Cycles in Zero-sum Differential Games and Biological Diversity

Ioannis Panageas (**CSAIL-MIT**)

joint work with **Tung Mai (Gatech), Milena Mihail (Gatech), Will Ratcliff (Gatech), Vijay Vazirani (UC Irvine), Peter Yunker (Gatech).**

Long-term behavior in Dynamics

Given a continuous or discrete time (learning) dynamics we are interested in its long-term behavior...

Long-term behavior in Dynamics

Given a continuous or discrete time (learning) dynamics we are interested in its long-term behavior...

- Convergence of time average ([FS99'], ...).

Long-term behavior in Dynamics

Given a continuous or discrete time (learning) dynamics we are interested in its long-term behavior...

- Convergence of time average ([FS99'], ...).
- Last iterate convergence ([PPP17'], [DP18']).

Long-term behavior in Dynamics

Given a continuous or discrete time (learning) dynamics we are interested in its long-term behavior...

- Convergence of time average ([FS99'], ...).
- Last iterate convergence ([PPP17'], [DP18']).
- Cycling behavior ([PPP17']) or *recurrent* behavior ([PS14'], [PNCS14'], [MPP18'], [BP18']).

Long-term behavior in Dynamics

Given a continuous or discrete time (learning) dynamics we are interested in its long-term behavior...

- Convergence of time average ([FS99'], ...).
- Last iterate convergence ([PPP17'], [DP18']).
- Cycling behavior ([PPP17']) or *recurrent* behavior ([PS14'], [PNCS14'], [MPP18'], [BP18']).
- Chaotic behavior ([PPP17']).

Long-term behavior in Dynamics

Given a continuous or discrete time (learning) dynamics we are interested in its long-term behavior...

- Convergence of time average ([FS99'], ...).
- Last iterate convergence ([PPP17'], [DP18']).
- Cycling behavior ([PPP17']) or *recurrent* behavior ([PS14'], [PNCS14'], [MPP18'], [BP18']).
- Chaotic behavior ([PPP17']).

Many applications to Game Theory, Optimization and Learning (GANs).

2-player zero sum games

Definition

Player $\mathbf{y}$ gets payoff $\mathbf{x}^T P\mathbf{y}$ and $\mathbf{x}$ gets $-\mathbf{x}^T P\mathbf{y}$. A Nash equilibrium is a solution to:

$$\min_{\mathbf{x}\in\Delta_n} \max_{\mathbf{y}\in\Delta_m} \mathbf{x}^T P\mathbf{y}.$$

Rock-Paper-Scissors

$$P_{RPS} = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}.$$

2-player zero sum games (cont.)

Time average converges to a Nash equilibrium when players update according to many learning dynamics including

- Fictitious play.
- MWUA and more generally for FTRL, FTPL.

2-player zero sum games (cont.)

Time average converges to a Nash equilibrium when players update according to many learning dynamics including

- Fictitious play.
- MWUA and more generally for FTRL, FTPL.

The same does not hold for last iterate. The system might exhibit "cycling" behavior e.g.,

[MPP18']

- Recurrent behavior for continuous time FTRL.

Dynamic payoff matrix P

Question: What if P changes with time? Can we show similarly "cycling" behavior (i.e., recurrent behavior persists)?

Definition (Differential Game)

A game the state space of which is described via a system of differential equations (continuous time dynamical system).[1]

For a zero sum game with payoff $P(t)$:

$$\frac{dP_{ij}}{dt} = f_{ij}(\mathbf{x}(t), t), \text{ for all } i, j^2.$$

[1]Stochastic games are the discrete time analogue.

[2]Time homogeneous for our purposes.

Our Model

- Symmetric zero sum game with n strategies (species).
- We use $\mathbf{x}$ to denote mixed strategy for both players (x_i fraction of species i).

Define the n-RPS game with payoff

$$P_{nRPS} = \begin{bmatrix} 0 & -\alpha & 0 & 0 & \dots & 0 & 0 & \alpha \\ \alpha & 0 & -\alpha & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \alpha & 0 & -\alpha \\ -\alpha & 0 & 0 & 0 & \dots & 0 & \alpha & 0 \end{bmatrix}.$$

Our Model (cont.)

Our *dynamic* payoff matrix $P^{\mathbf{w}}$ is a convex combination of n matrices P_i plus a matrix P_{nRPS}:

$$P^{\mathbf{w}} = w_1 P_1 + w_2 P_2 + \cdots + w_n P_n + P_{nRPS},$$

where

$$P_i = \begin{bmatrix} 0 & \dots & 0 & -\mu & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & -\mu & 0 & \dots & 0 \\ \mu & \dots & \mu & \underbrace{0}_{(i,i)} & \mu & \dots & \mu \\ 0 & \dots & 0 & -\mu & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & -\mu & 0 & \dots & 0 \end{bmatrix}$$

and $\mu, \alpha > 0$. The weights $\mathbf{w}$ change with time. P_i favors species i when competing with other species.

Our Model (cont.)

The dynamics can be described as follows:

$$\frac{dx_i}{dt} = x_i \cdot (\sum_j P^{\mathbf{w}}_{ij} x_j - \mathbf{x}^\top P^{\mathbf{w}} \mathbf{x}), \ \frac{dw_i}{dt} = w_i \cdot \sum_j w_j (x_j - x_i) \quad \forall i. \quad (1)$$

Remark 1.
x_i is increasing as long as average payoff of strategy i is higher than zero and decreasing otherwise. w_i is increasing as long as average frequency is higher than x_i and decreasing otherwise.

Remark 2.
Generalizes in higher dimensions the model of Weitz et. al. appeared in PNAS 16'.

Our Main theorem

Theorem (Recurrence)

For all but measure zero of initial positions in $\Delta_n \times \Delta_n$, the trajectories of the dynamics (1) return arbitrarily close to their initial position an infinite number of times.

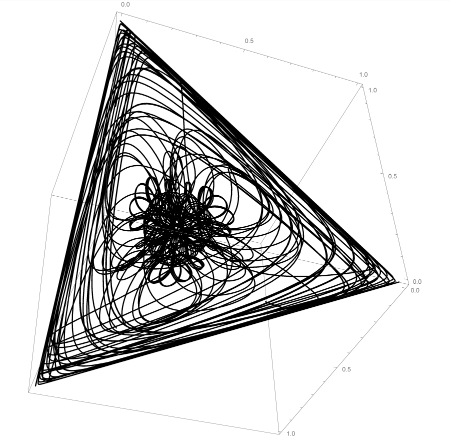

Figure 1: Trajectories of the vector **x** for different initial positions with $\mu = 0.1$, $\alpha = 1$. Trajectories intersect due to the fact 6 dimensions are projected to a 3D figure. The "cycling" behavior is observed.

Figures (cont.)

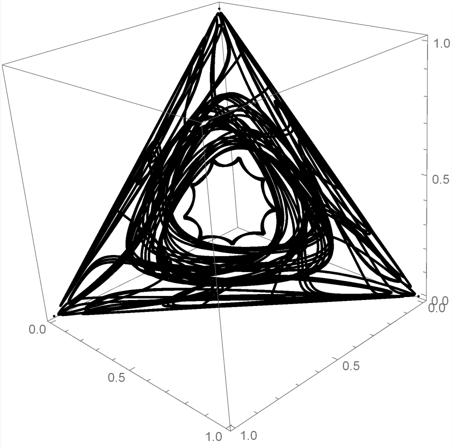

Figure 2: Trajectories of the vector **w** for different initial positions with $\mu = 0.1$, $\alpha = 1$. Trajectories intersect due to the fact 6 dimensions are projected to a 3D figure. The "cycling" behavior is observed.

Proof steps

We make use of the following important theorem.

Theorem (Poincaré Recurrence for continuous time)

If a flow preserves volume and has only bounded orbits then for each open set there exist orbits that intersect the set infinitely often.

- Flow: just the evolution of the dynamics.
- Bounded orbits: for each initial point, the trajectory does not diverge, inside a ball.

Volume preservation: Liouville's formula

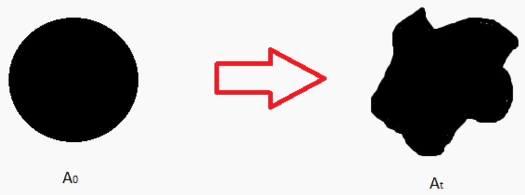

Figure 3: $\mu(A_0) = \mu(A_t)$ for all A_0, t where μ is the Lebesgue measure in $\mathbb{R}^{n-1} \times \mathbb{R}^{n-1}$ for this talk.

Theorem (Liouville theorem)

Let $\frac{d\mathbf{y}}{dt} = f(\mathbf{y})$ be an ode. It holds that $\frac{d\mu(A_t)}{dt} = \int_{A_t} (\nabla \cdot f) d\mu$ for each initial Lebesgue measurable set A_0. As long as $\nabla \cdot f = 0$, the flow preserves volume.

Coming up with a potential

We first project our dynamics to $\mathbb{R}^{2n-2}$ according to[3] $\Pi(\mathbf{y}) = \left(\log\left(\frac{y_1}{y_n}\right), ..., \log\left(\frac{y_{n-1}}{y_n}\right)\right)$. Boundary of simplex corresponds to vectors with infinity Euclidean norm in $\mathbb{R}^{2n-2}$.

Lemma (Constant motion of time)

$$\underbrace{\sum_{i=1}^{n} \log\left(\frac{1}{x_i}\right)}_{\geq 0} + \mu \underbrace{\sum_{i=1}^{n} \log\left(\frac{1}{w_i}\right)}_{\geq 0}$$

is independent of time (thus bounded).

[3] Map is bijective. $\Pi^{-1}(\mathbf{z}) = \left(\frac{e^{z_1}}{1+\sum_{j=1}^{n-1} e^{z_j}}, \ldots, \frac{e^{z_{n-1}}}{1+\sum_{j=1}^{n-1} e^{z_j}}, \frac{1}{1+\sum_{j=1}^{n-1} e^{z_j}}\right)$

Open Questions and Conclusion

- Provided a framework for proving recurrent behavior.
- Showed recurrent behavior for a class of differential games.
- Question: Generalize so that each strategy has different μ.
- Question: Different zero sum games?
- Question: Discrete time results?
- Question: Apply these techniques to other Learning dynamics.

Thank you!

Available Postdoc positions @SUTD!

Email us: ioannis@csail.mit.edu, georgios@sutd.edu.sg