Spectra and eigenvectors of scale-free networks

K.-I. Goh, B. Kahng, and D. Kim
School of Physics, Center for Theoretical Physics, Seoul National University, Seoul 151-747, Korea
(Received 16 March 2001; published 15 October 2001)

We study the spectra and eigenvectors of the adjacency matrices of scale-free networks when bidirectional interaction is allowed, so that the adjacency matrix is real and symmetric. The spectral density shows an exponential decay around the center, followed by power-law long tails at both spectrum edges. The largest eigenvalue \( \lambda_1 \) depends on system size \( N \) as \( \lambda_1 \sim N^{1/8} \) for large \( N \), and the corresponding eigenfunction is strongly localized at the hub, the vertex with largest degree. The component of the normalized eigenfunction at the hub is of order unity. We also find that the mass gap scales as \( N^{-0.69} \).

Complex systems such as social, biological, and economic systems consist of many constituents such as individuals, substrates, and companies, respectively [1]. Each constituent reacts and adapts to the pattern created in the system through diverse interactions. Cooperative phenomena between constituents in such complex systems may be described in terms of random graphs, consisting of vertices and edges, where vertices (edges) represent constituents (their interactions). The study of complex systems in terms of random graphs was initiated by Erdős and Rényi (ER) [2]. In the ER model, the number of vertices is fixed, and edges connecting one vertex to another occur with probability \( p \). Then there exists a probability threshold \( p_c \), above which the system is percolated. Recently Watts and Strogatz (WS) introduced the small-world network [3] to consider local clustering, while the number of vertices is also fixed. The WS model offers the first indication that real networks can be more complex than predicted by the ER model.

Recently Barabási and Albert (BA) [4] introduced an evolving model where the number of vertices increases linearly with time rather than being fixed, and a newly introduced vertex is connected to \( m \) already existing vertices, following the so-called preferential attachment rule that the vertices with more edges are preferentially selected for the connection to the new vertex. The number of edges \( k \) incident upon a vertex is the degree of the vertex. Then the degree distribution \( P(k) \) of vertices, equivalent to the connectivity distribution, follows a power law \( P(k) \sim k^{-3} \) for the BA model, while for the ER and WS models it follows a Poisson distribution. The BA model is interesting in the sense that a lot of complex interactive networks such as the World Wide Web [5,6], the actor network [4], the citation network of scientific papers and the author collaboration network [7], the metabolic network [8] and the food web [9] in biological systems all exhibit a power law in the degree distribution, implying that a characteristic degree is absent in such systems. Thus the BA model is called a scale-free network. In the BA network, one may assign a direction on each edge pointing from the younger vertex to the older one [10]. However, when the direction on each edge is ignored, allowing bidirectional interactions such as downloading and uploading communications on the Internet, the BA network may be regarded as a simple model for this topology [11].

While it is known that the BA network follows a power law in its degree distribution, further structural properties are not so well known. When \( m = 1 \), the BA network forms a tree structure without forming any loops, but for \( m > 1 \) loops are formed, and network topology becomes more complicated. So it would be interesting to investigate the spectrum of the BA network, because generally the spectrum of a random graph and the corresponding eigenvectors are closely related to topological features of the random graph [13–15]. In this paper, we study the spectrum and the corresponding eigenvectors of the adjacency matrix of the BA network, comparing spectral properties with structural features. When a BA network is composed of \( N \) vertices, the adjacency matrix \( A \) consists of \( N \times N \) elements \( a_{ij} \) \( (i,j = 1, \ldots, N) \), defined as \( a_{ij} = 1 \) if vertices \( i \) and \( j \) are connected via an edge, and \( a_{ij} = 0 \) otherwise. \( k_i = \sum_j a_{ij} \) is the degree of the vertex \( i \). In the BA model, the vertex with the largest degree is singled out and is called the hub, denoted by \( h \) hereafter, in this work. From previous studies [4], we know that \( k_h \), the degree of the hub, scales as \( k_h \sim m N^{1/2} \). The vertices are ordered in their ages; the vertex \( i \) is the one created at time \( i \).

Since we consider the bidirectional case (sometimes called the undirected case), the adjacency matrix \( A \) is real and symmetric, so that all eigenvalues are real and the largest eigenvalue is not degenerate. We obtained the spectrum of the BA network via exact diagonalization for \( N \) up to 5000 and for the first few largest eigenvalues, via the Lanczos method [12] for \( N \) up to 400 000. Throughout this paper, numerical simulations were carried out for \( m = 2 \), the simplest case including the loop structure.

Eigenvalues. We consider the distribution of the eigenvalues shown in Fig. 1. The analytic formula for the spectrum is not known yet, but one can easily see that the spectrum does not fit the semicircular equation derived by Wigner [13] appropriate to the ER random graph. The density of eigenvalues \( \rho(\lambda) \) in the middle part of the spectrum is likely to fit the formula \( \rho(\lambda) \sim \exp(-|\lambda|/a) \) where \( a \approx 1.25 \) (see the left inset of Fig. 1), while the density further out follows the power law \( \rho(\lambda) \sim \lambda^{-1} \) (see the right inset of Fig. 1). Since the power-law decay is much slower than the exponential one, the spectrum shows long tails at both edges (see Fig. 1). Such behavior has also been observed in the financial time series [16].

The size dependence of the largest eigenvalue \( \lambda_1 \) is of interest and we found numerically that \( \lambda_1 \) increases as \( \sim N^{1/4} \) for large systems (see Fig. 2). To consider the relation be-
network, we consider the following. Let \( x_j \) be \( x \) column vectors, related through

\[
y(n) = A^n x.
\]

Let \( v_{i,j} \) be the \( i \)th component of the normalized eigenvector corresponding to the \( i \)th eigenvalue of \( A \), \( \lambda_i \), \( (\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_N) \). Then Eq. (1) may be rewritten as

\[
y_j(n) = \sum_i \sum_j \lambda_i^n v_{i,j} x_j.
\]

When we set \( x_j = 1 \) for some \( j \) and \( x_r = 0 \) for \( r \neq j \), \( y_j(n) \) becomes the number of possible ways of \( n \)-step walks starting from the vertex \( j \) and terminating at the vertex \( i \). This will be denoted as \( y_{j\rightarrow i}(n) \) hereafter. In particular, when \( i = j \), \( y_{j\rightarrow j}(n) \) is the number of possible ways to return to the start-vertex \( j \) after \( n \)-step walks. When \( n \) is sufficiently large, \( y_{j\rightarrow i}(n) \) can be represented in terms of the largest eigenvalue \( \lambda_1 \) alone as

\[
y_{j\rightarrow i}(n) = \lambda_1^n v_{i,j} v_{j,1}.
\]

In particular, when \( i = j \),

\[
y_{j\rightarrow j}(n) = \lambda_1^n v_{j,1}^2.
\]

Even though the relation Eq. (4) holds for any vertex \( j \), we consider it particularly at the hub. Now let us consider the number of ways of \( n+2 \) steps returning to the starting vertex \( h \). It can be split into two parts,

\[
y_{h\rightarrow h}(n+2) = y_{h\rightarrow h}(n) y_{h\rightarrow h}(2) + \sum_{j \neq h} y_{h\rightarrow j}(n) y_{j\rightarrow h}(2).
\]

According to Eq. (4),

\[
\lambda_1^2 = \frac{y_{h\rightarrow h}(n+2)}{y_{h\rightarrow h}(n)} = \frac{\sum_{j \neq h} y_{h\rightarrow j}(n) y_{j\rightarrow h}(2)}{y_{h\rightarrow h}(n)},
\]

where \( y_{h\rightarrow h}(2) \) corresponds to the degree of the hub, \( y_{h\rightarrow h}(2) = m \sqrt{N} \), while \( y_{j\rightarrow h}(2) \) is found numerically to behave as \( \sim N^{0.05} j^{0.43} \), weakly depending on \( N \) (see Fig. 3). The second term of the right-hand side of Eq. (6) can be written using Eqs. (3) and (4) as \( \sum_{j \neq h} (v_{j,1}/v_{h,1}) y_{j\rightarrow h}(2) \). We will show later that \( v_{h,1} = 1/2 \), while \( v_{j,1} \sim (Nj)^{-1/4} \) for \( j \) away from \( h \), leading to the result that the second term becomes \( O(N^{0.1}) \). Consequently, \( \lambda_1 \) is contributed predominantly by the first term in the right-hand side of Eq. (6), leading to \( \sim m^{1/2} N^{1/4} \).

The result \( \lambda_1 \sim N^{1/4} \) can also be understood through two toy models. First, one vertex is located at the center, and all the other \( N-1 \) vertices are linked only to the center vertex [Fig. 4(a)]. This structure, called the radial structure, is an extreme case of “winner takes all.” The largest eigenvalue of this structure is \( \lambda_1 = \sqrt{N-1} \), while the largest degree is \( N-1 \). Second, we consider a two-level Cayley tree structure [Fig. 4(b)]. When the coordination number is chosen as \( k \),
The moment becomes zero, that is, \( M_z \) the spectrum. First, since the matrix \( A \) vertex, so that \( \lambda_1 \sim \sqrt{k_i} \).

The second largest eigenvalue in absolute magnitude is located on the negative side of the spectrum. The absolute magnitude of this eigenvalue \( |\lambda_N| \) is also found to scale as \( |\lambda_N| \sim N^{1/4} \) (see Fig. 2). The difference between the first two largest eigenvalues in their absolute magnitude \( \lambda_1 - |\lambda_N| \) is found to scale as \( \sim N^{-0.43} \) (see the inset of Fig. 2). Thus, the mass gap, defined as \( \ln(\lambda_1/|\lambda_N|) \) scales as \( \sim N^{-\zeta} \) with \( \zeta \approx 0.68 \).

Next, we consider the size dependence of the moments of the spectrum. First, since the matrix \( A \) is traceless, the first moment becomes zero, that is, \( \mathcal{M}_1 = \sum_i^N \lambda_i = 0 \). Second, for each vertex, the number of ways to return to the starting vertex by a two-step walk is the same as the degree of that vertex, \( \sum_i^N k_i = 2mN \). Third, since the path to return to the starting vertex by a three-step walk forms a triangle, the total number of triangles \( \mathcal{M}_3 = \sum_i^N \lambda_i^3 \sim 6T_N \). We found numerically that \( \mathcal{M}_3 \sim N^{0.40} \) (see Fig. 5). Note that while the first moment vanishes the third moment does not, implying that the spectrum is not completely symmetric with respect to \( \zeta = 0 \).

**Eigenvector.** Since the adjacency matrix \( A \) is real and symmetric, every component of the eigenvector corresponding to the largest eigenvalue is positive. Since the largest
eigenvalue and the corresponding eigenfunction are important, as in the case of Eq. (3), we focus on the eigenvector \( \{v_{j,1}\} \) for the largest eigenvalue. We consider the square of each component of the eigenvector \( \{v_{j,1}^2\} \) instead of \( \{v_{j,1}\} \) itself because of the normalization \( \sum j v_{j,1}^2 = 1 \), and compare it with the normalized degree of each vertex \( \{k_j/\Sigma_j k_j\} \). As seen in Fig. 6, the two quantities are correlated in such a way that up-down behavior occurs in a similar fashion.

The components \( \{v_{j,1}^2\} \) are strongly localized at the hub. This is in contrast to the ER case where the corresponding eigenfunction is extended over all vertices (see the inset of Fig. 6). We found that \( v_{h,1}^2 \) at the hub increases with increasing \( N \) for small \( N \), but converges to a constant close to 1/2 for large \( N \) (see Fig. 7). The value 1/2 can be obtained from the radial structure analytically. Since the radial structure is the extreme case where the connectivity is mostly localized at the hub, the value 1/2 can become the upper bound of \( v_{h,1}^2 \). The result that \( v_{h,1}^2 \) approaches a constant as \( N \to \infty \) is interesting, because \( v_{h,1}^2 \) behaves differently from the normalized degree at the hub, decreasing as \( \sim N^{-1/2} \) (see Fig. 7). Since the number of returning \( n \)-step walks from a vertex \( j \) is \( y_j \sim n^{5/2} v_{j,1}^2 \) from Eq. (4), \( v_{j,1}^2 \) can be considered as a measure of the contribution from the vertex \( j \) to transport processes in a network. The fact that \( v_{h,1}^2 \)

![FIG. 4. The radial structure (a) and the two-level Cayley tree structure with coordination number \( k = 4 \) (b).](image)

![FIG. 5. Double logarithmic plot of the second and third moments \( \mathcal{M}_2 \) and \( \mathcal{M}_3 \) versus system size \( N \). The dotted (solid) line has slope 1.0 (0.40), drawn for the eye.](image)

![FIG. 6. Plots of \( v_{j,1}^2 \) (○), the normalized degree \( k_j/\Sigma_j k_j \) (□), and the normalized involving number \( p_j \) (○) versus vertex index \( j \) for \( N = 10^3 \). Inset: The eigenfunction corresponding to the largest eigenvalue for the ER random network for \( N = 10^3 \), showing that the eigenfunction is extended.](image)

![FIG. 7. Double logarithmic plot of \( v_{h,1}^2 \) and the normalized degree \( k_j/\Sigma_j k_j \) at the hub versus system size \( N \). The dotted line at 0.5 is the asymptotic line of \( v_{h,1}^2 \). The dashed line for the normalized degree has slope \(-0.5\), drawn for the eye.](image)
\(\sim O(1)\) can be interpreted to mean that the hub plays a much more dominant role in transport on the BA network, compared with the contribution measured by the normalized degree at the hub, which is \(O(1/\sqrt{N})\). Thus, the presence of the hub in scale-free networks diversifies pathways, and enhances the efficiency of transport, while the network systems are vulnerable to attack on the hub [18]. On the other hand, for other components of the eigenfunction \(\{v_{j,i}^2\}\), \((j = 1, \ldots, N, \text{but } j \neq h)\), it was found numerically that \(v_{j,1}^2 \sim 1/(4\sqrt{N})\) (see Fig. 8).

Each component of eigenvectors can be obtained through diagonalization of the adjacency matrix \(A\). However, the eigenfunction for \(\lambda_1\) can be obtained using the relation Eq. (4). Knowledge of \(y_{j,j}(n)\) and \(\lambda_1\) enables one to obtain each component of the eigenfunction through \(v_{j,1}^2 \sim y_{j,j}(n)/\lambda_1^n\) for sufficiently large \(n\). However, when \(n\) is not large enough, \(y_{j,j}(n)\) for vertex \(j\) is affected by local topology around \(j\), so that the value of \(v_{j,1}^2\) obtained by this method is different from that by exact diagonalization. Thus, we can define a characteristic number of steps \(n_c\), such that for \(n > n_c\), Eq. (4) holds, while for \(n < n_c\), it breaks down. To find \(n_c\), we introduce the quantity,

\[
\delta_n = \left| \sum_j \frac{y_{j,j}(n)}{\lambda_1^n} - v_{j,1}^2 \right|
\]

which is found to decay exponentially as \(\delta_n \sim \exp(-n/n_c)\) (see the inset of Fig. 9). We found numerically that \(n_c\) shows a crossover at a characteristic system size \(N_c\) such that, for \(N < N_c\), \(n_c \sim N^{0.35}\), while for \(N > N_c\), \(n_c \sim N^{0.50}\) (see Fig. 9). On the other hand, Eq. (7) may be rewritten as \(\delta_n = |\Sigma_{j=1} y_{j,j}(n)/\lambda_1^n|\). Thus, for sufficiently large \(n\), the predominant contribution to \(\delta_n\) is from the term \((|\lambda_{j}|/\lambda_1)^n\) alone. Combined with \(\delta_n \sim \exp(-n/n_c)\), one obtains that \(n_c = \lambda_1 \lambda_{-1}/|\lambda_1^2|\), equivalent to the inverse of the mass gap. Thus, \(n_c \sim N^2\). However, the numerical value of the exponent \(z\) obtained in this way, \(z = 0.50\), deviates from the value \(z = 0.68\) obtained by the direct measurement of \(\lambda_1 - |\lambda_{-1}|\). This can be attributed to the fact that \(\delta_n\) includes contributions from other eigenvalues.

**Shortest paths.** Since the role of the hub in transport is much more dominant, compared with the contribution by the normalized degree, we also study topological features of the shortest paths between two vertices [19]. The transport from one position to another is mainly along the shortest path(s) between them, and is contributed predominantly by the largest eigenvalue of the adjacency matrix \(A\). We define a set composed of the vertices on the shortest path(s) from one position to another. Then, since there are \(N(N-1)/2\) pairs of vertices, the same number of sets exist in the system. Among them, we are interested in the number of different sets a certain vertex \(j\) belongs to. This number is called the involving number \(P_j\), while the normalized involving number is defined as \(p_j = P_j/\Sigma P_j\) for each vertex \(j\). Figure 6 also shows \(p_j\) versus \(j\). It behaves similarly to the up-down behavior of \(v_{j,1}\) and \(k_j\). In particular, the involving number at the hub \(P_h\) is found numerically to scale as \(P_h \sim N^2\), while the total involving number summed over all vertices is found numerically to scale as \(\Sigma_{j=1} p_j \sim N^2\) in \(N\) (see Fig. 10) [20]. So the normalized involving number at the hub, \(p_h = P_h/\Sigma P_j\), behaves as \(p_h \sim 1/\ln N\), decaying much more slowly compared with the normalized degree which decreases as \(\sim N^{-0.5}\). This weak dependence on \(N\) is compa-
rable to the result that $v_n \sim O(1)$ for large $N$. Therefore, the contribution of the hub to the shortest paths is much larger than that of the naive estimate based on the normalized degree at the hub, $O(1/\sqrt{N})$.

Very recent work by Farkas et al. [21] deals with a similar topic to ours. Here we summarize the overlaps and some discrepancies between the two works. Both analyzed, by numerical calculations, the spectra of the adjacency matrices of scale-free networks and found that the largest eigenvalue $\lambda_1$ grows as $\sim N^{1/4}$. Although the tail regions of the spectrum were interpreted to decay as a power law in both cases, Farkas et al. interpreted its central part as exhibiting a "triangle-like" shape, while our interpretation is that it decays exponentially. As for the eigenvalues other than $\lambda_1$, called the bulk part of the spectrum, Farkas et al. argued that they are distributed symmetrically. But we found that that is not the case. For example, the fact that $\lambda_1$ and $|\lambda_N|$ approach each other as $N$ increases (Fig. 2) illustrates this. Accordingly, the odd-power moments are not governed solely by $\lambda_1$ but by the entire spectrum, which leads, for example, to $M_5 \sim N^{0.40}$ rather than $\sim N^{0.75}$, as proposed by Farkas et al. In addition to these similarities and differences, we investigated the eigenfunction for the largest eigenvalue. In particular, we found that the component of the eigenfunction corresponding to the largest eigenvalue at the hub is independent of the system size. This result contains the important physical implication that the hub plays a much more important role in transport than expected according to the normalized degree, which was also checked by studying the involving number of each vertex. This finding is an important consequence of this paper and should be useful to help the understanding of transport properties in complex networks.

In conclusion, we have considered the spectrum and eigenvectors of the adjacency matrix of the BA network, when bidirectional interaction is allowed. The density of the eigenvalues decays exponentially for small $|\lambda|$, followed by power-law tails at both spectrum edges. This is different from the Wigner formula appropriate to random graphs. We found that the largest two eigenvalues $\lambda_1$ and $|\lambda_N|$ depend on system size $N$ as $\sim N^{1/4}$ for large $N$, and the mass gap scales as $N^{-0.68}$. The eigenfunction corresponding to the largest eigenvalue $\lambda_1$ is strongly localized at the vertex with the largest degree, called the hub. The component of the normalized eigenfunction for $\lambda_1$ at the hub is independent of $N$, implying that the role of the hub in transport on the scale-free network becomes much more important for a larger system, and its contribution becomes much more dominant than expected according to the normalized degree at the hub, which scales as $N^{-1/2}$. Therefore, it is very efficient in communication networks to construct central vertices, through which most of the information traffic passes.

We would like to thank S. Y. Park for providing us with the Lanczos algorithm codes. This work was supported by Grant No. 2000-2-11200-002-3 from the BRP program of the KOSEF.

[20] The numerical data may also be interpreted as following a power law $\Sigma P_j \sim N^{-2.15}$.