24

Rostock. Math. Kolloq. 41, 47 - 58 (1990) 05C30
05C80

Andrzej Ruciński

Maximal graphs with bounded maximum degree: structure, asymptotic enumeration, randomness

0. Introduction

A graph G is called a d-graph if it is not a spanning subgraph of any graph of maximum degree d. A vertex of degree less than d is called unsaturated. In this paper we examine the structure of d-graphs and enumerate them asymptotically with respect to the number of unsaturated vertices. We also discuss the question of Erdös about the limit distribution of the number of unsaturated vertices at the end of the graph process in which edges are added one by one, equiprobably, and so that the maximum degree does not exceed d.

1. Random graph models

By a random graph (a random graph process) we mean a probabilistic space whose elements are graphs (sequences of graphs) on the vertex set $V = \{1,2,\dots,n\}$. Most natural are equiprobable models: the Erdös-Rényi model $K_{n,N}$ consisting of all graphs with N edges and Bollobás' process Π_n formed by all sequences $(G_0,G_1,\dots,G_{\binom{n}{2}})$, where G_i has i edges and is contained in G_{i+1}, $i = 0,\dots,\binom{n}{2}-1$. Both can be thought of as results of random experiments: $K_{n,N}$ - drawing N out of $\binom{n}{2}$ pairs of vertices, Π_n - adding edges one by one to the empty graph. (See [4] for an extensive account on random graphs.)

In many applications in chemistry and physics these models are not satisfactory due to the lack of degree restrictions quite natural in the world of molecules. In the simplest case the condition that no vertex has degree larger than d is required. Then the appropriate model could be the equiprobable space of all graphs on V with N edges and maximum degree at most d - an analogue of $K_{n,N}$. However, in this paper we do not deal with this model. More appealing is an analogue of Π_n which we denote by $\Pi_{n,d}$ and define as follows.

Let $[V]^2$ be the set of 2-element subsets of V, $\Delta(G)$ and $E(G)$ stand for the maximum degree and edge set of a graph G, respectively, and $G \cup x = (V, E(G) \cup \{x\})$, where $x \in [V]^2 - E(G)$. The space $\Pi_{n,d}$ consists of all sequences $(G_o, G_1, \ldots, G_m)$ satisfying

(i) $|E(G_i)| = i$, $i = 0, \ldots, m$,

(ii) $\emptyset = E(G_o) \subset E(G_1) \subset \ldots \subset E(G_m)$,

(iii) $\Delta(G_i) \leqslant d$, $i = 0, \ldots, m$,

(iv) $A(G_m) = \emptyset$,

where $A(G) = \{ x \in [V]^2 - E(G) : \Delta(G \cup x) \leqslant d \}$.
The probability assigned to $(G_o, \ldots, G_m)$ equals to $\prod_{i=o}^{m-1} a_i^{-1}$, where $a_i = |A(G_i)|$.
More intuitively, one keeps adding edges one by one, each time choosing equiprobably a pair of vertices which are not yet joined and both have degree less than d.
Note that the length of the process varies, since it terminates when for the first time $A(G_i) = \emptyset$. At first glance we see two differences between Π_n and $\Pi_{n,d}$. The space $\Pi_{n,d}$ is not equiprobable and the last term, G_m, is not unique. Indeed, G_m ranges over all graphs G with $\Delta(G) = d$ and $\Delta(G \cup x) > d$ for all $x \in [V]^2 - E(G)$. Such graphs are called <u>d-graphs</u> here. This notion was introduced (under different name) by Kennedy and Quintas [5]. Some probabilistic spaces of d-graphs where also investigated by Balińska and Quintas [1].

In Section 2 we examine some structural properties of d-graphs, whereas Section 3 contains an asymptotic enumeration of n-vertex d-graphs (as $n \to \infty$) with respect to the number of vertices of degree less than d. Such vertices are called <u>unsaturated</u> here. It is an open problem posed by Erdös (a personal communication) to determine the limit distribution (as $n \to \infty$) of the number of unsaturated vertices at the end of the random process $\Pi_{n,d}$.

In Section 4 we present a simple procedure to find the distribution for $d = 2$ and any fixed n. For larger d, a computer simulation is all one is able to do at the moment. A simple procedure of generating a random process $\Pi_{n,d}$ goes as follows.

Procedure 1:

Set $E = \emptyset$, $c_1 = \ldots = c_n = n-1$, $d_1 = \ldots = d_n = 0$.
(*) Pick R = a random integer from $\{1,2,\ldots,c_1+\ldots+c_n\}$.
Set i for the smallest integer satisfying

$$R \leqslant c_1 + \ldots + c_i.$$

If $i = 1$ set $j = 0$ otherwise set j for the largest integer satisfying

$$R > c_1 + \ldots + c_j.$$

If $j = 0$ set $r = R$ otherwise set $r = R - (c_1 + \ldots + c_j)$.
Set $l = 0$.
For $k = 1,\ldots,n$, $k \neq i$ if $\{i,k\} \notin E$ and $d_k < d$ then $l = l+1$ until $l = r$.
Set $E = E \cup \{i,k\}$, $d_i = d_i+1$, $d_k = d_k+1$.
For $t = i,k$ if $d_t < d$ then $c_t = c_t-1$ otherwise set $c_t = 0$ and for $s = 1,\ldots,n$, $s \neq i,k$ if $\{t,s\} \notin E$ then $c_s = c_s-1$.
If all $c_i = 0$ then stop, otherwise go to (*).

Comment: At the end the set E is the set of edges of G_m. To gain the whole process one should keep the track of the order in which edges are included to E. Vertices i with $d_i < d$ are unsaturated.

Such simulation has been performed by K. Balińska and the data will be published in [2].

2. The structure of d-graphs

Let us denote by $U = U_G$ the set of unsaturated vertices of a d-graph G. Clearly, U induces a complete subgraph of G and therefore $u = |U| \leqslant d$. The unsaturated vertices have degrees between $u-1$ and $d-1$. Thus the number of edges in an n-vertex d-graph with $|U| = u$ varies from $\binom{u}{2} + \frac{1}{2}d(n-u)$ to $\frac{1}{2}(nd-u)$. It is not quite obvious that all theoretically possible cases are realisable. A trivial necessary condition is that $d \leqslant n-u-1$. It is satisfied in the whole range $1 \leqslant u \leqslant d$ only if $n \geqslant 2d+1$.
A more refined necessary condition for the existence of an n-vertex d-graph with $|U| = u$ and $2|E(G)| = L$, relevant when $n \leqslant 2d$, is

$$d-k \leqslant n-u-1, \text{ where } k = \lfloor (L - (n-u)d - u(u-1))/(n-u) \rfloor. \quad (1)$$

As it happens (1) is sufficient as well.

Proposition 1: For all $d \geq 2$, $n \geq d+1$, $1 \leq u \leq d$, and $u(u-1) + (n-u)d \leq L \leq nd-u$, L even, there exists an n-vertex d-graph G with $|U_G| = u$ and $2|E(G)| = L$ if and only if condition (1) is satisfied.

Proof: The difference $L - (n-u)d - u(u-1)$ is equal to the number of edges going from U to $V-U$. There must be a vertex in $V-U$ incident to at most k of them. Hence the necessity follows. Let x_o, y_o be the unique integer solution to the system of equations

$$\begin{cases} x + y = n-u, \\ kx + (k+1)y = L - (n-u)d - u(u-1). \end{cases}$$

It can be easily shown that there exists a graph H on $n-u$ vertices with x_o blue vertices of degree $d-k$ and y_o red vertices of degree $d-k-1$. Let K be a complete graph on u vertices disjoint from H. It is possible to join each blue vertex of H to k vertices of K and each red vertex of H to $k+1$ vertices of K and not produce a vertex in K of degree larger than $d-1$. ∎

As a consequence of the above result the doubled number of edges in an n-vertex d-graph may be as small as the smallest even integer not smaller than $nd - \frac{1}{4}(d^2+2d)$. The minimum is achieved when $u = \lfloor \frac{1}{2}(d+1) \rfloor$.

To avoid the parity problem we define a d-regular gaph as one with at most one vertex of degree $d-1$ and at least $n-1$ vertices of degree d. As we already know d-graphs are almost d-regular, especially when $n = |V|$ is large compared to d. In order to increase the number of edges in a non-d-regular d-graph one has to remove an edge and then add two new ones if possible. An edge of a d-graph which has the above property is called normal, the name is justified by the fact that every edge with neither endpoint joint to an unsaturated vertex is such and typically there are many such edges in a d-graph. Does every d-graph has a normal edge? It is not immediately seen that the answer is yes.

Proposition 2: Every d-graph is either d-regular or it contains a normal edge.

Proof: The assertion is trivial for $u = |U| = 1$. Assume, therefore, that $u \geq 2$ and set f for the number of edges with exactly one endpoint in U. Clearly $f \leq u(d-u)$.
Suppose, to the contrary, that there is no normal edge. We claim the existence of a pair of vertices $x \in U$, $y \notin U$ such that $N(y) \cap U = U-\{x\}$, where $N(y)$ is the set of neighbours of y. It follows from two facts:

(i) not all vertices of degree d are joined to all ones in U,
(ii) each vertex of degree d has at least $u-1$ unsaturated neighbours.

To prove (ii) suppose that for some $z \notin U$ $|N(z) \cap U| = 1 \leq u-2$. Then all its neighbours of degree d must be joined to all vertices in U and therefore $f \geq 1 + (d-1)u > u(d-u)$, a contradiction.
Each saturated (= of degree k) neighbour of y is joined either to all vertices in U or to all except x. Let l of them be of the second kind. If $l = 0$ then

$$f \geq (u-1) + u(d-u+1) > u(d-u).$$

If $l > 0$ then the degree of x is $d-1$ (otherwise there would be a normal edge from y to any of its neighbours not joined to x) and so x is joined to

$$(d-1) - ((u-1) + (d-u+l-1)) = l-1$$

saturated vertices not in $N(y) \cup \{y\}$. But each of these vertices is, in turn, joined to at least $u-1$ vertices in U. Altogether, we get

$$f \geq (u-1) + u(d-u+l-1) + l(u-1) + (l-1)(u-1) > u(d-u),$$

again a contradiction. ■

To make a d-graph d-regular one has to repeat the above operation $\lfloor \frac{1}{2}nd \rfloor - e(G)$ times, where $e(G) = |E(G)|$. This means that

$$f(G) = \min \{|E(G) \Delta E(F)| : \text{d-regular } F \text{ on vertex set } V\}$$
$$\leq 3\,(\lfloor \tfrac{1}{2}nd \rfloor - e(G)).$$

In fact, the equality holds and so the repetitive replacement of a normal edge by two new edges is the fastest way from a d-graph to a d-regular graph.

Proposition 3: For every d-graph on the vertex set V it holds

$$f(G) = 3\,(\lfloor \tfrac{1}{2}nd \rfloor - e(G)).$$

Proof: The sum of degrees of unsaturated vertices bears the whole deficit

$$\varepsilon = 2\lfloor \tfrac{1}{2}nd \rfloor - 2e(G).$$

Let F be a d-regular graph which minimizes $f(G)$. There has to be at least ε edges between U and $V-U$ in $E(F)-E(G)$. This, however, forces us to remove at least $\frac{1}{2}\varepsilon$ edges joining saturated vertices of G. ■

3. Asymptotic enumeration of d-graphs

Let $S = S_G$ be the set of vertices of degree k with at least one unsaturated neighbour. In this section we asymptotically enumerate n-vertex d-graphs ($n \to \infty$) with respect to the size of U and S. We consider separately the cases of dn odd and even. In the former cases almost all d-graphs are d-regular (under the broader definition of Section 2). For dn even, it turns out that for almost all d-graphs $0 \leq |U| \leq 2$ and $|S| = |U|(d-2)$, regardless the value of d. In both cases S is typically an independent set. For convenience, we present our results in the probabilistic form, associating to each n-vertex d-graph the same probability.

Proposition 4: Let P be the uniform probability measure on the set of all d-graphs on the vertex set $V_n = \{ 1,\dots,n \}$. Then

a) $\lim\limits_{\substack{n\to\infty \\ nd \text{ odd}}} P(|U| = 1,\ |S| = d-1) = 1,$

b) $\lim\limits_{\substack{n\to\infty \\ nd \text{ even}}} P(|U| = u,\ |S| = u(d-2)) = \dfrac{(2-\binom{u}{2})(d-1)^u}{d+1}, \quad u = 0,1,2,$

c) $\lim\limits_{n\to\infty} P(S \text{ is an independent set}) = 1.$

Proof: Throughout the proof we use the notations:

$$a_n \sim b_n \quad \text{if} \quad \lim_{n\to\infty} a_n/b_n = 1,$$

$$a_n = o(b_n) \quad \text{if} \quad \lim_{n\to\infty} a_n/b_n = 0,$$

$$a_n = O(b_n) \quad \text{if} \quad a_n \leqslant cb_n \text{ for some } c>0 \text{ and } n \text{ large enough,}$$

$$a_n \asymp b_n \quad \text{if} \quad a_n = O(b_n) \text{ and } b_n = O(a_n).$$

Let us set $u = |U|$, $s = |S|$, and let f be the number of edges from U to S. Denote by $\mathcal{A}(u,s,f)$ the set of all d-graphs on the vertex set V_n and with parameters u, s and f as above. Note that $0 \leqslant u \leqslant d$, $0 \leqslant s \leqslant u(d-u)$, $s \leqslant f \leqslant \min(sd, u(d-u))$. Bender and Canfield [3] proved that the number of graphs on V_n with the degrees $d_1,\dots,d_n$, $\max d_i \leqslant D$, is asymptotically equal to

$$\exp(-\alpha^2 - \alpha)\,(2q)!/(q!2^q \prod_{i=1}^{n} d_i!)$$

as $n \to \infty$, where $2q = \sum_{i=1}^{n} d_i$ and $\alpha = \frac{1}{4q}\sum_{i=1}^{n} d_i(d_i-1)$.

(Notice that $0 \leqslant \alpha \leqslant (D-1)/2$.)
Applying that result we get

$$|\mathcal{A}(u,s,f)| \asymp n^{u+s} \frac{(dn-du-f)!}{(\frac{1}{2}(dn-du-f)!2^{dn/2}(d!)^n}$$

and, for dn odd,

$$|\mathcal{A}(u,s,f)| \asymp n^{u+s-\frac{1}{2}(du+f+1)}|\mathcal{A}(1,d-1,d-1)|$$

$$= o(|\mathcal{A}(1,d-1,d-1)|)$$

unless $u = 1$ and $s = f = d-1$. This is because $s \leqslant f \leqslant u(d-2)$ for $u \geqslant 2$ and $s = f$ for $u = 1$.
Hence part a) is proved.
Now assumme that dn is even and denote $\mathcal{A}_0 = \mathcal{A}(0,0,0)$, $\mathcal{A}_1 = \mathcal{A}(1,d-2,d-2)$, $\mathcal{A}_2 = \mathcal{A}(2,2d-4,2d-4)$.

Then

$$|\mathcal{A}(u,s,f)|\cdot|\mathcal{A}_0|^{-1} \asymp n^{u+s-\frac{1}{2}(du+f)} = o(1)$$

unless $u \leqslant 2$ and $s = t = u(d-2)$. The reason is that for $u \geqslant 3$ we have $t \leqslant u(d-3)$ and therefore

$$2s \leqslant 2f \leqslant f + u(d-3) < f + u(d-2).$$

Moreover, if $u = 2$ then either of the inequalities $s < f$ and $f < u(d-2)$ implies that $2s < f + u(d-2)$.
Careful calculations show that

$$|\mathcal{A}_1|/|\mathcal{A}_o| \sim d-1 \quad \text{and} \quad |\mathcal{A}_2|/|\mathcal{A}_o| \sim (d-1)^2/2$$

completing the proof of b).
We are left with the proof of c). Let $\mathcal{A}'(u,s,f)$ be the set of all d-graphs belonging to $\mathcal{A}(u,s,f)$ and such that S is an independent set. Denote further by $\mathcal{H}$ the family of all graphs with the vertex set $\{1,2,\dots,s\}$, with at least one edge and maximum degree at most d-1. Given $H \in \mathcal{H}$, let R_H be the number of graphs G on vertices $\{1,\dots,n-u\}$ with $d_G(i) = d$ for $i = s+1,\dots,n-u$ and $d_G(i) = d-1-d_H(i)$ for $i = 1,\dots,s$, where $d_F(v)$ stands for the degree of vertex v in graph F. With the above notation

$$|\mathcal{A}'(u,s,s)| = O(n^{u+s} \sum_{H\in\mathcal{H}} R_H) = O(\tfrac{1}{n}|\mathcal{A}(u,s,s)|). \quad \blacksquare$$

A weaker version of the above result was proved for $d = 3$ in [1] using a recursion formula of Wormald.

4. 2-processes

Let us recall the question of Erdös:

For a random d-process $(G_o,G_1,\dots,G_m)$, what is the limit distribution of the number of unsaturated vertices in G_m as $n \to \infty$?

In this section we investigate the case $d = 2$. A connected component of a graph with the maximum degree at most 2 must be either a cycle or a path. We associate to each such graph a triple (a,b,c), called the type of a graph, where a, b, c are the number of isolated vertices, isolated edges, and components being paths of length at least 2, respectively.
Let $(G_o,\dots,G_m)$ be a 2-process and assume that G_i is of type (a,b,c). Then $a+b+c = n-i$ and G_{i+1} may be one of the following types:

(a-2, b+1, c), (a-1, b-1, c+1), (a, b-2, c+1), (a-1, b, c), (a, b-1, c), (a, b, c-1).

We write $(a,b,c) \to (a',b',c')$ if (a',b',c') is any of the six triples above. The transition probabilities multiplied by $\binom{a+2b+2c}{2} - b$ are $\binom{a}{2}$, $2ab$, $4\binom{b}{2}$, $2ac$, $4bc$, and $c + 4\binom{c}{2}$, respectively. Thus we have just defined a Markov process whose states are types of graphs and not graphs. Let us denote by $P(a,b,c)$ the probability that in the i-th step, $i = n-(a+b+c)$, the process is in state (a,b,c), equivalently that G_i is of the type (a,b,c).

In particular, $P(0,0,0)$, $P(1,0,0)$, and $P(0,1,0)$ are the probabilities we are interested in, i.e. they are equal to $P(|U_{G_m}| = k)$, $k = 0,1,2$.

There is a simple procedure of computing all $P(a,b,c)$ whose complexity is $O(n^3)$.

Procedure 2:

Set $P(n,0,0) = 1$. For $i = 1,\dots,n$ generate all triples (a',b',c'), $a'+b'+c' = n-i$ (with some further restrictions) and for each triple (a,b,c), $a+b+c = n-i+1$ check if $(a,b,c) \to (a',b',c')$.

If this is the case multiply $P(a,b,c)$ by the transition probability and add the outcome to the current value of $P(a',b',c')$.

Let us demonstrate how the procedure works in the case $n = 5$ (see Fig. 1).

A sample of the data obtained by the author using his ATARI 130 XE is given in Table 1 below. (The cases of $n = 30, 40, 48$ were supplied by K. Balinska.) The numbers are rounded to the fourth decimal position.

n	P(0,0,0)	P(1,0,0)
4	.7333	.2667
5	.6296	.2037
10	.7474	.1683
15	.7724	.1586
20	.7875	.1519
25	.7980	.1470
30	.8060	.1432
40	.8173	.1375
48	.8238	.1341

Table 1

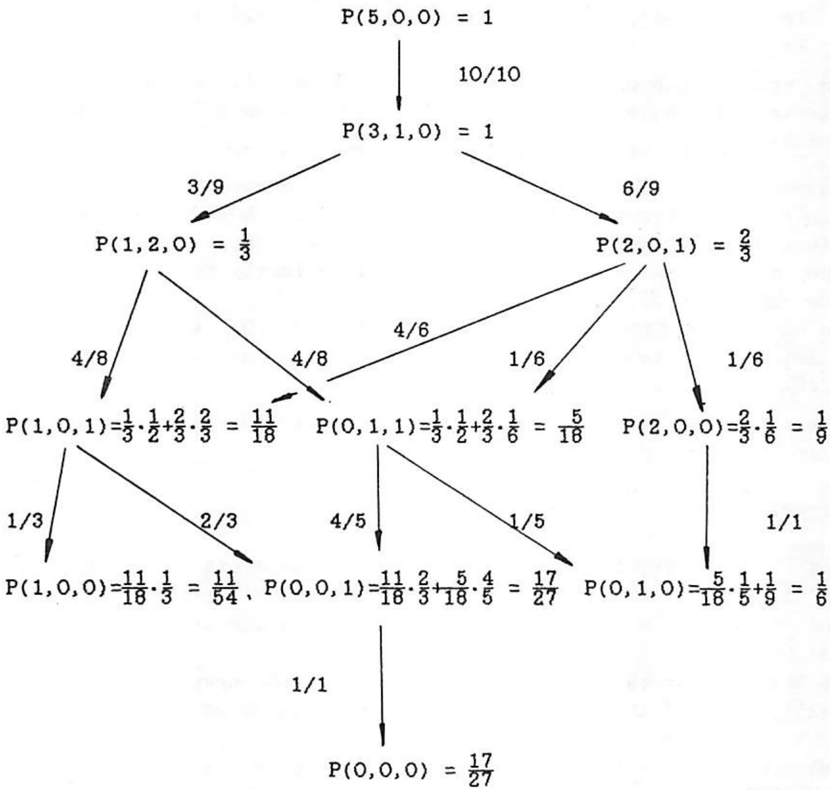
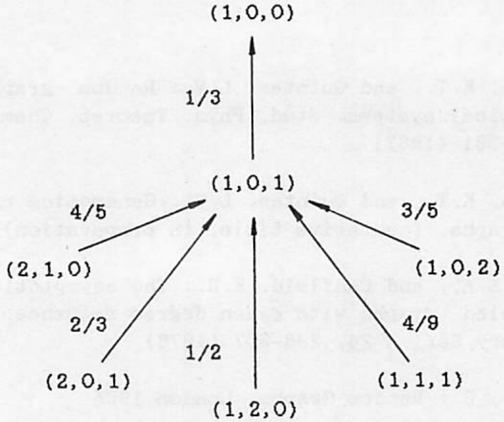


Fig. 1

But what is the limit distribution remains an open question By essentially the same approach but applied backward we get for all n

$$P(1,0,0) \leqslant \frac{4}{15} \quad \text{and} \quad P(0,1,0) \leqslant \frac{1}{5}.$$

The self-explanatory calculations leading to the first inequality together with the corresponding diagram are presented in Fig. 2. The second inequality can be derived similarly.



$$P(1,0,0) = \frac{1}{3} P(1,0,1) = \frac{1}{3} \left(\frac{4}{5} P(2,1,0) + \frac{2}{3} P(2,0,1) + \frac{1}{2} P(1,2,0) + \frac{4}{9} P(1,1,1) + \frac{3}{5} P(1,0,2) \right) \leq \frac{1}{3} \cdot \frac{4}{5}.$$

Fig. 2

If one allows parallel edges, the number of unsaturated vertices in G_m has two-point distribution (0 or 1). In such a case the process can be identified with a 2-dimensional random walk along a special lattice. A simple algorithm with complexity n^2 calculates the distribution. For instance, the probability of no unsaturated vertex is 2/3 for n = 3, 7/9 for n = 4, 116/150 for n = 5, and then it increases but rather slowly to reach, approximately, .835 for n = 20, .8597 for n = 50, .8973 for n = 500, .9087 for n = 1500 (the last took 30 hours on ATARI 130 XE).

Acknowledgments

This paper was born in New York in 1985 during inspiring discussions with J.W. Kennedy and L.V. Quintas. A preliminary version was presented to the participants of the weekly Seminar on Random Graphs in Poznań in January 1987. I appreciate the comments they made.
Finally, M. Truszczyński was of great help in proving Proposition 2.

References

1. Balińska, K.T., and Quintas, L.V.: Random graph models for physical systems. Stud. Phys. Theoret. Chem. 51, 349-361 (1987)

2. Balińska, K.T., and Quintas, L.V.: Generating random f-graphs. (tentative title, in preparation)

3. Bender, E.A., and Canfield, E.R.: The asymptotic number of labeled graphs with given degree sequence. J. Combin. Theory Ser. A 24, 296-307 (1978)

4. Bollobás, B.: Random Graphs. London 1985

5. Kennedy, J.W., and Quintas, L.V.: Probability models for random f-graphs. Third Internation. Conf. on Comb. Math., New York, June 10-14, 1985. Ann. New York Acad. Sci. (to appear)

received: September 17, 1988

Author:

Dr. Andrzej Ruciński
Department of Discrete Mathematics
Adam Mickiewicz University
Poznań
Poland