Online Learning in Periodic Zero-Sum Games: von Neumann vs Poincaré

Tanner Fiez [1] Ryann Sim [2] Stratis Skoulakis [2]
Georgios Piliouras [2] Lillian Ratliff [1]

[1]University of Washington

[2]Singapore University of Technology and Design

Introduction

Classical behaviors of online learning dynamics in zero-sum games:

- **Poincaré Recurrence** & **Time-Average Nash Equilibrium Convergence**

Focus of this work:

- Evaluating the robustness of online learning behaviors to **periodically evolving games**

Motivation:

- Model of game evolution which can capture economic and artificial intelligence problems

Problem Overview

Classes of Games and Dynamics

- Continuous Strategy Spaces: GDA Dynamics in Periodic Zero-Sum Bilinear Games
- Finite Strategy Spaces: FTRL Dynamics in Periodic Zero-Sum Polymatrix Games

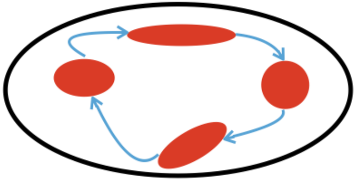
Question 1: Do the learning dynamics exhibit Poincaré recurrence?

Question 2: Do the time-average learning dynamics converge to a time-invariant Nash?

Background: Poincaré Recurrence

Poincaré Recurrence in Time-Invariant Systems

Definition: A dynamical system $\dot{x} = f(x)$ is called **Poincaré recurrent** if for almost all initial conditions $x(0)$, the dynamics return arbitrarily close to $x(0)$ infinitely often.



The accessible phase space is bounded & the trajectory cannot intersect itself except at the initial condition

Theorem[1] **(Informal):** If the flow of a system $\dot{x} = f(x)$ is volume preserving & has only bounded orbits, then it is Poincaré Recurrent.

[1] Ref: Poincaré, 1890.

Proof Method in Static Zero-Sum Games

Volume Preservation:

- Apply **Liouville's Theorem**: The flow of a dynamical system is volume preserving if and only if the divergence at any point equals zero.

Bounded Orbits:

- Find time-invariant function of the system state to imply orbits must be bounded.

Poincaré Recurrence:

- Apply Poincaré recurrence theorem using volume preservation & bounded orbits.

Proof method fails to generalize to time-dependent systems!

Poincaré Recurrence in Periodic Systems

The learning dynamics we study in periodic games form T-periodic systems

Definition: A system $\dot{x} = f(t, x)$ is T-perodic if $f(t + T, x) = f(t, x)$ for all (t, x).

Flow: The flow $\phi^t(x)$ of $\dot{x} = f(t, x)$ returns the system state at time t starting from x.

Poincaré map: The map ϕ^T is such that $\phi^{kT} = (\phi^T)^k$ for any integer k and induces a discrete-time (DT) dynamical system $x^+ = \phi^T(x)$.

Idea: Prove Poincaré recurrence by analyzing DT system induced by Poincaré map

Proof Method in Periodic Zero-Sum Games

Theorem[1] (Informal): If the DT system induced by the Poincaré map ϕ^T of a T-periodic system $\dot{x} = f(t, x)$ is volume preserving & has only bounded orbits, then $\dot{x} = f(t, x)$ is Poincaré recurrent.

Volume Preservation:[2]

- If the T-periodic system $\dot{x} = f(t, x)$ is divergence-free, then $x^+ = \phi^T(x)$ is volume preserving.

Bounded Orbits:

- Find time-invariant function to imply orbits of $x^+ = \phi^T(x)$ must be bounded.

Poincaré Recurrence:

- Apply Poincaré recurrence theorem using volume preservation & bounded orbits.

[1] Barreira, 2006; [2] Arnol'd, 2013.

Continuous Strategy Spaces: Periodic Zero-Sum Bilinear Games

Periodic Zero-Sum Bilinear Games & GDA Dynamics

Periodic Zero-Sum Bilinear Game:

- Player strategy spaces: $x_1 \in \mathbb{R}^{n_1}, x_2 \in \mathbb{R}^{n_2}$
- Time-dependent utility functions:

$$\text{Player 1:} \quad u_1(x_1, x_2, t) = x_1^\top A(t) x_2$$
$$\text{Player 2:} \quad u_2(x_1, x_2, t) = -x_2^\top A^\top(t) x_1$$

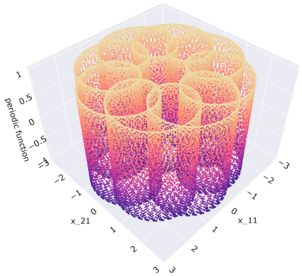
- Payoff matrix is periodic: $A(t) = A(t + T)$ for some finite period T
- $(x_1^*, x_2^*) = (0, 0)$ is a time-invariant Nash equilibrium of the game

Continuous-Time Gradient Descent Ascent (GDA) Dynamics:

$$\dot{x}_1 = A(t) x_2(t)$$
$$\dot{x}_2 = -A^\top(t) x_1(t)$$

Poincaré Recurrence

Theorem: The continuous-time gradient descent ascent learning dynamics are Poincaré recurrent in any periodic zero-sum bilinear game.

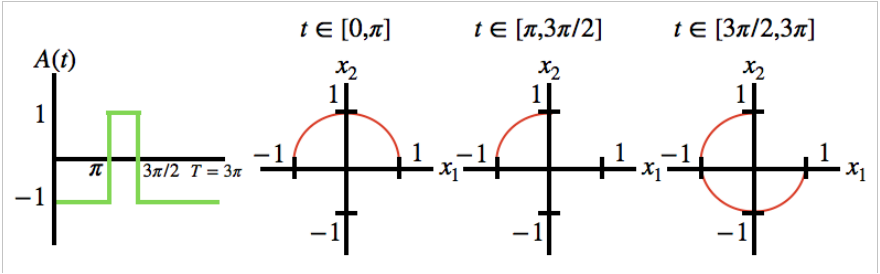


Proof Steps:

- **Volume Preservation:** Verify the vector field is divergence free
- **Bounded Orbits:** Show $\Phi(t) = \frac{1}{2}\left(x_1^\top(t)x_1(t) + x_2^\top(t)x_2(t)\right)$ is time-invariant
- **Poincaré Recurrence:** Apply Poincaré recurrence theorem

Time-Average Behavior

Theorem: There exists periodic zero-sum bilinear games where the time-average strategies of the continuous time gradient descent ascent dynamics fail to converge to the time-invariant equilibrium $(x_1^*, x_2^*) = 0$.



Finite Strategy Spaces:
Periodic Zero-Sum Polymatrix Games

Periodic Zero-Sum Polymatrix Games

A **Periodic Zero-Sum Polymatrix Game** is defined by a graph $G = (V, E)$:

- **Vertices** V represent **players** and **Edges** E represent T-**periodic bimatrix games** $(A^{ij}(t), A^{ji}(t)) = (A^{ij}(t+T), A^{ji}(t+T)) \in (\mathbb{R}^{n_i \times n_j}, \mathbb{R}^{n_j \times n_i})$ between players i and j
- Strategy space of any player $i \in V$ is $x_i \in \mathcal{X}_i = \Delta(\mathcal{A}_i)$ (simplex)
- Time-dependent utility functions: $u_i(x, t) = \sum_{j:(i,j)\in E} x_i^\top A^{ij}(t) x_j$

Assumption: There exists a time-invariant interior Nash equilibrium x^*

FTRL Dynamics

Follow-The-Regularized-Leader (FTRL) in Periodic Zero-Sum Polymatrix Games:

Cumulative Utility Space: $$y_i(t) = y_i(0) + \int_0^t \sum_{j:(i,j)\in E} A^{ij}(\tau) x_j(\tau) d\tau$$

Strategy Space: $$x_i(t) = \arg\max_{x_i \in \mathcal{X}_i} \{\langle x_i, y_i(t)\rangle - h_i(x_i)\}$$

Strongly convex regularization function $h_i : \mathcal{X}_i \to \mathbb{R}$ defines the dynamics:

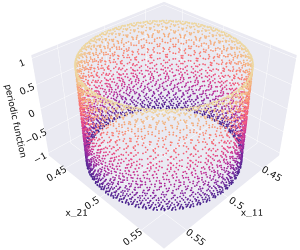
- $h_i(x_i) = \sum_{\alpha\in\mathcal{A}_i} x_{i\alpha} \log x_{i\alpha} \implies$ Replicator Dynamics
- $h_i(x_i) = \frac{1}{2}\|x_i\|_2^2 \implies$ Projected Gradient Descent

Poincaré Recurrence

Theorem: The Follow-The-Regularized-Leader learning dynamics are Poincaré recurrent in any periodic zero-sum polymatrix game.

Proof Steps

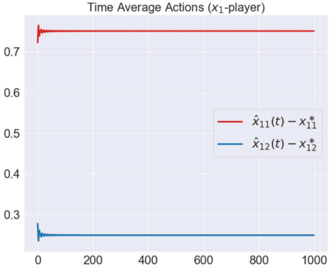
- **Transformation:** Strategies to utility differences
- **Volume Preservation:** Divergence free vector field
- **Invariant Function:** Fenchel coupling between $x^*, y(t)$
- **Poincaré Recurrence:** Apply Poincaré recurrence theorem



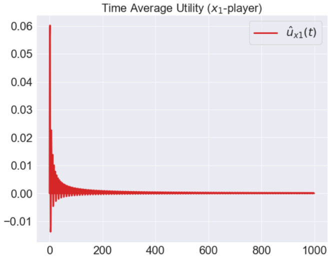
Time-Average Behavior

Theorem: There are periodic zero-sum bimatrix games where the time-average strategies of the FTRL dynamics fail to converge to the time-invariant equilibrium. However, the time-average utilities converge to the average game values over a period.

Example $\Longrightarrow$ Replicator in Periodic Matching Pennies: $A(t) = \sin(t)\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$



Time-Average Strategy



Time-Average Utility

Conclusion

Summary:

- **Poincaré recurrence** of learning dynamics persists with periodic game evolution
- **Time-Average Nash Equilibrium Convergence** fails in periodic zero-sum games

Practical Implications:

- This work showcases and proves unexpected failure modes that can happen if the underlying system parameters evolve over time.