Understanding graphs through spectral densities

David Bindel

Department of Computer Science
Cornell University

SCAN seminar, 29 Feb 2016
Can One Hear the Shape of a Drum?

\[-\nabla^2 u = \lambda u \text{ on } \Omega, \quad u = 0 \text{ on } \partial \Omega\]

Assume that for each \( n \) the eigenvalue \( \lambda_n \) for \( \Omega_1 \) is equal to the eigenvalue \( \mu_n \) for \( \Omega_2 \). Question: Are the regions \( \Omega_1 \) and \( \Omega_2 \) congruent in the sense of Euclidean geometry?

I first heard the problem posed this way some ten years ago from Professor Bochner. Much more recently, when I mentioned it to Professor Bers, he said, almost at once: “You mean, if you had perfect pitch could you find the shape of a drum.”

Mark Kac, American Math Monthly, 1966
Can One Hear the Shape of a Drum?

No in general (Gordon, Webb, Wolpert in 1992)
Yes with constraints (Zelditch in 2009)
What Do You Hear?

Size of bottlenecks (Cheeger inequality)

\[ h \leq 2\sqrt{\lambda_2} \]
Volume (Weyl law)

$$\lim_{x \to \infty} \frac{N(x)}{x^{d/2}} = (2\pi)^{-d} \omega_d \text{vol}(\Omega), \quad N(x) = \{\# \text{eigenvalues} \leq x\}$$

Can also tell lengths of geodesics for a closed Riemannian manifold.
Can One Hear the Shape of a Graph?

From eigenvalues of adjacency, Laplacian, normalized Laplacian?
A Bestiary of Matrices

- **Adjacency matrix:** \( A \)
- **Laplacian matrix:** \( L = D - A \)
- **Unsigned Laplacian:** \( L = D + A \)
- **Random walk matrix:** \( P = D^{-1} A \)
- **Normalized adjacency:** \( \bar{A} = D^{-1/2} A D^{-1/2} \)
- **Normalized Laplacian:** \( \bar{L} = I - \bar{A} = D^{-1/2} L D^{-1/2} \)
- **Modularity matrix:** \( B = A - \frac{dd^T}{2n} \)

All have examples of co-spectral graphs... through spectrum uniquely identifies *quantum graphs*
What Do You Hear?

Size of separators (Cheeger inequality) – $L$
What Do You Hear?

What information hides in the eigenvalue distribution?

1. Discretizations of Laplacian: something like Weyl’s law
2. Sparse random graphs: Wigner semicircular distribution
3. “Real” networks: less well understood

But computing all eigenvalues seems *expensive!*
Reminder: Spectral Mapping

Consider a matrix $H$, and let $f$ be analytic on the spectrum. Then if $H = V \Lambda V^{-1}$,

$$f(H) = V f(\Lambda)V^{-1}.$$  

(generalizes to non-diagonalizable case)
Spectra define a \textit{generalized function} (a \textit{density}): 

\[ \text{tr}(f(H)) = \int f(\lambda)\mu(\lambda)\,dx = \sum_{j=k}^{N} f(\lambda_k) \]

where $f$ is an analytic test function. Smooth out to get a picture: a \textit{spectral histogram} or \textit{kernel density estimate}. 

\[ \text{s}\int_{a}^{b} f(\lambda)\mu(\lambda)\,dx \]
Heat Kernels

DoS information equivalent to looking at the *heat kernel trace*:

\[ h(s) = \text{tr}(\exp(-sH)) = \mathcal{L}[\mu](s) \]

where \( H \) is a positive semi-definite operator.

\[ H = L \implies h(s)/N = \text{probability of self-return after time } s \text{ from uniform start} \]
Power Moments

DoS information equivalent to looking at the *power moments*:

$$\text{tr}(H^j).$$

Has a natural interpretation for matrices associated with graphs:

- $A$: number of length $k$ cycles.
- $\bar{A}$ or $P$: return probability for $k$-step random walk (times $N$).
- $L$: ??
Ordinary moments are not good for numerics – prefer Chebyshev:

\[ d_j = T_j(A) \]

where \( T_j(z) = \cos(j \cos^{-1}(z)) \) is the \( j \)th Chebyshev polynomial. Compute via three-term recurrence:

\[
\begin{align*}
T_0(z) &= 1 \\
T_1(z) &= z \\
T_{k+1}(z) &= 2zT_k(z) - T_{k-1}(z)
\end{align*}
\]
Exploring Spectral Densities

Kernel polynomial method (see Weisse, Reviews of Modern Physics)

- Think of spectral distribution on $[-1, 1]$ as a generalized function

$$\int_{-1}^{1} \mu(x) f(x) \, dx = \frac{1}{N} \sum_{k=1}^{N} f(\lambda_k)$$

- Write $f(x) = \sum_{j=1}^{\infty} c_j T_j(x)$ and $\mu(x) = \sum_{j=1}^{\infty} d_j \phi_j(x)$, where

$$\int_{-1}^{1} \phi_j(x) T_k(x) \, dx = \delta_{jk}$$

- Estimate $d_j = \text{tr}(T_j(H))$ by stochastic methods

- Truncate series for $\mu(x)$ and filter (avoid Gibbs)

*Much* cheaper than computing all eigenvalues!
$Z \in \mathbb{R}^n$ with independent entries, mean 0 and variance 1.

\[
E[(Z \odot HZ)_i] = \sum_j h_{ij}E[Z_iZ_j] = h_{ii}
\]

\[
\text{Var}[(Z \odot HZ)_i] = \sum_j h_{ij}^2.
\]

Serves as the basis for stochastic estimation of
- Trace (Hutchinson, others; review by Toledo and Avron)
- Diagonal (Bekas, Kokiopoulou, and Saad)

Power of diagonal estimation is under-appreciated...
Diagonal Estimation and LDoS

Diagonal estimation also useful for local DoS $\nu_k(x)$; in the symmetric case with $H = Q\Lambda Q^T$, have

$$\int f(x)\nu_k(x)\,dx = f(H)_{kk} = e_k^TQf(\Lambda)Q^Te_k$$

$$\nu_k(x) = \sum_{j=1}^{n} q_{kj}^2 \delta(x - \lambda_j)$$

DoS is sum of local densities of states:

$$\mu(x) = \sum_{k=1}^{n} \nu_k(x)$$
Can compute common *centrality measures* with LDoS

- **Estrada centrality**: \( \exp(\gamma A)_{kk} \)
- **Resolvent centrality**: \( [(I - \gamma \bar{A})^{-1}]_{kk} \)

Some motifs associated with localized eigenvectors at specific values:

- **Chief example**: Null vectors of \( \bar{A} \) supported on leaves.
- **Use** LDoS + topology to find motifs?

What else?
Phase Retrieval in Graph Reconstruction

Reconstruct graph from *completely resolved* LDoS at all nodes?

- Assume $H = Q\Lambda Q^T$
- No multiple eigenvalues $\implies$ know $|Q|$ and $\Lambda$
- Can we recover signs in $Q$?

Feels a little like phase retrieval...

Of course, we usually have noisy LDoS estimates!
Other methods

Golub and Meurant: Gauss quadrature for $\nu_k(x)$ via Lanczos
  - Good: No stochastic estimation error (vs KPM)
  - Bad: Separate Lanczos per node

Roder and Silver: Max entropy estimation
  - Good: Better resolution from Chebyshev moment estimates
  - Bad: More expensive computation per node

Under investigation: Hybrid approach.
Consider spectrum of normalized Laplacian (random walk matrix)
Approximate via KPM and compare to full eigencomputation

Things we know

- Eigenvalues in $[-1, 1]$; nonsymmetric in general
- Stability: change $d$ edges, have
  \[
  \lambda_{j-d} \leq \hat{\lambda}_j \leq \lambda_{j+d}
  \]

- $k$th moment = probability of return after $k$-step random walk
- Eigenvalue cluster near 1 ~ well-separated clusters
- Eigenvalue cluster near 0 ~ triangles connected by one node

What else can we “hear”?
Experimental setup

- **Global DoS**
  - 1000 Chebyshev moments
  - 10 probe vectors (componentwise standard normal)
  - Histogram with 50 bins

- **Local DoS**
  - 100 Chebyshev moments
  - 10 probe vectors (componentwise standard normal)
  - Plot smoothed density on $[-1, 1]$
  - Spectrally order nodes by density plot

Suggestions for better pics are welcome!
Internet topology (local)
Marvel characters (local)
Marvel comics
Marvel comics (local)
PGP (local)
Yeast (local)
And a few more...
Enron emails (SNAP)
US power grid (Pajek)
N = 326186, nnz = 1615400, 80 s (1000 moments, 10 probes)
$N = 1139905$, $nnz = 113891327$, 2093 s (1000 moments, 10 probes)