On the Shannon Capacity of a Graph

LÁSZLÓ LOVÁSZ

***Abstract*—It is proved that the Shannon zero-error capacity of the pentagon is $\sqrt{5}$. The method is then generalized to obtain upper bounds on the capacity of an arbitrary graph. A well-characterized, and in a sense easily computable, function is introduced which bounds the capacity from above and equals the capacity in a large number of cases. Several results are obtained on the capacity of special graphs; for example, the Petersen graph has capacity four and a self-complementary graph with n points and with a vertex-transitive automorphism group has capacity $\sqrt{n}$.**

I. Introduction

LET THERE BE a graph G, whose vertices are letters in an alphabet and in which adjacency means that the letters can be confused. Then the maximum number of one-letter messages which can be sent without danger of confusion is clearly $\alpha(G)$, the maximum number of independent points in the graph G. Denote by $\alpha(G^k)$ the maximum number of k-letter messages which can be sent without danger of confusion (two k-letter words are confoundable if for each $1 \leqslant i \leqslant k$, their ith letters are confoundable or equal). It is clear that there are at least $\alpha(G)^k$ such words (formed from a maximum set of nonconfoundable letters), but one may be able to do better. For example, if C_5 is a pentagon, then $\alpha(C_5^2)=5$. In fact, if $v_1,\cdots,v_5$ are the vertices of the pentagon (in this cyclic order), then the words v_1v_1, v_2v_3, v_3v_5, v_4v_2, and v_5v_4 are nonconfoundable.

It is easily seen that

$$\Theta(G)=\sup_k \sqrt[k]{\alpha(G^k)} = \lim_{k\to\infty} \sqrt[k]{\alpha(G^k)} .$$

This number was introduced by Shannon [6] and is called the *Shannon capacity* of the graph G. The previous consideration shows that $\Theta(G) \geqslant \alpha(G)$ and that, in general, equality does not hold.

The determination of the Shannon capacity is a very difficult problem even for very simple small graphs. Shannon proved that $\alpha(G)=\theta(G)$ for those graphs which can be covered by $\alpha(G)$ cliques (the best known such graphs are the so-called perfect graphs; see [1]). However, even for the simplest graph not covered by this result—the pentagon—the Shannon capacity was previously unknown.

Manuscript received February 23, 1978; revised June 20, 1978.

The author is with the Department of Combinatorics, University of Waterloo, Waterloo, ON, Canada, on leave from the Bolyai Institute, József Attila University, H-6720 Szeged, Aradi vértanúk t. 1, Hungary.

A general upper bound on $\Theta(G)$ was also given in [6] (this bound was discussed in detail by Rosenfeld [5]). We assign nonnegative weights $w(x)$ to the vertices x of G such that

$$\sum_{x\in C} w(x) \leqslant 1$$

for every complete subgraph C in G; such an assignment is called a *fractional vertex packing*. The maximum of $\Sigma_x w(x)$, taken over all fractional vertex packings, is denoted by $\alpha^*(G)$. It follows easily from the duality theorem of linear programming that $\alpha^*(G)$ can be defined dually as follows: we assign nonnegative weights $q(C)$ to the cliques C of G such that

$$\sum_{C\ni x} q(C) \geqslant 1$$

for each point x of G and minimize $\Sigma_C q(C)$.

With this notation Shannon's theorem states

$$\Theta(G) \leqslant \alpha^*(G).$$

For the case of the pentagon, this result and the remark above yield the bounds

$$\sqrt{5} \leqslant \Theta(C_5) \leqslant 5/2.$$

We shall prove that the lower bound is the precise value. This will be achieved by deriving a general upper bound on $\Theta(G)$. This upper bound is well characterized and in a sense easily computable. Our methods will enable us to determine or estimate the capacity of other graphs as well. For example, the Petersen graph has capacity four.

II. The Capacity of the Pentagon

Let G be a finite undirected graph without loops. We say that two vertices of G are *adjacent* if they are either connected by an edge or are equal.

The set of points of the graph G is denoted by $V(G)$. The *complementary graph* of G is defined as the graph $\overline{G}$ with $V(\overline{G})=V(G)$ and in which two points are connected by an edge iff they are not connected in G. A *k-coloration* of G is a partition of $V(G)$ into k sets independent in G. Note that this corresponds to a covering of the points of the complementary graph by k cliques. The least k for which G admits a k-coloration is called its *chromatic number*.

A permutation of $V(G)$ is an *automorphism* if it preserves adjacency of the points. The automorphisms of G

0018-9448/79/0100-0001$00.75

form a permutation group called the *automorphism group* of G. If for each pair of points $x, y \in V(G)$ there exists an automorphism mapping x onto y, then the automorphism group is called *vertex transitive*. *Edge transitivity* is defined in an analog manner. A graph is called *regular* of degree d if each point is incident with d edges. Note that graphs whose automorphism groups are vertex transitive are regular. This does not necessarily hold for edge transitivity (as, for example, in the case of a star).

If G and H are two graphs, then their *strong product* $G \cdot H$ is defined as the graph with $V(G \cdot H) = V(G) \times V(H)$, in which (x,y) is adjacent to (x',y') iff x is adjacent to x' in G and y is adjacent to y' in H. If we denote by G^k the strong product of k copies of G, then $\alpha(G^k)$ is indeed the maximum number of independent points in G^k.

We shall use linear algebra extensively. For various properties of (mostly semidefinite) matrices, see, for example, [4]. All vectors will be column vectors. We shall denote by I the identity matrix, by J the square matrix all of whose entries are ones, and by j the vector whose entries are ones (the dimension of these matrices and vectors will be clear from the context).

Besides the inner product of vectors v, w (denoted by $v^T w$, where T denotes transpose), we shall use the *tensor product*, defined as follows. If $v = (v_1, \cdots, v_n)$ and $w = (w_1, \cdots, w_m)$, then we denote by $v \circ w$ the vector $(v_1 w_1, \cdots, v_1 w_m, v_2 w_1, \cdots, v_n w_m)^T$ of length nm. A simple computation shows that the two kinds of vector multiplication are connected by

$$(x \circ y)^T (v \circ w) = (x^T v)(y^T w). \tag{1}$$

Let G be a graph. For simplicity we shall always assume that its vertices are $1, \cdots, n$. An *orthonormal representation* of G is a system $(v_1, \cdots, v_n)$ of unit vectors in a Euclidean space such that if i and j are nonadjacent vertices, then v_i and v_j are orthogonal. Clearly, every graph has an orthonormal representation, for example, by pairwise orthogonal vectors.

Lemma 1: Let $(u_1, \cdots, u_n)$ and $(v_1, \cdots, v_m)$ be orthonormal representations of G and H, respectively. Then the vectors $u_i \circ v_j$ form an orthonormal representation of $G \cdot H$.

The proof is immediate from (1).

Define the *value* of an orthonormal representation $(u_1, \cdots, u_n)$ to be

$$\min_{c} \max_{1 \leq i \leq n} \frac{1}{(c^T u_i)^2}$$

where c ranges over all unit vectors. The vector c yielding the minimum is called the *handle* of the representation. Let $\vartheta(G)$ denote the minimum value over all representations of G. It is easy to see that this minimum is attained. Call a representation *optimal* if it achieves this minimum value.

Lemma 2: $\vartheta(G \cdot H) \leq \vartheta(G)\vartheta(H)$.

Proof: Let $(u_1, \cdots, u_n)$ and $(v_1, \cdots, v_m)$ be optimal orthonormal representations of G and H, with handles c and d, respectively. Then $c \circ d$ is a unit vector by (1), and hence

$$\vartheta(G \circ H) \leq \max_{i,j} \frac{1}{\left((c \circ d)^T (u_i \circ v_j)\right)^2} = \max_{i,j} \frac{1}{(c^T u_i)^2} \cdot \frac{1}{(d^T v_j)^2} = \vartheta(G)\vartheta(H).$$

Remark: We shall see later that equality holds in Lemma 2.

Lemma 3: $\alpha(G) \leq \vartheta(G)$.

Proof: Let $(u_1, \cdots, u_n)$ be an optimal orthonormal representation of G with handle c. Let $\{1, \cdots, k\}$, for example, be a maximum independent set in G. Then $u_1, \cdots, u_k$ are pairwise orthogonal, and so

$$1 = c^2 \geq \sum_{i=1}^{k} (c^T u_i)^2 \geq \alpha(G)/\vartheta(G).$$

Theorem 1: $\Theta(G) \leq \vartheta(G)$.

Proof: By Lemmas 1 and 2, $\alpha(G^k) \leq \vartheta(G^k) \leq \vartheta(G)^k$.

Theorem 2: $\Theta(C_5) = \sqrt{5}$.

Proof: Consider an umbrella whose handle and five ribs have unit length. Open the umbrella to the point where the maximum angle between the ribs is $\pi/2$. Let u_1, u_2, u_3, u_4, u_5 be the ribs and c be the handle, as vectors oriented away from their common point. Then $u_1, \cdots, u_5$ is an orthonormal representation of C_5. Moreover, it is easy to compute from the spherical cosine theorem that $c^T u_i = 5^{-1/4}$, and hence

$$\Theta(C_5) \leq \vartheta(C_5) \leq \max_{i} \frac{1}{(c^T u_i)^2} = \sqrt{5}.$$

The opposite inequality is known, and hence the theorem follows.

III. Formulas for $\vartheta(G)$

To be able to apply Theorem 1 to estimate or calculate the Shannon capacity of other graphs we must investigate the number $\vartheta(G)$ in greater detail.

Theorem 3: Let G be a graph on vertices $\{1, \cdots, n\}$. Then $\vartheta(G)$ is the minimum of the largest eigenvalue of any symmetric matrix $(a_{ij})_{i,j=1}^{n}$ such that

$$a_{ij} = 1, \qquad \text{if } i = j \text{ or if } i \text{ and } j \text{ are nonadjacent.} \tag{2}$$

Proof:

1) Let $(u_1, \cdots, u_n)$ be an optimal orthonormal representation of G with handle c. Define

$$a_{ij} = 1 - \frac{u_i^T u_j}{(c^T u_i)(c^T u_j)}, \qquad i \neq j,$$

$$a_{ii} = 1,$$

and

$$A = (a_{ij})_{i,j=1}^{n}.$$

Then (2) is satisfied. Moreover,

$$-a_{ij}=\left(c-\frac{u_i}{(c^Tu_i)}\right)^T\left(c-\frac{u_j}{(c^Tu_j)}\right),\qquad i\neq j,$$

and

$$\vartheta(G)-a_{ii}=\left(c-\frac{u_i}{c^Tu_i}\right)^2+\left(\vartheta(G)-\frac{1}{(c^Tu_i)^2}\right).$$

These equations imply that $\vartheta(G)I-A$ is positive semidefinite, and hence the largest eigenvalue of A is at most $\vartheta(G)$.

2) Conversely, let $A=(a_{ij})$ be any matrix satisfying (2), and let λ be its largest eigenvalue. Then $\lambda I-A$ is positive semidefinite, and hence there exist vectors $x_1,\cdots,x_n$ such that

$$\lambda\delta_{ij}-a_{ij}=x_i^Tx_j.$$

Let c be a unit vector perpendicular to $x_1,\cdots,x_n$, and set

$$u_i=\frac{1}{\sqrt{\lambda}}(c+x_i).$$

Then

$$u_i^2=\frac{1}{\lambda}(1+x_i^2)=1,\qquad i=1,\cdots,n,$$

and for nonadjacent i and j,

$$u_i^Tu_j=\frac{1}{\lambda}(1+x_i^Tx_j)=0.$$

So $(u_1,\cdots,u_n)$ is an orthonormal representation of G. Moreover,

$$\frac{1}{(c^Tu_i)^2}=\lambda,\qquad i=1,\cdots,n,$$

and hence $\vartheta(G)\leqslant\lambda$. This completes the proof of the theorem.

Note that it also follows that among the optimal representations there is one such that

$$\vartheta(G)=\frac{1}{(c^Tu_1)^2}=\cdots=\frac{1}{(c^Tu_n)^2}.$$

The next theorem gives a good characterization of the value $\vartheta(G)$.

Theorem 4: Let G be a graph on the set of vertices $\{1,\cdots,n\}$, and let $B=(b_{ij})_{i,j=1}^n$ range over all positive semidefinite symmetric matrices such that

$$b_{ij}=0 \tag{3}$$

for every pair (i,j) of distinct adjacent vertices and

$$\operatorname{Tr} B=1. \tag{4}$$

Then

$$\vartheta(G)=\max_B \operatorname{Tr} BJ.$$

Note that Tr BJ is the sum of the entries in B.

Proof:

1) Let $A=(a_{ij})_{i,j=1}^n$ be a matrix satisfying (2) with largest eigenvalue $\vartheta(G)$, and let B be any symmetric matrix satisfying (3) and (4). Then using (2) and (3),

$$\operatorname{Tr} BJ=\sum_{i=1}^n\sum_{j=1}^n b_{ij}=\sum_{i=1}^n\sum_{j=1}^n a_{ij}b_{ij}=\operatorname{Tr} AB,$$

and so

$$\vartheta(G)-\operatorname{Tr} BJ=\operatorname{Tr}(\vartheta(G)I-A)B.$$

Here both $\vartheta(G)I-A$ and B are positive semidefinite. Let $e_1,\cdots,e_n$ be a set of mutually orthogonal eigenvectors of B, with corresponding eigenvalues $\lambda_1,\cdots,\lambda_n\geqslant 0$. Then

$$\operatorname{Tr}(\vartheta(G)I-A)B=\sum_{i=1}^n e_i^T(\vartheta(G)I-A)Be_i=\sum_{i=1}^n\lambda_ie_i^T(\vartheta(G)I-A)e_i\geqslant 0.$$

2) We have to construct a matrix B which satisfies the previous inequality with equality. For this purpose let $(i_1,j_1),\cdots,(i_m,j_m)(i_k<j_k)$ be the edges of G. Consider the $(m+1)$-dimensional vectors

$$\hat{h}=\left(h_{i_1}h_{j_1},\cdots,h_{i_m}h_{j_m},\left(\sum h_i\right)^2\right)^T$$

where $h=(h_1,\cdots,h_n)$ ranges through all unit vectors and

$$z=(0,0,\cdots,0,\vartheta(G))^T.$$

Claim: z is in the convex hull of the vectors $\hat{h}$. Suppose this is not the case. Since the vectors $\hat{h}$ form a compact set, there exists a hyperplane separating z from all the $\hat{h}$, i.e., there exists a vector a and a real number α such that $a^T\hat{h}\leqslant\alpha$ for all unit vectors h but $a^Tz>\alpha$.

Set

$$a=(a_1,\cdots,a_m,y)^T.$$

Then in particular $a^T\hat{h}\leqslant\alpha$, for $h=(1,0,\cdots,0)$; whence $y\leqslant\alpha$. On the other hand, $a^Tz>\alpha$ implies $\vartheta(G)y>0$. Hence $y>0$, and $\alpha>0$. We may suppose that $y=1$, and so $\alpha<\vartheta(G)$.

Now define

$$a_{ij}=\begin{cases}\frac{1}{2}a_k+1, & \text{if } \{i,j\}=\{i_k,j_k\}\\ 1, & \text{otherwise;}\end{cases}$$

then $a^T\hat{h}\leqslant\alpha$ can be written as

$$\sum_{i=1}^n\sum_{j=1}^n a_{ij}h_ih_j\leqslant\alpha.$$

Since the largest eigenvalue of $A=(a_{ij})$ is equal to

$$\max\{x^TAx:|x|=1\},$$

this implies that the largest eigenvalue of (a_{ij}) is at most α. Since (a_{ij}) satisfies (2), this implies $\vartheta(G)\leqslant\alpha$, a contradiction. This proves the claim.

By the claim, there exist a finite number of unit vectors $h_1,\cdots,h_N$ and nonnegative reals $\alpha_1,\cdots,\alpha_N$ such that

$$\alpha_1+\cdots+\alpha_N=1 \tag{5}$$

$$\alpha_1\hat{h}_1+\cdots+\alpha_N\hat{h}_N=z. \tag{6}$$

Set

$$h_p=(h_{p,1},\cdots,h_{p,n})^T$$
$$b_{ij}=\sum_{p=1}^{N}\alpha_p h_{pi}h_{pj}$$
$$B=(b_{ij}).$$

The matrix B is clearly symmetric and positive semidefinite. Further, (6) implies

$$b_{i_kj_k}=0,\qquad k=1,\cdots,m$$

and

$$\mathrm{Tr}\,BJ=\vartheta(G)$$

while (5) implies

$$\mathrm{Tr}\,B=1.$$

This completes the proof.

Lemma 4: Let $(\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n)$ be an orthonormal representation of G and $(\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n)$ be an orthonormal representation of the complementary graph $\overline{G}$. Moreover, let $\boldsymbol{c}$ and $\boldsymbol{d}$ be any vectors. Then

$$\sum_{i=1}^{n}(\boldsymbol{u}_i^T\boldsymbol{c})^2(\boldsymbol{v}_i^T\boldsymbol{d})^2\leqslant \boldsymbol{c}^2\boldsymbol{d}^2.$$

Proof: By (1), the vectors $\boldsymbol{u}_i\circ\boldsymbol{v}_i$ satisfy

$$(\boldsymbol{u}_i\circ\boldsymbol{v}_i)(\boldsymbol{u}_j\circ\boldsymbol{v}_j)=(\boldsymbol{u}_i^T\boldsymbol{u}_j)(\boldsymbol{v}_i^T\boldsymbol{v}_j)=\delta_{ij}.$$

Thus they form an orthonormal system, and we have

$$(\boldsymbol{c}\circ\boldsymbol{d})^2\geqslant\sum_{i=1}^{n}\left((\boldsymbol{c}\circ\boldsymbol{d})^T(\boldsymbol{u}_i\circ\boldsymbol{v}_i)\right)^2$$

which is just the inequality in Lemma 4.

Corollary 1: If $(\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n)$ is an orthonormal representation of $\overline{G}$ and $\boldsymbol{d}$ is any unit vector, then

$$\vartheta(G)\geqslant\sum_{i=1}^{n}(\boldsymbol{v}_i^T\boldsymbol{d})^2.$$

Corollary 2: $\vartheta(G)\vartheta(\overline{G})\geqslant n.$

We give now another minimax formula for the value $\vartheta(G)$, which shows a very surprising duality between G and its complementary graph $\overline{G}$.

Theorem 5: Let $(\boldsymbol{v}_1,\cdots,\boldsymbol{v}_m)$ range over all orthonormal representations of $\overline{G}$ and $\boldsymbol{d}$ over all unit vectors. Then

$$\vartheta(G)=\max\sum_{i=1}^{n}(\boldsymbol{d}^T\boldsymbol{v}_i)^2.$$

Proof: By Corollary 1 we already know that the inequality $\geqslant$ holds. We construct now a representation of $\overline{G}$ and a unit vector $\boldsymbol{d}$ with equality. Let $B=(b_{ij})$ be a positive semidefinite symmetric matrix satisfying (3) and (4) such that $\mathrm{Tr}\,BJ=\vartheta(G)$. Since B is positive semidefinite, we have vectors $\boldsymbol{w}_1,\cdots,\boldsymbol{w}_n$ such that

$$b_{ij}=\boldsymbol{w}_i^T\boldsymbol{w}_j. \tag{7}$$

Note that

$$\sum_{i=1}^{n}\boldsymbol{w}_i^2=1,\qquad\left(\sum_{i=1}^{n}\boldsymbol{w}_i\right)^2=\vartheta(G).$$

Set

$$\boldsymbol{v}_i=\boldsymbol{w}_i/|\boldsymbol{w}_i|\qquad \boldsymbol{d}=\left(\sum_{i=1}^{n}\boldsymbol{w}_i\right)\Big/\left|\sum_{i=1}^{n}\boldsymbol{w}_i\right|.$$

Then the vectors $\boldsymbol{v}_i$ form an orthonormal representation of $\overline{G}$ by (7) and (3). Moreover, using the Cauchy–Schwarz inequality we get

$$\sum_{i=1}^{n}(\boldsymbol{d}^T\boldsymbol{v}_i)^2=\left(\sum_{i=1}^{n}\boldsymbol{w}_i^2\right)\left(\sum_{i=1}^{n}(\boldsymbol{d}^T\boldsymbol{v}_i)^2\right)$$
$$\geqslant\left(\sum_{i=1}^{n}|\boldsymbol{w}_i|(\boldsymbol{d}^T\boldsymbol{v}_i)\right)^2=\left(\sum_{i=1}^{n}\boldsymbol{d}^T\boldsymbol{w}_i\right)^2$$
$$=\left(\boldsymbol{d}\sum_{i=1}^{n}\boldsymbol{w}_i\right)^2=\left(\sum_{i=1}^{n}\boldsymbol{w}_i\right)^2=\vartheta(G).$$

This completes the proof.

Note that since we have equality in the Cauchy–Schwarz inequality, it also follows that

$$(\boldsymbol{d}\boldsymbol{v}_i)^2=\vartheta(G)\boldsymbol{w}_i^2=\vartheta(G)b_{ii}. \tag{8}$$

Theorem 6: Let A range over all matrices such that $a_{ij}=0$ if i,j are adjacent in G, and let $\lambda_1(A)\geqslant\cdots\geqslant\lambda_n(A)$ denote the eigenvalues of A. Then

$$\vartheta(G)=\max_A\left\{1-\frac{\lambda_1(A)}{\lambda_n(A)}\right\}.$$

Proof:

1) Let A be any matrix such that $a_{ij}=0$ if i and j are adjacent. Let $\boldsymbol{f}=(f_1,\cdots,f_n)^T$ be an eigenvector belonging to $\lambda_1(A)$ such that $\boldsymbol{f}^2=-1/\lambda_n(A)$ (note that since $\mathrm{Tr}\,A=0$, the least eigenvalue of A is negative). Consider the matrices $F=\mathrm{diag}\,(f_1,\cdots,f_n)$ and

$$B=F(A-\lambda_n(A)I)F.$$

Obviously B is positive semidefinite. Moreover, $b_{ij}=0$ if i and j are distinct adjacent points, and

$$\mathrm{Tr}\,B=-\lambda_n(A)\,\mathrm{Tr}\,F^2=1.$$

So by Theorem 4,

$$\vartheta(G)\geqslant\mathrm{Tr}\,BJ=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}f_if_j-\lambda_n(A)\sum_{i=1}^{n}f_i^2$$
$$=\sum_{i=1}^{n}\left\{\lambda_1(A)f_i^2-\lambda_n(A)f_i^2\right\}=1-\frac{\lambda_1(A)}{\lambda_n(A)}.$$

2) The fact that equality is attained here follows by a more or less straightforward inversion of this argument and is omitted.

Corollary 3: (See Hoffman [3].) Let $\lambda_1\geqslant\cdots\geqslant\lambda_n$ be the eigenvalues of the adjacency matrix of a graph G. Then the chromatic number of G is at least

$$1-\frac{\lambda_1}{\lambda_n}.$$

Proof: The chromatic number of G is least $\vartheta(\overline{G})$. In fact, if $(\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n)$ is an orthonormal representation of G,

c is any unit vector, and $J_1, \cdots, J_k$ are the color classes in any k-coloration of G, then

$$\sum_{i=1}^{n} (c^T u_i)^2 = \sum_{m=1}^{k} \sum_{i \in J_m} (c^T u_i)^2 \leqslant \sum_{m=1}^{k} 1 = k$$

from which the assertion follows by Theorem 5. Now the adjacency matrix of G satisfies the condition in the theorem (with $\overline{G}$ instead of G), which implies the inequality in the corollary.

IV. Some Further Properties of $\vartheta(G)$

The results in the previous section make the value $\vartheta(G)$ quite easy to handle. Let us derive some consequences.

Theorem 7: $\vartheta(G \cdot H) = \vartheta(G)\vartheta(H)$.

Proof: We already know that

$$\vartheta(G \cdot H) \leqslant \vartheta(G)\vartheta(H).$$

To show the opposite inequality, let $(v_1, \cdots, v_n)$ be an orthonormal representation of $\overline{G}$, $(w_1, \cdots, w_m)$ be an orthonormal representation of $\overline{H}$, and c, d be unit vectors such that

$$\sum_{i=1}^{n} (v_i^T c)^2 = \vartheta(G) \qquad \sum_{i=1}^{m} (w_i^T d)^2 = \vartheta(H).$$

Then $v_i \circ w_j$ is an orthonormal representation of $\overline{G \cdot H}$ (this follows since it is an orthonormal representation of $\overline{G} \cdot \overline{H}$ and $\overline{G \cdot H} \supseteq \overline{G} \cdot \overline{H}$). Moreover, $c \circ d$ is a unit vector. So

$$\begin{aligned} \vartheta(G \cdot H) &\geqslant \sum_{i=1}^{n} \sum_{j=1}^{m} \left((v_i \circ w_j)^T (c \circ d)\right)^2 \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} (v_i^T c)^2 (w_j^T d)^2 \\ &= \sum_{i=1}^{n} (v_i^T c)^2 \sum_{j=1}^{m} (w_j^T d)^2 = \vartheta(G)\vartheta(H). \end{aligned}$$

Theorem 8: If G has a vertex-transitive automorphism group, then

$$\vartheta(G)\vartheta(\overline{G}) = n.$$

Corollary 4: If G has a vertex-transitive automorphism group, then

$$\Theta(G)\Theta(\overline{G}) \leqslant n.$$

Note that Theorem 8 and its corollary do not hold for all graphs because there are graphs with $\alpha(G)\alpha(\overline{G}) > n$ (for example, a star).

Proof: Let Γ be the automorphism group of G. We may consider the elements of Γ as $n \times n$ permutation matrices. Let $B = (b_{ij})$ be a matrix satisfying (3) and (4) such that $\operatorname{Tr} BJ = \vartheta(G)$. Consider

$$\overline{B} = (\bar{b}_{ij}) = \frac{1}{|\Gamma|} \left(\sum_{P \in \Gamma} P^{-1} B P \right).$$

Then trivially, $\overline{B}$ also satisfies (3), and

$$\operatorname{Tr} \overline{B} = 1 \qquad \operatorname{Tr} \overline{B} J = \vartheta(G)$$

(using $PJ = JP = J$). Also trivially, $\overline{B}$ is symmetric and positive semidefinite and satisfies $P^{-1}\overline{B}P = \overline{B}$, for all $P \in \Gamma$. Since Γ is transitive on the vertices, this implies $\bar{b}_{ii} = 1/n$, for all i. Constructing the orthonormal representation $(v_1, \cdots, v_n)$ and the unit vector d as in the proof of Theorem 5, we have

$$(d^T v_i)^2 = \frac{\vartheta(G)}{n}$$

by (8). So from the definition of $\vartheta(\overline{G})$,

$$\vartheta(\overline{G}) \leqslant \max_{1 \leqslant i \leqslant n} \frac{1}{(d^T v_i)^2} = \frac{n}{\vartheta(G)},$$

and hence

$$\vartheta(G)\vartheta(\overline{G}) \leqslant n.$$

Since we already know that the opposite inequality holds (Corollary 2), Theorem 8 is proved.

Theorem 9: Let G be a regular graph, and let $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ be the eigenvalues of its adjacency matrix A. Then

$$\vartheta(G) \leqslant \frac{-n\lambda_n}{\lambda_1 - \lambda_n}.$$

Equality holds if the automorphism group of G is transitive on the edges.

Corollary 5: For odd n,

$$\vartheta(C_n) = \frac{n \cos(\pi/n)}{1 + \cos(\pi/n)}.$$

Proof: Consider the matrix $J - xA$, where x will be chosen later. This satisfies condition (2) in Theorem 3, and hence its largest eigenvalue is at least $\vartheta(G)$. Let v_i denote the eigenvector of A belonging to λ_i. Then since A is regular, $v_1 = j$, and therefore, $j, v_2, \cdots, v_n$ are also eigenvectors of J. So the eigenvalues of $J - xA$ are $n - x\lambda_1, -x\lambda_2, \cdots, -x\lambda_n$. The largest of these is either the first or the last, and the optimal choice of x is $x = n/(\lambda_1 - \lambda_n)$ when they are both equal to $-n\lambda_n/(\lambda_1 - \lambda_n)$. This proves the first assertion.

Assume now that the automorphism group Γ of G is transitive on the edges. Let $C = (c_{ij})$ be a symmetric matrix such that $c_{ij} = 1$ if i and j are equal or nonadjacent and having largest eigenvalue $\vartheta(G)$. As in the proof of Theorem 8, consider

$$\overline{C} = \frac{1}{|\Gamma|} \sum_{P \in \Gamma} P^{-1} C P.$$

Then $\overline{C}$ also satisfies (2), and moreover, its largest eigenvalue is at most $\vartheta(G)$. By Theorem 3, it is equal to $\vartheta(G)$. Moreover, $\overline{C}$ is clearly of the form $J - xA$. Hence the second assertion follows.

V. Comparison with Other Bounds on Capacity

Theorem 10: $\vartheta(G) \leqslant \alpha^*(G)$.

Proof: We use Theorem 4. Let $(\boldsymbol{u}_i)$ be an orthonormal representation of $\overline{G}$ and $\boldsymbol{c}$ be a unit vector such that

$$\vartheta(G)=\sum_{i=1}^{n}\left(\boldsymbol{c}^T\boldsymbol{u}_i\right)^2.$$

Let C be any clique in G. Then $\{\boldsymbol{u}_i : i \in C\}$ is an orthonormal set of vectors, and hence

$$\sum_{i\in C}\left(\boldsymbol{c}^T\boldsymbol{u}_i\right)^2 \leqslant \boldsymbol{c}^2 = 1.$$

Hence the weights $(\boldsymbol{c}^T\boldsymbol{u}_i)^2$ form a fractional vertex packing, and so

$$\vartheta(G)=\sum_{i=1}^{n}\left(\boldsymbol{c}^T\boldsymbol{u}_i\right)^2 \leqslant \alpha^*(G).$$

A very simple upper bound on $\Theta(G)$ is the dimension of an orthonormal representation of G.

Theorem 11: Assume that G admits an orthonormal representation in dimension d. Then

$$\vartheta(G) \leqslant d.$$

Proof: Let $(\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n)$ be an orthonormal representation of G in d-dimensional space. Then $(\boldsymbol{u}_1\circ\boldsymbol{u}_1,\boldsymbol{u}_2\circ\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n\circ\boldsymbol{u}_n)$ is another orthonormal representation of G. Let $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_d\}$ be an orthonormal basis and

$$\boldsymbol{b}=\frac{1}{\sqrt{d}}\left(\boldsymbol{e}_1\circ\boldsymbol{e}_1+\boldsymbol{e}_2\circ\boldsymbol{e}_2+\cdots+\boldsymbol{e}_d\circ\boldsymbol{e}_d\right).$$

Then $\boldsymbol{b}^2=1$, and

$$(\boldsymbol{u}_i\circ u_i)^T\boldsymbol{b}=\frac{1}{\sqrt{d}}\sum_{k=1}^{d}(\boldsymbol{e}_k\circ\boldsymbol{e}_k)^T(\boldsymbol{u}_i\circ\boldsymbol{u}_i) =\frac{1}{\sqrt{d}}\sum_{k=1}^{d}\left(\boldsymbol{e}_k^T\boldsymbol{u}_i\right)^2=\frac{1}{\sqrt{d}}.$$

Therefore $\vartheta(G) \leqslant d$.

VI. Applications

We can use our methods to calculate the Shannon capacity of graphs other than the pentagon. We of course deal only with graphs G such that $\alpha(G)<\alpha^*(G)$, since if $\alpha(G)=\alpha^*(G)$, then $\Theta(G)=\alpha(G)$ by Shannon's theorem.

Theorem 12: If G has a vertex-transitive automorphism group, then $\Theta(G\cdot\overline{G})=|V(G)|$. If, in addition, G is self-complementary, then $\Theta(G)=\sqrt{|V(G)|}$.

Proof: The "diagonal" in $G\cdot\overline{G}$ is independent; hence

$$\Theta(G\cdot\overline{G}) \geqslant \alpha(G\cdot\overline{G}) \geqslant |V(G)|.$$

On the other hand, we have by Theorems 1, 6, and 7 that

$$\Theta(G\cdot\overline{G}) \leqslant \vartheta(G\cdot\overline{G})=\vartheta(G)\vartheta(\overline{G})=|V(G)|.$$

If G is self-complementary, then

$$\Theta(G\cdot\overline{G})=\Theta(G^2)=\Theta(G)^2.$$

This proves the theorem. The proof also shows that in these cases $\Theta=\vartheta$.

Theorem 13: Let $n \geqslant 2r$, and let the graph $K(n,r)$ be defined as the graph whose vertices are the r-subsets of an n-element set S, two subsets being adjacent iff they are disjoint. Then

$$\Theta(K(n,r))=\binom{n-1}{r-1}.$$

Corollary 6: The Petersen graph, which is isomorphic with $K(5,2)$, has capacity four.

Corollary 7: (See Erdös, Ko, and Rado [2].)

$$\alpha(K(n,r))=\binom{n-1}{r-1}.$$

Note that

$$\alpha^*(K(n,r))=\binom{n}{r}\Big/\left[\frac{n}{r}\right]$$

which is larger than $\binom{n-1}{r-1}$ unless r is a divisor of n.

Proof of Theorem 13: The r subsets containing a specified element of S form an independent set of points in $K(n,r)$; hence

$$\Theta(K(n,r)) \geqslant \alpha(K(n,r)) \geqslant \binom{n-1}{r-1}.$$

On the other hand, we calculate $\vartheta(K(n,r))$. Since the automorphism group of $K(n,r)$ is clearly transitive on the vertices and edges, we may use Theorem 9. So let us calculate the eigenvalues of $K(n,r)$. Clearly j is an eigenvector with eigenvalue $\binom{n-r}{r}$.

Let $1 \leqslant t \leqslant r$. For each $T\subset S$ such that $|T|=t$, let x_T be a real number such that for every $U \subset S$ with $|U|=t-1$,

$$\sum_{U\subset T} x_T=0. \tag{9}$$

There are $\binom{n}{t}-\binom{n}{t-1}$ linearly independent vectors (x_T) of this type. For each such vector, define

$$\bar{x}_A=\sum_{\substack{T\subseteq A\\|T|=t}} x_T$$

for every $A\subset S$, $|A|=r$. It is not difficult to see, and actually well-known, that the numbers x_T can be calculated from the numbers $\bar{x}_A$, whence there are $\binom{n}{t}-\binom{n}{t-1}$ linearly independent vectors of type $(\bar{x}_A)$.

Claim: Every $(\bar{x}_A)$ is an eigenvector of the adjacency matrix of $K(n,r)$ with eigenvalue $(-1)^t\binom{n-r-t}{r-t}$. In fact, for any $A_0\subset S$ such that $|A_0|=r$, we have

$$\sum_{A\cap A_0=\varnothing}\bar{x}_A=\sum_{T\cap A_0=\varnothing}\binom{n-r-t}{r-t}x_T=\binom{n-r-t}{r-t}\beta_0.$$

To determine this value we set

$$\beta_i=\sum_{|T\cap A_0|=i} x_T.$$

Then summing (9) for every $U \subset S$ such that $|U| = t-1$ and $|U \cap A_0| = i$, we get

$$(i+1)\beta_{i+1} + (t-i)\beta_i = 0.$$

This may be considered as a recurrence relation for the β_i and yields

$$\beta_i = (-1)^i \binom{t}{i} \beta_0$$

whence

$$\beta_0 = (-1)^t \beta_t = (-1)^t \bar{x}_{A_0}$$

which proves the claim.

By this construction we have found

$$1 + \sum_{t=1}^{r} \left(\binom{n}{t} - \binom{n}{t-1} \right) = \binom{n}{r}$$

linearly independent eigenvectors (there is no problem with the eigenvectors belonging to different values of t since they belong to different eigenvalues). Therefore, we have all eigenvectors, and it follows that the eigenvalues of $K(n,r)$ are the numbers

$$(-1)^t \binom{n-r-t}{r-t}, \qquad t = 0, 1, \cdots, r.$$

So the largest and smallest eigenvalues are $\binom{n-r}{r}$ and $-\binom{n-r-1}{r-1}$, respectively, and Theorem 9 yields

$$\vartheta(K(n,r)) = \frac{\binom{n-r-1}{r-1}\binom{n}{r}}{\binom{n-r}{r} + \binom{n-r-1}{r-1}} = \binom{n-1}{r-1}.$$

VII. Concluding Remarks

The purpose of introducing $\vartheta(G)$ has been to estimate $\Theta(G)$. So the obvious question is as follows.

Problem 1: Is $\vartheta = \Theta$? More modestly, find further graphs with $\vartheta(G) = \Theta(G)$. In particular, do odd circuits satisfy $\vartheta(G) = \Theta(G)$?

This last question pinpoints a difficulty which seems to be crucial. In all cases known to the author where $\Theta(G)$ is precisely determined, there is some k ($k=1$ or 2, in fact) such that $\alpha(G^k) = \Theta(G)^k$. But if $\Theta(G) = \vartheta(G)$ for the seven-circuit, for example, then no such k can exist, since no power of $\vartheta(C_7)$ is an integer.

Various properties of $\vartheta(G)$ established in this paper suggest further problems which would be solved by an affirmative answer to Problem 1.

Problem 2: Is $\Theta(G \cdot H) = \Theta(G)\Theta(H)$? (Note that $\Theta(G \cdot H) \geq \Theta(G)\Theta(H)$ is obvious.)

Problem 3: Is it true that $\Theta(G) \cdot \Theta(\bar{G}) \geq |V(G)|$?

Note that an affirmative answer to Problem 2 would imply an affirmative answer to Problem 3:

$$\Theta(G)\Theta(\bar{G}) = \Theta(G \cdot \bar{G}) \geq \alpha(G \cdot \bar{G}) \geq |V(G)|.$$

This, in turn, would imply an affirmative answer to the last question of Problem 1:

$$n \leq \Theta(C_n)\Theta(\bar{C}_n) \leq \vartheta(C_n)\vartheta(\bar{C}_n) = n;$$

hence $\Theta(C_n) = \vartheta(C_n)$ and $\Theta(\bar{C}_n) = \vartheta(\bar{C}_n)$).

Corollary 7 shows an example where the calculation of $\vartheta(G)$ helps to determine $\alpha(G)$ in a nontrivial way. Are there any further examples?

Acknowledgment

My sincere thanks are due to K. Vesztergombi and M. Rosenfeld for numerous conversations on the topic of this paper—also to T. Nemetz, A. Schrijver, and the referees of the paper for pointing out several errors in, and suggesting many improvements to, the original text. It is a pleasure to acknowledge that A. J. Hoffman and M. Rosenfeld have also found extensions of the original idea (in Section II) to other graphs. Among others Hoffman found Theorem 9, and Rosenfeld found $\vartheta(\bar{C}_n)$.

References

[1] C. Berge, *Graphs and Hypergraphs*. Amsterdam and London North-Holland; New York: American Elsevier, 1973.
[2] P. Erdös, C. Ko, and R. Rado, "Intersection theorems for systems of finite sets," *Quart. J. Math. Oxford*, vol. 12, pp. 313–320, 1961.
[3] A. J. Hoffman, "On eigenvalues and colorings of graphs," in B. Harris, Ed., *Graph Theory and its Applications. New York and London*: *Academic*, 1970, *pp*. 79–91.
[4] P. Lancaster, *Theory of Matrices*. New York and London: Academic, 1969.
[5] M. Rosenfeld, "On a problem of Shannon," *Proc. Amer. Math. Soc.*, vol 18, pp. 315–319, 1967.
[6] C. E. Shannon, "The zero-error capacity of a noisy channel," *IRE Trans. Inform. Thoery*, vol. IT-2, no. 3, pp. 8–19, Sept. 1956.