AN ASYMPTOTIC EXPRESSION FOR THE FIXATION PROBABILITY OF A MUTANT IN STAR GRAPHS

FABIO A. C. C. CHALUB
Departamento de Matemática and Centro de Matemática e Aplicações
Universidade Nova de Lisboa, Quinta da Torre, 2829-516, Caparica, Portugal

(Communicated by Jorge Zubelli)

Abstract. We consider the Moran process in a graph called the “star” and obtain the asymptotic expression for the fixation probability of a single mutant when the size of the graph is large. The expression obtained corrects the previously known expression announced in reference [E Lieberman, C Hauert, and MA Nowak. Evolutionary dynamics on graphs. Nature, 433(7023):312–316, 2005] and further studied in [M. Broom and J. Rychtar. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. Proc. R. Soc. A-Math. Phys. Eng. Sci., 464(2098):2609–2627, 2008]. We also show that the star graph is an accelerator of evolution, if the graph is large enough.

1. Introduction. The central question in the mathematical study of population genetics is to understand how gene frequencies vary in time. This topic has a long history, since at least the works of Wright and Fisher [15, 6]. A simpler model was introduced by Moran [10]: in that case a population of constant size $n$ of two types evolves in discrete time steps. To each type we attribute a positive number, called fitness. At each time one individual is selected to reproduce with probability proportional to its fitness and one to die (possibly the same one), with probability $1/n$. The so called Moran process continues until the entire population consists of individuals at a given type. We say that one type has fixed while the other has become extinct. One particularly important question is the fixation probability of a given type $A$, i.e., the probability that after a sufficiently long time, type $A$ reaches fixation; see [13, 11]

The Moran process is a particular case of the so called Birth-Death (BD) processes with selection on the birth. Birth-Death process are a particular case of stochastic process in which the number of individuals of a given type varies at most one per time step, i.e., transitions are allowed only between neighboring sites in an appropriate topology. The “selection at birth” indicates that the removed individual was selected using an equally distributed random variable, while the random variable indicating selection for reproduction is not necessarily identically distributed among all individuals. See, e.g., [11, 3] for further informations on BD processes.

In [9], topology, in the form of a graph, was introduced in the study of population dynamics. Individuals were represented by vertices, while edges represented possible positions for the offspring of individuals in a given vertex. In particular, consider a

2010 Mathematics Subject Classification. 92D25, 41A60.

Key words and phrases. Evolutionary graph theory, Moran process, fixation probability, asymptotic expansions, star graph.
generic graph where all vertices but one (selected at random) are occupied by a type with fitness 1 (the resident type). The remaining vertex is occupied by a mutant with fitness \( r > 0 \). After a certain fixed time, one of the individuals (resident or mutant) is selected to reproduce with probability proportional to fitness, and its offspring replaces one of the individuals occupying an adjacent vertex, selected with equal probability. This is the so called Moran process on graphs [10, 9]. The invasion probability \( \rho \) is given by the probability that after a sufficiently long time all vertices are occupied by the mutant type (i.e., the mutant has fixed and the resident has become extinct). For a complete graph (all pairs of vertices are connected) with \( n \) vertices, the invasion probability is given by

\[
\rho_C(r, n) := \frac{1 - r^{-1}}{1 - r^{1 - n}} \quad \text{if } r \neq 1 \text{ and } \rho_C(1, n) := \frac{1}{n}.
\]

This corresponds to the original problem studied in [10]. We say that a graph is an accelerator of evolution if \( \rho > \rho_C \) if and only if \( r > 1 \). If one of these inequalities is reversed, we say that the graph is a suppressor of evolution. See [11] for a more detailed explanation of concepts used in this article and also for examples of accelerators and suppressors of evolution.

In this article, we study a specific graph, called the “star”. The star is a graph with \( n + 1 \) vertices, labeled from 0 to \( n \), where vertex number 0 is called the center and all the others are the leaves. The only edges are between vertex 0 and all other vertices. These edges are bidirectional, see figure 1. In [9], the star graph was introduced as a particular case of a family of graphs, called the “superstar” and parameterized by two natural numbers: \( k \), the number of layers and \( n \), the number of vertices at layer \( k + 1 \), for any \( k' < k \), connected to any given vertex at layer \( k' \). The first layer has a single vertex, called the center. The star corresponds to \( k = 2 \).

According to the conjecture in [9], \( \rho \approx \frac{1 - r^{-1}}{1 - r^{1 - n}} \) for \( n \rightarrow \infty \). However, recently [5] showed that this conjecture is false for \( k \geq 5 \). See also [8] for results in the class of bithermal graphs, including the “star”, using fixed point techniques.

We start this article from a previous work (see [4]), where an explicit formula for the invasion probability in stars was derived and study the associated asymptotic expression when \( n \), the number of individuals, is large. The aim of this work is to show that the expression for the leading term for \( n \rightarrow \infty \) and fixed \( r \) in the asymptotic expression of the invasion probability is given by

\[
\rho_{\text{lo}}(r, n) := \frac{1 - r^{-2}}{1 - r^{-2n(e^{(r^2-1)/r})}}.
\]

(1)

A precise definition of \( \rho_{\text{lo}}(r, n) \) as the leading term of the exact expression in \( n \rightarrow \infty \) and fixed \( r \) is given in Eq. (3).

Note that this expression is different from the one conjectured in [9], discussed above, for the case \( k = 2 \).

In section 2 we discuss the results presented in [4]. In section 3, we rigorously derive Eq. (1) and in section 4, we prove that the star is an accelerator of evolution, if \( n \) is large enough. We study this problem numerically and present our conclusions in the final section 5.

2. Evolutionary dynamics in the star graph. In [4], the exact expression

\[
\rho_n(r) = \frac{r}{n + r} + \frac{n^2 r}{n r + 1} \left( 1 + \frac{n}{n + r} \sum_{j=1}^{n-1} \left( \frac{n + r}{r(n + 1)} \right)^j \right)^j
\]

(2)
for the invasion probability of mutant with fitness \( r \) in a star graph with \( n \) leaves was found.

We agree with formula (2), however the asymptotic expression found in the same reference when \( n \to \infty \) seems to not be correct. In particular, the asymptotic expression

\[
\frac{n \frac{n r}{n r+1} + \frac{r}{n+r}}{(n+1) \cdot \left(1 + \frac{n}{n+r} \sum_{j=1}^{n-1} \left(\frac{n+r}{r(n+r+1)}\right)^j \right)^\frac{1}{2}} \approx \frac{1 - \frac{r^{-2}}{1 - \frac{1}{r}}}{1 - \frac{r^{-2}}{n}}
\]

when \( n \to \infty \) is wrong (see [4, page 2616]).

3. Asymptotic expression. In this section, we will show that the associated asymptotic expression for \( \rho_n(r) \), when \( n \to \infty \), is given by

\[
\rho_n(r) = \rho_0(r,n) \left(1 + O \left(\frac{1}{n}\right)\right),
\]

where \( \rho(r,n) \) was defined in Eq. (1).

If \( r \neq 1 \), Eq. (2) can be simplified to

\[
\rho_n(r) = \frac{r}{n+1} \frac{n^3 + n^2r + nr + 1}{n^2r + nr^2 + n + r} \left\{1 + \frac{n}{n+r} \left[\frac{n+r}{r(n+r+1)} - \frac{n+r}{r(n+r+1)}\right]\right\}^{-1}.
\]

Now, we consider the asymptotic expression of \( \rho_n(r) \) when \( n \to \infty \) and \( r \neq 1 \). Note first that

\[
\frac{r}{n+1} \frac{n^3 + n^2r + nr + 1}{n^2r + nr^2 + n + r} = 1 + O \left(\frac{1}{n}\right).
\]

Furthermore

\[
1 + \frac{n}{n+r} \left[\frac{n+r}{r(n+r+1)} - \frac{n+r}{r(n+r+1)}\right] = 1 + \frac{1}{r^2 - 1} \left[1 - \left(\frac{n+r}{r(n+r+1)}\right)^{-1}\right]
\]
We also write
\[ \frac{n + r}{r(nr + 1)} = \frac{1}{r^2} \left( 1 + \frac{\frac{r-1}{n+1}}{1 + \frac{1}{r(nr + 1)}} \right) = \frac{1}{r^2} \left( 1 + \frac{r^2 - 1}{nr} + O(n^{-2}) \right). \]

Now, we use that
\[ \left( 1 + \frac{x}{n} + O(n^{-2}) \right)^n = e^{n \log(1 + x + O(n^{-2}))} = e^x + O(n^{-1}) = e^x \left[ 1 + O(n^{-1}) \right]. \]

and then
\[ \left( \frac{n + r}{r(nr + 1)} \right)^{n-1} = \frac{e^{(r^2 - 1)/r}}{r^{2(n-1)}} \left( 1 + O(n^{-1}) \right). \]

We put together Eqs. (5) and (6) and find
\[
1 + \frac{n}{n + r} \left[ \frac{n + r}{r(nr + 1)} - \left( \frac{n + r}{r(nr + 1)} \right)^n \right]
= 1 + \frac{1}{r^2} - 1 - \frac{1}{r^2 - 1} \frac{e^{(r^2 - 1)/r}}{r^{2(n-1)}} \left( 1 + O(n^{-1}) \right)
= 1 - \frac{r - 2}{1 - r^{-2}} - \frac{1}{1 - r^{-2}} \frac{e^{(r^2 - 1)/r}}{r^{2n}} \left( 1 + O(n^{-1}) \right)
= 1 - \frac{r - 2}{1 - r^{-2}} \frac{n e^{(r^2 - 1)/r}}{1 + O(n^{-1})}.
\]

Gathering all asymptotic expansions, and using the definition (3), we find expression (1).

4. The star is an accelerator of evolution. Now, let us compare the Moran process in the star and in the complete graph.

For neutral evolution \((r = 1)\), we have the following simple result:

**Lemma 1.** \( \rho_n(1) = \frac{1}{n+1} = \rho_C(1, n+1) \).

**Proof.** See Eq. (2) with \( r = 1 \). Equivalently, this result can be proved from Eq. (4) in the limit \( r \to 1 \). \( \square \)

A similar result is valid for a large class of graphs; in particular, it is valid for graphs with at most one root (a root is a vertex with no edge leading to it). However, a single mutant cannot invade a graph with two or more roots, and, in this case, the invasion probability is equal to zero. See [11].

Note that \( \lim_{r \to 1} \rho(r, n) = \frac{1}{n+1} \neq \frac{1}{n+1} \), showing that the large population limit \( n \to \infty \) and the weak selection limit \( r \to 1 \) are not interchangeable.

For \( r \neq 1 \), we show that the star is an accelerator of evolution:

**Theorem 1.** Consider \( r > 1 \) (\( r < 1 \)). Then, for \( n \) large enough, \( \rho_n(r) > \rho_C(r, n+1) \) (\(< \), respect.), and therefore the star is an accelerator of evolution.

**Proof.** We initially see that
\[ \lim_{n \to \infty} \frac{\rho_n(r, n)}{\rho_C(r^2, n+1)} = \lim_{n \to \infty} \frac{1 - r^{-2(n+1)}}{1 - r^{-2n}e^{(r^2 - 1)/r}} = \begin{cases} \frac{r^{-2}e^{-(r^2 - 1)/r}}{1 - r^{-2n}e^{(r^2 - 1)/r}} > 1, & r < 1, \\ 1, & r > 1. \end{cases} \]

On the other hand
\[ \frac{\rho_C(r^2, n+1)}{\rho_C(r, n+1)} = \frac{(1 - r^{-(n+1)})(1 - r^{-2})}{(1 - r^{-2(n+1)})(1 - r^{-1})}. \]
\[ \lim_{n \to \infty} \frac{\rho_n(r, n+1)}{\rho(r, n+1)} = \begin{cases} 0, & r < 1, \\ \frac{r+1}{r}, & r > 1. \end{cases} \]

Finally, we write

\[ \frac{\rho_n(r)}{\rho(r, n+1)} = (1 + O(n^{-1})) \times \frac{\rho_n(r, n)}{\rho(r, n+1)} \times \frac{\rho_n(r^2, n+1)}{\rho(r^2, n+1)} \]

and taking the limit \(n \to \infty\), we conclude that

\[ \lim_{n \to \infty} \frac{\rho_n(r)}{\rho(r, n+1)} = \begin{cases} 0, & r < 1, \\ \frac{r+1}{r}, & r > 1. \end{cases} \]

For \(n\) large enough \(\rho_n(r) > \rho(r, n+1)\) if and only if \(r > 1\).

5. **Discussion.** Despite the fact that the expression conjectured in [9] is wrong for \(k = 2\) and \(k \geq 5\), it has been widely used (e.g., [7, 14, 2]). However, exact or asymptotic expressions for \(k > 2\) have not been found and should be the subject of further investigations. See also [1, 12] for a review of evolutionary graph theory.

Differences between expression (1) and the expression \(\frac{1-r^{-2}}{1-r^{-2\times\pi}}\) given by [4] and [9] are hardly noticeable from the numerical point of view. In fact, for \(r > 1\), the difference between both expressions are exponentially small, as \(r^{-n} \to 0\) exponentially fast when \(n \to \infty\). However, for \(r < 1\) these differences, despite being small, cannot be neglected. See table 1 and figure 2 for numerical comparisons.

| \(r\) | \(n\) | \(\rho_n(r)\) | \(\rho(r, n)\) | \(\epsilon\) | \(\frac{1-r^{-2}}{1-r^{-2\times\pi}}\) | \(\epsilon\) |
|---|---|---|---|---|---|---|
| 0.800 | 500 | \(1.08 \times 10^{-47}\) | \(1.09 \times 10^{-47}\) | 0.000629 | \(6.92 \times 10^{-48}\) | 0.358 |
| 0.800 | 1000 | \(1.33 \times 10^{-47}\) | \(1.34 \times 10^{-47}\) | 0.003320 | \(8.51 \times 10^{-48}\) | 0.360 |
| 0.900 | 500 | \(5.04 \times 10^{-47}\) | \(5.06 \times 10^{-47}\) | 0.00509 | \(4.10 \times 10^{-48}\) | 0.186 |
| 0.900 | 1000 | \(8.83 \times 10^{-47}\) | \(8.85 \times 10^{-47}\) | 0.00255 | \(7.17 \times 10^{-48}\) | 0.188 |
| 0.950 | 1000 | \(3.34 \times 10^{-46}\) | \(3.35 \times 10^{-46}\) | 0.00224 | \(3.03 \times 10^{-46}\) | 0.0955 |
| 0.950 | 5000 | \(2.06 \times 10^{-224}\) | \(2.06 \times 10^{-224}\) | 0.000448 | \(1.86 \times 10^{-224}\) | 0.0971 |
| 0.990 | 1000 | \(3.85 \times 10^{-11}\) | \(3.86 \times 10^{-11}\) | 0.00203 | \(3.78 \times 10^{-11}\) | 0.0179 |
| 0.990 | 5000 | \(4.66 \times 10^{-46}\) | \(4.66 \times 10^{-46}\) | 0.000459 | \(4.57 \times 10^{-46}\) | 0.0195 |
| 1.01 | 1000 | 0.0197 | 0.0197 | 0.00204 | 0.0197 | 0.00204 |
| 1.05 | 1000 | 0.0928 | 0.0930 | 0.00197 | 0.0930 | 0.00197 |
| 1.01 | 5000 | 0.0197 | 0.0197 | 0.000484 | 0.0197 | 0.000484 |
| 1.05 | 5000 | 0.0929 | 0.0930 | 0.000410 | 0.0930 | 0.000410 |

*Table 1.* Invasion probabilities for different values of \(r\) and \(n \gg 1\): \(\rho_n\) indicates the exact value given by Eq. (2); in the third column, we compute the approximate value given by Eq. (1) and its associated error \(\epsilon = |\rho_n(r, n) - \rho_n(r)|/\rho_n(r)\); the last column indicates the approximation in references [9, 11, 4] and its associated error. Note that for \(r > 1\), the two approximations are essentially equivalent; for \(r < 1\) the first approximation is clearly superior. All calculations were performed in Sage v. 5.13.
Figure 2. (Color online) For different values of $n$ ($x$ axis), we plot $\rho_n(r)$ (blue dots), $\rho_0(n,r)$ (red line) and $1 - \frac{r-2}{r^2}$ (green line). We use a logarithmic scale in the $y$ axis. Note that the red line is consistently a better approximation than the green one with respect to the exact values.

Acknowledgments. This work was supported by CMA/FCT/UNL and FCT/Portugal, under the project UID/MAT/00297/2013. FACCC has also a “Investigador FCT” (FCT/Portugal) grant. FACCC also thanks the careful reading of the manuscript by the referees, in particular an inconsistency pointed out by one of the reviewers that allowed an improvement in the statement and proof of Theorem 1.

REFERENCES

[1] B. Allen and M. Nowak, Games on graphs, EMS Surv. Math. Sci., 1 (2014), 113–151.
[2] M. Broom, C. Hadjichrysanthou and J. Rychtar, Evolutionary games on graphs and the speed of the evolutionary process, Proc. R. Soc. A-Math. Phys. Eng. Sci., 466 (2010), 1327–1346.
[3] M. Broom and J. Rychtář, Game-theoretical Models in Biology, CRC Press, Boca Raton, FL, 2013.
[4] M. Broom and J. Rychtář, An analysis of the fixation probability of a mutant on special classes of non-directed graphs, Proc. R. Soc. A-Math. Phys. Eng. Sci., 464 (2008), 2609–2627.
[5] J. Diaz, L. A. Goldberg, G. B. Mertzios, D. Richerby, M. Serna and P. G. Spirakis, On the fixation probability of superstars, Proc. R. Soc. A-Math. Phys. Eng. Sci., 469 (2013), 20130193, 11 pp.
[6] R. A. Fisher, The Genetical Theory of Natural Selection, Clarendon Press, Oxford, 1999.
[7] M. Frean, P. B. Rainey and A. Traulsen, The effect of population structure on the rate of evolution, Proc. R. Soc. B-Biol. Sci., 280 (2013), 1762.
[8] B. Houchmandzadeh and M. Vallade, Exact results for fixation probability of bithermal evolutionary graphs, Biosystems, 112 (2013), 49–54.
[9] E. Lieberman, C. Hauert and M. A. Nowak, Evolutionary dynamics on graphs, Nature, 433 (2005), 312–316.
FIXATION IN STAR GRAPHS

[10] P. A. P. Moran, *The Statistical Processes of Evolutionary Theory*, Clarendon, Oxford, 1962.
[11] M. A. Nowak, *Evolutionary Dynamics: Exploring the Equations of Life*, The Belknap Press of Harvard University Press, Cambridge, MA, 2006.
[12] P. Shakarian, P. Roos and A. Johnson, *A review of evolutionary graph theory with applications to game theory*, *Biosystems*, 107 (2012), 66–80.
[13] A. Traulsen, M. A. Nowak and J. M. Pacheco, *Stochastic dynamics of invasion and fixation*, *Phys. Rev. E*, 74 (2006), 011909.
[14] C. Zhang, Y. Wu, W. Liu and X. Yang, Fixation probabilities on complete star and bipartite digraphs, *Discrete Dyn. Nat. Soc.*, (2012), Art. ID 940465, 21 pp.
[15] S. Wright, Evolution in mendelian populations, *Genetics*, 16 (1931), 97–159.

Received July 2015; revised February 2016.

E-mail address: chalub@fct.unl.pt