Remarks on an Edge Isoperimetric Problem

C. Bey

Abstract. Among all collections of a given number of k-element subsets of an n-element groundset find a collection which maximizes the number of pairs of subsets which intersect in $k-1$ elements.

This problem was solved for $k=2$ by Ahlswede and Katona, and is open for $k>2$.

We survey some linear algebra approaches which yield to estimations for the maximum number of pairs, and we present another short proof of the Ahlswede-Katona result.

1 Introduction

Let $G=(\mathcal{V},\mathcal{E})$ be a simple graph. Given a subset $\mathcal{M}$ of the vertex set $\mathcal{V}$, we put $B_G(\mathcal{M}) := \{\{u,v\} \in \mathcal{E} : u \in \mathcal{M}, v \notin \mathcal{M}\}$, the *edge-boundary* of $\mathcal{M}$, and $I_G(\mathcal{M}) := \{\{u,v\} \in \mathcal{E} : u,v \in \mathcal{M}\}$, the set of *inner edges* spanned by $\mathcal{M}$. Two edge isoperimetric problems (EIP's) for G are the determination of the numbers $B_G(m) := \min\{|B_G(\mathcal{M})| : \mathcal{M} \subseteq \mathcal{V}, |\mathcal{M}| = m\}$ and $I_G(m) := \max\{|I_G(\mathcal{M})| : \mathcal{M} \subseteq \mathcal{V}, |\mathcal{M}| = m\}$. If G is regular with degree d, then both problems are equivalent since $2I_G(\mathcal{M}) + B_G(\mathcal{M}) = d|\mathcal{M}|$ holds for every $\mathcal{M} \subseteq \mathcal{V}$. We refer to [6] for a survey on edge isoperimetric problems on graphs.

Here we consider the Johnson graph $G = J(n,k)$. Let $[n]$ denote the set $\{1,2,\ldots,n\}$ and $\mathcal{V}_k^n$ the set of all k-element subsets of $[n]$. The graph $J(n,k)$ has vertex set $\mathcal{V}_k^n$, and edge set $\{\{A,B\} : |A \cap B| = k-1\}$. Thus, looking at incidence vectors, $J(n,k)$ is the graph whose vertices are the $\{0,1\}$-sequences of length n and weight k, and whose edges are those pairs of sequences with Hamming distance 2. The Johnson graph $J(n,k)$ is an adjacency relation of the Johnson scheme, which is the natural setting for studying constant weight codes (cf. [10]).

Note that $J(n,k)$ is regular with degree $k(n-k)$.

We write $B_{n,k}(m)$ for $B_{J(n,k)}(m)$ and $B(\mathcal{M})$ resp. $I(\mathcal{M})$ for $B_{J(n,k)}(\mathcal{M})$ resp. $I_{G(n,k)}(\mathcal{M})$.

The EIP of the graph $J(n,2)$ was solved by Ahlswede and Katona in [3]. Solutions for the case $k=2$ also appeared in [1,7] and [16]. In order to state the result, we need the following definition.

Let $m = \binom{d}{2} + t$ with $0 \le t < d$. The *quasi-complete* graph C_n^m on n vertices is obtained from the complete graph on d vertices by adding a vertex of degree t and $n-1-d$ isolated vertices. The *quasi-star* S_n^m is the complement of the graph $C_n^{\binom{n}{2}-m}$.

Theorem 1 ([3]). *For every $0 \le m \le \binom{n}{2}$ the minimum boundary $B_{n,2}(m)$ is attained for the quasi-complete graph C_n^m or for the quasi-star S_n^m.*

R. Ahlswede et al. (Eds.): Information Transfer and Combinatorics, LNCS 4123, pp. 971–978, 2006.

The EIP for $J(n, k)$ with $k > 2$ is open. The papers [12] and [2] study continuous versions of the EIP, whose solutions yield the numbers $B_{n,k}(m)$ for certain values of m. A natural conjecture on the structure of optimal solutions for the EIP is also disproved in [2].

This note contains a survey on some estimations for the EIP based on eigenvalues, as well as a short and new proof for Theorem 1.

2 Estimations Via Eigenvalues

Given a set $\mathcal{M} \subseteq \mathcal{V}_k^n$ of vertices of $J(n, k)$ and a set $P \subseteq [n]$ we denote by

$$d(P) = d_{\mathcal{M}}(P) = |\{M \subseteq \mathcal{M} : P \subseteq M\}|$$

the *degree* of P in $\mathcal{M}$. If $p \in [n]$ we write also $d(p)$ for $d(\{p\})$.

Obviously,

$$\sum_{P \in \mathcal{V}_{k-1}^n} d(P)^2 = 2|I(\mathcal{M})| + k|\mathcal{M}| = k(n-k+1)|\mathcal{M}| - |B(\mathcal{M})| \,. \tag{1}$$

Thus, determining the maximum sum of squares of degrees $d_{\mathcal{M}}(P)$, $P \in \mathcal{V}_{k-1}^n$, over all subsets $\mathcal{M} \subseteq \mathcal{V}_k^n$ of a given size is equivalent to the EIP.

In [9], de Caen proves the following inequality for the case $k = 2$: Given $\mathcal{M} \subseteq \mathcal{V}_2^n$,

$$\sum_{P \in \mathcal{V}_1^n} d(P)^2 \leq \frac{2}{n-1}|\mathcal{M}|^2 + (n-2)|\mathcal{M}|. \tag{2}$$

Of course, this inequality can be checked using Theorem 1, but the calculations are somewhat involved. De Caen's inequality was generalized to arbitrary k in [5]: For every $0 \leq p \leq n$ and $\mathcal{M} \subseteq \mathcal{V}_k^n$ we have

$$\sum_{P \in \mathcal{V}_p^n} d(P)^2 \leq \frac{\binom{k}{p}\binom{k-1}{p}}{\binom{n-1}{p}}|\mathcal{M}|^2 + \binom{k-1}{p-1}\binom{n-p-1}{k-p}|\mathcal{M}| \,. \tag{3}$$

Equality holds in (3) for $0 < p < k < n$ if and only if $\mathcal{M} = \varnothing$ or $\mathcal{M} = \mathcal{V}_k^n$ or $\mathcal{M}$ is a star or a complement of a star, or $n = k+1$ and $\mathcal{M} \subseteq \mathcal{V}_k^{k+1}$ is arbitrary. Here a *star* is the set of all k-element subsets which contain a fixed element from $[n]$, and the complement of $\mathcal{M}$ is $\mathcal{V}_k^n \setminus \mathcal{M}$.

The proofs of the above estimations in [9,5] utilize a positive semidefinite matrix in the Bose-Mesner-algebra of the Johnson scheme (cf. [10]). This matrix is essentially the p-element versus k-element subsets incidence matrix multiplied with it transposed. A more transparent proof, using this matrix, can be given using the following theorem. Recall that a semiregular graph is a bipartite graph such that the degrees of the vertices are constant on each bipartition.

Theorem 2. *Let $G = (\mathcal{V}, \mathcal{E})$ be a connected semiregular graph with bipartition $\mathcal{V} = \mathcal{V}_0 \cup \mathcal{V}_1$ and degrees d_0, d_1. Let μ_2 be the second largest eigenvalue of the adjacency matrix of G. For $\mathcal{M} \subseteq \mathcal{V}_1$ and $v \in \mathcal{V}_0$ put*

$$d(v) = |\{v_1 \in \mathcal{M} \ : \{v, v_1\} \in \mathcal{E}\}| .$$

Then

$$\sum_{v \in \mathcal{V}_0} d(v)^2 \leq \left(\frac{d_0 d_1 - (\mu_2)^2}{|\mathcal{V}_1|} \right) |\mathcal{M}|^2 + (\mu_2)^2 |\mathcal{M}| .$$

Proof: Let A be the adjacency matrix of G. We have

$$A = \begin{pmatrix} 0 & W \\ W^\top & 0 \end{pmatrix},$$

where W is the $|V_0| \times |V_1|$-matrix describing adjacency between the bipartitions V_0 and V_1. It is known that the nonzero eigenvalues of A are exactly the (positive and negative) square roots of the nonzero eigenvalues of $W^\top W$, with equal multiplicities correspondingly. We denote the eigenvalues of the latter matrix by $\mu_1^2, \mu_2^2, \ldots$, with $\mu_1 > \mu_2 > \cdots \geq 0$. Since G is biregular and connected, the largest eigenvalue is $\mu_1^2 = d_0 d_1$, and the corresponding eigenspace is one-dimensional and generated by the all one vector which we denote by j. In particular, μ_2 is indeed the second largest eigenvalue.

Now let φ be the characteristic row vector of $\mathcal{M}$ (of length V_1), and $\varphi = \varphi_1 + \varphi_2 + \ldots$ be the orthogonal decomposition of φ according to the eigenspaces of $W^\top W$. Note that $|\mathcal{M}| = \varphi^\top j = \varphi_1^\top j$, thus φ_1 is the constant vector with entry $|\mathcal{M}|/|\mathcal{V}_1|$, and $\varphi_1^\top \varphi_1 = |\mathcal{M}|^2/|V_1|$. Now we have

$$\sum_{v \in V_0} d(v)^2 = (W\varphi)^\top W\varphi = \mu_1^2 \varphi_1^\top \varphi_1 + \sum_{i \geq 2} \mu_i^2 \varphi_i^\top \varphi_i$$
$$\leq d_0 d_1 \varphi_1^\top \varphi_1 + \mu_2^2 (|\mathcal{M}| - \varphi_1^\top \varphi_1) = \left(\frac{d_0 d_1 - \mu_2^2}{|\mathcal{V}_1|} \right) |\mathcal{M}|^2 + \mu_2^2 |\mathcal{M}| .$$

Inequality (3) now follows from Theorem 2. The computation of the corresponding eigenvalue μ_2 (cf. [5]) uses the known eigenvalues of the Johnson scheme (cf. [10]).

Using (1) and (3) with $p = k - 1$ we get the following estimation for the EIP of the graph $J(n, k)$:

$$B_{n,k}(m) \geq \frac{n}{\binom{n}{k}} m \left(\binom{n}{k} - m \right) . \tag{4}$$

Equality holds for $0 < m < \binom{n}{k}$ if and only if $m = \binom{n-1}{k-1}$ or $m = \binom{n-1}{k}$ or $n = k + 1$.

This recalls a general edge isoperimetric inequality using Laplace eigenvalues due to Alon and Milman:

Theorem 3 ([4]). *Let $G = (\mathcal{V}, \mathcal{E})$ be a simple graph, and λ_2 be the second smallest Laplace eigenvalue of G. Then, for every $0 \le m \le |\mathcal{V}|$,*

$$B_G(m) \ge \frac{\lambda_2}{|\mathcal{V}|} m\,(|\mathcal{V}| - m)\,.$$

Recall that the Laplace eigenvalues of a graph G are the eigenvalues of the difference of the degree matrix and the adjacency matrix of G, where the degree matrix is the diagonal matrix having the degrees of G on its diagonal. For regular graphs the Laplace eigenvalues are the differences of the degree and the adjacency eigenvalues.

The second smallest Laplace eigenvalue of the Johnson graph $J(n,k)$ is n. This is again an easy computation using the eigenvalues of the Johnson scheme. Thus, inequality (4) and hence also (3) follow from Theorem 3. In fact, it is also easy to deduce Theorem 2 from Theorem 3 (at least if the bipartite graph in Theorem 2 is strongly regular when viewed from the bipartition V_1).

We continue with an estimation for the EIP in the case $k = 2$. Our Theorem 2 can be considered as a bipartite version of the following result.

Theorem 4 ([13]). *Let $G = (\mathcal{V}, \mathcal{M})$ be a graph and μ_1 be the largest eigenvalue of the adjacency matrix of G. For $v \in \mathcal{V}$ let $d(v) = d_{\mathcal{M}}(v)$ be the degree of v. Then*

$$\sum_{v \in \mathcal{V}} d(v)^2 \le (\mu_1)^2 |\mathcal{V}|\,.$$

If G is connected then equality holds if and only if G is regular or semiregular.

In order to apply this theorem one needs bounds on the largest eigenvalue of a graph. Many such bounds have been established, and we refer to [8] for a survey.

For example, a result by Schwenk [19] and Hong [14] says that the largest eigenvalue μ_1 of a connected graph $G = (\mathcal{V}, \mathcal{M})$ satisfies

$$\mu_1 \le \sqrt{2|\mathcal{M}| - |\mathcal{V}| + 1}\,.$$

This bound together with Theorem 4 however yields a weaker estimation for the $k = 2$ case of the EIP than inequality (2). A best possible upper bound on the largest eigenvalue (called index) of a graph in terms of the number of edges was obtained by Rowlinson [17]:

Theorem 5 ([17]). *Among all graphs with n vertices and m edges exactly the quasi-complete graph C_n^m has largest index.*

Note that this result together with the Theorem 4 comes close to the optimum of the EIP of $J(n,2)$ in half of all cases (but does not yield to sharp estimations for all quasi-complete graphs due to the equality characterization in Theorem 4). Indeed, the two previous theorems yield a better estimation for the EIP of $J(n,2)$ than inequality (2). It seems worthwhile to study analogues of Theorem 5 for hypergraphs in order to improve the estimation (4).

3 A Short Proof for the Case $k = 2$

Recall the majorization (or dominance) order for sequences: If $d = (d_1, \ldots, d_n)$ and $e = (e_1, \ldots, e_n)$ are real vectors we say that d is majorized by e if for all $j = 1, \ldots, n$ the sum of the largest j entries of d is not larger than the corresponding sum for e, and if equality holds for $j = n$. Equivalently, d is majorized by e if d can be obtained from e by a sequence of alterations each of which replaces two entries d_i, d_j with $d_i < d_j$ by $d_i + x$, $d_j - x$ with $0 \leq 2x \leq d_j - d_i$.

We look for graphs with a fixed number of vertices and a fixed number of edges which maximize the sum of squares of degrees. For this it is sufficient to look among graphs having a degree sequence which is not majorized by any other degree sequence. Indeed, this follows by elementary means or by noting that the sum of squares of degrees is a symmetric and convex function of the degrees. A characterization of such degree sequences is known, and gives in fact a characterization of all degree sequences of graphs:

Given an integer vector $d = (d_1, \ldots, d_n)$ with nonincreasing entries (i.e. a partition of $d_1 + \cdots + d_n$), the number $r := \max\{i : d_i \geq i\}$ is called the *rank* of d. Further, the sequence $d' = (d'_1, \ldots, d'_{n-1}, \ldots)$ with $d'_i = |\{j : d_j \geq i\}|$ is called the *conjugate* sequence (or partition).

Theorem 6 (Ruch, Gutman [18]). *Let $d = (d_1, \ldots, d_n)$ be a rank r sequence of nonincreasing nonnegative integers. Then d is a degree sequence of a graph with n vertices and m edges if and only if $2m = d_1 + \cdots + d_n$ and*

$$d_1 + \cdots + d_i + i \leq d'_1 + \cdots + d'_i \quad \textit{holds for all} \quad 1 \leq i \leq r. \tag{5}$$

Degree sequences of graphs which satisfy equality in (5) for all $i = 0, \ldots, r$ are called *threshold sequences*, the corresponding graphs are called *threshold graphs.* These graphs were introduced in a different manner in [11] and have many characterizations. For example, threshold graphs are exactly those graphs whose degree sequences are not realized by any other nonisomorphic graph. Also, the threshold sequences and their permutations are exactly the extreme points of the convex hull of all degree sequences of graphs on a fixed number of vertices. We refer to [15] for more details.

Theorem 6 says that a partition $d = (d_1, \ldots, d_n)$ is the degree sequence of a graph on n vertices if and only if d is majorized by a threshold sequence.

Our proof of the Ahlswede-Katona theorem will proceed by some easy operations on the diagrams of a threshold sequences. Recall that the (Young) diagram of a sequence $(d_1, \ldots, d_n)$ of nonincreasing integers is the array of $d_1 + \cdots + d_n$ boxes having n left-justified rows with row i containing d_i boxes, $i = 1 \ldots, n$. The *square* of a rank r partition d is the square consisting of the r^2 boxes in the left upper corner of the diagram of d. We will identify sequences of nonincreasing integers with their diagrams. Figure 1 shows the degree sequences of a quasi-star and a quasi-complete graph. Both types of graphs are threshold graphs. Note also that a rank r diagram of a quasi-star has $n - 1$ boxes in each of its first $r - 1$ rows, and a rank r diagram of a quasi-complete graph has at most $r + 1$ columns.

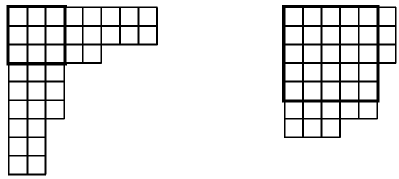

Fig. 1. Diagrams of S_9^{18} and C_9^{18}

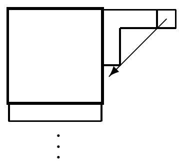

Fig. 2. Moving a box

A simple calculation shows that $\sum_i d_i'^2 = \sum_i d_i^2 + \sum_i d_i$ holds for all threshold sequences $(d_1, \ldots, d_n)$. Maximizing $\sum_i d_i^2$ over all threshold sequences with fixed $\sum_i d_i$ is thus equivalent to maximizing $\sum_i d_i'^2$. Let us place for each threshold sequence d the weight $i + j - 1$ in the box of row i and column j of the diagram of d. Then our maximization is equivalent to finding a threshold sequence which has largest sum of weights among all threshold sequences with a fixed number $2m = d_1 + \cdots + d_n$ of boxes.

Let us now turn to the proof of Theorem 1. We proceed by induction on the number m of edges. Suppose that G is an optimal threshold graph with m edges and nonincreasing degree sequence $(d_1, \ldots, d_n)$. We may assume that $d_1 = n - 1$. Indeed, in the other case the complement $\overline{G}$ of G has a vertex of degree $n - 1$, and is optimal among all graphs with $\binom{n}{2} - m$ edges (since the edge boundaries of G and its complement are equal). But if we can show that $\overline{G}$ can be taken to be a quasi-star or a quasi-complete graph, then G can be so too. The graph obtained from G by removing a vertex of degree $d_1 = n - 1$ is an optimal graph with $m - (n - 1)$ edges. By induction, we can assume that it is either a quasi-star or a quasi-complete graph. In the first case G itself is a quasi-star and we are done. Thus assume that the second case occurs. We will show how to transfer the diagram d of G's degree sequence into one of the diagrams of the quasi-star or the quasi-complete graph while maintaining the sum of weights. Since d is threshold, the boxes to the right of d's square have a mirror counterpart below the square. In the following, when we perform operations with the boxes lying to right of the square of a threshold diagram, we always assume that the same (transposed) operations are done with the boxes below the square, such that the resulting diagram will be again a threshold one.

Let r be the rank of d, and let $1 + s$ be the number of boxes in the $(r + 1)$-th column of d. We may assume that $r \geq 3$ and $n - 1 - r \geq 2$, since otherwise d is the diagram of a quasi-star resp. quasi-complete graph. We consider two cases.

First, let $r \geq n-1-r$. If $1+s \geq n-1-r$, then moving the last box of the first row to the first empty place in the $(r+1)$-th column (and doing the transposed operation below the square) will yield a threshold sequence with a larger sum of weights, contradicting the optimality of G (see Figure 2).

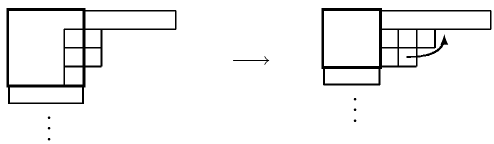

Fig. 3. Flipping and Moving a box

Thus $1 + s < n - 1 - r$. But then we flip the hook consisting of the $s + 1$ boxes in the $(r + 1)$-th column and the last $n - 1 - r$ boxes in the first row to obtain a threshold diagram d' with $n-1-r$ boxes in the $(r+1)$-th column and $r + s + 1$ boxes in the first row, while maintaining the sum of weights. If $s \geq 1$, then, as above, moving the last box from the first row of d' to the $(r + 1)$-th column yields a contradiction. Thus we have $s = 0$. But then d' is the diagram of a quasi-complete graph, and we are done with the case $r \geq n - 1 - r$. Let $r < n - 1 - r$. Then we flip the last s boxes of the $(r + 1)$-th column and the second up to r-th box of the r-th column to obtain a threshold sequence d'' with $2r - 2$ boxes in the second row and $r - 1 + s$ boxes in the third row, while maintaining the sum of weights (see Figure 3).

If $0 < s < r - 1$, then moving in the diagram d'' the last box of the third row to the first empty place of the second row increases the sum of weights, contradicting the optimality of G. If $s = r - 1$ (≥ 2), then moving the last two boxes of the third row to the first two empty places of the second row will increase the sum weights. Thus we have again $s = 0$. If $r \geq 4$, then d'' has rank $r-1$, and moving the box in row and column $r-1$ to the first empty place of the second row will increase the sum of weights. Thus we have $r = 3$. But then d'' is the diagram of a quasi-star, and we are also done with the case $r < n-1-r$.

References

1. R. Aharoni, A problem in rearrangements of $(0, 1)$-matrices. Discrete Math. 30, 191–201, 1980.
2. R. Ahlswede and N. Cai, On edge-isoperimetric theorems for uniform hypergraphs, Preprint 93-018, Diskrete Strukturen in der Mathematik SFB 343, Univ. Bielefeld, 1993; this volume.
3. R. Ahlswede and G.O.H. Katona, Graphs with maximal number of adjacent pairs of edges, Acta Math. Acad. Sci. Hungar. 32, 1-2, 97–120, 1978.
4. N. Alon and V.D. Milman, λ_1-isoperimetric inequalities for graphs, and superconcentrators, J. Combin. Theory Ser. B 38, 73–88, 1985.
5. C. Bey, An upper bound on the sum of squares of degrees in a hypergraph, Discrete Math. 269, 259–263, 2003.

6. S.L. Bezrukov, Edge isoperimetric problems on graphs, Graph theory and combinatorial biology (Balatonlelle, 1996), János Bolyai Math. Soc., Budapest, 157–197, 1999.
7. R.A. Brualdi, and E.S. Solheid, Some extremal problems concerning the square of a$(0,1)$-matrix, Linear Mult. Alg. 22, 57–73, 1987.
8. D. Cvetkovic and P. Rowlinson, The largest eigenvalue of a graph: A survey, Linear Mult. Alg. 28, 3–33, 1990.
9. D. de Caen, An upper bound on the sum of squares of degrees in a graph, Discrete Math. 185, 245–248, 1998.
10. P. Delsarte, An Algebraic Approach to the Association Schemes of Coding Theory, Philips Res. Rep. Suppl. 10, 1973.
11. P.L. Hammer, and V. Chvátal, Aggregation of inequalities in integer programming, Ann. Discrete Math. 1, 145–162, 1977.
12. L.H. Harper, On a problem of Kleitman and West, Discrete Math. 93, 2-3, 169–182, 1991.
13. M. Hofmeister, Spectral radius and degree sequence, Math. Nachr. 139, 37–44, 1988.
14. Y. Hong, A bound on the spectral radius of graphs, Linear Algebra Appl. 108, 135–139, 1988.
15. N.V.R. Mahadev and U.N. Peled, Threshold graphs and related topics, Ann. Discrete Math. 56, 1995.
16. D. Olpp, A conjecture of Goodman and the multiplicities of graphs. Austral. J. Comb. 14, 267–282, 1996.
17. P. Rowlinson, On the maximal index of graphs with a prescribed number of edges, Linear Algebra Appl. 110, 43–53, 1988.
18. E. Ruch and I. Gutman, The branching extent of graphs, J. Combinatorics, Inform. Syst. Sci. 4, 285–295, 1979.
19. A. Schwenk, New derivations of spectral bounds for the chromatic number, Graph Theory Newsletter 5, 77, 1975.