

Bounds on the Number of Isolated Vertices in Sum Graphs

Ronald J. Gould¹

Vojtech Rödl

Emory University
Atlanta, GA 30322

ABSTRACT

A graph G is called a sum graph if the vertices of G can be labeled with distinct positive integers so that $e = uv$ is an edge of G if, and only if, the sum of the labels on vertex u and vertex v is also a label in G . It is clear that if G is a properly labeled sum graph, then the vertex receiving the highest label cannot be adjacent to any other vertex. Thus, every sum graph must contain isolated vertices. We consider the problem of finding a general upper bound on the number of isolated vertices present in a sum graph, as well as the problem of finding a lower bound, at least for certain graphs.

1. Introduction

All terms not defined here can be found in [3]. Throughout this paper we will consider (p, q) -graphs, that is, graphs with p vertices and q edges. We denote the vertex set of G as $V(G)$, the edge set as $E(G)$ and a label attached to a vertex v as $L(v)$.

Many different graph labeling problems have been studied (see for example [1]). Recently, Harary introduced a new variation, called sum graphs. A *sum graph* is a graph G in which each vertex x is labeled with a distinct positive integer $L(x)$ such that $e = uv$ is an edge of G if, and only if, $L(u) + L(v) = L(w)$, for some $w \in V(G)$. It is obvious that if G is properly labeled as a sum graph, then the vertex assigned the highest label cannot be adjacent to any other vertex of G ; thus, every sum graph must

¹ Research supported in part by O.N.R. contract number N000014-88-K-0070.

contain at least one isolated vertex. We denote by $s(G)$ the number of isolated vertices we must add to a given graph G in order to be able to label it as a sum graph. Harary [4] posed the problem of finding an upper bound for $s(G)$. The object of this paper is to present a general upper bound and, at least for certain graphs, a lower bound on $s(G)$.

In studying sum graphs, it becomes obvious that if we wish to minimize the number of isolated vertices necessary for a proper labeling, then we must maximize the number of duplicate sums; that is, the number of edges $e = uv$, where $L(u) + L(v) = k$, for some fixed value of k .

2. Main Results

In this section we concentrate on obtaining general upper and lower bounds on $f(G)$, where $f(G) = |V(G)| + s(G)$. Bounds for $s(G)$ are then immediate. We note that it is shown in [2] that $s(K_n) \leq 2n - 3$, with equality holding when $n \geq 6$. This can be achieved by labeling the vertices of K_n using consecutive elements of an arithmetic sequence, a fact that will prove useful to us later. For example, the labeling of 11, 21, 31, 41 on K_4 produces the need for five additional labels, namely 32, 42, 52, 62, 72. A *clique cover* of G is a partition of $V(G)$ into cliques. The set of all clique covers for G will be denoted $Cl(G)$, while for a particular cover $\hat{P} \in Cl(G)$, the number of cliques in \hat{P} will be denoted $|\hat{P}|$.

Theorem 1 *Let G be a (p, q) -graph. Then G can be labeled as a sum graph with*

$$f(G) \leq q + 3p - 3 |\hat{P}| - \left| \bigcup_{i=1}^{|\hat{P}|} E(C_i) \right|,$$

where \hat{P} is a clique cover in $Cl(G)$ containing the maximum number of edges of G within its cliques $C_1, C_2, \dots, C_{|\hat{P}|}$.

Proof Let G be a (p, q) -graph. We begin by decomposing $V(G)$ into cliques. We denote the classes of vertices in this cover by C_1, C_2, \dots, C_k . Further, we assume that $|V(C_i)| = t_i$, $1 \leq i \leq k$. Thus, each class of vertices in this partition has order t_i and size $\binom{t_i}{2}$.

Select integers a_i and d_i and label vertex v_j^i of C_i with

$$L(v_j^i) = a_i + jd_i.$$

By labeling in this manner, we introduce at most $2t_i - 3$ new labels (hence, new isolated vertices) relative to each clique C_i (c.f. [2]). Further, the new labels necessary are

$$2a_i + 3d_i, 2a_i + 4d_i, \dots, 2a_i + (2t_i - 1)d_i.$$

In general, such a labeling is done for each class C_i , $1 \leq i \leq k$.

In performing this labeling, we select the integers a_i and a_j so that the labels assigned to the vertices along with their subsequent implied sums for the edges between cliques are all distinct. More precisely, for $i \neq j$, we want

$$a_i + xd_i \neq a_j + yd_j \tag{1}$$

whenever $x \leq 2t_i - 3$ and $y \leq 2t_j - 3$. Moreover, for $\{i, j\} \neq \{r, s\}$, we also want

$$(a_i + xd_i) + (a_j + yd_j) \neq (a_r + zd_r) + (a_s + wd_s). \tag{2}$$

To see that such a labeling is possible, we may for example, select the integers using the following rules:

1. $d_{i+1} > 2td_i$ for $i = 1, 2, \dots, k-1$ and $t = \max_{1 \leq i \leq k} t_i$.
2. $a_i - a_j \gg d_k$, for every $i, j, 1 \leq i, j \leq k$.
3. $a_1 \ll a_2 \ll a_3 \ll \dots \ll a_k$.

Rule 2 insures us that equation (1) holds and that we never have the situation that for $\{i, j\} \neq \{r, s\}$

$$a_i + a_j = a_r + a_s \text{ holds.}$$

Further, it follows from Rule 3 that

$$(a_i + xd_i) + (a_j + yd_j) = (a_r + zd_r) + (a_s + wd_s) \tag{3}$$

implies that

$$a_i + a_j = a_r + a_s.$$

Thus, by our previous statements, $i = r$ and $j = s$. But then,

$$xd_i + yd_j = zd_r + wd_s,$$

and in fact, we have that

$$(x - z)d_i = (w - y)d_j. \tag{4}$$

But, using Rule 1 and the fact that $1 \leq w - y \leq 2t$, and assuming without loss of generality that $i < j$, we see that

$$(w - y)d_j \geq d_j > 2td_i \geq (x - z)d_i,$$

contradicting equation (4). But, then (2) must hold. Thus, the labeling is as claimed.

Now, with this labeling in mind, how large is $f(G)$? The total number of vertices (and hence labels) necessary for this labeling is

$$p + q + \sum_{i=1}^k [(2t_i - 3) - \binom{t_i}{2}].$$

Thus, if $\hat{P} \in Cl(G)$ contains the maximum number of edges of G within its cliques, then

$$f(G) \leq q + 3p - 3|\hat{P}| - \left| \bigcup_{i=1}^{|\hat{P}|} E(C_i) \right|,$$

completing the proof. \square

We now turn our attention to obtaining a lower bound on $f(G)$. We begin with a Lemma.

Lemma 1 *Let p and d be positive integers satisfying*

$$\frac{\log(ep^3)}{\log(5/4)} < d < \frac{p}{10}$$

then

$$\frac{p}{2d} \left(\frac{1}{ep^3}\right)^{1/d} > 4. \quad (*)$$

Proof The fact that

$$\frac{\log(ep^3)}{\log(5/4)} < d$$

implies that

$$(5/4)^d > ep^3$$

and hence that

$$\frac{1}{ep^3} (10d)^d > (8d)^d.$$

Now, using the fact that $10d < p$, we infer that

$$\frac{1}{ep^3} p^d > (8d)^d$$

which is equivalent to (*). \square

Theorem 2 *Let*

$$\frac{\log(ep^3)}{\log(5/4)} < d < \frac{p}{10}$$

then there exists a (p, q) -graph G with $q = \binom{p}{2} - j$ where $j \leq pd$ and

$$s(G) \geq q - \frac{(p+1)(p-1)}{2d} \ln(ep^3) - p.$$

Proof Here we consider \hat{K} , the class of graph with p vertices and $q = \binom{p}{2} - j$ edges ($j \leq pd$). To each member of \hat{K} we assign the most economical labeling possible, that is, a labeling that minimizes $f(G)$. This labeling induces a system of equations of type

1. $L(v_i) + L(v_j) = L(v_k)$ and
2. $L(v_i) + L(v_j) = L(v_i') + L(v_j')$.

For otherwise, we would have $p + q$ labels, which is clearly not the best possible labeling.

Now, for each graph $G \in \hat{K}$, only certain of the equations of type (1) or (2) actually hold. Suppose we view this system of equations as a homogeneous system of linear equations for the variables $L(v_i)$. Then, it follows from the fact that this system has a nonzero solution, that there are at most $p - 1$ equations with the property that all others

are consequences of these equations (i.e. a linear combination of these equations). Thus, we may assign to each $G \in \hat{K}$, a system of at most $p - 1$ equations of type (1) and (2).

On the other hand, there are at most

$$\binom{p}{2} (p - 3)$$

equations of type (1) and at most

$$\frac{1}{2} \binom{p}{2} (p - 2)$$

equations of type (2). Thus, the total number of possible equations of type (1) and (2) is clearly less than p^4 . Hence, the number of ways in which we can choose this collection of at most $p - 1$ equations is

$$b = \sum_{j \leq p-1} \binom{p^4}{j} < \binom{p^4}{p} < (ep^3)^p. \tag{2.1}$$

We partition the graphs in \hat{K} into b classes, say

$$K_1, K_2, \dots, K_b$$

according to which tuple of equations was assigned to K_i ($1 \leq i \leq b$). As a notational convenience, in what follows let

$$m = \binom{p}{2},$$

and let

$$N = \sum_{i=1}^b |K_i|.$$

Since

$$\sum_{i=1}^b |K_i| = N = \sum_{j \leq pd} \binom{m}{j}, \tag{2.2}$$

then the average number of graphs in any of the classes is $\frac{N}{b}$. Without loss of generality, suppose K_1 is a class containing at least the average number of graphs. Then we have

$$|K_1| \geq \frac{N}{b}. \tag{2.3}$$

Equations of type (1) and (2) satisfied in K_1 induce an equivalence relation P_1, P_2, \dots, P_k on the set of all pairs of vertices. The edge set of each graph $G \in K_1$ is equal to

$$\bigcup_{i=1}^v P_{j_i}$$

where $P_{j_1}, P_{j_2}, \dots, P_{j_v}$ is a subset of P_1, \dots, P_k that satisfies

$$\sum_{i=1}^v |P_{j_i}| = q \geq \binom{p}{2} - pd. \tag{2.4}$$

As all labels of vertices of G need not correspond to sums of edge labels (for example, the minimum label), the number of additional isolated vertices needed to label G is at least $v - p$. If we set $|P_{j_i}| = t_i$ ($1 \leq i \leq k$), then equation (2.4) implies that

$$v = q - \sum_{i=1}^k (|P_{j_i}| - 1) \geq q - \sum_{i=1}^k t_i + k,$$

and hence,

$$f(G) = s(G) + p \geq v \geq q - \sum_{i=1}^k t_i + k = q - x \quad (2.5)$$

where, $x = \sum_{j=1}^k t_j - k$.

It is clear that

$$\sum_{j \leq pd} \binom{m-x}{j} \geq |K_1|,$$

and hence, from equation (2.3), it is clear that

$$\sum_{j \leq pd} \binom{m-x}{j} \geq |K_1| \geq \frac{N}{b}. \quad (2.6)$$

Using equations (2.1) and (2.2) we also see that

$$\frac{N}{b} \geq \sum_{j \leq pd} \binom{m}{j} (ep^3)^{-p} \geq \binom{m}{pd} (ep^3)^{-p}. \quad (2.7)$$

Therefore, (2.6) and (2.7) imply that

$$\sum_{j \leq pd} \binom{m-x}{j} \geq \binom{m}{pd} (ep^3)^{-p}. \quad (2.8)$$

Now, using the fact that $\binom{r}{s} > \left(\frac{r}{s}\right)^s$ and the fact that $\binom{p}{2} = \frac{p^2}{2}$, we see that

$$\binom{m}{pd} (ep^3)^{-p} \geq \left[\left(\frac{p}{2d}\right)^d \frac{1}{ep^3} \right]^p.$$

Comparing this and an obvious upper bound for the left hand side of equation (2.8), we infer that

$$2^{m-x} \geq \sum_{j \leq pd} \binom{m-x}{j} \geq \left[\left(\frac{p}{2d}\right)^d \frac{1}{ep^3} \right]^p,$$

and applying \log_2 , we see that

$$m-x \geq pd \log_2 \left(\frac{p}{2d} \left(\frac{1}{ep^3}\right)^{1/d} \right).$$

As the assumptions of Lemma 1 are satisfied, we infer that

$$\log_2 \left(\frac{p}{2d} \left(\frac{1}{ep^3}\right)^{1/d} \right) > 2$$

and hence, $m-x \geq 2pd$ and thus,

$$pd \leq 1/2(m - x).$$

Using this new upper bound for pd , we obtain a new upper bound for $\binom{m}{pd}(ep^3)^{-p}$, namely

$$(1 + pd)\binom{m-x}{pd} > \sum_{j \leq pd} \binom{m-x}{j} \geq \binom{m}{pd}(ep^3)^{-p}. \tag{2.9}$$

Dividing both sides of (2.9) by $(1 + pd)\binom{m}{pd}$ we see that

$$\left(1 - \frac{x}{m}\right)\left(1 - \frac{x}{m-1}\right) \dots \left(1 - \frac{x}{m-pd+1}\right) \geq \frac{1}{1+pd}(ep^3)^{-p}.$$

Then, using the largest of the factors, we obtain the fact that

$$\left(1 - \frac{x}{m}\right)^{pd} \geq \frac{1}{1+pd}(ep^3)^{-p}.$$

However, since $(1 - y)^z = e^{z \ln(1-y)} < e^{-zy}$, if we let

$$y = \frac{x}{m} \text{ and } z = pd,$$

we see that

$$e^{-pdx/m} > \left(1 - \frac{x}{m}\right)^{pd} \geq \frac{1}{1+pd}(ep^3)^{-p} > (ep^3)^{-p-1} = e^{-(p+1)\ln(ep^3)}.$$

Now since e^w is an everywhere increasing function, we see that

$$\frac{-pdx}{m} > -(p+1) \ln(ep^3)$$

or

$$\frac{pdx}{m} < (p+1) \ln(ep^3).$$

Thus,

$$x \leq \frac{p+1}{p} \frac{m}{d} \ln(ep^3). \tag{2.10}$$

Now, from (2.5) and (2.10) we obtain the desired bound. \square

Our next goal is to restate Theorem 1 in order to show that in many cases the bound of Theorem 2 is close to best possible. Recall that we denoted by $\hat{K} = \hat{K}(p, d)$, the class of all graphs with p vertices and $\binom{p}{2} - j$ edges, where $j \leq pd$.

In order to estimate the upper bound in Theorem 1, we will actually find a lower bound for

$$3|\hat{P}| + \left| \bigcup_{i=1}^{|\hat{P}|} E(C_i) \right|$$

where \hat{P} is a clique cover of G containing the maximum number of edges of G within its cliques. In order to do this, we find it convenient to consider \bar{G} , the complement of G , and we shall consider colorings of the vertices of G with color classes $C_1, C_2, \dots, C_{|\hat{P}|}$.

Let $g(p, d)$ be the minimum of

$$\max (3|\hat{P}| + \sum_{i=1}^{|\hat{P}|} |E(C_i)|),$$

where the minimum is taken over all graphs G with at least p vertices and at most pd edges; while the maximum is taken over all colorings of G with color classes $C_1, \dots, C_{|\hat{P}|}$.

We will show that

$$g(p, d) \geq \frac{p^2}{6d+6} + d - \frac{p}{2} = f(p, d) \quad (0)$$

Our proof is based on induction on p . For $p=2$ the statement trivially holds. Now, let G be a graph with p vertices and at most pd edges. Let the largest independent set in G be C_0 (where $|C_0| = x$) and consider the graph $G - C_0$. Since each vertex of $V(G) - C_0$ is joined to some vertex of C_0 (due to the maximality of C_0), the subgraph G' induced on $V(G) - C_0$ has at most $pd - (p-x)$ edges; hence its average degree is at most

$$d' = \frac{pd}{p-x} - 1.$$

Using C_0 as one partition class and partitioning G' according to the induction assumption we see that

$$g(G) \geq 3 + \binom{x}{2} + g(G') \geq 3 + \binom{x}{2} + f(p-x, \frac{pd}{p-x} - 1).$$

Thus, it remains to show that

$$3 + \binom{x}{2} + f(p-x, \frac{pd}{p-x} - 1) \geq f(p, d). \quad (1)$$

We will prove (1) by induction on x .

Note that $x \geq \frac{p}{2d+1}$ by Turan's Theorem [5] and thus, to verify the anchor of the induction, we will show that the inequality holds for $x = \frac{p}{2d+1}$. (Note that we do not assume that $\frac{p}{2d+1}$ is necessarily an integer.) Thus,

$$\begin{aligned} 3 + \binom{p/(2d+1)}{2} &\geq f(p, d) - f(p - \frac{p}{2d+1}, \frac{pd}{p - \frac{p}{2d+1}} - 1) \\ &= f(p, d) - f(\frac{2d}{2d+1}p, \frac{2d-1}{2}). \end{aligned} \quad (2)$$

Substituting we see that

$$\begin{aligned} \frac{p^2}{6d+6} - \frac{(\frac{2d}{2d+1}p)^2}{6d+6} + \frac{1}{2} - \frac{p}{2(2d+1)} &< \frac{p^2}{3} \left(\frac{1}{2d+1} - \frac{2d}{(2d+1)^2} \right) + \frac{1}{2} - \frac{p}{2(2d+1)} \\ &< \frac{p^2}{2} \frac{1}{(2d+1)^2} - \frac{1}{2} \frac{p}{2d+1} + \frac{1}{2}, \end{aligned}$$

which is obviously smaller than the left hand side of (2), verifying the anchor of the induction.

Now we show the inductive step. This amounts to showing that

$$\binom{x+1}{2} + f(p-x-1, \frac{pd}{p-x-1} - 1) \geq \binom{x}{2} + f(p-x, \frac{pd}{p-x} - 1). \quad (3)$$

Equation (3) is clearly equivalent to

$$x \geq f(p-x, \frac{pd}{p-x} - 1) - f(p-x-1, \frac{pd}{p-x-1} - 1), \quad (4)$$

which we now verify.

First, we estimate the right hand side of equation (4) as follows:

$$\begin{aligned} \frac{(p-x)^2}{6\frac{pd}{p-x}} - \frac{(p-x-1)^2}{6\frac{pd}{p-x-1}} + \frac{pd}{p-x} - \frac{pd}{p-x-1} - \frac{1}{2} &\leq \frac{(p-x)^3}{6pd} - \frac{(p-x-1)^3}{6pd} \\ &\leq \frac{3(p-x)^2}{6pd}. \end{aligned} \quad (5)$$

We will show that (5) is smaller than x . To this end it is enough to verify that

$$3(p-x)^2 \leq 6pdx. \quad (6)$$

Since $p \geq x \geq 0$ and $d \geq 0$, then the right hand side of equation (6) is increasing in x while the left hand side of equation (6) is decreasing in x . Thus, as $x \geq \frac{p}{2d+1}$, it is enough to verify equation (6) for $x = \frac{p}{2d+1}$. We leave this easy computation to the reader. This completes the induction on x and hence verifies equation (1). Therefore, our proof of the desired bound is completed.

Considering $G \in \hat{K}$, and substituting the just derived bound into the inequality of Theorem 1 yields

$$f(G) \leq q + 3p - \left(\frac{p^2}{6d+6} + d - \frac{p}{2} \right). \quad (7)$$

At the same time, Theorem 2 tells us that there is $G \in \hat{K}$ with

$$f(G) \geq q - \frac{3(p+1)(p-1)}{6d} \ln(ep^3). \quad (8)$$

Thus, we see these two bounds are reasonably close.

Finally, we set

$$f(p) = \max \{f(G) \mid G \text{ has } p \text{ vertices}\}.$$

Using (7) and Theorem 2 we infer the following.

Theorem 3

$$\binom{p}{2} - \sqrt{2 \ln ep^3} p^{3/2} \leq f(p) \leq \binom{p}{2} - \sqrt{2/3} p^{3/2} + \frac{9p}{2}.$$

Proof First we outline the upper bound. Noting that $q = \binom{p}{2} - pd$ and maximizing formula (7) over all d , we infer that the minimum is attained for

$$d+1 = \sqrt{\frac{p}{6}}.$$

On the other hand, using the fact that formula (8) attains its maximum for

$$d = \sqrt{\frac{(p+1)(p-1)}{2p} \ln(ep^3)} ,$$

we infer the lower bound. \square

REFERENCES

- [1] Chung, F.R.K., *Some Problems and Results in Labelings of Graphs*, The Theory and Applications of Graphs, ed. by Chartrand, Alavi, Goldsmith, Lesniak-Foster, and Lick, Wiley Interscience, New York, NY, 1981.
- [2] Faudree, R., Gould, R., Gyrafás, A., Lesniak, L., Schelp, R., *On Sum Graphs*, preprint.
- [3] Gould, R., *Graph Theory*, Benjamin/Cummings Publishing Co., Menlo Park, CA, 1988.
- [4] Harary, F., personal communication.
- [5] Turán, P., *On an Extremal Problem in Graph Theory*. Mat. Fiz. Lapok, 48 (1941), 436 - 452.