Poincaré-Bendixson Limit Sets in Multi-Agent Learning

Aleksander Czechowski
Delft University of Technology

Georgios Piliouras
Singapore University of Technology and Design

ABSTRACT

A key challenge of evolutionary game theory and multi-agent learning is to characterize the limit behaviour of game dynamics. Whereas convergence is often a property of learning algorithms in games satisfying a particular reward structure (e.g. zero-sum), it is well known, that for general payoffs even basic learning models, such as the replicator dynamics, are not guaranteed to converge. Worse yet, chaotic behavior is possible even in rather simple games, such as variants of Rock-Paper-Scissors games [35]. Although chaotic behavior in learning dynamics can be precluded by the celebrated Poincaré-Bendixson theorem, it is only applicable to low-dimensional settings. Are there other characteristics of a game, which can force regularity in the limit sets of learning?

In this paper, we show that behaviors consistent with the Poincaré-Bendixson theorem (limit cycles, but no chaotic attractor) follows purely based on the topological structure of the interaction graph, even for high-dimensional settings with arbitrary number of players and arbitrary payoff matrices. We prove our result for a wide class of follow-the-regularized leader (FoReL) dynamics, which generalize replicator dynamics, for games where each player has two strategies at disposal, and for interaction graphs where payoffs of each agent are only affected by one other agent (i.e. interaction graphs of indegree one). Since chaos has been observed in a game with only two players and three strategies, this class of non-chaotic games is in a sense maximal. Moreover, we provide simple conditions under which such behavior translates to social welfare guarantees, implying that FoReL learning achieves time average social welfare which is at least as good as that of a Nash equilibrium; and connecting the topology of the dynamics to the Price of Anarchy analysis.

KEYWORDS

Replicator Dynamics, Follow-the-regularized-leader, Regret Minimization, Poincaré-Bendixson Theorem, Polymatrix Games, Price of Anarchy

1 INTRODUCTION

Understanding and predicting the behavior of learning dynamics in normal form games has been a fundamental question that has attracted the question of researchers from diverse disciplines such as economics, optimization theory, artificial intelligence a.o. [6, 36, 37, 40]. Even in simple games, such as Rock-Paper-Scissors [25, 29, 35] models of evolution and learning are not guaranteed to converge; and even beyond cycles, long-term behavior may lead to chaotic behavior, known to the dynamical systems community, e.g., from weather models [22]. Not only does chaos manifest itself even in simple games with two players, but moreover, a string of recent results seems to suggest that such chaotic, unpredictable behavior may indeed be the norm across a variety of simple low dimensional game dynamics [1–3, 7, 8, 10, 13, 15, 27, 28, 33, 38]. Worse yet, the emergence of chaotic behavior has been connected with increased social inefficiency showing that even in games where with a unique socially optimum equilibrium, i.e. games with Price of Anarchy [31] equal to 1, chaotic dynamics [9] may lead to highly inefficient outcomes. Such profoundly negative results lead us to the following natural questions:

- Do there exist simple, robust conditions under which learning behaves well?
- Which type of games lie at the “edge of chaos"?
- Does dynamic simplicity translate to high efficiency, social welfare?

Traditionally, a lot of work has focused on showing that in specific classes of games (e.g., potential games), learning dynamics can lead to convergence and equilibration, (e.g. [6, 14, 34, 40] and references therein). One disadvantage of these approaches is that these classes of games are non-generic, i.e., the set of such games is of measure zero in the space of all games. There is a negligible chance that a game will satisfy any of these properties (e.g., be an exact potential game). Hence, such results are typically contingent on the stringent assumption that the agents must internalize an abstract game theoretic model to an arbitrary high degree of accuracy.

Another approach to multi-agent learning that has been relatively underexplored so far is the possibility of interpreting simple non-equilibrating behavior cyclic/periodic behavior as an example of a positive/helpful regularity. Instead, it is typically categorized together with chaos under the label of disequilibrium behavior despite the vast difference between them. In the closest work to our, recently, [26] explores the possibility of such non-equilibrating regularities, however, once again they do so by assuming non-generic structure on the set of allowable games (e.g., network of $2x2$ zero-sum/ coordination games). In terms of connections between cyclic behavior and the efficiency of learning dynamics, [20] has shown that for a class of three agent, two strategy games a cyclic attractor can result in social welfare (sum of payoffs) that can be better than the Nash equilibrium payoff [20], however, that result is once again constrained to exact game theoretic model.

The importance as well as the difficulty of understanding limit cycle behavior even in planar (i.e., two dimensional) dynamical systems with polynomial vector fields is captured by Hilbert's 16th open problem, which remains unresolved to this day [18]. Despite the hardness of such questions, progress in this direction is necessary as otherwise, this precludes us from studying even the simplest case of Matching Pennies under replicator dynamics, which is a polynomial, planar vector field which is everywhere periodic.

Our approach and results. To make progress along this hard task, we explore a different type of constraint in games. We constrain only the combinatorial structure of the game. How many strategies does each player have? We only allow two. What are the allowable interactions between the agents? Every agent is affected by the behavior of up to one other agent. Finally, we add a technical restriction that the game is connected – it cannot be decomposed into two subgames completely independent of each other. Under these assumptions, we prove our main contribution in form of Theorems 3, 4 – that the limit behavior of FoReL [25, 36] of this games is always consistent with the famous Poincaré-Bendixson theorem, which informally states that the system is either convergent or cyclic, and in particular no chaotic attractor is not possible.

Furthermore, under additional but structurally robust assumptions on the payoff matrices (i.e., assumptions that remain valid after small perturbations of the payoff matrices), we prove positive results about the efficiency of the time-average behavior of the dynamics regardless of whether they are convergent or not. As it is typically the case in the Price of Anarchy literature [21] we focus on the measure of *social welfare* – the sum of individual payoffs, but whereas the the typical PoA literature tries to argue that regret-minimizing dynamics (such as FoReL) are at most a constant factor worse than the behavior of the worst case Nash equilibrium [31, 32], we instead show that FoReL dynamics always at least as efficient as the worst case Nash equilibrium. Finally, in Section 5 we provide examples in form of sample trajectories in high-dimensional games, where limit behavior is non-chaotic, but cyclic.

1.1 Related work

Limit sets in two-player two-strategy games have been very well understood. In particular, the celebrated Poincaré-Bendixson theorem states that all smooth two-dimensional systems can have only stationary and cyclic limit sets [4]. Since the evolution of mixed strategies of both agents can be described by two variables only, the theorem implies that no chaotic behavior can emerge. Poincaré-Bendixson theorem has been successfully applied in the past also to higher-dimensional learning systems [12, 26]. The key technique of these papers was to show that the underlying dynamic is in fact two-dimensional, by finding constants of motion. Contrary to that, in this paper our analysis is not contigent on identifying any invariant function while at the same time exploring truly high-dimensional games without any restriction on the number of players. Moreover, our theorems apply to a wide class of learning models in games, so-called Follow-the-regularized-Leader (FoReL) systems, where agents evolve their mixed strategies in the direction of maximal reward, but with taking into account a regularizer term, which models exploratory behavior [24]. This class has a number of strong properties such as finite regret and contains as special cases e.g. variants of replicator dynamics [25].

Due to the chaotic example of Sato et al. [35] in two player, three action games as well another negative example by Plank [30] of complex quasi-cyclic behavior in three player games with two actions for each player but without a structured network of interaction both for replicator dynamics, (c.f. Figure 2), our topological results establish a maximal class of games for which such positive results are possible.

2 PRELIMINARIES

2.1 Normal form games

A *finite game in normal form* consists of a set of N players each with a finite set of *strategies* $\mathcal{A}_i$. The preferences of each player are represented by the payoff function $u_i : \prod_i \mathcal{A}_i \to \mathbb{R}$. To model behavior at scale, or probabilistic strategy choices, one assumes that players use *mixed strategies*, i.e. probability distributions $(x_{i\alpha_i})_{\alpha_i \in \mathcal{A}_i} \in \Delta(\mathcal{A}_i) =: \mathcal{X}_i$. With slight abuse of notation, the expected payoff of the i-th player in the profile $(x_{i\alpha_i})_{i,\alpha_i}$ is denoted by u_i again, and given by

$$u_i(x) = \Sigma_{\alpha_1 \in \mathcal{A}_1, \ldots \alpha_N \in \mathcal{A}_N} u_i(\alpha_1, \ldots, \alpha_N) x_{1\alpha_1} \ldots x_{N\alpha_N}. \quad (1)$$

A mixed strategy $\hat{x}$ is a *Nash equilibrium* iff $\forall i \forall x : x_j = \hat{x}_j,\ j \neq i$ we have $u_i(x) \leq u_i(\hat{x})$; in other words no player can unilaterally increase their payoff by changing their strategy distribution. The *minimax value* for player i is given by $\min_{x_{-i}} \max_{x_i} u_i(x)$, where $x_{-i} := (x_j)_{j \neq i}$. It is the smallest possible value player i can be forced to attain by other players, without them knowing player i strategy. We call a game *binary* iff each agent only has two strategies at their disposal, i.e. $|A_i| = 2$ for all i.

2.2 Graphical polymatrix games

To model the topology of interactions between players, we restrict our attention to a subset normal form games, where the structure of interactions between players can be encoded by a graph of two-player normal form subgames, leading us to consider so-called *graphical polymatrix games* (GPGs) [17, 19, 39]. A simple directed graph is a pair (V, E), where $V = \{1, \ldots, N\}$ is a finite set of *vertices* (representing the players), and E is a set ordered distinct vertex pairs (*edges*), where the first element is called predecessor, and the second is called successor. Each vertex (i, k) has an associated two-player normal form game, where only the successor k is assigned payoffs, and they are represented by a matrix $A^{i,k}$ with rows enumerating strategies of player k, and columns enumerating strategies of player i. For a given strategy profile $s = \{s_i\}_i \in \Pi_i S_i$ the payoffs for player k in the full game are then determined as the sum

$$u_k(s) = \sum_{i:(i,k) \in E} A^{i,k}(s_i, s_k) \quad (2)$$

The payoffs can be extended to mixed strategies in a usual multilinear fashion:

$$u_k(x) = \sum_{i:(i,k) \in E} \sum_{x_{s_i}, x_{s_k}} A^{i,k}(s_i, s_k) x_{s_i} x_{s_k}. \quad (3)$$

Note that a situation, where both the successor k and also the predecessor i obtain a reward can be modelled by including both edges (i, k) and (k, i) in the graph.

We say that a simple directed graph is weakly connected, if any two vertices can be connected by a set of edges, where the direction of the edges is not taken into account. This is a weaker condition than strong connectedness, where each pair of vertices needs to be connected by a *path*, i.e. a sequence of edges, together with associated vertices, where the successor in one edge needs to be the predecessor in the next one. The *indegree* of a vertex, is the amount of edges for which the vertex is the successor (in other words: the number of its predecessors). The *outdegree* is the amount of edges, for which the vertex is the predecessor, i.e. the number of its successors. A *cycle* is a path, where the predecessor in the first edge is the successor in the last edge. For our exposition we shall identify cycles modulo shifts, i.e. if two paths consist of the same edges in shifted order, then they form the same cycle.

In this paper we consider two types of connected GPGs:

(1) firstly, *cyclic* games, where the interaction between the agents forms a cycle, where each agent interacts only with the previous neighbor. We observe that in such a cyclic game the indegree and outdegree of each vertex is one.
(2) Secondly, a more general class of graphical games, where each player's payoffs depend on up to one other player, i.e. the indegree of each vertex is at most one. For a vertex $i \in V$, we will then denote the predecessor vertex by $\hat{i}$. For cyclic games we have $\hat{i} \mod N = i - 1 \mod N$.

Below, we state and prove a simple lemma, which characterizes the one-predecessor assumption in terms of graph topology (c.f. Figure 1).

Lemma 1. *Let (V, E) be a weakly connected, simple, directed graph. If the indegree of each vertex is at most one, then, the graph can have up to one cycle. If the graph has no cycle, then it has to have at most one* root *vertex, i.e. a vertex of indegree zero, such that all other vertices are connected to it by a unique, directed path.*

Proof. For the first part of the lemma, let us assume the contrary: that a_1, a_2 are nodes of two distinct cycles within the same weakly connected component. The edges between a_1 and a_2 need to form a path (otherwise there would be a vertex with two predecessors). Assume the path leads from a_1 to a_2, and let a_0 be the first vertex which is both on the path and on a_2 cycle. Then a_0 has two predecessors, which leads to a contradiction.

For the second part of the lemma, we can argue as follows. If any vertex would have a sequence of predecessors which would not form a cycle, then by backtracking through the predecessors we could identify an infinite collection of distinct vertices. Therefore, there needs to be at least one root node for each vertex. The path from such root node to the given vertex needs to be unique, otherwise one could identify a vertex along the path with two predecessors. Finally, if there were two root nodes, from connectedness it follows that there must be a node with two predecessors on the edges between them. □

Remark 1. *Under the assumptions of Lemma 1, if the graph has a cycle, then the cycle serves the role of the root node; i.e. there are no paths from outside of the cycle to it (otherwise one vertex in the cycle would have two predecessors), and all vertices outside of the cycle have to be connected by a path from one of the vertices of the cycle (unique, up to the starting point within the cycle). This follows from the same arguments as in the second part of the proof of Lemma 1. Further on we will refer to such cycle as the* root cycle.

2.3 Follow-the-regularized-leader equations

Denote by $v_{i\alpha_i}(x) := u_i(\alpha_i; x_{-i})$ and $v_i(x) = (v_{i\alpha_i}(x))_{\alpha_i \in \mathcal{A}_i}$. To model the dynamics of learning we use a class of learning systems known as *follow-the-regularized-leader* systems (FoReL) [6, 36]. This class encompasses a variety of models such as gradient and replicator, and allows for natural description of agent learning as regularized maximization of individual utilities.

FoReL dynamics for player i are defined by evolution of so-called *utilities* $y_i = \{y_{i\alpha_i}\}_{\alpha_i \in \mathcal{A}_i} \in \mathbb{R}^{|\mathcal{A}_i|}$ – that is real numbers representing a score each player assigns to each respective strategy

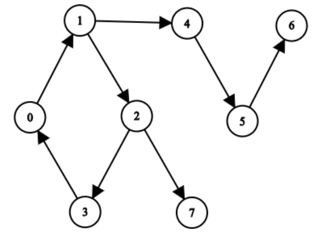

Figure 1: A connected graph where each vertex is of indegree at most one.

– by the integral equation

$$\begin{aligned} y_i(t) &= y_i(0) + \int_0^t v_i(x(s))ds, \\ x_i(t) &= Q_i(y_i(t)), \end{aligned} \tag{4}$$

where the the *choice map* $Q = (Q_1, \dots, Q_N)$, $Q_i : \mathbb{R}^{|\mathcal{A}_i|} \to \mathcal{X}_i$, which determines the evaluated strategy profile $x(t)$ is given on each coordinate by:

$$Q_i(y_i) = \text{argmax}_{x_i \in \mathcal{X}_i} \{ < y_i, x_i > - h_i(x_i) \}, \tag{5}$$

In the above $h_i : \mathcal{X}_i \to \mathbb{R} \cup \{-\infty, \infty\}$ is a convex regularizer function, representing a regularization/exploration term. The equation (4) represents the population adaptation to the perceived evolution of utility values for each respective strategy of each player.

In binary games, each agent has only two strategies at his disposal, say α_0, α_1. The variable x_i denotes then the proportion of time player i plays strategy α_0, and the proportion of α_1 is given by $1 - x_i$. Following [25], we introduce new variables $z_i \in \mathbb{R}$

$$z_i := y_{i\alpha_1} - y_{i\alpha_0}, \tag{6}$$

representing the difference in utilities between playing strategy α_1 and α_0. It was proved in [25] that $Q_i(z_i + c, c)$ is constant in c. Therefore, without loss of generality, we can choose $c := 0$ and restrict our considerations to a z-dependent choice map $\hat{Q}_i(z_i) := Q_i(z_i, 0)$. Provided that Q is sufficiently regular (e.g. continuous), the integral equation (4) can be converted to a system of differential equations

$$\dot{z} = V(z) \tag{7}$$

given coordinate-wise by

$$V_i(z) := v_{i\alpha_1}(\hat{Q}(z)) - v_{i\alpha_0}(\hat{Q}(z)), \tag{8}$$

for details again see [25].

Remark 2. *An intuitively obvious, but technically important observation is that evolution of ith coordinates of the system* (4)*, and, in turn* (8) *depends solely on the values of x_j/z_j, respectively, for nodes j that influence the payoffs of i. In particular, for GPGs we have $\partial V_i/\partial z_j \neq 0$ implies that there is an edge from j to i in the game graph; and for GPGs with up to one predecessor, without loss of generality we can rewrite* (7) *as*

$$\dot{z}_i = V_i(z_{\hat{i}}) = v_{i\alpha_1}(\hat{Q}(z_{\hat{i}})) - v_{i\alpha_0}(\hat{Q}(z_{\hat{i}})) \tag{9}$$

Perhaps the best known example of Follow-the-regularized-leader learning system are the *replicator equations*, where

$$h_i(x_i) := \sum_{\alpha_i} x_{i\alpha_i} \log x_{i\alpha_i}, \tag{10}$$

which yields the following equations for a binary GPG with up to one predecessor:

$$\dot{z}_i = \left(A^{\hat{i},i}(\alpha_1,\alpha_0) - A^{\hat{i},i}(\alpha_1,\alpha_1) - A^{\hat{i},i}(\alpha_0,\alpha_0) + A^{\hat{i},i}(\alpha_0,\alpha_1)\right) \frac{\exp(z_{\hat{i}})}{1+\exp(z_{\hat{i}})} + A^{\hat{i},i}(\alpha_1,\alpha_1) - A^{\hat{i},i}(\alpha_0,\alpha_1). \tag{11}$$

Firstly, we prove the following lemma on monotonicity and smoothness of the choice map, when a player has exactly two strategies at disposal (i.e. for $\mathcal{X}_i = [0,1]$), which will be used later in our other proofs.

Lemma 2. *Assume that the regularizer h_i satisfies the following conditions:*

(1) $h_i \in C^2((0,1)) \cap C^0([0,1])$ (smoothness),
(2) $h_i'(x) \to -\infty$ *as* $x \to 0$ *and* $h_i'(x) \to \infty$ *as* $x \to 1$ (steepness),
(3) $h_i''(x) > 0$ *for* $x \in (0,1)$ (strict convexity).

Then $Q_i \in C^1(\mathbb{R})$ *and* $Q_i'(z_i) > 0$.

Proof. For a given z_i, $\hat{Q}_i(z_i)$ is defined as the maximizer of $\langle (z_i,0),(x_i,1-x_i)\rangle - h_i(x_i)$ over $x_i \in [0,1]$. We have

$$\langle (z_i,0),(x_i,1-x_i)\rangle - h_i(x_i) = z_i x_i - h_i(x_i). \tag{12}$$

From steepness, continuity and strict convexity it follows that $h_i(0) = h_i(1) = \infty$ so the maximum cannot be attained there. A necessary condition for maximum to be attained within $(0,1)$ is

$$z_i = h_i'(x_i) \tag{13}$$

From steepness and strict convexivity it follows that equation (13) has a unique solution $x_i =: \hat{Q}_i(z_i)$ for any $z_i \in \mathbb{R}$. From the inverse function theorem we have

$$\frac{\partial x_i}{\partial z_i} = \hat{Q}_i'(z_i) = 1/h_i''(x_i) > 0 \tag{14}$$

which also implies that Q_i is C^1. □

2.4 Limit sets, periodic orbits and chaos

A differential equation $\dot{x} = F(x)$ given by a C^1 vector field $F : \Omega \to \mathbb{R}^n$ on a domain $\Omega \subset \mathbb{R}^n$ admits a unique solution on a maximal open interval $I = (I_l, I_r)$ $x(t) : I \to \mathbb{R}$ for any initial condition $x(0) = x_0 \in \Omega$. Among possible solutions to such equation, we distinguish particular types of solutions due to their qualitative properties: we say that a solution $x(t)$ is an *equilibrium* iff $x(t) =$ const for all $t \in I$. A solution is *periodic* iff $x(t) = x(t+T)$ for some $T > 0$ and all $t \in I$; and it is a connecting orbit between two equilibria (constant solutions) x_1 and x_2, iff $x(t) \to x_1$ as $t \to \infty$ and $x(t) \to x_2$ as $t \to -\infty$. If $x_1 = x_2$, such an orbit is called a *homoclinic orbit*, otherwise it is a *heteroclinic orbit*.

A set $\omega(x_0) \subset \Omega$ is a limit set for an initial condition $x_0 \in \Omega$, if $\forall x \in \omega(x_0)$ there exists a sequence $\{t_n\}_n \subset \mathbb{R}^+$, such that $x(t_n) \to x$, $n \to \infty$. Limit sets are *invariant* – they are formed by unions of solutions of the differential equation. They are also *compact* – bounded as subsets of $\mathbb{R}^n$, and closed under the limit operation for sequences from itself.

Fundamental research has been devoted to study the properties of solutions within limit sets, as they offer a qualitative description of long-term behavior of the system [16]. Since the discovery of chaotic attractors [22], it has become known that in the general setting, these solutions can have arbitrarily complicated shapes and exhibit seemingly random behavior, a clearly undesirable feature from the point of view of applications; and engineering systems with simple ω-limit sets became of particular interest.

Definition 1. *We say that a differential equation $\dot{x} = F(x)$, $x \in \Omega$ has the* Poincaré-Bendixson property *iff for all $x \in \Omega$ such that the solution $x(t)$ is bounded, each limit set $\omega(x)$ is either:*

- *an equilibrium;*
- *a periodic solution;*
- *a union of equilibria and connecting orbits between these equilibria.*

A well known result from the qualitative theory of differential equations shows that planar systems exhibit this trait.

Theorem 1. *(The Poincaré-Bendixson Theorem [4]) Let $F = F(x)$, $x \in \Omega \subset \mathbb{R}^2$ be a C^1 vector field with finitely many zeroes. Then, the differential equation $\dot{x} = F(x)$ has the Poincaré-Bendixson property.*

Already in $\mathbb{R}^3$ there are known examples of systems having complicated, chaotic attractors [22]. However, dimensionality is not the only factor which could determine potential shapes of limit sets. In particular, for certain systems of arbitrary dimension, with structured "previous-neighbor" interactions between the variables, the limit sets are as simple as in planar systems.

Theorem 2. *(Mallet-Paret & Smith [23])*
Let $(f_i(x_{i-1}, x_i))_{i=1}^n$, be a C^1 vector field on an open convex set $O \subset \mathbb{R}^n$, and let $x^0 := x^n$. Assume that $\delta^i \frac{\partial f^i}{\partial x^{i-1}} > 0$ for all $x \in O$, where $\delta^i \in \{-1,1\}$. Then, the system of differential equations

$$\dot{x}_i = f_i(x_{i-1}, x_i), \; i = 1, \ldots, n, \tag{15}$$

has the Poincaré-Bendixson property.

The above theorem is key to proving our further results above GPG games with one predecessor. We will refer to systems satisfying assumptions of the above theorem as *monotone cyclic feedback systems.*

3 THE POINCARÉ-BENDIXSON THEOREM FOR GAMES

In this section we will show that Follow-the-regularized-Leader systems of generic binary, cyclic games satisfy the Poincaré-Bendixson property. We will first state and prove the Poincaré-Bendixson theorem for cyclic games:

Theorem 3. *Let $\dot{z} = V(z)$ be a system of differential equations given by the vector field* (8) *the follow-the-regularized-leader learning dynamics of a binary, cyclic game. For any smooth, steep, strictly convex collection of regularizers $\{h_i\}_i$ and almost all values of payoffs – that is outside of a set of measure zero – such system possesses the Poincaré-Bendixson property.*

Proof. Since u_i depends only on Q_i and Q_{i-1}, we have

$$\begin{aligned} V_i(\hat{Q}(z)) &= V_i(\hat{Q}_{i-1}(z_{i-1})) \\ &= v_{i\alpha_1}(Q_{i-1}(z_{i-1}, 0)) - v_{i\alpha_0}(Q_{i-1}(z_{i-1}, 0)). \end{aligned} \tag{16}$$

Our goal is to employ Theorem 2 and show that the vector field V induces a monotone cyclic feedback system. Therefore, we would like to establish under which conditions

$$\delta_i \frac{\partial V_i}{\partial z_{i-1}} > 0. \tag{17}$$

for all i and any combination of $\delta_i \in \{-1, 1\}$. We have:

$$\frac{\partial V_i}{\partial z_{i-1}} = \frac{\partial v_{i\alpha_1}}{\partial x_{i-1}} \frac{\partial x_{i-1}}{\partial z_{i-1}} - \frac{\partial v_{i\alpha_0}}{\partial x_{i-1}} \frac{\partial x_{i-1}}{\partial z_{i-1}}. \tag{18}$$

Moreover,

$$\frac{\partial v_{i\alpha_1}}{\partial x_{i-1}} = A^{\hat{i},i}(\alpha_0, \alpha_1) - A^{\hat{i},i}(\alpha_1, \alpha_1), \tag{19}$$

and

$$\frac{\partial v_{i\alpha_0}}{\partial x_{i-1}} = A^{\hat{i},i}(\alpha_0, \alpha_0) - A^{\hat{i},i}(\alpha_1, \alpha_0). \tag{20}$$

Under assumptions from Lemma 2 we have $\frac{\partial x_{i-1}}{\partial z_{i-1}} > 0$, so the necessary condition to satisfy inequality (17) is:

$$\begin{aligned} &A^{\hat{i},i}(\alpha_0, \alpha_1) + A^{\hat{i},i}(\alpha_1, \alpha_0) \\ &\neq A^{\hat{i},i}(\alpha_0, \alpha_0) + A^{\hat{i},i}(\alpha_1, \alpha_1), \end{aligned} \tag{21}$$

which is generically satisfied for such normal form games.

□

Lemma 3. *Consider the following y-augmented system of differential equations*

$$\begin{aligned} &\dot{x} = f(x), \\ &\dot{y} = g(x_i), \\ &x = \{x_1, \ldots, x_n\} \in \mathbb{R}^n, \ y \in \mathbb{R}. \end{aligned} \tag{22}$$

for smooth f, g. If the original system

$$\dot{x} = f(x) \tag{23}$$

has the Poincaré-Bendixson property, then the augmented system 22 also has the Poincaré-Bendixson property.

Proof. Let Z be an ω-limit set corresponding to some solution $(x(t), y(t))$ to the system (22). Consider X – an ω-limit set to solution $x(t)$ of (23). From invariance of ω-limit sets it follows set Z consists of a union of solutions of (22). For any solution $\{x^*(t), y^*(t) : t \in \mathbb{R}\} \subset Z$, we have $\{x^*(t)\} \subset X$. By the Poincaré-Bendixson property of the original system, we can distinguish three cases:

(1) $x^*(t)$ is an equilibrium of (23),
(2) $x^*(t)$ is a periodic orbit of (23),
(3) $x^*(t)$ is a connecting orbit of (23) – a part of a cycle of connecting orbits.

In the rest of the proof we will frequently use the integral form of solutions $y(t)$ to (22), given by $y(t) = y(0) + \int_0^t g(x_i(s))ds$.

Case (1): We will prove that $(x^*(t), y^*(t))$ is stationary for (22). It is enough to show $g(x_i^*) = 0$. Assume otherwise. Then $y^*(t) = y(0) + \int_0^t g(x_i^*)ds = y(0) + tg(x_i^*) \to \infty$ as $t \to \infty$. This contradicts the boundedness of an ω-limit set.

Case (2) Let T be the period of $x^*(t)$. We will show that $(x^*(t), y^*(t))$ is a periodic solution of (22) of the same period. We have

$$\begin{aligned} \frac{d}{dt}(y^*(t+T) - y^*(t)) &= \frac{d}{dt} \int_t^{T+t} g(x_i^*(s))ds \\ &= g(x_i^*(T+t)) - g(x_i^*(t)) \\ &= 0, \end{aligned} \tag{24}$$

hence $y^*(t+T) - y^*(t) = const$. If this quantity would be non-zero, the diameter of the set $\{y^*(t) : t \in \mathbb{R}\}$ would be infinite. However, the set Z is bounded, and therefore $y^*(t+T) = y^*(t)$.

Case (3): We will show that $(x^*(t), y^*(t))$ is a connecting orbit between two equilibria for the full system (22). We shall only prove convergence with $t \to \infty$, the very same argument holds for $t \to -\infty$ and α-limit sets. The orbit $(x^*(t), y^*(t))$ is bounded and therefore it has an accumulation point as $t \to \infty$ given by $(x^{**}, y^{**}) \in \omega(x^*(0), y^*(0)$. The point x^{**} is an equilibrium for (23). We will show that (x^{**}, y^{**}) is an equilibrium. It is enough to show that $g(y^{**}) = 0$. Assume otherwise. Then $y^{**}(t) = y^{**} + tg(x_i^{**})$ which is unbounded. However, it is also a part of the $\omega(x^*(0), y^*(0)$, since ω-limit sets are invariant. Boundedness of $\omega(x^*(0), y^*(0)$ leads to a contradiction. The same process, repeated for all connecting orbits of (23), creates a cycle of connecting orbits for (22). □

Theorem 4. *Let $\dot{z} = V(z)$ be a system of differential equations given by the follow-the-regularized leader dynamics of a binary, connected, graphical polymatrix game, where each player has up to one predecessor. Then, for any smooth, steep, strictly convex collection of regularizers $\{h_i\}_i$ and almost all values of payoffs – that is outside of a set of measure zero – such system possesses the Poincaré-Bendixson property.*

Proof. By Lemma 1 and Remark 1 we know that the graph of the system has either a root vertex or a root cycle. We will first address the case of a root vertex. We will see that this case is somewhat degenerate. Without loss of generality let us assume that it is labelled as the 1st vertex, and that the other vertices are numbered in order of increasing path distance from vertex 1 (i.e. $j < i$ implies that the path from 1 to j is shorter than the path from 1 to i) – this is possible by Lemma 1.

The payoffs of the root node are only affected by its own choice of strategy. Therefore $\dot{z}_1 = u_1(\alpha_1) - u_1(\alpha_0)$, and therefore $z_1(t) = t(u_1(\alpha_1) - u_1(\alpha_0)) + z_1(0)$. This system constitutes an autonomous ODE, which trivially has the Poincaré-Bendixson property (as it is either all zeroes, or divergent). By adding the vertex 2, we again obtain an autonomous system, and again it is either stationary or divergent; and in the same manner the proof continues for all vertices. It should be noted that "divergence" in practice means that $z_i(t)$'s approach in the limit $t \to \infty$ to either ∞ or $infty$; the former implies that the player i is placing almost all probability mass on strategy α_1, and the latter – on α_0.

The more interesting scenario arises for the root cycle, where periodic limit sets are possible. Enumerate these vertices by $1, \ldots, N_0$, with $N_0 <= N$, and assume that the vertices from $N_0 + 1$ to N are arranged in the order of increasing path distance from vertices of the cycle (possible by Remark 1). Observe that the system

$$\begin{aligned} &\dot{z}_i = V_i(z_{\hat{i}}), \\ &i = 1, \ldots, N_0, \end{aligned} \tag{25}$$

is an autonomous system of differential equations (as there are no edges with successors in $\{1, \ldots, N_0\}$, and predecessors outside of this set), and forms a binary, cyclic game in the sense of Theorem 3. As such, this subsystem possesses the Poincaré-Bendixson property. From then on, the proof continues similarly as for the roof vertex – in an inductive way. We add the vertex $N_0 + 1$, and observe that the system

$$\begin{aligned} \dot{z}_i &= V_i(z_{\hat{i}}), \\ i &= 1, \ldots, N_0 + 1, \end{aligned} \tag{26}$$

is again autonomous, i.e. there are no edges with predecessors in $\{N_0 + 1, \ldots, N\}$ and successors in $\{1, \ldots, N_0 + 1\}$. By Lemma 2, this system then also possesses the Poincaré-Bendixson property. The proof continues inductively w.r.to the vertices, until we conclude that the full system $\dot{z} = V(z)$ has the Poincaré-Bendixson property. □

Remark 3. *Our theorems apply only to fully mixed initial strategy profiles, as the differences of utilities corresponding to pure strategies are infinite, and (FoReL) learning, as in Equation* (4) *is formally not defined in such situation. However, when one player has an initial pure strategy, the system can be suitably decomposed, and the Poincaré-Bendixson property still holds. More specifically, in a game where each agent has one predecessor, if agent i plays a pure strategy, then all the agents*

$$V_i := \{j : \text{ there exists a path from } i \text{ to } j\} \tag{27}$$

would eventually sequentially converge under all reasonable learning dynamics (including replicator) to their best response to strategy i. One can then apply Theorem 4 to the autonomous reduced system $V \backslash V_i$, where Equations (4) *are again well defined.*

Remark 4. *The assumption of connectedness is needed for the Poincaré-Bendixson property, as (by Lemma 1) it ensures that the graph of interactions has only one cycle. For games with multiple cycles, one can have yet another type of limit behavior. Consider a disjoint union of FoReL systems for two binary graphical games, both possessing the Poincaré-Bendixson property, such that the systems have non-resonant periods of periodic orbits; e.g. one of the systems has a periodic solution of period* 1 *and the other system has a periodic orbit of period $\sqrt{2}$. Such orbits can be easily obtained from replicator dynamics for appropriately scaled mismatched pennies games, c.f. Section 5. Let $(z_1(0), z_2(0))$ be a point belonging to the periodic solution of period* 1*, and $(z_2(0), z_3(0))$ be a point belonging to the periodic solution of period $\sqrt{2}$. Then the solution of the full system starting from $(z_1(0), z_2(0), z_3(0), z_4(0))$ forms a quasi-periodic motion, with ω-limit set of toroidal shape, see Figure 2, c.f. [5].*

4 FROM GEOMETRY TO EFFICIENCY: SOCIAL WELFARE ANALYSIS

In this section we will show that in the case of cyclic, binary games, under additional but structurally robust assumptions on the payoff matrices (i.e., assumptions that remain valid after small perturbations of the payoff matrices) the time-average social welfare of our FoReL dynamics is at least as large as the social welfare of the worst Nash equilibrium. As is typically the case the social welfare is defined as the sum of individual payoffs $SW = \sum_i u_i$.

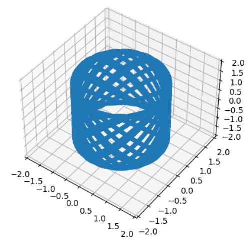

Figure 2: An invariant torus in a 4-dimensional dynamical system – a projection onto first three variables.

Theorem 5. *In any binary, cyclic game with the property that for any agent $k+1$, we have that the payoff entries are distinct and $[A^{k,k+1}(1,1) - A^{k,k+1}(2,1)][A^{k,k+1}(1,2) - A^{k,k+1}(2,2)] < 0$ then the time-average of the social welfare of FoReL dynamics is at least that of the social welfare of the worst Nash equilibrium. Formally,*

$$\liminf \frac{1}{T} \int_0^T \sum_k u_k(x(t))dt \geq \sum_k u_k(x_{NE})$$

where x_{NE} the worst case Nash equilibrium, i.e., a Nash equilibrium that minimizes the sum of utilities of all agents.

Proof. Lets consider the payoff matrix of each agent $k+1$. By assumption there is at most one agent i such that $A^{i,k}$ is a non-zero matrix, i.e., the unique predecessor of $k+1$, that for simplicity of notation we call k. By assumption the four entries will be considered distinct. Next, we break down the analysis into two cases:

As a first case, we consider the scenario where there exists at least one agent with a strictly dominant strategy. The FoReL dynamics of that agents will trivially converge to playing the strictly dominant strategy with probabilty one. Similarly, all agents reachable from agent k will similarly best response to it. This is clearly the unique NE for the binary cyclic game, so in this case the limit behavior of FTRL dynamics exactly corresponds to the unique Nash behavior and the theorem follows immediately.

Next, let's consider the case where no agent has a strictly dominant strategy. In this case, we will construct a specific Nash equilibrium for the cyclic game (although it may have more than one). In this Nash equilibrium every agent k plays the unique mixed strategy that makes its successor (agent $k+1$) indifferent between its two strategies. Such a strategy exists for each agent, because otherwise there would exist an agent with a strictly dominant strategy. In fact by the assumption $[A^{k,k+1}(1,1) - A^{k,k+1}(2,1)][A^{k,k+1}(1,2) - A^{k,k+1}(2,2)] < 0$ such a strategy would be the k agent's min-max strategy if they participated in a zero-sum game with agent $k+1$ defined by the payoff matrix of agent $k+1$. Indeed, this assumption along with the fact that agent $k+1$ does not have a dominant strategy exactly encodes the zero-sum game (defined by payoff matrix $A^{k,k+1}$) has an interior Nash. Given her predecessors behavior, agent $k+1$ will be receiving exactly her max-min payoff no matter which strategy they select, therefore this strategy profile where each agent k just plays the strategy that makes agent $k+1$ indifferent between their two options is a Nash equilibrium, where each agent receives exactly their max-min payoffs. However, by

[25], continuous-time FoReL dynamics are no-regret with their time-average regret converging to zero at an optimal rate of O(1/T). The sum of the time-average performance is thus at least the sum of the maxmin utilities minus a quickly vanishing term O(1/T) and the theorem follows. □

5 EXAMPLES

To illustrate our theoretical findings, we present the dynamical structure of two multidimensional binary, cyclic games, which exhibit non-convergence, and therefore non-trivial limit behavior. To determine the limit sets, we perform numerical integration of initial value problems for various starting conditions.

5.1 Matched-mismatched pennies game

Firstly, we analyze a four dimensional system of matched-mismatched pennies. Each player has a choice of two strategies, α_0 and α_1. The payoffs for players 0, 2 are given by matrices

$$A^{3,0} = A^{1,2} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \tag{28}$$

and the payoffs of players 1, 3 are given by

$$A^{0,1} = A^{2,3} = \begin{bmatrix} 1 & -1 \\ -1 & 1. \end{bmatrix} \tag{29}$$

where the rows correspond to the choice of $i-1$-st (previous) player strategy and columns corespond to the choice of i-th player strategy.

Simply put, players 0 and 2 will try to mismatch the strategy with players 1 and 3, and players 1 and 3 will try to match them.

The induced system of replicator equations is given by:

$$\begin{aligned} \dot{z}_i =& A^{\hat{i},i}(\alpha_1,\alpha_0)Q_{\hat{i}})(z) + A^{\hat{i},i}(\alpha_1,\alpha_1)(1-Q_{\hat{i}}(z)) \\ &- A^{\hat{i},i}(\alpha_0,\alpha_0)Q_{\hat{i}}(z) - A^{\hat{i},i}(\alpha_0,\alpha_1)(1-Q_{\hat{i}}(z)), \end{aligned} \tag{30}$$

with $Q_i(z) = x_i = \exp(z_i)/(1+\exp(z_i)) \in (0,1)$, and we recall that $\hat{i} \mod 4 = i-1 \mod 4$. To visualise the dynamics near boundary, we return to x coordinates, compactifying the state space in the following manner:

$$Q_i(z) = x_i \in 0,1 \Rightarrow z_i = 0. \tag{31}$$

This yields the original replicator dynamics, c.f. [25]), where $x = (x_i)_i$ denotes the frequencies with which the players play strategy α_0.

The system possesses three Nash equilibria, corresponding to the following strategy profiles: $(0,0,1,1)$, $(1,1,0,0)$, $(0.5,0.5,0.5,0.5)$, out of which the pure Nash equilibria are attracting, and the mixed Nash equilibrium has two center directions, one repelling direction, and one attracting direction. We will denote the mixed Nash by x^{MNE}.

Despite the nonlinear nature of the system, the dynamical situation in viccinity of x^{MNE} can be described exactly. Due to symmetry of the system, the plane

$$W^c(x^{MNE}) = \{(t,s,t,s),\ t,s \in [0,1]\} \tag{32}$$

is invariant, and consists purely of periodic orbits, which constitute a two-dimensional center manifold to the mixed Nash equilibrium. The parametrizations of stable/unstable manifold for the mixed Nash equilibrium are given by

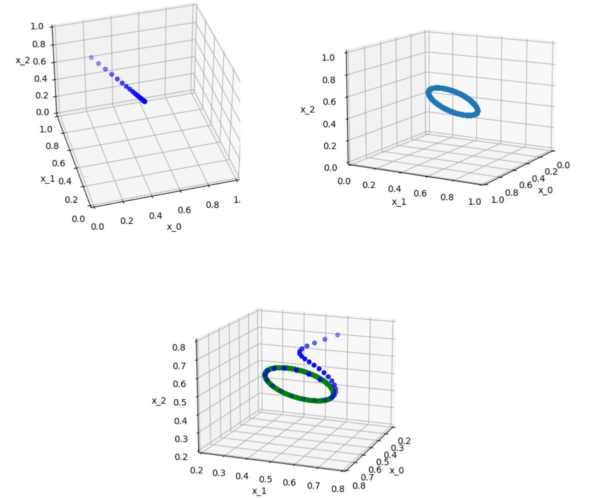

Figure 3: Limit sets in the matched-mismatched pennies system: a projection onto the first three variables. Top left: convergence towards mixed equilibrium along the one dimensional stable manifold, top right: a periodic orbit on the center manifold of the fixed equilibrium. Bottom: An orbit converging to a limit cycle.

$$\begin{aligned} W^s(x^{MNE}) &= \{(1-t,t,t,1-t), t \in [0,1]\}, \\ W^u(x^{MNE}) &= \{(t,t,1-t,1-t), t \in [0,1]\}, \end{aligned} \tag{33}$$

respectively.

The numerical results are in line with Theorems 3, 4. Indeed, the interior Nash equilibrium is an ω-limit set for its own stable manifold. Each periodic orbit is in particular an ω-limit set of any point lying on that orbit. From numerical simulations it appears that the periodic orbits are true limit cycles, in the sense that they themselves possess a stable manifold and an unstable manifold, and therefore are ω-limit sets for a collection of points outside of themselves, see Figure 3. Most crucially, more complicated behavior like chaos or invariant tori does not emerge, despite the system being nontrivially embedded in four dimensions.

The mixed Nash yields the minimax payoff vector $(0,0,0,0)$ for each player, and eventually social welfare of 0. The payoff matrices satisfy the assumptions of Theorem 5, and the average payoffs along solutions are therefore at least non-negative. In fact, almost all (as a set of full measure) initial conditions appear to converge to the pure equilibria at the boundary, and their time-average payoffs exceed the one of Nash equilibrium and converge to the maximal welfare of 4, see Figure 4.

5.2 Asymmetric N-pennies game

Our second system is a system of N-player asymmetric mismatched pennies, previously introduced in [20]. There are three players, and again each of them can choose between two strategies, α_0 and α_1. The payoffs for players i w.r.to the player $i-1$ are given by the

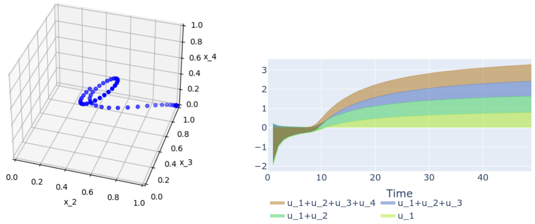

Figure 4: Time-average payoffs and social welfare of a sample orbit in the matched-mismatched pennies game.

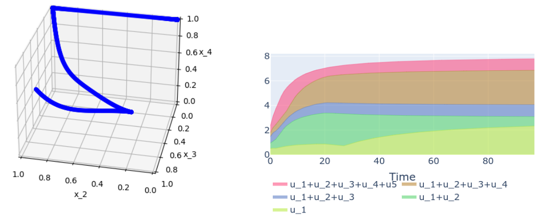

Figure 5: Left: projection of an orbit of the asymmetric mismatched pennies system for $N = 5$ and $p = 3$: onto first three variables. Note that the integration time approaches infinity as orbits approach the corners of the heteroclinic cycle, hence the simulated orbit appears to converge to one of the corners. Right: time-averaged payoffs and social welfare corresponding to this orbit.

matrix

$$A^{\hat{i},i} = \begin{bmatrix} 0 & 1 \\ p & 0 \end{bmatrix} \tag{34}$$

for $\hat{i} \mod N = i - 1 \mod N$, $p > 0$. The induced system of replicator equations is given by:

$$\dot{z}_i = Q_{\hat{i}}(z) - p(1 - Q_{\hat{i}}(z)), \tag{35}$$

with $x_i = Q_i(x)$ given by $\exp(z_i)/(1+\exp(z_i))$ as before. We extend the phase space by $x_i = 0, 1$, in the same manner as in our previous example.

For odd N there is no Nash equilibrium in pure strategies; instead best response dynamics in pure strategies eventually converge to cyclic behavior formed by mixtures of strategies α_0 and α_1 For the replicator system, the pure strategy profiles are saddle-type stationary points of the ODE, connected by heteroclinic orbits of mixed strategies. The system has a unique, mixed Nash equilibrium defined by $x_i = \frac{1}{p+1}$, $i \in \{1, \ldots, N\}$, where each player attains a payoff of $\frac{p}{p+1}$. Due to nonlinear nature, this time it is difficult to give exact formulas for its stable and unstable manifolds, however, linear stability analysis for various values of p, N (e.g. $p = 3$, $N = 5$) shows that the equilibrium is saddle-focus, i.e. have one attracting direction corresponding to a real negative eigenvalue of the Jacobian (corresponding to the diagonal direction), and multiple complex eigenvalues with non-zero real parts – some positive, yielding unstable directions.

The system was thoroughly analyzed in [20] and the main result provided therein is that for $N = 3$ and $p > 7$ all mixed strategies except for the diagonal converge to a sequence of heteroclinic orbits connecting boundary stationary points, Moreover, the social welfare attained close to the boundary exceeds the social welfare at the Nash equilibrium. We extend these results. From Theorem 3 we deduce that for all N, and for all $p \neq -1$, the only limit sets in the system are equilibria, periodic orbits, and cycles of connecting orbits to equilibria. The payoff matrices satisfy the assumptions of Theorem 5, and in particular for all $p > 0$, the mixed equilibrium yields the minimax payoff for each player, and time averages of payoffs along other orbits have to exceed the minimax payoffs. By numerical integration we observe that for almost all initial conditions the dynamics is attracted to the boundary cycle of average payoff $(p + 1)\frac{N-1}{2}$, see Figure 5, and indeed no chaotic emergent behavior is apparent.

6 DISCUSSION

Numerous recent results in learning in games have established a clear separation between the idealized behavior of equilibration and the erratic, unpredictable and typically chaotic behavior of learning dynamics even in simple games and domains [1–3, 8, 11, 13, 20, 27, 29, 35]. Although at a first glance, this realization might seem as a set-back, when viewed from the right perspective it opens up a new possibility, a new way of understanding learning dynamics in games that does not focus primarily on the vocabulary of solution concepts of game theory with its numerous notions of equilibration, but instead examines solution concepts from the topology of dynamical systems that are more native to the nature of game dynamics. Our results showcase the possibility of establishing links between the combinatorial structure of multi-agent games (e.g. game graph, number of actions) to understand and constrain the topological complexity of game dynamics (Poincaré-Bendixson behavior) and finally establish links back to more traditional game theoretic analysis such as understanding the social welfare, efficiency of the system. These connections showcase promising advantages of this approach and we hope that we encourage more work along these lines.

REFERENCES

[1] James P. Bailey and Georgios Piliouras. 2018. Multiplicative Weights Update in Zero-Sum Games. In *Proceedings of the 2018 ACM Conference on Economics and Computation, Ithaca, NY, USA, June 18-22, 2018.* 321–338. https://doi.org/10.1145/3219166.3219235

[2] James P. Bailey and Georgios Piliouras. 2019. Multi-Agent Learning in Network Zero-Sum Games is a Hamiltonian System. In *18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS).*

[3] Michel Benaïm, Josef Hofbauer, and Sylvain Sorin. 2012. Perturbations of set-valued dynamical systems, with applications to game theory. *Dynamic Games and Applications* 2, 2 (2012), 195–205.

[4] Ivar Bendixson. 1901. Sur les courbes définies par des équations différentielles. *Acta Mathematica* 24, 1 (1901), 1.

[5] Hendrik W Broer, George B Huitema, and Mikhail B Sevryuk. 2009. *Quasi-periodic motions in families of dynamical systems: order amidst chaos.* Springer.

[6] Nicolo Cesa-Bianchi and Gábor Lugosi. 2006. *Prediction, learning, and games.* Cambridge university press.

[7] Yun Kuen Cheung and Georgios Piliouras. 2020. Chaos, Extremism and Optimism: Volume Analysis of Learning in Games. *arXiv e-prints*, Article arXiv:2005.13996 (May 2020), arXiv:2005.13996 pages. arXiv:2005.13996 [cs.GT]

[8] Yun Kuen Cheung and Georgios Piliouras. 2019. Vortices Instead of Equilibria in MinMax Optimization: Chaos and Butterfly Effects of Online Learning in Zero-Sum Games. In *COLT*.

[9] Thiparat Chotibut, Fryderyk Falniowski, Michał Misiurewicz, and Georgios Piliouras. 2019. The route to chaos in routing games: When is Price of Anarchy too optimistic? *arXiv preprint arXiv:1906.02486* (2019).

[10] Thiparat Chotibut, Fryderyk Falniowski, Michał Misiurewicz, and Georgios Piliouras. 2020. Family of chaotic maps from game theory. *Dynamical Systems* (2020), 1–16.

[11] Constantinos Daskalakis and Ioannis Panageas. 2018. The Limit Points of (Optimistic) Gradient Descent in Min-Max Optimization. In *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, 3-8 December 2018, Montréal, Canada*. 9256–9266. http://papers.nips.cc/paper/8136-the-limit-points-of-optimistic-gradient-descent-in-min-max-optimization

[12] Lampros Flokas, Emmanouil-Vasileios Vlatakis-Gkaragkounis, and Georgios Piliouras. 2019. Poincar\'e Recurrence, Cycles and Spurious Equilibria in Gradient-Descent-Ascent for Non-Convex Non-Concave Zero-Sum Games. *arXiv preprint arXiv:1910.13010* (2019).

[13] Seth Frey and Robert L Goldstone. 2013. Cyclic game dynamics driven by iterated reasoning. *PloS one* 8, 2 (2013), e56416.

[14] Drew Fudenberg, Fudenberg Drew, David K Levine, and David K Levine. 1998. *The theory of learning in games*. Vol. 2. MIT press.

[15] Tobias Galla and J Doyne Farmer. 2013. Complex dynamics in learning complicated games. *Proceedings of the National Academy of Sciences* 110, 4 (2013), 1232–1236.

[16] J.K. Hale. 2009. *Ordinary Differential Equations*. Dover Publications.

[17] Joseph T Howson Jr. 1972. Equilibria of polymatrix games. *Management Science* 18, 5-part-1 (1972), 312–318.

[18] Yu Ilyashenko. 2002. Centennial history of Hilbert's 16th problem. *Bull. Amer. Math. Soc.* 39, 3 (2002), 301–354.

[19] Michael Kearns. 2007. Graphical games. *Algorithmic game theory* 3 (2007), 159–180.

[20] Robert D Kleinberg, Katrina Ligett, Georgios Piliouras, and Éva Tardos. 2011. Beyond the Nash Equilibrium Barrier.. In *ICS*. 125–140.

[21] Elias Koutsoupias and Christos Papadimitriou. 1999. Worst-case equilibria. In *Annual Symposium on Theoretical Aspects of Computer Science*. Springer, 404–413.

[22] Edward N Lorenz. 1963. Deterministic nonperiodic flow. *Journal of the atmospheric sciences* 20, 2 (1963), 130–141.

[23] John Mallet-Paret and Hal L Smith. 1990. The Poincaré-Bendixson theorem for monotone cyclic feedback systems. *Journal of Dynamics and Differential Equations* 2, 4 (1990), 367–421.

[24] Brendan McMahan. 2011. Follow-the-regularized-leader and mirror descent: Equivalence theorems and l1 regularization. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. 525–533.

[25] Panayotis Mertikopoulos, Christos Papadimitriou, and Georgios Piliouras. 2018. Cycles in adversarial regularized learning. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 2703–2717.

[26] Sai Ganesh Nagarajan, David Balduzzi, and Georgios Piliouras. 2020. From Chaos to Order: Symmetry and Conservation Laws in Game Dynamics. In *ICML*.

[27] Gerasimos Palaiopanos, Ioannis Panageas, and Georgios Piliouras. 2017. Multiplicative Weights Update with Constant Step-Size in Congestion Games: Convergence, Limit Cycles and Chaos. In *The Thirty-first Annual Conference on Neural Information Processing Systems (NIPS)*.

[28] Marco Pangallo, James Sanders, Tobias Galla, and Doyne Farmer. 2017. A taxonomy of learning dynamics in 2 x 2 games. *arXiv preprint arXiv:1701.09043* (2017).

[29] G. Piliouras and J. S. Shamma. 2014. Optimization Despite Chaos: Convex Relaxations to Complex Limit Sets via Poincaré Recurrence. In *SODA*.

[30] Manfred Plank. 1997. Some qualitative differences between the replicator dynamics of two player and n player games. *Nonlinear Analysis: Theory, Methods & Applications* 30, 3 (1997), 1411–1417.

[31] Tim Roughgarden. 2009. Intrinsic robustness of the price of anarchy. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*. 513–522.

[32] Tim Roughgarden. 2016. *Twenty lectures on algorithmic game theory*. Cambridge University Press.

[33] James BT Sanders, J Doyne Farmer, and Tobias Galla. 2018. The prevalence of chaotic dynamics in games with many players. *Scientific reports* 8, 1 (2018), 1–13.

[34] William H Sandholm. 2010. *Population games and evolutionary dynamics*. MIT press.

[35] Yuzuru Sato, Eizo Akiyama, and J Doyne Farmer. 2002. Chaos in learning a simple two-person game. *Proceedings of the National Academy of Sciences* 99, 7 (2002), 4748–4751.

[36] Shai Shalev-Shwartz et al. 2011. Online learning and online convex optimization. *Foundations and trends in Machine Learning* 4, 2 (2011), 107–194.

[37] Yoav Shoham and Kevin Leyton-Brown. 2008. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press.

[38] Sebastian van Strien. 2011. Hamiltonian flows with random-walk behaviour originating from zero-sum games and fictitious play. *Nonlinearity* 24, 6 (2011), 1715. http://stacks.iop.org/0951-7715/24/i=6/a=002

[39] Elena B Yanovskaya. 1968. Equilibrium points in polymatrix games. *Litovskii Matematicheskii Sbornik* 8 (1968), 381–384.

[40] H Peyton Young. 2004. *Strategic learning and its limits*. Oxford Univ. Press.