DIGRAPHS WHOSE NODES ARE MULTIGRAPHS HAVING EXACTLY TWO DEGREES f AND 2.

LOUIS V. QUINTAS AND JERZY SZYMAŃSKI

ABSTRACT. An (f,2)-graph is a multigraph G such that each vertex of G has degree either f or 2. Let S(n,f) denote the simple graph whose vertex set is the set of unlabeled (f,2)-graphs of order no greater than n and such that $\{G,H\}$ is an edge in S(n,f) if and only if H can be obtained from G by either an insertion or a suppression of a vertex of degree 2. We also consider digraphs whose nodes are labeled or unlabeled (f,2)-multigraphs and with arcs (G,H) defined as for $\{G,H\}$.

We study the structure of these graphs and digraphs. In particular, the diameter of a given component is determined. We conclude by defining a random process on these digraphs and derive some properties. Chemistry applications are suggested.

1. Introduction

By an (f,2)-graph we mean a multigraph G (with loops and multiple edges allowed) such that each vertex of G has degree either f or 2. By the insertion of a vertex of degree 2 in a graph we mean the replacement of an edge with a path of length 2 and the suppression of a vertex of degree 2 we mean replacing, with an edge, a pair of adjacent edges that meet at a vertex of degree 2. These operations will simply be called insertions and suppressions.

Let S(n, f) denote the simple graph whose vertex set is the set of unlabeled (f, 2)-graphs of order no greater than n and such that $\{G, H\}$ is an edge in S(n, f) if and only if H can be obtained from G by either an insertion or a suppression. We are interested in the properties of S(n, f) and relations between (f, 2)-graphs, the latter being viewed as vertices of S(n, f).

Let G^* be the (f, 2)-graph obtained from an (f, 2)-graph G after all possible suppressions of its vertices of degree 2. Note that G^* may have vertices of degree 2. For example, the vertex of degree 2 in a 1-cycle (also called a loop) cannot be suppressed, since a 1-cycle does not contain a pair

Date: March 15, 2002.

Key words and phrases. Random graph, bounded degree graph, graph process.

of adjacent edges that meet at a vertex of degree 2. The multigraph G^* is called the *suppressed graph associated with* G. The suppressed graph associated with G is unique up to isomorphism. Two (f,2)-graphs G and H are called *equivalent*, denoted by $G \equiv H$, if their suppressed graphs are isomorphic. A graph G is called *irreducible* if it is isomorphic to its suppressed graph (that is $G \sim G^*$), otherwise G is *reducible*.

A connected (1,2)-graph G is either a path or a cycle. It follows that the suppressed graph of G is either an edge or a loop, respectively. A connected (2,2)-graph is a cycle and its suppressed graph is a loop. Thus, our interest is in (f,2)-graphs with $f \geq 3$.

If G and H are equivalent (f,2)-graphs, the distance d(G,H) between G and H is defined as the least number of insertions and/or suppressions in G such that G is transformed into a graph isomorphic to H. If G and H are not equivalent, $d(G,H)=\infty$. Notice that the distance is equivalent to the graph theoretical distance in the S(n,f) graph.

The process considered above may be of interest in chemistry studies if an appropriate molecular graph interpretation of the vertex graphs of S(n, f) is introduced. That is, the vertices of degree f can correspond to chemical species of the same type and all vertices of degree 2 can correspond to species of a second type. For example, when f = 4, these types can be thought of as quaternary and secondary carbons, respectively, or the vertices of degree 2 as oxygen atoms. The insertions/suppressions of vertices of degree 2 represent transformations of these molecular graphs into other molecular graphs in which their number of quaternary types remains constant and the secondary species vary. For a model of this kind to be applicable it may be necessary to consider certain of these graphs as not being physically feasible and appropriate adjustments made to handle such a situation. The distance between molecular graphs can be defined as their graph theoretic distance and viewed as a measure of similarity. For a given component of molecular graphs one may be interested in, as in coding theory, finding graphs that are at a maximum distance apart. Such graphs are maximally dissimilar. Other graph theoretic invariants of S(n, f) and the probabilistic/statistical properties of the process when defined as a random process provide topics for further study and possible application.

2. Properties of S(n, f)

Let $\mathcal{E}_n(G)$ be the subgraph of S(n,f) induced by the (f,2)-graphs H equivalent to G and having order v(H) no greater than n. If G is an (f,2)-graph and G^* is the suppressed graph associated with G, then G is said to be on level $l(G) = v(G) - v(G^*)$ in $\mathcal{E}_n(G)$.

The insertion outdegree $d^+_{\mathcal{E}_n(G)}(G)$ of an (f,2)-graph G is the number of isomorphically distinct (f, 2)-graphs of order not greater then n that can be obtained from G by the insertion of one vertex of degree 2.

The suppression indegree $d_{\mathcal{E}_n(G)}^-(G)$ of an (f,2)-graph G is the number of isomorphically distinct (f, 2)-graphs that can be obtained from G by the suppression of one vertex of degree 2.

Given the (3,2)-graph G^* , an edge with a loop at each endvertex (i.e.,

 $G^* = \bigcirc$), the graph $\mathcal{E}_4(G^*)$ in S(4,3) is shown in Figure 2.1. Let G^* be an irreducible (f,2)-graph of order no greater than n. Then, $\mathcal{E}_n(G^*)$ can be generated from G^* , at level 0, to level $n-v(G^*)$ through successive levels using insertions.

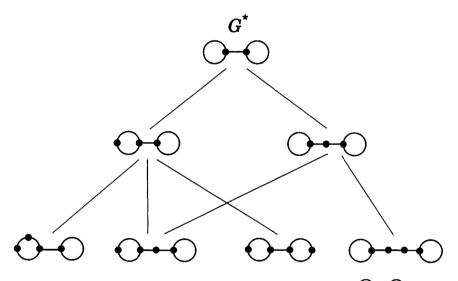


Figure 2.1. The graph $\mathcal{E}_4(G^*) \subset S(4,3)$, with $G^* = \bigcirc$

Observations (Properties of $\mathcal{E}_n(G)$)

- (1) $\mathcal{E}_n(G)$ is a connected bipartite simple graph.
- (2) $\mathcal{E}_n(G)$ can be generated by any graph H in (that is, any vertex H in) $\mathcal{E}_n(G)$ using insertions/suppressions.
- (3) $\mathcal{E}_n(\text{cycle})$ is a path of order n in S(n, f) with vertices corresponding to the k-cycles, k = 1 to n.
- (4) The number of vertices of degree f in each vertex of $\mathcal{E}_n(G)$ is the same.
- (5) The number of cycles in each vertex of $\mathcal{E}_n(G)$ is the same.
- (6) For any vertex H of $\mathcal{E}_n(G)$, the level of H is $v(H) v(H^*) = e(H) e(H)$ $e(H^*)$ (note, that $H^* = G^*$).

- (7) Let H be a vertex in $\mathcal{E}_n(G^*)$ at a maximum distance from G^* , then d(G,H) is defined to be the height of $\mathcal{E}_n(G^*)$ and is denoted $h(\mathcal{E}_n(G^*))$. It follows that $h(\mathcal{E}_n(G^*)) = v(H) v(G^*) = n v(G^*)$.
- (8) For any G^* , diam $\mathcal{E}_n(G^*) \geq h(\mathcal{E}_n(G^*))$. Moreover, there exist graphs H^* and K^* such that diam $\mathcal{E}_n(H^*) = h(\mathcal{E}_n(H^*))$ and diam $\mathcal{E}_n(K^*) > h(\mathcal{E}_n(K^*))$.

Proof.

- (1) Observe, that in $\mathcal{E}_n(G)$ edges can be only between vertices corresponding to two multigraphs H_1 and H_2 such that one of them has odd order and the second has even order.
- (2) If H and F are two vertices in $\mathcal{E}_n(G^*)$, then G^* can be obtained from H by successive suppressions and then F can be obtained from G^* by insertions.
- (3) Any insertion/suppression in a cycle generates a new cycle with the number of vertices increased/decreased by one.
- (4) Any vertex H in $\mathcal{E}_n(G^*)$ can be obtained from G^* by successive insertions. Each insertion does not change the number of vertices of degree f (but increases by one the number of vertices of degree 2).
- (5) Similary as in the previous observation, insertion does not change the number of cycles.
- (6) Notice, that multigraph H can be obtained from its suppressed graph H^* by the insertion of $v(H) v(H^*)$ vertices and that each insertion increases the number of vertices and the number of edges by one.
- (7) The height $h(\mathcal{E}_n(G))$ is the level of any vertex in $\mathcal{E}_n(G^*)$ corresponding to a multigraph with n vertices. Thus, from (6) it follows that $h(\mathcal{E}_n(G)) = n v(G^*)$.
- (8) Let G be a vertex in $\mathcal{E}_n(G)$ corresponding to a graph with n vertices. Then, from (7) $h(\mathcal{E}_n(G^*)) = d(G, G^*) \leq \dim \mathcal{E}_n(G^*)$. Let H^* be a loop and K^* the $G^* \subset \mathcal{E}(G)$ in Figure 2.1 with $n \geq 4$.

Theorem 2.1. Let G and H be equivalent (f, 2)-graphs $(G \equiv H)$, then

$$|v(G)-v(H)| \leq d(G,H) \leq v(G)+v(H)-2v(G^*)$$

and that bounds are sharp.

Proof. Let G^* be the suppressed graph of G (G^* is also the suppressed graph of H). Then H can be obtained from G by $v(G) - v(G^*)$ suppressions (resulting in the graph G^*) followed by $v(H) - v(G^*)$ insertions. Thus, H is obtained from G in $v(G) + v(H) - 2v(G^*)$ steps. So, $d(G, H) \leq v(G) + v(H) - 2v(G^*)$. Note that this bound is sharp. Consider, for example, the graph $G^* = \bigcirc$ in Figure 2.1. Let G be the graph obtained by

inserting vertices on the loops and H the graph obtained by only inserting vertices on the edge joining the loops. Then the equality holds.

Assume for simplicity, that v(G) > v(H). Then at least v(G) - v(H) vertices have to be supressed in G to obtain H. This bound is also sharp. For the above example, let both G and H be graphs obtained from G^* by insertions on the edge joining the loops.

Edges e_1 and e_2 in any suppressed graph G^* are equivalent if the insertion of a vertex of degree two in either e_1 or e_2 results in isomorphic multigraphs.

Theorem 2.2. For $f \geq 3$ the diameter of $\mathcal{E}_n(G)$ equals

$$\operatorname{diam} \mathcal{E}_n(G) = \begin{cases} 2\left(n - v(G^*)\right) & \text{if nonequivalent edges} \\ & \text{in G^* exist,} \\ 2\left(n - e(G^*) - \left\lfloor \frac{n}{e(G^*)} \right\rfloor + 1\right) & \text{otherwise.} \end{cases}$$

Proof. Let H_1 and H_2 be two vertices of $\mathcal{E}_n(G)$. Then from Theorem 2.1

$$d(H_1, H_2) \le v(H_1) + v(H_2) - 2v(G^*) \le 2n - 2v(G^*).$$

So,

$$\operatorname{diam} \mathcal{E}_n(G) \leq 2n - 2v(G^*).$$

If two nonequivalent edges in G^* exist (say, e_1 and e_2) then let H_1 (H_2) be a graph obtained from G^* by inserting $n - v(G^*)$ vertices in the edge e_1 (e_2). Then

$$\operatorname{diam} \mathcal{E}_n(G) \geq d(H_1, H_2) = 2n - 2v(G^*).$$

This complete the proof in the case of nonequivalent edges.

If all $e(G^*)$ edges of G^* are equivalent let H_1 be a graph obtained from G^* by inserting $n - e(G^*)$ vertices on one edge of G^* , while H_2 is a graph obtained by inserting $\left\lfloor \frac{n - e(G^*)}{e(G^*)} \right\rfloor$ or $\left\lceil \frac{n - e(G^*)}{e(G^*)} \right\rceil$ on each edge of G^* , such that the number of inserted vertices is $n - e(G^*)$. Then

$$d(H_1, H_2) = 2\left(n - \left(e(G^*) + \left\lfloor \frac{n - e(G^*)}{e(G^*)} \right\rfloor\right)\right)$$
$$= 2\left(n - e(G^*) - \left\lfloor \frac{n}{e(G^*)} \right\rfloor + 1\right).$$

So,

$$\operatorname{diam} \mathcal{E}_n(G) \geq 2\left(n - e(G^*) - \left|\frac{n}{e(G^*)}\right| + 1\right).$$

It is easy to see that this is the maximal distance between vertices of $\mathcal{E}_n(G^*)$ in this case.

Let $d_{\mathcal{E}_n(G^*)}^-(G)$ and $d_{\mathcal{E}_n(G^*)}^+(G)$ denote the suppression indegree and insertion outdegree of vertex G in the graph $\mathcal{E}_n(G^*)$, respectively.

Theorem 2.3. For any (f,2)-graph G which is a vertex of $\mathcal{E}_n(G)$

$$d_{\mathcal{E}_n(G^*)}^+(G) \le e(G^*)$$

and

$$d^-_{\mathcal{E}_n(G^*)}(G) \leq e(G^*)$$

where G^* is the suppressed graph of G.

Proof. Note, that the insertion of a new vertex on any edge in a path corresponding to a single edge of G^* results in isomorphic graphs. So, only insertions in paths corresponding to different edges of G^* can give different graphs. Thus, $d_{\mathcal{E}_n(G^*)}^+(G) \leq e(G^*)$.

Similarly, any vertex H in $\mathcal{E}_n(G)$ has a predecessor obtained by the suppression of a vertex of degree 2 in H. However, suppression of any vertex in a path corresponding to an edge in G^* will yield isomorphic graphs. Thus, $d_{\mathcal{E}_m(G^*)}(G) \leq e(G^*)$.

Let $\rho(G) = e(G)/v(G)$ denote the density of the graph G.

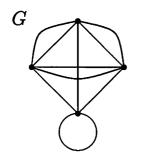
Theorem 2.4. For any (f,2)-graph G which is a vertex of $\mathcal{E}_n(G)$

$$\rho(G) = \frac{l(G) + e(G^*)}{l(G) + v(G^*)}$$

where l(G) is the level of G in $\mathcal{E}_n(G)$. Moreover, if $l(G) \to \infty$ then $\rho(G) \to 1$.

Proof. From observation (6) it follows that $v(G) = l(G) + v(G^*)$ and $e(G) = l(G) + e(G^*)$.

Notice, that any (f,2)-graph can be uniquely transformed into a (3,2)-graph by replacing all vertices of degree f by an f-cycle. Insertion and suppression operations can be done on the (3,2)-graph except for insertion in the edges of f-cycles corresponding to the initial vertices of degree f. Then the reverse process (that is, contractions of the f-cycles corresponding to the initial vertices of degree f) can be used to obtain an, equivalent to the original, (f,2)-graph. So, (3,2)-graphs are basic, in the sense that some properties of (f,2)-graphs $(f\geq 3)$ can be viewed as properties of (3,2)-graphs. An example of (5,2)-graph G and corresponding (3,2)-graph G is shown in Figure 2.2.



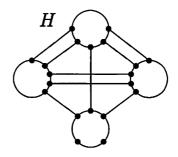


Figure 2.2. An example of (5,2)-graph G and its corresponding (3,2)-graph H.

3. RANDOM PROCESSES BASED ON S(n, f)

We shall now define a random process with the vertices of $\mathcal{E}_n(G)$ as the states of this process. Let G^* have order $v(G^*)$, size $e(G^*)$, k vertices of degree f and for technical reasons that follow, let the $e(G^*)$ edges of G^* be labeled.

If G is a vertex of $\mathcal{E}_n(G^*)$, there is a one-to-one correspondence between the labeled edges of G^* and the paths in G that start at a vertex of degree f, end at a vertex of degree f, and have no intermediate vertices of degree f. Let these paths be labeled with the same labels as the edges they correspond to in G^* . If G^* has an edge that is a loop, we shall consider this edge in correspondence with the closed path in G that starts and ends with the same vertex of degree f on which the original loop was based.

The transition digraph $D\mathcal{E}_n(G^*)$ for the random process on $\mathcal{E}_n(G^*)$ is defined as follows. An arc (G,H) is in $D\mathcal{E}_n(G^*)$ if and only if a path-labeled copy of H can be obtained from a path-labeled copy of G either by the insertion of a vertex of degree 2 into one of its labeled paths or by the suppression of a vertex of degree 2 in one of its x_G labeled paths that contain at least one vertex of degree 2 $(0 \le x_G \le e(G^*))$. Such a transition from G to H is called an admissible move.

For each G there are $e(G^*) + x_G$ possible admissible moves. Assuming each of these are selected with uniform probability $1/(e(G^*) + x_G)$ completes the definition of this random process on the vertices of $\mathcal{E}_n(G^*)$. Note that $x_{G^*} = 0$. If G has order n, no vertex of degree 2 can be inserted into G. Thus, an admissible move is a suppression that only involves paths with vertices of degree 2 and these are selected uniformly with probability $1/x_G$.

Since $D\mathcal{E}_n(G^*)$ is strongly connected, $D\mathcal{E}_n(G^*)$ is the transition digraph for an ergodic Markov chain (see [3]).

For any ergodic Markov chain two fundamental things to determine are:

- (a) the equilibrium vector, defined as the unique probability vector $\mathbf{w} = [w_G]$ such that $\mathbf{w}\mathbf{T} = \mathbf{w}$, where $\mathbf{T} = [p_{GH}]$ is the transition matrix for the chain, and
- (b) whether the chain is *time-reversible*, for which a necessary and sufficient condition is $w_G p_{GH} = w_H p_{HG}$, for all states G and H in the chain. For interpretations of the equilibrium vector \mathbf{w} and time-reversibility see [4].

Starting with G^* , we shall now define a random process with the path-labeled (f, 2)-graphs homeomorphic to G^* as the states of this process. The path-labeled (f, 2)-graphs are as defined in the definition of the random process $D\mathcal{E}_n(G^*)$, but remain labeled and distinct as path-labeled graphs not as graphs as is done in $D\mathcal{E}_n(G^*)$.

The transition digraph $d\mathcal{E}_n(G^*)$ for this random process is defined analogously to that for $D\mathcal{E}_n(G^*)$ as follows. An arc (G,H) is in $d\mathcal{E}_n(G^*)$ if and only if (a path-labeled copy of) H can be obtained from (a path-labeled copy of) G either by the insertion of a vertex of degree 2 into one of its labeled paths or by the suppression of a vertex of degree 2 in one of its x_G labeled paths that contain at least one vertex of degree 2. Here too we call such a transition from G to H an admissible move.

For each G there are $e(G^*) + x_G$ possible admissible moves. Assuming each of these moves is selected with uniform probability $1/(e(G^*) + x_G)$ completes the definition of the random process on the path-labeled (f, 2)-graphs homeomorphic to G^* . As in the Markov chain $D\mathcal{E}_n(G^*)$ we have $x_{G^*} = 0$ and if G has order n, no vertex of degree 2 can be inserted into G so that as before an admissible move can only be a suppression and only involves paths with vertices of degree 2 selected uniformly with probability $1/x_G$.

Since $d\mathcal{E}_n(G^*)$ is strongly connected, $d\mathcal{E}_n(G^*)$ is the transition digraph for an ergodic Markov chain which we denote $\mathcal{E}_n^d(G^*)$.

We note that the transition digraph $d\mathcal{E}_n(G^*)$ is equivalent to the simple graph with vertex set the same as the nodes of $d\mathcal{E}_n(G^*)$ and transitions from a vertex G to a neighboring vertex H are uniform with probability $p_{HG} = 1/\deg G$, where $\deg G = e(G^*) + x_G$. Thus, the analysis of the chain with transition digraph $d\mathcal{E}_n(G^*)$ can be pursued using the theory of random walks on graphs (see[4]). This observation yields the following theorem.

Theorem 3.1. The Markov chain $\mathcal{E}_n^d(G^*)$ is time-reversible and has equilibrium vector with components $w_G = \frac{\deg G}{2m}$, where m is the size of $d\mathcal{E}_n(G^*)$.

Proof. That $w_G = \frac{\deg G}{2m}$ is given in [4]. Since time-reversibility is equivalent to the condition $w_G p_{GH} = w_H p_{HG}$, it is sufficient to show that $w_G p_{GH}$ does not depend on G and H. We have $w_G p_{GH} = \frac{\deg G}{2m} \frac{1}{\deg G} = \frac{1}{2m}$.

The following, as noted in [4], is an immediate corollary of this theorem.

Corollary 3.1. If a random walk is at vertex G, then the expected number of steps before it returns to G is $\frac{1}{w_G} = \frac{2m}{\deg G}$.

Theorem 3.2. The order of $\mathcal{E}_n^d(G^*)$ is $\binom{n+e(G^*)-v(G^*)}{n-v(G^*)}$.

Proof. Let G^* be an f-regular multigraph of order $v = v(G^*)$ and size $e = e(G^*)$. Let the edges of G^* be labeled. Inserting t vertices of degree 2 into G^* yields a graph on level t. Noting that the t vertices can be considered as t indistinguishable objects and the labeled paths of G^* as distinguished boxes implies there are $\binom{e+t-1}{t}$ different path-labeled (f,2)-graphs on level t. Let t^* be the number of vertices of degree 2 that have to be inserted into G^* to obtain a graph of order n ($t^* = n - v(G^*)$). Thus the total number of such graphs up to and including those on level t^* is $\sum_{t=0}^{t^*} \binom{e+t-1}{t} = \binom{e+t^*}{t^*}$.

4. Some Examples

The graph $\mathcal{E}_4(G^*)$ with $G^* = K_2$ with a loop at each vertex is shown in Figure 2.1. The transition digraph $D\mathcal{E}_4(G^*)$ is shown in Figure 4.1.

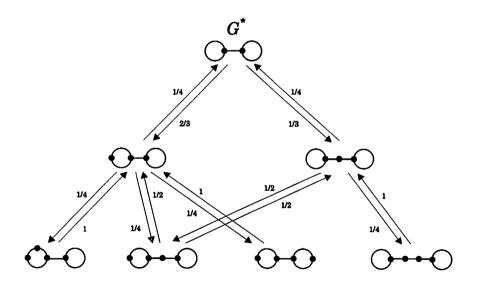


Figure 4.1. The digraph $D(\mathcal{E}_4(G^*))$ with $G^* = \bigcirc$

In Figure 4.2 we show the transition digraph $d\mathcal{E}_4(G^*)$.

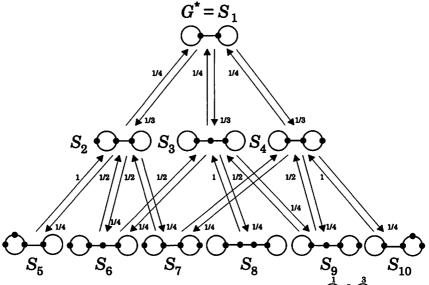


Figure 4.2. The digraph $d(\mathcal{E}_4(G^*))$ with $G^* = \bigcirc^2\bigcirc$

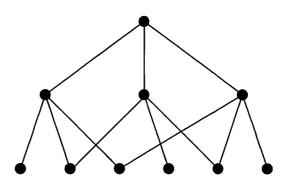


Figure 4.3. The simple graph associated with $d(\mathcal{E}_4(G^*))$ with $G^* = \bigcirc^1 \bigcirc^3$.

Analysis of these examples yields the following information. By Theorem 3.1 the equilibrium vector for the random process $\mathcal{E}_4^d(G^*)$ is

and $\mathcal{E}_4^d(G^*)$ is time-reversible.

Furthermore, we find that by combining components, that correspond to isomorphic graphs, in the equilibrium vector for $\mathcal{E}_4^d(G^*)$ and summing their values we can obtain the equilibrium vector for $D\mathcal{E}_4(G^*)$. The equilibrium vector for the random process $\mathcal{E}_4(G^*)$ is

$$\frac{1}{24}$$
[3 8 4 2 4 2 1]

and $D\mathcal{E}_4(G^*)$ is time-reversible. It is conjectured that a formalization of the above operations for the general case will yield a method for obtaining the equilibrium vector for $D\mathcal{E}_n(G^*)$ for any G^* . See [1] [2] for what was done for an analogous random process.

Problem: For $D\mathcal{E}_n(G^*)$ determine the equilibrium vector and whether time-reversibility holds.

Acknowledgments

LVQ thanks Pace University's Scholarly Research Committee, The Kenan Fund for Faculty Development, and the Department of Control, Robotics, and Computer Science of The Technical University of Poznań for partial support of the work.

REFERENCES

- [1] K. T. Balińska, M. L. Gargano and L. V. Quintas, Two models for random graphs with bounded degree, Croatia Chemica Acta, 74(2) (2001), 207-223.
- [2] K. T. Balińska and L. V. Quintas, Maximum outdegrees in a digraph whose nodes are graphs with bounded degree, Research Communications of the Conference held in the Memory of Paul Erdős (Eds. A. Sali, M. Simonovits, V. T. Sós), János Bolyai Mathematical Society, Budapest, (1999), 23-26.
- [3] P. G. Doyle and J. L. Snell, Random walks and Electric Networks, Carus Mathematical Monographs, No. 22, Mathematical Association of America (1984).
- [4] L. Lovász, Random walks on graphs, in: Combinatorics, Paul Erdős is Eighty (Volume 2), (Keszthely, Hungary, 1993), Bolyai Society Mathematical Studies (1996) 353-398.

LOUIS V. QUINTAS, MATHEMATICS DEPARTMENT, PACE UNIVERSITY, NEW YORK, NY 10038, U.S.A.

E-mail address: lquintas@pace.edu

JERZY SZYMAŃSKI, FACULTY OF MATHEMATICS AND COMPUTER SCIENCE, ADAM MICKIEWICZ UNIVERSITY, 61-614 POZNAŃ, POLAND

E-mail address: jesz@amu.edu.pl