Lecture IV: Evolution of social behavior in structured populations

F. Débarre

January 2017

1 Introduction

For this last lecture, we consider the evolution of social behavior in structured populations. The individuals that we call "social" interact with others and these interactions directly affect fitness components of actors (*i.e.*, those who engage in the social interaction) and recipients. We will have in mind the evolution of altruistic behavior, and say that social individuals provide others with benefits that directly increase their fecundity or survival probability, but that they even pay a fecundity/survival cost for being social. In a well-mixed population, this kind of behavior is counter-selected. In spatially structured populations however, the local environment that an individual experiences may be very different from the global composition of the population. We will see that spatial structure can, under certain circumstances, favor the evolution of social behavior.

2 Model

2.1 Types of individuals

There are N individuals in the population, and two types of individuals: individuals that are "social" (labeled C), and individuals that are not (labeled D).

We describe the state of the population at time t with a N-long column vector $\mathbf{X}(t)$, whose i^{th} element is an indicator variable, equal to 1 if the individual at site i at time t is of type C (social), and equal to 0 otherwise. A vector $\mathbf{x}$ corresponds to a given state of the population. The ensemble of all possible states is $\Omega = \{0, 1\}^N$. We are interested in the long-term behavior of the population, when $t \to \infty$, and will focus on the proportion of individuals of type C in the population (given by $\overline{X} = \sum_{i=1}^{N} X_i / N$).

2.2 Reproduction

Reproduction is clonal; mutations occur with probability μ, and a mutated offspring is of either type with probability $1/2$ (see figure 1). In other words, a parent produces an offspring with the same type with probability $1 - \mu/2$, and an offspring of the other type with probability $\mu/2$.

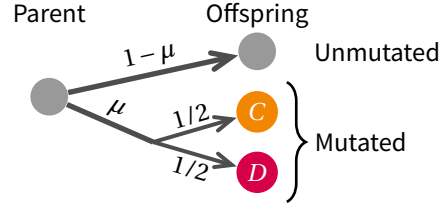In what follows, we will assume that mutation is rare ($\mu \ll 1$).

Figure 1: Reproduction and mutation

2.3 Spatial and social structure

To characterize population structure, we will say that there are N sites, *i.e.*, locations that can host an individuals. Since population size is constant and equal to N, each site is occupied by exactly one individual. Then we define two graphs, that describe the connexions between the different sites. These two graphs have N vertices, each vertex corresponding to a site (so we use the two words interchangeably).

Dispersal

A weighted graph $\mathscr{D}$ (with adjacency matrix $\mathbf{D} = \{d_{ij}\}_{1 \leq i, j \leq N}$) summarizes where individuals can send their offspring to; $\mathscr{D}$ is called *dispersal graph*. We consider graphs that are regular, scaled such that

$$\sum_{j=1}^{N} d_{ij} = \sum_{j=1}^{N} d_{ji} = 1, \tag{1a}$$

but also symmetric, *i.e.*,

$$\forall i, j \in \{1, \ldots, N\}, \quad d_{ij} = d_{ji}. \tag{1b}$$

We further assume that the graph is transitive, *i.e.*, looks the same from every node (/vertex/site). For simplicity, we also assume here that there are not self-loops on the graph, *i.e.*, that for all sites i, $d_{ii} = 0$. Examples of graphs satisfying these conditions are given in figure 2.

Social interactions

A weighted graph $\mathscr{E}$ (with adjacency matrix $\mathbf{E} = \{e_{ij}\}_{1 \leq i, j \leq N}$) summarizes who interacts with whom in the population. This graph, called the *interaction graph*, can be the same as the dispersal graph (which is often the case when considering for instance a regular grid), but can also be different (for instance in a subdivided population where individuals can disperse to other demes but only interact within their deme). We consider that interactions are with other individuals exclusively, not with oneself (*i.e.*, for all sites i, $e_{ii} = 0$).

Social individuals directly affect the fitness of individuals they interact with. Here, we assume that fecundity is the trait directly affected by social interactions (but we could instead consider effects on survival). Social individuals provide other individuals (social or not) with a benefit b, but pay a cost c. The overall effect of these interactions on fitness is modulated by a parameter ω, corresponding to the strength of
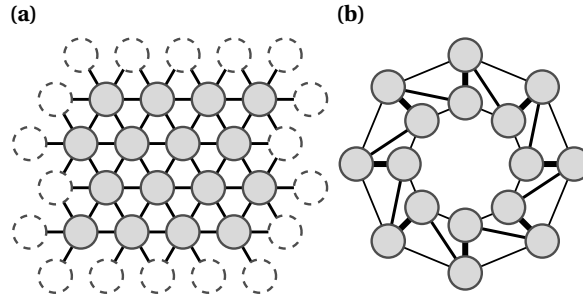
Figure 2: Examples of graphs that satisfy our assumptions of symmetry and transitivity. Graph (a) is a regular grid or lattice with $n = 6$ neighbors, graph (b) does not have a name as far as I know!

selection. As a result, with a population state $\mathbf{x}$, the fecundity of an individual living at site i is given by

$$f_i(\mathbf{x}) = 1 + \omega \left[\left(\sum_{l=1}^{N} e_{li}\, \mathsf{b}\, x_l \right) - \mathsf{c}\, x_i \right]. \tag{2}$$

In what follows, we will assume that selection is weak, *i.e.*, that $\omega \ll 1$.

2.4 Updating rules

The size of the population being fixed, at each time step the number of individuals dying is equal to the number of new individuals in the population. We will assume that one individual is replaced at each time step ("Moran model"), and we will consider two orders of events: Death-Birth (DB) and Birth-Death (BD). We denote by $w_{ji}(\mathbf{x})$ the probability that at the next time point, site j is occupied by a descendant of the individual currently living at site i, or is this individual (when $j = i$). Note that since each site is occupied by exactly one individual, we have

$$\sum_{i=1}^{N} w_{ji} = 1, \tag{3a}$$

and so in particular

$$w_{ii} = 1 - \sum_{\substack{j=1 \\ j \neq i}}^{N} w_{ij}, \tag{3b}$$

which is the probability that the individual at site i survives (since $d_{ii} = 0$, there is no reproduction to the parental site). (And yes, be careful, in (3b), it is w_{ij}).

Death-Birth

In a Death-Birth model, first the individual who dies is chosen uniformly at random (each individual has a probability $1/N$ of dying), and then one individual is chosen among all propagules that could reach the just emptied site:

$$\forall j \neq i, w_{ji}^{\mathrm{DB}}(\mathbf{x}) = \frac{1}{N} \frac{d_{ij} f_i(\mathbf{x})}{\sum_{k=1}^{N} d_{kj} f_k(\mathbf{x})}, \tag{4a}$$

and an individual survives if it is not the one chosen to die:

$$w_{ii}^{\text{DB}} = 1 - \frac{1}{N}. \tag{4b}$$

Using the formulas for f defined in equation (2), we obtain, at the first order in ω,

$$\forall j \neq i, w_{ji}^{\text{DB}}(\mathbf{x}) = \frac{d_{ij}}{N} \left(1 + \omega \left[\sum_{l=1}^{N} e_{li} \mathsf{b} x_l - \mathsf{c} x_i - \sum_{k=1}^{N} d_{kj} \left(\sum_{l=1}^{N} e_{lk} \mathsf{b} x_l - \mathsf{c} x_k \right) \right] + \mathcal{O}\left(\omega^2\right) \right). \tag{4c}$$

Birth-Death

In a Birth-Death model, first the individual who reproduces is chosen according to fecundities, and then the site where it reproduces – and kills the occupant – is chosen among the places where the offspring can disperse (the "neighbors"):

$$\forall j \neq i, w_{ji}^{\text{BD}}(\mathbf{x}) = \frac{f_i}{\sum_{k=1}^{N} f_k} \frac{d_{ij}}{\sum_{k=1}^{N} d_{ik}}. \tag{5a}$$

The difference between the two orders of events is reflected in the denominators of w_{ji}^{BD} and w_{ji}^{DB}. Given the regularity of the graph (equation (1a)) and the formulas for the fecundities (equation (2)), at the first order in ω, equation (5a) becomes

$$\forall j \neq i, w_{ji}^{\text{BD}}(\mathbf{x}) = \frac{d_{ij}}{N} \left(1 + \omega \left[\sum_{l=1}^{N} e_{li} \mathsf{b} x_l - \mathsf{c} x_i - \sum_{k=1}^{N} \frac{1}{N} \left(\sum_{l=1}^{N} e_{lk} \mathsf{b} x_l - \mathsf{c} x_k \right) \right] + \mathcal{O}\left(\omega^2\right) \right). \tag{5b}$$

The term w_{ii} is found using equation (3b):

$$w_{ii} = 1 - \sum_{\substack{j=1 \\ j \neq i}}^{N} w_{ij}$$

$$= 1 - \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{d_{ji}}{N} \left(1 + \omega \left[\sum_{l=1}^{N} e_{lj} \mathsf{b} x_l - \mathsf{c} x_i - \sum_{k=1}^{N} \frac{1}{N} \left(\sum_{l=1}^{N} e_{lk} \mathsf{b} x_l - \mathsf{c} x_k \right) \right] + \mathcal{O}\left(\omega^2\right) \right),$$

and given that $d_{ii} = 0$, this further simplifies in

$$w_{ii} = 1 - \frac{1}{N} - \omega \sum_{j=1}^{N} \frac{d_{ij}}{N} \left[\sum_{l=1}^{N} e_{lj} \mathsf{b} x_l - \mathsf{c} x_j - \sum_{k=1}^{N} \frac{1}{N} \left(\sum_{l=1}^{N} e_{lk} \mathsf{b} x_l - \mathsf{c} x_k \right) + \mathcal{O}\left(\omega^2\right) \right],$$

$$= 1 - \frac{1}{N} - \frac{\omega}{N} \sum_{j=1}^{N} d_{ij} \left(\sum_{l=1}^{N} e_{lj} \mathsf{b} x_l - \mathsf{c} x_j \right) + \frac{\omega}{N} \sum_{k=1}^{N} \frac{1}{N} \left(\sum_{l=1}^{N} e_{lk} \mathsf{b} x_l - \mathsf{c} x_k \right) + \mathcal{O}\left(\frac{\omega^2}{N}\right). \tag{6}$$

Now that we have defined the model, we are going to study how the proportion of social individuals in the population changes over time.

3 Change in the proportion of social individuals in the population

3.1 For any life-cycle

We denote by $\overline{\Delta}(\mathbf{x})$ the expected change in the frequency of social individuals in the population during one time step, given that the population is currently in state $\mathbf{x}$. We need to be careful to differentiate reproduction (during which mutation can occur) and survival. With our assumptions, we have

$$\overline{\Delta}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^{N} \left[\sum_{\substack{j=1 \\ j \neq i}}^{N} w_{ji}(\mathbf{x}) \left(x_i \left(1 - \frac{\mu}{2}\right) + (1 - x_i)\frac{\mu}{2} \right) + w_{ii}(\mathbf{x}) x_i \right] - \overline{x},$$

$$= \frac{1}{N} \sum_{i=1}^{N} \left[\sum_{\substack{j=1 \\ j \neq i}}^{N} w_{ji}(\mathbf{x}) x_i (1 - \mu) + w_{ii}(\mathbf{x}) x_i + \frac{\mu}{2} \sum_{\substack{j=1 \\ j \neq i}}^{N} w_{ji}(\mathbf{x}) \right] - \overline{x}. \tag{7}$$

To proceed further, we will specify the order of events (Death-Birth or Birth-Death) that we have chosen.

3.2 Death-Birth

Replacing w_{ji} and w_{ii} for this updating rule ((4c) et (4b)), from (7) we obtain

$$\overline{\Delta}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^{N} \left[\sum_{\substack{j=1 \\ j \neq i}}^{N} \left(\frac{d_{ij}}{N} \left(1 + \omega \left[\sum_{l=1}^{N} e_{li}\mathsf{b}x_l - \mathsf{c}x_i - \sum_{k=1}^{N} d_{kj} \left(\sum_{l=1}^{N} e_{lk}\mathsf{b}x_l - \mathsf{c}x_k \right) \right] \right) \right) x_i (1 - \mu) \right.$$

$$\left. + \left(1 - \frac{1}{N}\right) x_i \right] + \frac{\mu}{2N} - \overline{x} + \mathcal{O}\left(\frac{\omega^2}{N}\right). \tag{8a}$$

All the sums are getting a bit out of control, so we will simplify notation by omitting the ranges, keeping in mind that they are from 1 to N; also, since by assumption $d_{ii} = 0$, the sum involving j is the same whether we exclude i or not. Finally, we note that since x_i takes value in $\{0, 1\}$, $x_i^2 = x_i$. Reorganizing, we have

$$\overline{\Delta}(\mathbf{x}) = \frac{\mu}{N} \left(\frac{1}{2} - \overline{x} \right) + \frac{\omega}{N^2} \left[\mathsf{b} \left(\sum_i \sum_l e_{li} x_l x_i - \sum_i \sum_j \sum_k \sum_l d_{ij} d_{kj} e_{lk} x_l x_i \right) \right.$$

$$\left. + \mathsf{c} \left(\sum_i x_i - \sum_i \sum_j \sum_k d_{ij} d_{kj} x_k x_i \right) \right] + \mathcal{O}\left(\frac{\omega^2}{N}\right). \tag{8b}$$

We are now going to simplify this formula using matrix notation. To this end, we will introduce the matrix $\mathbf{p} = \mathbf{x} \cdot \mathbf{x}^T$, where ^{T} denotes transposition. In other words, $p_{ij} = x_i x_j$. The matrix $\mathbf{p}$ corresponds to a given population structure, $\mathbf{P}(t)$ is the corresponding random matrix. We denote by $\mathrm{Tr}(\mathbf{M})$ the trace of a matrix $\mathbf{M}$, *i.e.*, the sum of its diagonal elements. Finally, we need to remember that the dispersal graph is symmetric (equation (1b)), so that $\mathbf{D}^T = \mathbf{D}$. We are now equipped to rewrite our equation in a more compact way:

$$\overline{\Delta}(\mathbf{x}) = \frac{\mu}{N} \left(\frac{1}{2} - \overline{x} \right) + \frac{\omega}{N^2} \left[\mathsf{b} \, \mathrm{Tr}\left(\mathbf{E} \cdot (\mathbf{p} - \mathbf{D} \cdot \mathbf{D} \cdot \mathbf{p}) \right) + \mathsf{c} \, \mathrm{Tr}\left(\mathbf{p} - \mathbf{D} \cdot \mathbf{D} \cdot \mathbf{p} \right) \right] + \mathcal{O}\left(\frac{\omega^2}{N}\right). \tag{8c}$$

The first term in (8c), scaled by μ, corresponds to mutation, that brings the frequency of social individuals closer to 1/2, its value in the absence of selection. The second term, scaled by ω, corresponds to the effects of selection.

At the moment, to know the change in the *frequency* of social individuals in the population ($\overline{x}$), we need to know the state of the entire population (**x**), *i.e.*, not only how many social individuals there are, but also *where* they are. Instead of considering specific states of the population, we are now going to consider long-term expectations. If we observe our population for a very long time T, these long-term expectations correspond to the proportion of time spent in a give state. We denote by $\xi(\mathbf{x}, \omega, \mu)$ the long-term proportion of time that the population spends in state **x**, given a strength of selection ω and mutation probability μ; the expected state of the population is given by

$$\mathbb{E}_{\omega,\mu}[\mathbf{X}] = \sum_{\mathbf{x} \in \Omega} \mathbf{x}\, \xi(\mathbf{x}, \omega, \mu). \tag{9}$$

Now, because we are considering long-term expectations, $\mathbb{E}_{\omega,\mu}\left[\overline{\Delta}(\mathbf{X})\right] = 0$, and so we have, at the first orders in μ and in ω,

$$
\begin{aligned}
0 = {} & \frac{\mu}{N}\left(\frac{1}{2} - \mathbb{E}_{\omega,0}\left[\overline{X}\right]\right) \\
& + \frac{\omega}{N^2}\left[\mathsf{b}\,\mathrm{Tr}\left(\mathbf{E}\cdot\left(\mathbb{E}_{0,\mu}[\mathbf{P}] - \mathbf{D}\cdot\mathbf{D}\cdot\mathbb{E}_{0,\mu}[\mathbf{P}]\right)\right) + \mathsf{c}\,\mathrm{Tr}\left(\mathbb{E}_{0,\mu}[\mathbf{P}] - \mathbf{D}\cdot\mathbf{D}\cdot\mathbb{E}_{0,\mu}[\mathbf{P}]\right)\right] \\
& + \mathscr{O}\left(\frac{\omega^2}{N}\right) + \mathscr{O}\left(\mu^2\right).
\end{aligned}
$$

Solving for $\mathbb{E}_{\omega,0}\left[\overline{X}\right]$, we obtain

$$
\begin{aligned}
\mathbb{E}_{\omega,0}\left[\overline{X}\right] = {} & \frac{1}{2} + \frac{\omega}{\mu N}\left[\mathsf{b}\,\mathrm{Tr}\left(\mathbf{E}\cdot\left(\mathbb{E}_{0,\mu}[\mathbf{P}] - \mathbf{D}\cdot\mathbf{D}\cdot\mathbb{E}_{0,\mu}[\mathbf{P}]\right)\right)\right. \\
& \left. + \mathsf{c}\,\mathrm{Tr}\left(\mathbb{E}_{0,\mu}[\mathbf{P}] - \mathbf{D}\cdot\mathbf{D}\cdot\mathbb{E}_{0,\mu}[\mathbf{P}]\right)\right] + \mathscr{O}\left(\frac{\omega^2}{\mu}\right) + \mathscr{O}\left(\mu N\right).
\end{aligned}
\tag{10}
$$

This provides a formula for $\mathbb{E}_{\omega,0}\left[\overline{X}\right]$, the long-term frequency of social individuals in the population when mutation is vanishingly small, and with weak mutation (at the first order in ω). It involves $\mathbb{E}_{0,\mu}[\mathbf{p}]$, the long-term expected state of pairs of sites, evaluated in the absence of selection. This is the quantity that we are now going to concentrate on in the next section, after first dealing with the Birth-Death order of events.

3.3 Birth-Death

In the Birth-Death version of the model, the probability of dying is not constant, but depends on the composition of the population. Reorganizing (7), we have

$$\overline{\Delta}(\mathbf{x}) = \frac{1}{N}\sum_{i=1}^{N}\left[\sum_{\substack{j=1\\j\neq i}}^{N} w_{ji}(\mathbf{x})x_i + w_{ii}(\mathbf{x})x_i + \mu\sum_{\substack{j=1\\j\neq i}}^{N} w_{ji}(\mathbf{x})\left(\frac{1}{2} - x_i\right)\right] - \overline{x}. \tag{11}$$

We note that since exactly one individual reproduces at each time step, and since $d_{ii} = 0$ (*i.e.*, there is no reproduction from a given site i to the same site i),

$$\sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} w_{ji}(\mathbf{x}) = 1. \tag{12}$$

If we now use the formulas for w_{ji} et w_{ii} for the Birth-Death life-cycle (equations (5b) and (6)), we obtain:

$$
\begin{aligned}
\overline{\Delta}(\mathbf{x}) = {} & \frac{\omega}{N^2} \sum_{i=1}^{N} \sum_{l=1}^{N} (e_{li}\mathsf{b}x_l x_i - \mathsf{c}x_i) - \frac{\omega}{N^2} \sum_{i=1}^{N} \sum_{k=1}^{N} \frac{1}{N} \left(\sum_{l=1}^{N} e_{lk}\mathsf{b}x_l x_i - \mathsf{c}x_k x_i \right) \\
& - \frac{\omega}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} d_{ij} \left(\sum_{l=1}^{N} e_{lj}\mathsf{b}x_l x_i - \mathsf{c}x_j x_i \right) + \frac{\omega}{N^2} \sum_{i=1}^{N} \sum_{k=1}^{N} \frac{1}{N} \left(\sum_{l=1}^{N} e_{lk}\mathsf{b}x_l x_i - \mathsf{c}x_k x_i \right) \\
& + \frac{\mu}{N} \left(\frac{1}{2} - \overline{x} \right) \\
& - \frac{\omega\mu}{N^2} \sum_{i=1}^{N} \left[\sum_{l=1}^{N} e_{li}\mathsf{b}x_l x_i - \mathsf{c}x_i - \sum_{k=1}^{N} \frac{1}{N} \left(\sum_{l=1}^{N} e_{lk}\mathsf{b}x_l x_i - \mathsf{c}x_k x_i \right) \right] \\
& + \mathcal{O}\left(\frac{\omega^2}{N} \right),
\end{aligned}
\tag{13a}
$$

which simplifies into

$$
\begin{aligned}
\overline{\Delta}(\mathbf{x}) = {} & \frac{\omega}{N^2} \sum_{i=1}^{N} \left(\sum_{l=1}^{N} e_{li}\mathsf{b}x_l x_i - \mathsf{c}x_i \right) - \frac{\omega}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} d_{ij} \left(\sum_{l=1}^{N} e_{lj}\mathsf{b}x_l x_i - \mathsf{c}x_j x_i \right) \\
& + \frac{\mu}{N} \left(\frac{1}{2} - \overline{x} \right) \\
& - \frac{\omega\mu}{N^2} \sum_{i=1}^{N} \left[\sum_{l=1}^{N} e_{li}\mathsf{b}x_l x_i - \mathsf{c}x_i - \sum_{k=1}^{N} \frac{1}{N} \left(\sum_{l=1}^{N} e_{lk}\mathsf{b}x_l x_i - \mathsf{c}x_k x_i \right) \right] \\
& + \mathcal{O}\left(\frac{\omega^2}{N} \right).
\end{aligned}
\tag{13b}
$$

We can rewrite this expression using matrices like we did previously:

$$
\begin{aligned}
\overline{\Delta}(\mathbf{x}) = {} & \frac{\mu}{N} \left(\frac{1}{2} - \overline{x} \right) + \frac{\omega}{N^2} \left[\mathsf{b}\,\mathrm{Tr}\left(\mathbf{E} \cdot (\mathbf{p} - \mathbf{D} \cdot \mathbf{p}) \right) - \mathsf{c}(\mathbf{p} - \mathbf{D} \cdot \mathbf{p}) \right] \\
& - \frac{\omega\mu}{N^2} \sum_{i=1}^{N} \left[\sum_{l=1}^{N} e_{li}\mathsf{b}x_l x_i - \mathsf{c}x_i - \sum_{k=1}^{N} \frac{1}{N} \left(\sum_{l=1}^{N} e_{lk}\mathsf{b}x_l x_i - \mathsf{c}x_k x_i \right) \right] \\
& + \mathcal{O}\left(\frac{\omega^2}{N} \right).
\end{aligned}
\tag{13c}
$$

Taking the long-term expectation of this expression, we obtain

$$
\begin{aligned}
0 = {} & \frac{\mu}{N}\left(\frac{1}{2} - \mathbb{E}_{\omega,0}\left[\overline{X}\right]\right) \\
& + \frac{\omega}{N^2}\left[\mathsf{b}\,\mathrm{Tr}\left(\mathbf{E}\cdot\left(\mathbb{E}_{0,\mu}\left[\mathbf{P}\right] - \mathbf{D}\cdot\mathbb{E}_{0,\mu}\left[\mathbf{P}\right]\right)\right) + \mathsf{c}\,\mathrm{Tr}\left(\mathbb{E}_{0,\mu}\left[\mathbf{P}\right] - \mathbf{D}\cdot\mathbb{E}_{0,\mu}\left[\mathbf{P}\right]\right)\right] \\
& - \frac{\omega\mu}{N^2}\sum_{i=1}^{N}\left[\sum_{l=1}^{N}e_{li}\mathsf{b}\mathbb{E}_{0,0}\left[P_{li}\right] - \mathsf{c}\mathbb{E}_{0,0}\left[P_{ii}\right] - \sum_{k=1}^{N}\frac{1}{N}\left(\sum_{l=1}^{N}e_{lk}\mathsf{b}\mathbb{E}_{0,0}\left[P_{li}\right] - \mathsf{c}\mathbb{E}_{0,0}\left[P_{ki}\right]\right)\right] \\
& + \mathcal{O}\left(\frac{\omega^2}{N}\right) + \mathcal{O}\left(\mu^2\right).
\end{aligned}
\tag{14}
$$

In the limit of zero mutation and zero selection, the population is fixed for each type, and so

$$
\forall (i, j, k, l), \quad \mathbb{E}_{0,0}\left[P_{ij}\right] = \mathbb{E}_{0,0}\left[P_{k,l}\right].
$$

As a result, the term on the third line of equation (14) is null. Solving for $\mathbb{E}_{\omega,0}\left[\overline{X}\right]$, we finally obtain

$$
\begin{aligned}
\mathbb{E}_{\omega,0}\left[\overline{X}\right] = {} & \frac{1}{2} + \frac{\omega}{\mu N}\left[\mathsf{b}\,\mathrm{Tr}\left(\mathbf{E}\cdot\left(\mathbb{E}_{0,\mu}\left[\mathbf{P}\right] - \mathbf{D}\cdot\mathbb{E}_{0,\mu}\left[\mathbf{P}\right]\right)\right)\right. \\
& \left. + \mathsf{c}\,\mathrm{Tr}\left(\mathbb{E}_{0,\mu}\left[\mathbf{P}\right] - \mathbf{D}\cdot\mathbb{E}_{0,\mu}\left[\mathbf{P}\right]\right)\right] + \mathcal{O}\left(\frac{\omega^2}{\mu}\right) + \mathcal{O}\left(\mu N\right).
\end{aligned}
\tag{15}
$$

Equation (15) (Birth-Death) is almost the same as equation (10) (Death-Birth), except for the terms that correspond to competition ($-\mathbf{D}\cdot\mathbb{E}_{0,\mu}\left[\mathbf{P}\right]$ in the former case, $-\mathbf{D}\cdot\mathbf{D}\cdot\mathbb{E}_{0,\mu}\left[\mathbf{P}\right]$ in the latter). These terms come from the different denominators of w_{ji} before linearization (equations (5a) and (4a)). The difference can be further interpreted as follows.

With the Birth-Death updating, a focal individual competes with its direct neighbors (hence the $\mathbf{D}$ factor) for new sites: if a neighbor is chosen to reproduce, a focal individual is a risk of being killed. An increase (due to social interactions) in the fecundity of a direct neighbors therefore has indirect detrimental consequences on the focal individual; similarly, a decrease (the cost of sociality) in the fecundity of a focal individual has indirect beneficial consequences on its direct neighbors.

With the Death-Birth updating, a focal individual competes for new sites with its neighbors' neighbors (hence the $\mathbf{D}\cdot\mathbf{D}$ factor): if a direct neighbor is killed, a focal individual competes with that neighbor's neighbors to fill the newly emptied site. In this case, the "competition radius" is wider, two dispersal steps away, instead of one in the Birth-Death case.

Equations (10) and (15) are still implicit, because we need to evaluate the long-term expected state of pairs of sites in the absence of selection, *i.e.*, $\mathbb{E}_{0,\mu}\left[\mathbf{P}\right]$.

4 Long-term expected state of pairs of sites

4.1 Conditional expected change of a pair of sites

We denote by $\tilde{\Delta}_{ij}^{0}(\mathbf{x})$ the expected change in the state of the (i, j) pair of sites over one time step, given that the population is in state $\mathbf{x}$, in the absence of selection

(hence the "^{0}"). In this case, each individual of the pair is replaced with probability $1/N$; we have

$$\forall j \neq i, \tilde{\Delta}_{ij}^0(\mathbf{x}) = \frac{x_i}{N}\sum_k d_{kj}\left(x_k\left(1-\frac{\mu}{2}\right)+(1-x_k)\frac{\mu}{2}\right)$$
$$+ \frac{x_j}{N}\sum_k d_{ki}\left(x_k\left(1-\frac{\mu}{2}\right)+(1-x_k)\frac{\mu}{2}\right)$$
$$+ \left(1-\frac{2}{N}\right)x_i x_j - x_i x_j.$$
$$= (1-\mu)\sum_k \frac{d_{kj}x_i x_k + d_{ki}x_j x_k}{N} \quad + \frac{\mu}{2N}(x_i+x_j) - \frac{2}{N}x_i x_j. \qquad (16)$$

When $i=j$, we have

$$\tilde{\Delta}_{ii}^0(\mathbf{x}) = \left(1-\frac{1}{N}\right)x_i + \frac{1}{N}\sum_k d_{ki}\left(x_k\left(1-\frac{\mu}{2}\right)+(1-x_k)\frac{\mu}{2}\right) - x_i$$
$$= \frac{\mu}{2N} - \frac{x_i}{N} + \frac{1}{N}\sum_k d_{ki}x_k(1-\mu). \qquad (17)$$

4.2   Expectation of the state of a pair of sites at time t

In the previous section, we have been considering long-term expectations, *i.e.*, the fraction of the time spent in a given state over a long time scale. Here, we are going to consider the sequence of population states as a function of time. This sequence depends on the initial state of the population. Mutation being rare, we consider that the population is initially fixed for one type (it is the social type with probability $1/2$). We denote by $\mathbb{E}_{0,\mu,t}[\mathbf{X}]$ the expectation of the state of the population in the absence of selection, with mutation parameter μ, and at time t, given the scenario that we just described.

Since, in expectation, the population looks the same from every site, it does not matter which label we gave to a specific site. As a result, all sites have the same expected state, which we denote by $f(t)$. For the same reason, $\sum_{k=1}^N d_{ik}\mathbb{E}_{0,\mu,t}[P_{ki}]$ is the same for all sites i (recall that the random matrix $\mathbf{P}$ refers to pairs of sites; $\mathbf{P} = \mathbf{X}\cdot\mathbf{X}^T$).

To proceed further, we rewrite (16) in matrix form, and take its expectation:

$$\mathbb{E}_{0,\mu,t}\left[\tilde{\mathbf{\Delta}}^0(\mathbf{X})\right] = \frac{1-\mu}{N}\left(\mathbb{E}_{0,\mu,t}[\mathbf{P}]\cdot\mathbf{D} + \mathbf{D}\cdot\mathbb{E}_{0,\mu,t}[\mathbf{P}]\right) + \frac{\mu}{N}f(t)\mathbb{1}_{N\times N} - \frac{2}{N}\mathbb{E}_{0,\mu,t}[\mathbf{P}] + \mathbf{L}(t), \qquad (18)$$

where $\mathbf{L}$ is a diagonal matrix that ensures that the diagonal elements of $\mathbb{E}_{0,\mu,t}[\mathbf{P}]$, corresponding to the expected states of single sites, are all equal to $f(t)$. In other words, we can write $\mathbf{L}(t) = g(t,\mu)\mathbf{I}_N$.

We are now going to show recursively that at all times, $\mathbb{E}_{0,\mu,t}[\mathbf{P}]\cdot\mathbf{D} = \mathbf{D}\cdot\mathbb{E}_{0,\mu,t}[\mathbf{P}]$, that is, that the matrix of expected states of pairs of sites commutes with the adjacency matrix of the dispersal graph.

At time $t=0$, if the population is fixed for the social type, which occurs with probability $1/2$, the matrix of pairs is a N-by-N matrix of ones ($\mathbf{p} = \mathbb{1}_{N\times N}$); otherwise, the population is fixed for the non social type, and the matrix of pairs contains only zeros. Overall, $\mathbb{E}_{0,\mu,t=0}[\mathbf{P}] = \frac{1}{2}\mathbb{1}_{N\times N}$. Since the graph is regular (equation (1a)), $\mathbb{E}_{0,\mu,t=0}[\mathbf{P}]$ commutes with $\mathbf{D}$ ($\mathbb{1}_{N\times N}\cdot\mathbf{D} = \mathbf{D}\cdot\mathbb{1}_{N\times N} = \mathbb{1}_{N\times N}$).

Now assuming that $\mathbb{E}_{0,\mu,t}[\mathbf{P}]\cdot\mathbf{D} = \mathbf{D}\cdot\mathbb{E}_{0,\mu,t}[\mathbf{P}]$ holds, we have

$$\mathbb{E}_{0,\mu,t+1}[\mathbf{P}]\cdot\mathbf{D} = \mathbb{E}_{0,\mu,t}[\mathbf{P}]\cdot\mathbf{D} + \mathbb{E}_{0,\mu,t}\left[\tilde{\boldsymbol{\Delta}}^0(\mathbf{X})\right]\cdot\mathbf{D}$$

$$= \mathbf{D}\cdot\mathbb{E}_{0,\mu,t}[\mathbf{P}] + \frac{1-\mu}{N}2\,\mathbb{E}_{0,\mu,t}[\mathbf{P}]\cdot\mathbf{D}\cdot\mathbf{D} + \frac{\mu}{N}f(t)\mathbb{1}_{N\times N} - \frac{2}{N}\mathbb{E}_{0,\mu,t}[\mathbf{P}]\cdot\mathbf{D} + g(t,\mu)\mathbf{D}$$

$$= \mathbf{D}\cdot\mathbb{E}_{0,\mu,t+1}[\mathbf{P}].$$

$\square$

4.3 Long-term expectation

At last, we can consider the long-term expectations *i.e.*, the fraction of time spent in a given state:

$$\mathbb{E}_{0,\mu}\left[\tilde{\boldsymbol{\Delta}}^0(\mathbf{X})\right] = \lim_{T\to\infty}\frac{1}{T}\int_0^T\mathbb{E}_{0,\mu,t}\left[\tilde{\boldsymbol{\Delta}}^0(\mathbf{X})\right]dt.$$

Again we have $\mathbb{E}_{0,\mu}\left[\tilde{\boldsymbol{\Delta}}^0(\mathbf{X})\right] = 0$; solving for $\mathbb{E}_{0,\mu,t}[\mathbf{P}]$, from equation (18) we obtain

$$\mathbb{E}_{0,\mu}[\mathbf{P}] = (1-\mu)\mathbb{E}_{0,\mu,t}[\mathbf{P}]\cdot\mathbf{D} + \frac{\mu}{2}f^*\mathbb{1}_{N\times N} + \frac{N}{2}g^*(\mu)\mathbf{I}_N. \tag{19}$$

where f^* (resp. g^*) is the long-term average of f (resp. g). Given that $f(t)$ is the expected state of a single site at time t, in the absence of selection the two types are equiprobable and we have $f^* = \frac{1}{2}$. The other function was introduced to ensure that for all i and t, $\mathbb{E}_{0,\mu,t}[\mathbf{P}] = f(t)$. We note that when $\mu \to 0$, $\mathbb{E}_{0,\mu}[\mathbf{P}] \to f^*\mathbb{1}_{N\times N}$, which implies that $g^*(0) = 0$; at the first order in μ, (19) becomes

$$\mathbb{E}_{0,\mu}[\mathbf{P}] = \mathbb{E}_{0,\mu,t}[\mathbf{P}]\cdot\mathbf{D} + \mu\left(-\frac{1}{4}\mathbb{1}_{N\times N} + \frac{N}{2}h^*\mathbf{I}_N\right) + \mathcal{O}\left(\mu^2\right), \tag{20a}$$

and

$$h^* = \left.\frac{dg}{d\mu}\right|_{\mu=0}. \tag{20b}$$

In particular, (20a) implies that for all sites i,

$$\sum_{j=1}^N\mathbb{E}_{0,\mu}\left[P_{ij}\right] = \sum_{j=1}^N\sum_{k=1}^N\mathbb{E}_{0,\mu}[P_{ik}]\,d_{kj} - N\frac{\mu}{4} + N\frac{\mu}{2}h^* + \mathcal{O}\left(\mu^2\right)$$

$$= \sum_{k=1}^N\mathbb{E}_{0,\mu}[P_{ik}] - N\frac{\mu}{4} + N\frac{\mu}{2}h^* + \mathcal{O}\left(\mu^2\right)$$

so that

$$h^* = \frac{1}{2}. \tag{21}$$

We end up with

$$\mathbb{E}_{0,\mu}[\mathbf{P}] = \mathbb{E}_{0,\mu,t}[\mathbf{P}]\cdot\mathbf{D} + \mu\left(-\frac{1}{4}\mathbb{1}_{N\times N} + \frac{N}{4}\mathbf{I}_N\right) + \mathcal{O}\left(\mu^2\right). \tag{22}$$

Multiplying both sides by $\mathbf{D}$, we obtain

$$\mathbb{E}_{0,\mu,t}[\mathbf{P}]\cdot\mathbf{D} = \mathbb{E}_{0,\mu,t}[\mathbf{P}]\cdot\mathbf{D}\cdot\mathbf{D} + \mu\left(-\frac{1}{4}\mathbb{1}_{N\times N} + \frac{N}{4}\mathbf{D}\right) + \mathcal{O}\left(\mu^2\right),$$

and so

$$\mathbb{E}_{0,\mu}[\mathbf{P}] = \mathbb{E}_{0,\mu,t}[\mathbf{P}]\cdot\mathbf{D}\cdot\mathbf{D} + \mu\left(-\frac{1}{2}\mathbb{1}_{N\times N} + \frac{N}{4}\mathbf{D} + \frac{N}{4}\mathbf{I}_N\right) + \mathcal{O}\left(\mu^2\right). \tag{23}$$

5 Long-term expectation of the proportion of social individuals in the population

5.1 Death-Birth

We have a few calculations left before being able to evaluate equation (10), using the relationship given in equation (23). We recall that since there are no self-loops on the dispersal graph, $\text{Tr}(\mathbf{D}) = 0$. For the factor of the cost parameter c, we evaluate

$$\text{Tr}\left(\mathbb{E}_{0,\mu}[\mathbf{P}] - \mathbb{E}_{0,\mu,t}[\mathbf{P}] \cdot \mathbf{D} \cdot \mathbf{D}\right) = \frac{\mu}{4} N(N-2) + \mathscr{O}\left(()\,\mu^2\right), \tag{24}$$

and for the factor of the benefit parameter b, we evaluate

$$\text{Tr}\left(\mathbf{E} \cdot \left(\mathbb{E}_{0,\mu}[\mathbf{P}] - \mathbb{E}_{0,\mu,t}[\mathbf{P}] \cdot \mathbf{D} \cdot \mathbf{D}\right)\right) = \frac{\mu}{4}\left(-2\sum_{i=1}^{N}\sum_{k=1}^{N} e_{ik} + N\sum_{i=1}^{N}\sum_{k=1}^{N} d_{ik} e_{ki} + 0\right) + \mathscr{O}\left(\mu^2\right) \tag{25}$$

If we further assume that all individuals are involved in the same number of social interactions (giving and receiving), after scaling of the interaction graph we have

$$\forall i, \sum_{k=1}^{N} e_{ki} = \sum_{k=1}^{N} e_{ik} = 1; \tag{26}$$

then (25) becomes

$$\text{Tr}\left(\mathbf{E} \cdot \left(\mathbb{E}_{0,\mu}[\mathbf{P}] - \mathbb{E}_{0,\mu,t}[\mathbf{P}] \cdot \mathbf{D} \cdot \mathbf{D}\right)\right) = \frac{\mu}{4} N\left(\sum_{i=1}^{N}\sum_{k=1}^{N} d_{ik} e_{ki} - 2\right) + \mathscr{O}\left(\mu^2\right). \tag{27}$$

We can now go back to (10). In the limit of weak selection and rare mutation, the expected proportion of social individuals in the population ($\mathbb{E}_{\omega,0}\left[\overline{X}\right]$) is greater than what it would be in the absence of selection ($\frac{1}{2}$), *i.e.*, the social type is favored by selection, whenever

$$b\left(\sum_{i=1}^{N}\sum_{k=1}^{N} d_{ik} e_{ki} - 2\right) > c(N-2). \tag{28}$$

For instance, on a regular grid with n neighbors, social interactions occurring among the neighbors, this condition reduces to

$$b\frac{N}{n} > c(N-2),$$

which can be approximated as $b/c > n$ when the size N of the population is very large.

5.2 Birth-Death

To evaluate equation (15), we use equation (22). Starting with the factor for the cost c, we find

$$\text{Tr}\left(\mathbb{E}_{0,\mu,t}[\mathbf{P}] - \mathbb{E}_{0,\mu,t}[\mathbf{P}] \cdot \mathbf{D}\right) = \frac{\mu}{4} N(N-1). \tag{29}$$

11/12

For the benefits, we evaluate

$$\mathrm{Tr}\left(\mathbf{E}\cdot(\mathbb{E}_{0,\mu,t}\left[\mathbf{P}\right]-\mathbb{E}_{0,\mu,t}\left[\mathbf{P}\right]\cdot\mathbf{D})\right)=\mu\left(-\frac{1}{4}\sum_{i=1}^{N}\sum_{k=1}^{N}e_{ki}+0\right)+\mathcal{O}\left(\mu^2\right) \tag{30}$$

If we again assume that all individuals give or receive the same total amount of interactions (*i.e.*, that equation (26) holds), this simplifies into

$$\mathrm{Tr}\left(\mathbf{E}\cdot(\mathbb{E}_{0,\mu,t}\left[\mathbf{P}\right]-\mathbb{E}_{0,\mu,t}\left[\mathbf{P}\right]\cdot\mathbf{D})\right)=-\frac{\mu N}{4}. \tag{31}$$

Either way, this factor is negative: the benefits of being close to related individuals is more than compensated by the detriments of competing against relatives.

Going back to equation (15) – In the limit of weak selection and rare mutation, the expected proportion of social individuals in the population ($\mathbb{E}_{\omega,0}\left[\overline{X}\right]$) would be greater than what it is be in the absence of selection ($\frac{1}{2}$), *i.e.*, the social type is favored by selection, whenever

$$-\mathsf{b}-\mathsf{c}(N-1)>0, \tag{32}$$

which is not possible when $\mathsf{b}>0$ (actual benefits) and $\mathsf{c}>0$ (actual cost paid).

Bibliography

Débarre, F., Hauert, C. & Doebeli, M. (2014). Social evolution in structured populations. *Nature Communications*, 5.

Grafen, A. & Archetti, M. (2008). Natural selection of altruism in inelastic viscous homogeneous populations. *Journal of Theoretical Biology*, 252, 694–710.

Nowak, M.A. (2006). *Evolutionary dynamics*. Harvard University Press.

Ohtsuki, H., Hauert, C., Lieberman, E. & Nowak, M.A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441, 502–505.

Rousset, F. (2004). *Genetic Structure and Selection in Subdivided Populations*. vol. 40. Princeton University Press.

Taylor, P.D., Day, T. & Wild, G. (2007a). Evolution of cooperation in a finite homogeneous graph. *Nature*, 447, 469–472.

Taylor, P.D., Day, T. & Wild, G. (2007b). From inclusive fitness to fixation probability in homogeneous structured populations. *Journal of Theoretical Biology*, 249, 101–110.