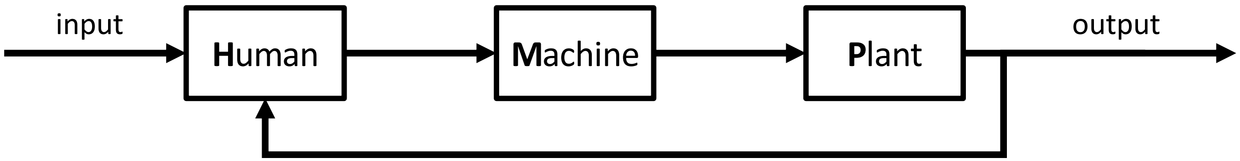
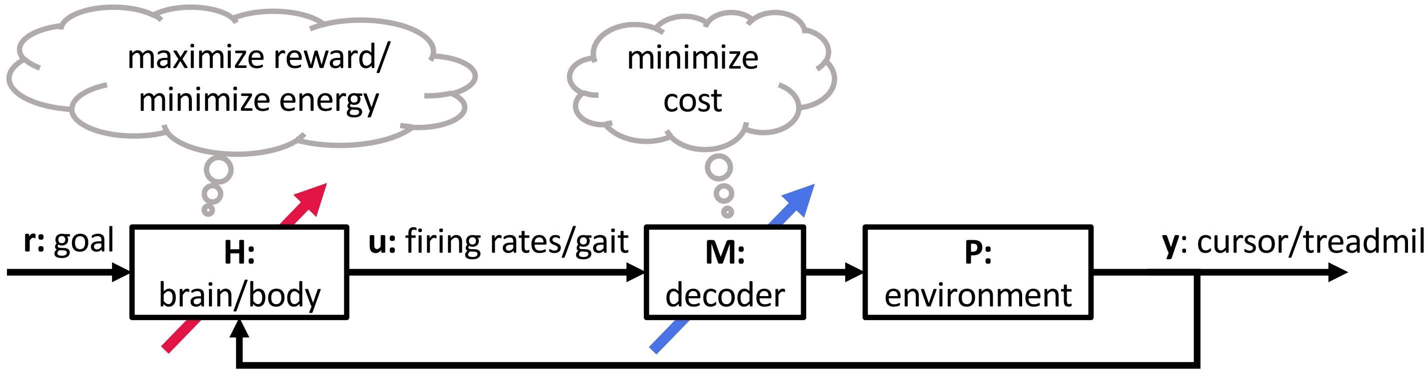
Human/machine interaction is a closed loop

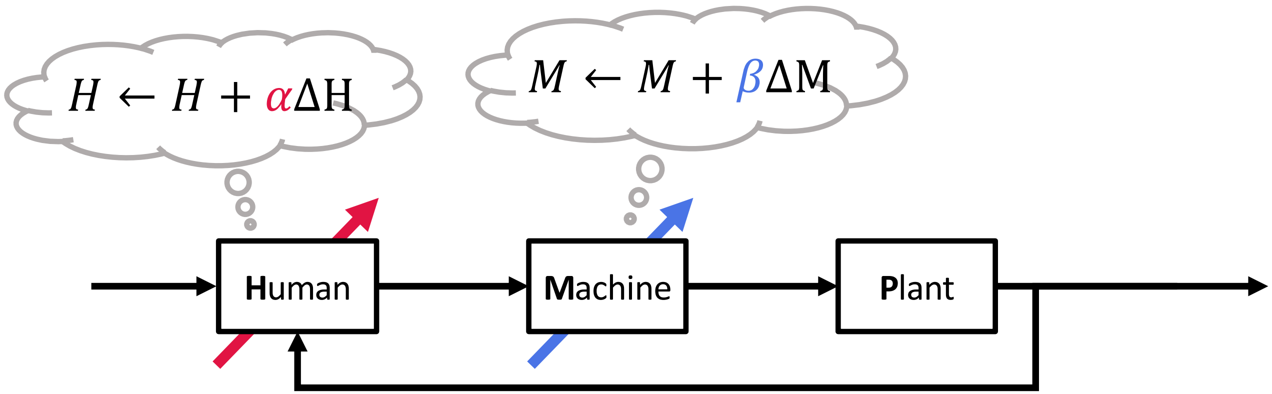


Zhou, Doyle, and Glover. *Robust and optimal control*. New Jersey: Prentice hall, 1996.
Åström and Murray. *Feedback systems.* Princeton university press. 2008.

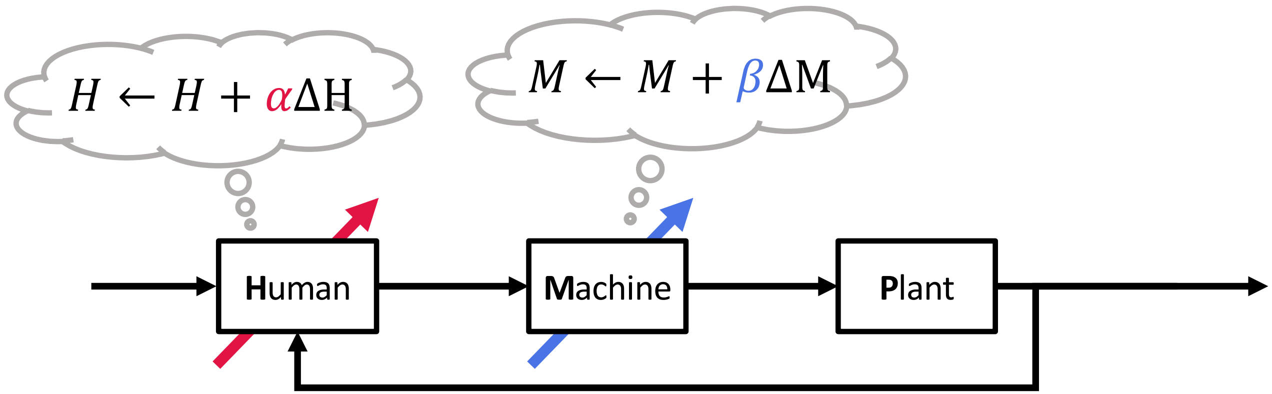
Humans and brains are constantly adapting too



Hypothesis: humans respond analogously



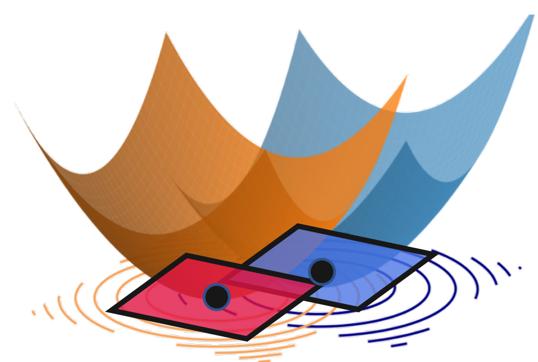
A solution concept of *continuous games*



Discrete-time dynamics

$$H \leftarrow H + \alpha \Delta \mathrm{H}$$
$$M \leftarrow M + \beta \Delta \mathrm{M}$$

Sufficient condition for a *differential Nash equilibrium*



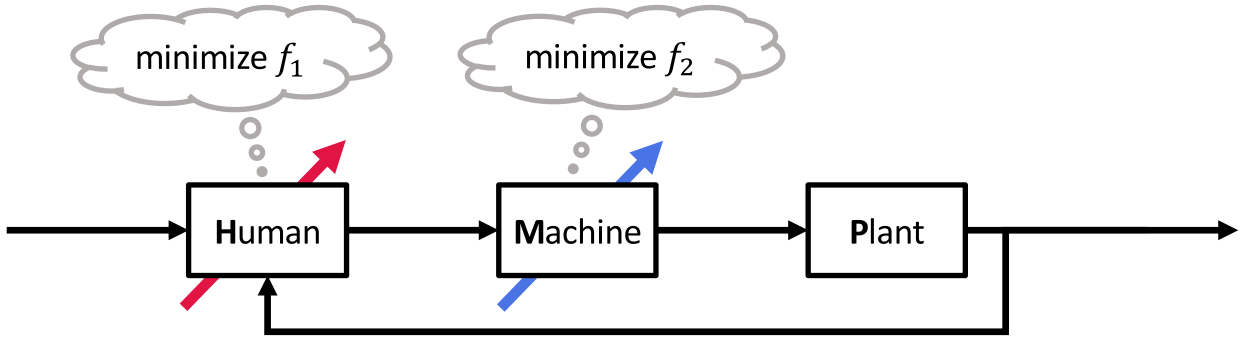
Derivatives:

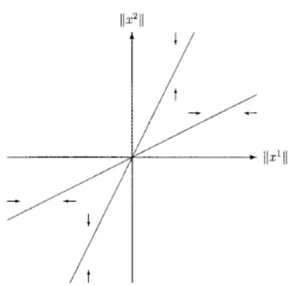
1^{st} order:	flat
2^{nd} order:	positive

Chasnov, Ratliff, Mazumdar, and Burden. "Convergence analysis of gradient-based learning in continuous games." In *Uncertainty in Artificial Intelligence,* PMLR, 2020.

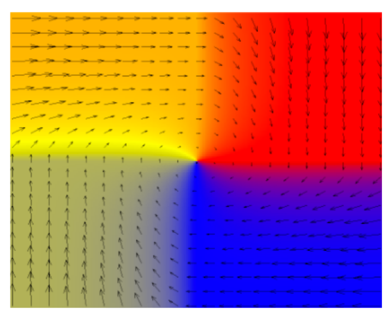
Ratliff, Burden, and Sastry. "Characterization and computation of local Nash equilibria in continuous games." In *Allerton*. IEEE, 2013.

Fundamental issues with learning in games

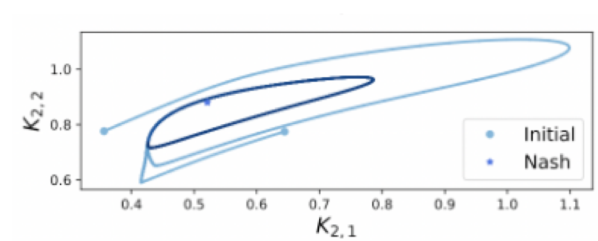




Hart and Mas-Colell. "Uncoupled dynamics do not lead to Nash equilibrium." *American Economic Review*, 2003.



Daskalakis, Goldberg, and Papadimitriou. “The complexity of computing a Nash equilibrium.” Electronic Colloquium on Computational Complexity, 2005.



Mazumdar, Ratliff, Jordan, and Sastry. "Policy-gradient algorithms have no guarantees of convergence in linear quadratic games." In *AAMAS,* 2020.

Theoretical questions on game dynamics

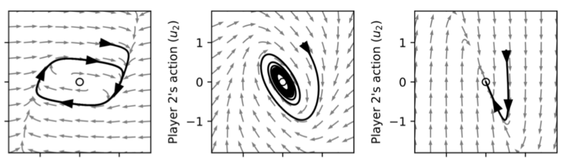
Discrete-time dynamics

$$H_{t+1} \leftarrow H_t + \alpha \Delta \mathrm{H_t}$$
$$M_{t+1} \leftarrow M_t + \beta \Delta \mathrm{M_t}$$

Gradients: $\Delta \mathrm{M_t}, \Delta H_t$

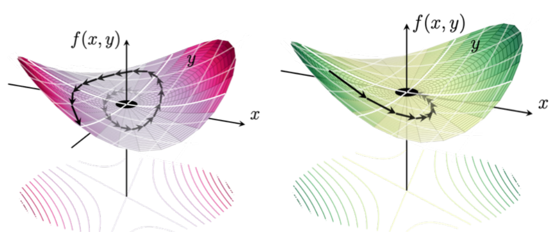
Learning rates: α, β

- **Q1** (*convergence*): as time t $\rightarrow \infty$, does $\Delta H_t \rightarrow 0, \Delta M_t \rightarrow 0$?

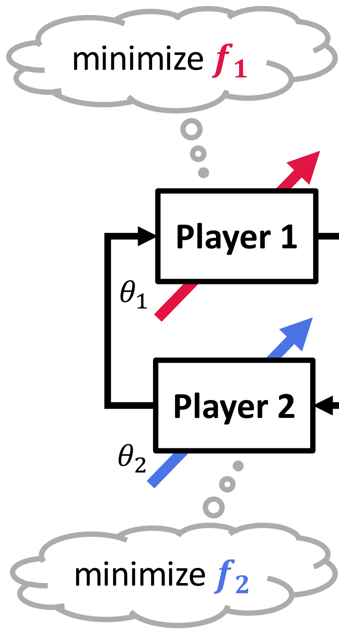


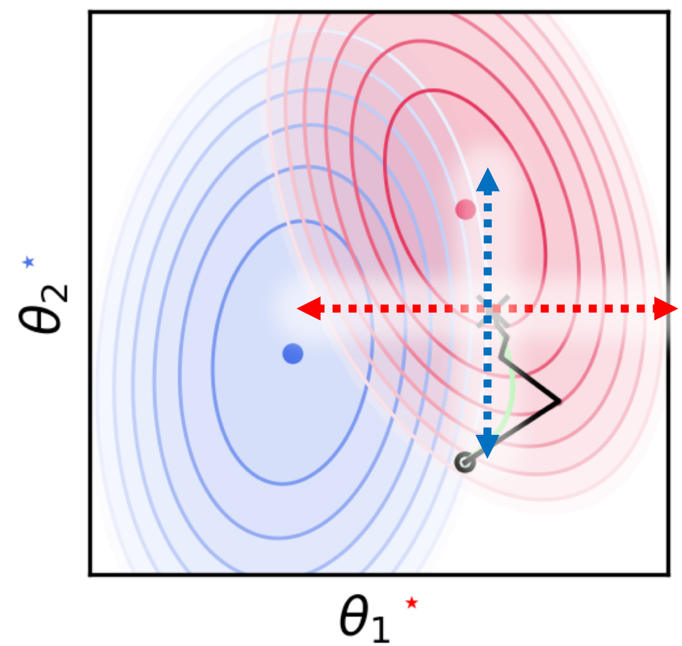
- **Q2** (*equilibrium*): when $\Delta H = 0, \Delta M = 0$, are the costs minimized?

- **Q3** (*learning rates*): as $\tau := \beta / \alpha \rightarrow \infty$, what happens to the equilibrium?



Example: simultaneous gradient descent



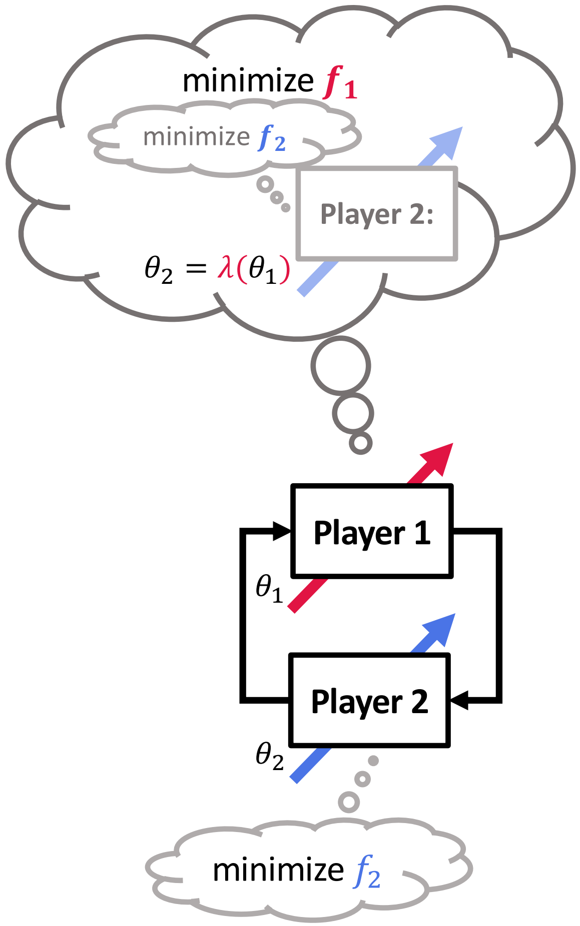


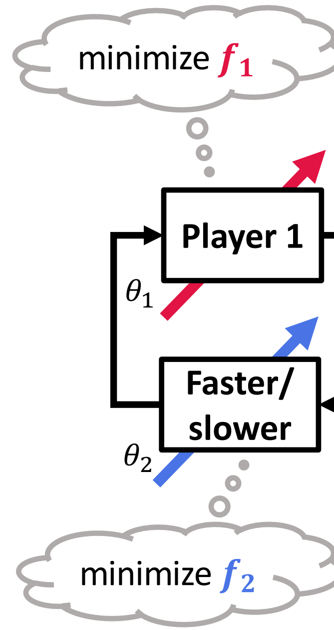
$$\theta_1 \leftarrow \theta_1 - \alpha \frac{\partial}{\partial \theta_1} f_1(\theta_1, \theta_2)$$

$$\theta_2 \leftarrow \theta_2 - \beta \frac{\partial}{\partial \theta_2} f_2(\theta_1, \theta_2)$$

Static webpage:
https://dynam.space/demo/

Modeling each other and time-scale separation





$$\theta_1 \leftarrow \theta_1 - \alpha \frac{d}{d\theta_1} f_1(\theta_1, \lambda(\theta_1))$$

$$\lambda(\theta_1) = \arg\min_{\theta} f_2(\theta_1, \theta)$$

$$\theta_2 \leftarrow \theta_2 - \beta \frac{\partial}{\partial\theta_2} f_2(\theta_1, \theta_2)$$

$$\theta_1 \leftarrow \theta_1 - \alpha \frac{\partial}{\partial\theta_1} f_1(\theta_1, \theta_2)$$

$$\theta_2 \leftarrow \theta_2 - \alpha\tau \frac{d}{d\theta_2} f_2(\theta_1, \theta_2)$$

Fiez, Chasnov, and Ratliff. "Implicit learning dynamics in Stackelberg games: Equilibria characterization, convergence analysis, and empirical study." In *ICML*. PMLR, 2020.

Chasnov, Calderone, Açıkmeşe, Burden, and Ratliff. "Stability of Gradient Learning Dynamics in Continuous Games: Scalar Action Spaces." In *CDC*. IEEE, 2020.

Game Jacobian of the learning dynamics

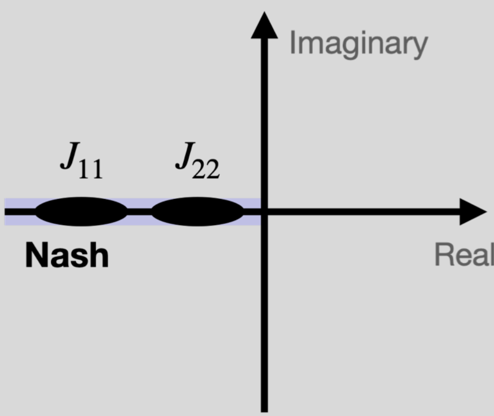
Linearized dynamics about an equilibrium:

$$\dot{\theta} = J(\tilde{\theta})\theta, \qquad g(\tilde{\theta}) = 0, \qquad J(\tilde{\theta}) = \begin{bmatrix} J_{11} & \boxed{J_{12}} \\ \boxed{J_{21}} & J_{22} \end{bmatrix}$$

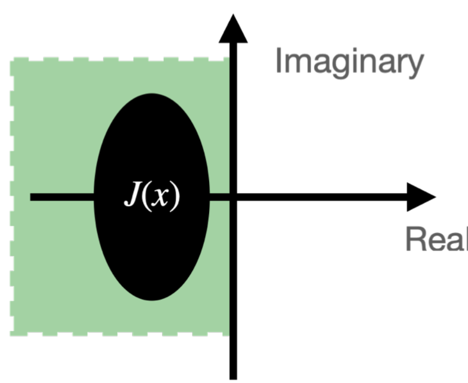
1st order derivative

2nd order derivatives

***Individual* terms:**
Nash



***Interaction* terms:**
stability



***Decomposition*:**
class

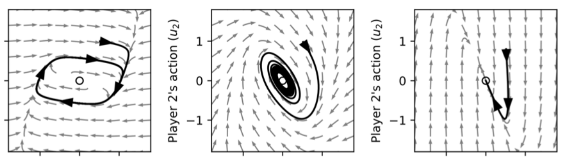
$$J = \begin{bmatrix} A & P \\ P^\top & D \end{bmatrix} + \begin{bmatrix} 0 & Z \\ -Z^\top & 0 \end{bmatrix}$$

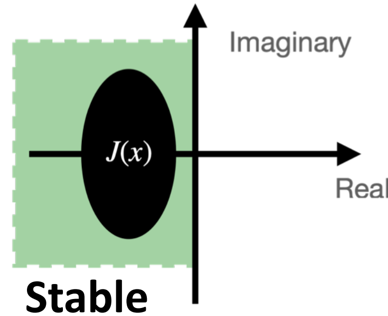
Chasnov, Calderone, Açıkmeşe, Burden, and Ratliff. "Stability of Gradient Learning Dynamics in Continuous Games: Vector Action Spaces." *arXiv: 2011.05562*, 2021.

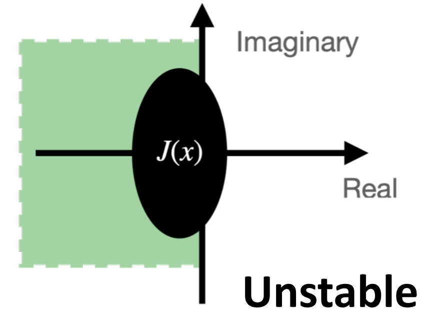
arXiv: 2011.05562

Theoretical results using the game Jacobian

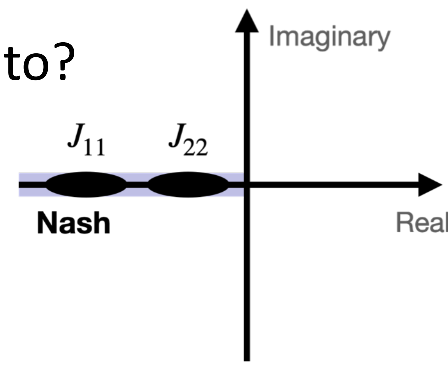
- **Q1:** does learning converge?

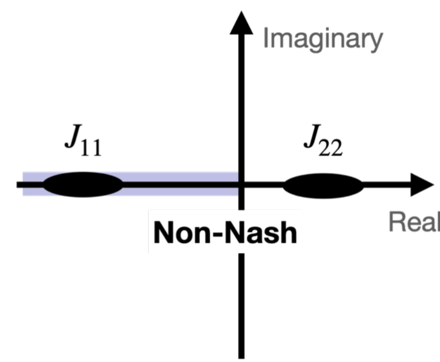




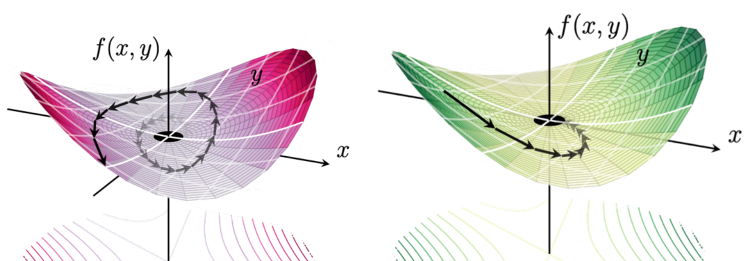


- **Q2:** what does it converge to?





- **Q3**: how do learning rates affect the dynamics?



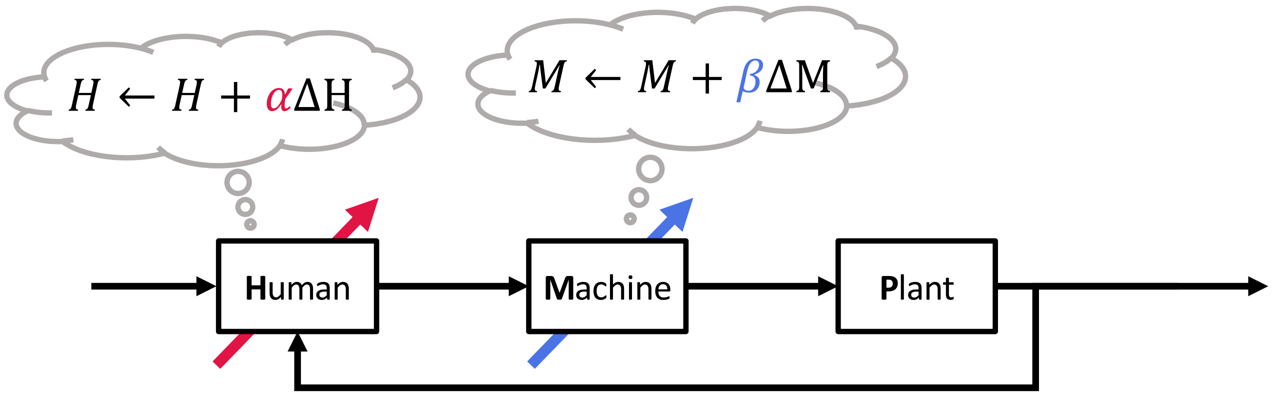
$$J = \begin{bmatrix} J_{11} & J_{12} \\ \tau J_{21} & \tau J_{22} \end{bmatrix}$$

$$S_1 = J_{11} - J_{12} J_{22}^{-1} J_{21}$$

Chasnov, Calderone, Açıkmeşe, Burden, and Ratliff. "Stability of Gradient Learning Dynamics in Continuous Games: Vector Action Spaces." *arXiv: 2011.05562*, 2021.

arXiv: 2011.05562

Continuous time approximation



Discrete-time dynamics

$$H \leftarrow H + \alpha \Delta \mathrm{H}$$
$$M \leftarrow M + \beta \Delta \mathrm{M}$$

for small α, β

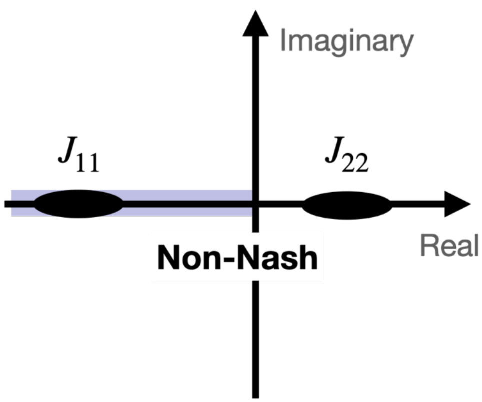
$$\Longrightarrow$$

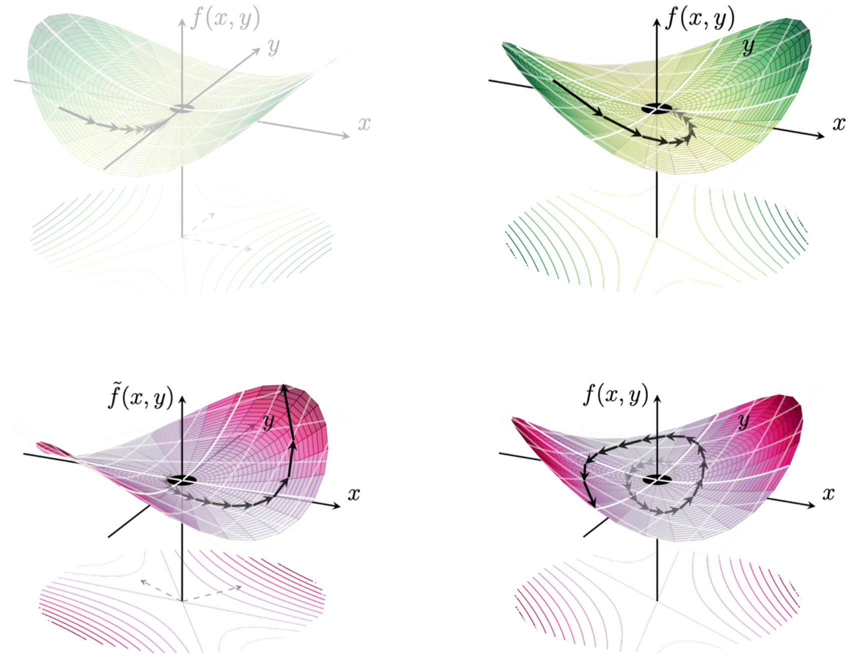
Continuous-time dynamics

$$\dot{H} = \Delta \mathrm{H}$$
$$\dot{M} = \tau\, \Delta \mathrm{M}$$

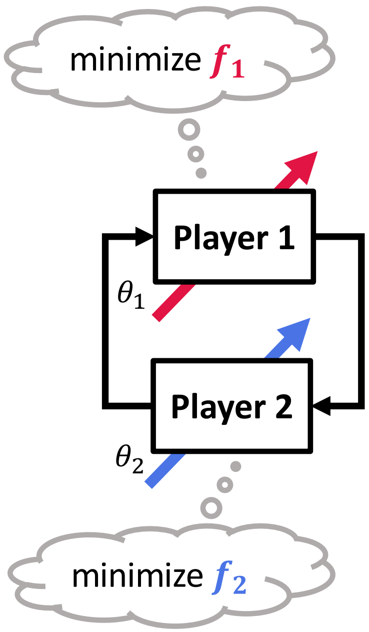
Chasnov, Ratliff, Mazumdar, and Burden. "Convergence analysis of gradient-based learning in continuous games." In *Uncertainty in Artificial Intelligence,* PMLR, 2020.

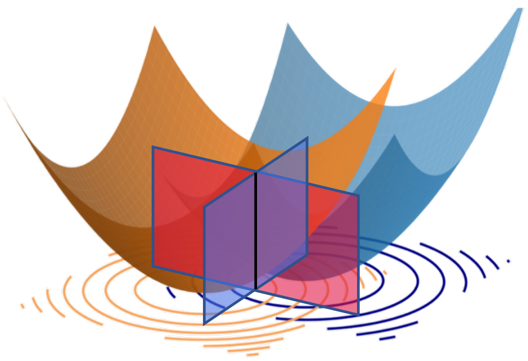
Non-Nash equilibria can be stable





Freezing the parameters of one player





Timescale separation is related to a leader-follower structure

The *Schur complement*:
timescale and order of play

$$\begin{bmatrix} \boxed{J_{11}} & \boxed{J_{12}} \\ \boxed{J_{21}} & \boxed{J_{22}} \end{bmatrix}$$

$$S_1 = \boxed{J_{11}} - \boxed{J_{12}}\,\boxed{J_{22}^{-1}}\,\boxed{J_{21}}$$

$$\theta_1^+ = \theta_1 - \gamma_1 \tfrac{d}{d\theta} f_1(\theta_1, \lambda(\theta_1))$$

$$\lambda(\theta_1) = \arg\min_{\theta} f_2(\theta_1, \theta)$$

Fiez, Chasnov, and Ratliff. "Implicit learning dynamics in Stackelberg games: Equilibria characterization, convergence analysis, and empirical study." In *ICML.* PMLR, 2020.

Chasnov, Burden, and Ratliff. ”Gradient-based learning with a finite timescale separation in continuous games " 2021. In progress.

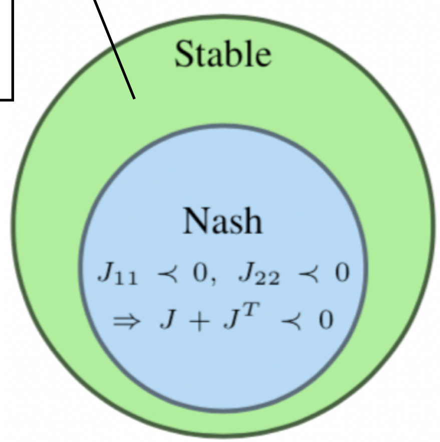
Implications of different classes of games

- Zero-sum game

$$\min_x \max_y R(x,y)$$

$$\begin{bmatrix} A & B \\ -\tau B^\top & \tau D \end{bmatrix}$$

Spurious attractors (at least one agents is not optimal)

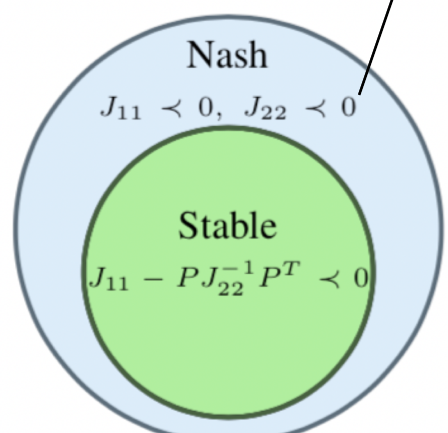


Potential game

$$\max_{x,y} \Phi(x,y)$$

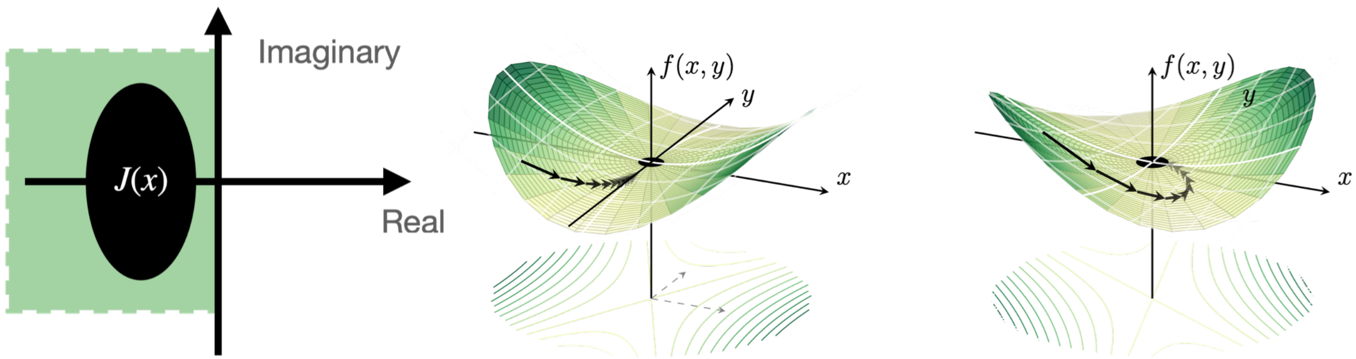
$$\begin{bmatrix} A & B \\ \tau B^\top & \tau D \end{bmatrix}$$

Not computable (using gradient dynamics)

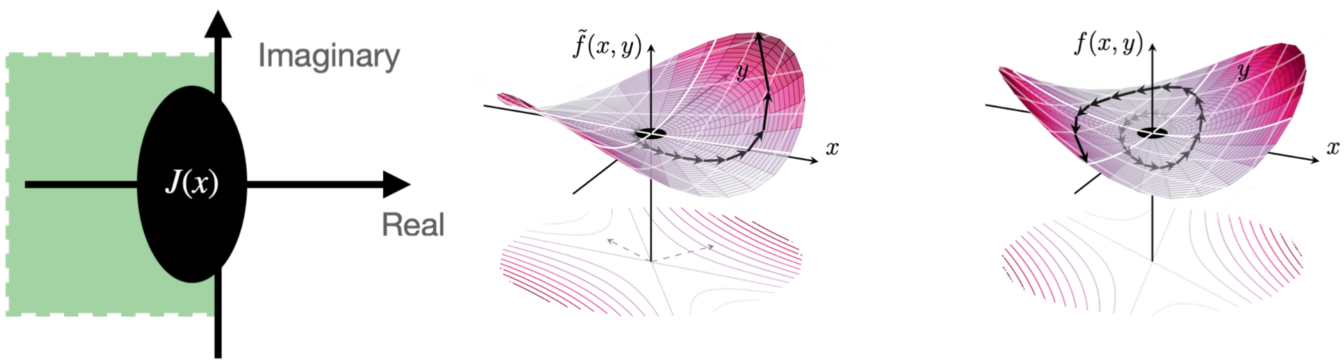


Stability of game dynamics near fixed points

Stable:



Unstable:



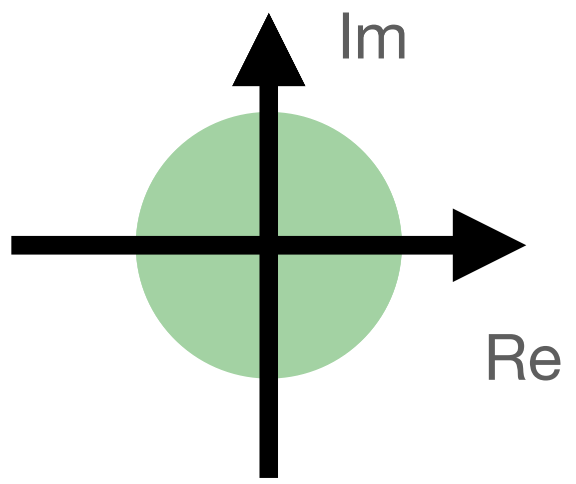
Two local equilibrium concepts

Definitions and known results

- Gradients: $g = [g_1, \ldots, g_n]$. Fixed point x^*: $g(x^*) = 0$. Jacobian: $J(x^*) = -Dg(x^*)$
- Locally exponentially stable equilibrium x^* (Khalil 2002)

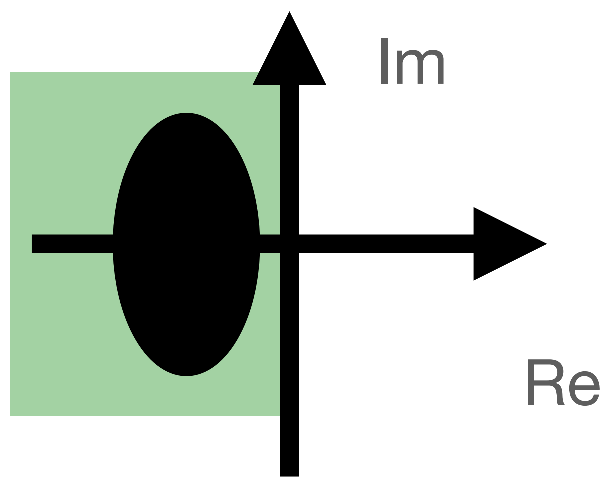
$$x(t+1) = x(t) - \gamma g(x(t))$$

$$\underbrace{\rho(I + \gamma J(x^*))}_{\textit{spectral radius}} < 1$$



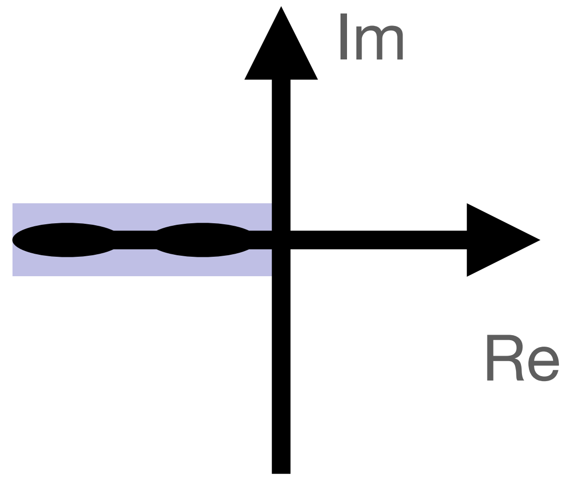
$$\dot{x} = g(x)$$

$$\boxed{\underbrace{\sigma(J(x^*))}_{\textit{spectrum}} \subset \mathbf{C}^\circ_-}$$

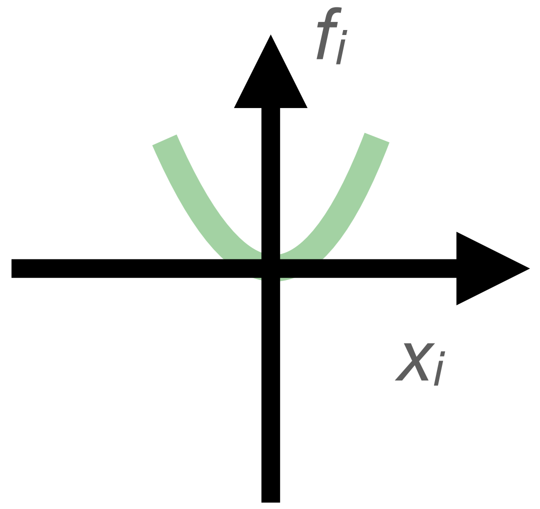


- Differential Nash equilibrium $x^* = (x_i^*, x_{-i}^*)$ (Ratliff, Burden, Sastry, 2014)

$$\boxed{\underbrace{D_i g_i(x_i^*, x_{-i}^*)}_{\textit{individual Hessian}} \succ 0,}$$



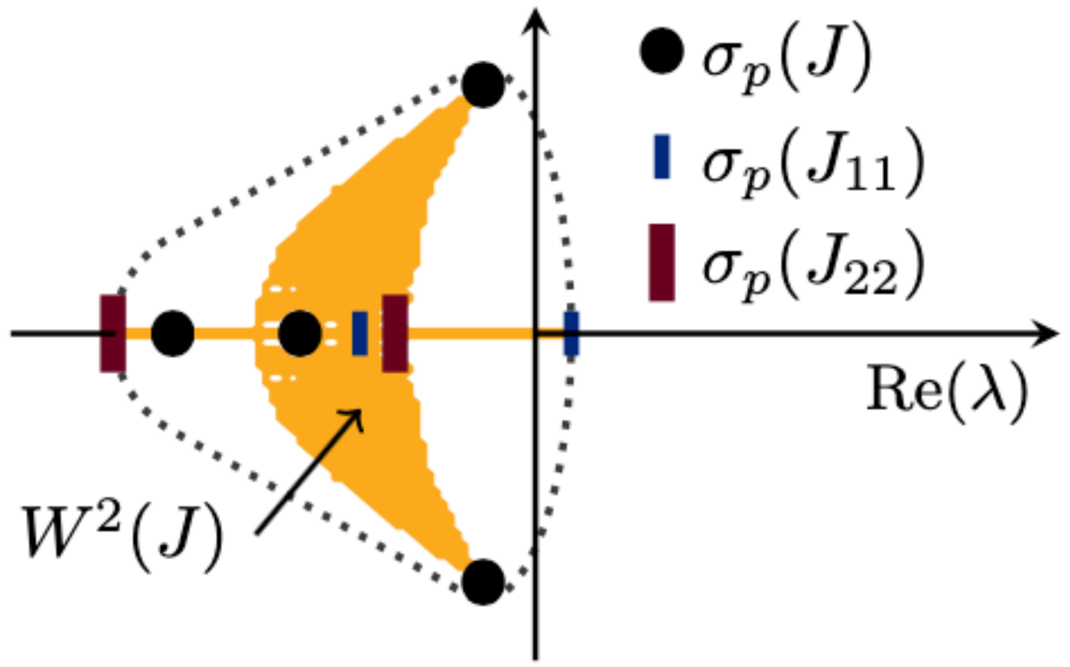
$$\underbrace{f_i(x_i^*, x_{-i}^*) < f_i(x_i, x_{-i}^*)}_{\textit{strict local Nash}}, \quad x_i \in U_i \backslash \{x_i^*\}, \quad \forall i$$

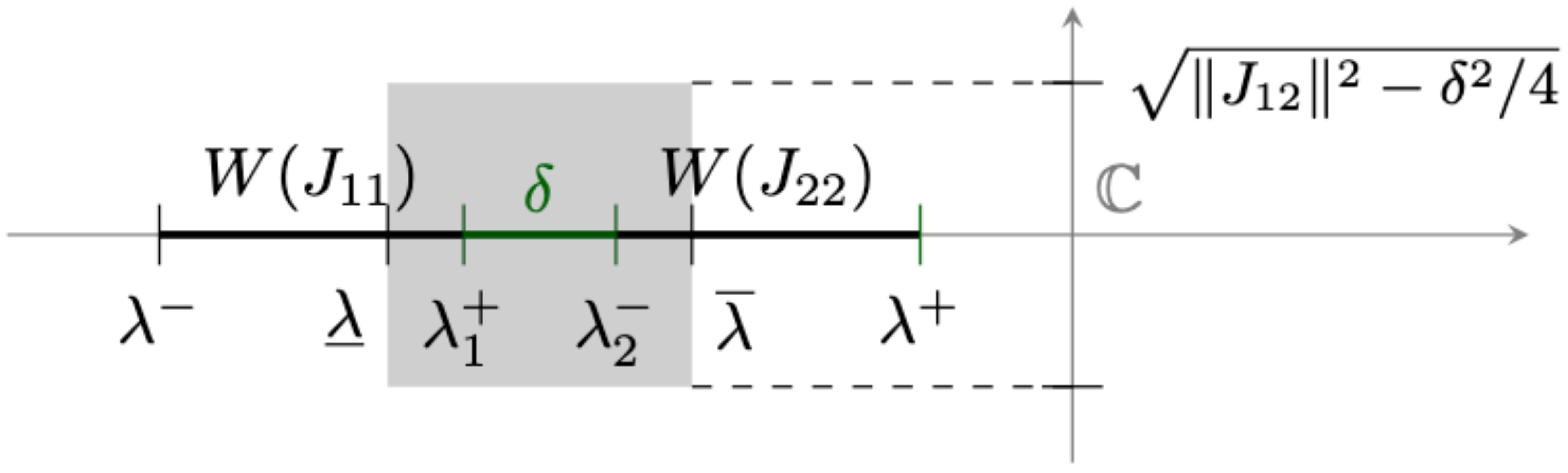


Spectrum of zero-sum and potential games

Results: vector actions

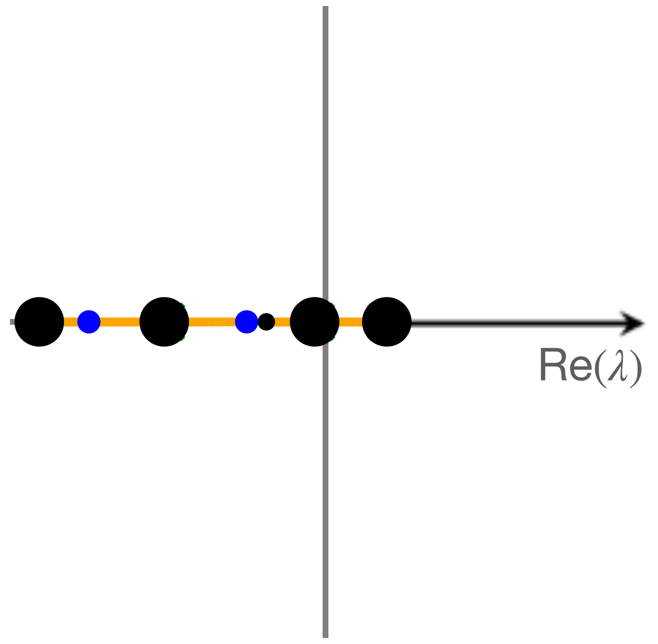
- Zero-sum game $J(x) = \begin{bmatrix} J_{11} & -Z \\ Z^\top & J_{22} \end{bmatrix}$

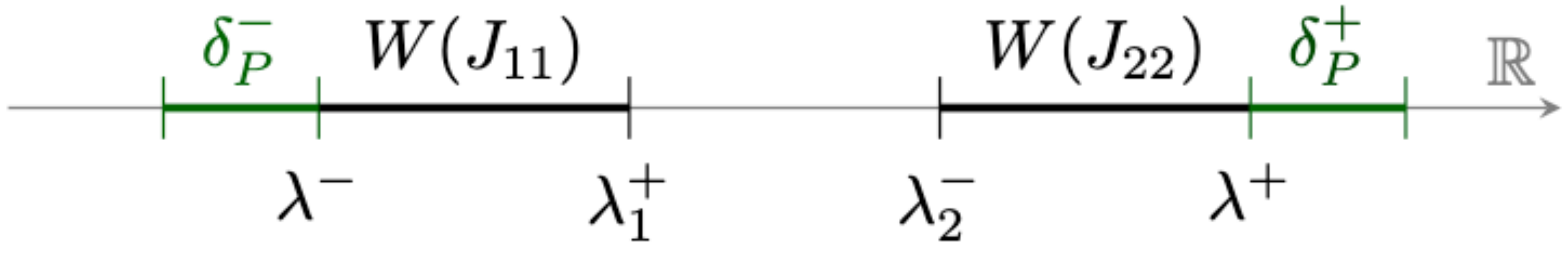




(a) Zero-sum game where $\delta = \lambda_2^- - \lambda_1^+ > 0$ and $\|J_{12}\| > \delta/2$

- Potential game $J(x) = \begin{bmatrix} J_{11} & P \\ P^\top & J_{22} \end{bmatrix}$





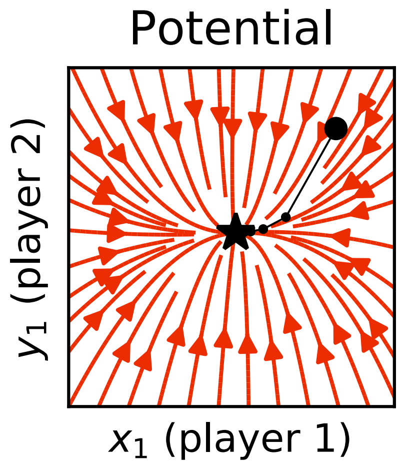
(b) Potential game where $\lambda_2^- - \lambda_1^+ > 0$.

Decomposition of the game Jacobian

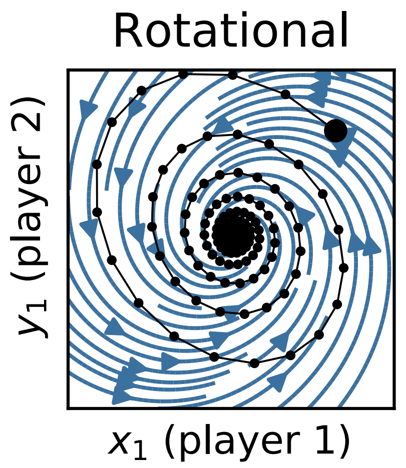
Game types

- $$J(x) = \overbrace{\underbrace{\begin{bmatrix} J_{11} & 0 \\ 0 & J_{22} \end{bmatrix}}_{individual} + \begin{bmatrix} 0 & P \\ P^\top & 0 \end{bmatrix}}^{symmetric} + \overbrace{\begin{bmatrix} 0 & -Z \\ Z^\top & 0 \end{bmatrix}}^{anti-symmetric}$$
  (the last two matrices together: $interaction$)

- Symmetric: $$\frac{1}{2}\left(J + J^\top\right) = \begin{bmatrix} J_{11} & \frac{1}{2}\left(J_{12} + J_{21}^\top\right) \\ \frac{1}{2}\left(J_{21} + J_{12}^\top\right) & J_{22} \end{bmatrix} = \begin{bmatrix} J_{11} & P \\ P^\top & J_{22} \end{bmatrix}$$



- Anti-symmetric: $$\frac{1}{2}\left(J - J^\top\right) = \begin{bmatrix} 0 & \frac{1}{2}\left(J_{12} - J_{21}^\top\right) \\ \frac{1}{2}\left(J_{21} - J_{12}^\top\right) & 0 \end{bmatrix} = \begin{bmatrix} 0 & -Z \\ Z^\top & 0 \end{bmatrix}$$



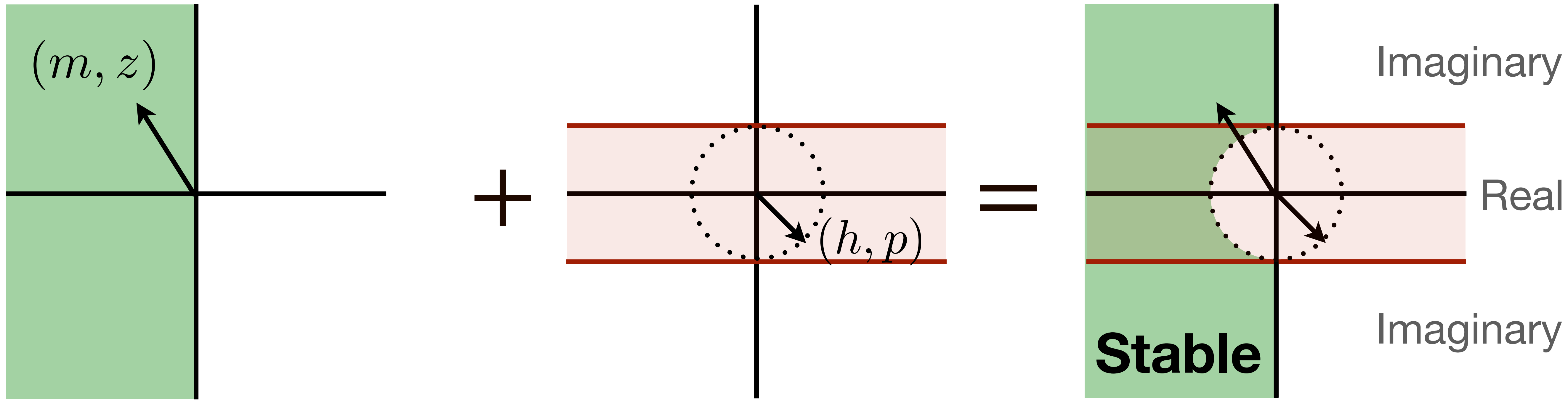
Stability of general-sum games

Results: scalar actions

$$J(x) = \underbrace{\begin{bmatrix} m & -z \\ z & m \end{bmatrix}}_{\textit{complex numbers}} + \underbrace{\begin{bmatrix} h & p \\ p & -h \end{bmatrix}}_{\textit{circle}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$m = \tfrac{1}{2}(a+d) \qquad h = \tfrac{1}{2}(a-d)$$

$$p = \tfrac{1}{2}(b+c) \qquad z = \tfrac{1}{2}(c-b)$$



$$y = m + zi$$

Decomposition of the game Jacobian

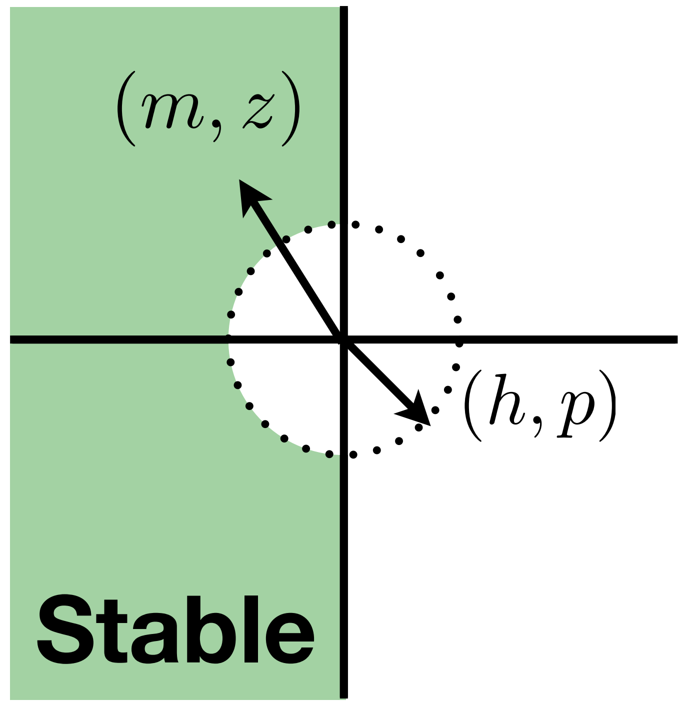
Game types

- $$J(x) = \overbrace{\underbrace{\begin{bmatrix} J_{11} & 0 \\ 0 & J_{22} \end{bmatrix}}_{individual} + \begin{bmatrix} 0 & P \\ P^\top & 0 \end{bmatrix}}^{symmetric} + \underbrace{\hphantom{\begin{bmatrix} 0 & P \\ P^\top & 0 \end{bmatrix} +} \overbrace{\begin{bmatrix} 0 & -Z \\ Z^\top & 0 \end{bmatrix}}^{anti-symmetric}}_{interaction}$$

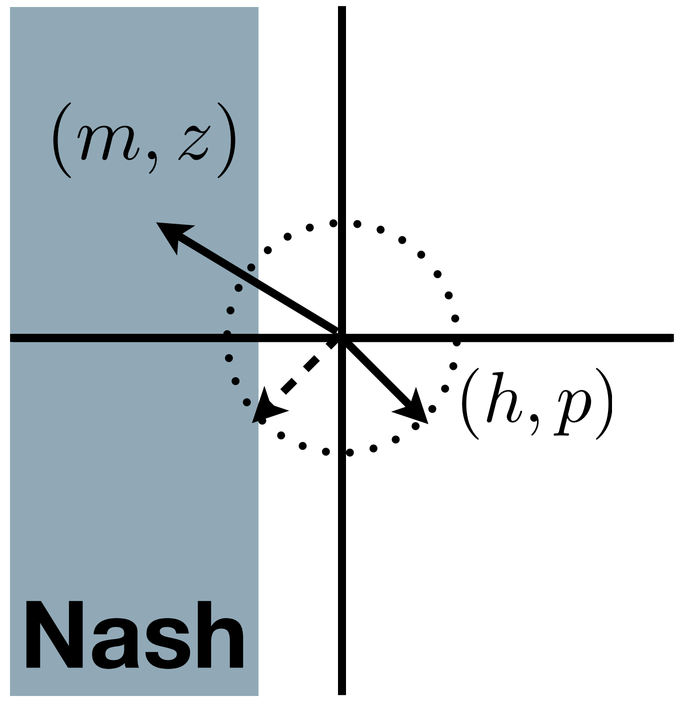
Necessary and sufficient conditions

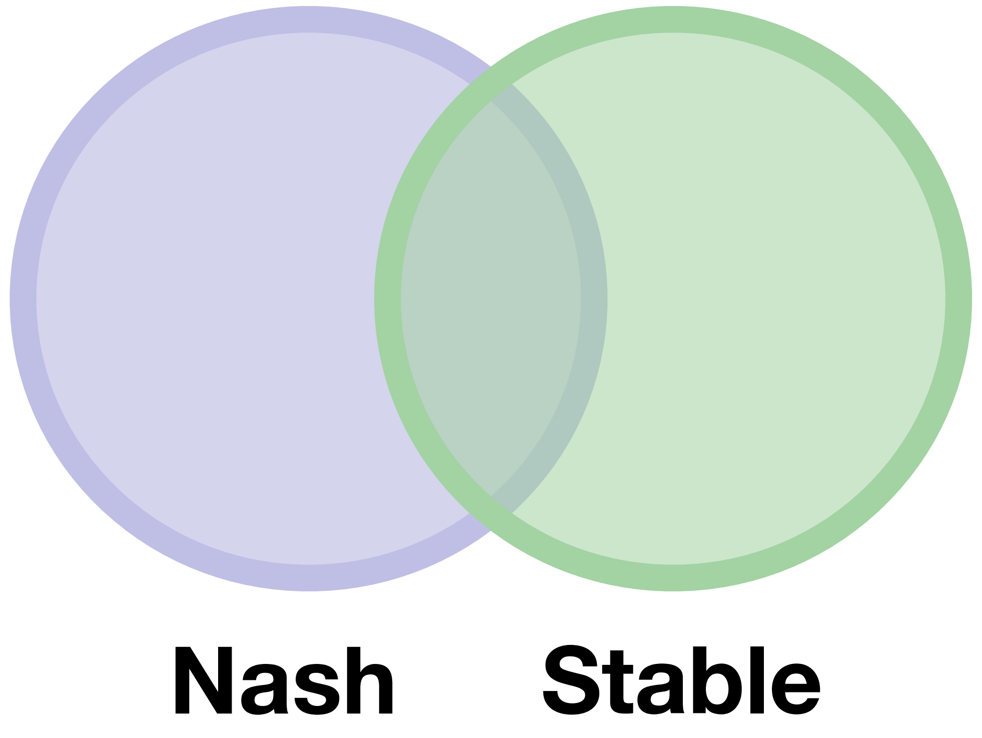
Results: scalar actions

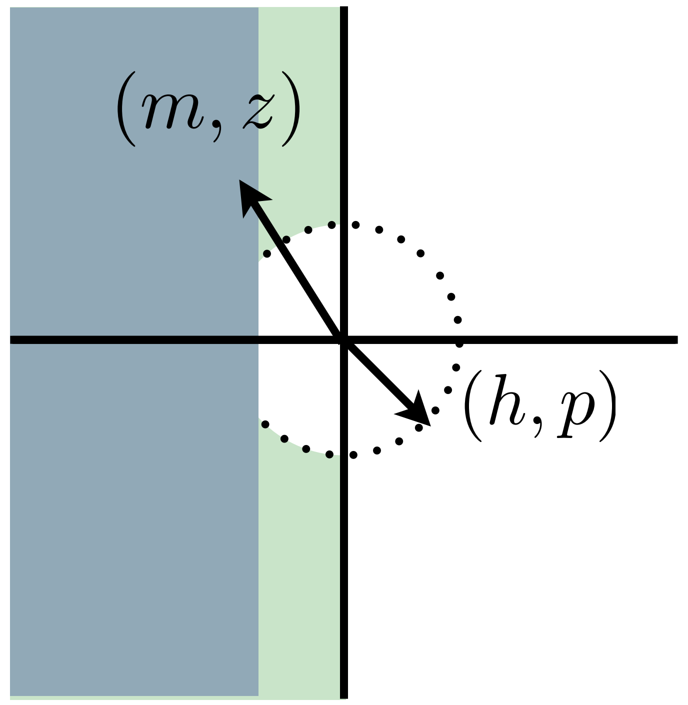
- Two-by-two matrix decomposition: $$J(x) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} m+h & p-z \\ p+z & m-h \end{bmatrix}$$
- Stability: $\mathrm{tr}(J) < 0$, $\det(J) > 0$



- Nash: $\underbrace{a, d}_{Hessians} < 0 \iff m < -|h|$



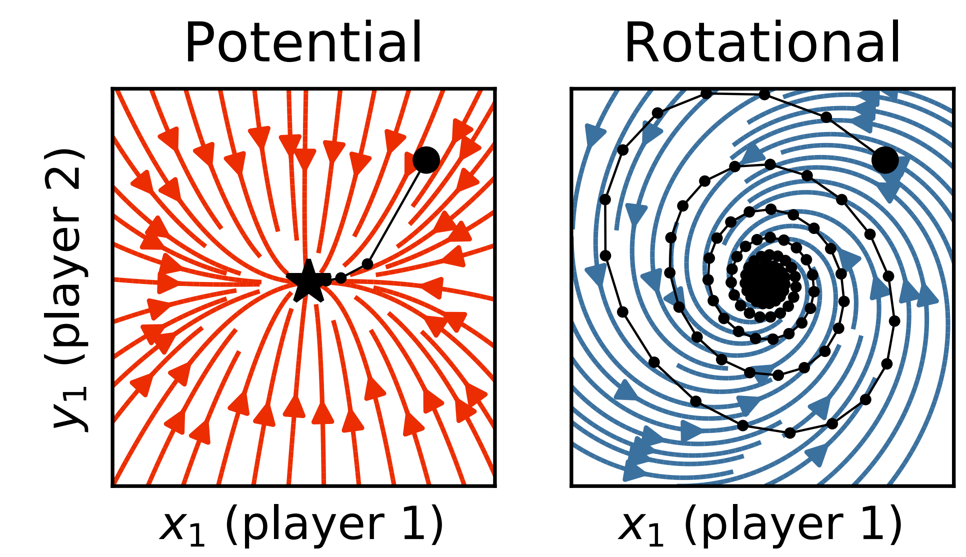


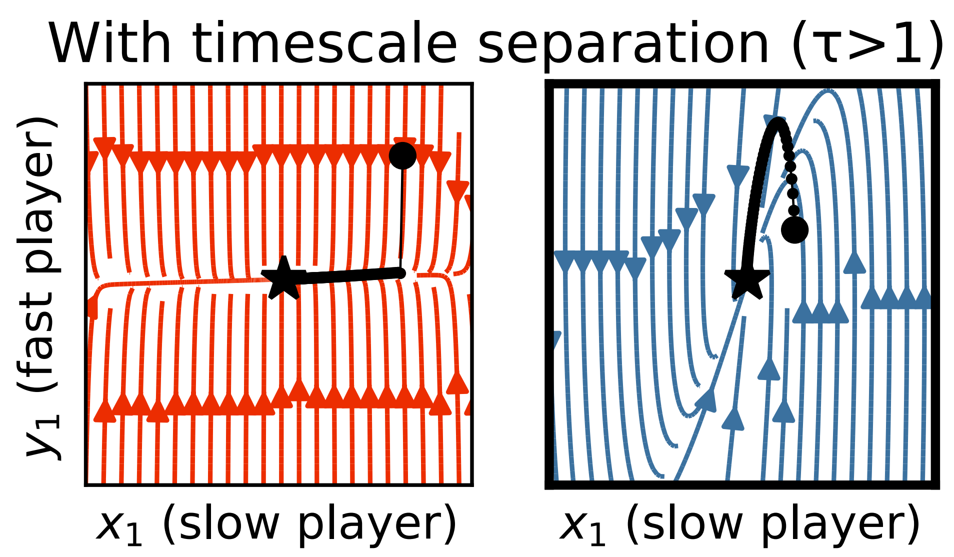


Potential and rotational vector fields

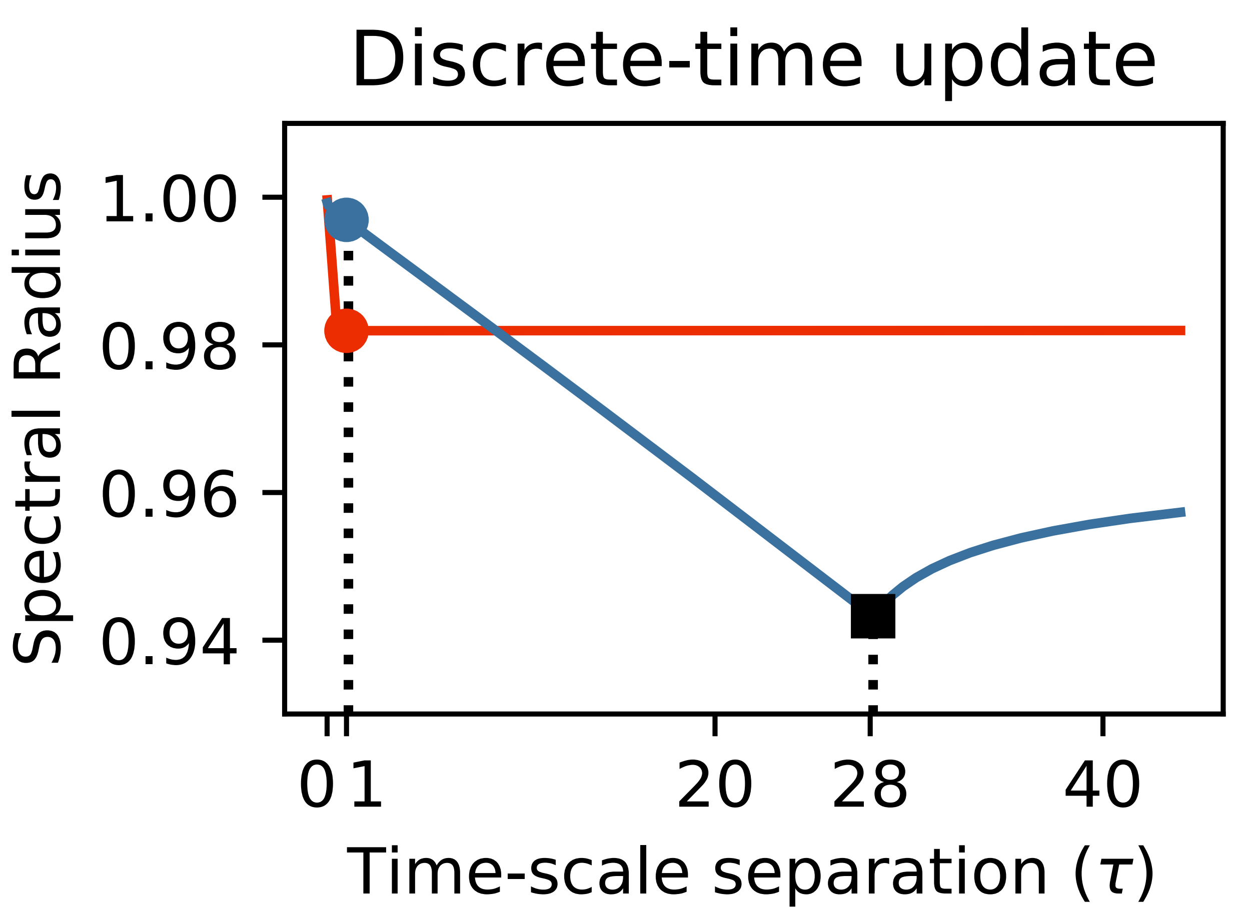
Numerical Experiments

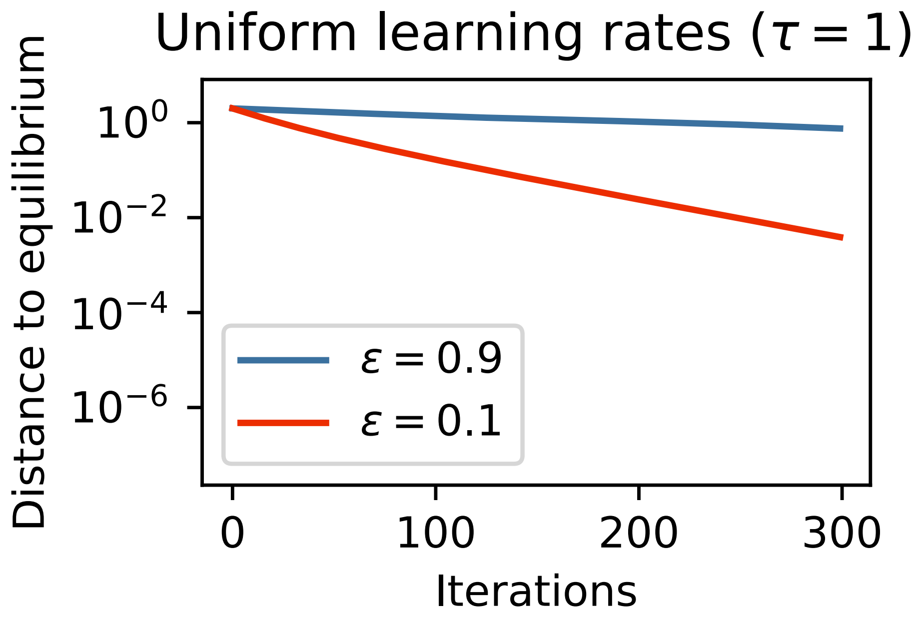
- Timescale separation $(\tau = \frac{\gamma_2}{\gamma_1} > 0)$: $\quad J(x) = \begin{bmatrix} J_{11} & J_{12} \\ \tau J_{21} & \tau J_{22} \end{bmatrix}$
- Symmetric/anti-symmetric Jacobian (zero-sum): $\quad J = (1-\varepsilon)S + \varepsilon A, \quad S = S^\top, A = -A^\top$

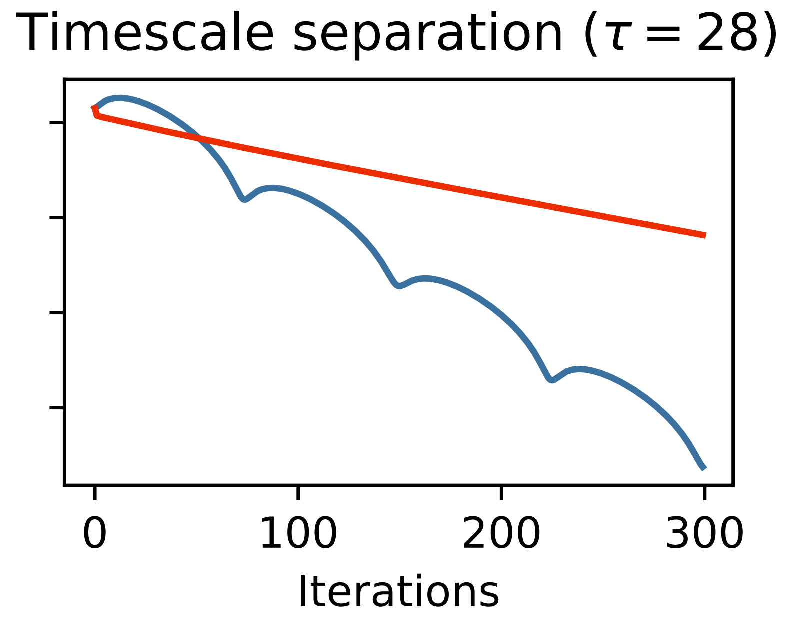




$$\rho(I + \gamma J(x^*))$$







$$\underbrace{\sigma\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)}_{spectrum} \subset \mathbf{C}_-^\circ \quad \Longleftrightarrow \quad \begin{aligned} \mathsf{Tr}(J) &= 2m < 0 \\ \det(J) &= m^2 + z^2 - p^2 - h^2 > 0 \end{aligned}$$

Types of games

Results: scalar actions

- Potential games (p)

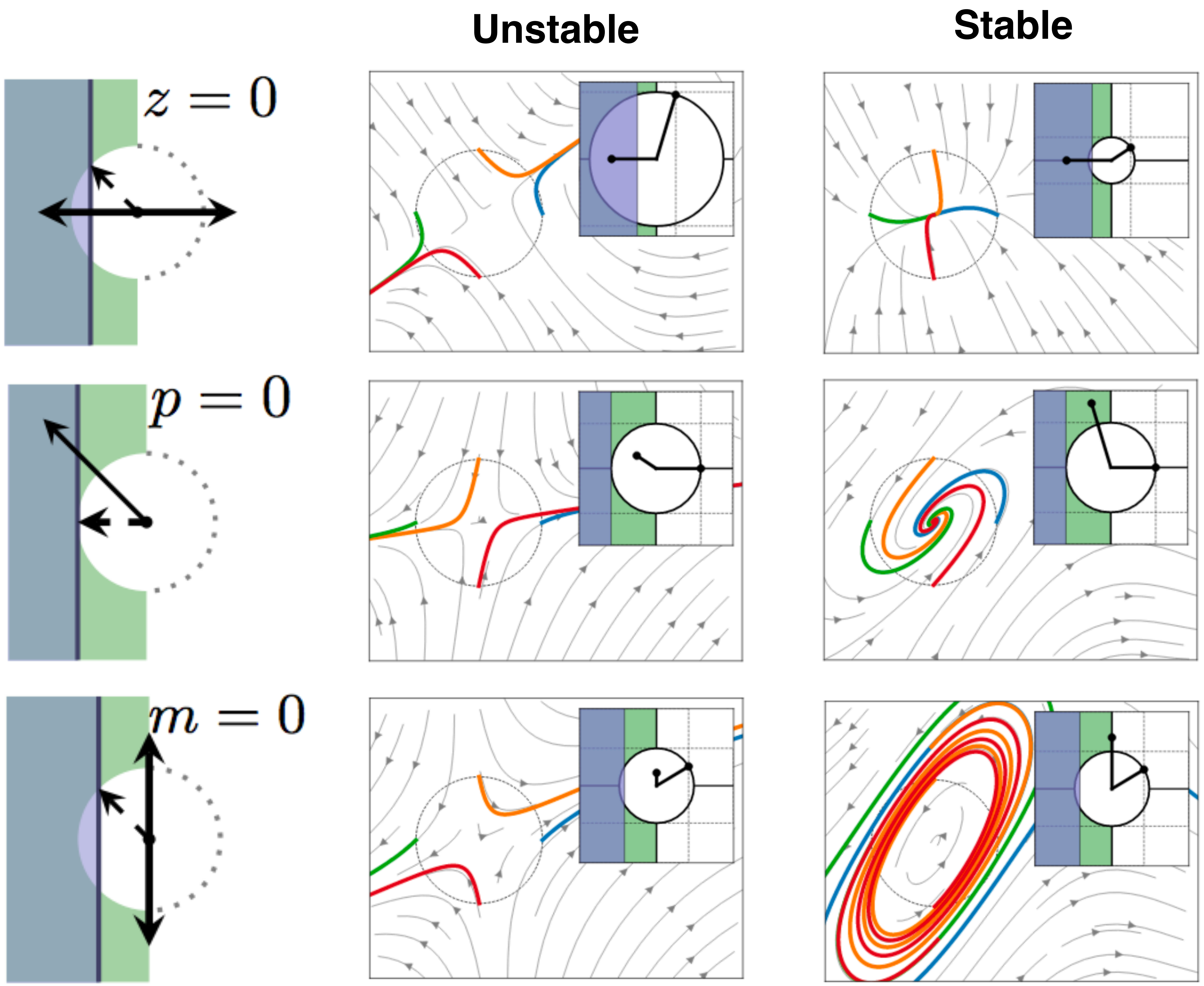
$$J = \begin{bmatrix} m+h & \mathbf{p} \\ \mathbf{p} & m-h \end{bmatrix}$$

- Zero-sum games (z)

$$J = \begin{bmatrix} m+h & \mathbf{z} \\ \mathbf{z} & m-h \end{bmatrix}$$

- Hamiltonian games (h)

$$J = \begin{bmatrix} h & p-\mathbf{z} \\ p+\mathbf{z} & -h \end{bmatrix}$$

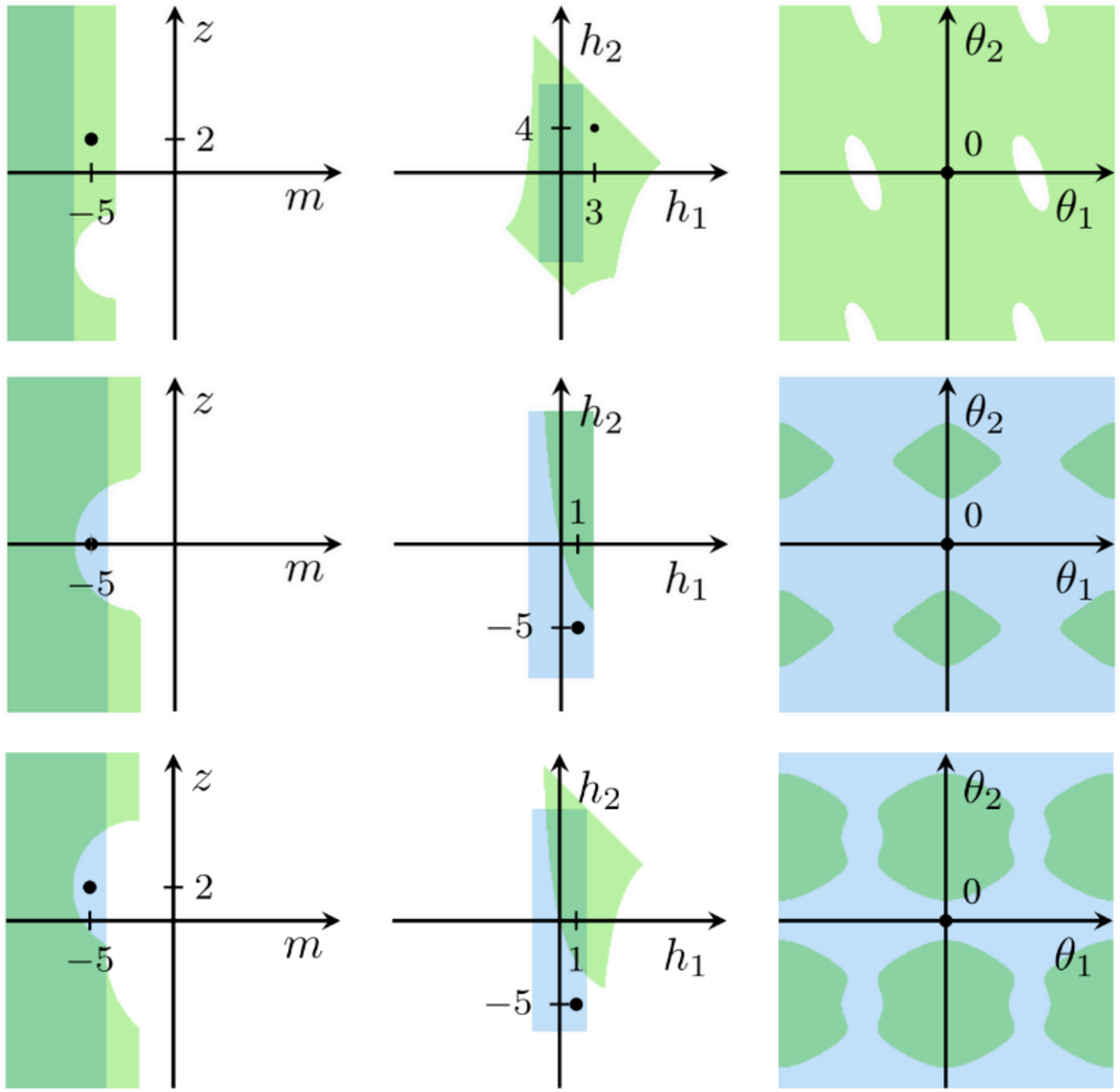


(marginally stable)

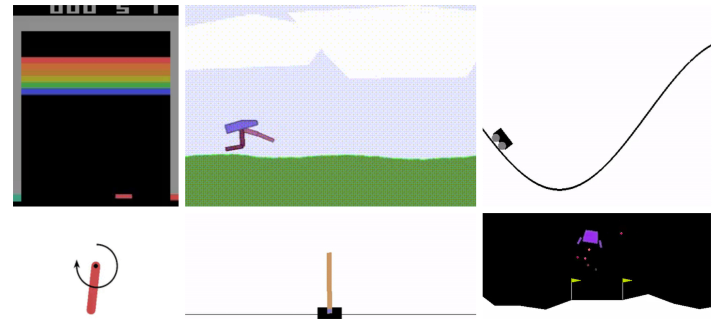
Overlapping regions of Nash and stability

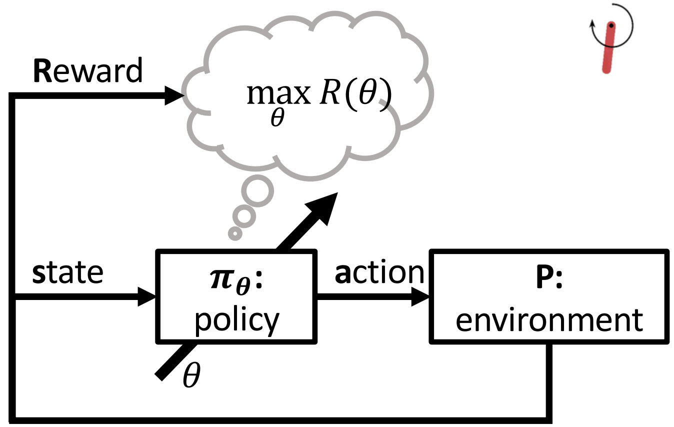
Numerical Experiments

- Zero-sum game
- Potential game
- General game



Reinforcement learning is optimization





Policy gradient: $\theta \leftarrow \theta + \alpha\Delta\theta$ $\tau = (s_0, a_0, s_1, a_1, \ldots, s_T, a_T)$

$$\Delta\theta = \mathrm{E}_{\tau\sim\pi_\theta}\left[\int_t \log \pi_\theta(a_t|s_t) R(\tau) dt\right]$$