The Replicator Dynamic (Draft)

June 10, 2021

We now turn to the main object of our study, the learning dynamics themselves. We begin with the *replicator dynamic*, which is a foundational concept in evolutionary game theory (EGT) and population dynamics [7].

In the following we look at the case where only a single matrix $\mathbf{A}$ defines the game. This corresponds to a symmetric game in which the agent is, in essence, playing against themselves. One could equivalently think of this as a two player game in which $\mathbf{B} = \mathbf{A}^T$, so we only need $\mathbf{A}$ to define the game. In this context an NE $\bar{\mathbf{x}}$ is a strategy for which

$$\mathbf{x} \cdot \bar{\mathbf{x}} \leq \bar{\mathbf{x}} \cdot \bar{\mathbf{x}} \tag{1}$$

The equivalent statement to Lemma 1.1 in this context is that $\bar{\mathbf{x}}$ is an NE if and only if $(\mathbf{Ax})_i = \mathbf{x} \cdot \mathbf{Ax}$ for all i such that $x_i > 0$.

1 Introducing the Replicator Dynamic

Consider a large population, in which individuals are identified according to some phenotype[1]. As an example, let us say that a population is divided into individuals with white fur, brown fur and black fur. In game theoretic terms, we can consider these as belonging to a set of pure strategies $\{W, Br, Bl\}$. Let x_i denote the proportion of the population who play with a given strategy i. In our example, x_1 denotes the proportion who have white fur. Then $\mathbf{x} = (x_1, \dots x_n)^T$ is a probability vector which takes values on the unit-simplex.

The question of evolutionary game theory is to ask, which of these phenotypes will thrive, and which have the potential to become extinct? EGT takes the established principle of evolution that of 'survival of the fittest', and places it into a game theoretic framework, which we can analyse mathematically. Specifically, let us consider that the fitness of an individual is encoded in a function $f : \Delta \to \mathbb{R}$. If this individual plays action i, then their payoff is $f_i(\mathbf{x})$. Further, the average payoff received by the entire population is $\sum_i x_i f_i(\mathbf{x})$. The thesis of evolutionary dynamics is that strategies which have higher fitness than the average should be more likely to survive and, therefore, their proportion should increase, whilst those who have a lower fitness should decrease over time. This is recorded in the continuous time replicator dynamics:

[1] a physical attribute of an individual which results as a combination of its genotype and the environment

$$\dot{x}_i(t) = x_i(t)(f_i(\mathbf{x}) - \sum_i x_i f_i(\mathbf{x}))$$

In particular, if f can be written in terms of a payoff matrix $\mathbf{A} \in M_n(\mathbb{R})$, then we have

$$\dot{x}_i(t) = x_i(t)((\mathbf{A}\mathbf{x})_i - \mathbf{x} \cdot \mathbf{A}\mathbf{x}) \tag{2}$$

As such, when the payoff received by an agent playing action i, given by $(\mathbf{A}\mathbf{x})_i$, is greater than the average payoff across the entire population, $\dot{x}_i$ is positive, and so the proportion playing i increases with time.

Remark. An important point to note regarding the replicator dynamic (RD) is that the unit-simplex is *invariant*[2] with respect to this equation. This means that $\mathbf{x}$ will always take values in the unit-simplex as they evolve using RD. This fact is easy to check.

Remark. Readers who are familiar with population dynamics may notice that the replicator dynamic shares a similar idea, and a similar form to the famous Lotka-Volterra dynamics [6], which similarly describe the evolution of competing species. We show in the Appendix that this similarity can indeed be formalised, as there is a diffeomorphic transformation which maps orbits of the replicator dynamic to those of the Lotka-Volterra dynamics.

Remark. Whilst the replicator dynamic speaks of proportions of phenotypes in a population, we can analyse it from the view of mixed strategies in a game. Instead of a population whose proportions are given by the vector $\mathbf{x} = (x_1, ..., x_n)^T$, we instead consider a single agent whose mixed strategy is given by the same vector.

Remark. Whilst its primary mode of application has historically been in understanding population dynamics, the replicator dynamic also turns out to be a continuous approximation to *Follow the Regularised Leader* [1], a particular type of online learning algorithm. We do not report on this relationship here, but it is well worth reading. As such, we point to [4] for a incredibly informative exposition into this relationship.

Another interesting point to note about the replicator dynamic is that its behaviour does not change under a certain transformation to the matrix $\mathbf{A}$.

Example 1. *Consider the symmetric matrix game* $\mathbf{A}$ *defined by*

$$\mathbf{A} = \begin{pmatrix} 2 & 5 \\ 4 & 3 \end{pmatrix} \tag{3}$$

The flow defined by the replicator dynamic (2) can be seen in Figure 1. It can be seen that there is an NE at the points $\mathbf{e}_1$, $\mathbf{e}_2$ *and* $\bar{\mathbf{x}} = (0.5, 0.5)^T$. *It can be seen that the point* $\bar{\mathbf{x}}$ *is a Nash Equilibrium of the game, whereas* $\mathbf{e}_1$ *and* $\mathbf{e}_2$ *are not.*

Lemma 1. *The transformation* $\mathbf{A} \to \tilde{\mathbf{A}}$ *given by* $\tilde{\mathbf{A}} = \mathbf{A} + (\mathbf{c}_1, \ldots, \mathbf{c}_n)$ *where* $\mathbf{c}_1, \ldots, \mathbf{c}_n \in \mathbb{R}^n$ *leaves the replicator dynamics unchanged.*

In words, if we were to add constants to each column of $\mathbf{A}$ does not alter the flow generated by RD. This fact is rather easy to check and so we leave it as an exercise.

With the replicator equation defined, we can begin to ask some central questions

[2]Formally, a set E is invariant with respect to a flow ϕ if, for any $\mathbf{x} \in E$, $\phi_t(x) \in E$ for all $t \geq 0$.

Figure 1: The flow defined by (2) on the game in Example 1

I How does the central idea of Game Theory, namely the Nash Equilibrium, fit into the replicator dynamic?

II Under what conditions does a given phenotype (or strategy) in the population go extinct in the long run?

III What sorts of dynamical behaviour does RD exhibit?

These questions are not unique to the Replicator Dynamic and, as we will see, form the basis for our analysis of the other learning algorithms. The first question is sometimes referred to as the 'folk theorem' [2], which frames the NE as a fixed point of the replicator dynamics, thereby giving it a dynamical interpretation. The second is referred to as the question of 'permanence', which we shall go on to describe. The final question, that of complex dynamics, is an area which is rich with open questions and active research. One of the most interesting results is that, even in the simplest settings, the replicator dynamic can lead to Hamiltonian chaos [5].

1.1 The Folk Theorem of Evolutionary Game Theory

Let us consider the first of our questions. For this, we have the following theorem.

Theorem 1. *Consider the replicator dynamics (2).*

I Any Nash equilibrium $\bar{\mathbf{x}}$ is an equilibrium of RD.

II If $\bar{\mathbf{x}}$ is the ω-limit of an orbit $\mathbf{x}(t)$ and $\bar{\mathbf{x}} \in int\Delta$, then $\bar{\mathbf{x}}$ is an NE.

III If $\bar{\mathbf{x}}$ is Lyapunov stable, then it is an NE.

Before we continue with the proof, we note that there is a brief introduction to concepts in dynamical systems provided in the Appendix, which includes the definition of an ω-limit set and Lyapunov stability.

Proof. (*I*) Let us first take the case in which $\bar{\mathbf{x}}_i = 0$. Then the right hand side of (2) is automatically zero. Now let us assume $\bar{\mathbf{x}}_i > 0$. We know from Lemma 1.1 that, for an NE $(\mathbf{A}\bar{\mathbf{x}})_i = \bar{\mathbf{x}} cdot \mathbf{A}\bar{\mathbf{x}}$. In this case the second term on the right hand side of (2) vanishes and we still attain $\dot{\bar{x}}_i = 0$. Therefore, $\bar{\mathbf{x}}$ is a fixed point of (2).

(*II*) Let us assume that $\bar{\mathbf{x}}$ is not an Nash Equilibrium. Then, there exists an i such that $e_i \cdot \mathbf{A}\bar{\mathbf{x}} > \bar{\mathbf{x}} \cdot A\bar{\mathbf{x}}$. Therefore, there is some ϵ for which $(\mathbf{A}\bar{\mathbf{x}})_i - \bar{\mathbf{x}} \cdot \mathbf{A}\bar{\mathbf{x}} > \epsilon$ and so $\dot{\bar{x}}_i > \epsilon \bar{\mathbf{x}}_i$ for all i. This, however, is impossible since we assumed $\bar{\mathbf{x}}$ was the ω-limit of some orbit.

(*III*) The argument works in a similar manner to that of (*II*). If we assume that $\bar{\mathbf{x}}$ is not an NE, then there exists some i and some ϵ for which $(\mathbf{A}\bar{\mathbf{x}})_i - \bar{\mathbf{x}} \cdot \mathbf{A}\bar{\mathbf{x}} > \epsilon$ for all $\mathbf{x}$ in a neighbourhood of $\bar{\mathbf{x}}$. In this case, $\mathbf{x}_i$ increases, which contradicts the assumption of Lyapunov stability. □

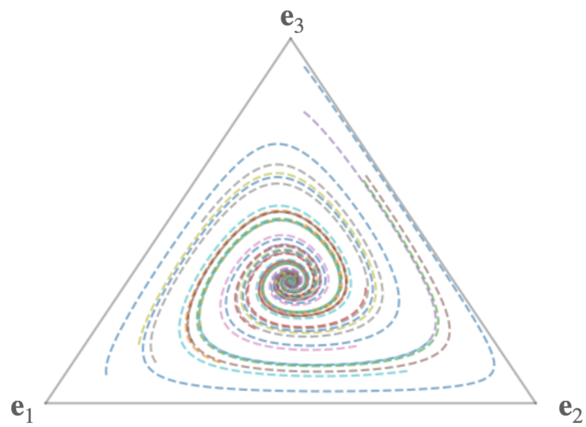


Figure 2: Trajectories of the Replicator Dynamic (2) under the symmetric game in Example 9.

Example 2. *Consider the symmetric game* $\mathbf{A}$ *defined by*

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & -0.5 \\ -0.5 & 0 & 1 \\ 1 & -0.5 & 0 \end{pmatrix} \tag{4}$$

We will see soon that this is an example of a 'monocylic' game, a class of games which shows some quite remarkable properties under the replicator dynamics. The flow under these dynamics is shown in Figure 2.

Notice that $\bar{\mathbf{x}} = (1/3, 1/3, 1/3)^T \in int\Delta$ *is the* ω*-limit of orbits through interior points. As such, we can see that it is an NE of the game* $\mathbf{A}$*. We can also verify this by seeing that the rows of* $\mathbf{A}$ *add to 1. Therefore* $(\mathbf{A}\bar{\mathbf{x}})_1 = (\mathbf{A}\bar{\mathbf{x}})_2 = (\mathbf{A}\bar{\mathbf{x}})_3$ *so* $\bar{\mathbf{x}}$ *is an NE.*

2 A brief excursion into Evolutionary Game Dynamics

We can, in fact, relate the equilibria of RD to the game by invoking a stronger notion of equilibrium.

Definition (Evolutionary Stable Strategy)**.** In a symmetric game $\mathbf{A}$, a strategy $\bar{\mathbf{x}}$ is an Evolutionary Stable Strategy (ESS) if, for all $\mathbf{x} \in \Delta$

$$\mathbf{x} \cdot \mathbf{A}(\epsilon \mathbf{x} + (1-\epsilon)\bar{\mathbf{x}}) < \bar{x} \cdot \mathbf{A}(\epsilon \mathbf{x} + (1-\epsilon)\bar{\mathbf{x}}) \tag{5}$$

From an EGT perspective, the ESS can be understood by considering a population of agents who play some mixed strategy $\bar{x}$. Suppose they were to be invaded by a small number of agents who play a 'mutant' strategy $\mathbf{x}$. We formalise that this is a small group of mutants by saying that their proportion ϵ in the population is sufficiently small. Now, if a randomly chosen agent is pitted against another random individual, the probability with which the opponent is playing $\mathbf{x}$ is ϵ, whilst the probability that they are a non-mutant is $1-\epsilon$. If the first player is a non-mutant, then the payoff they will receive is

$$\bar{x} \cdot \mathbf{A}(\epsilon\mathbf{x} + (1-\epsilon)\bar{\mathbf{x}}).$$

On the other hand, if they are a mutant, they will receive

$$\mathbf{x} \cdot \mathbf{A}(\epsilon\mathbf{x} + (1-\epsilon)\bar{\mathbf{x}})$$

The ESS condition, in some sense, says that a randomly chosen agent is better off in terms of payoff by being a non-mutant rather than a mutant. As such, the strategy $\bar{\mathbf{x}}$ is stable against invading groups. An equivalent definition of the ESS can be derived through the following Lemma.

Lemma 2. *The ESS condition is equivalent to the statement that, for all $\mathbf{x} \neq \bar{\mathbf{x}}$ sufficiently close to $\bar{\mathbf{x}}$*

$$\mathbf{x} \cdot \mathbf{A}\mathbf{x} < \bar{\mathbf{x}} \cdot \mathbf{A}\mathbf{x} \tag{6}$$

Moreover, if $\bar{\mathbf{x}} \in int\Delta$ is an ESS then this holds for all $\mathbf{x} \in \Delta \backslash \{\bar{\mathbf{x}}\}$

Proof. c.f. ([8] Lemma 1.3) □

The ESS relates to the NE in the following manner.

Lemma 3. *$\bar{\mathbf{x}}$ is a strict NE $\implies$ $\bar{\mathbf{x}}$ is an ESS $\implies$ $\bar{\mathbf{x}}$ is an NE*

Proof. If $\bar{\mathbf{x}}$ is a strict NE, then $\mathbf{x} \cdot \mathbf{A}\bar{\mathbf{x}} < \bar{\mathbf{x}} \cdot \mathbf{A}\bar{\mathbf{x}}$ for any $\mathbf{x} \in \Delta$. By continuity, the inequality must hold with small ϵ, i.e.

$$\mathbf{x} \cdot \mathbf{A}(\epsilon\mathbf{x} + (1-\epsilon)\bar{\mathbf{x}})$$

which is exactly the ESS condition. Now if we assume $\mathbf{x}$ is an ESS, then we can take the limit as $\epsilon \to 0$ to yield that $\mathbf{x} \cdot \mathbf{A}\bar{\mathbf{x}} \leq \bar{\mathbf{x}} \cdot \mathbf{A}\bar{\mathbf{x}}$ for all $\mathbf{x} \neq \bar{\mathbf{x}}$.

□

Lemma 4. *If $\bar{\mathbf{x}} \in int\Delta$ is an ESS, then there are no other NE.*

Proof. If $\bar{\mathbf{x}}$ is an ESS, then it is also an NE, so that $(\mathbf{A}\bar{\mathbf{x}})_i = \bar{\mathbf{x}} \cdot \mathbf{A}\bar{\mathbf{x}}$ for all i. We also have the ESS condition that $\mathbf{x} \cdot \mathbf{A}\mathbf{x} < \bar{\mathbf{x}} \cdot \mathbf{A}\mathbf{x}$ for all $\mathbf{x} \neq \bar{\mathbf{x}}$. In this case, $\mathbf{x}$ cannot be an NE, since it is not a best response to itself.

□

Example 3. *Consider the symmetric game $\mathbf{A}$ defined by*

$$\mathbf{A} = \begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix} \tag{7}$$

The rows of the matrix add to 1 and so we have that $\bar{\mathbf{x}} = (0.5, 0.5)^T$ is an NE of the game. This is not strict since we have that $\mathbf{x} \cdot \mathbf{A}\bar{\mathbf{x}} = \bar{\mathbf{x}} \cdot \mathbf{A}\bar{\mathbf{x}}$ so that $BR^A(\bar{\mathbf{x}}) = \Delta$.

Now let us check whether $\bar{\mathbf{x}}$ is also an ESS. Let us write $\mathbf{x} = \bar{\mathbf{x}} + \begin{pmatrix} \epsilon \\ -\epsilon \end{pmatrix}$. Then $\mathbf{A}\mathbf{x} = \begin{pmatrix} 1/2 - 3\epsilon \\ 1/2 - \epsilon \end{pmatrix}$ so that

$$\mathbf{x} \cdot \mathbf{A}\mathbf{x} = \left(\bar{\mathbf{x}} + \begin{pmatrix} \epsilon \\ -\epsilon \end{pmatrix}\right) \cdot \left(\bar{\mathbf{x}} + \begin{pmatrix} -3\epsilon \\ -\epsilon \end{pmatrix}\right)$$
$$= \frac{1}{2} - 2\epsilon - 2\epsilon^2.$$

On the other hand we have that

$$\bar{\mathbf{x}} \cdot \mathbf{A}\mathbf{x} = 1/2 - 2\epsilon.$$

So that

$$\mathbf{x} \cdot \mathbf{A}\mathbf{x} < \bar{\mathbf{x}} \cdot \mathbf{A}\mathbf{x}.$$

Therefore, by Lemma 2, $\bar{\mathbf{x}}$ is an ESS. Furthermore, by Lemma 4, $\bar{\mathbf{x}} \in int\Delta$ is the only NE.

3 RD and ESS

We now return to our analysis of the replicator dynamics with the following theorem

Theorem 2. *If $\bar{\mathbf{x}}$ is an ESS then it is asymptotically stable under RD. Moreover, if $\bar{\mathbf{x}} \in int\Delta$ then it is globally attracting for the entire simplex under the NE.*

Proof. Consider the function $P(\mathbf{x}) = \prod_i \mathbf{x}_i^{\bar{\mathbf{x}}_i}$. Our first claim is that this has a unique maximiser at $\bar{\mathbf{x}}$. This follows from a direct application of Jensen's inequality, namely that for a strictly convex function f defined on some interval I

$$f(\sum_i p_i x_i) \leq \sum_i p_i f(x_i)$$

for all $x_1, \ldots x_n \in I$ and all $p_1, \ldots, p_n \in \Delta$ with equality if and only if all x_i are equal. This gives the result that $\sum_i \bar{\mathbf{x}}_i log(\frac{x_i}{\bar{\mathbf{x}}_i}) \leq log\left(\sum_i x_i\right) = log1 = 0$. This holds since log is a strictly concave function and so the inequality is reversed. Therefore, $\sum_i \bar{\mathbf{x}}_i log\mathbf{x}_i < \sum_i \bar{\mathbf{x}}_i log\bar{\mathbf{x}}_i$. Since log is monotone, this means that $P(\mathbf{x}) \leq P(\bar{\mathbf{x}})$, with equality if and only if $\mathbf{x} = \bar{\mathbf{x}}$. This shows that the maximiser is unique.

We now show that P gives rise to a Lyapunov function. Notice

$$\frac{\dot{P}}{P} = \frac{d}{dt} logP = \frac{d}{dt} \sum_i \bar{\mathbf{x}}_i log\mathbf{x}_i = \sum_i \bar{\mathbf{x}}_i \frac{\dot{\mathbf{x}}_i}{\mathbf{x}_i}$$
$$= \sum_i \bar{\mathbf{x}}_i((\mathbf{A}\mathbf{x})_i - \mathbf{x} \cdot \mathbf{A}\mathbf{x}) = \bar{\mathbf{x}} \cdot \mathbf{A}\mathbf{x} - \mathbf{x} \cdot \mathbf{A}\mathbf{x}.$$

Since $\bar{\mathbf{x}}$ is assumed to be an ESS, we know that $\bar{\mathbf{x}} \cdot \mathbf{A}\mathbf{x} - \mathbf{x} \cdot \mathbf{A}\mathbf{x} > 0$ so that $\dot{P} > 0$. Therefore, P satisfies the conditions required for a Lyapunov function, and so $\bar{\mathbf{x}}$ is asymptotically stable (i.e. orbits starting near $\bar{\mathbf{x}}$ converge to $\bar{\mathbf{x}}$.)

□

4 Permanence

Now we turn to the second of our questions, namely: if a phenotype is present in the initial population, will it remain in the population for all time? We frame this question mathematically through the idea of *permanence* which allows us to determine which strategies and phenotypes are safe from extinction. Formally we write this as

Definition (Permanence)**.** The replicator equation (2) is permanent if there exists a compact set $K \subset int\Delta$ such that, for all interior initial conditions $\mathbf{x}(0) \in int\Delta$ there is a T such that, for all $t \geq T$, $\mathbf{x}(t) \in K$. Equivalently, RD is permanent if there is a $\delta > 0$ such that $\liminf_{t\to\infty} x_i(t) \geq \delta$ for all i whenever the initial condition $\mathbf{x}(0) \in int\Delta$.

A particular class of games which shows the permanence property is the symmetric Rock-Paper-Scissors game. In fact, we can generalise this to an even wider class of games.

Definition (Monocyclic games)**.** A symmetric game $\mathbf{A}$ is *monocyclic* if, for all i $(A)_{ii} = 0$, $(A)_{ij} > 0$ if $i = j + 1 (\text{mod } n)$ and $(A)_{ij} \leq 0$ otherwise.

Example 4. *Monocyclic Game Example*

Theorem 3 (Hofbauer and Sigmund '88 [2])**.** *If the symmetric game $\mathbf{A}$ is monocyclic, then the replicator dynamics (2) is permanent if and only if there is a rest point $\mathbf{z} \in int\Delta$ with $\mathbf{z} \cdot A\mathbf{z} > 0$.*

Theorem 4 (Hofbauer and Sigmund '88 [2])**.** *If RD is permanent, then there exists a unique rest point $\mathbf{z} \in int\Delta$. The time averages of the orbit converge to $\mathbf{z}$, that is*

$$\frac{1}{T}\int_0^T x_i(t)dt \to z_i \tag{8}$$

for $T \to \infty$ and $i = 1, \ldots, n$. Conversely, if there are are no rest points in $int\Delta$, then every orbit converges to the boundary of the simplex $\partial\Delta$

5 Complex Dynamics in the Replicator Equation

5.1 Two Player Replicator Dynamics

So far we have considered symmetric games, which only involve one player. Let us now return to our original study of two-player bimatrix games. In this setting, the Replicator Dynamic can be written as[3]

$$\begin{aligned} \dot{x}_i - x_i((\mathbf{A}\mathbf{y})_i - \mathbf{x} \cdot \mathbf{A}\mathbf{y}) \\ \dot{y}_i - y_i((\mathbf{B}^T\mathbf{x})_i - \mathbf{x} \cdot \mathbf{B}\mathbf{y}) \end{aligned} \tag{9}$$

[3] there are actually two different conventions for writing the replicator dynamic which stem from different conventions with which to write a bimatrix game. We choose the most commonly seen convention here and stick with it, since the other convention can be found by simply transposing $\mathbf{B}$. c.f. [8] Chapter 2 for a discussion on these conventions.

Example 5 (Iterated Prisoner's Dilemma)**.** *We cannot, in good conscience, report on a game theoretic concept without referring back to the Prisoner's Dilemma.*

Here, we consider a slight variation, known as the Iterated Prisoner's Dilemma (IPD). In this example, the prisoners are repeatedly arrested in their connection with a serious crime[4] *and must make the same decision as before. However, they will now base their decision on their prior experience in playing the game with their opponent.*

In particular, both agents will choose to either Always Cooperate (AllC) with the other prisoner (i.e. keep shut), Always Defect (AllD) (i.e. confess) or engage in Tit-for-Tat (TFT). The final of these states that, if an agent sees that their opponent cooperated in the last round of the game, then they will do so as well. However, if the opponent defected, then they will follow suit in this round. Note that the strategy AllD is what game theory predicts to be the natural outcome of the game, since the Nash Equilibrium of the game was to confess to the crime.

This game was studied experimentally in 1980 by Robert Axelrod. In this experiment, he asked participants to submit programs who would play the IPD game against each other. It was found that the agent which performed best was the TFT strategy. This came as something of a surprise since it went against the common notion that one should act 'selfishly', which is what AllD requires. Instead TFT encodes some form of altruism in the sense that agents will work with each other if they see that the other is cooperating. However, they will punish any defection by defecting in the next turn.

The IPD game is studied in [3] and we discuss some of the results here. Noting that the strategy set of the agents are {AllC, AllD, TFT}. Then the game is defined through the payoff matrix

$$\mathbf{A} = \begin{pmatrix} Rm & Sm & Rm \\ Tm & Pm & T + P(m-1) \\ Rm - c & S + P(m-1) - c & Rm - c \end{pmatrix} \tag{10}$$

In which $T > 4R > P > S$, $m > 0$ is the expected number of rounds in the game and $c > 0$ is a small cost for choosing the TFT strategy. In Figure 3 we show the action of the replicator dynamic as well as a variant known as the replicator-mutator equation. The latter considers the case in which the agent randomly changes their mixed strategy at some mutation rate u (c.f. [3]). This introduces some stochasticity into the strategy evolution. It can be seen that, by increasing u, the dynamics move from converging to the AllD strategy to a cycle near the TFT strategy.

5.2 Learning in Two-Player Zero Sum Games is a Hamiltonian System

In this section we consider what is, in the author's opinion, one of the most important results that bridge the gap between Game Theory and Dynamical Systems. In particular, we find that when agents update their action profiles through the replicator dynamic, the flow is akin to that of an incompressible fluid. Specifically, we will see that volume elements are conserved. In fact, we will go even further to learn that the replicator dynamic come under

[4] it never occurred to either to give up their life of crime since they get caught so often

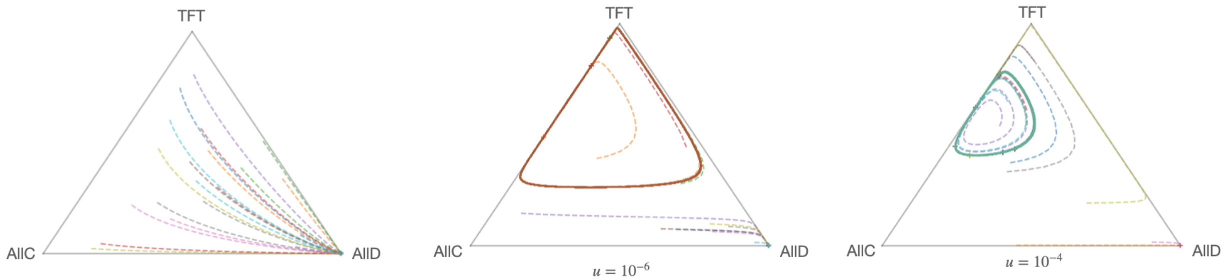


Figure 3: (Left) Trajectories of Replicator Equation (2) in the IPD Game. (Middle) Trajectories of the Replicator-Mutator Equation [3] for $u = 10^{-6}$ (Right) $u = 10^{-4}$.

the class of Hamiltonian Systems[5] [6]. We will also see that these facts mean that, in zero sum games, learning through replicator does not lead to an equilibrium. Instead, they give rise to periodic behaviour, and even chaos.

Example 6. *In Figures 4a and 4b we look at two zero-sum games - namely Matching Pennies and Rock-Paper-Scissors. These games are zero sum in the sense that* $\mathbf{B} = -\mathbf{A}$, *so we only need* $\mathbf{A}$ *to define the game. For Matching Pennies we have*

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \tag{11}$$

For Rock-Paper-Scissors we have

$$\mathbf{A} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix} \tag{12}$$

Since these matrices both have rows of constant sum, the NE lies in the centre of the simplex. From the Figure we can see that these take the form of topological centres. This means that orbits form closed ellipses around the NE. In particular, we notice that the flow is periodic.

Theorem 5. *Let* $(\mathbf{A}, -\mathbf{A})$ *be a zero-sum bimatrix game. Then, the flow generated by the replicator dynamics (9) is volume-preserving.*

Proof. We begin by making the transformation

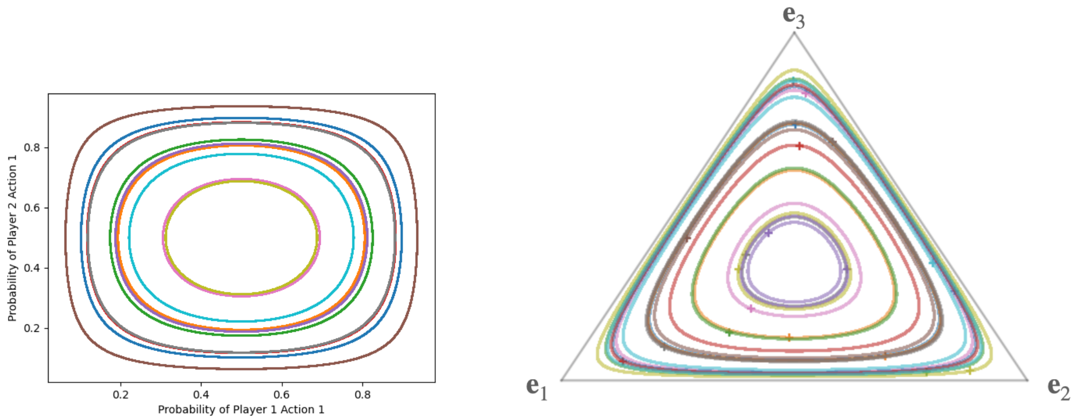
$$u_i = ln\frac{x_{i+1}}{x_1} \qquad v_i = ln\frac{y_{i+1}}{y_1}.$$

The inverse of this transformation is given by

$$x_{i+1} = \frac{e^{u_i}}{1 + \sum_{j=1}^{n-1} e^{u_j}}$$

Then

[5] Again, we discuss these concepts in the Appendix



(a) Matching Pennies Game (11) *(b) Rock-Paper-Scissors Game (12)*

Figure 4: Trajectories under Replicator Dynamic (9) for the (a) Matching Pennies Game and (b) Symmetric Rock-Paper-Scissors game. Both admit the centre of the simplex as an NE, and therefore a fixed point of (9). The nature of this fixed point is that of a topological centre [6] (i.e. closed orbits around the equilibrium).

$$\begin{aligned} \dot{u}_i &= \frac{1}{x_{i+1}}\dot{x}_{i+1} - \frac{1}{x_1}\dot{x}_1 & \dot{v}_i &= \frac{1}{y_{i+1}}\dot{y}_{i+1} - \frac{1}{y_1}\dot{y}_1 \\ &= (\mathbf{Ay})_{i+1} - (\mathbf{Ay})_1 & &= (\mathbf{B}^T\mathbf{x})_{i+1} - (\mathbf{B}^T\mathbf{x})_1 \\ &= \frac{\sum_i (a_{ij} - a_{1j})e^{v_j} + (a_{i1} - a_{11})}{1 + \sum_j e^{v_j}} & &= \frac{\sum_i (b_{ji} - b_{j1})e^{u_j} + (b_{1i} - b_{11})}{1 + \sum_j e^{u_j}} \end{aligned}$$

Now let us take the divergence of the vector field acting on $(\mathbf{u}, \mathbf{v})$

$$\nabla \cdot F = \sum_i \frac{\partial \dot{u}_i}{\partial \dot{u}_i} + \frac{\partial \dot{v}_i}{\partial \dot{v}_i} = 0.$$

Now, by Louiville's Formula (c.f. Appendix), we can say that the rate of change of volume $\frac{d}{dt} vol(\Omega(t)) = 0$ for any set of initial conditions $\Omega(0)$.

□

We now take this result even further with the following theorem.

Theorem 6 (Hofbauer '98)**.** *Let $(\mathbf{A}, -\mathbf{A})$ be a zero-sum bimatrix game. The replicator dynamic (9) is a Hamiltonian System. In particular, it can be written in the form*

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = J \nabla H \tag{13}$$

where J is a skew-symmetric matrix given by

$$\begin{pmatrix} 0 & A \\ -A & 0 \end{pmatrix} \tag{14}$$

and H is a Hamiltonian *function given by*

$$H(\mathbf{x}, \mathbf{y}) = D_{KL}(\bar{\mathbf{p}}||\mathbf{x}) + D_{KL}(\bar{\mathbf{q}}||\mathbf{y}) = \sum_i \bar{\mathbf{p}}_i \, ln\frac{\bar{\mathbf{p}}_i}{\mathbf{x}_i} + \sum_i \bar{\mathbf{q}}_i \, ln\frac{\bar{\mathbf{q}}_i}{\mathbf{y}_i}$$

Proof. Let us take the transformation

$$u_i = -lnx_i + N^{-1}\sum_k lnx_k \qquad v_i = -lny_i + N^{-1}\sum_k lny_k$$

Under this transformation the Replicator Dynamic becomes

$$u_i = (\mathbf{A}\mathbf{y})_i - \sum_j (\mathbf{A}\mathbf{y})_j$$
$$v_i = (\mathbf{B}^T\mathbf{x})_i - \sum_j (\mathbf{B}^T\mathbf{x})_j$$

Then, through some rote manipulation, we can find that

$$\begin{aligned} &\frac{\partial H}{\partial u_i} = (\mathbf{x}_i - \bar{\mathbf{p}}_i) \\ &\implies -\mathbf{A}\frac{\partial H}{\partial u_i} = -\mathbf{A}(\mathbf{x}_i - \bar{\mathbf{p}}_i) \\ &\implies -\mathbf{A}\frac{\partial H}{\partial u_i} = -\mathbf{A}\mathbf{x}_i. \end{aligned}$$

Similarly, $\mathbf{A}\frac{\partial H}{\partial v_i} = \mathbf{A}\mathbf{y}_i$. The final inequality holds since we can assume, without loss of generality, that $\sum_i a_{ij} = 0$. Finally, this means that

$$J\nabla H = \begin{pmatrix} \mathbf{A}\mathbf{y} \\ -\mathbf{A}\mathbf{x} \end{pmatrix} = \begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix}$$

□

References

[1] Cesa-Bianchi, N., and Orabona, F. Online learning algorithms, 3 2021.

[2] Hofbauer, J., and Sigmund, K. EVOLUTIONARY GAME DYNAMICS. Tech. Rep. 4, 2003.

[3] Imhof, L. A., Fudenberg, D., and Nowak, M. A. Evolutionary cycles of cooperation and defection. In *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102. 8 2005, pp. 10797–10800.

[4] Mertikopoulos, P., Papadimitriou, C., and Piliouras, G. Cycles in adversarial regularized learning. Tech. rep., 2018.

[5] Sato, Y., Akiyama, E., and Farmer, J. D. Chaos in learning a simple two-person game. *Proceedings of the National Academy of Sciences of the United States of America 99*, 7 (4 2002), 4748–4751.

[6] Shafer, D. S. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering (Steven H. Strogatz)*, vol. 37. Westview Press, 1995.

[7] Smith, J. M. *Evolution and the Theory of Games.* Cambridge University Press, 10 1982.

[8] Strien, S. V. Games and Dynamics : Alternative title : Dynamical learning models (in strategic environments) Lecture Notes, 2020.