Today.

Continue markov chain mixing analysis.

Prove "hard side" of Cheeger.

Analyzing random walks on graph.

Start at vertex, go to random neighbor.
For d-regular graph: eventually uniform.
if not bipartite. Odd /even step!

How to analyze?

Random Walk Matrix: M.
M - normalized adjacency matrix.
Symmetric, $\sum_j M[i,j] = 1$.
$M[i,j]$- probability of going to j from i.

Probability distribution at time t: v_t.
$v_{t+1} = Mv_t$ Each node is average over neighbors.

Evolution? Random walk starts at 1, distribution $e_1 = [1, 0, \ldots, 0]$.
$M^t v_1 = \frac{1}{N} v_1 + \sum_{i>1} \lambda_i^t \alpha_i v_i$.
$v_1 = [\frac{1}{N}, \ldots, \frac{1}{N}]$ →Uniform distribution.

Doh! What if bipartite?
Negative eigenvalues of value -1: $(+1, -1)$ on two sides.
Side question: Why the same size? Assumed regular graph.

Fix-it-up chappie!

"Lazy" random walk: With probability $1/2$ stay at current vertex.

Evolution Matrix: $\frac{I+M}{2}$
Eigenvalues: $\frac{1+\lambda_i}{2}$
$\frac{1}{2}(I+M)v_i = \frac{1}{2}(v_i + \lambda_i v_i) = \frac{1+\lambda_i}{2} v_i$
Eigenvalues in interval $[0, 1]$.

Spectral gap: $\frac{1-\lambda_2}{2} = \frac{\mu}{2}$.

Uniform distribution: $\pi = [\frac{1}{N}, \ldots, \frac{1}{N}]$
Distance to uniform: $d_1(v_t, \pi) = \sum_i |(v_t)_i - \pi_i|$

"Rapidly mixing": $d_1(v_t, \pi) \leq \epsilon$ in **poly**$(\log N, \log \frac{1}{\epsilon})$ time.
When is chain rapidly mixing?

Another measure: $d_2(v_t, \pi) = \sum_i ((v_t)_i - \pi_i)^2$.
Note: $d_1(v_t, \pi) \leq \sqrt{N} d_2(v_t, \pi)$
n – "size" of vertex, $\mu \geq \frac{1}{p(n)}$ for poly $p(n)$, $t = O(p(n) \log N)$.

$$d_2(v_t, \pi) = |A^t e_1 - \pi|^2 \leq \left(\frac{(1+\lambda_2)}{2}\right)^{2t} \leq (1 - \frac{1}{2p(n)})^{2t} \leq \frac{1}{\text{poly}(N)}$$

Rapidly mixing with big ($\geq \frac{1}{p(n)}$) spectral gap.

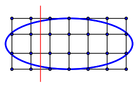
Rapid mixing, volume, and surface area..

Recall volume of convex body.
Grid graph on grid points inside convex body.

Recall Cheeger: $\frac{\mu}{2} \leq h(G) \leq \sqrt{2\mu}$.

Lower bound expansion → lower bounds on spectral gap μ
→ Upper bound mixing time.

$h(G) \approx \frac{\text{Surface Area}}{\text{Volume}}$



Isoperimetric inequality.

$\text{Vol}_{n-1}(S, \overline{S}) \geq \frac{\min(\text{Vol}(S), \text{Vol}(\overline{S}))}{diam(P)}$

Edges ∝ surface area, Assume $Diam(P) \leq p'(n)$
→ $h(G) \geq 1/p'(n)$
→ $\mu > 1/2p'(n)^2$
→ $O(p'(n)^2 \log N)$ convergence for Markov chain on BIG GRAPH.
→ Rapidly mixing chain:

Khachiyan's algorithm for counting partial orders.

Given partial order on $x_1, \ldots, x_n$.

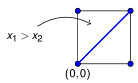
Sample from uniform distribution over total orders.

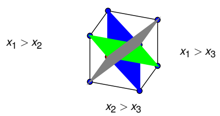
Start at an ordering.
Swap random pair and go if consistent with partial order.

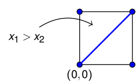
Rapidly mixing chain?

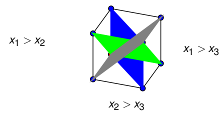
Map into d-dimensional unit cube.

$x_i < x_j$ corresponds to halfspace (one side of hyperplane) of cube.
"dimension i = dimension j"
total order is intersection of n halfspaces.
each of volume: $\frac{1}{n!}$.
since each total order is disjoint
and together cover cube.









Each order takes $\frac{1}{n!}$ volume.

Number of orders ≡ volume of intersection of partial order relations.

Diameter: $O(\sqrt{n})$

Isoperimetry:

$$Vol_{n-1}(S, \overline{S}) = \frac{E(S,\overline{S})}{(n-1)!} \geq \frac{|S|}{n!\sqrt{n}}$$

Edge Expansion: the degree d is $O(n^2)$,
$h(S) = \frac{|E(S,\overline{S})|}{d|S|} \geq \frac{1}{n^{7/2}}$ Mixes in time $O(n^7 \log N) = O(n^8 \log n)$.
Do the polynomial dance!!!

Summary.

Eigenvectors for hypercubes.
Tight example for LHI of Cheeger. Eigenvectors for cycle.
Tight example for RHI of Cheeger.

Random Walks and Sampling.

Eigenvectors, Isoperimetry of Volume, Mixing.

Partial Order Application.

Cheeger Hard Part.

Now let's get to the hard part of Cheeger $h(G) \le \sqrt{2(1-\lambda_2)}$.

Idea: We have $1-\lambda_2$ as a continuous relaxation of $\phi(G)$

Take the 2^{nd} eigenvector $x = argmin_{x \in \mathbb{R}^V - \text{Span}\{\mathbf{1}\}} \frac{\sum_{i,j} M_{ij}(x_i-x_j)^2}{\frac{1}{n}\sum_{i,j}(x_i-x_j)^2}$

Consider x as an embedding of the vertices to the real line.

Round x to get a $x \in \{0,1\}^V$

Rounding: Take a threshold t,

$$\begin{cases} x_i \ge t & \to x_i = 1 \\ x_i < t & \to x_i = 0 \end{cases}$$

What will be a good t?

We don't know. Try all possible thresholds ($n-1$ possibilities), and hope there is a t leading to a good cut!

Sweep Cut Algorithm

Input: $G=(V,E)$, $x \in \mathbb{R}^V, x \perp \mathbf{1}$

Sort the vertices in non-decreasing order in terms of their values in x
WLOG $V = \{1,\dots,n\}$ $\quad x_1 \le x_2 \le \dots \le x_n$

Let $S_i = \{1,\dots,i\}$ $\quad i = 1,\dots,n-1$

Return $S = argmin_{S_i} h(S_i)$

Main Lemma: $G=(V,E)$, d-regular
$x \in \mathbb{R}^V, x \perp \mathbf{1}, \mu = \frac{\sum_{i,j} M_{ij}(x_i-x_j)^2}{\frac{1}{n}\sum_{i,j}(x_i-x_j)^2}$
If S is the ouput of the sweep cut algorithm, then $h(S) \le \sqrt{2\mu}$

Note: Applying the Main Lemma with the 2^{nd} eigenvector v_2, we have $\mu = 1-\lambda_2$, and $h(G) \le h(S) \le \sqrt{2(1-\lambda_2)}$. Done!

Proof of Main Lemma

WLOG $V = \{1,\dots,n\}$ $\quad x_1 \le x_2 \le \dots \le x_n$

Want to show

$$\exists i \text{ s.t. } h(S_i) = \frac{\frac{1}{d}|E(S_i, V-S_i)|}{min(|S_i|, |V-S_i|)} \le \sqrt{2\mu}$$

Probabilistic Argument: Construct a distribution D over $\{S_1,\dots,S_{n-1}\}$ such that

$$\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|,|V-S|)]} \le \sqrt{2\mu}$$

$$\to \mathbb{E}_{S\sim D}[\tfrac{1}{d}|E(S,V-S)| - \sqrt{2\mu}\, min(|S|,|V-S|)] \le 0$$

$$\exists S \quad \tfrac{1}{d}|E(S,V-S)| - \sqrt{2\mu}\, min(|S|,|V-S|) \le 0$$

The distribution D

WLOG, shift and scale so that $x_{\lfloor \frac{n}{2} \rfloor} = 0$, and $x_1^2 + x_n^2 = 1$

Take t from the range $[x_1, x_n]$ with density function $f(t) = 2|t|$.

Check: $\int_{x_1}^{x_n} f(t)\mathrm{d}t = \int_{x_1}^{0} -2t\mathrm{d}t + \int_{0}^{x_n} 2t\mathrm{d}t = x_1^2 + x_n^2 = 1$

$S = \{i : x_i \le t\}$

Take D as the distribution over $S_1,\dots,S_{n-1}$ from the above procedure.

Goal: $\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|,|V-S|)]} \le \sqrt{2\mu}$

Denominator:

Let T_i = indicator for "i is in the smaller set of $S, V-S$"

Can check

$$\mathbb{E}_{S\sim D}[T_i] = Pr[T_i = 1] = x_i^2$$

$$\begin{aligned} \mathbb{E}_{S\sim D}[min(|S|,|V-S|)] &= \mathbb{E}_{S\sim D}[\sum_i T_i] \\ &= \sum_i \mathbb{E}_{S\sim D}[T_i] \\ &= \sum_i x_i^2 \end{aligned}$$

Goal: $\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|,|V-S|)]} \le \sqrt{2\mu}$

Numerator:

Let $T_{i,j}$ = indicator for i,j is cut by $(S, V-S)$

$$\begin{cases} x_i, x_j \text{ same sign:} & Pr[T_{i,j}=1] = |x_i^2 - x_j^2| \\ x_i, x_j \text{ different sign:} & Pr[T_{i,j}=1] = x_i^2 + x_j^2 \end{cases}$$

A common upper bound: $\mathbb{E}[T_{i,j}] = Pr[T_{i,j}=1] \le |x_i - x_j|(|x_i| + |x_j|)$

$$\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|] = \frac{1}{2}\sum_{i,j} M_{ij}\mathbb{E}[T_{i,j}]$$
$$\le \frac{1}{2}\sum_{i,j} M_{ij}|x_i - x_j|(|x_i| + |x_j|)$$

Cauchy-Schwarz Inequality

$|a \cdot b| \le \|a\|\|b\|$, as $a \cdot b = \|a\|\|b\|cos(a,b)$

Applying with $a, b \in \mathbb{R}^{n^2}$ with $a_{ij} = \sqrt{M_{ij}}|x_i - x_j|, b_{ij} = \sqrt{M_{ij}}|x_i| + |x_j|$

Numerator:

$$\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|] = \frac{1}{2}\sum_{i,j} M_{ij}\mathbb{E}[T_{i,j}]$$
$$\le \frac{1}{2}\sum_{i,j} M_{ij}|x_i - x_j|(|x_i| + |x_j|)$$
$$= \frac{1}{2} a \cdot b$$
$$\le \frac{1}{2}\|a\|\|b\|$$

Recall $\mu = \frac{\sum_{i,j} M_{ij}(x_i-x_j)^2}{\frac{1}{n}\sum_{i,j}(x_i-x_j)^2}$, $a_{ij} = \sqrt{M_{ij}}|x_i - x_j|, b_{ij} = \sqrt{M_{ij}}|x_i| + |x_j|$

$$\|a\|^2 = \sum_{i,j} M_{ij}(x_i - x_j)^2 = \frac{\mu}{n}\sum_{i,j}(x_i - x_j)^2$$
$$= \frac{\mu}{n}\sum_{i,j}(x_i^2 + x_j^2) - \sum_{i,j} 2x_i x_j$$
$$= \frac{\mu}{n}\sum_{i,j}(x_i^2 + x_j^2) - 2(\sum_i x_i)^2$$
$$\le \frac{\mu}{n}\sum_{i,j}(x_i^2 + x_j^2) = 2\mu\sum_i x_i^2$$

$$\|b\|^2 = \sum_{i,j} M_{ij}(|x_i| + |x_j|)^2$$
$$\le \sum_{i,j} M_{ij}(2x_i^2 + 2x_j^2)$$
$$= 4\sum_i x_i^2$$

Goal: $\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|,|V-S|)]} \le \sqrt{2\mu}$

Numerator:

$$\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|] =\le \frac{1}{2}\|a\|\|b\|$$
$$\le \frac{1}{2}\sqrt{2\mu\sum_i x_i^2}\sqrt{4\sum_i x_i^2} \qquad = \sqrt{2\mu}\sum_i x_i^2$$

Recall **Denominator:**

$$\mathbb{E}_{S\sim D}[min(|S|,|V-S|)] = \sum_i x_i^2$$

We get

$$\frac{\mathbb{E}_{S\sim D}[\frac{1}{d}|E(S,V-S)|]}{\mathbb{E}_{S\sim D}[min(|S|,|V-S|)]} \le \sqrt{2\mu}$$

Thus $\exists S_i$ such that $h(S_i) \le \sqrt{2\mu}$, which gives $h(G) \le \sqrt{2(1-\lambda)}$ □

Summary

Second largest eigenvlaue of matrix: λ_2.

Bounds mixing time.

Connected to "sparse" cuts.

Cheeger: $\frac{\mu}{2} \le h(G) \le \sqrt{2\mu}$.

Left hand tight: Hypercube.

Right hand tight: Cycle.

Left side proof: produce good Rayleigh quotient vector from sparse cut.

Right hand proof: produce sparse cut from good Rayleigh quotient.

Connect to bounding mixing time on Markov Chain.