Now we head for an upper bound "$\Theta(G) \leq \sigma^{-1}$" for the Shannon capacity of any graph G that has an especially "nice" orthonormal representation. For this let $T = \{v^{(1)}, \ldots, v^{(m)}\}$ be an orthonormal representation of G in $\mathbb{R}^s$, where $v^{(i)}$ corresponds to the vertex v_i. We assume in adition that all the vectors $v^{(i)}$ have the same angle ($\neq 90°$) with the vector $u := \frac{1}{m}(v^{(1)} + \ldots + v^{(m)})$, or equivalently that the inner product

$$\langle v^{(i)}, u \rangle = \sigma_T$$

has the same value $\sigma_T \neq 0$ for all i. Let us call this value σ_T the *constant* of the representation T. For the Lovász umbrella that represents C_5 the condition $\langle v^{(i)}, u \rangle = \sigma_T$ certainly holds, for $u = \overrightarrow{OS}$.

Now we proceed in the following three steps.

(**A**) Consider a probability distribution $x = (x_1, \ldots, x_m)$ on V and set

$$\mu(x) := \quad |x_1 v^{(1)} + \ldots + x_m v^{(m)}|^2,$$

and

$$\mu_T(G) := \inf_x \mu(x).$$

Let U be a largest independent set in G with $|U| = a$, and define $x_U = (x_1, \ldots, x_m)$ with $x_i = \frac{1}{\alpha}$ if $v_i \in U$ and $x_i = 0$ otherwise. Since all vectors $v^{(i)}$ have unit length and $\langle v^{(i)}, v^{(j)} \rangle = 0$ for any two non-adjacent vertices, we infer

$$\mu_T(G) \leq \mu(x_U) = |\sum_{i=1}^{m} x_i v^{(i)}|^2 = \sum_{i=1}^{m} {x_i}^2 = \alpha \frac{1}{\alpha^2} = \frac{1}{\alpha}.$$

Thus we have $\mu_T(G) \leq \alpha^{-1}$, and therefore

$$\alpha(G) \leq \frac{1}{\mu_T(G)}.$$

(**B**) Next we compute $\mu_T(G)$. We need the Cauchy-Schwarz inequality

$$\langle a, b \rangle^2 \leq |a|^2 |b|^2$$

for vectors $a, b \in \mathbb{R}^s$. Applied to $a = x_1 v^{(1)} + \ldots + x_m v^{(m)}$ and $b = u$, the inequality yields

$$\langle x_1 v^{(1)} + \ldots + x_m v^{(m)}, u \rangle^2 \leq \mu(x) |u|^2. \tag{1}$$

By our assumption that $\langle v^{(i)}, u \rangle = \sigma_T$ for all i, we have

$$\langle x_1 v^{(1)} + \ldots + x_m v^{(m)}, u \rangle = (x_1 + \ldots + x_m) \sigma_T = \sigma_T$$

for any distribution x. Thus, in particular, this has to hold for the uniform distribution $(\frac{1}{m}, \ldots, \frac{1}{m})$,which implies $|u|^2 = \sigma_T$. Hence (1) reduces to

$$\sigma_T^2 \leq \mu(x) \sigma_T \quad \text{or} \quad \mu_T(G) \geq \sigma_T.$$

On the other hand, for $x = (\frac{1}{m}, \dots, \frac{1}{m})$ we obtain

$$\mu_T(G) \le \mu(x) = |\frac{1}{m}(v^{(1)} + \ldots + v^{(m)})|^2 = |u|^2 = \sigma_T,$$

and so we have proved

$$\mu_T(G) = \sigma_T. \tag{2}$$

In summary, we have established the inequality

$$\alpha(G) \le \frac{1}{\sigma_T} \tag{3}$$

for any orthonormal respresentation T with constant σ_T.

(**C**) To extend this inequality to $\Theta(G)$, we proceed as before. Consider again the product $G \times H$ of two graphs. Let G and H have orthonormal representations R and S in $\mathbb{R}^r$ and $\mathbb{R}^s$, respectively, with constants σ_R and σ_S. Let $v = (v_1, \dots, v_r)$ be a vector in R and $w = (w_1, \dots, w_s)$ be a vector in S. To the vertex in $G \times H$ corresponding to the pair (v, w) we associate the vector

$$vw^T := (v_1w_1, \dots, v_1w_s, v_2w_1, \dots, v_2w_s, \dots, v_rw_1, \dots, v_rw_s) \in \mathbb{R}^{rs}.$$

It is immediately checked that $R \times S := \{vw^T : v \in R, w \in S\}$ is an orthonormal representation of $G \times H$ with constant $\sigma_R\sigma_S$. Hence by (2) we obtain

$$\mu_{R\times S}(G \times H) = \mu_R(G)\mu_S(H).$$

For $G^n = G \times \ldots \times G$ and the representation T with constant σ_T this means

$$\mu_{T^n}(G^n) = \mu_T(G)^n = \sigma_T^n$$

and by (3) we obtain

$$\alpha(G^n) \le \sigma_T^{-n}, \quad \sqrt[n]{\alpha(G^n)} \le \sigma_T^{-1}.$$

Taking all things together we have thus completed Lovász' argument:

Theorem. *whenever $T = \{v^{(1)}, \dots, v^{(m)}\}$) is an orthonormal representation of G with constant σ_T, then*

$$\Theta(G) \le \frac{1}{\sigma_T}. \tag{4}$$

Looking at the Lovász umbrella, we have $u = (0, 0, h = \frac{1}{\sqrt[4]{5}})^T$ and hence $\sigma = \langle v^{(i)}, u\rangle = h^2 = \frac{1}{\sqrt{5}}$, which yields $\Theta(C_5) \le \sqrt{5}$. Thus Shannon's problem is solved.

Let us carry our discussion a little further. We see from (4) that the larger σ_T is for a representation of G, the better a bound for $\Theta(G)$ we will get. Here is a method that gives us an orthonormal representation for *any* graph G. To $G = (V, E)$ we associate the adjacency matrix $A = (a_{ij})$, which is defined as follows: Let $V = \{v_1, \ldots, v_m\}$, then we set

$$a_{ij} := \begin{cases} 1 & \text{if } v_i v_j \in E \\ 0 & \text{otherwise.} \end{cases}$$

A is a real symmetric matrix with 0's in the main diagonal.

Now we need two facts from linear algebra. First, as a symmetric matrix, A has m real eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$ (some of which may be equal), and the sum of the eigenvalues equals the sum of the diagonal entries of A, that is, 0. Hence the smallest eigenvalue must be negative (except in the trivial case when G has no edges). Let $p = |\lambda_m| = -\lambda_m$ be the absolute value of the smallest eigenvalue, and consider the matrix

$$M := 1 + \frac{1}{p}A,$$

where I denotes the $(m \times m)$-identity matrix. This M has the eigenvalues $1 + \frac{\lambda_1}{p} \geq 1 + \frac{\lambda_2}{p} \geq \ldots \geq 1 + \frac{\lambda_m}{p} = 0$. Now we quote the second result (the principal axis theorem of linear algebra): If $M = (m_{ij})$ is a real symmetric matrix with all eigenvalues ≥ 0, then there are vectors $v^{(1)}, \ldots, v^{(m)} \in \mathbb{R}^s$ for $s = \text{rank}(M)$, such that

$$m_{ij} = \langle v^{(i)}, v^{(j)} \rangle \quad (1 \leq i, j \leq m).$$

In particular, for $M = I + \frac{1}{p}A$ we obtain

$$\langle v^{(i)}, v^{(i)} \rangle = m_{ii} = 1 \quad \text{for all } i$$

and

$$\langle v^{(i)}, v^{(j)} \rangle = \frac{1}{p} a_{ij} \quad \text{for } i \neq j.$$

Since $a_{ij} = 0$ whenever $v_i v_j \notin E$, we see that the vectors $v^{(1)}, \ldots, v^{(m)}$ form indeed an orthonormal representation of G.

Let us, finally, apply this construction to the m-cycles C_m for odd $m > 5$. Here one easily computes $p = |\lambda_{min}| = 2\cos\frac{\pi}{m}$ (see the box). Every row of the adjacency matrix contains two 1's, implying that every row of the matrix M sums to $1 + \frac{2}{p}$. For the representation $\{v^{(1)}, \ldots, v^{(m)}\}$ this means

$$\langle v^{(i)}, v^{(1)} + \ldots + v^{(m)} \rangle = 1 + \frac{2}{p} = 1 + \frac{1}{\cos\frac{\pi}{m}}$$

and hence

$$\langle v^{(i)}, u \rangle = \frac{1}{m}(1 + (\cos\frac{\pi}{m})^{-1}) = \sigma$$

for all i. We can therefore apply our main result (4) and conclude

$$\Theta(C_m) \leq \frac{m}{1 + (\cos\frac{\pi}{m})^{-1}} \qquad (\text{for } m \geq 5 \text{ odd}). \tag{5}$$

Notice that because of $\cos\frac{\pi}{m} < 1$ the bound (5) is better than the bound $\Theta(C_m) \le \frac{m}{2}$ we found before. Note further $\cos\frac{\pi}{5} = \frac{\tau}{2}$, where $\tau = \frac{\sqrt{5}+1}{2}$ is the golden section. Hence for $m = 5$ we again obtain

$$\Theta(C_5) \le \frac{5}{1 + \frac{4}{\sqrt{5}+1}} = \frac{5(\sqrt{5}+1)}{5+\sqrt{5}} = \sqrt{5}.$$

The orthonormal representation given by this construction is, of course, precisely the "Lovász umbrella."
And what about C_7, C_9, and the other odd cycles? By considering $\alpha(C_m^2)$, $\alpha(C_m^3)$ and other small powers the lower bound $\frac{m-1}{2} \le \Theta(C_m)$) can certainly be increased, but for no odd $m \ge 7$ do the best known lower bounds agree with the upper bound given in (4). So, twenty years after Lovász' marvelous proof of $\Theta(C_5) = \sqrt{5}$, these problems remain open and are considered very difficult — but after all we had this situation before.

The eigenvalues of C_m

Look at the adjacency matrix A of the cycle C_m. To find the eigenvalues (and eigenvectors) we use the m-th roots of unity. These are given by $1, \zeta, \zeta^2, \ldots, \zeta^{m-1}$ for $\zeta = e^{\frac{2\pi i}{m}}$ — see the box on page 25.
Let $\lambda = \zeta^k$ be any of these roots, then we claim that $(1, \lambda, \lambda^2, \ldots, \lambda^{m-1})^T$ is an eigenvector of A to the eigenvalue $\lambda + \lambda^{-1}$. In fact, by the set-up of A we find

$$A \begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \\ \vdots \\ \lambda^{m-1} \end{pmatrix} = \begin{pmatrix} \lambda + \lambda^{m-1} \\ \lambda^2 + 1 \\ \lambda^3 + \lambda \\ \vdots \\ 1 + \lambda^{m-2} \end{pmatrix} = (\lambda + \lambda^{-1}) \begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \\ \vdots \\ \lambda^{m-1} \end{pmatrix}$$

Since the vectors $(1, \lambda, \ldots, \lambda^{m-1})$ are independent (they form a so-called Vandermonde matrix) we conclude that for odd m

$$\begin{aligned} \zeta^k + \zeta^{-k} &= [(\cos(2k\pi/m)) + i\sin(2k\pi/m)] \\ &\quad + [\cos(2k\pi/m) - i\sin(2k\pi/m)] \\ &= 2\cos(2k\pi/m) \qquad (0 \le k \le \tfrac{m-1}{2}) \end{aligned}$$

are all the eigenvalues of A. Now the cosine is a decreasing function, and So

$$2\cos(\frac{(m-1)\pi}{m}) = -2\cos\frac{\pi}{m}$$

is the smallest eigenvalue of A.

References

[1] V. CHVÁTAL. *Linear Programming.* Freeman, New York, 1983.

[2] W. HAEMERS. Eigenvalue methods. *Packing and covering in combinatorics (A. Schrijver ed.), Math. Centre Tract*, 106:15–38, 1979.

[3] L. LOVÁSZ. On the shannon capacity of a graph. *IEEE Trans. Information theory*, 25(1):1–7, 1979.

[4] C. E. SHANNON. *The zero-error capacity of a noisy channel*, volume 3, pages 3–15. IRE Trans. Information Theory, 1956.