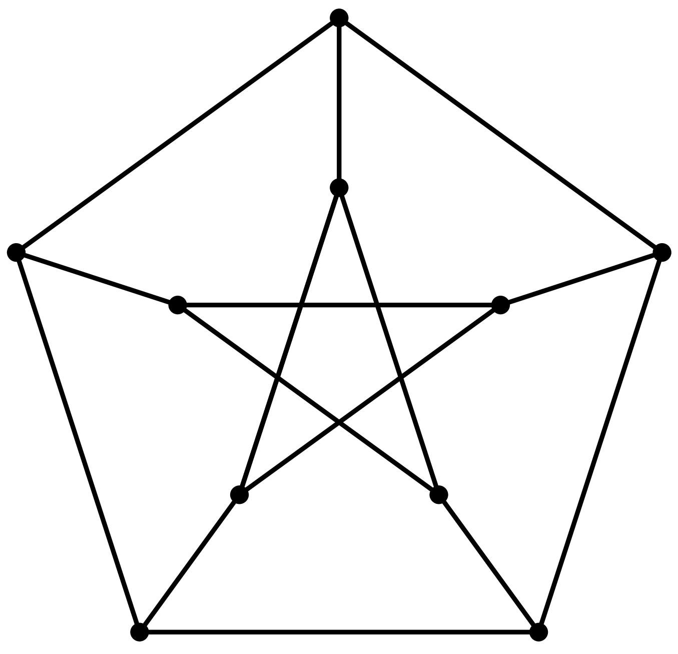
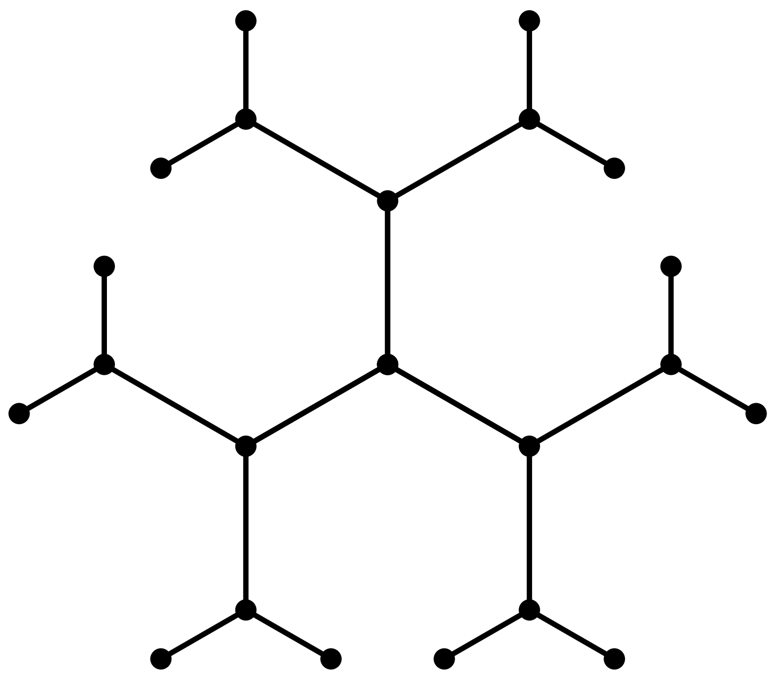
Superexpanders and Markov cotype in the work of Mendel and Naor

Keith Ball

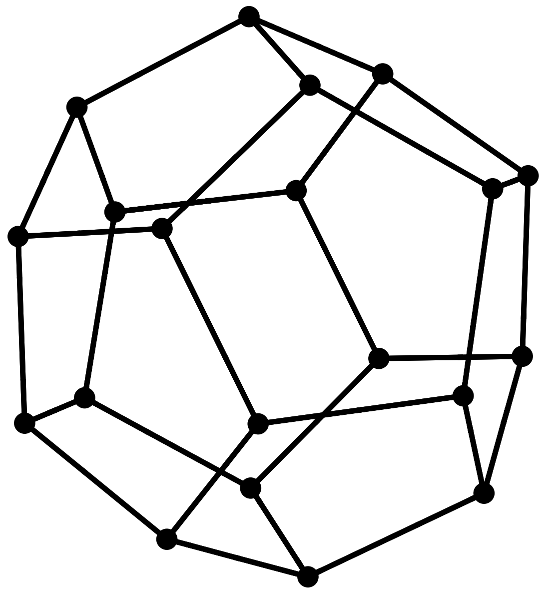


Expanders A graph is an expander if it has small degree but spreads information very fast: in a few steps you can go to most places.



Locally the graph looks like this. How do we join it up round the back?

We want something like a dodecahedron but on a much bigger scale.



For larger graphs the first constructions used groups/number theory: Margulis (1973).

Random graphs have the right properties but for most purposes these are a cheat.

How do we quantify expansion?

A set of vertices that is neither too large nor too small has many neighbours.

The graph has a large spectral gap.

Assume that the graph is d-regular: (each vertex has d neighbours): and consider the adjacency matrix:

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

The matrix A is stochastic: the transition matrix for the simple random walk on the graph.

If we have a random vertex with distribution

$$p = (p_1, p_2, \ldots, p_n)$$

on the vertices then after we take one step the new vertex has distribution Ap.

If we have a random vertex with distribution

$$p = (p_1, p_2, \ldots, p_n)$$

on the vertices then after we take one step the new vertex has distribution Ap.

The stationary distribution is uniform:

$$A \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

So the top eigenvalue is 1 and the other eigenvalues are smaller (assuming...). So whatever our initial distribution, we converge to the stationary distribution and the rate at which we do it is much faster if the remainder of the spectrum is close to 0.

The size of the next largest eigenvalue is $1-g$ if for every pair of sequences (x_i) and (y_i) of reals

$$\sum_{ij} a_{ij}(x_i - y_j)^2 \geq g \sum_{ij} \frac{1}{n}(x_i - y_j)^2.$$

The d-regular graph G on n vertices is an expander with constant C if its normalised adjacency matrix A satisfies

$$\frac{1}{n} \sum_{ij} (x_i - y_j)^2 \leq C \sum_{ij} a_{ij}(x_i - y_j)^2.$$

Superexpanders

In this form the condition can be tensored with the identity on a normed space X. G is an expander for X if

$$\frac{1}{n}\sum_{ij}\|x_i - y_j\|^2 \leq C\sum_{ij} a_{ij}\|x_i - y_j\|^2$$

for points (x_i) and (y_i) in X.

The same step was used many times: for Type, Cotype, Martingale transforms, the Rademacher projection, the Riesz projection and so on.

Are there expander graphs for spaces other than Hilbert space, a large class of spaces or even all spaces? (Order going to ∞, bounded degree and bounded C.)

You cannot find expanders for all spaces. This follows from Gromov's original idea for using non-linear spectral gaps to study coarse embeddings.

Suppose G is an expander for ℓ_∞. Equip G with its path metric and embed it into ℓ_∞.

$$\frac{1}{n}\sum_{ij}\|x_i - x_j\|^2 \leq C\sum_{ij} a_{ij}\|x_i - x_j\|^2$$

so

$$\frac{1}{n}\sum_{ij} d(x_i, x_j)^2 \leq C\sum_{ij} a_{ij} d(x_i, x_j)^2 = Cn.$$

So on average, vertices of the graph are only distance C apart: but if the graph has degree d only d^C vertices can have this property.

Why is this interesting? In connection with the Novikov conjecture, Kasparov and Yu asked:

Are there expanders for uniformly convex spaces?

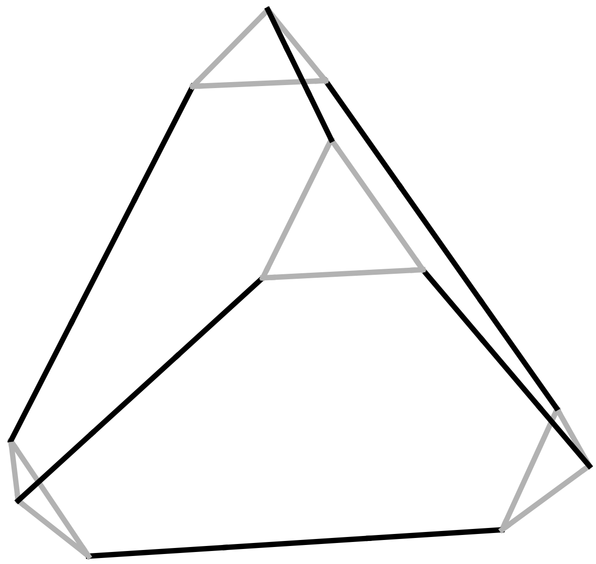
Lafforgue used a delicate construction with a strong property (T) for $SL_3(F)$ to give an example.

Mendel and Naor gave a simpler construction by adapting the Zig-Zag product construction of Reingold, Vadhan and Wigderson.

The real aim is to develop the non-linear spectral calculus initiated by Gromov.

The zig-zag construction depends upon using a pair of expanders G and H to generate a new, larger expander. The matching condition is that the number of vertices in H is equal to the degree of G.

For example, G is the tetrahedron and H the triangle. We start as follows:



The graph we create joins vertices if you can move between them grey-black-grey: so this is really the zig-zag-zig product.

If G has n vertices and degree m, and H has m vertices and degree d, the new graph has nm vertices and degree d^2. As long as $m = d^2$ we can zig-zag again with H and continue.

We need to relate the expansion of the new graph to that of G and H.

Let (x_{ir}) and (y_{js}) be sequences in X where i and j denote the cloud (vertex of G) and r and s denote the vertex in H. For each edge ij in G let $V(i,j)$ be the vertex in cloud i whose external edge goes to cloud j.

Then for each r and s

$$\frac{1}{n}\sum_{ij}\|x_{ir}-y_{js}\|^2 \leq C_G\sum_{ij}\frac{1}{m}\mathbf{1}_{(ij\in e(G))}\|x_{ir}-y_{js}\|^2.$$

Hence

$$\frac{1}{nm}\sum_{ijrs}\|x_{ir}-y_{js}\|^2 \leq \frac{C_G}{m^2}\sum_{ijrs}\mathbf{1}_{(ij\in e(G))}\|x_{ir}-y_{js}\|^2$$

$$=\frac{C_G}{m^2}\sum_{ijrstu}\mathbf{1}_{(t=V(i,j))}\mathbf{1}_{(u=V(j,i))}\|x_{ir}-y_{js}\|^2.$$

$$\frac{C_G}{m^2} \sum_{ijrstu} \mathbf{1}_{(t=V(i,j))} \mathbf{1}_{(u=V(j,i))} \|x_{ir} - y_{js}\|^2.$$

Fix i and s:

$$\frac{1}{m} \sum_{rtju} \mathbf{1}_{(t=V(i,j))} \mathbf{1}_{(u=V(j,i))} \|x_{ir} - y_{js}\|^2 =$$

$$\frac{1}{m} \sum_{rt} \|x_{ir} - y_{j_i(t)s}\|^2$$

where $j_i(t)$ is the cloud to which vertex t in the i^{th} cloud is joined.

This is at most

$$\frac{C_H}{d} \sum_{rt} \mathbf{1}_{(rt \in e(H))} \|x_{ir} - y_{j_i(t)s}\|^2.$$

So our expression is at most

$$\frac{C_G C_H}{md} \sum_{ijrstu} \mathbf{1}_{(rt \in e(H))} \mathbf{1}_{(t=V(i,j))} \mathbf{1}_{(u=V(j,i))} \|x_{ir} - y_{js}\|^2.$$

"Repeat" for the su sum except that now we can fix j but not r. We break the sum over r into d separate sums according to some indexing of the neighbours of $V(i_j(u), u)$. We get at most

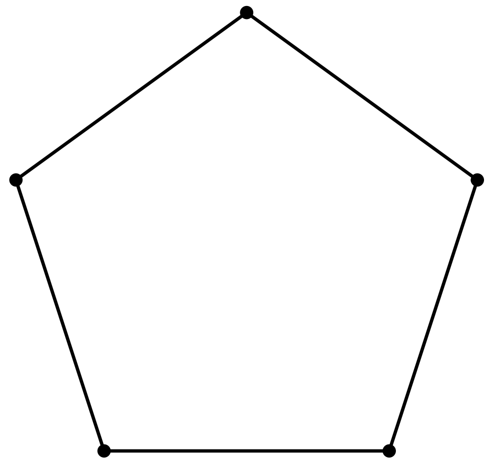
$$\frac{C_G C_H^2}{d^2} \sum_{ijrstu} \mathbf{1}_{(rt \in e(H))} \mathbf{1}_{(su \in e(H))} \times$$

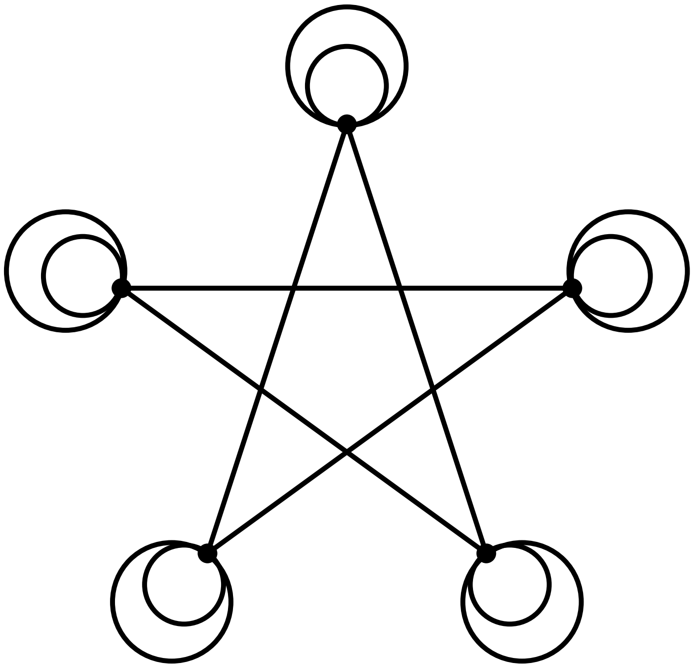
$$\mathbf{1}_{(t=V(i,j))} \mathbf{1}_{(u=V(j,i))} \|x_{ir} - y_{js}\|^2.$$

So the new graph is an expander with constant $C_G C_H^2$. This quantity is bigger than C_G so it looks as though we can't induct: the constant will grow as we keep zig-zag-zigging with H.

To avoid this we add a further step. You can always improve the spectral gap by a factor of almost two by squaring the adjacency matrix. If the next eigenvalue of A is λ then that of A^2 is λ^2. This changes the degree of G from m to m^2. So we choose H to have d^4 vertices instead of d^2.

Squaring also produces a matrix with entries that are not 0 and 1. The adjacency matrix of a graph with loops and multiple edges.





So the algorithm is this.

- Start with H: degree d and d^4 vertices.

- Let $G_0 = H^2$ which has degree d^2.

- Now move from G_n to G_{n+1} by squaring and then zig-zagging with H.

The inductive step multiplies the constant by $C_H^2/2$ and so as long as C_H is not too big we can continue.

What about superexpanders? The argument for the zig-zag product doesn't use any spectral theory so it works perfectly well in an arbitrary space.

Since we know that we cannot build expanders in arbitrary spaces, the problem must occur in the squaring step or the construction of H: actually both. These did use spectral theory.

For the squaring step we want to know that in our space X,

$$\sum_{ij} A_{ij} \|x_i - y_j\|^2$$

is genuinely less than

$$\sum_{ij} (A^2)_{ij} \|x_i - y_j\|^2.$$

We want to know that in our space X,

$$\sum_{ij} A_{ij} \|x_i - y_j\|^2$$

is genuinely less than

$$\sum_{ij} (A^2)_{ij} \|x_i - y_j\|^2.$$

In two steps you move further than in one.

This cannot be strictly true.

Markov type and cotype

A normed space X has Markov type 2 if for some K, for every $n \times n$ symmetric stochastic matrix $A = (a_{ij})$ and points (x_i) in X we have

$$\sum_{ij} (A^m)_{ij} \, \|x_i - x_j\|^2 \leq K^2 \, m \sum_{ij} a_{ij} \, \|x_i - x_j\|^2.$$

This says that in m steps you move at most $\sqrt{m}$ times as far as in one step.

We want the opposite, cotype, condition.

The cotype condition is more complicated: for every $n \times n$ symmetric stochastic matrix $A = (a_{ij})$ and points (x_i) in X there are points $(y_i = \sum b_{ij}x_j)$ in X so that

$$m \sum_{ij} a_{ij} \|y_i - y_j\|^2 \leq K^2 \sum_{ij} b_{ij} \|x_i - x_j\|^2$$

where B is roughly A^m.

Ball 2-uniformly convex spaces have Markov cotype 2.

Initial construction of H

This uses a quotient of the cube, Fourier analysis on the cube (Beckner's Inequality) and Pisier's analytic continuation of the heat semigroup for K-convex spaces.