Discussiones Mathematicae
Graph Theory 39 (2019) 23–29
doi:10.7151/dmgt.2057

THE BIPARTITE-SPLITTANCE OF A BIPARTITE GRAPH[1]

Jian-Hua Yin and Jing-Xin Guan

Department of Mathematics
College of Information Science and Technology
Hainan University, Haikou 570228, P.R. China

e-mail: yinjh@hainu.edu.cn

Abstract

A *bipartite-split graph* is a bipartite graph whose vertex set can be partitioned into a complete bipartite set and an independent set. The *bipartite-splittance* of an arbitrary bipartite graph is the minimum number of edges to be added or removed in order to produce a bipartite-split graph. In this paper, we show that the bipartite-splittance of a bipartite graph depends only on the degree sequence pair of the bipartite graph, and an easily computable formula for it is derived. As a corollary, a simple characterization of the degree sequence pair of bipartite-split graphs is also given.

Keywords: degree sequence pair, bipartite-split graph, bipartite-splittance.

2010 Mathematics Subject Classification: 05C07.

1. Introduction

Let G be a simple graph with vertex set $V(G)$ and edge set $E(G)$. A subset S of $V(G)$ is *complete* if the subgraph $G[S]$ induced by S is complete, and it is *independent* if $G[S]$ is a null graph (i.e., a graph without edges). A *split graph* is a graph whose set of vertices can be partitioned into a complete set and an independent set. Split graphs were introduced by Földes and Hammer [1], who proved that a graph is split if and only if it does not have an induced subgraph isomorphic to C_4, C_5 or $2K_2$, where C_k is a cycle on k vertices and $2K_2$ is the disjoint union of two complete graphs K_2. The *splittance* $\sigma(G)$ of an arbitrary graph G is the minimum number of edges to be added to, or removed from G

[1]Supported by National Natural Science Foundation of China (No. 11561017) and Natural Science Foundation of Hainan Province (No. 2016CXTD004).

in order to obtain a split graph. An explicit, easily computable formula for the splittance $\sigma(G)$ in terms of the degree sequence of G and a simple characterization of the degree sequence of split graphs were presented by Hammer and Simeone [3].

Analogous problem is also studied for bipartite graphs in this paper. The style of this paper closely follows that of [3]. Let G be a simple bipartite graph with two partite sets X and Y, where $|X| = m$ and $|Y| = n$. If $A = (a_1, \ldots, a_m)$ (respectively, $B = (b_1, \ldots, b_n)$) is the non-increasing sequence of vertex degrees for X (respectively, Y), then the pair $(A; B)$ is the *degree sequence pair* of G. We say that $(S_1; S_2)$ is a *subset pair* of $(X; Y)$ if $S_1 \subseteq X$ and $S_2 \subseteq Y$. A subset pair $(S_1; S_2)$ of $(X; Y)$ is *complete bipartite* if either $S_1 = S_2 = \emptyset$, or $S_1 \neq \emptyset$, $S_2 \neq \emptyset$ and $G[S_1 \cup S_2]$ is a complete bipartite graph with two partite sets S_1 and S_2. A subset pair $(S_1; S_2)$ of $(X; Y)$ is *independent* if $G[S_1 \cup S_2]$ is a null graph. A *bipartite-split graph* is a bipartite graph whose two partite sets can be partitioned into a complete bipartite subset pair and an independent subset pair. The *bipartite-splittance* $\tau(G)$ of an arbitrary bipartite graph G is the minimum number of edges to be added to, or removed from G in order to obtain a bipartite-split graph. Clearly, G is bipartite-split if and only if $\tau(G) = 0$. The main result of this paper is to give an easily computable formula for the bipartite-splittance $\tau(G)$ of a bipartite graph G in terms of the degree sequence pair of G (Theorem 5). As a corollary, a simple characterization of the degree sequence pair of bipartite-split graphs is also given (Corollary 6).

2. Main Result and Its Proof

Let $(S_1; S_2)$ be a subset pair of $(X; Y)$, and let

$$s_{(S_1;S_2)} = |S_1||S_2| - |E(G[S_1 \cup S_2])| + |E(G[V(G) \setminus (S_1 \cup S_2)])|.$$

It is easy to see that $|S_1||S_2| - |E(G[S_1 \cup S_2])|$ is the number of edges to be added to $G[S_1 \cup S_2]$ in order to make $G[S_1 \cup S_2]$ into a complete bipartite graph, and $|E(G[V(G) \setminus (S_1 \cup S_2)])|$ is the number of edges to be removed from $G[V(G) \setminus (S_1 \cup S_2)]$ in order to make $G[V(G) \setminus (S_1 \cup S_2)]$ into a null graph.

Lemma 1. *Let G be a bipartite graph with two partite sets X and Y. Then*

$$\tau(G) = \min_{S_1 \subseteq X, S_2 \subseteq Y} s_{(S_1;S_2)}.$$

Proof. Clearly $\tau(G) \leq s_{(S_1;S_2)}$ for all $S_1 \subseteq X$ and $S_2 \subseteq Y$. On the other hand, let G' be a bipartite-split graph, with the complete bipartite subset pair $(S_1'; S_2')$ and the independent subset pair $(X \setminus S_1'; Y \setminus S_2')$, obtained from G by the addition or the removal of a minimum number of edges. Because of the minimality assumption, no removed edge could have had an end-vertex in $S_1' \cup S_2'$

and no added edge could have had an end-vertex in $(X \setminus S_1') \cup (Y \setminus S_2')$, implying that $\tau(G) = s_{(S_1';S_2')}$. Therefore, $\tau(G) = \min_{S_1 \subseteq X, S_2 \subseteq Y} s_{(S_1;S_2)}$. ■

Let $A = (a_1, \ldots, a_m)$ and $B = (b_1, \ldots, b_n)$ be two integer sequences with $\sum_{i=1}^m a_i = \sum_{i=1}^n b_i$, $n \geq a_1 \geq \cdots \geq a_m \geq 0$ and $m \geq b_1 \geq \cdots \geq b_n \geq 0$. If there is a simple bipartite graph G such that $(A; B)$ is the degree sequence pair of G, then the pair $(A; B)$ is *bigraphic*. The following well-known result is the Gale-Ryser characterization of bigraphic pairs.

Theorem 2 [2, 4]**.** *The pair $(A; B)$ is bigraphic if and only if $\sum_{i=1}^k a_i \leq k\ell + \sum_{i=\ell+1}^n b_i$ for each $k = 1, \ldots, m$ and $\ell = 1, \ldots, n$.*

By the symmetry, Theorem 2 can be stated that the pair $(A; B)$ is bigraphic if and only if $\sum_{i=1}^\ell b_i \leq \ell k + \sum_{i=k+1}^m a_i$ for each $\ell = 1, \ldots, n$ and $k = 1, \ldots, m$. Therefore, we have the following.

Corollary 3. *If the pair $(A; B)$ is bigraphic, then $\sum_{i=1}^k a_i + \sum_{i=1}^\ell b_i \leq 2k\ell + \sum_{i=k+1}^m a_i + \sum_{i=\ell+1}^n b_i$ for each $k = 1, \ldots, m$ and $\ell = 1, \ldots, n$.*

Definition. Let $(A; B)$ be a bigraphic pair. Define

$$\tau_{k,\ell}^{(A;B)} = \frac{1}{2}\left(2k\ell - \sum_{i=1}^{k} a_i - \sum_{i=1}^{\ell} b_i + \sum_{i=k+1}^{m} a_i + \sum_{i=\ell+1}^{n} b_i\right)$$

for $1 \leq k \leq m$ and $1 \leq \ell \leq n$, $m_\ell = \max\{i | a_i \geq \ell\}$ for $1 \leq \ell \leq a_1$ and $n_1 = \max\{i | b_i \geq 1\}$. If

$$\min\left\{\tau_{m_\ell,\ell}^{(A;B)} | \ell = 1, \ldots, a_1\right\} \leq \tau_{1,n_1}^{(A;B)},$$

then we define $L \in \{1, \ldots, a_1\}$ so that $\tau_{m_L,L}^{(A;B)} = \min\{\tau_{m_\ell,\ell}^{(A;B)} | \ell = 1, \ldots, a_1\}$. If

$$\min\left\{\tau_{m_\ell,\ell}^{(A;B)} | \ell = 1, \ldots, a_1\right\} > \tau_{1,n_1}^{(A;B)},$$

then we define $L = n_1$ and $m_L = 1$.

Lemma 4. *If the pair $(A; B)$ is bigraphic, then*

(a) $\tau_{k,\ell}^{(A;B)} \geq 0$ *for each $k = 1, \ldots, m$ and $\ell = 1, \ldots, n$;*

(b) *for a given $\ell \in \{1, \ldots, a_1\}$, $\tau_{k,\ell}^{(A;B)}$ attains its minimum value when $k = m_\ell$;*

(c) *for a given $\ell \in \{a_1 + 1, \ldots, n\}$, $\tau_{k,\ell}^{(A;B)}$ attains its minimum value when $k = 1$;*

(d) $\min\left\{\tau_{k,\ell}^{(A;B)} | k = 1, \ldots, m \text{ and } \ell = 1, \ldots, n\right\} = \min\left\{\min\left\{\tau_{m_\ell,\ell}^{(A;B)} | \ell = 1, \ldots, a_1\right\}, \tau_{1,n_1}^{(A;B)}\right\} = \tau_{m_L,L}^{(A;B)}$.

Proof. (a) is a consequence of Corollary 3. As for (b), a direct computation shows that, for $1 \leq k \leq m$, we have that $\tau_{k,\ell}^{(A;B)} - \tau_{k-1,\ell}^{(A;B)} = \ell - a_k$. It is easy to see that $\tau_{1,\ell}^{(A;B)} \geq \tau_{2,\ell}^{(A;B)} \geq \cdots \geq \tau_{m_\ell,\ell}^{(A;B)}$ and $\tau_{m_\ell,\ell}^{(A;B)} \leq \tau_{m_\ell+1,\ell}^{(A;B)} \leq \cdots \leq \tau_{m,\ell}^{(A;B)}$. Hence $\tau_{k,\ell}^{(A;B)}$ attains its minimum value when $k = m_\ell$. As for (c), we have that $\tau_{k,\ell}^{(A;B)} - \tau_{k-1,\ell}^{(A;B)} = \ell - a_k$ for $1 \leq k \leq m$, implying that $\tau_{1,\ell}^{(A;B)} \leq \tau_{2,\ell}^{(A;B)} \leq \cdots \leq \tau_{m,\ell}^{(A;B)}$. Hence $\tau_{k,\ell}^{(A;B)}$ attains its minimum value when $k = 1$. As for (d), we have that $\tau_{1,\ell}^{(A;B)} - \tau_{1,\ell-1}^{(A;B)} = 1 - b_\ell$ for $a_1 + 1 \leq \ell \leq n$, implying that $\tau_{1,a_1+1}^{(A;B)} \geq \cdots \geq \tau_{1,n_1}^{(A;B)}$ and $\tau_{1,n_1}^{(A;B)} \leq \cdots \leq \tau_{1,n}^{(A;B)}$. Hence $\tau_{1,n_1}^{(A;B)} = \min\left\{\tau_{1,\ell}^{(A;B)} | \ell = a_1+1, \ldots, n\right\}$. Thus by (b) and (c), $\min\left\{\tau_{k,\ell}^{(A;B)} | k = 1, \ldots, m \text{ and } \ell = 1, \ldots, n\right\} = \min\left\{\min\left\{\tau_{m_\ell,\ell}^{(A;B)} | \ell = 1, \ldots, a_1\right\}, \min\left\{\tau_{1,\ell}^{(A;B)} | \ell = a_1+1, \ldots, n\right\}\right\} = \min\left\{\min\left\{\tau_{m_\ell,\ell}^{(A;B)} | \ell = 1, \ldots, a_1\right\}, \tau_{1,n_1}^{(A;B)}\right\} = \tau_{m_L,L}^{(A;B)}$. ■

We now give the main result of this paper as follows.

Theorem 5. *If G is a bipartite graph with two partite sets X and Y and $(A;B)$ is the degree sequence pair of G, then* $\tau(G) = \min\left\{\min\left\{\tau_{m_\ell,\ell}^{(A;B)} | \ell = 1, \ldots, a_1\right\}, \tau_{1,n_1}^{(A;B)}\right\} = \tau_{m_L,L}^{(A;B)}$.

Proof. Let $(S_1; S_2)$ be any subset pair of $(X;Y)$, where $|X| = m$ and $|Y| = n$. Then

$$\sum_{x \in S_1} d_G(x) + \sum_{y \in S_2} d_G(y) = 2|E(G[S_1 \cup S_2])| + e_G(S_1 \cup S_2, V(G) \setminus (S_1 \cup S_2)),$$

where $e_G(S_1 \cup S_2, V(G) \setminus (S_1 \cup S_2))$ denotes the number of edges in G having one end-vertex in $S_1 \cup S_2$ and the other end-vertex in $V(G) \setminus (S_1 \cup S_2)$. Thus,

$$\begin{aligned} s_{(S_1;S_2)} &= |S_1||S_2| - |E(G[S_1 \cup S_2])| + |E(G[V(G) \setminus (S_1 \cup S_2)])| \\ &= \tfrac{1}{2}\left(2|S_1||S_2| - 2|E(G[S_1 \cup S_2])| + 2|E(G[V(G) \setminus (S_1 \cup S_2)])|\right) \\ &= \tfrac{1}{2}\left(2|S_1||S_2| - \sum_{x \in S_1} d_G(x) - \sum_{y \in S_2} d_G(y) + \sum_{x \in X \setminus S_1} d_G(x) + \sum_{y \in Y \setminus S_2} d_G(y)\right). \end{aligned}$$

By putting $k = |S_1|$ and $\ell = |S_2|$, we have that $\sum_{x \in S_1} d_G(x) \leq \sum_{i=1}^{k} a_i$, $\sum_{y \in S_2} d_G(y) \leq \sum_{i=1}^{\ell} b_i$, $\sum_{x \in X \setminus S_1} d_G(x) \geq \sum_{i=k+1}^{m} a_i$ and $\sum_{y \in Y \setminus S_2} d_G(y) \geq \sum_{i=\ell+1}^{n} b_i$. It follows that

$$s_{(S_1;S_2)} \geq \frac{1}{2}\left(2k\ell - \sum_{i=1}^{k} a_i - \sum_{i=1}^{\ell} b_i + \sum_{i=k+1}^{m} a_i + \sum_{i=\ell+1}^{n} b_i\right) = \tau_{k,\ell}^{(A;B)}.$$

We notice that if we take S_1 to be the set of vertices with degree $a_1, \dots, a_k$ and S_2 to be the set of vertices with degree $b_1, \dots, b_\ell$, then $s_{(S_1;S_2)} = \tau_{k,\ell}^{(A;B)}$. Therefore, by Lemmas 1 and 4, we have that $\tau(G) = \min_{S_1 \subseteq X, S_2 \subseteq Y} s_{(S_1;S_2)} = \min_{1 \le k \le m, 1 \le \ell \le n} \tau_{k,\ell}^{(A;B)} = \min\left\{ \min\left\{ \tau_{m_\ell,\ell}^{(A;B)} | \ell = 1, \dots, a_1 \right\}, \tau_{1,n_1}^{(A;B)} \right\} = \tau_{m_L,L}^{(A;B)}$. The proof of Theorem 5 is completed. ∎

Theorem 5 yields an easily computable formula for the bipartite-splittance of a bipartite graph. For example, for $1 \le r \le m$, let G be an r-regular bipartite graph on $2m$ vertices with two partite sets X and Y, and let $(A;B)$ be the degree sequence pair of G. Then $|X| = |Y| = m$, $a_1 = \cdots = a_m = r$ and $b_1 = \cdots = b_m = r$. It is easy to compute that $m_\ell = m$ for $1 \le \ell \le r$ and $n_1 = m$, and so $\tau_{m_\ell,\ell}^{(A;B)} = \frac{1}{2}(2m\ell - mr - \ell r + (m - \ell)r) = (m - r)\ell$ for $1 \le \ell \le r$ and $\tau_{1,n_1}^{(A;B)} = \frac{1}{2}(2m - r - mr + (m-1)r) = m - r$. Thus, $\tau(G) = \min\left\{ \min\left\{ \tau_{m_\ell,\ell}^{(A;B)} | \ell = 1, \dots, r \right\}, \tau_{1,n_1}^{(A;B)} \right\} = m - r$.

Let G be a bipartite graph with two partite sets X and Y, where $|X| = m$ and $|Y| = n$. The proof of Theorem 5 yields a simple procedure (see Algorithm 1 on next page) for obtaining a bipartite-split graph from G with a minimum number of additions or removals of edges. Moreover, we can easily analyze the complexity of Algorithm 1 is $O(\max\{m \log m, n \log n, mn\})$.

By the fact that G is bipartite-split if and only if $\tau(G) = 0$, a simple characterization of the degree sequence pair of bipartite-split graphs is an immediate consequence of Theorem 5.

Corollary 6. *Let $(A;B)$ be a bigraphic pair, and let L and m_L be defined as in Definition. Then $(A;B)$ is the degree sequence pair of a bipartite-split graph G if and only if $\tau_{m_L,L}^{(A;B)} = 0$, that is, $\sum_{i=1}^{m_L} a_i + \sum_{i=1}^{L} b_i = 2m_L L + \sum_{i=m_L+1}^{m} a_i + \sum_{i=L+1}^{n} b_i$.*

The following Corollary 7 is an immediate consequence of Corollary 6.

Corollary 7. *If a bipartite graph G is bipartite-split, then every bipartite graph with the same degree sequence pair as G is also bipartite-split.*

Remark 8. The problem in this paper can directly be considered in general graphs and is clearly hard in general graphs (for instance using a reduction from minimum edge removing to make the graph bipartite). Tighter hardness results in super classes of bipartite graphs would provide a nice motivation of the explicit formula in the bipartite case.

Algorithm 1:

Input: Bipartite graph G;

Output: Bipartite-split graph from G;

Let two partite sets of G be X and Y; m = number of vertices in X; n = number of vertices in Y;

Determine the degree sequence pair $(A;B)$ of G so that $A=(a_1,\ldots,a_m)$ (respectively, $B=(b_1,\ldots,b_n)$) is the non-increasing sequence of vertex degrees for X (respectively, Y);

Index the vertices of G so that $X=\{x_1,\ldots,x_m\}$ with $d_G(x_i)=a_i$, for $1\le i\le m$ and $Y=\{y_1,\ldots,y_n\}$ with $d_G(y_j)=b_j$, for $1\le j\le n$;

$m_\ell=\max\{i|a_i\ge\ell\}$, for $1\le\ell\le a_1$; $n_1=\max\{i|b_i\ge 1\}$;

$\tau_{m_\ell,\ell}^{(A;B)}=\frac{1}{2}\left(2m_\ell\ell-\sum_{i=1}^{m_\ell}a_i-\sum_{i=1}^{\ell}b_i+\sum_{i=m_\ell+1}^{m}a_i+\sum_{i=\ell+1}^{n}b_i\right)$, for $1\le\ell\le a_1$; $\tau_{1,n_1}^{(A;B)}=\frac{1}{2}\left(2n_1-a_1-\sum_{i=1}^{n_1}b_i+\sum_{i=2}^{m}a_i+\sum_{i=n_1+1}^{n}b_i\right)$;

$\tau_{m_r,r}^{(A;B)}=\min\left\{\tau_{m_1,1}^{(A;B)},\ldots,\tau_{m_{a_1},a_1}^{(A;B)}\right\}$;

if $\tau_{m_r,r}^{(A;B)}\le\tau_{1,n_1}^{(A;B)}$ **then**

  $L=r$, $m_L=m_r$;

else

  $L=n_1$, $m_L=1$;

for $i=1,\ldots,m_L$ *and* $j=1,\ldots,L$ **do**

  Add edges to $E(G)$ so that x_i and y_j are adjacent;

for $i=m_L+1,\ldots,m$ *and* $j=L+1,\ldots,n$ **do**

  Remove edges from $E(G)$ so that x_i and y_j are not adjacent.

Acknowledgement

The authors would like to thank the referees for their helpful suggestions and comments.

References

[1] S. Földes and P.L. Hammer, *Split graphs*, in: Proc. 8th Sout-Eastern Conf. on Combinatorics, Graph Theory and Computing, F. Hoffman *et al.* (Ed(s)), (Baton Rouge, Lousiana State University, 1977) 311–315.

[2] D. Gale, *A theorem on flows in networks*, Pac. J. Math. **7** (1957) 1073–1082. doi:10.2140/pjm.1957.7.1073

[3] P.L. Hammer and B. Simeone, *The splittance of a graph*, Combinatorica **1** (1981) 275–284.
doi:10.1007/BF02579333

[4] H.J. Ryser, *Combinatorial properties of matrices of zeros and ones*, Canad. J. Math. **9** (1957) 371–377.
doi:10.4153/CJM-1957-044-3

Received 12 May 2016
Revised 18 April 2017
Accepted 18 April 2017