Chapter 1
Graph Theory and Small-World Networks

Dynamical and adaptive networks are the backbone of many complex systems. Examples range from ecological prey–predator networks to the gene expression and protein providing the grounding of all living creatures. The brain is probably the most complex of all adaptive dynamical systems and is at the basis of our own identity, in the form of a highly sophisticated neural network. On a social level we interact through social and technical networks like the Internet. Networks are ubiquitous through the domain of all living creatures.

A good understanding of network theory is therefore of basic importance for complex system theory. In this chapter we will discuss the most important notions of graph theory, like clustering and degree distributions, together with basic network realizations. Central concepts like percolation, the robustness of networks with regard to failure and attacks, and the “rich-get-richer” phenomenon in evolving social networks will be treated.

1.1 Graph Theory and Real-World Networks

1.1.1 The Small-World Effect

Six or more billion humans live on earth today and it might seem that the world is a big place. But, as an Italian proverb says,

Tutto il mondo è paese – The world is a village.

The network of who knows whom – the network of acquaintances – is indeed quite densely webbed. Modern scientific investigations mirror this century-old proverb.

Social Networks Stanley Milgram performed a by now famous experiment in the 1960s. He distributed a number of letters addressed to a stockbroker
in Boston to a random selection of people in Nebraska. The task was to send these letters to the addressee (the stockbroker) via mail to an acquaintance of the respective sender. In other words, the letters were to be sent via a social network.

The initial recipients of the letters clearly did not know the Boston stockbroker on a first-name basis. Their best strategy was to send their letter to someone whom they felt was closer to the stockbroker, socially or geographically: perhaps someone they knew in the financial industry, or a friend in Massachusetts.

**Six Degrees of Separation** About 20% of Milgram’s letters did eventually reach their destination. Milgram found that it had only taken an average of six steps for a letter to get from Nebraska to Boston. This result is by now dubbed “six degrees of separation” and it is possible to connect any two persons living on earth via the social network in a similar number of steps.

The Small-World Effect. The “small-world effect” denotes the result that the average distance linking two nodes belonging to the same network can be orders of magnitude smaller than the number of nodes making up the network.

The small-world effect occurs in all kinds of networks. Milgram originally examined the networks of friends. Other examples for social nets are the network of film actors or that of baseball players, see Fig. 1.1. Two actors are linked by an edge in this network whenever they co-starred at least once in the same movie. In the case of baseball players the linkage is given by the condition to have played at least once on the same team.
Networks are Everywhere  Social networks are but just one important example of a communication network. Most human communication takes place directly among individuals. The spreading of news, rumors, jokes and of diseases takes place by contact between individuals. And we are all aware that rumors and epidemic infections can spread very fast in densely webbed social networks.

Communication networks are ubiquitous. Well known examples are the Internet and the world-wide web, see Fig. 1.1. Inside a cell the many constituent proteins form an interacting network, as illustrated in Fig. 1.2. The same is of course true for artificial neural networks as well as for the networks of neurons that build up the brain. It is therefore important to understand the statistical properties of the most important network classes.

1.1.2 Basic Graph-Theoretical Concepts

We start with some basic concepts allowing to characterize graphs and real-world networks. We will use the terms graph and network as interchangeable, vertex, site and node as synonyms, and either edge or link.

Degree of a Vertex  A graph is made out of $N$ vertices connected by edges.

Degree of a Vertex. The degree $k$ of the vertex is the number of edges linking to this node.

Nodes having a degree $k$ substantially above the average are denoted “hubs”, they are the VIPs of network theory.

Coordination Number  The simplest type of network is the random graph. It is characterized by only two numbers: By the number of vertices $N$ and by the average degree $z$, also called the coordination number.

Coordination Number. The coordination number $z$ is the average number of links per vertex, viz the average degree.

A graph with an average degree $z$ has $Nz/2$ connections. Alternatively we can define with $p$ the probability to find a given edge.

Connection Probability. The probability that a given edge occurs is called the connection probability $p$.

Erdős–Rényi Random Graphs  We can construct a specific type of random graph simply by taking $N$ nodes, also called vertices and by drawing $Nz/2$ lines, the edges, between randomly chosen pairs of nodes, compare Fig. 1.3. This type of random graph is called an “Erdős–Rényi” random graph after two mathematicians who studied this type of graph extensively.

Most of the following discussion will be valid for all types of random graphs, we will explicitly state whenever we specialize to Erdős–Rényi graphs. In Sect. 1.2 we will introduce and study other types of random graphs.
Fig. 1.2 A protein interaction network, showing a complex interplay between highly connected hubs and communities of subgraphs with increased densities of edges (From ?)
For Erdős–Rényi random graphs we have
\[ p = \frac{Nz^2}{2N(N-1)} = \frac{z}{N-1} \] (1.1)
for the relation between the coordination number \(z\) and the connection probability \(p\).

**Ensemble Realizations** There are, for any given link probability \(p\) and vertex number \(N\), a large number of ways to distribute the links, compare Fig. 1.3. For Erdős–Rényi graphs every link distribution is realized with equal probability. When one examines the statistical properties of a graph-theoretical model one needs to average over all such “ensemble realizations”.

**The Thermodynamic Limit** Mathematical graph theory is often concerned with the thermodynamic limit.

The Thermodynamic Limit. The limit where the number of elements making up a system diverges to infinity is called the “thermodynamic limit” in physics. A quantity is extensive if it is proportional to the number of constituting elements, and intensive if it scales to a constant in the thermodynamic limit.

We note that \(p = p(N) \rightarrow 0\) in the thermodynamic limit \(N \rightarrow \infty\) for Erdős–Rényi random graphs and intensive \(z \sim O(N^0)\), compare Eq. (1.1).

There are small and large real-world networks and it makes sense only for very large networks to consider the thermodynamic limit. An example is the network of hyperlinks.

**The Hyperlink Network** Every web page contains links to other web pages, thus forming a network of hyperlinks. In 1999 there were about \(N \approx 0.8 \times 10^9\) documents on the web, but the average distance between documents was only about 19. The WWW is growing rapidly; in 2007 estimates for the total number of web pages resulted in \(N \approx (20 - 30) \times 10^9\), with the size...
of the Internet backbone, viz the number of Internet servers, being about \( \simeq 0.1 \times 10^9 \).

**Network Diameter and the Small-World Effect** As a first parameter characterizing a network we discuss the diameter of a network.

**Network Diameter.** The network diameter is the maximal separation between all pairs of vertices.

For a random network with \( N \) vertices and coordination number \( z \) we have

\[
zD \approx N, \quad D \propto \log N / \log z ,
\]

(1.2)

since any node has \( z \) neighbors, \( z^2 \) next-nearest neighbors and so on. The logarithmic increase in the number of degrees of separation with the size of the network is characteristic of small-world networks. \( \log N \) increases very slowly with \( N \) and the network diameter therefore remains small even for networks containing a large number of nodes \( N \).

**Average Distance.** The average distance \( \ell \) is the average of the minimal path length between all pairs of nodes of a network.

The average distance \( \ell \) is generally closely related to the diameter \( D \); it has the same scaling with the number of nodes \( N \).

**Clustering in Networks** Real networks have strong local recurrent connections, compare, e.g. the protein network illustrated in Fig. 1.2, leading to distinct topological elements, such as loops and clusters.

The **Clustering Coefficient.** The clustering coefficient \( C \) is the average fraction of pairs of neighbors of a node that are also neighbors of each other.

The clustering coefficient is a normalized measure of loops of length 3. In a fully connected network, in which everyone knows everyone else, \( C = 1 \).

In a random graph a typical site has \( z(z-1)/2 \) pairs of neighbors. The probability of an edge to be present between a given pair of neighbors is \( p = z/(N-1) \), see Eq. (1.1). The clustering coefficient, which is just the probability of a pair of neighbors to be interconnected is therefore

\[
C_{\text{rand}} = \frac{z}{N-1} \approx \frac{z}{N} .
\]

(1.3)

It is very small for large random networks and scales to zero in the thermodynamic limit. In Table 1.1 the respective clustering coefficients for some

<table>
<thead>
<tr>
<th>Network</th>
<th>( N )</th>
<th>( \ell )</th>
<th>( C )</th>
<th>( C_{\text{rand}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Movie actors</td>
<td>225,226</td>
<td>3.65</td>
<td>0.79</td>
<td>0.00027</td>
</tr>
<tr>
<td>Neural network</td>
<td>282</td>
<td>2.65</td>
<td>0.28</td>
<td>0.05</td>
</tr>
<tr>
<td>Power grid</td>
<td>4,941</td>
<td>18.7</td>
<td>0.08</td>
<td>0.0005</td>
</tr>
</tbody>
</table>
real-world networks and for the corresponding random networks are listed for comparison.

**Cliques and Communities** The clustering coefficient measures the normalized number of triples of fully interconnected vertices. In general, any fully connected subgraph is denoted a clique.

Clique. A clique is a set of vertices for which (a) every node is connected by an edge to every other member of the clique and (b) no node outside the clique is connected to all members of the clique.

The term “clique” comes from social networks. A clique is a group of friends where everybody knows everybody else. A clique corresponds, in terms of graph theory, to a maximal fully connected subgraph. For Erdős–Rényi graphs with $N$ vertices and linking probability $p$, the number of cliques is

$$\left(\frac{N}{K}\right) p^{K(K-1)/2} \left(1 - p^K\right)^{N-K},$$

for cliques of size $K$. Here

- $\left(\frac{N}{K}\right)$ is the number of sets of $K$ vertices,
- $p^{K(K-1)/2}$ is the probability that all $K$ vertices within the clique are mutually interconnected and
- $\left(1 - p^K\right)^{N-K}$ the probability that every of the $N-K$ out-of-clique vertices is not connected to all $K$ vertices of the clique.

The only cliques occurring in random graphs in the thermodynamic limit have the size 2, since $p = z/N$. For an illustration see Fig. 1.4.

Another term used is *community*. It is mathematically not as strictly defined as “clique”, it roughly denotes a collection of strongly overlapping cliques, viz of subgraphs with above-the-average densities of edges.
Clustering for Real-World Networks Most real-world networks have a substantial clustering coefficient, which is much greater than $O(N^{-1})$. It is immediately evident from an inspection, for example of the protein network presented in Fig. 1.2, that the underlying “community structure” gives rise to a high clustering coefficient.

In Table 1.1, we give some values of $C$, together with the average distance $\ell$, for three different networks:

- The network of collaborations between movie actors
- The neural network of the worm *C. Elegans*, and
- The Western Power Grid of the United States.

Also given in Table 1.1 are the values $C_{\text{rand}}$ that the clustering coefficient would have for random graphs of the same size and coordination number. Note that the real-world value is systematically higher than that of random graphs. Clustering is important for real-world graphs. These are small-world graphs, as indicated by the small values for the average distances $\ell$ given in Table 1.1.

Erdős–Rényi random graphs obviously do not match the properties of real-world networks well. In Sect. 1.2 we will discuss generalizations of random graphs that approximate the properties of real-world graphs much better. Before that, we will discuss some general properties of random graphs in more detail.

Correlation Effects The degree distribution $p_k$ captures the statistical properties of nodes as if they were all independent of each other. In general, the property of a given node will however be dependent on the properties of other nodes, e.g. of its neighbors. When this happens one speaks of “correlation effects”, with the clustering coefficient $C$ being an example.

Another example for a correlation effect is what one calls “ assortative mixing”. A network is assortatively correlated whenever large-degree nodes, the hubs, tend to be mutually interconnected and assortatively anti-correlated when hubs are predominantly linked to low-degree vertices. Social networks tend to be assortatively correlated, in agreement with the everyday experience that the friends of influential persons, the hubs of social networks, tend to be VIPs themselves.

Tree Graphs Most real-world networks show strong local clustering and loops abound. An exception are metabolic networks which contain typically only very few loops since energy is consumed and dissipated when biochemical reactions go around in circles.

For many types of graphs commonly considered in graph theory, like Erdős–Rényi graphs, the clustering coefficient vanishes in the thermodynamic limit, and loops become irrelevant. One denotes a loopless graph a “tree graph”.

Bipartite Networks Many real-world graphs have an underlying bipartite structure, see Fig. 1.1.
1.1 Graph Theory and Real-World Networks

Bipartite Graph. A bipartite graph has two kinds of vertices with links only between vertices of unlike kinds.

Examples are networks of managers, where one kind of vertex is a company and the other kind of vertex the managers belonging to the board of directors. When eliminating one kind of vertex, in this case it is customary to eliminate the companies, one retains a social network: the network of directors, as illustrated in Fig. 1.1. This network has a high clustering coefficient, as all boards of directors are mapped onto cliques of the respective social network.

1.1.3 Graph Spectra and Degree Distributions

So far we have considered mostly averaged quantities of random graphs, like the clustering coefficient or the average coordination number $z$. We will now develop tools allowing for a more sophisticated characterization of graphs.

Degree Distribution. The basic description of general random and non-random graphs is given by the degree distribution $p_k$.

**Degree Distribution**. If $X_k$ is the number of vertices having the degree $k$, then $p_k = X_k/N$ is called the degree distribution, where $N$ is the total number of nodes.

The degree distribution is a probability distribution function and hence normalized, $\sum_k p_k = 1$.

**Degree Distribution for Erdős–Rényi Graphs** The probability of any node to have $k$ edges is

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k},$$

for an Erdős–Rényi network, where $p$ is the link connection probability. For large $N \gg k$ we can approximate the degree distribution $p_k$ by

$$p_k \simeq e^{-pN} \frac{(pN)^k}{k!} = e^{-z} \frac{z^k}{k!},$$

where $z$ is the average coordination number. We have used

$$\lim_{N \to \infty} \left(1 - \frac{x}{N}\right)^N = e^{-x}, \quad \binom{N-1}{k} = \frac{(N-1)!}{k!(N-1-k)!} \simeq \frac{(N-1)^k}{k!},$$

and $(N-1)^kp^k = z^k$, see Eq. (1.1). Equation (1.5) is a Poisson distribution with the mean

$$\langle k \rangle = \sum_{k=0}^{\infty} k e^{-z} \frac{z^k}{k!} = e^{-z} \sum_{k=1}^{\infty} \frac{z^{k-1}}{(k-1)!} = z.$$
as expected.

**Ensemble Fluctuations** In general, two specific realizations of random graphs differ. Their properties coincide on the average, but not on the level of individual links. With “ensemble” one denotes the set of possible realizations.

In an ensemble of random graphs with fixed \( p \) and \( N \) the degree distribution \( X_k/N \) will be slightly different from one realization to the next. On the average it will be given by

\[
\frac{1}{N} \langle X_k \rangle = p_k .
\]  

(1.6)

Here \( \langle \ldots \rangle \) denotes the ensemble average. One can go one step further and calculate the probability \( P(X_k = R) \) that in a realization of a random graph the number of vertices with degree \( k \) equals \( R \). It is given in the large-\( N \) limit by

\[
P(X_k = R) = e^{-\lambda_k} \frac{(\lambda_k R)^R}{R!}, \quad \lambda_k = \langle X_k \rangle .
\]  

(1.7)

Note the similarity to Eq. (1.5) and that the mean \( \lambda_k = \langle X_k \rangle \) is in general extensive while the mean \( z \) of the degree distribution (1.5) is intensive.

**Scale-Free Graphs** Scale-free graphs are defined by a power-law degree distribution

\[
p_k \sim \frac{1}{k^\alpha}, \quad \alpha > 1 .
\]  

(1.8)

Typically, for real-world graphs, this scaling \( \sim k^{-\alpha} \) holds only for large degrees \( k \). For theoretical studies we will mostly assume, for simplicity, that the functional dependence Eq. (1.8) holds for all \( k \). The power-law distribution can be normalized if

\[
\lim_{K \to \infty} \sum_{k=0}^{K} p_k \approx \lim_{K \to \infty} \int_{0}^{K} \frac{dk}{k^\alpha} \propto \lim_{K \to \infty} K^{1-\alpha} < \infty ,
\]

i.e. when \( \alpha > 1 \). The average degree is finite if

\[
\lim_{K \to \infty} \sum_{k=0}^{K} k p_k \propto \lim_{K \to \infty} K^{2-\alpha} < \infty , \quad \alpha > 2 .
\]

A power-law functional relation is called scale-free, since any rescaling \( k \to a k \) can be reabsorbed into the normalization constant.

Scale-free functional dependencies are also called *critical*, since they occur generally at the critical point of a phase transition. We will come back to this issue recurrently in the following chapters.

**Graph Spectra** Any graph \( G \) with \( N \) nodes can be represented by a matrix encoding the topology of the network, the adjacency matrix.

The Adjacency Matrix. The \( N \times N \) adjacency matrix \( \hat{A} \) has elements \( A_{ij} = 1 \) if nodes \( i \) and \( j \) are connected and \( A_{ij} = 0 \) if they are not connected.
The adjacency matrix is symmetric and consequently has $N$ real eigenvalues.

The Spectrum of a Graph. The spectrum of a graph $G$ is given by the set of eigenvalues $\lambda_i$ of the adjacency matrix $\hat{A}$.

A graph with $N$ nodes has $N$ eigenvalues $\lambda_i$ and it is useful to define the corresponding “spectral density”

$$\rho(\lambda) = \frac{1}{N} \sum_j \delta(\lambda - \lambda_j), \quad \int d\lambda \rho(\lambda) = 1, \quad (1.9)$$

where $\delta(\lambda)$ is the Dirac delta function. In Fig. 1.5 typical spectral densities for Erdős–Rényi graphs are illustrated.

**Green’s Function**\(^1\) The spectral density $\rho(\lambda)$ can be evaluated once the Green’s function $G(\lambda)$,

$$G(\lambda) = \frac{1}{N} \text{Tr} \left[ \frac{1}{\lambda - \hat{A}} \right] = \frac{1}{N} \sum_j \frac{1}{\lambda - \lambda_j}, \quad (1.10)$$

is known. Here $\text{Tr}[\ldots]$ denotes the trace over the matrix $(\lambda - \hat{A})^{-1} \equiv (\lambda \hat{1} - \hat{A})^{-1}$, where $\hat{1}$ is the identity matrix. Using the formula

$$\lim_{\epsilon \to 0} \frac{1}{\lambda - \lambda_j + i\epsilon} = P \frac{1}{\lambda - \lambda_j} - i\pi \delta(\lambda - \lambda_j), \quad (1.11)$$

\(^1\) The reader without prior experience with Green’s functions may skip the following derivation and pass directly to the result, namely to Eq. (1.15).
where $P$ denotes the principal part, we find the relation
\[ \rho(\lambda) = -\frac{1}{\pi} \lim_{\varepsilon \to 0} \text{Im} G(\lambda + i\varepsilon). \]  

**The Semi-Circle Law** The graph spectra can be evaluated for random matrices for the case of small link densities $p = z/N$, where $z$ is the average connectivity. We note that the Green’s function (1.10) can be expanded into powers of $\hat{A}/\lambda$,
\[ G(\lambda) = \frac{1}{N\lambda} \text{Tr} \left[ 1 + \frac{\hat{A}}{\lambda} + \left(\frac{\hat{A}}{\lambda}\right)^2 + \ldots \right]. \]  

Traces over powers of the adjacency matrix $\hat{A}$ can be interpreted, for random graphs, as random walks which need to come back to the starting vertex.

Starting from a random site we can connect on the average to $z$ neighboring sites and from there on to $z - 1$ next-nearest neighboring sites, and so on. This consideration allows to recast the Taylor expansion (1.13) for the Green’s function into a recursive continued fraction,
\[ G(\lambda) = 1 - \frac{z}{\lambda} - \frac{z - 1}{\lambda} - \frac{z - 1}{\lambda} - \ldots \approx 1 - \frac{1}{\lambda} G(\lambda), \]  

where we have approximated $z - 1 \approx z$ in the last step. Equation (1.14) constitutes a self-consistency equation for $G = G(\lambda)$, with the solution
\[ G^2 - \frac{\lambda}{z} G + \frac{1}{z} = 0, \quad G = \frac{\lambda}{2z} - \sqrt{\frac{\lambda^2}{4z^2} - \frac{1}{z}}, \]  

since $\lim_{\lambda \to \infty} G(\lambda) = 0$. The spectral density Eq. (1.12) then takes the form
\[ \rho(\lambda) = \begin{cases} \sqrt{4z - \lambda^2}/(2\pi z) & \text{if } \lambda^2 < 4z \\ 0 & \text{if } \lambda^2 > 4z \end{cases} \]  

of a half-ellipse also known as “Wigner’s law”, or the “semi-circle law”.

**Loops and the Clustering Coefficient** The total number of triangles, viz the overall number of loops of length 3 in a network is, on the average, $C(N/3)(z - 1)z/2$, where $C$ is the clustering coefficient. This relation holds for large networks. One can relate the clustering coefficient $C$ to the adjacency matrix $A$ via
\[ C \frac{z(z - 1)N}{2} = \text{number of triangles} \]

---

2 Taking the principal part signifies that one has to consider the positive and the negative contributions to the $1/\lambda$ divergences carefully.
1.1 Graph Theory and Real-World Networks

\[ \sum_{i_1, i_2, i_3} A_{i_1 i_2} A_{i_2 i_3} A_{i_3 i_1} = \frac{1}{6} \text{Tr} \left[ A^3 \right], \]

since three sites \( i_1, i_2 \) and \( i_3 \) are interconnected only when the respective entries of the adjacency matrix are unity. The sum of the right-hand side of above relation is also denoted a “moment” of the graph spectrum. The factors 1/3 and 1/6 on the left-hand side and on the right-hand side account for overcountings.

**Moments of the Spectral Density** The graph spectrum is directly related to certain topological features of a graph via its moments. The \( l \)th moment of \( \rho(\lambda) \) is given by

\[
\int d\lambda \lambda^l \rho(\lambda) = \frac{1}{N} \sum_{j=1}^{N} (\lambda_j)^l = \frac{1}{N} \text{Tr} [A^l] = \frac{1}{N} \sum_{i_1, i_2, \ldots, i_l} A_{i_1 i_2} A_{i_2 i_3} \cdots A_{i_l i_1}, \tag{1.16}
\]

as one can see from Eq. (1.9). The \( l \)th moment of \( \rho(\lambda) \) is therefore equivalent to the number of closed paths of length \( l \), the number of all paths of length \( l \) returning to the starting point.

### 1.1.4 Graph Laplacian

Differential operators, like the derivative \( df/dx \) of a function \( f(x) \), may be generalized to network functions. We start by recalling the definitions of the first

\[
\frac{d}{dx} f(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x},
\]

and of the second derivative

\[
\frac{d^2}{dx^2} f(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) + f(x - \Delta x) - 2f(x)}{\Delta x^2}.
\]

**Graph Laplacians** Consider now a function \( f_i, i = 1, \ldots, N \) on a graph with \( N \) sites. One defines the graph Laplacian \( \hat{A} \) via

\[
A_{ij} = \left( \sum_l A_{il} \right) \delta_{ij} - A_{ij} = \begin{cases} k_i & i = j \\ -1 & i \text{ and } j \text{ connected} \\ 0 & \text{otherwise} \end{cases}, \tag{1.17}
\]

where the \( A_{ij} = (\hat{A})_{ij} \) are the elements of the Laplacian matrix, \( A_{ij} \) the adjacency matrix, and where \( k_i \) is the degree of vertex \( i \). \( \hat{A} \) corresponds,
apart from a sign convention, to a straightforward generalization of the usual
Laplace operator \( \Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2 \). To see this, just apply
the Laplacian matrix \( \Lambda_{ij} \) to a graph-function \( f = (f_1, \ldots, f_N) \). The graph
Laplacian is hence intrinsically related to diffusion processes on networks, as
discussed in Sect. ??.

Alternatively one defines by

\[
L_{ij} = \begin{cases} 
1 & i = j \\
-1/\sqrt{k_i k_j} & i \text{ and } j \text{ connected} \\
0 & \text{otherwise}
\end{cases},
\]

(1.18)

the “normalized graph Laplacian”, where \( k_i = \sum_j A_{ij} \) is the degree of vertex
\( i \). The eigenvalues of the normalized graph Laplacian have a straightforward
interpretation in terms of the underlying graph topology. One needs however
to remove first all isolated sites from the graph. Isolated sites do not generate
entries to the adjacency matrix and the respective \( L_{ij} \) are not defined.

**Eigenvalues of the Normalized Graph Laplacian** Of interest are the
eigenvalues \( \lambda_l, l = 0, \ldots, (N - 1) \) of the normalized graph Laplacian (1.18).

- The normalized graph Laplacian is positive semidefinite,

\[
0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 2.
\]

- The lowest eigenvalue \( \lambda_0 \) is always zero, corresponding to the eigenfunction

\[
e(\lambda_0) = \frac{1}{\sqrt{C}} \left( \sqrt{k_1}, \sqrt{k_2}, \ldots, \sqrt{k_N} \right),
\]

(1.19)

where \( C \) is a normalization constant and where the \( k_i \) are the respective
vertex-degrees.

- The degeneracy of \( \lambda_0 \) is given by the number of disconnected subgraphs
contained in the network. The eigenfunctions of \( \lambda_0 \) then vanish on all
subclusters beside one, where it has the functional form (1.19).

- The largest eigenvalue \( \lambda_{N-1} \) is \( \lambda_{N-1} = 2 \), if and only if the network is
bipartite. Generally, a small value of \( 2 - \lambda_{N-1} \) indicates that the graph
is nearly bipartite. The degeneracy of \( \lambda = 2 \) is given by the number of
bipartite subgraphs.

- The inequality

\[
\sum_l \lambda_l \leq N
\]

holds generally. The equality holds for connected graphs, viz when \( \lambda_0 \) has
degeneracy one.

**Examples of Graph Laplacians** The eigenvalues of the normalized graph
Laplacian can be given analytically for some simple graphs.
• For a complete graph (all sites are mutually interconnected), containing $N$ sites, the eigenvalues are
\[ \lambda_0 = 0, \quad \lambda_l = N/(N-1), \quad (l = 1, \ldots, N-1). \]

• For a complete bipartite graph (all sites of one subgraph are connected to all other sites of the other subgraph) the eigenvalues are
\[ \lambda_0 = 0, \quad \lambda_{N-1} = 2, \quad \lambda_l = 1, \quad (l = 1, \ldots, N-2). \]

The eigenfunction for $\lambda_{N-1} = 2$ has the form
\[ \mathbf{e}(\lambda_{N-1}) = \frac{1}{\sqrt{C}} (\sqrt{k_A}, \ldots, \sqrt{k_A}, -\sqrt{k_B}, \ldots, -\sqrt{k_B}) \] (1.20)

Denoting with $N_A$ and $N_B$ the number of sites in two sublattices $A$ and $B$, with $N_A + N_B = N$, the degrees $k_A$ and $k_B$ of vertices belonging to sublattice $A$ and $B$ respectively are $k_A = N_B$ and $k_B = N_A$ for a complete bipartite lattice.

A densely connected graph will therefore have many eigenvalues close to unity. For real-world graphs one may therefore plot the spectral density of the normalized graph Laplacian in order to gain an insight into its overall topological properties. The information obtained from the spectral density of the adjacency matrix and from the normalized graph Laplacian are distinct.

### 1.2 Percolation in Generalized Random Graphs

The most random of all graphs are Erdős–Rényi graphs. One can relax the extend of randomness somewhat and construct random networks with an arbitrarily given degree distribution. This procedure is also denoted “configurational model”.

When a graph contains only a few links it will decompose into a set of disconnected subgraphs; however for high link densities essentially all vertices will be connected. This property is denoted “percolation” and will be discussed within the configurational model.

#### 1.2.1 Graphs with Arbitrary Degree Distributions

In order to generate random graphs that have non-Poisson degree distributions we may choose a specific set of degrees.
Fig. 1.6 Construction procedure of a random network with nine vertices and degrees $X_1 = 2$, $X_2 = 3$, $X_3 = 2$, $X_4 = 2$. In step A the vertices with the desired number of stubs (degrees) are constructed. In step B the stubs are connected randomly.

The Degree Sequence. A degree sequence is a specified set $\{k_i\}$ of the degrees for the vertices $i = 1 \ldots N$.

The degree sequence is also the first information one obtains when examining a real-world network and hence the foundation for all further analysis.

Construction of Networks with Arbitrary Degree Distribution The degree sequence can be chosen in such a way that the fraction of vertices having degree $k$ will tend to the desired degree distribution

$$p_k, \quad N \to \infty$$

in the thermodynamic limit. The network can then be constructed in the following way:

1. Assign $k_i$ “stubs” (ends of edges emerging from a vertex) to every vertex $i = 1, \ldots, N$.
2. Iteratively choose pairs of stubs at random and join them together to make complete edges.

When all stubs have been used up, the resulting graph is a random member of the ensemble of graphs with the desired degree sequence. Figure 1.6 illustrates the construction procedure.

The Average Degree and Clustering The mean number of neighbors is the coordination number

$$z = \langle k \rangle = \sum_k k p_k .$$

The probability that one of the second neighbors of a given vertex is also a first neighbor, scales as $N^{-1}$ for random graphs, regardless of the degree distribution, and hence can be ignored in the limit $N \to \infty$.

Degree Distribution of Neighbors Consider a given vertex $A$ and a vertex $B$ that is a neighbor of $A$, i.e. $A$ and $B$ are linked by an edge.
Fig. 1.7 The excess degree distribution \( q_{k-1} \) is the probability of finding \( k \) outgoing links of a neighboring site. Here the site \( B \) is a neighbor of site \( A \) and has a degree \( k = 5 \) and \( k - 1 = 4 \) outgoing edges, compare Eq. (1.21). The probability that \( B \) is a neighbor of site \( A \) is proportional to the degree \( k \) of \( B \).

We are now interested in the degree distribution for vertex \( B \), viz in the degree distribution of a neighbor vertex of \( A \), where \( A \) is an arbitrary vertex of the random network with degree distribution \( p_k \). As a first step we consider the average degree of a neighbor node.

A high-degree vertex has more edges connected to it. There is then a higher chance that any given edge on the graph will be connected to it, with this chance being directly proportional to the degree of the vertex. Thus the probability distribution of the degree of the vertex to which an edge leads is proportional to \( kp_k \) and not just to \( p_k \).

**Excess Degree Distribution** When we are interested in determining the size of loops or the size of connected components in a random graph, we are normally interested not in the complete degree of the vertex reached by following an edge from \( A \), but in the number of edges emerging from such a vertex that do not lead back to \( A \), because the latter contains all information about the number of second neighbors of \( A \).

The number of new edges emerging from \( B \) is just the degree of \( B \) minus one and its correctly normalized distribution is therefore

\[
q_{k-1} = \frac{k p_k}{\sum_j j p_j}, \quad q_k = \frac{(k+1)p_{k+1}}{\sum_j j p_j},
\]

(1.21)
since \( kp_k \) is the degree distribution of a neighbor, compare Fig. 1.7. The distribution \( q_k \) of the outgoing edges of a neighbor vertex is also denoted “excess degree distribution”. The average number of outgoing edges of a neighbor vertex is then

\[
\sum_{k=0}^{\infty} k q_k = \sum_{k=0}^{\infty} \frac{k(k+1)p_{k+1}}{\sum_j j p_j} = \sum_{k=1}^{\infty} \frac{(k-1)kp_k}{\sum_j j p_j}
\]

\[
= \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle},
\]

(1.22)

where \( \langle k^2 \rangle \) is the second moment of the degree distribution.
Number of Next-Nearest Neighbors

We denote with \( z_m \), \( z_1 = \langle k \rangle \equiv z \)
the average number of \( m \)-nearest neighbors. Equation (1.22) gives the average
number of vertices two steps away from the starting vertex \( A \) via a particular
neighbor vertex. Multiplying this by the mean degree of \( A \), namely \( z_1 \equiv z \),
we find that the mean number of second neighbors \( z_2 \) of a vertex is
\[
z_2 = \langle k^2 \rangle - \langle k \rangle .
\]
(1.23)

Note that \( z_2 \) is not given by the variance of the degree distribution, which
would be \( \langle (k - \langle k \rangle)^2 \rangle = \langle k^2 \rangle - \langle k \rangle^2 \), compare Sect. ??.

\( z_2 \) for the Erdős–Rényi graph

The degree distribution of an Erdős–Rényi
graph is the Poisson distribution, \( p_k = e^{-z} z^k / k! \), see Eq. (1.5). We obtain
for the average number of second neighbors, Eq. (1.23),
\[
z_2 = \sum_{k=0}^{\infty} k^2 e^{-z} \frac{z^k}{k!} - z = z e^{-z} \sum_{k=1}^{\infty} (k - 1 + 1) \frac{z^{k-1}}{(k-1)!} - z
\]
\[
= z^2 = \langle k \rangle^2 .
\]

The mean number of second neighbors of a vertex in an Erdős–Rényi random
graph is just the square of the mean number of first neighbors. This is a special
case however. For most degree distributions, Eq. (1.23) will be dominated by
the term \( \langle k^2 \rangle \), so the number of second neighbors is roughly the mean square
degree, rather than the square of the mean. For broad distributions these two
quantities can be very different.

Number of Far Away Neighbors

The average number of edges emerging
from a second neighbor, and not leading back to where we came from, is
also given by Eq. (1.22), and indeed this is true at any distance \( m \) away from
vertex \( A \). The average number of neighbors at a distance \( m \) is then
\[
z_m = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} z_{m-1} - 1 = \frac{z_2}{z_1} z_{m-1} ,
\]
(1.24)
where \( z_1 \equiv z = \langle k \rangle \) and \( z_2 \) are given by Eq. (1.23). Iterating this relation we find
\[
z_m = \left[ \frac{z_2}{z_1} \right]^{m-1} z_1 .
\]
(1.25)

The Giant Connected Cluster

Depending on whether \( z_2 \) is greater than
\( z_1 \) or not, Eq. (1.25) will either diverge or converge exponentially as \( m \)
becomes large:
\[
\lim_{m \to \infty} z_m = \begin{cases} 
\infty & \text{if } z_2 > z_1 \\
0 & \text{if } z_2 < z_1 
\end{cases}
\]
(1.26)
\( z_1 = z_2 \) is the percolation point. In the second case the total number of neighbors

\[
\sum_m z_m = z_1 \sum_{m=1}^{\infty} \left[ \frac{z_2}{z_1} \right]^{m-1} = \frac{z_1}{1 - z_2/z_1} = \frac{z_1^2}{z_1 - z_2}
\]

is finite even in the thermodynamic limit, in the first case it is infinite. The network decays, for \( N \to \infty \), into non-connected components when the total number of neighbors is finite.

The Giant Connected Component. When the largest cluster of a graph encompasses a finite fraction of all vertices, in the thermodynamic limit, it is said to form a giant connected component (GCC).

If the total number of neighbors is infinite, then there must be a giant connected component. When the total number of neighbors is finite, there can be no GCC.

**The Percolation Threshold** When a system has two or more possibly macroscopically different states, one speaks of a phase transition.

Percolation Transition. When the structure of an evolving graph goes from a state in which two (far away) sites are on the average connected/not connected one speaks of a percolation transition.

This phase transition occurs precisely at the point where \( z_2 = z_1 \). Making use of Eq. (1.23), \( z_2 = \langle k^2 \rangle - \langle k \rangle \), we find that this condition is equivalent to

\[
\langle k^2 \rangle - 2\langle k \rangle = 0, \quad \sum_{k=0}^{\infty} k(k - 2)p_k = 0 . \tag{1.27}
\]

We note that, because of the factor \( k(k - 2) \), vertices of degree zero and degree two do not contribute to the sum. The number of vertices with degree zero or two therefore affects neither the phase transition nor the existence of the giant component.

- Vertices of degree zero are not connected to any other node, they do not contribute to the network topology.
- Vertices of degree one constitute the sole negative contribution to the percolation condition (1.27). No further path is available when arriving at a site with degree one.
- Vertices of degree two act as intermediators between two other nodes. Removing vertices of degree two does not change the topological structure of a graph.

One can therefore remove (or add) vertices of degree two or zero without affecting the existence of the giant component.
Clique Percolation  Edges correspond to cliques with $Z = 2$ sites (see page 7). The percolation transition can then also be interpreted as a percolation of cliques having size two and larger. It is then clear that the concept of percolation can be generalized to that of percolation of cliques with $Z$ sites, see Fig. 1.4 for an illustration.

The Average Vertex–Vertex Distance  Below the percolation threshold the average vertex–vertex distance $\ell$ is finite and the graph decomposes into an infinite number of disconnected subclusters.

Disconnected Subclusters. A disconnected subcluster or subgraph constitutes a subset of vertices for which (a) there is at least one path in between all pairs of nodes making up the subcluster and (b) there is no path between a member of the subcluster and any out-of-subcluster vertex.

Well above the percolation transition, $\ell$ is given approximately by the condition $z\ell \simeq N$:

\[
\log(N/z_1) = (\ell - 1)\log(z_2/z_1), \quad \ell = \frac{\log(N/z_1)}{\log(z_2/z_1)} + 1, \quad (1.28)
\]

using Eq. (1.25). For the special case of the Erdős–Rényi random graph, for which $z_1 = z$ and $z_2 = z^2$, this expression reduces to the standard formula (1.2),

\[
\ell = \frac{\log N - \log z}{\log z} + 1 = \frac{\log N}{\log z}.
\]

The Clustering Coefficient of Generalized Random Graphs  The clustering coefficient $C$ denotes the probability that two neighbors $i$ and $j$ of a particular vertex $A$ have stubs that do interconnect. The probability that two given stubs are connected is $1/(zN - 1) \approx 1/zN$, since $zN$ is the total number of stubs. We then have, compare Eq. (1.22),

\[
C = \frac{\langle k_i k_j \rangle_q}{Nz} = \frac{\langle k_i \rangle_q \langle k_j \rangle_q}{Nz} = \frac{1}{Nz} \left[ \sum_k kq_k \right]^2
\]

\[
= \frac{1}{Nz} \left[ \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \right]^2 = \frac{z}{N} \left[ \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle^2} \right]^2, \quad (1.29)
\]

since the distributions of two neighbors $i$ and $j$ are statistically independent. The notation $\langle \ldots \rangle_q$ indicates that the average is to be take with respect to the excess degree distribution $q_k$, as given by Eq. (1.21).

The clustering coefficient vanishes in the thermodynamic limit $N \to \infty$, as expected. However, it may have a very big leading coefficient, especially for degree distributions with fat tails. The differences listed in Table 1.1, between the measured clustering coefficient $C$ and the value $C_{\text{rand}} = z/N$ for Erdős–Rényi graphs, are partly due to the fat tails in the degree distributions $p_k$ of the corresponding networks.
1.2.2 Probability Generating Function Formalism

Network theory is about the statistical properties of graphs. A very powerful method from probability theory is the generating function formalism, which we will discuss now and apply later on.

**Probability Generating Functions**  We define by

\[ G_0(x) = \sum_{k=0}^{\infty} p_k x^k \]  \hspace{1cm} (1.30)

the generating function \(G_0(x)\) for the probability distribution \(p_k\). The generating function \(G_0(x)\) contains all information present in \(p_k\). We can recover \(p_k\) from \(G_0(x)\) simply by differentiation:

\[ p_k = \frac{1}{k!} \left. \frac{d^k G_0}{dx^k} \right|_{x=0} . \]  \hspace{1cm} (1.31)

One says that the function \(G_0\) “generates” the probability distribution \(p_k\).

**The Generating Function for Degree Distribution of Neighbors**  We can also define a generating function for the distribution \(q_k\), Eq. (1.21), of the other edges leaving a vertex that we reach by following an edge in the graph:

\[ G_1(x) = \sum_{k=0}^{\infty} q_k x^k = \frac{\sum_{k=0}^{\infty} (k+1)p_{k+1}x^k}{\sum_j j p_j} = \frac{\sum_{k=0}^{\infty} kp_k x^{k-1}}{\sum_j j p_j} = \frac{G'_0(x)}{z} , \]  \hspace{1cm} (1.32)

where \(G'_0(x)\) denotes the first derivative of \(G_0(x)\) with respect to its argument.

**Properties of Generating Functions**  Probability generating functions have a couple of important properties:

1. Normalization: The distribution \(p_k\) is normalized and hence

\[ G_0(1) = \sum_k p_k = 1 . \]  \hspace{1cm} (1.33)

2. Mean: A simple differentiation

\[ G'_0(1) = \sum_k k p_k = \langle k \rangle \]  \hspace{1cm} (1.34)

yields the average degree \(\langle k \rangle\).

3. Moments: The \(n\)th moment \(\langle k^n \rangle\) of the distribution \(p_k\) is given by
\[ \langle k^n \rangle = \sum_k k^n p_k = \left[ \left( \frac{d}{dx} \right)^n G_0(x) \right]_{x=1} . \] (1.35)

The Generating Function for Independent Random Variables Let us assume that we have two random variables. As an example we consider two dice. Throwing the two dice are two independent random events. The joint probability to obtain \( k = 1, \ldots, 6 \) with the first die and \( l = 1, \ldots, 6 \) with the second die is \( p_k p_l \). This probability function is generated by

\[ \sum_{k,l} p_k p_l x^{k+l} = \left( \sum_k p_k x^k \right) \left( \sum_l p_l x^l \right) , \]

i.e. by the product of the individual generating functions. This is the reason why generating functions are so useful in describing combinations of independent random events.

As an application consider \( n \) randomly chosen vertices. The sum \( \sum_i k_i \) of the respective degrees has a cumulative degree distribution, which is generated by

\[ \left[ G_0(x) \right]^n . \]

The Generating Function of the Poisson Distribution As an example we consider the Poisson distribution \( p_k = e^{-z} z^k / k! \), see Eq. (1.5), with \( z \) being the average degree. Using Eq. (1.30) we obtain

\[ G_0(x) = e^{-z} \sum_{k=0}^{\infty} \frac{z^k}{k!} x^k = e^z(x-1) . \] (1.36)

This is the generating function for the Poisson distribution. The generating function \( G_1(x) \) for the excess degree distribution \( q_k \) is, see Eq. (1.32),

\[ G_1(x) = \frac{G'_0(x)}{z} = e^z(x-1) . \] (1.37)

Thus, for the case of the Poisson distribution we have, as expected, \( G_1(x) = G_0(x) \).

Further Examples of Generating Functions As a second example, consider a graph with an exponential degree distribution:

\[ p_k = \left( 1 - e^{-1/\kappa} \right) e^{-k/\kappa} , \quad \sum_{k=0}^{\infty} p_k = \frac{1 - e^{-1/\kappa}}{1 - e^{-1/\kappa}} = 1 , \] (1.38)

where \( \kappa \) is a constant. The generating function for this distribution is

\[ G_0(x) = \left( 1 - e^{-1/\kappa} \right) \sum_{k=0}^{\infty} \frac{e^{-k/\kappa}}{x^k} x^k = \frac{1 - e^{-1/\kappa}}{1 - xe^{-1/\kappa}} . \] (1.39)
and
\[ z = G_0'(1) = \frac{e^{-1/\kappa}}{1 - e^{-1/\kappa}}, \quad G_1(x) = \frac{G_0'(x)}{z} = \left[ \frac{1 - e^{-1/\kappa}}{1 - xe^{-1/\kappa}} \right]^2. \]

As a third example, consider a graph in which all vertices have degree 0, 1, 2, or 3 with probabilities \( p_0 \ldots p_3 \). Then the generating functions take the form of simple polynomials
\begin{align*}
G_0(x) &= p_3x^3 + p_2x^2 + p_1x + p_0, \\
G_1(x) &= q_2x^2 + q_1x + q_0 = \frac{3p_3x^2 + 2p_2x + p_1}{3p_3 + 2p_2 + p_1}. 
\end{align*}

Stochastic Sum of Independent Variables Let’s assume we have random variables \( k_1, k_2, \ldots \), each having the same generating functional \( G_0(x) \). Then
\[ G_0^2(x), \quad G_0^3(x), \quad G_0^4(x), \ldots \]
are the generating functionals for
\[ k_1 + k_2, \quad k_1 + k_2 + k_3, \quad k_1 + k_2 + k_3 + k_4, \ldots. \]

Now consider that the number of times \( n \) this stochastic process is executed is distributed as \( p_n \). As an example consider throwing a dice several times, with a probability \( p_n \) of throwing exactly \( n \) times. The distribution of the results obtained is then generated by
\[ \sum_p p_n G_0^n(x) = G_N(G_0(x)), \quad G_N(z) = \sum_n p_n z^n. \]

We will make use of this relation further on.

### 1.2.3 Distribution of Component Sizes

**The Absence of Closed Loops** We consider here a network below the percolation transition and are interested in the distribution of the sizes of the individual subclusters. The calculations will crucially depend on the fact that the generalized random graphs considered here do not have any significant clustering nor any closed loops.

Closed Loops. A set of edges linking vertices
\[ i_1 \to i_2 \ldots i_n \to i_1 \]
is called a closed loop of length \( n \).
In physics jargon, all finite components are tree-like. The number of closed loops of length 3 corresponds to the clustering coefficient $C$, viz to the probability that two of your friends are also friends of each other. For random networks $C = [(k^2) - \langle k \rangle]^2/(z^3 N)$, see Eq. (1.29), tends to zero as $N \to \infty$.

**Generating Function for the Size Distribution of Components** We define by

$$H_1(x) = \sum_m h^{(1)}_m x^m$$

the generating function that generates the distribution of cluster sizes containing a given vertex $j$, which is linked to a specific incoming edge, see Fig. 1.8. That is, $h^{(1)}_m$ is the probability that the such-defined cluster contains $m$ nodes.

**Self-Consistency Condition for $H_1(x)$** We note the following:

1. The first vertex $j$ belongs to the subcluster with probability 1, its generating function is $x$.
2. The probability that the vertex $j$ has $k$ outgoing stubs is $q_k$.
3. At every stub outgoing from vertex $j$ there is a subcluster.
4. The total number of vertices consists of those generated by $[H_1(x)]^k$ plus the starting vertex.

The number of outgoing edges $k$ from vertex $j$ is described by the distribution function $q_k$, see Eq. (1.21). The total size of the $k$ clusters is generated by $[H_1(x)]^k$, as a consequence of the multiplication property of generating functions discussed in Sect. 1.2.2. The self-consistency equation for the total number of vertices reachable is then

$$H_1(x) = x \sum_{k=0}^{\infty} q_k [H_1(x)]^k = x G_1(H_1(x)),$$  \hspace{1cm} (1.43)

where we have made use of Eqs. (1.32) and (1.42).
The Embedding Cluster Distribution Function The quantity that we actually want to know is the distribution of the sizes of the clusters to which the entry vertex belongs. We note that

1. The number of edges emanating from a randomly chosen vertex is distributed according to the degree distribution $p_k$.
2. Every edge leads to a cluster whose size is generated by $H_1(x)$.

The size of a complete component is thus generated by

$$H_0(x) = x \sum_{k=0}^{\infty} p_k [H_1(x)]^k = x G_0(H_1(x)),$$  \hspace{1cm} (1.44)

where the prefactor $x$ corresponds to the generating function of the starting vertex. The complete distribution of component sizes is given by solving Eq. (1.43) self-consistently for $H_1(x)$ and then substituting the result into Eq. (1.44).

The Mean Component Size The calculation of $H_1(x)$ and $H_0(x)$ in closed form is not possible. We are, however, interested only in the first moment, viz the mean component size, see Eq. (1.34).

The component size distribution is generated by $H_0(x)$, Eq. (1.44), and hence the mean component size below the percolation transition is

$$\langle s \rangle = H_0'(1) = \left[ G_0(H_1(x)) + x G_0'(H_1(x)) H_1'(x) \right]_{x=1} = 1 + G_0'(1) H_1'(1),$$  \hspace{1cm} (1.45)

where we have made use of the normalization

$$G_0(1) = H_1(1) = H_0(1) = 1.$$

of generating functions, see Eq. (1.33). The value of $H_1'(1)$ can be calculated from Eq. (1.43) by differentiating:

$$H_1'(x) = G_1(H_1(x)) + x G_1'(H_1(x)) H_1'(x),$$  \hspace{1cm} (1.46)

$$H_1'(1) = \frac{1}{1 - G_1'(1)}.$$

Substituting this into (1.45) we find

$$\langle s \rangle = 1 + \frac{G_0'(1)}{1 - G_1'(1)}. \hspace{1cm} (1.47)$$

We note that

$$G_0'(1) = \sum_k k p_k = \langle k \rangle = z_1,$$  \hspace{1cm} (1.48)
where we have made use of Eq. (1.23). Substitution into (1.47) then gives the average component size below the transition as

$$\langle s \rangle = 1 + \frac{z_1^2}{z_1 - z_2}.$$  

(1.49)

This expression has a divergence at $z_1 = z_2$. The mean component size diverges at the percolation threshold, compare Sect. 1.2, and the giant connected component forms.

**Size of the Giant Component** Above the percolation transition the network contains a giant connected component, which contains a finite fraction $S$ of all sites $N$. Formally we can decompose the generating functional for the component sizes into a part generating the giant component, $H_0^*(x)$, and a part generating the remaining components,

$$H_0(x) \to H_0^*(x) + H_0(x), \quad H_0^*(1) = S, \quad H_0(1) = 1 - S,$$

and equivalently for $H_1(x)$. The self-consistency Eqs. (1.43) and (1.44),

$$H_0(x) = x G_0(H_1(x)), \quad H_1(x) = x G_1(H_1(x)),$$

then take the form

$$1 - S = H_0(1) = G_0(u), \quad u = H_1(1) = G_1(u).$$  

(1.50)

For Erdős-Rényi graphs we have $G_1(x) = G_0(x) = \exp(z(x - 1))$, see Eq. (1.37), and hence

$$1 - S = u = e^{z(u-1)} = e^{-zS}, \quad S = 1 - e^{-zS}.$$  

(1.51)

One finds then, using the Taylor expansion for $\exp(-zS)$ in (1.51), that $\lim_{z \to 1+} S(z) = 2(z - 1)/z^2 = 0$. The size of the giant component vanishes hence linearly at the percolation transition $z_c = 1$, as shown in Fig. 1.9, approaching unity exponentially fast for large degrees $z$.

### 1.3 Robustness of Random Networks

Fat tails in the degree distributions $p_k$ of real-world networks (only slowly decaying with large $k$) increase the robustness of the network. That is, the network retains functionality even when a certain number of vertices or edges is removed. The Internet remains functional, to give an example, even when a substantial number of Internet routers have failed.
1.3 Robustness of Random Networks

**Fig. 1.9** The size $S(z)$ of the giant component, as given by Eq. (1.51), vanishes linearly, for Erdős-Rényi graphs, at the percolation transition occurring at the critical coordination number $z_c = 1$

**Removal of Vertices** We consider a graph model in which each vertex is either “active” or “inactive”. Inactive vertices are nodes that have either been removed, or are present but non-functional. We denote by

$$b(k) = b_k$$

the probability that a vertex is active. The probability can be, in general, a function of the degree $k$. The generating function

$$F_0(x) = \sum_{k=0}^{\infty} p_k b_k x^k, \quad F_0(1) = \sum_k p_k b_k \leq 1,$$  \hspace{1cm} (1.52)

generates the probabilities that a vertex has degree $k$ and is present. The normalization $F_0(1)$ is equal to the fraction of all vertices that are present. Alternatively one could work with the normalized generating function

$$\left(1 - \sum_k p_k b_k\right) + \sum_{k=0}^{\infty} p_k b_k x^k,$$

since all inactive nodes are equivalent to nodes with degree zero, with identical results. By analogy with Eq. (1.32) we define by

$$F_1(x) = \frac{\sum_k k p_k b_k x^{k-1}}{\sum_k k p_k} = \frac{F_0'(x)}{z}$$  \hspace{1cm} (1.53)

the (non-normalized) generating function for the degree distribution of neighbor sites.

**Distribution of Connected Clusters** The distribution of the sizes of connected clusters reachable from a given vertex, $H_0(x)$, or from a given edge, $H_1(x)$, is generated respectively by the normalized functions
\[ H_0(x) = 1 - F_0(1) + x F_0(H_1(x)), \quad H_0(1) = 1, \]
\[ H_1(x) = 1 - F_1(1) + x F_1(H_1(x)), \quad H_1(1) = 1, \quad (1.54) \]

which are the logical equivalents of Eqs. (1.43) and (1.44).

**Random Failure of Vertices** First we consider the case of random failures of vertices. In this case, the probability
\[ b_k \equiv b \leq 1, \quad F_0(x) = b G_0(x), \quad F_1(x) = b G_1(x) \]
of a vertex being present is independent of the degree \( k \) and just equal to a constant \( b \), which means that
\[ H_0(x) = 1 - b + bx G_0(H_1(x)), \quad H_1(x) = 1 - b + bx G_1(H_1(x)), \quad (1.55) \]
where \( G_0(x) \) and \( G_1(x) \) are the standard generating functions for the degree of a vertex and of a neighboring vertex, Eqs. (1.30) and (1.32). This implies that the mean size of a cluster of connected and present vertices is
\[ \langle s \rangle = H_0'(1) = b + b G_0'(1) H_1'(1) = b + \frac{b^2 G_0'(1)}{1 - b G_1'(1)} = b \left[ 1 + \frac{b G_0'(1)}{1 - b G_1'(1)} \right], \]
where we have followed the derivation presented in Eq. (1.46) in order to obtain \( H_1'(1) = b / (1 - b G_1'(1)) \). With Eq. (1.48) for \( G_0'(1) = z_1 = z \) and \( G_1'(1) = z_2 / z_1 \) we obtain the generalization
\[ \langle s \rangle = b + \frac{b^2 z_2^2}{z_1 - b z_2} \quad (1.56) \]
of Eq. (1.49). The model has a phase transition at the critical value of \( b \)
\[ b_c = \frac{z_1}{z_2} = \frac{1}{G_1'(1)}. \quad (1.57) \]
If the fraction \( b \) of the vertices present in the network is smaller than the critical fraction \( b_c \), then there will be no giant component. This is the point at which the network ceases to be functional in terms of connectivity. When there is no giant component, connecting paths exist only within small isolated groups of vertices, but no long-range connectivity exists. For a communication network such as the Internet, this would be fatal.

For networks with fat tails, however, we expect that the number of next-nearest neighbors \( z_2 \) is large compared to the number of nearest neighbors \( z_1 \) and that \( b_c \) is consequently small. The network is robust as one would need to take out a substantial fraction of the nodes before it would fail.

**Random Failure of Vertices in Scale-Free Graphs** We consider a pure power-law degree distribution
1.3 Robustness of Random Networks

\[ p_k \sim \frac{1}{k^\alpha}, \quad \int \frac{dk}{k^\alpha} < \infty, \quad \alpha > 1, \]

see Eq. (1.8) and also Sect. 1.5. The first two moments are

\[ z_1 = \langle k \rangle \sim \int dk \frac{k}{k^\alpha}, \quad \langle k^2 \rangle \sim \int dk \frac{k^2}{k^\alpha}. \]

Noting that the number of next-nearest neighbors

\[ z_2 = \langle k^2 \rangle - \langle k \rangle, \quad \text{Eq. (1.23)}, \]

we can identify three regimes:

- \(1 < \alpha \leq 2\): \(z_1 \to \infty, z_2 \to \infty b_c = z_1/z_2\) is arbitrary in the thermodynamic limit \(N \to \infty\).
- \(2 < \alpha \leq 3\): \(z_1 < \infty, z_2 \to \infty b_c = z_1/z_2 \to 0\) in the thermodynamic limit. Any number of vertices can be randomly removed with the network remaining above the percolation limit. The network is extremely robust.
- \(3 < \alpha\): \(z_1 < \infty, z_2 < \infty b_c = z_1/z_2\) can acquire any value and the network has normal robustness.

**Biased Failure of Vertices**

What happens when one sabotages the most important sites of a network? This is equivalent to removing vertices in decreasing order of their degrees, starting with the highest degree vertices. The probability that a given node is active then takes the form

\[ b_k = \theta(k_c - k), \quad (1.58) \]

where \(\theta(x)\) is the Heaviside step function

\[ \theta(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}. \quad (1.59) \]

This corresponds to setting the upper limit of the sum in Eq. (1.52) to \(k_c\). Differentiating Eq. (1.54) with respect to \(x\) yields

\[ H_1'(1) = F_1(H_1(1)) + F_1'(H_1(1)) H_1'(1), \quad H_1'(1) = \frac{F_1(1)}{1 - F_1'(1)}, \]

as \(H_1(1) = 1\). The phase transition occurs when \(F_1'(1) = 1\),

\[ \sum_{k=1}^\infty k(k-1)p_kb_k = \frac{\sum_{k=1}^{k_c} k(k-1)p_k}{\sum_{k=1}^{\infty} kp_k} = 1, \quad (1.60) \]

where we used the definition Eq. (1.53) for \(F_1(x)\).

**Biased Failure of Vertices for Scale-Free Networks**

Scale-free networks have a power-law degree distribution, \(p_k \propto k^{-\alpha}\). We can then rewrite Eq. (1.60) as

\[ H_{k_c}^{(\alpha-2)} - H_{k_c}^{(\alpha-1)} = H_{\infty}^{(\alpha-1)}, \quad (1.61) \]
The critical fraction $f_c$ of vertices, Eq. (1.63). Removing a fraction greater than $f_c$ of highest degree vertices from a scale-free network, with a power-law degree distribution $p_k \sim k^{-\alpha}$ drives the network below the percolation limit. For a smaller loss of highest degree vertices (shaded area) the giant connected component remains intact.

\[
H_n^{(r)} = \sum_{k=1}^{n} \frac{1}{k^r}. \tag{1.62}
\]

The number of vertices present is $F_0(1)$, see Eq. (1.52), or $F_0(1)/\sum_k p_k$, since the degree distribution $p_k$ is normalized. If we remove a certain fraction $f_c$ of the vertices we reach the transition determined by Eq. (1.61):

\[
f_c = 1 - \frac{F_0(1)}{\sum_k p_k} = 1 - \frac{H_k^{(\alpha)}}{H_\infty^{(\alpha)}}. \tag{1.63}
\]

It is impossible to determine $k_c$ from (1.61) and (1.63) to get $f_c$ in closed form. One can, however, solve Eq. (1.61) numerically for $k_c$ and substitute it into Eq. (1.63). The results are shown in Fig. 1.10, as a function of the exponent $\alpha$. The network is very susceptible with respect to a biased removal of highest-degree vertices.

- A removal of more than about 3% of the highest degree vertices always leads to a destruction of the giant connected component. Maximal robustness is achieved for $\alpha \approx 2.2$, which is actually close to the exponents measured in some real-world networks.
- Networks with $\alpha < 2$ have no finite mean, $\sum_k k/k^2 \to \infty$, and therefore make little sense physically.
- Networks with $\alpha > \alpha_c = 3.4788\ldots$ have no giant connected component. The critical exponent $\alpha_c$ is given by the percolation condition $H_\infty^{(\alpha-2)} = 2H_\infty^{(\alpha-1)}$, see Eq. (1.27).

**Fig. 1.10** The critical fraction $f_c$ of vertices, Eq. (1.63). Removing a fraction greater than $f_c$ of highest degree vertices from a scale-free network, with a power-law degree distribution $p_k \sim k^{-\alpha}$ drives the network below the percolation limit. For a smaller loss of highest degree vertices (shaded area) the giant connected component remains intact (From ?)
1.4 Small-World Models

Random graphs and random graphs with arbitrary degree distribution show no clustering in the thermodynamic limit, in contrast to real-world networks. It is therefore important to find methods to generate graphs that have a finite clustering coefficient and, at the same time, the small-world property.

Clustering in Lattice Models

Lattice models and random graphs are two extreme cases of network models. In Fig. 1.11 we illustrate a simple one-dimensional lattice with connectivity $z = 2, 4$. We consider periodic boundary conditions, viz the chain wraps around itself in a ring. We then can calculate the clustering coefficient $C$ exactly.

- The One-Dimensional Lattice: The number of clusters can be easily counted. One finds

$$C = \frac{3(z - 2)}{4(z - 1)},$$

which tends to $3/4$ in the limit of large $z$.

- Lattices with Dimension $d$: Square or cubic lattices have dimension $d = 2, 3$, respectively. The clustering coefficient for general dimension $d$ is

$$C = \frac{3(z - 2d)}{4(z - d)},$$

which generalizes Eq. (1.64). We note that the clustering coefficient tends to $3/4$ for $z \gg 2d$ for regular hypercubic lattices in all dimensions.

Distances in Lattice Models

Regular lattices do not show the small-world effect. A regular hypercubic lattice in $d$ dimensions with linear size $L$ has $N = L^d$ vertices. The average vertex–vertex distance increases as $L$, or equivalently as

$$\ell \approx N^{1/d}.$$ 

The Watts and Strogatz Model

Watts and Strogatz have proposed a small-world model that interpolates smoothly between a regular lattice and an Erdős–Rényi random graph. The construction starts with a one-


**Fig. 1.12** Small-world networks in which the crossover from a regular lattice to a random network is realized. (a) The original Watts–Strogatz model with the rewiring of links. (b) The network with the addition of shortcuts (From ?)

 dimensional lattice, see Fig. 1.12a. One goes through all the links of the lattice and rewires the link with some probability $p$.

**Rewiring Probability.** We move one end of every link with the probability $p$ to a new position chosen at random from the rest of the lattice.

For small $p$ this process produces a graph that is still mostly regular but has a few connections that stretch long distances across the lattice as illustrated in Fig. 1.12a. The average coordination number of the lattice is by construction still the initial degree $z$. The number of neighbors of any particular vertex can, however, be greater or smaller than $z$.

**The Newman and Watts Model** A variation of the Watts–Strogatz model has been suggested by Newman and Watts. Instead of rewiring links between sites as in Fig. 1.12a, extra links, also called “shortcuts”, are added between pairs of sites chosen at random, but no links are removed from the underlying lattice, see Fig. 1.12b. This model is somewhat easier to analyze than the original Watts and Strogatz model, because it is not possible for any region of the graph to become disconnected from the rest, whereas this can happen in the original model.

The small-world models illustrated in Fig. 1.12, have an intuitive justification for social networks. Most people are friends with their immediate neighbors. Neighbors on the same street, people that they work with or their
Fig. 1.13 The clustering coefficient $C(p)$ and the average path length $L(p)$, as a function of the rewiring probability for the Watts and Strogatz model, compare Fig. 1.12 (From ?)

relatives. However, some people are also friends with a few far away persons. Far away in a social sense, like people in other countries, people from other walks of life, acquaintances from previous eras of their lives, and so forth. These long-distance acquaintances are represented by the long-range links in the small-world models illustrated in Fig. 1.12.

**Properties of the Watts and Strogatz Model** In Fig. 1.13 the clustering coefficient and the average path length are shown as a function of the rewiring probability $p$. The key result is that there is a parameter range, say $p \approx 0.01 - 0.1$, where the network still has a very high clustering coefficient and already a small average path length, as observed in real-world networks. Similar results hold for the Newman–Watts model.

1.5 Scale-Free Graphs

**Evolving Networks** Most real-world networks are open, i.e. they are formed by the continuous addition of new vertices to the system. The number of vertices, $N$, increases throughout the lifetime of the network, as it is the case for the WWW, which grows exponentially by the continuous addition of new web pages. The small world networks discussed in Sect. 1.4 are, however, constructed for a fixed number of nodes $N$, growth is not considered.

**Preferential Connectivity** Random network models assume that the probability that two vertices are connected is random and uniform. In contrast, most real networks exhibit the “rich-get-richer” phenomenon.
Fig. 1.14 Illustration of the preferential attachment model for an evolving network. At $t = 0$ the system consists of $N_0 = 3$ isolated vertices. At every time step a new vertex (shaded circle) is added, which is connected to $m = 2$ vertices, preferentially to the vertices with high connectivity, determined by the rule Eq. (1.66)

Preferential Connectivity. When the probability for a new vertex to connect to any of the existing nodes is not uniform for an open network we speak of preferential connectivity.

A newly created web page, to give an example, will include links to well-known sites with a quite high probability. Popular web pages will therefore have both a high number of incoming links and a high growth rate for incoming links. The growth of vertices in terms of edges is therefore in general not uniform.

Barabási–Albert Model We start at time $t = 0$ with $N(0) = N_0$ unconnected vertices. The preferential attachment growth process can then be carried out in two steps:

- Growth: At every time step a new vertex connected with $m \leq N_0$ links to the existing network is added.
- Preferential Attachment: The new links are determined by connecting the $m$ stubs of the new vertex to the vertices present at $t-1$ with the probability

$$
\Pi(k_i) = \frac{k_i + C}{a(t)} \quad a(t) = 2m(t-1) + C(N_0 + t - 1),
$$

(1.66)

where the normalization $a(t)$ has been chosen such that the overall probability to connect is unity, $\sum \Pi(k_i) = 1$. The attachment probability $\Pi(k_i)$ is linearly proportional to the number of links already present, modulo an offset $C$. Other functional dependencies for $\Pi(k_i)$ are of course possible, but not considered here.

After $t$ time steps this model leads to a network with $N(t) = t + N_0$ vertices and $mt$ edges, see Fig. 1.14. We will now show that the preferential rule leads to a scale-free degree distribution

$$
p_k \sim k^{-\gamma} \quad \gamma = 3 + \frac{C}{m}.
$$

(1.67)

The scaling relation Eq. (1.66) is valid in the limit $t \to \infty$ and for large degrees $k_i$. 
Time-Dependent Connectivities The time dependence of the degree of a given vertex can be calculated analytically using a mean-field approach. We are interested in vertices with large degrees \( k \); the scaling relation Eq. (1.67) is defined asymptotically for the limit \( k \to \infty \). We may therefore assume \( k \) to be continuous:

\[
\Delta k_i(t) \equiv k_i(t+1) - k_i(t) \approx \frac{\partial k_i}{\partial t} = A \Pi(k_i) = A \frac{k_i + C}{a(t)}, \tag{1.68}
\]

where \( \Pi(k_i) \) is the attachment probability. The overall number of new links is proportional to a normalization constant \( A \), which is hence determined by the sum rule

\[
\sum_i \Delta k_i(t) \equiv m = A \sum_i \Pi(k_i) = A ,
\]

since the overall probability to attach, \( \sum_i \Pi(k_i) \), is unity. Making use of the normalization \( a(t) \), Eq. (1.66), we obtain

\[
\frac{\partial k_i}{\partial t} = \frac{m}{2m+C} \frac{k_i + C}{t+a_0} \approx \frac{m}{2m+C} \frac{k_i + C}{t} \tag{1.69}
\]

for large times \( t \), with \( a_0 = C(N_0 - 1) - 2m \). The solution of (1.69) is given by

\[
\frac{k_i}{k_i + C} = \frac{m}{2m+C} \frac{1}{t}, \quad k_i(t) = (C + m) \left( \frac{t}{t_i} \right)^{m/(2m+C)} - C , \tag{1.70}
\]

where \( t_i \) is the time at which the vertex \( i \) had been added, \( k_i(t) = 0 \) for \( t < t_i \).

Adding Times We now specialize to the case \( C = 0 \), the general result for \( C > 0 \) can then be obtained from scaling considerations. When a vertex is added at time \( t_i = N_i - N_0 \) it has initially \( m = k_i(t_i) \) links, in agreement with (1.70), which reads

\[
k_i(t) = m \left( \frac{t}{t_i} \right)^{0.5}, \quad t_i = t m^2 / k_i^2 \tag{1.71}
\]

for \( C = 0 \). Older nodes, i.e. those with smaller \( t_i \), increase their connectivity faster than the younger vertices, viz those with bigger \( t_i \), see Fig. 1.15. For social networks this mechanism is dubbed the rich-gets-richer phenomenon.

The number of nodes \( N(t) = N_0 + t \) is identical to the number of adding times,

\[ t_1, \ldots, t_{N_0} = 0, \quad t_{N_0+j} = j, \quad j = 1, 2, \ldots, \]

where we have defined the initial \( N_0 \) nodes to have adding times zero.

Integrated Probabilities Using (1.71), the probability that a vertex has a connectivity \( k_i(t) \) smaller than a certain \( k \), \( P(k_i(t) < k) \) can be written as
Graph Theory and Small-World Networks

The adding times are uniformly distributed, compare Fig. 1.15, and the probability \( P(t_i) \) to find an adding time \( t_i \) is then

\[
P(t_i) = \frac{1}{N_0 + t}, \quad (1.73)
\]

just the inverse of the total number of adding times, which coincides with the total number of nodes. \( P(t_i > m^2 t/k^2) \) is therefore the cumulative number of adding times \( t_i \) larger than \( m^2 t/k^2 \), multiplied with the probability \( P(t_i) \) (Eq. 1.73) to add a new node:

\[
P \left( t_i > \frac{m^2 t}{k^2} \right) = \left( t - \frac{m^2 t}{k^2} \right) \frac{1}{N_0 + t}. \quad (1.74)
\]

**Scale-Free Degree Distribution** The degree distribution \( p_k \) then follows from Eq. (1.74) via a simple differentiation,

\[
p_k = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{\partial P(t_i > m^2 t/k^2)}{\partial k} = \frac{2m^2 t}{N_0 + t} \frac{1}{k^3}, \quad (1.75)
\]

in accordance with Eq. (1.67). The degree distribution Eq. (1.75) has a well defined limit \( t \to \infty \), approaching a stationary distribution. We note that \( \gamma = 3 \), which is independent of the number \( m \) of added links per new site. This result indicates that growth and preferential attachment play an important role for the occurrence of a power-law scaling in the degree distribution. To verify that both ingredients are really necessary, we now investigate a variant of above model.

One can repeat this calculation for a finite set-off \( C > 0 \). The exponent \( \gamma \) is identical to the inverse of the scaling power of \( k_i \) with respect to time \( t \) in Eq. (1.70), plus one, viz \( \gamma = (2m + C)/m + 1 = 3 + C/m \), see Eq. (1.67).
1.5 Scale-Free Graphs

Growth with Random Attachment We examine then whether growth alone can result in a scale-free degree distribution. We assume random instead of preferential attachment. The growth equation for the connectivity \( k_i \) of a given node \( i \), compare Eqs. (1.68) and (1.73), then takes the form

\[
\frac{\partial k_i}{\partial t} = \frac{m}{N_0 + t}.
\]

(1.76)

The \( m \) new edges are linked randomly at time \( t \) to the \((N_0 + t - 1)\) nodes present at the previous time step. Solving Eq. (1.76) for \( k_i \), with the initial condition \( k_i(t_i) = m \), we obtain

\[
k_i = m \left[ \ln \left( \frac{N_0 + t}{N_0 + t_i} \right) + 1 \right],
\]

(1.77)

which is a logarithmic increase with time. The probability that vertex \( i \) has connectivity \( k_i(t) \) smaller than \( k \) is then

\[
P(k_i(t) < k) = P \left( t_i > (N_0 + t) e^{1-k/m} - N_0 \right)
= \left[ t - (N_0 + t) e^{-k/m} + N_0 \right] \frac{1}{N_0 + t},
\]

(1.78)

where we assumed that we add the vertices uniformly in time to the system. Using

\[
p_k = \frac{\partial P(k_i(t) < k)}{\partial k}
\]

and assuming long times, we find

\[
p_k = \frac{1}{m} e^{1-k/m} = \frac{e}{m} \exp \left( -\frac{k}{m} \right).
\]

(1.79)

Thus for a growing network with random attachment we find a characteristic degree

\[
k^* = m,
\]

(1.80)

which is identical to half of the average connectivities of the vertices in the system, since \( \langle k \rangle = 2m \). Random attachment does not lead to a scale-free degree distribution. Note that \( p_k \) in Eq. (1.79) is not properly normalized, nor in Eq. (1.75), since we used a large-\( k \) approximation during the respective derivations.

Internal Growth with Preferential Attachment The original preferential attachment model yields a degree distribution \( p_k \sim k^{-\gamma} \) with \( \gamma = 3 \). Most social networks such as the WWW and the Wikipedia network, however, have exponents \( 2 < \gamma < 3 \), with the exponent \( \gamma \) being relatively close to 2. It is also observed that new edges are mostly added in between existing nodes, albeit with (internal) preferential attachment.
We can then generalize the preferential attachment model discussed above in the following way:

- **Vertex Growth**: At every time step a new vertex is added.
- **Link Growth**: At every time step \( m \) new edges are added.
- **External Preferential Attachment**: With probability \( r \in [0, 1] \) any one of the \( m \) new edges is added between the new vertex and an existing vertex \( i \), which is selected with a probability \( \propto \Pi(k_i) \), see Eq. (1.66).
- **Internal Preferential Attachment**: With probability \( 1 - r \) any one of the \( m \) new edges is added in between two existing vertices \( i \) and \( j \), which are selected with a probability \( \propto \Pi(k_i) \Pi(k_j) \).

The model reduces to the original preferential attachment model in the limit \( r \to 1 \). The scaling exponent \( \gamma \) can be evaluated along the lines used above for the case \( r = 1 \). One finds

\[
\rho_k \sim \frac{1}{k^\gamma}, \quad \gamma = 1 + \frac{1}{1 - r/2} \quad \text{.} \tag{1.81}
\]

The exponent \( \gamma = \gamma(r) \) interpolates smoothly between 2 and 3, with \( \gamma(1) = 3 \) and \( \gamma(0) = 2 \). For most real-world graphs \( r \) is quite small; most links are added internally. Note, however, that the average connectivity \( \langle k \rangle = 2m \) remains constant, since one new vertex is added for \( 2m \) new stubs.