Rev. Semin. Iberoam. Mat. **4** fasc. II (2013) 3–17

EVOLUTIONARY DYNAMICS ON GRAPHS

FERNANDO ALCALDE CUESTA, PABLO GONZÁLEZ SEQUEIROS,
AND ÁLVARO LOZANO ROJO

Abstract. In this paper, we summarize some ideas and results about evolutionary dynamics on graphs. This theory is illustrated with a concrete example, the so-called *star graph*, for which we calculate the average fixation probability.

1. Introduction and motivation

Population genetics studies the genetic composition of biological populations, and the changes in this composition that result from the action of four different process: *natural selection*, *random drift*, *mutation* and *migration*. The *modern evolutionary synthesis* combines Darwin's thesis on the natural selection and Mendel's theory of inheritance. According to this synthesis, the central object of study in evolutionary dynamics is the frequency distribution of the alternative forms (*allele*) that a hereditary unit (*gene*) can take in a population evolving under these four forces.

Many mathematical models have been proposed to understand the evolutionary biological processes. For example, the *Wright-Fisher model* (stated explicitly by S. Wright [8], but present in the work of and R. A. Fisher [4]) describes the change of gene frequency by random drift on a population of finite fixed size N. For simplicity, the involved organisms are assumed to be *haploids* (containing only one set of chromosomes) with only two possible alleles a and A for a given locus, although the Wright-Fisher model can be extended to multiple alleles in diploids organisms. Then there are only $N+1$ possible gene frequencies i/N for $0 \leq i \leq N$. Assume that in some population there are exactly i copies of the allele A (and therefore $N-i$ copies of a). If each of the N offspring contains a copy of a randomly chosen allele from the present generation, then the gene frequency in the next generation could assume any of the $N+1$ possible values, except when $i = 0$ or $i = N$.

The *Moran model* (introduced by P. A. P. Moran in [6]) shows a particular equilibrium between natural selection and random drift. This model has many variants, but we will consider the variant which is closest to the Wright-Fisher model. We have a haploid population of N individuals having only two possible

2010 *Mathematics Subject Classification.* 05C81, 60J20 92D15.

Key words and phrases. Evolutionary dynamics, Moran process, fixation probability, star graph.

Partially supported by the Ministry of Science and Innovation - Government of Spain (Grant MTM2010-15471).

alleles a and A for a given locus. In the Moran model, instead of all individuals dying simultaneously upon the birth of the next generation, at each unit of time, one individual is chosen at random for reproduction and its clonal offspring replaces another individual chosen at random to die. To model natural selection, it suffices to assume that the parent individuals with allele A have relative fitness $r > 1$, as compared to those with allele a whose fitness is 1. In this case, individuals with the advantageous allele A have a certain chance of *fixation* generating a lineage that takes over the whole population, whereas individuals with the disadvantageous allele a are likely to become extinct, although it will be never guaranteed.

As in the previous models, evolutionary dynamics has been usually studied for homogeneous populations. But it is a natural question to ask how non-homogeneous structures affects the dynamics. The study of evolutionary dynamics on directed graphs was initiated by E. Liberman, C. Hauert and M. A. Novak [5] (see also [7]). Now each vertex represents an individual in the population, and the offspring of each individual only replace direct successors, i.e. end-points of edges with origin in this vertex. The fitness of an individual represents again its reproductive rate which determines how often offspring takes over its neighbor vertices, although these vertices do not have to be replaced in a equiprobable way. In other words, the evolutionary process is given by the choice of stochastic matrix $W = (w_{ij})$ where w_{ij} denotes the probability that individual i places its offspring into vertex j. In fact, further generalizations of evolutionary graphs are considered in [5] assuming simply that the probability above is proportional to the product of a weight w_{ij} and the fitness of the individual i. In this case, the matrix W does not need to be stochastic, but non-negative.

In this context, several interesting and important results have been shown by Liberman, Hauert and Novak:

- Different graph structures support different dynamical behaviors amplifying or suppressing the reproductive advantage of *mutant* individuals (having the advantageous allele A) over to the *resident* individuals (having the disadvantageous allele a).
- An *'isothermal theorem'* which states that an evolutionary process on a graph is *equivalent to a Moran process* (in the sense that there is a well-defined fixation probability which coincides with the fixation probability for an homogeneous population) if and only if it is defined by a doubly stochastic matrix W. More generally, a non-negative matrix W defines an evolutionary process equivalent to a Moran process if and only if it is a *circulation*, i.e. each vertex i has the same entering and leaving weight $w_-(i) = \sum_{j=1}^N w_{ji} = \sum_{j=1}^N w_{ij} = w_+(i)$, which is equal to 1 in the stochastic case.

However, for evolutionary processes on graphs, the fixation probability depends usually on the starting position of the mutant. The effect of the initial placement on the mutant spread has been discussed by M. Broom, J. Rychtář and B. Stadler in the case of undirected graphs, see [2] and [3].

The aim of this paper is summarize some fundamental ideas and results on evolutionary dynamics on graphs. This theory will be illustrated with a concrete example, the so-called *star graph*, for which we calculate the (average) fixation probability outlined in [5] (see also [1]).

2. Moran process

The *Moran process* was introduced by Moran [6] to model random drift and natural selection for finite homogeneous populations. As indicated in the introduction, we consider a haploid population of N individuals having only two possible alleles a and A for a given locus. At the beginning, all individuals have the allele a, then one resident individual is chosen at random and replaced by a mutant having the neutral or advantageous allele A. At successive steps, one randomly chosen individual replicates with probability proportional to the relative fitness and its offspring replaces another individual randomly chosen to be eliminated. Since the future states of the process depend only on the present state, and not on the sequence of events that preceded it, the Moran process is defined by the Markov chain

$$X_n = \text{number of mutant individuals with the allele } A \text{ at the step } n$$

with state space $\mathcal{S} = \{0, \dots, N\}$. Moreover, this process is *stationary* because the probability to pass from i to j mutant individuals

$$P_{i,j} = \mathbb{P}[X_{n+1} = j | X_n = i] = \mathbb{P}[X_{n+1} = j | X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i]$$

does not depend on the time. But the number of mutant individuals can change at most by one at each time step and therefore a non-trivial transition exists only between state i and state $i-1$, i or $i+1$. Then, the *transition matrix* of the stochastic process is a tridiagonal matrix

$$P = \begin{pmatrix} P_{0,0} & P_{0,1} & \dots & P_{0,N} \\ P_{1,0} & P_{1,1} & \dots & P_{1,N} \\ \vdots & \vdots & \ddots & \vdots \\ P_{N,0} & P_{N,1} & \dots & P_{N,N} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ \delta_1^- & 1-\delta_1^- - \delta_1^+ & \delta_1^+ & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \delta_{N-1}^+ \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

where $P_{i,i-1} = \delta_i^-$, $P_{i,i+1} = \delta_i^+$ and $P_{i,i} = 1 - \delta_i^- - \delta_i^+$. The states $i = 0$ and $i = N$ are *absorbing* while the other states are *transient*.

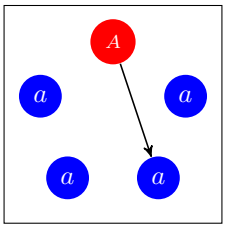

Figure 1. Moran process on a homogeneous population

For a general *birth-death process*, defined by a tridiagonal matrix P, the chance that any set of i mutant individuals spread taking over the whole population is denoted by

$$x_i = \mathbb{P}[\exists n : X_n = N | X_0 = i].$$

In particular, the probability of one mutant individual to reach fixation

$$x_1 = \mathbb{P}[\exists n : X_n = N | X_0 = 1]$$

is called *fixation probability* and also denoted by Φ_A. Now we consider the following system of linear equations

$$\begin{array}{l} x_0 = 0 \\ x_i = \delta_i^- x_{i-1} + (1 - \delta_i^- - \delta_i^+) x_i + \delta_i^+ x_{i+1} \\ x_N = 1 \end{array} \tag{2.1}$$

To find a solution $x = (x_0, x_1, \ldots, x_N)$ of the linear equation $Px = x$ with the conditions $x_0 = 0$ and $x_N = 1$, it is useful to define

$$y_i = x_i - x_{i-1}$$

verifying $\sum_{i=1}^{N} y_i = x_N - x_0 = 1$. Then, dividing each side of (2.1) by δ_i^+, we have:

$$y_{i+1} = \gamma_i y_i \tag{2.2}$$

where $\gamma_i = \delta_i^- / \delta_i^+$ is the *death-birth rate* (which is reciprocal to the reproductive advantage of any set of i mutant individuals). It follows:

$$y_i = x_1 \prod_{j=1}^{i-1} \gamma_j.$$

Therefore, the fixation probability is given by

$$x_1 = \frac{1}{1 + \sum_{i=1}^{N-1} \prod_{j=1}^{i} \gamma_j} \tag{2.3}$$

Random drift. If none of alleles a and A is reproductive advantageous, the random drift phenomenon can be modeled by the Moran process with relative fitness $r = 1$. In this case, the transition probabilities are given by

$$\begin{array}{rcl} P_{i,i-1} & = & \dfrac{N-i}{N} . \dfrac{i}{N-1} \\ P_{i,i+1} & = & \dfrac{i}{N} . \dfrac{N-i}{N-1} \\ P_{i,i} & = & \dfrac{i}{N} . \dfrac{i-1}{N-1} + \dfrac{N-i}{N} . \dfrac{N-i-1}{N-1} \end{array} \tag{2.4}$$

Since $\gamma_i = 1$, the fixation probability $\Phi_A = 1/N$. As for every birth-death process, if the population reaches one of the absorbing states, then it stays there forever. In the other states, the population of mutant individuals randomly evolves, but eventually these individuals will either become extinct or take over the whole population.

Natural selection. The effect of fitness on the evolutionary dynamics of a population is described by the Moran process provided mutant individuals with the allele A have relative fitness $r > 1$. Now, the transition probabilities are given by

$$\begin{aligned} P_{i,i-1} &= \frac{N-i}{ri+N-i}.\frac{i}{N-1} \\ P_{i,i+1} &= \frac{ri}{ri+N-i}.\frac{N-i}{N-1} \\ P_{i,i} &= \frac{ri}{ri+N-i}.\frac{i-1}{N-1} + \frac{N-i}{ri+N-i}.\frac{N-i-1}{N-1} \end{aligned} \tag{2.5}$$

Since the death-birth rate $\gamma_i = 1/r$, the fixation probability

$$\Phi_A = \frac{1}{1+\sum_{j=1}^{N-1} r^{-i}} = \frac{1-r^{-1}}{1-r^{-N}} \geq 1 - \frac{1}{r}.$$

Thus, an advantageous mutation with $r >$ reaches fixation with positive probability but this is not always guaranteed, because this probability is strictly less than 1.

3. Evolutionary processes on graphs

Evolutionary graph theory was introduced by Liberman, Hauert and Novak [5]. Like for homogenous populations, the first natural question is to determine the chance that the offspring of a mutant individual having an advantageous allele spreads through the graph reaching any vertex. But this chance depends obviously on the initial position of the individual (see [2] and [3]) and the global graph structure may significantly modify the equilibrium between random drift and natural selection observed in homogeneous populations (as proved in [5]; see also [2] and [3]).

Let $G = (V, E)$ be a directed graph, where V is the set of vertices and E is the set of edges. We assume G is finite, connected and simple graph (without loop or multiple edges). Thus, E identifies to a subset of $V \times V$ which does not meet the diagonal. Any graph structure on the vertex set $V = \{1, \ldots, N\}$ is completely determined by the adjacency matrix (a_{ij}) where $a_{ij} = \mathbb{1}_E(i, j)$ for each pair $(i, j) \in V \times V$. An *evolutionary process on* G is also given by a Markov chain, but each state is now described by a set of vertices $S \in \mathcal{S} = \mathcal{P}(V)$ inhabited by mutant individuals having an advantageous allele A. This reproductive advantage is measured by the fitness $r \geq 1$. The transition probabilities of this Markov chain are defined from a non-negative matrix $W = (w_{ij})$ whose entries are edge weights satisfying $w_{ij} = 0 \Leftrightarrow a_{ij} = 0$. So evolutionary process on G can be identified with the elements of the set $\mathcal{W}$ of such matrices. The transition probability between two states $S, S' \in \mathcal{S} = \mathcal{P}(V)$ (which is still time-independent)

is given by

$$P_{S,S'} = \begin{cases} \dfrac{r\sum_{i\in S} w_{ij}}{r\sum_{i\in S}\sum_{j\in V} w_{ij} + \sum_{i\in V\setminus S}\sum_{j\in V} w_{ij}} & \text{if } S'\setminus S = \{j\} \\ \dfrac{\sum_{i\in V\setminus S} w_{ij}}{r\sum_{i\in S}\sum_{j\in V} w_{ij} + \sum_{i\in V\setminus S}\sum_{j\in V} w_{ij}} & \text{if } S\setminus S' = \{j\} \\ \dfrac{r\sum_{i,j\in S} w_{ij} + \sum_{i,j\in V\setminus S} w_{ij}}{r\sum_{i\in S}\sum_{j\in V} w_{ij} + \sum_{i\in V\setminus S}\sum_{j\in V} w_{ij}} & \text{if } S = S' \\ 0 & \text{otherwise} \end{cases} \tag{3.6}$$

where $r\sum_{i\in S}\sum_{j\in V} w_{ij} + \sum_{i\in V\setminus S}\sum_{j\in V} w_{ij}$ is the sum of the reproductive weight of the mutant and resident individuals (equal to $r\#S + N - \#S = N + (r-1)\#S$ when the matrix W is stochastic). In other words, the process is defined by a $2^N \times 2^N$ stochastic matrix $P = (P_{S,S'})$. As for the Moran process, $S = \emptyset$ and $S = V$ are absorbing states, but there may exist other absorbing states, as well as recurrent states, so the probability that resident or mutant individuals become extinct can be strictly less than 1. Anyway, the fixation probability of any other set S inhabited by mutant individuals

$$\Phi_S = \mathbb{P}[\exists n : X_n = V | X_0 = S]$$

can be obtained as the solution of a linear equation, which is analogous to (2.1) for the classical Moran process. Assuming the absorbing states $S = \emptyset$ and $S = V$ are connected with other states and using that P is stochastic, it is possible to prove this equation has always a unique solution. Details will be reported elsewhere. By simplifying Φ_S terms, this equation reduces to the following equation:

$$\Phi_S = \frac{\sum_{i\in S}\sum_{j\in V\setminus S}\left(rw_{ij}\Phi_{S\cup\{j\}} + w_{ji}\Phi_{S\setminus\{i\}}\right)}{\sum_{i\in S}\sum_{j\in V\setminus S}\left(rw_{ij} + w_{ji}\right)}$$

(with $P_\emptyset = 0$ and $P_V = 1$) used in [1], [2] and [3]. In particular, for $S = \{i\}$, we have the equation:

$$\Phi_{\{i\}} = \frac{\sum_{j\neq i} rw_{ij}\Phi_{\{i,j\}}}{\sum_{j\neq i}\left(rw_{ij} + w_{ji}\right)}.$$

Contrary to the case of homogeneous populations, the fixation probability depends on the starting position of the mutant in the graph. This fact justifies the following definition:

Definition 3.1. For any matrix $W \in \mathcal{W}$, we define the *average fixation probability* on G as the average

$$\Phi_A = \frac{1}{N}\sum_{i=1}^{N}\Phi_{\{i\}}.$$

The definitions and results above can be illustrated by some examples:

Moran process. The classical Moran process coincides with the evolutionary process on the complete graph $G = K_N$ (where $V = \{1, \ldots, N\}$ and $E = V \times V \backslash \Delta$) defined by the stochastic matrix $W = (w_{ij})$ where $w_{ij} = \frac{1}{N-1}$ if $i \neq j$. Since G is *symmetric* (i.e. its automorphism group acts transitively on the vertex and edge sets) and W is preserved by the action of the automorphism group of G, the fixation probability $\Phi_{\{i\}} = \Phi_{\{j\}}$ for all $i \neq j$. Therefore the average probability fixation $\Phi_A = \Phi_{\{i\}}$ for all i.

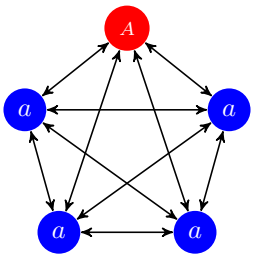

FIGURE 2. Evolutionary process on a complete graph

Directed line graph. This graph is described in the figure below, and the process is given by the adjacency matrix

$$W = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

If the starting position of the mutant individual coincide with the root, then this mutant generates with probability 1 a lineage that will take the whole population. But this will never be possible in other positions. In other words,

$$\Phi_{\{i\}} = \begin{cases} 1 & \text{if } i = 1 \\ 0 & \text{if } i \neq 1 \end{cases}$$

According to [5], such a graph structure is be said to be a *suppressor of selection* since the average fixation probability $\Phi_A = 1/N$ is the same that of the random drift for a homogeneous population independently of the mutant fitness.

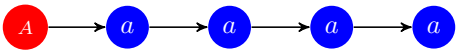

FIGURE 3. The line graph

Cycle graph. As before, this graph is described in the figure below, but now the process is given by the stochastic matrix

$$W = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \dots & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

Since the graph is still symmetric and W is preserved by the action of its automorphism group, even if the population is not yet homogeneous, the starting position of a mutant individual does not have any effect on the process. Thus, we can assume the starting state is $S = \{1\}$, which only may evolve to $S = \emptyset$, $S = \{1, 2\}$ or $S = \{N, 1\}$. Arguing by recurrence, we see that any accessible state is a connected subset, and the non-trivial transition probabilities depend only on its cardinal number i. In more precise way, these probabilities are given by

$$\begin{aligned} P_{i,i-1} &= \frac{1}{2}\frac{1}{ri+N-i} + \frac{1}{2}\frac{1}{ri+N-i} &&= \frac{1}{ri+N-i} \\ P_{i,i+1} &= \frac{1}{2}\frac{r}{ri+N-i} + \frac{1}{2}\frac{r}{ri+N-i} &&= \frac{r}{ri+N-i} \\ P_{i,i} &= \frac{r(i-1)}{ri+N-i} + \frac{N-i-1}{ri+N-i} &&= 1 - P_{i,i-1} - P_{i,i+1} \end{aligned}$$

for $1 \leq i < N$.

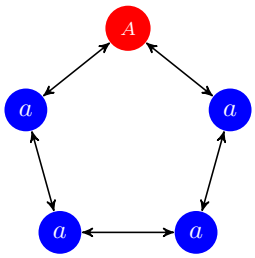

FIGURE 4. Cycle graph

4. CIRCULATION THEOREM

Complete graphs and cycle graphs show the same equilibrium between random drift and natural selection that a homogeneous population. Now, according to [5], it is natural to adopt the following definition:

Definition 4.1 ([5])**.** An evolutionary process on a graph G defined by a matrix $W = (w_{ij}) \in \mathcal{W}$ is said to be *equivalent to the Moran process* if the fixation probability of a single copy of a mutant allele A having fitness $r > 1$ is well defined (that is, it does not depend on the initial placement of the mutant allele)

and equal to the fixation probability

$$\Phi_A = \frac{1 - r^{-1}}{1 - r^{-N}}$$

of the Moran process, where N is the number of vertices of G.

Our next aim is to recall the *circulation theorem* proved by Liberman, Hauert and Novak [5], where they give some necessary and sufficient conditions for this equivalence. We start by recalling the circulation condition:

Definition 4.2 ([5])**.** A matrix $W = (w_{ij}) \in \mathcal{W}$ defines a *circulation* on G if for any vertex $i \in V$ the entering weight

$$w_-(i) = \sum_{j=1}^{N} w_{ji}$$

and the leaving weight

$$w_+(i) = \sum_{j=1}^{N} w_{ij}$$

are equal. The weighted graph (G, W) is also said to be *weight-balanced*.

In the case where W is stochastic, the entering weight $w_-(i) = \sum_{j=1}^{N} w_{ji}$ is also called the *temperature* of the vertex i and denote by T_i, while the leaving weight $w_+(i) = \sum_{j=1}^{N} w_{ij}$ is always equal to 1.

Circulation Theorem [5]**.** *For any matrix $W = (w_{ij}) \in \mathcal{W}$, the following conditions are equivalent:*

(1) W defines an evolutionary process equivalent to the Moran process.
(2) The probability that a initial population of n mutant individuals having fitness $r > 1$ reaches a mutant population of m individuals is given by

$$\Phi_A(r, W, n, m) = \frac{1 - r^{-n}}{1 - r^{-m}}.$$

(3) W defines a circulation on G.
(4) The number of elements of a state S performs a biased random walk on the integer interval $[0, N]$ with forward bias $r > 1$ and absorbing states 0 and N.

Proof. We prove a cycle of implications (1) ⇒ (2) ⇒ (3) ⇒ (4) ⇒ (1). To verify (1) ⇒ (2), it suffices to remark:

$$\Phi_A(r, W, n, N) = \Phi_A(r, W, n, m)\Phi_A(r, W, m, N) \quad , \quad \forall m \geq n$$

since the probability of reaching one state from another state depends only on their number of vertices.

Now we prove (2) ⇒ (3). First, for each state $S \neq \emptyset, V$, we define entering and leaving weights

$$w_-(S) = \sum_{i \in S} w_-(i) = \sum_{i \in S} \sum_{j=1}^{N} w_{ji}$$

and

$$w_+(S) = \sum_{i\in S} w_+(i) = \sum_{i\in S}\sum_{j=1}^{N} w_{ij}.$$

The probability that the mutant population S increases or decreases of one individual is given by

$$\delta^+(S) = \frac{rw_+(S)}{rw_+(S)+w_-(S)} \qquad \text{or} \qquad \delta^-(S) = \frac{w_-(S)}{rw_+(S)+w_-(S)}.$$

Then the birth-death rate is equal to

$$\frac{\delta^+(S)}{\delta^-(S)} = \frac{rw_+(S)}{w_-(S)}. \tag{4.7}$$

By hypothesis, we know:

$$\Phi_A(r, W, 1, 2) = \frac{1-r^{-1}}{1-r^{-2}} = \frac{r}{r+1}.$$

In particular, this means that the evolutionary process does not depend on the initial placement of the mutant individual. Then, writing $\delta^\pm = \delta^\pm(\{i\})$ for any $i \in V$, we have also:

$$\Phi_A(r, W, 1, 2) = \sum_{k=0}^{\infty} \delta^+(1-\delta^+-\delta^-)^k = \frac{\delta^+}{\delta^++\delta^-}.$$

We deduce:

$$\frac{\delta^+}{\delta^-} = r.$$

Combining this equality with (4.7) for $S = \{i\}$, we obtain $w_+(i) = w_-(i)$ for all $i \in V$, that is, W defines a circulation on G

To show (3) $\Rightarrow$ (4), we need to prove that the number of individuals $k = \#S$ of a mutant population S defines a Markov chain X_n verifying

$$\mathbb{P}[X_{n+1} = k+1 | X_n = k] = \delta^+(S) \quad \text{and} \quad \mathbb{P}[X_{n+1} = k-1 | X_n = k] = \delta^-(S)$$

with forward bias

$$\frac{\delta^+(S)}{\delta^-(S)} = r$$

and absorbing states 0 and N. Since W defines a circulation, we have:

$$w_+(S) - w_-(S) = \sum_{i\in S} w_+(i) - w_-(i) = 0$$

for every state $S \neq \emptyset, V$ in $\mathcal{S} = \mathcal{P}(V)$. Using (4.7), we deduce that the Markov chain X_n verifies:

$$\frac{\delta^+(S)}{\delta^-(S)} = \frac{rw_+(S)}{w_-(S)} = r$$

for all $S \neq \emptyset, V$.

Finally, we prove (4) $\Rightarrow$ (1). By hypothesis, the fixation probability of a single mutant individual having fitness $r > 1$ does not depend on its initial placement. More generally, the probability of reaching the whole population V from one

state S depends only on its number of vertices $k = \#S$. Thus, W defines a Markov chain with transition matrix

$$P = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ \delta_1^- & 1-\delta_1^- - \delta_1^+ & \delta_1^+ & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \delta_{N-1}^+ \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

where

$$\frac{\delta_k^+}{\delta_k^-} = r \quad , \quad \forall k = 1, \dots, N-1$$

by hypothesis. Arguing as for the case of a Moran process, we obtain:

$$\Phi_A(r, W, 1, N) = \frac{1 - r^{-1}}{1 - r^{-N}}.$$

This completes the proof. □

As a corollary of this theorem, we have:

Isothermal Theorem [5]**.** *For any stochastic matrix $W = (w_{ij}) \in \mathcal{W}$, the following conditions are equivalent:*

(1) W defines an evolutionary process equivalent to the Moran process.

(2) W defines an isothermal process *on G, i.e. all vertices $i \in V$ have the same temperature $T_i = \sum_{j=1}^N w_{ji} = T$.*

(3) W is doubly stochastic*, i.e. $T_i = \sum_{j=1}^N w_{ji} = 1$ for all $i \in V$.*

Proof. Firstly, we have:

$$\sum_{i=1}^n T_i = \sum_{i=1}^N \sum_{j=1}^N w_{ji} = \sum_{j=1}^N \sum_{i=1}^N w_{ji} = N$$

if W is stochastic. To see (1) ⇒ (2), it suffices to apply the circulation theorem, so the temperature $T_i = w_-(i)$ of each vertex i is equal to its leaving weight $w_+(i) = 1$. The implication (2) ⇒ (3) follows from the equation above because $\sum_{i=1}^n T_i = NT = N$ and hence $T = 1$ when W defines an isothermal process. Finally, according to the circulation theorem, any doubly stochastic matrix W defines a circulation (and hence an isothermal process) equivalent to the Moran process. □

5. Star graph

In [5], Liberman, Hauert and Novak showed that there are some graph structures, called *star structures*, which act as evolutionary amplifiers favoring advantageous alleles in non-homogeneous populations. These structures have also been studied in [1]. We will explicitly describe the asymptotic behavior of the average fixation probability.

A *star graph* consists of $N = m+1$ vertices labelled $0, 1, \dots, m$ where only the center 0 is connected with the peripheral vertices $1, \dots, m$, see the figure below.

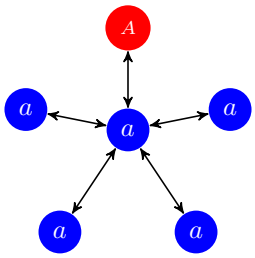

FIGURE 5. Star graph

Since the automorphism group is isomorphic to the symmetric group acting on the peripheral vertices, the state space reduces to the subsets $\emptyset$ and $\{1,\ldots,i\}$, $\{0\}$ and $\{0,1,\ldots,i\}$, $1 \leq i \leq m$, which can be described using ordered pairs. In the first entry, we write the number i of peripheral vertices inhabited by mutant individuals. In the second one, we use 1 or 0 to indicate whether or not there is a mutant individual at the center. Thus, the fixation probabilities will be denoted by

$$\Phi_{i,1} = \mathbb{P}[\exists n : X_n = (m,1) | X_0 = (i,1)]$$

and

$$\Phi_{i,0} = \mathbb{P}[\exists n : X_n = (m,1) | X_0 = (i,0)].$$

As for the Moran process, the evolutionary dynamics of the star structure is described by the system of linear equations

$$\begin{aligned}
\Phi_{0,0} &= 0 \\
\Phi_{i,1} &= \delta^+_{i,1}\Phi_{i+1,1} + \delta^-_{i,1}\Phi_{i,0} + (1-\delta^+_{i,1}-\delta^-_{i,1})\Phi_{i,1} && (5.8)\\
\Phi_{i,0} &= \delta^+_{i,0}\Phi_{i,1} + \delta^-_{i,0}\Phi_{i-1,0} + (1-\delta^+_{i,0}-\delta^-_{i,0})\Phi_{i,0} && (5.9)\\
\Phi_{m,1} &= 1
\end{aligned}$$

since non-trivial transition exists only between state $(i,1)$ (resp. $(i,0)$) and states $(i+1,1)$, $(i,0)$ and $(i,1)$ (resp. $(i-1,0)$, $(i,1)$ and $(i,0)$), see the figure below. The non-trivial entries in the transition matrix are given by

$$\delta^+_{i,1} = \mathbb{P}[X_{n+1} = (i+1,1) | X_n = (i,1)] = \frac{r}{r(i+1)+m-i} \cdot \frac{m-i}{m} \quad (5.10)$$

$$\delta^-_{i,1} = \mathbb{P}[X_{n+1} = (i,0) | X_n = (i,1)] = \frac{m-i}{r(i+1)+m-i} \quad (5.11)$$

$$\delta^+_{i,0} = \mathbb{P}[X_{n+1} = (i,1) | X_n = (i,0)] = \frac{ri}{ri+m+1-i} \quad (5.12)$$

$$\delta^-_{i,0} = \mathbb{P}[X_{n+1} = (i-1,0) | X_n = (i,0)] = \frac{1}{ri+m+1-i} \cdot \frac{i}{m} \quad (5.13)$$

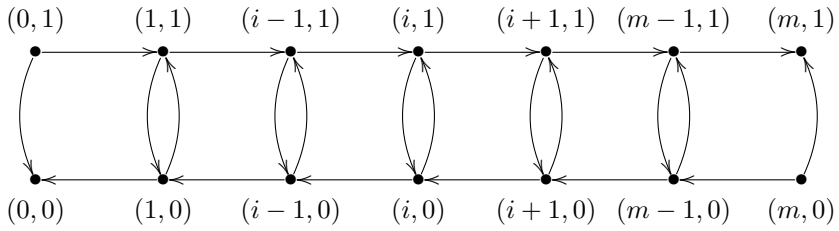

FIGURE 6. State space of a star graph

and

$$\begin{aligned} 1-\delta_{i,1}^{+}-\delta_{i,1}^{-} &= \frac{m+1}{m}.\frac{ri}{r(i+1)+m-i} \\ 1-\delta_{i,0}^{+}-\delta_{i,0}^{-} &= \frac{m+1}{m}.\frac{m-i}{ri+m+1-i} \end{aligned}$$

In particular, we have:

$$\Phi_{0,1}=\frac{r}{r+m}\Phi_{1,1} \quad \text{and} \quad \Phi_{1,0}=\frac{rm}{rm+1}\Phi_{1,1}. \tag{5.14}$$

Thus, the death-birth rates are given by

$$\gamma_{i,1} = \frac{\delta_{i,1}^{-}}{\delta_{i,1}^{+}} = \frac{m}{r} \tag{5.15}$$

$$\gamma_{i,0} = \frac{\delta_{i,0}^{-}}{\delta_{i,0}^{+}} = \frac{1}{rm} \tag{5.16}$$

Arguing as for (2.1) and dividing each side of Equations (5.8) and (5.9) by $\delta_{i,1}^{+}$ and $\delta_{i,0}^{+}$ respectively, we obtain equations

$$\Phi_{i+1,1}-\Phi_{i,1} = \gamma_{i,1}(\Phi_{i,1}-\Phi_{i,0}) = \frac{m}{r}(\Phi_{i,1}-\Phi_{i,0}) \tag{5.17}$$

$$\Phi_{i,1}-\Phi_{i,0} = \gamma_{i,0}(\Phi_{i,0}-\Phi_{i-1,0}) = \frac{1}{rm}(\Phi_{i,0}-\Phi_{i-1,0}) \tag{5.18}$$

analogous to (2.2). From (5.18), we prove inductively the following lemma:

Lemma 5.1. *For each $i=1,\dots,m$, the fixation probability*

$$\Phi_{i,0}=\sum_{j=1}^{i}\Big(\frac{1}{rm}\Big)^{i-j}\Big(\frac{rm}{rm+1}\Big)^{i-j+1}\Phi_{j,1}.$$

Proof. For $i=1$, the identity reduces to the second identity in (5.14). Assuming it is true for $i-1\leq 1$, from (5.18), we deduce:

$$(1+\frac{1}{rm})\Phi_{i,0}=\Phi_{i,1}+\frac{1}{rm}\Phi_{i-1,0}=\Phi_{i,1}+\frac{1}{rm}\sum_{j=1}^{i-1}\Big(\frac{1}{rm}\Big)^{i-1-j}\Big(\frac{rm}{rm+1}\Big)^{i-1-j+1}\Phi_{j,1}.$$

This implies the formula. □

Now, using (5.17), we obtain the following equation:

$$\begin{aligned}\Phi_{i+1,1}-\Phi_{i,1} &= \frac{m}{r}\big[\Phi_{i,1}-\frac{rm}{rm+1}\Phi_{i,1}-\frac{1}{rm}.(\frac{rm}{rm+1})^2\Phi_{i-1,1}\\ &- \sum_{j=1}^{i-2}(\frac{1}{rm})^{i-j}(\frac{rm}{rm+1})^{i-j+1}\Phi_{j,1}\big]\\ &= \frac{m}{r(rm+1)}\Phi_{i,1}-(\frac{m}{rm+1})^2\Phi_{i-1,1}\\ &- \sum_{j=1}^{i-2}\frac{m}{r}(\frac{1}{rm})^{i-j}(\frac{rm}{rm+1})^{i-j+1}\Phi_{j,1}\end{aligned}$$

where

$$\lim_{m\to+\infty}\sum_{j=1}^{i-2}\frac{m}{r}(\frac{1}{rm})^{i-j}(\frac{rm}{rm+1})^{i-j+1}\Phi_{j,1}=0.$$

Thus, when the number of peripheral vertices m tends to $+\infty$, the peripheral process whose fixation probabilities are equal to $\Phi_{i,1}$ becomes more and more close to the Moran process determined by the system of linear equations

$$\Phi_{i+1,1}-\Phi_{i,1}=\frac{1}{r^2}(\Phi_{i,1}-\Phi_{i-1,1}) \tag{5.19}$$

Even though there is no limit process, we say that the peripheral process is *asymptotically equivalent* to the Moran process determined by (5.19) and whose relative fitness is quadratically amplified.

On the other hand, according to (5.14), the average fixation probability

$$\Phi_A=\frac{1}{m+1}\Phi_{0,1}+\frac{m}{m+1}\Phi_{1,0}$$

is equal to

$$(\frac{1}{m+1}.\frac{r}{r+m}+\frac{m}{m+1}.\frac{rm}{rm+1})\Phi_{1,1}$$

and therefore Φ_A also becomes more and more close to the fixation probability of the Moran process determined by (5.19), which has relative fitness $r^2>1$. We can resume this discussion in the following statement:

Star Theorem [5]**.** *The star structure is a quadratically amplifier of selection in the sense that the average fixation probability of a mutant individual with relative fitness $r>1$ is asymptotically equivalent to the fixation probability*

$$\Phi_A=\frac{1-r^{-2}}{1-r^{-2m}}$$

for the Moran process with relative fitness $r^2>1$

Actually, there are super-star structures which act as amplifiers of selection of arbitrary polynomial degree [5].

6. Conclusion

In this paper, we have described some basic ideas of evolutionary graph theory sketched by Liberman, Hauert and Novak [5] by focusing in the study of the average fixation probability of a randomly arising mutation. The effect of the starting position of the mutant individual has been discussed by Broom, Rychtář and Stadler in [2] and [3] for small-world networks and other small-order graphs. For example, like they observed, regular graphs have the worst structure for the mutant spread. These authors were also interested in the time to fixation of a advantageous allele for strongly connected directed graphs and undirected graphs, and how it depends on the graph structure and the initial placement of the mutant. In a forthcoming paper, we will study the average fixation probability and the expected fixation time in different types of graphs and complex networks.

References

[1] M. Broom, J. Rychtář. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. *Proc. R. Soc. A*, **464** (2008), 2609–2627.
[2] M. Broom, J. Rychtář, B. Stadler. Evolutionary dynamics on small-order graphs. *J. Intesdiscip. Math.*, **12** (2009), 129–140.
[3] M. Broom, J. Rychtář, B. Stadler. Evolutionary dynamics on graphs - the effect of graph structure and initial placement on mutant spread. *J. Stat. Theory Pract.*, **5** (2011), 369–381.
[4] R.A. Fisher. *The genetical theory of natural selection.* Clarendon Press, Oxford, 1930.
[5] E. Lieberman, C. Hauert, M. A. Novak. Evolutionary dynamics on graphs. *Nature*, **433** (2005), 312–316.
[6] P. A. P. Moran. Random processes in genetics. *Proc. Camb. Phil. Soc.*, **54** (1958), 60–71.
[7] M. A. Novak. *Evolutionary dynamics: exploring the equations of life.* Harvard University Press, Cambridge, 2006.
[8] S. Wright. Evolution in Mendelian populations. *Genetics*, **16** (1931), 97–159.

Departamento de Xeometría e Topoloxía, Facultade de Matemáticas, Universidade de Santiago de Compostela, Rúa Lope Gómez de Marzoa s/n, E-15782 Santiago de Compostela (Spain)
E-mail address: fernando.alcalde@usc.es

Departamento de Didáctica das Ciencias Experimentais, Facultade de Formación do Profesorado, Universidade de Santiago de Compostela, Avda. Ramón Ferreiro, 10, E-27002 Lugo (Spain)
E-mail address: pablo.gonzalez.sequeiros@usc.es

Centro Universitario de la Defensa - IUMA Universidad de Zaragoza, Academia General Militar, Ctra. Huesca s/n, E-50090 Zaragoza (Spain)
E-mail address: alvarolozano@unizar.es