (2001) 211–249.

[5] A. A. Dobrynin, I. Gutman, S. Klavžar and P. Žigert, Wiener index of hexagonal systems, *Acta Appl. Math.* **72** (2002) 247–294.

[6] R. C. Entringer, D. E. Jackson and D. A. Snyder, Distance in graphs, *Czech. Math. J.* **26** (1976) 283–296.

[7] I. Gutman, A property of the Wiener number and its modifications, *Indian J. Chem.* **36A** (1997) 128–132.

[8] I. Gutman, D. Vidović and L. Popović, On graph representation of organic molecules - Cayley's plerograms vs. his kenograms, *J. Chem. Soc. Faraday Trans.* **94** (1998) 857–860.

[9] D. J. Klein and I. Gutman, Wiener–number–related sequences *J. Chem. Inf. Comput. Sci.* **39** (1999) 534–536.

[10] D. J. Klein, I. Lukovits and I. Gutman, On the definition of the hyper–Wiener index for cycle–containing structures, *J. Chem. Inf. Comput. Sci.* **35** (1995) 50–52.

[11] D. Plavšić, S. Nikolić, N. Trinajstić and Z. Mihalić, On the Harary index for the characterization of chemical graphs, *J. Math. Chem.* **12** (1993) 235–250.

[12] M. Randić, Novel molecular descriptor for structure–property studies, *Chem. Phys. Lett.* **211** (1993) 478–483.

[13] S. S. Tratch, M. I. Stankevich and N. S. Zefirov, Combinatorial models and algorithms in chemistry. The expanded Wiener number – A novel topological index, *J. Comput. Chem.* **11** (1990) 899–908.

[14] R. Todeschini and V. Consonni, *Handbook of Molecular Descriptors*, Wiley–VCH, Weinheim, 2000.

[15] H. Wiener, Structural determination of paraffin boiling points, *J. Am. Chem. Soc.* **69** (1947) 17–20.

[16] H. Y. Zhu, D. J. Klein and I. Lukovits, Extensions of the Wiener number, *J. Chem. Inf. Comput. Sci.* **36** (1996) 420–428.