Generalized Hypercubes.

We already saw that H_d was a Cayley graph with $\Gamma = (\mathbb{Z}/2\mathbb{Z})^d$ and generator set $S = \{b_1 \ldots b_d \mid \text{exactly one } b_i = 1\}$. We noticed then that (since each element is its own inverse), we could take <u>any</u> $S \subset \Gamma$ and still get a Cayley graph.

So suppose $S = \{g_1, \ldots, g_k\}$ and let G be the Cayley graph with these generators.

We previously computed the eigenvectors of H_d. We now check these are <u>also</u> eigenvectors of G.

For each $b = (b_1 \dots b_d) \in \{0,1\}^d$, define

$$\psi_b(\vec{x}) = (-1)^{\langle \vec{b}, \vec{x} \rangle}$$

Lemma. The vector $\vec{\psi}_b$ is a Laplacian matrix eigenvector with eigenvalue

$$K - \sum_{i=1}^{K} (-1)^{\langle \vec{b}, \vec{g}_i \rangle}$$

Proof. Observe that

$$\vec{\psi}_b(\vec{x}+\vec{y}) = (-1)^{\langle \vec{b}, \vec{x}+\vec{y} \rangle} = (-1)^{\langle \vec{b}, \vec{x} \rangle} (-1)^{\langle \vec{b}, \vec{y} \rangle}$$

$$= \vec{\psi}_b(\vec{x})\, \vec{\psi}_b(\vec{y}).$$

Now we compute

$$(L\vec{\Psi}_b)(\vec{x}) = \sum_{i=1}^{K} \vec{\Psi}_b(\vec{x}) - \vec{\Psi}_b(\vec{x}+\vec{g}_i)$$

$$= K\,\vec{\Psi}_b(\vec{x}) - \sum_{i=1}^{K} \vec{\Psi}_b(\vec{x}+\vec{g}_i)$$

$$= K\,\vec{\Psi}_b(\vec{x}) - \sum_{i=1}^{K} \vec{\Psi}_b(\vec{x})\,\vec{\Psi}_b(\vec{g}_i)$$

$$= \vec{\Psi}_b(\vec{x}) \left[K - \sum_{i=1}^{K} \vec{\Psi}_b(\vec{g}_i) \right],$$

as claimed!

Now we show that if we choose random generators $g_1, \ldots, g_K$ for $K = cd$ where c is constant, it will be very likely that all eigenvalues will be close to K. (We will need $c > 1$; maybe $c = 10$, $c = 2$ or $c = 1 + \epsilon$.)

Lemma. If $b = (b_1 \ldots b_d)$ contains at least one 1, and g is chosen at random (uniformly) from $\{0,1\}^d$,

$$\langle \vec{b}, \vec{g} \rangle \bmod 2$$

is uniformly distributed between $\{0,1\}$.

~~Proof. Homework.~~

Proof. Homework.

Given the lemma,

$$\lambda_b = K - \sum_{i=1}^{K} (-1)^{\langle \vec{b}, \vec{g}_i \rangle}$$

$$= K - (\text{a sum of indepent} \pm 1 \text{ random variables})$$

We expect sums of i.i.d. ± 1 r.v.s to be concentrated around zero. Precisely,

Theorem (Simple Chernoff bounds).
Let $x_1, \ldots, x_K$ be i.i.d. ± 1 random variables (with equal probability of $+1, -1$). For all $t > 0$, we have

$$\Pr\left[\left|\sum_i x_i\right| \geq t\right] \leq 2e^{-t^2/2K}.$$

Thus, if $t \sim cK$, this becomes very small for large K.

We can combine the Chernoff bound with the union bound:

Theorem. (Boole's inequality)

For any finite (or countable) set of events $A_1, \ldots, A_i, \ldots$ we have

$$\Pr\left[\bigcup_i A_i\right] = \Pr\left[\begin{array}{c}\text{at least one of} \\ \text{the } A_i \text{ occurs}\end{array}\right]$$

$$\leq \sum_i \Pr[A_i].$$

Notice that this is true whether or not the events A_i are correlated.

We now consider the eigenvalues of the generalized hypercube.

Suppose the dimension is d, and we choose K generators at random.

There are $2^d - 1$ nonzero eigenvectors each indexed by some $b = (b_1 \ldots b_d)$ containing at least one 1.

$$\lambda_b = K - \sum_{i=1}^{K} (-1)^{\langle b, \vec{g}_i \rangle}$$

$$= K - \sum_{i=1}^{K} x_i$$

where x_i is a Bernoulli ± 1 random variable.

So for each K, t we have

$$\Pr\left[|K-\lambda_b| \geq t\right] = \Pr\left[\left|\sum_{i=1}^{K} x_i\right| \geq t\right] \leq 2e^{-t^2/2K}$$

by the Chernoff bound. Let's assume that $K = cd$ (for some c) and $t = K\sqrt{2/c}$.

$$\Pr\left[|K-\lambda_b| \geq t\right] \leq 2e^{-K^2 \frac{\cancel{2}}{c}/\cancel{2}K}$$

$$\leq 2e^{-K/c}$$

$$\leq 2e^{-d}$$

Now there are $2^d - 1$ eigenvalues, so

~~the~~

$$\Pr\left[\exists b \,\middle|\, |K-\lambda_b| \geq K\sqrt{\tfrac{2}{c}}\right] \leq (2^d - 1)\, 2e^{-d}$$

by the union bound.

We observe that

$$(2^d - 1)\, 2e^{-d} \leq 2 \cdot \frac{2^d}{e^d} = 2\left(\frac{2}{e}\right)^d.$$

Since $2/e < 1$, the right hand side approaches 0 exponentially fast (in d).

We have proved:

Theorem. If $K = cd$, ~~and~~ then

$$\Pr\left[\exists b \,\middle|\, |K - \lambda_b| \geq K\sqrt{\tfrac{2}{c}}\right] \leq 2\left(\frac{2}{e}\right)^d$$

Notice that $|K - \lambda_b| \leq K$ (always) because $\lambda_b \in [0, 2K]$, so this bound is useful only for $c > 2$.

Now let's ~~add~~ weight each edge of our generalized hypercube with $2^d/K$. Weighting edges ~~and~~ scales eigenvalues, so

Theorem. If G is the graph with vertex set $\{0,1\}^d$ and edges determined by $K = cd$ random generators, each edge with weight $2^d/K$,

$$\Pr\left[\exists b \,\middle|\, |2^d - \lambda_b| \geq 2^d \sqrt{\tfrac{2}{c}}\right] \leq 2\left(\tfrac{2}{e}\right)^d.$$

or, if v is # of vertices,

$$\Pr\left[\exists b \,\middle|\, |v - \lambda_b| \geq \sqrt{\tfrac{2}{c}}\, v\right] \leq 2\left(\tfrac{2}{e}\right)^d$$

Now the complete graph K_v has eigenvalues $v, \ldots, v$, so we can rewrite our theorem as

Theorem. If G is the graph with vertex set $\{0,1\}^d$ of $v = 2^d$ vertices, and edges determined by $K = cd = c\log_2 v$ random generators, each ~~[illegible]~~ edge with weight $v/K = 2^d/K$,

$$\Pr\left[G \text{ is a } \left(1+\sqrt{2/c}\right)\text{-approximation of } K_v\right]$$

is $\geq 1 - 2\left(\frac{2}{e}\right)^d$.

This is a surprise! These graphs have degree $K = c\log_2 v \ll$ the degree $(v-1)$ of the complete graphs K_v.

Further, the graph G has only

$$\frac{vk}{2} = \frac{c}{2} v \log_2 v$$

edges - much less than the

$$\binom{v}{2} = \frac{v(v-1)}{2}$$

edges in K_v.