ON A RELATION BETWEEN SZEGED AND WIENER INDICES OF BIPARTITE GRAPHS

L. CHEN, X. LI∗, M. LIU AND I. GUTMAN

Communicated by Alireza Abdollahi

Abstract. Hansen et. al., using the AutoGraphiX software package, conjectured that the Szeged index \( Sz(G) \) and the Wiener index \( W(G) \) of a connected bipartite graph \( G \) with \( n \geq 4 \) vertices and \( m \geq n \) edges, obeys the relation \( Sz(G) - W(G) \geq 4n - 8 \). Moreover, this bound would be the best possible. This paper offers a proof to this conjecture.

1. Introduction

All graphs considered in this paper are finite, undirected and simple. We refer the readers to [3] for terminology and notation. Let \( G \) be a connected graph with vertex set \( V(G) \) and edge set \( E(G) \). For \( u, v \in V(G) \), \( d(u, v) \) denotes the distance between \( u \) and \( v \). If the graph \( G \) is connected, then its Wiener index is defined as

\[
W(G) = \sum_{\{u,v\} \subseteq V(G)} d(u, v) .
\]

This topological index has been extensively studied in the mathematical literature; see, e.g., [4, 9, 10, 6]. Let \( e = uv \) be an edge of \( G \). Define three sets as follows:

\[
\begin{align*}
N_u(e) &= \{ w \in V(G) : d(u, w) < d(v, w) \} \\
N_v(e) &= \{ w \in V(G) : d(v, w) < d(u, w) \} \\
N_0(e) &= \{ w \in V(G) : d(u, w) = d(v, w) \} .
\end{align*}
\]
Thus, \( \{N_u(e), N_v(e), N_0(e)\} \) is a partition of the vertex set of \( G \) with regard to \( e \in E(G) \). The number of elements of \( N_u(e) \), \( N_v(e) \), and \( N_0(e) \) will be denoted by \( n_u(e) \), \( n_v(e) \), and \( n_0(e) \), respectively. Evidently, if \( n \) is the number of vertices of the graph \( G \), then \( n_u(e) + n_v(e) + n_0(e) = n \).

If \( G \) is bipartite, then the equality \( n_0(e) = 0 \) holds for all \( e \in E(G) \). Therefore, for any edge \( e \) of a bipartite graph, \( n_u(e) + n_v(e) = n \).

A long time known property of the Wiener index is the formula \([4, 11, 20]\):

\[
W(G) = \sum_{e = uv \in E(G)} n_u(e) n_v(e)
\]

which is applicable for trees. Motivated by the above formula, one of the present authors \([7]\) introduced a graph invariant, named as the \textit{Szeged index}, defined by

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

where \( G \) is any graph, not necessarily connected. Evidently, the Szeged index is defined as a proper extension of the formula (1.1) for the Wiener index of trees.

Details of the theory of the Szeged index can be found in \([8]\) and in the recent papers \([1, 13, 2, 5, 13, 14, 15, 16, 17, 21]\).

In \([12]\) Hansen et. al. used the AutoGraphiX software package and made the following conjecture:

**Conjecture 1.1.** Let \( G \) be a connected bipartite graph with \( n \geq 4 \) vertices and \( m \geq n \) edges. Then

\[
Sz(G) - W(G) \geq 4n - 8.
\]

Moreover the bound is best possible as shown by the graph composed of a cycle \( C_4 \) on 4 vertices and a tree \( T \) on \( n - 3 \) vertices sharing a single vertex.

This paper offers a confirmative proof to this conjecture.

2. Main Results

In \([19]\), another expression for the Szeged index was put forward, namely

\[
Sz(G) = \sum_{e = uv \in E(G)} n_u(e) n_v(e) = \sum_{e = uv \in E(G)} \sum_{\{x, y\} \subseteq V(G)} \mu_{x,y}(e)
\]

where \( \mu_{x,y}(e) \), interpreted as the contribution of the vertex pair \( x \) and \( y \) to the product \( n_u(e) n_v(e) \), is defined as:

\[
\mu_{x,y}(e) = \begin{cases} 
1 & \text{if } \begin{cases} d(x, u) < d(x, v) \text{ and } d(y, v) < d(y, u) \\
& \text{or } \begin{cases} d(x, v) < d(x, u) \text{ and } d(y, u) < d(y, v) 
\end{cases}
\end{cases} \\
0 & \text{otherwise.}
\end{cases}
\]

We first show that for a 2-connected bipartite graph Conjecture 1.1 is true.
Lemma 2.1. Let $G$ be a 2-connected bipartite graph of order $n \geq 4$. Then

$$Sz(G) - W(G) \geq 4n - 8$$

with equality if and only if $G \cong C_4$.

Proof. From Eq. (2.1), we know that

$$Sz(G) - W(G) = \sum_{\{x,y\} \subseteq V(G)} \sum_{e \in E(G)} \mu_{x,y}(e) - \sum_{\{x,y\} \subseteq V(G)} d(x, y)$$

$$= \sum_{\{x,y\} \subseteq V(G)} \left[ \sum_{e \in E(G)} \mu_{x,y}(e) - d(x, y) \right].$$

Claim: For every pair $\{x, y\} \subseteq V(G)$, we have

$$\sum_{e \in E(G)} \mu_{x,y}(e) - d(x, y) \geq 1.$$

In fact, if $xy \in E(G)$, that is $d(x, y) = 1$, then we can find a shortest cycle $C$ containing $x$ and $y$ since $G$ is 2-connected. Then, $G[C]$ has no chord. Since $G$ is bipartite, the length of $C$ is even. There is an edge $e'$ which is the antipodal edge of $e = xy$ in $C$. It is easy to check that $\mu_{x,y}(e') = \mu_{x,y}(e) = 1$. So the claim is true.

If $d(x, y) \geq 2$, let $P_1$ be a shortest path from $x$ to $y$ and $P_2$ be a second-shortest path from $x$ to $y$, that is, $P_2 \neq P_1$ and $|P_2| = \min \{|P| | P$ is a path from $x$ to $y$ and $P \neq P_1\}$. Since $G$ is 2-connected, $P_2$ always exists. If there is more than one path satisfying the condition, we choose $P_2$ as a one having the greatest number of common vertices with $P_1$.

If $E(P_1) \cap E(P_2) = \emptyset$, let $P_1 \cup P_2 = C$, and then $|E(P_2)| \geq |E(P_1)|$ and all the antipodal edges of $P_1$ in $C$ make $\mu_{x,y}(e) = 1$. We also know that $\mu_{x,y}(e) = 1$ for all $e \in E(P_1)$. Hence, $\sum_{e \in E(G)} \mu_{x,y}(e) - d(x, y) \geq d(x, y) > 1$.

If $E(P_1) \cap E(P_2) \neq \emptyset$, then $P_1 \triangle P_2 = C$, where $C$ is a cycle. Let $P'_1 = P_i \cap C = xP_1y'$. It is easy to see that $|E(P'_1)| \geq |E(P_1)|$, and the shortest path from $x$ (or $y$) to the vertex $v$ in $P'_2$ is $xP_2x'$ (or $yP_2y'$) together with the shortest path from $x'$ (or $y'$) to $v$ in $C$. So, all the antipodal edges of $P'_1$ in $C$ make $\mu_{x,y}(e) = 1$. We also know that $\mu_{x,y}(e) = 1$ for all $e \in E(P_1)$. Hence, $\sum_{e \in E(G)} \mu_{x,y}(e) = |E(P_1)| + d(x', y') \geq d(x, y) + 1$, which proves the claim.

Now, let $C = v_1v_2 \ldots v_pv_1$ be a shortest cycle in $G$, where $p$ is even and $p \geq 4$. Actually, for every $e \in E(C)$ we have that $\mu_{v_i,v_{p/2+i}}(e) = 1$ for $i = 1, 2, \ldots, \frac{p}{2}$. Then $\sum_{e \in E(G)} \mu_{v_i,v_{p/2+i}}(e) = |C| = p$, that is,

$$\sum_{e \in E(G)} \mu_{v_i,v_{p/2+i}}(e) - d(v_i, v_{p/2+i}) = p/2 \geq 2. \text{ Combining this with the claim, we have that}$$

$$Sz(G) - W(G) \geq \left(\frac{n}{2}\right) + \frac{p}{2} \left(\frac{p}{2} - 1\right) \geq \left(\frac{n}{2}\right) + 2 \geq 4n - 8.$$
The last two equalities hold if and only if $p = 4$, $n = 4$ or 5. If $n = 4, p = 4$, then $G \cong C_4$. If $n = 5, p = 4$, then $G \cong K_{2,3}$, and in this case we can easily calculate that $S_z(G) - W(G) > 12$. Thus, the equality holds if and only if $G \cong C_4$.

We now complete the proof of Conjecture 1.1 in the general case.

**Theorem 2.2.** Let $G$ be a connected bipartite graph with $n \geq 4$ vertices and $m \geq n$ edges. Then

$$S_z(G) - W(G) \geq 4n - 8.$$ 

Equality holds if and only if $G$ is composed of a cycle $C_4$ on 4 vertices and a tree $T$ on $n - 3$ vertices sharing a single vertex.

**Proof.** We have proved that the conclusion is true for a 2-connected bipartite graph. Now suppose that $G$ is a connected bipartite graph with blocks $B_1, B_2, \ldots, B_k$, where $k \geq 2$. Let $|B_i| = n_i$. Then, $n_1 + n_2 + \cdots + n_k = n + k - 1$. Since $m \geq n$ and $G$ is bipartite, there exists at least one block, say $B_1$, such that $n_1 \geq 4$. Consider a pair $\{x, y\} \subseteq V$. We have the following four cases:

**Case 1:** $x, y \in B_i$, and $n_i \geq 4$. Then for every $e \in E(B_i)$, we have $u_{x,y}(e) = 0$, which combined with Lemma 2.1 yields

$$\sum_{\{x, y\} \subseteq B_i} \left[ \sum_{e \in E(G)} u_{x,y}(e) - d(x, y) \right] = \sum_{\{x, y\} \subseteq B_i} \left[ \sum_{e \in E(B_i)} u_{x,y}(e) - d(x, y) \right] \geq 4n_i - 8.$$  

**Case 2:** $x, y \in B_i$, and $n_i = 2$. In this case,

$$\sum_{\{x, y\} \subseteq B_i} \left[ \sum_{e \in E(G)} u_{x,y}(e) - d(x, y) \right] = 0 = 4n_i - 8.$$  

**Case 3:** $x \in B_1$, $y \in B_i$, $i \neq 1$. Let $P$ be a shortest path from $x$ to $y$, and let $w_1, w_i$ be the cut vertices in $B_1$ and $B_i$, such that every path from a vertex in $B_1$ to $B_i$ must go through $w_1, w_i$. By the proof of Lemma 2.1, we can find an edge $e' \in E(B_1) \setminus E(P)$, such that $u_{x,w_1}(e') = 1$. Because every path from a vertex in $B_1$ to $y$ must go through $w_1$, we have $u_{x,y}(e') = 1$. We also know that $u_{x,y}(e) = 1$ for all $e \in E(P)$. Hence, $\sum_{e \in E(G)} u_{x,y}(e) - d(x, y) \geq 1$.

We are now in the position to show that for all $y \in B_1 \setminus \{w_1\}$, we can find a vertex $z \in B_1 \setminus \{w_1\}$ such that $\sum_{e \in E(G)} u_{z,y}(e) - d(z, y) \geq 2$. Since $B_1$ is 2-connected with $n_1 \geq 4$, there is a cycle containing $w_1$. Let $C$ be a shortest cycle containing $w_1$, say $C = v_1v_2 \cdots v_pv_1$, where $v_1 = w_1$ and $p$ is even. Set $z = v_{p/2+1}$. By the proof of Lemma 2.1, we have that $\sum_{e \in E(B_1)} u_{z,w_1}(e) - d(z, w_1) \geq p/2 \geq 2$. It follows that there are two edges $e', e''$, that are not in the shortest path from $z$ to $w_1$, such that $u_{z,w_1}(e') = 1$ and $u_{z,w_1}(e'') = 1$. Thus, $u_{z,y}(e') = 1$ and $u_{z,y}(e'') = 1$. Hence, $\sum_{e \in E(G)} u_{z,y}(e) - d(z, y) \geq 2$. 

If we fix $B_i$, we obtain that

$$\sum_{x \in B_i \setminus \{w_i\}} \left[ \sum_{y \in B_i \setminus \{w_i\}} \mu_{x,y}(e) - d(x, y) \right] \geq (n_1 - 1)(n_i - 1) + (n_i - 1) = n_1(n_1 - 1).$$

**Case 4:** $x \in B_i$, $y \in B_j$, $i \geq 2$, $j \geq 2$, $i \neq j$. Let $P$ be a shortest path between $x$ and $y$. If $P$ passes through a block $B_{\ell}$ with $n_\ell \geq 4$, and $|B_{\ell} \cap P| \geq 2$, then we have that $\sum_{e \in E(G)} \mu_{x,y}(e) - d(x, y) \geq 1$.

Otherwise, $\sum_{e \in E(G)} \mu_{x,y}(e) - d(x, y) \geq 0$. So,

$$\sum_{x \in B_i \setminus \{w_i\}} \left[ \sum_{y \in B_i \setminus \{w_i\}} \mu_{x,y}(e) - d(x, y) \right] \geq 0.$$

Equality holds if and only if $P$ passes through a block $B_{\ell}$ with $n_\ell = 2$ or $n_\ell \geq 4$, and $|B_{\ell} \cap P| = 1$.

From the above four cases it follows that

$$Sz(G) - W(G) = \sum_{(x,y) \subseteq V(G)} \sum_{e \in E(G)} \mu_{x,y}(e) - \sum_{(x,y) \subseteq V(G)} d(x, y)$$

$$= \sum_{(x,y) \subseteq V(G)} \left[ \sum_{e \in E(G)} \mu_{x,y}(e) - d(x, y) \right]$$

$$= \sum_{i=1}^k \sum_{(x,y) \subseteq B_i} \left[ \sum_{e \in E(G)} \mu_{x,y}(e) - d(x, y) \right] + \sum_{j=2}^k \sum_{x \in B_i \setminus \{w_i\}} \sum_{y \in B_j \setminus \{w_j\}} \left[ \sum_{e \in E(G)} \mu_{x,y}(e) - d(x, y) \right]$$

$$+ \frac{1}{2} \sum_{i \neq j} \sum_{x \in B_i \setminus \{w_i\}} \sum_{y \in B_j \setminus \{w_j\}} \left[ \sum_{e \in E(G)} \mu_{x,y}(e) - d(x, y) \right] \geq \sum_{i=1}^k (4n_i - 8) + n_1 \sum_{j=2}^k (n_j - 1)$$

$$= 4(n + k - 1) - 8k + n_1(n - n_1) = 4n - 4k - 4 + n_1(n - n_1).$$

Since $n_1 + n_2 + \cdots + n_k = n + k - 1$, $n_1 \geq 4$, $n_i \geq 2$, for $2 \leq i \leq k$, we have that $4 \leq n_1 \leq n - k + 1$, and $2 \leq k \leq n - 3$.

If $k \geq 5$, then $n_1(n - n_1) \geq 4(n - 4)$. Thus,

$$4n - 4k - 4 + n_1(n - n_1) \geq 8n - 4k - 20 \geq 8n - 4(n - 3) - 20 = 4n - 8.$$  

Equality holds if and only if $n_1 = 4$, $n_2 = n_3 = \cdots = n_{n-3} = 2$ i.e., if $B_2, B_3, \ldots, B_{n-3}$ form a tree $T$ on $n - 3$ vertices, that shares a single vertex with $B_1$.

If $2 \leq k \leq 4$, then $n_1(n - n_1) \geq (n - k + 1)(k - 1)$. 


If \( k = 2 \), then \( 4n - 4k - 4 + (n - k + 1)(k - 1) = 5n - 13 \geq 4n - 8 \). Equality holds if and only if \( n = 5 \), \( G \) is a graph composed of a cycle on 4 vertices and a pendant edge.

If \( k = 3 \), then \( 4n - 4k - 4 + (n - k + 1)(k - 1) = 6n - 20 \geq 4n - 8 \). Equality holds if and only if \( n = 6 \), \( G \) is a graph composed of a cycle on 4 vertices and a tree on 3 vertices sharing a single vertex.

If \( k = 4 \), then \( 4n - 4k - 4 + (n - k + 1)(k - 1) = 7n - 29 \geq 4n - 8 \). Equality holds if and only if \( n = 7 \), \( G \) is a graph composed of a cycle on 4 vertices and a tree on 4 vertices sharing a single vertex.

By this, the proof of Theorem 2.2 is completed. \( \square \)

**Remark 2.3.** The method used in the proof of Theorem 2.2 is not applicable to non-bipartite graphs. This is because given a 2-connected non-bipartite graph \( G \), for any two vertices \( x, y \in V(G) \), if \( C \) is an odd cycle, where \( C \) is defined as in Lemma 2.1, we cannot get \( \sum_{e \in E(G)} \mu_{x,y}(e) - d(x,y) \geq 1 \). Hence, for non-bipartite graphs we do not have an auxiliary result like Lemma 2.1.

**Acknowledgments**

This work was supported by the “973” program and NSFC.

**References**


Lily Chen
Center for Combinatorics, LPMC-TJKLC, Nankai University, Tianjin 300071, China
Email: lily60612@126.com

Xueliang Li
Center for Combinatorics, LPMC-TJKLC, Nankai University, Tianjin 300071, China
Email: lxl@nankai.edu.cn

Mengmeng Liu
Center for Combinatorics, LPMC-TJKLC, Nankai University, Tianjin 300071, China
Email: liumm05@163.com

Ivan Gutman
Faculty of Science, University of Kragujevac, P. O. Box 60, 34000 Kragujevac, Serbia
Email: gutman@kg.ac.rs