Kragujevac Journal of Mathematics
Volume 35 Number 1 (2011), Pages 165–172.

A NEW PROOF OF THE SZEGED–WIENER THEOREM

H. KHODASHENAS [1], M. J. NADJAFI–ARANI [1], A. R. ASHRAFI [1,2], AND I. GUTMAN [3]

Abstract. The Wiener index $W(G)$ is the sum of distances between all pairs of vertices of a connected graph G. For an edge e of G, connecting the vertices u and v, the set of vertices lying closer to u than to v is denoted by $N_e(u)$. The Szeged index, $Sz(G)$, is the sum of products $|N_u(e)| \times |N_v(e)|$ over all edges of G. A block graph is a graph whose every block is a clique. The Szeged–Wiener theorem states that $Sz(G) = W(G)$ holds if and only if G is a block graph. A new proof of this theorem if offered, by means of which some properties of block graphs could be established.

1. Introduction

Throughout this article G stands for a simple connected graph with vertex and edge sets $V(G)$ and $E(G)$, respectively. The distance between the vertices u and v of the graph G (= the number of edges in a shortest path connecting u and v) [3] will be denoted by $d(u, v|G)$.

There is a large number of distance–based graph invariants that have attracted the attention of, and that have been extensively studied by, mathematicians. Of these, the Wiener index $W(G)$ is the oldest [14], defined as the sum of distances between all pairs of vertices of G:

$$W(G) = \sum_{\{x,y\} \subseteq V(G) \times V(G)} d(x, y|G).$$

The Wiener index has noteworthy applications in chemistry and the interested readers are referred to the reviews [4, 5] and references therein for details.

Key words and phrases. Szeged index, Wiener index, Block graph.
2010 *Mathematics Subject Classification.* Primary: 05C12, Secondary: 05C05 .
Received: December 14, 2010.
Revised: January 29, 2011.

Suppose that G is a connected graph and $e = uv \in E(G)$. Define:

$$\begin{aligned} N_u(e) &= \{x \in V(G) \mid d(x,u|G) < d(x,v|G)\} \\ N_v(e) &= \{x \in V(G) \mid d(x,v|G) < d(x,u|G)\} \\ N_0(e) &= \{x \in V(G) \mid d(x,u|G) = d(x,v|G)\}. \end{aligned}$$

Define $n_u(e)$ to be the number of vertices of G lying closer to u than to v, and define $n_v(e)$ analogously. Thus $n_u(e) = |N_u(e)|$ and $n_v(e) = |N_v(e)|$. Notice that vertices equidistant from both ends of the edge $e = uv$, i.e., the vertices belonging to $N_0(e)$, are not counted in $n_u(e)$ and $n_v(e)$.

The Szeged index of G is defined as [8]

$$Sz(G) = \sum_{e=uv \in E(G)} n_u(e)\, n_v(e).$$

Details of the theory of this distance–based graph invariant can be found in the survey [9] as well as in the recent articles [15–17].

Lukovits [13] introduced an all-path version of the Wiener index, denoted by $P(G)$. To explain, we assume that $V(G) = \{1, 2, \ldots, n\}$. Then

$$P(G) = \sum_{i<j} \sum_{P \in \pi_{i,j}} \ell(P)$$

where $\ell(P)$ denotes the length of the path P, i.e., the number of edges in P, and where $\pi_{i,j}$ is the set of all path connecting the vertices i and j. Thus the summations in the above formula embrace all paths contained in G.

In [13] some mathematical properties of $P(G)$ were established, in particular its extremal values. In the next section we present a "path–edge" matrix aimed at studying the Wiener and Szeged indices of graphs, simultaneously. This matrix is defined in a similar way as the "all-path" index of Lukovits.

Throughout this paper our notation is standard and taken mainly from the standard textbooks of graph theory. Thus, K_n, P_n, and C_n denote the complete graph, path, and cycle on n vertices, respectively.

2. Preliminaries

The block graphs are natural generalization of trees. They are the connected graphs in which every block (i.e., every maximal 2-connected subgraph) is a clique. Of the several known characterizations of block graphs [10] we mention the following:

Lemma 2.1. *Let G be a connected graph. The following conditions are equivalent:*

(a) *G is a block graph.*

(b) *For every four vertices u, v, x, y of G, the greatest two among $d(u, v|G) + d(x, y|G)$, $d(u, x|G) + d(v, y|G)$, and $d(u, y|G) + d(v, x|G)$ are always mutually equal (the so-called "four–point condition")* [11].

(c) *G does not have induced subgraphs isomorphic to $K_4 - e$ (the "diamond graph") or C_n, $n \geq 4$* [1].

In [6], Dobrynin and one of the present authors studied the structure of a connected graph G with the property that $Sz(G) = W(G)$. They conjectured that $Sz(G) = W(G)$ if and only if G is a block graph. A year later, the conjecture was proved by the same authors [7]. In what follows we refer to it as the *Szeged–Wiener theorem*. Quite recently, apparently unaware of the works [6, 7], Behtoei et al. [2] presented another proof of the Szeged–Wiener theorem. In this paper, a third proof of this result will be communicated, as well as a new characterization of block graphs.

Let G be a connected graph. A set $Y = \left\{P_1, P_2, \ldots, P_{\binom{n}{2}}\right\}$ of shortest paths in G, such that for every pair of vertices $a, b \in V(G)$, $a \neq b$, there exists a unique path $P \in Y$ connecting vertices a and b, is called a *complete set of shortest paths* of G (CSSP for short). In what follows, $P_G(u, v)$ denotes the set of all shortest paths connecting vertices u and v of G and CSSP(G) denotes the set of all CSSP's of G.

Define the matrix $A_Y = [a_{ij}]$, as follows:

$$a_{ij} = \begin{cases} 1 & e_j \in E(P_i) \\ 0 & e_j \notin E(P_i) \end{cases} .$$

Clearly, if P_i is a path connecting the vertices x and y then $d(x, y|G)$ is the number of non-zero entries in the i-th row of A_Y. Thus the sum of entries of the matrix A_Y is equal to the Wiener index of G.

Lemma 2.2. *Let $e = uv \in E(G)$ and a and b be arbitrary vertices of G. If there exists a path $P \in P_G(a, b)$, such that $e \in E(P)$, then one of the following is satisfied:*

(i) *$a \in N_u(e)$ and $b \in N_v(e)$,*

(ii) *$a \in N_v(e)$ and $b \in N_u(e)$.*

Proof. Suppose that P is a shortest path containing the edge $e = uv$. Traverse the path P from the source vertex a to the destination vertex b. If we traverse the vertex

u before v then $d(a,v|G) = d(a,u|G) + d(u,v|G)$. This implies that $a \in N_u(e)$ and $b \in N_v(e)$, proving claim (i). If the vertex v is before u then similarly $a \in N_v(e)$ and $b \in N_u(e)$, as desired. □

The converse of Lemma 2.2 is not generally valid. To see this, it is enough to consider the case $G \cong C_n$ for $n \geq 4$.

In what follows, by $P_G(e)$ we denote the set of all shortest paths through the edge e.

Corollary 2.1. *For each edge $e = uv$ of a connected graph G, $|P_G(e)| \leq n_u(e)\, n_v(e)$.*

Proof. Apply Lemma 2.2. □

3. Main results

Suppose that G is a connected graph, $Y \in CSSP(G)$ and $A_Y = [a_{ij}]$. The sum of entries of the j-th column of A_Y is the number of shortest paths containing e_j. Thus, for each j, $1 \leq j \leq |E(G)|$,

$$\sum_i a_{ij} \leq |P_G(e_j)|$$

and therefore

$$W(G) = \sum_j \sum_i a_{ij} \leq \sum_j n_u(e_j)\, n_v(e_j) = Sz(G).$$

This presents a new proof of the following result of Klavžar et al. [12]:

Theorem 3.1. *For every connected graph G, $W(G) \leq Sz(G)$.*

Theorem 3.2. *Let G be a graph containing a non-complete block. Then the following is satisfied:*

(i) *G has an induced subgraph isomorphic either to $K_4 - e$ or to C_n, $n \geq 4$.*

(ii) *If G does not have an induced subgraph isomorphic to $K_4 - e$, then in the smallest induced cycle C_n, $n \geq 4$, the following condition is satisfied:*

$$\forall x, y \in V(C_n):\ d(x,y|C_n) = d(x,y|G).$$

Proof. The statement of Theorem 3.2 is a direct consequence of Lemma 2.1. In order that this paper be self–contained, we nevertheless provide its proof.

(i) Suppose that B is a non-complete block graph and a and b are its two non-adjacent vertices. Choose C to be the smallest cycle of B containing the vertices a and b. Then C contains two paths $P_1:\ a = x_0, x_1, \ldots, x_n = b$ and $P_2:\ a =$

$y_0, y_1, \ldots, y_m = b$, $m, n \geq 2$, such that $V(P_1) \cap V(P_2) = \{a, b\}$. Since C has the minimum size among the cycles of B containing a and b, a is not adjacent to x_i's and y_j's, $1 < i \leq n$ and $1 < j \leq m$. Suppose that the induced subgraph of G generated by $V(C)$ does not have an induced cycle C_n, $n \geq 4$. We claim that G has an induced subgraph isomorphic to $K_4 - e$. If $\ell(P_1) = \ell(P_2) = 2$, then C has size 4 and since C is not an induced cycle, x_1 is adjacent to y_1. But a and b are not adjacent, so we find an induced subgraph isomorphic to $K_4 - e$, as desired. Therefore, without loss of generality we can assume that $\ell(P_1) > 2$. Since G does not have an induced cycle of size ≥ 4, x_1 is again adjacent to y_1. On the other hand, by assumption x_1 is adjacent to y_2 or y_1 is adjacent to x_2. In each case, we will find an induced subgraph isomorphic to $K_4 - e$, which completes our argument.

(ii) Suppose that C is a smallest induced cycle such that for two vertices $x, y \in V(C)$, $d(x, y|G) < d(x, y|C)$. Choose Q to be a shortest path connecting x and y in G. Using C and Q one can obtain a new cycle C' of size at least four, smaller than C. By our assumption, C' is not an induced cycle. Applying an argument similar to case (i), we obtain an induced subgraph isomorphic to $K_4 - e$, a contradiction. □

Theorem 3.3. *Let G be a connected graph. Then $Sz(G) > W(G)$ holds if and only if G has an induced subgraph isomorphic either to a cycle of size ≥ 4 or to $K_4 - e$.*

Proof. ($\Rightarrow$) Suppose that $Sz(G) > W(G)$, $Y \in CSSP(P)$ and $A_Y = [a_{ij}]$. Then there exists some j, such that $\sum_i a_{ij} < n_u(e_j)\, n_v(e_j)$, where $e_j = uv$. This means that we can choose $a \in N_u(e_j)$, $b \in N_v(e_j)$ such that $e_j \notin P_{(a,b)}$, where $P_{(a,b)}$ is the unique path of Y connecting a and b. We consider three separate cases as follows:

Case 1. $a = u, b \neq v$. Suppose that Q is a shortest path connecting b and v. Let x be the first common vertex of $P_{(a,b)}$ and Q in traversing from v to b. Thus $x \in N_v(e_j)$. Since $e_j = uv \notin P_{(a,b)}$, $x \neq v$. So, $d(x, v|G) \geq 1$, $d(x, u|G) \geq 2$ and the size of the cycle C containing x, u, and v is at least 4. Thus the block of G containing this cycle is not complete. Then by Theorem 3.2, G has an induced subgraph isomorphic to $K_4 - e$ or a cycle C_n, $n \geq 4$.

Case 2. $a \neq u$, $b = v$. It is enough to apply a similar argument as that given in the Case 1.

Case 3. $a \neq u$, $b \neq v$. Let Q_1 and Q_2 be the shortest paths connecting a, u and b, v, respectively. Suppose that $x \in V(Q_1) \cap V(Q_2)$ and $d(x, u|G) \leq d(x, v|G)$. Then $d(b, u|G) \leq d(b, x|G) + d(x, u|G) \leq d(b, x|G) + d(x, v|G) = d(b, v|G)$. Thus

$b \notin N_v(e_j)$, a contradiction. If $x \in V(Q_1) \cap V(Q_2)$ and $d(x,u|G) > d(x,v|G)$ then $d(a,v|G) \leq d(a,x|G)+d(x,v|G) < d(a,x|G)+d(x,u|G) = d(a,u|G)$ and so $a \in N_v(e_j)$ and we arrive at another contradiction.

Therefore, $V(Q_1) \cap V(Q_2) = \emptyset$. Suppose that x is the last common vertex of $P_{(a,b)}$ and Q_1 and y is the first common vertex of $P_{(a,b)}$ and Q_2 when traversing the path $P_{(a,b)}$ from a to b. By our assumption, $d(x,y|G) \geq 1$. If $x = u$, then $v \notin V(P_{(a,b)})$ whereas if $y = v$, then $x \neq u$. In each case, a similar argument as in Cases 1 or 2, shows that G contains an induced subgraph isomorphic to either $K_4 - e$ or to C_n, $n \geq 4$. Therefore, we may assume that $x \neq u$ and $y \neq v$. Consider the cycle C containing x, u, v, and y. Since the size of C is at least 4, and x, v are not adjacent, the block B containing C is not complete and by Theorem 3.2, G has an induced subgraph isomorphic to $K_4 - e$ or to C_n, $n \geq 4$.

($\Leftarrow$) $Y \in CSSP(G)$. We first assume that G has an induced subgraph H isomorphic either to a cycle of size ≥ 4 or to $K_4 - e$. It is enough to show that there exists a j such that $\sum_i a_{ij} < n_u(e_j)\, n_v(e_j)$.

We first assume that $H \cong K_4 - e$. Suppose that $V(H) = \{v_1, v_2, v_3, v_4\}$ such that v_1 and v_3 are not adjacent. Without loss of generality, we can assume that $P_r : v_1v_2v_3$ is an element of Y connecting v_1 and v_3. Suppose that $e_j = v_1v_4$. Then $v_1 \in N_{v_1}(e_j)$ and $v_3 \in N_{v_4}(e_j)$. Thus $a_{rj} = 0$ and so $\sum_i a_{ij} < n_{v_1}(e_j)\, n_{v_4}(e_j)$, as desired.

If G does not have an induced subgraph isomorphic to K_4-e, then G has an induced cycle C_n, $n \geq 4$. Let $C : v_1, v_2, \ldots, v_{n+1} = v_1$ be an induced cycle of minimum size. Then by Theorem 3.2 (ii), for each vertex $x, y \in C$, $d(x,y|C) = d(x,y|G)$.

We separately consider two cases, namely when n is odd and n is even.

If n is odd, then we assume that $t = (n+1)/2$ and $e_j = v_1v_2$. By Lemma 2.2, $v_t \in N_{v_2}(e_j)$ and $v_{t+2} \in N_{v_1}(e_j)$. Since C has a minimum size, $d(v_t, v_{t+2}) = 2$ and $\ell(v_t \cdots v_2v_1 \cdots v_{t+2})$ has minimum length 3. Thus e_j is outside the shortest path $P_r : v_tv_{t+1}v_{t+2}$. Therefore, $a_{rj} = 0$ and $\sum_i a_{ij} < n_{v_1}(e_j)n_{v_2}(e_j)$, as desired.

If n is even, then by choosing the edge v_tv_{t+1}, $t = n/2+1$ and by a similar argument as above, we see that $v_t \in N_{v_2}(e_j)$ and $v_{t+1} \in N_{v_1}(e_j)$. So, $\sum_i a_{ij} < n_{v_1}(e_j)\, n_{v_4}(e_j)$, which completes our argument. ☐

Corollary 3.1 ([6, Corollary 3]). *If G is a connected non-acyclic bipartite graph, then $W(G) < Sz(G)$.*

Corollary 3.2. *For any (connected) graph G, the following conditions are equivalent:*

(a) $W(G) = Sz(G)$.

(b) *G does not have induced subgraphs isomorphic to* $K_4 - e$ *or* C_n, $n \geq 4$.

(c) *G is a block graph.*

Proof. Apply Theorems 3.1–3.3. □

Corollary 3.3 (Wiener [14]). *For every tree* T*, the equality* $W(T) = Sz(T)$ *holds.*

Acknowledgement: The work of A. R. A. was supported in part by a grant from IPM (No. 89050111). I. G. thanks the Serbian Ministry of Science, for support through Grant no. 144015G.

References

[1] H. J. Bandelt and H. M. Mulder, *Distance–hereditary graphs*, J. Combin. Theory **B 41** (1986), 182–208.

[2] A. Behtoei, M. Jannesari and B. Taeri, *A characterization of block graphs*, Discrete Appl. Math. **158** (2010), 219–221.

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

[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] A. Dobrynin and I. Gutman, *On a graph invariant related to the sum of all distances in a graph*, Publ. Inst. Math. (Beograd) **56** (1994), 18–22.

[7] A. Dobrynin and I. Gutman, *Solving a problem connected with distances in graphs*, Graph Theory Notes New York **28** (1995), 21–23.

[8] I. Gutman, *A formula for the Wiener number of trees and its extension to graphs containing cycles*, Graph Theory Notes New York **27** (1994), 9–15.

[9] I. Gutman and A. A. Dobrynin, *The Szeged index – a success story*, Graph Theory Notes New York **34** (1998), 37–44.

[10] F. Harary, *A characterization of block graphs*, Canad. Math. Bull. **6** (1963), 1–6.

[11] E. Howorka, *On metric properties of certain clique graphs*, J. Combin. Theory **B 27** (1979), 67–74.

[12] S. Klavžar, A. Rajapakse and I. Gutman, *The Szeged and the Wiener index of graphs*, Appl. Math. Lett. **9** (1996), 45–49.

[13] I. Lukovits, *An all-path version of the Wiener index*, J. Chem. Inf. Comput. Sci. **38** (1998), 125–129.

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

[15] S. Yousefi, H. Yousefi–Azari, A. R. Ashrafi and M. H. Khalifeh, *Computing Wiener and Szeged indices of a polyhex Nanotorus*, J. Sci. Univ. Tehran **33** (2008), 7–11.

[16] H. Yousefi–Azari, A. R. Ashrafi and N. Sedigh, *On the Szeged index of some benzenoid graphs applicable in nanostructures*, Ars Combin. **90** (2009), 55–64.

[17] H. Yousefi–Azari, B. Manoochehrian and A. R. Ashrafi, *PI and Szeged indices of some benzenoid graphs related to nanostructures*, Ars Combin. **84** (2007), 255–267.

[1] Department of Mathematics, Statistics and Computer Science,
Faculty of Science,
University of Kashan,
Kashan 87317–51167,
I. R. Iran
E-mail address: khodashenas@kashanu.ac.ir, nadjafi-arani@grad.kashanu.ir

[2] School of Mathematics,
Institute for Research in Fundamental Sciences (IPM),
P. O. Box 19395-5746,
Tehran,
I. R. Iran
E-mail address: ashrafi@kashanu.ac.ir

[3] Faculty of Science,
University of Kragujevac,
P. O. Box 60, 34000 Kragujevac,
Serbia
E-mail address: gutman@kg.ac.rs