Appendix 2

Generating functions formalism

The percolation theory in networks with arbitrary degree distribution developed in Chapter 6 can be treated using the generating function formalism, as developed by Callaway *et al*. (2000) and Newman *et al.* (2001).

Let us note for simplicity $P_k = P(k)$, the probability that a randomly chosen node has degree k. The *generating function* of the distribution P_k is

$$G_0(x) = \sum_k P_k x^k. \tag{A2.1}$$

The name "generating function" comes from the fact that each P_k can be recovered from the knowledge of $G_0(x)$ through the following formula:

$$P_k = \frac{1}{k!} \frac{\mathrm{d}^k}{\mathrm{d}x^k} G_0 \bigg|_{x=0}, \tag{A2.2}$$

i.e. by taking the kth derivative of G_0, while the moments of the distribution are given by

$$\langle k^n \rangle = \left(x \frac{\mathrm{d}}{\mathrm{d}x} \right)^n G_0 \bigg|_{x=1}. \tag{A2.3}$$

For example, it is easy to see that the average degree $\langle k \rangle = \sum_k k P_k$ is equal to $G_0'(1)$.

Another important property of G_0 is the following: the distribution of the sum of the degrees of m vertices taken at random is generated by the mth power of G_0, G_0^m. This is easy to see in the case $m = 2$:

$$G_0^2(x) = \left(\sum_k P_k x^k \right)^2 = \sum_k \left(\sum_{k_1, k_2} P_{k_1} P_{k_2} \delta_{k_1+k_2,k} \right) x^k, \tag{A2.4}$$

where the Kronecker symbol is such that $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ if $i \neq j$. The coefficient of x^k thus the sum of all terms $P_{k_1} P_{k_2}$ such that the sum of k_1 and k_2 is k, i.e. exactly the probability that the sum of the degrees of two independent vertices is k.

While P_k is the probability that a randomly chosen node has degree k, the probability that a node reached *by following a randomly chosen edge* has degree k is $k P_k / \langle k \rangle$. The probability that such a node has k "outgoing" links, i.e. k other than the edge followed to

https://doi.org/10.1017/CBO9780511791383.016 Published online by Cambridge University Press

reach it, is thus $q_k = (k+1)P_{k+1}/\langle k \rangle$, and the corresponding generating function therefore reads

$$G_1(x) = \sum_k \frac{(k+1)P_{k+1}}{\langle k \rangle} x^k = \frac{1}{\langle k \rangle} G_0'(x). \tag{A2.5}$$

The definition and properties of generating functions allow us now to deal with the problem of percolation in a random network: let us call $H_1(x)$ the generating function for the distribution of the sizes of the connected components reached by following a randomly chosen edge. Note that $H_1(x)$ considers only *finite* components and therefore excludes the possible giant cluster. We neglect the existence of loops, which is indeed legitimate for such finite components. The distribution of sizes of such components can be visualized by a diagrammatic expansion as shown in Figure A2.1: each (tree-like) component is composed by the node initially reached, plus k other tree-like components, *which have the same size distribution*, where k is the number of outgoing links of the node, whose distribution is q_k. The probability that the global component Q_S has size S is thus

$$Q_S = \sum_k q_k \text{ Prob(union of } k \text{ components has size } S-1) \tag{A2.6}$$

(counting the initially reached node in S). The generating function H_1 is by definition

$$H_1(x) = \sum_S Q_S x^S, \tag{A2.7}$$

and the distribution of the sum of the sizes of the k components is generated by H_1^k (as previously explained for the sum of degrees), i.e.

$$\sum_S \text{Prob(union of } k \text{ components has size } S) \cdot x^S = (H_1(x))^k. \tag{A2.8}$$

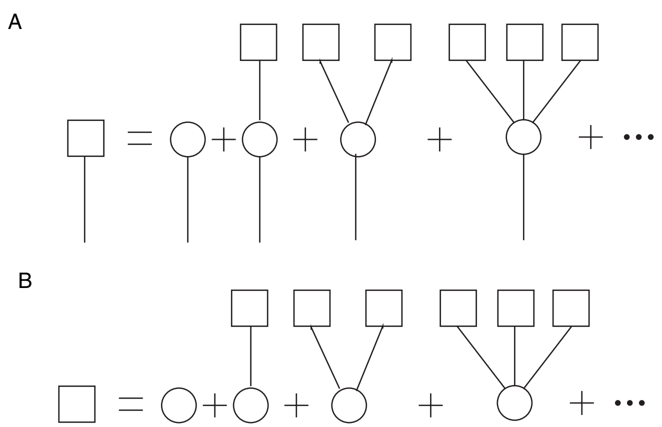

Fig. A2.1. Diagrammatic visualization of (A) Equation (A2.9) and (B) Equation (A2.10). Each square corresponds to an arbitrary tree-like cluster, while the circle is a node of the network.

https://doi.org/10.1017/CBO9780511791383.016 Published online by Cambridge University Press

Putting together equations (A2.6), (A2.7), and (A2.8), one obtains the following self-consistency equation for H_1:

$$\begin{aligned} H_1(x) &= xq_0 + xq_1 H_1(x) + xq_2(H_1(x))^2 + \cdots + xq_k(H_1(x))^k + \cdots \\ &= xG_1(H_1(x)). \end{aligned} \quad \text{(A2.9)}$$

We now define $H_0(x)$ as the generating function of the sizes of connected components obtained starting from a randomly chosen vertex. A similar diagrammatic expansion can be written (Figure A2.1): a component either is restricted to the vertex if it has degree 0, or is the union of k components reached through the k links of the vertex. As previously the size of this union is generated by H_1^k, so that:

$$\begin{aligned} H_0(x) &= xP_0 + xP_1 H_1(x) + xP_2(H_1(x))^2 + \cdots + xP_k(H_1(x))^k + \cdots \\ &= xG_0(H_1(x)). \end{aligned} \quad \text{(A2.10)}$$

Equation (A2.9) is in general complicated and cannot be easily solved. Together with Equation (A2.10), however, it allows us to obtain the average size of the connected clusters. This size is given by $\langle s \rangle = H_0'(1)$, i.e. by deriving Equation (A2.10) in $x = 1$:

$$H_0'(1) = G_0(H_1(1)) + 1 \cdot H_1'(1) \cdot G_0'(H_1(1)). \quad \text{(A2.11)}$$

If no giant component is present, normalization of the size distribution ensures that $H_1(1) = 1$, and $G_0(1) = 1$ as well by normalization of P_k. The derivative $H_1'(1)$ can also be computed by deriving Equation (A2.9) in $x = 1$, yielding

$$H_1'(1) = \frac{1}{1 - G_1'(1)}. \quad \text{(A2.12)}$$

Moreover, $G_0'(1) = \langle k \rangle$, and $G_1'(1) = G_0''(1)/\langle k \rangle = \kappa - 1$ where $\kappa = \langle k^2 \rangle / \langle k \rangle$ is the heterogeneity parameter of the network, so that one finally obtains

$$\langle s \rangle = 1 + \frac{\langle k \rangle}{2 - \kappa} \quad \text{(A2.13)}$$

which diverges when $\langle k^2 \rangle = 2\langle k \rangle$. We recover the Molloy–Reed criterion for the transition percolation, Equation (6.13). When a giant component is present, i.e. for $\langle k^2 \rangle > 2\langle k \rangle$, the computations can still be carried out, with slight modifications since the size distribution is no longer normalized (Newman *et al.*, 2001).