Strategy Evolution on Graphs

Rafael Montoya Robledo
Graduate Project
Universidad de los Andes

2015

Contents

Chapter 1

Introduction

This project is meant for an audience in a graduate level either in mathematics or biology. This means that the reader should have a basic formation in mathematics, including basic probability, advanced calculus and basic game theory. Also, it will be useful for the reader to have a basic intuition in biomathematical modelling. The purpose of this project is to delve into the topic of evolutionary strategies on graphs. Mainly, we aim to fully understand the paper written in 2006 by Martin Nowak *et al* titled "A Simple Rule for the Evolution of Cooperation on Graphs and Social Networks". This paper is the base for an article also published by the same authors in Nature Magazine. The paper has an important deal of mathematical content in the areas of game theory, probability theory, dynamical systems and real analysis. Our goal is to fully understand the content and to clear any results that are not fully explained due to the conciseness of the text. We want to fill in all the details and steps in the paper. This will be done in order to attempt to generalize some of these results.

In order to fulfil this goal, we will first give some preliminaries on dynamical systems, showing some of the most important classic results in the area. Then we will jump to game theoretical areas. Mainly we will look to some of the most important concepts in population dynamics. We will focus on replicator dynamics and evolutionary stable strategies. After that, we will observe a deduction of some mathematical models in genetics, that will be very important for the understanding of the paper. Mainly we are going to deduce the backward Kolmogorov equation, which is an ordinary differential equation that can be used to model fixation probabilities, in the context of evolving populations.

After that, we embrace in filling up all the details of the paper. This will take a complete chapter since the paper mainly names the results but most of the calculations, deductions and key ideas are omitted. It is important to clarify that we will not fully justify the soundness of the mathematical calculations that are, most of the time, assymptotical approximations. At the end of the chapter we will present some Monte Carlo simulations that will help us verify the validity of the obtained results. Since most of the obtained results are assymptotical, but do not specify the speed of convergence, we will like to give approximate values on the different parameters upon which the rules work.

Finally, we will try to generalize a similar model to the backward Kolmogorov equation but in a two variable case, which will yield a second order PDE. This will be intended for developing some further theory on evolutionary strategies on graphs with more than two possible strategies.

Chapter 2

Preliminaries

For the purpose of this project we are going to use tools from two areas of mathematics. These are: dynamical systems (ordinary differential equations on a time parameter) and game theory (mainly evolutionary games). This will let us pass on to our main goal that is to work in evolutionary games on graphs.

2.1 On Dynamical Systems

We assume that the reader is quite familiar with the basics of the theory of dynamical systems. Therefore we shall not delve deeply in the topic. We will only name some of the most important terminology and briefly explain the main theorems and how they are used. Proofs of the theorems are quite technical and will be omitted most of the time. We also intend to define some main terminology that will be of great use throughout the text. We will also deduce some important models that will have applications in the topic of evolutionary games on graphs.
So, let us start with the basics:

Definition 1. A **dynamical system** defined on $\mathbb{R}^d$ is a system of equations of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

where $\dot{\mathbf{x}}$ is the derivative of $\mathbf{x}(t)$ with respect to time t and $\mathbf{f} : A \subseteq \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is a continuously differentiable function and A is open. In this setting, an **equilibrium point** of the system is a point $\mathbf{x}$ where $\mathbf{f}(\mathbf{x}) = 0$.

In the case where the dynamical system is linear, i.e. can be written in the form $\dot{\mathbf{x}} = A\mathbf{x}$ where A is a $d \times d$ matrix, we basically know everything about the system. In this situation we just need to find the Jordan normal form, since the solution to the system will be given by

$$\mathbf{x}(t) = e^{At}\mathbf{x}(0).$$

This theory for linear systems is usually covered in a basic differential equation calculus class and is mainly a work on linear algebra. On the other hand, we are interested in the general case where $f(\mathbf{x})$ is not a linear function of $\mathbf{x}$. In real life applications, the dynamical models are not generally linear. In this case the work becomes much more complicated, to the point that there is no general method for solving or understanding the system. Nevertheless, there are two theorems that are quite useful under certain circumstances.

2.1.1 Linearization

Usually, a dynamical system which is not linear is intractable. Nevertheless, there is an intuition based on Taylor's theorem that can help us understand such system. When analysing a system, we are mainly concerned with the equilibrium points, and their behaviour as sinks, sources or saddles. Let $\mathbf{z}$ be an equilibrium point (so that $\mathbf{f}(\mathbf{z}) = 0$). Then, we can build a Jacobian matrix based on $\mathbf{f}$ and evaluate it at $\mathbf{z}$. Call this matrix A and consider the system

$$\dot{\mathbf{y}} = A\mathbf{y}.$$

which we know how to solve. From here, and using the previous notation, we can enunciate the Hartman-Grobman theorem:

Theorem 1. *Let $\mathbf{z}$ be an equilibrium point and let A be it's corresponding Jacobian matrix. Suppose that none of the eigenvalues of A have a vanishing real part, then there exists a neighbourhood U and a homeomorphism $\mathbf{h}$ from U to some neighbourhood V of the origin $\mathbf{0}$ such that $\mathbf{y} = \mathbf{h}(\mathbf{x})$ implies $\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t))$ for all $t \in \mathbb{R}$ with $\mathbf{x}(t) \in U$.*

We must recall that a homeomorphism is a continuous map with a continuous inverse. Now, what this theorem says is that, locally, the orbits of the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ around $\mathbf{z}$ look like those of $\dot{\mathbf{y}} = A\mathbf{y}$ near $\mathbf{0}$. From this theorem, we can interpret the stability of the equilibrium points, and classify them. Basically we can understand the system's local main features by applying all the theory of linear systems, which is fully understood.

2.1.2 The Lyapunov Criterion

Another important tool is based in this subsection's theorem. The theorem uses the notation $\dot{V}(\mathbf{x})$. In this context, $\dot{V}(\mathbf{x})$ is the derivative of the function $V : E \longrightarrow \mathbb{R}$, where E is an open set, along a solution. i.e. $\dot{V}(x) = \frac{d}{dt}V(x(t))$. This results, by the chain rule, in the following:

$$\dot{V}(\mathbf{x}) = DV(\mathbf{x})f(\mathbf{x})$$

where DV stands for the gradient of V. Now the theorem:

Theorem 2. *Let E be an open subset of $\mathbb{R}^n$ containing the point $\mathbf{z}$. Suppose the $\mathbf{f} \in C^1(E)$ and that $\mathbf{z}$ is an equilibrium point. Suppose that there exists a real valued function $V \in C^1(E)$ that satisfies $V(\mathbf{z}) = 0$ and $V(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{z}$. Then:*

- *if* $\dot{V}(\mathbf{x}) \leq 0$ *for all* $\mathbf{x} \in E$*,* $\mathbf{z}$ *is stable.*
- *if* $\dot{V}(\mathbf{x}) < 0$ *for all* $\mathbf{x} \in E - \{\mathbf{z}\}$*,* $\mathbf{z}$ *is asymptotically stable.*
- *if* $\dot{V}(\mathbf{x}_0) > 0$ *for all* $\mathbf{x} \in E - \{\mathbf{z}\}$*,* $\mathbf{z}$ *is unstable.*

This is a very important tool in the sense that it allows us to determine the behaviour of an equilibrium point by means of suitable functions. This means that by looking at a dynamical system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$$

and by cleverly building a function V with the hypotheses requirements, one can immediately determine the nature of the system at the desired point. This type of function is known as the **Lyapunov function**. The selection of such function can be quite an excruciating task. Nevertheless, the derivative $\dot{V}(\mathbf{x})$ can be calculated no matter that explicit solutions are not found. Keeping this in mind, it will be useful to observe a simple example when the criterion can be used.

Example 1. Consider the system:

$$\dot{x}_1 = -x_2^3$$

$$\dot{x}_2 = x_1^3$$

The origin is an equilibrium point of the system. Now, the Jacobian matrix is given by:

$$J = \begin{bmatrix} 0 & -3x_2 \\ 3x_1 & 0 \end{bmatrix}.$$

By finding the eigenvalues of the linearized system, one can observe that the origin is also non hyperbolic, since none of the eigenvalues have a positive real part. If it were hyperbolic, we would be done studying the local behaviour by using the Hartman-Grobman theorem. Now, we consider the following Lyapunov function:

$$V(\mathbf{x}) = x_1^4 + x_2^4.$$

With this function we get that $\dot{V}(\mathbf{x}) = 4x_1^3\dot{x}_1 + 4x_2^3\dot{x}_2 = 0$. Hence, the origin is a stable point but it is not asymptotically stable since the solution curves for the system lie in the closed curves

$$x_1^4 + x_2^4 = c^2$$

where c is a constant. It is important to see that if the linearization procedure had been implemented one would have obtained two 0 eigenvalues, which suggests to us that the point is stable, but hides the fact that the solutions are periodical orbits. This shows the importance of Lyapunov criterion, since it allows, sometimes, for a deeper understanding of the system than with the linearization approach.

The two previous theorems help us understand the local behaviour of equilibrium points of a dynamical system. Still, there is a different kind of stability which is linked to game theory and which we will see later.

2.1.3 On $\omega - limit$ Sets

For the purpose of this work, we are going to include some not so basic terminology that will be needed later. Since the asymptotic features of a system are going to be of great importance, we need to define the following:

Definition 2. The $\omega - limit$ of $\mathbf{x}(t)$ is defined as $\omega(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^n : \mathbf{x}(t_k) \to \mathbf{y}\; k \to \infty \text{ for a sequence } (t_k)_k\}$.

Some important features about the $\omega - limit$ set of $\mathbf{x}(t)$ are that:

- it is closed.
- it can be empty.
- its points have the property that all their neighbourhoods keep being visited by the solution $\mathbf{x}(t)$ after arbitrarily long periods of time.
- It is invariant: if $\mathbf{x} \in \omega(\mathbf{x})$ then $\mathbf{x}(t) \in \omega(\mathbf{x})$ for all $t \in \mathbb{R}$..

Also, it is important to say that there is a version of Lyapunov's theorem in terms of $\omega - limits$. Meaning that the $\omega - limits$ can be classified in terms of Lyapunov functions.

Theorem 3. *Let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ ODE defined on $U \subset \mathbb{R}^d$. Let $V : U \to \mathbb{R}$ be C^1. If for some solution $\mathbf{x}(t)$, $\dot{V}(\mathbf{x}(t)) \geq 0$ then $\omega(\mathbf{x}) \cap U$ is contained in the set $\{\mathbf{x} \in U : \dot{V}(\mathbf{x}) = 0\}$.*

In the context of example (1) the curves

$$x_1^4 + x_2^4 = c^2$$

are the $\omega - limits$ of the system and for the same Lyapunov function, we obtain that $\omega(\mathbf{x}) \cap U$ is contained in the set $\{\mathbf{x} \in U : \dot{V}(\mathbf{x}) = 0\}$ which in this case is all $\mathbb{R}^2$, which implies that the solutions are periodic.

2.2 Fast and Slow Dynamics

Dynamical systems can have fast or slow dynamics. Meaning that the system can converge, in one or more of its variables, to a stable manifold. For instance one of its variables can stabilize pretty fast in comparison to the others, so that the system can be thought of as with less variables. This is a reasonable approach when dealing with multivariate complicated dynamical systems. In this section we will look towards two main examples that reflect the concepts of fast and slow dynamics. The importance of this concepts will be highly relevant in the next chapter, were the considered system has a variable with a fast dynamic which will be therefore ignored.

In many fields, dynamical systems appear. When a dynamical system can be expressed as

$$\dot{x} = f(x, y) \tag{2.1}$$
$$\dot{y} = \epsilon g(x, y) \tag{2.2}$$

where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, and ϵ is very small we say that $x = (x_1, ..., x_n)$ are the **fast dynamic** variables, whereas $y = (y_1, ..., y_m)$ are the **slow dynamic** variables. This is way easier to understand when $m = n = 1$. The key aspect here will be that

$$|\dot{y}| << |\dot{x}|.$$

The effect that the ϵ is having on time makes the convergence of y to go much slower. In such case we can assume that the system will be given in terms of $\dot{y}$ and that $f(x, y) = 0$. The relation $f(x, y) = 0$, in the case $n = m = 1$ may let us find y as a function of x which may help us solve for y in the system.

Now we are going to show a famous model that exemplifies this situations, but first, notice that if we let $\tau = \epsilon t$ the previous system (2.1) can be written as

$$\epsilon \frac{dx}{d\tau} = f(x, y) \tag{2.3}$$
$$\frac{dy}{d\tau} = g(x, y). \tag{2.4}$$

2.2.1 FitzHughNagumo model (FHN)

Consider the system described by

$$\dot{v} = v - \frac{v^3}{3} - w + I_{\text{ext}}$$
$$\tau \dot{w} = v + a - bw$$

Where I_{ext} is an external stimulus that can be fixed in order to cause different stabilizing reactions on v and w. This is known as the **FitzHughNagumo model (FHN).** Notice that it has the same form as system (2.3.) It is an example of a fast-slow dynamic model that appears commonly in neuroscience and in electronics.

Now we are ready to pass on to the next section that deals with game theory and its relationship with dynamical systems.

2.3 On Game Dynamics

In this section we begin by introducing some of the basic concepts of dynamical game theory in population dynamics. Again, it is assumed that the reader has some knowledge about the basic concepts of game theory. That includes the basic concepts of cooperative games, players, strategies, pure and mixed strategies, payoff matrices, expected payoffs, Nash equilibrium, etc.. We aim to link some game theoretic tools with dynamical systems including Nash equilibrium and evolutionary stable strategy.

We will try to do this in a way in which it results intuitive, or at least with some motivational background. This intuition will come basically from biology examples. This means that we will be dealing with populations, which are groups of individuals which interact during time periods. Most of this section is based on the second chapter of [1]. It is basically a summary that includes only some of the proofs.

The importance of these concepts lie in the fact that an evolutionary stable population, which we will define soon, is proof against invading minorities. That means it will be more likely, for an population that has an evolutionary stable strategy, to stay still. By staying still (not changing their strategy) they will be stronger against invading minorities with different strategies. Throughout this section, some times the terms; strategy, behaviour and type will mean the same.

2.3.1 Evolutionary Stable Strategies

First we must define the concept of evolutionary stability.

Definition 3. Let I be a type of individual or type of behaviour and Q a population composition. This means that in a population there are types $I_1, I_2, ..., I_n$ with a composition proportional to $Q = (q_1, ..., q_n)$ such that $q_1 + ... + q_n = 1$. Let $W(I, Q)$ denote the payoff of the individual of type I in a population of composition Q. A population consisting of $I-$types is **evolutionary stable** if for all $J \neq I$ and for all ϵ sufficiently small

$$W(J, Q_\epsilon) < W(I, Q_\epsilon) \tag{2.5}$$

holds, where $Q_\epsilon = (0, 0, ..., \underbrace{1-\epsilon}_{I\,position}, 0, ..., 0, \underbrace{\epsilon}_{J\,position}, 0, ..., 0)$.

What the definition is saying is that a type of behaviour is evolutionary stable if whenever all the members of the population adopt such behaviour, no dissident behaviour could invade the population under the influence of natural selection. In this context the payoff is what encourages individuals to change a behaviour, so if the payoff can only decrease by changing the actual behaviour, the composition tends to remain still in his strategy. In the context of the previous definition, the ϵ refers to the frequency of the $J-$types. In the following, when we write $W(X,Y)$ we will mean that the composition has all zero entries except the one that represents Y which has a one.

Now, it is reasonable to assume that W is continuous in the second component, since a small change in the population composition will have a small effect in the payoff of the individual behaviour. So, by letting $\epsilon \to 0$ we obtain from (2.5) that

$$W(J,I) \leq W(I,I).$$

This shows that an individual with an evolutionary stable behaviour I has the best payoff when the population composition has all its weight on that strategy. Also, that an invader in a evolutionary stable population has a worse payoff compared to the members of the stable population. The converse is not true in general, that is $W(J,I) \leq W(I,I)$ does not imply (2.5).

2.3.2 Normal Form Games

For the following, a strategy can be expressed as a point $\mathbf{p}$ in the simplex:

$$S_n = \left\{ \mathbf{p} \in \mathbb{R}^n : p_i \geq 0, \sum_{i=1}^{n} p_i = 1 \right\} \tag{2.6}$$

In the previous background we can introduce a **normal form game** as the game the payoff for a strategy is linear. If we assume that the game involves two players only, we can denote by u_{ij} the payoff of a player playing the pure strategy R_i against the other player playing the strategy R_j. Here, the matrix $U = (u_{ij})$ is the payoff matrix. Then, if a player plays strategy R_i, his expected payoff against a player with a strategy $\mathbf{q} \in S_n$ is given by $(U\mathbf{q})_i = \sum_{i=1}^{n} u_{ij} q_j$.Similarly, the payoff of a player with strategy $\mathbf{p}$ against a player with strategy $\mathbf{q}$ is given by

$$\mathbf{p}^t U \mathbf{q} = \sum_{ij} u_{ij} p_i q_j \tag{2.7}$$

With this in mind we recall two important definitions

Definition 4. A strategy $\mathbf{q}$ is said to be a **Nash equilibrium** if

$$\mathbf{p}^t U \mathbf{q} \leq \mathbf{q}^t U \mathbf{q}$$

for all $\mathbf{p} \neq \mathbf{q}$.

Definition 5. A strategy $\mathbf{q}$ is said to be a **strict Nash equilibrium** if

$$\mathbf{p}^t U \mathbf{q} < \mathbf{q}^t U \mathbf{q}$$

for all $\mathbf{p} \neq \mathbf{q}$.

Every normal form game admits at least one Nash equilibrium, which is basically a strategy that is a best reply to itself.
In general, games involve a large population of players, not just two. This is what pushes the need of a generalization of the previous concept to multi-player games and the definition of the evolutionary stable strategies, which has a lot to do with (2.5).

2.3.3 Evolutionary Stable Strategies Revisited

If we consider a large population of players who randomly encounter and play, and we know that if there is a strict Nash equilibrium, by penalizing any different strategy, we can assure no dissident behaviour to spread. In this case, stability would have been achieved. Nevertheless, we cannot always guarantee the existence of a strict Nash equilibrium. First, we define the following

Definition 6. A strategy $\hat{\mathbf{p}} \in S_n$ is said to be **evolutionary stable** if for all $\mathbf{p} \in S_n$ with $\mathbf{p} \neq \hat{\mathbf{p}}$

$$\mathbf{p}^t U(\epsilon\mathbf{p} + (1-\epsilon)\hat{\mathbf{p}}) < \hat{\mathbf{p}}^t U(\epsilon\mathbf{p} + (1-\epsilon)\hat{\mathbf{p}}) \tag{2.8}$$

holds for all $\epsilon > 0$ that are sufficiently small, that is smaller than a number $\bar{\epsilon}(\mathbf{p}) > 0$, which is called **invasion barrier.**

Notice that (2.8) can be rewritten as

$$(1-\epsilon)(\hat{\mathbf{p}}^t U\hat{\mathbf{p}} - \mathbf{p}^t U\hat{\mathbf{p}}) + \epsilon(\hat{\mathbf{p}}^t U\mathbf{p} - \mathbf{p}^t U\mathbf{p}) > 0$$

Going back to (2.5) where we had behaviours I and J, we can thought of these as corresponding to strategies $\mathbf{p}$ and $\hat{\mathbf{p}}$ respectively, so that the similarity between this concept and the one defined previously can be observed. Continuing, it can be shown, by letting $\epsilon \longrightarrow 0$ and setting $\epsilon = 1 respectively$, that $\hat{\mathbf{p}}$ is an evolutionary stable strategy(ESS) if and only if the two following conditions hold:

1. Equilibrium condition

$$\hat{\mathbf{p}}^t U\hat{\mathbf{p}} \geq \mathbf{p}^t U\hat{\mathbf{p}} \text{ for all } \mathbf{p} \in S_n \tag{2.9}$$

2. Stability condition

$$\text{if } \mathbf{p} \neq \hat{\mathbf{p}} \,\text{and}\, \mathbf{p}^t U\hat{\mathbf{p}} = \hat{\mathbf{p}}^t U\hat{\mathbf{p}}, \text{ then } \mathbf{p}^t U\mathbf{p} < \hat{\mathbf{p}}^t U\mathbf{p} \tag{2.10}$$

(2.9) is just the definition of Nash equilibrium, whereas (2.10) means that when compared, $\hat{\mathbf{p}}$ fares better against $\mathbf{p}$ than $\mathbf{p}$ itself. What is concluded is that strict Nash equilibrium implies ESS but ESS only implies Nash equilibrium. Now, we are ready for a theorem that gives us necessary and sufficient conditions for ESS.

Theorem 4. *A strategy* $\hat{\mathbf{p}} \in S_n$ *is an ESS if and only if*

$$\hat{\mathbf{p}}^t U \mathbf{p} > \mathbf{p}^t U \mathbf{p}$$

for all $\mathbf{p} \neq \hat{\mathbf{p}}$ *in some neighbourhood of* $\hat{\mathbf{p}}$ *in* S_n.

Proof. Suppose that $\hat{\mathbf{p}}$ is ESS. Now, let $supp(\hat{\mathbf{p}}) := \{i : 1 \leq i \leq n,\, p_i > 0\}$. We first want to show that for every $\mathbf{p}$ close to $\hat{\mathbf{p}}$, there exists $\mathbf{q}$ such that $\mathbf{p}$ can be written as $\epsilon\mathbf{q} + (1-\epsilon)\hat{\mathbf{p}}$, for ϵ sufficiently small. In fact, we can choose $\mathbf{q}$ in the set $C = \{\mathbf{x} \in S_n : x_i = 0 \text{ for some } i \in supp(\hat{\mathbf{p}})\}$, which is the set of faces of the simplex that do not contain $\hat{\mathbf{p}}$. Since for every $\mathbf{q} \in C$ (2.8) holds for all $0 < \epsilon < \bar{\epsilon}(\mathbf{q})$, we can observe that $\bar{\epsilon}(\mathbf{q})$ can be chosen to be continuous on $\mathbf{q}$. Then, by compactness of C we obtain that $\bar{\epsilon} = \min\{\bar{\epsilon}(\mathbf{q}) : \mathbf{q} \in C\}$ is strictly positive and (2.8) holds for all $\epsilon \in (0, \bar{\epsilon})$. So, if we multiply bot sides of (2.8) by ϵ, then add

$$(1-\epsilon)\hat{\mathbf{p}}^t U((1-\epsilon)\hat{\mathbf{p}} + \epsilon\mathbf{q})$$

and then take $\mathbf{p} := (1-\epsilon)\hat{\mathbf{p}} + \epsilon\mathbf{q}$ we get the desired inequality. The converse is almost immediate from the previous idea. □

From this theorem, (2.9) and (2.10) it follows that if $\hat{\mathbf{p}} \in int(S_n)$ is ESS then there is no other ESS. This is due to the fact that the theorem guarantees that there will be a neighbourhood around $\hat{\mathbf{p}}$ where $\hat{\mathbf{p}}^t U \mathbf{p} > \mathbf{p}^t U \mathbf{p}$ holds, hence combining this with (2.10) and supposing that there is another ESS different from $\hat{\mathbf{p}}$ one gets a contradiction, since the neighbourhood cannot touch the frontier of S_n. Still, there are games with no ESS, games with one ESS in the interior and games with multiple ESS all in the boundary.

Until now, we have considered games that deal with pairwise encounters. We now need to generalize ESS in order to take into consideration the cases where the success of a strategy depends on the population success rate and not only the opponents success rate.

Definition 7. Let f_i be the payoff obtained from strategy R_i, and assume that it is a function of the frequencies m_j of the pure strategies R_j. In this case $\mathbf{m} \in S_n$ is called the **population strategy mix.** Now, the fitness for a strategy $\mathbf{p}$ is given by

$$\sum p_i f_i(\mathbf{m}) = \mathbf{p}^t \mathbf{f}(\mathbf{m}).$$

We define $\hat{\mathbf{p}}$ to be a **local ESS** if $\hat{\mathbf{p}}^t \mathbf{f}(\mathbf{q}) > \mathbf{q}^t \mathbf{f}(\mathbf{q})$ holds for all $\mathbf{q} \neq \hat{\mathbf{p}}$ in some neighbourhood of $\hat{\mathbf{p}}$.

It is important to observe the relation between the terms payoff and fitness. The payoff is the what the player receives in every play time step, and in the game theoretical context it is given by a matrix, and it represents the payoff received by a player with a certain strategy when playing against a player with another (not necessarily different) strategy. In this context fitness is the same as payoff. It is important to remark that this will not be the case in the next chapter where the fitness will be a state of wealth of a player that is affected during each time step by his previous state and the payoff given by interacting with his neighbours.

Notice that if the payoff functions are linear in $\mathbf{m}$ we are again in the normal form games. Still, this may not be the case in general. We would also want an analogue in this case for (2.9) and (2.10). The analogue is given by: $\hat{\mathbf{p}}$ is a local ESS if and only if

1. $\mathbf{q}^t\mathbf{f}(\hat{\mathbf{p}}) \leq \hat{\mathbf{p}}^t\mathbf{f}(\hat{\mathbf{p}})$ for all $\mathbf{q} \in S_n$.
2. if $\mathbf{q}^t\mathbf{f}(\hat{\mathbf{p}}) = \hat{\mathbf{p}}^t\mathbf{f}(\hat{\mathbf{p}})$ and $\mathbf{q} \neq \hat{\mathbf{p}}$ is sufficiently close to $\hat{\mathbf{p}}$, then $\hat{\mathbf{p}}^t\mathbf{f}(\mathbf{q}) > \mathbf{q}^t\mathbf{f}(\mathbf{q})$.

With this in mind, we are ready to begin with the topic of population dynamics. We are going to explore some of the theory on replicator dynamics and general game dynamics.

2.4 Replicator Equations and Other Game Dynamics

Having set the preliminaries, we are now ready to follow our path into linking game theory and dynamical systems in a way that will be useful for our goal. We are going to focus on replicator dynamics, which is basically the description of the evolution of the frequencies of strategies in a population. We aim to search for a representation of Nash equilibria and ESS in the language of replicator dynamics.

2.4.1 The Replicator Equation

As before, we are going to work on the simplex S_n. What we will see is evolutionary stability relies on dynamical systems in some sense which can be modelled by differential equations on S_n.
Formally speaking, let the population be divided into n types $E_1, ..., E_n$ with frequencies $x_1, ..., x_n$ respectively. The fitness f_i of E_i will be considered to be a function of $\mathbf{x} = (x_1, ..., x_n) \in S_n$. Under the assumptions that the population is large and that the strategy frequencies change continuously, we can suppose that $\mathbf{x}(t)$ evolves in S_n as a differentiable function of t. For instance, in the case of a parliamentary population such as the Colombian Congress, this assumptions will not be unrealistic. The senate for instance has more than a hundred senators, which can be thought of as a large population. Also, every time that there is a new election, the "new generation" blends continuously, since there is usually some renewal, but not a complete new senate that will change the course of the system. The time steps can also be thought of as continuous and the renewal will take place every time a senator changes his strategy (making an alliance or lobbying). Any of these two approaches can be fitted into the

previous theory. The rate of increase $\frac{\dot{x_i}}{x_i}$ of type E_i will be used as a measure of the evolutionary success of the type E_i. Both a biological Darwinian approach and the replicator approach, suggests that the rate of increase can be modelled as the difference between the fitness $f_i(\mathbf{x})$ of E_i and the weighted average fitness $\hat{f}(\mathbf{x}) = \sum_{i=1}^{n} x_i f_i(\mathbf{x})$. Hence a population does not always behave in a rational way, but it imitates the strategies of the individuals that are temporarily succeeding. This yields the following the **replicator equation** which is given by:

$$\dot{x_i} = x_i(f_i(\mathbf{x}) - \hat{f}(\mathbf{x})) \quad i = 1, ..., n \tag{2.11}$$

S_n is invariant under (2.11): if $\mathbf{x} \in S_n$ then $\mathbf{x}(t) \in S_n$ for all $t \in \mathbb{R}$. This follows form the fact that $S = x_1 + ... + x_n$ satisfies

$$\dot{S} = (1 - S)\bar{f}$$

(which is just adding all the values of i of (2.11)) This system has solution $S(t) = 1$, an equilibrium point, which makes the invariance clear. Also, if $x_i(0) = 0$ then $x_i(t) = 0$ for $t \in \mathbb{R}$. This makes the faces of S_n invariant, which implies the invariance of S_n. Here it becomes relevant to recall the fundamental theorem for ordinary differential equations which states the local uniqueness of first-order differential equation solutions:

Theorem 5. *Consider the initial value problem* $y'(t) = f(t, y(t)), y(t_0) = y_0, t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ *.Suppose* f *is Lipschitz continuous in* y *and continuous in* t*. Then, for some value* $\epsilon > 0$, *there exists a unique solution* $y(t)$ *to the initial value problem on the interval* $[t_0 - \varepsilon, t_0 + \varepsilon]$ *.*

This gives us a justification for only considering (2.11) restricted to S_n, since the faces of S_n (frontiers or boundaries) will not be crossed from one side to another.
Following this track of mind, it will appear useful to consider the case when the fitness functions are linear. This means, when there exists an $n \times n$ matrix A such that $f_i(\mathbf{x}) = (A\mathbf{x})_i$. Hence, the replicator equation (2.11) now becomes

$$\dot{x_i} = x_i((A\mathbf{x})_i - \mathbf{x}^t A\mathbf{x}) \quad i = 1, ..., n. \tag{2.12}$$

Notice that solutions of

$$(A\mathbf{x})_1 = ... = (A\mathbf{x})_n \tag{2.13}$$
$$x_1 + ... + x_n = 1 \tag{2.14}$$

satisfying $x_i > 0$ for $i = 1, ..., n$ are equilibrium points of (2.12) in $int(S_n)$. The existence of such solutions is not always true. By restricting to simplexes of a smaller dimension, one can obtain similar solutions. In fact, what appears on (2.13) and (2.14) is typically a more friendly version of the system that one aims to solve(system (2.11)).

2.4.2 Nash Equilibria and Evolutionary Stable States Revisited

With the previous section in mind, we are now ready to link the replicator dynamics and the game theory concepts we saw in the previous chapter. As before, there is a normal form game with N pure strategies $R_1, ..., R_N$ and a payoff matrix $U\,(N \times N)$. A strategy is defined to be a point in S_N, so that the types $E_1, ..., E_n$ correspond to points $\mathbf{p}^1, ..., \mathbf{p}^n \in S_N$. The state of the population is given by the frequencies of each type, which is a point $\mathbf{x} \in S_n$. Hence, the payoff obtained by a player of type E_i against a player of type E_j is given by $a_{ij} = {\mathbf{p}^i}^t U \mathbf{p}^j$. So, the fitness of the type E_i is given by

$$f_i(\mathbf{x}) = \sum_j a_{ij} x_j = (A\mathbf{x})_i$$

so that we are in the linear case given by (2.12).
With this in mind we are now ready to delve into game theory topics.

Definition 8. A point $\hat{\mathbf{x}} \in S_n$ is a **Nash equilibrium** if

$$\mathbf{x}^t A \hat{\mathbf{x}} \leq \hat{\mathbf{x}}^t A \hat{\mathbf{x}} \tag{2.15}$$

for all $\mathbf{x} \in S_n$

Definition 9. A point $\hat{\mathbf{x}} \in S_n$ is called **evolutionary stable state** if

$$\hat{\mathbf{x}}^t A \mathbf{x} > \mathbf{x}^t A \mathbf{x} \tag{2.16}$$

for all $\mathbf{x} \neq \hat{\mathbf{x}}$ in a neighbourhood of $\hat{\mathbf{x}}$

It is important to point out that here we are defining Nash equilibrium and evolutionary stability in terms of states and not of strategies, like we did before. Continuing, we would be interesting in obtaining results that allow us to understand when is there a Nash equilibrium. We want this result to be in terms of the dynamical system tools that we introduced in the previous chapter. The following theorem does exactly that:

Theorem 6. *1. If $\hat{\mathbf{x}} \in S_n$ is a Nash equilibrium of the game described by the payoff matrix A, then $\hat{\mathbf{x}}$ is a rest point of the linear replicator equation (2.12).*

2. If $\hat{\mathbf{x}}$ is the $\omega-$limit of an orbit $\mathbf{x}(t) \in int(S_n)$, then $\hat{\mathbf{x}}$ is a Nash equilibrium.

3. If $\hat{\mathbf{x}}$ is Lyapunov stable, then it is a Nash equilibrium.

Proof. 1. If $\hat{\mathbf{x}}$ is a Nash equilibrium, then there is a constant c such that $(A\hat{\mathbf{x}})_i = c$ for all i with $x_i > 0$. The reason for this is due to Lagrange multipliers. If we define $F(\mathbf{x}) = \mathbf{x}^t A \hat{\mathbf{x}}$ and we consider finding the point that maximizes this function restricted to $x_1+...+x_n=1$ we should obtain $\hat{\mathbf{x}}$ (the Nash equilibrium). Then we get, using Lagrange multipliers, that:

$$F(\mathbf{x}) = \mathbf{x}^t A\hat{\mathbf{x}} = \sum_i x_i \underbrace{\left(\sum_j A_{ij}\hat{x_j}\right)}_{c_i}$$

From this, it follows that

$$\nabla F(\mathbf{x}) = c = (c_i)_i = (\lambda, ..., \lambda).$$

Hence, since $\hat{\mathbf{x}}$ is a Nash equilibrium, it must happen that c is the multiplier that solves the previous optimization problem. This idea will be key in the next two proofs. Now, we get that that $\hat{\mathbf{x}}$ satisfies (2.13) and (2.14) for a rest point spanned by the canonical vectors $\mathbf{e}_i$ with $i \in supp(\hat{\mathbf{x}})$. Thus, it must be a rest point of (2.12).

2. Suppose by a contradiction that $\mathbf{x}(t_k)$ converges to $\hat{\mathbf{x}}$ but $\hat{\mathbf{x}}$ is not a Nash equilibrium. Then there exists an i and an ϵ such that $\frac{\dot{x_i}}{x_i} = \mathbf{e}_i^t A\hat{\mathbf{x}} - \hat{\mathbf{x}}^t A\hat{\mathbf{x}} > \epsilon$ for a large enough t, which yields a contradiction to the definition of $\omega-$limit.

3. Suppose for a contradiction that $\hat{\mathbf{x}}$ is not a Nash equilibrium. Then by continuity, there exists an index i and $\epsilon > 0$ such that $(A\mathbf{x})_i - \mathbf{x}^t A\mathbf{x} > \epsilon$ for all $\mathbf{x}$ in a neighbourhood of $\hat{\mathbf{x}}$. Now, since $\hat{\mathbf{x}}$ is Lyapunov stable, we have that all the $\mathbf{x}$ in the just described neighbourhood tend to stabilize. But the $i-$th component of such $\mathbf{x}$ increase exponentially due to $(A\mathbf{x})_i - \mathbf{x}^t A\mathbf{x} > \epsilon$. This contradicts the Lyapunov stability.

□

Theorem 7. *If $\hat{\mathbf{x}} \in S_n$ is an evolutionary stable state for the game with payoff matrix A, then it is an asymptotically stable rest point of (2.12).*

Proof. We want to show that $P(\mathbf{x}) \leq P(\hat{\mathbf{x}})$ where

$$P(\mathbf{x}) = \prod x_i^{\bar{x_i}}.$$

In the following we extend the log function to the point to 0 by setting $0 \log 0 = 0 \log \infty = 0$. By applying log we obtain that

$$\begin{aligned}
\sum_{i=1}^n \hat{x_i} \log \frac{x_i}{\hat{x_i}} &= \sum_{\hat{x_i}>0} \hat{x_i} \log \frac{x_i}{\hat{x_i}} \\
&\leq \log \sum_{\hat{x_i}>0} x_i \ \ (\mathrm{by Jensen's inequality}) \\
&= \log \sum_{i=1}^n x_i \\
&= \log 1 \\
&= 0
\end{aligned}$$

Equality holds if and only if $\mathbf{x} = \hat{\mathbf{x}}$. If $P > 0$ we have that

$$\begin{aligned}\frac{\dot{P}}{P} &= (\dot{\log P}) \\ &= \frac{d}{dt}\left(\sum \hat{x_i}\log(x_i)\right) \\ &= \sum_{\hat{x_i}>0} \hat{x_i}\frac{\dot{x_i}}{x_i} \\ &= \sum \hat{x_i}((A\mathbf{x})_i - \mathbf{x}^t A\mathbf{x}) \\ &= \hat{\mathbf{x}}^t A\mathbf{x} - \mathbf{x}^t A\mathbf{x}\end{aligned}$$

Which by evolutionary stability (2.16) implies that $\dot{P} > 0$ for all $\mathbf{x} \neq \hat{\mathbf{x}}$ in some neighbourhood of $\hat{\mathbf{x}}$. Then, we have that P is a strict local Lyapunov function for (2.12), which therefore confirms that all orbits near $\hat{\mathbf{x}}$ converge to $\hat{\mathbf{x}}$. □

Now, we will stop with the game dynamics to pass on to a related topic.

2.5 On Population Genetics and Game Dynamics

In this section we are going to observe some results and deductions that will be useful in the context of strategy evolution on graphs. We will focus on the fixation probability which will be a very important concept in the latter chapters. The fixation probability is the probability that a strategy, type or, in this case, gene starting with a certain frequency takes over the entire population. This gives rise to a function that begins with a frequency p and a type and returns the probability that the type takes over the entire population . We will deduce the backward Kolmogorov equation which nicely fits the fixation probability function for some cases that we will encounter in the next chapter. So first we begin by defining some of the key concepts both in a mathematical and biological framework.

2.5.1 The Kolmogorov Backward Equation

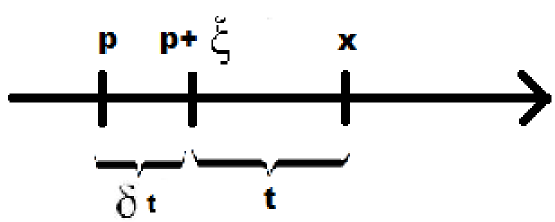
We will be working in the context of the description of the genetic composition of a Mendelian population (this an interbreeding population that shares a genetic pool). The process of change in gene (we will call it gene instead of type or strategy since the origin of this theory is genetics) frequency will be regarded as a continuous stochastic process. This means that as the observed time interval becomes smaller, the change in the gene frequency will not explode i.e. for all $\epsilon > 0$ the probability that the change in the frequency exceeds ϵ during the time interval $(t, t + \delta t)$ is $o(\delta t)$. Also, we will assume that that the process is Markovian, i.e. that the frequency depends only on the previous time but not on all the past history. First we suppose that A_1 and A_2 are segregating alleles in a population. Now, define $\phi(p,x;t)$ as the probability density that the frequency of A_1 lies between x and $x+dx$ at time t, given that the initial frequency($t=0$) is p. We assume thet the gene frequency x is fixed at time t and that p, the initial frequency, is a random variable. This is known as a reversed process which satisfies the following equation which is known as the **Kolmogorov backward equation:**

$$\frac{\partial\phi(p,x;t)}{\partial t} = \frac{1}{2}V_{\delta t}\frac{\partial^2}{\partial p^2}\phi(p,x;t) + M_{\delta t}\frac{\partial}{\partial p}\phi(p,x;t) \tag{2.17}$$

where $M_{\delta t}$ and $V_{\delta t}$ are respectively the mean and variance of the amount of change in gene frequency per time step δt, when the gene frequency is set to start at p. We will restrict to the case where the process is time homogeneous, that is when if x_{t_1} and x_{t_2} are gene frequencies at times t_1 and t_2, then the distribution of x_{t_2} given x_{t_1} only depends on $t_2 - t_1$. Then for such a time homogeneous Markov process:

$$\phi(p,x;t+\delta t) = \int g(p,\xi;\delta t)\phi(p+\xi,x,;t)d\xi \tag{2.18}$$

where $g(p,\xi;\delta t)$ is the probability density that the gene frequency changes from x to $x+\xi$ during the time interval of length δt (since the process is time homogeneous only this length matters). The above result follows from the fact that the process is Markovian. Equation (2.18) is just saying that the probability that the gene frequency that started at p and is x at time $t+\delta t$ is the sum (integral) over the cases of the density of going from frequency p to $p+\xi$ in an initial time interval of length δt and then of starting with frequency $p+\xi$ and ending up in x after time t. If δt is chosen to be very small and because of assumption the we are in a continuous stochastic process, the change ξ during this time interval is practically restricted to very small values. This can be better explained by the following picture



Then by Taylor expansion of $\phi(p+\xi,x;t)$ inside the integral we get that:

$$\phi(p,x;t+\delta t) = \int (\phi g) + g\xi\frac{\partial(\phi)}{\partial p} + g\frac{\xi^2}{2!}\frac{\partial^2(\phi)}{\partial p^2} + O(\xi^3)d\xi$$

Where ϕ and g stand for $\phi(p,x;t)$ and $g(p,\xi;\delta t)$ respectively. Hence, using the fact that g is a density (hence $\int g d\xi = 1$ for a fixed p and a small δt), ignoring the $O(\xi^3)$ terms and subtracting the first term on the right we obtain that

$$\phi(p,x;t+\delta t) - \phi(p,x;t) = \int g\xi \frac{\partial(\phi)}{\partial p} + g\frac{\xi^2}{2!}\frac{\partial^2(\phi)}{\partial p^2}$$

Then dividing both sides by δt and letting $\delta t \longrightarrow 0$ we get that:

$$\frac{\partial \phi(p,x;t)}{\partial t} = \frac{V(p)}{2}\frac{\partial^2 \phi(p,x;t)}{\partial p^2} + M(p)\frac{\partial \phi(p,x;t)}{\partial p}$$

where

$$M(p) = \lim_{\delta t \to 0} \frac{1}{\delta t}\int \xi g(p,\xi;\delta t) d\xi$$

and

$$V(p) = \lim_{\delta t \to 0} \frac{1}{\delta t}\int \xi^2 g(p,\xi;\delta t) d\xi$$

Then, by substituting the mean and the variance of the amount of change per generation, $M_{\delta t}$ and $V_{\delta t}$ for $M(p)$ and $V(p)$ we obtain the Kolmogorov backward equation.

Now, we are interested in the application of this model to the fixation problem. This is the case where $x = 1$, and the population is invaded by a single gene or specie. In this case, $u(p,t) := \phi(p,1;t)$ of the backward Kolmogorov equation can be interpreted as the probability of gene or specie fixation at time t if the initial frequency was p. Thus, $u(p,t)$ satisfies the differential equation:

$$\frac{\partial u(p,t)}{\partial t} = \frac{V_{\delta t}}{2}\frac{\partial^2 u(p,t)}{\partial p^2} + M_{\delta t}\frac{\partial u(p,t)}{\partial p}$$

The fixation probability will be found by solving the differential equation with boundary conditions

$$u(0,t) = 0, \ u(1,t) = 1$$

which means that a gene with 0 frequency stays with 0 frequency and that a gene with full frequency will remain that way. Now, what is more important is the ultimate fixation probability which is just the limit as $t \to \infty$

$$u(p) = \lim_{t \to \infty} u(p,t).$$

This satisfies

$$\frac{\partial u}{\partial t} = 0$$

since the probability stabilizes, so it becomes, in the limit, constant, thus the derivative is equal to 0. So we obtain that the ultimate fixation probability satisfies the ordinary differential equation:

$$0 = \frac{V_{\delta t}}{2}\frac{d^2 u(p)}{dp^2} + M_{\delta t}\frac{du(p)}{dp} \tag{2.19}$$

with boundary conditions $u(0) = 0$ and $u(1) = 1$.

This whole deduction will be of great importance in the next chapter.

It is very important to point out that this deduction is an approximation. The step in which the terms of order higher than 3 on ξ (noted as $O(\xi^3)$) is an approximation in which the assymptotical convergence. Still, this deduction is the one considered by Nowak in his article, and is directly pulled of from [2]. Although this deduction is nos at all satisfactory, it is the one that was used by Nowak. It is still important to point out that since we are in a diffussion process (markovian time homogeneous), the best way to obtain a differential ecuation that is not approximate is to use a stochastic differential equation that properly models the allele invasion problem, and try to solve it by usin Ito's lemma or other stochastic calculus tool. Although it is not the main concern of this project, we want to give a look at the stochastic calculus deduction for the desired equation in order to give an insight to a more rigorous deduction of (2.19).

Informally, we want to give rise to a SDE that models the allele invasion problem, and then using Ito's lemma, we want to solve the SDE and obtain a PDE that can be treated to become 2.19. In population dynamics, there is a differential population model that describes the population frequency at time t when knowing the inithial frequency p. This model is known as the *Malthusian* model for population change and is given by

$$\frac{dX_t}{dt} = \mu X_t$$

where $X_0 = p$ and μ is the mean of the population frecuency at time t. Then by assuming changes in enviromental conditions (stochastic changes) the mean can be replaced by a Gaussian random variable (with non-zero non-constant mean $\mu(X_t)$ and variance $(\sigma(x_t))$) to obtain the stochastic differential ecuation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t$$

(with the usual notation for the brownian movement of B_t.) Now using the integrating factor $\phi(p, X_t, t)$ for Ito's lemma we obtain that

$$-\frac{\partial}{\partial t}\phi(p, X_t, t) = \mu(X_t)\frac{\partial}{\partial X_t}\phi(p, X_t, t) + \frac{1}{2}\sigma^2(X_t)\frac{\partial^2}{\partial X_t^2}\phi(p, X_t, t)$$

which we can now interpret as

$$-\frac{\partial u(p,t)}{\partial t} = \frac{V_{\delta t}}{2}\frac{\partial^2 u(p,t)}{\partial p^2} + M_{\delta t}\frac{\partial u(p,t)}{\partial p}$$

using the above notation. From, here the deduction is the same as before, but now it has no approximate. Although the steps are not fully developted in this informal deduction, it gives a clear outlook onto why (2.19) is a valid ecuation for our model.

Chapter 3

Strategy Evolution on Graphs

This chapter will be based on the 2006 paper written by Martin Nowak, *et al* called "A simple rule for the evolution of cooperation on graphs and social networks". We will try to present, as thoroughly as possible the contents of the paper. That is, we will try to fill in the blanks for all the missing calculations in the article. It is important to note that our goal is not to test the soundness of the mathematical calculations, but rather try to deeply understand the calculations that lead to the desired result. First, we need to establish the background that we will be working with. Consider a game between two strategies, A and B, with payoff matrix

$$\begin{array}{ccc} & A & B \\ A & a & b \\ B & c & d \end{array}$$

This means we are in a population that has two types or two strategies. Also we assume that a population of size N is distributed over the vertices of a graph. That is, all vertices of the graph are either of type A or of type B. The payoff of each individual is the sum of the interactions with its neighbours in the graph. We will also assume that the graph has certain restrictions. For instance, we will suppose that each vertex in the graph is connected to exactly k other vertices. Thus, the degree of the graph will be k.

In order for this to be in the dynamical system background, we will need to introduce some change during time. This means that the graph will have to update upon some changes in time. Since we are in the context of graphs, time will be considered as discrete, so we will work with time steps (unitary time steps). We will look three different update rules: the "Death-Birth" (DB), the "Imitation" (IM) and the "Birth-Death"(BD). DB means that in each time step a random individual is chosen to die, then the neighbours will compete to colonize that vertex with their type. This competition for colonizing will be based

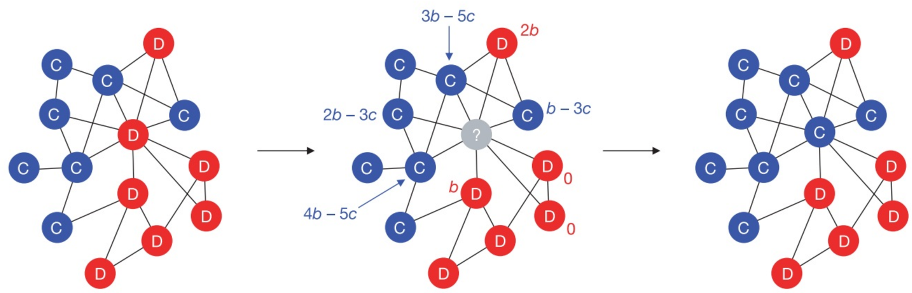


Figure 3.1: The rules of the game. Each individual occupies the vertex of a graph and derives a payoff P from interactions with adjacent individuals. A cooperator (blue) pays a cost, c, for each neighbour to receive a benefit, b. A defector (red) pays no cost and provides no benefit. The fitness of a player is given by $1 - w + wP$, where w measures the intensity of selection (described later). Strong selection means $w = 1$. Weak selection means $w << 1$. For deathbirth updating, at each time step, a random individual is chosen to die (grey); subsequently the neighbours compete for the empty site in proportion to their fitness. In this example, the central, vacated vertex will change from a defector to a cooperator with a probability $\frac{F_C}{(F_C+F_D)}$, where the total fitness of all adjacent cooperators and defectors is $F_C = 4(1 - w) + (10b - 16c)w$ and $F_D = 4(1-w) + 3bw$, respectively. This figure and the caption was taken from the article in nature magazine ([3])

on the fitness of such neighbours. On the other hand, IM will mean that in every time step a random individual will evaluate his strategy and decide if it stays or change by imitating a neighbour. This decision will take place based upon fitness of the neighbours too. Finally, BD means that in each time step an individual is chosen for reproduction based on fitness, and then the offspring takes a random neighbour's vertex. This can be explained in a more detailed way by a picture that shows a time step in the rule in a game (prisoner's dilemma) where the values of the matrix are $a = b - c$, $b = -c, c = b$ and $d = 0$. This can be seen and better explained in figure (3.1) which was directly taken from Nowak's article ([3])

One important aspect of this algorithm is that it may not work under some circumstances. For instance, if the values b and c are equal and w is close to 1, the probability coefficient $\frac{f_C}{(f_c+f_d)}$ can be negative, which totally ruins the procedure.

It is important to explain that we are going to use *pair approximation* and *diffusion approximation* to derive a fixation probability, ρ_A, which is the probability that a single player A starting at a random position on the graph with the rest of vertices being occupied by $B-$types, generates a lineage of players that takes over the entire population. If $\rho_A > \frac{1}{N}$, then selection tends to favour the fixation of A. This follows from the fact that in genetics, a neutral mutant gene's (can be interpreted as a strategy) fixation probability is proportional to the frequency which it began with. That is, if a population has $N-1$ Y-types and only one $X-$type which is neutral, then $\rho_X = \frac{1}{N}$. Hence, if the $X-$type was not neutral but was instead favoured for fixation one would expect that $\rho_X > \frac{1}{N}$. It will also be of help to compare ρ_A to ρ_B. *Pair approximation* means that the frequencies of large clusters, can be derived from the frequencies of pairs. We are interested in this because cluster frequencies can give us a great insight about the whole population behaviour. On the other hand, *diffusion approximation* works as a method to deal with intractable discrete processes, like the updating in the graph, by changing it into a diffusion process which has easier properties and can be tackled with dynamical systems theory.

The course of action in this chapter will be targeted toward finding conditions on the values a, b, c and d of the payoff matrix and on the degree of the graph k for a certain strategy to be favoured upon fixation. Then we apply these results to the prisoner's dilemma (which we explain later) where we get specific decision rules for the values of the . This will be done for each of the updating rules, and in each case we will get different results. In each case we will follow a similar approach, in which the Kolmogorov backward equation models the fixation probability function, and we attempt to approximate a solution for the differential equation by a third degree polynomial. In order to get the particular Kolmogorov backward equation we need to obtain a mean and variance to fit in. These are obtained by approximations that usually remove terms that asymptotically vanish. This is important in the sense that the calculations in this chapter are not exact, but are asymptotically accurate enough.

Now lets take a look a closer look at these updating rules.

3.1 DB-Updating

First, we will establish some notation. Let p_A and p_B denote the frequencies of A and B respectively. Let p_{AA}, p_{BB}, p_{AB}, and p_{BA} denote the frequencies of AA, BB, AB and BA pairs respectively. These are probabilities on the edges of the graph. That is, p_{XY} is the frequency of edges that connect $X-$types to $Y-$types. Let $q_{X|Y}$ denote the conditional probability to find an $X-$type given that the adjacent node has a $Y-$type, where X and Y can stand for both A and B. We then have the following:

$$p_A + p_B = 1 \tag{3.1}$$

$$q_{A|X} + q_{B|X} = 1 \tag{3.2}$$

$$p_{XY} = q_{X|Y} p_Y \tag{3.3}$$

$$p_{AB} = p_{BA} \tag{3.4}$$

This implies, after some straightforward calculations, that the system can be fully described in terms of two variables, namely p_A and $q_{A|A}$. So we are interested in calculating the probabilities of change of these two variables during a time step. As explained before, in each time step a random player is eliminated and the neighbours of his vertex compete for taking it based upon their fitness.

So, suppose in a time step a $B-$type is eliminated. This happens with probability p_B. Then, his k neighbours compete for the now empty vertex. We can write $k = k_A + k_B$ where k_A and k_B are the amount of A and B neighbours respectively. The frequency of such configuration is given by

$$\left(\frac{k!}{k_A!k_B!}\right) q_{A|B}^{k_A} q_{B|B}^{k_B}.$$

Also, the fitness of the players is given by

$$\begin{aligned} f_A &= (1-w) + w\left[(k-1)q_{A|A}a + ((k-1)q_{B|A}+1)b\right] && (3.5)\\ &= 1 + w\underbrace{\left[(k-1)q_{A|A}a + ((k-1)q_{B|A}+1)b - 1\right]}_{:=\bigstar_A} && (3.6)\end{aligned}$$

$$\begin{aligned} f_B &= (1-w) + w\left[(k-1)q_{A|B}c + ((k-1)q_{B|B}+1)d\right] && (3.7)\\ &= 1 + w\underbrace{\left[(k-1)q_{A|B}c + ((k-1)q_{B|B}+1)d - 1\right]}_{:=\bigstar_B} && (3.8)\end{aligned}$$

Here the parameter $w \in [0,1]$ represents the intensity of the selection. This intensity reflects how high is the payoff relative to the fitness. For instance, if there is a strong selection, say $w = 1$ then the fitness is identical to the payoff. On the other hand, if $w << 1$ then the payoff is really small compared to the fitness. This case is called **weak selection.** It is important to point out that (3.5) and (3.7) mean that depending on w the fitness can be highly or lowly contributed by the payoff.

The probability that one of the $A-$type neighbours takes the vacant vertex is then given by

$$\frac{k_A f_A}{k_A f_A + k_B f_B}.$$

Then, p_A increases by $\frac{1}{N}$, i.e. the number of $A-$types increases by 1, with probability

$$P(\Delta p_A = \frac{1}{N}) = p_B \sum_{k_A+k_B=k} \left(\frac{k!}{k_A!k_B!}\right) q_{A|B}^{k_A} q_{B|B}^{k_B} \frac{k_A f_A}{k_A f_A + k_B f_B}. \qquad (3.9)$$

Following this train of thought, the number of $AA-$pairs increases by k_A and therefore p_{AA} increases by $\frac{2k_A}{kN} = \frac{k_A}{(\frac{kN}{2})}$ with probability

$$P(\Delta p_{AA} = \frac{2k_A}{kN}) = p_B \left(\frac{k!}{k_A!k_B!}\right) q_{A|B}^{k_A} q_{B|B}^{k_B} \frac{k_A f_A}{k_A f_A + k_B f_B}. \tag{3.10}$$

Pretty similar arguments can be used if instead of a $B-$type, an $A-$type is eliminated. With the same notation one obtains that the fitness will be given by

$$g_A = (1-w) + w\left[((k-1)q_{A|A} + 1)a + ((k-1)q_{B|A})b\right] \tag{3.11}$$
$$g_B = (1-w) + w\left[((k-1)q_{A|B} + 1)c + ((k-1)q_{B|B})d\right] \tag{3.12}$$

Let's take a break to observe that with the previous notation of $\bigstar_A$ and $\bigstar_B$ we obtain that

$$f_A = 1 + \bigstar_A w \tag{3.13}$$
$$f_B = 1 + \bigstar_B w \tag{3.14}$$
$$g_A = 1 + (\bigstar_A - b + a)w \tag{3.15}$$
$$g_B = 1 + (\bigstar_B - d + c)w \tag{3.16}$$

Also, that the probability that one of the $B-$type neighbours takes the vacant vertex is given by

$$\frac{k_B g_B}{k_A g_A + k_B g_B}.$$

And we obtain that

$$P(\Delta p_A = -\frac{1}{N}) = p_A \sum_{k_A+k_B=k} \left(\frac{k!}{k_A!k_B!}\right) q_{A|A}^{k_A} q_{B|A}^{k_B} \frac{k_B g_B}{k_A g_A + k_B g_B}. \tag{3.17}$$

$$P(\Delta p_{AA} = -\frac{2k_A}{kN}) = p_A \left(\frac{k!}{k_A!k_B!}\right) q_{A|A}^{k_A} q_{B|A}^{k_B} \frac{k_B g_B}{k_A g_A + k_B g_B}. \tag{3.18}$$

3.1.1 Diffusion Approximation

If we suppose that one replacement event takes place in one unit of time, we can do some calculations that will lead us to a tractable problem. Under that hypothesis we obtain that the time derivatives are given by

$$\dot{p_A} = \frac{1}{N} P(\Delta p_A = \frac{1}{N}) - \frac{1}{N} P(\Delta p_A = -\frac{1}{N}) \tag{3.19}$$
$$= \frac{1}{N}\left(p_B \sum_{k_A+k_B=k} \left(\frac{k!}{k_A!k_B!}\right) q_{A|B}^{k_A} q_{B|B}^{k_B} \frac{k_A f_A}{k_A f_A + k_B f_B}\right) \tag{3.20}$$
$$- \frac{1}{N}\left(p_A \sum_{k_A+k_B=k} \left(\frac{k!}{k_A!k_B!}\right) q_{A|A}^{k_A} q_{B|A}^{k_B} \frac{k_B g_B}{k_A g_A + k_B g_B}\right) \tag{3.21}$$
$$= w\frac{k-1}{N} p_{AB}(I_a a + I_b b - I_c c - I_d d) + O(w^2). \tag{3.22}$$

Where

$$I_a = \frac{k-1}{k} q_{A|A}(q_{A|A} + q_{B|B}) + \frac{1}{k} q_{A|A}) \tag{3.23}$$
$$I_b = \frac{k-1}{k} q_{B|A}(q_{A|A} + q_{B|B}) + \frac{1}{k} q_{B|B}) \tag{3.24}$$
$$I_c = \frac{k-1}{k} q_{A|B}(q_{A|A} + q_{B|B}) + \frac{1}{k} q_{A|A}) \tag{3.25}$$
$$I_d = \frac{k-1}{k} q_{B|B}(q_{A|A} + q_{B|B}) + \frac{1}{k} q_{B|B}). \tag{3.26}$$

This calculation is not only long but it also involves some algebraic tricks, hence we will show it step by step. So, first note that since

$$(x+y)^k = \sum_{k_A+k_B=k} \frac{k!}{k_A!k_B!} x^{k_A} y^{k_B}$$

then, if we differentiate both sides with respect to x and y, and then take double derivatives, with respect to x, y and crossed double derivatives we obtain

$$k(x+y)^{k-1} = \sum_{k_A+k_B=k} \frac{k!}{k_A!k_B!} k_A x^{k_A-1} y^{k_B} \tag{3.27}$$
$$k(x+y)^{k-1} = \sum_{k_A+k_B=k} \frac{k!}{k_A!k_B!} k_B x^{k_A} y^{k_B-1} \tag{3.28}$$
$$k(k-1)(x+y)^{k-2} = \sum_{k_A+k_B=k} \frac{k!}{k_A!k_B!} k_A(k_A-1) x^{k_A-2} y^{k_B} \tag{3.29}$$
$$k(k-1)(x+y)^{k-2} = \sum_{k_A+k_B=k} \frac{k!}{k_A!k_B!} k_B(k_B-1) x^{k_A} y^{k_B-2} \tag{3.30}$$
$$k(k-1)(x+y)^{k-2} = \sum_{k_A+k_B=k} \frac{k!}{k_A!k_B!} k_B k_A x^{k_A-1} y^{k_B-1}. \tag{3.31}$$

Now, we must make some manipulations to some of the terms in (3.19) in order to make them tractable. First,

$$k_A f_A + k_B f_B = k + (k_A \bigstar_A + k_B \bigstar_B) w \tag{3.32}$$

$$\tag{3.33}$$

$$\implies \frac{1}{k_A f_A + k_B f_B} = \frac{1}{k} - \frac{(k_A \bigstar_A + k_b \bigstar_B)}{k^2} w + O(w^2) \tag{3.34}$$

$$\tag{3.35}$$

$$\implies \frac{k_A f_A}{k_A f_A + k_B f_B} \tag{3.36}$$

$$\tag{3.37}$$

$$= \frac{k_A}{k} + \left[- \frac{k_A (k_A \bigstar_A + k_B \bigstar_B)}{k^2} + \frac{k_A \bigstar_A}{k} \right] w + O(w^2) \tag{3.38}$$

$$\tag{3.39}$$

$$= \frac{k_A}{k} + \frac{1}{k^2} \left[k_A k_B (\bigstar_A - \bigstar_B) \right] w + O(w^2). \tag{3.40}$$

Similarly, we obtain that

$$\frac{k_B g_B}{k_A g_A + k_B g_B} = \frac{k_B}{k} \tag{3.41}$$

$$+ \frac{1}{k^2} \left[k_A k_B (\bigstar_B - \bigstar_A - a - d + c + b) \right] w \tag{3.42}$$

$$+ O(w^2) \tag{3.43}$$

Then, to obtain the desired result, we first need that the constant parts that appear from the last equality in (3.19) cancel out, so that only the term multiplying by w and the $O(w^2)$ remain. Observe that by (3.27) and by only taking the constant parts at the end of (3.32) we get that

$$\frac{1}{N}\left(p_B \sum_{k_A+k_B=k}\left(\frac{k!}{k_A!K_B!}\right) q_{A|B}^{k_A} q_{B|B}^{k_B}\frac{k_A}{k}\right)$$

$$-\frac{1}{N}\left(p_A \sum_{k_A+k_B=k}\left(\frac{k!}{k_A!k_B!}\right) q_{A|A}^{k_A} q_{B|A}^{k_B}\frac{k_B}{k}\right)$$

$$=\frac{1}{N}\left(\sum_{k_A+k_B=k} p_B\left(\frac{k!}{k_A!k_B!}\right) q_{A|B}^{k_A} q_{B|B}^{k_B}\frac{k_A}{k} - \frac{1}{N}p_A\left(\frac{k!}{k_A!k_B!}\right) q_{A|A}^{k_A} q_{B|A}^{k_B}\frac{k_B}{k}\right)$$

$$=\frac{1}{N}\left(\sum_{k_A+k_B=k} p_B q_{A|B}\left(\frac{k!}{k_A!k_B!}\right) q_{A|B}^{k_A-1} q_{B|B}^{k_B}\frac{k_A}{k} - \frac{1}{N}p_A q_{B|A}\left(\frac{k!}{k_A!k_B!}\right) q_{A|A}^{k_A} q_{B|A}^{k_B-1}\frac{k_B}{k}\right)$$

$$=\frac{1}{N}(p_{AB}-p_{BA})$$
$$=0.$$

Now lets focus on the term that is multiplied by w. The trick will be, again, to use (3.27), but this time also considering the double derivatives.

$$\frac{1}{N}\left(p_B \sum_{k_A+k_B=k}\left(\frac{k!}{k_A!k_B!}\right) q_{A|B}^{k_A} q_{B|B}^{k_B}\frac{1}{k^2}\left[k_A k_B(\bigstar_A-\bigstar_B)\right]\right)$$

$$-\frac{1}{N}\left(p_A \sum_{k_A+k_B=k}\left(\frac{k!}{k_A!k_B!}\right) q_{A|A}^{k_A} q_{B|A}^{k_B}\frac{1}{k^2}\left[k_A k_B(\bigstar_B-\bigstar_A-a-d+c+b)\right]\right)$$

$$=\frac{1}{N}\frac{k(k-1)}{k^2}\left[p_B q_{A|B} q_{B|B}(\bigstar_A-\bigstar_B)\right]$$

$$-\frac{1}{N}\frac{k(k-1)}{k^2}\left[p_A q_{A|A} q_{B|A}(\bigstar_B-\bigstar_A-a-d+c+b)\right]$$

$$= \frac{1}{N}\frac{k(k-1)}{k^2}\left[p_{AB}q_{B|B}(\bigstar_A - \bigstar_B)\right]$$

$$- \frac{1}{N}\frac{k(k-1)}{k^2}\left[p_{AB}q_{A|A}(\bigstar_B - \bigstar_A - a - d + c + b)\right]$$

$$= \frac{1}{N}\frac{(k-1)}{k}p_{AB}\left[q_{B|B}(\bigstar_A - \bigstar_B) - q_{A|A}(\bigstar_B - \bigstar_A - a - d + c + b)\right]$$

$$= \frac{1}{N}\frac{(k-1)}{k}p_{AB}\left[q_{B|B}((k-1)q_{A|A}a + \{(k-1)q_{B|A}+1\}b - (k-1)q_{A|B}c - \{(k-1)q_{B|B}+1\}d)\right]$$

$$- \frac{1}{N}\frac{(k-1)}{k}p_{AB}\left[q_{A|A}(-\{(k-1)q_{A|A}+1\}a - (k-1)q_{B|A}b + \{(k-1)q_{A|B}+1\}c + (k-1)q_{B|B}d)\right]$$

$$= \frac{k-1}{N}p_{AB}\left[\frac{1}{k}\left((k-1)q_{B|B}q_{A|A} + (k-1)q_{A|A}q_{A|A} + q_{A|A}\right)\right]a$$

$$+ \frac{k-1}{N}p_{AB}\left[\frac{1}{k}\left((k-1)q_{B|B}q_{B|A} + (k-1)q_{A|A}q_{B|A} + q_{B|B}\right)\right]b$$

$$- \frac{k-1}{N}p_{AB}\left[\frac{1}{k}\left((k-1)q_{B|B}q_{A|B} + (k-1)q_{A|A}q_{A|B} + q_{A|A}\right)\right]c$$

$$- \frac{k-1}{N}p_{AB}\left[\frac{1}{k}\left((k-1)q_{B|B}q_{B|B} + (k-1)q_{A|A}q_{B|B} + q_{B|B}\right)\right]d$$

$$= \frac{k-1}{N}p_{AB}(I_a a + I_b b - I_c c - I_d d).$$

For the pairs, we get that:

$$\dot{p_{AA}} = \sum_{k_A=0}^{k}\frac{2k_A}{kN}P(\Delta p_{AA} = \frac{2k_A}{kN}) \tag{3.44}$$

$$+ \sum_{k_A=0}^{k} -\frac{2k_A}{kN}P(\Delta p_{AA} = -\frac{2k_A}{kN}) \tag{3.45}$$

$$= \frac{2}{kN}p_{AB}\left[1 + (k-1)(q_{A|B} - q_{A|A})\right] + O(w). \tag{3.46}$$

This result follows in a similar way as for $\dot{p_A}$. Let us look at it:

$$\dot{p_{AA}} = \sum_{k_A=0}^{k} \frac{2k_A}{kN} P(\Delta p_{AA} = \frac{2k_A}{kN}) \tag{3.47}$$

$$+ \sum_{k_A=0}^{k} -\frac{2k_A}{kN} P(\Delta p_{AA} = -\frac{2k_A}{kN}) \tag{3.48}$$

$$\tag{3.49}$$

$$= \sum_{k_A=0}^{k} \frac{2k_A}{kN} p_B \frac{k!}{k_A! k_B!} q_{A|B}^{k_A} q_{B|B}^{k_B} (\frac{k_A}{k} + O(w)) \tag{3.50}$$

$$+ \sum_{k_A=0}^{k} -\frac{2k_A}{kN} p_A \frac{k!}{k_A! k_B!} q_{A|A}^{k_A} q_{B|A}^{k_B} (\frac{k_B}{k} + O(w)) \tag{3.51}$$

$$\tag{3.52}$$

$$= \sum_{k_A=0}^{k} \frac{2p_B}{k^2 N} \frac{k!}{k_A! k_B!} q_{A|B}^{k_A} q_{B|B}^{k_B} \underbrace{k_A^2}_{=k_A^2 - k_A + k_A} \tag{3.53}$$

$$- \sum_{k_A=0}^{k} -\frac{2p_A}{k^2 N} \frac{k!}{k_A! k_B!} q_{A|A}^{k_A} q_{B|A}^{k_B} k_A k_B + O(w) \tag{3.54}$$

$$\tag{3.55}$$

$$= \frac{2p_B}{k^2 N}(k(k-1) q_{A|B} q_{B|A} + k q_{A|B}) \tag{3.56}$$

$$- \frac{2p_B}{k^2 N} k(k-1) q_{A|A} q_{B|A} + O(w) \tag{3.57}$$

$$\tag{3.58}$$

$$= \frac{2}{kN} p_{AB}(((k-1) q_{A|B} + 1 - (k-1) q_{A|A}) + O(w) \tag{3.59}$$

$$\tag{3.60}$$

$$= \frac{2}{kN} \left[1 + (k-1)(q_{A|B} - q_{A|A})\right] + O(w). \tag{3.61}$$

From (3.44) and (3.1) we get that

$$\dot{q_{A|A}} = \frac{d}{dt}\left(\frac{p_{AA}}{p_A}\right)$$

$$= \frac{\dot{p_{AA}} p_A - p_{AA} \dot{p_A}}{p_A^2}$$

$$= \frac{\left[\frac{2}{kN} p_{AB} \left[1 + (k-1)(q_{A|B} - q_{A|A})\right] + O(w)\right] p_A}{p_A^2}$$

$$- \frac{p_{AA}\left[w\frac{k-1}{N}p_{AB}(I_a a + I_b b - I_c c - I_d d) + O(w^2)\right]}{p_A^2}$$

$$= \frac{\left[\frac{2}{kN}p_{AB}\left[1+(k-1)(q_{A|B}-q_{A|A})\right]+O(w)\right]}{p_A} - \underbrace{\frac{O(w^2)}{p_A^2}}_{=O(w^2)}$$

$$= \frac{\left[\frac{2}{kN}p_{AB}\left[1+(k-1)(q_{A|B}-q_{A|A})\right]\right]}{p_A} + \underbrace{\frac{O(w)}{p_A}}_{=O(w)} + O(w^2)$$

$$= \frac{2}{kN}\frac{p_{AB}}{p_A}\left[1+(k-1)(q_{A|B}-q_{A|A})\right] + \underbrace{O(w)+O(w^2)}_{=O(w)}$$

$$= \frac{2}{kN}\frac{p_{AB}}{p_A}\left[1+(k-1)(q_{A|B}-q_{A|A})\right] + O(w).$$

Summarizing we have obtained so far that

$$\dot{p_A} = \frac{k-1}{N}p_{AB}(I_a a + I_b b - I_c c - I_d d) + O(w^2)$$

and

$$\dot{q_{A|A}} = \frac{2}{kN}\frac{p_{AB}}{p_A}\left[1+(k-1)(q_{A|B}-q_{A|A})\right] + O(w)$$

Since the system is described by p_A and $q_{A|A}$ we can rewrite the previous results as

$$\dot{p_A} = wF_1(p_A, q_{A|A}) + O(w^2) \tag{3.62}$$

$$\dot{q_{A|A}} = F_2(p_A, q_{A|A}) + O(w). \tag{3.63}$$

So, in the case $w << 1$(weak selection)$q_{A|A}$ equilibrates faster than p_A. Hence, the dynamical system converges to the manifold defined by $F_2(p_A, q_{A|A}) = 0$, which is equivalent to

$$q_{A|A} = p_A + \frac{1}{k-1}(1-p_A)$$

then, using (3.1) we obtain that

$$q_{A|A} - q_{A|B} = \frac{1}{k-1} \tag{3.64}$$

$$q_{B|B} - q_{B|A} = \frac{1}{k-1} \tag{3.65}$$

We will show the first equality because it is not obvious at all. First, notice that the equality holds if and only if

$$q_{A|A} - q_{A|B} = \frac{q_{A|A} - p_A}{1 - p_A}.$$

Also, notice that

$$q_{A|A} - q_{A|B} = \frac{p_{AA}}{p_A} - \frac{p_{AB}}{p_B} = \frac{p_{AA}}{p_A} - \frac{p_{AB}}{1 - p_A} \tag{3.66}$$

$$= \frac{p_{AA}(1 - p_A) - p_{BA}p_A}{p_A(1 - p_A)} \tag{3.67}$$

so we get the following chain of algebraic equivalences

$$q_{A|A} - q_{A|B} = \frac{q_{A|A} - p_A}{1 - p_A} \tag{3.68}$$

$$\iff \frac{q_{A|A} - p_A}{1 - p_A} = \frac{p_{AA}(1 - p_A) - p_{BA}p_A}{p_A(1 - p_A)} \tag{3.69}$$

$$\iff q_{A|A} - p_A = \frac{p_{AA}(1 - p_A) - p_{BA}p_A}{p_A} \tag{3.70}$$

$$\iff \frac{p_{AA} - p_A^2}{p_A} = \frac{p_{AA} - p_{AA}p_A - p_{BA}p_A}{p_A} \tag{3.71}$$

$$\iff p_{AA} - p_A^2 = p_{AA} - p_{AA}p_A - p_{BA}p_A \tag{3.72}$$

$$\iff p_A^2 = p_{AA}p_A + p_{BA}p_A \tag{3.73}$$

$$\iff p_A^2 = p_A(p_{AA} + p_{BA}) \tag{3.74}$$

$$\iff p_A = p_{AA} + p_{BA} \tag{3.75}$$

$$\iff p_A = q_{A|A}p_A + q_{B|A}p_A \tag{3.76}$$

$$\iff 1 = q_{A|A} + q_{B|A}. \tag{3.77}$$

Where the last term is stated in (3.1). The second equality in (3.64) follows similarly.

We now should give some interpretation to (3.64). So, among the $k-1$ neighbours, an A type has on average one more A neighbours than a B type has A neighbours. This relationship will lead us to a rule to decide when the game is favouring cooperation.

To make things a little bit more tractable, we assume that the relation

$$q_{A|A} = p_A + \frac{1}{k-1}(1 - p_A)$$

holds, and we then study the diffusion process over p_A only. This relation was established previously by assuming that $q_{A|A}$ stabilizes much more faster than p_A so that $\dot{q_{A|A}} = 0$ when we assume w small. This is now a one dimensional problem. This assumption, combined with (3.1), leads to the following identities that will be of great use in the next step

$$q_{B|A} = \frac{k-2}{k-1}(1-p_A) \tag{3.78}$$

$$\tag{3.79}$$

$$q_{A|B} = \frac{k-2}{k-1}p_A \tag{3.80}$$

$$\tag{3.81}$$

$$q_{B|B} = 1 - \frac{k-2}{k-1}p_A \tag{3.82}$$

$$\tag{3.83}$$

$$p_{AB} = p_{BA} = \frac{k-2}{k-1}p_A(1-p_A) \tag{3.84}$$

$$\tag{3.85}$$

$$q_{A|A} + q_{B|B} = \frac{k}{k-1}. \tag{3.86}$$

By considering a short time interval, Δt, we obtain that

$$\mathbb{E}(\Delta p_A) \approx w\frac{k-2}{k(k-1)N}p_A(1-p_A)(\alpha p_A + \beta)\Delta t := m(p_A)\Delta t \tag{3.87}$$

$$Var(\Delta p_A) \approx \frac{2(k-2)}{N^2(k-1)}p_A(1-p_A)\Delta t := v(p_A)\Delta t \tag{3.88}$$

Where

$$\alpha = (k+1)(k-2)(a-b-c+d)$$

$$\beta = (k+1)a + (k^2-k-1)b - c - (k^2-1)d.$$

Let us now justify the approximations for $\mathbb{E}(\Delta p_A)$ and $Var(\Delta p_A)$. Both are the result of using the definition of expected value and variance of a random variable, and ignoring the quadratic terms ($O(w^2)$) for the mean and the linear terms for the variance ($O(w)$).The reason for neglecting higher order terms will become apparent later on. So, the corresponding calculations for the mean are

$$
\begin{aligned}
\mathbb{E}(\Delta p_A) &= \frac{1}{N}P(\Delta p_A = \frac{1}{N}) - \frac{1}{N}P(\Delta p_A = -\frac{1}{N}) \\
&= w\frac{k-1}{N}p_{AB}(I_a a + I_b b - I_c c - I_d d) + O(w^2) \\
&\approx w\frac{k-1}{N}p_{AB}(I_a a + I_b b - I_c c - I_d d) \\
&= w\frac{k-1}{N}p_{AB}\left[\frac{1}{k}\left((k-1)q_{B|B}q_{A|A} + (k-1)q_{A|A}q_{A|A} + q_{A|A}\right)\right]a \\
&+ w\frac{k-1}{N}p_{AB}\left[\frac{1}{k}\left((k-1)q_{B|B}q_{B|A} + (k-1)q_{A|A}q_{B|A} + q_{B|B}\right)\right]b \\
&- w\frac{k-1}{N}p_{AB}\left[\frac{1}{k}\left((k-1)q_{B|B}q_{A|B} + (k-1)q_{A|A}q_{A|B} + q_{A|A}\right)\right]c \\
&- w\frac{k-1}{N}p_{AB}\left[\frac{1}{k}\left((k-1)q_{B|B}q_{B|B} + (k-1)q_{A|A}q_{B|B} + q_{B|B}\right)\right]d \\
&= w\frac{k-1}{N}p_A(1-p_A)\frac{k-2}{k-1}\left[\frac{k+1}{k}(p_A + \frac{1}{k+1}(1-p_A))\right]a \\
&+ w\frac{k-1}{N}p_A(1-p_A)\frac{k-2}{k-1}\left[\frac{k-2}{k-1}(1-p_A) + \frac{1}{k}(1 - \frac{k-2}{k-1}p_A)\right]b \\
&- w\frac{k-1}{N}p_A(1-p_A)\frac{k-2}{k-1}\left[\frac{k-2}{k-1}p_A + \frac{1}{k}(p_A + \frac{1}{k-1}(1-p_A))\right]c \\
&- w\frac{k-1}{N}p_A(1-p_A)\frac{k-2}{k-1}\left[\frac{k+1}{k}(1 - \frac{k-2}{k-1}p_A)\right]d \\
&= w\frac{k-2}{N}p_A(1-p_A)\left[\frac{k+1}{k(k-1)}((k-2)p_A + 1)\right]a \\
&+ w\frac{k-2}{N}p_A(1-p_A)\left[\frac{1}{k(k-1)}((k(k-2)(1-p_A)) + (k-1) - (k-2)p_A)\right]b
\end{aligned}
$$

$$- w\frac{k-2}{N}p_A(1-p_A)\left[\frac{(k+1)(k-2)p_A+1}{k(k-1)}\right]c$$

$$- w\frac{k-2}{N}p_A(1-p_A)\left[\frac{(k+1)(k-1)-(k+1)(k-2)p_A}{k(k-1)}\right]d$$

$$= w\frac{k-2}{N}p_A(1-p_A)\left[\frac{k+1}{k(k-1)}((k-2)p_A+1)\right]a$$

$$+ w\frac{k-2}{N}p_A(1-p_A)\left[\frac{1}{k(k-1)}(k^2-k-1-(k+1)(k-2)p_A)\right]b$$

$$+ w\frac{k-2}{N}p_A(1-p_A)\left[\frac{1}{k(k-1)}-(k+1)(k-2)p_A-1\right]c$$

$$+ w\frac{k-2}{N}p_A(1-p_A)\left[\frac{1}{k(k-1)}((k+1)(k-2)p_A-(k^2-1))\right]d$$

$$= w\frac{k-2}{k(k-1)N}p_A(1-p_A)\left[(k+1)a((k-2)p_A+1)\right]$$

$$+ w\frac{k-2}{k(k-1)N}p_A(1-p_A)\left[-(k+1)(k-2)p_Ab+(k^2-k-1)b\right]$$

$$+ w\frac{k-2}{k(k-1)N}p_A(1-p_A)\left[-(k+1)(k-2)p_A-1\right]c$$

$$+ w\frac{k-2}{k(k-1)N}p_A(1-p_A)\left[((k+1)(k-2)p_A-(k^2-1))\right]d$$

$$= w\frac{k-2}{k(k-1)N}p_A(1-p_A)(\alpha p_A+\beta)$$

On the other hand, the variance is obtained in the following way:

$$Var(\Delta p_A) = \frac{1}{N^2} P(\Delta p_A = \frac{1}{N}) + \frac{1}{N^2} P(\Delta p_A = -\frac{1}{N}) \tag{3.89}$$

$$\tag{3.90}$$

$$= \frac{2(k-2)}{N^2(k-1)} p_A(1-p_A) + O(w) \tag{3.91}$$

$$\tag{3.92}$$

$$\approx \frac{2(k-2)}{N^2(k-1)} p_A(1-p_A). \tag{3.93}$$

This calculations basically are the same as in (3.19) but instead of $\frac{1}{N}$ and $-\frac{1}{N}$ we use $\frac{1}{N^2}$ in both cases respectively. Hence, we obtain that the constant part is given by

$$\frac{1}{N^2}(p_{AB} + p_{BA}) = \frac{2}{N^2} p_{AB} = \frac{2(k-2)}{N^2(k-1)} p_A(1-p_A)$$

by using the identities in (3.78). We are going to see that for the purpose of this project, the part of the variance which has w in front is not necessary for the approximation, since we will need it for a differential equation in which this term is not relevant. This fixation probability fits in the population genetics model described in the previous chapter. Hence, the fixation probability, $\phi_A(y)$ of a type A with initial frequency $p_A(t=0) = y$ satisfies the Kolmogorov backward equation:

$$0 = \mathbb{E}(y) \frac{d}{dy} \phi_A(y) + \frac{Var(y)}{2} \frac{d^2}{dy^2} \phi_A(y) \tag{3.94}$$

Where $\mathbb{E}(y) = w \frac{k-2}{k(k-1)N} y(1-y)(\alpha y + \beta) + O(w^2)$ and $Var(y) = \frac{2(k-2)}{N^2(k-1)} y(1-y) + O(w)$. From the approximations previously enunciated, the equation can be replaced by:

$$0 = m(y) \frac{d}{dy} \phi_A(y) + \frac{v(y)}{2} \frac{d^2}{dy^2} \phi_A(y) \tag{3.95}$$

This approximation makes sense since w is chosen to be small. What we are doing is to ignore all the terms that are $O(w^2)$. The previous ODE is separable. The problem is that when we try and solve it by algebraic calculations, we encounter the integral $int_0^x e^{t^2} dt$ which we know does not qualify for the fundamental theorem of calculus. Even if we approximate it, we would not get a good approximation for obtaining a decision rule in terms of the parameters a, b, c, d, k and N. Hence, we will again need an approximate solution. We are mainly interested in a polinomial approximation that very well takes into consideration the boundary conditions, and works for interpolation.

So, by assuming weak selection (very small w) we have the following as an approximate solution

$$\phi_A(y) = y + w\frac{N}{6k}y(1-y)((\alpha+3\beta)+\alpha y). \tag{3.96}$$

This approximation will now be deduced. First of all, we want to justify why the $O(w)$ terms in (3.89) are omitted in that approximation that leads to $v(y)$. We may assume, for practical reasons, that $\phi_A(y)$ can be approximated by $\overline{\phi_A^{(w)}}(y)$, its third order Taylor approximation in the polynomial basis

$$B = \{1, y, y(1-y), y^2(1-y)\}.$$

This basis is quite suitable when using the boundary conditions. The notation $\phi_A^{(w)}(y)$ means that the fixation probability depends both on y and on w. So we suppose that

$$\phi_A^{(\bar{w})}(y) = a_0^{(w)} + a_1^{(w)}y + a_2^{(w)}y(1-y) + a_3^{(w)}y^2(1-y).$$

Here

$$a_i^{(w)} = a_{i0} + a_{i1}w + a_{i2}w^2 + ...$$

Now, because of the boundary conditions of $\phi_A(y)$, for all w we have that $\phi_A^{(w)}(0) = 0$ and that $\phi_A^{(w)}(1) = 1$, which implies that $a_0^{(w)} = 0$ and that $a_1^{(w)} = 1$ for all w. Also, notice that for $w = 0$ in (3.95) we obtain that

$$0 = \underbrace{\frac{v(y)}{2}}_{\neq 0}\frac{d^2}{dy^2}\phi_A^{(0)}(y)$$

which implies that

$$\frac{d^2}{dy^2}\phi_A^{(0)}(y) = 0,$$

so that $\phi_A^{(0)}(y)$ is a polynomial of degree 1. Hence, $a_{i0} = 0$ for all $i \geq 2$ (This means that $a_i^{(w)} = O(w)$ for all $i \geq 2$). Thus, we obtain that

$$\phi_A^{(w)}(y) = y + w(a_{21}y(1-y) + a_{31}y^2(1-y) + ...) + O(w^2)$$

This implies that $\frac{d^2\phi_A(y)}{dy^2} = O(w)$. So, if we place this result in (3.94) we would get

$$0 = \mathbb{E}(y)\frac{d}{dy}\phi_A(y) + \frac{Var(y)}{2}\frac{d^2}{dy^2}\phi_A(y)$$

$$\Longrightarrow 0 = \mathbb{E}(y)\frac{d}{dy}\phi_A(y) + \frac{v(y)+O(w)}{2}\frac{d^2}{dy^2}\phi_A(y)$$

$$\Longrightarrow 0 = \mathbb{E}(y)\frac{d}{dy}\phi_A(y) + \frac{v(y)}{2}\frac{d^2}{dy^2}\phi_A(y) + \frac{O(w)}{2}\underbrace{\phi_A(y)}_{=O(w)}\frac{d^2}{dy^2}\phi_A(y)$$

$$\Longrightarrow 0 = (m(y)+O(w^2))\frac{d}{dy}\phi_A(y) + \frac{v(y)}{2}\frac{d^2}{dy^2}\phi_A(y) + O(w^2)$$

$$\Longrightarrow 0 = m(y)\frac{d}{dy}\phi_A(y) + \frac{v(y)}{2}\frac{d^2}{dy^2}\phi_A(y)$$

This is enough justification to solve (3.95) instead of (3.94), when w is small. Now we are ready to show where the approximation (3.96) comes from. So, as said before and considering some of the previous calculations, we can estimate $\phi_A^{(w)}(y)$ with

$$\phi_A^{\overline{(w)}}(y) = y + a_2^{(w)}y(1-y) + a_3^{(w)}y^2(1-y).$$

From now on, we will denote $\phi_A^{\overline{(w)}}(y)$ by $\phi_A^{(w)}(y)$. Also, we can notice that

$$\frac{d}{dy}\phi_A^{(w)}(y) = 1 + a_2^{(w)} + (-2a_2^{(w)} - a_3^{(w)})y + 3a_3^{(w)}y(1-y)$$

$$\frac{d^2}{dy^2}\phi_A^{(w)}(y) = -2a_2^{(w)} + 2a_3^{(w)} - 6a_3^{(w)}y$$

So, the differential equation (3.95) becomes

$$\begin{aligned}0 &= Mwy(1-y)(\alpha y+\beta)(1 + a_2^{(w)} + (-2a_2^{(w)} - a_3^{(w)})y + 3a_3^{(w)}y(1-y)) \\ &+ Vy(1-y)(-2a_2^{(w)} + 2a_3^{(w)} - 6a_3^{(w)}y)\end{aligned}$$

Where $M = \frac{k-2}{k(k-1)N}$ and $V = \frac{k-2}{N^2(k-1)}$. Notice that $\frac{M}{V} = \frac{N}{k}$. Now, since this equation is not so easy to deal with, we will aim to get the approximate solution by plugging in some evenly spread values for $y \in (0,1)$ and then solving for both $a_2^{(w)}$ and $a_3^{(w)}$. Since the equation has degree 3 we will need to plug in four values. We will let $y = 0$, $y = 1$, $y = \frac{1}{3}$ and $y = \frac{2}{3}$.

For $y = \frac{1}{3}$ we get

$$
\begin{aligned}
&0 = Mw\frac{1}{3}\frac{2}{3}(\frac{\alpha}{3}+\beta)(1+\underbrace{a_2^{(w)}+(-2a_2^{(w)}-a_3^{(w)})\frac{1}{3}+3a_3^{(w)}\frac{1}{3}\frac{2}{3}}_{=O(w)})\\
&\quad + V\frac{1}{3}\frac{2}{3}(-2a_2^{(w)}+2a_3^{(w)}-6a_3^{(w)}\frac{1}{3})\\
\Longrightarrow &0 = Mw\frac{2}{9}(\frac{\alpha}{3}+\beta)+\underbrace{Mw\frac{2}{9}O(w)}_{=O(w^2)}\\
&\quad - V\frac{2}{9}(2a_2^{(w)})\\
\Longrightarrow &0 = Mw(\alpha+3\beta)-6Va_2^{(w)}\\
\Longrightarrow &a_2^{(w)} = \frac{Nw}{6k}(\alpha+3\beta)
\end{aligned}
$$

So, we have established an approximation for $a_2^{(w)}$. By taking $y = \frac{2}{3}$ we hope to obtain an approximation for $a_3^{(w)}$. By taking $y = \frac{2}{3}$ we get

$$0 = Mw\frac{2}{3}\frac{1}{3}(\frac{2\alpha}{3} + \beta)(1 + \underbrace{a_2^{(w)} + (-2a_2^{(w)} - a_3^{(w)})\frac{2}{3} + 3a_3^{(w)}\frac{2}{3}\frac{1}{3}}_{=O(w)})$$

$$+ V\frac{2}{3}\frac{1}{3}(-2a_2^{(w)} + 2a_3^{(w)} - 6a_3^{(w)}\frac{2}{3})$$

$$\Longrightarrow 0 = Mw\frac{2}{9}(\frac{2\alpha}{3} + \beta) + \underbrace{Mw\frac{2}{9}O(w)}_{=O(w^2)}$$

$$- V\frac{2}{9}(-2a_2^{(w)} - 2a_3^{(w)})$$

$$\Longrightarrow 2a_2^{(w)} + 2a_3^{(w)} = \frac{Mw}{V}(\frac{2\alpha}{3} + \beta)$$

$$\Longrightarrow 2a_3^{(w)} = \frac{Mw}{V}(\frac{2\alpha}{3} + \beta) - \frac{Mw}{2V}(\frac{\alpha}{3} + \beta)$$

$$\Longrightarrow a_3^{(w)} = \frac{Mw}{V}(\frac{\alpha}{6})$$

$$\Longrightarrow a_3^{(w)} = \frac{Nw}{6k}\alpha$$

Thus, the deduction of (3.96) is now complete. We are now interested in how good the approximation is. Hence we have the following theorem:

Theorem 8. *Let $\phi_A(y)$ be the exact solution of the differential equation*

$$0 = m(y)\frac{d}{dy}\phi_A(y) + \frac{v(y)}{2}\frac{d^2}{dy^2}\phi_A(y)$$

and boundary conditions $\phi_A(0) = 0$ and $\phi_A(1) = 1$ and where $m(y)$ and $v(y)$ are as described above. Let $\hat{\phi}_A(y) = y + w\frac{N}{6k}y(1-y)((\alpha + 3\beta) + \alpha y)$. Then for $y = \frac{1}{N}$ we have that $|\phi_A(y) - \hat{\phi}_A(y)| \longrightarrow 0$ as $w \longrightarrow 0$ and $N \longrightarrow \infty$.

Proof. We first notice that

$$0 = m(y)\frac{d}{dy}\phi_A(y) + \frac{v(y)}{2}\frac{d^2}{dy^2}\phi_A(y)$$

is a separable differential equation if we use the substitution $u(y) = \frac{d}{dy}\phi_A(y)$. Through simple algebra one obtains that the exact solution to the ODE is given by

$$\phi_A(y) = \frac{\int_0^y e^{-\frac{w\alpha}{2kN}t^2 - \frac{w\beta}{kN}t} dt}{C}$$

where $C = \int_0^1 e^{-\frac{w\alpha}{2kN}t^2 - \frac{w\beta}{kN}t} dt$. Hence,

- $\frac{d}{dt}\phi_A(y) = \frac{e^{-\frac{w\alpha}{2kN}y^2 - \frac{w\beta}{kN}y}}{C}$
- $\frac{d^2}{dt^2}\phi_A(y) = \frac{-e^{-\frac{w\alpha}{2kN}y^2 - \frac{w\beta}{kN}y}(\frac{w\alpha y}{kN} + \frac{w\beta}{kN})}{C}$
- $\frac{d^n}{dt^n}\phi_A(y) = O(w^{n-1})$ for $n > 2$

So, if we assume that $\phi_A(y)$ is analytic at 0 and we take the Taylor series at 0, we can choose N large enough such that $y = \frac{1}{N}$ is in the interval of convergence of the series. In this case we have that

$$\phi_A(y) = \frac{d}{dt}\phi_A(0)y + \frac{d^2}{dt^2}\phi_A(0)\frac{y^2}{2} + O(w^2).$$

It is important to point out that $\frac{d}{dt}\phi_A(0) = \frac{1}{C}$ and $\frac{d^2}{dt^2}\phi_A(0) = -\frac{w\beta}{kNC}$. With these calculations in mind we are now ready to begin our proof. Let $y = \frac{1}{N}$, then

$$|\phi_A(y) - \hat{\phi_A}(y)| = |\frac{d}{dt}\phi_A(0)y + \frac{d^2}{dt^2}\phi_A(0)\frac{y^2}{2} + O(w^2) - y + w\frac{N}{6k}y(1-y)((\alpha+3\beta)+\alpha y)|$$

$$\leq |\frac{y}{C} - \frac{w\beta y^2}{2kNC} - y + \frac{w}{6k}\underbrace{(1-y)}_{\leq 1}((\alpha+3\beta)+\alpha y)| + |O(w^2)|$$

$$\leq \frac{1}{NC} + \frac{w\beta}{2kN^3C} + \frac{1}{N} + \frac{w}{6k}(\alpha+3\beta) + \frac{\alpha}{N}) + O(w^2) \longrightarrow 0$$

when $w \longrightarrow 0$ and $N \longrightarrow \infty$. This finishes our proof.

□

3.1.2 Fixation Probabilities

Now, the probability that in a population a single individual (type A) takes the entire population ($N-1$ players of type b) by is given by $\rho_A = \phi_A(\frac{1}{N})$. By (3.96) we obtain that for large N, $\rho_A > \frac{1}{N} \iff \alpha + 3\beta > 0$ which is equivalent to

$$(k^2+2k+1)a + (2k^2-2k-1)b > (k^2-k+1)c + (2k^2+k-1)d. \quad (3.97)$$

Similarly, $\rho_B > \frac{1}{N}$ is equivalent to

$$(k^2+2k+1)d + (2k^2-2k-1)c > (k^2-k+1)b + (2k^2+k-1)a. \quad (3.98)$$

So, if we consider a large enough k, then (3.97) can be approximated by $a+2b > c+2d$. Hence, if we are in a game where strategies A and B are best replies to themselves: $a > c$ and $d > b$, we obtain that the **replicator equation** for the frequency of A has an unstable equilibrium at the point $\frac{(d-b)}{(a-b-c+d)}$. This was obtained by plugging in the replicator equation (2.12) the payoff matrix and using the relation $x = 1-y$ where $\mathbf{x} = (x, y)$. Then the Jacobian was obtained and Hartman-Grobman's theorem as used to determine that such equilibrium point is unstable. Notice that this equilibrium point only makes sense when $(a-b-c+d) \neq 0$.

3.1.3 The Rule $\frac{b}{c} > k$

In Nowak's words: "A fundamental aspect of all biological systems is cooperation. Cooperative interactions are required for many levels of biological organization ranging from single cells to groups of animals. Human society is based to a large extent on mechanisms that promote cooperation. It is well known that in unstructured populations, natural selection favours defectors over cooperators. There is much current interest, however, in studying evolutionary games in structured populations and on graphs. These efforts recognize the fact that who-meets-whom is not random, but determined by spatial relationships or social networks. Here we describe a surprisingly simple rule that is a good approximation for all graphs that we have analysed, including cycles, spatial lattices, random regular graphs, random graphs and scale free networks: natural selection favours cooperation, if the benefit of the altruistic act, b, divided by the cost, c, exceeds the average number of neighbours, k, which means $\frac{b}{c} > k$."
Now let's consider a cooperative game, also known as the Prisoner's dilemma, where A cooperates and B is a defector. So we have the payoff matrix given by

$$\begin{array}{ccc} & A & B \\ A & b-c & -c \\ B & b & 0 \end{array}$$

This matrix describes a situation where cooperators give up c whenever they play and receive b when they encounter another cooperator. Hence, when a defector encounters a cooperator he earns b but does not pay anything. This is known as the prisoner's dilemma because it models a game where two prisoners are held captive and one is innocent and the other one is guilty. Here the guilty one has no reason to cooperate, since either way he does not loose, whereas the innocent one will be willing to cooperate in order to get a smaller sentence. This is a very famous game in game theory.

Continuing, we substitute these values in (3.97) and (3.98) and conclude that if $\frac{b}{c} > k$ then $\rho_A > \frac{1}{N} > \rho_B$ and vice versa, if $\frac{b}{c} < k$ then $\rho_A < \frac{1}{N} < \rho_B$. Thus, in a large population with weak selection, cooperation is favoured if and only if

$$\frac{b}{c} > k.$$

Notice that in this case $a-b-c+d=0$, hence the game theoretical approach will be useless in this particular problem, still the diffusion approximation gave us a rule with a while different idea.

3.2 IM-Updating

We will now begin our study based on a different update rule. Here, a random player is chosen to compare his payoff with his neighbours, then he chooses if he keeps his strategy or imitates a neighbour depending on which has the highest payoff. The main difference with the previous updating rule is that in IM updating, the payoff of the random selected players matters, whereas in DB updating did not, it was solely based on the fitness of his neighbours. This rule is expected to work in the benefit of defectors, since defectors at the boundary of a cluster have a higher payoff than cooperators, therefore defectors have no incentive to change their strategy.

So, this is how the conditions change subject to this updating rule. The fitness of a $B-$type player with k_A many $A-$type and k_B many $B-$type neighbours is given by

$$f_0 = 1 - w + w(k_A c + k_B d).$$

The probability that this B player will adopt strategy A is given by

$$\frac{k_A f_A}{k_A f_A + k_B f_B + f_0}.$$

This is just the split of the two case, either he stays in B or he changes to A, and his own fitness matters, so the f_0 must be taken into account. On the other hand, the fitness of an $A-$type player, under the same notation, is given by

$$g_0 = 1 - w + w(k_A a + k_B b).$$

Similarly, the probability that the $A-$player will adopt strategy B is given by

$$\frac{k_B g_B}{k_A g_A + k_B g_B + g_0}.$$

Thus, using the exact same reasoning as in the beginning of the previous section, we obtain that

$$P(\Delta p_A = \frac{1}{N}) = p_B \sum_{k_A+k_B=k} \left(\frac{k!}{k_A!k_B!}\right) q_{A|B}^{k_A} q_{B|B}^{k_B} \frac{k_A f_A}{k_A f_A + k_B f_B + f_0}. \quad (3.99)$$

$$P(\Delta p_{AA} = \frac{2k_A}{kN}) = p_B \left(\frac{k!}{k_A!k_B!}\right) q_{A|B}^{k_A} q_{B|B}^{k_B} \frac{k_A f_A}{k_A f_A + k_B f_B + f_0}. \quad (3.100)$$

$$P(\Delta p_A = -\frac{1}{N}) = p_A \sum_{k_A+k_B=k} \left(\frac{k!}{k_A!k_B!}\right) q_{A|A}^{k_A} q_{B|A}^{k_B} \frac{k_B g_B}{k_A g_A + k_B g_B + g_0}. \quad (3.101)$$

$$P(\Delta p_{AA} = -\frac{2k_A}{kN}) = p_A \left(\frac{k!}{k_A!k_B!}\right) q_{A|A}^{k_A} q_{B|A}^{k_B} \frac{k_B g_B}{k_A g_A + k_B g_B + g_0}. \qquad (3.102)$$

Now we are going to proceed in a similar manner as in the previous section in order to deduce similar results using the diffusion approximation technique. First we will establish some results in order to proceed to deeper calculations of the transition probabilities of p_A and p_{AA}.

Notice that since $k_A f_A + k_B f_B + f_0 = k + 1 + w(k_A(\bigstar_A + c) + k_B(\bigstar_B + d) - 1$ we have that

$$\frac{k_A f_A}{k_A f_A + k_B f_B + f_0}$$

$$= \frac{k_A}{k+1}$$

$$+ w\frac{k_A}{(k+1)^2}((k_B+1)\bigstar_A - k_A c - k_B(\bigstar_B + d) + 1)$$

$$+ O(w^2)$$

Similarly, we obtain that

$$\frac{k_B g_B}{k_A g_A + k_B g_B + g_0} = \frac{k_B}{k+1} +$$

$$w\frac{k_B}{(k+1)^2}((k_A+1)(\bigstar_B - d + c) - k_A(\bigstar_A - b + 2a) - k_B(\bigstar_B - d + c + b) + 1)$$

$$+ O(w^2)$$

We are now set and ready to deal with the diffusion approximation approach. It is important to point out that since the calculations are so similar to the ones in the previous section, the level of detail will decrease a bit.

3.2.1 Diffusion Approximation

$$\dot{p_A} = \frac{1}{N}P(\Delta p_A = \frac{1}{N}) - \frac{1}{N}P(\Delta p_A = -\frac{1}{N})$$

$$= \frac{1}{N}p_B \sum_{k_A+k_B=k} \frac{k!}{k_A!k_B!} q_{A|B}^{k_A} q_{B|B}^{k_B} \{ \frac{k_A}{k+1}$$

$$+ w \frac{((k_A(k_B+1)(\bigstar_A - k_A^2 c - k_A k_B(\bigstar_B + d) + k_A))}{(k+1)^2} \}$$

$$- \frac{1}{N}p_A \sum_{k_A+k_B=k} \frac{k!}{k_A!k_B!} q_{A|A}^{k_A} q_{B|A}^{k_B} (\frac{k_B}{k+1}$$

$$+ w \frac{((k_B(k_A+1)(\bigstar_B - d + c) - k_B k_A^2(\bigstar_A - b + 2a) - k_A^2 b + k_B))}{(k+1)^2}$$

$$+ O(w^2)$$

$$= \frac{w}{N(k+1)^2} p_B (k(k-1) q_{A|B} q_{B|B} \bigstar_A + k q_{A|B} \bigstar_A$$

$$- k(k-1) q_{A|B}^2 c - k(k-1) q_{A|B} q_{B|B} (\bigstar_B + d + k q_{A|B})))$$

$$- \frac{w}{N(k+1)^2} p_A (k(k-1) q_{A|A} q_{B|A} (\bigstar_B - d + c) + k q_{B|A} (\bigstar_B - d + c)$$

$$- k(k-1) q_{B|A}^2 b - k(k-1) q_{A|A} q_{B|A} (\bigstar_A - b + 2a) + k q_{B|A})))$$

$$+ O(w^2)$$

$$= w \frac{k}{N(k+1)^2} p_{AB} ((k-1) q_{B|B} ((k-1) q_{A|A} a + ((k-1) q_{B|A} + 1) b - 1)$$

$$+ ((k-1) q_{A|A} a + ((k-1) q_{B|A} + 1) b - 1)$$

$$- (k-1) q_{A|B} c$$

$$- (k-1)q_{B|B}((k-1)q_{A|B}c + ((k-1)q_{B|B}+2)d - 1)$$

$$- (k-1)q_{A|A}(((k-1)q_{A|B}+1)c + (k-1)q_{B|B}d - 1)$$

$$- ((k-1)q_{A|B}+1)c + (k-1)q_{B|B}d - 1)$$

$$+ (k-1)q_{A|A}(((k-1)q_{A|A}+2)a + (k-1)q_{B|A}b - 1)$$

$$+ (k-1)q_{B|A}b + O(w^2)$$

$$= \frac{k}{N(k+1)^2} p_{AB}(a((k-1)^2 q_{A|A}(q_{A|A}+q_{B|B}) + 3(k-1)q_{A|A})$$

$$+ b((k-1)^2 q_{B|A}(q_{A|A}+q_{B|B}) + (k-1)q_{B|B} + 2(k-1)q_{B|A} + 1)$$

$$- c((k-1)^2 q_{A|B}(q_{A|A}+q_{B|B}) + (k-1)q_{A|A} + 2(k-1)q_{A|B} + 1)$$

$$- d((k-1)^2 q_{B|B}(q_{A|A}+q_{B|B}) + 3(k-1)q_{B|B})) + O(w^2)$$

On the other hand, we get

$$\dot{p_{AA}} = \sum_{k_A+k_B=k} \frac{2k_A}{kN} P(\Delta p_{AA} = \frac{2k_A}{kN}) - \frac{2k_A}{kN} P(\Delta p_{AA} = -\frac{2k_A}{kN})$$

$$= \sum_{k_A+k_B=k} \frac{2k_A}{kN} p_B \frac{k!}{k_A!k_B!} q_{A|B}^{k_A} q_{B|B}^{k_B} \frac{k_A}{k+1}$$

$$- \sum_{k_A+k_B=k} \frac{2k_A}{kN} p_A \frac{k!}{k_A!k_B!} q_{A|A}^{k_A} q_{B|A}^{k_B} \frac{k_B}{k+1} + O(w)$$

$$= \frac{2}{k(k+1)N}(p_B k(k-1)q_{A|B}^2 + p_B k q_{A|B}$$

$$- p_A k(k-1) q_{A|A} q_{B|A}) + O(w)$$

$$= \frac{2}{(k+1)N} p_{AB}(1 + (k-1)(q_{A|B} - q_{A|A})) + O(w)$$

We are then ready to calculate $\dot{q_{A|A}}$.

$$\dot{q_{A|A}} = \frac{d}{dt}\left(\frac{p_{AA}}{p_A}\right) = \frac{p_A\dot{p_{AA}} - p_{AA}\dot{p_A}}{p_A^2}$$

$$= \frac{\dot{p_{AA}}}{p_A} + O(w)$$

$$= \frac{2}{(k+1)N}\frac{p_{AB}}{p_A}\big(1 + (k-1)(q_{A|B} - q_{A|A})\big) + O(w)$$

Since the system is described by p_A and $q_{A|A}$ we can rewrite the previous results as

$$\dot{p_A} = wF_1(p_A, q_{A|A}) + O(w^2)$$
$$\dot{q_{A|A}} = F_2(p_A, q_{A|A}) + O(w).$$

So, in the case $w << 1$(weak selection)$q_{A|A}$ equilibrates faster than p_A. Hence, the dynamical system converges to the manifold defined by $F_2(p_A, q_{A|A}) = 0$, which is equivalent to

$$q_{A|A} = p_A + \frac{1}{k-1}(1 - p_A).$$

Here the identities in (3.78) hold also and we are now able to begin our calculation of $\mathbb{E}(\Delta p_A)$ and $Var(\Delta p_A)$.

$$\mathbb{E}(\Delta p_A) = \frac{1}{N}P(\Delta p_A = \frac{1}{N}) - \frac{1}{N}P(\Delta p_A = -\frac{1}{N})$$

$$\approx \frac{k}{N(k+1)^2}p_{AB}\{a((k-1)^2 q_{A|A}(q_{A|A} + q_{B|B}) + 3(k-1)q_{A|A})$$

$$+ b((k-1)^2 q_{B|A}(q_{A|A} + q_{B|B}) + (k-1)q_{B|B} + 2(k-1)q_{B|A} + 1)$$

$$- c((k-1)^2 q_{A|B}(q_{A|A} + q_{B|B}) + (k-1)q_{A|A} + 2(k-1)q_{A|B} + 1)$$

$$- d((k-1)^2 q_{B|B}(q_{A|A} + q_{B|B}) + 3(k-1)q_{B|B})\}$$

$$= \frac{k}{N(k+1)^2}\frac{(k-2)}{(k-1)}p_A(1-p_A)\{$$

$$a((k-1)^2(p_A + \frac{1}{k-1}(1-p_A))(\frac{k}{k-1}) + 3(k-1)(p_A + \frac{1}{k-1}(1-p_A)))$$

$$+ b((k-1)^2\frac{k-2}{k-1}(1-p_A)(\frac{k}{k-1}) + (k-1)(1-\frac{k-2}{k-1}p_A) + 2(k-1)\frac{k-2}{k-1}(1-p_A) + 1)$$

$$- c((k-1)^2\frac{k-2}{k-1}p_A(\frac{k}{k-1}) + (k-1)(p_A + \frac{1}{k-1}(1-p_A)) + 2(k-1)\frac{k-2}{k-1}p_A + 1)$$

$$- d((k-1)^2(1-\frac{k-2}{k-1}p_A)(\frac{k}{k-1}) + 3(k-1)(1-\frac{k-2}{k-1}p_A))\}$$

$$= \frac{k}{N(k+1)^2}\frac{(k-2)}{(k-1)}p_A(1-p_A)\{$$

$$a((k+3)(k-2)p_A + (k+3))$$

$$+ b(-(k+3)(k-2)p_A + (k^2+k-4))$$

$$- c((k+3)(k-2)p_A + 2)$$

$$- d(-(k+3)(k-2)p_A + (k^2+2k-3))\}$$

$$= \underbrace{\frac{k}{N(k+1)^2}\frac{(k-2)}{(k-1)}p_A(1-p_A)\left[\alpha p_A + \beta\right]}_{:=m(p_A)}$$

where $\alpha = (k+3)(k-2)(a-b-c+d)$ and $\beta = (k+3)a + (k^2+k-4)b - 2c - (k^2+2k-3)d$.

In the same way, we get

$$Var(\Delta p_A) = \frac{1}{N^2}P(\Delta p_A = \frac{1}{N}) + \frac{1}{N^2}P(\Delta p_A = -\frac{1}{N})$$

$$= \frac{2}{N^2}p_{AB}\frac{k}{k+1} + O(w)$$

$$\approx \underbrace{\frac{2}{N^2}\frac{k}{k+1}\frac{k-2}{k-1}p_A(1-pa)}_{:=v(p_A)}$$

We are now ready to give an approximate solution for (3.95) in order to obtain the fixation probability function $\phi_A(y)$. Using exactly the same notation as in the previous section (and only changing the constant values of α and β) and the values of M and V. We can proceed in the exact same way, except taking $M = \frac{k(k-2)}{N(k+1)^2(k-1)}$ and $V = \frac{k(k-2)}{N^2(k+1)(k-1)}$ (which gives that $\frac{M}{V} = \frac{N}{k+1}$). We then obtain the following approximation, when taking a small value of w, for (3.95)

$$\phi_A(y) = y + w\frac{N}{6(k+1)}y(1-y)((\alpha+3\beta)+\alpha y).$$

3.2.2 Fixation Probabilities

Now, the probability that in a population a single individual (type A) takes the entire population ($N-1$ players of type b) by is given by $\rho_A = \phi_A(\frac{1}{N})$. By the previous approximation for $\phi_A(y)$ we obtain that for large N, $\rho_A > \frac{1}{N} \iff \alpha + 3\beta > 0$ which is equivalent to

$$(k^2+4k+3)a + (2k^2+2k-3)b > (k^2+k+3)c + (2k^2+5k-3)d.$$

Similarly, $\rho_B > \frac{1}{N}$ is equivalent to

$$(k^2+4k+3)d + (2k^2-2k-3)c > (k^2+k+3)b + (2k^2+5k-3)a.$$

3.2.3 The Rule $\frac{b}{c} > k+2$

Let's consider the same cooperative game as in the previous sections, where A cooperates and B is a defector. So we have the payoff matrix given by

	A	B
A	$b-c$	$-c$
B	b	0

If we substitute these values in the previous results we obtain that $\rho_C > \frac{1}{N} > \rho_D$ if

$$\frac{b}{c} > k+2.$$

Conversely, if $\frac{b}{c} < k + 2$, then $\rho_D > \frac{1}{N} > \rho_C$. Here the rule is saying that the altruistic value b, given by the cooperators over the cost c must exceed the degree of the graph by two in order for the fixation of a strategy to be favoured.

3.3 BD-Updating

We are now set to look at the third kind of update rule. Here, in each time step an individual is selected for reproduction proportional to fitness, then his offspring randomly chooses a neighbour and replaces it. Since, we are interested in deriving the transition probabilities of both p_A and p_{AA}, we need to specify who gets to reproduce and also the local configuration (in the graph) of this individual who was just chosen to reproduce. The probability that an $A-$type player with k_A $A-$type and k_B $B-$type neighbours is selected for reproduction is given by

$$\left[p_A \frac{k!}{k_A!k_B!} q_{A|A}^{k_A} q_{B|A}^{k_B}\right] [1 - w + w(k_A a + k_B b)] .$$

The term in the left corresponds with the frequency of such configuration whereas the term in the right is the fitness of the $A-$type player. Now, if one of the $B-$type players is chosen (which happens with probability $\frac{k_B}{k}$) then p_A increases by 1 (hence, Δp_A increases by $\frac{1}{N}$). In this case, p_{AA} increases by $1 + (k-1)q_{A|B}$(meaning Δp_{AA} in creases by $\frac{2(1+(k-1)q_{A|B})}{kN}$).

In the same way, we obtain that the probability that a $B-$type player is chosen for reproduction is given by

$$\left[p_B \frac{k!}{k_A!k_B!} q_{A|B}^{k_A} q_{B|B}^{k_B}\right] [1 - w + w(k_A c + k_B d)] .$$

In the case that one of the $A-$type neighbours gets to be replaced (which happens with probability $\frac{k_A}{k}$) then p_A decreases by one and p_{AA} decreases by $(k-1)q_{A|A}$. With these probabilities in mind, we will now make some similar calculations as in the previous sections to obtain rules for the cooperator-defector game. First, we will establish the transition probabilities of p_A and p_{AA}.

$$P(\Delta p_A = \frac{1}{N}) = \sum_{k_A+k_B=k} p_A \frac{k!}{k_A!k_B!} q_{A|A}^{k_A} q_{B|A}^{k_B} \frac{k_B}{k} \left[1 + w(k_A a + k_B b - 1)\right]$$

$$P(\Delta p_A = -\frac{1}{N}) = \sum_{k_A+k_B=k} p_B \frac{k!}{k_A!k_B!} q_{A|B}^{k_A} q_{B|B}^{k_B} \frac{k_A}{k} \left[1 + w(k_A c + k_B d - 1)\right]$$

$$P(\Delta p_{AA} = \frac{2(1 + (k-1)q_{A|B})}{kN}) = p_A \frac{k!}{k_A!k_B!} q_{A|A}^{k_A} q_{B|A}^{k_B} \frac{k_B}{k} (1 + w(k_A a + k_B b - 1)$$

$$P(\Delta p_{AA} = \frac{-2(k-1)q_{A|A})}{kN}) = p_B \frac{k!}{k_A!k_B!} q_{A|B}^{k_A} q_{B|B}^{k_B} \frac{k_A}{k} (1 + w(k_A c + k_B d - 1).$$

With these identities in mind, we are now ready to calculate $\dot{p_A}$ and $\dot{p_{AA}}$ in the diffusion approximation approach.

3.3.1 Diffusion Approximation

This will be done in an analogous way to the previous sections. The trick, as before, is to use the binomial identities in (3.27) and the identities (3.1).

$$\dot{p_A} = \frac{1}{N}P(\Delta p_A = \frac{1}{N}) - \frac{1}{N}P(\Delta p_A = -\frac{1}{N})$$

$$= \frac{1}{N}\sum_{k_A+k_B=k} p_A \frac{k!}{k_A!k_B!} q_{A|A}^{k_A} q_{B|A}^{k_B} \frac{k_B}{k}\left[1 + w(k_A a + k_B b - 1)\right]$$

$$- \frac{1}{N})\sum_{k_A+k_B=k} p_B \frac{k!}{k_A!k_B!} q_{A|B}^{k_A} q_{B|B}^{k_B} \frac{k_A}{k}\left[1 + w(k_A c + k_B d - 1)\right]$$

$$= \frac{p_A}{kN}\left[kq_{B|A} + w(k(k-1)q_{A|A}q_{B|A}a + k(k-1)q_{B|A}q_{B|A}b + kq_{B|A}b - kq_{B|A})\right]$$

$$- \frac{p_B}{kN}\left[kq_{A|B} + w(k(k-1)q_{A|B}q_{A|B}c + kq_{A|B}c + k(k-1)q_{A|B}q_{B|B}d - kq_{A|B})\right]$$

$$= \frac{p_{AB}}{N}\left[1 + w((k-1)q_{A|A}a) + ((k-1)q_{B|A} + 1)b - 1)\right]$$

$$- \frac{p_{AB}}{N}\left[1 + w(((k-1)q_{A|B} + 1)c) + (k-1)q_{B|B}d - 1)\right]$$

$$= \frac{wp_{AB}}{N}((k-1)q_{A|A}a + ((k-1)q_{B|A} + 1)b - ((k-1)q_{A|B} + 1)c - (k-1)q_{B|B}d).$$

On the other hand,

$$\dot{p_{AA}} = \sum_{k_A+k_B=k} \frac{2(1 + (k-1)q_{A|B})}{kN} P(\Delta p_{AA} = \frac{2(1 + (k-1)q_{A|B})}{kN})$$

$$- \sum_{k_A+k_B=k} \frac{2(k-1)q_{A|B}}{kN} P(\Delta p_{AA} = -\frac{2(k-1)q_{A|B}}{kN})$$

$$= \sum_{k_A+k_B=k} \frac{2(1 + (k-1)q_{A|B})}{p_A} \frac{k!}{k_A!k_B!} q_{A|A}^{k_A} q_{B|A}^{k_B} \frac{k_B}{k}\left[1 + w(k_A a + k_B b - 1)\right]$$

$$- \sum_{k_A+k_B=k} \frac{2(k-1)q_{A|B}}{kN} p_B \frac{k!}{k_A!k_B!} q_{A|B}^{k_A} q_{B|B}^{k_B} \frac{k_A}{k} \left[1 + w(k_A c + k_B d - 1)\right]$$

$$= \frac{2(1+(k-1)q_{A|B}}{k^2 N} p_A (kq_{B|A} + w(k(k-1)q_{A|A}q_{B|A}a + k(k-1)q_{B|A}q_{B|A}b + kq_{B|A}b - kq_{B|A}))$$

$$- \frac{2((k-1)q_{A|A}}{k^2 N} p_B (kq_{A|B} + w(k(k-1)q_{A|B}q_{A|B}c + kq_{A|B}c + k(k-1)q_{A|B}q_{B|B}d - kq_{A|B}))$$

$$= \frac{2(1+(k-1)q_{A|B}}{kN} p_{AB}(1+O(w))$$

$$- \frac{2((k-1)q_{A|A}}{kN} p_{AB}(1+O(w))$$

$$= \frac{2}{kN} p_{AB}(1+(k-1)q_{A|B} - (k-1)q_{A|A}) + O(w)$$

$$= \frac{2}{kN} p_{AB}(1+(k-1)(q_{A|B} - q_{A|A})) + O(w)$$

This last results makes our calculation quite simple, since it implies that

$$\dot{q_{A|A}} = \frac{\dot{p_{AA}}}{p_A} + O(w)$$

where

$$\frac{\dot{p_{AA}}}{p_A} = \frac{2}{kN}\frac{p_{AB}}{p_A}(1+(k-1)(q_{A|B} - q_{A|A})).$$

Thus, when assuming that $q_{A|A}$ equilibrates much more quickly for weak selection, we obtain the manifold defined by

$$\frac{2}{kN}\frac{p_{AB}}{p_A}(1+(k-1)(q_{A|B} - q_{A|A})) = 0,$$

which is equivalent to

$$q_{A|A} = p_A + \frac{1}{k-1}(1-p_A).$$

This is the same manifold that appears for the previous update rules. Hence all the identities in (3.78) are still valid, and therefore our calculation towards estimating a solution for the Kolmogorov backward equation is ready to begin. As before, this relation let us study the whole system by just describing p_A. We are just left to find $\mathbb{E}(\Delta p_A)$ and $Var(\Delta p_A)$.

$$\mathbb{E}(\Delta p_A) = \frac{1}{N}P(\Delta p_A = \frac{1}{N}) - \frac{1}{N}P(\Delta p_A = -\frac{1}{N})$$

$$= \frac{wp_{AB}}{N}((k-1)q_{A|A}a + ((k-1)q_{B|A}+1)b - ((k-1)q_{A|B}+1)c - (k-1)q_{B|B}d) + O(w^2)$$

$$\approx \frac{w(k-2)}{N(k-1)}p_A(1-p_A)\left[(k-1)(p_A + \frac{1}{k-1}(1-p_A)\right]a$$

$$+ \frac{w(k-2)}{N(k-1)}p_A(1-p_A)\left[(k-1)(\frac{k-2}{k-1}(1-p_A)) + 1\right]b$$

$$- \frac{w(k-2)}{N(k-1)}p_A(1-p_A)\left[(k-1)(\frac{k-2}{k-1}(1-p_A)) + 1\right]c$$

$$- \frac{w(k-2)}{N(k-1)}p_A(1-p_A)(k-1)\left[1 - \frac{k-2}{k-1}p_A\right]d$$

$$= \frac{w}{N}\frac{k-2}{k-1}p_A(1-p_A)(p_A((k-2)a - (k-2)b - (k-2)c + (k-2)d)$$
$$+ a + (k-1)b - c - (k-1)d)$$

$$= \underbrace{\frac{w}{N}\frac{k-2}{k-1}p_A(1-p_A)(\alpha p_A + \beta)}_{:=m(p_A)}$$

where $\alpha = (k-2)(a-b-c+d)$ and $\beta = a + (k-1)b - c - (k-1)d$. For the variance we get

$$\begin{aligned} Var(\Delta p_A) &= \frac{1}{N^2}P(\Delta p_A = \frac{1}{N}) + \frac{1}{N^2}P(\Delta p_A = -\frac{1}{N}) \\ &= \frac{p_A}{kN^2}\left[kq_{B|A} + O(w)\right] + \left[kq_{A|B} + O(w)\right] + O(w^2) \\ &= \frac{2p_{AB}}{N^2} + O(w) \approx \frac{2p_{AB}}{N^2} \\ &= \underbrace{\frac{2}{N^2}\frac{k-2}{k-1}p_A(1-p_A)}_{:=v(p_A)} \end{aligned}$$

Similar calculations to the ones in the DB updating section for the Kolmogorov backward equation hold. Here the α and β change, but remain constants. Here

$M = \frac{k-2}{N(k-1)}$, $V = \frac{k-2}{N^2(k-1)}$ and $\frac{M}{V} = N$. So we obtain that for a small enough w the fixation probability has as an approximate solution

$$\phi_A(y) = y + w\frac{N}{6}y(1-y)((\alpha + 3\beta) + \alpha y).$$

3.3.2 Fixation Probabilities

Now, the probability that in a population a single individual (type A) takes the entire population ($N-1$ players of type b) by is given by $\rho_A = \phi_A(\frac{1}{N})$. By (3.96) we obtain that for large N, $\rho_A > \frac{1}{N} \iff \alpha + 3\beta > 0$ which is equivalent to

$$(k+1)a + (2k-1)b > (k+1)c + (2k-1)d.$$

Similarly, $\rho_B > \frac{1}{N}$ is equivalent to

$$(2k+1)d + (2k-1)c > (k+1)b + (2k-1)a.$$

3.3.3 Never Favoured Cooperators

Now let's consider the same cooperative game as in the previous sections, where A cooperates and B is a defector. So we have the payoff matrix given by

	A	B
A	$b-c$	$-c$
B	b	0

If we substitute these values in the previous results we obtain that $\rho_C > \frac{1}{N} > \rho_D$ can only hold for negative values of either b or c, hence, $\rho_D > \frac{1}{N} > \rho_C$ for any choice of $b > c > 0$. This implies that selection does not favour cooperators in the BD update rule. This should not come as a surprise. Since the individual who gets to reproduce tends to be the one with the best fitness, and since cooperators get beaten when playing in a defector cluster, then there is no reason for the cooperator to invade a graph mostly made up of defectors.

3.3.4 Brief Summary

We are almost done now with this chapter. We will now present a table that shows the chapter's main results. Since, in each of the three updating rules the relation given by $\dot{q_{A|A}} = 0$ was the same, then the fixation probability was always approximated by

$$\phi_A(y) = y + w\frac{N}{6}y(1-y)((\alpha + 3\beta) + \alpha y)$$

where the α and β were the only values that changed.

We would like to think that these results can be generalized to a game with more strategies. This will give, instead of a backward Kolmogorov equation,

Update Scheme	α	β	Rule that favours cooperators
DB-Updating	$(k+1)(k2)(abc+d)$	$(k+1)a+(k^2-k-1)b$ $-c-(k^2-1)d$	$\frac{b}{c}>k$
IM-Updating	$(k+3)(k-2)(a-b-c+d)$	$(k+3)a+(k^2+k-4)b$ $-2c-(k^2+2k-3)d$	$\frac{b}{c}>k+2$
BD-Updating	$(k-2)(a-b-c+d)$	$a+(k-1)b-c-(k-1)d$	Cooperators are not favoured

a problem based upon a partial differential equation. This might be a quite difficult task, but we will like to have an equation to solve in the case where there are three strategies. This will be discussed briefly in the next chapter.

3.4 Computational Approach

It is very important to point out that since our results are approximate, we would be interested in concluding, or at least get a practical insight, for what sizes of N, k and w does the rules derived for cooperation to be favoured are useful. Both for DB and IM updating schemes. Therefore, we must enrol in a Monte Carlo simulation based on the description of figure (3.1) to obtain a value for ρ_C. The procedure her will be to set the parameters $N, k, w, \frac{b}{c}$ create a random regular graph and run the game starting with a single cooperator, stop when the cooperators either vanish or invade the graph. Run this many times and obtain a frequency of how many invasions there were. This will be a good estimate of ρ_C if iterated enough times. As an informative fact, we want to comment that we worked in R statistics. We will first comment some of the problems and results that we encountered in the process of these simulations.

3.4.1 The Regular Graph

Building a random regular graph is not such a simple task. First of all there are two requirements for a random regular graph to exist.

- That Nk is even.
- That $N > k$

At a first glance, one would simply think that in order to generate a random regular graph one would take the space of graphs of size N and degree k and then just randomly pick one. But the size of this sets grows so much so fast on both parameters that this method becomes unbearable for the storage of modern computers. Then, we searched for an algorithm and found Bollobas algorithm, a commonly accepted and used algorithm to create these graphs when the value of k is small. It also requires for $N >> k$. In fact, the probability to obtain a regular graph with this algorithm has probability $e^{-\frac{k^2}{4}}$. This was quite disappointing.

Then we were able to find a 2006 paper named "Generating Random Regular Graphs" which included an algorithm that is $O(Nk^2)$ and works (the probability of obtaining a regular graph tends to 1) for values of $k = O(n^{\frac{1}{18}})$. This is the best possible algorithm which we could find and we don not know if there is a better one. This makes the simulations quite long since the condition $k = O(n^{\frac{1}{18}})$ for the algorithm to run in $O(Nk^2)$ is quite restrictive. Still, We choose to simulate on parameters $k \leq 8$ and $N \leq 200$ for which the random graph generation is a feasible task. Now we can pass onto the next subject.

3.4.2 The parameter w

At some point in the simulations we were getting a program error. The error was due to the fact that depending on the values of w, b and c we can obtain that the quotients

$$\frac{f_C}{f_C + f_D}$$

which determines the probability that a vertex obtains the strategy C can be negative, which cannot be a probability. Hence, in order for the algorithm to run properly we need that when the values b, c and k grow and the rate $\frac{b}{c}$ is smaller, the value of w becomes smaller. For instance when $b = 5$, $c = 1, k = 4$, 8 and $w = 1$, 0.1 the error occurred every time. So, it is not only that the the parameter must be rather small for the cooperation to be favoured, but also for the game to be feasible.

3.4.3 Results

We ran simulations with parameters $k = 2, 4, 8,$ $w = 0.01, 0.001,$ $\frac{b}{c} = 5$ and $N = 10, 50, 100$. The simulations were quite long (at least a week for around 40 different tasks in a highly vectorized encoding). We wanted to see how the clusters evolved, but mainly to get an insight on the level of accuracy of the rules given in the previous sections.

	$k=2$	$k=4$	$k=8$
$N=10$	0.899	0.764	0.506
$N=50$	0.092	0.017	0.004
$N=100$	0.027	0.008	0

Table 3.1: Table for the **DB-updating** simulation when starting with a random regular graph of size N and degree k, and with a single cooperator. The simulation was done using $\mathbf{w = 0.01}$ and $b = 5$, $c = 1$ and $\frac{b}{c} = 5$. Invasion was considered when more than 70% of the population was taken by cooperators. The values in this table are simulations for ρ_C.

Based on the previous tables we can conclude the following

- The rules that were previously stated are reasonable for $N > 50$.

	$k = 2$	$k = 4$	$k = 8$
$N = 10$	0.845	0.779	0.674
$N = 50$	0.122	0.028	0.013
$N = 100$	0.029	0.012	0

Table 3.2: Table for the **DB-updating** simulation when starting with a random regular graph of size N and degree k, and with a single cooperator. The simulation was done using $\mathbf{w = 0.001}$ and $b = 5$, $c = 1$ and $\frac{b}{c} = 5$. Invasion was considered when more than 70% of the population was taken by cooperators. The values in this table are simulations for ρ_C.

	$k = 2$	$k = 4$	$k = 8$
$N = 10$	0.544	0.347	0.782
$N = 50$	0.023	0.031	0.040
$N = 100$	0.012	0.017	0.009

Table 3.3: Table for the **DB-updating** simulation when starting with a random regular graph of size N and degree k, and with a single defector. The simulation was done using $\mathbf{w = 0.01}$ and $b = 5$, $c = 1$ and $\frac{b}{c} = 5$. Invasion was considered when more than 70% of the population was taken by cooperators. The values in this table are simulations for ρ_D.

- Simulations for small populations can give odd results, so the size is an important factor.
- The change from $w = 0.01$ to $w = 0.001$ does not give drastically different results, hence, taking w of the order of $\frac{1}{N}$ or smaller should suffice for the rules to hold (this is a careful guess based on the simulations).
- The rules do not work as strictly as one would desire, but still give a good approach for estimating if cooperation is favoured.

It is important to share that when running some simulations for the third updating rule (BD-updating) the results were not indicators of any relation between the rate $\frac{b}{c}$ and k. In the simulations, fixation of cooperators was sometimes heavily favoured and sometimes, not favoured at all. This would indicate, as was seen before, that the previous approach is useless for this updating rule.

	$k=2$	$k=4$	$k=8$
$N=10$	0.477	0.481	0.655
$N=50$	0.019	0.024	0.201
$N=100$	0	0	0.023

Table 3.4: Table for the **DB-updating** simulation when starting with a random regular graph of size N and degree k, and with a single defector. The simulation was done using $\mathbf{w} = \mathbf{0.001}$ and $b = 5$, $c = 1$ and $\frac{b}{c} = 5$. Invasion was considered when more than 70% of the population was taken by cooperators. The values in this table are simulations for ρ_D.

	$k=2$	$k=4$	$k=8$
$N=10$	0.844	0.344	0.231
$N=50$	0.097	0.010	0.007
$N=100$	0.012	0.008	0

Table 3.5: Table for the **IM-updating** simulation when starting with a random regular graph of size N and degree k, and with a single cooperator. The simulation was done using $\mathbf{w} = \mathbf{0.01}$ and $b = 5$, $c = 1$ and $\frac{b}{c} = 5$. Invasion was considered when more than 70% of the population was taken by cooperators. The values in this table are simulations for ρ_C.

	$k=2$	$k=4$	$k=8$
$N=10$	0.774	0.285	0.211
$N=50$	0.027	0.018	0.005
$N=100$	0.012	0	0

Table 3.6: Table for the **IM-updating** simulation when starting with a random regular graph of size N and degree k, and with a single cooperator. The simulation was done using $\mathbf{w} = \mathbf{0.001}$ and $b = 5$, $c = 1$ and $\frac{b}{c} = 5$. Invasion was considered when more than 70% of the population was taken by cooperators. The values in this table are simulations for ρ_C.

	$k=2$	$k=4$	$k=8$
$N=10$	0.332	0.227	0.302
$N=50$	0.017	0.031	0.044
$N=100$	0.009	0.017	0.020

Table 3.7: Table for the **IM-updating** simulation when starting with a random regular graph of size N and degree k, and with a single defector. The simulation was done using $\mathbf{w} = \mathbf{0.01}$ and $b = 5$, $c = 1$ and $\frac{b}{c} = 5$. Invasion was considered when more than 70% of the population was taken by cooperators. The values in this table are simulations for ρ_D.

	$k=2$	$k=4$	$k=8$
$N=10$	0.470	0.419	0.297
$N=50$	0.019	0.029	0.298
$N=100$	0	0.019	0.024

Table 3.8: Table for the **IM-updating** simulation when starting with a random regular graph of size N and degree k, and with a single defector. The simulation was done using $\mathbf{w} = \mathbf{0.001}$ and $b = 5$, $c = 1$ and $\frac{b}{c} = 5$. Invasion was considered when more than 70% of the population was taken by cooperators. The values in this table are simulations for ρ_D.

3.4.4 On the Behaviour of Clusters

On the other hand we wanted to track how the clusters of an invading strategy evolved in time. We did this for ρ_C for both the DB and IM updating rules. The results appear in the following tables and are based upon a single simulation of the game were invasion (over 70%) was obtained. The game with $N = 50$ and $k = 2$ was considered since it was the one giving the most consistent results. Again, $b = 5$ and $c = 1$. Also, $w = 0.001$.

Time	Number of cooperators	Size of biggest cluster
$t = 0$	1	1
$t = 1000$	4	4
$t = 2000$	6	6
$t = 3000$	9	9
$t = 4000$	11	11
$t = 5000$	13	13
$t = 6000$	13	13
$t = 7000$	12	12
$t = 8000$	16	16
$t = 9000$	17	17
$t = 10000$	20	20
$t = 11000$	17	17
$t = 12000$	19	19
$t = 13000$	21	21
$t = 14000$	24	24
$t = 15000$	22	22
$t = 16000$	22	22
$t = 17000$	25	25
$t = 18000$	26	26
$t = 19000$	25	25
$t = 20000$	27	27
$t = 21000$	32	32
$t = 22000$	31	31
$t = 23000$	28	28
$t = 24000$	29	29
$t = 25000$	33	33
$t = 25906$	35	35

Table 3.9: Table for the **IM-updating** simulation when starting with a random regular graph of size 50 and degree 2, and with a single cooperator. The simulation was done using $\mathbf{w = 0.001}$ and $b = 5$, $c = 1$ and $\frac{b}{c} = 5$. Invasion was considered when more than 70% of the population was taken by cooperators.

In the previous tables one can observe that what commonly happens is that a cluster of cooperators starts to grow until it takes control over the entire population. This happens probably because cooperation is favoured in a cooperating population, hence cooperators need clusters in order to succeed in invading. The

Time	Number of cooperators	Size of biggest cluster
$t = 0$	1	1
$t = 5000$	14	10
$t = 10000$	17	17
$t = 11000$	16	16
$t = 12000$	17	17
$t = 13000$	19	19
$t = 14000$	24	24
$t = 15000$	22	20
$t = 16000$	25	25
$t = 17000$	25	25
$t = 18000$	24	24
$t = 19000$	25	25
$t = 20000$	26	26
$t = 21000$	29	29
$t = 22000$	30	30
$t = 23000$	28	28
$t = 24000$	30	30
$t = 25000$	31	31
$t = 26000$	29	29
$t = 27000$	30	30
$t = 28000$	32	32
$t = 29000$	31	31
$t = 30000$	33	33
$t = 31000$	34	34
$t = 31112$	35	35

Table 3.10: Table for the **DB-updating** simulation when starting with a random regular graph of size 50 and degree 2, and with a single cooperator. The simulation was done using $\mathbf{w = 0.001}$ and $b = 5$, $c = 1$ and $\frac{b}{c} = 5$. Invasion was considered when more than 70% of the population was taken by cooperators.

number of steps that was taken in each updating rule gives us no insight since we were just taking a game simulation were the cooperators actually invaded, but in order to determine which of the two rules makes invasion fastest, one would need to run the same simulation for at least a thousand times (a quite long computation).

Chapter 4

Three Strategies Equation

Based on the discussion that appeared in the previous chapter, we are now interested in a game with three strategies. This leads us to a deduction of an equation similar to the Kolmogorov backward equation which was deduced in the preliminaries, but that works for additional variables. It will also be a PDE instead of an ODE.

4.1 The Equation

We will use a notation similar to the one used in the preliminaries for the deduction of the Kolmogorov backward equation. In a similar context, we will consider having three alleles A_1, A_2 and A_3. Let $\phi(p_1, p_2, x_1, x_2, t + \delta t)$ denote the probability density that the frequency of A_1 lies between x_1 and $x_1 + dx_1$ and the frequency of A_2 lies between x_2 and $x_2 + dx_2$ at time t; given that the initial frequency ($t = 0$) of the alleles A_1 and A_2 is p_1 and p_2 respectively. Let $g(x_1, x_2, \xi_1, \xi_2, t, \delta t)$ be the probability density that the frequency changes from x_1 and x_2 to $x_1+\xi_1$ and $x_2+\xi_2$ respectively, during the time interval $(t, t+\delta t)$.It is clear that the frequency of the first two alleles completely determines the third one, so we will only be concerned with A_1 and A_2. Again, we will consider a time homogeneous process, hence the value of t will not be of relevance, so the t in g will be omitted. It is also important to consider the case when x_1 and x_2 are fixed and p_1 and p_2, the initial frequencies, are random variables.

Analogously as in Chapter 2, we obtain that:

$$\phi(p_1, p_2, x_1, x_2, t+\delta t) = \int\int g(p_1, p_2, \xi_1, \xi_2, \delta t)\phi(p_1+\xi_1, p_2+\xi_2, x_1, x_2, t)d\xi_1 d\xi_2.$$

This means that the probability that A_1 lies between x_1 and $x_1 + dx_1$ and the frequency of A_2 lies between x_2 and $x_2 + dx_2$ at time t is the same as the all possible ways of starting at (p_1, p_2) and changing to $(p_1 + \xi_1, p_2, \xi_2)$ in time δt and from here on, going to (x_1, x_2) in time t. We will now use Taylor's polynomial approximation to deduce the equation. In the following the notation g and ϕ will stand for $g(x_1, x_2, \xi_1, \xi_2, t, \delta t)$ and $\phi(p_1, p_2, x_1, x_2, t)$ respectively. Also, $O(\xi^3)$ will mean polynomial terms of ξ_1 and ξ_2 of degree greater than 3.

$$\phi(p_1, p_2, x_1, x_2, t + \delta t)$$

$$= \int\int g(p_1, p_2, \xi_1, \xi_2, \delta t)\phi(p_1 + \xi_1, p_2 + \xi_2, x_1, x_2, t)d\xi_1 d\xi_2$$

$$= \int\int \{(g\phi) + g\xi_1 \frac{\partial}{\partial \xi_2}\phi + g\xi_2 \frac{\partial}{\partial \xi_2}\phi$$

$$+ g\frac{\xi_1^2}{2}\frac{\partial^2}{\partial \xi_1^2}(\phi) + g\xi_1\xi_2 \frac{\partial^2}{\partial \xi_1 \partial \xi_2}(\phi) + g\frac{\xi_2^2}{2}\frac{\partial^2}{\partial \xi_2^2}(\phi) + O(\xi^3)\}d\xi_1 d\xi_2$$

Now, using that g is a density ($\int\int g = 1$) and assuming that addition, differentiation and integration can be freely interchanged, we obtain that:

$$\phi(p_1, p_2, x_1, x_2, t + \delta t)$$

$$\approx \phi(p_1 + \xi_1, p_2 + \xi_2, x_1, x_2, t) + \frac{\partial}{\partial \xi_1}(\phi)\int\int \xi_1 g d\xi_1 d\xi_2 + \frac{\partial}{\partial \xi_2}(\phi)\int\int \xi_2 g d\xi_1 d\xi_2$$

$$+ \frac{1}{2}\{\frac{\partial^2}{\partial \xi_1^2}(\phi)\int\int \xi_1^2 g d\xi_1 d\xi_2 + \frac{\partial^2}{\partial \xi_2^2}(\phi)\int\int \xi_2^2 g d\xi_1 d\xi_2\} + \frac{\partial^2}{\partial \xi_1 \partial \xi_2}(\phi)\int\int \xi_1\xi_2 g d\xi_1 d\xi_2.$$

Then, subtracting the first term on the right from both sides and dividing by δt we obtain:

$$\frac{\phi(p_1, p_2, x_1, x_2, t + \delta t) - \phi(p_1, p_2, x_1, x_2, t)}{\delta t}$$

$$= \frac{\partial}{\partial \xi_1}(\frac{\phi}{\delta t})\int\int \xi_1 g d\xi_1 d\xi_2 + \frac{\partial}{\partial \xi_2}(\frac{\phi}{\delta t})\int\int \xi_2 g d\xi_1 d\xi_2$$

$$+ \frac{1}{2}\{\frac{\partial^2}{\partial \xi_1^2}(\frac{\phi}{\delta t})\int\int \xi_1^2 g d\xi_1 d\xi_2 + \frac{\partial^2}{\partial \xi_2^2}(\frac{\phi}{\delta t})\int\int \xi_2^2 g d\xi_1 d\xi_2\} + \frac{\partial^2}{\partial \xi_1 \partial \xi_2}(\frac{\phi}{\delta t})\int\int \xi_1\xi_2 g d\xi_1 d\xi_2$$

Now, letting $\delta t \longrightarrow 0$ we obtain:

$$\frac{\partial}{\partial t}\phi(p_1 + \xi_1, p_2 + \xi_2, x_1, x_2, t) = \frac{\partial}{\partial \xi_1}(\phi)M_1 + \frac{\partial}{\partial \xi_2}(\phi)M_2$$

$$+ \frac{1}{2}\{\frac{\partial^2}{\partial \xi_1^2}(\phi)V_{11} + \frac{\partial^2}{\partial \xi_2^2}(\phi)V_{22}\} + \frac{\partial^2}{\partial \xi_1 \partial \xi_2}(\phi)V_{12}$$

where

$$M_1(p_1,p_2,x_1,x_2) = \lim_{\delta t \longrightarrow 0} \frac{1}{\delta t} \int \int \xi_1 g d\xi_1 d\xi_2$$

$$M_2(p_1,p_2,x_1,x_2) = \lim_{\delta t \longrightarrow 0} \frac{1}{\delta t} \int \int \xi_2 g d\xi_1 d\xi_2$$

$$V_{11}(p_1,p_2,x_1,x_2) = \lim_{\delta t \longrightarrow 0} \frac{1}{\delta t} \int \int \xi_1^2 g d\xi_1 d\xi_2$$

$$V_{22}(p_1,p_2,x_1,x_2) = \lim_{\delta t \longrightarrow 0} \frac{1}{\delta t} \int \int \xi_2^2 g d\xi_1 d\xi_2$$

$$V_{12}(p_1,p_2,x_1,x_2) = \lim_{\delta t \longrightarrow 0} \frac{1}{\delta t} \int \int \xi_1 \xi_2 g d\xi_1 d\xi_2$$

Now, we can replace M_1, M_2, V_{11}, V_{22} and V_{12} by $\mathbb{E}(X), \mathbb{E}(Y), Var(X), Var(Y)$ and $\mathbb{E}(XY)$ respectively, where X and Y are the respective amounts of change of each allele per generation or time step. This would yields the following second order PDE:

$$\frac{\partial}{\partial t}\phi(p_1+\xi_1, p_2+\xi_2, x_1, x_2, t) = \mathbb{E}(X)\frac{\partial}{\partial \xi_1}(\phi) + \mathbb{E}(Y)\frac{\partial}{\partial \xi_2}(\phi)$$

$$+\frac{1}{2}\nabla^t Cov(X,Y)\nabla\phi$$

where $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})^t$ and $Cov(X,Y)$ the covariance matrix of the random vector (X,Y). Now, proceeding as in the preliminaries, we would want to relate this model to the fixation probability problem. This can happen when $x_1 = 1$ (this will make $x_2 = 0$ automatically) or when $x_2 = 1$ (hence, $x_1 = 0$). We will deal only with the first case since the other one in analogous. In this case $u(p_1,p_2,t) := \phi(p_1,p_2,1,0,t)$ can be interpreted as the probability of fixation at time t when the initial frequency was (p_1,p_2). Thus, $u(p_1,p_2,t)$ satisfies the differential equation:

$$\frac{\partial u(p_1,p_2,t)}{\partial t} = \mathbb{E}(X)\frac{\partial}{\partial \xi_1}(u(p_1,p_2,t)) + \mathbb{E}(Y)\frac{\partial}{\partial \xi_2}(u(p_1,p_2,t))$$

$$+\frac{1}{2}\nabla^t Cov(X,Y)\nabla u(p_1,p_2,t)$$

In order to make the fixation problem we put the boundary conditions

$$u(1,0,t) = 1,\ u(0,p_2,t) = 0.$$

We proceed now to obtain the ultimate fixation probability which is defined by $u(p_1, p_2) = \lim_{t\longrightarrow\infty} u(p_1, p_2, t)$. By an analogous argument as in the preliminaries we obtain that $\frac{\partial u}{\partial t} = 0$ so we obtain the following second order PDE:

$$0 = \mathbb{E}(X)\frac{\partial}{\partial \xi_1}(u(p_1, p_2)) + \mathbb{E}(Y)\frac{\partial}{\partial \xi_2}(u(p_1, p_2))$$

$$+ \frac{1}{2}\triangledown^t Cov(X, Y)\triangledown u(p_1, p_2)$$

Notice that this is a conic semi-linear PDE (as appears in [5]). This type of PDE is quite common in the field of PDEs and is tractable to some level. For instance numerical approximations using finite elements can be quite precise when trying to find an approximate solution. We will now like to discuss how this can be related to the models of evolutionary dynamics on graphs discussed in the previous chapter.

4.2 Further Discussion

In the previous section we obtained a second order PDE which models the fixation probability function. We will now like to discuss why is it that this PDE can help us generalize to three strategies some of the ideas proposed in [3] that were treated in the previous chapter. The main idea in that chapter was to use the variable of change in the frequency of each strategy and of the pairs to obtain a dynamical system that let the whole system be expressed in terms of a single frequency, using that the fast dynamic of $q_{A|A}$ will make it stabilize faster than p_A making $\dot{q_{A|A}} = 0$, which let the system be expressed in terms of a single variable. Then, the mean and variance of the change in frequency per time step could be approximated when working with a small w. These results were then introduced in the Kolmogorov backward equation. This lead to an approximation of the fixation probability function which finally led to the decision rule for the cooperator/defector game. Now, having that in mind, we could use the same updating rules to model the dynamical game, but this time, with three strategies (A, B and C). Hence, using the same notation as in the previous chapter, the identities appearing in (3.1) can be replaced by

$$\begin{aligned}
&p_A + p_B + p_C = 1 \\
&q_{A|X} + q_{B|X} + q_{C|X} = 1 \\
&p_{XY} = q_{X|Y}p_Y \\
&p_{XY} = p_{YX}
\end{aligned}$$

and the payoff matrix is now given by

$$\begin{array}{cccc}
 & A & B & C \\
A & a & b & c \\
B & d & e & f \\
C & g & h & i
\end{array}.$$

Before, all the system could be expressed in terms of p_A and $q_{A|A}$. Now, we have introduced more variables and thus, have make the system way more complex. Still, two of the frequencies determine the third one ($p_C = 1 - p_A - p_B$). The idea will be to proceed to obtain the system in terms of the least possible amount of variables including p_A and p_B and the rest in terms of $q_{X|Y}$s. Here the fast dynamic approach will be again held effective, letting the three strategy problem be expressed in terms of p_A and p_B, and letting us use the just deduced PDE. Then we must try to calculate $\mathbb{E}(\Delta p_A), \mathbb{E}(\Delta p_B), Var(\Delta p_A), Var(\Delta p_B)$ and $\mathbb{E}(\Delta p_A \Delta p_B)$ and plug them into the PDE. This is the subject of a work in progress

Chapter 5

Appendix

5.1 Code

The following is a version of the code that was used for the simulation of the two strategy game for obtaining a Monte Carlo approximation of the $\phi_A(y)$ probability function. The code was slightly modified in order to obtain different results but it is the backbone of the general procedure. The language of the code is R-statistics. Here the lines that start with # are documentation.

```
rest=function(full_vector,searched_vector){

      found=c()

      for(i in full_vector){

        if(any(is.element(searched_vector,i))){
          searched_vector[(which(searched_vector==i))[1]]=NA
        }
        else{
          found=c(found,i)
        }
    }

    return(found)
}

#----------------------------------------------------------------
#N es el numero de nodos y k el grado del grafo
graReg=function(N,k)
{
ns=seq(1:N)
buc=c()
```

```
for (i in 1:k){
    buc=c(buc,ns)
}

    ari=matrix(rep(0,N*k),N*k/2,2)
    for(i in 1:(N*k/2)){
    a=sample(buc,1)
    if(length(setdiff(buc,a))!=0){
    d=sample(setdiff(buc,a),1)
    ari[i,]=c(a,d)
    buc=rest(buc,c(a,d))
    }
    else {d=a
    ari[i,]=c(a,d)
    buc=rest(buc,c(a,d))
    }
}
return(ari)
}
#------------------------------------------------------------

cont=function(M){
cont=0
for(i in 1:(dim(M)[1])){
if(M[i,1]==M[i,2]){cont=cont+1}
}
return(cont)
}
#-------------------------------------------------------
repetidos=function(M){
n=dim(M)[1]
cont=0
for(i in 1:n){
pareja=M[i,]
for (j in setdiff(seq(1:n),i)){
if(isTRUE(all.equal(pareja,M[j,]))){cont=cont+1}
else if ((pareja[1]==M[j,2] & pareja[2]==M[j,1])){cont=cont+1}
    }
}
return(cont)
}


graReg=function(N,k)
{
  ns=seq(1:N)
```

```
  buc=c()
  for (i in 1:k){
    buc=c(buc,ns)
  }
  #matriz de aristas
  ari=c()
  for(i in 1:(N*k/2)){
    c=TRUE
    while(c){
    arista=c()
    a=sample(buc,1)
    #si quedan nodos tipo a en el bucket
    if(length(setdiff(buc,a))!=0){

      t=seq(1:length(buc))
      #se escoge un elemento del bucket distinto de a
      d=sample(buc[-t[buc[t]==a]],1)
      arista=c(a,d)
    }else {d=sample(setdiff(buc,a),1)
    arista=c(a,d)
    }
    #si arista no habia sido escogida
    if(isTRUE(repetidos(rbind(ari,arista))==0)){
      c=FALSE
      ari=rbind(ari,arista)
      buc=rest(buc,arista)
      }

    }
  }
  return(ari)
}

#--------------------------------------------------------------------
acDBIM=function(N,k,w,bene=5,cost=1,tiempo=100000,ite=1000){
  nodos=seq(1:N)
  grafo=graReg(N,k)
  #contador de invasiones
  coinvaDB=0
  coinvaIM=0
  for(i in 1:ite){
    #vector de estrategias
    est=rep(0,N)
    #los 0 son defractores y los 1 cooperan
    est[sample(seq(1:N),1)]=1
    #frecuencia de cooperacion
```

```
freco=1/N
contador=0
while(isTRUE(freco != 0) && isTRUE(freco != 1) && isTRUE(contador<tiempo) ){
  contador=contador+1
  nodoazar=sample(nodos,1)
  t=seq(1:(N*k/2))
  #se obtienen los nodos vecinos del nodo escogido
  vecinos=c(grafo[t[grafo[t,1]==nodoazar],2],grafo[t[grafo[t,2]==nodoazar],1])
  #vectores nulos para almacenar el pago de los vecinos cooperadores y defractores
  pC=c()
  pD=c()
  #si el nodo eliminado era defractor
  if(isTRUE(est[nodoazar]==0)){
    for(j in vecinos){
      #vecinos del j esimo vecino
      vecij=c(grafo[t[grafo[t,1]==j],2],grafo[t[grafo[t,2]==j],1])
      #si el vecino j esimo es cooperador
      if(isTRUE(est[j]==1)){
        #pago del j-esimo vecino
        pC=c(pC,((bene-cost)*(sum(est[vecij]==1)))-(cost*(sum(est[vecij]==0)-1)))
        #si el jesimo vecino es defractor
      } else {pD=c(pD,(bene*(sum(est[vecij]==1))))}
    }
    #se cierra el if
  }
  #si nodo azar era cooperador
  if (isTRUE(est[nodoazar]==1)) {
    for(j in vecinos)
    {
      #vecinos del j esimo vecino
      vecij=c(grafo[t[grafo[t,1]==j],2],grafo[t[grafo[t,2]==j],1])
      #si el vecino j esimo es cooperador
      if(isTRUE(est[j]==1)){
        #pago del j-esimo vecino
        pC=c(pC,((bene-cost)*(sum(est[vecij]==1)-1))-(cost*(sum(est[vecij]==0))))
        #si el jesimo vecino es defractor
      } else {pD=c(pD,(bene*(sum(est[vecij]==1)-1)))}
    }
  }

  fC=1-w+w*(sum(pC))
  fD=1-w+w*(sum(pD))
  est[nodoazar]=sample(c(0,1),1,prob=c(fD/(fC+fD),fC/(fC+fD)))
  #se actualiza la frecuencia
  freco=mean(est)
  if(isTRUE(freco==1)){coinvaDB=coinvaDB+1}
```

```
      #se cierra el while
    }
    #se cierra el for de ITE
  }
  for(i in 1:ite){
    #vector de estrategias
    est=rep(0,N)
    #los 0 son defractores y los 1 cooperan
    est[sample(seq(1:N),1)]=1
    #frecuencia de cooperacion
    freco=1/N
    contador=0
    while(isTRUE(freco != 0) && isTRUE(freco != 1) && isTRUE(contador<tiempo) ){
      contador=contador+1
      nodoazar=sample(nodos,1)
      t=seq(1:(N*k/2))
      #se obtienen los nodos vecinos del nodo escogido
      vecinos=c(grafo[t[grafo[t,1]==nodoazar],2],grafo[t[grafo[t,2]==nodoazar],1])
      #vectores nulos para almacenar el pago de los vecinos cooperadores y defractores resp
      p0=c()
      pC=c()
      pD=c()
      #si el nodo eliminado era defractor
      if(isTRUE(est[nodoazar]==0)){
        #pago del nodo azar
        p0=bene*(sum(est[vecinos]==1))
        for(j in vecinos){
          #vecinos del j esimo vecino
          vecij=c(grafo[t[grafo[t,1]==j],2],grafo[t[grafo[t,2]==j],1])
          #si el vecino j esimo es cooperador
          if(isTRUE(est[j]==1)){
            #pago del j-esimo vecino
            pC=c(pC,((bene-cost)*(sum(est[vecij]==1)))-(cost*(sum(est[vecij]==0))))
            #si el jesimo vecino es defractor
          } else {pD=c(pD,(bene*(sum(est[vecij]==1))))}
        }
        #se cierra el if
      }
      #si nodo azar era cooperador
      if (isTRUE(est[nodoazar]==1)) {
        #pago del nodo azar
        p0=((bene-cost)*(sum(est[vecinos]==1)))-(cost*(sum(est[vecinos]==0)))
        for(j in vecinos)
        {
          #vecinos del j esimo vecino
          vecij=c(grafo[t[grafo[t,1]==j],2],grafo[t[grafo[t,2]==j],1])
```

```
        #si el vecino j esimo es cooperador
        if(isTRUE(est[j]==1)){
          #pago del j-esimo vecino
          pC=c(pC,((bene-cost)*(sum(est[vecij]==1)))-(cost*(sum(est[vecij]==0))))
          #si el jesimo era defractor
        } else {pD=c(pD,(bene*(sum(est[vecij]==1))))}
      }
    }
    f0=1-w+w*p0
    fC=1-w+w*(sum(pC))
    fD=1-w+w*(sum(pD))
    if (isTRUE(est[nodoazar]==1)){
      est[nodoazar]=sample(c(0,1),1,prob=c(fD/(fC+fD+f0),(fC+f0)/(fC+fD+f0)))
    } else {est[nodoazar]=sample(c(0,1),1,prob=c(fC/(fC+fD+f0),(fD+f0)/(fC+fD+f0)))}
    #se actualiza la frecuencia
    freco=mean(est)
    if(isTRUE(freco==1)){coinvaIM=coinvaIM+1}
    #se cierra el while
  }
  #se cierra el for de I
 #se retorna la frecuencia de fijacion
 return(c((coinvaDB/ite),(coinvaIM/ite)))
}
```

Chapter 6

Bibliography

[1] Hopfbauer, J., and Sigmund,K. "Evolutionary Games and Population Dynamics",Cambridge University Press. pp. 1-140. June 13, 1998.

[2] Crow, J.F., Kimura, M., "An Introduction to Population Genetics Theory", Burgess Publishing Company, pp 371-382, 1970.

[3] Ohtsuki,H., Hauert,C., Lieberman, E. and Nowak,M.A."A Simple Rule for the Evolution of Cooperation on graphs and Social Networks",nature Magazine. Vol 441—25, pp 502-505, May 2006.

[4] Perko, L., "Differential Equations and Dynamical Systems", Texts in Applied Mathematics 7, Third Edition, Springer,pp 1-192, 1991.

[5] Roubicek, T., "Nonlinear Partial Differential Equations with Applications", Second Edition, Birkhauser, 2005.