Learning in Games with Conflicting Interests

D. Legacci, P. Mertikopoulos, C. H. Papadimitriou, G. Piliouras, B. Pradelski

June 19, 2024 – POLARIS Seminar

2024-06-19

Learning in Games with Conflicting Interests

Learning in Games with Conflicting Interests

D. Legacci, P. Mertikopoulos, C. H. Papadimitriou, G. Piliouras, B. Pradelski
June 19, 2024 – POLARIS Seminar

Question

What is the long-run behavior of learning dynamics in games where players have conflicting interests?

How

1. Conflicting Interests and Harmonic Games

   1.1 *Finite Games*

   1.2 *Strategic decomposition of games*

   1.3 *Mixed characterization of harmonic games*

2. Learning in Harmonic Games

   2.1 *Follow The Regularized Leader (FTRL)*

   2.2 *Proof Sketch - Recurrence of Replicator Dynamics*

   2.3 *Proof Sketch - Recurrence of FTRL*

Question

What is the long-run behavior of learning dynamics in games where players have conflicting interests?

How

1. Conflicting Interests and Harmonic Games
1.1 Finite Games
1.2 Strategic decomposition of games
1.3 Mixed characterization of harmonic games
2. Learning in Harmonic Games
2.1 Follow The Regularized Leader (FTRL)
2.2 Proof Sketch - Recurrence of Replicator Dynamics
2.3 Proof Sketch - Recurrence of FTRL

2024-06-19

Learning in Games with Conflicting Interests
└─Conflicting Interests and Harmonic Games

1. Conflicting Interests and Harmonic Games
1.1 Finite Games
1.2 Strategic decomposition of games
1.3 Mixed characterization of harmonic games
2. Learning in Harmonic Games

Conflicting Interests and Harmonic Games

Finite Games

Conflicting Interests and Harmonic Games
Finite Games

Finite normal form game $\Gamma = (\mathcal{N}, \mathcal{A}, u)$

- Finite set of **players** $\mathcal{N} = \{1, 2, \dots, N\}$, generic $i \in \mathcal{N}$

For each player $i \in \mathcal{N}$:

- Finite set of **pure strategies** $\mathcal{A}_i = \{1, 2, \dots, A_i\}$
- **Payoff function** $u_i : \mathcal{A} := \prod_{j \in \mathcal{N}} \mathcal{A}_j \to \mathbb{R}$

Reward $u_i(\alpha) \in \mathbb{R}$ of player $i \in \mathcal{N}$ at action profile $\alpha \in \mathcal{A}$

Standard notational convention

$$\mathcal{A} \ni \alpha = (\alpha_1, \dots, \alpha_N) = (\alpha_i, \alpha_{-i}) \quad \text{for any } i \in \mathcal{N}$$

Finite normal form game $\Gamma = (\mathcal{N}, \mathcal{A}, u)$
- Finite set of players $\mathcal{N} = \{1, 2, \dots, N\}$, generic $i \in \mathcal{N}$
For each player $i \in \mathcal{N}$:
- Finite set of pure strategies $\mathcal{A}_i = \{1, 2, \dots, A_i\}$
- Payoff function $u_i : \mathcal{A} := \prod_{j \in \mathcal{N}} \mathcal{A}_j \to \mathbb{R}$
Reward $u_i(\alpha) \in \mathbb{R}$ of player $i \in \mathcal{N}$ at action profile $\alpha \in \mathcal{A}$
Standard notational convention
$\mathcal{A} \ni \alpha = (\alpha_1, \dots, \alpha_N) = (\alpha_i, \alpha_{-i})$ for any $i \in \mathcal{N}$

- think of $\mathcal{A}$ as the space of states of the game; an element is a tuple that contains one strategy for each player
- given a strategy profile - a state - each player gets some payoff...
- ...and putting these together we get the global payoff

Example – 2×2 game

- $\mathcal{N} = \{1, 2\}$
- $\mathcal{A}_1 = \mathcal{A}_2 = \{A, B\}$
- $\mathcal{A} = \{AA, AB, BA, BB\}$

$$u : \mathcal{A} \to \mathbb{R}^2$$

$$\begin{aligned} (AA) &\longmapsto (1, -1) \\ (AB) &\longmapsto (-1, 1) \\ (BA) &\longmapsto (0, 0) \\ (BB) &\longmapsto (0, -1) \end{aligned} \quad \overset{\text{bimatrix notation}}{\longleftrightarrow} \quad \begin{pmatrix} 1, -1 & -1, 1 \\ 0, 0 & 0, -1 \end{pmatrix}$$

Example – 2×2 game
- $\mathcal{N} = \{1, 2\}$
- $\mathcal{A}_1 = \mathcal{A}_2 = \{A, B\}$
- $\mathcal{A} = \{AA, AB, BA, BB\}$
$u : \mathcal{A} \to \mathbb{R}^2$
$(AA) \longmapsto (1, -1)$
$(AB) \longmapsto (-1, 1)$
$(BA) \longmapsto (0, 0)$
$(BB) \longmapsto (0, -1)$
bimatrix notation
$\begin{pmatrix} 1, -1 & -1, 1 \\ 0, 0 & 0, -1 \end{pmatrix}$

- will be running example, keep an eye!

Visualize the strategic structure of a game: Preference graph

- Draw a node for each action profile $\alpha \in \mathcal{A}$
- Draw an edge between unilateral deviations, action profiles that differ only in the strategy of one player:

$$\alpha = (\alpha_i, \alpha_{-i}), \quad \beta = (\beta_i, \alpha_{-i})$$

- Weight each edge (α, β) by the payoff difference of the deviating player:

$$\mathrm{dev}_i(\alpha, \beta) = u_i(\alpha_i, \alpha_{-i}) - u_i(\beta_i, \alpha_{-i})$$

Visualize the strategic structure of a game: Preference graph

- Draw a node for each action profile $\alpha \in \mathcal{A}$
- Draw an edge between unilateral deviations, action profiles that differ only in the strategy of one player:

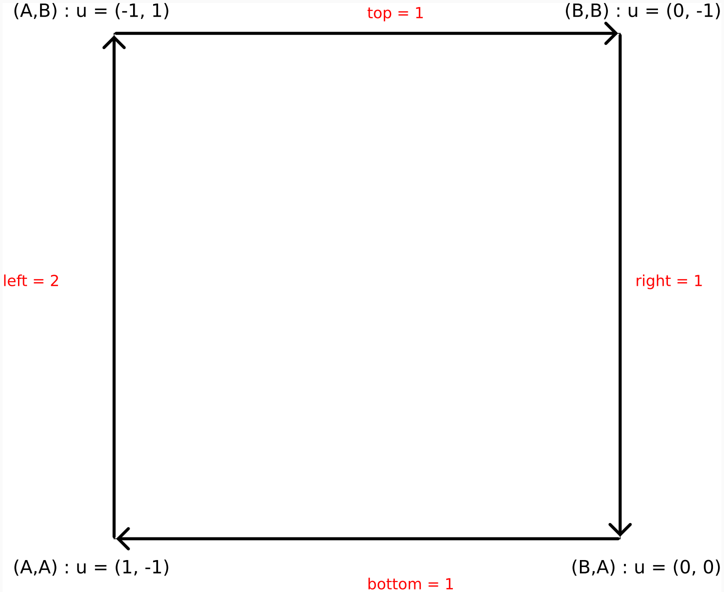
$$\alpha = (\alpha_i, \alpha_{-i}), \quad \beta = (\beta_i, \alpha_{-i})$$

- Weight each edge (α, β) by the payoff difference of the deviating player:

$$\mathrm{dev}_i(\alpha, \beta) = u_i(\alpha_i, \alpha_{-i}) - u_i(\beta_i, \alpha_{-i})$$

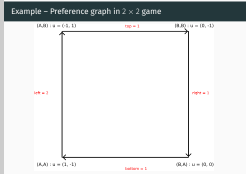
- Now there's a more intuitive way of visualizing the strategic structure of a game..

Example – Preference graph in 2×2 game



2024-06-19

Learning in Games with Conflicting Interests
└Conflicting Interests and Harmonic Games
└Finite Games
└Example – Preference graph in 2×2 game



- Looks rather empty now, but we will gradually populate
- nodes = action profiles
- edges = unilateral deviations

What do we do with a preference graph?

The preference graph of a game captures its strategic structure

- The orientation of the edges describes the interest of the players at each action profile $\alpha \in \mathcal{A}$
- The strategy $\alpha_i^* \in \mathcal{A}_i$ is a best response to $\alpha_{-i} \in \mathcal{A}_{-i}$ if

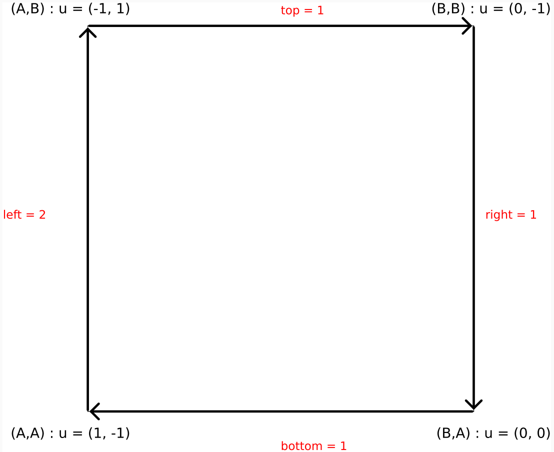
$$u_i(\alpha_i^*, \alpha_{-i}) \geq u_i(\alpha_i, \alpha_{-i}) \quad \text{for all } \alpha_i \in \mathcal{A}_i$$

Let's have a look at the best responses in the previous example.

What do we do with a preference graph?
The preference graph of a game captures its strategic structure
- The orientation of the edges describes the interest of the players at each action profile $\alpha \in \mathcal{A}$
- The strategy $\alpha_i^* \in \mathcal{A}_i$ is a best response to $\alpha_{-i} \in \mathcal{A}_{-i}$ if
$$u_i(\alpha_i^*, \alpha_{-i}) \geq u_i(\alpha_i, \alpha_{-i}) \quad \text{for all } \alpha_i \in \mathcal{A}_i$$
Let's have a look at the best responses in the previous example.

- Why do we bother?
- It captures the interest of a player at any state of the game
- we can make this more precise with the notion of best response…

Example – Preference graph and best responses



Player 1: *You A? Me A!*

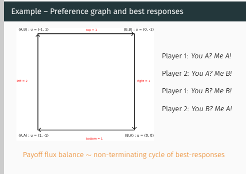
Player 2: *You A? Me B!*

Player 1: *You B? Me B!*

Player 2: *You B? Me A!*

Payoff flux balance ∼ non-terminating cycle of best-responses

2024-06-19



- BR for player 1 if player 2 plays A: player 2 plays A down, so player 1 wants A
- BR for player 2 if player 1 plays A: player 1 plays A left, so player 2 wants B
- etc
- flux balance = net difference between payoff at action and payoff of all possible deviations
- Building on this idea, we model strategic interaction based on conflict

Conflicting Interests and Harmonic Games

Strategic decomposition of games

Conflicting Interests and Harmonic Games
Strategic decomposition of games

Harmonic Games [Can+11; Abd+22]

A finite normal form game $(\mathcal{N}, \mathcal{A}, u)$ is called harmonic if

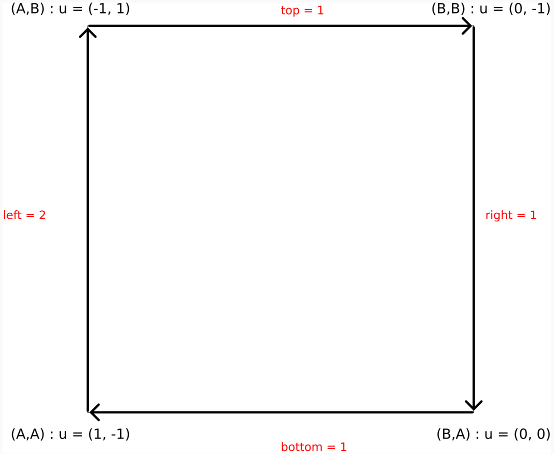
- there exists a collection of weights $\mu_{i\alpha_i} \in (0, \infty)$ such that
- for every action profile $\alpha \in \mathcal{A}$
- the weighted net payoff difference between α and all the possible unilateral deviations (β_i, α_{-i})
- vanishes identically:

$$\sum_{i \in \mathcal{N}} \sum_{\beta_i \in \mathcal{A}_i} \mu_{i\beta_i} \left[u_i(\alpha) - u_i(\beta_i, \alpha_{-i})\right] = 0 \quad \text{for all } \alpha \in \mathcal{A}$$

Harmonic Games [Can+11; Abd+22]
A finite normal form game $(\mathcal{N}, \mathcal{A}, u)$ is called harmonic if
- there exists a collection of weights $\mu_{i\alpha_i} \in (0, \infty)$ such that
- for every action profile $\alpha \in \mathcal{A}$
- the weighted net payoff difference between α and all the possible unilateral deviations (β_i, α_{-i})
- vanishes identically:
$\sum_{i \in \mathcal{N}} \sum_{\beta_i \in \mathcal{A}_i} \mu_{i\beta_i} [u_i(\alpha) - u_i(\beta_i, \alpha_{-i})] = 0$ for all $\alpha \in \mathcal{A}$

- harmonic games introduced by Candogan et al in seminal work
- rather algebraic approach; here motivated by strategical viewpoint
- meaning: for every action profile there are players interested in deviating towards and players interested in deviating away, this interest being weighted by some parameters

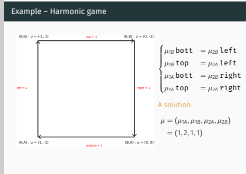
Example – Harmonic game



$$\begin{cases} \mu_{1B}\,\texttt{bott} &= \mu_{2B}\,\texttt{left} \\ \mu_{1B}\,\texttt{top} &= \mu_{2A}\,\texttt{left} \\ \mu_{1A}\,\texttt{bott} &= \mu_{2B}\,\texttt{right} \\ \mu_{1A}\,\texttt{top} &= \mu_{2A}\,\texttt{right} \end{cases}$$

A solution:

$$\mu = (\mu_{1A}, \mu_{1B}, \mu_{2A}, \mu_{2B}) = (1, 2, 1, 1)$$



- go back to our example and try to find weights solving the harmonic equation
- Oss: with this measure player 1 assigns more weight to their second strategy

Strategic decomposition of games

Harmonic games: For any player considering deviation **towards** any action profile, there exist other players with incentive to deviate **away** from said profile → anti-aligned interests

Potential games [MS96][1]: There exist action profile(s) every player has incentive to deviate **towards**→ aligned interests

Theorem ([Can+11; Abd+22])

For any choice of measure μ, every game in normal form can be uniquely decomposed as the sum of a potential game and a harmonic game.*

[1] $u_i(\alpha_i, \alpha_{-i}) - u_i(\beta_i, \alpha_{-i}) = \phi(\alpha_i, \alpha_{-i}) - \phi(\beta_i, \alpha_{-i})$ with $\phi : \mathcal{A} \to \mathbb{R}$

- so, harmonic games are strategic counterpart of potential ones
- if you know def of potential games good, else no problem; suffices to say that aligned interests
- Uniquely: Up to affine transformations that do not alter the strategic structure of the game.

Strategic relevance of harmonic games

- Natural complement to potential games from a strategic viewpoint
- Archetypal model for strategic interaction with conflicting interests

What is the long-run behavior of learning dynamics in games where players have conflicting interests?

→ Games mixed extension

Strategic relevance of harmonic games
- Natural complement to potential games from a strategic viewpoint
- Archetypal model for strategic interaction with conflicting interests
What is the long-run behavior of learning dynamics in games where players have conflicting interests?
→ Games mixed extension

- In light of this: harmonic games are the canonical complement to potential games, and standard to model conflicting interactions.
- There is also potential zero-sum decomposition but nor orthogonal, in particular zero-sum and potential intersect non trivially. Conversely (up to non strategic games) harmonic and potential intersect only at the zero game.
- since the behavior of learning dynbamics in potential games is well understoof, this brings weights to the overarching question:...

Conflicting Interests and Harmonic Games

Mixed characterization of harmonic games

Conflicting Interests and Harmonic Games
Mixed characterization of harmonic games

Mixed extension of a game

Mixed strategy x_i = probability distribution over pure strategies:

$$x_i \in \mathcal{X}_i = \Delta(\mathcal{A}_i) \subset \mathbb{R}^{A_i} \quad \text{for all } i \in \mathcal{N}$$

Expected payoff $u_i : \mathcal{X} := \prod_{j \in \mathcal{N}} \mathcal{X}_j \to \mathbb{R}$,

$$u_i(x) = \mathbb{E}_{\alpha \sim x}[u_i(\alpha)] = \sum_{\alpha \in \mathcal{A}} u_i(\alpha)\, x_{1,\alpha_1} \dots x_{N,\alpha_N}$$

Payoff field $v_i : \mathcal{X} \to \mathbb{R}^{A_i}$ = individual payoff gradient:

$$v_i(x) := \nabla_i u_i(x) \equiv \left(\frac{\partial u_i(x)}{\partial x_{i,\alpha_i}} \right)_{\alpha_i \in \mathcal{A}_i}$$

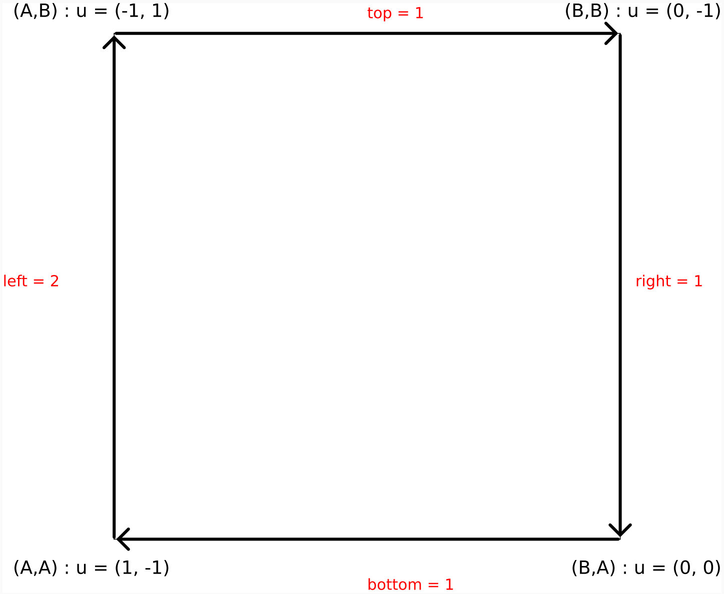
Mixed extension of a game
Mixed strategy x_i = probability distribution over pure strategies:
$x_i \in \mathcal{X}_i = \Delta(\mathcal{A}_i) \subset \mathbb{R}^{A_i}$ for all $i \in \mathcal{N}$
Expected payoff $u_i : \mathcal{X} := \prod_{j \in \mathcal{N}} \mathcal{X}_j \to \mathbb{R}$,
$u_i(x) = \mathbb{E}_{\alpha \sim x}[u_i(\alpha)] = \sum_{\alpha \in \mathcal{A}} u_i(\alpha)\, x_{1,\alpha_1} \dots x_{N,\alpha_N}$
Payoff field $v_i : \mathcal{X} \to \mathbb{R}^{A_i}$ = individual payoff gradient:
$v_i(x) := \nabla_i u_i(x) \equiv \left(\frac{\partial u_i(x)}{\partial x_{i,\alpha_i}} \right)_{\alpha_i \in \mathcal{A}_i}$

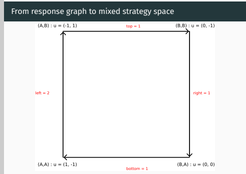
- **Expected payoff**: expectation value of $u_i(a)$ where the pure strategy profile a is drawn according to the probability distribution x

$$u_i : \mathcal{X} := \prod_{i \in \mathcal{N}} \mathcal{X}_i \to \mathbb{R}, \quad \underbrace{(x_1, \dots, x_N)}_{\text{mixed strategy profile}} \mapsto \mathbb{E}_{a \sim x}[u_i(a)] = \sum_{a \in \mathcal{A}} u_i(a) \underbrace{\prod_{j \in \mathcal{N}} x_{j,a_j}}_{P_x(a)}$$

- Take $\mathcal{V}_i^*$ as notation to distinguish strategy space from payoff space; deeper reason why $v_i \in \mathcal{V}_i^*$ is that v_i is actually a differential, not a gradient, and as such it lives in the dual space

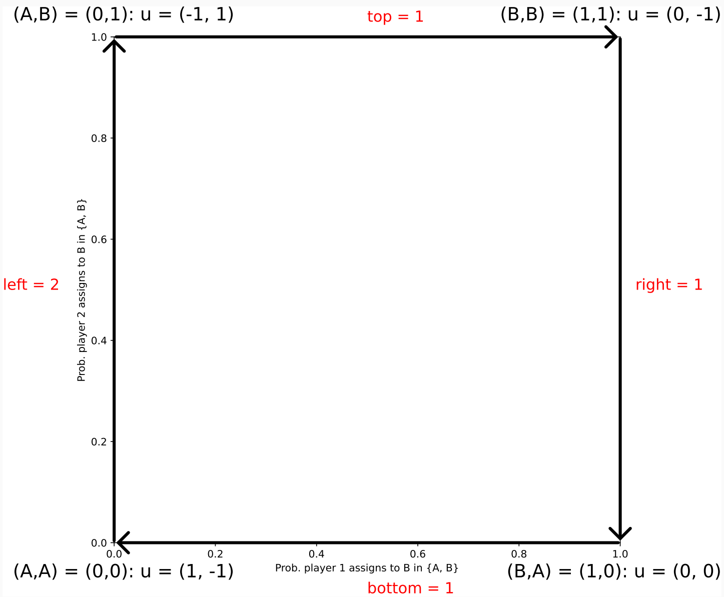
From response graph to mixed strategy space

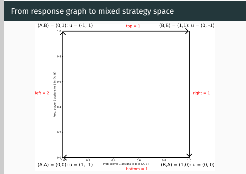




- the response graph can accomodate for these notions:

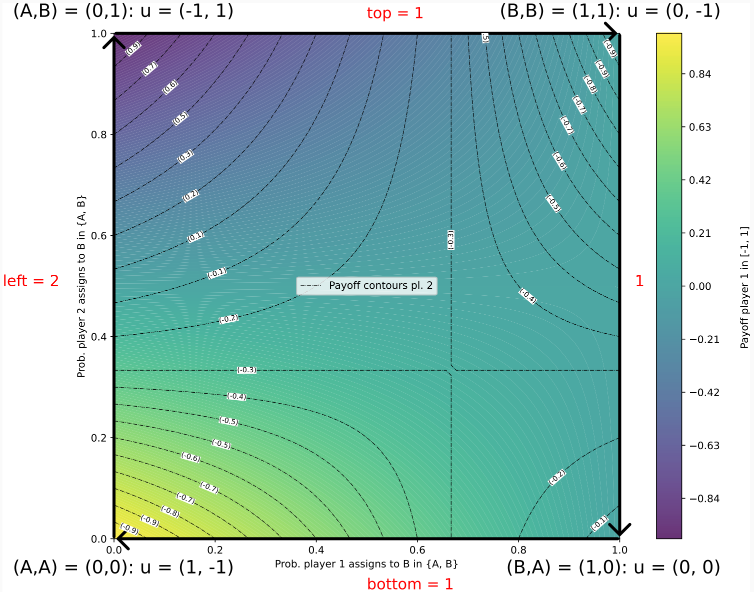
From response graph to mixed strategy space





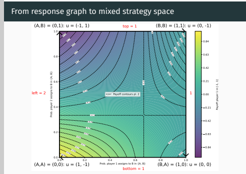
- first, let axes denote mixed strategies to play, second strategy; and note mixed representation of pures

From response graph to mixed strategy space



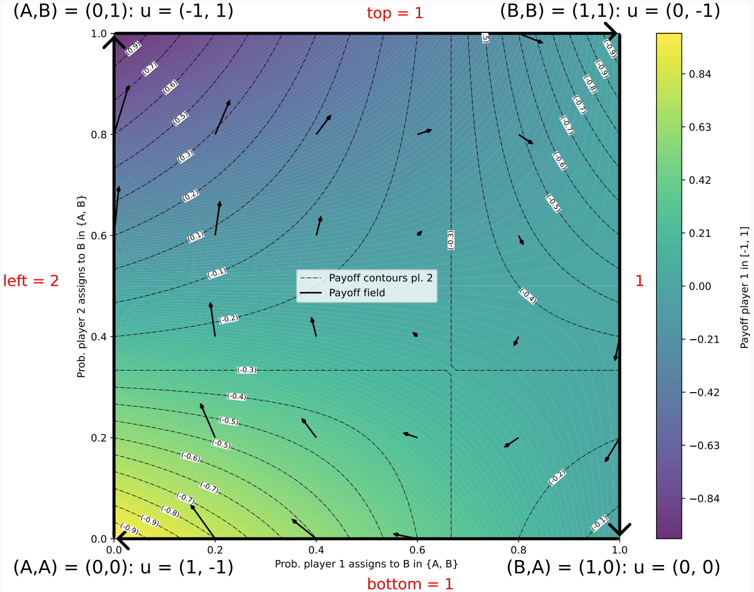
2024-06-19

Learning in Games with Conflicting Interests
└Conflicting Interests and Harmonic Games
└Mixed characterization of harmonic games
└From response graph to mixed strategy space

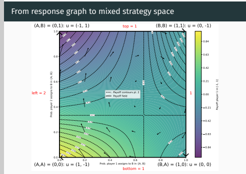


- then, include contours of payoff functions. Color = player 1, lines = player 2
- player 1 moves horizontally; on the bottom wants to go left (towards yellow); on top wants to go right (towards non-purple)
- player 2 moves vertically; on the left wants to go up; on the right wants to go down (contour labels)
- next step: direction of maximal individual payoff increase i.e. individual gradients

From response graph to mixed strategy space



2024-06-19



- individual payoff gradients; note again right left up down pattern and circular pattern
- next, characterize harmonic games in this language, via payoff field

Mixed characterization of harmonic games – Strategic center

Proposition ([LMP24b])

A finite game $\Gamma = \Gamma(\mathcal{N}, \mathcal{A}, u)$ is harmonic if and only if it admits a strategic center (m, q), i.e., if there exist a vector $m \in \mathbb{R}^N_{++}$ and a fully mixed strategy $q \in \mathcal{X}$ such that

$$\sum_{i \in \mathcal{N}} m_i \langle v_i(x), x_i - q_i \rangle = 0 \quad \textit{for all } x \in \mathcal{X}.$$

Proof sketch. Denote $|\mu_i| := \sum_{\alpha_i} \mu_{i\alpha_i}$

- By definition of harmonic games and multilinearity, $\sum_{i \in \mathcal{N}} |\mu_i| \left\langle v_i(x), x_i - \frac{\mu_i}{|\mu_i|} \right\rangle = 0 \quad \text{for all } x \in \mathcal{X}$
- $m_i = |\mu_i|$ and $q_i = \mu_i / |\mu_i|$ $\square$

Mixed characterization of harmonic games – Strategic center
Proposition ([LMP24b])
A finite game $\Gamma = \Gamma(\mathcal{N}, \mathcal{A}, u)$ is harmonic if and only if it admits a strategic center (m, q), i.e., if there exist a vector $m \in \mathbb{R}^N_{++}$ and a fully mixed strategy $q \in \mathcal{X}$ such that
$\sum_{i \in \mathcal{N}} m_i \langle v_i(x), x_i - q_i \rangle = 0$ for all $x \in \mathcal{X}$.
Proof sketch. Denote $|\mu_i| := \sum_{\alpha_i} \mu_{i\alpha_i}$
By definition of harmonic games and multilinearity, $\sum_{i \in \mathcal{N}} |\mu_i| \left\langle v_i(x), x_i - \frac{\mu_i}{|\mu_i|} \right\rangle = 0$ for all $x \in \mathcal{X}$
$m_i = |\mu_i|$ and $q_i = \mu_i / |\mu_i|$ $\square$

- center = one number per player, and one fully mixed strategy per player

Mixed characterization of harmonic games – Strategic center

Harmonic game: every player has a fully mixed strategy q_i s.t.,

- for all $x \in \mathcal{X}$
- the payoff vector $v(x)$
- points in the direction that is perpendicular
- with respect to a m-weighted inner product
- to $x - q$, with $q_i = \mu_i / |\mu_i|$

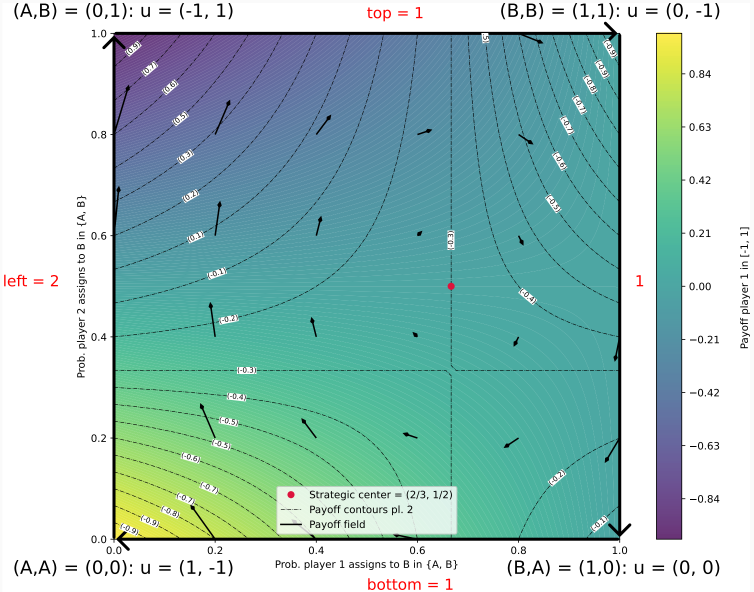
Recall harmonic measure of running example:

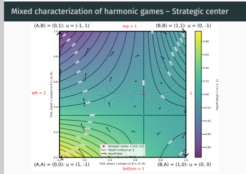
$$\mu = (\mu_{1A}, \mu_{1B}, \mu_{2A}, \mu_{2B}) = (1, 2, 1, 1) \implies q = \left[\left(\frac{1}{3}, \frac{2}{3}\right), \left(\frac{1}{2}, \frac{1}{2}\right)\right]$$

Mixed characterization of harmonic games – Strategic center
Harmonic game: every player has a fully mixed strategy q_i s.t.,
- for all $x \in \mathcal{X}$
- the payoff vector $v(x)$
- points in the direction that is perpendicular
- with respect to a m-weighted inner product
- to $x - q$, with $q_i = \mu_i / |\mu_i|$
Recall harmonic measure of running example:
$\mu = (\mu_{1A}, \mu_{1B}, \mu_{2A}, \mu_{2B}) = (1, 2, 1, 1) \implies q = \left[\left(\frac{1}{3}, \frac{2}{3}\right), \left(\frac{1}{2}, \frac{1}{2}\right)\right]$

- meaning of strategic center: perpendicularity and circular structure

Mixed characterization of harmonic games – Strategic center





- recall example; I know it's disappointing the red dot is not at the crossing of the contour lines, but that point does not have a particular meaning.
- Conversely from the measure (1, 2, 1, 1) we see that the center is 2/(1+2), 1/(1+1)
- connect center to base of payfield with segment
- finally, move to dynamics; one last remark if time
- if short on time skip next slide

Harmonic games and zero-sum games

Lemma ([LMP24b])

- *A harmonic game with measure μ such that $\sum_{\alpha_i} \mu_{i\alpha_i} = 1$ is (strategically equivalent to) a zero-sum game.*

- *Every two-player zero-sum game with interior Nash equilibrium x^* is harmonic, with measure $\mu = x^*$.*

Take-away: harmonic games generalize two-player zero-sum games with interior equilibrium

Harmonic games and zero-sum games
Lemma ([LMP24b])
A harmonic game with measure μ such that $\sum_{\alpha_i} \mu_{i\alpha_i} = 1$ is (strategically equivalent to) a zero-sum game.
Every two-player zero-sum game with interior Nash equilibrium x^* is harmonic, with measure $\mu = x^*$.
Take-away: harmonic games generalize two-player zero-sum games with interior equilibrium

- Traditionally, zero-sum games are used to model conflict
- but zero-sum and potential intersect non trivially
- and zero-sum is too loose in N-player games
- harmonic games do include an important class of zero-sum games, and are themselves zero-sum in special circumstances

So far: Only strategic considerations

Next: Dynamical consequences of circular strategic structure in harmonic games

So far: Only strategic considerations

Next: Dynamical consequences of circular strategic structure in harmonic games

2024-06-19

1. Conflicting Interests and Harmonic Games
2. Learning in Harmonic Games
2.1 Follow The Regularized Leader (FTRL)
2.2 Proof Sketch - Recurrence of Replicator Dynamics
2.3 Proof Sketch - Recurrence of FTRL

Learning in games

In which games and through which adaptation processes do players learn to emulate rational behaviour through repeated interactions?

- Continuous-time, deterministic, multi-agent decision processes
- Agents aim at maximizing their payoff
- Minimal requirement: minimization of individual regret

$$\mathrm{Reg}_i(T) = \max_{p \in \mathcal{X}_i} \int_0^T [u_i(p_i, x_{-i}(t)) - u_i(x(t))] \, dt$$

- A player has **no-regret** if $\mathrm{Reg}_i(T) = o(T)$ as $T \to \infty$
- Standard scheme to achieve no-regret: **FTRL**

Learning in games

In which games and through which adaptation processes do players learn to emulate rational behaviour through repeated interactions?

- Continuous-time, deterministic, multi-agent decision processes
- Agents aim at maximizing their payoff
- Minimal requirement: minimization of individual regret

$$\mathrm{Reg}_i(T) = \max_{p \in \mathcal{X}_i} \int_0^T [u_i(p_i, x_{-i}(t)) - u_i(x(t))] \, dt$$

- A player has no-regret if $\mathrm{Reg}_i(T) = o(T)$ as $T \to \infty$
- Standard scheme to achieve no-regret: FTRL

- player's regret = diff between payoff of best strategy in hindsight, and incurred payoff along trajectory of play until now
- no regret, sublinear in time
- look at standard class with strong regret guarantees

Learning in Harmonic Games

Follow The Regularized Leader (FTRL)

Learning in Harmonic Games
Follow The Regularized Leader (FTRL)

Follow The Regularized Leader

- Each player tracks cumulative incurred payoff…
- …and updates strategy according to this information via so-called choice map:

$$\text{choice map } Q : \text{cumulative payoff} \mapsto \text{next strategy}$$

Follow The Regularized Leader
- Each player tracks cumulative incurred payoff…
- …and updates strategy according to this information via so-called choice map:
choice map Q : cumulative payoff ↦ next strategy

- we need of course to define this choice map

Follow The Regularized Leader

$$\begin{cases} y_i(t) &= \int_0^t v_i\,(x(s))\;ds \\ x_i(t) &= Q_i\,(y_i(t)) \end{cases} \Rightarrow \begin{cases} \dot{y}_i(t) &= v_i\,(x(t)) \\ x_i(t) &= Q_i\,(y_i(t)) \end{cases} \qquad \text{(FTRL)}$$

Player's set of optimal strategies given mixed strategy $x \in \mathcal{X}$:

$$\arg\max_{x_i \in \mathcal{X}_i} \{v_i(x) \cdot x_i\}$$

For each player consider strongly convex regularizer $h_i : \mathcal{X}_i \to \mathbb{R}$ and define the choice map:

$$Q_i : \mathbb{R}^{\mathcal{A}_i} \to \mathcal{X}_i$$

$$y_i \longmapsto \arg\max_{x_i \in \mathcal{X}_i} \{y_i \cdot x_i - h_i(x_i)\}$$

"Soft" arg max correspondence, single valued.

Follow The Regularized Leader
Player's set of optimal strategies given mixed strategy $x \in \mathcal{X}$:
For each player consider strongly convex regularizer $h_i : \mathcal{X}_i \to \mathbb{R}$ and define the choice map:
"Soft" arg max correspondence, single valued.

- $x_i(t) \in \mathcal{X}_i$ is mixed strategy of player i at time t
- $y_i(t) \in \mathcal{V}_i^*$ aggregates payoffs of player i until time t
- Aggregate payoff used to update strategy via choice map Q

Example - Exponential Weights and Replicator Dynamics

- Entropic regularizer $h_i(x_i) = \sum_{\alpha_i \in \mathcal{A}_i} x_{i\alpha_i} \log x_{i\alpha_i}$
- Induces **logit** choice map

$$Q_i(y_i) = \frac{(e^{y_{i\alpha_i}})_{\alpha_i \in \mathcal{A}_i}}{\sum_{\beta_i} e^{y_{i\beta_i}}}$$

For each player $\dot{y}_i = v_i(x)$ and $x_i = Q_i(y_i)$ gives

$$\boxed{\dot{x}_{i,\alpha_i} = x_{i,\alpha_i}\Big(u_i(\alpha_i, x_{-i}) - u_i(x)\Big)} \tag{RD}$$

Example - Exponential Weights and Replicator Dynamics

- Taylor Jonker 1978
- the probability to use a pure strategy is exponentially proportional to cumulative payoff
- terminology: from mathematical biology, used to model species evolution
- action share grows if at the current game state the payoff of using such pure strategy is higher than the expected payoff
- we are finally in the position to state the main result of this talk

Main result – FTRL is Poincaré recurrent in harmonic games

Theorem ([LMP24a; LMP24b], first announced [PP23])

Suppose Γ is harmonic. Then FTRL is recurrent, i.e., almost every orbit $x(t)$ of returns arbitrarily close to its starting point infinitely often.

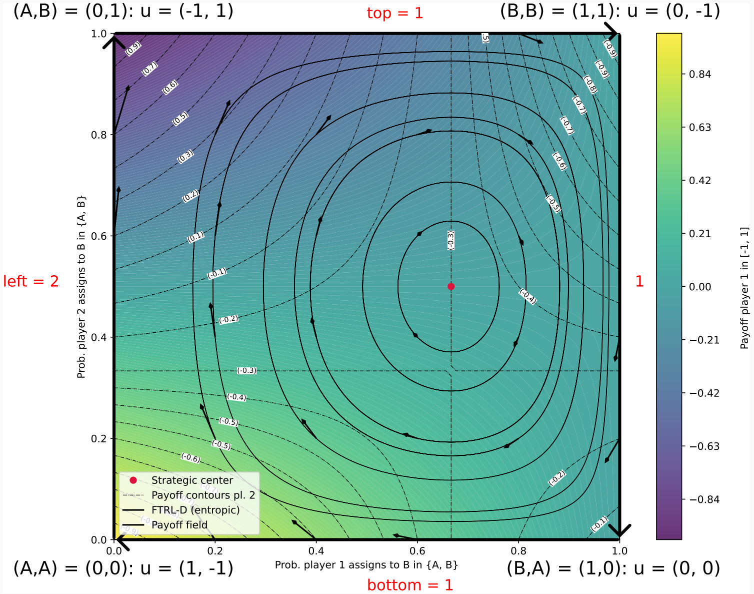
Main result – FTRL is Poincaré recurrent in harmonic games

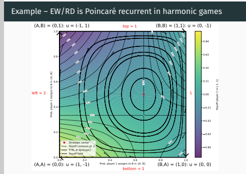
Theorem ([LMP24a; LMP24b], first announced [PP23])

Suppose Γ is harmonic. Then FTRL is recurrent, i.e., almost every orbit x(t) of returns arbitrarily close to its starting point infinitely often.

- Before commenting, let's go back to the familiar example

Example – EW/RD is Poincaré recurrent in harmonic games





- in this case recurrence reduced to periodicity because of the phase space dimension; in higher dimension trajecgories are actually recurrent, not periodic.
- NE is stable, not AS stable

FTRL is Poincaré recurrent in harmonic games – Remarks

- FTRL in continuous time has no hope to converge in harmonic games
- Nest known result [MPP18] for 2-player zero-sum games with interior equilibrium, and generalize to N-player games

---

- FTRL converges globally in potential games
- Harmonic games complement potential games not only from strategic but also from dynamic viewpoint

- if converged, would not come back to initial point
- generalizes important family of non-convergence results
- harmonic are not only strategic but also dynamic potentila compolement

Proof sketch - Recurrence of FTRL in harmonic games

Tools: Dynamical systems theory

$$\dot{x} = f(x), \quad f \text{ vector field } : M \text{ open } \subseteq \mathbb{R}^n \to \mathbb{R}$$

- **Liouville's theorem**

$$\operatorname{div} f = 0 \implies \text{volum-preserving system}$$

- **Poincaré's theorem**

$$\begin{cases} \text{volum-preserving system} \\ \text{bounded orbits} \end{cases} \implies \text{recurrent system}$$

Proof sketch - Recurrence of FTRL in harmonic games
Tools: Dynamical systems theory
$\dot{x} = f(x), \quad f$ vector field $: M$ open $\subseteq \mathbb{R}^n \to \mathbb{R}$
Liouville's theorem
$\operatorname{div} f = 0 \implies$ volum-preserving system
Poincaré's theorem
volum-preserving system, bounded orbits $\implies$ recurrent system

- div as trace of Jacobian
- volume preserving means that the volume of an initial set of initial condition does not change as the system evolves, neither shrinking nor expanding
- Liouville: sufficient condition for volume preservation
- poincare: sufficient condition for recurrence

If time, sketch proofs, else...

Thanks for your attention

If time, sketch proofs, else...
Thanks for your attention

Learning in Harmonic Games

Proof Sketch - Recurrence of Replicator Dynamics

Learning in Harmonic Games
Proof Sketch - Recurrence of Replicator Dynamics

Replicator dynamics as Riemannian individual payoff gradient

Recall: payoff field is individual payoff Euclidean gradient

$$v_i(x) = \nabla_i u_i(x)$$

As it turns out: replicator field is individual payoff gradient under non-Euclidean geometry[2] g^*:

$$RD_i(x) = \nabla_i^{g^*} u_i(x)$$

Define divergence operator with respect to geometry g^*

[2] cf. Shahshahani [Sha79]

Replicator dynamics as Riemannian individual payoff gradient

Recall: payoff field is individual payoff Euclidean gradient

$v_i(x) = \nabla_i u_i(x)$

As it turns out: replicator field is individual payoff gradient under non-Euclidean geometry[2] g^*:

$RD_i(x) = \nabla_i^{g^*} u_i(x)$

Define divergence operator with respect to geometry g^*

[2] cf. Shahshahani [Sha79]

. .

Equivalence between harmonic and divergence-free games

Theorem ([LMP24a])

A finite game $\Gamma = \Gamma(\mathcal{N}, \mathcal{A}, u)$ *is harmonic with uniform measure* $\mu_{i\alpha_i} = 1$ *if and only if its associated replicator vector field* $\nabla_i^{g^*} u_i(x)$ *has zero divergence under the geometry* g^*.

- By Liouville's theorem, RD on harmonic games is volume-preserving in strategy space;
- RD has only bounded orbits in all games;
- Recurrence follows by Poincaré's theorem. □

Equivalence between harmonic and divergence-free games

Theorem ([LMP24a])

A finite game $\Gamma = \Gamma(\mathcal{N}, \mathcal{A}, u)$ is harmonic with uniform measure $\mu_{i\alpha_i} = 1$ if and only if its associated replicator vector field $\nabla_i^{g^*} u_i(x)$ has zero divergence under the geometry g^*.

- By Liouville's theorem, RD on harmonic games is volume-preserving in strategy space;
- RD has only bounded orbits in all games;
- Recurrence follows by Poincaré's theorem. □

Riemannian approach: Pros and cons

Pros

- Surprising connection between Riemannian construction and uniform harmonic games
- Fine understanding of dynamics-geometry interplay in strategy space

Cons

- Harmonic / divergence-free equivalence fails changing metric
- Need to change approach for general FTRL case.

For general FTRL adapt standard method [MPP18; BP19]

$\rightarrow$ relatively easy result, but lose geometrical interpretation of what happens in strategy space

Riemannian approach: Pros and cons

Pros
- Surprising connection between Riemannian construction and uniform harmonic games
- Fine understanding of dynamics-geometry interplay in strategy space

Cons
- Harmonic / divergence-free equivalence fails changing metric
- Need to change approach for general FTRL case.

For general FTRL adapt standard method [MPP18; BP19]

$\rightarrow$ relatively easy result, but lose geometrical interpretation of what happens in strategy space

Learning in Harmonic Games

Proof Sketch - Recurrence of FTRL

Learning in Harmonic Games
Proof Sketch - Recurrence of FTRL

FTRL in harmonic games admits a constant of motion

Proposition ([LMP24b], first announced [PP23])

Let $\Gamma = \Gamma(\mathcal{N}, \mathcal{A}, u)$ be a finite game and consider a vector $m \in \mathbb{R}^N_{++}$ and a fully mixed strategy $q \in \mathcal{X}$. Then the function defined by

$$F_{m,q}(y) := \sum_i m_i \left[h_i(q_i) + h_i^*(y_i) - \langle q_i, y_i \rangle\right]$$

is a constant of motion under FTRL if and only if Γ is harmonic with strategic center (m, q).

FTRL in harmonic games admits a constant of motion
Proposition ([LMP24b], first announced [PP23])
Let $\Gamma = \Gamma(\mathcal{N}, \mathcal{A}, u)$ be a finite game and consider a vector $m \in \mathbb{R}^N_{++}$ and a fully mixed strategy $q \in \mathcal{X}$. Then the function defined by
$F_{m,q}(y) := \sum_i m_i [h_i(q_i) + h_i^*(y_i) - \langle q_i, y_i \rangle]$
is a constant of motion under FTRL if and only if Γ is harmonic with strategic center (m, q).

- h convex conjugate $h^*(y) = max_x\{y \cdot x - h(x)\}$
- known as Fenchek coupling
- standard technique to exhibit bounded orbits, study level sets of constant of motion since trajectories are constrained therein

FTRL is divergence-free in all games in payoff space

$$\begin{cases} \dot{y}_i &= v_i(x) \\ x_i &= Q_i(y_i) \end{cases} \implies \dot{y}_i = v_i(Q(y)) \qquad \text{(FTRL)}$$

$$\frac{d\dot{y}_{i\alpha_i}}{dy_{j\beta_j}} \equiv 0 \quad \text{by multilinearitity of the payoff functions}$$

- By Liouville's theorem, FTRL in payoff space is volume-preserving in all games ;
- the constant of motion can be used to show that FTRL in harmonic games has only bounded orbits;
- Recurrence follows by Poincaré's theorem. □

FTRL is divergence-free in all games in payoff space

$$\begin{cases} \dot{y}_i &= v_i(x) \\ x_i &= Q_i(y_i) \end{cases} \implies \dot{y}_i = v_i(Q(y)) \qquad \text{(FTRL)}$$

$$\frac{d\dot{y}_{i\alpha_i}}{dy_{j\beta_j}} \equiv 0 \quad \text{by multilinearitity of the payoff functions}$$

- By Liouville's theorem, FTRL in payoff space is volume-preserving in all games ;
- the constant of motion can be used to show that FTRL in harmonic games has only bounded orbits;
- Recurrence follows by Poincaré's theorem. □

- NB in Riemannian case, the special harmonic property is volume preservation; bounded orbits always true
- here, volume preservation in payoff space always true; special harmonic property is bounded orbits, by const of motion

Bibliography

[Abd+22] Abdou, Joseph et al. “Decomposition of Games: Some Strategic Considerations”. In: *Mathematics of Operations Research* 47.1 (Feb. 2022), pp. 176–208.

[BP19] Boone, Victor and Piliouras, Georgios. *From Darwin to Poincaré and von Neumann: Recurrence and Cycles in Evolutionary and Algorithmic Game Theory*. Oct. 2019. arXiv: `1910.01334 [cs, math, q-bio]`.

[Can+11] Candogan, Ozan et al. “Flows and Decompositions of Games: Harmonic and Potential Games”. In: *Mathematics of Operations Research* 36.3 (Aug. 2011), pp. 474–503.

[LMP24a] Legacci, Davide, Mertikopoulos, Panayotis, and Pradelski, Bary. “A Geometric Decomposition of Finite Games: Convergence vs. Recurrence under Exponential Weights”. In: *Forty-First International Conference on Machine Learning*. June 2024.

[LMP24b] Legacci, Davide, Mertikopoulos, Panayotis, and Pradelski, Bary. “No-Regret Learning in Harmonic Games: Extrapolation in the Face of Conflicting Interests”. In: (May 2024).

[MPP18] Mertikopoulos, Panayotis, Papadimitriou, Christos, and Piliouras, Georgios. “Cycles in Adversarial Regularized Learning”. In: *Proceedings of the 2018 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. Proceedings. Society for Industrial and Applied Mathematics, Jan. 2018, pp. 2703–2717.

[MS96] Monderer, Dov and Shapley, Lloyd S. “Potential Games”. In: *Games and Economic Behavior* 14.1 (May 1996), pp. 124–143.

[PP23] Papadimitriou, Christos and Piliouras, Georgios. “Towards a Unified Theory of Learning in Games: Chaos, Regret and Equilibration.”. In: *The 2023 Alpine Game Theory Symposium*. June 2023.

[Sha79] Shahshahani, Siavash. *A New Mathematical Framework for the Study of Linkage and Selection*. American Mathematical Soc., 1979.