The Physics of Networks

INSNA Sunbelt XVII

San Diego, February 13-17, 1997

Andrew Seary and William Richards

School of Communication

Simon Fraser University

Burnaby, BC Canada

The Physics of Networks

Introduction

We present some recent interesting material from mathematical physics and graph theory that has not appeared in social network literature. We hope that this material will

- help in understanding the roles of eigenvalues and eigenvectors in network analysis
- suggest new directions and applications in network analysis

The Laplacian Operator in Physics

The Laplacian first appeared in solutions to the wave equation, and soon found application in many other areas of physics, such as the flow of heat (and other forms of diffusion). Many of the methods and functions of mathematical physics were developed to deal specifically with solutions to this operator, undoubtedly the most important and best understood mathematical idea in physics. For example, Fourier analysis was first developed to deal with the heat equation. [CH]

The Laplacian may be constructed from two simpler operators, each a generalisation of the derivative operator:

The gradient

$$grad\,(f) = \nabla f = \partial f/\partial x,\ \partial f/\partial y, \ldots \qquad \text{(a vector)}$$

measures the maximum rate of change with direction (Fig 1)

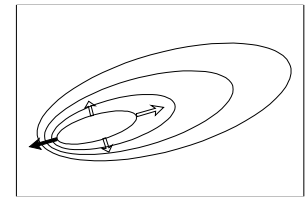

Figure 1

The divergence

$$div\,(f) = \nabla \cdot f = \partial f/\partial x + \partial f/\partial y + \ldots \qquad \text{(a scalar product)}$$

measures the net outflow of *f* in a unit of volume (fig. 2).

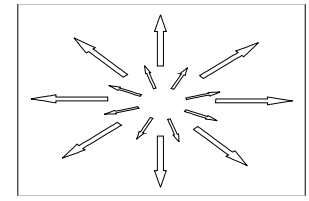

Figure 2

Both these names are very descriptive.

The Laplacian is

$$grad\,(f) = \nabla \cdot \nabla f = \nabla^2 f = \partial^2 f/\partial x^2 + \partial^2 f/\partial y^2$$

It is a divergence, and so a scalar product and a measure of flow. It may also be thought of as a measure of the difference between *f* at a point R, and the *average* of *f* in a region around R. This is possibly easier to understand for the discrete Laplacian, which we consider later (fig. 3).

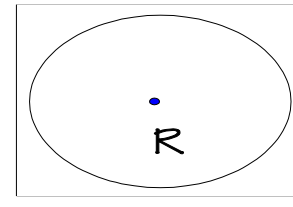

Figure 3

Solving the Laplacian

We present two typical examples of problems involving the Laplacian:

Wave Equation	**Heat Equation**
$\nabla^2 f = \partial^2 f / \partial t^2$	$\nabla^2 f = \partial f / \partial t$

The standard method for solving these equations since the 18th century is called "separation of variables"[CH], and assumes a solution of the form

$f = u(x,y,...)\, v(t)$	$f = u(x,y,...)\, v(t)$

that is, the solution is the product of two functions one of which depends only on x,y,... and the other only on t. Solutions are:

$\sum_k a_k \varphi_k(x, y, ...)\, e^{\sqrt{-\lambda_k}t}$	$\sum_k a_k \varphi_k(x, y, ...)\, e^{-\lambda_k t}$

(We are using special units to simplify these expressions). Notice that for the heat equation, the function of t is an exponential decay (which satisfies intuition for such a diffusive process), while for the wave equation, the function of t is a complex exponential, i.e. sines and cosines, (which also makes sense for an oscillatory process). [Cha],[Be]

In each case, the $\varphi_k(x, y, ...)$ are eigenfunctions of the Laplacian, and the λ_k are the eigenvalues. These depend only on the *geometry* of the "surface" under consideration.

For example, if we are solving either equation for a 2-dimensional rectangular surface, the $\varphi_k(x, y, ...)$ could be *cos(mx)cos(ny)* or *sin(mx)sin(ny)*. Many special functions were discovered by looking for the solution to the Laplacian in different geometries (e.g., Bessel functions in cylindrical coordinates, spherical harmonics in spheres, etc...). [CH]

Notice that we have completely separated the properties of the Laplacian to be properties of the *space* in which it acts. Some of the spaces considered in exploring the properties of the Laplacian can be very complex indeed (having curved "surfaces", possibly with "holes" and "bridges" fig. 4). In an important sense we can explore the properties of these "surfaces" (the general term in n-dimensions is "manifold") by exploring the properties of the Laplacian on these "surfaces".

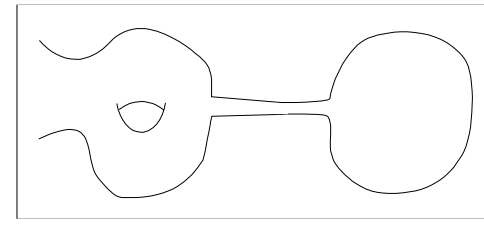

Figure 4

Properties of the Laplacian

We are interested in three properties that can also be applied to the discrete Laplacian of a graph: Boundary Conditions, Hilbert Nodal Theorem, and Cheeger's constant.

Boundary conditions:

We fudged a bit when stating above that the $\phi_k(x, y, \dots)$ could be $cos(mx)cos(nx)$ or $sin(mx)sin(nx)$ — which is it? The answer depends on whether conditions are specified for the boundary of the surface. We will consider two types of conditions: the boundary is fixed to some value (e.g., a vibrating string or drum), or there is no boundary condition at all (this will generally be the case for graphs). [CH]

It is probably easiest to understand the role of boundary conditions by considering the very simple case of a vibrating string. Here the difference in the two cases is captured by the difference between the sine (which goes to zero at both ends of the string) and the cosine (which goes to one at each end). Thus the eigenfunctions for a string with ends fixed are

$$\sin(x\sqrt{\lambda_k}); \quad 0 < \lambda_1 < \lambda_2 \leq , \dots \infty \quad \text{(fig. 5)}$$

and without ends fixed are

$$\cos(x\sqrt{\lambda_k}); \quad 0 = \lambda_1 < \lambda_2 \leq , \dots \infty \quad \text{(fig. 6)}$$

The "no boundary" solution has an eigenvalue of 0, and the associated eigenfunction is a constant.

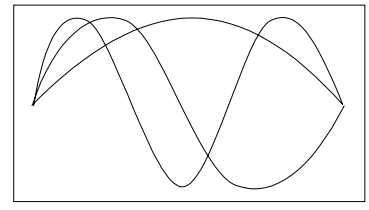

Fig. 5

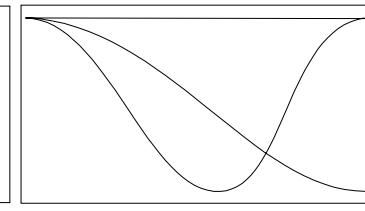

Fig. 6

Hilbert Nodal Theorem

A node is a point at which an eigenfunction goes to zero. On either side of the node, it has opposite sign and between nodes, the sign is constant. Regions between nodes are called *nodal regions*. The nodal theorem states that there are no more than k nodal regions for the kth eigenfunction. Again, the vibrating string shows this. The lowest eigenfunction (in either case) has constant sign, and so one nodal region. The second eigenfunction (in either case) has one node and two nodal regions and so on. (See figs. 5 and 6) Note that nodal regions can be much more complex for dimensions greater than one, where nodal lines can be formed by linear combinations of eigenfunctions with multiple eigenvalues. See CH p. 302 for examples of this. Nodal regions are also closely connected to boundary conditions. If we apply the nodal boundary conditions of eigenfunction k to a nodal region, then the corresponding node of that eigenfunction will be the "first" eigenfunction of the region with eigenvalue λ_k. In addition, some of the higher eigenfunctions must conform to the boundary condition imposed by the nodal region. This is true whether we have an initial boundary condition or not. In either case, it is "boundary conditions all the way up" (to higher eigenvalues and associated eigenfunctions). This will be important when we consider different types of graph decomposition.[CH]

Cheeger's Constant

Cheeger's constant [Che] is an isoperimetric property of a "surface" or manifold. An example of an isoperimetric problem is "what shape contains the largest volume for the least surface area"? In Euclidean space, the answer is a sphere (a circle in 2 dimensions, a hyper-sphere in n dimensions). [Ba]

Cheeger's constant is derived from the following problem:

Let E be a surface with area A, which divides a manifold into two parts with volumes V_1 and V_2.

Define: $h(M) = inf_E A(E)/\min(V_1,V_2)$

where the infimum is over all possible E (and the resulting V_1 and V_2). (Note: these "areas" may be n-1 dimensional with n dimensional volumes)

Cheeger showed that

$$h^2/4 \le \lambda_2$$

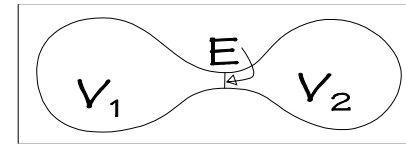

Figure 7

where λ_2 is the second-smallest Laplacian eigenvalue (for either type of boundary condition). To picture this problem, consider the "dumb-bell" shape in fig. 7. Obviously, E will be a circle somewhere on along the "bottleneck" between V_1 and V_2. Also, obviously the second eigenfunction will be positive in V_1 and negative in V_2. The inequality is not difficult to prove for either type of boundary condition. This lower bound on the second eigenvalue has important consequences in the discrete case, as we will see.

The discrete Laplacian

We will derive the discrete Laplacian from discrete versions of the gradient and divergence. It will turn out to have a very simple form for graphs. [Bi],[Mo],[Me]

Discrete gradient

Here we want to model first differences, but our geometry is very simple: all distances are 1 (so the /dx term is not needed), and we only have differences between nodes that are connected. We can model the gradient by constructing a *signed incidence matrix* (or matrix of dyads) with edges represented by rows, and nodes represented by columns. We assign a pair of 1 and -1 for each edge arbitrarily (this will make no difference for the discrete Laplacian). Example: a path of 5 nodes has as gradient the matrix:

```
      nodes
e     1  2  3  4  5
d 1   1 -1  0  0  0
g 2   0  1 -1  0  0
e 3   0  0  1 -1  0
s 4   0  0  0  1 -1
```

edges and nodes are numbered arbitrarily

The discrete Laplacian is formed from the scalar products matrix of this array (post-multiply the matrix by its transpose) and is

```
 1 -1  0  0  0
-1  2 -1  0  0
 0 -1  2 -1  0
 0  0 -1  2 -1
 0  0  0 -1  1
```

We see that the arbitrariness of assigning edges, nodes and signs does not matter. In fact the Laplacian can always be formed from the adjacency matrix by subtracting it from a matrix with node degree along its diagonal: L = D - A where

$$D = diag \sum_j a_{ij}$$

Note that this means the rows (and columns) of L sum to 0. Also, there is nothing in the definition L = D - A that requires A to be binary. We can see that L does a form of averaging in the neighbourhood of each node.

Now we will examine how the three properties of the continuous Laplacian mentioned above apply to the discrete Laplacian.

Boundary conditions

Generally, we do not consider that a graph has a boundary, but there is a famous example in graph theory where this idea is used: Tutte's algorithm for drawing a 3-connected planar graph[T]. Tutte's idea is based on finding the largest non-separating cycle, and then fixing the positions of those nodes on a regular polygon: this is the boundary condition. The positions of the rest of the nodes are determined by solving the a set of linear equations involving the Laplacian. If the graph is planar, this always gives a solution with no edge crossings. (fig. 8)

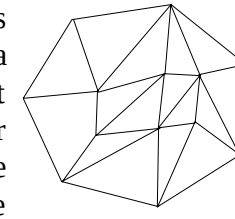

Figure 8

Graphs with mixed boundary conditions may be defined by considering *some* of the nodes as "inputs" and "outputs". The idea of graphs with boundaries is examined in some detail in (Friedman, 1993)[Fr]. He shows that with a suitable definition of boundary conditions, the nodal theorem holds as well. For the Laplacian of a graph with no boundary, we have:

$$0 = \lambda_1 \le \lambda_2 \le , \ldots , \lambda_n \le n$$

where $\lambda_2 = 0$ only for graphs with more than one component. The eigenvectors are orthonormal, with $(1,1,...,1)/\sqrt{n}$ as the eigenvector belonging to eigenvalue 0. Thus all the other eigenvectors must have zero means.

Nodal Theorem

The discrete version of the nodal theorem is states that there are no more than k *connected* regions for the *k*th eigenvector[Fr]. Here a nodal region is one in which all nodes must have the same sign. The signs of the *k*th eigenvector induce no more than *k* disconnected sets of nodes. (fig. 9) This important idea, along with the idea that boundary conditions hold "all the way up" to the highest eigenvalue, is useful in examining the results of two common methods for decomposing graphs based on eigenvectors: global and recursive. We know that the global method satisfies nodal boundary conditions up to the highest eigenvalue. A recursive method (such as Recursive Spectral Bisection)[PSL] uses the partition induced by the *second* eigenvector (see below) to divide a graph into two parts, and each part is then independantly sub-divided. This, in effect, starts with a new set of (no) boundary conditions with each bisection, and may not produce the same set of partitions as the global method.

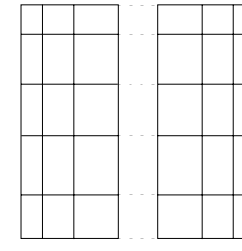

Figure 9

There is a more serious problem with a method like CONCOR[BBA] which (in effect) uses the second eigenvector of the *adjacency* matrix (A, not D-A) which may belong to a *negative* eigenvalue. If the graph is nearly bipartite, the result will be a partition into two sets of nodes *which are not graphs*: they will have few or no connections at all. This is not a problem for the global method, since we know that a partition based on a large eigenvalue must conform to boundary conditions of the underlying graph.

Cheeger's Constant

Here we will consider why the *second* Laplacian eigenvector is useful in partitioning graphs. Cheeger's constant has the following discrete analogue:

$$h(G) = inf_E(\#E)/\min(\#A, \#B)$$

where E runs over collections of edges such that G-E decomposes into two pieces A and B, and #E counts the number of edges, while #A and #B count the number of nodes in A and B. The Cheeger inequality then becomes:[Br]

$$h(G)^2 / 4k^2 \leq \lambda_2$$

where k is the largest degree of the graph G There is also a useful upper bound on the second eigenvalue:[Br]

$$2h(G) \geq \lambda_2$$

(the continuous version is not as useful) We can use h(M) to bound λ_2 in the continuous case, since the latter can be difficult to calculate. The opposite holds true for the discrete version: calculating h(G) is co-NP complete, while λ_2 is polynomial.

We can interpret λ_2 as a measure of how connected a graph is[Fe]. Recall that λ_2 can only be 0 if the graph is not connected. A graph is connected if and only if every set of vertices S has at least one neighbour outside of S. We can define an *expander graph* as one in which *every* set S (up to 1/2 the number of nodes) has *more* than #S neighbours outside of S, and we would expect that λ_2 would be large for such a graph, since the Cheeger constant would also be large. This insight led to explicit constructions of expander graphs, which are of great theoretical importance in Computer Science and Algorithmics [A,AM,D]. We can generally typify such graphs as having many paths between nodes, and will consider them again below. Also, we can see that graphs with small λ_2 (and Cheeger constant) will have "bottlenecks" and thus well defined nodal regions induced by the second eigenvector. Conversely, we can typify graphs with *small* λ_2 as having *few* paths between *some* nodes. The eigenvector belonging to a small λ_2 can be used to partition a graph, and it can also contain some useful structural information (fig. 10)[H]

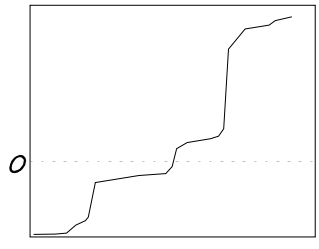

Figure 10

As an example of an application of the Laplacian to network analysis, Maas (1987)[Ma] presents a number of interesting results based on the second Laplacian eigenvalue and eigenvector of a transportation network, with a diffusive time-dependent part of $f(x,t)=u(x)v(t)=L(G)e^{-kt}$. He considers

three properties of the network: permeability, connection, and position. For example, he defines the half-life of the diffusive process as

$$\tau = -ln2 / (-k + \sqrt{k^2 + \lambda_2})$$

and examines different graphs with the same number of nodes. Large λ_2 results in small τ and greater permeability.

Node *i* is considered to have better position than node *j* if

$$\sqrt{\sum_k (u_{2k}, v_i)^2} \quad < \quad \sqrt{\sum_k (u_{2k}, v_j)^2}$$

where u_{2k} are the eigenvectors of (possibly multiple) λ_2 and where v_i and v_j have 1 in the *i,j* positions, zero elsewhere. (This is yet another definition of centrality).

Two nodes *i,j* are considered to be better connected than two nodes *k,l* if

$$\sqrt{\sum_k (u_{2k}, v_{ij})^2} \quad < \quad \sqrt{\sum_k (u_{2k}, v_{kl})^2}$$

where v_{ij} has 1 in position i and -1 in position j, 0 elsewhere, and similarly for v_{kl}.

He then goes on to show how to estimate upper and lower bounds on λ_2 (and therefore permeability) as connections are added and removed.

Early theoretical work on expander and random graphs concentrated on *regular* graphs. This allowed easy translation between the second-smallest Laplacian eigenvalue and the second-*largest* adjacency eigenvalue. Recently, the theory has been extended to random walks on non-regular graphs.

Random walks on a graph and Markov Chains

An obvious model for social networks is a Markov chain[W], where a probablility transition matrix M is built from the underlying graph by dividing each row of the (symmetric) adjacency matrix by the node degree D_r

$$M = D_r^{-1} A(G)$$

The matrix M has eigenvalues $1 = \mu_1 \geq \mu_2 \geq \ldots \geq \mu_n \geq -1$ where $\mu_2 = 1$ only if the graph has more than one component, and where $\mu_n = -1$ only if the graph is bipartite.[DS]

The left and right eigenvectors of M are biorthogonal, with $(\psi_i, \varphi_j) = \delta_{ij}$. Here, $\varphi_1 = (1,1,\ldots,1)/\sqrt{n}$ and ψ_1, the *stationary distribution*, is proportional to the node degrees.

We may also consider ϕ to be orthonormal in the metric defined by D_r:

$$\varphi_i D_r \varphi_j = \delta_{ij}$$

If we do this, then we can write M as a Laplacian in the D_r metric as:[DS]

$$L = I_n - M$$

Since this is a subtraction from the identity matrix, the eigenvectors do not change, and the new eigenvalues are

$$\lambda_i = 1 - \mu_i \text{ and } 0 = \lambda_1 \le \lambda_2 \le \ldots \lambda_n \le 2$$

As usual, $\lambda_2 = 0$ only if G has more than one component. In fact, the three properties of the Laplacian that we have been considering follow through. The only one we need to work on is the Cheeger inequality, which can be expressed as follows:

$$\Phi^2/4 \le \lambda_2 = 1 - \mu_2 \le \Phi$$

where $\Phi(G)$ is the *conductance* of a graph, calculated as

$$\Phi(G) = inf_E \, n\#E / d\#A\#B$$

where *d*=maximum degree and #E, #A, and #B are as above. The minimum is taken over all sets of edges.

The *mixing rate*[DS],[LS] of a Markov chain is a measure of how fast the chain converges to its limiting stationary distribution, and can be expressed as:

$$\lim_{t \to \infty} |M - 1/n|^{1/t} = (1 - \Phi^2/4)$$

Thus Markov chains with large λ_2 reach stationarity very quickly, and if the stationary state has roughly equal values (G is nearly regular), this stationary state is an almost uniform distribution. This property makes graphs with large λ_2 valuable as probability amplifiers in constructing randomised algorithms. This is an application to construction of algorithms which can give good approximate answers to "hard" problems.

As usual, graphs with "bottlenecks" have small λ_2 (and μ_2 near 1). Some other useful measures from random walks are (with $(u,v) = uD_rv$):[MR], [L]

- *mean access time* (expected time to reach *j* from *i*):

$$h_{ij} = n \sum_{k=2}^{n} (v_{kj}^2 - (v_{ki}, v_{kj}))/1 - \mu_k$$

  where access time will be large when μ_2 is near 1 (λ_2 near 0).

- *mean commute time* (expected time to reach *j* from *i* and then return to *i*):

$$= n \sum_{k=2}^{n} (v_{ki} - v_{kj})^2/1 - \mu_k$$

- *mean cover time* (expected time to visit *all* nodes):
  Let C_u be the mean access time to all other nodes from u then mean cover time is

$$\max_u C_u \le (1 + 1/2 + \ldots + 1/n) \max_{ij} h_{ij}$$

We may also use the definitions of Maas, again with the inner product (u,v) defined as $u\,D_r v$.[SR]

Finally, we note that the Matrix M is exactly the matrix used in Correspondence Analysis [G] of symmetric graphs, with the same eigenvalues (although most implementations of CA ignore eigenvalue sign) and eigenvectors (at least in direction: CA multiplies the eigenvectors by the eigenvalues).[SR] The clustering seen in the 1, 2 and 3-dimensional displays of CA are nodal regions which are "locally" rapidly mixing (fig. 11).

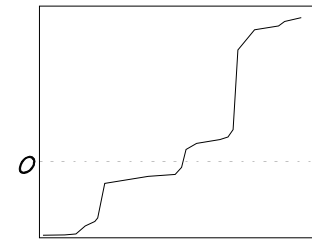

Figure 11

Conclusions

We have presented some of the work done in 1970 in mathematical physics, as well as some recent results obtained in graph theory and probabilistics which show the usefulness of spectral methods in getting good approximate answers to hard problems.

References:

[AM] Alon, N. and Millman, V. (1985). λ_1, Isoperimetric Inequalities for Graphs, and Superconcentrators. *J. Comb. Theory B*. 38: 73-88.

[A1] Alon, N. (1986). Eigenvalues and Expanders. *Combinatorica* 6:2,73-88.

[Ba] Bandle, C. (1980). *Isoperimetric inequalities and applications.* Boston: Pitman.

[Be] Berard, P.H. (1986). *Spectral geometry: direct and inverse problemsI.* New York: Springer-Verlag.

[Bi] Biggs, N. (1993). *Algebraic Graph Theory*. New York: Cambridge University Press.

[BBA] Breiger, R., Boorman, S. and Arabie, P. (1975). An Algorithm for Clustering Relational Data with Applications to Social Network Analysis and Comparison with Multidimensional Scaling. *J. Math. Psych*. 12:3, 328-382.

[Br] Brooks R. (1993). Spectral Geometry and the Cheeger Constant, in *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* Vol. 10, J. Friedman ed. p. 5-19

[Cha] Chavel, I. (1984) Eigenvalues in Riemannian Geometry, Academic Press: New York.

[Che] Cheeger, J. (1970). A lower Bound for the Lowest Eigenvalue of the Laplacian. *Problems in Analysis* (ed. R. C. Gunning) Princeton Univ. Press, 195-199.

[CH] Courant, R. and Hilbert, D. (1966). *Methods of Mathematical Physics*. Interscience Publishers.

[DS] Diaconis, P. and Stroock, D. (1991). Geometric Bounds for Eigenvalues of Markov Chains. *Ann. Appl. Prob*. 1: 36-61.

[D] Dodziuk, J. (1984). Difference Equations, Isoperimetric Inequality and the Transience of Certain Random Walks. *American Mathematical Society*. 284:2, 787-794.

[Fi] Fiedler, M. (1975). A Property of Eigenvectors of Non-negative Symmetric Matrices and its Application to Graph Theory. *Czech. Math. J.* 85: 619-633.

[Fr] Friedman, J. (1993). Some Geometrical Aspects of Graphs and their Eigenfunctions. *Duke Mathematical Journal* 69,487-525.

[H] Hagen, L. (1992). New Spectral Methods for Ratio Cut Partitioning and Clustering. *IEEE Trans. CAD*, 11:9,1074-1085.

[LS] Lawler, G. F. and Sokal, A. D. (1988). Bounds on the L^2 Spectrumn for Markov Chains and Markov Processes: a Generalization of Cheeger's Inequality. *Trans. Amer. Math. Soc*. 309: 557-580.

[L] Lovasz, L. (1995). Random Walks, Eigenvalues and Resistance, in Handbook of Combinatorics, Vol. II, Graham, Grotschel and Lovasz eds. MIT Press. p. 1740-1748

[Ma] Mass, C. H. (1985). Transportation in Graphs and the Admittance Spectrum, *Disc. Appl. Math.* 16: 31-49.

[Me] Merris, R. (1994). Laplacian Matrices of Graphs: A Survey. *Linear Alg. App.* 198,199: 143-176.

[Mo] Mohar, B. (1991). The Laplacian Spectrum of Graphs, in Graph Theory. *Combinatorics and Applications* (ed. Alavi, Chartrand, Ollermann and Schwenk) Wiley, pp. 871-898.

[MP] Motwani, R. and Prabhakar, R. (1996). *Randomized Algorithms*, Cambridge Univ. Press.

[PSL] Pothen, A., Simon, H, and Liou, K-P., (1990). Partitoning Sparse Matrices with Eigenvalues of Graphs. *SIAM J. Matrix Anal. App*. 11:3, 430-452.

[SR] Seary, A.J. and Richards, W.D. (1995). Partitioning Networks by Eigenvectors. Presented to European Netowrk Conference, London. Published in Everett, M.G. and Rennolls, K. (eds). (1996). *Proceedings of the International Conference on Social Networks, Volume 1: Methodology*. 47-58.

[T] Tutte, W. (1963). How to draw a graph, Proc. London Math. Soc. 52: 743-767

[W] Wasserman, S. (1980). Analyzing Social Networks as Stochastic Processes. *J. Am. Statistical Association.* 75, 280-294.