Time evolution of a random graph

2005 J. Phys. A: Math. Gen. 38 L777

(http://iopscience.iop.org/0305-4470/38/46/L02)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **38** (2005) L777–L782 doi:10.1088/0305-4470/38/46/L02

LETTER TO THE EDITOR

Time evolution of a random graph

A A Lushnikov

Department of Physical Sciences, University of Helsinki, PO Box 64, FIN-00014, Helsinki, Finland

E-mail: alex.lushnikov@helsinki.fi

Received 8 September 2005, in final form 3 October 2005
Published 2 November 2005
Online at stacks.iop.org/JPhysA/38/L777

Abstract

An initially empty (no edges) graph of the order of M is assumed to evolve by adding one edge at a time. This edge can connect either two linked components and form a new component of a larger order (coalescence of graphs) or increase (by one) the number of edges in a given linked component. The evolution equation for the generating functional of the probability to find in the system a given set of occupation numbers (the numbers of graphs of the order of g having exactly v edges) at time t is formulated and solved exactly. The expression for the graph composition spectrum is derived and analysed in the limit of large M.

PACS numbers: 02.50.−r, 05.90.+m, 89.75.He

Let us imagine a system comprising M functional units connected with $\mathcal{N} \leqslant M(M-1)/2$ links allowing for an exchange of the information between them. One easily recognizes a schematic model of a computer, the Internet, a living organism or the human brain. Other (more physical examples) of such systems are polymers, disordered materials, random electric chains. The structures of all these systems can be modelled by a random graph in which $\mathcal{N}$ edges randomly distributed among M vertices form clusters (linked components). Any realization of the random graph can be given by $\{n_{g,v}\}$, a set of the numbers of linked components of the order of g with v edges. The problem is to find the probability for the realization of a given partition, once $\mathcal{N}$ edges are randomly distributed among M vertices. In this statement the problem had been introduced by Erdös and Renyi (1960) and then considered by many authors (see recent review article by Albert and Barabási (2002) and references therein).

All commonly accepted approaches to the problem of the evolution of random graphs rely upon rather sophisticated combinatorial considerations (see, e.g., Bollobás (1985), Janson *et al* (2000)). In this letter, I propose another approach based on the formulation and the solution of an equation describing the time evolution of the generating functional for the probability to find a given set of occupation numbers at time t. The idea of this approach is taken from my recent articles Lushnikov (2004, 2005a, 2005b, 2005c), where the sol-to-gel transition was investigated in finite coagulating systems. Being applied to the problem of the evolution of random graphs this approach gives very impressive and elegant results that I

never met before. Although the analogy between the evolution of a coagulating system and a random graph is quite straightforward and had been noticed long ago (Leyvraz 2003, Spouge 1985), the approach proposed below allows one not only to trace it in detail, but also to use it for obtaining new interesting results in theory of random graphs. The results presented in this letter are exact and valid for any *finite* random graph. It is shown that the total number of all linked components of the order of g exactly repeats the fate of the particle mass spectrum of a finite coagulating system with the kernel proportional to the product of the masses of coalescing particles.

Let there be a graph of the order of M comprising N linked components. Each linked component can be characterized with its order g, the number of vertices, and the degree of filling ν (the number of edges in the component). It is clear that $g-1 \leqslant \nu \leqslant g(g-1)/2$. The minimal value of ν corresponds to a tree of the order of g and the maximal one is the number of edges in the complete graph of the order of g. A bare vertex is also regarded as a linked component of the order of 1.

Any state of the graph can be given by the set of the population numbers

$$Q = \{n_{g,\nu}\}, \tag{1}$$

where $n_{g,\nu}$ is the number of linked components of the order of g having exactly ν edges (g, ν-component).

Let us consider an initially empty graph of the order of M (simply M bare vertices) and begin to add to this graph the edges (one edge at a time) linking two valent vertices. This process gives rise either to a coalescence of two linked components,

$$(l, \lambda) + (m, \mu) \longrightarrow (l+m, \lambda+\mu+1) \tag{2}$$

or to filling a given linked component with one extra-edge,

$$(g, \nu) \longrightarrow (g, \nu+1). \tag{3}$$

The graph thus evolves due to changing the number of linked components, their order and their degree of filling. The efficiency of the coalescence process is proportional to lm, the number of ways to connect two linked components by an extra-edge. The efficiency for filling a linked component is proportional to $g(g-1)/2-\nu$, the number of possible vacant placements of the extra-edge.

Now we introduce the probability $W(Q,t)$ to find the graph in the state Q at time t and its generating functional,

$$\Psi(X,t) = \sum_{Q} W(Q,t) \prod_{g,\nu} x_{g,\nu}^{n(g,\nu|Q)}, \tag{4}$$

where $n(g,\nu|Q)$ stands for the occupation number $n_{g,\nu}$ belonging to the given state Q and $X = \{x_{g,\nu}\}$ is the set of independent *formal* variables. The functional Ψ obeys the equation,

$$V\frac{\partial \Psi}{\partial t} = (\hat{L}_f + \hat{L}_c)\Psi, \tag{5}$$

where the multiplier V defines the scale of time. In theory of coagulation V is the total volume of the system. In theory of random graph the choice $V = M$ is preferable.

The right-hand side of this equation contains two differential operators, $\hat{L}_f$ and $\hat{L}_c$. The operator $\hat{L}_f$ is responsible for the evolution of the filling of the linked clusters

$$\hat{L}_f = \sum_{l,\lambda} \left[\left(\frac{l(l-1)}{2} - \lambda + 1 \right) x_{l,\lambda} \frac{\partial}{\partial x_{l,\lambda-1}} - \left(\frac{l(l-1)}{2} - \lambda \right) x_{l,\lambda} \frac{\partial}{\partial x_{l,\lambda}} \right]. \tag{6}$$

The coalescence operator has the form

$$\hat{L}_c = \frac{1}{2}\sum_{l,\lambda;m,\mu} lm(x_{l+m,\lambda+\mu+1} - x_{l,\lambda}x_{m,\mu})\frac{\partial^2}{\partial x_{l,\lambda}\partial x_{m,\mu}}. \tag{7}$$

Equation (5) replaces the master equation governing the time evolution of the probability $W(Q,t)$. Because the order of the initial graph conserves

$$M = \sum_{l,\lambda} ln_{l,\lambda}, \tag{8}$$

the solution to the evolution equation (5) can be found in the form:

$$\Psi = M!\mathbf{Coef}_z z^{-(M+1)} \exp\left(\sum_{g,\nu} a_{g,\nu}(t)x_{g,\nu}z^g\right), \tag{9}$$

where the multiplier $M!$ provides the correct normalization of the generating functional, $\Psi(1,t) = 1$. The operation **Coef** introduced in (Egorychev 1977) replaces the contour integration when one deals with formal series. By definition

$$\mathbf{Coef}_z \sum_k b_k z^k = b_{-1}.$$

The operation **Coef** displays many features of ordinary residues.

On substituting equation (9) into equation (5) gives the system of equations for $a_{g,\nu}(t)$,

$$V\frac{\mathrm{d}a_{g,\nu}}{\mathrm{d}t} = \left[\frac{1}{2}g(g-1) - \nu + 1\right]a_{g,\nu-1} - \left[\frac{1}{2}g(g-1) - \nu\right]a_{g,\nu} + \frac{1}{2}\sum_{l,\lambda;m,\mu} lma_{l,\lambda}a_{m,\mu}\delta_{g,l+m}\delta_{\nu,\lambda+\mu+1} - \frac{1}{2}Mga_{g,\nu} + \frac{g^2}{2}a_{g,\nu}. \tag{10}$$

Here $\delta_{\alpha,\beta}$ stands for Kroneker's delta. The initial condition to equation (10) is chosen in the form

$$a_{g,\nu} = \delta_{g,1}\delta_{\nu,0}. \tag{11}$$

It is easy to check that this initial condition corresponds to $\Psi(X,0) = x_{1,0}^M$, i.e., to the initially empty graph comprising only M bare vertices. Next, all $a_{g,\nu}(t) = 0$, once ν lies beyond the permitted interval $(g-1) \leqslant \nu \leqslant g(g-1)/2$.

Now let us try to solve equation (10). To this end we introduce the bivariate generating function for $a_{g,\nu}(t)$,

$$G(z,\zeta;t) = \sum_{g,\nu} z^g\zeta^\nu a_{g,\nu}(t). \tag{12}$$

The summation on the right-hand side of the above equation goes over all g and ν. The equation for $G(z,\zeta;t)$ immediately follows from equation (10),

$$V\frac{\partial G}{\partial t} = \frac{\zeta}{2}\left[\left(z\frac{\partial G}{\partial z}\right)^2 + z\frac{\partial}{\partial z}z\frac{\partial G}{\partial z}\right] - (\zeta-1)\left(\zeta\frac{\partial G}{\partial \zeta} + \frac{1}{2}z\frac{\partial G}{\partial z}\right) - \frac{M}{2}z\frac{\partial G}{\partial z}. \tag{13}$$

This equation reduces to a linear one for the function $D(z,\zeta;t)$

$$V\frac{\partial D}{\partial t} = \frac{\zeta}{2}z\frac{\partial}{\partial z}z\frac{\partial D}{\partial z} - (\zeta-1)\left(\zeta\frac{\partial D}{\partial \zeta} + \frac{1}{2}z\frac{\partial D}{\partial z}\right), \tag{14}$$

where

$$D(z,\zeta;t) = \exp[G(z\,\mathrm{e}^{Mt/2V},\zeta;t)].$$

The initial condition to this equation follows from equation (11),

$$D(z,\zeta;0)=\mathrm{e}^{z}. \tag{15}$$

Equation (14) is readily solved by separating variables. Let

$$D(z,\zeta;t)=\sum_{n,\kappa} b_{n,\kappa}T_{\kappa}(t,)Z_{n,\kappa}(\zeta)z^{n}.$$

Then $T_{\kappa}(t)=\mathrm{e}^{\kappa t}$ and

$$\kappa Z_{n,\kappa}=\frac{\zeta}{2}n^{2}Z_{n,\kappa}+(1-\zeta)\left(\zeta\frac{\mathrm{d}Z_{n,\kappa}}{\mathrm{d}\zeta}+\frac{n}{2}Z_{n,\kappa}\right), \tag{16}$$

where κ is a separation constant. The solution to this equation is

$$Z_{n,\kappa}(\zeta)=b_{n,\kappa}(1-\zeta)^{n^{2}/2-\kappa}\zeta^{\kappa-n/2}. \tag{17}$$

The function D should be analytical at $\zeta=0$. Hence, $\kappa-n/2=s$, where s is a nonnegative integer. Next, the coefficients $b_{n,\kappa}$ should be chosen from the initial condition (15). It is easy to see that $b_{n,\kappa}=\binom{(n^{2}-n)/2}{s}$. We then come to the result

$$D(z,\zeta;t)=\sum_{n=0}^{\infty}\frac{z^{n}}{n!}\mathrm{e}^{nt/2V}[\zeta\,\mathrm{e}^{t/V}+(1-\zeta)]^{(n^{2}-n)/2}. \tag{18}$$

In order to return to $a_{g,\nu}(t)$ we use the Knuth identity (Knuth 1998)

$$\ln\sum_{n=1}^{\infty}x^{n(n-1)/2}\frac{z^{n}}{n!}=\sum_{n=1}^{\infty}(x-1)^{n-1}F_{n-1}(x)\frac{z^{n}}{n!}. \tag{19}$$

Here $F_{n}(x)$ are the Mallows–Riordan polynomials (details and earlier references see Lushnikov 2005a).

Equation (19) allows us to restore $A_{g}(\zeta,t)=\sum_{\nu}a_{g,\nu}(t)\zeta^{\nu}$,

$$A_{g}(\zeta,t)=\frac{1}{g!}\mathrm{e}^{-g(M-1)t/2V}(\mathrm{e}^{t/V}-1)^{g-1}\zeta^{g-1}F_{g-1}(\zeta\,\mathrm{e}^{t/V}+1-\zeta). \tag{20}$$

Next, we use the expansion (Knuth 1998)

$$w^{g-1}F_{g-1}(1+w)=\sum_{\nu=g-1}^{g(g-1)/2}C_{g,\nu}w^{\nu}, \tag{21}$$

where $C_{g,\nu}$ is the number of labelled linked graphs of the order of g having ν edges. Equation (21) is readily applied for restoring $a_{g,\nu}(t)$. The result is

$$a_{g,\nu}(t)=\frac{1}{g!}\mathrm{e}^{-g(M-1)t/2V}(\mathrm{e}^{t/V}-1)^{\nu}C_{g,\nu}. \tag{22}$$

Now we are ready to find the average number of linked components of the order of g having exactly ν edges. From equation (9) we have

$$\bar{n}_{g,\nu}(t)=M!a_{g,\nu}(t)\,\mathbf{Coef}_{z}z^{-M+g-1}D(z,1,t). \tag{23}$$

At $\zeta=1$ equation (18) gives

$$D(z,1;t)=\sum_{n=0}^{\infty}\frac{z^{n}}{n!}\mathrm{e}^{n^{2}t/2V}. \tag{24}$$

We thus come to the result

$$\bar{n}_{g,\nu}(t)=\binom{M}{g}\mathrm{e}^{(g^{2}-2Mg+g)t/2V}(\mathrm{e}^{t/V}-1)^{\nu}C_{g,\nu}. \tag{25}$$

Because the function $D(z, 1; t)$ coincides with the D-function corresponding to the spectrum of coagulating particles in the system with the kernel $K(g, l) = gl$ considered in Lushnikov (1978, 2004, 2005a, 2005b) we can easily derive the expression for the distribution of linked components over their orders (numbers of vertices), $\bar{n}_g = \sum_\nu \bar{n}_{g,\nu}$. The result has the same form as the average particle mass spectrum in the coagulating system with the coagulation kernel $K(g, l) = gl$,

$$\bar{n}_g(t) = \binom{M}{g} \mathrm{e}^{(g^2-2Mg+g)t/2V} (\mathrm{e}^{t/V} - 1)^{g-1} F_{g-1}(\mathrm{e}^{t/V}). \tag{26}$$

In coagulating systems this spectrum is known to conserve the total mass $M = \sum_g g\bar{n}_g$ and after the critical time $t = t_c = 1$ a giant particle forms with the mass comparable to M. For details see in Lushnikov (2005a). Of course, the spectrum (26) contains this giant component. However, the time dependence of partial distributions given by equation (25) differs from that given by equation (26) and the questions come up what does it go on to the partial distributions? Does a giant component present in them?

In order to answer these questions let us apply the result (25) to the simplest situation when $\nu = g - 1$, i.e. all linked components are trees. In this case $C_{g,g-1} = g^{g-2}$. The distribution of the trees is then given by the formula

$$\bar{n}_{g,\nu}(t) = \binom{M}{g} \mathrm{e}^{(g^2-2Mg+g)t/2V} (\mathrm{e}^{t/V} - 1)^{g-1} g^{g-2}. \tag{27}$$

In the thermodynamic limit ($M, V \longrightarrow \infty$, their ratio is constant, $M/V = 1$) equation (27) gives

$$\bar{n}^s_{g,g-1}(t) = M \frac{g^{g-2}}{g!} t^{g-1} \mathrm{e}^{-gt}. \tag{28}$$

This is exactly the mass spectrum derived from the Smoluchowskii equation with the coagulation kernel $K(g, l) = gl$. This spectrum does not contain a giant particle at $t > t_c$ and does not conserve the total mass. The problem is then how to reconcile this fact with the statement that the giant component should appear?

The answer to this question is simple. We loose the giant tree in the thermodynamic limit, but it presents in the exact spectrum equation (27). However, the probability to find this tree is small as $C_{g,g-1} \big/ \sum_\nu C_{g,\nu}$. Nevertheless, it is possible to see this giant tree in the exact spectrum (27). The idea is simple. If the exact spectrum (28) contains a small hump at $g \propto M$ we can detect it against the background of the approximate spectrum (28) which is an exponentially dropping function at such large g. We thus consider the ratio $\bar{n}_g \big/ \bar{n}^s_g$ and exponentiate it,

$$\frac{\bar{n}_g}{\bar{n}^s_g} = \exp[M\Omega(\mu, t)], \tag{29}$$

where $\mu = g/M$ and

$$\Omega(\mu, t) = -(1 - \mu)\ln(1 - \mu) - \mu + \tfrac{1}{2}\mu^2 t. \tag{30}$$

Differentiating Ω over μ and putting the result equal to zero we find the position μ_c of the maximum,

$$t = \frac{1}{\mu_c} \ln \frac{1}{1 - \mu_c}. \tag{31}$$

This is exactly the location of the giant component of the random graph (Erdös and Renyi 1960) as well as the gel particle in the respective coagulating system (Lushnikov 1978).

Similar results can be derived for the bipartite graph. In this case two kinds of vertices are randomly connected with edges. The edges connecting the vertices of the same sort are forbidden. In this case the number of ways for two linked components containing m_1, n_1 and m_2, n_2 vertices of the first and the second kinds to coalesce is proportional to $m_1 n_2 + m_2 n_1$. The number of vacant places for edges in m, n-component is $mn - \nu$. The minimal value of ν is $\nu = m + n - 1$.

Using the results of Lushnikov (2005c) it is possible to write the expressions for the average numbers of $m, n; \nu$-components,

$$\bar{n}_{m,n;\nu}(t) = \binom{M_g}{m}\binom{M_r}{n} \mathrm{e}^{mnt/V - m\mathcal{M}_r t - n\mathcal{M}_g t}(\mathrm{e}^{t/V} - 1)^{\nu} C_{m,n;\nu}, \tag{32}$$

and for the number of m, n components,

$$\bar{n}_{m,n}(t) = \binom{M_g}{m}\binom{M_r}{n} \mathrm{e}^{mnt/V - m\mathcal{M}_r t - n\mathcal{M}_g t}(\mathrm{e}^{t/V} - 1)^{m+n-1} F_{m-1,n-1}(\mathrm{e}^{t/V}). \tag{33}$$

Here $\bar{n}_{m,n}(t) = \sum_{\nu} \bar{n}_{m,n;\nu}(t)$, M_g, M_r are the numbers of vertices of the first and the second kinds in the graph, $\mathcal{M}_{g,r} = M_{g,r}/V$, $C_{m,n;\nu}$ is the number of $m, n; \nu$-linked components, and the polynomials $F_{m,n}(x)$ introduced in Lushnikov (2005c) play the same role as Mallows–Riordan polynomials $F_g(x)$ in the expression for the number of linked components of the order of g in equation (26). In particular, the polynomials $w^{m+n-1} F_{m-1,n-1}(1 + w)$ generate $C_{m,n;\nu}$,

$$w^{m+n-1} F_{m-1,n-1}(1 + w) = \sum_{\nu=m+n-1}^{mn} C_{m,n;\nu} w^{\nu}. \tag{34}$$

The main results of this letter can be summarized as follows. The time evolution of the spectrum of linked components in random graphs has been shown to be described by the same formulae as the mass spectrum in coagulating systems with the kernel proportional to the product of the masses of coalescing particles. The partial distributions of (g, ν)- (or $(m, n; \nu)$-) components are expressed in terms of the numbers of linked components with a given number of edges. Equations (25), (26), (32), (33) are the central results of this letter.

References

Albert R and Barabási A-L 2002 Statistical mechanics of complex networks *Rev. Mod. Phys.* **74** 47

Bollobás B 1985 *Random Graphs* (London: Academic)

Egorychev G P 1977 *Integral Representation and Counting of Combinatorial Sums* (Siberian branch, Novosibirsk: Nauka) (in Russian)

Erdös P and Renyi A 1960 On the evolution of random graphs *Publ. Math. Inst. Hung. Acad. Sci.* **5** 17

Han D 2003 The asymptotic distributions of the size of the largest length of a reversible Markov process of polymerization *J. Phys. A: Math. Gen.* **36** 7485

Janson S, Luczak T and Rucinski A 2000 *Random Graphs* (New York: Wiley)

Knuth D 1998 Linear probing and graphs *Algorithmica* **22** 561

Leyvraz F 2003 Scaling theory and exactly solved models in the kinetics of irreversible aggregation *Phys. Rep.* **383** 95

Lushnikov A A 1978 Coagulation in finite systems *J. Colloid Interface Sci.* **65** 276

Lushnikov A A 2004 From sol to gel exactly *Phys. Rev. Lett.* **93** 198303

Lushnikov A A 2005a Exact kinetics of the sol–gel transition *Phys. Rev.* E **71** 046129

Lushnikov A A 2005b Exact particle mass spectrum in a gelling system *J. Phys. A: Math. Gen.* **38** L35

Lushnikov A A 2005c Sol–gel transition in a coagulating mixture *J. Phys. A: Math. Gen.* **38** L383

Spouge J L 1985 Polymers and random graphs: asymptotic equivalence to branching processes *J. Stat. Phys.* **38** 573