The Invisible Hand of Stability

When Collective Rationality Emerges without Coordination

Geelon So | ETH Zürich | Cambridge Math+ML Seminar

Recent interest in multi-agent systems

Multi-Agent Systems for ML Problems

"We found that **a multi-agent system** with Claude Opus 4 as the lead agent and Claude Sonnet 4 subagents **outperformed single-agent** Claude Opus 4 **by 90.2% on our internal research eval.**"

***How we built our multi-agent research system*, Anthropic (2025)**

Multi-Agent Systems for ML Problems

"**Traditional AI often relies on a single, general-purpose AI that struggles with complex workflows. Multi-agent systems** change the game by orchestrating specialized agents, each focused on a specific task, to **automate more complex problems.**"

Multiagent Systems: A New Era in AI-Driven Enterprise Automation, **Gartner (2026)**

Multi-Agent Systems for ML Problems

Published as a conference paper at COLM 2025

MALT: Improving Reasoning with Multi-Agent LLM Training

Sumeet Ramesh Motwani[1], **Chandler Smith**[2], **Rocktim Jyoti Das**[3], **Rafael Rafailov**[4], **Ivan Laptev**[3], **Philip Torr**[1], **Fabio Pizzati**[3], **Ronald Clark**[1], **Christian Schroeder de Witt**[1]

[1]University of Oxford [2]Cooperative AI Foundation [3]MBZUAI [4]Stanford University

Correspondence: sumeet.motwani@eng.ox.ac.uk, cs@robots.ox.ac.uk
Project page: https://multiagentllmtraining.com/

Abstract

Large Language Models (LLMs) often produce answers with a single chain-of-thought, which restricts their ability to explore reasoning paths or self-correct flawed outputs in complex tasks. In this paper, we introduce MALT (Multi-Agent LLM Training), a novel post-training strategy that divides the reasoning process into generation, verification, and refinement steps using a sequential pipeline of heterogeneous agents. During data generation, each agent is repeatedly sampled to form a multi-agent search tree, where final outputs are graded against ground-truth data. We then apply value iteration to propagate reward signals back to each role-conditioned model, automatically producing multi-agent post-training data without human or teacher-model supervision. Our off-policy approach allows each agent to specialize by learning from correct and incorrect trajectories, ultimately improving the end-to-end reasoning chain. On MATH, GSM8K, and CSQA, MALT surpasses the same baseline LLM with relative improvements of 15.66%, 7.42%, and 9.40%. It also generalizes to more challenging benchmarks, marking an early advance in multi-agent cooperative training.

Economy of Minds: Emerging Multi-Agent Intelligence with Economic Interactions

Zhenting Qi[*] **Huangyuan Su**[*] **Ao Qu** **Chenyu Wang**
Yu Yao **Han Zheng** **Kushal Chattopadhyay** **Guowei Xu**
Zihan Wang **Weirui Ye** **Vijay Janapa Reddi** **Ju Li** **Paul Pu Liang**
Himabindu Lakkaraju **Sham Kakade** **Yilun Du**[*]

Harvard MIT 2077AI Kempner Institute

*Corresponding to: zhentingqi@g.harvard.edu, ydu@seas.harvard.edu

GitHub Project Page

Abstract

How can a population of agents self-orchestrate and self-adapt into stronger collective intelligence without centralized control? Inspired by Friedrich Hayek's economic theory of decentralized coordination in markets, we study this question through an *agent economy* in which agents compete via auctions for the right to act, exchange payments, and accumulate wealth from environmental rewards. These simple local incentives induce decentralized credit assignment, driving global planning without explicit orchestration or communication protocols. The population evolves through economic selection: effective agents accumulate wealth and are mutated via exploitation, while ineffective ones go bankrupt and are replaced via exploration. Using modern Large Language Model (LLM) based agents, we show that, initialized with weak agents, the agent economy produces emergent multi-step reasoning strategies and outperforms stronger monolithic baselines across five agentic tasks, including mathematical reasoning, financial research, scientific research, accelerator design, and distributed-system optimization. Our results suggest a new path to multi-agent intelligence: rather than engineering coordination, we can design decentralized incentive structures under which it automatically emerges.

Multi-agent systems: an old problem?

“The greatest improvement in the productive powers of labour, and the greater part of the skill, dexterity, and judgment, with which it is anywhere directed, or applied, seem to have been the effects of the division of labour.”

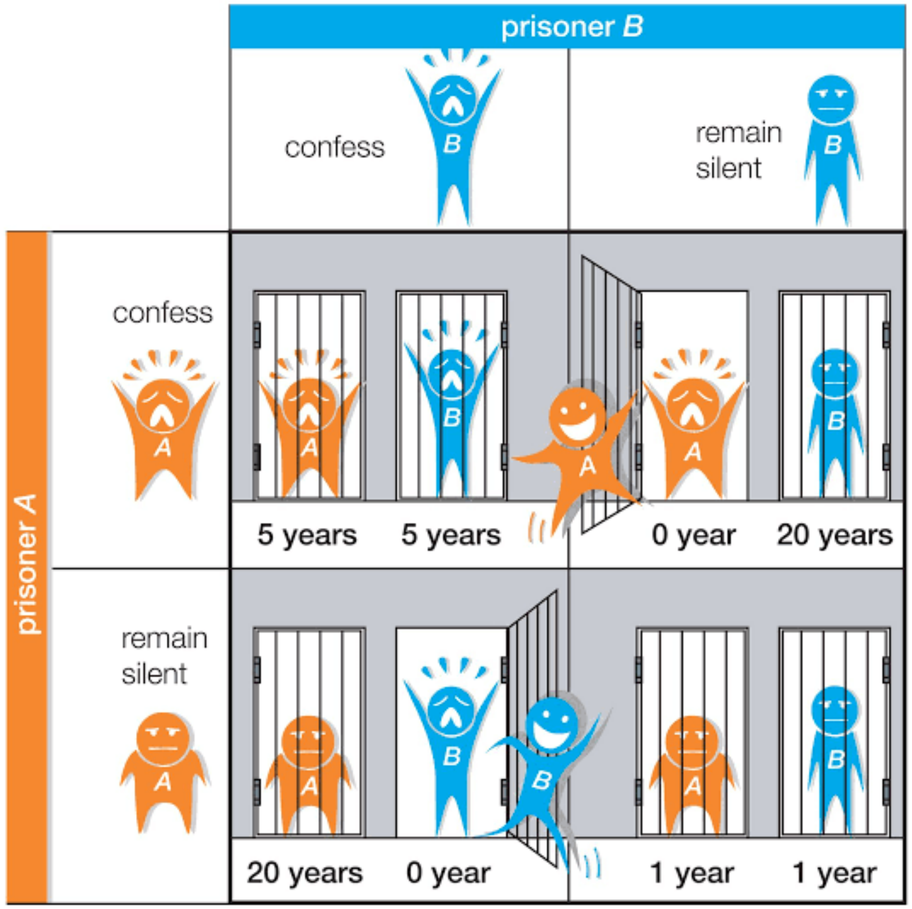
***The Wealth of Nations*, Adam Smith (1776)**

Why is the division of labor so effective?

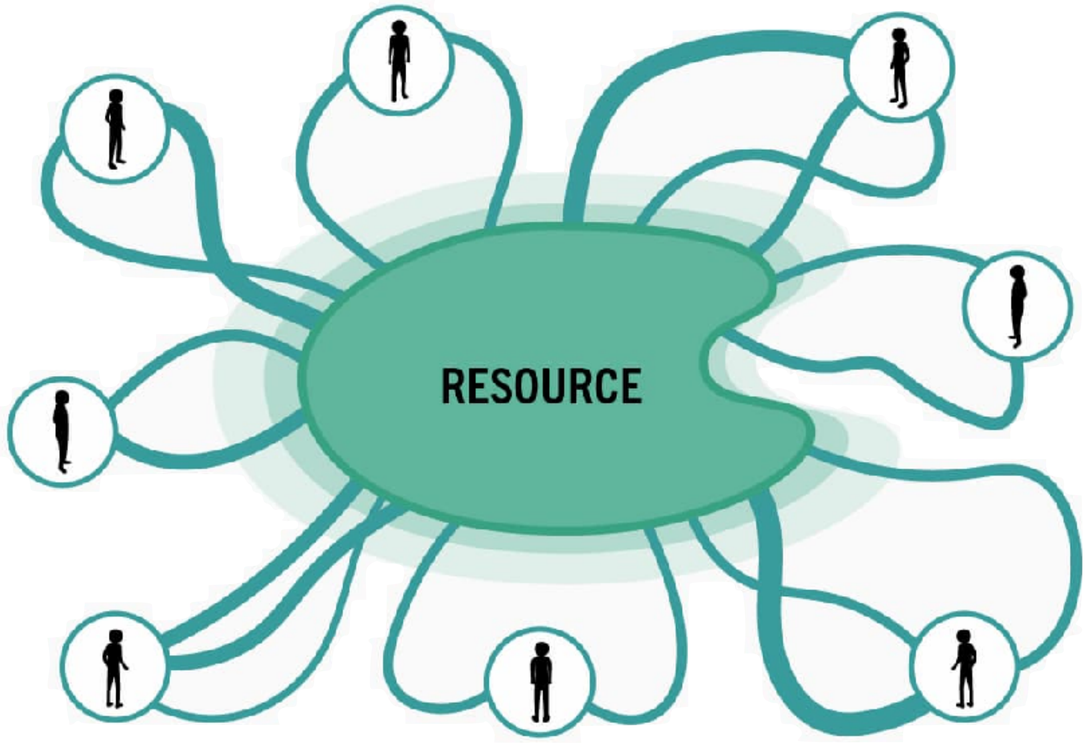
Why is the division of labor so effective?

Can it come with a price of anarchy?

Price of Anarchy



Prisoner's Dilemma



Tragedy of the Commons

Price of Anarchy

Decentralization can engender situations where everyone is worse off.

Prisoner's Dilemma

Tragedy of the Commons

This talk: why can individually-learned behavior find jointly rational solutions?

I. **Optimization**

II. Definitions

III. Existing Results

IV. Uniform Stability

V. Conclusion

Learning Dynamics in Optimization

- Let $f : \Omega \to \mathbb{R}$ be an **objective** that we want to maximize.

Learning Dynamics in Optimization

- Let $f : \Omega \to \mathbb{R}$ be an **objective** that we want to maximize.
- The **fixed points** of many natural learning dynamics are the *critical points*

$$\left\{x \in \Omega : \nabla f(x) = 0\right\}.$$

Learning Dynamics in Optimization

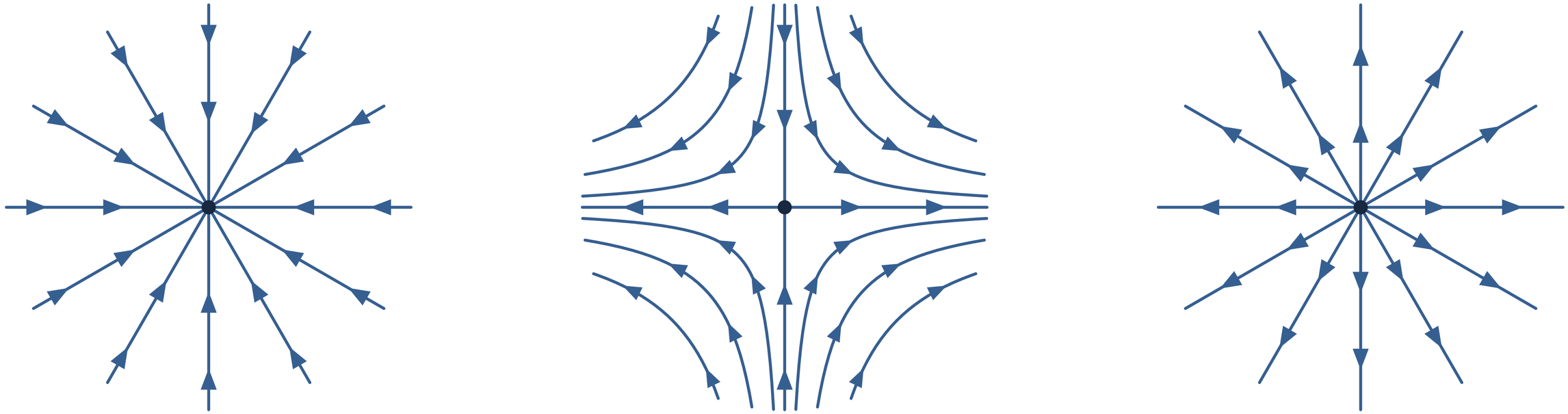
- Let $f : \Omega \to \mathbb{R}$ be an **objective** that we want to maximize.
- The **fixed points** of many natural learning dynamics are the *critical points*

$$\left\{x \in \Omega : \nabla f(x) = 0\right\}.$$

Question: why do these dynamics generally converge to *local maxima*, instead of saddle points or local minima?

Stability of Fixed Points

The **stability** of a fixed point $x^\star$ tells us how robust the qualitative behavior of the dynamics are under small perturbations.



Classification of gradient-based dynamics

Analysis of Stability

- For concreteness, let's consider continuous-time **gradient ascent**

$$\dot{x}_t = \nabla f(x_t).$$

Analysis of Stability

- For concreteness, let's consider continuous-time **gradient ascent**

$$\dot{x}_t = \nabla f(x_t).$$

- We can **linearize** the dynamics about a fixed point $x^\star$

$$\dot{x}_t = \nabla f(x^\star) + J(x^\star)(x_t - x^\star) + o(\|x - x^\star\|).$$

where $J = \nabla^2 f$ is the Jacobian of the dynamics.

Analysis of Stability

- For concreteness, let's consider continuous-time **gradient ascent**

$$\dot{x}_t = \nabla f(x_t).$$

- We can **linearize** the dynamics about a fixed point $x^\star$

$$\dot{x}_t = \cancelto{0}{\nabla f(x^\star)} + J(x^\star)(x_t - x^\star) + o(\|x - x^\star\|).$$

where $J = \nabla^2 f$ is the Jacobian of the dynamics.

Analysis of Stability

- For concreteness, let's consider continuous-time **gradient ascent**

$$\dot{x}_t = \nabla f(x_t).$$

- We can **linearize** the dynamics about a fixed point $x^\star$

$$\dot{z}_t = J(x^\star)\, z_t,$$

where $J = \nabla^2 f$ is the Jacobian of the dynamics and $z = x - x^\star$.

Analysis of Stability

- For concreteness, let's consider continuous-time **gradient ascent**

$$\dot{x}_t = \nabla f(x_t).$$

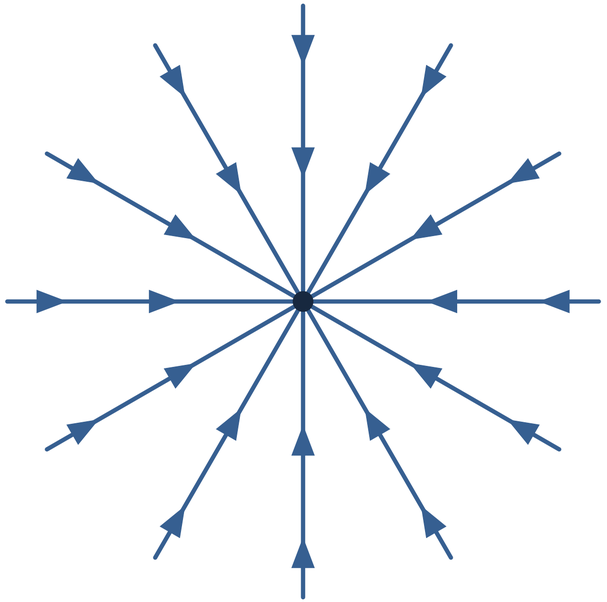
- We can **linearize** the dynamics about a fixed point $x^\star$

$$\dot{z}_t = J(x^\star)\, z_t,$$

where $J = \nabla^2 f$ is the Jacobian of the dynamics and $z = x - x^\star$.

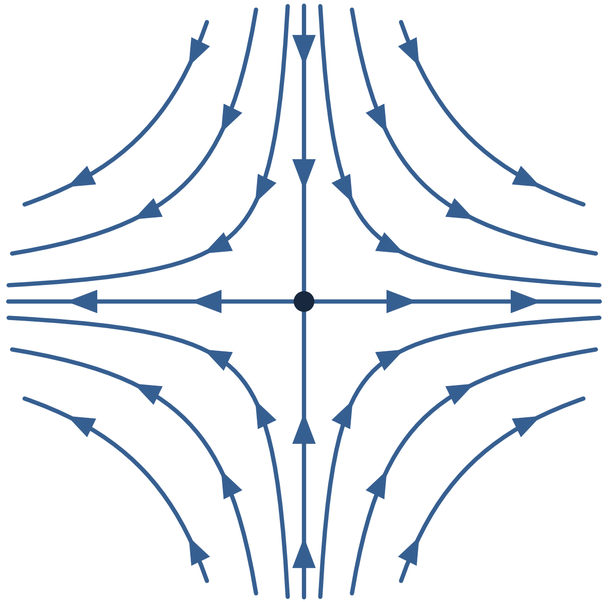
- The qualitative dynamics depend on the eigenvalues of $J^\star = J(x^\star)$.

Characterization of Linear Stability

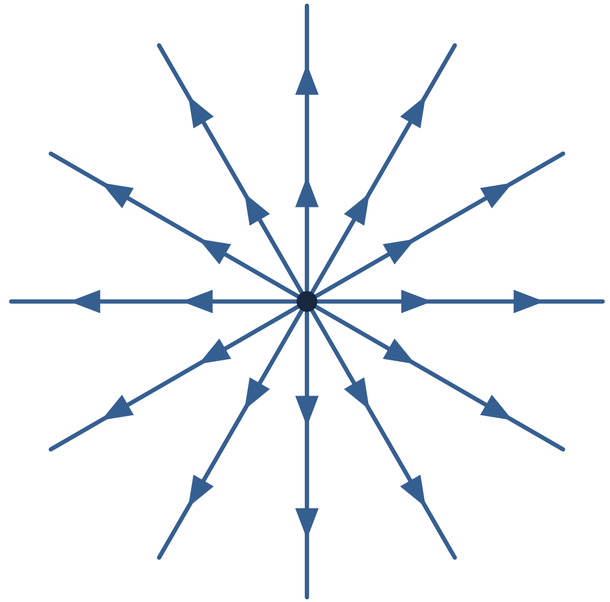


All eigenvalues negative

$\mathrm{spec}(J^\star) \subset \mathbb{R}_{<0}$



$J^\star$ has both positive and negative eigenvalues



All eigenvalues positive

$\mathrm{spec}(J^\star) \subset \mathbb{R}_{>0}$

Characterization of Linear Stability

Economic Property	Dynamical Property	Algebraic Characterization
Local maxima	Stable	Negative eigenvalues
Saddle point	Unstable	Mixed eigenvalues
Local minima	Unstable	Positive eigenvalues

Gradient Descent Converges to Minimizers

Jason D. Lee[♯], Max Simchowitz[♯], Michael I. Jordan[♯†], and Benjamin Recht[♯†]
[♯]Department of Electrical Engineering and Computer Sciences
[†]Department of Statistcs
University of California, Berkeley

March 7, 2016

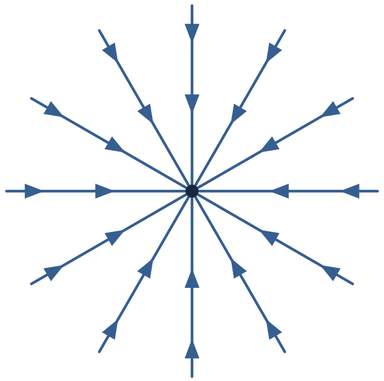
Abstract

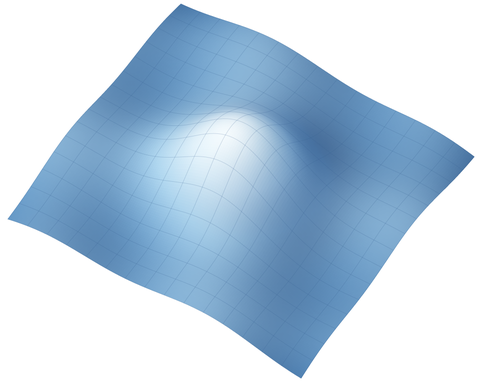
Gradient descent converges to a local minimizer, almost surely with random initialization.

Classic Story from Optimization

Gradient ascent dynamics around fixed points enjoy the following property:

Dynamical Stability $\Longleftrightarrow$ Economic Optimality

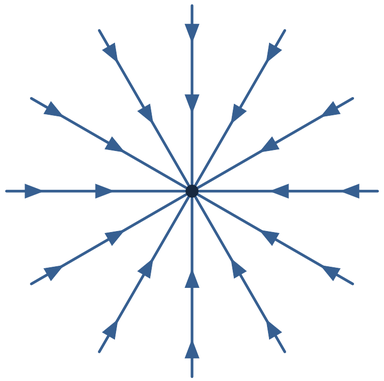


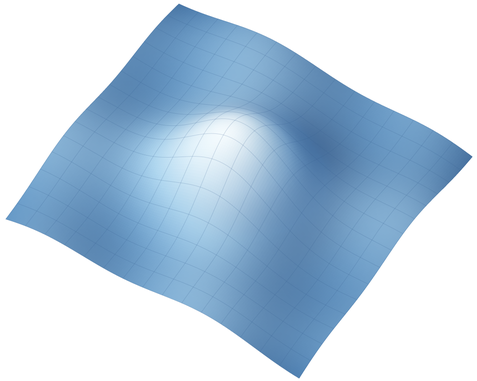


Classic Story from Optimization

Many other learning dynamics around fixed points enjoy the following property:

Dynamical Stability $\Longleftrightarrow$ Economic Optimality





Stability Across Optimization Geometry

- Learning dynamics that incorporate additional **geometric** information:
  - mirror ascent, gradient ascent with preconditioning, Riemannian ascent…

$$\dot{x}_t = H(x_t)^{-1} \nabla f(x_t), \qquad \text{where } H \succ 0.$$

Stability Across Optimization Geometry

$$\dot{x}_t = H(x_t)^{-1} \nabla f(x_t), \qquad \text{where } H \succ 0.$$

Why this form?

Stability Across Optimization Geometry

Claim. Every two non-zero vectors $u, v \neq 0$ satisfy

$$u^\top v > 0$$

Stability Across Optimization Geometry

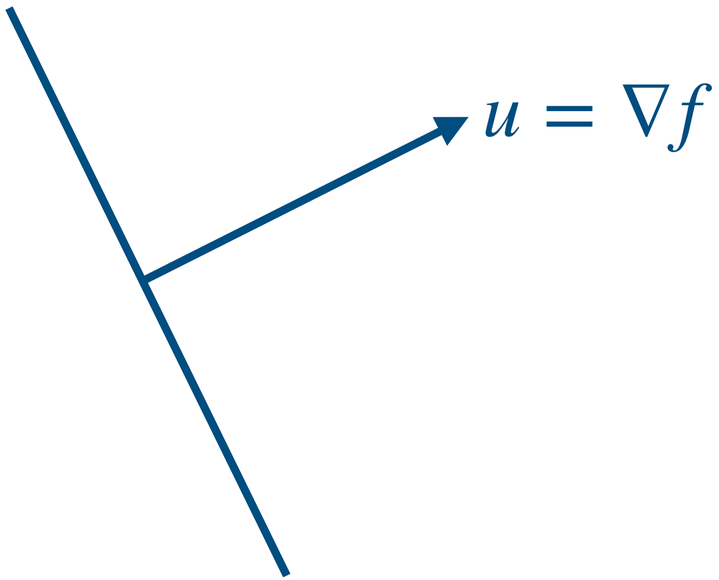
Claim. Every two non-zero vectors $u, v \neq 0$ satisfy

$$u^\top v > 0 \qquad \Longleftrightarrow \qquad \exists H \succ 0 \;\text{ s.t. }\; v = H^{-1}u.$$

Stability Across Optimization Geometry

Claim. Every two non-zero vectors $u, v \neq 0$ satisfy

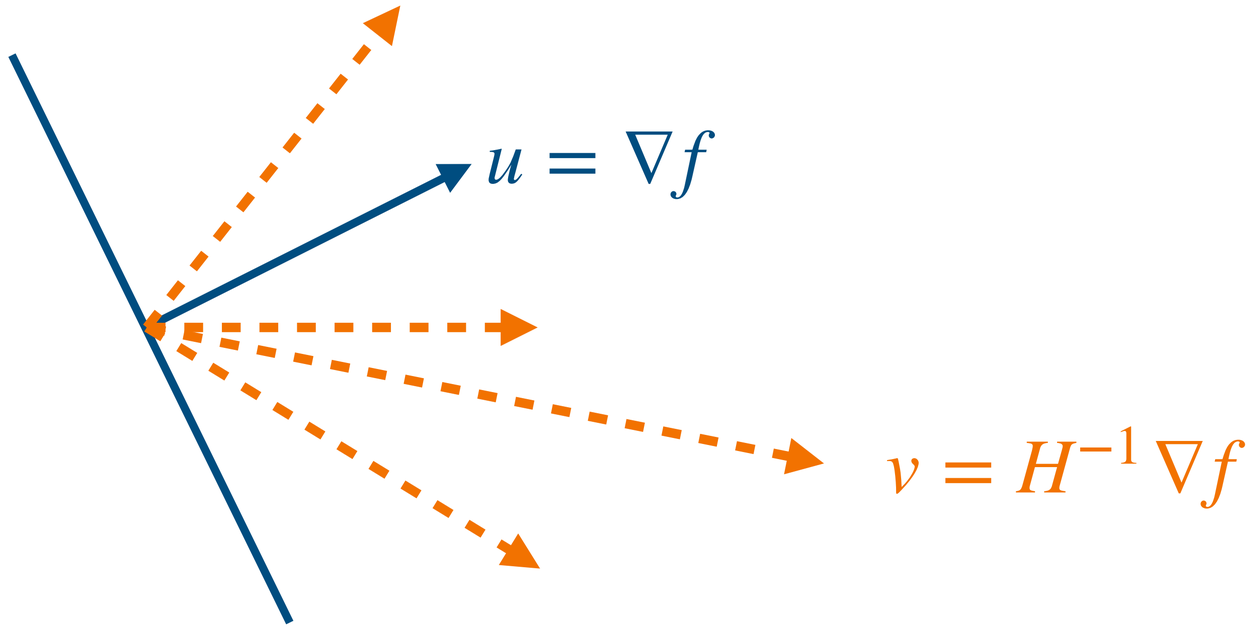
$$u^\top v > 0 \qquad \Longleftrightarrow \qquad \exists H \succ 0 \;\text{ s.t. }\; v = H^{-1}u.$$



Stability Across Optimization Geometry

Claim. Every two non-zero vectors $u, v \neq 0$ satisfy

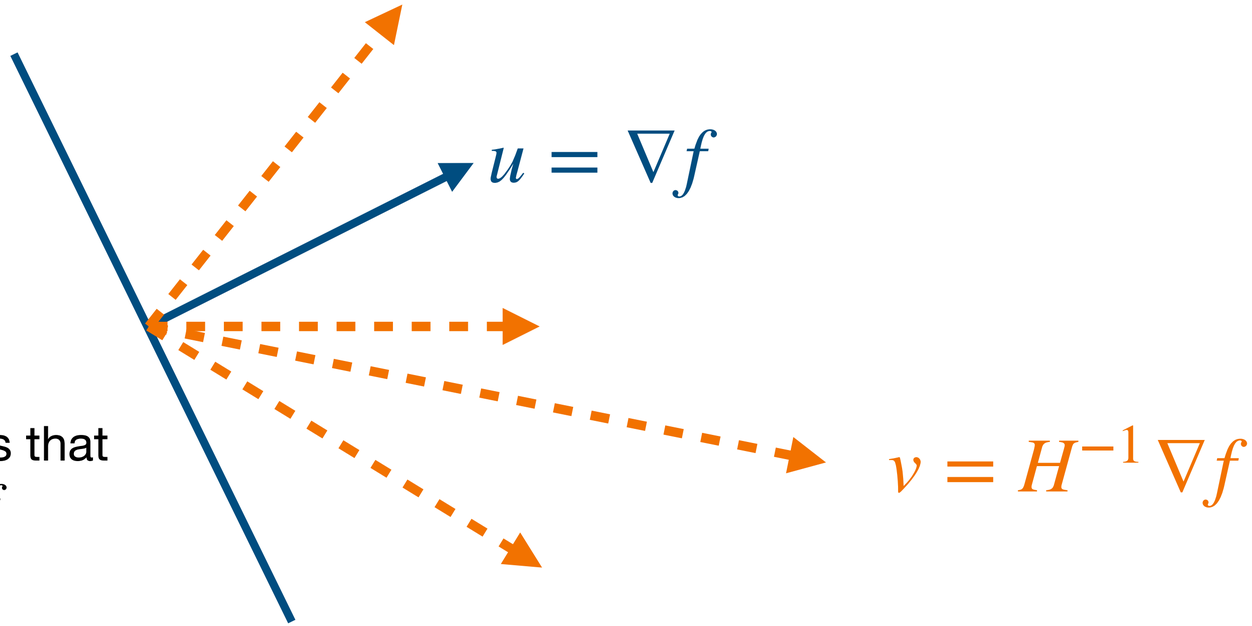
$$u^\top v > 0 \qquad \Longleftrightarrow \qquad \exists H \succ 0 \;\text{ s.t. }\; v = H^{-1}u.$$



Stability Across Optimization Geometry

Claim. Every two non-zero vectors $u, v \neq 0$ satisfy

$$u^\top v > 0 \qquad \Longleftrightarrow \qquad \exists H \succ 0 \;\; \text{s.t.} \;\; v = H^{-1} u.$$



$H^{-1} \nabla f$ are all the update directions that can make infinitesimal progress in f

$$(\nabla f)^\top H^{-1} \nabla f > 0.$$

Stability Across Optimization Geometry

Claim. Every two non-zero vectors $u, v \neq 0$ satisfy

$$u^\top v > 0 \qquad \Longleftarrow \qquad \exists H \succ 0 \;\text{ s.t. }\; v = H^{-1}u.$$

Proof.

$$u^\top v = u^\top H^{-1} u > 0.$$

$\uparrow$

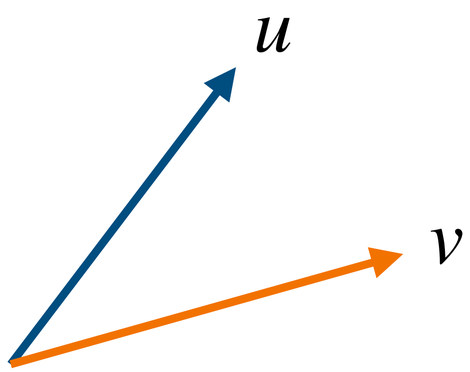
H^{-1} is positive-definite

Stability Across Optimization Geometry

Claim. Every two non-zero vectors $u, v \neq 0$ satisfy

$$u^\top v > 0 \qquad \Longrightarrow \qquad \exists H \succ 0 \;\text{ s.t. }\; v = H^{-1}u.$$

Proof.

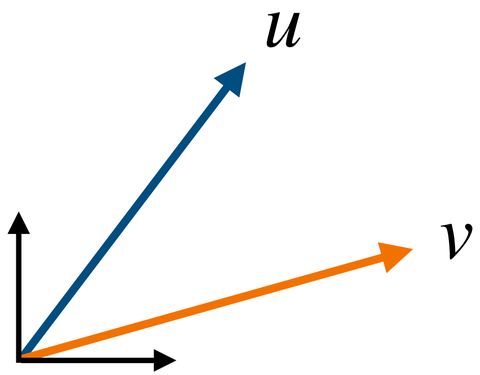


Stability Across Optimization Geometry

Claim. Every two non-zero vectors $u, v \neq 0$ satisfy

$$u^\top v > 0 \qquad \Longrightarrow \qquad \exists H \succ 0 \;\text{ s.t. }\; v = H^{-1}u.$$

Proof.



$\exists B$ an orthonormal basis s.t. $\alpha_i, \beta_i > 0$ and

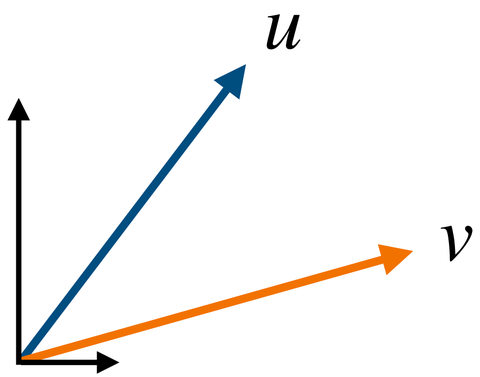
$$B^\top u = (\alpha_1, \alpha_2, 0, \ldots, 0) \;\text{ and }\; B^\top v = (\beta_1, \beta_2, 0, \ldots, 0).$$

Stability Across Optimization Geometry

Claim. Every two non-zero vectors $u, v \neq 0$ satisfy

$$u^\top v > 0 \qquad \Longrightarrow \qquad \exists H \succ 0 \;\text{ s.t. }\; v = H^{-1}u.$$

Proof.



$\exists B$ an orthonormal basis s.t. $\alpha_i, \beta_i > 0$ and

$$B^\top u = (\alpha_1, \alpha_2, 0, \ldots, 0) \;\text{ and }\; B^\top v = (\beta_1, \beta_2, 0, \ldots, 0).$$

We can stretch along the basis directions so that

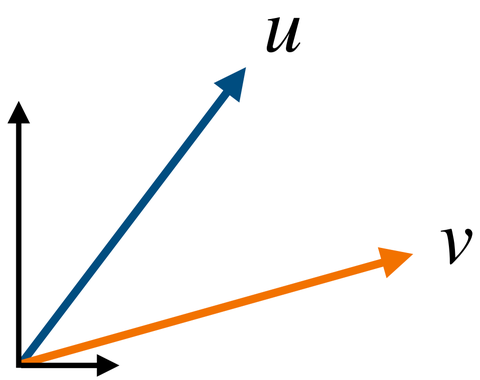
$$\Lambda(B^\top v) = B^\top u \qquad \Lambda = \begin{bmatrix} \alpha_1/\beta_1 & & \\ & \alpha_2/\beta_2 & \\ & & \ddots \end{bmatrix}$$

Stability Across Optimization Geometry

Claim. Every two non-zero vectors $u, v \neq 0$ satisfy

$$u^\top v > 0 \qquad \Longrightarrow \qquad \exists H \succ 0 \;\text{ s.t. }\; v = H^{-1}u.$$

Proof.



By construction, $B\Lambda B^\top v = u$.

$B\Lambda B^\top$ is positive-definite □

Stability Across Optimization Geometry

- Learning dynamics that incorporate additional **geometric** information:
  - mirror ascent, gradient ascent with preconditioning, Riemannian ascent…

$$\dot{x}_t = H(x_t)^{-1} \nabla f(x_t), \qquad \text{where } H \succ 0.$$

Stability Across Optimization Geometry

- Learning dynamics that incorporate additional **geometric** information:
  - mirror ascent, gradient ascent with preconditioning, Riemannian ascent…

$$\dot{x}_t = H(x_t)^{-1} \nabla f(x_t), \qquad \text{where } H \succ 0.$$

- The linearized dynamics about a fixed points $x^\star$ all have the form

$$\dot{z}_t = H(x^\star)^{-1} J(x^\star)\, z_t,$$

  - where $J = \nabla^2 f$ is the Hessian of the objective.

Stability Across Optimization Geometry

$$\dot{z}_t = H(x^\star)^{-1} J(x^\star)\, z_t,$$

Stability depends on the eigenvalues of this matrix.

Stability Across Optimization Geometry

Claim. The signs of the eigenvalues of $H^{-1}J$ coincide with those of J.

Stability Across Optimization Geometry

Claim. The signs of the eigenvalues of $H^{-1}J$ coincide with those of J.

Proof.

$$H^{-1}J =$$

Stability Across Optimization Geometry

Claim. The signs of the eigenvalues of $H^{-1}J$ coincide with those of J.

Proof.

$$H^{-1}J = H^{-1}\ U\Lambda U^\top$$

$J \equiv \nabla^2 f$ is symmetric

$\Longrightarrow$ has spectral decomposition $J = U\Lambda U^\top$

Stability Across Optimization Geometry

Claim. The signs of the eigenvalues of $H^{-1}J$ coincide with those of J.

Proof.

$$H^{-1}J \sim H^{1/2}\, H^{-1}\, U\Lambda U^{\top}\, H^{-1/2}$$

H is positive-definite

$\Longrightarrow$ has square root

Stability Across Optimization Geometry

Claim. The signs of the eigenvalues of $H^{-1}J$ coincide with those of J.

Proof.

$$H^{-1}J \sim H^{-1/2}\, U\Lambda U^{\top}\, H^{-1/2}$$

Stability Across Optimization Geometry

Claim. The signs of the eigenvalues of $H^{-1}J$ coincide with those of J.

Proof.

$$H^{-1}J \sim H^{-1/2}\, U\Lambda U^\top\, H^{-1/2}$$

Let $M = H^{-1/2}U$.

Stability Across Optimization Geometry

Claim. The signs of the eigenvalues of $H^{-1}J$ coincide with those of J.

Proof.

$$H^{-1}J \sim M \Lambda M^{\top}$$



The signs of the eigenvalues of $H^{-1}J$ depend only on the eigenvalues Λ of J. $\square$

Stability Across Optimization Geometry

- Learning dynamics that incorporate additional **geometric** information:
  - mirror ascent, gradient ascent with preconditioning, Riemannian ascent…

$$\dot{x}_t = H(x_t)^{-1} \nabla f(x_t), \qquad \text{where } H \succ 0.$$

- The linearized dynamics about a fixed points $x^\star$ all have the form

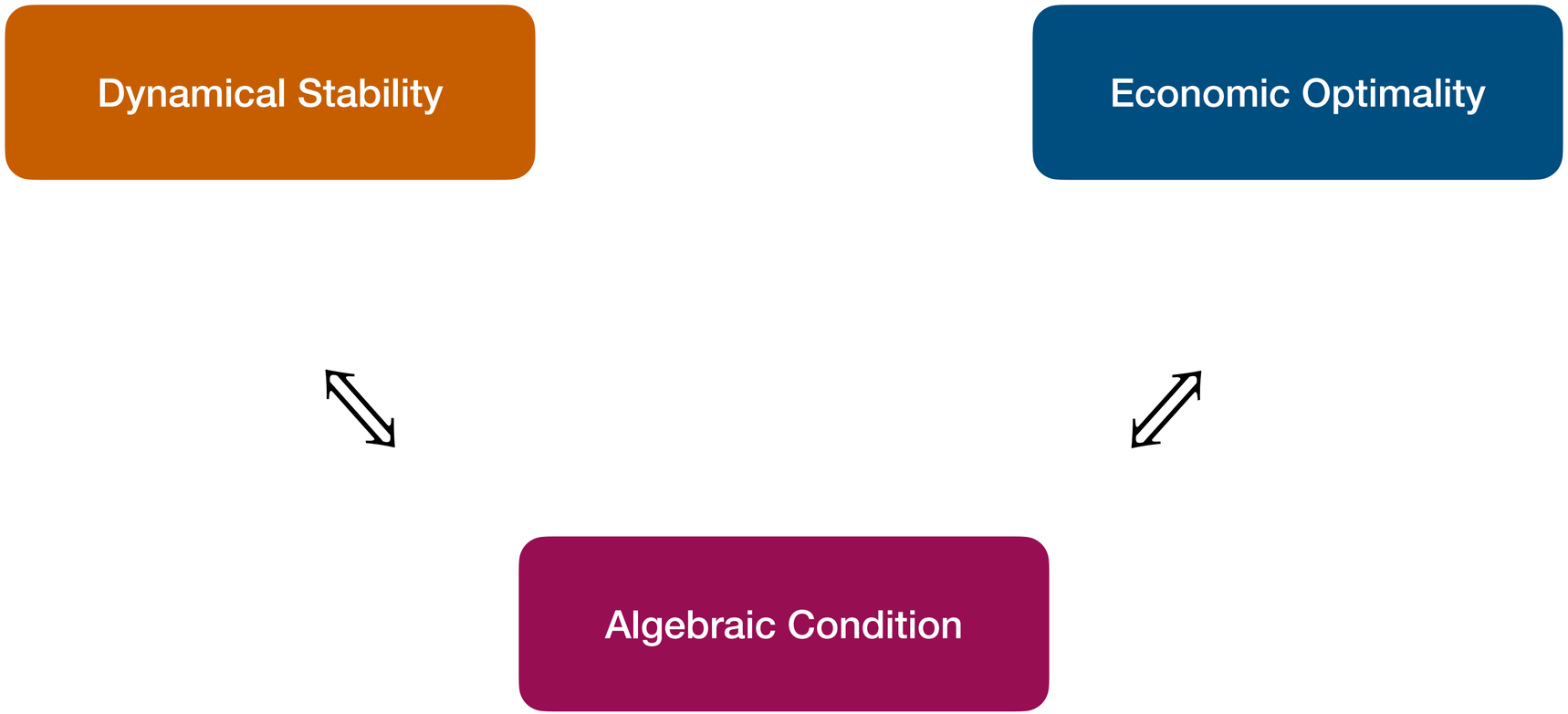
$$\dot{z}_t = H(x^\star)^{-1} J(x^\star)\, z_t,$$

  - where $J = \nabla^2 f$ is the Hessian of the objective.

- Stability of **any** learning dynamics $\iff$ stability of **all** learning dynamics.

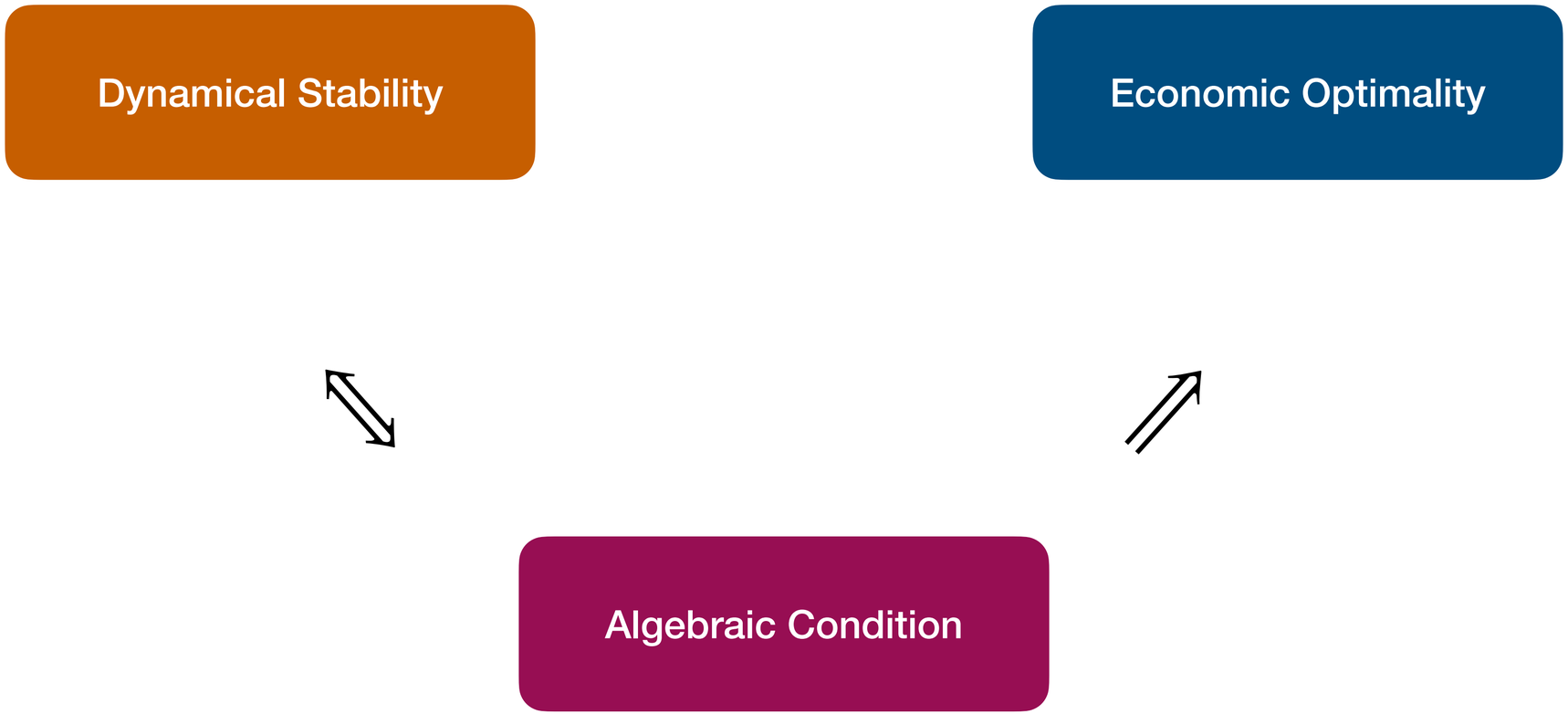
Uniform Stability for Optimization

If **all reasonable learning dynamics** converge, then $x^\star$ is locally optimal.

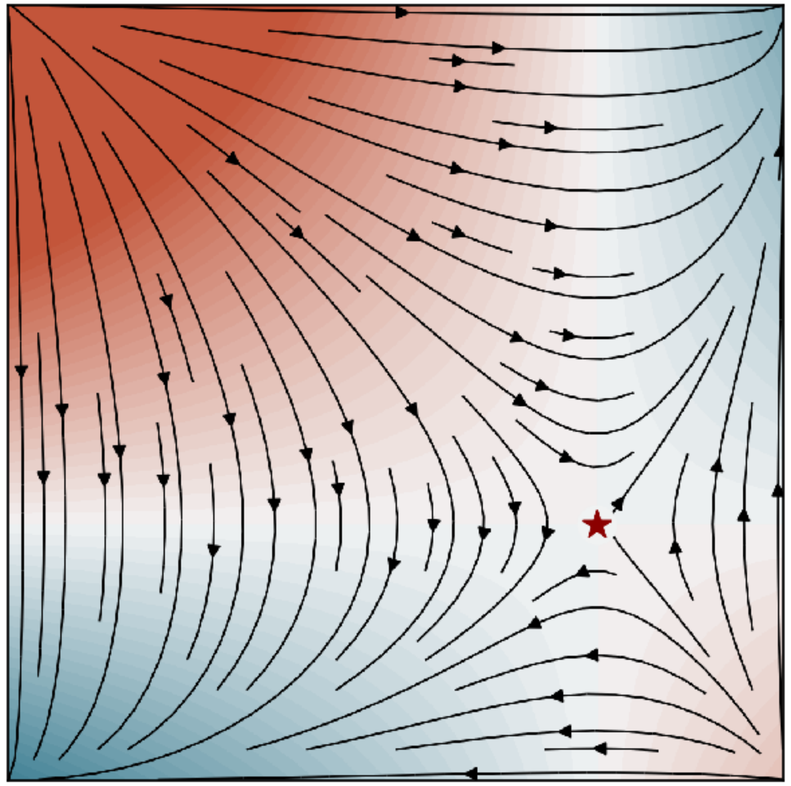


Preview: Uniform Stability for Games

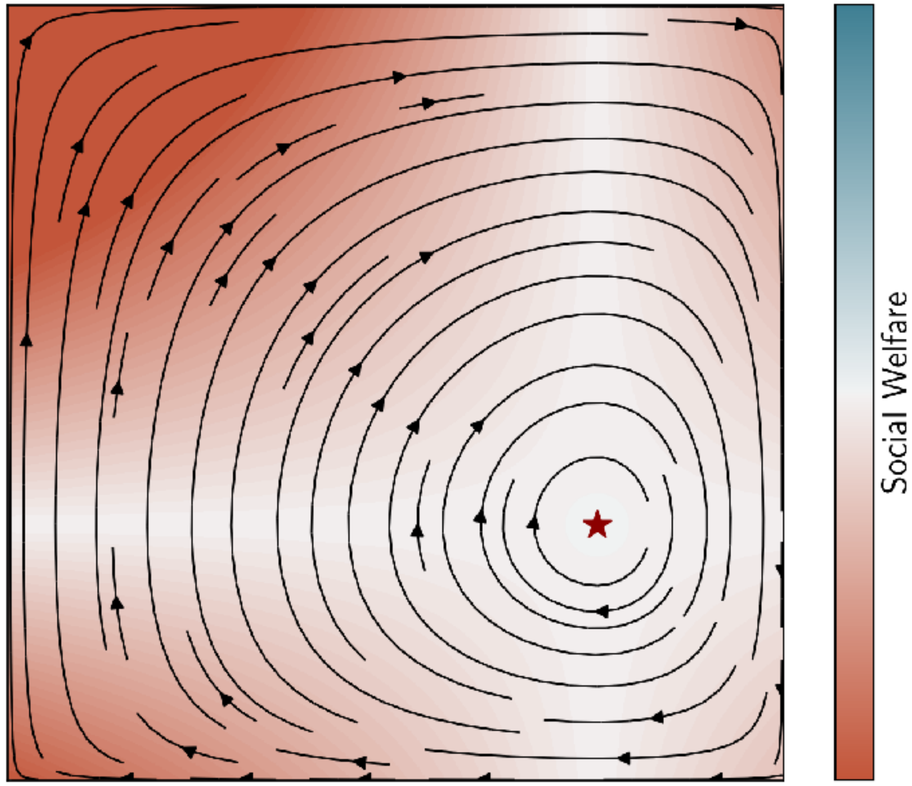
If **all reasonable learning dynamics** converge to an interior Nash equilibrium (it is uniformly stable), then it is collectively rational (strategically Pareto optimal).



Comparison of Dynamics Around Non-Strict NE



Strategically dominated solutions are not uniformly stable.



Uniformly stable Nash equilibria are strategically Pareto optimal.

I. Optimization

II. Definitions

III. Existing Results

IV. Uniform Stability

V. Conclusion

Hallmark of a Game

A game is a multi-agent system where:

- the **outcome** for each agent depends on the collective, but
- each agent **controls** only a subset of all decision variables.

Normal-Form Games

- Consider a general-sum, N-player **normal-form game.**

Normal-Form Games

- Consider a general-sum, N-player **normal-form game.**
- Each player/agent $n \in [N]$ chooses to (randomly) play one of k_n actions

$$x_n \in \Omega_n \equiv \Delta^{k_n - 1}.$$

Probability simplex over agent n's action set.

Normal-Form Games

- Consider a general-sum, N-player **normal-form game**.
- Each player/agent $n \in [N]$ chooses to (randomly) play one of k_n actions

$$x_n \in \Omega_n \equiv \Delta^{k_n - 1}.$$

Probability simplex over agent n's action set.

  - Interpretation: fraction of the nth population choosing each action.

Normal-Form Games

- Consider a general-sum, N-player **normal-form game**.
- Each player/agent $n \in [N]$ chooses to (randomly) play one of k_n actions

$$x_n \in \Omega_n \equiv \Delta^{k_n - 1}.$$

Probability simplex over agent n's action set.

  - Interpretation: fraction of the nth population choosing each action.
- Let $f_1, \ldots, f_N$ be the set of **objectives** (expected utilities) of each player

$$f_n : \Omega_1 \times \cdots \times \Omega_N \to \mathbb{R}\,.$$

f_n is multilinear: it is linear in each set of decision variables x_i.

Learning Setting

- **Infinitely repeated game:** For $t = 1,2,\ldots$:
  - Each player $n \in [N]$ plays a strategy $x_n(t)$
  - Everyone observes the joint strategy played $x(t)$

Learning Setting

- **Infinitely repeated game:** For $t = 1,2,\ldots$:*
  - Each player $n \in [N]$ plays a strategy $x_n(t)$
  - Everyone observes the joint strategy played $x(t)$

*The continuous time version is also sensible.

Learning Setting

- **Infinitely repeated game:** For $t = 1,2,\ldots$:*
  - Each player $n \in [N]$ plays a strategy $x_n(t)$
  - Everyone observes the joint strategy played $x(t)$
- **Last-iterate behavior:** how does $x(t)$ evolve in the long run?

*The continuous time version is also sensible.

Learning in Games: What Behavior Emerges

Players enter a game with limited knowledge:

- They know what set of actions they/others can take.
- They know their own utilities.

Learning in Games: What Behavior Emerges

Players enter a game with limited knowledge:

- They know what set of actions they/others can take.
- They know their own utilities.

They do not know:

- What other players want.
- How other players learn.

] The players are **uncoupled**.

Learning in Games: What Behavior Emerges

Players enter a game with limited knowledge:

- They know what set of actions they/others can take.
- They know their own utilities.

They do not know:

- What other players want.
- How other players learn.

Players will repeatedly play the game, observe what other do, and adjust their own behaviors for future rounds.

Simple Learning Rules

- **Best-response dynamics**

$$x_n(t+1) = \arg\max_{z_n \in \Omega_n} f_n\left(z_n; x_{-n}(t)\right)$$

- **Smoothed best-response dynamics**
- **Gradient ascent**
- **Positive-definite dynamics**

Simple Learning Rules

- **Best-response dynamics**
- **Smoothed best-response dynamics**

$$x_n(t+1) = \arg\max_{z_n \in \Omega_n} f_n\left(z_n; x_{-n}(t)\right) + \text{regularizer}(z_n)$$

- **Gradient ascent**
- **Positive-definite dynamics**

Simple Learning Rules

- **Best-response dynamics**
- **Smoothed best-response dynamics**
- **Gradient ascent-ascent**

$$\dot{x}_n(t) = \eta \nabla_n f_n\Big(x_n(t); x_{-n}(t)\Big)$$

- **Positive-definite dynamics**

Simple Learning Rules

- **Best-response dynamics**
- **Smoothed best-response dynamics**
- **Gradient ascent**
- **Positive-definite dynamics**

$$\dot{x}(t) = H_n^{-1} \nabla_n f_n\Big(x(t)\Big)$$

where H_n is positive-definite (also, preconditioned gradient ascent, mirror ascent,…).

Qualitative Behaviors of Game Dynamics

Linearizing Gradient Ascent Dynamics

The gradient ascent dynamics:

$$\dot{x}_n(t) = \nabla_n f_n\Big(\mathbf{x}(t)\Big)$$

Linearizing Gradient Ascent Dynamics

The gradient ascent dynamics:

$$\dot{x}_n(t) = \nabla_n f_n\Big(\mathbf{x}(t)\Big)$$

- The Jacobian of this dynamics is given by:

$$\mathbf{J} = \begin{bmatrix} \nabla^2_{11} f_1 & \nabla^2_{12} f_1 & \cdots & \nabla^2_{1N} f_1 \\ \nabla^2_{21} f_2 & \nabla^2_{22} f_2 & & \\ \vdots & & \ddots & \\ \nabla^2_{N1} f_N & & & \nabla^2_{NN} f_N \end{bmatrix}$$

Linearizing Gradient Ascent Dynamics

The gradient ascent dynamics:

$$\dot{x}_n(t) = \nabla_n f_n\Big(\mathbf{x}(t)\Big)$$

- The Jacobian of this dynamics is given by:

$$\mathbf{J} = \begin{bmatrix} 0 & \nabla^2_{12} f_1 & \cdots & \nabla^2_{1N} f_1 \\ \nabla^2_{21} f_2 & 0 & & \\ \vdots & & \ddots & \\ \nabla^2_{N1} f_N & & & 0 \end{bmatrix}$$

  - The diagonal is zero by multilinearity.

Linearizing Gradient Ascent Dynamics

The gradient ascent dynamics:

$$\dot{x}_n(t) = \nabla_n f_n\Big(\mathbf{x}(t)\Big)$$

- The Jacobian of this dynamics is given by:

$$\mathbf{J} = \begin{bmatrix} 0 & \nabla^2_{12} f_1 & \cdots & \nabla^2_{1N} f_1 \\ \nabla^2_{21} f_2 & 0 & & \\ \vdots & & \ddots & \\ \nabla^2_{N1} f_N & & & 0 \end{bmatrix}$$

  - This object is called the **game Jacobian**…

Linearizing Gradient Ascent Dynamics

The gradient ascent dynamics:

$$\dot{x}_n(t) = \nabla_n f_n\Big(\mathbf{x}(t)\Big)$$

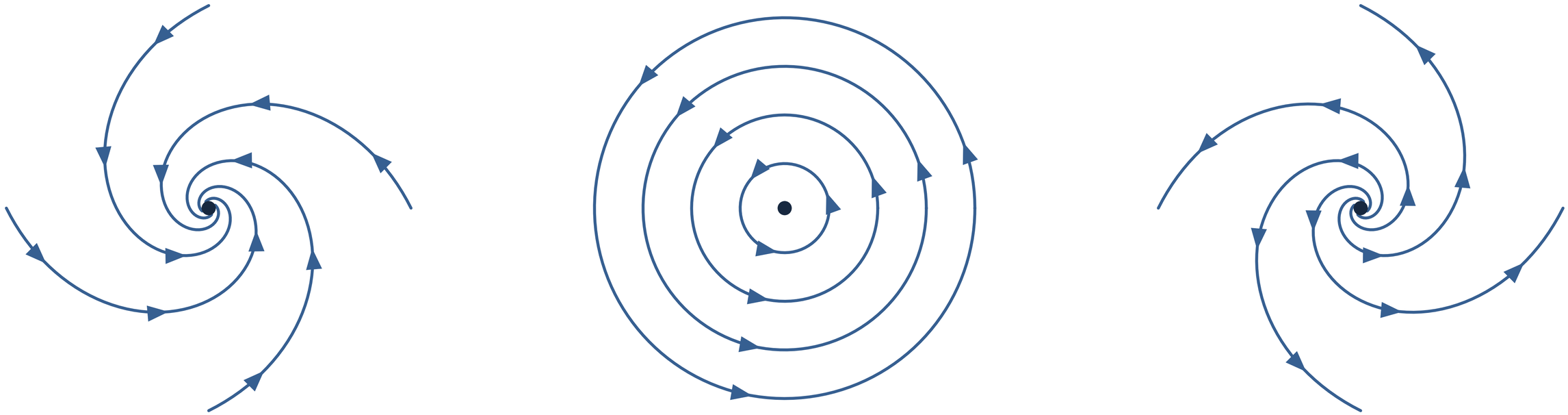
- The Jacobian of this dynamics is given by:

$$\mathbf{J} = \begin{bmatrix} 0 & \nabla^2_{12} f_1 & \cdots & \nabla^2_{1N} f_1 \\ \nabla^2_{21} f_2 & 0 & & \\ \vdots & & \ddots & \\ \nabla^2_{N1} f_N & & & 0 \end{bmatrix}$$

  - This object is called the **game Jacobian**; it does not have to be symmetric.

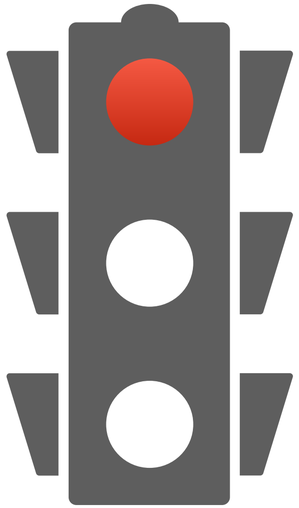
Stability of Fixed Points

Game dynamics can be more complicated than optimization dynamics.



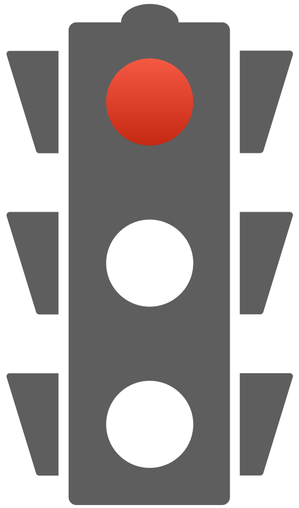
Additional types of non-gradient dynamics

Prior Work: Stability of Social Conventions



Dynamics of multi-agent systems around strict Nash equilibria are well-understood.

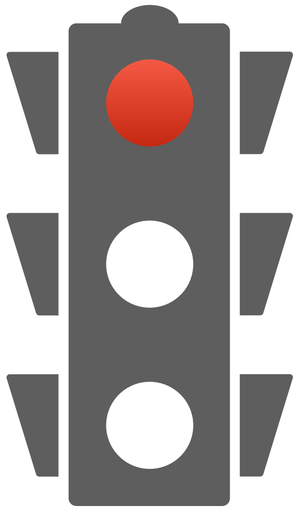
Prior Work: Stability of Social Conventions



Dynamics of multi-agent systems around strict Nash equilibria are well-understood.

There is a clearly optimal, deterministic action; $x^\star$ on the boundary of Ω.

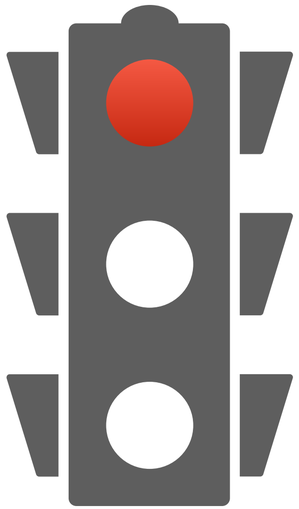
Prior Work: Stability of Social Conventions



Dynamics of multi-agent systems around strict Nash equilibria are well-understood.

Example. Once enough people agree **RED/GREEN** means **STOP/GO**, everyone else is forced to adopt this convention.

Prior Work: Stability of Social Conventions



Dynamics of multi-agent systems around strict Nash equilibria are well-understood.

Example. Once enough people agree **RED/GREEN** means **STOP/GO**, everyone else is forced to adopt this convention.

Stability does not guarantee collective rationality.

Stability in Non-Strict Equilibria



What about dynamics around non-strict Nash equilibria?

Example. When the job market is at equilibrium, many career options are viable, leading to a mixed population.

Stability in Non-Strict Equilibria



What about dynamics around non-strict Nash equilibria?

Example. When the job market is at equilibrium, many career options are viable, leading to a mixed population.

Stability guarantees collective rationality.

This work

Solution Concepts in Games

Individual Rationality: Nash Equilibria

A Nash equilibrium $x^\star$ is a solution where each player plays an optimal strategy from the individual standpoint

$$x_n^\star \in \arg \max_{x_n \in \Omega_n} f_n(x_n; x_{-n}^\star) \, .$$

Individual Rationality: Nash Equilibria

A Nash equilibrium $x^\star$ is a solution where each player plays an optimal strategy from the individual standpoint

$$x_n^\star \in \arg \max_{x_n \in \Omega_n} f_n(x_n; x_{-n}^\star) \, .$$

- This is a “minimal” property of a persistent solution.

Individual Rationality: Nash Equilibria

A Nash equilibrium $x^\star$ is a solution where each player plays an optimal strategy from the individual standpoint

$$x_n^\star \in \arg \max_{x_n \in \Omega_n} f_n(x_n; x_{-n}^\star) \, .$$

- This is a “minimal” property of a persistent solution.
- Local version:

$$\nabla_{x_n} f_n(x^\star) = 0.$$

Collective Rationality: Pareto Optimality

There is no alternative joint strategy x that is strictly better for everyone

$$f_n(x) > f_n(x^\star) \text{ for all } n \in [N] \,.$$

Collective Rationality: Pareto Optimality

There is no alternative joint strategy x that is strictly better for everyone

$$f_n(x) > f_n(x^\star) \text{ for all } n \in [N] \,.$$

- Weak Pareto optimality is a “minimal condition” of collective rationality.

Learnability: Dynamical Stability

A learning dynamics is **asymptotically stable** around $x^\star$ if convergence is robust against small perturbations.

Learnability: Dynamical Stability

A learning dynamics is **asymptotically stable** around $x^\star$ if convergence is robust against small perturbations.

- For smooth, continuous-time dynamics $\dot{x}_t = T(x_t)$ and $\nabla T(x^\star) = 0$:

Learnability: Dynamical Stability

A learning dynamics is **asymptotically stable** around $x^\star$ if convergence is robust against small perturbations.

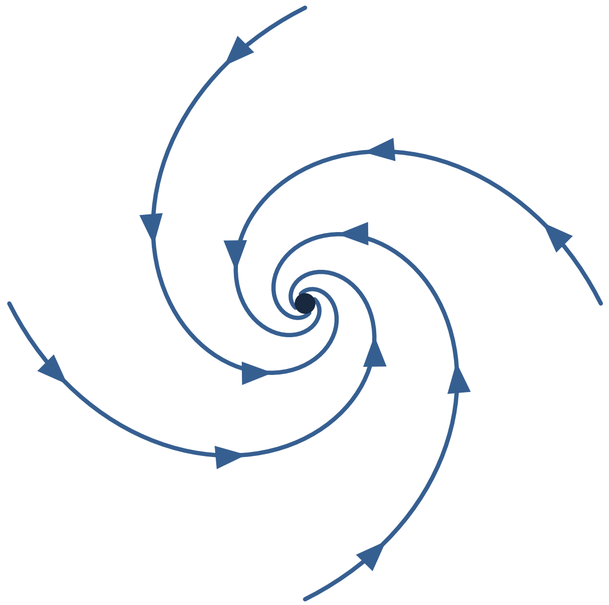
- For smooth, continuous-time dynamics $\dot{x}_t = T(x_t)$ and $\nabla T(x^\star) = 0$:
  - a sufficient condition: all eigenvalues of $\nabla T(x^\star)$ have negative real part.

Learnability: Dynamical Stability

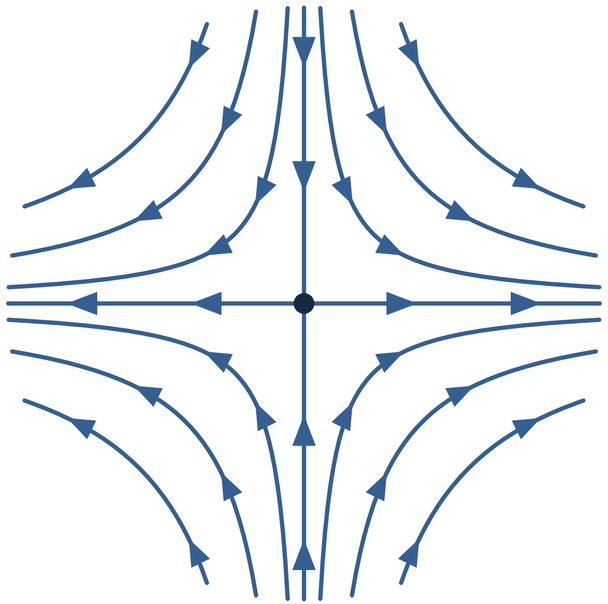
A learning dynamics is **asymptotically stable** around $x^\star$ if convergence is robust against small perturbations.

- For smooth, continuous-time dynamics $\dot{x}_t = T(x_t)$ and $\nabla T(x^\star) = 0$:
  - a sufficient condition: all eigenvalues of $\nabla T(x^\star)$ have negative real part.
  - a necessary condition: all eigenvalues of have non-positive real part.

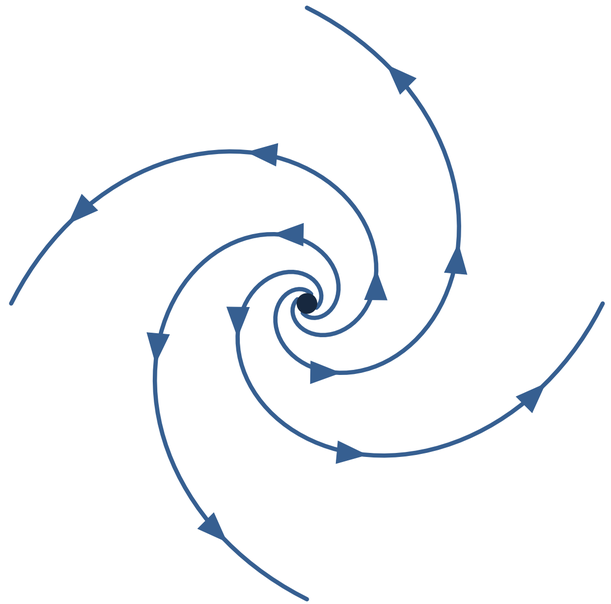
Characterization of Linear Stability



All eigenvalues have negative real part



Eigenvalues have positive and negative real parts



All eigenvalues have positive real part

Jacobian of Learning Dynamics

It turns out that the Jacobian of many learning dynamics is of the form:

$$\nabla T = \mathbf{H}^{-1}\mathbf{J}$$

where $\mathbf{J}$ is the game Jacobian and $\mathbf{H}$ is block-diagonal and positive definite:

$$\mathbf{H} = \begin{bmatrix} H_1 & & & \\ & H_2 & & \\ & & \ddots & \\ & & & H_N \end{bmatrix} \quad \text{and} \quad \mathbf{J} = \begin{bmatrix} 0 & \nabla^2_{12} f_1 & \cdots & \nabla^2_{1N} f_1 \\ \nabla^2_{21} f_2 & 0 & & \\ \vdots & & \ddots & \\ \nabla^2_{N1} f_N & & & 0 \end{bmatrix}$$

Jacobian of Learning Dynamics

- **Preconditioned gradient ascent**

$$\dot{x}_n(t) = H_n^{-1} \nabla_n f_n\Big(\mathbf{x}(t)\Big)$$

Jacobian of Learning Dynamics

- **Smoothed best-response dynamics**

$$x_n(t+1) = \arg\max_{z_n \in \Omega_n} f_n\left(z_n; \mathbf{x}_{-n}(t)\right) - h_n(z_n)$$

  - Usually h_n is a smooth, strictly convex regularizer
    - Entropy map, negative-definite quadratic.
    - It defines the “local geometry” on Ω_n.

Jacobian of Learning Dynamics

- **Smoothed best-response dynamics**

$$x_n(t+1) = \arg\max_{z_n \in \Omega_n} f_n\left(z_n; \mathbf{x}_{-n}(t)\right) - h_n(z_n)$$

  - Usually h_n is a smooth, strictly convex regularizer

$$\nabla T(\mathbf{x}) = \mathbf{H}^{-1}\mathbf{J} \qquad \text{where} \quad \mathbf{H} = \begin{bmatrix} \nabla^2 h_1 & & & \\ & \nabla^2 h_2 & & \\ & & \ddots & \\ & & & \nabla^2 h_N \end{bmatrix}$$

I. Optimization

II. Definitions

III. Existing Results

IV. Uniform Stability

V. Conclusion

An Impossibility Result on Finding Nash Equilibria

Uncoupled dynamics do not lead to Nash equilibrium, Hart and Mas-Colell 2003

Main result: there is a family of normal-form games $\mathscr{F}$ such that:

An Impossibility Result on Finding Nash Equilibria

Uncoupled dynamics do not lead to Nash equilibrium, Hart and Mas-Colell 2003

Main result: there is a family of normal-form games $\mathscr{F}$ such that:

1. Each game has a unique, interior Nash equilibrium.

An Impossibility Result on Finding Nash Equilibria

Uncoupled dynamics do not lead to Nash equilibrium, Hart and Mas-Colell 2003

Main result: there is a family of normal-form games $\mathscr{F}$ such that:

1. Each game has a unique, interior Nash equilibrium.
2. For every uncoupled dynamics, the dynamics will fail to converge to the Nash equilibrium with asymptotic stability for at least one game in this family.

Significance of Result

Uncoupled dynamics do not lead to Nash equilibrium, Hart and Mas-Colell 2003

Raises questions on the theoretical viability of the Nash equilibrium

- Is the Nash equilibrium meaningful if players cannot find it?

Some Foreshadowing

Uncoupled dynamics do not lead to Nash equilibrium, Hart and Mas-Colell 2003

Asymptotic stability is too strong of a criterion of learnability

Some Foreshadowing

Uncoupled dynamics do not lead to Nash equilibrium, Hart and Mas-Colell 2003

Asymptotic stability is too strong of a criterion of learnability

- We will introduce a non-asymptotic notion called uniform stability.

Some Foreshadowing

Uncoupled dynamics do not lead to Nash equilibrium, Hart and Mas-Colell 2003

Asymptotic stability is too strong of a criterion of learnability

- We will introduce a non-asymptotic notion called uniform stability.
- Perhaps it is also not so bad we cannot find certain Nash equilibria.

Which Nash Equilibria Can We Find?

Convergent dynamics exist for certain important subclasses of equilibria:

Which Nash Equilibria Can We Find?

Convergent dynamics exist for certain important subclasses of equilibria:

- For many standard learning dynamics, there are result of the form:

An equilibrium admits asymptotic convergence
if and only if it is a strict Nash equilibrium.

Which Nash Equilibria Can We Find?

Convergent dynamics exist for certain important subclasses of equilibria:

- For many standard learning dynamics, there are result of the form:

An equilibrium admits asymptotic convergence
if and only if it is a strict Nash equilibrium.

- Strictness: each player has a deterministic strategy that is locally optimal.

Which Nash Equilibria Can We Find?

Convergent dynamics exist for certain important subclasses of equilibria:

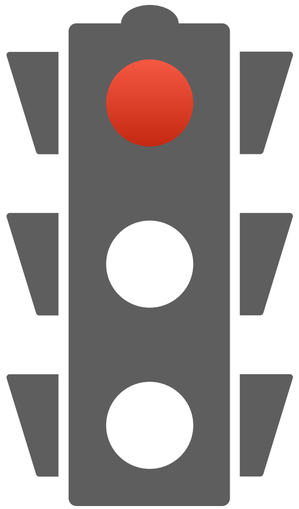
- For many standard learning dynamics, there are result of the form:

An equilibrium admits asymptotic convergence
if and only if it is a strict Nash equilibrium.

- Strictness: each player has a deterministic strategy that is locally optimal.
- Fictitious play, replicator dynamics, regularized learning, no-regret learning…

Samuelson and Zhang (1992); Vlatakis-Gkaragkounis et. al. 2020; Giannou et. al. (2021) …

Prior Work: Stability of Social Conventions



Dynamics of multi-agent systems around strict Nash equilibria are well-understood.

Example. Once enough people agree **RED/GREEN** means **STOP/GO**, everyone else is forced to adopt this convention.

Stability does not guarantee collective rationality.

Which Nash Equilibria Can We Find?

Interior Nash equilibria do not usually admit convergence, but convergent dynamics have been discovered for zero-sum games:

Which Nash Equilibria Can We Find?

Interior Nash equilibria do not usually admit convergence, but convergent dynamics have been discovered for zero-sum games:

- Optimistic gradient ascent
- Extragradient method

Which Nash Equilibria Can We Find?

Interior Nash equilibria do not usually admit convergence, but convergent dynamics have been discovered for zero-sum games:

- Optimistic gradient ascent
- Extragradient method

Further foreshadowing: our work gives some explanation as to why.

Barrier to Asymptotic Stability

Linear Stability from Spectral Analysis

Consider the eigenvalues of the Jacobian of some learning dynamics.

- If all eigenvalues have **negative real parts** $\Longrightarrow$ asymptotic stability
- If there is an eigenvalue with **positive real part** $\Longrightarrow$ instability
- If all eigenvalues are **purely imaginary** $\Longrightarrow$ neutral stability

The Jacobian of the Dynamics are Traceless

For learning dynamics whose Jacobians are of the form:

$$\nabla T = \mathbf{H}^{-1}\mathbf{J} = \begin{bmatrix} 0 & H_1^{-1}J_{12} & \cdots & H_1^{-1}J_{1N} \\ H_2^{-1}J_{21} & 0 & & \\ \vdots & & \ddots & \\ H_N^{-1}J_{N1} & & & 0 \end{bmatrix}$$

the trace is zero, $\mathrm{tr}\left(\mathbf{H}^{-1}\mathbf{J}\right) = 0$.

Barrier to Asymptotic Stability

- Suppose the dynamics around a fixed point is **traceless:** $\mathrm{tr}\Big(\nabla T(x^\star)\Big) = 0.$

Barrier to Asymptotic Stability

- Suppose the dynamics around a fixed point is **traceless:** $\mathrm{tr}\Big(\nabla T(x^\star)\Big) = 0.$
  - Recall that the trace is equal to the sum of the eigenvalues:

$$\mathrm{tr}\Big(\nabla T(x^\star)\Big) = \lambda_1 + \cdots + \lambda_N\,.$$

Barrier to Asymptotic Stability

- Suppose the dynamics around a fixed point is **traceless:** $\mathrm{tr}\Big(\nabla T(x^\star)\Big) = 0.$
  - Recall that the trace is equal to the sum of the eigenvalues:

$$\mathrm{tr}\Big(\nabla T(x^\star)\Big) = \lambda_1 + \cdots + \lambda_N\,.$$

- Asymptotic stability requires that the real part $\Re(\lambda_n) < 0$ for all $n \in [N]$.

Barrier to Asymptotic Stability

- Suppose the dynamics around a fixed point is **traceless:** $\mathrm{tr}\Big(\nabla T(x^\star)\Big) = 0.$
  - Recall that the trace is equal to the sum of the eigenvalues:

$$\mathrm{tr}\Big(\nabla T(x^\star)\Big) = \lambda_1 + \cdots + \lambda_N\,.$$

- Asymptotic stability requires that the real part $\Re(\lambda_n) < 0$ for all $n \in [N]$.
  - If some λ_n has a negative real part, then another λ_m must have positive real part. Then, the dynamics are **unstable**.

Barrier to Asymptotic Stability

- Suppose the dynamics around a fixed point is **traceless:** $\mathrm{tr}\Big(\nabla T(x^\star)\Big) = 0.$
  - Recall that the trace is equal to the sum of the eigenvalues:

$$\mathrm{tr}\Big(\nabla T(x^\star)\Big) = \lambda_1 + \cdots + \lambda_N\,.$$

- Asymptotic stability requires that the real part $\Re(\lambda_n) < 0$ for all $n \in [N]$.
  - If some λ_n has a negative real part, then another λ_m must have positive real part. Then, the dynamics are **unstable**.

This is the very high-level argument used across many works showing that certain dynamics fail to converge stably to interior Nash equilibria.

Linear Stability from Spectral Analysis

Consider the eigenvalues of the Jacobian of a learning dynamics.

- ~~If all eigenvalues have **negative real parts** $\Longrightarrow$ asymptotic stability~~
- If there is an eigenvalue with **positive real part** $\Longrightarrow$ instability
- If all eigenvalues are **purely imaginary** $\Longrightarrow$ neutral stability

What Remains?

Neutral stability (purely imaginary eigenvalues)

This talk

arXiv > cs > arXiv:2510.14907

Learnable Mixed Nash Equilibria are Collectively Rational

Geelon So, Yi-An Ma

We extend the study of learning in games to dynamics that exhibit non-asymptotic stability. We do so through the notion of uniform stability, which is concerned with equilibria of individually utility-seeking dynamics. Perhaps surprisingly, it turns out to be closely connected to economic properties of collective rationality. Under mild non-degeneracy conditions and up to strategic equivalence, if a mixed equilibrium is not uniformly stable, then it is not weakly Pareto optimal: there is a way for all players to improve by jointly deviating from the equilibrium. On the other hand, if it is locally uniformly stable, then the equilibrium must be weakly Pareto optimal. Moreover, we show that uniform stability determines the last-iterate convergence behavior for the family of incremental smoothed best-response dynamics, used to model individual and corporate behaviors in the markets. Unlike dynamics around strict equilibria, which can stabilize to socially-inefficient solutions, individually utility-seeking behaviors near mixed Nash equilibria lead to collective rationality.

I. Optimization

II. Definitions

III. Existing Results

IV. Uniform Stability

V. Conclusion

Uniform Stability: Non-Asymptotic Stability

Definition. The game Jacobian $\mathbf{J}(\mathbf{x})$ is **uniformly stable** at $\mathbf{x}$ if the eigenvalues of $\mathbf{H}^{-1}\mathbf{J}(\mathbf{x})$ are purely imaginary for all positive-definite block-diagonal matrix $\mathbf{H}$,

$$\operatorname{spec}\left(\mathbf{H}^{-1}\mathbf{J}(\mathbf{x})\right) \subset i\mathbb{R}.$$

Uniform Stability: Non-Asymptotic Stability

Definition. The game Jacobian $\mathbf{J}(\mathbf{x})$ is **uniformly stable** at $\mathbf{x}$ if the eigenvalues of $\mathbf{H}^{-1}\mathbf{J}(\mathbf{x})$ are purely imaginary for all positive-definite block-diagonal matrix $\mathbf{H}$,

$$\operatorname{spec}\left(\mathbf{H}^{-1}\mathbf{J}(\mathbf{x})\right) \subset i\mathbb{R}.$$

- A Nash equilibrium $\mathbf{x}^\star$ is pointwise uniformly stable if $\mathbf{J}(\mathbf{x}^\star)$ is uniformly stable.

Uniform Stability: Non-Asymptotic Stability

Definition. The game Jacobian $\mathbf{J}(\mathbf{x})$ is **uniformly stable** at $\mathbf{x}$ if the eigenvalues of $\mathbf{H}^{-1}\mathbf{J}(\mathbf{x})$ are purely imaginary for all positive-definite block-diagonal matrix $\mathbf{H}$,

$$\operatorname{spec}\left(\mathbf{H}^{-1}\mathbf{J}(\mathbf{x})\right) \subset i\mathbb{R}.$$

- A Nash equilibrium $\mathbf{x}^\star$ is pointwise uniformly stable if $\mathbf{J}(\mathbf{x}^\star)$ is uniformly stable.
- It is locally uniformly stable if there is an open set U containing $\mathbf{x}^\star$ such that $\mathbf{J}(\mathbf{x})$ is uniformly stable for all $\mathbf{x} \in U$.

Uniform Stability: Motivation

Definition. The game Jacobian $\mathbf{J}(\mathbf{x})$ is **uniformly stable** at $\mathbf{x}$ if the eigenvalues of $\mathbf{H}^{-1}\mathbf{J}(\mathbf{x})$ are purely imaginary for all positive-definite block-diagonal matrix $\mathbf{H}$,

$$\operatorname{spec}\left(\mathbf{H}^{-1}\mathbf{J}(\mathbf{x})\right) \subset i\mathbb{R}.$$

- The imaginary spectrum captures the only chance for the dynamics to exhibit neutral stability (instead of instability).

Uniform Stability: Motivation

Definition. The game Jacobian $\mathbf{J}(\mathbf{x})$ is **uniformly stable** at $\mathbf{x}$ if the eigenvalues of $\mathbf{H}^{-1}\mathbf{J}(\mathbf{x})$ are purely imaginary for all positive-definite block-diagonal matrix $\mathbf{H}$,

$$\operatorname{spec}\left(\mathbf{H}^{-1}\mathbf{J}(\mathbf{x})\right) \subset i\mathbb{R}.$$

- The imaginary spectrum captures the only chance for the dynamics to exhibit neutral stability (instead of instability).
- The quantifier "for all $\mathbf{H}$" expands the notion of *uncoupledness*: players do not know how others learn from their experiences.

Technical Detail: Strategic Equivalence

All of the dynamics mentioned factor through *strategic equivalence*:

- Take two utilities f_n and $f'_n = f_n + g_n$ where $g_n(\mathbf{x}_{-n})$ does not depend on x_n. The dynamics for player n are the same for both. For example:

$$\nabla_n f_n \equiv \nabla_n \left(f_n + g_n \right).$$

- Parts of the utility can be “external” to the strategic component of the game.
  - We focus on **purely strategic** utilities (or, define strategic Pareto optimality to depend only on the strategic component of the utility).

The Meaning of Uniform Stability

Local Uniform Stability implies Pareto Optimality

Theorem. If $\mathbf{x}^\star$ be a mixed Nash equilibrium, then:

local uniform stability $\Longrightarrow$ strategic Pareto optimality.

A Simpler Class: Polymatrix Games

In a polymatrix game, each utility f_n decomposes as:

$$f_n(x) - f_n(x^\star) = \sum_{m \neq n} (x_n - x_n^\star)^\top J_{nm} (x_m - x_m^\star).$$

A Simpler Class: Polymatrix Games

In a polymatrix game, each utility f_n decomposes as:

$$f_n(x) - f_n(x^\star) = \sum_{m \neq n} (x_n - x_n^\star)^\top J_{nm}(x_m - x_m^\star).$$

- Players are simply playing multiple two-player games at once.

A Simpler Class: Polymatrix Games

In a polymatrix game, each utility f_n decomposes as:

$$f_n(x) - f_n(x^\star) = \sum_{m \neq n} (x_n - x_n^\star)^\top J_{nm} (x_m - x_m^\star).$$

- Players are simply playing multiple two-player games at once.
- There are only bilinear interactions, and no higher-order interactions.

Uniform Stability implies Pareto Optimality

Theorem (polymatrix version). Consider a purely strategic, polymatrix game. Let $\mathbf{x}^\star$ be a mixed Nash equilibrium. Then:

uniform stability $\Longrightarrow$ Pareto optimality.

Proof (By Contradiction)

1. **Suppose that $x^\star$ is not weakly Pareto optimal, so there exists x so that**

$$f_n(x) - f_n(x^\star) = \sum_{m \neq n} (x_n - x_n^\star) J_{nm} (x_m - x_m^\star) > 0.$$

Proof (By Contradiction)

1. **Suppose that $x^\star$ is not weakly Pareto optimal, so there exists x so that**

$$f_n(x) - f_n(x^\star) = \sum_{m \neq n} (x_n - x_n^\star) J_{nm} (x_m - x_m^\star) > 0.$$

2. **Let $u_n = x_n - x_n^\star$ and $v_n = \sum_{m \neq n} J_{nm}(x_m - x_m^\star)$.**

Proof (By Contradiction)

1. **Suppose that $x^\star$ is not weakly Pareto optimal, so there exists x so that**

$$f_n(x) - f_n(x^\star) = \sum_{m \neq n} (x_n - x_n^\star) J_{nm} (x_m - x_m^\star) > 0.$$

2. **Let $u_n = x_n - x_n^\star$ and $v_n = \sum_{m \neq n} J_{nm}(x_m - x_m^\star)$.**

3. **There exists H_n such that $u_n = H_n^{-1} v_n$.**

Proof (By Contradiction)

1. **Suppose that $x^\star$ is not weakly Pareto optimal, so there exists x so that**

$$f_n(x) - f_n(x^\star) = \sum_{m \neq n} (x_n - x_n^\star) J_{nm} (x_m - x_m^\star) > 0.$$

2. **Let $u_n = x_n - x_n^\star$ and $v_n = \sum_{m \neq n} J_{nm}(x_m - x_m^\star)$.**

3. **There exists H_n such that $u_n = H_n^{-1} v_n$.**

4. **It follows that $\mathbf{H}^{-1}\mathbf{J}u = u$ where $\mathbf{H} = \mathrm{diag}(H_1, \ldots, H_N)$.**

Proof (By Contradiction)

1. **Suppose that $x^\star$ is not weakly Pareto optimal, so there exists x so that**

$$f_n(x) - f_n(x^\star) = \sum_{m \neq n} (x_n - x_n^\star) J_{nm} (x_m - x_m^\star) > 0.$$

2. **Let $u_n = x_n - x_n^\star$ and $v_n = \sum_{m \neq n} J_{nm}(x_m - x_m^\star)$.**

3. **There exists H_n such that $u_n = H_n^{-1} v_n$.**

4. **It follows that $\mathbf{H}^{-1}\mathbf{J}u = u$ where $\mathbf{H} = \mathrm{diag}(H_1, \ldots, H_N)$.**

5. **The spectrum of $\mathbf{H}^{-1}\mathbf{J}$ contains the non-imaginary eigenvalue 1.** □

Thus, it is not uniformly stable.

Uniform Stability implies Pareto Optimality

Theorem (polymatrix version). Consider a purely strategic, polymatrix game. Let $\mathbf{x}^\star$ be a mixed Nash equilibrium. Then:

uniform stability $\Longrightarrow$ Pareto optimality.

On General Games

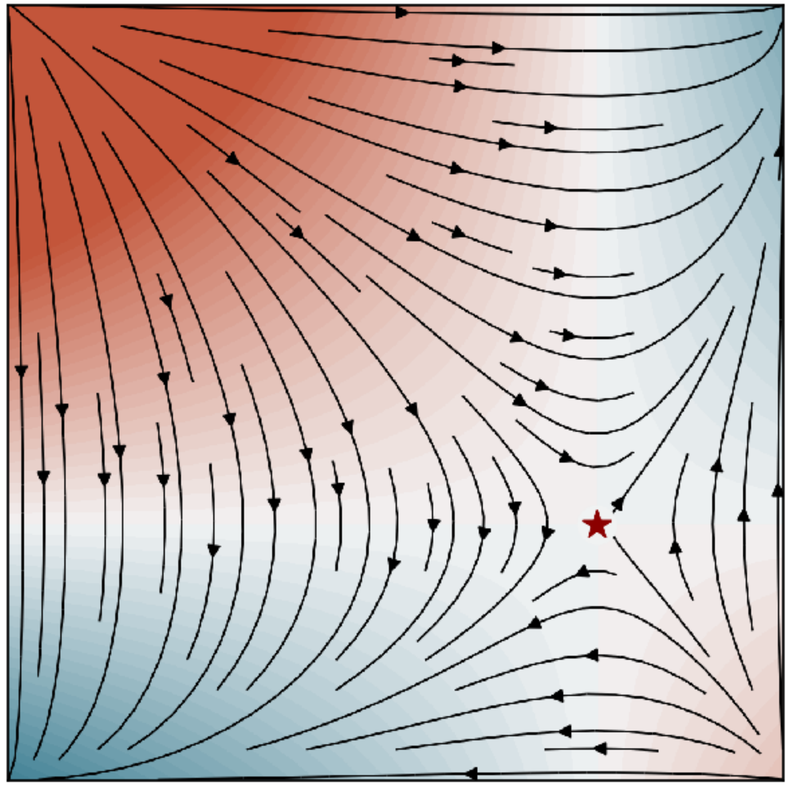
In general-sum games, there may be higher-order interactions.

- It turns out that local uniform stability imposes a lot of structure on the utilities.
- Proof requires quite a bit more machinery.

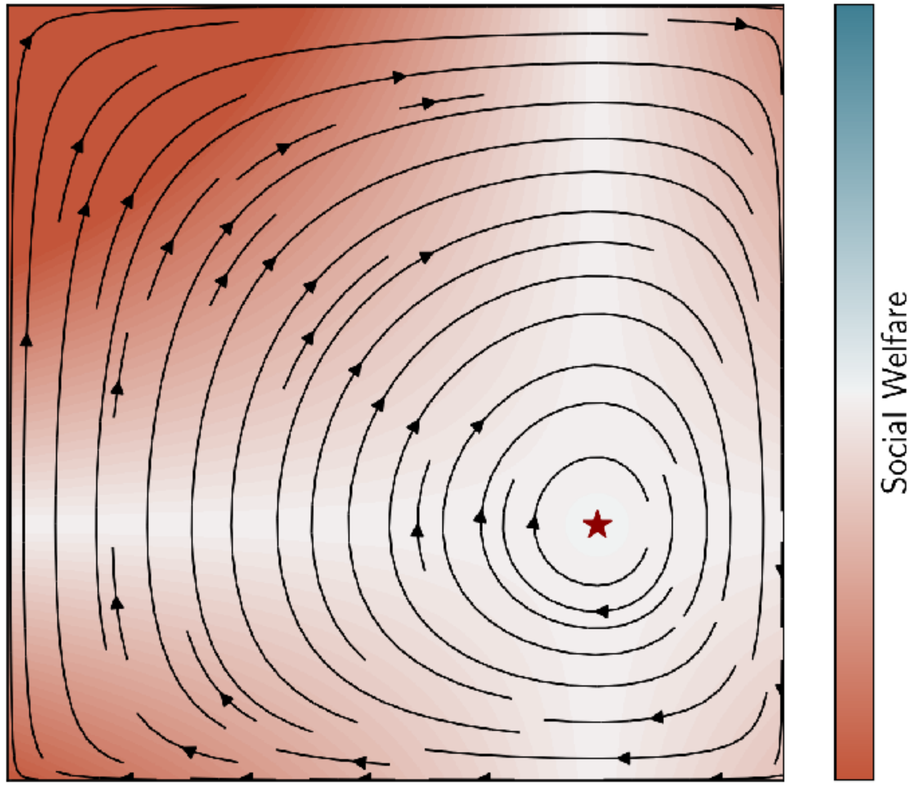
An Important Class: Zero-Sum Games

- Zero-sum games (and generalizations beyond 2 players) are locally uniformly stable.
- This helps explain why convergent dynamics could be designed for zero-sum games, despite general negative results.

Comparison of Dynamics Around Non-Strict NE



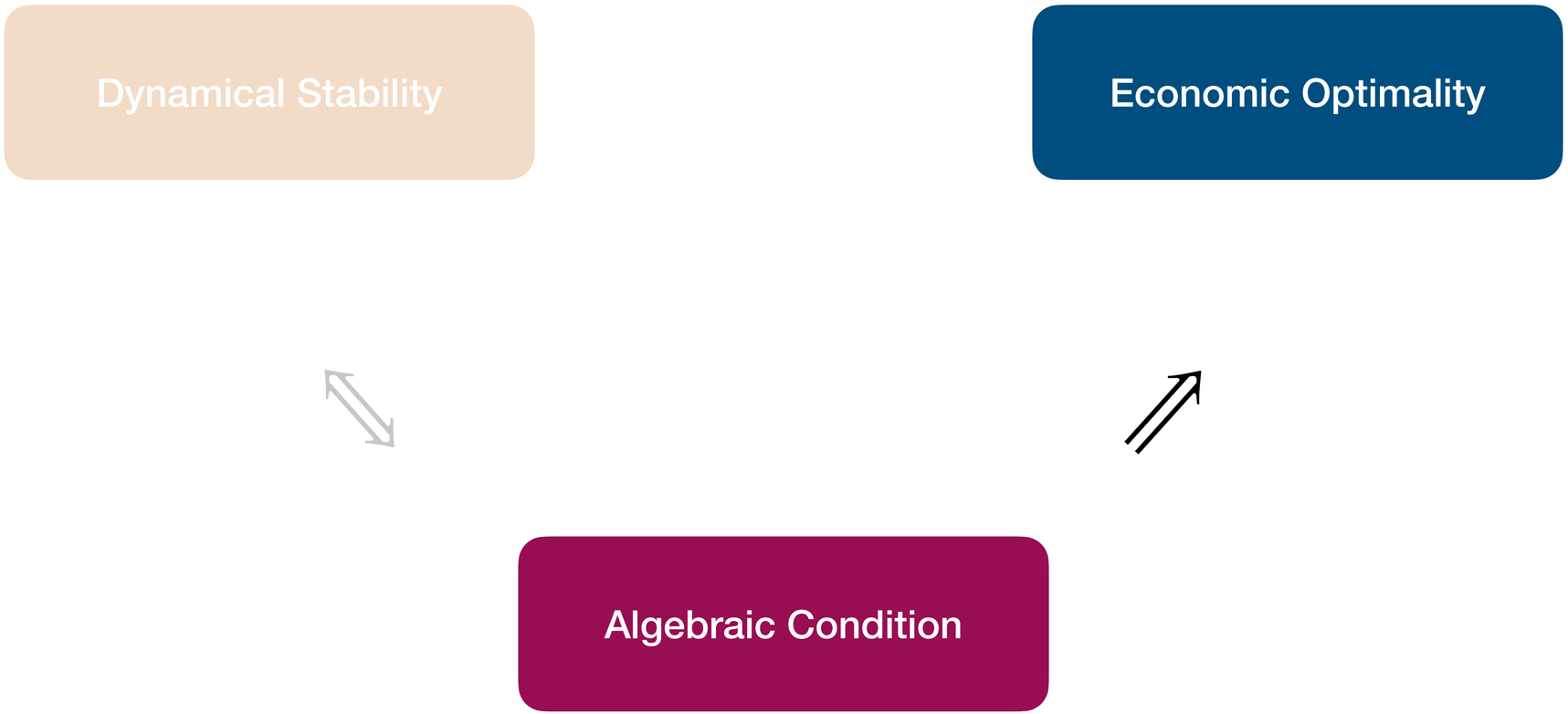
Strategically dominated solutions are not uniformly stable.



Uniformly stable Nash equilibria are strategically Pareto optimal.

Uniform Stability for Games

If a Nash equilibrium is locally uniformly stable, then it is collectively rational.



Convergence Under Uniform Stability

Smoothed Best-Response Dynamics

- Smoothed best-response dynamics

$$x_n(t+1) = \mathrm{SBR}_n^{\beta}\big(x(t)\big)$$

Smoothed Best-Response Dynamics

- Smoothed best-response dynamics

$$x_n(t+1) = \mathrm{SBR}_n^{\beta}\big(x(t)\big)$$

$$\mathrm{SBR}_n^{\beta}\big(x\big) = \arg\max_{z_n \in \Omega_n} f_n(z_n; x_{-n}) - \beta h_n(z_n)$$

Smoothed Best-Response Dynamics

- Smoothed best-response dynamics

$$x_n(t+1) = \mathrm{SBR}_n^{\beta}\big(x(t)\big)$$

  - h_n is a convex regularization term (e.g. negative entropy leads to softmax)

$$\mathrm{SBR}_n^{\beta}\big(x\big) = \arg\max_{z_n \in \Omega_n} f_n(z_n; x_{-n}) - \beta h_n(z_n)$$

Smoothed Best-Response Dynamics

- Smoothed best-response dynamics

$$x_n(t+1) = \mathrm{SBR}_n^{\beta}\big(x(t)\big)$$

  - h_n is a convex regularization term (e.g. negative entropy leads to softmax)
  - β is a smoothing/temperature parameter ($\beta \to 0$ gives less smoothing)

$$\mathrm{SBR}_n^{\beta}\big(x\big) = \arg\max_{z_n \in \Omega_n} f_n(z_n; x_{_n}) - \beta h_n(z_n)$$

Smoothed Best-Response Dynamics

- Smoothed best-response dynamics

$$x_n(t+1) = \mathrm{SBR}_n^{\beta}\big(x(t)\big)$$

  - h_n is a convex regularization term (e.g. negative entropy leads to softmax)
  - β is a smoothing/temperature parameter ($\beta \to 0$ gives less smoothing)
    - Each β gives rise to an $O(\beta)$-approximate Nash equilibrium

$$\mathrm{SBR}_n^{\beta}\big(x\big) = \arg\max_{z_n \in \Omega_n} f_n(z_n; x_{-n}) - \beta h_n(z_n)$$

Smoothed Best-Response Dynamics

- Smoothed best-response dynamics

$$x_n(t+1) = \mathrm{SBR}_n^{\beta}\big(x(t)\big)$$

  - h_n is a convex regularization term (e.g. negative entropy leads to softmax)
  - β is a smoothing/temperature parameter ($\beta \to 0$ gives less smoothing)
    - Each β gives rise to an $O(\beta)$-approximate Nash equilibrium
  - Smoothed best response is a model of “bounded rationality” in economics

$$\mathrm{SBR}_n^{\beta}\big(x\big) = \arg\max_{z_n \in \Omega_n} f_n(z_n; x_{-n}) - \beta h_n(z_n)$$

Smoothed Best-Response Dynamics

- Incremental smoothed best-response dynamics

$$x_i(t+1) = (1-\eta_t)x_i(t) + \eta_t \cdot \mathrm{SBR}_i^{\beta_t}(x(t))\,.$$

Smoothed Best-Response Dynamics

- Incremental smoothed best-response dynamics

$$x_i(t+1) = (1-\eta_t)x_i(t) + \eta_t \cdot \mathrm{SBR}_i^{\beta_t}(x(t))\,.$$

  - Two parameters:
    - β_t is the amount of smoothing
    - η_t is the learning rate

Smoothed Best-Response Dynamics

- Incremental smoothed best-response dynamics

$$x_i(t+1) = (1-\eta_t)x_i(t) + \eta_t \cdot \mathrm{SBR}_i^{\beta_t}(x(t))\,.$$

  - Two parameters:
    - β_t is the amount of smoothing
    - η_t is the learning rate
  - Population view: an η_t-fraction of the population updates their strategy

Convergence Under Uniform Stability

Theorem. Let $\mathbf{x}^\star$ be an isolated, non-degenerate, locally uniformly stable mixed Nash equilibrium. There is a open set U such that if $x(0) \in U$, then the following holds. There is a sequence of $(\beta_t)_t$ and $(\eta_t)_t$ such that iterates of the smoothed best-response dynamics $x(t)$ are ε-approximate Nash equilibria when

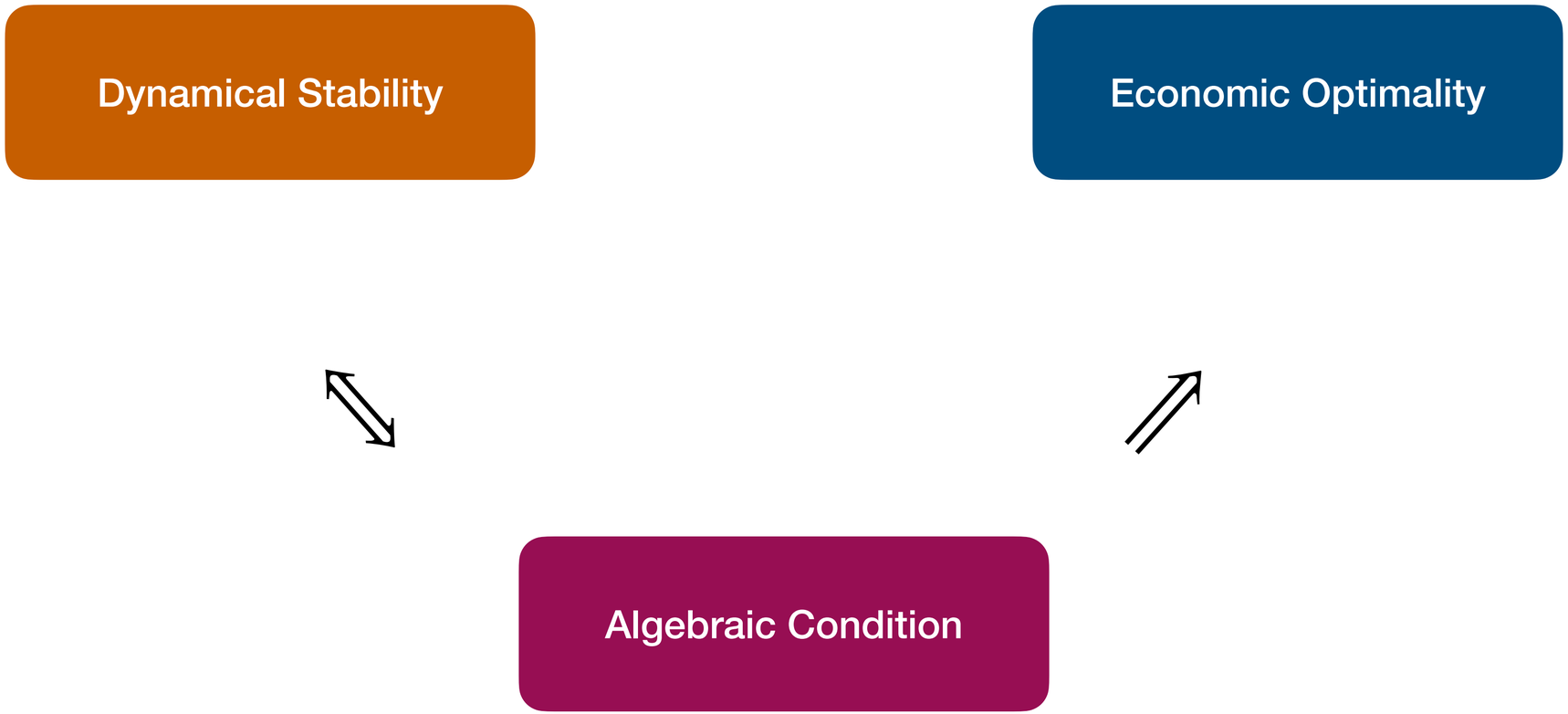
$$t = \Omega\left(\frac{1}{\varepsilon^3}\log\frac{1}{\varepsilon}\right).$$

Partial Converse

Proposition. Suppose $\mathbf{x}^\star$ is not uniformly stable. Then, there are no convergent smoothed best-response dynamics where $\beta_t \to 0$.

Uniform Stability for Games

If a Nash equilibrium is locally uniformly stable, then it is collectively rational.



I. Optimization

II. Definitions

III. Existing Results

IV. Uniform Stability

V. Conclusion

Main Takeaways

- Asymptotic stability is perhaps too coarse to analyze game dynamics.
  - The interesting phenomena is on the boundary of stability and instability.
  - This work introduces the non-asymptotic stability analysis for games.
- What sort of mathematical language will help us design and control complex, multi-agent systems with learned behavior?

Thanks!