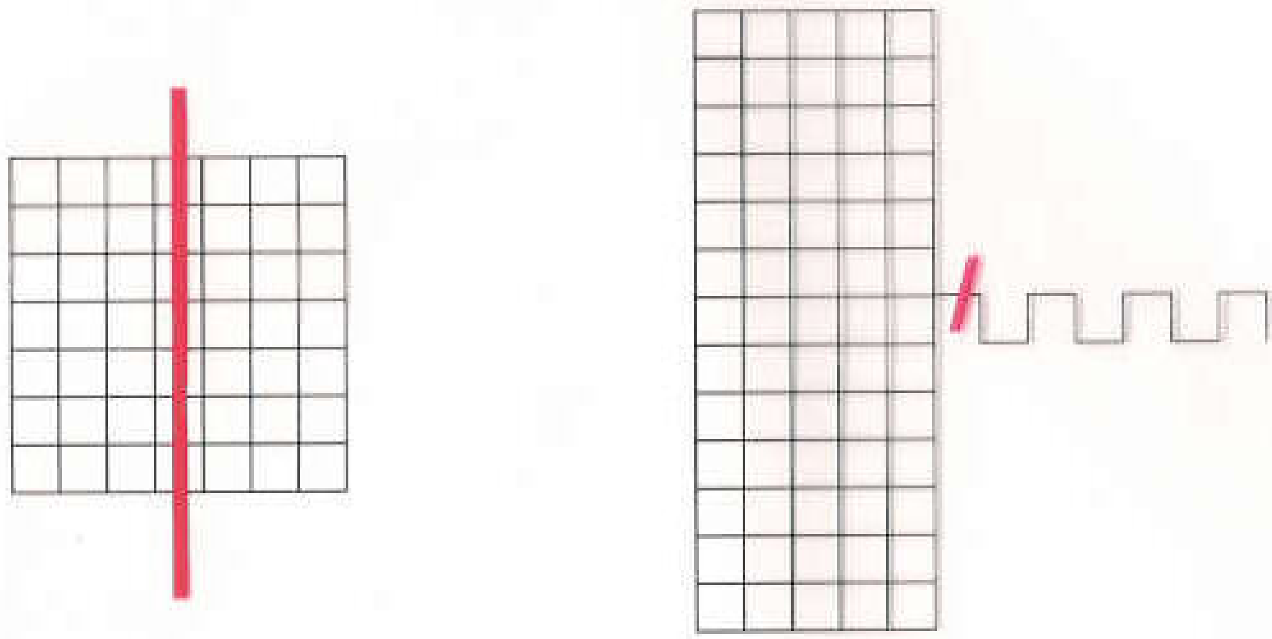
Spectral Methods, Graph Partitioning, and Clustering

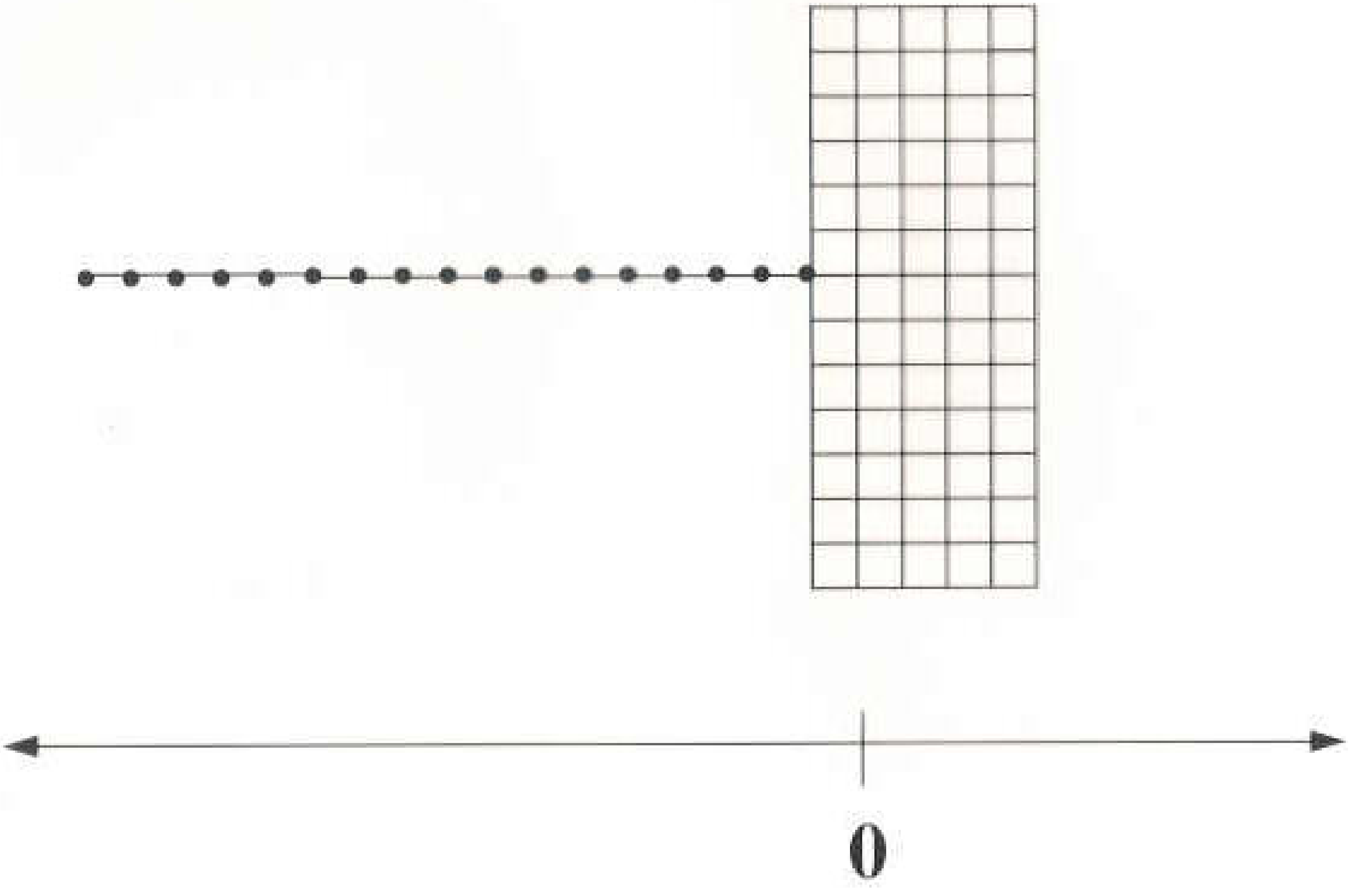
Shang-Hua Teng
Boston University/Akamai

Joint work with Daniel Spielman (MIT)



$$\text{Ratio} = \frac{\text{edges cut}}{\text{vertices removed}}$$

Eigenvector tries for good ratio cut



Planar graph eigenvalue bound

$$\lambda < \frac{8\,\Delta}{n}$$

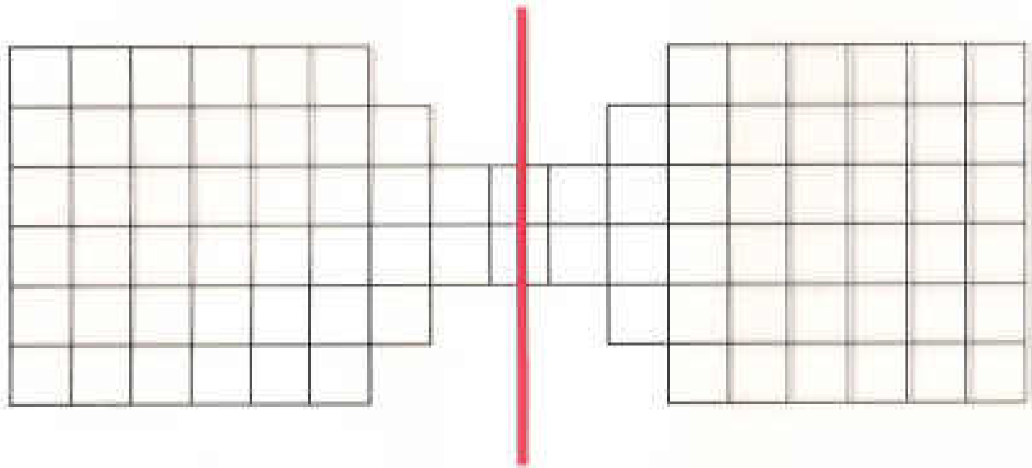
Δ = max degree

$$\lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} & & \\ & L & \\ & & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$\Downarrow$$

$$\lambda = \frac{\sum_{(i,j) \in E} (x_i - x_j)^2}{\sum_i x_i^2}$$

Graph Partitioning



Bisection

Motivation

- **Classical idea** (Donath-Hoffman; 1972)
- **Works well experimentally**
- **WHY? And ALWAYS?**
- **Graphs that arise in practice:**
  - **planar graphs**
  - **meshes, N-body graphs,**
  - **nearest neighbor graphs**
- **Other Applications: Data Clustering**

Future Research and Open Questions

- Constant–factor approximation of bisection
- Eigenvectors and multicommodity flow
- Spectral methods for combinatorial problems:
  coloring, clustering, ordering, independent sets
- Graph embedding and geometry of graphs

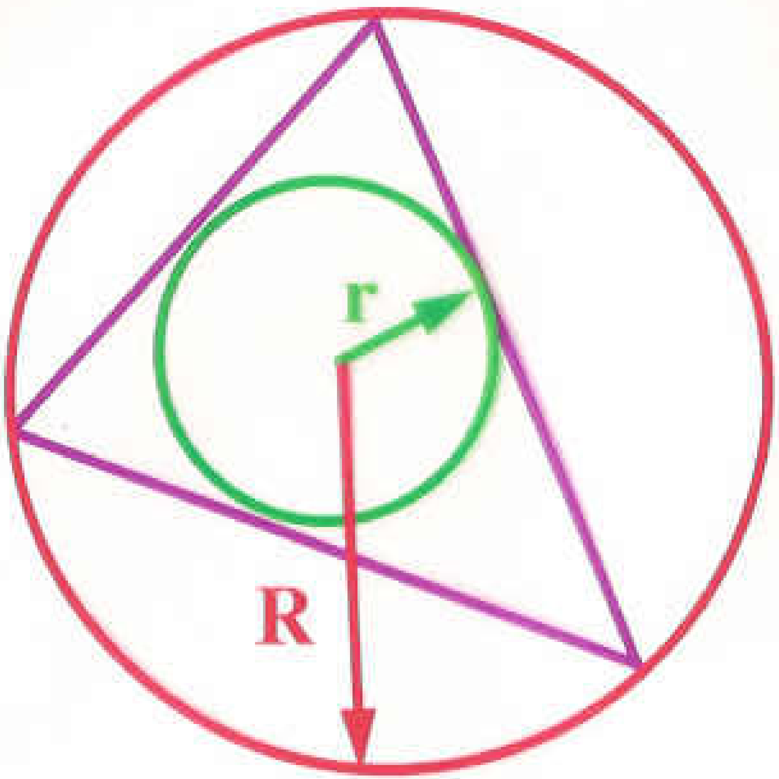
Small Separator Theorems

Trees	(1, 1:2)	**Jordan**
Planar Graphs	$(\sqrt{n}, 1:2)$	**Lipton–Tarjan**
Bounded–Genus Graphs	$(\sqrt{gn}, 1:2)$	**Gilbert–Hutchinson–Tarjan**
Bounded–Minor Graphs	$(h^{3/2}\sqrt{n}, 1:2)$	**Alon–Seymour–Thomas**
Nearest Neighbor Graphs	$(n^{1-1/d}, 1:d+1)$	**Miller–Teng–Thurston–Vavasis**
Finite–Element Meshes	$(n^{1-1/d}, 1:d+1)$	**Miller–Teng–Thurston–Vavasis**
N–body Graphs	$(n^{1-1/d} \lg n, 1:d+1)$	**Teng**

Spectral Separator Theorems

- Planar Graphs (Bounded Degree)
- Well–Shaped Meshes
- N–Body Graphs
- Nearest Neighbor Graphs

Well Shaped Mesh



R:r is bounded

Convergence of Kleinberg Algorithms

- Eigenvalues and Eigenvectors
  - $Ax = \lambda x$

- Related with Spectral method for graph partitioning (Spielman-Teng)
  - Principle eigenvector projects good localities.
  - Eigenvector can be used for partitioning and clustering

Akamai

Kleinberg's Algorithm

- Hubs: pages with links to many quality authorities
- Authorities: pages with links from many quality hubs
- Hubs (imagine a good text book and survey paper) and authorities (imagine Karp's first paper on NP complete problem).
  - $A(q) \sim \Sigma_{p \text{ in } IN(q)} H(p)$
  - $H(q) \sim \Sigma_{p \text{ in } OUT(q)} A(p)$

Akamai

Ranking Relevant Web-pages

- Use Link structures (Web-Graph)
  - Pages with high in-degree are important
  - Pages has links from important pages are important
- Model Web-graph as Markov Chains
  - Model random surfers
  - Roughly, let R(q) be the rank of a page q and let IN(q) be the set of page that refer q, then

$$R(q) \sim \Sigma_{p \text{ in } IN(q)} R(p) / N_p$$

- The rank is related with the singular vector of the web-matrix.

Akamai

Search Engines

Google

altavista:

Inktomi

infoseek

Northern Light

IBM

CLEVER Searching

IBM

QBIC™

© IBM Corporation

ResearchIndex

The NECI Scientific Literature Digital Library

YAHOO! MAPS

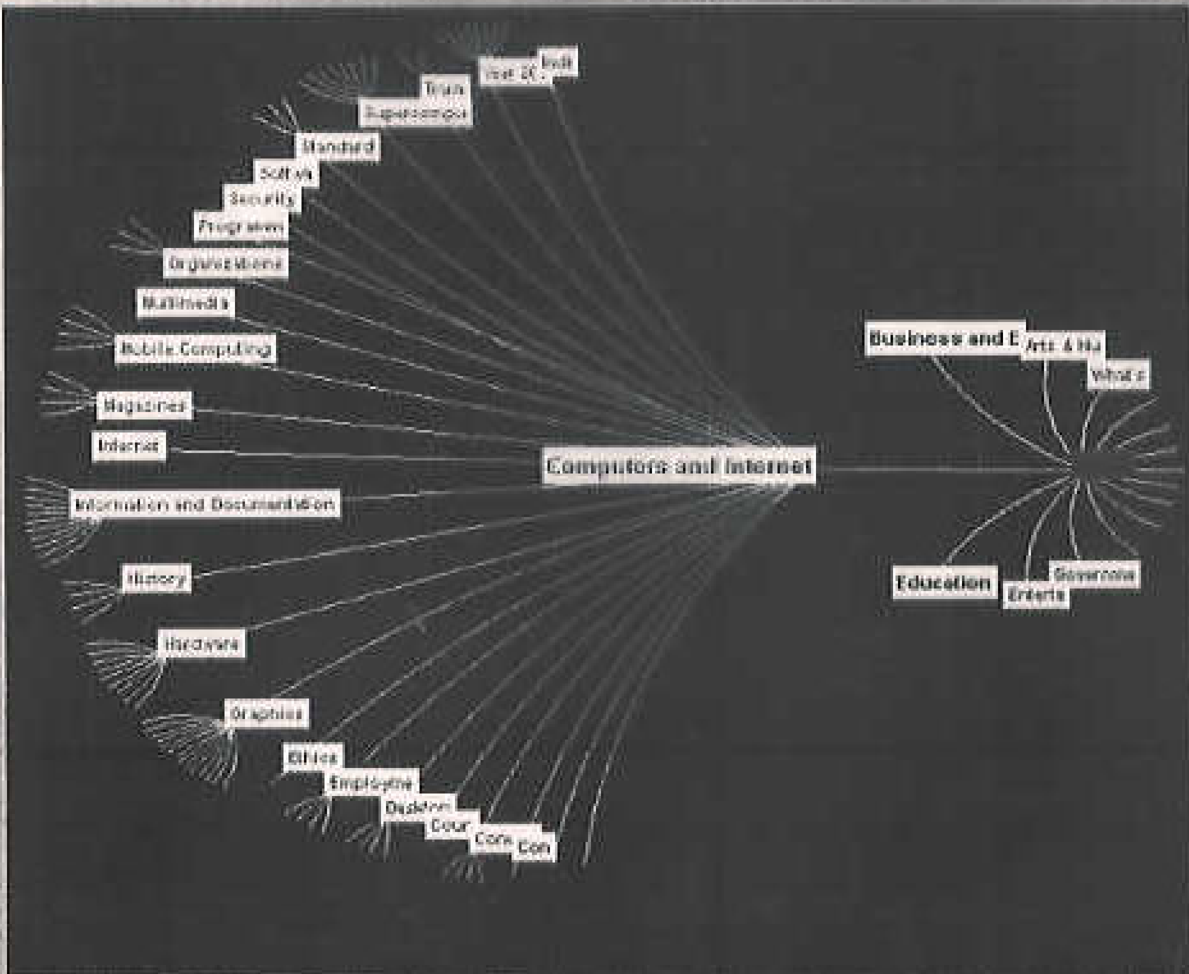
Akamai

Challenging Problems

- Searching Relevant Information
- Fast Delivery of Contents
- Secure Communication and Transaction
- Very Very Large Scale
- User Pattern Detection, and Profile Generation

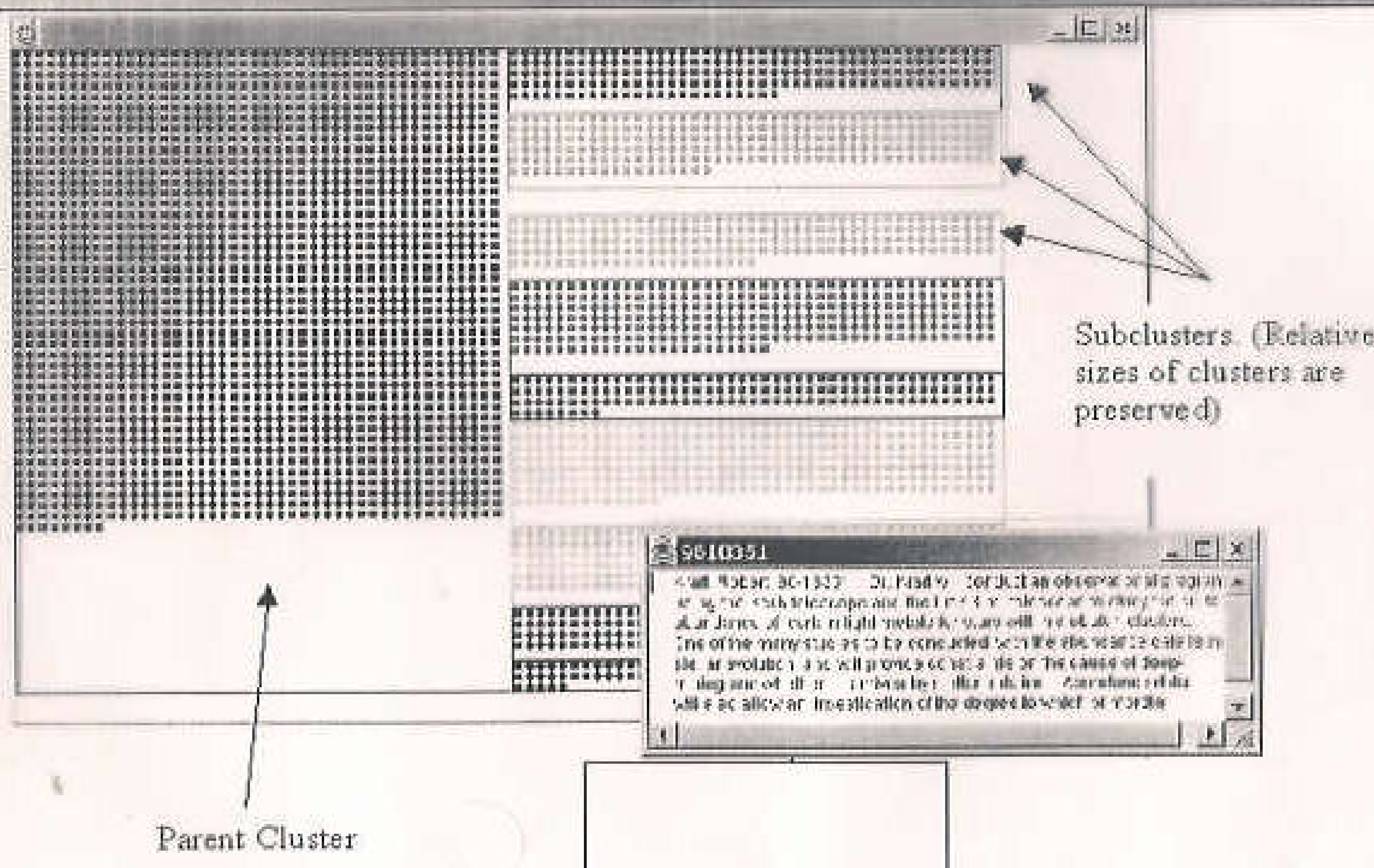
Akamai

Clustering and Hierarchy

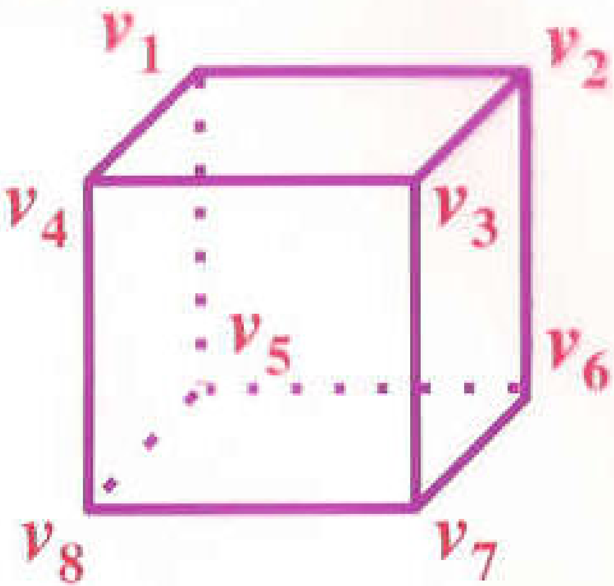


Akamai

Clustering and Hierarchy



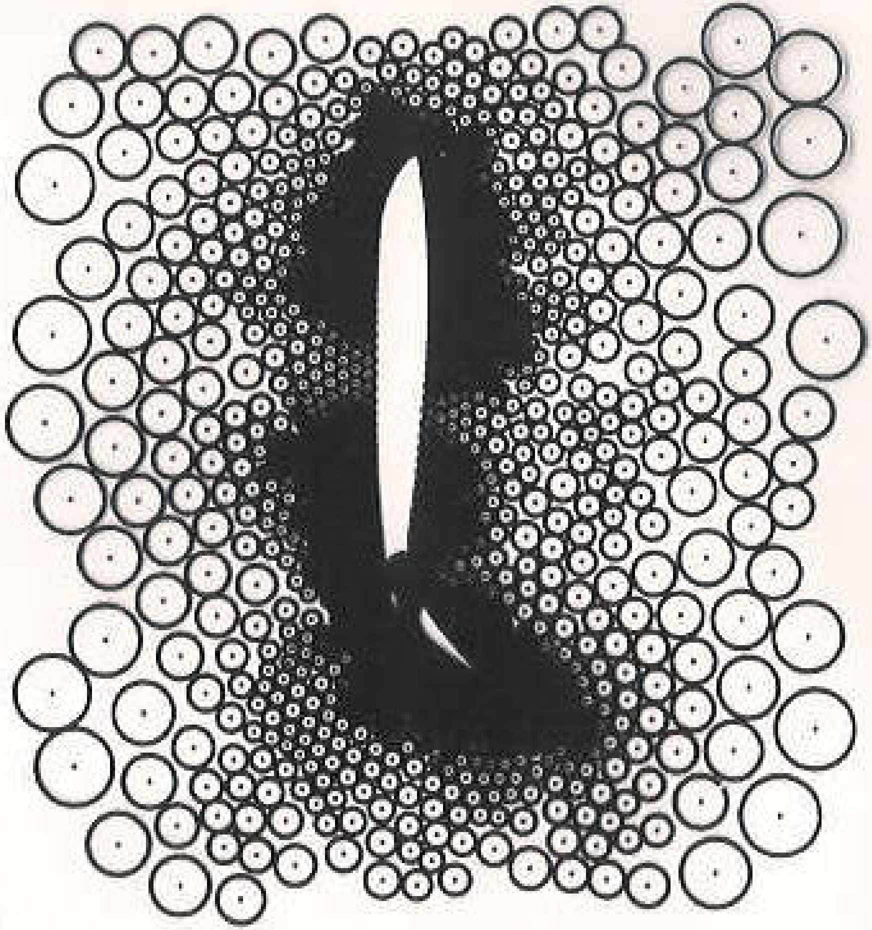
Embedding Lemma:



$$\lambda = \min_{\substack{v_1,\ldots,v_n \\ \sum v_i = 0}} \frac{\sum_{(i,j) \text{ in } E} \text{dist}(v_i, v_j)^2}{\sum \| v_i \|^2}$$

- Donath-Hoffman
- Fiedler
- Cheeger, Alon, Sinclair-Jerrum
- Pothen-Simon-Liou
- Guattery-Miller

	λ	cut size
2D mesh	$O(1/n)$	$O(n^{1/2})$
3D mesh	$O(1/n^{2/3})$	$O(n^{2/3})$

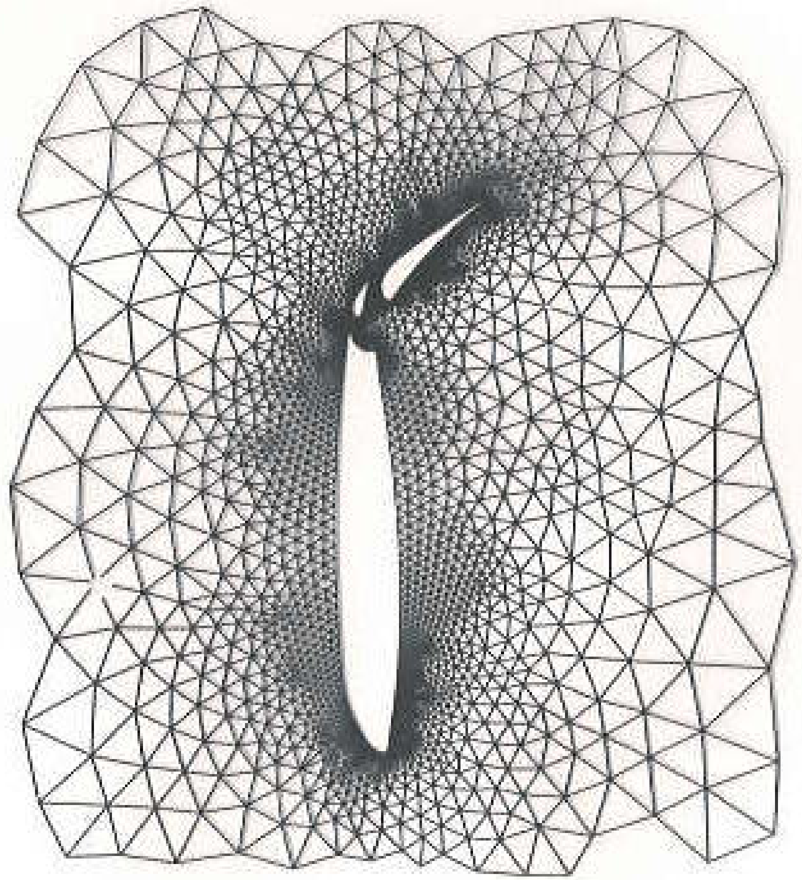


Well-Shaped Meshes and Sphere-Packings

Miller-Teng-Thurston-Vavasis

Well-Shaped Meshes ⟶ **Sphere-Packings**
Well-Shaped Meshes ⟵ **Sphere-Packings**

Miller-Talmor-Teng



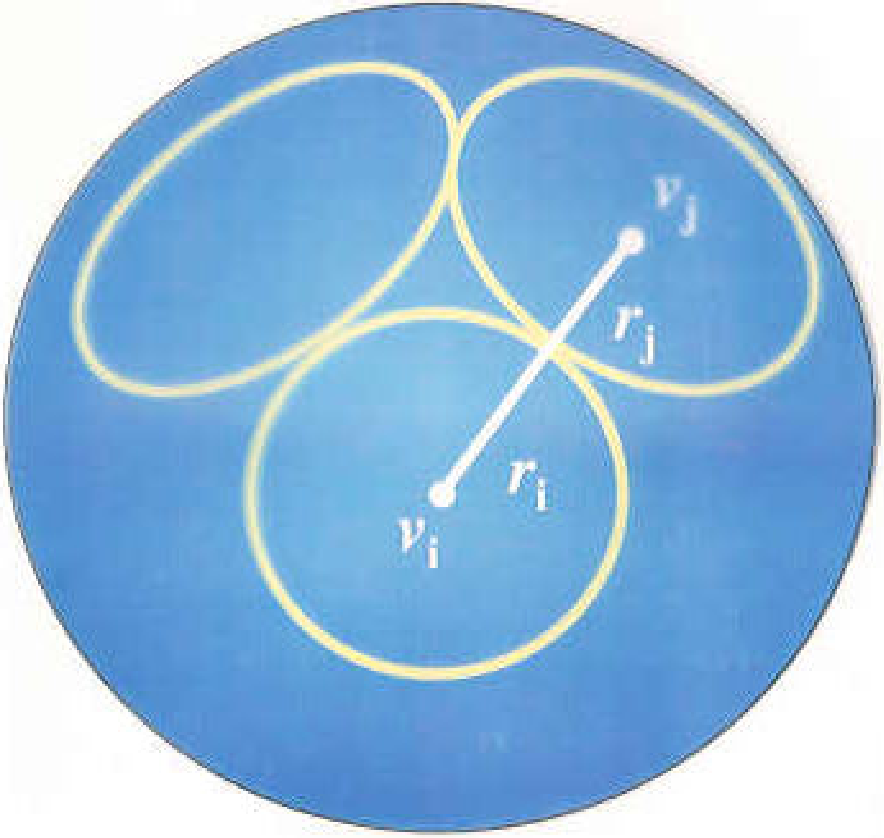
Planar graph eigenvalue bound

$$\sum_{(i,j) \text{ in } E} \text{dist}(v_i, v_j)^2 < 2\,\Delta \sum r_i^2 < 8\,\Delta$$

$$\lambda < \frac{8\,\Delta}{n}$$

Δ = **max degree**

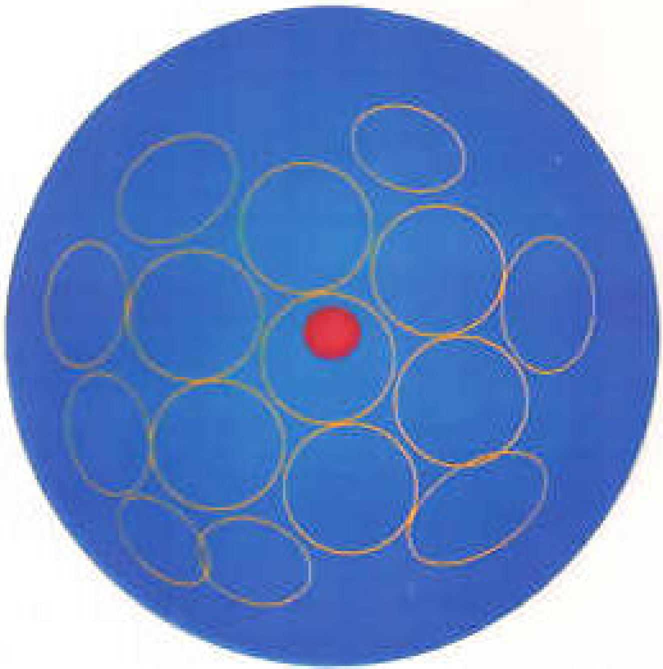
$$\Sigma \pi r_i^2 < 4\pi$$



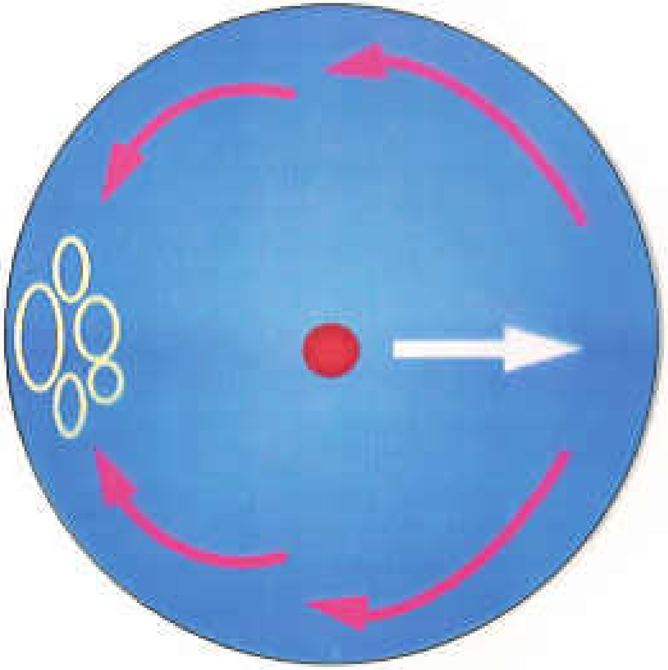
$$\mathrm{dist}(v_i, v_j)^2 < 2\, r_i^2 + 2\, r_j^2$$

Center of gravity
at
sphere center

Center of gravity at sphere center

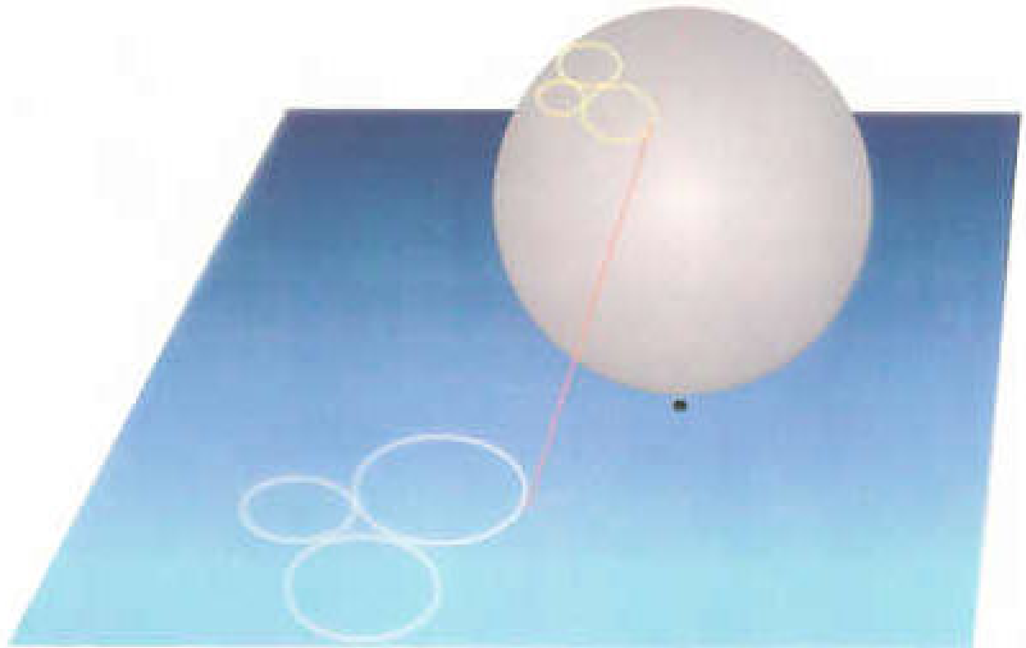


proof: **Brouwer's fixed point theorem.**



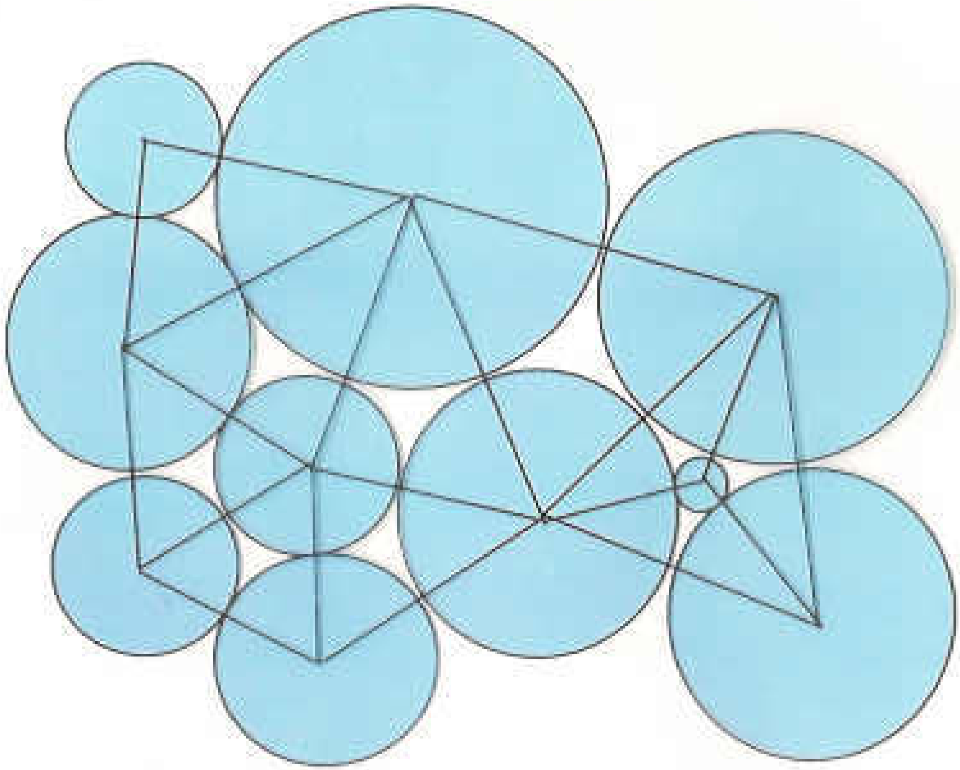
Use Brouwer's Fixed Point Theorem

Clustering



sphere-preserving map

Koebe-Andreev-Thurston Embedding Theorem:



kissing disks for planar graphs

Proof Outline

1. relate λ to quality of embedding

2. prove graphs have good embeddings

Rayleigh Quotient

$$\lambda_x = \frac{x^T L x}{x^T x}$$

⬇

cut ratio $< (2 \Delta \lambda_x)^{1/2}$

[Cheeger, Alon, Sinclair-Jerrum]

[Mihail]

Results:

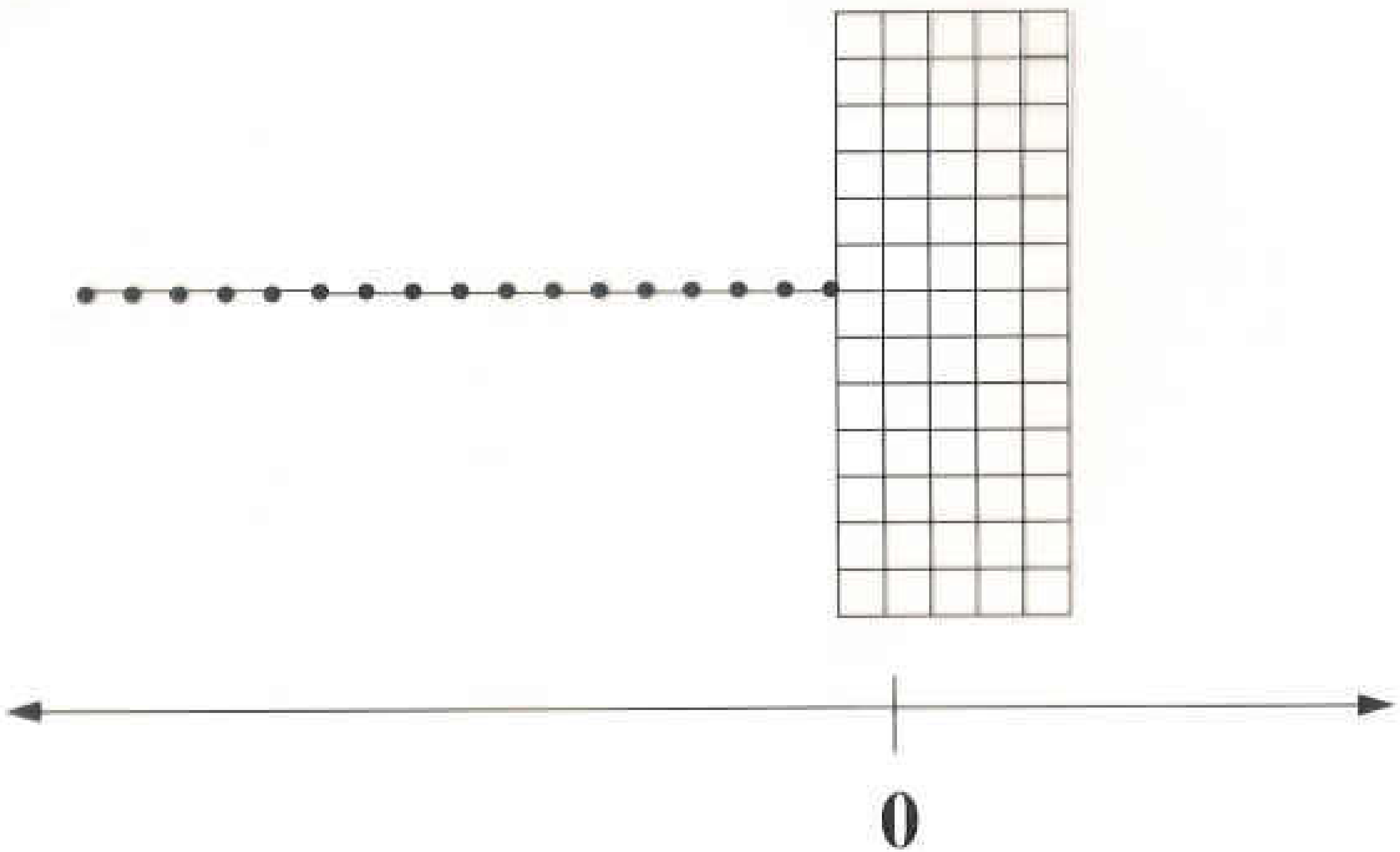
graph	λ	ratio	cut size
planar	$O(1/n)$	$O(1/n^{1/2})$	$O(n^{1/2})$
2D mesh	$O(1/n)$	$O(1/n^{1/2})$	$O(n^{1/2})$
3D mesh	$O(1/n^{2/3})$	$O(1/n^{1/3})$	$O(n^{2/3})$

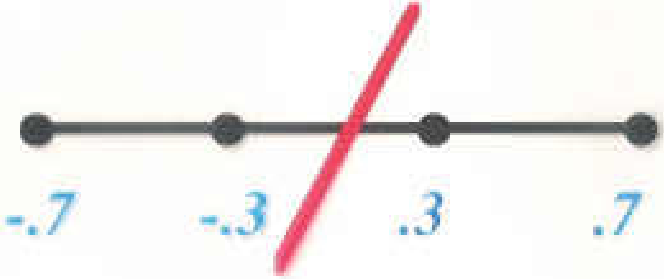
Spectral Partitioning

Spectral Methods Always Work

Myth

Bisection may fail:
eigenvector tries for good ratio cut

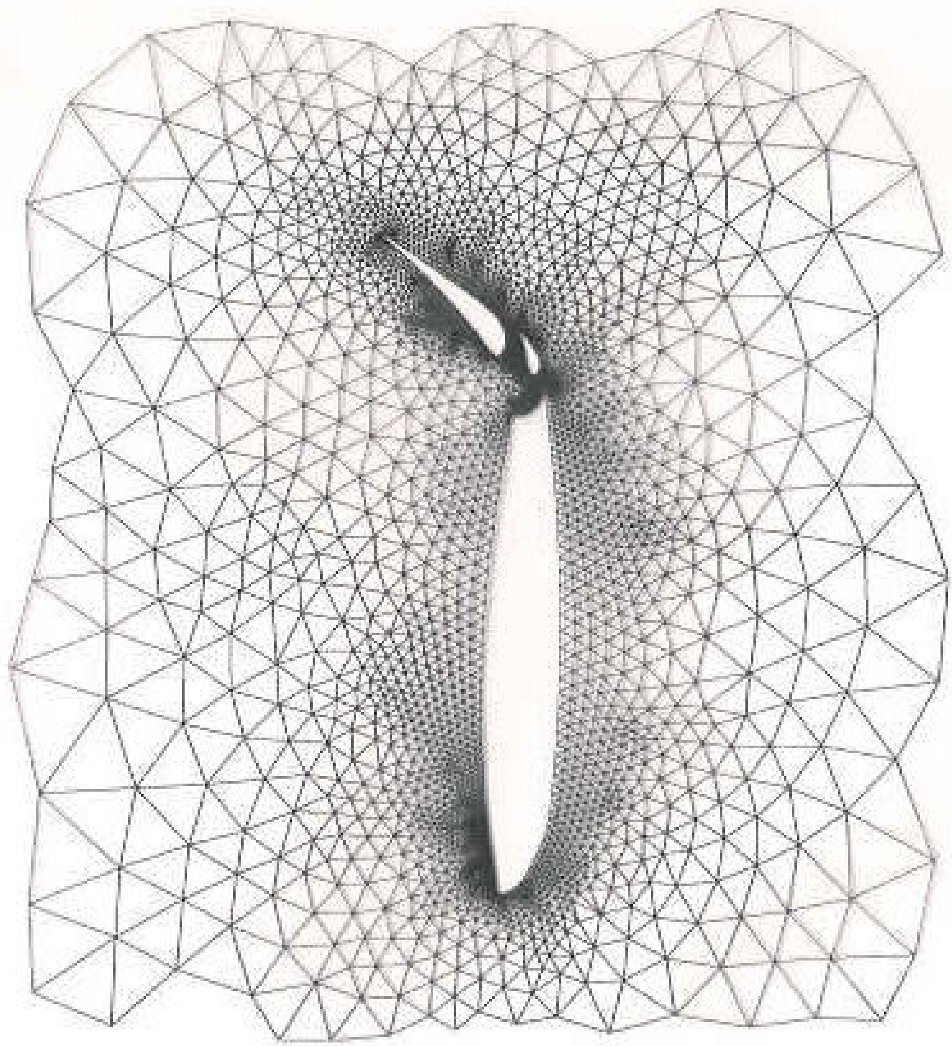


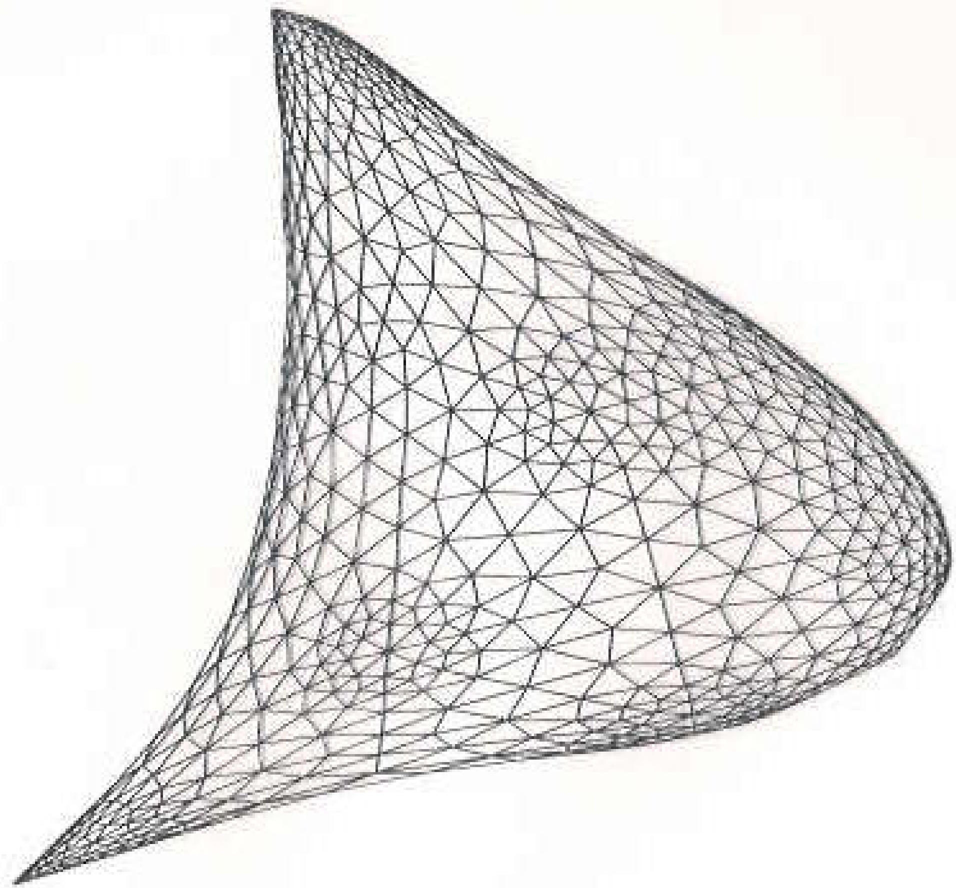


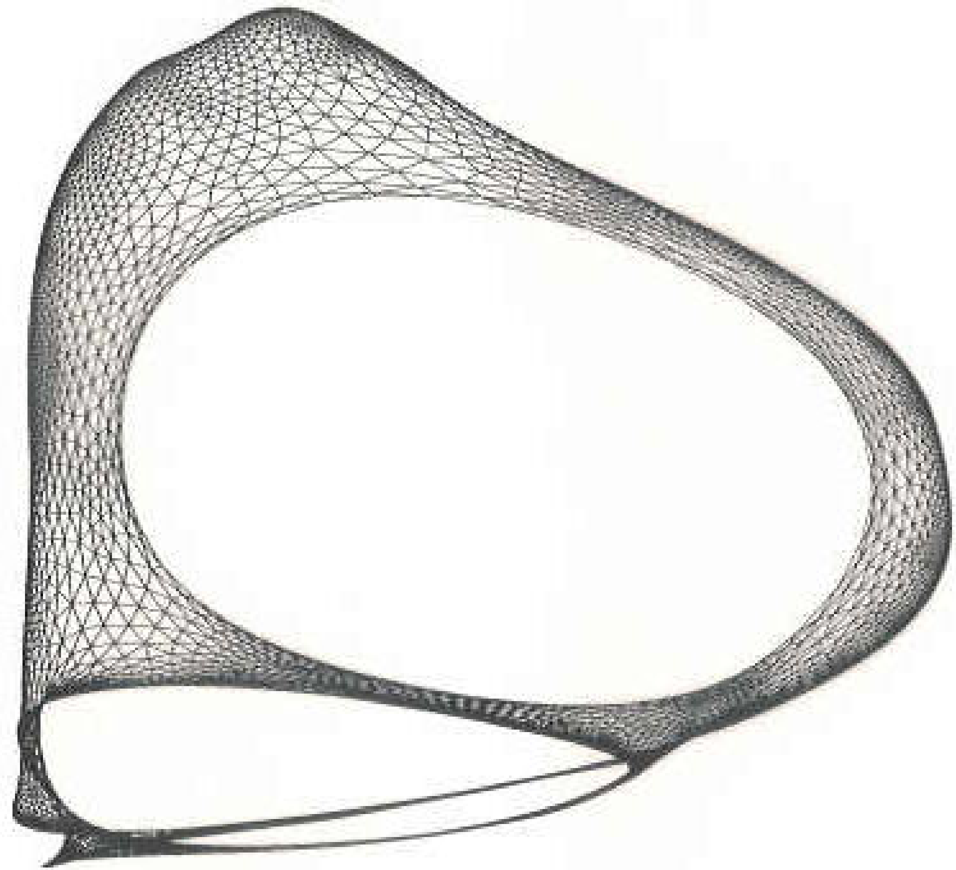
λ small ➡ cut of small ratio

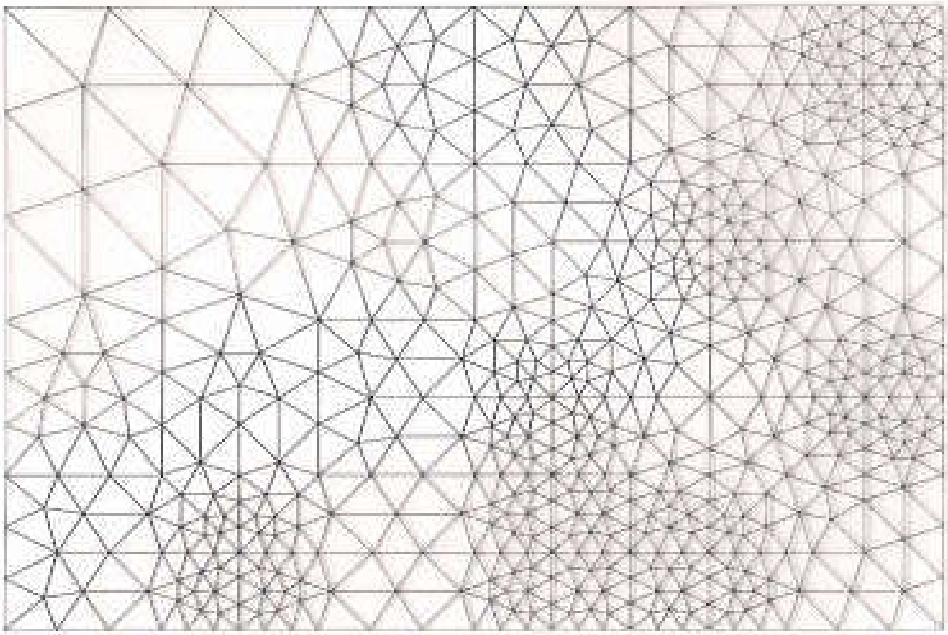
[Cheeger, Alon, Sinclair-Jerrum]

$\text{ratio} \sim \sqrt{\lambda}$









Spectral Embedding

X: Second Eigenvector

Y: Third Eigenvector

Smaller Eigenvalue, better locality

Rayleigh Quotient

$$\lambda_{\mathbf{x}} = \frac{\mathbf{x}^T L\,\mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{\sum_{(i,j)\ in\ E} (x_i - x_j)^2}{\sum x_i^2}$$

$$\lambda_2 = \min_{\mathbf{x} \perp \vec{1}} \lambda_{\mathbf{x}}$$

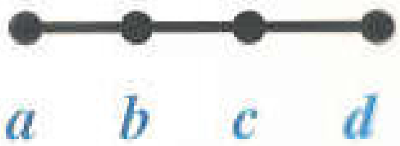
Small eigenvalues imply locality

Properties of Laplacian

(Assume G is connected)

- Symmetric
- $\lambda_1 = 0$
- $\lambda_2, \lambda_3, \ldots \lambda_n > 0$
- $\mathbf{x}^T L \mathbf{x} = \sum_{(i,j) \text{ in } E} (x_i - x_j)^2$

Laplacian of a Graph



$$
\begin{array}{c} \\ a \\ b \\ c \\ d \end{array}
\begin{array}{c}
\begin{array}{cccc} a & b & c & d \end{array} \\
\begin{pmatrix}
1 & -1 & 0 & 0 \\
-1 & 2 & -1 & 0 \\
0 & -1 & 2 & -1 \\
0 & 0 & -1 & 1
\end{pmatrix}
\end{array}
$$

$$L$$

Eigenvalue and Eigenvector

$$\boldsymbol{A}\,\boldsymbol{x} = \lambda\,\boldsymbol{x}$$

Partitioning Methods

Local Improvement	(Kernighan–Lin)
Multicommodity Flows	(Leighton–Rao)
Multilevel	(Bui–Jones: Chaco; MeTiS)
Geometric	(Miller–Teng–Thurston–Vavasis)
Spectral (Eigenvector–Based)	(Donath–Hoffman)

From Sparsest Partition to Bisection

If every subgraph of G of size x has a partition of sparsity

$$O(1/x^{\alpha})$$

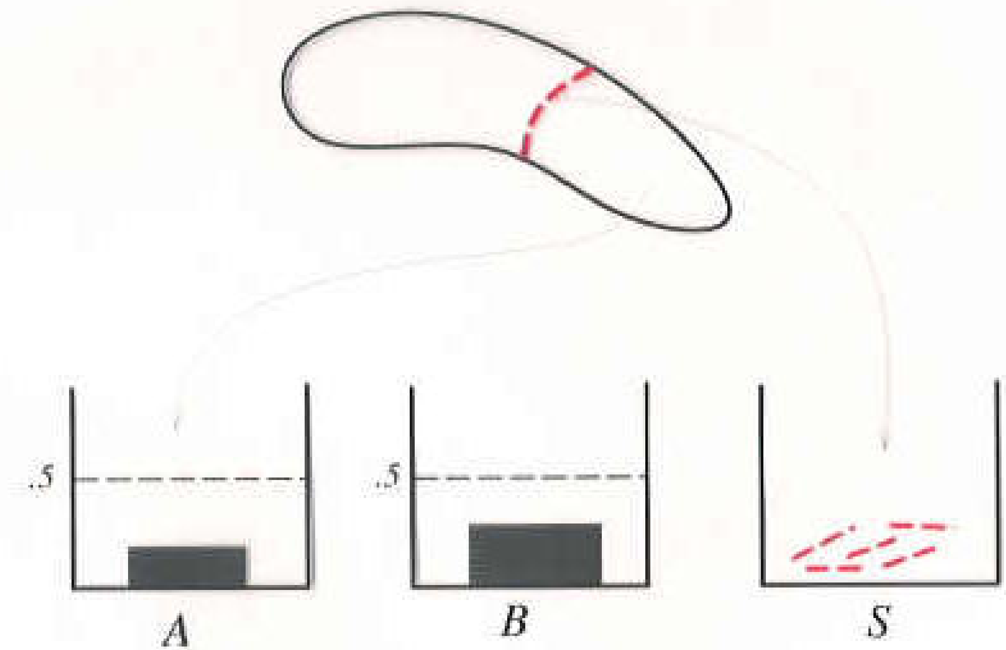
then G has a bisection of cut size

$$O(n^{1-\alpha}).$$

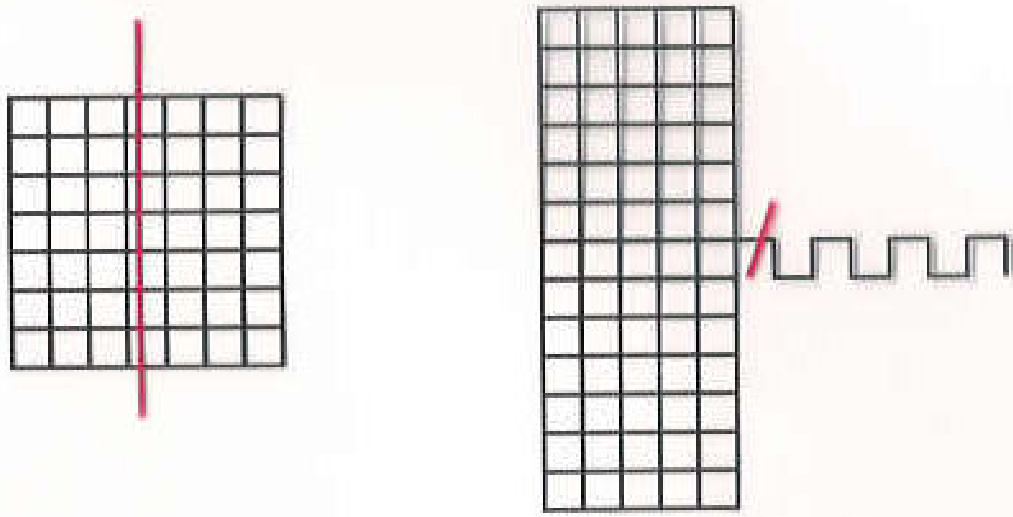
$$\phi(x) \rightarrow \int_1^n \phi(x)dx$$

$$O(1/x^\alpha) \rightarrow O(n^{1-\alpha})$$

From Sparse Cut to Bisection



Sparsity: Reduce Two Parameters to One



$$\text{Sparsity} = \frac{\text{Cut–Size}}{\text{"Volume" of the Smaller Side}}$$

Surface–to–Volume Ratio   Isoperimetric Number   Cut–Ratio

Applications

VLSI Design

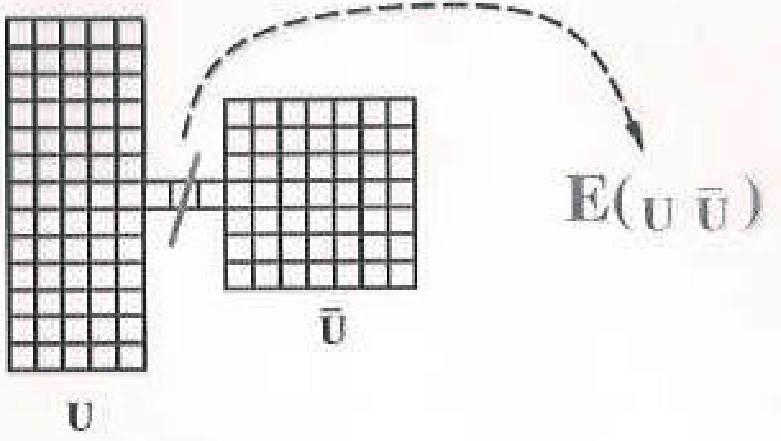
Parallel Processing

Scientific Computing

Information Organization

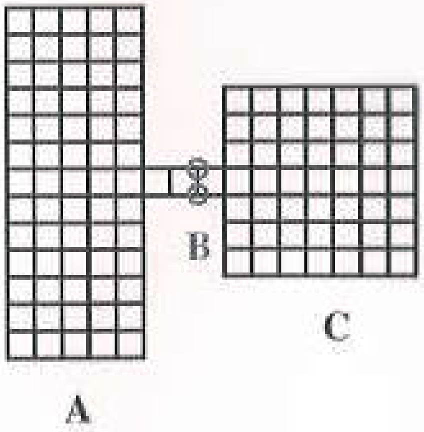
Efficient Search Structure

Partition and Bisection

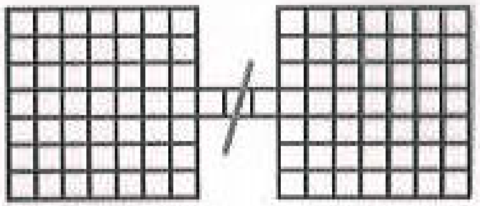


Cut–Size
Splitting–Ratio

Separator



Bisection



Graph Partitioning

Spectral Methods

Eigenvector and Eigenvector

Underlying Matrices

Classical Method (70's)

Many Variations and Software

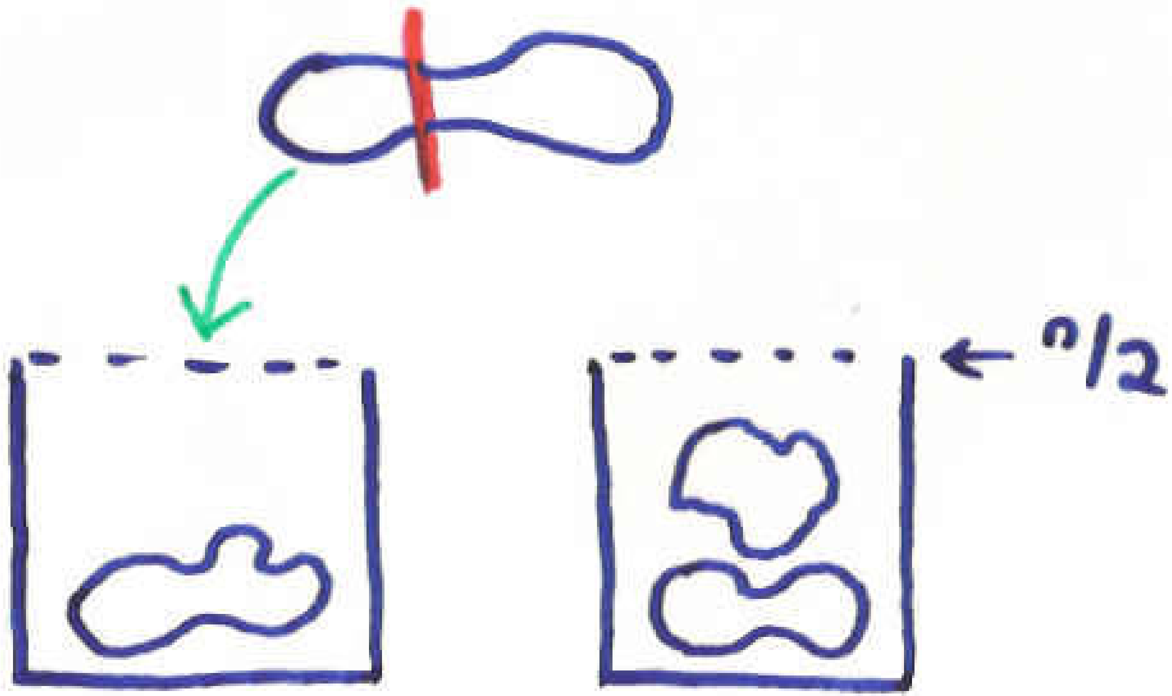
Great Experimental Results

Lack of Mathematical Justification

cut ratio → bisection

$$\phi(x) \rightarrow \int_1^n \phi(x)\,dx$$

$$O(n^{-\alpha}) \rightarrow O(n^{1-\alpha})$$



Eigenvector of a Graph

$$\lambda \begin{pmatrix} -.7 \\ -.3 \\ .3 \\ .7 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} -.7 \\ -.3 \\ .3 \\ .7 \end{pmatrix}$$

Eigenvector with min non-zero λ