Chapter 1
Network Theory

Dynamical and adaptive networks are the backbone of many complex systems. Examples range from ecological prey–predator networks to the gene- and protein expression networks on which all living creatures are based. The basis of our own identity is provided by the highly sophisticated neural network we carry in our heads. On a social level we interact through social and technical networks like the Internet. Networks are ubiquitous throughout the domain of all living creatures.

A good understanding of network theory is of basic importance for complex systems 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 Properties of Real-World Networks

The Small World Effect Eight or more billion humans live on earth and it might seem that the world is a big place. But, as an Italian proverb states,

> 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.

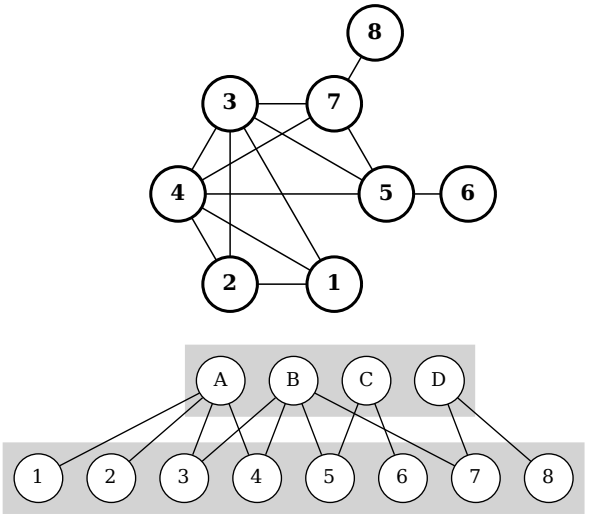

Fig. 1.1 Bottom: A network with two kind of nodes, a bipartite networks. The letters A-D may stand for distinct companies, the numerals 1-8 for managers serving on the respective boards of directors. **Top:** Managers serving on the same boards of directors are acquainted, forming a densely webbed social 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, dubbed "six degrees of separation", suggests that it should be possible to connect any two persons on earth via social networks in a similar number of steps.

> 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 examined the networks of friends, which may originate from participating in common activities. This is the case also for other social networks, as illustrated in Fig. 1.1. An example are managers serving together on the board of directors of the same company, which results in a dense network of acquaintances between the managers.

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 contacts between individuals. And we are all aware that rumors and epidemic infections can spread fast in densely webbed social networks.

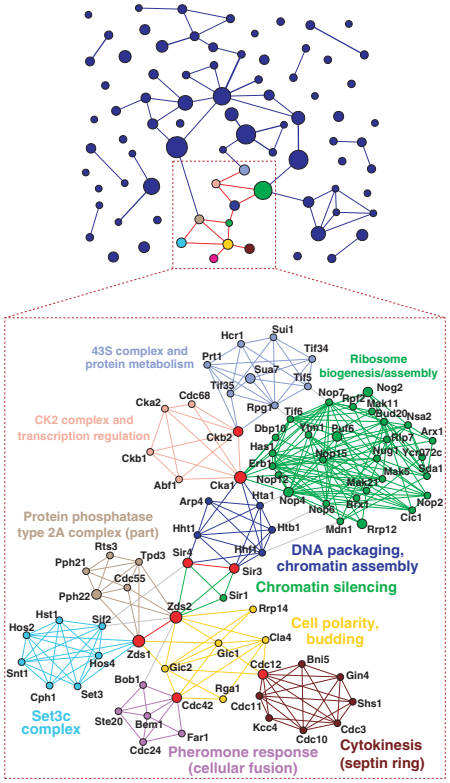

Fig. 1.2 A protein interaction network, showing a complex interplay between highly connected hubs and communities of subgraphs with increased densities of edges. Reprinted from **?** with permissions, © 2005 Nature Publishing Group.

Communication networks are ubiquitous. Well known examples are the Internet and the world-wide web. Inside a cell the constituent proteins form an interacting network, as illustrated in Fig. 1.2. The same holds for artificial neural networks and for the scaffholding of our brain, neural networks. It is therefore important to understand the statistical properties of core network classes.

1.1.1 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, with the degree k of a vertex being given by the number of edges linking to it. Nodes having a degree k substantially above the average are denoted "hubs", they are the VIPs of network theory.

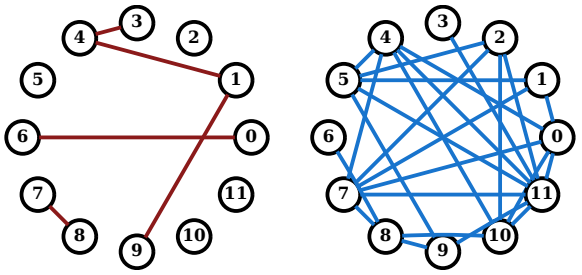

Fig. 1.3 Random graphs with $N = 12$ nodes and connection probabilities $p = 0.0758$ and $p = 0.3788$ (left/right). Three mutually connected vertices, such (0,1,7), contribute to the clustering coefficient.

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.

Erdös–Rényi Graphs We can construct a specific type of a random graph with N nodes 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 "Erdös–Rényi" random graph after two mathematicians who studied this type of graph extensively. In Sect. 1.3 we will introduce and study other types of random graphs.

For Erdös–Rényi random graphs we have

$$p = \frac{Nz}{2}\frac{2}{N(N-1)} = \frac{z}{N-1} \tag{1.1}$$

for the relation between the coordination number z and the connection probability p. In (1.1) we divided the total number of active links, $Nz/2$, by the total number of possible links, $N(N-1)/2$.

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 $p = z/(N-1) \to 0$ in the thermodynamic limit $N \to \infty$ for Erdös–Rényi random graphs when the coordination number is intensive, viz when $z \sim O(N^0)$.

For real-world networks, the thermodynamic limit makes sense for very large numbers of nodes. This is the case for the world-wide web, the network of hyperlinks, which is so large that its size can only be estimated. In order of magnitude, the number of documents available online did grow from $N \simeq 0.8 \times 10^9$ in 1999 to $N \simeq 3.4 \times 10^{15}$ in 2023.

Network Diameter and the Small-World Effect A basic parameter characterizing networks is the diameter D, which specifies the maximal separation between all possible pairs of vertices. For a random network with N vertices and coordination number z we have,

$$z^D \approx N, \qquad D \propto \log N / \log z\,, \tag{1.2}$$

in order of magnitude, since nodes have 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. A logarithm increases very slowly with its argument, which implies that network diameters remains small even for networks containing large numbers of nodes.

For large networks, the diameter is closely related to the average distance.

Average Distance The average distance ℓ is the average of the minimal path length between all pairs of nodes of a network.

Stochastic consideration will use in general the average distance ℓ.

Clustering in Networks Real-world networks, as the protein network reproduced in Fig. 1.2, contain substantial local recurrent connections. Distinct topological elements and loops are hence frequent.

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 three. 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 $p =$

Table 1.1 The number of nodes N, average degree of separation ℓ, and clustering coefficient C, for four real-world networks. The last column is the value $C_{\rm rand} = z/(N-1)$ the clustering coefficient would take for a random graph with the same size and coordination number. From **?** and **?**.

Network	N	ℓ	C	$C_{\rm rand}$
Movie actors	225,226	3.65	0.79	0.00027
Neural network	282	2.65	0.28	0.05
Power grid	4,941	18.70	0.08	0.0005
Internet domains	6,242,195	–	0.20	0.000006

$z/(N-1)$, see (1.1). The clustering coefficient, which is just the probability of a pair of neighbors to be interconnected is therefore

$$C_{\text{rand}} = \frac{z}{N-1} \approx \frac{z}{N} \,. \tag{1.3}$$

The clustering coefficient scales to zero in the thermodynamic limit, viz for large networks. In Table 1.1 the respective clustering coefficients for some real-world networks and for the corresponding random networks are listed for comparison.

Cliques The clustering coefficient measures the normalized number of triples of fully interconnected vertices. Cliques generalize the notion of clustering to larger subgraphs, see Fig. 1.4.

> CLIQUES 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 Erdös–Rényi graphs with N vertices and linking probability p, the number of cliques is

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

for cliques of size K. Here

- $\binom{N}{K}$ 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 two, since $p = z/N$.

Communities Another term used is "community". It is mathematically not as strictly defined as 'clique', denoting subgraphs with above-the-average densities of edges. Communities naturally emerge when decimating a bipartite graph stochastically. For the example shown in Fig. 1.1 one would assume that managers are not automatically acquainted when serving on the same board of directors, but with a certain probability, say 80%. Instead of cliques, the respective network of acquaintances between managers will be characterized by a dense community structure.

Clustering for Real-World Networks Most real-world networks have clustering coefficients that are substantially larger than the one a random network of the same size would have. This is immediately evident for the

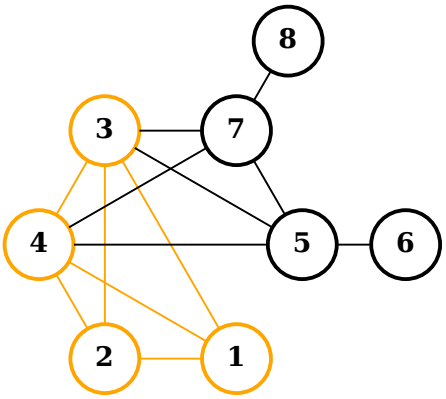

Fig. 1.4 The social graph of Fig. 1.1, with the four-site clique (1,2,3,4) highlighted. The network contains three additional cliques, (3,4,5,7), (5,6) and (7,8).

protein network shown in Fig. 1.2, which is characterized by a rich community structure and an elevated clustering coefficient.

In Table 1.1, we list C and the average distance ℓ for four different networks.

- A network of collaborations between movie actors.
- The neural network of the worm *C. Elegans*.
- A power grid of the United States.
- The networks of hyperlinks between domains (not webpages).

Also given in Table 1.1 are the clustering coefficients C_{rand} of random graphs of identical size and coordination number. Note that the real-world value is systematically higher than that of random graphs. The small values for the average distances ℓ given in Table 1.1 for three of the four networks relects their small-world nature.

Erdös–Rényi random graphs obviously do not match the properties of real-world networks well. In Sect. 1.3 we will discuss generalizations of random graphs that approximate the properties of real-world graphs better. Before, we will discuss several general properties of random graphs.

Correlation Effects The degree distribution p_k captures the statistical properties of nodes as if they where 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 Erdös–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.

> 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 nodes is a company and the other kind of nodes the managers belonging to the board of directors. When eliminating one kind of vertices, 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.

1.1.2 Network Degree Distribution

So far we 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 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 Erdös–Rényi Graphs For an Erdös–Rényi network, the probability of a node to have k edges is given by the Binomial distribution,

$$p_k = \binom{N-1}{k} p^k \, (1-p)^{N-1-k} \,, \tag{1.4}$$

where p is the link connection probability. For large $N \gg k$ we can approximate the degree distribution p_k by

$$p_k \simeq \mathrm{e}^{-pN} \, \frac{(pN)^k}{k!} = \mathrm{e}^{-z} \, \frac{z^k}{k!} \,, \tag{1.5}$$

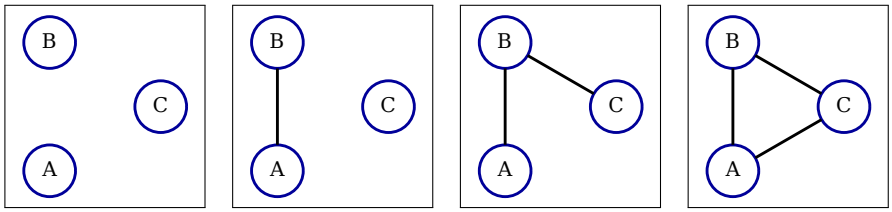

Fig. 1.5 Four networks with three nodes exist, respectively with 0/1/2/3 edges (from left to right). Given a link probability p, the probability to generate networks with 0/1/2/3 edges is $(1-p)^3$, $3p(1-p)^2$, $3p^2(1-p)$ and p^3. One speaks of ensemble flucuations, compare (1.7).

where z is the average coordination number. We have used

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

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

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

The variance of the Poisson distribution is z, which implies that the relative width $\langle k\rangle/\sigma = 1/\sqrt{k}$ becomes progressively smaller when the coordination number z increases.

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. The set of all possible realization of three-site networks is shown in Fig. 1.5.

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

$$\frac{1}{N}\langle X_k\rangle = p_k\,. \tag{1.6}$$

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

$$P(X_k = R) = \mathrm{e}^{-\lambda_k}\,\frac{(\lambda_k)^R}{R!}, \qquad \lambda_k = \langle X_k\rangle\,. \tag{1.7}$$

Note the similarity to (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.

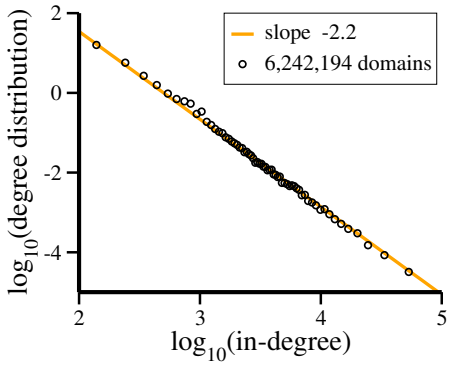

Fig. 1.6 Log-log plot of the distribution of the in-degrees of the hyperlink network between 6,242,194 domains. To a fair degree, the degree distribution is scale invariant, $p_k \sim 1/k^\alpha$. The value of the exponent, $\alpha \approx 2.2$, implies that the mean degree is finite, with a diverging variance. At the time of the crawl, $\langle k \rangle = 35.8$ was found. Data from **?**.

Power Law Distribution A special case are degree distributions decaying for large degrees as a power law,

$$p_k \;\sim\; \frac{1}{k^\alpha}, \qquad\qquad k \to \infty\,. \tag{1.8}$$

Typically, for real-world graphs, the scaling $\sim k^{-\alpha}$ holds only for large degrees k. The divergence at $k \to 0$ is avoided by the "Tsallis–Pareto distribution" $p_k = (\alpha - 1)(1 + k)^{-\alpha}$. For the normalization on $\boldsymbol{R}^+$ we consider $k \in [0, K]$, with $K \to \infty$,

$$\sum_{k=0}^{K} p_k \;\approx\; (\alpha - 1) \int_0^K \frac{dk}{(1+k)^\alpha} = \frac{-1}{(1+k)^{\alpha-1}}\Big|_{k=0}^{K} = 1 - \frac{1}{(1+K)^{\alpha-1}}\,.$$

Powerlaw distributions are hence normalizable only if $\alpha > 1$. The average degree of the Tsallis–Pareto distribution is

$$\langle k \rangle = (\alpha - 1) \int_0^K dk\, \frac{k + 1 - 1}{(1+k)^\alpha} = \frac{\alpha - 1}{2 - \alpha}(1+k)^{2-\alpha}\Big|_0^K - 1\,,$$

which diverges when $\alpha \leq 2$. Otherwise, when $\alpha > 2$, one has $\langle k \rangle = 1/(\alpha - 2)$. The variance is finite in analogy when $\alpha > 3$.

Scale-Free Graphs Power-law functional relations are "scale free", since any rescaling $k \to a\,k$ can be reabsorbed into the normalization constant. A famous example of a scale-invariant degree distribution is that of the in-degree of the hyperlink network between domains, as shown in Fig. 1.6.

Scale-free functional dependencies are considered to be '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.

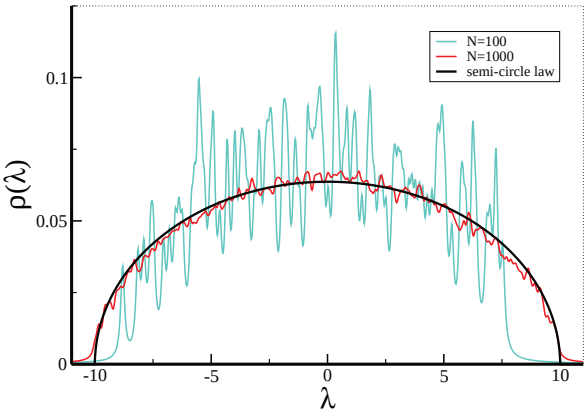

Fig. 1.7 The spectral densities (1.9) for Erdös–Rényi graphs with a coordination number $z = 25$ and $N = 100/1{,}000$ nodes respectively. For comparison the expected result for the thermodynamic limit $N \to \infty$, the semi-circle law (1.15), is given. A broadening of $\epsilon = 0.1$ has been used, as defined by (1.11).

1.2 Spectral Properties

In order to understand a given real-world network it is desirable to develop characterization schemes relying just on a few key observables, e.g. like the coordination number z or the clustering coefficient C. As a further step in this direction we study now spectral properties.

Adjacency Matrix Any graph G with N nodes can be represented by a matrix encoding the topology of the network, the adjacency matrix.

ADJACENCY MATRIX The $N \times N$ adjacency matrix $\hat{A}$ has elements $A_{ij} = 1$ if nodes i and j are connected, with $A_{ij} = 0$ if they are not connected.

The adjacency matrix is symmetric and consequently has N real eigenvalues.

SPECTRUM OF A GRAPH The spectrum of a graph G is given by the set of eigenvalues λ_i of the adjacency matrix $\hat{A}$.

A graph with N nodes has N eigenvalues λ_i and it is useful to define the corresponding "spectral density"

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

where $\delta(\lambda)$ is the Dirac delta function. In Fig. 1.7 the spectral density of an Erdös–Rényi graph is given.

Green's Function The spectral density $\rho(\lambda)$ can be evaluated once the Green's function[1] $G(\lambda)$,

[1] The reader without prior experience with Green's functions may skip the following derivation and pass directly to the result, namely to (1.15).

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

is known. Here Tr[...] denotes the trace over the matrix $(\lambda - \hat{A})^{-1} = (\lambda \mathbf{1} - \hat{A})^{-1}$, where $\mathbf{1}$ is the identity matrix. In complex plane, we can decompose the inverse,

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

where P denotes the principal part, which takes into account that the positive and the negative contributions of the $1/\lambda$ divergence cancel. We find the relation

$$\rho(\lambda) = -\frac{1}{\pi} \lim_{\varepsilon \to 0} \mathrm{ImG}(\lambda + i\varepsilon) \,. \tag{1.12}$$

Semi-Circle Law The graph spectra can be evaluated for random matrices[2] 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 $\hat{A}/\lambda$,

$$G(\lambda) = \frac{1}{N\lambda} \mathrm{Tr} \left[1 + \frac{\hat{A}}{\lambda} + \left(\frac{\hat{A}}{\lambda} \right)^2 + \ldots \right] . \tag{1.13}$$

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,

$$G(\lambda) = \cfrac{1}{\lambda - \cfrac{z}{\lambda - \cfrac{z-1}{\lambda - \cfrac{z-1}{\lambda - \ldots}}}} \approx \frac{1}{\lambda - z\, G(\lambda)} \,, \tag{1.14}$$

where we used the approximation $z - 1 \approx z$ in the last step. Eq. (1.14) constitutes a self-consistency equation for $G = G(\lambda)$, with the solution

$$G^2 - \frac{\lambda}{z} G + \frac{1}{z} = 0, \qquad G = \frac{\lambda}{2z} - \sqrt{\frac{\lambda^2}{4z^2} - \frac{1}{z}} \,,$$

since $\lim_{\lambda \to \infty} G(\lambda) = 0$. The spectral density (1.12) then takes the form

$$\rho(\lambda) = \begin{cases} \sqrt{4z - \lambda^2}/(2\pi z) & \text{if} \quad \lambda^2 < 4z \\ 0 & \text{if} \quad \lambda^2 > 4z \end{cases} \tag{1.15}$$

[2] Random matrix theory will be further discussed in Sect. ?? of Chap. ??, "??".

of a half-ellipse also known as "Wigner law", or the "semi-circle law".

Loops and the Clustering Coefficient The total number of triangles, viz the overall number of loops of length three 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

$$C\frac{z(z-1)}{2}\frac{N}{3} = \text{number of triangles} \\ = \frac{1}{6}\sum_{i_1,i_2,i_3} A_{i_1 i_2}A_{i_2 i_3}A_{i_3 i_1} = \frac{1}{6}\mathrm{Tr}\left[A^3\right] ,$$

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 a moment of the graph spectrum. The factors $1/3$ and $1/6$ account for overcountings. Above formula holds only when all node have identical degrees $k_i \equiv z$. Otherwise one needs to take an average, $z(z-1) \to \sum_i k_i(k_i-1)/N$.

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

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

as one can see from (1.9). The lth 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.

1.2.1 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 a function f_i on a graph with N sites, with $i = 1, \ldots, N$. One defines the graph Laplacian $\hat{\Lambda}$ via

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

where the $\Lambda_{ij} = (\hat{\Lambda})_{ij}$ are the elements of the Laplacian matrix, A_{ij} the adjacency matrix, and where k_i is the degree of vertex i.

Apart from a sign convention, $\hat{\Lambda}$ corresponds 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 Λ_{ij} to a graph-function $\mathbf{f} = (f_1, \ldots, f_N)$. The graph Laplacian is hence intrinsically related to diffusion processes on networks[3]. Alternatively one defines by

$$L_{ij} \;=\; \begin{cases} 1 & i = j \\ -1/\sqrt{k_i\, k_j} & i \text{ and } j \text{ connected} \\ 0 & \text{otherwise} \end{cases} , \tag{1.18}$$

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, which do not generate entries to the adjacency matrix. The respective L_{ij} would be ill defined.

Eigenvalues of the Normalized Graph Laplacian Of interest are the eigenvalues λ_l 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 λ_0 is always zero, corresponding to the eigenfunction

$$\mathbf{e}(\lambda_0) = \frac{1}{\sqrt{C}} \left(\sqrt{k_1}, \sqrt{k_2}, \ldots, \sqrt{k_N} \right) , \tag{1.19}$$

where C is a normalization constant and where the k_i are the respective vertex degrees.

– The degeneracy of λ_0 is given by the number of disconnected subgraphs contained in the network. The eigenfunctions of λ_0 vanish on all subclusters beside one, where it has the functional form (1.19).

[3] See Sect. **??** of Chap. **??**, "**??**", for an indepth discussion of diffusion processes.

– The largest eigenvalue λ_{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 λ_0 has degeneracy one.

Examples of Graph Laplacians The eigenvalues of the normalized graph Laplacian can be evaluated analytically for some reference graphs.

– For a complete N-site graph, viz when all sites are mutually interconnected, the eigenvalues are

$$\lambda_0 = 0, \qquad \lambda_l = N/(N-1), \qquad l = 1, \ldots, N-1\,.$$

– For a complete bipartite graph (all sites of one subgraph are connected to all other sites of the other subgraph) the eigenvalues are

$$\lambda_0 = 0, \qquad \lambda_{N-1} = 2, \qquad \lambda_l = 1, \qquad l = 1, \ldots, N-2\,.$$

The eigenfunction for $\lambda_{N-1} = 2$ has the form

$$\mathbf{e}(\lambda_{N-1}) = \frac{1}{\sqrt{C}}\big(\underbrace{\sqrt{k_A}, \ldots, \sqrt{k_A}}_{\text{A sublattice}}, \underbrace{-\sqrt{k_B}, \ldots, -\sqrt{k_B}}_{\text{B sublattice}}\big)\,. \tag{1.20}$$

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 a substantial number of 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 somewhat distinct.

1.3 Percolation in Generalized Random Graphs

The most random of all graphs are Erdös–Rényi graphs. One can relax the extent of randomness somewhat and construct random networks with an

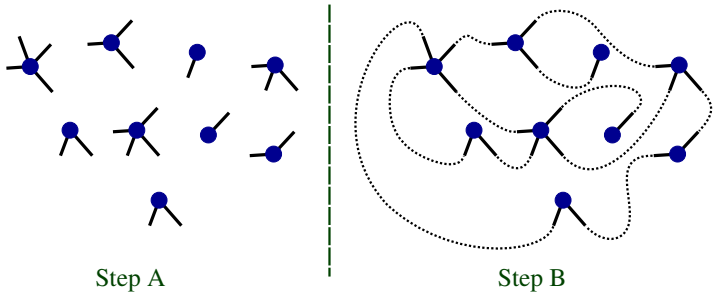

Fig. 1.8 Construction procedure of a random network with nine vertices and degrees $X_1 = 2$, $X_2 = 3$, $X_3 = 2$, $X_4 = 2$. In step A (left) the vertices with the desired number of stubs (degrees) are constructed. In step B (right) the stubs are connected randomly.

arbitrarily given degree distribution. This procedure is also denoted "configurational model". We will use the configurational model to examine the "percolation transition", which occurs as a function for link probability p. Graphs decompose into a set of disconnected subgraphs when containing only a few links, but not for high link densities. The respective transition is the percolation transition.

1.3.1 Graphs with Arbitrary Degree Distributions

In order to generate random graphs having non-Poisson degree distributions, we may choose a specific set of degrees.

DEGREE SEQUENCE A degree sequence is a specified set $\{k_i\}$ of the degrees for the vertices $i = 1 \dots N$.

The degree sequence is also the first information one obtains when examining real-world networks 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 in the thermodynamic limit,

$$p_k, \qquad N \to \infty\,.$$

For the construction, two steps suffice.

- Assign k_i "stubs" (ends of edges emerging from a vertex) to all vertices, $i = 1, \dots, N$.
- Iteratively choose pairs of stubs at random and join them together to make complete edges.

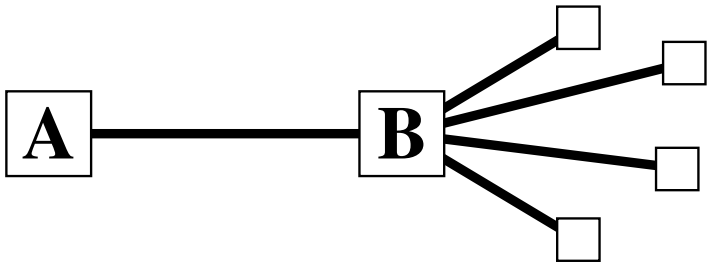

Fig. 1.9 The excess degree distribution q_{k-1} is the probability of finding k outgoing links of a neighboring site. Here the site B is a neighbor of site A and has a degree $k = 5$ and $k - 1 = 4$ outgoing edges, compare (1.21). The probability that B is a neighbor of site A is proportional to the degree k of B.

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 1.8 illustrates the construction procedure.

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. Loops 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.

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 node of a 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. The probability distribution of the degree of the vertex to which an edge leads is thus 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

$$q_{k-1} = \frac{k p_k}{\sum_j j p_j}, \qquad q_k = \frac{(k+1) p_{k+1}}{\sum_j j p_j}\,, \tag{1.21}$$

since $k p_k$ is the degree distribution of a neighbor, compare Fig. 1.9. The distribution q_k of the outgoing edges of a neighbor vertex is the "excess degree distribution". The average number of outgoing edges of a neighbor vertex is then

$$\begin{aligned}\sum_{k=0}^{\infty} k q_k &= \frac{\sum_{k=0}^{\infty} k(k+1) p_{k+1}}{\sum_j j p_j} = \frac{\sum_{k=1}^{\infty} (k-1) k p_k}{\sum_j j p_j} \\ &= \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}\,,\end{aligned} \tag{1.22}$$

where $\langle k^2 \rangle$ is the second moment of the degree distribution.

Number of Next-Nearest Neighbors We denote with

$$z_m, \qquad z_1 = \langle k \rangle \equiv z$$

the average number of m-nearest neighbors. Eq. (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

$$z_2 = \langle k^2 \rangle - \langle k \rangle\,. \tag{1.23}$$

Note that z_2 is not given by the variance of the degree distribution,[4] which would be $\langle (k - \langle k \rangle)^2 \rangle = \langle k^2 \rangle - \langle k \rangle^2$.

z_2 for the Erdös–Rényi graph The degree distribution of an Erdös–Rényi graph is the Poisson distribution, $p_k = \mathrm{e}^{-z} z^k / k!$, see (1.5). For the average number of second neighbors, Eq. (1.23), we obtain

$$\begin{aligned} z_2 &= \sum_{k=0}^{\infty} k^2 \mathrm{e}^{-z} \frac{z^k}{k!} - z = z \mathrm{e}^{-z} \sum_{k=1}^{\infty} (k - 1 + 1) \frac{z^{k-1}}{(k-1)!} - z \\ &= z^2 = \langle k \rangle^2\,. \end{aligned}$$

The mean number of second neighbors of a vertex in an Erdös–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 $\langle k^2 \rangle$, which implies that 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.

[4] Compare Sect. **??** of Chap. **??**, "**??**".

Number of Far Away Neighbors The average number of edges emerging from a second neighbor, and not leading back to where it came from, is also given by (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

$$z_m = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} z_{m-1} = \frac{z_2}{z_1} z_{m-1} \,, \tag{1.24}$$

where $z_1 \equiv z = \langle k \rangle$ and z_2 are given by (1.23). Iterating this relation we find

$$z_m = \left[\frac{z_2}{z_1} \right]^{m-1} z_1 \,. \tag{1.25}$$

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.

$$\lim_{m \to \infty} z_m = \begin{cases} \infty & \text{if} \quad z_2 > z_1 \\ 0 & \text{if} \quad z_2 < z_1 \end{cases} \,. \tag{1.26}$$

The percolation point is therefore $z_1 = z_2$. Below the transition, the total number of neighbors is

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

The total number of sites connected to the starting node is finite below the percolation transition, which implies that the network decomposes into non-connected components. Beyond the percolation transition one or more clusters with macroscopic numbers of nodes form.

> 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.

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 giant connected component.

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

> Percolation Transition A percolation transition occurs when the structure of an evolving graph transits from a state in which two (far away) sites are on the average connected/not connected.

This phase transition occurs when $z_2 = z_1$. Making use of (1.23), $z_2 = \langle k^2 \rangle - \langle k \rangle$, we find that this condition is equivalent to

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

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 a 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, compare Fig. 1.4. The percolation transition can hence 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.

Average Vertex–Vertex Distance Below the percolation threshold the average vertex–vertex distance ℓ 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, ℓ is given approximately by the condition $z_\ell \simeq N$,

$$\log(N/z_1) = (\ell - 1)\log(z_2/z_1), \qquad \ell = \frac{\log(N/z_1)}{\log(z_2/z_1)} + 1\,, \tag{1.28}$$

when using $z_l = z_1(z_2/z_1)^{l-1}$, as given by (1.25). For Erdös–Rényi random graphs one has $z_1 = z$ and $z_2 = z^2$, which leads to

$$\ell = \frac{\log N - \log z}{\log z} + 1 = \frac{\log N}{\log z}\,,$$

as derived prviously, see (1.2).

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 have

$$C = \frac{\langle k_i \rangle_q \langle k_j \rangle_q}{Nz} = \frac{1}{Nz} \left[\sum_k k q_k \right]^2 \tag{1.29}$$
$$= \frac{1}{Nz} \left[\frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \right]^2 = \frac{z}{N} \left[\frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle^2} \right]^2 ,$$

where the notation $\langle \ldots \rangle_q$ indicates that the average is to be taken with respect to the excess degree distribution $q = q_k$, as given by (1.21).

As expected, the clustering coefficient vanishes in the thermodynamic limit $N \to \infty$. 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 Erdös–Rényi graphs, are partly due to the fat tails in the degree distributions p_k of the corresponding networks.

1.3.2 Probability Generating Function Formalism

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

Probability Generating Functions We define with

$$G_0(x) = \sum_{k=0}^{\infty} p_k x^k \tag{1.30}$$

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 $G_0(x)$ by differentiation,

$$p_k = \frac{1}{k!} \frac{\mathrm{d}^k G_0}{\mathrm{d}x^k} \bigg|_{x=0} . \tag{1.31}$$

One says that G_0 "generates" the probability distribution p_k.

Generating Function for the Excess Degree Distribution For later use we define the generating function $G_1(x)$ for the distribution $q_k = (k+1)p_{k+1}/z$ measuring the number of non-returning edges leaving a neighboring vertex,

$$G_1(x) = \sum_{k=0}^{\infty} q_k x^k = \frac{\sum_{k=0}^{\infty} (k+1) p_{k+1} x^k}{z} = \frac{\sum_{k=1}^{\infty} k p_k x^{k-1}}{z}$$
$$= \frac{G_0'(x)}{z}, \qquad z = \langle k \rangle , \tag{1.32}$$

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.

– Normalization
The normalization of p_k implies

$$G_0(1) = \sum_k p_k = 1\,. \tag{1.33}$$

– Mean
Straighforward differentiation,

$$G_0'(1) = \sum_k k\, p_k = \langle k \rangle\,, \tag{1.34}$$

yields the average degree $\langle k \rangle$.
– Moments
The nth moment $\langle k^n \rangle$ of the distribution p_k is given by

$$\langle k^n \rangle = \sum_k k^n p_k \;=\; \left[\left(x \frac{\mathrm{d}}{\mathrm{d}x} \right)^n G_0(x) \right]_{x=1}\,. \tag{1.35}$$

Generating Function for Independent Random Variables Let us assume that we have two random variables, such as two dice. Throwing the two dice correspond to two independent random events. The joint probability to obtain $k = 1, \ldots, 6$ with the first dice and $l = 1, \ldots, 6$ with the second dice 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

$$\Big[\, G_0(x) \,\Big]^n\,.$$

Generating Function of the Poisson Distribution The generating function $G_0 = \sum_k p_k x^k$ of the Poisson distribution $p_k = \mathrm{e}^{-z}\, z^k/k!$ is

$$G_0(x) = \mathrm{e}^{-z} \sum_{k=0}^{\infty} \frac{z^k}{k!}\, x^k = \mathrm{e}^{z(x-1)}\,,$$

with z being the average degree. Compare (1.5). The generating function $G_1(x)$ for the excess degree distribution q_k is

$$G_1(x) = \frac{G_0'(x)}{z} = \mathrm{e}^{z(x-1)}\,, \tag{1.36}$$

compare (1.32). As expected, we find $G_1(x) = G_0(x)$ for the Poisson distribution.

Generating Function of the Exponential Distribution As a second example, consider a graph with an exponential degree distribution:

$$p_k = (1-\mathrm{e}^{-1/\kappa})\,\mathrm{e}^{-k/\kappa}, \qquad \sum_{k=0}^{\infty} p_k = \frac{1-\mathrm{e}^{-1/\kappa}}{1-\mathrm{e}^{-1/\kappa}} = 1\,, \tag{1.37}$$

where κ is a constant.The generating function is

$$G_0(x) = (1-\mathrm{e}^{-1/\kappa}) \sum_{k=0}^{\infty} \mathrm{e}^{-k/\kappa} x^k = \frac{1-\mathrm{e}^{-1/\kappa}}{1-x\mathrm{e}^{-1/\kappa}}\,,$$

together with

$$z = G_0'(1) = \frac{\mathrm{e}^{-1/\kappa}}{1-\mathrm{e}^{-1/\kappa}}, \qquad G_1(x) = \frac{G_0'(x)}{z} = \left[\frac{1-\mathrm{e}^{-1/\kappa}}{1-x\mathrm{e}^{-1/\kappa}}\right]^2.$$

Stochastic Sum of Independent Variables Let's assume that the random variables $k_1,\ k_2, \ldots$ are identically distributed, which implies that they have the same generating functions $G_0(x)$. The powers

$$G_0^2(x), \qquad G_0^3(x), \qquad G_0^4(x)\,,$$

are the generating functions for

$$k_1 + k_2, \qquad k_1 + k_2 + k_3, \qquad k_1 + k_2 + k_3 + k_4\,,$$

and so on. Next 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 to throw n times. The distribution of the results obtained is then generated by

$$\sum_n p_n G_0^n(x) = G_N\left(G_0(x)\right), \qquad G_N(z) = \sum_n p_n z^n\,. \tag{1.38}$$

We will make use of this relation further on.

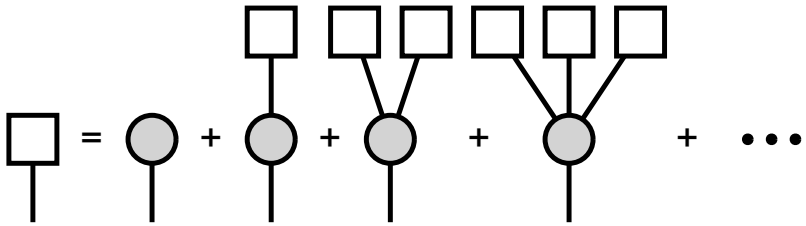

Fig. 1.10 Graphical representation of the self-consistency (1.39) for the generating function $H_1(x)$, represented by the open box. A single vertex is represented by a hashed circle. The subcluster connected to an incoming vertex can be either a single vertex or an arbitrary number of subclusters of the same type connected to the first vertex. In analogy to **?**.

1.3.3 Distribution of Component Sizes

Absence of Closed Loops For networks below the percolation transition we are interested in the distribution of the sizes of the individual subclusters. The calculations will depend on the absense of closed loops, or any other kind of clustering, for generalized random graphs in the thermodynamic limit. In physics jargon, lareg random graphs are "tree-like".

We recall that the number of closed loops of length three corresponds to the clustering coefficient C, viz to the probability that two of your friends are also friends of each other. For random networks $C = [\langle k^2\rangle - \langle k\rangle]^2/(z^3 N)$, see (1.29), which tends to zero as $N \to \infty$.

Generating Function for the Size Distribution of Components We define by

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

the generating function for the distribution of cluster sizes containing a given vertex j, with the condition that j is is linked to a specific incoming edge, see Fig. 1.10. That is, $h_m^{(1)}$ is the probability that the such-defined cluster contains m nodes.

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

- The first vertex j belongs to the subcluster with probability 1, its generating function is x.
- The probability that the vertex j has k outgoing stubs is q_k.
- At every stub outgoing from vertex j there is a subcluster.
- 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 (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. 1.3.2. The self-consistency equation for the total number of vertices reachable is then

$$H_1(x) = x \sum_{k=0}^{\infty} q_k \, [H_1(x)]^k = x \, G_1(H_1(x)) \,, \tag{1.39}$$

where we made use of (1.32) and (1.38).

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.

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

The size of a complete component is thus generated by

$$H_0(x) = x \sum_{k=0}^{\infty} p_k \, [H_1(x)]^k = x \, G_0(H_1(x)) \,, \tag{1.40}$$

where the prefactor x corresponds to the generating function of the starting vertex. The complete distribution of component sizes is given by solving (1.39) self-consistently for $H_1(x)$ and then substituting the result into (1.40).

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 (1.34).

The component size distribution is generated by $H_0(x)$, Eq. (1.40). The the mean component size below the percolation transition is therefore

$$\begin{aligned} \langle s \rangle = H_0'(1) &= \Big[G_0(H_1(x)) + x \, G_0'(H_1(x)) \, H_1'(x) \Big]_{x=1} \\ &= 1 + G_0'(1) H_1'(1) \,, \end{aligned} \tag{1.41}$$

where we made use of the normalization

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

of generating functions, see (1.33). The value of $H_1'(1)$ can be calculated from (1.39) by differentiating:

$$\begin{aligned} H_1'(x) &= G_1(H_1(x)) + x \, G_1'(H_1(x)) \, H_1'(x), \\ H_1'(1) &= \frac{1}{1 - G_1'(1)} \,. \end{aligned} \tag{1.42}$$

Substituting this into (1.41) we find

$$\langle s \rangle = 1 + \frac{G_0'(1)}{1 - G_1'(1)} \,. \tag{1.43}$$

Percolation transition On a closer look, above expression for the mean component size reproduces the percolation condition $z_2 = z_1$ derived in Sect. 1.3. We make use of (1.23) and

$$G_0'(1) = \sum_k k\, p_k = \langle k \rangle = z_1, \tag{1.44}$$
$$G_1'(1) = \frac{\sum_k k(k-1)p_k}{\sum_k k p_k} = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} = \frac{z_2}{z_1} \,,$$

where z_1 and z_2 are the mean numbers of nearest- and next-nearest neighbors. Substitution into (1.43) gives the average component size below the transition as

$$\langle s \rangle = 1 + \frac{z_1^2}{z_1 - z_2} \,. \tag{1.45}$$

This expression has a divergence at $z_1 = z_2$. The mean component size diverges at the percolation threshold, compare Sect. 1.3, beyond which 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,

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

and equivalently for $H_1(x)$. The self-consistency (1.39) and (1.40),

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

then take the form

$$1 - S = H_0(1) = G_0(u), \qquad u = H_1(1) = G_1(u) \,. \tag{1.46}$$

For Erdös–Rényí graphs we have $G_1(x) = G_0(x) = \exp(z(x-1))$, see (1.36). This leads to $1 - S = u$ in (1.46), which takes with

$$1 - S = u = \mathrm{e}^{z(u-1)} = \mathrm{e}^{-zS}, \qquad S = 1 - \mathrm{e}^{-zS} \tag{1.47}$$

the form of a self-consistency condition for the mean size S of the giant component. Percolation takes place for $z \geq 1$.

Critical Behaviour The Taylor expansion for the exponential,

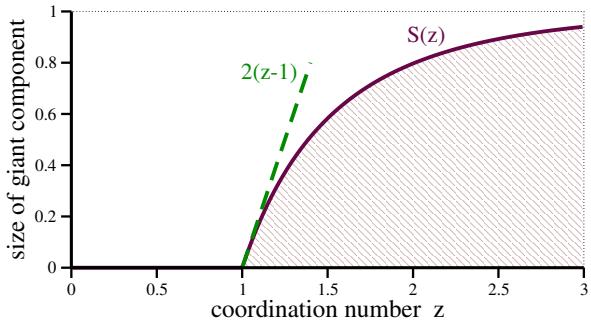

Fig. 1.11 For Erdös–Rényí graphs, the size $S(z)$ of the giant component vanishes linearly, at the percolation transition occurring at the critical coordination number $z_c = 1$. Compare (1.47) and (1.48).

$$S = 1 - \mathrm{e}^{-zS} \approx zS - \frac{(zS)^2}{2}, \qquad S(z) \approx \frac{2(z-1)}{z^2} \tag{1.48}$$

shows that the $S(z)$ increases linearly in the percolating regime, as shown in Fig. 1.11. It follows from (1.47) that the giant component engulfs the entire net exponentially fast for large degrees z, namely that that $\lim_{z\to\infty} S(z) \to 1$.

1.4 Robustness of Random Networks

Degree distributions of real-world networks often have "fat tails", a vaguely defined notion that a given distribution decays only slowly when the degree becomes large. In general, fat tails increase the robustness of a 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 fail.

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 with

$$b = b_k$$

the probability that a vertex of degree k is active. The generating function

$$F_0(x) = \sum_{k=0}^{\infty} p_k\, b_k\, x^k, \qquad F_0(1) = \sum_k p_k\, b_k \;\le\; 1\,, \tag{1.49}$$

generates the probabilities that a vertex has degree k and is present. Strictly speaking, $F_0(x)$ is not a proper generating function because the normalization $F_0(1)$ is not unity, but the fraction of all vertices that are active. This shortcoming could be rectified by adding via

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

a constant to $F(x)$, which corresponds to the number of inactive nodes, which are in turn equivalent to nodes with degree zero. The normalization can hence be disregarded. By analogy with (1.32) we define by

$$F_1(x) = \frac{\sum_k k\, p_k\, b_k\, x^{k-1}}{\sum_k k\, p_k} = \frac{F_0'(x)}{z} \tag{1.50}$$

the generating function for the degree distribution of neighbor sites. Also $F_1(x)$ is not proper normalized.

Distribution of Connected Clusters The distributions of the sizes of connected clusters reachable from a given vertex, $H_0(x)$, or from a given edge, $H_1(x)$, are generated respectively by the normalized functions

$$\begin{aligned} H_0(x) &= 1 - F_0(1) + xF_0(H_1(x)), \qquad H_0(1) = 1, \\ H_1(x) &= 1 - F_1(1) + xF_1(H_1(x)), \qquad H_1(1) = 1\,, \end{aligned} \tag{1.51}$$

which are the logical equivalents of (1.39) and (1.40).

Random Failure of Vertices First we consider the case of random failures of vertices. In this case, the probability

$$b_k \equiv b \ \leq 1, \qquad F_0(x) = b\, G_0(x), \qquad F_1(x) = b\, G_1(x)$$

of a vertex being present is independent of the degree k and just equal to a constant, b, which means that

$$H_0(x) = 1 - b + bxG_0(H_1(x)), \qquad H_1(x) = 1 - b + bxG_1(H_1(x)), \tag{1.52}$$

where $G_0(x)$ and $G_1(x)$ are the standard generating functions for the degree of a vertex and of a neighboring vertex, see (1.30) and (1.32). This implies that the mean size of a cluster of connected and present vertices is

$$\begin{aligned} \langle s \rangle &= H_0'(1) = b + bG_0'(1)\, H_1'(1) \\ &= b + \frac{b^2 G_0'(1)}{1 - bG_1'(1)} = b\left[1 + \frac{bG_0'(1)}{1 - bG_1'(1)}\right], \end{aligned}$$

where we followed the derivation presented in (1.42) in order to derive the expression $H_1'(1) = b/(1 - bG_1'(1))$, which we used. With (1.44) for $G_0'(1) = z_1 = z$ and $G_1'(1) = z_2/z_1$, we obtain the generalization

$$\langle s \rangle = b + \frac{b^2 z_1^2}{z_1 - bz_2}$$

of (1.45). The model has a phase transition at the critical value of b,

$$b_c = \frac{z_1}{z_2} = \frac{1}{G_1'(1)}\,. \tag{1.53}$$

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

$$p_k \;\sim\; \frac{1}{k^\alpha}, \qquad \int \frac{\mathrm{d}k}{k^\alpha} \;<\; \infty, \qquad \alpha > 1\,,$$

see (1.8) and Sect. 1.6. The first two moments are

$$z_1 = \langle k\rangle \sim \int \mathrm{d}k\,(k/k^\alpha), \qquad \langle k^2\rangle \sim \int \mathrm{d}k\,(k^2/k^\alpha)\,.$$

Noting that the number of next-nearest neighbors $z_2 = \langle k^2\rangle - \langle k\rangle$, we can identify three regimes:

- $1 < \alpha \le 2$, with $z_1 \to \infty$ and $z_2 \to \infty$.
  $b_c = z_1/z_2$ is arbitrary in the thermodynamic limit $N \to \infty$.
- $2 < \alpha \le 3$, with $z_1 < \infty$ and $z_2 \to \infty$.
  $b_c = z_1/z_2 \to 0$ 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.
- $3 < \alpha$, with $z_1 < \infty$ and $z_2 < \infty$.
  $b_c = z_1/z_2$ can acquire any value and the network has normal robustness.

Biased Failure of Vertices What happens when somebody 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

$$b_k = \theta(k_c - k)\,, \tag{1.54}$$

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

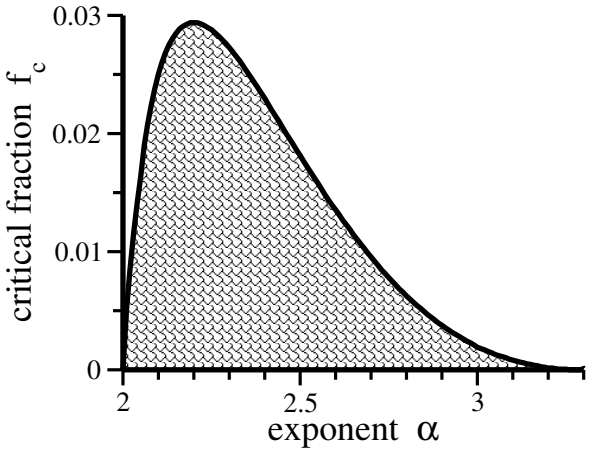

Fig. 1.12 For a scale-free network with a power-law degree distribution $p_k \sim 1/k^\alpha$ the critical fraction f_c of vertices, as defined by (1.59). Removing a fraction greater than f_c of high degree vertices drives the network below the percolation limit. For a smaller loss of highest degree vertices the giant connected component remains intact (shaded area).

$$\theta(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases} . \tag{1.55}$$

This form of targeted attacts correspond to setting the upper limit of the sum in (1.49) to k_c. Differentiating (1.51) with respect to x yields

$$H_1'(1) = F_1(H_1(1)) + F_1'(H_1(1))\, H_1'(1), \qquad H_1'(1) = \frac{F_1(1)}{1 - F_1'(1)} ,$$

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

$$\frac{\sum_{k=1}^{\infty} k(k-1)p_k b_k}{\sum_{k=1}^{\infty} k p_k} = \frac{\sum_{k=1}^{k_c} k(k-1)p_k}{\sum_{k=1}^{\infty} k p_k} = 1 , \tag{1.56}$$

where we used the definition (1.50) 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 (1.56) as

$$H_{k_c}^{(\alpha-2)} - H_{k_c}^{(\alpha-1)} = H_{\infty}^{(\alpha-1)} , \tag{1.57}$$

where $H_n^{(r)}$ is the nth harmonic number of order r,

$$H_n^{(r)} = \sum_{k=1}^{n} \frac{1}{k^r} . \tag{1.58}$$

The number of vertices present is $F_0(1)$, see (1.49), 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 (1.57),

$$f_c = 1 - \frac{F_0(1)}{\sum_k p_k} = 1 - \frac{H_{k_c}^{(\alpha)}}{H_\infty^{(\alpha)}} \,. \tag{1.59}$$

It is impossible to determine k_c from (1.57) and (1.59) to get f_c in closed form. However one can, solve (1.57) numerically for k_c and substitute it into (1.59). The results are shown in Fig. 1.12, as a function of the exponent α. 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. Compare Fig. 1.6.
- Networks with $\alpha > \alpha_c = 3.4788$ never have have a giant connected component. The critical exponent α_c is given by the percolation condition $H_\infty^{(\alpha-2)} = 2H_\infty^{(\alpha-1)}$, see (1.27).

1.5 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. 1.13 we illustrate a simple one-dimensional lattice with connectivity z. For periodic boundary conditions, viz when the chain wraps around itself in a ring, one can evaluate the clustering coefficient C analytically.

- One dimension
  For number of clusters one finds

$$C = \frac{3(z-2)}{4(z-1)} \,, \tag{1.60}$$

  which tends to 3/4 in the limit of large z.
- General dimension
  Square or cubic lattices have dimension $d = 2, 3$, respectively. The clustering coefficient for general dimensions d is

$$C = \frac{3(z-2d)}{4(z-d)} \,, \tag{1.61}$$

  which generalizes (1.60). We note that the clustering coefficient tends to 3/4 for $z \gg 2d$ for regular hypercubic lattices in all dimensions.

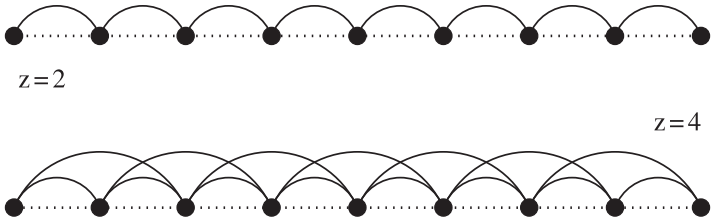

Fig. 1.13 Regular linear graphs with connectivities $z = 2$ (top) and $z = 4$ (bottom).

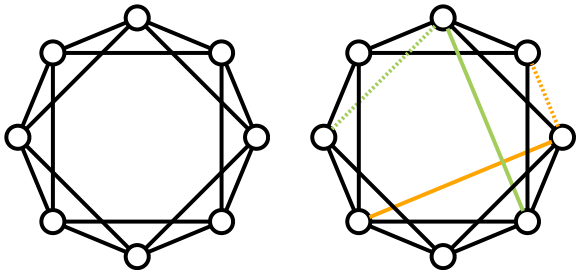

Fig. 1.14 Crossover from a regular lattice to a small-world network via rewiring. **Left:** Regular lattice with periodic boundary conditions. **Right:** Two rewired links, original and new links are color coded (dashed/solid).

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 proposed a small-world model that interpolates smoothly between a regular lattice and an Erdös–Rényi random graph. The construction starts with a one-dimensional periodic lattice, see Fig. 1.14. There are several possibilities

- One goes through all the links of the lattice and rewires the link with a given probability p_{new}.
- A single edge is selected and rewired, with the procedure repeated N_{new} times.
- As a variation one may add links, instead of rewiring. The advantage is that networks may not become disconnected.

For small p_{new} and/or N_{new}, this process produces a graph that is still mostly regular, possessing however a few connections stretching long distances across the lattice, as illustrated in Fig. 1.14. The average coordination number z of the lattice remains constant. The number of neighbors of any particular vertex can, however, be greater or smaller than z.

Social Network Interpretation The small-world models such as the ones illustrated in Fig. 1.14, 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 relatives. However, some people

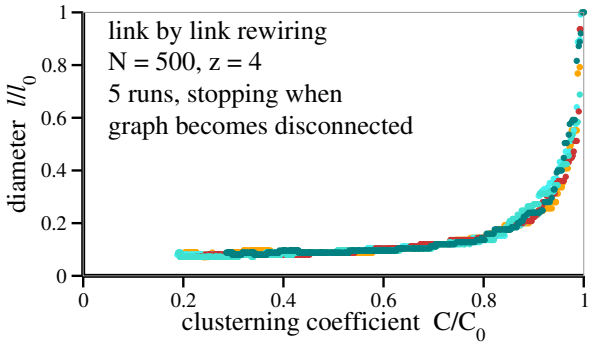

Fig. 1.15 The scaled clustering coefficient C/C_0 as as a function of the scaled diameter l/l_0, where $C_0 = 3/4$ and $l = N/4$ are the initial values, for a one-dimensional ring with $N = 500$ sites and $z = 4$, compare Fig. 1.14. Points correspond to the consecutive random rewiring of a single edge.

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 small-world models.

In Fig. 1.15 the clustering coefficient is shown as a function of network diameter when rewiring is performed edge by edge. The key result is that a few steps are sufficient to achieve a drastic reduction of intra-network distances, as measured by the graph diameter. At the same time, the clustering coefficient remains high, as observed in many real-world networks.

1.6 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 world wide web, which grows exponentially by the continuous addition of new web pages. The small world networks discussed in Sect. 1.5 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.

> Preferential Connectivity New vertices prefer to connect to exisiting nodes which already have high degrees.

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 is added and connected with $m \leq N_0$ links to exisiting nodes.
- Preferential Attachment
  The new links are determined by connecting the m stubs of the new vertex to nodes with degrees k_i with probability

$$\Pi(k_i) = \frac{k_i + C}{a(t)} \qquad a(t) = 2m(t-1) + C(N_0 + t - 1)\,, \tag{1.62}$$

  where the normalization $a(t)$ has been chosen such that the overall probability to connect is unity, $\sum_i \Pi(k_i) = 1$.

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 above two steps leads to a network with $N(t) = t + N_0$ vertices and mt edges, see Fig. 1.16. We will now show that preferential attachment leads to a scale-free degree distribution

$$p_k \;\sim\; k^{-\gamma} \qquad \gamma = 3 + \frac{C}{m}\,. \tag{1.63}$$

This scaling relation is valid for large degrees k_i in the limit $t \to \infty$.

Time-Dependent Connectivities The time dependence of the degree of a given vertex can be calculated analytically using a mean-field approach. Only nodes with large degrees matter, as the scaling relation (1.63) applies asymptotically in the limit $k \to \infty$, We may therefore assume k_i to be continuous,

$$\Delta k_i(t) \equiv k_i(t+1) - k_i(t) \approx \frac{\partial k_i}{\partial t} = A\,\Pi(k_i) = A\,\frac{k_i + C}{a(t)}\,, \tag{1.64}$$

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 $a(t) = 2m(t-1) + C(N_0 + t - 1)$, we obtain

$$\frac{\partial k_i}{\partial t} = m\frac{k_i + C}{(2m + C)t + a_0} \;\approx\; \frac{m}{2m + C}\frac{k_i + C}{t}\,, \tag{1.65}$$

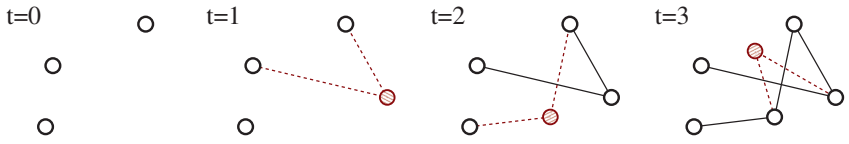

Fig. 1.16 Illustration of the preferential attachment model for an evolving network. At $t = 0$ the system consists of $N_0 = 3$ isolated vertices. At every time step a new vertex (shaded circle) is added, which is connected to $m = 2$ vertices, preferentially to the vertices with high connectivity, determined by the rule (1.62).

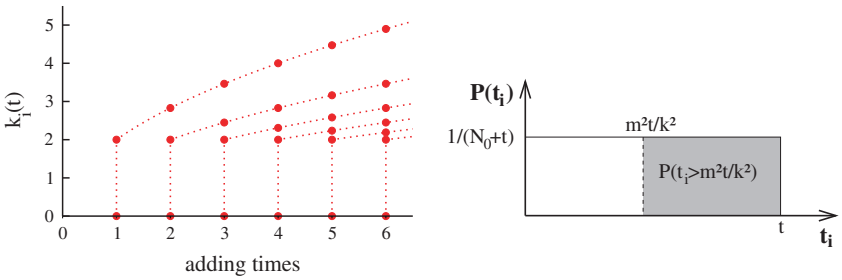

Fig. 1.17 Left: Time evolution of the connectivities for vertices with adding times $t = 1, 2, 3, \ldots$ and $m = 2$, following (1.67). **Right:** The integrated probability, $P(k_i(t) < k) = P(t_i > tm^2/k^2)$, see (1.68).

where did neglect $a_0 = C(N_0 - 1) - 2m$ for large times t. The solution of (1.65) is given by

$$\frac{\dot{k}_i}{k_i + C} = \frac{m}{2m + C}\,\frac{1}{t}, \qquad k_i(t) = (C + m)\left(\frac{t}{t_i}\right)^{m/(2m+C)} - C\,, \tag{1.66}$$

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 specialize to the case $C = 0$, the general result for $C > 0$ can be obtained subsequently from scaling considerations. A vertex added at time $t_i = N_i - N_0$ has initially $m = k_i(t_i)$ links, in agreement with (1.66), which reads

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{0.5}, \qquad t_i \;= t\,m^2/k_i^2 \tag{1.67}$$

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. 1.17. 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,

$$t_1, \ldots, t_{N_0} = 0, \qquad t_{N_0+j} \;= j, \qquad j = 1, 2, \ldots\,,$$

where we did define the initial N_0 nodes to have adding times zero.

Integrated Probabilities Using (1.67), 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

$$P(k_i(t) < k) = P\left(t_i > \frac{m^2 t}{k^2}\right) . \tag{1.68}$$

Adding times are uniformly distributed, compare Fig. 1.17, which implies that the probability $P(t_i)$ to find an adding time t_i is

$$P(t_i) = \frac{1}{N_0 + t} , \tag{1.69}$$

viz just the inverse of the total number of adding times, which coincides in turn 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 $P(t_i)$ to add a new node,

$$P\left(t_i > \frac{m^2 t}{k^2}\right) = \left(t - \frac{m^2 t}{k^2}\right) \frac{1}{N_0 + t} . \tag{1.70}$$

Scale-Free Degree Distribution The degree distribution p_k follows from (1.70) via a simple differentiation,

$$p_k = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{\partial P(t_i > m^2 t/k^2)}{\partial k} = \frac{2m^2 t}{N_0 + t} \frac{1}{k^3} , \tag{1.71}$$

in accordance with (1.63).

The degree distribution (1.71) has a well defined limit $t \to \infty$, approaching a stationary distribution. We note that the exponent $\gamma = 3$ 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 investigate variants of above model.

One can repeat above calculation for a finite offset $C > 0$. The exponent γ is identical to the inverse of the scaling power of k_i with respect to time t in (1.66), plus one, which leads to $\gamma = (2m + C)/m + 1 = 3 + C/m$.

Growth with Random Attachment We examine whether growth alone can result in a scale-free degree distribution, assuming random instead of preferential attachment. The growth equation for the connectivity k_i of a given node i, compare (1.65), then takes the form

$$\frac{\partial k_i}{\partial t} = \frac{m}{N_0 + t} . \tag{1.72}$$

The m new edges are linked randomly at time t to the $(N_0 + t - 1)$ nodes present at the previous time step. Solving (1.72) for k_i, with the initial condition $k_i(t_i) = m$, we obtain

$$k_i = m \left[\ln \left(\frac{N_0 + t}{N_0 + t_i} \right) + 1 \right] , \tag{1.73}$$

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

$$\begin{aligned} P(k_i(t) < k) &= P\left(t_i > (N_0 + t)\, \mathrm{e}^{1-k/m} - N_0 \right) \\ &= \left[t - (N_0 + t)\, \mathrm{e}^{1-k/m} + N_0 \right] \frac{1}{N_0 + t} , \end{aligned} \tag{1.74}$$

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

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

for large times t, we find

$$p_k = \frac{1}{m}\, \mathrm{e}^{1-k/m} = \frac{\mathrm{e}}{m}\, \exp\left(-\frac{k}{m} \right) . \tag{1.75}$$

Growing networks with random attachment are characterized hence by

$$k^* = m , \tag{1.76}$$

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 (1.75) is not properly normalized, nor in (1.71), 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 Internet and the Wikipedia network, however, have exponents $2 < \gamma < 3$, with the exponent γ being relatively close to two. 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.

- At every time step a new vertex is added.
- At every time step m new edges are added.
- 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 (1.62).
- 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 γ can be evaluated along the lines used above for the case $r = 1$. One finds

$$p_k \sim \frac{1}{k^\gamma}, \qquad \gamma = 1 + \frac{1}{1 - r/2}\,. \tag{1.77}$$

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