



An impossibility theorem in game dynamics

Jason Milionis^{a,1}, Christos Papadimitriou^{a,1}, Georgios Piliouras^{b,1}, and Kelly Spendlove^{c,1}

Contributed by Christos Papadimitriou; received April 8, 2023; accepted July 31, 2023; reviewed by J. Doyne Farmer and Vasilis Syrgkanis

The Nash equilibrium—a, combination of choices by the players of a game from which no self-interested player would deviate—is the predominant solution concept in game theory. Even though every game has a Nash equilibrium, it is not known whether there are deterministic behaviors of the players who play a game repeatedly that are guaranteed to converge to a Nash equilibrium of the game from all starting points. If one assumes that the players' behavior is a discrete-time or continuous-time rule whereby the current mixed strategy profile is mapped to the next, this question becomes a problem in the theory of dynamical systems. We apply this theory, and in particular Conley index theory, to prove a general impossibility result: There exist games, for which all game dynamics fail to converge to Nash equilibria from all starting points. The games which help prove this impossibility result are degenerate, but we conjecture that the same result holds, under computational complexity assumptions, for nondegenerate games. We also prove a stronger impossibility result for the solution concept of approximate Nash equilibria: For a set of games of positive measure, no game dynamics can converge to the set of approximate Nash equilibria for a sufficiently small yet substantial approximation bound. Our results establish that, although the notions of Nash equilibrium and its computation-inspired approximations are universally applicable in all games, they are fundamentally incomplete as predictors of long-term player behavior.

game theory | dynamical systems | Nash equilibrium | solution concept

The Nash equilibrium, defined and shown universal by John F. Nash in 1950 (1), is paramount in game theory, routinely considered as the default solution concept—the “meaning of the game.” Over the years—and especially in the past three decades during which game theory has come under intense computational scrutiny—the Nash equilibrium has been noted to suffer from a number of disadvantages of a computational nature. There are no known efficient algorithms for computing the Nash equilibrium of a game, and in fact, the problem has been shown to be intractable (2–4). In addition, there are typically many Nash equilibria in a game, and the selection problem leads to conceptual complications and further intractability; see, for example, ref. 5.

A common defense of the Nash equilibrium in the face of such shortcomings is the informal argument that “the players will eventually get there.” However, no learning behavior has been shown to converge to equilibrium in all games (6–13). In fact, refs. 8 and 9 show that for wide classes of games, numerous learning dynamics fail to converge to a fixed point of any kind, and instead exhibit chaotic dynamics. The celebrated work of Hart and Mas-Colell (14) shows that even in games with a single Nash equilibrium, learning dynamics which are uncoupled (players only know their own payoff function) fail to converge to the Nash equilibrium. However, the uncoupledness assumption is strong and crucial for this result, and the problem remains open for general dynamics.

Given a game, the deterministic way players move from one mixed strategy profile to the next can be formalized in terms of dynamical systems as a continuous function φ assigning to each point x in the strategy space and each time $t \geq 0$ another point $\varphi(t, x)$: the point where the players will be after time t ; naturally, this function must satisfy $\varphi(t', \varphi(t, x)) = \varphi(t + t', x)$. If $t \in \mathbb{R}$, this setup is called continuous time dynamics, while if $t \in \mathbb{Z}$, it is referred to as discrete time dynamics. For φ to qualify as game dynamics for a game g , it must satisfy Nash stationarity: The Nash equilibria are a subset of the fixed points of φ . This is a reasonable assumption, since game dynamics should encode the strategic stability of Nash equilibria, i.e., no rational player will ever move away from a Nash equilibrium.

Having defined what we mean by “game dynamics,” we turn to defining “convergence.” In the theory of dynamical systems, a standard tool for convergence is Lyapunov functions, which provide an intuitive optimization point of view in the analysis of long-term behavior. Informally, these are real-valued functions that are always decreasing except when the dynamics have reached a final set of configurations. The standard

Significance

The Nash equilibrium is the fundamental solution concept of game theory, playing a central role in fields beyond economics. Nash's existence proof is famously nonconstructive, and the dynamics employed in the proof fail to converge to a Nash equilibrium. We show that this failure is a special case of a much more sweeping impossibility result for any game dynamics satisfying a minimal set of desiderata. Thus, the solution concept of Nash equilibria, while universally applicable, is also simultaneously a fundamentally incomplete description of the possible long-term agent behavior.

Author affiliations: ^aDepartment of Computer Science, Columbia University, New York, NY 10027; ^bGoogle DeepMind, London EC4A 3TW, United Kingdom; and ^cGoogle, Mountain View, CA 94043

Author contributions: C.P., G.P., and K.S. designed research; J.M., C.P., G.P., and K.S. performed research; and J.M., C.P., G.P., and K.S. wrote the paper.

Reviewers: J.D.F., University of Oxford and V.S., Stanford University.

The authors declare no competing interest.

Copyright © 2023 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

¹To whom correspondence may be addressed. Email: christos@columbia.edu, jm@cs.columbia.edu, gpil@deepmind.com, or spendlove@google.com.

Published October 5, 2023.

example of such dynamics is gradient flows, where the dynamics move in the direction of the steepest descent, and the final configurations are the critical points of the function. The surprising power of Lyapunov arguments lies in explaining the behavior of dynamics that are not gradient-like. Specifically, an arbitrary dynamical system can be described by a decomposition of its domain into the chain recurrent components, on which the dynamics stay within the component, and outside of which the dynamics flows from one component to another (15), guided by a Lyapunov function: a function that is constant along orbits within the chain recurrent components, and strictly decreasing outside of them.

In view of this mathematical framework, we are interested in this question: Are there game dynamics which converge to Nash equilibria for all games and from all starting points? Our main result is an impossibility theorem:

Main Result (informal statement) there are games in which any game dynamics will fail to converge to the set of Nash equilibria.

That is to say, we exhibit games in which any game dynamics must fail to converge to the set of Nash equilibria—or even approximate Nash equilibria—from certain starting points. This suggests that the Nash equilibrium concept is plagued by a form of incompleteness: It is incapable of capturing the long-term behaviors of the players in all games.

This result complements existing literature, in which the nonexistence of Nash-convergent dynamics is shown from stronger assumptions (14, 16). Also note that our impossibility result is complementary to the known intractability of the Nash equilibrium (2–4) and cannot be derived from it. Indeed, it is well known that there are dynamical systems whose next-step function can be computed efficiently and which converge to points of extremely high computational complexity (PSPACE-complete); see, for example, ref. 17; therefore, computational complexity in and by itself does not preclude convergent dynamics.

Two objections can be raised to the impossibility result: Degenerate games (as we use in our example) are known to have measure zero—so are there dynamics that work for almost all games? (As we shall see soon, the answer is “yes, but.”) Second, in view of the known intractability of the Nash equilibrium, exact equilibria may be asking too much; are there dynamics that converge to an arbitrarily good approximation of the Nash equilibrium? We show how our result can be extended so that it addresses these questions.

For the first question, it turns out that, in some sense, degeneracy is required for the impossibility result: We give an algorithm which, given any nondegenerate game, specifies somewhat trivial game dynamics which are Nash convergent. The downside of this positive result is that the algorithm requires exponential time unless P = PPAD—a complexity collapse that is considered unlikely; see, for example, ref. 2. In fact, we conjecture that such intractability is inherent. In other words, we suspect that, in nondegenerate games, it is complexity theory, and not topology, that provides the proof of the impossibility result. Proving this conjecture would require the development of a complexity-theoretic treatment of dynamics, which seems to us a very attractive research direction.

Second, we exhibit a family of games, with nonzero measure (in particular, perturbations of the game used for our main result), for which any game dynamics will fail to converge (in the above sense) to the set of ϵ -approximate Nash equilibria, for some fixed $\epsilon > 0$. Note that, because of the nonzero measure claim, this result appears to be stronger than our main result; however, this is not true because it relies on the stronger assumption that all approximate equilibria are stationary.

Our impossibility result also applies to player behaviors with memory, that is, one in which the players maintain and update quantities which affect their choices. If memory is modeled by expanding the space to include memory components, then impossibility can be recovered under assumptions, even though this requires a more detailed treatment. An interesting example is recent work on optimistic dynamics, where agents keep track of the last two periods of play (18–20). Exploring the limits of the applicability of the impossibility result to such dynamics is an intriguing direction for future work.

Finally, we note that our impossibility results do not apply to stochastic dynamics—e.g., discrete-time dynamics in which $\varphi(x)$ is a distribution of possible next points. In fact, it is easy to see that even trivial such schemes—e.g., “from any point x jump to a uniformly random point in the domain”—succeed in eventually reaching an arbitrarily good approximation of a Nash equilibrium, albeit in exponential time.

Game Dynamics and Nash Convergence

Before turning to our notion of convergence, we shall discuss what it means for a dynamical system φ to be compatible with a game g . Let us denote the Nash equilibria of g as $NE(g)$ or simply NE when the context is clear. For nondegenerate games, there are an odd number of isolated Nash equilibria. In general, however, the Nash equilibria of a game are not isolated, and instead form closed, connected components, of which there are finitely many (21). That is, NE is the disjoint union $NE = \bigsqcup_{1 \leq i \leq n} NE_i$, where NE_i is a component. Recall that a Nash equilibrium is by definition a point from which no player will want to deviate. Hence, for a dynamics φ to be compatible with game g , we only require that φ satisfy Nash stationarity: $NE(g) \subseteq Fix(\varphi)$, where by $Fix(\varphi)$, we denote the fixed points of φ , the set $\{x \in X : \varphi(t, x) = x \text{ for all } t \geq 0\}$. That is, the dynamics is compatible with a game if it does not move away from any of the Nash equilibria of the game. Namely, game dynamics encode the strategically stability of all Nash equilibria, i.e., that self-interested agents would not deviate from them. This is a very natural definition, and game dynamics have been studied extensively under this assumption, or variants; see for instance, refs. 21–27; these studies often result in a characterization of Nash components in terms of fixed-point index theory.

Now, given such a dynamical system, what is an appropriate notion of convergence? This is a deep question within topological dynamical systems theory, and a classical tool used to characterize convergence is that of Lyapunov functions. In fact, Lyapunov function arguments are the archetypal strategy for proving convergence to Nash equilibria in games, cf. refs. 28–34. Furthermore, the topological theory of dynamical systems implies that any dynamical system has an appropriate Lyapunov function, as the existence of a complete Lyapunov function is guaranteed by Conley’s fundamental theorem of dynamical systems (15): This theorem states that any dynamical system can be decomposed into the so-called chain recurrent set—the set to which the dynamics converges and on which the Lyapunov function is constant—and a gradient-like part, upon which a Lyapunov function strictly decreases along trajectories; see refs. 15 and 35. This leads us to our definition:

Definition 1: Given game dynamics φ for a game g , we say that φ is Nash convergent if there exists a continuous function $V : X \rightarrow [0, 1]$ such that

1. if $x \notin NE$ and $t > 0$ then $V(\varphi(t, x)) < V(x)$,
2. if $x, y \in NE_i$ for some i , then $V(x) = V(y)$.

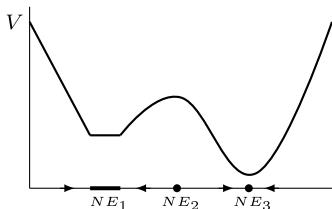


Fig. 1. Rendering of Lyapunov function V for Nash convergent game dynamics φ . For game dynamics, the components of Nash equilibria are invariant sets. The function V is constant on the components while points which are not Nash equilibria lie on trajectories between components.

See Fig. 1 for an illustration. That is to say, V is a continuous function on X which is strictly decreasing on trajectories outside NE and constant on trajectories within any Nash component; cf. ref. 15.

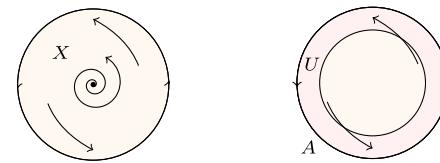
Now that we know what convergence means, the statement of our main result becomes clearer: There is a game g which does not admit any Nash convergent game dynamics. The techniques we use to prove this result come from Conley index theory (15) and require minimal technical assumptions on the space and dynamics. These techniques have not been previously applied to the setting of game dynamics, and provide a powerful framework for analysing the global dynamics of games, the implications of which will be explored in future work. Moreover, up to a basic knowledge of algebraic topology, our paper is self-contained, as all the necessary results are provided in the Notes.

Dynamical Systems Theory

To prove our result, we make use of the relationship between Lyapunov functions, attractors, and Conley indices. To this end, we introduce all these mathematical concepts and their connections to the extent that they are needed in the proof. We will consider a compact metric space X (for game dynamics, X is the product of the players' strategy simplices) and let $\mathbb{T}^+ = \{t \in \mathbb{T} : t \geq 0\}$ where \mathbb{T} is either \mathbb{R} or \mathbb{Z} . A dynamical system is a continuous map $\varphi : \mathbb{T}^+ \times X \rightarrow X$, such that i) $\varphi(0, x) = x$ for all $x \in X$ and ii) $\varphi(t + t', x) = \varphi(t', \varphi(t, x))$ for all $t, t' \in \mathbb{T}^+$ and $x \in X$. When $\mathbb{T} = \mathbb{R}$, i.e., for continuous-time, φ is called a semiflow; if $\mathbb{T} = \mathbb{Z}$, i.e., discrete time, φ is often called a map or discrete time dynamical system.

A basic notion in dynamical systems is that of an attractor, which intuitively is a set to which trajectories converge asymptotically. For a subset $Y \subset X$, the asymptotic behavior of Y under the dynamics φ is captured by the omega-limit set, formally defined as $\omega(Y) = \bigcap_{t>0} \text{cl}(\varphi([t, \infty), Y))$, where cl denotes topological closure. A set $A \subset X$ is an attractor if there exists a neighborhood U of A such that $\omega(U) = A$. An attracting block U is a closed neighborhood of A such that $\omega(U) = A$ and $\varphi(t, U) \subset \text{int } U$ for all $t > 0$. Attracting blocks always exist (36). Finally, the maximal attractor is $\omega(X)$.

Conley Index Theory for Attractors. Conley index theory deals with the more general concept of an isolated invariant set; however, for our purposes, we need only introduce the Conley index in the limited setting of attractors. The Conley index takes a different form depending upon whether the time is discrete or continuous, and we introduce each as well as prove, separately, both corresponding theorems. We defer the formal exposition of Conley index theory to the Notes, providing for now an intuition for the continuous-time case.



$$CH_i(\omega(X)) = \begin{cases} \mathbb{Z} & i = 0 \\ 0 & i \geq 1 \end{cases} \quad CH_i(A) = \begin{cases} \mathbb{Z} & i = 0, 1 \\ 0 & i \geq 2 \end{cases}$$

Fig. 2. A disk X with semiflow φ , attractor A (stable periodic orbit) and an unstable fixed point in the interior (Left). Here, $\omega(X) = X$, and the Conley index of $\omega(X)$ is the homology of a point. Attracting block U for A which is homotopy equivalent to a circle (Right). The Conley index of A is the homology of a circle.

Let $\varphi : \mathbb{T}^+ \times X \rightarrow X$ be a semiflow, A be an attractor for φ , and U an attracting block of A . The Conley index of A is defined as:

$$CH_\bullet(A) := H_\bullet(U).$$

Here, H_\bullet denotes singular homology with integer coefficients. As shown in Proposition 1 of the Notes, the Conley index is independent of the particular choice of U and, thus, provides an algebraic topological invariant for A . As X is an attracting block for the maximal attractor, $\omega(X)$, its Conley index is defined as:

$$CH_\bullet(\omega(X)) := H_\bullet(X).$$

Fig. 2 provides elementary Conley index computations for a semiflow.

Impossibility in Game Dynamics

In this section, we prove our main result, the impossibility theorem, splitting the proof into the cases of continuous time (Theorem 1) and discrete time (Theorem 2).

Theorem 1. *There exists a game g that does not admit continuous-time Nash convergent game dynamics.*

Proof: Define g to be the following bimatrix game as in (21, 25):

$$(R, C) = \begin{pmatrix} 1, 1 & 0, -1 & -1, 1 \\ -1, 0 & 0, 0 & -1, 0 \\ 1, -1 & 0, -1 & -2, -2 \end{pmatrix}. \quad [1]$$

By way of contradiction, suppose that there exist game dynamics φ which are Nash convergent. For g given in Eq. 1, the set of Nash equilibria forms a topological circle, i.e., S^1 (21). Thus, there is a single component of Nash equilibria, which is homeomorphic to S^1 . As φ is Nash convergent, there exists a Lyapunov function $V : X \rightarrow [0, 1]$ such that for any $x \notin NE$ and any $t > 0$, $V(x) > V(\varphi(t, x)) > 0$. Without loss of generality, we may assume that $V^{-1}(0) = NE$. For any $\lambda \in (0, 1]$, setting $U_\lambda = V^{-1}([0, \lambda])$, we have that $NE = \omega(U_\lambda)$. Thus, NE is an attractor for φ . In fact, as $X = V^{-1}([0, 1])$, and $NE = \omega(X)$, NE is the maximal attractor. It follows from the continuity of V that U_λ is a closed attracting block for NE for any $\lambda \in (0, 1]$. In particular, we can first calculate the Conley index of NE as:

$$CH_i(NE) = H_i(X) = \begin{cases} \mathbb{Z} & i = 0 \\ 0 & i > 0 \end{cases} \quad [2]$$

Now, the inclusion $i : NE \hookrightarrow U_\lambda$ induces the homomorphism on homology $i_\bullet : H_\bullet(NE) \rightarrow H_\bullet(U_\lambda)$. Intuitively, we can

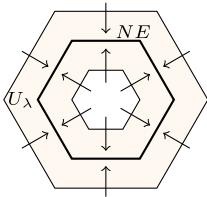


Fig. 3. Rendering of circle of Nash equilibria (NE lives in four dimensions); $NE = V^{-1}(0)$. Attracting block U_λ for attractor NE and Lyapunov function V ensures that the dynamics converges to NE .

choose λ small enough so that U_λ is a tight enough neighborhood of NE so that the circle of Nash equilibria which generates the 1-dimensional hole in $H_1(NE)$ still exists as a 1-dimensional hole in U_λ , as in Fig. 3. More formally, if w is the generator of $H_1(NE) = \mathbb{Z}$, for $\lambda > 0$, sufficiently small U_λ is a small neighborhood of NE and the image $\iota_1(w)$ cannot be the boundary of any 2-chain in U_λ . Thus, the map $\iota_1 : H_1(NE) \rightarrow H_1(U_\lambda)$ is injective. As U_λ is an attracting block, we can compute the Conley index of NE as:

$$CH_\bullet(NE) = H_\bullet(U_\lambda). \quad [3]$$

However, as the map ι_1 is injective, we have that

$$CH_1(NE) = H_1(U_\lambda) \neq 0. \quad [4]$$

However, by Proposition 1 the Conley index* is an invariant of NE and does not depend on the choice of attracting block. Therefore, this is a contradiction and such a φ cannot exist. \square

We now consider the discrete time case, which is proved in the same fashion, using the Conley index for maps, the background for which is provided in the Notes.

Theorem 2. *There exists a game g that does not admit discrete time Nash convergent game dynamics.*

Proof: Consider again the game of Eq. 1. Assuming that there are Nash convergent discrete time game dynamics for g , there exists a Lyapunov function $V : X \rightarrow [0, 1]$, and without loss of generality, we may assume $V^{-1}(0) = NE$. Once again, it follows that NE is an attractor, and in fact the maximal attractor of φ and for any $\lambda \in (0, 1]$, $U_\lambda = V^{-1}([0, \lambda])$ is an attracting block for NE . We will reach the desired contradiction by comparison of the Conley indices computed via X and U_λ for sufficiently small λ . First, as X is an attracting block for NE , we can use it to compute the Conley index for NE . Note that $H_\bullet(X)$ is computed in Eq. 2 and induced map $f_X : H_\bullet(X) \rightarrow H_\bullet(X)$ is the identity map (taking the generator of the connected component to itself). The Conley index of NE is given by $(L_\bullet(X), \chi_X)$, which is:

$$L_\bullet(X) = M_\bullet(X) = H_\bullet(X),$$

and the automorphism χ_X is the identity map. In particular, $L_1(X) = H_1(X) = 0$. Again, the inclusion $\iota : NE \hookrightarrow U_\lambda$ induces a map on homology $\iota_\bullet : H_\bullet(NE) \rightarrow H_\bullet(U_\lambda)$ and when λ is sufficiently small, the map $\iota_1 : H_1(NE) \rightarrow H_1(U_\lambda)$ is injective. Thus, $H_1(U_\lambda) \neq 0$. By Nash stationarity, NE is a circle of fixed points, and thus, the map $f(x) := \varphi(1, x)$ induces the identity map $f|_{NE} : H_\bullet(NE) \rightarrow H_\bullet(NE)$. In particular, there is a commutative diagram:

$$\begin{array}{ccc} H_1(NE) & \xrightarrow{id} & H_1(NE) \\ \downarrow \iota_1 & & \downarrow \iota_1 \\ H_1(U_\lambda) & \xrightarrow{f_{U_\lambda}} & H_1(U_\lambda). \end{array}$$

Letting w be the generator of $H_1(NE) = \mathbb{Z}$, and $w' = \iota_1(w)$, the diagram implies $f_{U_\lambda}(w') = w'$. It is elementary that this implies that $L_1(U_\lambda) \neq 0$. However, Proposition 2 implies that the Conley index is independent of attracting block. Thus, this is a contradiction, and there cannot exist such a φ . \square

Remark 1: An inspiration for our results is the work of Benaïm et al. (25), wherein they make the statement, without complete proof, that there are games g and such that for all φ within a class of multivalued dynamical systems, it cannot be the case that $NE(g) = CR(\varphi)$, where $CR(\varphi)$ is the chain recurrent set (see ref. 15) of φ . The argument sketched in ref. 25 ultimately rests on the development of a fixed point index for components of Nash equilibria and its comparison to the Euler characteristic of the component. However, the argument is incomplete (indeed, the reader is directed to two papers, one of which is a preprint that seems to have not been published). Instead, for our result, we assume Nash convergence, which is in fact a weaker condition as the chain recurrent set has such a Lyapunov function obeying (1) and (2), and thus, if $CR(\varphi) = NE$, φ is Nash convergent. Thus, as a corollary of Theorem 1, there exists a game which does not admit game dynamics for which the chain recurrent set coincides with the set of Nash equilibria.

Remark 2: In discussions with Gil Kalai and, independently, Venkat Anantharam, the following dynamics came up: In the two-dimensional disk $D = \{x, y : x^2 + y^2 \leq 1\}$, consider the continuous dynamics, described by the vector field $f(x, y) = (0, 1 - x^2 - y^2)$. Obviously, the only fixed points are the boundary of the disk, and there are no cycles. Now, homeomorphically map this to a topological disk T whose boundary is the hexagonal Nash component of g and move every other point of the domain toward T , smoothening appropriately the transition to the motion of T . Is this not then a dynamical system that converges to the Nash equilibria of g , contradicting our main result?

The topological theory of dynamical systems provides a negative answer: There can be no Lyapunov function supporting this dynamics, and the chain recurrent set of the dynamics is all of T , instead of its boundary. Indeed, a similar example is contained in ref. 15.

Nondegenerate Games and Approximate Equilibria

Nondegenerate Games. Our impossibility result in the previous section is constructed around a degenerate normal-form game with a continuum of equilibria. What if the game is nondegenerate?

Theorem 3. *For any nondegenerate game g , there is dynamical system φ_g satisfying Nash stationarity which is Nash-convergent.*

Proof: Since g is nondegenerate, it has an odd number of isolated Nash equilibria (37). Thus, in this case, the Nash equilibria themselves are the maximal components. Fix one such equilibrium and call it y ; e.g., $y = NE_0$. We next define φ_g in terms of y . We shall define it at point x implicitly, in terms of the direction of motion, and the speed of motion; if this is done, $\varphi_g(t, x)$ is easily computed through integration on t . The direction of motion is the unit vector of $y - x$: The dynamics

*For dynamics with memory, the impossibility result follows from the inequality between the Conley index of their statespace (e.g., two copies of a product of simplices in ref. 19) and their respective target set (e.g., two copies of the same Nash equilibrium).

heads to y . The speed of motion is defined to be $c \cdot D_g(x)$, where $c > 0$ is a constant, and by $D_g(x)$, we denote the deficit at x : the sum, over all players, of the difference between the best-response utility at x , and the actual utility at x . It is clear that $D_g(x) \geq 0$, and it becomes zero precisely at any Nash equilibrium; thus, φ_g has the property of Nash stationarity. Define $V(x) = \|x - y\|_2$ as the Lyapunov function. It is clear that V strictly decreases along trajectories off of NE (such trajectories head to y). Therefore, φ is Nash convergent. \square

Remark 3: It's worth remarking on why φ_g as constructed above is not Nash-convergent for the degenerate game in Eq. 1. In this case, there are points $x \notin NE$ that belong on orbits converging to the chosen Nash equilibrium y but come from a Nash equilibrium $z \neq y$. However, the Lyapunov function V must have $V(z) = V(y)$ as both z and y belong to the same component. Therefore, V could not be strictly decreasing along such trajectories. In fact, these dynamics are similar to the example dynamics discussed in the previous remark.

Now note that the algorithm for specifying φ_g requires finding y , a PPAD-complete (and FIXP-complete for more than two players) problem. We believe that the dependence is inherent:

Conjecture 1. *The computational task of finding, from the description of a game g , either a degeneracy of g or an algorithm producing the direction of motion and speed of a dynamical system satisfying Nash stationarity and Nash convergence is PPAD-hard (and FIXP-hard for three or more players).*

We believe this is an important open question in the boundary between computational complexity, game theory, and the topology of dynamical systems, whose resolution is likely to require the development of complexity-theoretic techniques pertinent to dynamical systems. We further suspect that, if the sought dynamics are constrained to be utility-increasing, i.e., that given stationary opponents the agents move in the direction of increasing utility whenever possible, the difficulty of finding such a Nash convergent dynamics is increased to NP-hard.

Approximate Nash Equilibria. Next, it may seem plausible that the difficulty of designing Nash convergent game dynamics can be overcome when only ϵ -approximation is sought, for some $\epsilon > 0$. Recall that an ϵ -Nash equilibrium is a mixed strategy profile in which all players' utilities are within an additive $\epsilon > 0$ of their respective best response. We show that our impossibility theorem extends to this case as well. Let us denote by $NE_\epsilon(g)$ the set of ϵ -approximate Nash equilibria of a game g ; there are a finite number of components of ϵ -approximate Nash, so that once again, $NE_\epsilon(g) = \bigsqcup_{1 \leq i \leq n} NE_{\epsilon,i}$. We say that a dynamical system φ has the property of ϵ -Nash stationarity with respect to a game g if $NE_\epsilon(g) \subset \text{Fix}(\varphi)$, and that φ is ϵ -Nash convergent if there is a Lyapunov function V that is constant on these components, and strictly decreasing along trajectories if $x \notin NE_\epsilon$. We begin with a lemma.

Lemma 1. *Consider the bimatrix game $g = (R, C)$ and $\epsilon > 0$. For any $\epsilon' > 0$ and game $g' = (R', C')$ such that $\|R - R'\|_F \leq \epsilon'$ and $\|C - C'\|_F \leq \epsilon'$ (where $\|\cdot\|_F$ denotes the Frobenius norm), we have $NE_\epsilon(g) \subset NE_{2\epsilon' + \epsilon}(g')$.*

Proof: The argument for the first player follows, as for any $(x, y) \in NE_\epsilon(g)$ any i :

$$\begin{aligned} e_i^T R' y^T - x^T R' y &= (e_i - x)^T (R' - R)y + e_i^T Ry - x^T Ry \\ &\leq \|e_i - x\|_2 \|R' - R\|_F \|y\|_2 + \epsilon \leq 2\epsilon' + \epsilon. \end{aligned}$$

The inequality for the second player is proved similarly. \square

Theorem 4. *There is a game g and an $\epsilon > 0$ so that g does not admit any dynamics that are both ϵ -Nash stationary and ϵ -Nash convergent. In fact, the set of such games has positive measure.*

Proof: Consider again the Kohlberg–Mertens game Eq. 1. Define the function

$$h_g(z, y) = \max_i \max_j (e_i^T Ry - z^T Ry, z^T Ce_i - z^T Cy).$$

Then, h_g is a continuous function such that for any $\epsilon > 0$, we have that $NE_\epsilon(g) = h_g^{-1}([0, \epsilon])$; in particular, $NE(g) = h_g^{-1}(0)$. For a sufficiently small $\epsilon > 0$, the inclusion $NE(g) \subset NE_\epsilon(g)$ induces an injection on homology $\iota_1 : H_1(NE(g)) \rightarrow H_1(NE_\epsilon(g))$. Assuming that there exists ϵ -Nash convergent ϵ -Nash stationary dynamics φ , the arguments in the proofs of Theorem 1 and Theorem 2 can be repeated to show that $NE_\epsilon(g)$ is a maximal attractor for φ , and that the Conley index for $NE_\epsilon(g)$ computed via X and a small neighborhood of $NE_\epsilon(g)$ do not agree, which is a contradiction. See Fig. 4 for a projection of $NE_\epsilon(g)$.

We now show that the set of such games which admit no $\frac{\epsilon}{2}$ -Nash convergent $\frac{\epsilon}{2}$ -game dynamics has positive measure. Consider perturbations of the game g where each utility value of the normal form of g is perturbed independently, so that there is an 18-dimensional vector of perturbations whose 2-norm is no more than $\frac{\epsilon}{4}$. It is clear that this set of games has positive measure, and for any such perturbed game $g' = (R', C')$, we have $\|R - R'\|_F \leq \frac{\epsilon}{4}$ and $\|C - C'\|_F \leq \frac{\epsilon}{4}$. We can apply Lemma 1 twice to get the containment $NE(g) \subset NE_{\frac{\epsilon}{2}}(g') \subset NE_\epsilon(g)$. Consider the associated induced maps on homology

$$H_*(NE(g)) \xrightarrow{\iota_1} H_*(NE_{\frac{\epsilon}{2}}(g')) \xrightarrow{\iota_2} H_*(NE_\epsilon(g)).$$

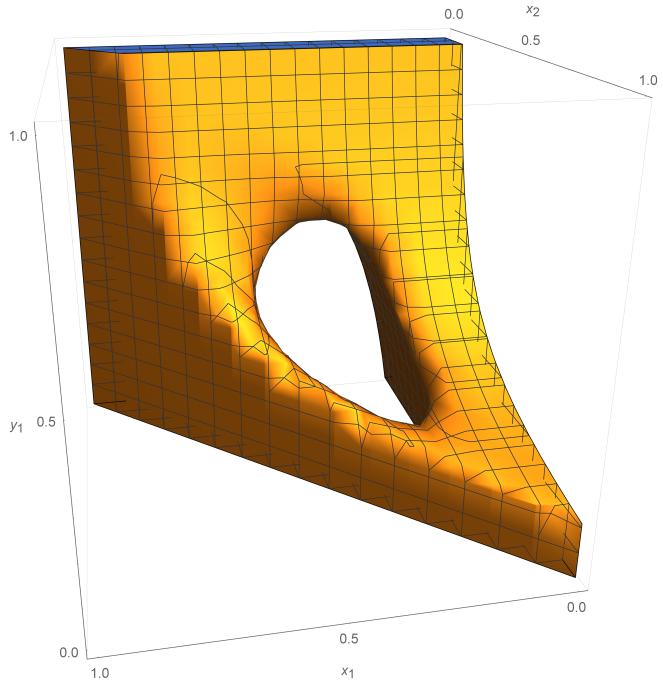


Fig. 4. Three-dimensional projection of $NE_\epsilon(g)$ when g is the Kohlberg–Mertens game Eq. 1, and for utility-normalized $\epsilon = 0.09$. The object depicted is homotopy equivalent to S^1 (a circle).

The map $H_\bullet(\iota_2) \circ H_\bullet(\iota_1) = H_\bullet(\iota_2 \circ \iota_1)$ is an injection, which implies that $H(\iota_1)$ is injective and therefore $H_\bullet(NE_{\frac{\epsilon}{2}}(g')) \neq H_\bullet(X)$. Therefore, a similar argument as above shows that if $NE_{\frac{\epsilon}{2}}(g')$ is a maximal attractor, the Conley indices computed via a small neighborhood and by X do not agree, which completes the proof. \square

Conclusion

We have argued that the notion of the Nash equilibrium is fundamentally incomplete for describing the global dynamics of games. More precisely, we have shown that there are games in which no dynamics can be Nash convergent, and thus, the concept of the Nash equilibrium cannot completely account for the long-term dynamical behavior in games. Moreover, this is true even when one relaxes the focus from Nash equilibria to approximate Nash equilibria.

Ultimately, in view of the present results about the limitations of the Nash equilibrium, the meaning of the game, as well as the meaning of general economic systems, should be sought elsewhere, perhaps with a closer focus on agent dynamics as echoed in ref. 38. Indeed, detailed studies of several learning dynamics in ref. 10 suggest that the structure of best-reply paths and cycles holds important clues for making progress in this direction. Specifically, the existence of best-reply cycles predicts nonconvergence of six different learning dynamics supported by experiments with human subjects. See also refs. 39 and 40 for proposed general solution concepts in game theory, alternatives to the Nash equilibrium, that are inspired by game dynamics and the topological theory of dynamical systems. Moving beyond game theory, a natural and exciting direction for future work is the pursuit of such dynamics-inspired solution concepts more generally in the social and even the natural sciences.

Regarding a more robust understanding of game dynamics, Conley theory provides a natural choice. For instance, as player utilities are not derived from first principles, but are presumably rough approximations of reality, it is natural that the appropriate mathematical objects for the analysis of game dynamics should be robust to perturbation—in the words of Conley (15), “...if such rough equations are to be of use it is necessary to study them in rough terms.” This suggests the use of further ideas from dynamical systems theory, for instance, the concept of Morse decomposition, which allows for a coarse, multiscale description of global dynamics. Morse decompositions are partially ordered sets of isolated invariant sets which possess a duality theory with lattices of attractors and in addition have an associated homological theory using the Conley index (15). We believe that these ideas could provide a robust theory of the global dynamics of games.

Notes: Conley Index Theory

The classical reference for Conley index theory is Conley’s original monograph (15). A standard reference for homotopy theory and singular homology theory is ref. 41.

Continuous Time. Let $\varphi : \mathbb{T}^+ \times X \rightarrow X$ be a semiflow, A an attractor for φ , and U an attracting block for A . The Conley index of the attractor A is defined as follows:

$$CH_\bullet(A) := H_\bullet(U),$$

where H_\bullet denotes singular homology with integer coefficients. Most importantly for this paper, the Conley index is an invariant of A and is independent of the particular choice of U (15, 42). An elementary proof of this is as follows.

Proposition 1. *If U and U' are attracting blocks for attractor A , then U and U' are homotopy equivalent. In particular, $H_\bullet(U) \cong H_\bullet(U')$, and thus, the Conley index is well defined.*

Proof: It is elementary that if U and U' are attracting blocks, $U \cap U'$ is an attracting block; see also ref. 36. Thus, we may assume without loss of generality that $U \subset U'$. Now, since U and U' are attracting blocks for the same attractor A , there exists some $T > 0$ such that $\varphi(T, U') \subset U$. It is straightforward to check that the pair of maps $\varphi(T, \cdot) : U' \rightarrow U$ and inclusion $\iota : U \rightarrow U'$ form a homotopy equivalence between U and U' , as for any $t^* > 0$, the map $\varphi(t^*, \cdot)$ is homotopic to the identity map via the homotopy $(x, s) \mapsto \varphi(s \cdot t^*, x)$. \square

Discrete Time. In the discrete time case, the Conley index takes a slightly more complex form, as the dynamical system itself no longer provides a homotopy equivalence. Let $\varphi : \mathbb{T}^+ \times X \rightarrow X$ be a discrete time dynamical system, A be an attractor for φ , and U be an attracting block for A . Define $f : X \rightarrow X$ via $f(x) := \varphi(1, x)$, then $f(U) \subset U$, and, with slight abuse of notation, there is an induced map $f_U : H_\bullet(U) \rightarrow H_\bullet(U)$, i.e., the well-defined homomorphism given by $f_U([x]) = [f(x)]$. The construction, introduced in ref. 43, proceeds in two stages and depends on the notions of generalized kernel and generalized image. Define the generalized kernel of f_U as:

$$\text{gker}(f_U) := \bigcup_{n>0} \ker f_U^n, \quad \text{and} \quad M_\bullet(U) := H_\bullet(U)/\text{gker}(f_U).$$

The map f_U induces an injective map $f'_U : M_\bullet(U) \rightarrow M_\bullet(U)$. Now consider the generalized image of f'_U :

$$L_\bullet(U) := \bigcap_{n>0} f'_U(M_\bullet(U)).$$

The map f'_U induces a surjection on the generalized image: $\chi_U : L_\bullet(U) \rightarrow L_\bullet(U)$. Moreover, as the restriction of f'_U , χ_U is an injection. Thus, χ_U is an automorphism. The Conley index of U is defined as the pair:

$$\text{Con}(A) := (L_\bullet(U), \chi_U),$$

and is independent of choice of attracting block U .

Proposition 2. *Let φ be a discrete dynamical system, A be an attractor for φ and U and U' attracting blocks for A . The Conley index is independent of choice of U . In other words, there exists an isomorphism $\psi : L_\bullet(U) \rightarrow L_\bullet(U')$ so that the following diagram commutes:*

$$\begin{array}{ccc} L_\bullet(U) & \xrightarrow{\chi_U} & L_\bullet(U) \\ \downarrow \psi & & \downarrow \psi \\ L_\bullet(U') & \xrightarrow{\chi_{U'}} & L_\bullet(U'). \end{array}$$

Proof: We may again assume $U \subset U'$ without loss of generality. Moreover, since U and U' are attracting blocks for A , there is some $n > 0$ we have $f^n(U') \subset U$. It is straightforward that the inclusion $U \subset U'$ induces a map $\iota : L_\bullet(U) \rightarrow L_\bullet(U')$ and

the map f^n induces a map $s: L_\bullet(U') \rightarrow L_\bullet(U)$ so the following diagram commutes:

$$\begin{array}{ccccccc} L_\bullet(U) & \xrightarrow{\chi_U} & L_\bullet(U) & \xrightarrow{\chi_U} & \dots & \xrightarrow{\chi_U} & L_\bullet(U) \\ \downarrow \iota & & \downarrow \iota & s & & & \downarrow \iota \\ L_\bullet(U') & \xrightarrow{\chi_{U'}} & L_\bullet(U') & \xrightarrow{\chi_{U'}} & \dots & \xrightarrow{\chi_{U'}} & L_\bullet(U') \end{array} .$$

Now, χ_U and χ'_U are automorphisms, and as $\chi_U^n = f^n \circ \iota$, we have that ι is a injection, and further as $\chi_{U'}^n = \iota \circ f^n$, ι is a

surjection. Thus, ι is an isomorphism, and the Conley index is independent of choice of attracting block. \square

Data, Materials, and Software Availability. There are no data underlying this work.

ACKNOWLEDGMENTS. K.S. would like to thank Rob Vandervorst for numerous enlightening discussions regarding Conley theory. The work of J.M. and C.P. was supported by NSF grants CCF-1763970 and CCF-2212233.

1. J. F. Nash, Equilibrium points in n -person games. *Proc. Natl. Acad. Sci. U.S.A.* **36**, 48–49 (1950).
2. C. Daskalakis, P. W. Goldberg, C. H. Papadimitriou, "The complexity of computing a Nash equilibrium" in *STOC '06* (ACM, 2006), pp. 71–78.
3. X. Chen, X. Deng, Sh. Teng, "Computing Nash equilibria: Approximation and smoothed complexity" in *FOCS'06* (IEEE Computer Society, 2006), pp. 603–612.
4. K. Eftessami, M. Yannakakis, "On the complexity of Nash equilibria and other fixed points" in *FOCS '07* (IEEE Computer Society, USA, 2007), pp. 113–123.
5. P. W. Goldberg, C. H. Papadimitriou, R. Savani, The complexity of the homotopy method, equilibrium selection, and Lemke–Howson solutions. *ACM Trans. Econ. Comput.* **1**, 1–25 (2013).
6. Y. Sato, E. Akiyama, J. D. Farmer, Chaos in learning a simple two-person game. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4748–4751 (2002).
7. M. Benaim, J. Hofbauer, E. Hopkins, Learning in games with unstable equilibria. *J. Econ. Theory* **144**, 1694–1709 (2009).
8. T. Galla, J. D. Farmer, Complex dynamics in learning complicated games. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1232–1236 (2013).
9. J. B. Sanders, J. D. Farmer, T. Galla, The prevalence of chaotic dynamics in games with many players. *Sci. Rep.* **8**, 4902 (2018).
10. M. Pangallo, T. Heinrich, J. Doyne Farmer, Best reply structure and equilibrium convergence in generic games. *Sci. Adv.* **5**, eaat1328 (2019).
11. P. Mertikopoulos, C. Papadimitriou, G. Piliouras, "Cycles in adversarial regularized learning" in *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms* (SIAM, 2018), pp. 2703–2717.
12. W. B. Arthur, Foundations of complexity economics. *Nat. Rev. Phys.* **3**, 136–145 (2021).
13. J. Bielawski *et al.*, "Follow-the-regularized-leader routes to chaos in routing games" in *International Conference on Machine Learning* (PMLR, 2021), pp. 925–935.
14. S. Hart, A. Mas-Colell, Uncoupled dynamics do not lead to Nash equilibrium. *Am. Econ. Rev.* **93**, 1830–1836 (2003).
15. C. Conley, *Isolated Invariant Sets and the Morse Index*, *Regional Conference Series in Mathematics* (American Mathematical Society, Providence, RI, 1978), vol. 38.
16. S. Hart, A. Mas-Colell, Stochastic uncoupled dynamics and Nash equilibrium. *Games Econ. Behav.* **57**, 286–303 (2006).
17. C. Papadimitriou, N. Vishnoi, "On the computational complexity of limit cycles in dynamical systems" in *ITCS 16* (2016), p. 403.
18. C. Daskalakis, A. Ilyas, V. Syrgkanis, H. Zeng, "Training GANs with optimism" in *The International Conference on Learning Representations (ICLR)* (2018).
19. C. Daskalakis, I. Panageas, "Last-iterate convergence: Zero-sum games and constrained min-max optimization" in *Innovations in Theoretical Computer Science (ITCS)* (2019).
20. P. Mertikopoulos *et al.*, "Optimistic mirror descent in saddle-point problems: Going the extra (gradient) mile" in *The International Conference on Learning Representations (ICLR)* (2019).
21. E. Kohlberg, J. F. Mertens, On the strategic stability of equilibria. *Econometrica* **54**, 1003 (1986).
22. S. Govindan, R. Wilson, Essential equilibria. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15706 (2005).
23. S. Govindan, R. Wilson, Uniqueness of the index for Nash equilibria of two-player games. *Econ. Theory* **10**, 541–549 (1997).
24. S. Govindan, R. Wilson, Maximal stable sets of two-player games. *Int. J. Game Theory* **30**, 557–566 (2002).
25. M. Benaim, J. Hofbauer, S. Sorin, Perturbations of set-valued dynamical systems, with applications to game theory. *Dyn. Games Appl.* **2**, 195–205 (2012).
26. S. DeMichelis, F. Germano, On the indices of zeros of Nash fields. *J. Econ. Theory* **94**, 192–217 (2000).
27. S. Demichelis, K. Ritzberger, From evolutionary to strategic stability. *J. Econ. Theory* **113**, 51–75 (2003).
28. V. Losert, E. Akin, Dynamics of games and genes: Discrete versus continuous time. *J. Math. Biol.* **17**, 241–251 (1983).
29. N. Nisan, T. Roughgarden, E. Tardos, V. V. Vazirani, Eds., *Algorithmic Game Theory* (Cambridge University Press, Cambridge, 2007).
30. D. Monderer, L. S. Shapley, Potential games. *Games Econ. Behav.* **14**, 124–143 (1996).
31. J. Hofbauer, W. H. Sandholm, On the global convergence of stochastic fictitious play. *Econometrica* **70**, 2265–2294 (2002).
32. J. S. Shamma, G. Arslan, Dynamic fictitious play, dynamic gradient play, and distributed convergence to Nash equilibria. *IEEE Trans. Autom. Control* **50**, 312–327 (2005).
33. J. Hofbauer, W. H. Sandholm, Stable games and their dynamics. *J. Econ. Theory* **144**, 1665–1693 (2009).
34. W. H. Sandholm, *Population Games and Evolutionary Dynamics, Economic Learning and Social Evolution* (MIT Press, Cambridge, MA, 2010).
35. C. Robinson, *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos* (CRC Press, 1998).
36. W. D. Kalies, K. Mischaikow, R. C. Vandervorst, Lattice structures for attractors III. *J. Dyn. Differ. Eq.* **34**, pages 1729–1768 (2021).
37. L. S. Shapley, *A Note on the Lemke–Howson Algorithm* (Springer Berlin Heidelberg, Berlin/Heidelberg, 1974), pp. 175–189.
38. J. D. Farmer, D. Foley, The economy needs agent-based modelling. *Nature* **460**, 685–686 (2009).
39. S. Omidshafiei *et al.*, α -rank: Multi-agent evaluation by evolution. *Sci. Rep.* **9**, 1–29 (2019).
40. C. Papadimitriou, G. Piliouras, Game dynamics as the meaning of a game. *SIAGeom Exch.* **16**, 53–63 (2019).
41. A. Hatcher, *Algebraic Topology* (Cambridge University Press, Cambridge, NY, 2002).
42. D. Salamon, Connected simple systems and the Conley index of isolated invariant sets. *Trans. Am. Math. Soc.* **291**, 1–41 (1985).
43. M. Mrozek, Leray functor and cohomological Conley index for discrete dynamical systems. *Trans. Am. Math. Soc.* **318**, 149–178 (1990).