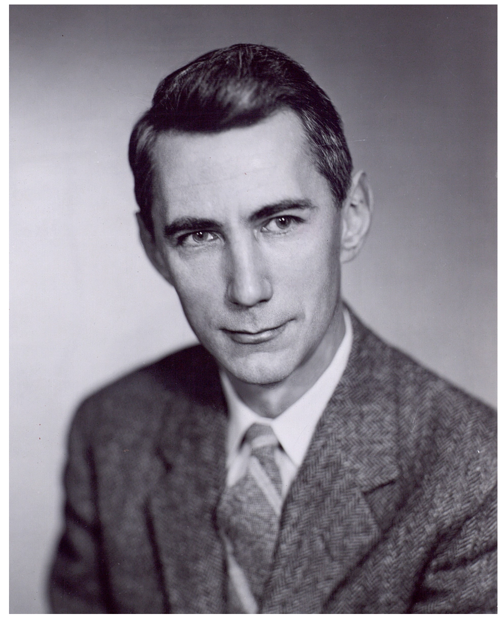
Chapter 33

Communicating without errors

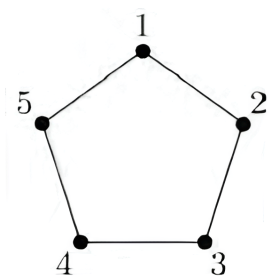


Claude Shannon

In 1956 Claude Shannon, the founder of information theory, posed the following very interesting question:

> *Suppose we want to transmit messages across a channel (where some symbols may be distorted) to a receiver: What is the maximum rate of transmission such that the receiver may recover the original message without errors?*

Let us see what Shannon meant by "channel" and "rate of transmission." We are given a set V of symbols, and a message is just a string of symbols from V. We model the channel as a graph $G = (V, E)$ where V is the set of symbols, and E the set of edges between unreliable pairs of symbols, that is, symbols which may be confused during transmission. For example, communicating over a phone in everyday language, we connnect the symbols B and P by an edge since the receiver may not be able to distinguish them. Let us call G the *confusion graph.* The 5-cycle C_5 will play a prominent role in our discussion. In this example, 1 and 2 may be confused, but not 1 and 3, etc. Ideally we would like to use all 5 symbols for transmission, but since we want to communicate error-free we can if we only send single symbols use only one letter from each pair that might be confused. Thus for the 5-cycle we can use only two different letters (any two that are not connected by an edge). In the language of information theory, this means that for the 5-cycle we achieve an information rate of $\log_2 2 = 1$ (instead of the maximal $\log_2 5 \approx 2.32$). It is clear that in this model, for an arbitrary graph $G = (V, E)$, the best we can do is to transmit symbols from a largest independent set. Thus the information rate, when sending single symbols, is $log_2\alpha(G)$, where $\alpha(G)$ is the *independence number* of G.



Let us see whether we can increase the information rate by using larger strings in place of single symbols. Suppose we want to transmit strings of length 2. The strings ulu2 and vlv, can only be confused if one of the following three cases holds:

- $u_1 = v_1$ and u_2 can be confused with v_2,
- $u_2 = v_2$ and u_1 can be confused with v_1, or
- $u_1 \neq v_1$ can be confused and $u_2 \neq v_2$ can be confused.

In graph-theoretic terms this amounts to considering the product $G_1 \times G_2$ of two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_1, E_1)$. $G_1 \times G_2$ has the vertex set $V_1 \times V_2 = \{(u_1, u_2) : u_1 \in V_1, u_2 \in V_2\}$, with $(u_1, u_2) \neq (v_1, v_2)$ connected by an edge if and only if $u_i = v_i$ or $u_i v_i \in E$ for $i = 1, 2$. The confusion graph for strings of length 2 is thus $G^2 - G \times G$, the product of the confusion graph G for single symbols with itself. The information rate of strings of length 2 *per symbol* is then given by

$$\frac{\log_2 \alpha(G^2)}{2} = \log_2 \sqrt{\alpha(G^2)}.$$

Now, of course, we may use strings of any length n. The n-th confusion graph $G^n = G \times G \times \cdots \times G$ has vertex set $V^n = \{(u_1, \cdots, u_n) : u_i \in V\}$ with $(u_1, \cdots, u_n) \neq (v_1, \cdots, v_n)$ being connected by an edge if $u_i = v_i$ or $u_i v_i \in E$ for all i. The rate of information per symbol determined by strings of length n is

$$\frac{\log_2 \alpha(G^n)}{n} = \log_2 \sqrt[n]{\alpha(G^n)}.$$

What can we say about $\alpha(G^n)$? Here is a first observation. Let $U \subseteq V$ be a largest independent set in G, $|U| = \alpha$. The an vertices in G^n of the form $(u_1, \cdots, u_n)$, $u_i \in U$ for all i, clearly form an independent set in G^n. Hence

$$\alpha(G^n) \geq \alpha(G)^n$$

and therefore

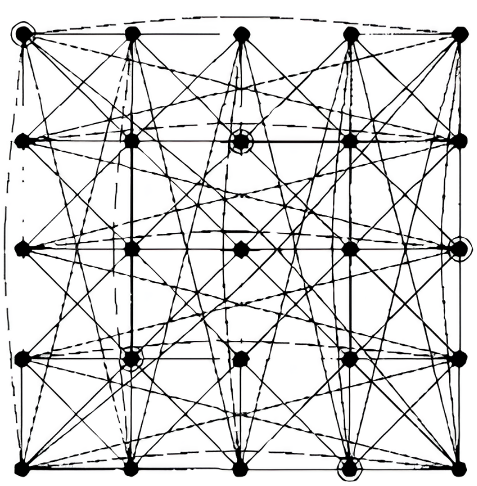
$$\sqrt[n]{\alpha(G^n)} \geq \alpha(G),$$

meaning that we never decrease the information rate by using longer strings instead of single symbols. This, by the way, is a basic idea of coding theory: By encoding symbols into longer strings we can make error-free communi- cation more efficient. Disregarding the logarithm we thus arrive at Shannon's fundamental definition: The *zero-error capacity* of a graph G is given by

$$\Theta(G) := \sup_{n \geq 1} \sqrt[n]{\alpha(G^n)},$$

and Shannon's problem was to compute $\Theta(G)$, and in particular $\Theta(C_5)$. Let us look at C_5. So far we know $\alpha(C_5) = 2 \leq \Theta(C_5)$. Looking at the 5-cycle as depicted earlier, or at the product $C_5 x C_5$ as drawn on the left, we see that the set $\{(1,1), (2,3), (3,5), (4,2), (5,4)\}$ is independent in C_5^2. Thus we have $\alpha(C_5^2) \geq 5$. Since an independent set can contain only two vertices from any two consecutive rows we see that $\alpha(C_5^2) = 5$. Hence, by using strings of length 2 we have increased the lower bound for the capacity to $\Theta(C_5) \geq \sqrt{5}$.

So far we have no upper bounds for the capacity. To obtain such bounds we again follow Shannon's original ideas. First we need the dual definition of an independent set. We recall that a subset $C \subseteq V$ is a *clique* if any two vertices of C are joined by an edge. Thus the vertices form trivial



The graph $C_5 \times C_6$.

cliques of size 1, the edges are the cliques of size 2, the triangles are cliques of size 3, and so on. Let C be the set of cliques in G. Consider an arbitrary probability distribution $x = (x, : v \in V)$ on the set of vertices, that is, $x_v \geq 0$ and $\sum_{v \in V} x_v = 1$. To every distribution x we associate the "maximal value of a clique"

$$\lambda(x) = \max_{C \in \mathcal{C}} \sum_{v \in C} x_v.$$

and finally we set

$$\lambda(G) = \min_x \lambda(x) = \min_x \max_{C \in \mathcal{C}} \sum_{v \in C} x_v.$$

To be precise we should use inf instead of min, but the minimum exists because $\lambda(x)$ is continuous on the compact set of all distributions. Consider now an independent set $U \subseteq V$ of maximal size $\alpha(G) = \alpha$. Associated to U we define the distribution $x_U = (x_v : v \in V)$ by setting $x_v = \frac{1}{\alpha}$ if $v \in U$ and $x_v = 0$ otherwise. Since any clique contains at most one vertex from U, we infer $\lambda(x_U) = \frac{1}{\alpha}$, and thus by the definition of $\lambda(G)$

$$\lambda(G) \leq \frac{1}{\alpha(G)} \quad \text{or} \quad \alpha(G) \leq \lambda(G)^{-1}.$$

What Shannon observed is that $\lambda(G)^{-1}$ is, in fact, an upper bound for all $\sqrt[n]{\alpha(G^n)}$, and hence also for $\Theta(G)$. In order to prove this it suffices to show that for graphs G, H

$$\lambda(G \times H) = \lambda(G)\lambda(H) \tag{1}$$

holds, since this will imply $\lambda(G^n) = \lambda(G)^n$ and hence

$$\alpha(G^n) \leq \lambda(G^n)^{-1} = \lambda(G)^{-n}$$
$$\sqrt[n]{\alpha(G^n)} \leq \lambda(G)^{-1}.$$

To prove (1) we make use of the duality theorem of linear programming (see [I]) and get

$$\lambda(G) = \min_x \max_{C \in \mathcal{C}} \sum_{v \in C} x_v = \max_y \min_{v \in V} \sum_{C \ni v} y_C, \tag{2}$$

where the right-hand side runs through all probability distributions $y = (y_C : C \in \mathcal{C})$ on $\mathcal{C}$.

Consider $G \times H$, and let x and x' be distributions which achieve the minima, $\lambda(x) = \lambda(G)$, $\lambda(x') = \lambda(H)$. In the vertex set of $G \times H$ we assign the value $z_{(u,v)} = x_u x'_v$ to the vertex (u, v). Since $\sum_{(u,v)} z_{(u,v)} = \sum_u x_u \sum_v x'_v, = 1$, we obtain a distribution. Next we observe that the maximal cliques in $G \times H$ are of the form $C \times D = \{(u, v) : u \in C, v \in D\}$ where C and D are cliques in G and H, respectively. Hence we obtain

$$\lambda(G \times H) \leq \lambda(z) = \max_{C \times D} \sum_{(u,v) \in C \times D} z_(u, v)$$
$$- \max_{C \times D} \sum_{u \in C} x_u \sum_{v \in D} x'_v = \lambda(G)\lambda(H)$$

by the definition of $\lambda(G \times H)$. In the same way the converse inequality $\lambda(G \times H) \geq \lambda(G)\lambda(H)$ is shown by using the dual expression for $\lambda(G)$ in (2). In summary we can state:

$$\Theta(G) \leq \lambda(G)^{-1},$$

for any graph G. Let us apply our findings to the 5-cycle and, more generally, to the m-cycle C_m. By using the uniform distribution $(\frac{1}{m}, \cdots, \frac{1}{m})$ on the vertices, we obtain $\lambda(C_m) \leq \frac{2}{m}$, since any clique contains at most two vertices. Similarly, choosing $\frac{1}{m}$ for the edges and 0 for the vertices, we have $\lambda(C_m) \geq \frac{2}{m}$ by the dual expression in (2). We conclude that $\lambda(C_m) = \frac{2}{m}$ and therefore

$$\Theta(C_m) \leq \frac{m}{2}$$

for all m. Now, if m is even, then clearly $\alpha(C_m) = \frac{m}{2}$ and thus also $\Theta(C_m) = \frac{m}{2}$. For odd m, however, we have $\alpha(C_m) = \frac{m-1}{2}$. For $m = 3$, C_3 is a clique, and so is every product C_3^n, implying $a(C_3) = \Theta(C_3) = 1$. So, the first interesting case is the 5-cycle, where we know up to now

$$\sqrt{5} \leq \Theta(C_m) \leq \frac{5}{2} \tag{3}$$

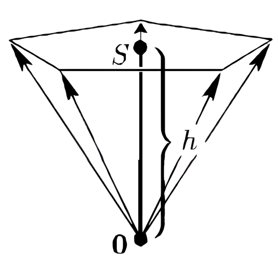
Using his linear programming approach (and some other ideas) Shannon was able to compute the capacity of many graphs and, in particular, of all graphs with five or fewer vertices — with the single exception of C_5, where he could not go beyond the bounds in (3). This is where things stood for more than 20 years until László Lovász showed by an astonishingly simple argument that indeed $\Theta(C_5) = \sqrt{5}$. A seemingly very difficult combinatorial problem was provided with an unexpected and elegant solution.

Lovász' main new idea was to represent the vertices v of the graph by real vectors of length 1 such that any two vectors which belong to non-adjacent vertices in G are orthogonal. Let us call such a set of vectors an orthonormal representation of G. Clearly, such a representation always exists: just take the unit vectors $(1, 0, 0, \cdots, 0)^T$, $(0, 1, 0, \cdots, 0)^T$, $\cdots$,(0,0,0,$\cdots$, 1)^T of dimension $m = |V|$.

For the graph C_5 we may obtain an orthonormal representation in $\mathbb{R}^3$ by considering an "umbrella" with five ribs $v_1, \cdots, v_5$ of unit length. Now open the umbrella (with tip at the origin) to the point where the angles between alternate ribs are 90°.
Lovász then went on to show that the height h of the umbrella, that is, the distance between 0 and S, provides the bound

$$\Theta(C_5) \leq \frac{1}{h^2}. \tag{4}$$

A simple calculation yields $h^2 = \frac{1}{5}$; see the box on the next page. From this $\Theta(C_5) \leq \sqrt{5}$ follows, and therefore $\Theta(C_5) = \sqrt{5}$.



The Lovász umbrella

Let us see how Lovász proceeded to prove the inequality (4). (His results were, in fact, much more general.) Consider the usual inner product

$$\langle x, y\rangle = x_1y_1 + \cdots + x_8y_8$$

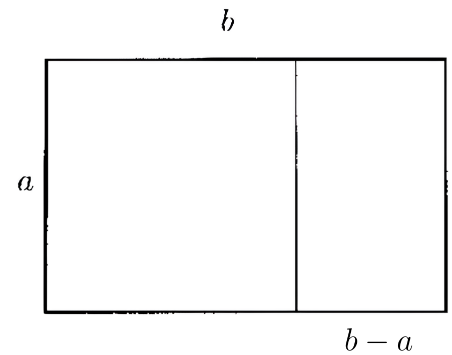
of two vectors $x = (x_1, \cdots, x_8)$, $y = (y_1, \cdots, y_8)$ in $\mathbb{R}^8$. Then $|x|^2 = \langle x, x\rangle = x_1^2 + \cdots + x_8^2$ is the square of the length $|x|$ of x, and the angle γ between x and y is given by

$$\cos\gamma = \frac{\langle x, y\rangle}{|x||y|}.$$

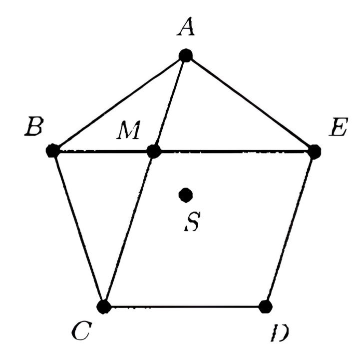
Thus $\langle x, y\rangle = 0$ if and only if x and y are orthogonal.

Pentagons and the golden section

Tradition has it that a rectangle was considered aesthetically pleasing if, after cutting off a square of length a, the remaining rectangle had the same shape as the original one. The side lengths a, b of such a rectangle must satisfy $\frac{b}{a} = \frac{a}{b-a}$. Setting $\tau := \frac{b}{a}$ for the ratio, we obtain $\tau = \frac{1}{\tau-1}$ or $\tau^2 - \tau - 1 = 0$. Solving the quatratic equation yields the *golden section* $\tau = \frac{1+\sqrt{5}}{2} \approx 1.6180$.



Consider now a regular pentagon of side length a, and let d be the length of its diagonals. It was already known to Euclid (Book XIII,8) that $\frac{d}{a} = \tau$, and that the intersection point of two diagonals divides the diagonals in the golden section.



Here is Euclid's Book Proof. Since the total angle sum of the pentagon is 3π, the angle at any vertex equals $\frac{3\pi}{5}$. It follows that $\sphericalangle ABE = \frac{\pi}{5}$, since ABE is an isosceles triangle. This, in turn, implies $\sphericalangle AMB = \frac{3\pi}{5}$, and we conclude that the triangles ABC and AMB are similar. The quadrilateral $CMED$ is a rhombus since opposing sides are parallel (look at the angles), and so $|MC| = a$ and thus $|AM| = d - a$. By the similarity of ABC and AMB we conclude

$$\frac{d}{a} = \frac{|AC|}{|AB|} = \frac{|AB|}{|AM|} = \frac{a}{b-a} = \frac{|MC|}{|MA|} = \tau.$$

There is more to come. For the distance s of a vertex to the center of the pentagon S, the reader is invited to prove the relation $s^2 = \frac{d^2}{\tau+2}$ (note that BS cuts the diagonal AC at a right angle and halves it). To finish our excursion into geometry, consider now the umbrella with the regular pentagon on top. Since alternate ribs (of length 1) form a right angle, the theorem of Pythagoras gives us $d = sqrt2$, and hence $s^2 = \frac{2}{\tau+2} = \frac{4}{\sqrt{5}+5}$. So, with Pythagoras again, we find for the height $h = |OS|$ our promised result

$$h^2 = 1 - s^2 = \frac{1+\sqrt{5}}{\sqrt{5}+5} = \frac{1}{\sqrt{5}}.$$

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 $T = \{v^{(1)}, \dots, v^{(m)}\}$ be an orthonormal representation of G in $\mathbb{R}^s$, where $v^{(i)}$ corresponds to the vertex. We assume in adition that all the vectors $v^{(i)}$ have the same angle ($\neq 90°$) with the vector $u := \frac{1}{m}(v^{(1)} + \dots + v^{(m)})$, or equivalently that the inner product

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

has the same value $\sigma_T \neq 0$ for all i. Let us call this value σ the constant of the representation T. For the Lovisz umbrella that represents Cs 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 $x = (x, \dots, x)$ on V and set

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

and

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

Let U be a largest independent set in G with $|U| = a$, and define $\mathrm{x}_U = (\mathrm{x}_1, \dots, x_m)$ 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 $\in \mathbb{R}^s. Appliedto$a $= \mathrm{x}_1 v^{(1)} + \dots + x_m v^{(m)}$ and $b = u$, the inequality yields

$$\langle x_1 v^{(1)} + \dots + x_m v^{(m)}, u \rangle^2 \leq \mu(x)|u|^2. \tag{5}$$

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

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

for any distribution x. Thus, in particular, this has to hold for the uniform distribution $(\frac{1}{m}, \dots, \frac{1}{m})$,which implies $|u|^2 = \sigma^2$.Hence (5) 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 $x = (\frac{1}{m}, \ldots, \frac{1}{m})$ 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

$$\mu_T(G) = \sigma_T \tag{6}$$

In summary, we have established the inequality

$$\alpha(G) \le \frac{1}{\sigma_T} \tag{7}$$

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 $v = (v_1, \ldots, v_r)$ be a vector in R and $w = (w_1, \ldots, w_s)$ be a vector in S. To the vertex in $G \times H$ corresponding to the pair (v, w) we associate the vector

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

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 (6) 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 (7) 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 Lovlsz' argument:

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

$$\Theta(G) \le \frac{1}{\sigma_T} \tag{8}$$

Looking at the Lovlsz 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.

"Umbrellas with five ribs"

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}$$

The adjacency matrix for the 5-cycle C_5

Let us carry our discussion a little further. We see from (8) 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 $s = rank(M)$, 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 $p = |\lambda_{min}| = 2 \cos fracpim$ (see the box). Every row of the adjacency matrix contains two I '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

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

and hence

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

for all i. We can therefore apply our main result (8) and conclude

$$\Theta(C_m) \leq \frac{m}{1 + (\cos fracpim)^{-1}} \tag{9}$$

Notice that because of $\cos fracpim < 1$ the bound (9) is better than the bound $\Theta(C_m) \leq \frac{m}{2}$ we found before. Note further $\cos fracpi5 = fractau2$, where $\tau = \frac{\sqrt{5}+1}{2}$ is the golden section. Hence for $m = 5$ we again obtain

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

The orthonormal representation given by this construction is, of course, precisely the "Lovisz umbrella."

And what about C_7, C_9, and d the other odd cycles? By considering $\alpha(C_m^2)$, $\alpha(C_m^3)$ and other small powers the lower bound $\frac{m-1}{2} \leq \Theta(C_m)$) can cer- tainly be increased, but for no odd $m \geq 7$ do the best known lower bounds agree with the upper bound given in (8). So, twenty years after Lovasz' which is 3.2141 5 Q(C7) 5 3.3177. 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.

For example, for $m = 7$ all we know is

$$\sqrt[5]{344} \leq \Theta(C_7) \leq \frac{7}{1+(\cos\frac{\pi}{7}^{-1})},$$

which is $3.2141 \leq \Theta(C_7) \leq 3.3177.$

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 $1, \zeta,^2, \ldots, \zeta^{m-1}$ 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

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

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

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

is the smallest eigenvalue of A.