

Disc Game Dynamics: A Latent Space Perspective on Selection and Learning in Games

Pablo Lechon-Alonso ¹, Andrew Dennehy ¹, Ruizheng Bai ², Nicolas Sanchez ³,
 Derek K. Wise ⁴ , David Sewell ⁴, David Rosenbluth ⁴, and Alexander Strang ³⁺

¹ University of Chicago ² Texas A&M ³ University of California, Berkeley ⁴ Lockheed Martin.
 + Correspondence: alexstrang@berkeley.edu

November 25, 2025

Contents

1 Abstract	3
2 Introduction	3
2.1 Paradigms in Evolutionary Game Theory	3
2.2 Game Formalism	6
2.3 Objectives	6
3 Constructing a Latent Space	8
3.1 Desiderata	8
3.2 Bilinear Games and a Canonical Coordinate System	9
3.3 Disc Games: Interpretation and Geometry	11
3.4 Disc Game Embedding	12
3.4.1 Existence and Approximation	14
3.4.2 Consistency and Uniqueness	17
3.4.3 Regularity	18
3.5 Disc Game Embeddings as a Candidate Latent Space for Learning	18
4 Learning in the Latent Space	20
4.1 The Replicator Dynamic	20
4.2 Replicator Parameter Dynamics Via Disc Game-Embedding	22
4.2.1 Objectives and Special Cases	22
4.2.2 Closure	24
4.2.3 Parameterization	26
4.2.4 Hamiltonian Parameter Dynamics	28
4.2.5 Dynamical Characterization	31
4.2.6 Solutions for Simplifying Geometries	40
4.3 Generalization	44
4.3.1 Additional Frequency Dependence	44
4.3.2 Metapopulations and Inhomogeneous Mixing	47
4.3.3 Discussion	49
4.4 Efficient Simulation Leveraging the Latent Space	49
5 Conclusions	50
6 Appendix	61
6.1 An Example Game	61
6.2 Proofs of Embedding Results	62
6.2.1 Proof of Lemma 1	62
6.2.2 Proof of Lemma 2	62
6.2.3 Proof of Theorem 3	63

6.2.4	Proof of Lemma 6	64
6.3	Alternate Derivation of the Disc Game Embedding	64
6.4	Additional Learning Dynamics	66
6.4.1	Self Play and Adaptive Dynamics	66
6.4.2	Fictitious Self-Play	66
6.4.3	Simultaneous Gradient Ascent	66
6.5	Explicit Solutions	67
6.5.1	Advection Dynamics	67
6.5.2	Replicator Dynamics: The Transitive Case	70
6.6	Generic Parameter Dynamics Proofs	71
6.6.1	Proof of Result 1	71
6.6.2	Properties of the Hamiltonian Dynamic	72
6.6.3	Proof of Corollary 5.3	74
6.6.4	Proof of Theorem 6	74
6.6.5	Proof of Lemma 20	75
6.6.6	Proof of Corollary 6.1	76
6.6.7	Proof of Corollaries 7.1 - 7.2	76
6.7	Solutions for Special Geometries	77
6.7.1	Decoupled Parameter Dynamics	77
6.7.2	Rotationally Symmetric Base Measures	78
6.7.3	Separable Base Measures	78
6.7.4	Heteroclinic Dynamics Near Boundary	81

1 Abstract

Evolutionary game theory studies populations that change in response to an underlying game. Often, the functional form relating outcome to player attributes or strategy is complex, preventing mathematical progress. In this work, we axiomatically derive a latent space representation for pairwise, symmetric, zero-sum games by seeking a coordinate space in which the optimal training direction for an agent responding to an opponent depends only on their opponent’s coordinates. The associated embedding represents the original game as a linear combination of copies of a simple game, the disc game, in a new coordinate space. In this article, we show that disc-game embedding is useful for studying learning dynamics. We demonstrate that a series of classical evolutionary processes simplify to constrained oscillator equations in the latent space. In particular, the continuous replicator equation reduces to a Hamiltonian system of coupled oscillators that exhibit Poincaré recurrence. This reduction allows exact, finite-dimensional closure when the underlying game is finite-rank, and optimal approximation otherwise. It also establishes an exact equivalence between the continuous replicator equation and adaptive dynamics in the transformed coordinates. By identifying a minimal rank representation, the disc game embedding offers numerical methods that could decouple the cost of simulation from the number of attributes used to define agents. These results generalize to metapopulation models that mix inhomogeneously, and to any time-differentiable dynamic where the rate of growth of a type, relative to its expected payout, is a nonnegative function of its frequency. We recommend disc-game embedding as an organizing paradigm for learning and selection in response to symmetric two-player zero-sum games.

2 Introduction

2.1 Paradigms in Evolutionary Game Theory

Evolutionary game theory studies the dynamics of populations adapting in response to an underlying game [201]. Evolutionary game theory has a rich history in biology. Example applications include the evolution of sex-ratios [91, 216], ritual fighting [106, 201], mate choice [109], sibling rivalry [149], social habitat selection and dispersal [64], cooperative behavior [176, 220], honest and dishonest signalling, parasite-host interactions [3], and infection virulence [8]. For a survey see [162]. Evolutionary game theory is also important in learning theory, where it presents the generic framework for studying multi-agent optimization, mini-max optimization, and adversarial learning [20, 22, 82, 108, 194].

Before proceeding, we emphasize two contrasts that position our narrative: evolutionary processes need not behave as noisy optimizers, and, co-evolutionary dynamics require dynamic, not static, solution concepts.

First, evolution is not, generically, optimization. In the folk narrative, random perturbation and selection combine to drive populations upwards on a fitness landscape [76, 231]. If taken literally, this model presupposes the existence of a fitness function, i.e. the stochastic process specifying evolution over possible types can be shown to, in some sense, optimize a function that assigns “fitness” to types. No such function need exist [115, 198]. Moreover, even if such a function exists, it may not provide a useful dynamical summary [39, 60], as when it is highly rough [61, 118, 179, 184, 211], quickly varying due to environmental perturbations [217], or when the dynamics are dominated by noise, as in neutral theory [44, 155, 166, 193, 204, 227].

If taken figuratively, the optimization perspective should, at least, provide generalizable intuition. This point is also circumspect. Evolution is an emergent feature of ecological dynamics, which are almost universally frequency-dependent. When equated to per-capita growth rate, the fitness of a type often depends on the current frequency of other types. Then, even if the population adapts to increase the frequency of fit types, it does not optimize any fixed objective, since the objective changes in tandem with the population. To quote Nowak and Sigmund [162], “The landscape metaphor... neglects one-half the evolutionary mechanism.” Namely, co-evolution.

Frequency-dependent selection can produce evolutionary dynamics that are incompatible with any optimization process [20]. Optimization implies constant improvement, and directed motion towards, or convergence to, local optima [162]. More subtly, instantaneous step-wise improvement ensures cumulative progress. In contrast, when frequency-dependent, evolutionary dynamics need not admit any notion of improvement, need not admit any notion of direction, and need not converge, since they may circulate periodically, recur, or mix chaotically [7, 32, 72, 102, 125, 144, 190].

So, in the generic co-evolutionary setting, optimization, at best, provides an instantaneous intuition via a constantly shifting, population-dependent, fitness “seascape” [143, 156]. Even proponents of the seascape model admit that it is difficult to analyze and is hard to communicate as a minimal mental model [39].

Collected, these observations beg a critical question; *what is the utility of an optimization metaphor whose objective is ill-defined or constantly changing in response to the system’s state?*

In response, we will seek alternate paradigms. What paradigm should complement optimization?

To scope alternatives, let’s step back to a more fundamental problem frame. Population dynamics are often parameterized by a fixed set of rules governing interactions between individuals and types. Formally, these rules fix, or are fixed by, an underlying game [201]. Accordingly, game theory provides a candidate frame [162].

Formally, a game is a function that accepts the strategies selected by a group of players, and returns a utility to each based on the collection of chosen strategies. Thus, games model multi-agent interactions. For ecological, or evolutionary, realizations, exchange strategy for type, and utility for growth rate. Importantly, only narrow classes of games, e.g. potential games [150], admit an exact analogy to optimization. Instead, generic games exhibit a mixture of potential and harmonic components [40, 125]. If the harmonic component is non-zero, then the game does not admit an exact reduction to optimization. Generic evolutionary dynamics in response to harmonic games are poorly understood [125].

While game theory is a classically static theory, evolutionary game theory is *dynamic*. Classical game theory studies equilibria as solution concepts for games. While an equilibria is commonly understood as a fixed point of some dynamic [105], the dynamic in traditional game theory is usually implicit, if not unspecified [137]. Equilibria are typically justified by rationality criteria, most famously, that a set of strategies cannot be stable if any agent could improve their utility by adjusting their strategy unilaterally [159]. That said, while stationarity conditions are often considered necessary for rationality (c.f. [137]), they are not sufficient to show that equilibria can be, will be, or frequently are discovered or approached by evolving populations. Nor can they identify which equilibrium a population will select.

Evolutionary game theory aims to solve the equilibrium selection problem by posing specific evolutionary dynamics [137]. Equilibria are often fixed points for these dynamics. Indeed, all asymptotically stable rest points of a monotone dynamic are Nash equilibria [137]. However, not all equilibria are asymptotically stable or attracting. Additional criteria are almost always needed [78]. For example, inefficient equilibria are often unstable and invasible under simple dynamics [137]. Since generic equilibria are not attracting for all game categories or all reasonable dynamics [160], evolutionary game theory has produced a refined taxonomy of equilibrium classes that satisfy increasingly stringent conditions. In some cases, as for evolutionary-stable strategies, simple non-invasibility conditions suffice, and ensure global attraction for a wide class of games and dynamics [104, 201]. In other settings, stronger conditions are required (c.f. [9, 62, 70, 126, 164]). For extensions to stochastic processes see [38, 74, 78, 107, 111].

The diversity of solution concepts reflects a key impossibility result; there is no uncoupled learning dynamic that converges to a Nash equilibrium in all games from all initial conditions [94, 95]. For example games whose Nash equilibria are not approached by any uncoupled dynamic, or by any adjustment dynamic, see [94, 103]. In retrospect, it is not surprising that reasonable learning dynamics, which typically react myopically to limited game information with boundedly rational updates, do not have the power to discover generic equilibria. Nash equilibria are hard to compute [47, 59, 142], to approximate [186, 187, 199], and to communicate [19, 92]. Moreover, even when players can communicate, Nash equilibria need not be self-enforcing [16, 50]. Experimental evidence for Nash play is similarly qualified [48, 105], even in foundational studies on canonical examples [73, 181]. For laboratory studies, see [36, 127, 139, 151, 167, 163, 171, 182]. For observational studies in real-world games, see [18, 13, 65, 79, 81, 121, 170]. For example controversies, see [167] versus [36], or [170, 171] versus [121, 127].

Convergence time is also problematic. Even if an equilibrium is attracting, convergence may be too slow to plausibly restrict observed populations near it. This concern mirrors a broader doubt that dynamical systems theories in ecology have focused too narrowly on linearized dynamics near steady-states (c.f. [138]), and have not fully addressed the possibility that real systems are better characterized by slowly decaying transients [98, 99, 100].

These concerns raise another critical question; *are equilibria sufficient solution concepts for evolutionary games?*

At strongest, all generic no-regret learning dynamics converge to equilibria in the sense that the time-averaged population converges to a subset of the game's coarse correlated equilibria [85, 93, 90]. In general two-player games, the coarse correlated equilibria may be exclusively supported on strictly dominated strategies [222], so fail basic rationality criteria [72]. However, in constant-sum two-player games, all coarse correlated equilibria are Nash equilibria [136]. So, generic no-regret learners may discover Nash equilibria. Importantly, this convergence result only holds for the time-averaged population. So, hindsight rationality only guarantees that long trajectories are rational when taken as a whole. The population at any particular time is only justified by the adaption process that produced it, not by any static notion independent of the trajectory.¹

At weakest, while equilibria necessarily exist, and can be computed using linear programming, in two-player zero-sum games [225], they are not attracting under the most basic evolutionary game dynamics unless they correspond to pure strategies that dominate all other strategies [7, 32, 102]. Instead, even in the two-player zero-sum setting, basic evolutionary dynamics, including a broad class of no-regret dynamics, are recurrent [20, 32, 72, 125, 144]. Recurrent dynamics orbit persistently, thus fail to converge. For instance, the replicator dynamic, which is the most widely used model in evolutionary game theory, either retains a constant KL divergence away from a fully-mixed equilibria, or diverges from that equilibria when simulated in discrete time [7, 32]. Recurrent dynamics return arbitrarily close to any initial condition arbitrarily often, so do not satisfy any notion of directed learning. While these results only apply when a fully

¹Whether convergence of the time average is sufficient or insufficient depends on the application. In artificial settings, the simulator can time-average post hoc. In real settings, time averaging can be performed if a long trajectory is recorded, but longitudinal data is expensive to collect, and does not characterize the population at any fixed time. If the population can be subdivided into parts that evolve independently, then a time average may be realized by an average over separate populations. This, however, either requires an ergodic evolutionary dynamic, or, in the absence of ergodicity, specially chosen initial conditions. The latter explanation begs the question: what process would initialize the populations appropriately?

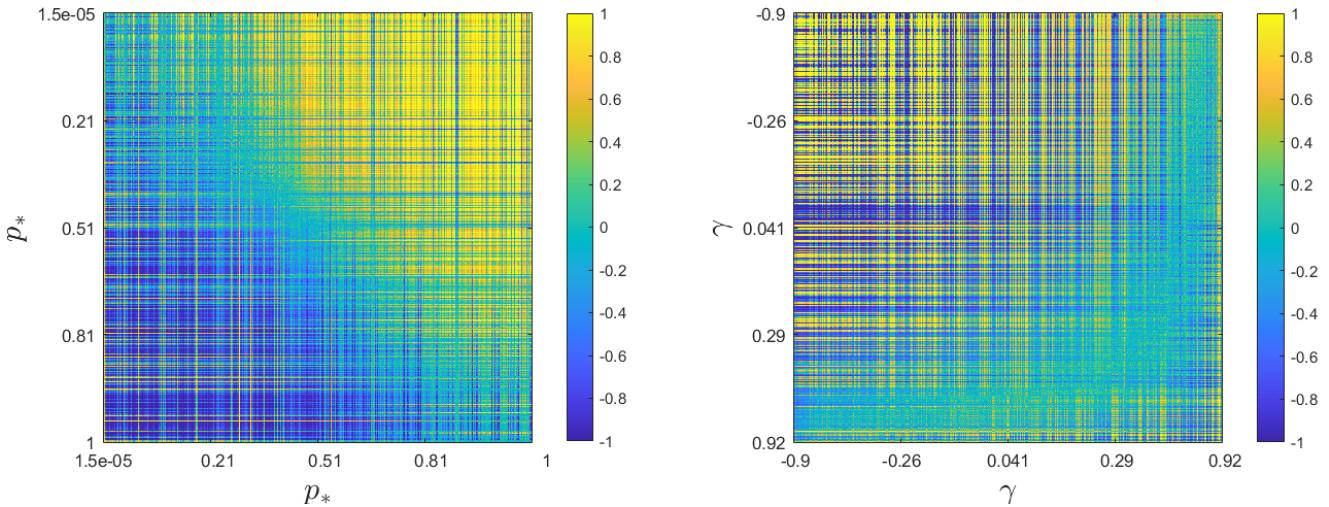


Figure 1: Payout matrices for an population implementation of the iterated prisoner’s dilemma (See Appendix Section 6.1). The color of the i, j entry represents $f(x(i), x(j))$ for a population of 800 randomly drawn agents with attributes $\{x(i)\}_{i=1}^{800}$. The matrices represent the same data, but with the population ordered according to different attributes. Note that, while performance could be roughly approximated as a simple function of the attributes p_* and γ , in both cases an exact description is difficult. For example, the left panel shows that agents with large p_* (rows near the bottom), tend to lose to agents with a small p_* (columns near the left). However, this game is not transitively ordered by p_* alone, as there exist upsets against this prediction. The upper left half is mostly, but not entirely yellow, and the bottom half is mostly, but not entirely, blue. The upsets correspond to particular choices of the pair (p_*, γ) that reverse the general trend in p_* . Even in this two-parameter case, it is not immediately apparent how to extract a simple strategic description from this data, despite the apparent structure.

mixed equilibrium exists, if no such equilibrium exists, then at least one strategy is dominated, and will be eliminated by the dynamic [5]. Thus, in two-player zero-sum games, the canonical no-regret learning dynamic only succeeds in iteratively eliminating dominated strategies, before orbits recurrently among the surviving strategies [7].

Both critical questions interrogate widely used, if often criticized (c.f. [84]), paradigms. Both paradigms remain attractive mental models because they provide clear solution concepts, simplify complex dynamics, and promise guiding intuition by analogy to familiar processes. Both have directed large bodies of successful work.² A compelling alternative needs to offer a similar suite of advantages [68, 123].

We argue, in the restricted setting of two-player, symmetric, constant-sum games, for an alternative that addresses each of the shortcomings of the standard paradigms while encompassing them as special cases and critical points. Our alternative is either exact or allows arbitrarily accurate approximation in essentially all constant-sum games. It allows precise mathematical conversion between learning dynamics and physical processes. These conversions allow visualization, and supply clear dynamical intuition which, unlike optimization, correctly captures the *generic* behavior of evolution subject to zero-sum games. Importantly, the paradigm produces *global, dynamical*, rather than local, static, intuition. The paradigm admits solutions that converge and that orbit persistently, and provides a clean geometric criteria separating the convergent and recurrent cases.

Advocating for a paradigm requires detailed exposition. In this pursuit, we adopt Polya’s imperative regarding vague analogies: “clarify them” [80]. We aim to show that our paradigm can be derived from simple aims regarding the choice of coordinates used to represent the game, that the resulting representation is general, exact, unique, and interpretable. Finally, we show that it is *useful* through an in-depth case-study. En route we recover and generalize a representation first proposed in [20, 21, 22, 23], then use it to extend classical results constraining the behavior of the replicator dynamic subject to zero-sum games to continuous trait spaces [7, 102, 32]. In addition, we demonstrate an exact equivalence between the replicator dynamic and adaptive dynamics. We show that our key dynamical conclusions generalize to metapopulations, and to any dynamic where the rate of growth in the density of a type, relative to its expected payout, is a nonnegative function of its density. These results strongly recommend the paradigm.

2.2 Game Formalism

This paper pursues a canonical coordinate system for learning and selection dynamics driven by a pairwise, symmetric, zero-sum, functional form game. A functional form game, (Ω, f) , is defined by a space of attributes Ω , where $x \in \Omega$ is collection of attributes used to characterize agents, and $f : \Omega \times \Omega \rightarrow \mathbb{R}$ is a payout function where $f(x, x')$ returns the payout to an agent with attributes x when paired against an opponent with attributes x' . If the game is symmetric, then agents do not take on distinguishable roles in the game, so share the same payout function. If the game is symmetric and zero-sum, then the payout function f must satisfy $f(x, x') + f(x', x) = 0$, so f must be skew-symmetric; $f(x, x') = -f(x', x)$.

The attributes x are any list of agent characteristics that determine payout [205, 206]. Attributes are not limited to physical characteristics, play statistics, decision probabilities, etc. Instead, attributes differ by context. Since we will seek a procedure for standardizing agent description, the generality of this initial description is a strength of the subsequent analysis. Some examples follow:

1. In biology, attributes may include phenotypic, morphological, or behavioral characteristics. Common examples include body size, aggressiveness, and testosterone levels [45, 232]. As specific examples, [26, 86, 208] study the relationship between attributes and competitive success in pumpkinseed sunfish, elephant seals, and Cape dwarf chameleon. For sunfish, relevant attributes include body size and experience [26]. For elephant seals, relevant attributes include body mass, length, age, and time of beach arrival [86]. For chameleons, relevant attributes include body mass, morphology (torso, tail, and jaw length, head width, casque size), and the size of a signaling patch [208]. Additional examples are studied in [110, 115, 188, 202, 221].
2. In sports analytics, attributes may include physical characteristics and play statistics. Example predictive performance models that attempt to relate play statistics to performance are reviewed in [46, 112, 120, 189, 218, 223, 229].
3. In artificial settings, attributes are typically parameters. For example, in traditional game theory, mixed strategies are parameterized by decision probabilities. A collection of decision probabilities, or policy, may be the output of a more complicated model, as is usually the case in artificial intelligence, where neural network or decision tree architectures may be employed to produce a policy, or to predict value, based on game state (c.f. [37, 192, 197, 196, 213, 153]). In this setting, the explicit parameters of the policy and value networks are a valid, if exhaustive, collection of attributes. These attributes often carry little to no meaning when isolated from the full list of parameters, which is typically too long to interpret without tools that can reorganize and compress the agent description. At a higher level the hyperparameters used for tuning training may be used as attributes.

2.3 Objectives

Suppose you can observe a population evolving in response to a functional form game. The game is defined by a payout function f whose specific form depends on the attributes used to describe the agents. You can choose what attributes to measure. Different attributes may be easier or harder to estimate, and may be more or less related to competitive success. Often, it is not clear *a priori* *what* variables are relevant, that is, which variables are sufficient to predict advantage. Often, it is also unclear which *combination* of these variables provides the simplest coordinate space to represent competition.

For example, is body mass, experience, or priority more important among elephant seals [86]? What combination of play statistics are most predictive of performance in sports [218]? What changes in the weights of a policy network correspond to interpretable changes in strategy?

Both problems matter. Identifying a minimal, sufficient set of attributes is important since measurements require effort. Organizing those attributes to simplify the mapping from attributes to outcome is important when extracting a strategic interpretation or predicting selection for specific traits. This organization problem is closely related to feature engineering in artificial intelligence [27, 75, 235, 236] and credit assignment in reinforcement learning [14, 71, 147, 209, 210, 219]. Identifying a latent space that organizes parameters is a longstanding area of research in both fields [4, 177, 209]. Organizing attributes by importance can also guide attribute selection, so can help solve the first problem.

What, then, should you measure and in what combination? We will attempt to answer the second question.

Suppose that you can measure all relevant attributes, but do not know, *a priori*, how they determine competitive outcome. For example, consider an artificial setting where agents compete in a defined game via a simulator, and act according to a policy that involves many parameters. We will distinguish extrinsically motivated coordinates, those chosen based on considerations separate from the payout, f , such as ease of measurement or convenience of parameterization, from intrinsically motivated coordinates, those chosen based on f alone. Intrinsically motivated coordinates should be selected to simplify the functional form of f . For consistency, these coordinates should only depend on f and the frequency of agent types, not the initial attributes used for description.

²Examples include robust linearized analyses in certain ecological systems [1, 134, 161, 183], empirical efforts to measure fitness landscapes with applications to drug sequencing, the evolution of antibiotic resistance, and HIV treatment [83, 87, 118, 168, 172, 179, 184, 211], and observational demonstration of equilibrium play in real games [18, 48, 79, 170].

Identifying an intrinsically meaningful coordinate system is a classical problem in model reduction and exploratory data analysis. For example, principal component analysis, multi-dimensional scaling, diffusion mapping, sparse dictionary learning, and variational auto-encoding all seek a mapping into a latent space that simplifies a data set by encoding some of its structure in an embedded geometry [2, 51, 88, 116, 117, 122, 174, 214]. That geometry is typically equipped with a canonical interpretation that gives meaning to the coordinates, and provides a frame for deriving a specific representation from an arbitrary initial coordinate system. The striking success of transformer models suggests that complex functional relationships can be effectively learned from sample data by first embedding entirely arbitrary token representations into such latent spaces [114, 128, 215].

We seek a meaningful latent space by mapping from an arbitrary, sufficient trait space into a new trait space, selected to simplify the relationship between agent description and agent performance. We pursue a coordinate change that:

1. simplifies f so that the selective response of a population is intuitive, analytically tractable, and/or efficiently computable when working in the latent space, and
2. uses only as many variables as needed, and allows optimal approximation upon truncation.

To select a canonical coordinate system, we introduce a constraint on the desired representation. The constraint is chosen to plausibly accomplish the first goal for a range of selection dynamics. Section 3.1 specifies the constraint. Section 3.2 explores a special case when the constraint is satisfied, then shows that, if satisfied, the chosen constraint fixes the desired representation. Namely, if the constraint is satisfied, then the game may be expressed as a combination of simple games after changing coordinates. These are disc games. Disc games encode strategic relations in a simultaneously expressive and intuitive geometry [23, 21, 206]. In Section 3.3 we list sample relations between strategy and geometry to show that the coordinates in a disc game are interpretable. These relations recommend the disc game as a natural, desired representation. Section 3.4 shows that such a representation exists generically, and is unique in the sense that it produces a consistent representation that is invariant to changes in agent description. We specify the method needed to perform the coordinate change, and show that it allows optimal model reduction, thus satisfies our second objective.

To show that the new coordinates are actually useful for particular dynamics we focus on generalizations of the continuous replicator equation. Other sample dynamics are discussed in the appendix. In Section 4 we show that the coordinate change can be used to reduce the continuous replicator dynamic to a Hamiltonian system of coupled oscillators that is equivalent to adaptive dynamics in the new coordinates. These results show that the intuition suggested by the representation is accurate. In Section 4.4 we show that the coordinate change also allows faster computation by reducing the dimension used for agent description. Collectively, we argue that disc game dynamics offer a compelling complement to optimization as a paradigm for evolution in response to pairwise, symmetric, zero-sum games.

3 Constructing a Latent Space

3.1 Desiderata

Let Ω represent a set of possible agent attribute where $x \in \Omega$ represents a particular collection of attributes. Unless stated otherwise, we will assume that Ω may be represented as a subset of a finite-dimensional vector space. We assume that Ω is sufficient to predict payout so there exists a function $f : \Omega \times \Omega \rightarrow \mathbb{R}$ where $f(x, x')$ returns the competitive advantage (payout) an agent of type x receives when interacting with an opponent x' . We assume that f is deterministic. If the game is stochastic, or there are relevant agent attributes outside Ω that can be sensibly averaged given the attributes $x \in \Omega$, then f is defined as an expected payout. We restrict our attention to symmetric, zero-sum games, so assume that f is skew-symmetric: $f(x, x') = -f(x', x)$ for all pairs $(x, x') \in \Omega \times \Omega$.

Consider an agent with attributes x , paired against an opponent with attributes x' . How should the agent x adapt, to best respond to x' ? To focus our construction, we will start by assuming that Ω is a Euclidean vector space, and f is differentiable. We will relax both constraints later.

At simplest, the first agent might adapt their attributes to maximize $f(x, x')$ over all possible choices of x . This would define a best-response dynamic [78]. If maximization is impossible, then the agent might simply aim to change their attributes so as to increase $f(x, x')$. If the agent is restricted to adapt continuously, or via small steps, then the agent should adapt along a direction which maximizes $f(\cdot, x')$ over all possible choices near x . If f is sufficiently regular, then, in the limit of small steps, the first agent should adapt along the direction of fastest ascent about x, x' . That is, along the gradient of f with respect to the first input, evaluated at x : $\nabla_w f(w, x')|_{w=x}$. Let $v_f(x, x')$ denote the optimal local training direction:

$$v_f(x, x') = \nabla_w f(w, x')|_{w=x} \quad (1)$$

The optimal local training vector plays a fundamental role in many selection dynamics including self-play, fictitious self-play, simultaneous gradient ascent, and adaptive dynamics [78, 162]. See Appendix Section 6.4 for details. These dynamics need not assume rational agents, nor rational designers, as the optimal local training vector fixes the local linear approximation to g about y and y' . Consequently, models that implicitly enforce a form of myopic gradient ascent may be parameterized by v without requiring explicit calculation of gradients [3]. In Section 4.1 we show that v remains relevant for an essential ecological dynamic that does not require local perturbations to agents' types.

The optimal training vector v_f is defined on the product space consisting of pairs of agent attributes. Selection is easier to study if the vector field is constant in one of its arguments. Since v_f is defined as the optimal local training vector for x when paired with x' we will consider x the student and x' their opponent, who acts as a teacher in the sense of a sparring partner. Since we can't guarantee that v_f is constant in one of its arguments, we will seek a procedure for changing coordinates such that the vector field in the new coordinates, v_g , is constant in one of its arguments.

Let $T : \Omega \rightarrow \Psi$ denote a one-to-one transformation mapping from Ω to Ψ . If T is diffeomorphic, then $y = T(x)$ is the new representation of an agent with attributes x after a change of coordinates. Let $g(y, y') = f(T^{-1}(y), T^{-1}(y'))$, where T^{-1} denotes the inverse transformation over Ψ . Then $g(y, y') = f(x, x')$ when $y = T(x)$ and $y' = T(x')$. We will constrain T by enforcing simplifying criteria on v_g , which will, in turn, enforce simplifying criteria on g . Since we seek constraints on v_g we will, from here on out, suppress the subscript, and let v denote $v_g = \nabla_w g(w, y)|_{w=y}$.

Desiderata: Identify a transform T such that:

- D1. $v(y, y')$ depends only on the student's attributes, y , and is independent of the opponent.
- D2. $v(y, y')$ depends only on the opponent's attributes, y' , and is independent of the student.

Surprisingly, desideratum (D1) is only achievable for an extremely limited subset of games, while desideratum (D2) is achievable without placing any substantive restrictions on f . We will show that desideratum (D2) can be accomplished in general (see Section 3.4) and uniquely specifies T up to linear transformation and a choice of measure on Ω .

Desideratum (D1) implies that the optimal training direction for y depends on y alone and is independent of their opponent. Then selection and learning reduce to optimization of an opponent-independent fitness or rating. Not all payout functions can be reduced to a fitness comparison. Those that can are perfectly transitive [42]:

Definition: A skew-symmetric function $f : \Omega \times \Omega \rightarrow \mathbb{R}$ is **perfectly transitive** if there exists a rating function $r : \Omega \rightarrow \mathbb{R}$ such that $f(x, x') = r(x) - r(x')$ for all pairs $(x, x') \in \Omega \times \Omega$.

Perfectly transitive payout functions induce well-ordered hierarchies and do not admit any cyclic (intransitive) competitive relationships. Generic payouts need not be perfectly transitive, as they may admit cycles [205].

If f is perfectly transitive, then $g(y, y') = r(T^{-1}(y)) - r(T^{-1}(y'))$, so is also perfectly transitive. If $g(y, y') = r(y) - r(y')$ for some r , then $f(x, x') = g(T(x), T(x')) = r(T(x)) - r(T(x'))$, so is also perfectly transitive. If desideratum (D1) is satisfied, then the fundamental theorem of calculus (FTC) implies that g must be perfectly transitive, so f must also be perfectly transitive.

Lemma 1: [Opponent-Independent Learning Implies Transitivity] If there exists a transformation T such that g satisfies desideratum (D1) and Ψ is connected³, then f is perfectly transitive.

Proof Outline: Set $r(y) = \int_{w=y_0}^y v(w, z) dy$ for any choice of y_0 and z in Ψ . Then, by the FTC, $g(y, y') = r(y) - r(y')$. See Appendix Section 6.2.1 for details. \square

Therefore, desideratum (D1) cannot be achieved in general. Consider desideratum (D2) instead.

Lemma 2: [Student-Independent Learning Implies Bi-Affine] The vector field $v(y, y')$ does not depend on y if and only if $g(y, y')$ is a bi-affine function of y and y' of the form:

$$g(y, y') = \hat{y}^\top \hat{G} \hat{y}' = [y^\top, 1] \begin{bmatrix} G & a \\ -a^\top & 0 \end{bmatrix} \begin{bmatrix} y' \\ 1 \end{bmatrix} = y^\top G y' + a^\top (y - y') \quad (2)$$

where G is skew-symmetric.

Proof Outline: If g is bi-affine, then v is student-independent since, for each fixed y' , $g(y, y')$ is an affine function of y . If $v(y, y')$ does not depend on y , then, for fixed y' , $g(y, y')$ must be an affine function of y . By skew-symmetry, if g is an affine function of y for fixed y' , it is also an affine function of y' for fixed y . Thus, g is bi-affine and admits the general form (2). Expanding to \mathbb{R}^d , by replacing y with $[y, 1]$, expresses the bi-affine function as a bilinear function in \mathbb{R}^{d+1} , restricted to a \mathbb{R}^d dimensional affine subspace. See Appendix Section 6.2.2 for details. \square

The appended coordinate fixed to 1 in Equation (2) contributes a perfectly transitive component with rating function $r(y) = a^\top y$. We will, from now on, expand the coordinates y to include the appended 1 whenever $a \neq 0$, so that bi-affine functions may be expressed as restrictions of bilinear functions. So, with some abuse of notation, we discard the hats, and use y to mean \hat{y} and G to mean \hat{G} . This reduction is worthwhile since all subsequent statements hold with an appended transitive term if expressed explicitly, and we will subsequently focus on a class of functional form games that can be expressed in a bilinear form, restricted to a subset of a larger Euclidean space.

Not all bilinear payouts are perfectly transitive. For example, the simple bilinear game $g(y, y') = y_1 y'_2 - y_2 y'_1$ allows cyclic advantage relationships if Ψ includes the origin in its convex hull. Thus, while desideratum (D1) requires perfect transitivity, desideratum (D2) does not. Desideratum (D2) is more general. It retains some coupling between student and opponent, so learning and selection in games that satisfy desideratum (D2) remain a multi-agent dynamics.

Bilinearity is attractive for two reasons. First, if bilinear, then the optimal training vector $v(y, y') = v(y')$ completely parameterizes performance for all agents paired against the opponent y' . In particular $v(y') = Gy'$ and $g(y, y') = y \cdot v(y')$. In this sense, the local linearization of payout g defined by the gradient v extends globally when g is bilinear. Third, many selective dynamics update agents or populations based on either expected best responses or expected payouts when sampling an opponent y' from a distribution π . Expectations are linear operations, so:

$$\bar{v}_\pi = \mathbb{E}_{y' \sim \pi}[v(y, y')] = \mathbb{E}_{y' \sim \pi}[v(y')] = \mathbb{E}_{y' \sim \pi}[Gy'] = G\mathbb{E}_{y' \sim \pi}[y'] = v(\bar{y}_\pi). \quad (3)$$

Therefore, when desideratum (D2) holds, the averaged optimal training vector can always be reduced to the optimal training vector evaluated at an average opponent. Then, by the second property listed above, the expected payout to a student y against a pool of opponents $y' \sim \pi$ equals the payout of the student y against the average opponent \bar{y}_π :

$$\mathbb{E}_{y' \sim \pi}[g(y, y')] = \mathbb{E}_{y' \sim \pi}[y \cdot v(y')] = y \cdot \bar{v}_\pi = y^\top G \bar{y}_\pi. \quad (4)$$

These properties can dramatically simplify selection dynamics that depend on weighted averages against the current and past population. They also explain why the pursuit of desideratum (D2) will simplify dynamics that do not depend on local perturbations (c.f. Section 4). Once bilinear, local linearization predicts payouts globally and averages over a population reduce to the local best response at an average agent.

3.2 Bilinear Games and a Canonical Coordinate System

To start, we consider a special case where desideratum (D2) can be accomplished trivially; if f is bilinear in x and x' , then any invertible affine transformation produces bi-affine g that can be expressed as a bilinear function restricted to an affine subspace.

Bilinear payouts occur naturally in normal form games with finite strategy sets. For example, consider a simultaneous, symmetric, zero-sum, normal form game. Then, both players may select a pure strategy from a finite set of available strategies, or, more generally, may pick a strategy at random from a distribution over possible strategies. In the latter case, the agents use a mixed strategy. A mixed strategy assigns a distribution p over the possible strategies. Then, given s possible strategies, the agents may be parameterized by selecting $d = s - 1$ probabilities which act as traits.

Let f_{ij} represent the payout received by the first agent if they use pure strategy i when the second agent uses pure strategy j . Let F denote the matrix storing all the payouts. When symmetric and zero-sum, F is skew-symmetric. Let p represent the mixed strategy for the first agent, and p' the mixed strategy for the second agent. Then:

$$f(p, p') = \mathbb{E}_{I \sim p, J \sim p'}[f_{IJ}] = p^\top F p'. \quad (5)$$

Equation (5) is bilinear. So, all symmetric, zero-sum, normal-form games satisfy desideratum (D2) before transformation. Nevertheless, it may still be possible to select an affine coordinate transformation that simplifies the normal form game by replacing F with G , where G is sparser than F .

In particular, since F is real, skew-symmetric, it has even rank, r , and admits a real Schur-form [233, 237]:

$$F = QWQ^\top, \quad W = \begin{bmatrix} \omega_1 R & 0 & \dots \\ 0 & \omega_2 R & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}, \quad R = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (6)$$

where Q is $d \times r$ with orthonormal columns $\{q_j\}_{j=1}^r$, W is $r \times r$ block-diagonal with 2×2 diagonal blocks of the form $\omega_j R$, and $\{\omega_j\}_{j=1}^{r/2}$ are strictly positive and decreasing. Notice that multiplication by R performs a ninety-degree rotation so the quadratic form $v^\top R w$ evaluates a cross-product: $v^\top R w = v_1 w_2 - v_2 w_1 = v \times w$.

The components of real Schur form can be recovered from the eigenvalue decomposition of F . Set $\omega_k = |\lambda_{2k-1}|$, and $[q_{2k-1}, q_{2k}] \propto [\text{Real}(v_{2k-1}), \text{Imag}(v_{2k-1})]$ where $\{(\lambda_j, v_j)\}_j$ are the eigenvalues and vectors of F listed so that the eigenvalues decrease in magnitude, and so that the eigenvalue with positive imaginary part appears first in every pair.

Let $T(x) = Mx$. Then, since f is bilinear, and T is linear, g is bilinear. Next, let $M = D_\omega^{1/2} Q^\top$ where $D_\omega = \text{diag}(\omega_1, \omega_1, \omega_2, \omega_2, \dots)$. If $r < d$, then append $d - r$ rows to the bottom of M , chosen to span the component of \mathbb{R}^d perpendicular to the range of F . Then:

$$g(y, y') = T^{-1}(y)^\top FT^{-1}(y') = y^\top U y' \text{ where } U = \begin{bmatrix} R & 0 & \dots \\ 0 & R & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}. \quad (7)$$

In this coordinate system, the normal form game decouples blockwise. In fact:

$$g(y, y') = y^\top U y' = \sum_{k=1}^{r/2} [y_{2k-1}, y_{2k}]^\top R \begin{bmatrix} y'_{2k-1} \\ y'_{2k} \end{bmatrix} = \sum_{k=1}^{r/2} y_{2k-1} y'_{2k} - y_{2k} y'_{2k-1} = \sum_{k=1}^{r/2} y^{(k)} \times y'^{(k)} \quad (8)$$

where \times denotes the cross product and $y^{(k)} = [y_{2k-1}, y_{2k}]$ represents the k^{th} consecutive pair of coordinates [23, 21, 206]. The cross-product defines a canonical functional form game.

Definition: A functional form game (f, Ω) with $\Omega \subseteq \mathbb{R}^2$ and $f(x, x') = \text{disc}(x, x') = x \times x'$ is a **disc-game** [21].

Equation (8) states that any symmetric, zero-sum normal form game can be represented with a sum of disc-games, in the coordinate system produced by applying $T(x) = D_\omega^{1/2} Q^\top x$:

$$g(y, y') = \sum_k \text{disc}(y^{(k)}, y'^{(k)}). \quad (9)$$

Notice that $Q^\top x$ is the orthonormal change of coordinates associated with the projection onto the eigenvectors of F . We will use this observation to generalize from the spectral decomposition of a skew-symmetric matrix to the spectral decomposition of a skew-symmetric function f .

Equation (9) is an example of a disc-game decomposition [23]. It represents a game as a linear combination of disc-games, each acting on separate pairs of coordinates. Disc game decomposition has been proposed as a tool for expanding player evaluation [23], designing population level learning algorithms which seek diverse populations [21], and as an exploratory tool that visualizes strategic trade-offs [206]. In the next section, we explain how geometry encodes strategy in a disc game. This geometry recommends the disc game as an interpretable representation of otherwise complicated games.

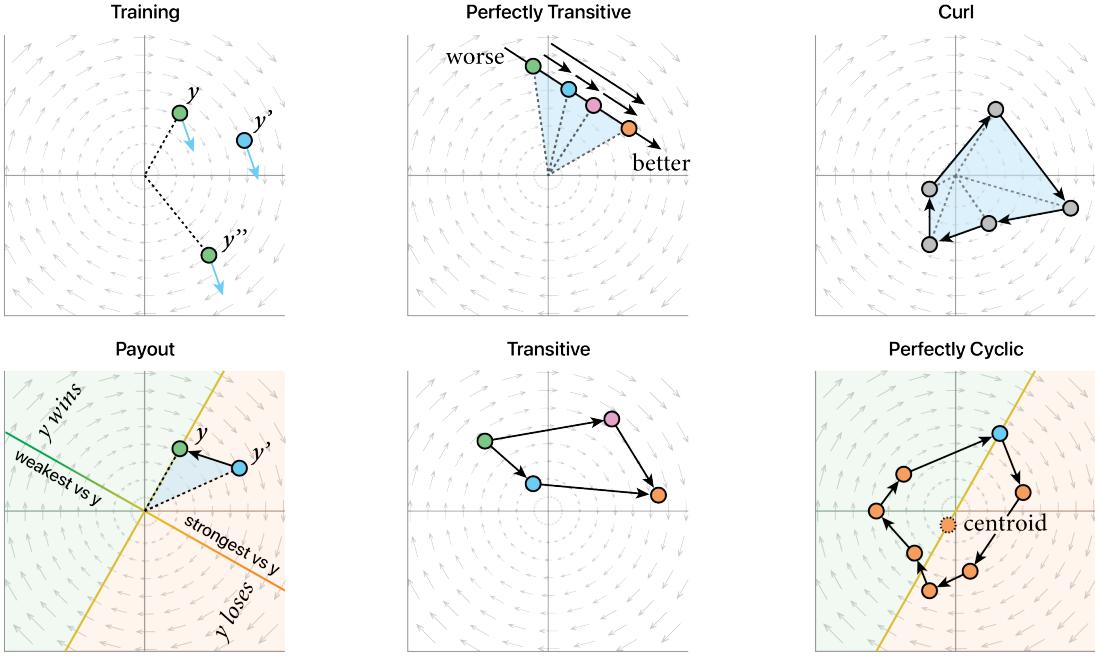


Figure 2: Disc game geometries. In all panels, the two coordinates correspond to y_1 and y_2 , and the circulating grey vector field represents the optimal training response to an agent at each possible location in the disc-game. This is the optimal self-play vector field, $v(y, y)$. It represents the direction in which agents should move when training against themselves.

3.3 Disc Games: Interpretation and Geometry

Disc-game decompositions are useful since each individual disc-game is simple, yet can encode rich strategic information in its geometry [21, 206]. Since the disc-game determines payout via a cross-product, strategy is intimately related to agents' coordinates $y(x)$. Figure 2 illustrates the geometric features associated with training, payout, transitivity, intransitity, and cyclic competition. The geometric features are explained in Box 1.

Box 1. Disc Game Geometry

- Training:** optimal training vectors depend only on the opponent. Opponent, y' : small blue circle. Optimal training vector, $v(y')$: small black arrow pointing parallel to the background vector field. Students, y and y'' : green circles. Students should both move in the direction $v(y')$ to optimize their payout against y' : light blue arrows.
- Payout:** advantage circulates clockwise. Agents counterclockwise from y lose to y , while agents clockwise from y beat y . Advantage is maximized at 90° off-sets. Agents lying along any line passing through the origin are equally matched. The payout disc(y, y') equals twice the signed area of the blue triangle $[0, 0] \rightarrow y' \rightarrow y \rightarrow [0, 0]$. Equivalently, disc(y, y') equals the path integral over the line from y' to y against the optimal training vector field.
- Perfectly Transitive:** Competition among a set of agents is perfectly transitive [205] if and only if the agents are colinear. Ratings equal signed distance along the line. Rating is represented by color: green (weak) to orange (strong).
- Transitive:** Agents may be assigned a partial order, so that $g(y, y') \geq 0$ if $y \succeq y'$. A population is transitive if and only if the convex hull of the agents (outlined with black arrows) does not contain the origin (see Section 4.2.5).
- Curl:** The sum of the payout between pairs of agents, moving around a cycle of agents: arrows pointing between agents [205]. In a disc game, the curl equals twice the signed area of the region enclosed by the loop: blue polygon.
- Perfectly Cyclic:** Every agent receives zero payout, on average, against a random opponent [205]. A population is perfectly cyclic if and only if the centroid, after removing any agent, lies along the line connecting the origin to that agent.
- Origin:** The origin marks a competitor who ties against all opponents.

By encoding strategic information in geometry, disc-game decompositions imbue their coordinates y_1, y_2 with strategic meaning. For example, in Kuhn poker, the axes associated with a disc-game decomposition represent interpretable betting behaviors such as honesty and skepticism [206].

The geometric relations listed in Box 1 suggest alternative axioms. For example, advantage in a disc-game is perfectly transitive along any line. Thus, training advantages are additive along lines; the running sum of the payout between a sequence of agents along a line segment adds to the payout between the agents at the endpoints. This observation suggests an alternate desideratum; pursue coordinates such that payout compounds predictably along training paths. Alternatively, one could pursue coordinates where the sum of performance around any loop (the curl) equals its embedded area. Each approach points back to disc-game embedding (See Appendix Section 6.3).

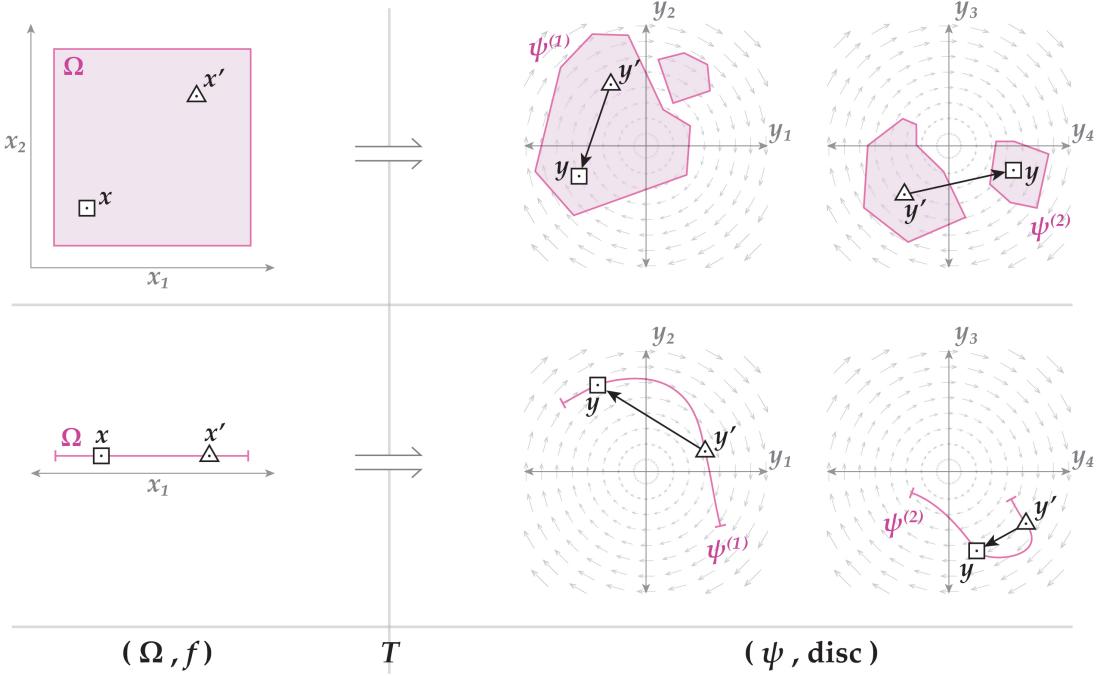


Figure 3: Disc game embedding for a two-dimensional trait space (top row) and a one-dimensional trait space (bottom row). The left-hand column shows the original trait spaces, Ω (magenta region), and a pair of agents in each, with attribute vectors x and x' (marked with a square and a triangle). Competitive advantage is evaluated using f in the original space. In the original attribute space, f may be arbitrarily complicated. The arrow represents the transformation T which maps Ω to Ψ . The right-hand column represents the disc game embedding. The coordinates, y are broken into consecutive pairs, each composing a separate game. The image of the original attribute space under the embedding, Ψ , is shown in magenta. The image of x and x' , $y = T(x)$ and $y' = T(x')$ are shown with matching markers. In the embedding space, the competitive advantage between y and y' is simple, and is determined by a disc game, or, equivalently, cross-product between the embedding coordinates. The value of the cross-product equals the value of the line integral from y' to y against the optimal local training vector field, v (shown in grey in each disc game). Note: although we motivated the disc game embedding assuming smooth T , disc game embeddings may use nondifferentiable and discontinuous T . The regularity of T is constrained by the regularity of f (see Section 3.4.3).

3.4 Disc Game Embedding

The existence of a disc-game decomposition for all symmetric, zero-sum, normal form games, suggests an additional desideratum for bilinear games:

Desideratum: Identify a transform T such that:

3. $g(y, y')$ admits a disc-game decomposition of the form (9).

Notice that, if desideratum (D2) is possible, then g is bilinear, so desideratum (3) is also possible. Thus, any f that admits a transform satisfying desideratum (D2), also admits a coordinate transformation such that:

$$f(x, x') = g(y, y') = \sum_k \text{disc}(y^{(k)}, y'^{(k)}), \text{ where } y = T(x). \quad (10)$$

Equation (10) is a disc-game embedding. The desired transformation is represented schematically in Figure 3. The left-hand column of Figure 3 illustrate simple two-dimensional (top-row), and one-dimensional (middle-row), attribute spaces. These are intentionally drawn as simple regions since the attribute spaces are usually defined by the user. For example, the possible parameter values defining a value or policy network would usually be a set of real-valued weights and biases. While the attribute space is simple, its relation to performance is opaque, and, for any example with more than two attributes cannot be visualized directly since f is a function of four or more variables. The disc-game embedding exchanges a simple attribute space Ω with a, possibly complicated, performance function, for a, possibly complicated, embedded attribute space Ψ , and a simple, indeed canonical, performance function: $\text{disc}(y, y')$. In essence, any complexity in the original performance function f is absorbed into the transform T , and the image of the original attribute space under the embedding, Ψ .

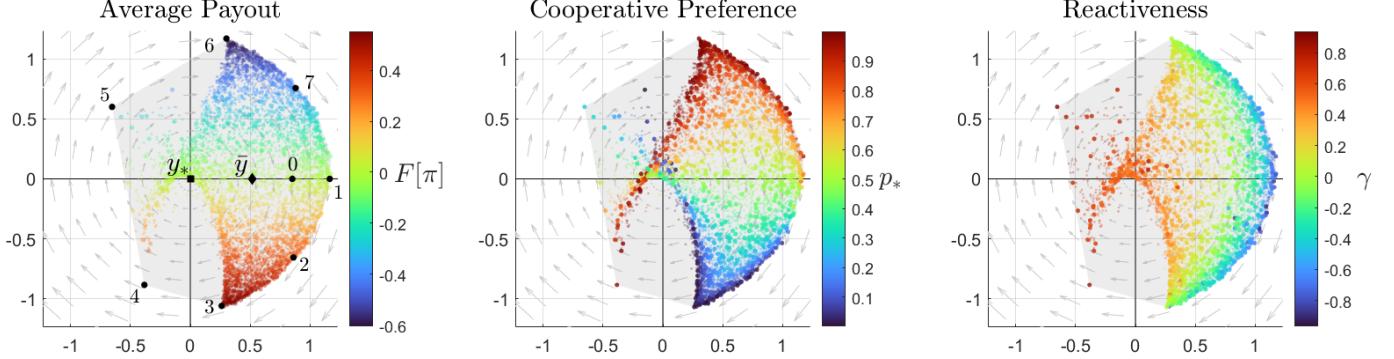


Figure 4: The disc game representation of the IPD payout matrices presented in Figure 1. For the full game specification, see Appendix Section 6.1. All three panels plot the coordinates, $[y_1^{(1)}(x), y_2^{(1)}(x)]$ for each agent of type x in the sampled population. The first 800 agents are bolded. Small scatter points correspond to an interpolation of the embedded 800 agents using 3,000 agents. The interpolation is added to smooth the visual trends, and to highlight the apparently simple underlying functional relation mapping from x to y . Here T is only solved for pointwise. Only the first disc game is shown since it accounts for 90% of the total variance in performance. **Left:** Agents are colored according to their average performance, in the first disc game, against the 800 agent population. Since the disc game is bilinear, the average payout is determined by a cross-product against the centroid in the disc game space, marked with a black diamond and denoted \bar{y} . Agents with embedded attributes y_* , marked with a black square, performs neutrally in the disc game. They correspond to a Nash Equilibrium policy as their disc-game payout against any opponent is zero, and their opponent's payout is also zero. The remaining marked agents correspond to agents with distinct behaviors. For an interpretive key, see Box 1. **Middle:** Agents are colored according to their innate preference for cooperation, p_* . The value of p_* sets the probability with which they cooperate before interacting with an opponent, and is the policy they would return to in absence of information about their opponent. Notice that phase around the origin is a smooth function of p_* , with two fan-shaped lobes. Within each lobe, agents with smaller p_* posses an advantage over agents with a larger p_* . That is, within each lobe, it is better to be distrustful. **Right:** Agents are colored by an attribute, γ , which controls how they react to their opponent's actions. Agents with large positive γ imitate their opponent, agents with γ near zero mostly don't respond to their opponent, and agents with large negative γ play the opposite action their opponent last played. Again, note that position in the disc game is a smooth function of the attribute γ . In particular, the horizontal coordinate y_1 is approximately monotonic in γ . The two fans correspond to $\gamma < 0.4$ (right fan), and $\gamma > 0.4$ (left fan).

Box 2. Example Disc Game Interpretation

Figure 4 illustrates the first pair of disc-game coordinates produced by embedding the payout matrices illustrated in Figure 1. Position in the disc game space is related to an agent's policies by smooth functions (see the middle and left most panel). 8 embedded policies are numbered in the left most panel. These form an advantage cycle in which $1 \succ 7 \succ 6 \succ \dots \succ 2 \succ 1$.

Each agent is an IPD agent. The numbered policies are interpreted below:

- 0: plays a uniform random strategy (coin toss) and does not respond to their opponent (stoic/stubborn),
- 1: prefers a coin toss, but responds in opposition to their opponent by doing whatever the opponent didn't do (deceptive)
- 2: usually defects (distrusting) and is deceptive
- 3: defects deterministically (distrusting, stoic),
- 4: prefers to cooperate (trusting) and attempts to enforce a social-contract by imitating their opponent (law-enforcing). 4 is the famous "tit-for-tat" agent [17].
- 5: is distrusting and law-enforcing,
- 6: deterministically cooperates (trusting, stoic),
- 7: usually cooperates (trusting) and is deceptive.

Figure 4 illustrates a point-wise disc-game embedding constructed using the Schur decomposition, as described in Section 3.2 and in [206]. The input matrix F used to create the embeddings is shown in Figure 1, and the underlying game is described in Appendix Section 6.1. Box 2 offers a simple interpretation based on the relations between strategy and embedded geometry presented in Box 1.

It is unclear whether such an embedding is possible in general. Can we identify a disc-game embedding for all f ? If not, what properties must f satisfy to admit a disc game embedding? Are those properties satisfied by most games of interest? Alternately, if a game cannot be embedded, can it be approximated to sufficient accuracy via a disc-game embedding?

We will show that essentially all payout functions f admit a disc-game embedding, a function f is disc-game embeddable if and only if desideratum (D2) is achievable, and, all transforms T satisfying desideratum (D2) are affine transformations of a disc-game embedding.

3.4.1 Existence and Approximation

Theorem 1: [Essentially All Games are Disc Game Embeddable] If there exists a finite, positive measure ν that is absolutely continuous with respect to the Lebesgue measure on Ω such that $\|f\|_{\nu \times \nu} < \infty$ where $\|f\|_{\nu \times \nu}^2 = \iint_{x,x' \in \Omega \times \Omega} f(x, x')^2 d\nu(x) d\nu(x')$ then f admits a disc-game embedding which can be constructed using the real Schur form of the integral operator:

$$F_\nu[h](x) = \int_{y \in \Omega} f(x, y) h(y) d\nu(y). \quad (11)$$

Remark: The measure ν is not prescribed by the game. Rather, it is chosen by the user. Technically, the measure is needed to define an inner product on the space of functions used for constructing the coordinate transformation. This inner product induces notions of orthogonality, orthogonal projection, and optimal approximation with respect to an associated norm. For example, we will see that, after choosing ν , the embedding coordinates $\{y_j(X)\}_j$ are uncorrelated if X is drawn from ν , and have variances associated with the importance of each embedding coordinate in the overall reconstruction of f . Formally, the measure is used to define a Hilbert space, within which functions on the product space, such as f and g , may be treated like matrices. This analogy follows from the fact that matrices are scalar-valued functions of pairs of index-valued inputs, whereas f and g are scalar-valued functions of pairs of generic inputs.

The measure weights Ω . Regions where the measure places a large mass are important, and regions where the measure places a small mass are not important. Since the measure of the full space does not matter, we will assume that the measure is normalized, and can be treated as a probability measure when convenient. Then, the measure could be chosen to preferentially focus on certain trait combinations of interest, to reflect the distribution of a particular population, to reflect the empirical distribution of some observed set of types, or to represent a generating distribution producing observed agents. Note that the uniform measure is, itself, a finite measure in compact spaces, so replacing with $d\nu(y)$ with dy corresponds to a more restricted theory.

Even though the embedding depends on the measure, we will see that some crucial characteristics, namely, the number of embedding coordinates required, are independent of the choice of measure. Moreover, the introduction of a measure will prove useful since it fixes the embedding in terms of a reference population, allowing consistency across different initial choices of Ω . The latter property is essential. If the embedding depends on the initial choice of Ω , then it is extrinsic. \square

Proof: The proof is a special application of the Hilbert-Schmidt theorem [185] to skew-symmetric linear operators. The operator $F_\nu[\cdot]$ is skew-symmetric, so $iF_\nu[\cdot]$ is Hermitian and thus is self-adjoint. If $\|f\|_\nu$ is finite, then the operator is also compact so is a Hilbert operator [31].

Hilbert operators obey the Hilbert-Schmidt theorem, namely, they admit an expansion onto a sequence of eigenpairs, $\{(\lambda_j, \phi_j)\}_{j=1}^{\text{rank}(F)}$ [185]. When real, skew-symmetric, the eigenvalues λ_j are purely imaginary, come in complex conjugate pairs, and, by convention, are listed in non-increasing order, with positive imaginary eigenvalues listed first. The eigenfunctions $\phi_j : \Omega \rightarrow \mathbb{C}$ also come in complex conjugate pairs and are orthonormal under the inner product induced by ν , $\langle u, w \rangle_\nu = \int_{x \in \Omega} \bar{u}(x) w(x) d\nu(x)$. The eigenfunctions and eigenvalues satisfy the equation:

$$F_\nu[\phi_j](x) = \lambda_j \phi_j(x) \quad (12)$$

and depend on the choice of ν .

The rank of F is the total number of eigenpairs such that $\lambda_j \neq 0$. Unlike the specific eigenvalues and eigenfunctions, the rank is independent of ν provided ν is continuous and is supported everywhere on Ω . Thus, the rank is uniquely defined by the payout f and the support of the reference measure. For details, see Lemma 5. Then, we can write $\text{rank}(f)$ unambiguously. If the rank is finite, then f is degenerate.

The Hilbert-Schmidt theorem expresses F_ν in terms of its eigenfunctions:

$$F_\nu[h](x) = \sum_{j=1}^{\text{rank}(f)} \lambda_j \phi_j(x) \langle \phi_j, h \rangle_\nu \quad (13)$$

for all functions h that are square integrable with respect to ν .

By selecting appropriate test functions h , the payout function also admits an expansion onto the eigenfunctions:

$$f(x, x') = \sum_{j=1}^{\text{rank}(f)} \lambda_j \phi_j(x) \bar{\phi}_j(x') \quad (14)$$

for almost all (x, x') with respect to the product measure $\nu \times \nu$.

Since the eigenfunctions and eigenvalues form complex conjugate pairs, group the sum in consecutive pairs. Then, setting:

$$y^{(k)}(x) = \sqrt{|\lambda_{2k-1}|} [\text{Real}(\phi_{2k-1}(x)), \text{Imag}(\phi_{2k-1}(x))] \quad (15)$$

yields:

$$f(x, x') = g(T(x), T(x')) = \sum_{j=1}^{\text{rank}(f)} \text{disc}(T(x), T(x')) \text{ where } T(x) = [y^{(1)}(x), y^{(2)}(x), \dots, y^{(\text{rank}(f)/2)}(x)]. \quad (16)$$

Thus, if $\|f\|_{\nu \times \nu} < \infty$, f admits a disc-game embedding. \square

Remark: The decomposition is unique up to arbitrary rotations of each pair of coordinates $y^{(k)} = [y_1^{(k)}, y_2^{(k)}]$ provided we require that each coordinate $y_j(X)$ has equal variance when the input X is drawn from the measure ν , and all of the eigenvalues are distinct. If some of the eigenvalues are identical, then the decomposition is unique up to arbitrary rotations among the set of coordinates corresponding to matching eigenvalues (c.f. [206]). \square

Remark: Since the eigenfunctions of a Hermitian operator are orthonormal, the disc-game embedding functions $\{y_j\}_{j=1}^{\text{rank}(F)}$ are mutually orthogonal with respect to the inner product induced by ν . That is, all distinct embedding coordinates $y_i(X)$, $y_j(x)$ are uncorrelated when X is drawn from ν . \square

Remark: The disc-game embedding introduced in Theorem 1 is an example of a spectral embedding. Spectral embeddings represent objects via the eigenvectors or eigenfunctions of an operator. Other popular examples include principal component analysis, multi-dimensional scaling, diffusion maps, and spectral graph embeddings [2, 51, 88, 117, 122, 174, 214]. Like other spectral embeddings, disc-game embedding may be defined by minimizing a squared error between a reconstruction of some reference object, in our case, f , and its approximation, where the approximation is bilinear (a quadratic form), in some nonlinear transformation of the input coordinates.

Here, the complexity of the original object is absorbed into the nonlinear coordinate transformation, while its essence is distilled by the quadratic form. Often, the quadratic form is simply an inner product. For example, large-language models and natural language processing frequently use token-embeddings where a comparison is reduced to an inner product after some change of coordinates. Similar spectral approaches are popular in methods that leverage reproducing kernel Hilbert spaces [28, 173], where either the Hilbert-Schmidt theorem [185], or Mercer's theorem [230], are used to decompose a kernel into its eigenfunctions, the eigenfunctions are used for embedding, and the kernel is replaced with an inner product in the embedding space [230]. The disc-game embedding used here simply replaces the inner product with a cross-product. For an in-depth discussion of the significance of this exchange, see [206]. \square

Many important examples of functional form game satisfy the requirements of Theorem 1. For example, Theorem 1 applies to all bounded functions f defined over compact Ω . This case applies to all extensive form games with finite payouts with traits equal to action probabilities. It requires only that the variance in payout, when sampling X and X' independently and identically from ν is finite for some continuous ν supported everywhere on Ω .

Importantly, the finite norm requirement introduced in Theorem 1 can be dropped when the rank of f is finite. In particular, all degenerate f are also disc-game embeddable.

Definition: A function $f : \Omega \times \Omega \rightarrow \mathbb{R}$ is **separable** if it can be expressed as a finite sum of **product** functions:

$$f(x, x') = \sum_{j=1}^r g_j(x) h_j(x') \text{ where } r < \infty. \quad (17)$$

Separable skew-symmetric functions are degenerate by construction. They may be expressed:

$$f(x, x') = b(x)^\top F b(x') \quad (18)$$

where $F \in \mathbb{R}^{m \times m}$ is skew-symmetric, $b(x)^\top = [b_1(x), b_2(x), \dots, b_m(x)]$ is a vector-valued function whose entries $\{b_j\}_{j=1}^m$ each map from Ω to \mathbb{R} , and m is finite. The collection of functions $\mathcal{B} = \{b_j\}_{j=1}^m$ is a function basis. If the basis functions are linearly independent, then the rank of the linear operator F_ν coincides with the rank of the coefficient matrix F .

Theorem 2: [Separable Functions are Disc Game Embeddable] If there exists a function basis \mathcal{B} containing finitely many functions $\{b_j\}_{j=1}^m$, $b_j : \Omega \rightarrow \mathbb{R}$, such that $f(x, x') = b(x)^\top F b(x')$ for some skew-symmetric $F \in \mathbb{R}^{m \times m}$, then f is disc-game embeddable using $r = \text{rank}(F)$ coordinates.

Proof: Given a payout function of the form (18), the real Schur form of F offers a disc-game decomposition. If $F = Q D_\omega^{1/2} U D_\omega^{1/2} Q^\top$, then setting $T(x) = D_\omega^{1/2} Q^\top b(x)$ gives:

$$f(x, x') = b(x)^\top F b(x') = T(x)^\top U T(x') = \sum_{k=1}^{\text{rank}(F)} \text{disc}(y^{(k)}, y'^{(k)}) \text{ where } y^{(k)} = [T_{2k-1}(x), T_{2k}(x)]. \quad (19)$$

Thus, any separable function is disc-game embeddable. \square

Since \mathcal{B} and m were arbitrary, the class of separable payouts of the form (18) is quite large. It includes all skew-symmetric polynomials, as well as all piecewise polynomials with finitely many pieces. Since the payout function of any finite extensive form game is a polynomial in the decision probabilities assigned to each branch in the game tree, the payout for all finite extensive form games are disc-game embeddable. Indeed, any finite combination of any classical function bases, whether polynomial, piecewise polynomial, Fourier, wavelet, or kernel are disc-game embeddable.

Theorem 1 and Theorem 2 are related. When the linear operator F_ν has finite rank, then the Hilbert-Schmidt theorem guarantees that f can be expressed as a separable function, where each basis function is an eigenfunction of the operator. Thus, when F_ν is degenerate, f is disc-game embeddable using the eigenfunctions as a function basis. If all of the eigenfunctions themselves can be expressed as a linear combination of other basis functions \mathcal{B} , then f is also a finite linear combination of product functions built from the basis \mathcal{B} .

Conversely, when $\|f\|_\nu$ is finite, then Theorem 2 can be extended to recover Theorem 1 by considering the closure of the set of separable functions as the order m increases. This is, in essence, an approximation argument. If f cannot be expressed as a finite sum of product functions, but can be approximated to arbitrary accuracy by sufficiently many product functions, then f is in the closure of the set of separable functions. For example, consider analytic functions, which may be expressed locally as power series, thus approximated to arbitrary accuracy by finite linear combinations of monomials of increasing degree.

The closure of a set of separable functions depends on the form of convergence enforced. The form of convergence requires a norm on functions used to measure the error in an approximation. Enter ν .

Let ν denote a non-negative measure supported on Ω . Let $\langle \cdot, \cdot \rangle_\nu$ be the inner product with respect to ν , and $\|\cdot\|_\nu$ the induced two-norm. Similarly, let $\nu \times \nu$ be the product measure on $\Omega \times \Omega$, $\langle \cdot, \cdot \rangle_{\nu \times \nu}$ the associated inner product, and $\|\cdot\|_{\nu \times \nu}$ the associated norm. Then, let $\mathcal{B}^{(m)}$ be a collection of m basis functions that are orthogonal with respect to $\langle \cdot, \cdot \rangle_\nu$, and normalized with respect to $\|\cdot\|_\nu$. This basis can be extended to a product basis by letting $\mathcal{B} \times \mathcal{B}$ be the set of all product functions of the form $b_i(x)b_j(x')$. Then the product basis is $\langle \cdot, \cdot \rangle_{\nu \times \nu}$ orthogonal and all of product basis functions are $\|\cdot\|_{\nu \times \nu}$ normalized. Let $\text{Proj}_{\nu \times \nu, \mathcal{B} \times \mathcal{B}}$ be the projection operator such that:

$$\text{Proj}_{\nu \times \nu, \mathcal{B} \times \mathcal{B}}[f] = \sum_{i,j=1}^m F_{ij} b_i(x) b_j(x') = \underset{\hat{f} \in \text{span}(\mathcal{B} \times \mathcal{B})}{\text{argmin}} \{\|f - \hat{f}\|_{\nu \times \nu}\}. \quad (20)$$

When the basis functions are orthonormal:

$$F_{ij} = \langle b_i b_j, f \rangle_{\nu \times \nu} = \iint_{x, x' \in \Omega \times \Omega'} b_i(x) f(x, x') b_j(x') d\nu(x) d\nu(x'). \quad (21)$$

Otherwise, the coefficient matrix F can be recovered by either orthonormalizing the basis via a Gram-Schmidt procedure, or by inverting the Gram matrix storing all the inner products of the basis functions. Notice that these equations only make sense when working in the space of functions whose $\|\cdot\|_\nu$ and $\|\cdot\|_{\nu \times \nu}$ norms are finite. That is, within the Hilbert space associated with the inner product defined by the product measure [185].

Since any separable payout function is disc-game embeddable, $\hat{f}^{(m)} = \text{Proj}_{\nu \times \nu, \mathcal{B} \times \mathcal{B}}[f]$ is disc-game embeddable. So, if $\hat{f}^{(m)}$ converges to f in $\|\cdot\|_{\nu \times \nu}$, then there exists a sequence of disc-game embeddings that converge to f , and there exists a disc-game embedding that recovers f to arbitrarily small error. Importantly, the disc-game embedding produced by embedding $\hat{f}^{(m)}$ provides the closest approximation to f in $\|\cdot\|_{\nu \times \nu}$ among all transformations $T \in \text{span}(\mathcal{B}^{(m)})$.

Theorem 1 and Theorem 2 are unified by searching for a basis \mathcal{B} such that $\hat{f}^{(r)}$ is the closest approximation to f in $\|\cdot\|_{\nu \times \nu}^2$ among all separable functions of rank r . In particular, by extending the Eckart-Mirsky-Young theorem [66, 148, 207], the best rank $r < \text{rank}(F)$ embedding is recovered by truncating spectrum of F_ν :

Lemma 3: [Optimal Rank- $2r$ Approximation] The rank $2r \leq \text{rank}(F)$ approximation:

$$\hat{f}^{(2r)}(x, x') = \sum_{k=1}^r \text{disc}(y^{(k)}, y'^{(k)}) \quad (22)$$

defined by projection onto the eigenbasis $\mathcal{B} = \{\text{Real}(\phi_1), \text{Imag}(\phi_1), \dots, \text{Real}(\phi_{2r-1}), \text{Imag}(\phi_{2r-1})\}$ is equivalent to truncating the exact disc-game embedding, (16), after r terms, and minimizes the approximation error $\|\hat{f}^{(2r)} - f\|_{\nu \times \nu}$ among all rank $2r$ separable functions (all combinations of $2r$ linearly independent, product functions).

As long as $\|f\|_{\nu \times \nu} < \infty$, the sequence of approximations converges, and the error in the approximation is:

$$\|\hat{f}^{(m)} - f\|_{\nu \times \nu}^2 = 2 \sum_{k>r} |\lambda_{2k}|^2. \quad (23)$$

Thus, the sequence of eigenvalues of F_ν control the rate of convergence of all separable approximations to f by setting an achievable lower bound on the $\|\cdot\|_{\nu \times \nu}$ error over all rank $2r$ functions.

Together, Theorems 1 and 2 establish that essentially all games of interest are either exactly disc-game embeddable, or may be approximated to arbitrary accuracy with a disc-game embedding. To approximate, either start by identifying the eigenfunctions under a particular choice of ν , then truncate, or, choose a complete function basis (e.g. polynomials), project onto the function basis, then embed the projection.

More strongly still:

Theorem 3: [Disc Game Embeddings Are Sufficient] Any coordinate transformation satisfying desideratum (D2) using finitely many coordinates is related to a disc-game embedding by a linear transformation.

Proof Outline: Any coordinate change satisfying desideratum (D2) with finitely many coordinates produces bilinear g , and all finite-dimensional bilinear g are separable, thus disc-game embeddable. See Appendix Section 6.2.3 for details \square

It follows, that any coordinate transformation T that satisfies desideratum (D2) and uses finitely many coordinates, may be reduced to disc-game embedding by a linear mapping. Since the disc-game embedding separates the coordinates into non-interacting pairs, there is little reason to adopt a transform satisfying desideratum (D2) that is not a disc-game embedding.

3.4.2 Consistency and Uniqueness

To isolate an intrinsically meaningful coordinates from any set of sufficient coordinates, the coordinate system produced cannot depend on the initial coordinate system chosen. Suppose that $w = M(x)$ for some invertible mapping M . Then an embedding is intrinsic if and only if the coordinates assigned to any agent after the embedding are the same no matter whether the agent was initially described using x or w .

Let $M : \Omega \rightarrow \Phi$ denote a one-to-one map and $w = M(x)$ the change of coordinates associated with M . Let $h(w, w') = f(M^{-1}(w), M^{-1}(w'))$. Let ν denote a measure over Ω , and μ denote the associated measure over Φ where $\mu(A) = \nu(M^{-1}(A))$ for any set A .

Definition: [Embedding Equivalency] Two disc-game embeddings are equivalent if the coordinates assigned to all agents are equal up to a rotation within each set of coordinates sharing an eigenvalue.

Lemma 4: [Composition Consistency] Given $M : \Omega \rightarrow \Phi$, the disc-game embedding constructed using F_ν is equivalent to the disc-game embedding constructed using H_μ in the sense that $\{y_j(x)\} \sim \{z_j(M(x))\}$ if y and z are the disc-game embeddings for (f, Ω, ν) and (h, Φ, μ) respectively, and $\mu(A) = \nu(M^{-1}(A))$ for any $A \subseteq \Phi$.

Remark: Lemma 4 shows that, by introducing a measure over Ω , it is possible define an embedding procedure that is entirely intrinsic, that is, whose output coordinates are independent of the choice of initial attributes up to invertible transformations. In other words, the embedding of the payout is unique up to the choice of measure.

Note that this consistency is impossible in the absence of the measure, since it is the consistency of the measures across representations that ensures a consistent geometry in the Hilbert space used to construct the embedding. More direct methods based on the expansion of a separable function do not satisfy Lemma 4, since they depend on the specific functional form of f , and thus, on the particular choice of Ω .

In this sense, Theorem 1 and Lemma 3 provide transformations into an intrinsic coordinate system. Lemma 4 shows that the embedding produced by Theorem 1 or Lemma 3 depends only on payout f and the measure ν , not the initial representation selected. This is natural when ν models a population, or distribution generating samples. The embedding is a function of *who* makes up a population and *how* they interact, not the attributes used to *describe* the members of the population. \square

Since the embedding is defined with respect to a user-chosen reference measure, ν , it is essential to distinguish which features of the embedding depend on ν and which do not. In particular:

Lemma 5: [Range and Rank Invariance] If ν and μ are each continuous reference measures, both supported on the same set $S \subseteq \Omega$, then $\text{range}(F_\nu) = \text{range}(F_\mu)$ so $\text{rank}(F_\nu) = \text{rank}(F_\mu)$ and the embeddings are related by invertible linear transformation.

Proof: Suppose that $h(x) \in \text{range}(F_\nu)$. Then there exists a function g such that $h(x) = \int_{x' \in S} f(x, x') g(x') d\nu(x')$. Since ν and μ are both supported on S , $h(x) = \int_{x' \in S} f(x, x') \left(g(x') \frac{d\nu}{d\mu}(x') \right) d\mu(x')$, where $\frac{d\nu}{d\mu}(x')$ is the Radon-Nikodym derivative of ν with respect to μ . The derivative is finite since μ and ν are both supported on S , and are continuous, so the derivative equals a ratio of finite, non-zero densities. Then $h(x) \in \text{range}(F_\mu)$ since $h(x) = F_\mu[g(\cdot) \frac{d\nu}{d\mu}(\cdot)](x)$.

The rank of F_ν is the dimension of its range. Since the range of F_ν equals the range of F_μ , their ranks also match. The eigenfunctions under either measure form orthonormal bases for this function space, so must be related by an invertible linear transformation. \square

Lemma 5 establishes that both the range and rank of the operator F_ν depend only on f and $S = \text{support}(\nu)$, and are otherwise invariant to changes in ν that preserve the support set S . This establishes, that the rank is well defined by f and S alone. We will use this fact when study learning dynamics to show that changing the reference measure while preserving the support set S produces topologically equivalent dynamics. This will allow more flexible analysis, since the reference measure can be adjusted strategically.

3.4.3 Regularity

Recognizing the disc game embedding functions as eigenfunctions of the operator F_ν ensures that the embeddings inherit regularity properties of f . For instance, if f is Lipschitz continuous, then, since the embeddings are scaled versions of the eigenfunctions of f , and since the eigenfunctions of f live inside the range of F_ν , all of the embedding functions will be Lipschitz continuous with Lipschitz constants determined by the eigenvalues of F_ν and the Lipschitz constant of f . Similar statements hold for differentiability. It also follows that, if f is restricted to a particular function class, then its embeddings may be as well. For example, if f is a polynomial, as for finite extensive form games, then its range contains polynomials of matching order, so the embeddings must be polynomials of matching order. Symmetries of f also constrain the behavior of its eigenfunctions. For example, if f is stationary in the sense that it is translationally invariant, $f(x, x') = s(x - x')$ for some $s : \Delta\Omega \rightarrow \mathbb{R}$, then the eigenfunctions of f may be restricted to Fourier modes [206]. The last example arises naturally for certain auctions and bidding games. Observations of this kind can be used to guess a relevant function basis when estimating the embeddings from data, or, to ensure that the embeddings are not exceedingly bizarre functions of the initial traits. Conceptually, regularity results of this kind ensure that the embeddings are only as exotic as f , so only as exotic as need be given the original agent descriptions.

One regularity statement is relevant for the following analysis. Namely, the embeddings are continuous functions of f in the sense that, if two distinct agents x and x' , share similar outcomes for all opponents x'' , then they will share similar embeddings y and y' . In short, agents who play together, stay together once embedded.

Lemma 6: [Outcome Continuity] If $|f(x, x'') - f(x', x'')| \leq \delta$ for almost all x'' with respect to ν , or $\|f(x, \cdot) - f(x', \cdot)\|_\nu = \mathbb{E}_{X'' \sim \nu}[(f(x, x'') - f(x', x''))^2] \leq \delta^2$, then $\|y^{(k)}(x) - y^{(k)}(x')\|_2 \leq |\lambda_{2k-1}|^{-1/2}\delta$.

For the proof, see Appendix section 6.2.4.

This is a desirable property for an intrinsic set of coordinates. It ensures that agents who perform similarly are embedded near each other. Lipschitz continuity and differentiability follow if f is Lipschitz or differentiable since any smoothness statement relating perturbations in traits to perturbations in outcome can be extended, via Lemma 6, to perturbations in embedding. It also ensures that agents who cannot be distinguished by their payout alone are assigned the same coordinates once embedded (see Lemma 11). This reduction by outcome equivalency will be useful when analyzing selection dynamics, since agents who are outcome equivalent are subject to the same selective pressures in canonical evolutionary dynamics.

3.5 Disc Game Embeddings as a Candidate Latent Space for Learning

Since disc-games satisfy desideratum (D2), allow easy manipulation of averages via bilinearity, decouple blockwise, represent strategy with intuitive geometry, and can be defined uniquely up to the choice of a reference measure, we propose the disc-game coordinates as a candidate latent space for learning and selection dynamics. For an alternate derivation based on accumulated training advantages, see Appendix Section 6.3. By adopting a canonical coordinate transformation, we, in essence, adopt a canonical game. Since essentially all symmetric, zero-sum games admit a disc-game embedding, up to a change of coordinates, essentially all symmetric, zero-sum games *are* combinations of disc-games.

In this sense, the disc-game payout replaces other canonical payout models, such as differences in fitness [23]. Unlike fitness landscapes, which only describe transitive competition, the disc-game payout is generic. Moreover, unlike fitness “seascapes” [143, 156], or frequency-dependent fitness landscapes [162], which are generic, but change as a population changes, the disc-game payout is fixed. Instead, individual games differ only through the set of admissible agents, Ψ . In turn, fixing the payout function will simplify dynamical analysis, since, up to a change of coordinates, all selection dynamics respond to the same payout. Based on its generality and interpretability, we contend that selection in response to a disc-game payout should replace optimization of a fitness landscape as the canonical mental model for evolution in response to a skew-symmetric payout.

A successful mental model should guide intuition and ease communication through analogy or extension of familiar ideas. This process is often aided by visualization, especially when the analogy builds on real-world processes. To work as a successful analogy the model should accurately map the structural relations between components of the unfamiliar and the familiar system [80].

Does optimization satisfy these criteria? It is familiar, visualizable, even concrete when related to “hill-climbing.” Yet, the intuition it provides is inaccurate, even misleading. Optimization suggests directed progress towards optima. As discussed in Section 2.1, and shown in Section 4.2.5, this behavior is not generic, indeed is atypical, in simple settings.

Do disc-games offer a compelling alternative? They are strictly more general (see Section 3.4), visualizable, and interpretable (see Section 3.3). What intuition do they provide?

Disc games suggest recurrent motion, even periodic circulation, unless barred by the boundaries of Ψ . The basic physical processes suggested by disc-games are all oscillatory - planetary motion, pendulum motion - etc. In the next section, we demonstrate that this intuition is not only qualitatively correct, it is mathematically precise.

4 Learning in the Latent Space

To demonstrate the utility of the disc-game embedding, we turn to selection and learning dynamics. For classical reviews, see [77, 78, 103, 228]. We focus on the replicator dynamic [212] as a case study. Sample results for other dynamics are presented in Box 3. These dynamics are detailed in Appendix Section 6.4, and solved in Appendix Section 6.5.

Box 3. Explicit Solutions

Explicit solutions are described below for three standard learning dynamics. These all set the rate of change in the attributes of an agent, or set of agents, proportional to the average optimal training vector for the agent in response to a reference population. In particular, all of the dynamics take the form:

$$\frac{d}{dt}y(j, t) \propto \mathbb{E}_{y' \sim \pi(t)}[v(y, y')] \quad (24)$$

where j indexes a particular agent, and $\pi(t)$ is the reference population at time t .

These dynamics all admit explicit solutions when posed in the disc-game space, and when the set of agents are contained in the interior of Ψ . The following solutions all leverage the simplifying characteristics used to design the disc-game embedding. Namely, since disc-games satisfy desideratum (D2), the optimal response directions are student-independent, and linear in the opponent attributes. Since averages are linear operations, the average optimal response direction equals the optimal response direction at the reference population centroid. Then, since the disc-game is defined by a block-diagonal matrix, consecutive pairs of disc-game coordinates form separate blocks that evolve independently. Thus, the problem can be solved blockwise. That is, a single disc-game at a time. Accordingly, all solutions are presented for one pair of coordinates. Finally, since the disc-game is bilinear, and the optimal training vector field is a rotationally symmetric linear function, all three dynamics are highly symmetric linear systems differential equations which can be solved by finding a particular solution for one initial condition. For detailed solution techniques, see Appendix Section 6.5.

Self-Play [78]

Reference Population: The agent

Solution: Simple harmonic motion. Coordinate pairs orbit on circular paths at a fixed rate. Orbits move clockwise.



Fictitious Self-Play [78]

Reference Population: All past agents.

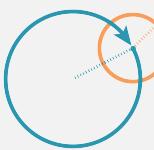
Solution: Asymptotic to logarithmic spirals. The phase converges to $-\sqrt{2t}$, while the radius grows exponentially in the phase. Exact solutions are Kelvin spirals. The figure (left) is log-scaled.



Simultaneous Gradient Ascent [22]

Reference Population: All current agents.

Solution: Epicycles. The population centroid (blue) orbits as in self-play. Each agent (orange) orbits backwards about the centroid at rate $1/n$, where n is the number of agents.



The replicator dynamic admits multiple reasonable derivations [169]. These include mean-field approximations, and infinite population limits, for imitation learning in finite-population models [29, 33], for bandits [191], and for selection in population genetics [162], as a limited rationality model for inference [113], as a self-play dynamic in the space of mixed strategy distributions equipped with an appropriate metric [125, 145], or as the time evolution of mixed strategies under an exponential/multiplicative-weight learning rule [15, 130, 226, 72] which is the gold-standard example of a no-regret learning rule since it is both easy to implement and satisfies a constant regret bound [12, 124, 125]. Indeed, the replicator dynamic satisfies a *constant* cumulative regret upper-bound, with incremental regret converging to zero faster than t^{-1} [32]. See [58, 162] for reviews.

The replicator dynamic sets the per-capita growth rate of a population of type x proportional to the expected payout of type x against an opponent drawn at random from the current population [162]. In particular, we focus on the continuous replicator equation [53, 55], which extends the classical replicator dynamic to a measure-valued dynamic on continuous trait spaces [53, 54, 56, 55, 57]. We focus on the continuous setting since it is less well explored, results in the continuum

All of the dynamics described in Box 3 are based on advective models in which agents change type according to the best-response gradient v , but without changing the overall population size. In this sense, these are explicitly learning dynamics. They consist of a fixed population of agents each moving to optimize their payout against an average over a population. In these cases, the optimal training vector field enters explicitly, so a coordinate system that simplifies v simplifies the dynamics.

In contrast, ecological population dynamics typically start from a birth-death process, in which individuals are produced and destroyed, but do not explicitly change types. Here, selection replaces training as the primary mechanism. Nevertheless, we will show that the disc-game embedding is a useful tool even when the dynamic of interest does not explicitly reference the training vector field.

4.1 The Replicator Dynamic

The replicator dynamic [103, 212] is the canonical ecological dynamic of evolutionary game theory. As the Lotka-Volterra equations [131, 132, 224] set baseline expectations in ecology, the replicator equation sets baseline expectations in evolutionary game theory. In fact, the two dynamics are equivalent up to an appropriate transformation [101, 103, 169].

often provide useful analytic limits for discrete spaces, and continuous trait spaces are natural in learning problems that involve a continuous range of possible parameter values or action probabilities. For analogous results in the discrete setting see [7, 102, 32].

Formally, a measure-valued dynamic uses a time-evolving measure $\Pi[\cdot](t)$, where $\Pi[S](t)$ is the number of individuals with traits in the set $S \subset \Omega$ at time t . The existence and uniqueness of measure-valued solutions to the continuous replicator equation is established in [164, 165] provided f is continuous.

If the population distribution is continuous, then the measure admits a density π , where $\pi(x, t)$ is the population density in trait space at type x and time t . Let:

$$P(t) = \Pi[\Omega](t) = \int_{x \in \Omega} \pi(x, t) dx \quad (25)$$

denote the total population size. Note that $P(t)$ need not equal 1. That is π need not be normalized.

Then, given skew-symmetric f , and a continuous measure, the replicator equation sets:

$$\partial_t \pi(x, t) = \frac{\rho(P(t), t)}{P(t)} \mathbb{E}_{X' \sim \pi(\cdot, t)} [f(x, X')] \pi(x, t) \quad (26)$$

where $\rho(P, t) \in \mathbb{R}^+$ represents the per capita interaction rate between individuals given a total population of size P at time t in an underlying agent based model [7]. For agent based derivations of the replicator equation, see [29, 33].

The standard replicator equation studies the change in the relative frequency of types implied by the dynamic (26). Since we assume that f is skew-symmetric, and ignore scalar factors, there is no need to distinguish between the frequency of types, and the distribution representing the population. When f is skew-symmetric, the two dynamics differ by a constant factor equal to the total population size. Scaling the right-hand side of the differential equation by a constant is equivalent to scaling the units used to measure time, so the trajectories generated by the traditional replicator dynamic and the form given above are equivalent up to a change of time units.

Notice that, since f is skew-symmetric, the total population is necessarily conserved by the dynamic:

$$\frac{d}{dt} P(t) \propto \int_{x \in \Omega} \mathbb{E}_{X' \sim \pi(\cdot, t)} [f(x, X')] \pi(x, t) dx = \mathbb{E}_{X, X' \sim \pi \times \pi} [f(X, X')] = 0$$

where the latter inequality follows since $\mathbb{E}_{X, X' \sim \pi \times \pi} [f(X, X')] = -\mathbb{E}_{X, X' \sim \pi \times \pi} [f(X', X)]$ since f is skew-symmetric, but $\mathbb{E}_{X, X' \sim \pi \times \pi} [f(X, X')] = \mathbb{E}_{X, X' \sim \pi \times \pi} [f(X', X)]$ since X and X' are drawn identically so are exchangeable. It follows that $P(t) = P(0) = P$ is constant.

To standardize the analysis, we will make two assumptions, one without loss of generality.

First, we will restrict our attention to solutions that can be expressed with time-evolving densities. This is sufficient to explain transient finite-time dynamics and mixing, though may not be sufficient to explain all stability concerns for singular (monomorphic) steady-state distributions associated with equilibria. For stability considerations see [54, 57]. Note that not all static equilibria (e.g. Nash equilibria) are stable under the replicator dynamic, nor do sufficient conditions (interior evolutionary stability [200]) developed in discrete trait spaces necessarily extend to continuous trait spaces unless the topology on measures is chosen appropriately [164, 165]. Moreover, no-regret learning dynamics, which include the replicator dynamic, fail to converge to interior, fully-mixed equilibria in zero-sum normal form games with finitely many strategies. So, unless there exists a single dominating strategy, no-regret learning dynamics wander persistently [7, 20, 102, 144, 72]. Since we are concerned with a dynamical analysis, rather than a limiting analysis, we leave further discussion of measure-theoretic complications to the literature.

Second, we will also assume that the interaction rate is linear in the total population size: $\rho(P(t), t) = P(t)$. This assumption is made without loss of generality. If $\rho(P, t)$ is not equal to P , then all solutions to the replicator equation are equal to solutions of the replicator equation with a linear reaction rate after transforming the time coordinate.

In particular, let $\tau(t)$ represent a monotonically increasing function of time. Then $\tau(t)$ is invertible, so we can exchange time coordinates freely. Write $\pi(\cdot, t) = \pi(\cdot, t(\tau))$. Then,

$$\begin{aligned} \pi(x, t) &= \tilde{\pi}(x, \tau(t)) \text{ where } \begin{cases} \partial_t \pi(x, t) = \frac{\rho(P(t), t)}{P(t)} \mathbb{E}_{X' \sim \pi(\cdot, t)} [f(x, X')] \pi(x, t) \\ \partial_\tau \tilde{\pi}(x, \tau) = \mathbb{E}_{X' \sim \tilde{\pi}(\cdot, \tau)} [f(x, X')] \tilde{\pi}(x, \tau) \end{cases} \\ \text{and: } \tau(t) &= \int_{s=0}^t \frac{\rho(P(s), s)}{P(s)} ds, \quad t(\tau) = \int_{s=0}^\tau \frac{P(s)}{\rho(P(s), \tau)} ds. \end{aligned} \quad (27)$$

Thus, all solutions to (26) can be recovered from the solution when $\rho(P, t) = P$ by changing the time coordinate. From now on, we will write:

$$\partial_t \pi(x, t) \propto \mathbb{E}_{X' \sim \pi(\cdot, t)} [f(x, X')] \pi(x, t) \quad (28)$$

with the understanding that the proportionality represents an unspecified interaction rate ρ , and will solve all systems using the convention that $\rho(P, t) = P$ so that the proportionality is an equality.

4.2 Replicator Parameter Dynamics Via Disc Game-Embedding

Here, we demonstrate that the disc-game embedding proposed in Section 3 is a useful coordinate change when studying parametric solutions to the replicator dynamic.

4.2.1 Objectives and Special Cases

A parametric solution to a density-valued dynamic is defined by fixing a parametric family of densities that is invariant under the dynamic, then solving for the parameter dynamics. Specifically, we write $\pi(x; \theta(t))$ for the density at traits x given a parameter vector θ that evolves in time. We seek solutions for families of densities with finitely many degrees of freedom. Typically, we will seek $\theta(t) \in \mathbb{R}^r$ for some integer-valued $r < \infty$. When no such solution is available, we seek consistent approximation methods that converge in the limit as r diverges. The latter problem is an example of a closure problem (e.g. [158]). Exact closure is possible if the replicator dynamics can be reduced to the dynamics of a finite set of parameters without any approximation error for any initial distribution.

The formal problem follows. Let $\Theta \subseteq \mathbb{R}^r$ be an admissible set of parameters. Let $\mathcal{R}(\Theta) = \text{range}_{\theta \in \Theta}(\pi(\cdot; \theta)) = \{\pi(\cdot; \theta) \mid \theta \in \Theta\}$ denote the set of all density functions that can be expressed by choosing parameters from Θ . A set of densities is invariant under the replicator dynamic (26) if, given $\pi(\cdot, t) \in \mathcal{R}(\Theta)$ at some time t , then $\pi(\cdot, t') \in \mathcal{R}(\Theta)$ for all future times $t' \geq t$. We aim to find $(\Omega, f, \mathcal{R}(\Theta))$ that are invariant under the replicator dynamic.

Problem: Given a functional form game (Ω, f) , does the replicator dynamic (26) admit an exact closure using $r < \infty$ degrees of freedom, and, if so, for what range of parametric densities $\mathcal{R}(\theta) = \{\pi(\cdot; \theta) \mid \theta \in \Theta\}$?

We will use the disc-game embedding to answer both questions. To build intuition for the solution, and the ensuing parameter dynamics, we consider two special cases.

1. **Perfectly Transitive Functions:** Suppose that $f(x, x') = r(x) - r(x')$ for some rating function $r : \Omega \rightarrow \mathbb{R}$. Then, f is a perfectly transitive function [205]. When f is perfectly transitive, the replicator Equation (28) can be solved exactly for any initial distribution. In particular (see Appendix Section 6.5.2),

$$\pi(x, t) \propto e^{tPr(x)} \pi(x, 0), \text{ where } P = P(0). \quad (29)$$

Equation (29) is an example of a parametric solution with two degrees of freedom that evolve in time. Given $f(x, x') = r(x) - r(x')$ consider parametric distributions of the form $\pi(x, t) = \pi(x, 0) \exp(\theta_0(t) + \theta_1(t)r(x))$. Then the solution (29) sets $\theta_1(t) = Pt$, where $\theta_0(t)$ is implied by $\theta_1(t)$ via population conservation. More generally, we could write, $\pi(x, t) = e^{u(x, t)}$ where $u(x, t)$ is the log-density at traits x and time t . Then, the transitive solution spans two-dimensional affine subspace of functions of the form $u(\cdot, t) \in u(\cdot, 0) + \text{span}\{1, r(\cdot)\}$ where $\text{span}\{1, r(\cdot)\}$ is all possible linear combinations of the constant function and the rating function.

The constant function and rating function correspond to the disc-game embedding functions when f is perfectly transitive since they correspond to the real and imaginary parts of the eigenfunctions of the operator F . In fact, $f(x, y) = r(x)1 - 1r(y) = [r(x), 1] \times [r(y), 1]$ is already in the form of a disc-game decomposition with embedding functions $r(\cdot)$ and 1 . We will show that the parameters θ_{2k-1} and θ_{2k} corresponding to paired embedding coordinates are always related by a Hamiltonian system of ODE's.

The transitive solution (29) also illustrates a generic reduction technique that will be accomplished automatically by disc-game embedding. Notice that if x and x' satisfy $r(x) = r(x')$ then $\pi(x, t)/\pi(x', t) = \pi(x, 0)/\pi(x', 0)$ is constant. It follows that the conditional population distribution remains constant along every level set of the rating function r . This is an example of a reduction by outcome equivalency.

Definition: [Outcome Equivalency] Trait vectors x and x' are **outcome equivalent**, $x \sim x'$, if $f(x, y) = f(x', y)$ for almost all $y \in \Omega$.

When $\pi(x, t)$ obeys the replicator dynamic, the population at outcome equivalent traits evolve in tandem. Indeed, for any $x \sim x'$, $\pi(x, t)/\pi(x', t) = \pi(x, 0)/\pi(x', 0)$ for all t . It follows that the conditional population distribution on each set of outcome equivalent traits remains constant under the replicator dynamic. Thus, the dynamic can be reduced by grouping outcome equivalent traits into a single representative trait for each equivalency class, then studying the replicator equation on the equivalency classes.

We will show that all solutions generated using a disc-game embedding automatically reduce by equivalency class since all outcome equivalent trait vectors are mapped to the same set of embedded coordinates. To retain the information needed to separate distinct but outcome equivalent types, we can append additional coordinates to the disc-game embedding that do not contribute to the dynamics. In practice, this is never necessary since the

conditional distributions within each equivalency class match the initial conditionals. Solutions can always be expressed by first solving for the time evolution of the marginal densities, then redistributing the marginal density of each class proportionally according to the initial conditional distribution over outcome equivalent types.

2. **Quadratic Functions:** The transitive solution solves the replicator Equation (26) for all linear f since any linear, skew-symmetric function is perfectly transitive [42]. Consider quadratic f next.

When f is quadratic, the family of Gaussian distributions is invariant under the replicator dynamics [43, 53, 56]. Proofs are provided in [54]. This result is weaker than the result presented for the linear case (transitive case), as it only applies if $\pi(x, 0)$ is Gaussian, and provides no guarantee regarding the stability of the manifold of Gaussian distributions. For example, it does not guarantee that the family of Gaussian distributions is attracting in some neighborhood of the family. Using the disc-game embedding, we will show that the family of Gaussian distributions is neutrally stable (Lyapunov stable) when f is quadratic, so is neither attracting or repelling. For now, we recapitulate the extant result.

Any Gaussian distribution is parameterized by its mean and covariance, so, we assume the parametric form:

$$\pi(x; \theta) \propto \exp\left(-\frac{1}{2}(x - \bar{x})^\top \Sigma_x^{-1}(x - \bar{x})\right) \quad (30)$$

Then, θ consists of the independent entries of the centroid \bar{x} and covariance Σ_x . Since the family of Gaussian distributions is invariant under the replicator dynamic when f is quadratic, we can replace the density-valued dynamic with a finite dimensional system of ODE's relating the rate of change in \bar{x} and Σ_x to f .

To express this ODE, we need a canonical form for quadratic f . All skew-symmetric quadratic functions can be expressed:

$$f(x, x') = r(x) - r(x') + x^\top H_{xx'} x' \text{ where } r(x) = g^\top x + \frac{1}{2}x^\top H_{xx}x \quad (31)$$

where g is a $T \times 1$ vector corresponding to the gradient of $f(x, x')$ with respect to x at $x = x' = 0$, and H_{xx} and $H_{x'x'}$ are $T \times T$ matrices corresponding to the on-diagonal and off-diagonal blocks of the Hessian of f evaluated at $x = x' = 0$. The Hessian blocks are symmetric and skew-symmetric respectively. Then $J = H_{xx} + H_{x'x'}$ is the Jacobian of the self-play vector field $v(z) = \nabla_x f(x, x')|_{x=x'=z} = g + Jz$ [42].

Lemma 7: [Gaussian Invariance] [43] If f is a quadratic function, then the manifold of multivariate normal distributions is invariant under the continuous replicator dynamic (26) so admits an exact closure in terms of its mean and covariance. Given (28), the mean and covariance satisfy the closed system of ODE's:

$$\begin{aligned} \frac{d}{dt}\bar{x}(t) &= \Sigma_x(t)(g + J\bar{x}(t)) \\ \frac{d}{dt}\Sigma(t) &= \Sigma_x(t)^\top H_{xx}\Sigma_x(t) \end{aligned} \quad (32)$$

Equation (32) is an example of a set of parameter dynamics. It replaces the continuous replicator equation, which is density-valued, so operates in an infinite dimensional space, with a closed system of ODE's in finitely many parameters. In this case, $r = T + T(T + 1)/2$ degrees of freedom are needed. We will show, using the disc-game embedding, that r corresponds with the rank(F).

The covariance evolves independently of the centroid, and can be solved explicitly:

$$\Sigma_x(t) = (\Sigma_x(0) - H_{xx}t)^{-1} \quad (33)$$

If H_{xx} is negative definite, as when near a local maximum in the rating function, then the solution is valid for all times, and the population distribution approaches a delta distribution at the maximizer of the rating function $g^\top x + \frac{1}{2}x^\top H_{xx}x$. If H_{xx} has any positive eigenvalue, then the solution only exists for t up until one of the eigenvalues of $\Sigma_x(0) - H_{xx}t$ crosses zero [43]. Then, the population density escapes all bounded neighborhoods in finite time. This illustrates that, while working with densities is sufficient for transient dynamics, it does not explain the full measure-valued dynamic, even for finite time.

Given, $\Sigma_x(t)$ satisfying (33), the centroid obeys the equation:

$$\frac{d}{dt}\bar{x}(t) = \Sigma_x(t)(g + J\bar{x}(t)) = \Sigma_x(t)\nabla_x f(x, \bar{x}(t))|_{x=\bar{x}(t)}. \quad (34)$$

Equation (34) is an example of an *adaptive dynamics* equation [3, 141], which is itself a generalization of self-play [78] that incorporates scaling by a positive definite matrix (see Appendix 6.4). In adaptive dynamics a (typically

concentrated) population is approximated with a monomorphism at its centroid. The motion of the centroid is directed by the self-play vector field $v(z) = \nabla_x f(x, z)|_{x=z}$ [162]. In particular, $\frac{d}{dt}\bar{x}(t) = C(\bar{x}(t), t)v(\bar{x}(t))$ where $\bar{x}(t)$ is the population centroid, $v(z)$ is the self-play vector field, and $C(x, t)$ is a symmetric positive semi-definite matrix representing either Σ_x , variation introduced by mutation, or a metric on Ω [3, 53, 62, 126, 141, 146, 234].

So, as illustrated in [53], when f is quadratic, $\pi(x, 0)$ is Gaussian, and $\pi(x, t)$ obeys the replicator dynamic, then the centroid $\bar{x}(t) = \mathbb{E}_{X \sim \pi(\cdot, t)}[X]$ obeys adaptive dynamics. Unfortunately, this result does not hold for arbitrary initial distributions, nor does it hold for arbitrary f . We will show that, for any separable f , that is, any f that admits a finite-dimensional disc-game embedding, the centroid of the population distribution in the disc-game coordinates obeys adaptive dynamics, no matter the initial distribution. More strongly, we will show that this adaptive dynamics equation fully specifies the original replicator dynamic, so is sufficient to recover the full distributional dynamic.

The quadratic-Gaussian case also illustrates a generic dichotomy in the dynamical behavior of replicator systems. If H_{xx} is nonzero then the distribution either converges to a delta or escapes to infinity. In either case, the density function converges to a limiting function that is ill-defined or is only supported on a subset of Ω . If H_{xx} is zero, then $\Sigma_x(t) = \Sigma_x(0)$ is constant in time, and the centroid obeys a time-autonomous affine ODE. Moreover, there is either a unique solution to $g + H_{xx'}\bar{x} = 0$ corresponding to an interior, fully mixed Nash equilibrium, or there is not. Setting \bar{x} equal to the equilibrium produces a steady-state distribution. The dynamics of \bar{x} are linear when centered about the equilibrium, with all imaginary eigenvalues, so the equilibrium is a center. If no such equilibrium exists, then there is a component of g in the nullspace of $H_{xx'}$, and $\bar{x}(t)$ will diverge to infinity along that direction. Thus, given quadratic f , and Gaussian π_0 , $\pi(x, t)$ either escapes the space of well-defined densities, or orbits about a fully mixed Nash equilibrium. This dichotomy has been shown for normal form games on finite strategy space and for more general no-regret dynamics [5, 32, 72, 144, 178]. We will illustrate the same dichotomy for all symmetric, constant-sum, functional form games.

In the Section 4.2.2 we will show that the embedding maps used for disc-game embedding provide a natural function space for answering the closure problem. Using that function space, we will solve the closure problem in complete generality. We will show that the replicator dynamic admits an exact finite-dimensional closure if and only if f is separable, in which case solutions to the replicator equation require $r = \text{rank}(F)$ dynamic degrees of freedom. Moreover, we will show that the parameter dynamics are block Hamiltonian, so take the generic form of nonlinear coupled oscillators. This generalizes the special solutions developed for the transitive and quadratic cases. It also generalizes the observation that simultaneous gradient ascent is Hamiltonian given a zero-sum normal form game on a finite strategy space to gradient-free learning dynamics and to functional form games on continuous trait spaces [22].

4.2.2 Closure

To simplify the replicator dynamic, recall a standard result (c.f. [7]):

Lemma 8: [Support Preservation] Let $S(t) = \{x \mid \pi(x, t) > 0\}$ denote the support of π at time t . If $\pi(\cdot, t)$ obeys the continuous replicator dynamic, then $S(t) = S(0)$ for all t on the interval containing 0 for which the continuous replicator equation admits unique, density-valued, solutions.

Proof Outline: If $t' > t$ then $S(t') \subseteq S(t)$, since, if $\pi(x, t) = 0$ then $\partial_t \pi(x, s) = 0$. That is, the support can never expand. The reverse statement is established by repeating the same argument but differentiating with respect to time running in reverse. Sufficient conditions for existence and uniqueness are provided in [164, 165]. For instance, if f is continuous and $S(0)$ is compact, or if f is bounded, then the solutions to the continuous replicator equation exist and are unique for all finite time. \square

Since the support is preserved for all t such that solutions to the continuous replicator equation remain density valued, we can, without loss of generality, restrict our attention to solutions on the domain $S(t) = S(0) = S$. In particular, dividing Equation (28) by the population density at each $x \in S$ yields:

$$\partial_t \log(\pi(x, t)) = \frac{\partial_t \pi(x, t)}{\pi(x, t)} = \mathbb{E}_{x' \sim \pi(\cdot, t)}[f(x, x')] \text{ for all } x \in S. \quad (35)$$

Let:

$$u(x, t) = \log(\pi(x, t)) \quad (36)$$

denote the log population density at all $x \in S$.

Since the logarithm is invertible, any parametric solution to the continuous replicator equation $\pi(x; \theta(t))$ can be expressed via a parametric sequence of log-densities $u(x; \theta(t))$. By the chain rule, the continuous replicator equation admits an exact solution in terms of a parametric form if and only if:

$$\nabla_\theta u(\cdot; \theta(t)) \cdot \frac{d}{dt}\theta(t) = \mathbb{E}_{x' \sim \pi(\cdot; \theta(t))}[f(\cdot, x')] \quad (37)$$

for some vector $\frac{d}{dt}\theta(t)$ and where $\pi(\cdot; \theta) = \exp(u(\cdot; \theta))$. The closure equation, (37), must hold for all $x \in S$.

Notice that $\mathbb{E}_{x' \sim \pi}[f(x, x')] = F[\pi]$ where F is the integral operator with a uniform reference measure. Therefore, the right-hand side of the closure Equation (37) is contained in the range of the integral operator F restricted to all densities in $\mathcal{R}(\Theta)$. Thus, a parametric family of densities $\mathcal{R}(\Theta)$ is invariant under the replicator dynamic if and only if:

$$F[\pi(\cdot; \theta)] \in \text{span}(\nabla_\theta \log(\pi(\cdot; \theta))) \text{ for all } \theta \in \Theta \quad (38)$$

and admits a unique parametric solution if the set of functions $\{\partial_{\theta_j} u(\cdot; \theta)\}_{j=1}^r$ are linearly independent for all $\theta \in \Theta$. A simpler sufficient condition is $\text{range}(F) \subseteq \text{span}(\nabla_\theta \log(\pi(\cdot; \theta)))$ for all $\theta \in \Theta$ which can easily be ensured if $u(\cdot; \theta)$ is parametrized using a linear combination of basis functions that span the range(F).

These observations suggest the following parametric form:

$$u(\cdot; \theta) = u_0(\cdot) + \sum_{j=1}^r \theta_j b_j(\cdot), \quad u_0(x) = \log(\pi(x, 0)) \quad (39)$$

where $\mathcal{B} = \{b_j\}_{j=1}^r$ is a collection of linearly independent basis functions $b_j : \Omega \rightarrow \mathbb{R}$ whose span contains the range of the operator F . If F is not degenerate, choose a complete basis \mathcal{B} so that the sequence of partial sums of the projection of any function in the range of F onto a truncation of \mathcal{B} converges to that function as the dimension of the basis diverges. Equation (39) defines an exponential family [67, 69] with sufficient statistic $b(\cdot)$, natural parameters θ , and base measure $\pi(x, 0)$. Like any exponential family, Equation (39) expresses a collection of distributions that can be formed by an exponential tilting of some base measure.

After adopting (39), the closure equations reduce to:

$$\sum_{j=1}^r b_j(x) \frac{d}{dt} \theta_j(t) = F[\pi(\cdot; \theta(t))](x) \text{ for all } x \in S. \quad (40)$$

Any parameterization of u of the form (39) is an affine parameterization. Since the offset u_0 is determined by the initial distribution, the parameterization itself is determined by the choice of basis \mathcal{B} . We will say that an affine parameterization with basis \mathcal{B} is minimal for S if it uses as few dynamic degrees of freedom as possible, that is, if the span of \mathcal{B} could not be made smaller while allowing exact closure (Equation (40)) for all initial distributions with support S .

Theorem 4: [Invariant Families] Let (Ω, f) denote a functional form game, $S \subseteq \Omega$ a support set, and $\pi(x, t)$ is a density over Ω with support S that evolves according to the continuous replicator equation, (26). Then an affine parameterization of the log-density $u(\cdot; \theta) = \log(\pi(\cdot; \theta))$ defines a minimal invariant family of densities for initial distributions with support S if and only if $\text{range}(F) = \text{span}(\mathcal{B})$ where \mathcal{B} is the basis used in the affine parameterization and F is the integral operator defined by the kernel f and support set S . The parameters of the affine parameterization provide an exact closure of the replicator dynamic. The corresponding parameter dynamics are uniquely defined if and only if the basis functions are linearly independent.

Proof Outline: Sufficiency is established by (38). Necessity is established by varying the initial distribution. By choosing a fixed reference distribution with support S , then perturbing about the reference, it is easy to show that the range of F over all distributions with a fixed support has dimension equal to the rank of F on that support. Thus, if we seek an affine parameterization that admits an exact closure for all initial distributions with shared support, then the affine closure must have dimension at least equal to the rank of F . Minimality is established by showing that, if the range of F is a proper subset of the span of \mathcal{B} , then there exists a choice of basis functions such that at least one parameter remains constant over time. Then, since the affine subspace is shifted by the initial distribution, that parameter will remain zero for all time. Thus, the associated direction can be dropped from the parameterization. Uniqueness is established by linear independence of the basis functions defining the linear closure equation, (40). \square

Theorem 4 provides clear instructions for selecting an affine parameterization given the range of F . For example, if $f(x, x')$ is a polynomial function of x and x' with maximal degree m in x , then the range of F consists of polynomial functions of x with maximal degree m and those polynomials provide a minimal, sufficient, affine parameterization of the log-density dynamics. Then, $\pi(x, t)$ is constrained to the form $\pi(x, 0) \exp\left(\sum_{j=0}^m \sum_{|\alpha| < j} \theta_j(t)x^\alpha\right)$ where α is a multi-index. If $f(x, x')$ is linear, $m = 1$, and the time-varying exponential term must be linear in x . In particular, using $f(x, x') = x - x'$ recovers the transitive solution presented in Section 4.2.1. Similarly, if f is quadratic, then $m = 2$, so the log-density is quadratic in x at all times t . The only family of densities whose support is unbounded, and whose logarithms are quadratic functions of their argument are Gaussian distributions. Thus, as stated in Section 4.2.1, when f is quadratic, the family of Gaussian distributions is invariant under the replicator dynamic.

Corollary 4.1: [Rank and Dynamic Dimension] If (Ω, f) is a functional form game, and $\pi(x, t)$ is a density over Ω that evolves according to the continuous replicator equation, (26), with support S , then the replicator equation admits an exact closure using $r \geq \text{rank}(F)$ degrees of freedom, where the rank is evaluated over all distributions supported on S .

Corollary 4.1 establishes the minimal number of degrees of freedom needed to express all solutions to the replicator dynamic with a fixed support. In particular, given a support set S , exact closure is possible using as many degrees of freedom as the rank of the integral operator F over S . If F is degenerate on support S , then this rank is finite. If F is not degenerate on S , then the rank is infinite, so exact finite dimensional closure is impossible. In that case, any attempt at closure will require an approximation.

Corollary 4.2: [Finite Dimensional Closure] If (Ω, f) is a functional form game, and $\pi(x, t)$ is a density over Ω that evolves according to the continuous replicator Equation (26), with support S , then the replicator equation admits an exact finite-dimensional closure if and only if the integral operator $F[\cdot]$ has finite rank on the domain S .

Theorem 4, and Corollaries 4.1 and 4.2 suggest a canonical choice of basis for a minimal affine parameterization of the continuous replicator equation. Recall that, if $\text{rank}(F)$ is finite given support S , then the performance function f admits an expansion onto a finite linear combination of disc games and each disc game expands into a combination of separable functions of the form $\text{disc}(y^{(k)}(x), y^{(k)}(x'))$ where $\{y^{(k)}\}_{k=1}^{\text{rank}(F)/2}$ are the disc-game embedding functions. The embedding functions are orthogonal with respect to the reference measure used for embedding, so are uncorrelated when applied to draws from the reference measure. Therefore, the disc-game embedding functions are linearly independent. Moreover, since f is a linear combination of separable functions formed as an outer product of the disc-game embedding functions, the range of $F[\cdot]$ must be contained in the span of the embedding functions.

Therefore, the embedding functions form a linearly independent basis for the range of F and the affine parameterization:

$$\pi(x; \theta) = e^{u(x; \theta)}, \quad u(\cdot; \theta) = u_0(x) + \sum_{k=1}^{\text{rank}(F)/2} \theta_1^{(k)} y_1^{(k)}(x) + \theta_2^{(k)} y_2^{(k)}(x), \quad u_0(x) = \log(\pi(x, 0)) \quad (41)$$

is minimal, allows exact closure with respect to the parameters θ , and admits uniquely defined parameter dynamics. We will show that the geometric interpretation of the disc-game will provide useful insight into the parameter dynamics implied by Equation (41).

If, given the support set S , F is not degenerate, then no such finite-dimensional closure exists. Nevertheless, the affine parameterization proposed in Equation (41) still provides optimal approximation to the replicator dynamic for any finite number of parameters r . Truncate the expansion at some even $r < \infty$. This parameterization would provide exact closure for the truncated disc-game approximation to f , $f^{(r)} = \sum_{k=1}^r \text{disc}(y^{(k)}, y^{(k)})$. Recall that any truncated disc-game approximation provides an optimal approximation to f among all rank r functions (see Lemma 3). The error in the approximation to f is exactly equal to the sum of squares of the dropped eigenvalues, so is finite, and arbitrarily small for sufficiently large r . Therefore, while exact, finite-dimensional closure is impossible for all f , exact finite-dimensional closure is possible for an optimal sequence of approximations to f which converge to f in the limit as r diverges.

So, from now on, we will assume that f is either finite rank, or, has been replaced by a sufficiently accurate, finite rank, approximation before applying the replicator dynamic.

4.2.3 Parameterization

Equation (39) selects a parameterization that is both sufficient and minimal. This parameterization required two choices. First, the disc game embeddings depend on an arbitrarily selected reference measure. Second, the parameterization is centered at the initial distribution. Both choices require some discussion since the first is arbitrary, and the second is confounding. The chosen parameterization also maps from a set of bounded functions to an unbounded domain. It is important to distinguish how boundaries of the original set of densities are encoded via diverging sequences of parameters. We will address each point in turn.

To start, let:

$$\mathcal{U}(\nu, u_0) = u_0 + \text{range}(F_\nu) \quad (42)$$

denote the affine subspace of log-densities associated with an initial distribution $\exp(u_0)$ and reference measure ν . Figure 5 illustrates the subspace, and its dependence on u_0 and ν . The choice of u_0 fixes the offset of the affine subspace from the zero function. It centers the parameterization, since, when $\theta = 0$, $u(\cdot; 0) = u_0(\cdot)$. We will show that \mathcal{U} is invariant to changes in the reference measure that preserve its support and to changes in the offset parallel to the range of F_ν . The component of the offset perpendicular to the range of F_ν encodes the conditional distribution over each outcome equivalent class.

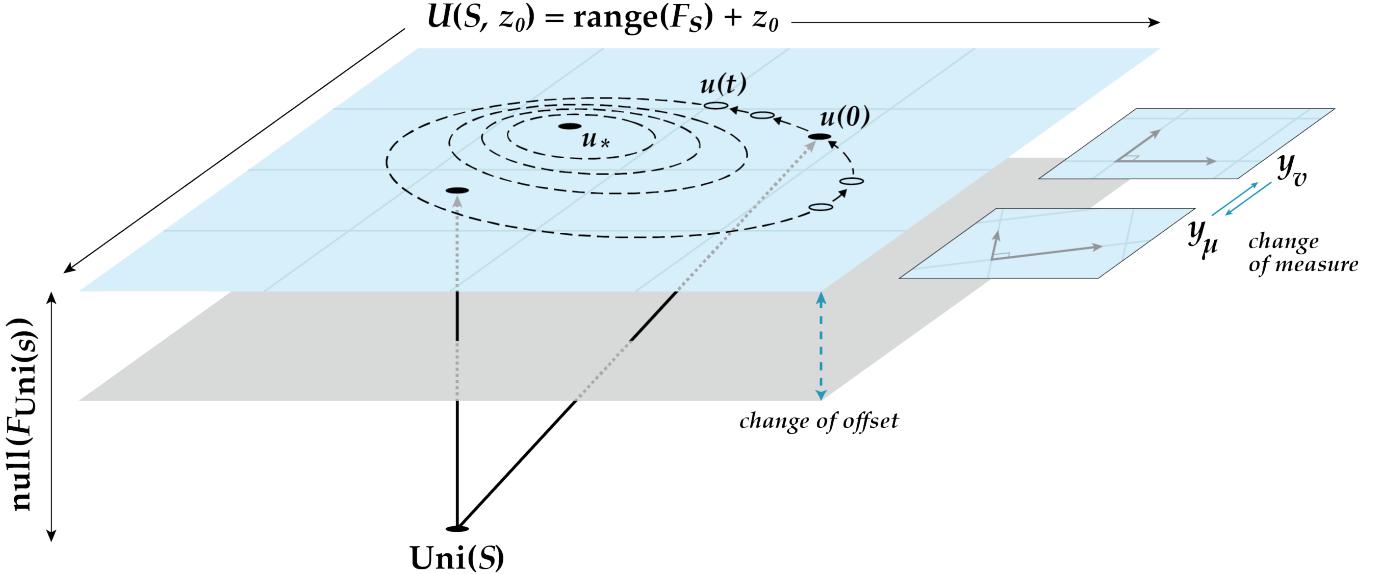


Figure 5: The invariant subspace $\mathcal{U}(\nu, u_0)$. The blue shaded parallelogram represents the subspace. The range of F_S is represented by the pair of vectors spanning the subspace. The null space of F given the uniform measure on S is represented by the vertical direction perpendicular to the subspace. The oval marked $\text{Uni}(S)$ represents the constant function on S . The vector labeled z_0 is the null vector fixing the offset of \mathcal{U} . Changing the offset shifts the subspace along a direction contained in the null space. The vector labeled u_0 points to an initial log-density, $u(0)$. The concentric ovals represent sample orbits of the log density over time, $u(t)$. These remain within the invariant subspace. They orbit an equilibrium u_* corresponding to a Nash equilibrium that is fully mixed over S . The offset fixing the position of the subspace can be specified by u_0 , x_0 , or u_* (when it exists). Distributions with support contained in, but not equal to S correspond to infinite limits where u diverges. The dashed grey lines represent the coordinates imposed by the choice of reference measure ν , and the ensuing embedding functions y_ν . Changing reference measure amounts to changing this basis via an invertible linear transformation, as illustrated in the two small panels to the right of the main panel.

Lemma 9: [Change of Reference Measure] $\mathcal{U}(\nu, u_0)$ is invariant to changes in ν that preserve the support of ν when ν is continuous. Moreover, two affine parameterizations (39) defined with respect to continuous measures ν and μ that share the same support are equivalent up to an invertible linear transformation.

Proof: By Lemma 5, $\text{range}(F_\nu) = \text{range}(F_\mu)$ if ν and μ are both continuous and share the same support set. Therefore, $\mathcal{U}(\nu, u_0) = \mathcal{U}(\mu, u_0)$. Let y_ν and y_μ denote the embedding functions for the reference measures ν and μ . Since both span the same subspace, any linear combination of y_ν can be expressed with a linear combination of y_μ . Since both sets are linearly independent the two affine parameterizations are equivalent up to an invertible linear transformation. \square

Since changing measure amounts to an invertible linear transformation of the parameter space, parameter dynamics associated with two measures that share the same support are topologically equivalent. We will use this fact to adapt the choice of reference measure to best suit the dynamical question of interest.

In contrast, changing u_0 changes the distribution associated with the origin of the parameter space. Changing the origin preserves the affine subspace if the log ratio of the corresponding distributions lies in the range of F .

Lemma 10: [Change of Center] $\mathcal{U}(\nu, u_0) = \mathcal{U}(\nu, v_0)$ if and only if $u_0 - v_0 \in \text{range}(F_\nu)$.

Proof: The subspaces $\mathcal{U}(\nu, u_0)$ and $\mathcal{U}(\nu, v_0)$ are parallel to $\text{range}(F)$. \square

Since the affine subspace $\mathcal{U}(\nu, u_0)$ is invariant to shifts in u_0 parallel to the range of F_ν , the particular affine subspace can be represented by the choice of an element in the nullspace of the adjoint of F_ν . Under the inner product $\langle \cdot, \cdot \rangle_\nu$ the adjoint of F_ν is $-F_\nu$, so the nullspace of the adjoint is the nullspace of F_ν . Thus, it is sufficient to fix $u_0 \setminus \text{range}(F_\nu) = \text{Proj}_{\text{null}(F_\nu)}(u_0)$. In other words, $\mathcal{U}(\nu, u_0) = \mathcal{U}(\nu, v_0)$ if $\text{Proj}_{\text{null}(F_\nu)}(u_0) = \text{Proj}_{\text{null}(F_\nu)}(v_0)$.

Since \mathcal{U} is invariant to changes in measure with fixed support S , and changes in u_0 parallel to the range of F_ν , we can instead write $\mathcal{U}(S, z_0)$ where S is the support of some continuous measure ν , and z_0 is a null-vector for F_ν . Since the particular choice of ν does not matter, we could, for example, adopt the uniform measure over S , $\text{Uni}(S)$. Adopting the uniform measure is suggestive. All steady-state distributions with support S must be contained in $\text{null}(F_{\text{Uni}(S)})$. We will

show that, each $\mathcal{U}(S, z_0)$ either admits a unique steady-state distribution with support S , or does not. Thus, we could index invariant families by the steady-states they admit and the choice of offset amounts to a choice of steady state.

The offset used to fix a particular invariant family is also closely related to the information discarded when reducing by outcome equivalency. Recall that x is outcome equivalent to x' if $f(x, x'') = f(x', x'')$ for almost all $x'' \in \Omega$. We can extend the same definition over any set S ; $x \sim_S x'$ if x and x' return the same payout against almost all $x'' \in S$.

Lemma 11: [Outcome Equivalent Types are Embedded Together] Let $y(\cdot)$ denote a disc-game embedding defined with a reference measure supported on S . Then $y(x) = y(x')$ if and only if $x \sim_S x'$

Proof: Lemma 11 is a special case of Lemma 6 with $\delta = 0$. If $x \sim_S x'$ then $f(x, x'') = f(x', x'')$, so $g(x) = g(x')$ for all functions $g \in \text{range}(F)$. Since the embedding functions span the range of F , each embedding function is contained in the range of F . This establishes the forward direction. For the reverse, note that, $f(x, x') = \text{disc}(y(x), y(x'))$. So, if $y(x) = y(x')$, then $x \sim_S x'$. \square

Lemma 11 shows that disc game embedding automatically reduces by outcome equivalency. All outcome equivalent types are mapped to the same disc game coordinate and each disc game coordinate specifies a unique equivalency class. The disc game coordinates index the equivalency classes. Therefore, given a distribution π_X over traits x , the corresponding distribution π_Y over the embedded traits, $y(X)$ automatically marginalizes over each equivalence class.

We will see that the parameter dynamics depend only on the marginal distribution over each equivalency class, i.e. the distribution of the disc game coordinates $y(X)$. Therefore, the parameter dynamics are invariant to changes in π_X that preserve the marginal distribution over the classes, π_Y . Thus, the parameter dynamics are both independent of the conditional distribution within each equivalency class, and preserve those conditional distributions as the parameters evolve. It follows that, all of the information specifying the conditional distributions over each equivalency class is absorbed in the static component of the affine parameterization, i.e. the offset z_0 .

To make this notion precise, notice that, any two distributions with logarithms in the same affine subspace must have a log-ratio in the range of F . Since the range of F is constant within each equivalency class, they must also share the same conditional distributions over each equivalency class.

Lemma 12: [Invariant Affine Subspaces] Two initial distributions can only have logarithms in the same affine subspace if they share the same support, and conditional distributions in each outcome equivalency class.

So, each invariant family of affine log-densities corresponds to the choice of a support S , and an offset z_0 in the nullspace of F for a reference measure supported on S . The choice of the offset fixes a unique steady state distribution within the class, when it exists. It also fixes the conditional distribution over each outcome equivalency class. These conditionals do not need a dynamical degree of freedom since they have no influence on, and are preserved by, the dynamic.

Finally, consider the nonlinear relationship between the original space of possible densities, and the set of possible parameters. The space of densities with supports contained in S is an infinite-dimensional simplex. The set of densities with support equal to S lies on the interior of that simplex. If the parameters are finite, then the corresponding density is nonzero everywhere in S . Therefore, the set of densities with finite parameters, $\{\pi(x; \theta) \mid \|\theta\| < \infty\}$, is contained in the interior of the set of densities. Accordingly, we will call all finite parameter vectors interior. The boundary of the parameter space corresponds to limiting densities produced by a sequence of diverging parameters. These include densities with smaller support, and densities that diverge on sets of measure zero. This language extends the standard terminology on the simplex (c.f. [7]).

Definition: [Interior] A parameter vector θ is *interior* if $\|\theta\| < \infty$.

4.2.4 Hamiltonian Parameter Dynamics

Suppose that (Ω, f) is a functional form game, that $\pi(x, 0)$ is a density over Ω with support S and that f is finite rank on S . Finally, suppose that $\pi(x, t)$ is subject to the continuous replicator Equation (28). Parameterize $\pi(x, t)$ according to the affine parameterization (41) with basis functions set to the disc-game embedding functions given a reference measure ν supported on S . Then, the parameter dynamics must solve the closure equation (40). Solving the closure equation for $\frac{d}{dt}\theta(t)$ produces a system of ODE's that can be elegantly expressed in terms of either the best response direction to the population centroid in disc-game space, or, the total population size as a function of the parameters.

Theorem 5: [Replicator Parameter Dynamics] Given (Ω, f) , a support set S , a reference measure ν supported on S , and an initial density $\pi(\cdot, 0)$ supported on S , define:

1. The rank $r = \text{rank}(F) < \infty$ given the operator F over support S .
2. The embedding functions $y(x) = [y_1(x), y_2(x), \dots, y_r(x)]$ where $y(\cdot)$ is a column-vector valued function, with pairwise block-indexing $y^{(k)}(x) = [y_1^{(k)}(x), y_2^{(k)}(x)] = [y_{2k-1}(x), y_k(x)]$, and their corresponding parameters $\theta = [\theta_1, \theta_2, \dots, \theta_r]$.
3. The exponential family $\pi(x; \theta) = \exp(u(x; \theta))$, $u(x; \theta) = \log(\pi(x, 0)) + \theta \cdot y(x)$.

Then, the replicator equation admits an exact r -dimensional closure. The corresponding parameter dynamics can be expressed using:

1. The rotation matrices R , the 2×2 , 90° rotation matrix as defined in (6), and U , the $r \times r$, block diagonal with diagonal blocks equal to R as defined in (7).
2. The population centroid in disc-game space:

$$\bar{y}[\pi] = \mathbb{E}_{X \sim \pi}[y(X)] \quad (43)$$

3. The self-play vector field which specifies the optimal training direction in response to the opponent y :

$$v(y) = Uy \quad (44)$$

4. The total population size as a function of the parameters:

$$P(\theta) = \int_{x \in \Omega} \pi(x; \theta) dx. \quad (45)$$

Then, the parameters obey the system of ODE's:

$$\frac{d}{dt} \theta = v(\bar{y}[\pi(\cdot; \theta)]) = U \nabla_\theta P(\theta) \xrightarrow{\text{blockwise, } k \leq r/2} \begin{cases} \frac{d}{dt} \theta_1^{(k)} = \bar{y}_2^{(k)}[\pi(\cdot; \theta)] = \partial_{\theta_2^{(k)}} P(\theta) \\ \frac{d}{dt} \theta_2^{(k)} = -\bar{y}_1^{(k)}[\pi(\cdot; \theta)] = -\partial_{\theta_1^{(k)}} P(\theta) \end{cases} \quad (46)$$

Proof Outline: The first equality in Equation (46) is established by substituting the affine parameterization of u into the replicator equation, and simultaneously expanding f into a linear combination of outer products of the embedding functions. Since each outer product is separable, and expectations are linear, the expectation $\mathbb{E}_{X \sim \pi}[f(x, X)]$ decomposes into the sum $y(x)^\top U \bar{y}[\pi]$. Then, since the embedding functions are independent, the closure Equation (40) is solved by matching the coefficients of the expansion of each side on the basis formed by the embedding functions.

The second form involving the gradient of the total population size with respect to the parameters follows from the observation that the total population size, as a function of the parameters, is the Laplace transform of the initial population distribution in the disc-game space. The Laplace transform of a distribution is its moment-generating function:

$$P(\theta) = \int_{x \in \Omega} \pi(x; \theta) dx = \int_{x \in \Omega} \pi(x, 0) e^{\theta \cdot y(x)} dx = \mathbb{E}_{X \sim \pi_X(\cdot, 0)}[e^{\theta \cdot y(X)}] = \mathbb{E}_{Y \sim \pi_Y(\cdot, 0)}[e^{\theta \cdot Y}]. \quad (47)$$

Derivatives of moment-generating functions recover moments, i.e. expectations, against the original density. Straightforward calculus establishes that:

$$\bar{y}[\pi(\cdot; \theta)] = \mathbb{E}_{X \sim \pi(\cdot; \theta)}[y(X)] = \nabla_\theta P(\theta) \quad (48)$$

thus establishing the equivalence of the two forms in Equation (46). Appendix Section 6.6.1 provides a detailed proof. \square

Equation (46) is disarmingly simple. No matter the specific game or trait space, the parameters θ simply evolve in the best response direction to the centroid \bar{y} . Equivalently, the parameter dynamics are entirely defined by their constant of motion, the scalar-valued function $P(\theta)$ which returns the total population size. In other words, Equation (46) is a normal form dynamic. It reduces the space of possible dynamics by grouping together functional form games, and invariant subspaces, that admit equivalent total population functions P .

The simplicity of (46) highlights characteristics of the replicator dynamic that are shared by all choices of f and Ω , so allows highly generic analysis. It also obscures behaviors that depend on the specific choices of f and Ω . While Equation (46) may appear entirely independent and invariant to game specific details, it is not independent of f or Ω . Rather, it expresses the game specific dependences in the structure of the function $P(\cdot)$, or, in the parametric relation between θ and \bar{y} . This approach is useful as it provides alternative interpretations of the dynamic that highlight characteristic behaviors. For example, we will use Equation (46) to show that $\bar{y}(t)$ always obeys an adaptive dynamics equation, no matter the choice of f and Ω , and without any approximation.

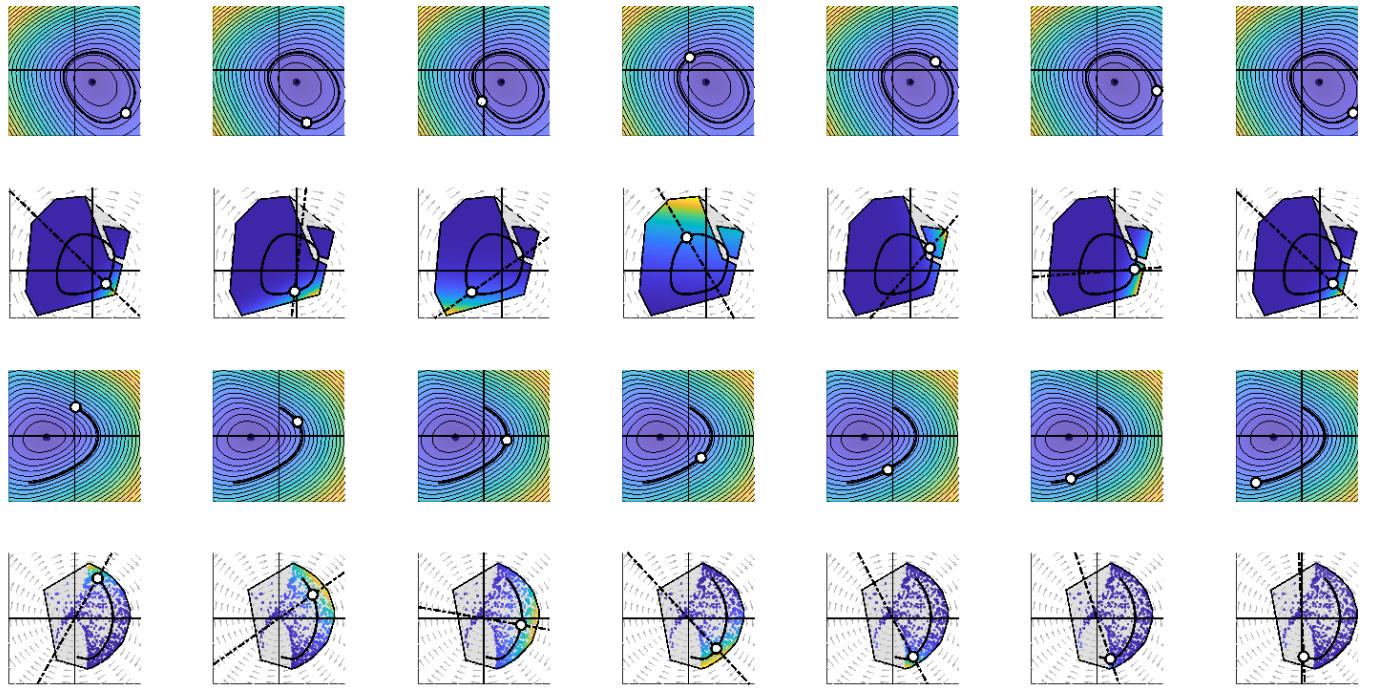


Figure 6: Autonomous centroid and Hamiltonian parameter dynamics for two example games. The top two rows correspond to the example game illustrated in Figures 3, 6, and 8. The bottom two rows correspond to the IPD example illustrated in Figure 4. The first and third rows show orbits in the parameter space. The second and fourth rows show the corresponding sequence of population distributions in the embedding space, formed by exponentially tilting the base measure on Ψ . Each column represents the population at a different, equally spaced time points. The heatmap in the first and third rows show the total population function $P(\theta)$. Solutions to the replicator equation orbit on level sets of the total population. A sample orbit is shown. The white circle represents the parameters at different time points. In the second and fourth rows, the axes correspond to disc game coordinates, the colored region is Ψ , and the grey dashed region is its convex hull. The grey vector field is the optimal local training vector field v . Color represents the density of $y(X)$ given $X \sim \pi(\cdot, t)$. The white circle represents the population centroid. The dashed line passing through the centroid and the origin separates types whose frequency is increasing (on the downstream side of the dashed line), and whose frequency is decreasing (on the upstream side of the dashed line).

To start, we offer some intuition. The first form of the parameter dynamics specified by (46) shows that the parameters, θ always move in the best response direction to the embedded population centroid. In a sense, this equation imitates simultaneous gradient ascent in the disc-game space (see Appendix Section 6.5). In simultaneous gradient ascent, individual agents shift in the best response direction to the centroid [22]. In the replicator dynamic, the *parameters* shift in the best response direction to the centroid.

Like simultaneous gradient ascent, Equation (46) encodes feedbacks that produce oscillation. Increasing $\theta_1^{(k)}$ multiplies the population distribution in disc-game space by the exponential function $\exp(\theta_1^{(k)} y_1^{(k)})$, so biases the distribution to move more mass to larger $y_1^{(k)}$. Types with large $y_1^{(k)}$ are best countered by types with large, negative, $y_2^{(k)}$. So, the best response direction to a type with large $y_1^{(k)}$ points in a direction that rapidly decreases $y_2^{(k)}$. Then, following Equation (46), increasing $\theta_1^{(k)}$ leads to decreasing $\theta_2^{(k)}$. Repeating the same argument shows that decreasing $\theta_2^{(k)}$ decreases $\theta_1^{(k)}$. This establishes a self-regulating feedback loop. Increasing $\theta_1^{(k)}$ decreases $\theta_2^{(k)}$, which decreases $\theta_1^{(k)}$, which increases $\theta_2^{(k)}$, and so on. The result is a nonlinear oscillator equation, with restoring feedbacks mediated by the relationship between the parameters θ , ensuing distribution over disc-game space, $\pi_Y(\cdot; \theta)$, and resulting centroid $\bar{y}[\pi(\cdot; \theta)] = \mathbb{E}_{Y \sim \pi_Y(\cdot; \theta)}[Y]$.

A word on coupling. The parameter dynamics nearly decouple by block, since the update to $\theta^{(k)}$ only depends on the k^{th} pair of coordinates \bar{y} . The blocks don't fully decouple since $\bar{y}^{(k)}$ is a function of the distribution $\pi(x; \theta)$, which depends on all θ for each x . When the reference distribution used in the disc-game embedding is chosen appropriately, then linearizing Equation (46) about the reference distribution decouples the blocks. In this case, the cross-block coupling arises in third and higher-order terms of the Taylor expansion of the exponential function used to convert from the log-density u back to the density π .

Oscillators are frequently Hamiltonian. Simultaneous gradient ascent in response to a zero-sum normal form game on finite strategy spaces is always a Hamiltonian dynamic [22], as is the replicator equation for generic bimatrix games on

finite strategy spaces [7, 102]. The second equality in Equation (46) establishes that the replicator parameter dynamics are also Hamiltonian.

Hamiltonian dynamics were introduced to study mechanical systems, where the state of a system can be specified by a list of position and momenta coordinates, x, p , whose dynamics conserve a Hamiltonian function $H(x, p)$ (c.f. [102, 175]). At its simplest, a Hamiltonian dynamic sets $\frac{d}{dt}x = \partial_p H(x, p)$ and $\frac{d}{dt}p = -\partial_x H(x, p)$ [89, 135]. Equation (46) assumes this form with Hamiltonian equal to the total population size, P , “position” variables, $\theta_1^{(k)}$, and “momentum” variables, $\theta_2^{(k)}$. In particular, Equation (46) is block Hamiltonian, with separate blocks corresponding to separate disc-games.

The fact that the parameter dynamics use the total population, $P(\cdot)$, as their Hamiltonian distinguishes our analysis from standard analyses. For example, Akin, Losert, and Hofbauer demonstrated that the replicator dynamic is Hamiltonian given a zero-sum normal form game with finitely many strategies, with Hamiltonian equal to the Kullback-Liebler divergence from an interior equilibrium [7, 102]. See [32] for an extension to polymatrix games and [20] for an extension to generic follow-the-regularized-leader dynamics. While more complex constants of motion are a powerful analytic tools (c.f. [144]), the simplicity of $P(\cdot)$ is appealing. The total population size has immediate biological relevance, especially when payout is interpreted as per-capita growth rate, and is the most obvious quantity conserved by the replicator dynamic. It is surprising that, the total population is, as a function of the parameters, not only conserved, but also encodes all the game information needed to define the parameter dynamics. The total population size is also analytically convenient since, like the log partition function of any exponential family, it carries rich information about the distribution in its derivatives. Figure 7 presents four sample Hamiltonians.

4.2.5 Dynamical Characterization

Here we show that the Hamiltonian form for the parameter dynamics admits strong global analysis. Each fact can be easily established using the disc game representation. For all proofs, see Appendix Section 6.6.2.

Lemma 13: [Properties of the Hamiltonian] The Hamiltonian function $P(\theta)$ defined in Equation (45) satisfies the following properties:

1. The mixed partial derivatives of the Hamiltonian correspond to raw moments of the embedded population distribution. Let $p(\cdot; \theta) = \pi(\cdot; \theta)/P(\theta)$. Let $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_r] \in \mathbb{Z}^r$ be a multi-index. Let $|\alpha| = \sum_j \alpha_j$ and $x^\alpha = \prod_j x_j^{\alpha_j}$. Then:

$$\partial_\theta^\alpha P(\theta) = P(\theta) \mathbb{E}_{X \sim p(\cdot; \theta)}[y(X)^\alpha] = P(\theta) \mathbb{E}_{Y \sim p_Y(\cdot; \theta)}[Y^\alpha]. \quad (49)$$

In particular, the Hessian of $P(\theta)$ equals both the Gram matrix for the embedding functions $y(\cdot)$ with respect to the inner product $\langle \cdot, \cdot \rangle_{\pi(\cdot; \theta)}$, and, equivalently, the covariance in the embedding of a randomly selected agent:

$$[\partial_\theta^2 P(\theta)]_{ij} = \langle y_i(\cdot), y_j(\cdot) \rangle_{\pi(\cdot; \theta)} = P(\theta) (\text{Cov}_{Y \sim p_Y(\cdot; \theta)}[Y_i, Y_j] + \mathbb{E}_{Y \sim p_Y(\cdot; \theta)}(\theta)[Y_i] \mathbb{E}_{Y \sim p_Y(\cdot; \theta)}[Y_j]). \quad (50)$$

If we adopt a constant interaction rate rather than a linear interaction rate model, then the Hamiltonian is $\log(P(\theta))$ and its partials are the centered moments of $Y \sim p_Y$ without any extra scaling against the population size.

2. The Hamiltonian is strictly convex.
3. The Hamiltonian either admits a unique global minimizer at some finite θ_* , or no such minimizer exists. If such a minimizer exists, then $P(\theta)$ is radially unbounded.

Proof Outline: The relation between partial derivatives of the Hamiltonian and the moments follows by directly differentiating Equation (45) with respect to θ . The Hamiltonian is convex since its Hessian at θ is a covariance matrix. All covariance matrices are positive semi-definite, so the Hessian is positive definite everywhere. Strict convexity follows from the orthogonality of the embedding functions under an appropriate change of reference measure ν . \square

Remark: The convexity of the $P(\theta)$ is useful since Hamiltonian dynamics mix on level sets of their Hamiltonian. Since the level sets of a convex function are boundaries of convex regions, solutions to the parameter dynamics mix on the boundaries of convex regions. If a level set is bounded, then the parameters remain bounded for all time. If a level set is unbounded, then the parameters may escape to infinity (as observed in the transitive and quadratic special cases).

When $P(\theta)$ admits a unique global minimizer, it is radially unbounded, so its level sets are bounded. Then, the parameters mix on the boundaries of compact sets. These observations will be used to show that all interior equilibria for the parameter dynamics are isolated, Lyapunov stable centers, so the parameters behave as coupled oscillators. More precisely, Equation (49) will be used to show that, up to linearization, parameters near an interior equilibria satisfy a harmonic oscillator equation while higher-order terms couple the oscillators. \square

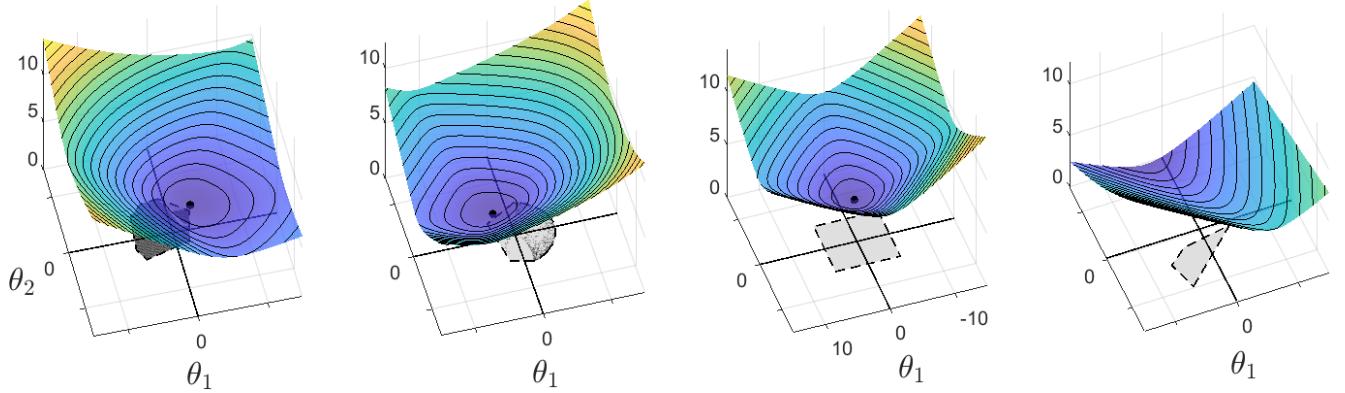


Figure 7: Four sample Hamiltonians, $P(\theta)$. The convex hull of the range of the embedding map, Ψ , is shown in light grey with dashed boundaries. The base measure ν is represented by dark grey scatter points in the first two panels. The last two use a uniform measure. The vertical axis represents the Hamiltonian on a log base 10 scale. Notice that the Hamiltonians are all convex functions. The Hamiltonians grow slowest in the direction along which the boundary of Ψ is closest to the origin, and grow exponentially (linearly in the log scale) away from their minima. The first example matches the examples provided in Figures 3, 6, 8, and 9. The second matches the IPD embedding illustrated in Figure 4. The third example shows that the Hamiltonian develops approximately polygonal level sets away from its minimizer when Ψ is polygonal. The fourth example uses a Ψ whose convex hull does not enclose the origin. In this case, the Hamiltonian does not admit an interior minimizer, so sample trajectories may escape to infinity (see Figure 8). In this case there exist dominated types that will be driven to extinction.

Corollary 5.1: [Dynamics] The parameters $\theta(t)$ satisfy the following dynamical (non-steady state) properties:

1. The total population size $P(\theta(t))$ is conserved, $P(\theta(t)) = P(\theta(0)) = P$.
2. If the interaction rate is time-autonomous, $\rho(P, t) = \rho(P)$, then changes of interaction rate can be encoded via a transformation of the Hamiltonian. Let $R(P) = \int \rho(P)/P dP$. Then, for any time-autonomous interaction rate:

$$\frac{d}{dt}\theta(t) = U\nabla_\theta R(P(\theta)). \quad (51)$$

In particular, if the interaction rate is constant, then $\frac{d}{dt}\theta(t) = \rho U\nabla_\theta \log(P(\theta))$.

3. The parameter dynamics are incompressible, i.e. volume-preserving. Let $\Theta(0)$ denote a closed set of initial parameters θ , and let $V(0)$ equal the volume enclosed by $\Theta(0)$. Let $\Theta(t)$ denote the set of parameters at time t after applying the replicator dynamic. Let $V(t)$ represent volume enclosed by $\Theta(t)$. Then, $V(t) = V(0)$.
4. If $P(\cdot)$ admits a unique global minimizer, and $\theta(0)$ is finite, then $\theta(t)$ remains finite for all times $t \geq 0$. That is, the interior is invariant under the replicator dynamic.
5. If $P(\cdot)$ admits a unique global minimizer, then the parameter dynamics are Poincaré recurrent, and, if $r \leq 2$, then the dynamics are periodic.

Proof Outline: All Hamiltonian dynamics preserve their Hamiltonian. In this case, since the Hamiltonian is the total population size, the total population size is conserved. The change of interaction rate formula is established via the chain rule. Volume preservation is established by showing that the divergence of the right-hand-side of Equation (137) is zero at all θ . The interior is time-invariant under the replicator dynamic when $P(\theta)$ admits a global minimizer since, when $P(\theta)$ admits a global minimizer it is radially unbounded, so has bounded level sets. If the level sets are bounded, then trajectories are bounded and volume preserving, so Poincaré recurrent [24, 180]. If $r = 2$, then, the dynamics are periodic since the level sets of the Hamiltonian are closed, nested, orbits that fill a plane. \square

Remark: Fact 3 is a statement of Liouville's theorem for Hamiltonian systems [129]. A similar result is established in [72] for finite strategy sets (discrete trait space). The converse - the replicator is volume preserving if and only if the game is zero-sum - is established in [32]. For related discussions on constants of motion and volume preservation in evolutionary game theory see [103, 125, 228].

Volume preservation is interesting since it bans the existence of any bounded attractor [32]. A bounded attractor is a

set of parameters that can be enclosed in a bounded basin of attraction. A basin of attraction is an open set containing the attractor such that any trajectory initialized in the basin converges onto the attractor. Since the dynamic is volume preserving, no such basin can exist. Otherwise, any set that covers the attractor that is contained inside the basin would need to preserve volume while contracting.

It follows that there are no isolated, attracting equilibria of the parameter dynamics with finite parameter values. Running the same argument in reverse time implies that there are also no isolated, interior, unstable equilibria. In other words, there are no isolated sources or sinks with finite parameter values.

Thus, all isolated attractors in the space of densities, in both forward and reverse time, must exist as limiting densities that can be approached via a diverging sequence of parameters. The closure of the set of densities with finite parameters includes both densities with smaller support than the initial support S , and singular densities such as delta distributions (monomorphic populations). Trajectories that diverge along an unbounded level set can converge to attracting singular densities. For example, if f is transitive with a rating function r that admits a unique global maximizer, then the delta distribution at that maximizer is globally attracting. For the stability analysis of monomorphic distributions, see [54, 57, 164, 165]. More generally, diverging parameter trajectories can represent a process of iterated dominance, in which dominated types are sequentially eliminated from the support [78]. \square

Remark: Facts 3 and 4 also establish Fact 5; when the Hamiltonian admits a global minimizer, the corresponding parameter dynamics are Poincaré recurrent. A dynamical system is Poincaré recurrent if solution trajectories return arbitrarily close to almost all initial condition infinitely often. Equivalently, a solution trajectory initialized in any open set will return to that set (recur) in finite time [180, 41]. Thus, Poincaré recurrent systems exhibit irregular oscillations. If $r = 2$ then those oscillations are periodic [32]. If $r > 2$, then the recurrent cycles may mix chaotically [144, 190]. \square

Let's return, for a moment, to paradigm comparison. At the outset, we argued that the disc-game would provide a more accurate intuition for learning in zero-sum games than either optimization, or equilibria. Fact 5 offers strong support for this argument.

First, it ensures that, unless initialized at a steady state, almost all initial conditions lead to solutions that wander forever without converging. Second, it ensures that motion along solution trajectories fail to optimize any continuous function, indeed, fail basic notions of directed adaptation.

Often learning or selective dynamics are attributed direction by identifying some property of the population that changes consistently over time. For example, one might argue that adapting populations will exhibit increasing specialization, complexity, or rational behavior (c.f. [137]). To formalize this idea, let h represent a continuous functional that accepts distributions and returns a real number. If $h[\pi(\cdot, t)]$ is either monotonically increasing or decreasing in time, then the dynamics are plainly directed by h . More weakly, h can lend a sense of directed adaptation if $\inf_{s>t}\{h[\pi(\cdot, s)]\}$ is increasing, or $\sup_{s>t}\{h[\pi(\cdot, s)]\}$ is decreasing, along some sequence of times t . In the first case, $h(t) = h[\pi(\cdot, t)]$ does not increase monotonically, but does increase in the sense that, there exists a sequence of times, $\{t_j\}_{j=1}^\infty$ chosen so that $h[\pi(\cdot, t_j)] \leq h[\pi(\cdot, s)]$ for all $s > t_j$, and such that $h[\pi(\cdot, t_j)]$ is monotonically increasing. In the second case the same statements hold, only for a sequence of decreasing upper bounds.

When the dynamics are Poincaré recurrent, no such continuous function can exist, since the distribution will return arbitrarily close to its initial condition in finite time. Thus, $h[\pi(\cdot, t)]$ will return arbitrarily close to $h[\pi(\cdot, 0)]$ infinitely often. It follows that, when an interior equilibrium exists, the replicator dynamics do not optimize *any* continuous function of the population distribution.

Lemma 14: [No Objective] If $\pi(\cdot, t)$ is Poincaré recurrent, then there is no continuous, non-constant functional h such that:

1. $h(t) = h[\pi(\cdot, t)]$ is monotonically increasing or decreasing,
2. $\inf_{s>t}\{h(s)\}$ is monotonically increasing, or $\sup_{s>t}\{h(s)\}$ is monotonically decreasing.

Remark: The restriction to continuous functionals in Lemma 14 is necessary if the dynamics are recurrent but aperiodic. For example, if we do not require continuity, then we could define a monotonically increasing functional by choosing a fixed initial distribution for every solution path, then setting $h[\pi]$ equal to the time it would require for a distribution initialized at the corresponding π_0 to reach π . Nevertheless, the restriction to continuous functions does not discard much, since any function that is monotonic along recurrent trajectories must span its full range of possible values on every open set containing any point on the trajectory. It is hard to imagine such a function offering a substantive notion of direction. To adopt such a function would assign meaning to arbitrarily small perturbations of the distribution that preserve its support. Such sensitivity should disqualify any associated interpretation. \square

These conclusions are, perhaps, surprising, given that the replicator dynamic can be derived from a multiplicative weight updating scheme, so corresponds to a continuous implementation of the canonical no-regret learning process [125]. However, the fact that the replicator equation fails to converge to all interior equilibria is actually a generic characteristic

of no-regret learning dynamics shared by all follow-the-regularized-leader dynamics [72]. For similar proofs of recurrence in zero-sum games with fully mixed equilibria, and a more general proof for incompressible games, see [32, 144, 178].

The fact that no-regret dynamics circulate about fully mixed equilibria in zero-sum games emphasizes a subtler distinction between the optimization paradigm and learning in multi-agent settings. In optimization problems, step-wise guarantees ensure that the optimizer moves towards its goal. Sufficiently strong step-wise guarantees ensure pointwise guarantees for the limit of the path. These relations are more than ancillary features; they are often the reason optimization problems can be solved iteratively.

The replicator dynamic satisfies both a step-wise and a path-wise rationality guarantee. It is “step-wise” rational since the distribution changes in the self-play direction subject to a Shashahani, or Fisher information, metric [145]. In this sense, it always adapts optimally in response to the current population. The replicator dynamic is also path-wise rational since it is a no-regret dynamic [72, 125]. No-regret is a path-wise guarantee. It ensures that a trajectory is, in the long run, rational, since the learner does not, in hindsight, regret following it. If the trajectory converges to a fixed distribution, then that distribution must satisfy a static, pointwise, rationality criteria. Namely, there cannot exist an alternate distribution with positive expected payout against the fixed distribution. Otherwise, the learner would regret settling there. Thus, convergence converts a path-wise guarantee into a pointwise guarantee at the end of the path. In contrast, when a no-regret learner fails to converge, then the guarantee remains path-wise. There is no alternate fixed distribution that would have a positive expected payout against a distribution selected at random from the *entire* trajectory.

In this case, step-wise, and path-wise, rationality, offer no guarantee that any distribution reached by the process is itself rational in a static, pointwise sense. Nor is any distribution reached more rational than those visited before. Rather, when there exists an interior equilibria, no type is dominated, so every type can be justified as an exploiter of some other type [5]. More strongly, every distribution can be justified by the path it lies on. It is an optimal response to some other distribution (its immediate antecedent) in a set of distributions that is collectively not exploitable. Yet, it is not, individually, better justified than any other distribution on the path. Learning does not order specific distributions as more or less rational since the trajectory recurs. The dynamic will return arbitrarily close to every distribution it visits infinitely often, so every distribution visited can be viewed as both a predecessor, and a successor, to every other distribution visited. There exist rational learning trajectories moving in either direction between any pair of distributions sharing a path. So, rather than progressing towards an increasingly optimal distribution, the dynamic orbits ambivalently among a set of distributions that are collectively unexploitable. The process could be said to be constantly learning but to never learn, since the sequence of distributions visited do not acquire any property that distinguishes progress along the trajectory. The resulting solution concept, then, is the entire orbit [20]. Since almost all initial conditions produce a recurrent orbit, the resulting solution concept should consist of the foliation of the space, that is, the collection of all orbits, not a collection of equilibria. This is categorically *not* the intuition conveyed by optimization. It is the intuition suggested by the disc-game, whose essence suggests circulation about a center.

These observations also impact the classification and interpretation of equilibria. They are not characteristic solutions since they are neither attracting nor optimal. Equilibria are special only in that they are balance points where the dynamic rests. That said, the equilibria still provide a useful window into the shape of orbits, since each equilibria offers the simplest possible orbit - a point - and, linearization about the equilibria characterizes orbits near that point. Using Lemmas 12 and 13, we will show that every interior equilibria corresponds to a center, all the other solution trajectories orbit that center on compact, concentric shells, and, near the equilibria, the replicator dynamic behaves like a system of independent simple harmonic oscillators.

Corollary 5.2: [Steady-States] Let π_* denote a steady-state distribution of the continuous replicator dynamic and let θ_* denote an equilibrium of the corresponding parameter dynamics. A distribution $\pi(\cdot; \theta)$ with finite θ is a steady-state π_* if and only if θ is an equilibrium θ_* . The equilibria obey the following properties:

1. All interior equilibria θ_* are critical points of the Hamiltonian, $\nabla_\theta P(\theta)|_{\theta=\theta_*} = 0$.
2. The parameter dynamics either admit no interior equilibrium, or admits a unique, isolated, interior equilibrium corresponding to the global minimizer of the Hamiltonian.
3. If an interior equilibrium exists, then it is a globally Lyapunov stable center for the parameter dynamics.
4. Up to linearization, the parameter dynamics about any interior equilibrium are topologically equivalent to a collection of $r/2$ independent harmonic oscillators with frequencies $\{P\omega_k\}_{k=1}^{r/2}$ where ω_k equals the magnitude of the $2k-1^{\text{st}}$ eigenvalue of F_{p_*} , for $p_* = \pi(\cdot; \theta_*)/P(\theta_*)$.

Proof Outline: The embedding functions are linearly independent so $\frac{d}{dt}\pi(x, t) = 0$ if and only if $\frac{d}{dt}\theta(t) = 0$, and $\frac{d}{dt}\theta(t) = 0$ if and only if $\nabla_\theta P(\theta) = 0$. Since the Hamiltonian is strictly convex, all critical points correspond to minimizers, and the Hamiltonian either admits a unique global minimizer or no critical point exists. If an interior equilibrium exists,

then $P(\theta)$ admits a finite global minimizer so is radially unbounded. So, any perturbation off the equilibrium produces a trajectory that mixes on a the boundary of a compact neighborhood containing the equilibrium. Thus, the equilibrium is neutrally stable, i.e. is a center for the dynamic. Linearization is discussed in a subsequent remark. \square

Remark: When $r = 2$, minima of $P(\theta)$ correspond to centers, since the level sets of $P(\theta)$ are nested circuits in \mathbb{R}^2 enclosing the equilibrium. Rotating the gradient vector field about the minimizer by ninety-degrees produces a circulating vector field that circulates in the direction of the applied rotation around these loops. \square

Remark: For general r , we linearize to investigate the parameter dynamics near an equilibrium. Suppose that θ_* is a interior equilibrium for the parameter dynamics. The Jacobian of the right-hand side of Equation (137) at θ_* equals $P(\theta_*)$ times the row-rotated Hessian of the total population function, or, equivalently, the row-rotated covariance in $y(X)$ when $X \sim p_* = p(\cdot; \theta_*)$:

$$J(\theta) = P(\theta_*) U \text{Cov}_{Y \sim p_Y(\cdot; \theta)}[Y]. \quad (52)$$

Therefore, the linearized dynamic is:

$$\frac{d}{dt}(\theta(t) - \theta_*) = J(\theta_*)(\theta(t) - \theta_*) + \mathcal{O}(\|\theta(t) - \theta_*\|^2) \simeq P(\theta_*) U \text{Cov}_{Y \sim p_Y(\cdot; \theta_*)}[Y](\theta(t) - \theta_*). \quad (53)$$

By a standard result, the matrix J is similar to a skew-symmetric matrix, so has purely imaginary eigenvalues [10, 102]. Here, we can adopt a more immediately contextual argument. The linearized dynamic (53) is equivalent, under a change of parameterization, to a decoupled system of simple harmonic oscillator equations with oscillation frequency determined by the eigenvalues of the operator $F_{p_*}[\cdot]$.

Recall that the set of embedding functions used as a function basis for the range of F depends on a choice of reference measure. As long as two reference measures share the same support, then the range of F is fixed, so changing the reference measure changes the particular basis used to span the range of F , but does not change the range. Accordingly, any pair of affine parameterizations arising from different choices of reference measure correspond to a different choice of basis. Then, the parameters associated with two reference measures that share the same support, are related by an invertible linear mapping (see Section 4.2.3). It follows that, parameter dynamics associated with two different choices of reference measure are topologically equivalent provided the reference measures have the same support.

If π_* is an interior steady-state, then it has support S , as does the reference measure ν . Since parameter dynamics whose reference measures share their support are topologically equivalent we are free to change the reference measure. Set $\nu = \pi_*$. Then, applying Equation (50), and recalling that the eigenfunctions y are orthogonal with respect to the reference measure yields:

$$J(\theta_*) = UD_\omega = P(\theta_*)W \quad (54)$$

where W is the matrix used in the real Schur form of F_{p_*} (see Equation (6)), and $\omega_k = \|\lambda_{2k}(F_{p_*})\|$ equal the magnitudes of the eigenvalues of F_{p_*} .

Therefore, when $\nu = \pi_*$, the linearized parameter dynamics decouple blockwise, and each block satisfies a harmonic oscillator equation with frequency ω . Let $\Delta\theta = \theta - \theta_*$. Rescaling time by $P/P(\theta_*)$ to recover the correct interaction rate for initial population P yields:

$$\frac{d}{dt}\Delta\theta_1^{(k)}(t) \simeq P\omega_k\Delta\theta_2^{(k)}(t), \quad \frac{d}{dt}\Delta\theta_2^{(k)}(t) \simeq -P\omega_k\Delta\theta_1^{(k)}(t) \quad (55)$$

Equation (55) makes the analogy between the replicator dynamic and mechanical oscillators concrete. Up to linearization, the parameter dynamics near any interior equilibrium behave as decoupled harmonic oscillators, with one oscillator per pair of embedding coordinates. Collectively, solutions to the linearized equations orbit on $r/2$ -dimensional tori. The solutions are periodic if the eigenvalues ω_k are all integer multiples of the smallest eigenvalue. Otherwise, generic solutions are quasi-periodic. \square

Remark: Coupling between the oscillators enters via higher order derivatives. These correspond to higher order moments of the embedded population distribution (see Equation (49)). Therefore, the oscillators are coupled by the cross-block third and higher order moments of the distribution $\pi_Y(\cdot; \theta_*)$. When these moments vanish, the oscillators decouple to the corresponding order. For example, if at π_* , $Y^{(k)}$ and $Y^{(k')}$ are independent for all $k \neq k'$, then the blocks evolve independently. The same holds in reverse since, if the coordinates are not dependent, then there exists a nonzero moment coupling the parameter dynamics. So, given an interior equilibrium θ_* , $\theta^{(k)}(t)$ decouple if and only if, when setting $\nu = \pi_*$, the embedding coordinates corresponding to distinct blocks are independent when sampling from π_* . Otherwise, nonzero coupling produces rich dynamics, including chaos [190].

All of the eigenvalues of $J(\theta_*)$ are purely imaginary, so θ_* is not a hyperbolic equilibrium. Then, the Hartman-Grobman theorem [96, 97] does not apply, so solutions to the exact dynamic need not be topologically equivalent to the linearized dynamic near interior equilibria. In principle, even small higher-order coupling terms could change the dynamics arbitrarily close to an interior equilibrium. The Lyapunov-center theorem [133] ensures that the linearization is characteristic. Namely, there exist $r/2$, two-dimensional manifolds, each composed of periodic orbits, intersecting

the equilibrium. The k^{th} manifold is tangent to $2k - 1^{st}$ and $2k^{th}$ unit directions, and contains orbits with period approximately equal to $2\pi/\omega_k$ [102]. More strongly still, the Kolmogorov-Arnold-Moser theorem ensures that most of the toroidal solutions of the linearized dynamic survive nonlinear coupling up to diffeomorphism, provided the linearized trajectory is sufficiently far from periodic, and the system is initialized close to the equilibrium [11, 102, 154]. \square

We have established that the parameter dynamics obey a clean dichotomy. Given f and support S , the Hamiltonian either admits a unique global minimizer or does not. If such a minimizer exists, it corresponds to an isolated interior equilibrium and a steady-state supported on all of S . The equilibrium is a globally Lyapunov stable center about which all perturbations mix on compact shells and is locally approximated by a system of decoupled simple harmonic oscillators. Higher-order terms couple the oscillators without eliminating the oscillating character of solutions, which return arbitrarily close to all initial conditions infinitely often. Figures 6, 8, and 9 illustrate a recurrent examples. If there is no interior minimizer then all steady-state distributions correspond to parameter sequences that diverge to a boundary [32, 144]. Given a finite trait space, these are the only generic options. Either, there exists a fully mixed equilibria, which the replicator dynamics orbit but fail to approach [32], or there is no fully mixed equilibria, in which case there exists at least some type that is dominated by all other types and that is driven to extinction under the replicator dynamic [5, 7].

There is a straightforward geometric test that distinguishes these two cases after embedding. Recall that the origin plays a special role in a disc-game since it corresponds to a neutral type that has no advantage or disadvantage over any other type (see Box 1). When included in the convex hull of Ψ , it corresponds to a fully mixed evolutionary stable strategy (ESS), since any ESS in a symmetric two-player game must receive the same payout against all opponents, including itself [30, 201], against which it must receive a payout of zero. Recall also that any population embedded in a single disc game competes transitively if the convex hull of the embedded population does not contain the origin. Here, we show that the same test can be extended to arbitrary r , and distinguishes the form of the parameter dynamics.

Corollary 5.3: [Interior Equilibria, Invariant Interior, and the Range of the Embedding] Suppose that (Ω, f) is a functional form game with skew-symmetric f , S is a support set, and $\pi(x, t)$ is a population distribution supported on S evolving according to the replicator equation, (28). Let ν denote an arbitrary reference measure that has support equal to S . Let $\Psi = \{y_\nu(x) | x \in S\} \subseteq \mathbb{R}^r$ denote the image of the set S under the disc-game embedding. Suppose that Ψ is bounded.

Then, the replicator parameter dynamics admit an interior equilibrium if and only if the convex hull of Ψ contains the origin in its interior.

Proof Outline: First, notice that θ_* is an equilibrium if and only if $v(\bar{y}[\pi(\cdot; \theta_*)]) = U\bar{y}[\pi(\cdot; \theta_*)] = 0$. The rotation matrix U is invertible, so θ_* is an equilibrium if and only if the corresponding centroid, $\bar{y}_* = \bar{y}[\pi(\cdot; \theta_*)] = 0$.

Necessity is straightforward. If Ψ does not contain the origin inside its convex hull, then, there is no distribution on Ψ with centroid at the origin. To show sufficiency, we demonstrate that $P(\theta)$ is radially unbounded when the origin is in the interior. For details, see Appendix Section 6.6.3. \square

Figure 8 illustrates the geometric test. The geometric test formalizes the intuition that, a population evolving in response to any combination of disc games, will orbit recurrently unless it is restricted to a subset of disc game space that excludes competitive cycles.

Importantly, the test is invariant to changes in the reference measure ν provided those changes preserve the support S . Changes of reference measure that preserve the support transform Ψ by an invertible linear mapping (see Lemma 5 and 8). Any linear transformation of the origin is the origin. Open neighborhoods remain open under invertible linear transformations. Therefore, the origin is either inside, or outside, the interior of the convex hull of Ψ_ν for all ν with support S .

To conclude our analysis, we show that the centroid \bar{y} in disc-game space obeys an adaptive dynamics equation:

Corollary 5.4: [Adaptive Dynamics] Suppose that (Ω, f) is a functional form game with skew-symmetric f , S is a support set, and $\pi(x, t)$ is a population distribution supported on S evolving according to the replicator equation, (28). Then, for any reference measure ν supported on S , the trait centroid of the embedded population, $\bar{y}(t) = \mathbb{E}_{X \sim \pi(\cdot, t)}[y(X)]$, obeys the adaptive dynamics equation:

$$\frac{d}{dt}\bar{y}(t) = \text{Cov}_{Y \sim \pi_Y(\cdot, t)}[Y]v(\bar{y}(t)) \quad (56)$$

where $v(y) = Wy$ is the self-play vector field for the disc-game.

Proof: Proceed directly.

$$\begin{aligned} \frac{d}{dt}\bar{y}(t) &= \frac{d}{dt} \int_{x \in \Omega} y(x) \exp(\theta(t) \cdot y(x)) \pi(x, 0) dx = \int_{x \in \Omega} y(x) \left(y(x) \cdot \frac{d}{dt} \theta(t) \right) \exp(\theta(t) \cdot y(x)) \pi(x, 0) dx \\ &= \left[\int_{x \in \Omega} y(x) y^\top(x) \pi(x; \theta(t)) dx \right] \frac{d}{dt} \theta(t) = \text{Cov}_{Y \sim \pi_Y(\cdot, t)}[Y]v(\bar{y}(t)) \end{aligned}$$

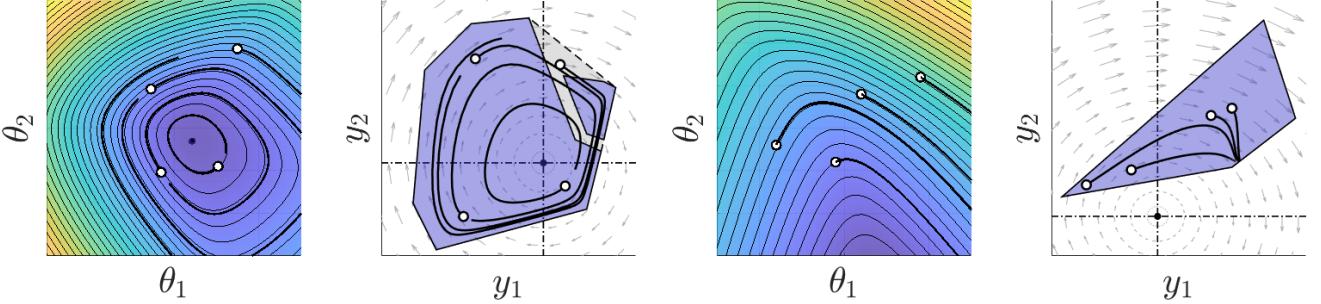


Figure 8: Hamiltonian parameter dynamics, and the associated centroid dynamics, for two example games. The examples match the first and last Hamiltonians shown in Figure 7. The odd panels illustrate the parameter dynamics. The coordinates, θ , correspond to the parameters responsible for fixing the distribution over Ψ . The base measure is uniform on ψ . The four trajectories shown are each initialized at the white scatter points lying at the start of each curve. The even panels illustrate the centroid dynamics. The coordinates, y , represent a pair of disc game coordinates. The blue shaded region Ψ represents the range of the embedding. The grey shaded region enclosing it is its convex hull. The grey circulating vector field represents the optimal local training vector field v . In the first example, the parameters orbit an equilibrium, θ_* , along the level sets of the Hamiltonian function $P(\theta)$. This equilibrium corresponds to a fully mixed Nash equilibrium, and it exists since the convex hull of ψ includes the origin (see Corollary 5.3). In the second example the parameters follow level sets of the Hamiltonian to infinity. As the parameters diverge, the centroid approaches the boundary of Ψ . In the first case, the convex hull of ψ contains the origin. In the second it does not.

where the last equality follows from the parameter dynamics and the fact that $v(\bar{y})$ is orthogonal to \bar{y} . \square

Equation (56) establishes that the centroid of the embedded population obeys a non-autonomous adaptive dynamics equation. Notice that this adaptive dynamics equation does not require any approximation, time-scale separation, does not restrict f or π to limited parametric forms, and does not require a monomorphic limit. In contrast [162], it is both exact and general.

The equation appears non-autonomous since the positive definite matrix $\text{Cov}_{Y \sim \pi_Y(\cdot, t)}[Y]$ depends on the embedded distribution, which is changing in time, as specified by the parameters $\theta(t)$. Thus, as written, it is not clear whether Equation (56) can be solved without first solving for the parameter dynamics. Equation (56) is, in fact, autonomous, since the embedded centroid \bar{y} uniquely determines the parameters θ [69].

Lemma 15: [Parameters From Centroid] Given an affine parameterization (39) the map from θ to \bar{y} is invertible.

Proof: By definition, $\bar{y} = \mathbb{E}_{X \sim \pi_X(\cdot; \theta)}[Y(X)] = \mathbb{E}_{Y \sim \pi_Y(\cdot; \theta)}[Y]$. Therefore, the parameters θ uniquely determine a centroid \bar{y} . To show that the mapping, $\bar{y}(\theta)$ is invertible, recall that $\bar{y}(\theta) = \nabla_\theta P(\theta)$ (see Equation (48) in Theorem 5). Lemma 13 establishes that $P(\theta)$ is strictly convex. For any strictly convex function, the mapping from argument to gradient is invertible, so, \bar{y} uniquely determines the gradient $\nabla_\theta P(\theta)$, and the gradient uniquely determines the parameters θ . \square

Remark: Therefore, for a fixed parameterization, we can write $\theta(\bar{y})$, and:

$$\frac{d}{dt} \bar{y}(t) = \text{Cov}_{Y \sim \pi_Y(\cdot; \theta(\bar{y}))}[Y]v(\bar{y}(t)) \quad (57)$$

which is an *autonomous* adaptive dynamics equation.

Since the density $\pi_X(\cdot; \theta)$ is uniquely determined by the parameters, the parameters are uniquely determined by the embedded centroid \bar{y} , and the embedded centroid obeys an autonomous system of equations, the centroid dynamics could, in principle, be solved independently, then inverted to recover the parameter and distributional dynamics. Solving for the inverse map is equivalent to solving a maximum likelihood estimation problem for all possible observed sufficient statistics [69]. In practice, this is infeasible since the inverse map $\theta(\bar{y})$ is intractable. However, the fact that such a map exists implies that the centroid dynamics are a sufficient and minimal representation of the distributional dynamics. Therefore, any continuous replicator dynamic with separable f is fully specified by an autonomous adaptive dynamics equation in the disc game space. In this sense, disc-game embedding radically simplifies the replicator dynamic. In the disc-game space, the continuous replicator dynamic *is* adaptive dynamics! \square

Remark: The adaptive dynamics equation, closely mirrors a self-play dynamic in the disc-game space since the initial direction of motion proposed on the right-hand side is $v(\bar{y})$, the best response direction when training in response to \bar{y} . In fact, Equation (57) *is* a self-play dynamic, with respect to a Riemannian metric whose metric tensor is the covariance matrix appearing in (57) [145].

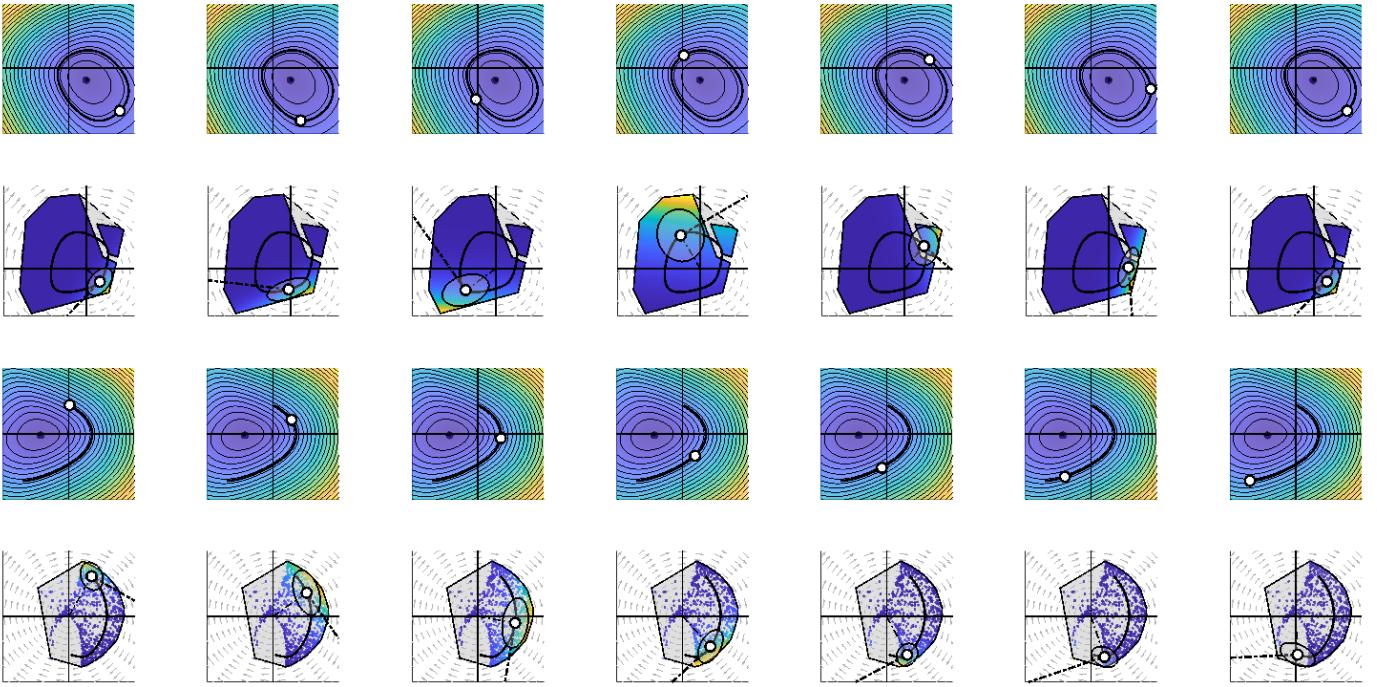


Figure 9: Autonomous centroid and Hamiltonian parameter dynamics for two example games. The top two rows correspond to the example game illustrated in Figures 3, 6, and 8. The bottom two rows correspond to the IPD example illustrated in Figure 4. The first and third rows show orbits in the parameter space. The second and fourth rows show the corresponding sequence of population distributions in the embedding space. The white circles mark the population centroids and parameters at a sequence of evenly spaced times (separate columns). The shaded oval centered at each centroid represents the population covariance. The bold dashed line leaving the centroid marks the optimal local training vector field, v , evaluated at the centroid. The solid black line represents the realized orbit formed by multiplying the covariance against v (see Equation (57)). Notice that, as the centroid approaches the boundary of $\text{conv}(\Psi)$, the covariance compresses so that its major axes are tangent to the boundary, while its minor axes are perpendicular to the boundary. Then, applying the covariance to v compresses the components of v pointing out of the boundary, and elongates those parallel to the boundary. This bends the self-play orbit so that the centroid remains within the convex hull of Ψ . For examples, see the second and fifth columns in the second row. It also slows the orbit near corners in the convex hull of Ψ (see the last three columns in the bottom row). When the centroid is well removed from the boundary, the population distribution is approximately equivariant, so the centroid dynamics align to the self-play orbit (align to v).

The covariance in Equation (57) is responsible for restricting solutions to the adaptive dynamics equation to the interior of the convex hull of Ψ . Consider a sequence of \bar{y} approaching the boundary of the convex hull of Ψ . Recall that $\pi_Y(y; \theta) = \pi_Y(y; 0)e^{\theta \cdot y}$. To move the centroid arbitrarily close to a point on the boundary, y_* , the parameters θ must be chosen to concentrate the mass of π_Y along the boundary of the convex hull of Ψ passing through the y_* . Thus, as \bar{y} approaches y_* , the covariance converges to a singular matrix, whose nullspace contains all directions normal to the boundary of the convex hull of Ψ at y_* . As a result, as \bar{y} approaches the boundary of the convex hull of Ψ , multiplication by the covariance increasingly suppresses components of $v(\bar{y})$ pointing out of the convex hull. Instead, the product skews $v(\bar{y})$ towards directions tangent to the boundary of the convex hull of Ψ (see Figure 9). \square

Remark: Like the parameter dynamics, the centroid dynamics match the intuition implied by the disc-game model. Namely, that solutions to the replicator equation should circulate unless barred by the boundaries of Ψ .

If the convex hull of Ψ includes the origin, then \bar{y} can approach it. A distribution whose centroid is near the origin is near to an interior equilibrium, π_* . If we set the reference measure to π_* , then the disc-game coordinates are uncorrelated at the equilibrium, so $\text{Cov}_{Y \sim \pi_Y(\cdot; \theta(\bar{y}))}[Y] = D_\omega$, the diagonal matrix with diagonal entries $\omega_1, \omega_1, \omega_2, \omega_2, \dots, \omega_{r/2}, \omega_{r/2}$. Then, linearizing equation (57) about $\bar{y} = 0$ gives the familiar system of decoupled oscillator equations corresponding to self-play in a disc game: $\frac{d}{dt}\bar{y}(t) \simeq D_\omega v(\bar{y}(t)) = D_\omega U\bar{y}(t) = W\bar{y}(t)$. Solutions to the self-play equation circulate independently in each pair of disc game coordinates at rates specified by ω and P . If the convex hull of Ψ does not include the origin, then the solutions to the adaptive dynamics equation are non-recurrent since the parameter dynamics are non-recurrent. Indeed, solutions to the adaptive dynamics equation satisfy exactly the same dichotomy that solutions to the parameter dynamics obey since the mapping between θ and \bar{y} is differentiable and invertible. Thus, the collection of solution trajectories to equations (137) and (57) are diffeomorphic (see Figure 8). \square

Why does the population centroid obey an autonomous adaptive dynamics equation in the disc game space?

We present two arguments. The first shows that the desiderata chosen to specify the disc game embedding are precisely the constraints needed to ensure that the centroid in the transformed space obeys an autonomous adaptive dynamics equation. The second shows that the equivalence between the replicator dynamic and adaptive dynamics is an immediate consequence of a more familiar equivalence between the replicator dynamic and self-play in population games.

First, since the conversion into the disc game space preserves outcome distinct agents, the continuous replicator equation in the original trait space induces a continuous replicator equation over the population distribution in the disc game space:

$$\begin{aligned}\partial_t \pi_Y(y, t) &= \int_{x \in T^{-1}(y)} \partial_t \pi(x, t) dx = \int_{x \in T^{-1}(y)} \mathbb{E}_{X' \sim \pi_x} [f(x, X')] \pi(x, t) dx \\ &= \mathbb{E}_{X' \sim \pi_x} [f(x(y), X')] \int_{x \in T^{-1}(y)} \pi(x, t) dx = \mathbb{E}_{Y' \sim \pi_Y} [g(y, Y')] \pi_Y(y, t)\end{aligned}$$

where $x(y)$ is any member of the equivalence class $T^{-1}(y)$.

Under a disc game embedding, $g(\cdot, \cdot)$ is a bilinear function, so:

$$\partial_t \pi_Y(y, t) = g(y, \bar{y}(t)) \pi_Y(y, t) \quad (58)$$

It follows that, if $\pi_Y(y, 0)$ is known, then the population density $\pi(y, t)$ can be recovered at any time using $\bar{y}(t)$ alone:

$$\pi_Y(y, t) = \exp \left(\int_{s=0}^t g(y, \bar{y}(s)) ds \right) \pi_Y(y, 0). \quad (59)$$

That is, given an initial distribution π_Y in the disc game space, the trajectory of the centroid fully specifies the distribution at all future times. Since the current centroid fixes the rate of change in the distribution, and the rate of change in the distribution fixes the centroid, $\frac{d}{dt} \bar{y}(t) = \int \partial_t \pi(y, t) dy$, the centroid in the disc game space must obey an autonomous equation. Notice that this property is shared by any coordinate transformation that preserves outcome distinct classes, and that satisfies desideratum 1.

Next, recall that, if g is bilinear then $g(y, y')$ can always be expressed $g(y, y') = y \cdot v(y')$ where $v(y')$ is the optimal training vector field evaluated at y' (see Lemma 2). Then, under any coordinate transformation that satisfies desideratum 1 and preserves the distinctions between outcome distinct classes:

$$\partial_t \pi_Y(y, t) = (y \cdot v(\bar{y})) \pi_Y(y, t). \quad (60)$$

Integrating to recover the rate of change in the centroid gives the autonomous adaptive dynamics equation (57):

$$\frac{d}{dt} \bar{y}(t) = \int_{y \in \Psi} yy^\top v(\bar{y}(t)) \pi_Y(y, t) dy = \left(\int_{y \in \Psi} yy^\top \pi_Y(y, t) dy \right) v(\bar{y}(t)) = \text{Cov}_{Y \sim \pi_Y(\cdot; \theta(\bar{y}))} [Y] v(\bar{y}(t)).$$

Lemma 16: [Bilinearity and Autonomous Adaptive Dynamics] If T is a mapping such that outcome distinct classes remain distinct after applying T , and that satisfies desideratum 1, then, given a population evolving according to a replicator dynamic, the centroid in the transformed coordinates must obey an autonomous adaptive dynamics equation.

Thus, the desiderata used to reach the disc game embedding coincide with the desiderata needed to ensure that the centroid, after the transformation, obeys an autonomous adaptive dynamics equation. This result realizes the promise made at the end of Section 3.1 and helps justify our choice of desiderata.

Next, consider the bilinear population game with strategy space equal to all distributions supported on Ω , and payout $f[\pi, \pi'] = \mathbb{E}_{X \sim \pi, X' \sim \pi'} [f(X, X')] = \int f(x, x') \pi(x) \pi'(x) dx$. Consider a self-play dynamic defined by setting $\partial_t \pi(x, t)$ proportional to $\nabla_\nu f[\nu, \pi(\cdot, t)]|_{\nu=\pi(\cdot, t)}$. The functional gradient is defined by maximizing the increase in payout constrained to a neighborhood of $\pi(\cdot, t)$ [145]. In the Shashahani metric, self-play recovers the replicator dynamic [6, 145, 195]. Since the exponential family formed by the embedding functions is a sufficient and minimal parameterization, the parameters of the exponential family, θ , must also move along the self-play gradient of the population payout in the metric formed by applying the Shashahani metric to distributions parameterized by Equation 41.

Since the mapping from parameters to centroid is diffeomorphic for exponential families [69], the population centroid in the disc game space also provides a sufficient and minimal parameterization, so must also obey some form of self-play dynamic. To recover its form, note that the population payout can also be expressed directly using the population centroids in the disc game space. By bilinearity, $f[\pi, \pi'] = \mathbb{E}_{X \sim \pi, X' \sim \pi'} [f(X, X')] = \mathbb{E}_{Y \sim \pi_Y, Y' \sim \pi'_Y} [Y^\top U Y] = \bar{y} \cdot U \bar{y}' = \bar{y} \cdot v(\bar{y}')$. It follows that, no matter the base measure, the self-play gradient with respect to the centroid is $\nabla_{\bar{y}} f[\pi_y, \pi'_y]|_{\pi=\pi'} = v(\bar{y})$. This gradient was computed using the Euclidean metric in the disc game space. The centroid dynamics must be a self-play dynamic for the population game, but in the metric implied by mapping from centroid back to distribution, then

applying the Shashahani metric. Changing the metric used to compute a gradient transforms it by a positive definite linear transformation, so, under the replicator dynamic, the centroid dynamics in the disc game space must take the form:

$$\frac{d}{dt} \bar{y}(t) = C(\bar{y}(t))v(\bar{y}(t)) \quad (61)$$

where $C(\cdot)$ is a matrix valued function that returns positive definite matrices for \bar{y} in the interior of $\text{conv}(\Psi)$. Equation (61) takes the generic adaptive dynamics form. Thus, the equivalence between the replicator and adaptive dynamics represents the generic equivalence between replicator dynamics and self-play dynamics in the population game.

4.2.6 Solutions for Simplifying Geometries

The disc game approach converts strategic problems into geometric problems by changing the game representation. In this section, we will consider example cases with embedded geometries Ψ that simplify the continuous replicator dynamic.

To initialize a continuous replicator dynamic in the embedding space we need to fix the number of disc games and the embedding of the initial population distribution, $\pi_Y(\cdot, 0)$. Since we are now working exclusively in the disc game space, we can drop the extra subscript Y , and will use π_0 to represent the initial population distribution. Since the continuous replicator dynamic is support preserving, we will assume, without loss of generality, that Ψ equals the support of π_0 . Then, the choice of Ψ may be expressed implicitly by the choice of π_0 .

Like any multivariate dynamical system, the parameter dynamics associated with the continuous replicator equation are easiest to solve when the parameters decouple. When they decouple, the system may be broken into a series of independent, lower-dimensional systems. These may be solved independently.

Let $y^{(k)} = [y_{2k-1}, y_{2k}]$ denote the pair of disc game coordinates assigned to the k^{th} game. Similarly, let $\theta^{(k)} = [\theta_{2k-1}, \theta_{2k}]$ denote the parameters associated with the k^{th} game. When do the parameter dynamics, $\theta(t)$ decouple?

Lemma 17: [Decoupled Parameter Dynamics] Consider $Y \sim \pi_0$. If the set of component disc games $D = \{1, 2, \dots, r/2\}$ admits a partition $\mathcal{K} = \{K_j\}$ into disjoint subsets such that $Y^{(k)}$ is independent of $Y^{(k')}$ when k and k' are members of different sets in the partition, then each set of parameters, $\Theta_j = \{\theta^{(k)}, k \in K_j\}$ evolves independently of each other set of parameters, and, if $Y \sim \pi(\cdot, t)$, then $\{Y^{(k)}, k \in K_j\}$ remain independent of $\{Y^{(k')}, k \in K_{j'}\}$ for all t .

Proof Outline: Under the independence assumptions the base measure factors into a product of marginal measures. The exponential tilt also factors into marginal terms, so the integral defining the Hamiltonian factors. As a result, the gradient of the log Hamiltonian separates, decoupling the parameter dynamics. For details see Appendix section 6.7.1. \square

Corollary 17.1: [Fully Decoupled Parameter Dynamics] Consider $Y \sim \pi_0$. If $Y^{(k)}$ is independent of $Y^{(k')}$ for all $k \neq k'$, then each pair of parameters $\theta^{(k)}$ evolves independently of each other pair of parameters, and, if $Y \sim \pi(t)$, then $Y^{(k)}$ remains independent of $Y^{(k')}$ for all t .

Corollary 17.1 relates problems with many disc game components to replicator dynamics on a single disc game. If we construct π_0 as a product of marginal measures assigned to each disc game, then the parameters in each disc game evolve independently, so may be solved without studying the other games. For the remainder of this section, we will focus on replicator dynamics in a single disc game. These may be lifted to quasi-periodic dynamics in more complex games by Corollary 17.1.

At simplest, $\pi_0(y)$ is an exponential tilt of a rotationally symmetric base measure, $\pi_0(y) = \exp(\theta(0) \cdot y)\nu(y)$, where $\nu(y)$ is only a function of $\|y\|_2$. For example, consider ν uniform on a ball centered at zero, or ν set equal to a centered normal with covariance proportional to an identity. Allowing exponential tilts includes all π_0 that are non-centered normals with covariance proportional to an identity.

Lemma 18: [Rotational Symmetry] Let $\Psi \in \mathbb{R}^2$ be the domain for a single disc game, and suppose that π_0 is an initial population distribution that is an exponential tilt of a base measure ν that is symmetric under rotations. If the population evolves according to a continuous replicator dynamic, then its parameters, using base measure ν , orbit on concentric circles at a constant rate with period equal to $2\pi\|\theta(0)\|_2/\|\nabla_\theta P(\theta(0))\|_2$. In the parametrization with base measure π_0 , the parameters orbit on concentric ellipses with period $2\pi\|\theta(0)\|_2/\|\nabla_\theta P(\theta(0))\|_2$.

Proof: If π_0 is an exponential tilt of ν where ν is rotationally symmetric then, with ν as the base measure, $P(\theta)$ must be invariant to rotations of θ , so is also only a function of $\|\theta\|_2$. Then, since $P(\theta)$ is the Hamiltonian, and Hamiltonian dynamics follow level sets of their Hamiltonian function, the parameters must evolve along concentric circles. For remaining details, see Appendix section 6.7.2. \square

The continuous replicator dynamics simplify further in the case when all components of $Y \sim \pi_0$ are independent. Then, π_0 fully factors into a product of marginals so each separate pair of parameters evolves independently. The parameters associated with the same disc game remain coupled through the dynamics, however, the coupling simplifies.

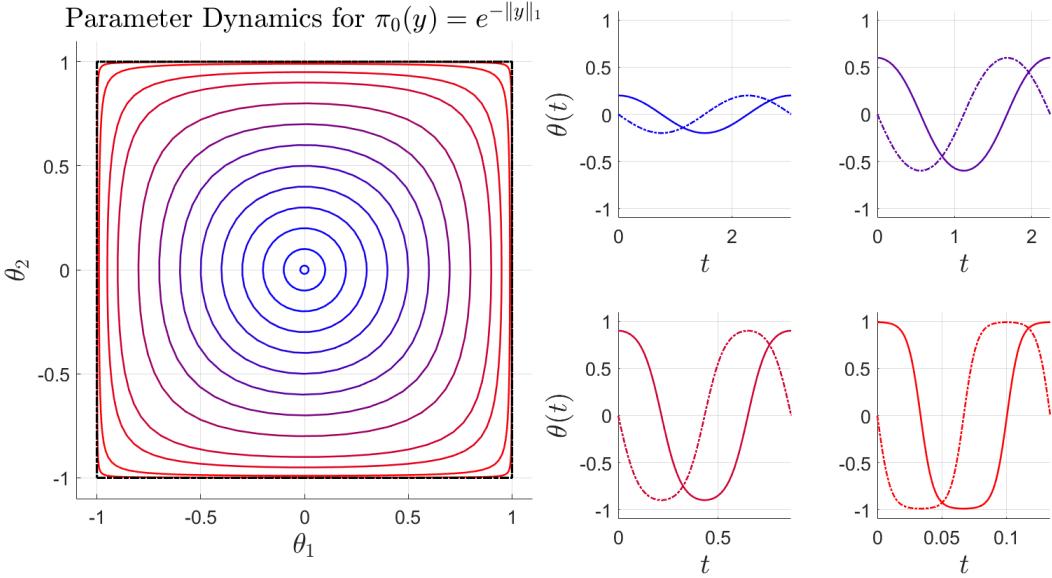


Figure 10: Parameter dynamics for a continuous replicator equation defined in a single disc game, with a Laplace initial distribution. The time traces correspond to initial amplitudes $a = [0.2, 0.6, 0.9, 0.99]$. Solid lines are θ_1 and dashed are θ_2 . Each time trace covers one full period. Notice that, for small amplitudes, the parameters orbit on approximately circular level sets, but for large amplitudes the level sets approach the square defining the boundary of valid parameters. Also notice that, for small amplitude, the parameters move at a constant angular rate while, for large amplitude, each orbit dwells near the corners of the square, before transitioning rapidly to a new corner when near an edge of the square. In all cases the parameters orbit clockwise, i.e. in the direction of advantage in a disc game.

Consider a single disc game, with a base measure that factors into a product of marginal measures, $\pi_0(y) = \pi_0^1(y_1)\pi_0^2(y_2)$. Then, the Hamiltonian fully factors, $P(\theta_1, \theta_2) = \prod_{j=1}^2 P_j(\theta_j)$ where $P_j(s) = \int_{y_j} e^{sy_j} d\pi_0^j(y_j)$ is the moment generating function (Laplace transform) of the marginal measure π_0^j . Then, the Hamiltonian parameter dynamics reduce to:

$$\frac{d}{dt}\theta_1(t) = Pg_2(\theta_2), \quad \frac{d}{dt}\theta_2(t) = -Pg_1(\theta_1) \quad \text{where} \quad g_j(s) = \frac{d}{ds} \log(P_j(s)) \quad (62)$$

is the slope of the cumulant generating function and $P = P(\theta(0))$ is the initial population size. For a detailed derivation see Appendix section 6.7.3.

Equation (62) is especially simple when the marginals for Y_1 and Y_2 are identical. Then we can suppress the subscripts, and, scaling time by the total population size, reduce to a normal form nonlinear oscillator equation:

$$\frac{d}{dt}\theta_1(t) \propto g(\theta_2(t)), \quad \frac{d}{dt}\theta_2(t) \propto -g(\theta_1(t)). \quad (63)$$

Suppose, for example, that π_0 is a Laplace distribution, so has marginals of the form $\frac{1}{2} \exp(-|y_j|)$. Then:

$$\frac{d}{dt}\theta_1(t) \propto \left(\frac{2}{(1 - \theta_2(t)^2)} \right) \theta_2(t), \quad \frac{d}{dt}\theta_2(t) \propto -\left(\frac{2}{(1 - \theta_1(t)^2)} \right) \theta_1(t). \quad (64)$$

Equation (64) defines a separable elliptic oscillator, that approaches a simple harmonic oscillator for small θ . It is well defined for all θ in the unit square $[-1, 1]^2$. The solution trajectories are the level sets of the Hamiltonian function $((1 - \theta_1^2)(1 - \theta_2^2))^{-1}$. The dynamical system admits an analytic solution for all initial conditions.

Since the solutions are all periodic orbits in the square $[-1, 1]^2$ we can adopt convenient initial conditions. Set $\theta_1(0) = 0$ and $\theta_2(0) = a > 0$, where a represents the amplitude of the oscillator when it crosses a coordinate axis. Then, every solution to equation (161) lies on an orbit of the form:

$$\theta_1(t) = \operatorname{asn}(2Pt|a^2), \quad \theta_2(t) = a \frac{\operatorname{cn}(2Pt|a^2)}{\operatorname{dn}(2Pt|a^2)} \quad (65)$$

where sn , cn , and dn are the elliptic sine, elliptic cosine, and delta amplitude functions. Each is an example of a Jacobi elliptic function [63, 119]. For the detailed derivation, calculation of the period, and solution for arbitrary initial conditions, see Appendix section 6.7.3.

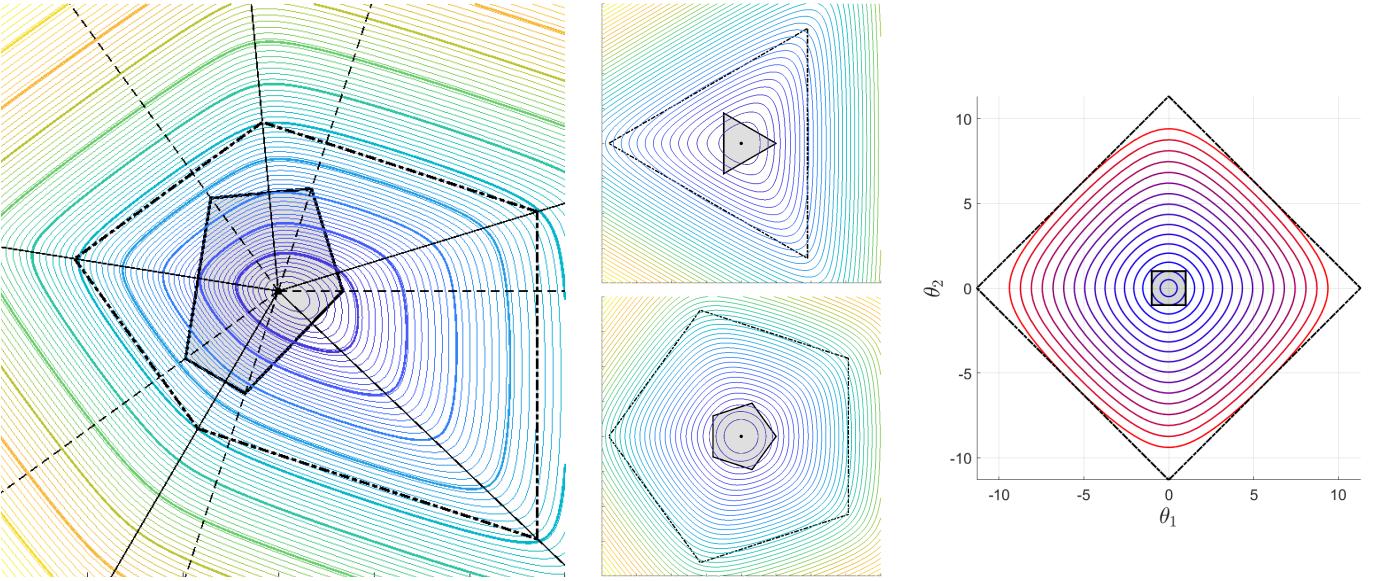


Figure 11: **Left:** Solution trajectories for the parameter dynamics of a population evolving under the continuous replicator dynamic (colored curves, colored by value of the Hamiltonian) with initial measure equal to a uniform distribution over a polygon Ψ (shaded grey region). The dot-dashed outer polygon is dual to Ψ . Its vertices lie on rays normal to the boundaries of Ψ . Its boundaries are perpendicular to rays passing through the vertices of Ψ . Notice that trajectories with large amplitudes (large Hamiltonian) converge to forms similar to the dual polygon. **Middle:** Solutions for two regular polygonal regions (triangles and pentagons). **Right:** Solution trajectories for π_0 uniform on the square $[-1, 1]^2$.

The corresponding orbits are shown in Figure 10 for amplitudes ranging from 0.1 to 0.999. Notice that, small amplitudes produce approximately circular orbits, while large amplitudes produce approximately square orbits. Also notice that, when the amplitude is small, the parameters orbit at an approximately constant velocity while, when the amplitude is large, the parameters dwell near the corners of the square.

This behavior is generic when the convex hull of Ψ is polygonal. For example, if the base measure separates into a product of semi-circle measures, then Ψ is square, $g(s) = I_2(s)/I_1(s)$ where $I_j(s)$ is the modified Bessel function of the first kind, and the oscillator equations produce polygonal parameter orbits similar to heteroclinic motion about the boundary of a square. The Hamiltonians plotted in Figure 7 show that the level sets of the Hamiltonian approach polygons when far from the minimizer of the Hamiltonian for a variety of initial domains. Since the parameter orbits, and centroid orbits, are diffeomorphic, and large exponential tilts move the centroid near to the boundary of Ψ , these limits correspond to the limiting orbits of the adaptive dynamics equation, in the limit where the centroid approaches the boundary of the convex hull of Ψ (c.f. Figure 8). We will show that, when the convex hull of Ψ is polygonal, then the parameters approach polygonal orbits at large Hamiltonian values, and, the corners of the convex hull of the embedding space, Ψ , are key geometric features that generate the limiting polygonal orbit. Moreover, we will show that, as the centroid approaches the boundary, then the centroid dynamics slow down near corners of Ψ . Therefore, the centroid behaves like a heteroclinic oscillator near the boundaries of Ψ .

To study these orbits, we focus on a single disc game, assume that Ψ is radially bounded so that its convex hull is compact, that the convex hull contains the origin, and that the convex hull is a polygon. Polygonal Ψ occur in any normal form game with a finite strategy space since, in a normal form game, the payout is bilinear, so the embedding maps are linear functions, and Ψ is a linear function of the space of mixed strategies, which is itself a convex polytope [206]. In this case, the vertices of the embedding correspond to pure strategies. To simplify the geometry of the parameter orbits we will also assume that the base measure is uniform over Ψ .

The orbits of interest correspond to solutions of the parameter dynamics in a limit where the value of the Hamiltonian on the orbit diverges, or, equivalently, where $\inf_t(\|\theta(t)\|_2)$ diverges. We will also work, without loss of generality, after scaling time by the initial population size, or, equivalently, with a constant interaction rate model. Then, $\frac{d}{dt}\theta(t) = U\bar{y}(\theta)$ where $\bar{y}(\theta)$ is the population centroid in the disc game space associated with the normalized distribution $p(y, \theta) = \pi(y, \theta)/P(\theta)$.

Let $\text{conv}(\Psi)$ denote the convex hull of Ψ , and $C = \{c_1, c_2, \dots, c_m\}$ denote the corners of the convex hull. Consider $\frac{d}{dt}\theta(t)$ at $t = 0$ if $\theta(0) = a\hat{\theta}_0$ in the limit as the amplitude, a , diverges. Then $p(y; a\hat{\theta}_0)$ approaches a measure that places all of its mass on the set of $y \in \Psi$ that maximize $\hat{\theta}_0 \cdot y$. When Ψ is polygonal, the maxima are solutions to a linear programming problem. These are restricted to the vertices of the domain for generic $\hat{\theta}_0$, i.e. the corners C . More generally, since Ψ has a polygonal convex hull, Ψ is contained inside its convex hull, and all vertices of the convex hull

are elements of Ψ , the set of y that maximize $\hat{\theta}_0 \cdot y$ is restricted to the corners C of $\text{conv}(\Psi)$ for almost all $\hat{\theta}_0$.

If Ψ includes a linear boundary segment connecting two corners of its convex hull, then all y on the boundary maximize $\hat{\theta}_0 \cdot y$ for $\hat{\theta}_0$ normal to the boundary. Since the normal to the boundary is uniquely defined, and Ψ has, at most m linear boundary segments that match a boundary of its convex hull, there are, at most, m values of $\hat{\theta}_0$ for which the argument maximizer of $\hat{\theta}_0 \cdot y$ is non-unique. Let $\hat{N} = \{\hat{n}_1, \hat{n}_2, \dots, \hat{n}_m\}$ denote the collection of normals to the boundaries of $\text{conv}(\Psi)$.

List the boundaries of the convex hull sequentially. Then, the sequence of normal directions $\{\hat{n}_1, \hat{n}_2, \dots, \hat{n}_m\}$ partition \mathbb{R}^2 into m angular regions, bounded by the rays $\{\{s\hat{n}_j\}_{s \geq 0}\}_{j=1}^m$. If the ray $\{s\hat{\theta}_0\}_{s > 0}$ points into the region bounded by $\{s\hat{n}_j\}_{s \geq 0}$ and $\{s\hat{n}_{j+1}\}_{s \geq 0}$, where indices are counted mod m , then $\hat{\theta}_0 \cdot y$ is maximized at the vertex contained between edges j and $j + 1$.

Then, for each corner c_j , we can identify an angular region, bounded by the rays parallel to \hat{n}_{j-1} and \hat{n}_j , such that, if $\hat{\theta}_0$ lies inside the region, then $p(y, a\hat{\theta}_0)$ approaches a delta distribution at the corner c_j . In this case, $\lim_{a \rightarrow \infty} \bar{y}(a\hat{\theta}_0) = c_j$. Since this limit holds for all rays pointing into the j^{th} region, the level sets of the Hamiltonian approach straight line segments parallel to U_{c_j} inside the angular region. Since the angular regions generated by the corners partition all angles about the origin, and since all level sets of the Hamiltonian form a closed orbit, in the limit of large initial θ , the parameter orbits must approach a series of similar polygons.

Lemma 19: [Dual Polygons] Suppose that $\text{conv}(\Psi)$ is a bounded polygon, P , and π_0 is uniform on Ψ . Then, the parameters orbits approach, in the limit of large amplitude, a dual polygon, P' , with an edge for every vertex of P , and a vertex for every edge of P . The dual polygon is defined by the orthogonality constraints:

1. The vector from the origin to the j^{th} vertex of P' is perpendicular to the j^{th} edge of P .
2. The vector from the origin to the j^{th} vertex of P is perpendicular to the j^{th} edge of P' .

Such a dual exists for all P . It equals the polar reciprocal of P with respect to the unit circle [25, 35]. Figure 11 shows two polygons related by polar reciprocation, and an example set of parameter orbits converging to the dual of an irregular polygonal domain Ψ .

If Ψ is a regular polygon, then the dual polygon associated with limiting orbits is a regular polygon of the same degree, rotated so that its edges are bisected by the vertices of Ψ . For example, the square boundary observed for a Laplace base measure is the dual to the diamond-shaped level sets of the Laplace density. The middle column of Figure 11 shows example trajectories for a uniform measure on a regular triangle and pentagon.

The time it takes to traverse each edge of the limiting polygon diverges as the initial parameter amplitude, a , diverges since the length of the segment diverges linearly in the initial amplitude a , but, the rate of travel approaches the distance from the associated corner of Ψ to the origin, which is bounded. As a consequence, the centroid dynamics slow down near the corners of $\text{conv}(\Psi)$ with the time spent near a corner diverging linearly in the parameter amplitude a . In the limit, each corner acts as an equilibrium, since, if the normalized population is initialized with centroid at the corner, it must be monomorphic, so, by support conservation, cannot change over time.

For example, consider a uniform base measure on the square $[-1, 1]^2$. Then, the Hamiltonian factors, $P(\theta) = (\sinh(\theta_1)/\theta_1)(\sinh(\theta_2)/\theta_2)$ so the parameter dynamics simplify (See Appendix section 6.7.4). For large θ_1 and θ_2 , the dynamics are approximated by:

$$\frac{d}{dt}\theta_1(t) = \text{sign}(\theta_2(t)) + \mathcal{O}(|\theta_2(t)|^{-1}), \quad \frac{d}{dt}\theta_2(t) = -\text{sign}(\theta_1(t)) + \mathcal{O}(|\theta_1(t)|^{-1}). \quad (66)$$

Then, ignoring the small corrections, $\theta_1(t)$ and $\theta_2(t)$, follow diamond-shaped orbits at a constant speed equal to $\sqrt{2}$ (see the right-most column of Figure 11). Each side of the diamond corresponds to a corner of the original square domain. If the parameters are initialized with amplitude a (e.g. they pass through the point $[a, 0]$), then each side of the diamond has length $\sqrt{2}a$ and is traversed at rate $\sqrt{2}$. Then, the time spent on any side is asymptotic to its amplitude, a . Therefore, the time the centroid \bar{y} spends near each corner diverges as the centroid orbit approaches the boundary.

These observations suggest that the boundary of $\text{conv}(\Psi)$ should play an organizing role in explaining continuous replicator dynamics. When the initial distribution is concentrated near a boundary of Ψ , or places its centroid near a boundary of $\text{conv}(\Psi)$, then the boundary of the embedding space determines the shape of the parameter dynamics, and, the corners of the embedding space act like the semistable equilibria of a heteroclinic orbit.

Extending the intuition developed with polygons, we should expect regions where the boundary is sharply curved to act like corners whose incident edges form an acute angle. These produce long edges in the dual polygon, so require a long time to transit. Therefore, the centroid will slow down when it approaches a point on the boundary of the convex hull of ψ with large curvature. In contrast, regions with low curvature will act like edges in the polygonal approximation, so will not slow the centroid.

4.3 Generalization

A successful paradigm should provide accurate intuition for a family of related problems. Here, we show that the conclusions drawn in the context of the continuous replicator dynamic extend to a family of dynamics that: (a) are equivalent to self-play in the population game using a different metric over the space of unnormalized densities, (b) allow additional frequency dependence in the per-capita growth rates of types, and (c) extend to metapopulation models that mix inhomogeneously.

4.3.1 Additional Frequency Dependence

Following [145], we will derive a general class of evolutionary dynamics by introducing a Lagrangian function, then selecting the dynamic to instantaneously maximize the Lagrangian. We will construct the Lagrangian as a mixture of two terms. The first is a fitness flux [157]. Consider a time evolving population $\pi(\cdot, t)$, changing at rate $\dot{\pi}(\cdot, t) = \frac{d}{dt}\pi(\cdot, t)$. Then the fitness flux is defined by the bilinear form:

$$\phi[\pi, \dot{\pi}; f] = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} f[\pi + \Delta t \dot{\pi}, \pi] = \langle \dot{\pi}(\cdot), F[\pi](\cdot) \rangle = \int_{x, x' \in \Omega \times \Omega} \dot{\pi}(x) f(x, x') \pi(x) dx' dx. \quad (67)$$

The second is a cost-of-motion that penalizes rapid changes in the distribution. It evaluates the rate of change, squared, of the distribution given an implicitly defined Riemannian metric over the space of distributions. We will consider costs of motion of the form:

$$K[\pi, \dot{\pi}; g] = \frac{1}{2} \langle \dot{\pi}(\cdot), G_\pi[\pi](\cdot) \rangle = \frac{1}{2} \int_{x, x' \in \Omega \times \Omega} \dot{\pi}(x) \frac{\delta(x - x')}{g(\pi(x))} \dot{\pi}(x') dx dx' = \frac{1}{2} \int_{x \in \Omega} \frac{\dot{\pi}(x)^2}{g(\pi(x))} dx \quad (68)$$

for $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, and where $g(\pi) = 0$ if and only if $\pi = 0$. We will restrict our attention to g continuous, monotonically increasing, $\mathcal{O}(\pi)$ as $\pi \rightarrow 0$ and as $\pi \rightarrow \infty$, possibly with different bounding constants. These ensure that the ensuing dynamic is support preserving in finite time for any game with bounded payouts.

Changing g varies the metric structure imposed on the space of distributions. For example, setting $g(\pi) = \pi$ returns the Shahshahani, or Fisher information, metric [144]. Then $K[\pi, \dot{\pi}]$ is the squared rate of change in the symmetrized Kullbach-Liebler divergence (Jensen-Shannon distance) between $\pi + \dot{\pi}\Delta t$ and π .

Next, construct the Lagrangian:

$$L[\pi, \dot{\pi}; f, g, \lambda] = \phi[\pi, \dot{\pi}; f] - \lambda K[\pi, \dot{\pi}; g]. \quad (69)$$

Then, since the cost of motion is strictly convex in $\dot{\pi}(x)$ over the support of $\pi(x)$, maximizing the Lagrangian with respect to $\dot{\pi}$ given π is equivalent to computing the self-play gradient of the population payout in the metric induced by the cost of motion [144]. To maximize the Lagrangian, set its variational gradient with respect to $\dot{\pi}$ to zero:

$$\nabla_{\dot{\pi}} L[\pi, \dot{\pi}_*(\pi); f, g, \lambda](x) = F[\pi](x) - \lambda \frac{1}{g(\pi(x))} \dot{\pi}_*(\pi) = 0. \quad (70)$$

The ensuing generalized replicator dynamic takes the form:

$$\partial_t \pi(x, t) \propto \mathbb{E}_{X' \sim \pi(\cdot, t)} [f(x, X')] g(\pi(x, t)) = \mathbb{E}_{X' \sim \pi(\cdot, t)} [f(x, X')] \frac{g(\pi(x, t))}{\pi(x, t)} \pi(x, t). \quad (71)$$

Varying the metric over the space of distributions by varying g amounts to varying the formula for the per capita growth rate of type x as a function of its expected payout and frequency. The function g controls the growth rate. Setting $g(\pi) = \pi$ recovers the standard replicator dynamic. Otherwise, the per-capita rates scale in the frequency of each type according to the ratio $g(\pi(x, t))/\pi(x, t)$. A dynamic of the form (71) applies whenever the ratio of the growth rate of a type to its expected payout is a nonnegative function of the frequency of the type. This is a useful generalization since saturating per capita rates are a common feature in ecological models.

Choosing $g(\pi) = \mathcal{O}(\pi)$ ensures that these per capita rates remain finite for all π , although they may limit to different constants as π approaches zero or approaches infinity. In particular, if $g(\pi) = \mathcal{O}(\pi)$ and f is bounded then $F[\pi](x)$ is bounded, so $\pi(x, t)$ can be bounded above and below by an envelope that grows exponentially in t . Since no exponential function can reach zero in finite time, or diverge in finite time, the dynamic will preserve the support of π , and will not introduce singularities that place finite mass at a single type in finite time.

Allowing nonlinear $g(\cdot)$ allows additional modeling freedom. For example, if g is monotonically increasing and concave down, then less occupied types will change more rapidly than highly occupied types in response to game outcomes. This could model a setting where low density types can change abundance more rapidly than high density types, when rates of change are measured per capita.

Given g , define h^{-1} and h :

$$\frac{d}{du} h^{-1}(u) = g(h^{-1}(u)), \quad h(\pi) = \int_1^\pi \frac{1}{g(s)} ds. \quad (72)$$

Since $g(\pi) = \mathcal{O}(\pi)$, $h(\pi) \rightarrow -\infty$ as $\pi \rightarrow 0$ and $h(\pi) \rightarrow \infty$ as $\pi \rightarrow \infty$. Since g is nonnegative, and positive for all positive entries h and h^{-1} are monotonically increasing. Since g is finitely valued and continuous, h and h^{-1} are continuously differentiable. It follows that h defines an diffeomorphic map from the positive half-line to the real line, and h^{-1} maps from the whole real line to the positive half line. If, in addition, g is monotonically increasing, then h is concave and h^{-1} is convex.

Let $u(x, t) = h(\pi(x, t))$. In the standard replicator equation g is the identity function, so $1/g(\pi) = 1/\pi$ and $h(\cdot) = \log(\cdot)$. Alternately, the saturating function $g(\pi) = \pi/(1 + \pi)$ yields $h(\pi) = \int_1^\pi 1 + 1/s ds = (\pi - 1) + \log(\pi)$ so $h^{-1}(u) = W(e^{u+1})$ where W is the Lambert W , or product-log function. The function $g(\pi) = \pi^2/(1 + \pi)$ yields $h^{-1}(u) = 1/W(e^{1-u})$. This models a situation where per capita growth rates are asymptotically proportional to expected payout, but are subject to an Allee effect that suppresses growth at low densities [34, 140, 203]. Classic examples include increased difficulty finding mates, decreased chance of pollination, loss of genetic diversity, or the collapse of social groups [52, 203]. In the following theory, h will replace the log, and h^{-1} the exponential, in the exponential family parameterization used to study the replicator equation.

Then:

$$\partial_t u(x, t) = \mathbb{E}_{X' \sim h^{-1}(u(\cdot, t))}[f(x, X')] = F[h^{-1}(u(\cdot, t))] \in \text{range}(F). \quad (73)$$

It follows that we can build an invariant family of densities by expressing $u(x, t)$ as a linear combination of the embedding functions:

$$u(x, t) = \theta(t) \cdot y(x) + u_0(x), \quad \pi(x, t) = h^{-1}(u(x, t)) \quad (74)$$

where $u_0(x)$ is the projection of $u(x, 0) = h(\pi(x, 0))$ onto the orthogonal complement of the range of F with respect to the $\langle \cdot, \cdot \rangle_\nu$ inner product:

$$u_0(x) = h(\pi(x, 0)) - \text{Proj}_{\text{range}(F)}[h(\pi(x, 0))]. \quad (75)$$

Theorem 6: Suppose that π obeys the generalized replicator dynamic specified in equation (71). Let h^{-1} be the inverse of $h = \int_1^\pi 1/g(s) ds$. Then the parameters θ of the parameterization (74) obey a Hamiltonian dynamic of the form:

$$\frac{d}{dt} \theta(t) = U \mathbb{E}_{X \sim \pi(\cdot; \theta(t))}[y(X)] = U \nabla_\theta H(\theta) \quad (76)$$

with Hamiltonian:

$$H(\theta) = \int_{x \in \Omega} \int_0^{\theta \cdot y(x) + u_0(x)} h^{-1}(u) du \quad (77)$$

Proof Outline: The proof follows immediately from the construction of the Hamiltonian, using the same techniques applied to prove Theorems 4 and 5 for the standard case when $h^{-1}(u(x; \theta, u_0))$ defined an exponential family. See Appendix Section 6.6.4. \square

Equation (77) can be used to compute the Hamiltonian for different functions g . Setting g to the identity recovers the total population size as illustrated before: $1/g(s) = 1/s$ so $h(\pi) = \log(\pi)$ so $h^{-1}(u) = \exp(u)$ so $\int^P h^{-1}(u) = \exp(u)$. Importantly, $H(\theta)$ need not equal $P(\theta)$ for generic g , so the relationship between the total population size and the Hamiltonian is a special feature of the standard replicator dynamic. Indeed, the generalized dynamic need not conserve the total population size. Figure 12 illustrated the four example Hamiltonians for $g(\pi) = \pi$ and $g(\pi) = \pi/(1 + \pi)$ for the same cases illustrated in Figure 7.

We can reuse the same sequence of arguments established in the standard case to prove essentially the same sequence of results for the general case.

First, the gradient and Hessian of the Hamiltonian are:

$$\nabla_\theta H(\theta) = \mathbb{E}_{X \sim \pi(\cdot; \theta)}[y(X)], \quad \mathbf{H}_{ij}(\theta) = [\partial_\theta^2 H(\theta)]_{i,j} = \int_{x \in \Omega} g(\pi(x; \theta)) y_i(x) y_j(x) dx. \quad (78)$$

The Hamiltonian is evidently convex since $v^\top \mathbf{H}(\theta)v = \int_{x \in \Omega} g(\pi(x; \theta))(v \cdot y(x))^2 dx$, and the integral of any nonnegative function is greater than or equal to zero. It is also strictly convex given some constraints on g .

Lemma 20: [Strict Convexity] If f is bounded, $0 \leq g(\pi)$ with equality if and only if $\pi = 0$, and $g(\pi) = \mathcal{O}(\pi)$ as $\pi \rightarrow 0$ and $\pi \rightarrow \infty$, then the Hamiltonian function $H(\cdot)$ is a strictly convex, second-differentiable function of θ for all interior θ .

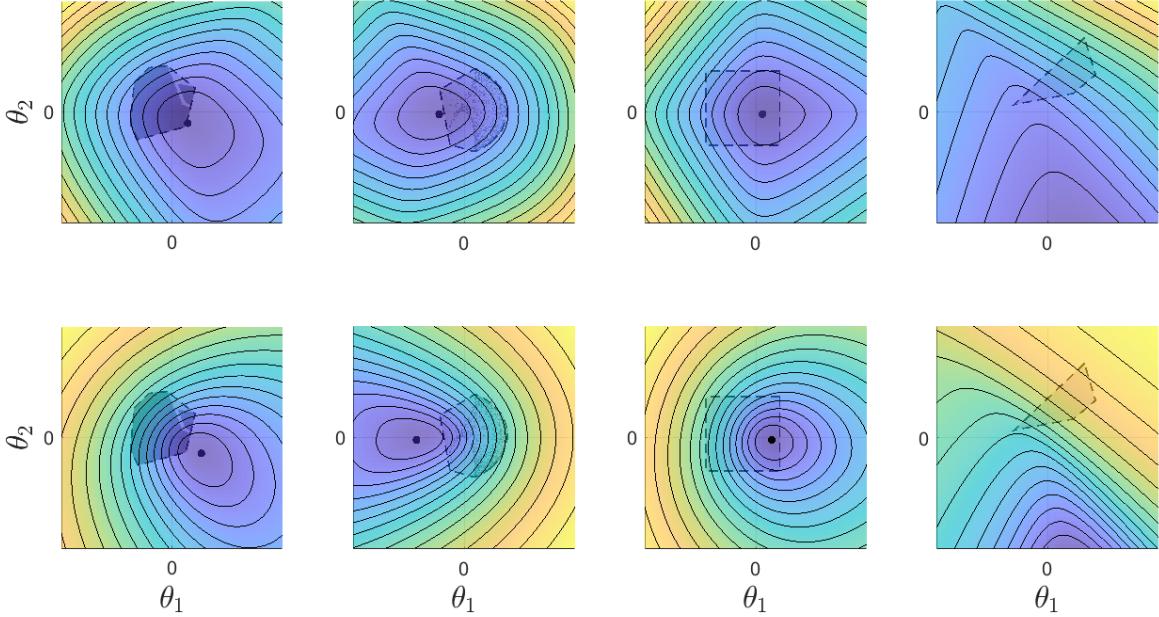


Figure 12: The Hamiltonian functions for four different games (columns), and for $g(p) = p$ (top row), versus $g(p) = p/(1+p)$ (bottom row). The games are represented by the image of Ω under the embedding map, Ψ , whose boundary is shown as the dashed region in each plot. The second column corresponds to the IPD example (see Figure 4). The Hamiltonians are similar near their minimizers, since $p \simeq p/(1+p)$ for small p , but differ in shape for large θ . The level sets of the Hamiltonians generated by $g(p) = p$ are more polygonal, with sharper corners, while the level sets of the Hamiltonians generated by $g(p) = p/(1+p)$ are smoother, with fewer pronounced corners. This follows since $g(p) = p/(1+p)$ saturates for large p , so $h^{-1}(u)$ is nearly linear in u for large u .

Proof Outline: Convexity is established by nonnegativity of the integrand. Existence is guaranteed by bounding f , as, if f is bounded, then the embedding functions are bounded. If $g(\pi) = \mathcal{O}(\pi)$, then the dynamic is support-preserving, so π is supported on Ω for all $\theta(t)$ reachable in finite time from a density with initial support Ω . Then, $g(\pi(x; \theta))$ is nonzero for all x and any $\theta(t)$ reachable in finite time. If $g(\pi) = \mathcal{O}(\pi)$ then $g(\pi) \leq C\pi$ for some C , so the integral $\int_{x \in \Omega} g(\pi(x; \theta))$ is finite. Let $\pi^g(x; \theta) \propto g(\pi(x; \theta))$ denote a normalized density supported on all of Ω . Then, the Hessian is proportional to the second moment tensor for $y(X)$ when $X \sim \pi^g$. This tensor is positive definite by construction of the embedding map. See Appendix Section 6.6.5 for details and proof with a more general theorem statement. \square

If the Hamiltonian is strictly convex, then it is radially unbounded if it admits an isolated interior equilibrium, so the parameter dynamics satisfy the same dichotomy between recurrent and boundary-seeking dynamics observed in the standard case when g is the identity function.

Corollary 6.1: [Recurrence] Suppose that the conditions of Lemma 20 hold. Then, the parameter dynamics either admit an isolated, interior equilibrium that corresponds to the choice of θ such that $\mathbb{E}_{X \sim \pi(\cdot; \theta)}[y(X)] = 0$, so correspond to a fully mixed NE, or no such equilibrium exists. The former case requires $0 \in \text{conv}(\Psi)$ and is impossible if the convex hull of Ψ does not contain the origin. If an isolated interior equilibrium exists then it is a center and the parameter dynamics are Poincaré recurrent.

Proof Outline: Under the conditions of Lemma 20, the Hamiltonian is strictly convex so either admits a unique, isolated, interior minima, which is an isolated equilibrium for the parameter dynamics, or it does not, and all parameter orbits escape to infinity. The remaining statements follow by the same arguments used for the standard replicator dynamic. For details, see Appendix Section 6.6.6 \square

Suppose that such an equilibrium exists and denote it θ_* . The Jacobian at θ_* is $U\mathbf{H}(\theta_*)$. Then, up to linearization, the parameter dynamics obey a simple Harmonic oscillator equation for a sequence of $\lfloor r/2 \rfloor$ independent oscillators. To find the frequency/period of each oscillator, recall that changing the reference measure without changing its support transforms the parameter dynamics by an invertible linear transformation, so preserves the topology of the orbits and their period. It follows that, up to linearization, the parameters obey the approximate dynamic:

$$\frac{d}{dt} \Delta\theta_1^{(k)}(t) \simeq P(\theta_*) \omega_k \Delta\theta_2^{(k)}(t), \quad \frac{d}{dt} \Delta\theta_2^{(k)}(t) \simeq -P(\theta_*) \omega_k \Delta\theta_1^{(k)}(t) \quad (79)$$

where $\Delta\theta = \theta - \theta_*$, and where ω_k is the magnitude of the $2k - 1^{st}$ eigenvalue of F_ν for $\nu = \frac{1}{P(\theta_*)}\pi(x; \theta_*)$.

Since the generalized dynamic is a self-play dynamic for the population game, it can be reduced to adaptive dynamics in the disc game coordinates. Corollary 6.2 establishes this fact for the generalized replicator dynamic using the same arguments developed for the standard case. Any population evolving under a generalized replicator dynamic, also obeys an adaptive dynamics equation in the disc game space, and, the centroid dynamics in the disc game coordinates are an equivalent specification of the full distributional dynamic given the initial distribution.

Corollary 6.2: [Adaptive Dynamics] Let let $\bar{y}(\theta) = \mathbb{E}_{X \sim \pi(\cdot; \theta)}[y(X)]$.

$$\frac{d}{dt} \bar{y}(t) = \mathbf{H}(\theta(\bar{y}(t); u_0))v(\bar{y}(t)) \quad (80)$$

where $\theta(\bar{y}; u_0)$ is the unique solution to the equation $\mathbb{E}_{X \sim \pi(\cdot; \theta, u_0)}[y(X)] = \bar{y}$ given $u_0(x) = h(\pi(x, 0))$.

Proof: Given $\bar{y}(\theta) = \mathbb{E}_{X \sim \pi(\cdot; \theta)}[y(X)]$,

$$\frac{d}{dt} \bar{y}(\theta(t)) = \frac{d}{dt} \mathbb{E}_{X \sim \pi(\cdot; \theta)}[y(X)] = \frac{d}{dt} \nabla_\theta H(\theta(t)) = \mathbf{H}(\theta(t)) \frac{d}{dt} \theta(t) = \mathbf{H}(\theta(t))v(\bar{y}(t))$$

where $\mathbf{H}(\theta)$ is the Hessian of the Hamiltonian and where $v(y) = Uy$ is the optimal self-play vector field. Since the Hessian is strictly convex, the matrix $\mathbf{H}(\theta(t))$ is positive definite, so the centroid $\bar{y}(t)$ obeys an adaptive dynamics equation. Since $H(\theta)$ is strictly convex, the map between θ and $\nabla_\theta H(\theta) = \bar{y}(\theta)$ is a bijection. Therefore, the parameters obey an autonomous adaptive dynamics equation. \square

4.3.2 Metapopulations and Inhomogeneous Mixing

Real populations may not mix homogeneously. Consider a metapopulation model involving l patches. Patches need not represent spatial clusters of agents. Rather, we will assume that patches represent any partitioning of the total population that satisfies two conditions. First, payout is conditionally independent of patch membership given agent attributes. This can be accomplished trivially by adding all attributes used to assign patch membership (e.g. all covariates that influence interaction rates) into the list of attributes Ω , however this approach is semantic, since agents in different patches must be drawn from distributions with disjoint support. More practically, it means that the covariates responsible for determining interaction rate have no impact on the outcome of an interaction, conditional on a set of attributes Ω that *can be shared by agents in different patches*. Second, the interaction probabilities between agents is constant in agent type conditional on their patch membership.

Let $\pi_i(x, t)$ denote the density of type x at time t in patch i , and $P_i(t)$ the total population of the i^{th} patch. The population within a patch may change over time.

Suppose that individuals are paired at random, where the frequency of interactions between individuals in patch i with individuals in patch j , per unit time, is given by $m_{ij}P_i(t)P_j(t)$ for some $m_{ij} > 0$. Since the rate cannot depend on the ordering of the indices, the mixing matrix M must be symmetric. The diagonal entries are intra-patch mixing rates, and the off-diagonal entries are inter-patch mixing rates.

Suppose that the per capita growth rate of type x in patch i is proportional to the average total payout received by an individual of type x , in patch i , averaging over the interaction rate with each possible opponent, times a frequency dependent term $g(\pi)/\pi$. Then, the metapopulation will obey a block-structured, generalized replicator dynamic of the form:

$$\partial_t \pi_i(x, t) = \left[\sum_{j=1}^l m_{ij} \mathbb{E}_{X' \sim \pi_j(\cdot, t)}[f(x, X')] \right] g(\pi_i(x, t)) \quad (81)$$

This model collapses back to the homogeneous case if $m_{ij} = m_{lk}$ for all i, j, k, l , with interaction rate $\rho(P) = P$. If M is an identity, then it reduces to l independent populations.

Since the rates are all defined per-capita, or with growth rate function $g(\pi) = \mathcal{O}(\pi)$, the dynamic is support preserving within each patch. Then, dividing across by the density yields the equivalent dynamic:

$$\partial_t h(\pi_i(x, t)) = \sum_{j=1}^l m_{ij} \mathbb{E}_{X' \sim \pi_j(\cdot, t)}[f(x, X')] \quad (82)$$

where $h = \log$ when $g(\pi) = \pi$.

To make use of common embeddings, assume that for all $i \leq l$, $\pi_i(x, 0)$ share support S , and, without loss of generality, assume that $S = \Omega$. Then, parameterize the distribution in each patch separately:

$$\pi_i(x, t) = h^{-1}(u_0^{(i)}(x) + \theta^{(i)}(t) \cdot y(x)). \quad (83)$$

Theorem 7: Suppose that $\{\pi_i(\cdot, t)\}_{i=1}^l$ obey a metapopulation, generalized replicator dynamic of the form specified in equation (81), and each population is initialized with the same support. Then, the parameters $\theta = [\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(l)}]$ of the parameterization (83), obey a Hamiltonian dynamic of the form:

$$\frac{d}{dt} \theta(t) = (M \otimes U) \nabla_{\theta} H(\theta(t)) \quad (84)$$

where \otimes is the Kronecker product, and H is the Hamiltonian function:

$$H(\theta) = \sum_{i=1}^l H^{(i)}(\theta^{(i)}, u_0^{(i)}), \quad H^{(i)}(\theta^{(i)}) = \int_{x \in \Omega} \int_0^{\theta^{(i)}.y(x) + u_0^{(i)}(x)} h^{-1}(u). \quad (85)$$

Proof: The reduction to equation (84) follows the same steps used before. The Kronecker product is a bookkeeping artifact. It remains to show that the specified dynamic is Hamiltonian.

Any dynamic of the form $\frac{d}{dt} \theta = A \nabla_{\theta} H(\theta)$ is Hamiltonian if A is a square, real-valued, skew-symmetric matrix. Indeed, expanding A into its Schur form, then changing coordinates via the orthonormal matrices Q , recovers a block Hamiltonian dynamic of the kind we derived for the homogeneous population case. In this setting, block-diagonalizing the matrix $M \otimes U$ is unhelpful, since it introduces a linear transformation that mixes parameters assigned to separate patches.

So, to show that the dynamic is Hamiltonian it suffices to show that $M \otimes U$ is skew symmetric. This follows immediately since M is symmetric, and U is skew symmetric:

$$(M \otimes U)^T = M^T \otimes U^T = M \otimes (-U) = -(M \otimes U). \quad (86)$$

It follows that the parameter dynamics are Hamiltonian. \square

Since the metapopulation Hamiltonian is a sum of Hamiltonian functions that all share the same characterization when there exists, or does not exist, a fully mixed NE (when the origin is contained, or is not contained, inside the convex hull of Ψ), every qualitative result established before holds for the metapopulation extension. In particular,

Corollary 7.1: [Recurrence] The parameter dynamics either admit an isolated, interior equilibrium that corresponds to a fully mixed NE, or no such equilibrium exists. If an isolated interior equilibrium exists, then it is a center and the parameter dynamics are Poincaré recurrent. An isolated interior equilibrium exists if and only if the origin is contained inside the convex hull of Ψ .

Proof Outline: All arguments follow exactly as before. If such an equilibrium exists, then the Hamiltonian is a sum of radially unbounded, strictly convex functions, so is radially unbounded and strictly convex. In that case, the dynamics are volume preserving and bounded, so are recurrent. For details, see Appendix Section 6.6.7 \square

As in the homogeneous case, the metapopulation parameter dynamics can also be reduced by studying the collection of population centroids in the embedding space.

Corollary 7.2: [Mixed Adaptive Dynamics] Let $\bar{y}_i(t) = \mathbb{E}_{X \sim \pi_i(\cdot, t)}[y(X)]$. Let $\bar{y}(t) = [\bar{y}_1(t), \bar{y}_2(t), \dots, \bar{y}_l(t)]$ denote the concatenation of all the centroids, $\theta(\bar{y})$ the map from the centroid back to the parameters, and \mathbf{H} the Hessian of the Hamiltonian. Let $v(y) = Uy$ denote the optimal training vector field that directs self-play. Let $m_i = \sum_{j=1}^k m_j$ denote the total per capita interaction rate for patch i . Let $\hat{M} = \text{diag}(m_1^{-1}, m_2^{-1}, \dots, m_l^{-1})M$ denote the row-stochastic version of M . Then, the centroids obey an autonomous dynamic of the form:

$$\frac{d}{dt} \bar{y}_i(t) = m_i \mathbf{H}(\theta(\bar{y}(t))) v([\hat{M} \bar{y}(t)]_i). \quad (87)$$

Proof Outline: See Appendix Section 6.6.7. \square

Remark: In the standard case when $g(\pi) = \pi$, \mathbf{H} can be replaced with the raw second moment tensor $\mathbb{E}_{Y \sim \pi_y(\cdot; \theta(y))}[YY^T]$. In this case, the positive definite matrix entering the mixed adaptive dynamics equation is closely related to the population covariance in after embedding. \square

Remark: Equation (87) is to simultaneous gradient ascent as adaptive dynamics is to self-play. The right most term is the self-play gradient, or optimal training vector, evaluated at a weighted average of the centroids of each embedded population. The weights in the average differ by the focal patch i , and are specified by the terms in the mixing matrix. Since the optimal training vector field is linear, this is equal to a weighted average of the optimal training vectors associated with each patch. This averaged training vector is then multiplied by the covariance of the embedded population in the i^{th} patch to recover the rate of change of the i^{th} centroid. \square

The notation in equation (87) hides two nuances. First, the centroids are really the expected value of $y(X)$ when X is drawn from π_i times the population P_i . This is kept implicit since we adopted the definition that $\mathbb{E}_{X \sim \pi}[f(X)] = \int_x \pi(x)f(x)dx$ regardless the normalization of π . Second, each $\pi_i(\cdot; \theta)$ function is different depending on $\pi_i(x, 0)$. It follows that the inverse $\theta^{(i)}(\bar{y})$ is also a different function for each patch i .

Since the dynamic is autonomous, is uniquely defined up to the initial distributions in each patch, and since we can recover the parameters in every patch from their centroids, the metapopulation version of the replicator equation also admits an exact representation as simultaneous (mixed) adaptive dynamics equation in the embedded centroids, for any payout function f .

4.3.3 Discussion

Theorems 6 and 7, and their corollaries, show that the essential qualitative description of evolutionary population dynamics in symmetric, zero-sum games, as systems of coupled oscillators, holds beyond the special cases of self-play, adaptive dynamics, simultaneous gradient ascent, or the standard replicator equation. This characterization does not require homogeneous mixing, nor does it require that per capita growth rates are strictly proportional to expected payouts. Indeed, generic no-regret learning dynamics are Hamiltonian, and Poincaré recurrent, in zero sum, symmetric games with fully mixed equilibria (c.f. [20, 32, 72, 125, 144]). Our results are consistent with this more general collection of theories. They differ in the use of the disc game embedding to recover a useful parameterization of densities, to provide a geometric condition distinguishing recurrent dynamics from iterated dominance dynamics, and a generic, yet explicit conversion to adaptive dynamics in a coordinate system intrinsic to the game.

The generality of these conclusions in the setting of symmetric, two-player, zero-sum games suggests that the key phenomena, iterated domination followed by recurrent dynamics, are essential properties of the class of games, not the specification of the learning dynamic. All of these dynamics are Hamiltonian because, after an appropriate transformation, the right hand side of the dynamic is a skew-symmetric linear transformation of an input that depends on the current population. In all cases, the skew symmetry of the transformation was a direct consequence of the skew-symmetry of the payout, and would not hold if the payout was not skew-symmetric, that is, given a variable-sum game.

A paradigm should provide incisive intuition within a clearly defined class of problems. A paradigm may be misused if it is unclear which problems it does not address. Barring adversarially constructed games with infinite rank, and infinite variance under all continuously distributed populations, a disc game embedding exists if and only if the underlying game is skew symmetric. This symmetry is the only substantive restriction on its existence. The results illustrated here suggest that the embedding provides qualitatively accurate intuition for a broad class of problems that share the same sharply defined scope: learning dynamics in symmetric, two-player, *constant-sum* games. The embedding does not exist for variable-sum games, where its misapplication as a mental guide may not provide accurate intuition.

4.4 Efficient Simulation Leveraging the Latent Space

The latent space form, (46), offers promising new avenues for simulating the continuous replicator equation, (26).

Direct simulation of the continuous replicator equation is expensive since, to evaluate the rate of change in the population density at each supported x requires performing an integral over the full support of the population. Let t denote the number of traits. Then the cost of evaluating the rate of change of the density at any x will grow exponentially in t . This exponential cost is incurred at every quadrature node where the density is computed. So, for general f , the cost for direct implementation of the continuous replicator equation in the original trait space is exponential in $2t$. Consequently, even relatively low dimensional settings exceed reasonable simulation budgets.

If the rank, r , of f is less than t , then simulation in the latent space may be cheaper. What is the cost of simulating the parameter dynamics, $\frac{d}{dt}\theta = U\nabla_\theta P(\theta)$?

First, at each time step, we only need to update r parameters to update the density. In contrast, without a functional form, we need to update the density at each quadrature node explicitly. Second, to update each parameter, we need to evaluate the partial derivative of $P(\theta)$ with respect to one parameter. In total, the cost per time step is equivalent to computing $\nabla_\theta P(\theta)$, then permuting and negating the entries of the gradient.

Recall that $\nabla_\theta P(\theta) = \mathbb{E}_{Y \sim \pi_Y(\cdot; \theta)}[Y]$ where $\pi_Y(y; \theta) = \exp(\theta \cdot y)\pi_Y(y, 0)$ and $\pi_Y(y, 0)$ is the distribution of $y(X)$ for $X \sim \pi(x, 0)$. The cost of evaluating the expectation of each coordinate is equivalent to the cost of performing quadrature in the parameter space. Quadrature incurs an exponential cost in r . Therefore, instead of paying an exponential cost in $2r$, we pay an exponential cost in r , times r . The required integrals can be all conveniently collected into a single quadrature call using an auto-differentiation routine.

All that remains is evaluation of the base measure $\pi_Y(y; 0)$ given by embedding the initial distribution $\pi(x; 0)$, or some other base measure in x . The population distribution in the embedding space is a simple exponential tilting of the base measure pushed forward by the embedding. Since the base measure remains fixed, evaluating the base measure is a one-time cost. The additional cost to push the base measure forward is distributed across each time step.

So, disc game embedding promises a simple method for fast simulation of the continuous replicator equation whenever the rank r is sufficiently small. Push the chosen base measure forward. Then, evolve equation (46) using:

1. A symplectic integrator for each time step,
2. An auto-differentiator to evaluate $\nabla_\theta P(\theta)$ from a single evaluation of $P(\theta)$, and
3. A chosen quadrature scheme for evaluating $P(\theta)$.

Notice that only the push forward requires integrals over the original t dimensional space. Thus, after pushing forward, the simulation cost is essentially independent of t . As a back of the envelope comparison, an n -step simulation in the original space has cost $\mathcal{O}(\exp(2t))$ per time step, while simulation in the latent space has cost $\mathcal{O}(\exp(t+r))/n + \mathcal{O}(r\exp(r))$. So, in the limit of large n , simulation in the latent space can be performed at a cost that is entirely independent of the dimension of the original space. We leave detailed numerical development and testing to separate work.

In practice, it is more stable to apply the same algorithm using $\log(P(\theta))$ as the Hamiltonian, then scaling time by the initial population size to recover the desired dynamic.

5 Conclusions

This paper offers three contributions. First, we have established an axiomatic justification for disc-game embedding that motivate its use for studying learning dynamics. Second, we have shown that most skew-symmetric games admit a disc-game embedding and that the embedding allows accurate approximation for essentially all games. All previous work leveraging disc game embeddings either assumed a bilinear payout function or only applied the embedding pointwise to sampled data [21, 22, 23, 206]. The generality and interpretability of the embedding recommend its use as a mental model for evolution subject to a zero-sum game.

Third, we have shown that the intuition provided by disc-game embedding accurately characterizes an important class of learning dynamics in zero-sum games. Specifically, the disc-game embedding functions offer a useful function basis when studying generalized, continuous replicator dynamics in populations that could mix inhomogeneously. This function basis offers analytical insight by identifying a normal form dynamic and speeds computation. Using the disc-game embedding basis, we derived the necessary conditions for exact finite-dimensional closure, showed that the number of dynamical degrees of freedom needed corresponds to the rank of the payout, and derived two intuitive forms for the associated parameter dynamics. These results completely generalize the restricted parameteric solutions presented in [43, 53, 56]. We showed that the parameter dynamics are highly structured, so admit strong global analysis. In particular, the replicator dynamic reduces to a Hamiltonian system of coupled oscillators, driven by adaptive dynamics in the embedding space. These results directly recover parallel results for finite strategy sets, such as Poincaré recurrence in the presence of a fully mixed equilibria [32, 144, 178]. In addition, they establish an exact correspondence between replicator/multiplicative weight dynamics and adaptive dynamics in the disc game space.

For the reasons demonstrated in this paper, we recommend disc-game embedding as a canonical latent space representation for populations evolving in response to a symmetric, zero-sum, game.

Acknowledgements

The authors thank Jade Sheng for her assistance creating schematic figures, as well as Stefano Allesina and Mary Silber for their insightful support. The authors also thank Yucong Liu for his contributions applying the Hilbert-Schmidt theorem to games.

References

- [1] K. C. ABBOTT, Theoretical Ecology: Concepts and Applications, Oxford University Press, Oxford, 2020, ch. 4, pp. 40–52.
- [2] H. ABDI AND L. J. WILLIAMS, Principal component analysis, Wiley interdisciplinary reviews: computational statistics, 2 (2010), pp. 433–459.
- [3] P. A. ABRAMS, Modelling the adaptive dynamics of traits involved in inter-and intraspecific interactions: an assessment of three methods, Ecology Letters, 4 (2001), pp. 166–175.
- [4] A. K. AGOGINO AND K. TUMER, Unifying temporal and structural credit assignment problems, in Autonomous agents and multi-agent systems conference, 2004.

- [5] E. AKIN, Domination or equilibrium, Mathematical Biosciences, 50 (1980), pp. 239–250.
- [6] ———, The geometry of population genetics, vol. 31, Springer Science & Business Media, 2013.
- [7] E. AKIN AND V. LOSERT, Evolutionary dynamics of zero-sum games, Journal of mathematical biology, 20 (1984), pp. 231–258.
- [8] R. ANDERSON AND R. MAY, Infectious diseases of humans, 1991.
- [9] J. APALOO, Revisiting strategic models of evolution: The concept of neighborhood invader strategies, Theoretical Population Biology, 52 (1997), pp. 71–77.
- [10] V. I. ARNOLD, Geometrical methods in the theory of ordinary differential equations, vol. 250, Springer Science & Business Media, 2012.
- [11] ———, Mathematical methods of classical mechanics, vol. 60, Springer Science & Business Media, 2013.
- [12] S. ARORA, E. HAZAN, AND S. KALE, The multiplicative weights update method: a meta-algorithm and applications, Theory of computing, 8 (2012), pp. 121–164.
- [13] A. A. ARRIETA, Professionals don't always play minimax: evidence from latin american soccer leagues, Cuadernos de Economía, 43 (2020), pp. 305–324.
- [14] D. ARUMUGAM, P. HENDERSON, AND P.-L. BACON, An information-theoretic perspective on credit assignment in reinforcement learning, arXiv preprint arXiv:2103.06224, (2021).
- [15] P. AUER, N. CESA-BIANCHI, Y. FREUND, AND R. E. SCHAPIRE, Gambling in a rigged casino: The adversarial multi-armed bandit problem, in Proceedings of IEEE 36th annual foundations of computer science, IEEE, 1995, pp. 322–331.
- [16] R. AUMANN, Nash equilibria are not self-enforcing, Economic decision making: Games, econometrics and optimisation, (1990), pp. 201–206.
- [17] R. AXELROD AND W. D. HAMILTON, The evolution of cooperation, science, 211 (1981), pp. 1390–1396.
- [18] O. H. AZAR AND M. BAR-ELI, Do soccer players play the mixed-strategy nash equilibrium?, Applied Economics, 43 (2011), pp. 3591–3601.
- [19] Y. BABICHENKO AND A. RUBINSTEIN, Communication complexity of approximate nash equilibria, in Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, 2017, pp. 878–889.
- [20] J. P. BAILEY AND G. PILIOURAS, Multi-agent learning in network zero-sum games is a hamiltonian system, arXiv preprint arXiv:1903.01720, (2019).
- [21] D. BALDUZZI, M. GARNELO, Y. BACHRACH, W. CZARNECKI, J. PEROLAT, M. JADERBERG, AND T. GRAEPEL, Open-ended learning in symmetric zero-sum games, in International Conference on Machine Learning, PMLR, 2019, pp. 434–443.
- [22] D. BALDUZZI, S. RACANIÈRE, J. MARTENS, J. FOERSTER, K. TUYLS, AND T. GRAEPEL, The mechanics of n-player differentiable games, in International Conference on Machine Learning, PMLR, 2018, pp. 354–363.
- [23] D. BALDUZZI, K. TUYLS, J. PEROLAT, AND T. GRAEPEL, Re-evaluating evaluation, Advances in Neural Information Processing Systems, 31 (2018).
- [24] L. BARREIRA, Poincaré recurrence: old and new, in XIVth International Congress on Mathematical Physics, World Scientific, 2006, pp. 415–422.
- [25] A. BARVINOK, A course in convexity, vol. 54, American Mathematical Society, 2025.
- [26] J. L. BEACHAM, The relative importance of body size and aggressive experience as determinants of dominance in pumpkinseed sunfish, Lepomis gibbosus, Animal Behaviour, 36 (1988), pp. 621–623.
- [27] Y. BENGIO, A. COURVILLE, AND P. VINCENT, Representation learning: A review and new perspectives, IEEE transactions on pattern analysis and machine intelligence, 35 (2013), pp. 1798–1828.
- [28] A. BERLINET AND C. THOMAS-AGNAN, Reproducing kernel Hilbert spaces in probability and statistics, Springer Science & Business Media, 2011.

- [29] K. BINMORE AND L. SAMUELSON, Muddling through: Noisy equilibrium selection, Journal of Economic Theory, 74 (1997), pp. 235–265.
- [30] D. BISHOP AND C. CANNINGS, A generalized war of attrition, Journal of theoretical biology, 70 (1978), pp. 85–124.
- [31] V. I. BOGACHEV AND O. G. SMOLYANOV, Real and functional analysis, Springer, 2020.
- [32] V. BOONE AND G. PILIOURAS, From darwin to poincaré and von neumann: Recurrence and cycles in evolutionary and algorithmic game theory, in Web and Internet Economics: 15th International Conference, WINE 2019, New York, NY, USA, December 10–12, 2019, Proceedings 15, Springer, 2019, pp. 85–99.
- [33] T. BÖRGERS AND R. SARIN, Learning through reinforcement and replicator dynamics, Journal of economic theory, 77 (1997), pp. 1–14.
- [34] D. S. BOUKAL AND L. BEREC, Single-species models of the allee effect: extinction boundaries, sex ratios and mate encounters, Journal of theoretical biology, 218 (2002), pp. 375–394.
- [35] A. BRONDSTED, An introduction to convex polytopes, vol. 90, Springer Science & Business Media, 2012.
- [36] J. N. BROWN AND R. W. ROSENTHAL, Testing the minimax hypothesis: A re-examination of o’neill’s game experiment, Econometrica: Journal of the Econometric Society, (1990), pp. 1065–1081.
- [37] N. BROWN AND T. SANDHOLM, Superhuman ai for multiplayer poker, Science, 365 (2019), pp. 885–890.
- [38] A. CABRALES, Stochastic replicator dynamics, International Economic Review, 41 (2000), pp. 451–481.
- [39] J. CAIRNS, F. BORSE, T. MONONEN, T. HILTUNEN, AND V. MUSTONEN, Strong selective environments determine evolutionary outcome in time-dependent fitness seascapes, Evolution Letters, 6 (2022), pp. 266–279.
- [40] O. CANDOGAN, I. MENACHE, A. OZDAGLAR, AND P. A. PARRILO, Flows and decompositions of games: Harmonic and potential games, Mathematics of Operations Research, 36 (2011), pp. 474–503.
- [41] C. CARATHÉODORY, Über den Wiederkehrsatz von Poincaré, De Gruyter., 1919.
- [42] C. CEBRA AND A. STRANG, Similarity suppresses cyclicity: Why similar competitors form hierarchies, SIAM Journal on Applied Mathematics, 83 (2023), pp. 2027–2051.
- [43] ———, The almost sure evolution of hierarchy among similar competitors, arXiv preprint arXiv:2402.06005, (2024).
- [44] H. S. CHAN AND E. BORNBERG-BAUER, Perspectives on protein evolution from simple, Applied bioinformatics, 50 (2002), pp. 121–144.
- [45] C. CHANG, C. Y. LI, R. L. EARLEY, AND Y. HSU, Aggression and related behavioral traits: The impact of winning and losing and the role of hormones, Integrative and Comparative Biology, 52 (2012), pp. 801–813.
- [46] V. CHAZAN-PANTZALIS, Sports analytics algorithms for performance prediction, na, (2020).
- [47] X. CHEN, X. DENG, AND S.-H. TENG, Settling the complexity of computing two-player nash equilibria, Journal of the ACM (JACM), 56 (2009), pp. 1–57.
- [48] P.-A. CHIAPPORI, S. LEVITT, AND T. GROSECLOSE, Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer, American Economic Review, 92 (2002), pp. 1138–1151.
- [49] S. Y. CHONG, J. HUMBLE, G. KENDALL, J. LI, X. YAO, ET AL., The iterated prisoner’s dilemma: 20 years on, Advances in Natural Competition, 4 (2007), pp. 1–22.
- [50] K. CLARK, S. KAY, AND M. SEFTON, When are nash equilibria self-enforcing? an experimental analysis, International Journal of Game Theory, 29 (2001), pp. 495–515.
- [51] R. R. COIFMAN AND S. LAFON, Diffusion maps, Applied and computational harmonic analysis, 21 (2006), pp. 5–30.
- [52] F. COURCHAMP, T. CLUTTON-BROCK, AND B. GRENFELL, Inverse density dependence and the allee effect, Trends in ecology & evolution, 14 (1999), pp. 405–410.
- [53] R. CRESSMAN, Coevolution, adaptive dynamics, and the replicator equation for a single species with a continuous trait space, Proceedings, International Society of Dynamic Games, Tucson, Arizona, USA, (2004).
- [54] ———, Dynamic stability of the replicator equation with continuous strategy space, na, (2004).

- [55] R. CRESSMAN AND J. HOFBAUER, Measure dynamics on a one-dimensional continuous trait space: theoretical foundations for adaptive dynamics, *Theoretical Population Biology*, 67 (2005), pp. 47–59.
- [56] ———, Measure dynamics on a one-dimensional continuous trait space: Theoretical foundations for adaptive dynamics, *Theoretical Population Biology*, 67 (2005), pp. 47–59.
- [57] R. CRESSMAN, J. HOFBAUER, AND F. RIEDEL, Stability of the replicator equation for a single species with a multi-dimensional continuous trait space, *Journal of Theoretical Biology*, 239 (2006), pp. 273–288.
- [58] R. CRESSMAN AND Y. TAO, The replicator equation and other game dynamics, *PNAS*, 111 (2014), pp. 10810–10817.
- [59] C. DASKALAKIS, P. W. GOLDBERG, AND C. H. PAPADIMITRIOU, The complexity of computing a nash equilibrium, *Communications of the ACM*, 52 (2009), pp. 89–97.
- [60] J. A. G. DE VISSER, S. F. ELENA, I. FRAGATA, AND S. MATUSZEWSKI, The utility of fitness landscapes and big data for predicting evolution, *Heredity*, 121 (2018), pp. 401–405.
- [61] J. A. G. DE VISSER AND J. KRUG, Empirical fitness landscapes and the predictability of evolution, *Nature Reviews Genetics*, 15 (2014), pp. 480–490.
- [62] U. DIECKMANN AND R. LAW, The dynamical theory of coevolution: a derivation from stochastic ecological processes, *Journal of mathematical biology*, 34 (1996), pp. 579–612.
- [63] A. C. DIXON, The Elementary Properties of the Elliptic Functions: With Examples, Macmillan & Company, 1894.
- [64] M. DOEBELI AND G. D. RUXTON, Evolution of dispersal rates in metapopulation models: branching and cyclic dynamics in phenotype space, *Evolution*, 51 (1997), pp. 1730–1741.
- [65] T. DOHMHEN AND H. SONNABEND, Further field evidence for minimax play, *Journal of sports economics*, 19 (2018), pp. 371–388.
- [66] C. ECKART AND G. YOUNG, The approximation of one matrix by another of lower rank, *Psychometrika*, 1 (1936), pp. 211–218.
- [67] B. EFRON, The geometry of exponential families, *The Annals of Statistics*, (1978), pp. 362–376.
- [68] ———, Why isn't everyone a bayesian?, *The American Statistician*, 40 (1986), pp. 1–5.
- [69] ———, Exponential families in theory and practice, Cambridge University Press, 2022.
- [70] I. ESHEL, Evolutionary and continuous stability, *Journal of theoretical Biology*, 103 (1983), pp. 99–111.
- [71] J. FERRET, R. MARINIER, M. GEIST, AND O. PIETQUIN, Self-attentional credit assignment for transfer in reinforcement learning, arXiv preprint arXiv:1907.08027, (2019).
- [72] L. FLOKAS, E. V. VLATAKIS-GKARAGKOUNIS, T. LIANEAS, P. MERTIKOPOULOS, AND G. PILIOURAS, No-regret learning and mixed nash equilibria: They do not mix, arXiv preprint arXiv:2010.09514, (2020), p. 10.
- [73] M. M. FLOOD, Some experimental games, *Management Science*, 5 (1958), pp. 5–26.
- [74] D. FOSTER AND P. YOUNG, Stochastic evolutionary game dynamics, *Theoretical population biology*, 38 (1990), pp. 219–232.
- [75] V. FRANÇOIS-LAVET, Y. BENGIO, D. PRECUP, AND J. PINEAU, Combined reinforcement learning via abstract representations, in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 3582–3589.
- [76] H. FRAUENFELDER, A. BISHOP, A. GARCIA, AND A. S. PERELSON, P., sherrington, d. and swart, pj (eds.), landscape paradigms in physics and biology, *Physica*, 107 (1997), pp. 117–436.
- [77] D. FRIEDMAN, Evolutionary games in economics, *Econometrica: journal of the econometric society*, (1991), pp. 637–666.
- [78] D. FUDENBERG AND D. LEVINE, Learning in games, *European economic review*, 42 (1998), pp. 631–639.
- [79] R. GAURIOT, L. PAGE, AND J. WOODERS, Nash at wimbledon: evidence from half a million serves, Available at SSRN 2850919, (2016).

- [80] D. GENTNER AND M. JEZIORSKI, The shift from metaphor to analogy in western science, *Metaphor and thought*, 447 (1993).
- [81] G. T. GILBERT AND R. L. HATCHER, Wagering in final jeopardy!, *Mathematics Magazine*, 67 (1994), pp. 268–277.
- [82] I. GOODFELLOW, J. POUGET-ABADIE, M. MIRZA, B. XU, D. WARDE-FARLEY, S. OZAIR, A. COURVILLE, AND Y. BENGIO, Generative adversarial nets, *Advances in neural information processing systems*, 27 (2014).
- [83] C. P. GOULART, M. MAHMUDI, K. A. CRONA, S. D. JACOBS, M. KALLMANN, B. G. HALL, D. C. GREENE, AND M. BARLOW, Designing antibiotic cycling strategies by determining and understanding local adaptive landscapes, *PloS one*, 8 (2013), p. e56040.
- [84] S. J. GOULD AND R. C. LEWONTIN, The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptionist programme, *Proceedings of the Royal Society B: Biological Sciences*, 205 (1979).
- [85] A. GREENWALD AND A. JAFARI, A general class of no-regret learning algorithms and game-theoretic equilibria, in *Learning Theory and Kernel Machines: 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003*, Washington, DC, USA, August 24-27, 2003. *Proceedings*, Springer, 2003, pp. 2–12.
- [86] M. P. HALEY, C. J. DEUTSCH, AND B. J. LE BOEUF, Size, dominance and copulatory success in male northern elephant seals, Mirounga angustirostris, *Animal Behaviour*, 48 (1994), pp. 1249–1260.
- [87] B. G. HALL, Predicting evolution by in vitro evolution requires determining evolutionary pathways, *Antimicrobial agents and chemotherapy*, 46 (2002), pp. 3035–3038.
- [88] K. M. HALL, An r-dimensional quadratic placement algorithm, *Management science*, 17 (1970), pp. 219–229.
- [89] W. R. HAMILTON, On a General Method of Expressing the Paths of Light, & of the Planets, by the Coefficients of a Characteristic Function, PD Hardy, 1833.
- [90] J. HANNAN, Approximation to bayes risk in repeated play, *Contributions to the Theory of Games*, 3 (1957), pp. 97–139.
- [91] I. C. HARDY, Sex ratios, 2002.
- [92] S. HART AND Y. MANSOUR, How long to equilibrium? the communication complexity of uncoupled equilibrium procedures, *Games and Economic Behavior*, 69 (2010), pp. 107–126.
- [93] S. HART AND A. MAS-COLELL, A simple adaptive procedure leading to correlated equilibrium, *Econometrica*, 68 (2000), pp. 1127–1150.
- [94] ———, Uncoupled dynamics do not lead to nash equilibrium, *American Economic Review*, 93 (2003), pp. 1830–1836.
- [95] ———, Stochastic uncoupled dynamics and nash equilibrium, *Games and economic behavior*, 57 (2006), pp. 286–303.
- [96] P. HARTMAN, A lemma in the theory of structural stability of differential equations, *Proceedings of the American Mathematical Society*, 11 (1960), pp. 610–620.
- [97] ———, On local homeomorphisms of euclidean spaces, *Bol. Soc. Mat. Mexicana*, 5 (1960), pp. 220–241.
- [98] A. HASTINGS, Transient dynamics and persistence of ecological systems, *Ecology Letters*, 4 (2001), pp. 215–220.
- [99] ———, Transients: the key to long-term ecological understanding?, *Trends in ecology & evolution*, 19 (2004), pp. 39–45.
- [100] A. HASTINGS, K. C. ABBOTT, K. CUDDINGTON, T. FRANCIS, G. GELLNER, Y.-C. LAI, A. MOROVZOV, S. PETROVSKII, K. SCRANTON, AND M. L. ZEEMAN, Transient phenomena in ecology, *Science*, 361 (2018), p. eaat6412.
- [101] J. HOFBAUER, On the occurrence of limit cycles in the volterra-lotka equation, *Nonlinear Analysis: Theory, Methods & Applications*, 5 (1981), pp. 1003–1007.
- [102] ———, Evolutionary dynamics for bimatrix games: A hamiltonian system?, *Journal of mathematical biology*, 34 (1996), pp. 675–688.
- [103] J. HOFBAUER AND K. SIGMUND, Evolutionary games and population dynamics, Cambridge university press, 1998.
- [104] ———, Evolutionary game dynamics, *Bulletin of the American Mathematical Society*, 40 (2003), pp. 479–519.

- [105] C. A. HOLT AND A. E. ROTH, The nash equilibrium: A perspective, Proceedings of the National Academy of Sciences, 101 (2004), pp. 3999–4002.
- [106] A. I. HOUSTON AND J. M. McNAMARA, Models of adaptive behaviour: an approach based on state, Cambridge University Press, 1999.
- [107] L. A. IMHOF, The long-run behavior of the stochastic replicator dynamics, The Annals of Applied Probability, 15 (2005), pp. 1019–1045.
- [108] N. IMMORLICA, A. T. KALAI, B. LUCIER, A. MOITRA, A. POSTLEWAITE, AND M. TENNENHOLTZ, Dueling algorithms, in Proceedings of the forty-third annual ACM symposium on Theory of computing, 2011, pp. 215–224.
- [109] Y. IWASA AND A. POMIANKOWSKI, Continual change in mate preferences, Nature, 377 (1995), pp. 420–422.
- [110] W. M. JACKSON AND R. L. WINNEGARD, Linearity in dominance hierarchies: a second look at the individual attributes model, Animal Behaviour, (1988).
- [111] M. KANDORI, G. J. MAILATH, AND R. ROB, Learning, mutation, and long run equilibria in games, Econometrica: Journal of the Econometric Society, (1993), pp. 29–56.
- [112] K. KAPADIA, H. ABDEL-JABER, F. THABTAH, AND W. HADI, Sport analytics for cricket game results using machine learning: An experimental study, Applied Computing and Informatics, 18 (2022), pp. 256–266.
- [113] G. P. KAREV, Replicator equations and the principle of minimal production of information, Bulletin of mathematical biology, 72 (2010), pp. 1124–1142.
- [114] A. KARVONEN, Emergent world models and latent variable estimation in chess-playing language models, arXiv preprint arXiv:2403.15498, (2024).
- [115] B. KERR, M. A. RILEY, M. W. FELDMAN, AND B. J. BOHANNAN, Local dispersal promotes biodiversity in a real-life game of rock–paper–scissors, Nature, 418 (2002), pp. 171–174.
- [116] D. P. KINGMA, M. WELLING, ET AL., An introduction to variational autoencoders, Foundations and Trends® in Machine Learning, 12 (2019), pp. 307–392.
- [117] Y. KOREN, On spectral graph drawing, in International Computing and Combinatorics Conference, Springer, 2003, pp. 496–508.
- [118] R. D. KOUYOS, G. E. LEVENTHAL, T. HINKLEY, M. HADDAD, J. M. WHITCOMB, C. J. PETROPOULOS, AND S. BONHOEFFER, Exploring the complexity of the hiv-1 fitness landscape, PLoS genetics, 8 (2012), p. e1002551.
- [119] I. KOVACIC, L. CVETICANIN, M. ZUKOVIC, AND Z. RAKARIC, Jacobi elliptic functions: a review of nonlinear oscillatory application problems, Journal of Sound and Vibration, 380 (2016), pp. 1–36.
- [120] S. A. KOVALCHIK, Searching for the goat of tennis win prediction, Journal of Quantitative Analysis in Sports, 12 (2016), pp. 127–138.
- [121] K. KOVASH AND S. D. LEVITT, Professionals do not play minimax: Evidence from major league baseball and the national football league, tech. rep., National Bureau of Economic Research, 2009.
- [122] J. B. KRUSKAL AND M. WISH, Multidimensional scaling, Sage, 1978.
- [123] T. S. KUHN, The structure of scientific revolutions, vol. 962, University of Chicago press Chicago, 1997.
- [124] T. LATTIMORE AND C. SZEPESVÁRI, Bandit algorithms, Cambridge University Press, 2020.
- [125] D. LEGACCI, P. MERTIKOPOULOS, AND B. PRADELSKI, A geometric decomposition of finite games: Convergence vs. recurrence under exponential weights, in ICML 2024-41st International Conference on Machine Learning, 2024.
- [126] O. LEIMAR, Multidimensional convergence stability and the canonical adaptive dynamics, Elements of adaptive dynamics, (2005), pp. 117–128.
- [127] S. D. LEVITT, J. A. LIST, AND D. H. REILEY, What happens in the field stays in the field: Exploring whether professionals play minimax in laboratory experiments, Econometrica, 78 (2010), pp. 1413 – 1434.
- [128] K. LI, A. K. HOPKINS, D. BAU, F. VIÉGAS, H. PFISTER, AND M. WATTENBERG, Emergent world representations: Exploring a sequence model trained on a synthetic task, arXiv preprint arXiv:2210.13382, (2022).

- [129] J. LIOUVILLE, Note sur la théorie de la variation des constantes arbitraires, Journal de mathématiques pures et appliquées, 3 (1838), pp. 342–349.
- [130] N. LITTLESTONE AND M. K. WARMUTH, The weighted majority algorithm, Information and computation, 108 (1994), pp. 212–261.
- [131] A. J. LOTKA, Analytical note on certain rhythmic relations in organic systems, Proceedings of the National Academy of Sciences, 6 (1920), pp. 410–415.
- [132] ———, Undamped oscillations derived from the law of mass action., Journal of the american chemical society, 42 (1920), pp. 1595–1599.
- [133] A. M. LYAPUNOV, The general problem of the stability of motion, International journal of control, 55 (1992), pp. 531–534.
- [134] R. MACARTHUR AND R. LEVINS, The limiting similarity, convergence, and divergence of coexisting species, The american naturalist, 101 (1967), pp. 377–385.
- [135] R. MACKAY AND J. MEISS, Survey of hamiltonian dynamics, in Hamiltonian Dynamical Systems, CRC Press, 2020, pp. 3–19.
- [136] R. MACQUEEN, A proof that coarse correlated equilibrium implies nash equilibrium in two-player zero-sum games, arXiv preprint arXiv:2304.07187, (2023).
- [137] G. J. MAILATH, Do people play nash equilibrium? lessons from evolutionary game theory, Journal of Economic Literature, 36 (1998), pp. 1347–1374.
- [138] R. M. MAY, Stability and complexity in model ecosystems, Princeton university press, 2019.
- [139] K. A. MCCABE, A. MUKHERJI, AND D. E. RUNKLE, An experimental study of information and mixed-strategy play in the three-person matching-pennies game, Economic Theory, 15 (2000), pp. 421–462.
- [140] M. McCARTHY, The allee effect, finding mates and theoretical models, Ecological Modelling, 103 (1997), pp. 99–102.
- [141] B. J. MCGILL AND J. S. BROWN, Evolutionary game theory and adaptive dynamics of continuous traits, Annu. Rev. Ecol. Evol. Syst., 38 (2007), pp. 403–435.
- [142] R. MEHTA, Constant rank bimatrix games are ppad-hard, in Proceedings of the forty-sixth annual ACM symposium on Theory of computing, 2014, pp. 545–554.
- [143] D. J. MERRELL, The adaptive seascape: the mechanism of evolution, U of Minnesota Press, 1994.
- [144] P. MERTIKOPOULOS, C. PAPADIMITRIOU, AND 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.
- [145] P. MERTIKOPOULOS AND W. H. SANDHOLM, Riemannian game dynamics, Journal of Economic Theory, 177 (2018), pp. 315–364.
- [146] G. MESZNA, E. KISDI, U. DIECKMANN, S. A. GERITZ, AND J. A. METZ, Evolutionary optimisation models and matrix games in the unified perspective of adaptive dynamics, Selection, 2 (2002), pp. 193–220.
- [147] M. MINSKY, Steps toward artificial intelligence, Proceedings of the IRE, 49 (1961), pp. 8–30.
- [148] L. MIRSKY, Symmetric gauge functions and unitarily invariant norms, The quarterly journal of mathematics, 11 (1960), pp. 50–59.
- [149] D. W. MOCK AND G. A. PARKER, The evolution of sibling rivalry, Oxford University Press, 1997.
- [150] D. MONDERER AND L. S. SHAPLEY, Potential games, Games and economic behavior, 14 (1996), pp. 124–143.
- [151] D. MOOKHERJEE AND B. SOPHER, Learning behavior in an experimental matching pennies game, Games and Economic Behavior, 7 (1994), pp. 62–91.
- [152] P. A. P. MORAN, Random processes in genetics, in Mathematical proceedings of the cambridge philosophical society, vol. 54, Cambridge University Press, 1958, pp. 60–71.

- [153] M. MORAVČÍK, M. SCHMID, N. BURCH, V. LISÝ, D. MORRILL, N. BARD, T. DAVIS, K. WAUGH, M. JOHANSON, AND M. BOWLING, Deepstack: Expert-level artificial intelligence in heads-up no-limit poker, *Science*, 356 (2017), pp. 508–513.
- [154] J. MOSER, Stable and random motions in dynamical systems: With special emphasis on celestial mechanics, vol. 1, Princeton university press, 2001.
- [155] S. A. MUÑOZ-GÓMEZ, G. BILOLIKAR, J. G. WIDEMAN, AND K. GEILER-SAMEROTTE, Constructive neutral evolution 20 years later, *Journal of molecular evolution*, 89 (2021), pp. 172–182.
- [156] V. MUSTONEN AND M. LÄSSIG, From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation, *Trends in genetics*, 25 (2009), pp. 111–119.
- [157] ———, Fitness flux and ubiquity of adaptive evolution, *Proceedings of the National Academy of Sciences*, 107 (2010), pp. 4248–4253.
- [158] M. NAGHNAEIAN AND D. DEL VECCHIO, Robust moment closure method for the chemical master equation, in *2017 IEEE Conference on Control Technology and Applications (CCTA)*, IEEE, 2017, pp. 967–972.
- [159] J. F. NASH JR, Equilibrium points in n-person games, *Proceedings of the national academy of sciences*, 36 (1950), pp. 48–49.
- [160] J. NEWTON, Evolutionary game theory: A renaissance, *Games*, 9 (2018), p. 31.
- [161] A. NICHOLSON, V. BAILEY, ET AL., The balance of animal populations, na, (1935).
- [162] M. A. NOWAK AND K. SIGMUND, Evolutionary dynamics of biological games, *science*, 303 (2004), pp. 793–799.
- [163] J. OCHS, Games with unique, mixed strategy equilibria: An experimental study, *Games and Economic Behavior*, 10 (1995), pp. 202–217.
- [164] J. OECHSSLER AND F. RIEDEL, Evolutionary dynamics on infinite strategy spaces, *Economic theory*, 17 (2001), pp. 141–162.
- [165] ———, On the dynamic foundation of evolutionary stability in continuous models, *Journal of Economic Theory*, 107 (2002), pp. 223–252.
- [166] T. OHTA, The nearly neutral theory of molecular evolution, *Annual review of ecology and systematics*, (1992), pp. 263–286.
- [167] B. O’NEILL, Nonmetric test of the minimax theory of two-person zerosum games., *Proceedings of the national academy of sciences*, 84 (1987), pp. 2106–2109.
- [168] J. OTWINOWSKI AND I. NEMENMAN, Genotype to phenotype mapping and the fitness landscape of the e. coli lac promoter, *PloS one*, 8 (2013), p. e61570.
- [169] K. M. PAGE AND M. A. NOWAK, Unifying evolutionary dynamics, *Journal of theoretical biology*, 219 (2002), pp. 93–98.
- [170] I. PALACIOS-HUERTA, Professionals play minimax, *The Review of Economic Studies*, 70 (2003), pp. 395–415.
- [171] I. PALACIOS-HUERTA AND O. VOLJI, Professionals play minimax in laboratory experiments, *Econometrica*, 76 (2008), pp. 71–115.
- [172] A. C. PALMER AND R. KISHONY, Understanding, predicting and manipulating the genotypic evolution of antibiotic resistance, *Nature Reviews Genetics*, 14 (2013), pp. 243–248.
- [173] V. I. PAULSEN AND M. RAGHUPATHI, An introduction to the theory of reproducing kernel Hilbert spaces, vol. 152, Cambridge university press, 2016.
- [174] K. PEARSON, Liii. on lines and planes of closest fit to systems of points in space, *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, 2 (1901), pp. 559–572.
- [175] A. M. PERELOMOV, Integrable systems of classical mechanics and lie algebras, na, (1990).
- [176] T. PFEIFFER, S. SCHUSTER, AND S. BONHOEFFER, Cooperation and competition in the evolution of atp-producing pathways, *Science*, 292 (2001), pp. 504–507.

- [177] E. PIGNATELLI, J. FERRET, M. GEIST, T. MESNARD, H. VAN HASSELT, AND L. TONI, A survey of temporal credit assignment in deep reinforcement learning, arXiv preprint arXiv:2312.01072, (2023).
- [178] G. PILIOURAS AND J. S. SHAMMA, Optimization despite chaos: Convex relaxations to complex limit sets via poincaré recurrence, in Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms, SIAM, 2014, pp. 861–873.
- [179] J. N. PITT AND A. R. FERRÉ-D'AMARÉ, Rapid construction of empirical rna fitness landscapes, Science, 330 (2010), pp. 376–379.
- [180] H. POINCARÉ, Sur le problème des trois corps et les équations de la dynamique, Mémoire couronné du prix de SM le Roi Oscar II, 1889.
- [181] H. RAIFFA, Game theory at the university of michigan, 1948–1952, Toward a history of game theory, (1992), pp. 165–176.
- [182] A. RAPOPORT AND R. B. BOEBEL, Mixed strategies in strictly competitive games: A further test of the minimax hypothesis, Games and Economic Behavior, 4 (1992), pp. 261–283.
- [183] M. L. ROSENZWEIG, Paradox of enrichment: destabilization of exploitation ecosystems in ecological time, Science, 171 (1971), pp. 385–387.
- [184] W. ROWE, M. PLATT, D. C. WEDGE, P. J. DAY, D. B. KELL, AND J. KNOWLES, Analysis of a complete dna–protein affinity landscape, Journal of The Royal Society Interface, 7 (2010), pp. 397–408.
- [185] H. L. ROYDEN AND P. FITZPATRICK, Real analysis., Featured Titles for Real Analysis Series, (2010), p. 26.
- [186] A. RUBINSTEIN, Inapproximability of nash equilibrium, in Proceedings of the forty-seventh annual ACM symposium on Theory of computing, 2015, pp. 409–418.
- [187] ———, Settling the complexity of computing approximate two-player nash equilibria, ACM SIGecom Exchanges, 15 (2017), pp. 45–49.
- [188] R. M. SAPOLSKY AND L. J. SHARE, Rank-related differences in cardiovascular function among wild baboons: Role of sensitivity to glucocorticoids, American Journal of Primatology, 32 (1994), pp. 261–275.
- [189] V. SARLIS AND C. TJORTJIS, Sports analytics—evaluation of basketball players and team performance, Information Systems, 93 (2020), p. 101562.
- [190] Y. SATO, E. AKIYAMA, AND J. D. FARMER, Chaos in learning a simple two-person game, Proceedings of the National Academy of Sciences, 99 (2002), pp. 4748–4751.
- [191] K. H. SCHLAG, Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits, Journal of economic theory, 78 (1998), pp. 130–156.
- [192] J. SCHRITTWIESER, I. ANTONOGLOU, T. HUBERT, K. SIMONYAN, L. SIFRE, S. SCHMITT, A. GUEZ, E. LOCKHART, D. HASSABIS, T. GRAEPEL, ET AL., Mastering atari, go, chess and shogi by planning with a learned model, Nature, 588 (2020), pp. 604–609.
- [193] P. SCHUSTER, Prediction of rna secondary structures: from theory to models and real molecules, Reports on Progress in Physics, 69 (2006), p. 1419.
- [194] D. SCHUURMANS AND M. A. ZINKEVICH, Deep learning games, Advances in neural information processing systems, 29 (2016).
- [195] S. SHAHSHAHANI, A new mathematical framework for the study of linkage and selection, American Mathematical Soc., 1979.
- [196] D. SILVER, T. HUBERT, J. SCHRITTWIESER, I. ANTONOGLOU, M. LAI, A. GUEZ, M. LANCTOT, L. SIFRE, D. KUMARAN, T. GRAEPEL, ET AL., A general reinforcement learning algorithm that masters chess, shogi, and go through self-play, Science, 362 (2018), pp. 1140–1144.
- [197] D. SILVER, J. SCHRITTWIESER, K. SIMONYAN, I. ANTONOGLOU, A. HUANG, A. GUEZ, T. HUBERT, L. BAKER, M. LAI, A. BOLTON, ET AL., Mastering the game of go without human knowledge, nature, 550 (2017), pp. 354–359.
- [198] B. SINERVO AND C. M. LIVELY, The rock–paper–scissors game and the evolution of alternative male strategies, Nature, 380 (1996), pp. 240–243.

- [199] A. SKOPALIK AND B. VÖCKING, Inapproximability of pure nash equilibria, in Proceedings of the fortieth annual ACM symposium on Theory of computing, 2008, pp. 355–364.
- [200] J. M. SMITH, The theory of games and the evolution of animal conflicts, Journal of theoretical biology, 47 (1974), pp. 209–221.
- [201] J. M. SMITH, Evolution and the theory of games, in Did Darwin get it right? Essays on games, sex and evolution, Springer, 1982, pp. 202–215.
- [202] E. J. SOLBERG AND T. H. RINGSBY, Does male badge size signal status in small island populations of house sparrows, *passer domesticus*?, Ethology, 103 (1997), pp. 177–186.
- [203] P. A. STEPHENS, W. J. SUTHERLAND, AND R. P. FRECKLETON, What is the allee effect?, Oikos, (1999), pp. 185–190.
- [204] A. STOLTZFUS, On the possibility of constructive neutral evolution, Journal of molecular evolution, 49 (1999), pp. 169–181.
- [205] A. STRANG, K. C. ABBOTT, AND P. J. THOMAS, The network HHD: Quantifying cyclic competition in trait-performance models of tournaments, SIAM Review, 64 (2022), pp. 360–391.
- [206] A. STRANG, D. SEWELL, A. KIM, K. ALCEDO, AND D. ROSENBLUTH, Principal trade-off analysis, Information Visualization, (2022), p. 14738716241239018.
- [207] G. STRANG, Linear Algebra and Learning from Data, vol. 4, Wellesley-Cambridge Press Cambridge, 2019.
- [208] D. M. STUART-FOX, D. FIRTH, A. MOUSSALLI, AND M. J. WHITING, Multiple signals in chameleon contests: designing and analysing animal contests as a tournament, Animal Behaviour, 71 (2006), pp. 1263–1271.
- [209] R. S. SUTTON, Temporal credit assignment in reinforcement learning, University of Massachusetts Amherst, 1984.
- [210] R. S. SUTTON, Learning to predict by the methods of temporal differences, Machine learning, 3 (1988), pp. 9–44.
- [211] I. G. SZENDRO, M. F. SCHENK, J. FRANKE, J. KRUG, AND J. A. G. DE VISSER, Quantitative analyses of empirical fitness landscapes, Journal of Statistical Mechanics: Theory and Experiment, 2013 (2013), p. P01005.
- [212] P. D. TAYLOR AND L. B. JONKER, Evolutionary stable strategies and game dynamics, Mathematical biosciences, 40 (1978), pp. 145–156.
- [213] G. TESAURO, Neurogammon: A neural-network backgammon program, in 1990 IJCNN international joint conference on neural networks, IEEE, 1990, pp. 33–39.
- [214] W. S. TORGESSON, Multidimensional scaling: I. theory and method, Psychometrika, 17 (1952), pp. 401–419.
- [215] S. TOSHNIWAL, S. WISEMAN, K. LIVESCU, AND K. GIMPEL, Chess as a testbed for language model state tracking, in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36, 2022, pp. 11385–11393.
- [216] R. TRIVERS, The evolution of sex, 1983.
- [217] B. TRUBENOVA, M. S. KREJCA, P. K. LEHRE, AND T. KÖTZING, Surfing on the seascape: Adaptation in a changing environment, Evolution, 73 (2019), pp. 1356–1374.
- [218] C. S. VALERO, Predicting Win-Loss outcomes in MLB regular season games – A comparative study using data mining methods, International Journal of Computer Science in Sport, 15 (2016), pp. 91–112.
- [219] H. VAN HASSELT, S. MADJIHEUREM, M. HESSEL, D. SILVER, A. BARRETO, AND D. BORSA, Expected eligibility traces, in Proceedings of the AAAI conference on artificial intelligence, vol. 35, 2021, pp. 9997–10005.
- [220] G. J. VELICER AND Y.-T. N. YU, Evolution of novel cooperative swarming in the bacterium *myxococcus xanthus*, Nature, 425 (2003), pp. 75–78.
- [221] M. C. VIEIRA AND P. E. PEIXOTO, Winners and losers: a meta-analysis of functional determinants of fighting ability in arthropod contests, Functional Ecology, 27 (2013), pp. 305–313.
- [222] Y. VIOSSAT AND A. ZAPECHELNYUK, No-regret dynamics and fictitious play, Journal of Economic Theory, 148 (2013), pp. 825–842.

- [223] D. M. VISTRO, F. RASHEED, AND L. G. DAVID, The cricket winner prediction with application of machine learning and data analytics, International journal of scientific & technology Research, 8 (2019), pp. 21–22.
- [224] V. VOLTERRA, Fluctuations in the abundance of a species considered mathematically, Nature, 119 (1927), pp. 12–13.
- [225] J. VON NEUMANN AND O. MORGENSTERN, Theory of games and economic behavior: 60th anniversary commemorative edition, in Theory of games and economic behavior, Princeton university press, 2007.
- [226] V. G. VOVK, Aggregating strategies, in Proceedings of the third annual workshop on Computational learning theory, 1990, pp. 371–386.
- [227] A. WAGNER, Neutralism and selectionism: a network-based reconciliation, Nature Reviews Genetics, 9 (2008), pp. 965–974.
- [228] J. W. WEIBULL, Evolutionary game theory, MIT press, 1997.
- [229] S. WILKENS, Sports prediction and betting models in the machine learning age: The case of tennis, Journal of Sports Analytics, 7 (2021), pp. 99–117.
- [230] C. K. WILLIAMS AND C. E. RASMUSSEN, Gaussian processes for machine learning, vol. 2, MIT press Cambridge, MA, 2006.
- [231] S. WRIGHT ET AL., The roles of mutation, inbreeding, crossbreeding, and selection in evolution, na, (1932).
- [232] J. L. YAN, N. M. SMITH, D. C. FILICE, AND R. DUKAS, Winner and loser effects: a meta-analysis, Animal Behaviour, 216 (2024), pp. 15–22.
- [233] D. YOULA, A normal form for a matrix under the unitary congruence group, Canadian Journal of Mathematics, 13 (1961), pp. 694–704.
- [234] E. C. ZEEMAN, Dynamics of the evolution of animal conflicts, Journal of theoretical Biology, 89 (1981), pp. 249–270.
- [235] A. ZHANG, R. MCALLISTER, R. CALANDRA, Y. GAL, AND S. LEVINE, Learning invariant representations for reinforcement learning without reconstruction, arXiv preprint arXiv:2006.10742, (2020).
- [236] A. ZHENG AND A. CASARI, Feature engineering for machine learning: principles and techniques for data scientists, ” O'Reilly Media, Inc.”, 2018.
- [237] B. ZUMINO, Normal forms of complex matrices, Journal of Mathematical Physics, 3 (1962), pp. 1055–1057.

6 Appendix

6.1 An Example Game

Consider the following population extension of the iterated prisoner’s dilemma game, chosen for illustration:

1. **Pairwise, single-round game:** The game is subdivided into individual rounds between individual pairs of agents. Each individual round, between each individual pair, is an instance of the canonical prisoner’s dilemma game (PD). Each agent must choose to either cooperate or defect. If both agents cooperate, then they both receive 0 units of utility, if both defect, they receive the -1 units of utility, if one defects and the other cooperates, the defecting agent receives 1 unit of utility, while the cooperating agent receives -2 unit of utility. The agents make their choices simultaneously and cannot communicate during each individual round.
2. **Iterated game:** Instead of playing a single round, each pair of agents play multiple rounds. We adopt a variant of the iterated prisoner’s dilemma (IPD) in which the total number of rounds is random (and drawn from a geometric distribution). We set the expected number of rounds to 50. The final payout to a pair after playing IPD is the sum of the utilieis they received in each round. IPD admits rich strategies since agents who can effectively cooperate, and punish defection, may outcompete dishonest opponents. However, dishonest opponents may exploit cooperative opponents if they can mislead the cooperative opponent. Thus, IPD provides an interesting example of decision problems involving cooperation and deception [17, 49].
3. **Sufficient Parameterization:** If agents can remember all actions played over m rounds, then, since each round has $2^2 = 4$ possible action combinations, the number of total strings an agent may remember is of size 4^m . A complete policy specification for agents with perfect memory would consist of a function mapping from each of the 4^m possible histories to a probability of cooperating. This would represent a sufficient parameterization, with attribute space $\Omega = [0, 1]^{4^m}$. While this attribute space is sufficiently rich to express all perfect memory agents, it is evidently opaque. Agents with memories of length 10 would have policies defined by 10^6 attributes.
4. **Auto-regressive Policies:** For interpretability, we adopt a simpler policy parameterization. Agents policies are parameterized by three numbers: an innate preference for cooperation over defection, p_* , a memory or learning rate parameter, α , which controls how quickly agents adapt their policies and the time-scale with which the influence of old information decays, and a reactive parameter, γ , which controls how, and with what severity, agents react to their opponents actions. At each round, an agent possesses a probability of cooperating, $p(j)$. On the first round, the agent uses their innate preference $p(1) = p_*$. On subsequent rounds, their policies update autoregressively:

$$p(k+1) = \alpha p(k) + (1 - \alpha) ((1 - |\gamma|)p_* + |\gamma|\frac{1}{2}(1 + \text{sign}(\gamma)\Delta p(a_-(k)))$$

where $a_-(k)$ is the opponents action on round k , and $\Delta p(a)$ returns +1 for cooperative actions and -1 for defections.

5. **Parameters:** The attribute space Ω is $[0, 1] \times [0, 1] \times [-1, 1]$.

Agents with a large innate preference p_* are inherently cooperative or trusting. Agents with a small p_* are inherently uncooperative or distrustful.

Agents with large α react slowly to new information. Agents with $\gamma = 0$ do not react to their opponents actions. Agents with large positive γ react by imitating their opponent. The larger γ , the more they imitate, and the less strongly they revert towards their innate preference. $\gamma = 1$ returns “tit-for-tat” like strategies [17], where an agent imitates their opponent in order to reward sustained cooperation, but defects immediately following any defection to punish their opponent for defection. Agents of this kind attempt to enforce a cooperative social contract with their opponent, so are “law-enforcing”.

Agents with negative γ react to their opponent by adopting the opposing action. If their opponent cooperates, they see an opportunity, and defect. If their opponent defects, they attempt to re-establish trust, and cooperate. This strategy aims, in essence, to trick the opponent. This strategy has advantages. For example, when an inherently trusting law-enforcing agent ($p_* = 1, \gamma = 1$) is paired with an inherently distrusting and deceptive agent ($p_* = 1, \gamma = -1$), the sequence of plays will proceed deterministically according to repetitions fo the sequence (C,D), (D,D), (D,C), (C,C). On the whole, this balances the payouts received to both agents. However, the deceptive agent is always ahead or tied, while the law-enforcing agent is always behind or tied. The cumulative payout sequences are: (-2,1), (-3,0), (-2,-2), (-1,-1). This means that, if the game is stopped randomly, the deceptive agent is expected to receive a higher payout than their law-enforcing opponent.

6. **Payout:** All of the methods developed in this paper treat zero-sum, or constant-sum games. The cumulative IPD payout is not constant-sum, as illustrated in the example discussed above. Indeed, skew-symmetrizing the IPD payout directly by setting payout equal to the difference in cumulative utilities eliminates the component of the

game that contributes its strategic interest. Namely, the shared reward for cooperating, and the shared cost of defection.

IPD was introduced to study the evolution of cooperative behavior. So, we consider an evolutionary instance of the game. When pairing types x and x' , we initialize a well-mixed population, half of whose agents are of type x , and half of type x' . The population updates through a Moran process [152]. At each time step of the Moran process, the members of the population are broken into randomly assigned pairs. They play the IPD game, and are assigned a fitness equal to their cumulative payouts. On the next round, we draw a new population of the same size, where each child is randomly assigned a parent from the previous population. The probability distribution over parents is set to a soft-max of their cumulative fitness. This has the effect of rewarding cooperation since agents who effectively cooperate with their opponent both receive large cumulative payouts, so produce more children in the subsequent round. The process terminates when the one type goes extinct, and the dominating type is declared the winner. The probability that the dominating type is x , when paired with x' is the fixation probability, $p_{\text{fix}}(x, x'; N, \lambda)$ where N and λ are the total initial population size and the softmax parameter (20, and 1 in the experiments reported).

Then, the payout $f(x, x')$ is set equal to $2 \times (p_{\text{fix}}(x, x'; N, \lambda) - 1/2)$. This payout is skew-symmetric by construction, but encodes the evolutionary reward for cooperative behavior in addition to the evolutionary reward for successful defection.

We adopted this example for demonstration since it illustrates a number of the key points of the paper. First, it does not reduce to a simple hierarchy. Second, while it uses a low-dimensional attribute space, with interpretable parameters, the exact mapping between attributes and outcome is opaque. Direct calculation of f is nontrivial, since it requires accounting for all possible play trajectories, and all possible trajectories of the Moran process. Direct simulation is easy to construct, but is similarly limited since it requires repeatedly sampling outcomes between populations of types x and x' to estimate the fixation probability, each outcome requires running an evolutionary process, and each time step of the evolutionary process encodes many repetitions of the IPD process, one for every pair of agents. Even the IPD process itself is stochastic under the autoregressive policies. As such, while it seems plausible that, given the relatively simple individual components, a clear strategic interpretation should exist, it is non-obvious how to recover or represent it.

Figure 1 shows the resulting payout matrices for a sample population of 800 agents, with population distribution π given by independently sampling $p_* \sim \text{Beta}(0.7, 0.7)$, $\alpha \sim \text{Beta}(1, 1)$ and $\gamma = 2 \times (\eta - 1/2)$ where $\eta \sim \text{Beta}(3, 3)$.

Figure 4 illustrates the leading disc game for the resulting embedding. This disc game accounts for 90% of the overall variance in performance. The next two disc games account for 6% and 2% of the variance respectively.

6.2 Proofs of Embedding Results

6.2.1 Proof of Lemma 1

Lemma 1: [Opponent Independent Learning Implies Transitivity] If there exists a transformation T such that g satisfies desiderata (1) and Ψ is connected⁴, then f is perfectly transitive.

Proof: Suppose that desiderata (1) applies, and Ψ is connected. Then $v(y, y') = v(y)$, and, by the fundamental theorem of calculus,

$$\begin{aligned} g(y, y') - g(y_0, y') &= \int_{w=y_0}^y v(w) dw \\ g(y, y') - g(y, y'_0) &= - \int_{w=y'_0}^{y'} v(w) dw. \end{aligned} \tag{88}$$

The latter equality holds since f is skew-symmetric, thus g must also be skew-symmetric. If g is skew-symmetric, then reversing the order of the arguments converts the first expression into the second.

Also, by skew-symmetry $g(y_0, y'_0) = 0$ if $y_0 = y'_0$ for any y_0 . Therefore, adding zero:

$$g(y, y') = g(y, y') - g(y_0, y_0) = (g(y, y') - g(y_0, y')) + (g(y_0, y') - g(y_0, y_0)) = \int_{w=y_0}^y v(w) dw - \int_{w=y_0}^{y'} v(w) dw. \tag{89}$$

Let $r(y) = \int_{w=y_0}^y v(w) dw$. Then $g(y, y') = r(y) - r(y')$ for all pairs $(y, y') \in \Psi \times \Psi$, so g is perfectly transitive. If g is perfectly transitive, then f is perfectly transitive. \square

6.2.2 Proof of Lemma 2

Lemma 2: [Student Independent Learning Implies Bilinearity] If $v(y, y')$ does not depend on y , then $g(y, y')$ is a bilinear function of y and y' of the form:

$$g(y, y') = [y^\top, 1] G \begin{bmatrix} y' \\ 1 \end{bmatrix} \quad (90)$$

where G is skew-symmetric.

Proof: Suppose that the optimal training vector is student independent. Then, $v(y, y')$ can be rewritten, $v(y')$, since v does not depend on y . By definition, $\nabla_w g(w, y')|_{w=y} = v(y, y') = v(y')$. That is, the gradient of $g(y, y')$ with respect to y depends only on y' . So, for each fixed y' , $g(y, y')$ is an affine function of y . Since g is skew symmetric, it must also be affine in y' for any fixed y . It follows that g must take the form:

$$g(y, y') = y^\top G y' + a^\top y + b^\top y' + c \quad (91)$$

for some choice of the matrix G , vectors a and b , and vector c . Since g is skew-symmetric, $g(0, 0) = 0$. It follows that $c = 0$.

Then, by skew symmetry:

$$y^\top G y' + a^\top y + b^\top y' = g(y, y') = -g(y', y) = -y'^\top G y - a^\top y' - b^\top y. \quad (92)$$

Subtracting the left-hand side from the right-hand side leaves:

$$y^\top (G + G^\top) y' + (a + b)^\top (y + y') = 0 \text{ for all } y, y' \quad (93)$$

Picking $y = y' = tz$ such that $(a + b)^\top z \neq 0$, then varying t produces a quadratic function of t . A quadratic function of t is only zero for all t if its coefficients equal zero. This requires that the projection of $a + b$ onto the range of possible y is 0, so the projection of a onto the range of possible y equals the projection of $-b$ onto the range of possible y . Any remaining component vanishes in the inner-product, thus may be dropped. Therefore, $a = -b$.

Then, matching the bilinear terms requires:

$$y^\top (G + G^\top) y' = 0 \quad (94)$$

for all pairs y and y' . This requires $G = -G^\top$, so the matrix G must be skew-symmetric. Then:

$$g(y, y') = y^\top G y' + a^\top (y - y'). \quad (95)$$

We can extend this to a bilinear form when $a \neq 0$ by expanding G . Set:

$$G \leftarrow \begin{bmatrix} G & -a \\ a^\top & 0 \end{bmatrix} \quad (96)$$

Then, G is skew-symmetric, and $g(y, y')$ takes the form (90). \square

Notice that the linear term $a^\top (y - y') = a^\top y - a^\top y'$ is transitive with $r(y) = a^\top y$. Thus, if the bilinear term is zero, then the game is perfectly transitive. When the bilinear term is not zero, then the game is not perfectly transitive, thus the second desiderata is more general than the first, which requires perfect transitivity.

6.2.3 Proof of Theorem 3

Theorem 3: [Disc Game Embeddings Are Sufficient] Any coordinate transformation satisfying desiderata (2) using finitely many coordinates is related to a disc game embedding by a linear transformation.

Proof: Suppose that there exists a transform T satisfying desiderata (2). Then $g(y, y') = f(T^{-1}(y), T^{-1}(y'))$ is bilinear. All bilinear functions, skew-symmetric functions admit a disc-game embedding as, all finite-dimensional, skew-symmetric, bilinear functions may be expressed as a linear combination of finitely many separable functions, thus are degenerate. All degenerate functions are disc-game embeddable.

To find a particular embedding, suppose that:

$$g(y, y') = [y^\top, 1] G \begin{bmatrix} y' \\ 1 \end{bmatrix}. \quad (97)$$

for some skew-symmetric G . Then, Schur-decompose G , and regroup terms:

$$\begin{aligned} f(x, x') &= g(T(x), T(x')) = [T(x)^\top, 1] G \begin{bmatrix} T(x') \\ 1 \end{bmatrix} = [T(x)^\top, 1] Q D_\omega U D_\omega^{1/2} Q^\top \begin{bmatrix} T(x') \\ 1 \end{bmatrix} \\ &= \left(D_\omega^{1/2} Q^\top \begin{bmatrix} T(x) \\ 1 \end{bmatrix} \right)^\top U \left(D_\omega^{1/2} Q^\top \begin{bmatrix} T(x') \\ 1 \end{bmatrix} \right) \end{aligned} \quad (98)$$

If we let:

$$z(x) = D_\omega^{1/2} Q^\top \begin{bmatrix} T(x') \\ 1 \end{bmatrix} \quad (99)$$

then:

$$f(x, x') = \sum_k \text{disc}(z^{(k)}(x), z^{(k)}(x')) \quad (100)$$

where $z(x)$ is a linear transformation of $y(x)$. When this transform is invertible, then y may be recovered from z by a linear transformation. If not, then G is not full rank, so has a nontrivial nullspace, and any restriction of y to a subspace perpendicular to the nullspace of G , produces a valid coordinate system satisfying desiderata (2), that may be recovered from z via a linear transformation. \square

6.2.4 Proof of Lemma 6

Lemma 6: [Outcome Continuity] If $|f(x, x'') - f(x', x'')| \leq \delta$ for almost all x'' with respect to ν , or $\|f(x, \cdot) - f(x', \cdot)\|_\nu = \mathbb{E}_{X'' \sim \nu}[(f(x, x'') - f(x', x''))^2] \leq \delta^2$ then $\|y^{(k)}(x) - y^{(k)}(x')\|_2 \leq |\lambda_{2k-1}|^{-1/2}\delta$.

Proof: First, if $|f(x, x'') - f(x', x'')| \leq \delta$ for almost all x'' with respect to ν , then $\mathbb{E}[(f(x, X'') - f(x', X''))^2] \leq \delta^2$ when $X'' \sim \nu$. Therefore, if the lemma statement holds for the two-norm $\|f(x, \cdot) - f(x', \cdot)\|_\nu$, it will also hold for the infinity norm. Both cases can be shown by writing $y^{(k)}$ in terms of the eigenfunctions of F_ν .

Recall that $y^{(k)}(x)$ is the vector in \mathbb{R}^2 whose coordinates correspond to $\sqrt{|\lambda_{2k-1}|}\phi_k(x)$ in the complex plane. Therefore, $\|y^{(k)}(x) - y^{(k)}(x')\|_2 = \sqrt{|\lambda_{2k-1}|}\|\phi_{2k-1}(x) - \phi_{2k-1}(x')\|_2$. We state our bounds blockwise, but not coordinate-wise, since the embedding maps are only uniquely defined up to rotation within each block.

Now, since $\phi_j = \lambda_j F_\nu[\phi_j]$:

$$\begin{aligned} \|\phi_j(x) - \phi_j(x')\|^2 &= \frac{1}{|\lambda_j|^2} \|\mathbb{E}_{X'' \sim \nu}[(f(x, X'') - f(x', X''))\phi_j(X'')] \|^2 \\ &\leq \frac{1}{|\lambda_j|^2} \mathbb{E}_{X'' \sim \nu} [|f(x, X'') - f(x', X'')|^2] \mathbb{E}_{X'' \sim \nu} [\|\phi(X'')\|^2] \leq \frac{\delta^2}{|\lambda_j|^2} \end{aligned} \quad (101)$$

where the second step follows by the Cauchy-Schwarz inequality for the inner product $\langle \cdot, \cdot \rangle_\nu$, and the last step follows from the facts that the eigenfunctions are normalized with respect to $\|\cdot\|_\nu^2$.

Then:

$$\|y^{(k)}(x) - y^{(k)}(x')\|_2 = \sqrt{|\lambda_{2k-1}|} \|\phi_{2k-1}(x) - \phi_{2k-1}(x')\| \leq \frac{\delta}{\sqrt{|\lambda_{2k-1}|}}. \quad \blacksquare \quad (102)$$

6.3 Alternate Derivation of the Disc Game Embedding

We derived the disc-game embedding by seeking a coordinate transformation that satisfied desiderata (2). Since we will propose disc-game embedding as a canonical representation for learning dynamics, it is important to highlight other possible rationales for disc-game embedding.

For example, consider a single agent involved in a training process. Let $\{x(j)\}_{j=1}^n$ represent the attributes of the agent at some sequence of times. With each step, the agent measures their relative improvement in performance by evaluating $f(x(j+1), x(j))$. At the end of the training process, they evaluate their actual improvement via $f(x(n), x(1))$. It is natural to expect that, their accumulated incremental advantage, $\sum_{j=1}^{n-1} f(x(j+1), x(j))$ might approximate their ultimate advantage, $f(x(n), x(1))$. However, there is no guarantee that this approximation is accurate. Moreover, there is no obvious way to predict the size of the error given the sequence of agent types.

The error in the approximation, $\sum_{j=1}^{n-1} f(x(j+1), x(j)) - f(x(n), x(1))$ is an example of a curl. By convention, set $x(n+1) = x(1)$. Then, the sequence $\{x(j)\}_{j=1}^n$ form a cycle, \mathcal{C} of agent types. By the skew-symmetry of f , the difference between accumulated incremental advantage, and ultimate advantage is $\text{curl}_{\mathcal{C}}[f] = \sum_{j=1}^n f(x(j+1), x(j))$.

Let T denote an invertible coordinate change. Then $\text{curl}_{\mathcal{C}}[f] = \text{curl}_{T(\mathcal{C})}[g]$ where $y = T(x)$, and $g(y, y') = f \cdot (T^{-1} \times T^{-1})$. Let's choose T so that the curl is easy to predict from the sequence of embedded coordinates $\{y(j)\}_{j=1}^n$. For

example, we might seek a coordinate change such that, the curl is zero, for all sets of colinear agents. Then, when $\{y(j)\}_{j=1}^n$ all lie on some line in the embedding space, the cumulative incremental advantage received by stepping along the line will correctly predict the ultimate advantage an agent at one end possesses over the other.

Desiderata: Identify a transform T such that:

4. $\text{curl}_C[g] = 0$ for all cycles $C = \{y(j)\}_{j=1}^{n+1}, y(n+1) = y(1)$ such that $\{y(j)\}_{j=1}^{n+1}$ are colinear .

Like desiderata (2), desiderata (4) restricts the class of possible g .

Lemma A.1: A symmetric, zero-sum, differentiable functional form game satisfies desiderata (4) if and only if it is a vector field game. A symmetric, zero-sum, differentiable functional form game is a vector field game if and only if its payout function g can be satisfies:

$$g(y, y') = \int_{z=y'}^y v_{\text{self}}(z) \cdot dz \quad (103)$$

where the path integral is evaluated over the line segment from y' to y .

In a vector-field game, the payout y receives against y' can be expanded as a path-integral against the optimal self-play gradient along the line segment linking y' to y . Note that this is, in essence, a fundamental theorem of calculus. The ultimate advantage gained by training along a line equals the accumulated advantage gained by each training step. Indeed, when y is close to y' and g is differentiable, then the path-integral provides an asymptotically accurate approximation to $g(y, y')$ by linearization. However, although the path integral is a fundamental theorem in essence, it is not the fundamental theorem, and does not apply in general. It fails in general since $g(y, y')$ is simply a function evaluation in the product space $\Psi \times \Psi$, while the path integral is evaluated over a sequence of different local neighborhoods of (z, z) in the product space, where z ranges along the line from y' to y . When y' is far from y , the set of local neighborhoods along the line segment from y' to y may be far removed from the point (y, y') .

Theorem A.2: [Vector Field Games] The set of vector-field games is:

1. a vector space of functions on the product space,
2. non-empty and isomorphic the set of continuous vector fields on Ψ ,
3. a proper subset of the set of all symmetric, zero-sum games on $\Psi \times \Psi$ (there exist non-vector-field games).

Moreover, an analytic symmetric, zero-sum, functional form game is not a vector-field game if, for any $n \neq m$, both greater than zero, there exists a directional derivative $\partial_t^n \partial_s^m g(z + tv, z + sv) \neq 0$ at some $z \in \Psi \subseteq \mathbb{R}^d$, for all directions $v \in \mathbb{R}^d$.

It follows that, all quadratic polynomial payouts define vector-field games. Generic third-order polynomials do not define vector-field games. The disc game is bilinear, so is a vector-field game. The disc game is a particularly convenient vector-field game since by Stoke's theorem, the curl over any cycle of competitors equals twice the sum over each pair of consecutive coordinates, of the signed-area traced out by the cycle in that pair of coordinates. Thus, loops with large areas correspond to sequences of competitors where the accumulated incremental advantage is a poor predictor of ultimate advantage.

A functional form game (Ω, f) is vector-field representable if there exists an invertible, differentiable coordinate transformation T such that (Ψ, g) is a vector-field game. Then, accumulated incremental advantage equals true advantage accrued over paths corresponding to straight lines in Ψ . These paths are geodesics with respect to Euclidean distance in the embedding coordinates. Then, $f(x, x')$ equals the path integral over the geodesic linking x' to x against the self-play vector field v_{self} in Ω . Note that there may not be a geodesic that remains in Ω connecting all pairs x, x' if Ψ is not connected and convex.

A priori, it is unclear which games are vector-field representable. This class is at least as large as the set of all continuous vector fields on Ω , and should be much larger since we did not enforce many constraints on T .

Here, the observation that essentially all games of interest are disc-game embeddable offers a clear answer. If a game is disc-game embeddable, then, since the disc-game is a vector-field game, the original game is vector-field representable, and the disc-game embedding provides a sufficient transform T . Therefore, the disc-game embedding accomplishes desiderata (4) for essentially all games of interest. Note that desiderata (4) is less restrictive than desiderata (2), so there may be other, transforms that satisfy desiderata (4) that are not affine transformations of a disc game embedding. Nevertheless, the disc-game embedding offers a sufficient construction for almost all games of interest, and, as shown in the main text, offers more than a vector-field representation alone.

6.4 Additional Learning Dynamics

Evolutionary game theory studies the dynamics of populations evolving subject to an underlying game. For classical surveys, focusing primarily on economics, see [77, 78, 228].

6.4.1 Self Play and Adaptive Dynamics

The simplest learning dynamic is self-play. Self-play assumes a monomorphic population. In a monomorphic population, all individuals have the same type, i.e. share the same traits. Self-play may be used to study learning when the population consists of a single agent that trains against copies of itself. Thus, assuming self-play, we can replace the study of a time-evolving distribution $\pi(\cdot, t)$ over Ω , with the time evolution of the traits of the single individual, $x(t)$.

Let $v_{\text{self}}(x) = v(x, x) = \nabla_w f(w, x)|_{w=x}$ return the direction of fastest increase in f for an agent of type x paired against themselves. The vector field $v_{\text{self}}(x)$ gives the optimal training direction for an individual of type x learning under self-play, when constrained to incremental adjustments. A self-play dynamic sets:

$$\frac{d}{dt}x(t) \propto v_{\text{self}}(x(t)) = \nabla_w f(w, x(t))|_{w=x(t)}. \quad (104)$$

Here, \propto allows time-dependent proportionality, as when a designer adopts a decaying learning rate. We will generally ignore any time-dependent scaling factor since all such dynamics are equivalent under nonlinear transformations of time.

In a training example, the gradient $\nabla_w f(w, x)|_{w=x}$ may be approximated using automatic differentiation, or by simulating many competition events between an agent of type x , and opponents of type x' for $x' \approx x$.

Self-play is a form of adaptive dynamics. Like self-play, adaptive dynamics studies the change in a trait vector, $x(t)$, that acts as a proxy for a distribution. The form of the proxy, and specific dynamics, depend on the approximations used. In adaptive dynamics,

$$\frac{d}{dt}x(t) \propto C(x, t)v_{\text{self}}(x(t)) \quad (105)$$

for some positive semi-definite matrix $C(x, t)$.

The choice of $C(x, t)$ depends on the method used to derive the adaptive dynamics equation. For example, if mutations are rare on the time scale of fixation, and only produce small changes in traits, then an evolving population will be near to monomorphic most of the time. Then, $x(t)$ may be chosen to represent the traits of the majority of the population or the centroid of the distribution of traits. Then $x(t)$ changes due to a combination of random mutation and the fixation of mutant types. By adopting a long time scale relative to both fixation and mutation, an analyst can predict the expected dynamics of $x(t)$ by averaging over many possible mutations and fixation events. Then $C(x, t)$ depends on the distribution of mutations.

Alternately, adaptive dynamics equations can be derived by assuming a highly concentrated population. That is, a population consisting of mostly similar individuals. In this case, the dynamics of $x(t)$ can be approximated using closure approximations. For example, if the original population obeys a replicator dynamic, is Gaussian distributed, and is driven by a quadratic payout function f , then the time-evolution of the centroid of the population obeys an adaptive dynamics equation. In this setting, $C(x, t)$ is the covariance in the traits of the population.

6.4.2 Fictitious Self-Play

Fictitious self-play adapts self-play by recording the history of an evolving agent with type $x(t)$. Instead of adapting $x(t)$ along the direction of best response to the current type, fictitious self-play adapts along the average best response direction against the full past of $x(t)$. By incorporating the history of $x(t)$, fictitious self-play can avoid training in cycles when presented with rock-paper-scissors-type games that involve advantage cycles.

Specifically, fictitious self-play sets:

$$\frac{d}{dt}x(t) \propto \int_{s=0}^t \kappa(t, s)v(x(t), x(s))ds \quad (106)$$

where $\kappa(t, s)$ is a kernel used to weight the contribution of time s to the update of the agent at time t .

Different choices of κ produce different dynamics. For example, if $\kappa = \frac{1}{t}$ weights all past agent types uniformly when choosing the update at time t , while $\kappa(t, s) \propto k(t-s)$ for some monotonically decreasing function k weights the average to attend more to recent types than past types.

6.4.3 Simultaneous Gradient Ascent

Fictitious self-play updates an agent by adapting in the average best-response direction with respect to all the past variants of the agent. Given a population containing multiple types, the average over past types can be replaced with an average against all types contained in the current population.

Consider a discrete population containing n individuals of types $\{x(j, t)\}_{j=1}^n$. Then, the population obeys simultaneous gradient-ascent if every individual adapts in their average best-response direction against opponents drawn from the current population:

$$\frac{d}{dt}x(j, t) \propto \frac{1}{n} \sum_{i \neq j} v(x(j, t), x(i, t)) \quad (107)$$

Alternately, consider a continuous population represented by a density function $\pi(\cdot, t)$. Then, if the population obeys simultaneous gradient ascent:

$$\partial_t \pi(x, t) \propto -\nabla_x \cdot (\mathbb{E}_{Y \sim \pi(\cdot, t)}[v(x, Y)]\pi(x, t)) = -\mathbb{E}_{Y \sim \pi(\cdot, t)}[\nabla_x \cdot (v(x, Y)\pi(x, t))]. \quad (108)$$

Like fictitious self-play, simultaneous gradient ascent can be extended to incorporate past agents by extending the average to run over the full population and over all past times $s \leq t$.

6.5 Explicit Solutions

6.5.1 Advective Dynamics

Here, we present explicit solutions for self-play, fictitious self-play, and simultaneous gradient ascent on discrete populations. Our solutions are restricted to the interior of the set of admissible agents Ψ .

First, a note of caution. The learning dynamics that advect agents along average optimal training vectors (self-play, fictitious self-play, simultaneous gradient ascent), cannot, barring very restrictive assumptions, be reduced to transformations of a set of canonical solutions in an intrinsic space since the transformed dynamics retain a nontrivial dependence on distances in the original (extrinsic) space. Thus, even if essentially all payouts can be replaced with the disc game, different choices of the original coordinate system lead to different dynamics in the latent space.

Let $J(x)$ denote the Jacobian of the transformation from x to y . Then, by the chain rule

$$\begin{aligned} \frac{d}{dt}y(t) &= \frac{d}{dt}T(x(t)) = J(x(t))\frac{d}{dt}x(t) \\ v(x, x') &= \nabla_w f(w, x')|_{w=x} = \nabla_w T(w)^\top U T(x')|_{w=x} = J(x)^\top U T(x') = J(x)^\top U y' = J(x)^\top v_{\text{self}}(y'). \end{aligned} \quad (109)$$

Therefore, any learning dynamic that equates $\frac{d}{dt}x(t)$ to a weighted average of $v(x, x')$ over some set of x' of the form:

$$\frac{d}{dt}x(t) \propto \mathbb{E}_{x'}[v(x, x')] \quad (110)$$

sets:

$$\frac{d}{dt}y(t) = [J(x(t))J^\top(x(t))] \mathbb{E}_{y'}[v_{\text{self}}(y')] = C(x(t))\bar{v}_{\text{self}}(t) = C(x(t))v_{\text{self}}(\bar{y}(t)) \quad (111)$$

where $C(x) = J(x)J(x)^\top$ is the metric tensor relating differential distances in the embedded coordinate system to differential distances in the original coordinate system, $x(t) = T^{-1}(y(t))$, and $\bar{y}(t) = \mathbb{E}_{y'}[y'] = \mathbb{E}_{x'}[T(x')]$ is the centroid in the embedded coordinates of the distribution of agents used in the average (110). The metric tensor $C(x)$ is positive definite when T is invertible, so Equation (111) generalizes the adaptive dynamics equation by averaging the optimal self-play training field over a population of possible opponents.

Since (111) depends on the metric tensor C , the solutions to (111) depend on the specific mapping T from the original coordinates x into the embedded coordinates y . Therefore, even though the coordinate change T replaces arbitrary f with a fixed payout function (the disc game), the solutions to (111) still vary with the choice of the original coordinate system. The dependence on the metric tensor C vanishes if $J(x)$ has orthogonal columns, and consecutive pairs of columns of $J(x)$ share a constant norm for all x . In that case, C is diagonal, with constant diagonal entries grouped in consecutive pairs, and, since $v_{\text{self}}(y) = Uy$ where U is block-diagonal, equation (111) can be broken into a set of decoupled two-dimensional learning dynamics, each of the same form, with time-constant determined by the norm of the columns of J . Unfortunately, the set of transformations T whose Jacobians are orthogonal is very small. For example, Louiville's theorem states that, the set of smooth conformal maps (transformations whose Jacobians are orthogonal matrices) in $d > 2$ is restricted to Möbius transformations (affine transformations, and transformations between circles and lines).⁵

Therefore, learning dynamics of the form (110) do not admit a canonical solution up to a coordinate change for a reasonable generic set of payout functions. Therefore, when seeking explicit solutions, we will focus on the case when learning dynamics of the form (110) are applied in the disc game space directly. This significantly restricts the generality of the following explicit solutions. In particular, they limit the use of these results to artificial training settings where a rational designer can choose the learning dynamic.

⁵Is there a reasonable symmetry of f that implies the embedding is conformal?

1. **Self-Play:** Suppose that $y(t) = T(x(t))$ obeys a self-play ODE in the embedded coordinate space:

$$\frac{d}{dt}y(t) = \lambda(t)v_{\text{self}}(y(t)) \quad (112)$$

for some $\lambda(t) > 0$.

Without loss of generality, we assume that $\lambda(t) = 1$. All solutions to (112) are equivalent to the solution for $\lambda(t) = 1$ up to a monotonic change in the time variable. Since $v_{\text{self}}(y) = Uy$ when $f(y, y') = \text{disc}(y, y')$, equation (112) decouples blockwise:

$$\frac{d}{dt}y_{2k-1}(t) = y_{2k}(t), \quad \frac{d}{dt}y_{2k}(t) = -y_{2k-1}(t). \quad (113)$$

Equation (113) is the phase space equation for a simple harmonic oscillator. As long as y_{2k-1} and y_{2k} remain on the interior of Ψ , it is solved by:

$$y^{(k)}(t) = [y_{2k-1}(t), y_{2k}(t)] = \|y^{(k)}(0)\|[\cos(t-s), -\sin(t-s)] \quad (114)$$

for the phase s satisfying the initial conditions. In other words, each pair of consecutive disc-game coordinates evolve independently, and follow circular path segments at a unit rate as long as the circular path segment does not intersect the boundary of Ψ .

2. **Fictitious Self-Play:** Suppose that $y(t) = T(x(t))$ obeys a fictitious self-play dynamic in the embedded coordinate space. For example, consider the dynamic:

$$\frac{d}{dt}y(t) = \frac{1}{t} \int_{s=0}^t v_{\text{self}}(y(s)). \quad (115)$$

Since $v_{\text{self}}(y) = Uy$ is linear:

$$\frac{d}{dt}y(t) = U\bar{y}(t), \quad \bar{y}(t) = \frac{1}{t} \int_{s=0}^t y(s)ds \quad (116)$$

where $\bar{y}(t)$ is the centroid of the full past trajectory $\{y(s)\}_{s=0}^t$. In, other words, if $y(t)$ evolves according to fictitious self-play in the latent space, then $y(t)$ always moves along the optimal training response direction to the average past agent.

The lag equation 116 can be written as a first-order system of coupled ODE's by differentiating $\bar{y}(t)$:

$$\frac{d}{dt}y(t) = U\bar{y}(t), \quad \frac{d}{dt}\bar{y}(t) = \frac{1}{t}(y(t) - \bar{y}(t)), \quad \bar{y}(0) = y(0). \quad (117)$$

The factor of $1/t$ in Equation (117) accounts for the fact that, as t increases, the average agent $\bar{y}(t)$ over the full past history becomes less sensitive to the new agent added at time t . In other words, as t increases, $\bar{y}(t)$ moves more slowly.

Differentiating (117) with respect to t a second-time produces a single second-order amplified oscillator equation:

$$\frac{d^2}{dt^2}y(t) = U\frac{d}{dt}\bar{y}(t) = \frac{1}{t}U(y(t) - \bar{y}(t)) = \frac{1}{t}\left(Uy(t) - \frac{d}{dt}y(t)\right). \quad (118)$$

Rearranging, and separating blockwise yields the system inhomogeneous of second-order equations:

$$\begin{aligned} t\frac{d^2}{dt^2}y_{2k-1}(t) &= y_{2k}(t) - \frac{d}{dt}y_{2k-1}(t) \\ t\frac{d^2}{dt^2}y_{2k}(t) &= -y_{2k-1}(t) - \frac{d}{dt}y_{2k}(t) \end{aligned} \quad (119)$$

Equation 119 is almost solved by a logarithmic spiral with phase increasing proportional to $\sqrt{2t}$. That is $\theta_k(t) = \tan^{-1}(y_{2k-1}(t)/y_{2k}(t)) \simeq \sqrt{2t} + \theta_k(0)$, and $r_k(t) = \|y^{(k)}(t)\| \simeq c_3 \exp(\theta_k(t))$ for some choice of constants. This solution is inexact, since, for small t , $\bar{y}(t) \approx y(t)$. Thus, for small t , the fictitious play solution imitates the self-play solution. Indeed, at $t = 0$, the two solutions must align, which is impossible if the radius is increasing at $t = 0$. The radius only starts to grow exponentially in the phase once $y(t)$ has separated from $\bar{y}(t)$. For

example, if initialized from $y_k(0) = \bar{y}_k(0) = [1, 0]$, then numerically solving the dynamic produces $\theta_k(t) \approx \sqrt{t/2}$ and $\log(r_k(t)) \approx (\theta_k(t)^2 - \theta_k(t)/\sqrt{2} - \sqrt{2})/(\theta_k(t) + 2)$ to small errors. Both approximations are essentially exact for $t > 10$. Notice that $\log(r_k(t))$ is very nearly proportional to $\theta_k(t)$ for large t , but, for small t , is approximately constant in t . The solution initialized at $[1, 0]$ is generic since the initial average $\bar{y}(0)$ will always match the initial agent type $y(0)$, the disc game is rotationally symmetric, so changing the initial phase only rotates the solution, and, the dynamic is linear, so scaling by some scalar r_0 only scales the corresponding solution by r_0 .

The solution spirals outwards since the average $\bar{y}(t)$ lags behind $y(t)$, and, as y increases, responds more slowly to $y(t)$. As $\bar{y}(t)$ slows in response, $y(t)$ travels along longer path segments that are approximately linear, with direction fixed by $\bar{y}(t)$.

Like our solution for self-play, this solution is only valid until $y(t)$ intersects the boundary of Ψ . Notice that, fictitious self-play leads the sequence of agents to explore outwards, so is likely to encounter the boundary if Ψ is bounded, as is usually the case when f is bounded.

Equation 119 can be solved exactly using power series expansion. Without loss of generality, set $y^{(k)}(0) = [1, 0]$ and $\bar{y}^{(2k)}(0) = y^{(k)}(0)$. Then, $\frac{d}{dt}y^{(k)}(0) = [0, -1]$. Let $z(t) = y_{2k-1}(t)$ and $w(t) = y_{2k}(t)$. Then, if we expand:

$$w(t) = \sum_{n=0}^{\infty} \frac{w_0^{(n)}}{n!} t^n, \quad z(t) = \sum_{n=0}^{\infty} \frac{z_0^{(n)}}{n!} t^n \quad (120)$$

substituting into equation (116), and matching coefficients yields the recursion:

$$w_0^{(n+1)} = \frac{1}{n+1} z_0^{(n)}, \quad z_0^{(n+1)} = -\frac{1}{n+1} w_0^{(n)}. \quad (121)$$

Substituting $w_0^{(0)} = 1$, $z_0^{(0)} = 0$ gives the correct first order derivatives at $t = 0$, $w_0^{(1)} = 0$, $z_0^{(0)} = -1$. Then, closing the recursion yields the power series expansion:

$$\begin{aligned} w(t) &= 1 + 2 \sum_{k=1}^{\infty} \frac{(-1)^k}{(2k+1)(2k)!^2} t^{2k} \\ z(t) &= 2 \sum_{k=1}^{\infty} \frac{(-1)^k}{(2k)(2k-1)!^2} t^{2k-1} \end{aligned} \quad (122)$$

These power series close, producing the explicit solution for $y^{(k)}(0) = [1, 0]$:

$$\begin{aligned} y_{2k-1}(t) &= \sqrt{\frac{1}{2t}} \left(\text{ber}_1(2\sqrt{t}) - \text{bei}_1(2\sqrt{t}) \right) \\ y_{2k}(t) &= \sqrt{\frac{1}{2t}} \left(\text{ber}_1(2\sqrt{t}) + \text{bei}_1(2\sqrt{t}) \right) \end{aligned} \quad (123)$$

where $\text{bei}_1(x)$ and $\text{ber}_1(x)$ are the Kelvin bei and ber functions. These are, respectively, the real and imaginary parts of the Bessel function of the first kind, J_1 . Specifically $\text{bei}_1(2\sqrt{t}) = \text{Imag}(J_1((-1+i)\sqrt{2t}))$ and $\text{ber}_1(2\sqrt{t}) = \text{Real}(J_1((-1+i)\sqrt{2t}))$. All other solutions can be found by first scaling the solution to match the initial radius, then rotating to match the initial phase.

- 3. Simultaneous Gradient Ascent:** Suppose that a population of n agents obeys simultaneous gradient ascent in the latent space. Let $y(j, t)$ denote the coordinates of the j^{th} agent at time t . Then, for all j :

$$\frac{d}{dt}y(j, t) = \lambda(t)\mathbb{E}_{I \neq j} [v(y(I, t))] = \frac{\lambda(t)}{n-1} \sum_{i \neq j} v(y(i, t)) = \frac{\lambda(t)}{n-1} U \sum_{i \neq j} y(i, t). \quad (124)$$

Without loss of generality, let $\lambda(t) = \frac{n-1}{n}$. Then:

$$\frac{d}{dt}y(j, t) = U \left(\bar{y}(t) - \frac{1}{n} y(j, t) \right) \quad (125)$$

where $\bar{y}(t) = \frac{1}{n} \sum_{i=1}^n y(i, t)$. This equation can be solved efficiently by noting that the motion of the average \bar{y} , and the difference between each agent and the average, $w(j, t) = y(j, t) - \bar{y}$, decouple, and can be solved independently.

Solving for the average:

$$\frac{d}{dt}\bar{y}(t) = \frac{1}{n} \sum_{j=1}^n \frac{d}{dt}y(j,t)) = \frac{1}{n}U \sum_{j=1}^n \left(\bar{y}(t) - \frac{1}{n}y(j,t) \right) = \left(1 - \frac{1}{n} \right) U\bar{y}(t). \quad (126)$$

Then the centroid obeys the self-play equation with time constant $1 - \frac{1}{n}$. Therefore, $\bar{y}(t)$ traces a circular path at rate $1 - \frac{1}{n}$.

Let $w(j,t) = y(j,t) - \bar{y}(t)$. Then:

$$\frac{d}{dt}w(j,t) = \frac{d}{dt}y(j,t) - \frac{d}{dt}\bar{y}(t) = U \left(\bar{y}(t) - \frac{1}{n}y(j,t) \right) - \left(1 - \frac{1}{n} \right) U\bar{y}(t) = -\frac{1}{n}U(y(j,t) - \bar{y}(t)) = -\frac{1}{n}Uw(j,t). \quad (127)$$

Therefore, $w(j,t)$ obeys a time-reversed self-play equation with time constant $1/n$.

Combining these results produces an epicyclic solutions where the centroid of the population moves clockwise at rate $1 - \frac{1}{n}$, and where each individual orbits counterclockwise about the centroid at rate $\frac{1}{n}$:

$$y^{(k)}(j,t) = \begin{bmatrix} y_{2k-1}(j,t) \\ y_{2k}(j,t) \end{bmatrix} = \|\bar{y}^{(k)}(0)\| \begin{bmatrix} \cos(\frac{n-1}{n}(t - \bar{s}^{(k)})) \\ -\sin(\frac{n-1}{n}(t - \bar{s}^{(k)})) \end{bmatrix} + \|y^{(k)}(j,0) - \bar{y}^{(k)}(0)\| \begin{bmatrix} \cos(\frac{1}{n}(t - s^{(k)}(j))) \\ \sin(\frac{1}{n}(t - s^{(k)}(j))) \end{bmatrix} \quad (128)$$

for appropriate phase offsets \bar{s} and s .

The continuous population equation can be solved by taking the limit as n approaches infinity. As n approaches infinity the motion of the centroid approaches the self-play solution, while the rate of backward rotation about the centroid converges to zero. Thus, in the limit as n approaches infinity, every agent in the population obeys a self-play solution centered at their initial offset from the population centroid. Equivalently, the entire distribution translates in a circular orbit at a unit rate. So, while averaging against the past, as in fictitious self play, leads to rapid outward pressure towards the boundary of Ψ , averaging over a current population leads to a bulk dynamic that is, in essence, equivalent to self-play.

Like the self-play, and fictitious self-play solutions, the epicyclic solution (128) only applies when all the individuals are on the interior of Ψ .

The three solutions developed above are each special cases of more general dynamics. For example, changing the kernel used for time-averaging would change the fictitious self-play solution, as would any interaction with the boundary of Ψ . Nevertheless, these results provide simple intuition, and demonstrate some of the analytic advantages of the disc game payout. In each case, we used bilinearity to replace an average over optimal response directions with the optimal response direction to an average agent, used the block structure of the disc game to isolate separate pairs of coordinates that evolve independently, then used the linearity of the optimal response vector field, to solve the associated dynamical system. More realistic dynamics may not admit such clean analysis, however, the advantages leveraged here remain useful, even when the dynamic cannot be explicitly solved, or decoupled between pairs of coordinates. We use a detailed case study of the replicator equation to illustrate the efficacy of the disc-game embedding for dynamics that do not admit such easy solutions.

Our choice to focus on interior solutions is limiting since the particulars of different games are encoded in the boundaries of Ψ , and, if the dynamic was defined in a different initial coordinate system, in the metric tensor associated with the coordinate change. For example, if the underlying game is perfectly transitive, then Ψ is a line segment, in which case Ψ has no proper interior. Developing explicit solutions for different boundary geometries is an interesting avenue for future work.

6.5.2 Replicator Dynamics: The Transitive Case

Suppose that $f(x, x') = r(x) - r(x')$ for some rating function $r : \Omega \rightarrow \mathbb{R}$. Then, f is a perfectly transitive function [205]. When f is perfectly transitive, the replicator equation can be solved exactly for any initial distribution. In particular (see Appendix section 6.5.2),

$$\pi(x, t) \propto e^{tPr(x)}\pi(x, 0), \text{ where } P = P(0). \quad (129)$$

To confirm the solution, differentiate with respect to time. Let $Z(t) = \int_{x \in \Omega} e^{tPr(x)}\pi(x, 0)dx$ be the proportionality factor needed to hold the population fixed. Then:

$$\partial_t \pi(x, t) = \partial_t \frac{e^{tPr(x)}}{Z(t)} \pi(x, 0) = (r(x) - \partial_t \log(Z(t))) \frac{e^{tPr(x)}}{Z(t)} \pi(x, 0) = (r(x) - \partial_t \log(Z(t))) \pi(x, t).$$

Expanding the logarithmic derivative:

$$\partial_t \log(Z(t)) = \frac{\partial_t Z(t)}{Z(t)} = \int_{x \in \Omega} \frac{\partial_t e^{tPr(x)}}{Z(t)} \pi(x, 0) = \int_{x \in \Omega} r(x) \frac{e^{tPr(x)}}{Z(t)} \pi(x, 0) = \mathbb{E}_{X' \sim \pi(\cdot, t)}[r(X')].$$

Then:

$$\partial_t \pi(x, t) = (r(x) - \partial_t \log(Z(t))) \pi(x, t) = \mathbb{E}_{X' \sim \pi(\cdot, t)}[r(x) - r(X')] \pi(x, t) = \mathbb{E}_{X' \sim \pi(\cdot, t)}[f(x, X')] \pi(x, t)$$

so the proposed solution satisfies the replicator equation. \square

6.6 Generic Parameter Dynamics Proofs

6.6.1 Proof of Result 1

Suppose that (Ω, f) is a functional form game, that $\pi(x, 0)$ is a density over Ω with support S . Suppose that f is of finite rank on S . Finally, suppose that $\pi(x, t)$ is subject to the continuous replicator equation. Parameterise $\pi(x, t)$ according to the exponential family parameterisation with basis functions set to the disc game embedding functions for f and some reference measure ν supported on S . Then, the parameter dynamics must solve the closure equation. Simplifying the closure equation gives:

$$\sum_{k=1}^{\text{rank}(F)} y_1^{(k)}(x) \frac{d}{dt} \theta_1^{(k)}(t) + y_2^{(k)}(x) \frac{d}{dt} \theta_2^{(k)}(t) = \sum_{k=1}^{\text{rank}(F)/2} y_1^{(k)}(x) \mathbb{E}_{X \sim \pi(\cdot; \theta(t))}[y_2^{(k)}(X)] - y_2^{(k)}(x) \mathbb{E}_{X \sim \pi(\cdot; \theta(t))}[y_1^{(k)}(X)] \text{ for all } x \in S. \quad (130)$$

Notice that both the left and right hand sides are a linear combination of the disc game embedding functions. Since the disc game embedding functions are linearly independent, the left and right hand sides match for all $x \in S$ if and only if the coefficients of the combination on the left and right-hand sides match.

Therefore, the parameters obey the closed system of $r = \text{rank}(F)$ ODE's:

$$\text{for all } k \in [1, 2, \dots, \text{rank}(F)/2] \quad \begin{cases} \frac{d}{dt} \theta_1^{(k)}(t) = \mathbb{E}_{X \sim \pi(\cdot; \theta(t))}[y_2^{(k)}(X)] \\ \frac{d}{dt} \theta_2^{(k)}(t) = -\mathbb{E}_{X \sim \pi(\cdot; \theta(t))}[y_1^{(k)}(X)] \end{cases} \quad (131)$$

Expressed vector-wise:

$$\frac{d}{dt} \begin{bmatrix} \theta_1^{(k)}(t) \\ \theta_2^{(k)}(t) \end{bmatrix} = R \mathbb{E}_{X \sim \pi(\cdot; \theta(t))} \begin{bmatrix} y_1^{(k)}(X) \\ y_2^{(k)}(X) \end{bmatrix} \text{ for all } k \in [1, 2, \dots, \text{rank}(F)/2]. \quad (132)$$

where R is the two-by-two ninety degree rotation matrix.

Let $\theta^{(k)} = [\theta_1^{(k)}, \theta_2^{(k)}]$ and $y^{(k)} = [y_1^{(k)}, y_2^{(k)}]$. Then, recall that $v(y) = Ry$ is the optimal self-play vector field for a disc game. Then, the parameter dynamics can be expressed blockwise:

$$\frac{d}{dt} \theta^{(k)}(t) = \mathbb{E}_{X \sim \pi(\cdot; \theta(t))}[v(y^{(k)}(X))] = v(\bar{y}^{(k)}(t)) \text{ where } \bar{y}^{(k)}(t) = \mathbb{E}_{X \sim \pi(\cdot; \theta(t))}[y^{(k)}(X)] \quad (133)$$

for all $k \in [1, 2, \dots, \text{rank}(F)/2]$.

Equation (133) shows that the parameters, θ always move in the best response direction to the centroid of the population in the disc game coordinates. In a sense, this equation imitates simultaneous gradient ascent in the disc-game space, whose right-hand side is also equal to the optimal self-play vector field of the population centroid in the disc game space (see Appendix section 6.5). We will show that, like simultaneous gradient ascent for normal form games on finite strategy spaces, equation (133) defines a Hamiltonian dynamic, so may be interpreted as a system of coupled nonlinear oscillator equations.

To show that the parameter dynamics are Hamiltonian, define the total population function:

$$P(\theta) = \int_{x \in \Omega} \pi(x; \theta) dx = \int_{x \in \Omega} \pi(x, 0) \exp\left(\sum_{k=1}^r \theta^{(k)} \cdot y^{(k)}(x)\right) = \mathbb{E}_{y^{(k)}(X) | X \sim \pi(\cdot, 0)}[\exp\left(\sum_{k=1}^r \theta^{(k)} \cdot y^{(k)}(X)\right)] = M_{\pi_y(\cdot, 0)}[\theta]. \quad (134)$$

The total population function is the Laplace transform of the initial population distribution in the disc game coordinates. The Laplace transform of a distribution is its moment generating function, M . Derivatives of a moment generating function recover moments of the corresponding distribution.

In this case, the centroid of the population distribution in the disc-game space can be expressed as the gradient of the total population function with respect to the parameters θ :

$$\begin{aligned}\bar{y}_j(t) &= \mathbb{E}_{X \sim \pi(\cdot, t)}[y_j(X)] = \int_{x \in \Omega} y_j(x) \pi(x, t) dx = \int_{x \in \Omega} y_j(x) e^{\sum_i \theta_i(t) y_i(x)} \pi(x, 0) dx \\ &= \int_{x \in \Omega} \partial_{\theta_j} e^{\sum_i \theta_i(t) y_i(x)} \pi(x, 0) dx = \int_{x \in \Omega} \partial_{\theta_j} \pi(x; \theta(t)) dx = \partial_{\theta_j} P(\theta).\end{aligned}$$

Therefore:

$$\bar{y}(t) = \mathbb{E}_{X \sim \pi(\cdot; \theta(t))} = \nabla_{\theta} P(\theta(t)). \quad (135)$$

Then, equation (133) admits the block Hamiltonian form:

$$\frac{d}{dt} \theta_1^{(k)}(t) = \partial_{\theta_2^{(k)}} P(\theta(t)), \quad \frac{d}{dt} \theta_2^{(k)}(t) = -\partial_{\theta_1^{(k)}} P(\theta(t)) \quad (136)$$

for all $k \in [1, 2, \dots, \text{rank}(F)/2]$.

In vector form:

$$\frac{d}{dt} \theta(t) = U \nabla_{\theta} P(\theta(t)) \quad (137)$$

where U is the block diagonal matrix with two-by-two diagonal blocks equal to the ninety-degree rotation matrix R . \square

6.6.2 Properties of the Hamiltonian Dynamic

1. **Population Conservation:** Multiplication by U rotates by 90 degrees. So, $\frac{d}{dt} \theta = U \nabla_{\theta} P(\theta)$ is orthogonal to $\nabla_{\theta} P(\theta)$. Therefore, the parameters always move tangent to the level sets of P , thus conserve P .

2. **Interaction Rate and Hamiltonian:** Let $R(P) = \int \rho(P) / P dP$. Then, for any time-autonomous interaction rate:

$$U \nabla_{\theta} R(P(\theta)) = U \frac{\rho(P)}{P} \nabla_{\theta} P(\theta) = \frac{\rho(P)}{P} U \nabla_{\theta} P(\theta) = \frac{d}{dt} \theta(t). \quad (138)$$

3. **Volume Preservation:** The Hamiltonian dynamics are also volume preserving. Let $\Theta(0)$ denote a closed set of parameters θ , and let $V(0)$ equal the volume of the region enclosed by $\Theta(0)$. Let $\Theta(t)$ denote the set of parameters formed by applying the parameter dynamic to every initial parameter vector in $\Theta(0)$. Let $V(t)$ represent volume of the region enclosed by $\Theta(t)$. Then, $V(t) = V(0)$.

The proof follows via the divergence theorem. The rate of change in the volume at any given time is the integral over the interior of $\Theta(t)$ of the divergence of the vector field defining the rate of change of the parameters. Hamiltonian dynamics are divergence free. In particular,

$$\nabla_{\theta} \cdot \frac{d}{dt} \theta(\theta) = \nabla_{\theta} \cdot U \nabla_{\theta} P(\theta) = \sum_{k=1}^{r/2} \partial_{\theta_1^{(k)}, \theta_2^{(k)}}^2 P(\theta) - \partial_{\theta_2^{(k)}, \theta_1^{(k)}}^2 P(\theta) = 0.$$

Notice that, the divergence vanishes since the divergence restricted to each pair of coordinates vanishes. It follows that, the parameter dynamics are also volume preserving for any set of parameters $\Theta(0)$ that only vary over subsets of the parameter pairs. Suppose, for instance, that the $\theta \in \Theta(0)$ only differ in a subset of the coordinates $J \subseteq [1, 2, \dots, r]$. Let $|J|$ denote the number of coordinate along which the parameters vary, and $V(t)$ denote $2|J|$ dimensional volume. Then, $V(t)$ is also conserved, $V(t) = V(0)$, provided $2k \in J$ if $2k-1 \in J$. That is, if J is a set of consecutive coordinate pairs. In particular, if $\Theta(0)$ is constant in all parameters except for one parameter pair, then $V(t)$ is an area. Then, the parameter dynamics are area preserving in each pair of parameters corresponding to paired disc-game embedding coordinates.

4. **No Bounded Attractors:** A bounded attractor is a set of finite parameters that can be enclosed in an open neighborhood of finite parameters such that any parameter trajectory initialized in the neighborhood converges onto the attractor. This neighborhood is the basin of attraction of the attractor. Since the dynamic is volume preserving, no such attractor can exist.

The proof proceeds by contradiction. Suppose such an attractor existed. Initialize a set of parameters inside the basin of attraction of the attractor, $\Theta(0)$, such that $\Theta(0)$ encloses the entire attractor. Then, by definition of the basin of attraction, $\Theta(t)$ converges to the attractor as time progresses. However, since the attractor is contained inside $\Theta(0)$, $\Theta(t)$ cannot converge onto the attractor without losing volume.

It follows that no isolated, finite attractor exists. If such an attractor existed it would admit a Lyapunov function, and level sets of the Lyapunov function would enclose a volume that would have to both decrease over time, and remain constant.

- 5. Interior Steady States are Critical Points:** The rotation matrix U is unitary, so its kernel is trivial. Therefore, $U\nabla_\theta P(\theta) = 0$ if and only if $\nabla_\theta P(\theta) = 0$. So, all interior equilibria correspond to critical points of the total population function.

All interior steady state distributions correspond to interior equilibria since the set of basis functions are linearly independent. Therefore, if $\frac{d}{dt}\theta(t) \neq 0$, then $\partial_t\pi(x, t) \neq 0$ for some x .

By manipulating the Jacobian of $U\nabla_\theta P(\theta)$ about a critical point, we can show that maxima and minima of $P(\theta)$ correspond to centers in the parameter dynamics, while saddle points remain saddles. Minima correspond to centers whose direction of rotation aligns with the self-play vector field. Maxima correspond to centers whose direction of rotation runs backwards against the self-play vector field.

- 6. Convexity:** The total population function $P(\theta)$ is convex.

Proof: The entries of the Hessian are,

$$\partial_{\theta_i, \theta_j}^2 P(\theta) = \int_{x \in \Omega} \partial_{\theta_i, \theta_j}^2 e^{\theta \cdot y(x)} \pi(x, 0) dx = \int_{x \in \Omega} y_i(x) y_j(x) e^{\theta \cdot y(x)} \pi(x, 0) dx = \int_{x \in \Omega} y_i(x) y_j(x) \pi(x; \theta) dx.$$

Therefore, the Hessian of the total population function at θ equals both the Gram matrix for the embedding functions $y(\cdot)$ with respect to the inner product $\langle \cdot, \cdot \rangle_{\pi(\cdot; \theta)}$, and, equivalently, the covariance in the embedding coordinates of an agent drawn at random from the population specified by the parameters θ :

$$[\partial_\theta^2 P(\theta)]_{ij} = \langle y_i(\cdot), y_j(\cdot) \rangle_{\pi_Y(\cdot; \theta)} = \text{Cov}_{Y \sim \pi_Y(\cdot; \theta)}[Y_i, Y_j]. \quad (139)$$

Both Gram matrices and covariance matrices are positive semi-definite. It follows that the Hessian of the total population function is positive semi-definite at all θ . Any function whose Hessian is positive semi-definite at all inputs is convex. Therefore, $P(\theta)$ is a convex function. \square

- 7. Strict Convexity:** More strongly, the total population function is strictly convex.

Proof: Recall, a function is strictly convex if its Hessian is positive definite everywhere. Since P is convex, it is sufficient to show that the Hessian is invertible everywhere.

The Hessian at θ equals the Gram matrix $\langle y(\cdot), y(\cdot) \rangle_{\pi_Y(\cdot; \theta)}$. Recall that, for any finite θ , $\pi(\cdot; \theta)$ is supported everywhere on S , the support of π_0 . Recall also that the embedding functions y depend on the reference measure ν , but changing the reference measure amounts to a linear change of basis when the support of the measure is fixed. Since ν was chosen with support S , we can change measure from ν to $\pi(\cdot; \theta)$.

Let ϕ denote the coordinates of the new parameterization. Then, $\theta(\phi) = (Y_{\pi(\cdot; \theta)}^{(\nu)})^\top \phi$. Since the matrix Y_μ^ν responsible for the change of basis is invertible, the Hessian of $P(\theta)$ is invertible at $\theta(\phi)$ if and only if the Hessian of $P(\phi)$ is invertible.

The Hessian of $P(\phi)$ is the Gram matrix for the embedding functions defined with the reference measure $\pi(\cdot; \theta)$. The embedding functions are always orthogonal under the inner product defined by their reference measure. So, the Hessian for $P(\phi)$ is diagonal entries, with diagonal entries equal to the norms of the embedding functions. Since none of the embedding functions are identically zero, the Hessian of $P(\phi)$ is diagonal with nonzero entries, so is invertible. It follows that the Hessian of $P(\theta)$ is invertible for all finite θ , so P is strictly convex. \square

- 8. All Interior Equilibria are Isolated:** The total population function is strictly convex. Any interior critical point of a strictly convex function is a unique, isolated, global minimizer of the function.

- 9. Radially Unbounded:** A function is radially unbounded if it diverges whenever the norm of the argument diverges. So, to show that $P(\cdot)$ is radially unbounded, it is enough to show that $P(\theta)$ diverges if $\|\theta\|$ diverges. We will show that, if an interior equilibrium exists, then $P(\cdot)$ is radially unbounded.

Proof: If an interior equilibrium exists, then $P(\cdot)$ is strictly convex and admits a global minimizer. A strictly convex function is strictly convex along any line segment. Consider a ray in parameter space, of the form $\theta(s) = \theta_* + \phi s$ for $s \geq 0$ and any $\phi \in \mathbb{R}^r$. Since θ_* is an isolated equilibrium, it is the unique global minimizer of $P(\cdot)$. It follows that $P(\theta(s))$ must be monotonically increasing in s . Any monotonically increasing, strictly convex function of s must diverge as s approaches infinity. Therefore, $P(\cdot)$ diverges along every ray leaving the equilibrium. Since the equilibrium is on the interior, and thus $\|\theta_*\| < \infty$, $P(\cdot)$ also diverges along every ray leaving the origin. Thus, $P(\cdot)$ is radially unbounded. \square

- 10. The Interior is Invariant:** If $P(\cdot)$ is radially unbounded, then its level sets are all bounded, as, for any $p \in [P(\theta_*), \infty)$, there exists a minimal radius $R(p)$ such that $P(\theta) \leq p$ for any θ such that $\|\theta\| \leq R(p)$. Bounded convex sets are compact, so, when an interior equilibrium exists, the level sets of P are boundaries of nested compact regions. It follows that, if an interior equilibrium exists, then the parameters mix on compact sets.

11. **All Interior Equilibria are Centers:** The total population function $P(\theta)$ is strictly convex so does not admit any maxima or saddles. Moreover, any interior local minimizer of $P(\theta)$ is an isolated global minimizer. Minima of $P(\theta)$ correspond to centers, since rotating the gradient vector field about a minimizer by ninety-degrees produces a circulating vector field that circulates in the direction of the applied rotation. More strongly, if an interior equilibrium exist, then perturbations off the equilibrium mix on the boundaries of nested compact sets containing the equilibrium, so neither converge to or diverge from the equilibrium. This remains true for perturbations of any finite size. In other words, all interior equilibria are globally Lyapunov stable centers.
12. **Poincaré Recurrence:** If $P(\theta)$ admits a global minimizer then, as it is strictly convex, it is radially unbounded, so its level sets are bounded. Any volume-preserving dynamic whose trajectories are restricted to bounded sets is Poincaré recurrent [41, 180]. \square

6.6.3 Proof of Corollary 5.3

Corollary 5.3: [Interior Equilibria, Invariant Interior, and the Range of the Embedding] Suppose that (Ω, f) is a functional form game with skew-symmetric f , S is a support set, and $\pi(x, t)$ is a population distribution supported on S evolving according to the replicator equation. Let ν denote an arbitrary reference measure that has support equal to S . Let $\Psi = \{y_\nu(x) | x \in S\} \subseteq \mathbb{R}^r$ denote the image of the set S under the disc-game embedding.

Suppose that Ψ is bounded. Then, the replicator parameter dynamics admit an interior equilibrium, and the parameters mix on compact sets, if and only if the convex hull of Ψ contains the origin in its interior.

Proof: Note that θ_* is an equilibrium if and only if $v(\bar{y}[\pi(\cdot; \theta_*)]) = U\bar{y}[\pi(\cdot; \theta_*)] = 0$. The rotation matrix U is invertible, so θ_* is an equilibrium if and only if the corresponding centroid, $\bar{y}_* = \bar{y}[\pi(\cdot; \theta_*)] = 0$. In other words, θ_* is an equilibrium if and only if the corresponding distribution of traits is centered after embedding in the disc-game space. Recall that the origin plays a special role in disc-games since it corresponds to a neutral type that has no advantage or disadvantage over any other type.

Let $\Psi = \{y(x) | x \in S\}$ denote the image of the support set S after embedding. Let $\text{Conv}(\Psi)$ denote the convex hull of the image of S after embedding. If $\text{Conv}(\Psi)$ does not contain the origin, then no distribution over Ψ is centered. More strongly, since, for finite θ , $\pi_Y(\cdot; \theta)$ is supported on all of S , $\bar{y}(\theta)$ is contained in the interior of the convex hull of Ψ for all finite θ . Therefore, if the interior of the convex hull of Ψ does not contain the origin, then the parameter dynamics do not admit interior equilibrium.

To show the converse, we will show that $P(\cdot)$ is radially unbounded if the origin is contained inside the interior of the convex hull of Ψ . If $P(\cdot)$ is radially unbounded, then it cannot admit an infimum at infinity, must have compact level sets, and must admit an interior global minimizer.

First, $P(\cdot)$ is radially unbounded if and only if $\log(P(\cdot))$ is radially unbounded since $P(\cdot)$ is positive valued, and $\log(P)$ is monotonically increasing and unbounded as P diverges.

To show that $\log(P)$ is radially unbounded, note that $\nabla_\theta \log(P(\theta)) = \frac{1}{P(\theta)} \mathbb{E}_{Y \sim \pi(\cdot; \theta)}[Y]$ is the centroid of the normalized trait distribution in the embedding space. Let $\{\theta(s) = s\phi\}_{s>0}$ denote a ray leaving the origin. Then, let $p(\cdot; s) = \frac{1}{P(\theta(s))} \pi_Y(\cdot; \theta(s))$. Let $\bar{y}(s) = \mathbb{E}_{Y \sim p(\cdot; s)}[Y]$. Then, since $\pi_Y(y; \theta) = \pi_Y(y, 0) \exp(\theta \cdot y)$, $\lim_{s \rightarrow \infty} \bar{y}(s) = y_*(\Psi, \phi)$ where $y_*(\Psi, \phi)$ is the intersection of the boundary of the convex hull of Ψ with the ray $\{y(s) = \phi s\}_{s>0}$. Then, if ϕ is normalized, $y_*(\Psi, \phi) = d(\Psi, \phi)\phi$ where $d(\Psi, \phi)$ is the distance from the origin to the boundary of Ψ along the direction ϕ . By assumption, $d(\Psi, \phi) > 0$ for all ϕ .

Choose ϕ normalized. Then, for any ray leaving the origin of parameter space along the direction ϕ , $\lim_{s \rightarrow \infty} \partial_s \log(P(\theta(s))) = \lim_{s \rightarrow \infty} \bar{y}(s) \cdot \phi = d(\Psi, \phi) > 0$. Then, along every ray leaving the origin, there exists a sufficiently large distance $s_*(\Psi, \phi)$ such that, $\partial_s \log(P(\theta(s))) > 0$ for all $s > s_*(\Psi, \phi)$. Therefore, $\log(P(\theta))$ diverges along all rays leaving the origin, so is radially unbounded. It follows that P is radially unbounded. \square

6.6.4 Proof of Theorem 6

Theorem 6: Suppose that π obeys a generalized replicator dynamic with frequency dependent growth rates $g(\pi)$. Let h^{-1} be the inverse of $h = \int^\pi 1/g(s)ds$. Then, $h^{-1}(u) = \sum_{n=0}^{\infty} h_n u^n$, the parameters θ obey a Hamiltonian dynamic of the form:

$$\frac{d}{dt} \theta(t) = U \mathbb{E}_{X \sim \pi(\cdot; \theta(t))}[y(X)] = U \nabla_\theta H(\theta) \quad (140)$$

with Hamiltonian:

$$H(\theta) = \int_{x \in \Omega} \int_0^{\theta \cdot y(x) + u_0(x)} h^{-1}(u) du \quad (141)$$

Proof: Expanding f using the disc game decomposition gives:

$$y(x) \cdot \frac{d}{dt} \theta(t) = y(x) \cdot U \int_{x' \in \Omega} y(x') h^{-1}(\theta(t) \cdot y(x') + u_0(x)) dx'$$

Since the disc game embedding functions are linearly independent, the coefficients defining the combination of embedding functions on each side must match. Therefore:

$$\frac{d}{dt} \theta(t) = U \int_{x'} y(x') h^{-1}(\theta(t) \cdot y(x') + u_0(x)) dx' = U \int_{x'} y(x') \pi(x'; \theta(t)) dt = U \mathbb{E}_{X' \sim \pi(\cdot; \theta(t))}[y(X')] = v(\bar{y}(\theta(t))).$$

So, just as before, the rate of change in the parameters equal the optimal response vector evaluated at the centroid of the embedded population. This result will always hold when $\mathbb{E}_{X' \sim \pi}[f(x, X')]$ is expanded onto the embedding functions.

It remains to show that the expectation can be expressed as the gradient of a scalar-valued function of θ . Define the function:

$$H(\theta) = \int_{x \in \Omega} \int_0^{\theta \cdot y(x) + u_0(x)} h^{-1}(u) du \quad (142)$$

Then:

$$\partial_{\theta_j} H(\theta) = \int_{x \in \Omega} \partial_p \left[\int_0^p h^{-1}(u) du \right]_{p=\theta \cdot y(x) + u_0(x)} \partial_{\theta_j} \theta \cdot y(x) dx = \int_{x \in \Omega} h^{-1}(\theta \cdot y(x) + u_0(x)) y_j(x) dx = \mathbb{E}_{X \sim \pi(\cdot; \theta(t))}[y(X)].$$

It follows that the generalized replicator dynamic is Hamiltonian in its parameters with Hamiltonian H . \square

6.6.5 Proof of Lemma 20

Lemma 20: [Strict Convexity] If f is bounded, $0 \leq g(\pi)$ with equality if and only if $\pi = 0$, and $g(\pi) = \mathcal{O}(\pi)$ as $\pi \rightarrow 0$ and $\pi \rightarrow \infty$, then the Hamiltonian function $H(\cdot)$ is a strictly convex, second-differentiable function of θ for all interior θ .

Proof: The Hamiltonian has Hessian entries are:

$$\begin{aligned} \partial_{\theta_i}^2 H(\theta) &= \partial_{\theta_i} \int_{x \in \Omega} h^{-1}(\theta \cdot y(x) + u_0(x)) y_j(x) dx \\ &= \int_{x \in \Omega} \frac{d}{du} h^{-1}(u)|_{u=\theta \cdot y(x) + u_0(x)} y_i(x) y_j(x) dx = \int_{x \in \Omega} \left[\frac{d}{dp} h(p) \right]_{p=h^{-1}(\theta \cdot y(x) + u_0(x))}^{-1} y_i(x) y_j(x) \\ &= \int_{x \in \Omega} [1/g(p)]_{p=h^{-1}(\theta \cdot y(x) + u_0(x))}^{-1} y_i(x) y_j(x) dx = \int_{x \in \Omega} g(h^{-1}(\theta \cdot y(x) + u_0(x))) y_i(x) y_j(x) \\ &= \int_{x \in \Omega} g(\pi(x; \theta)) y_i(x) y_j(x). \end{aligned}$$

It follows that, the Hessian of the Hamiltonian is:

$$\mathbf{H}_{ij}(\theta) = [\partial_{\theta}^2 H(\theta)]_{i,j} = \int_{x \in \Omega} g(\pi(x; \theta)) y_i(x) y_j(x) dx. \quad (143)$$

if the integral converges.

If f is bounded, then the embedding functions are bounded so $y(X)$ has finite moments of all order, for any distribution defined on Ω . It remains to show that the integral defines a finite, positive-definite matrix for all interior θ reachable by the dynamic in finite time. Showing that the integral is finite establishes second-differentiability. Showing that it is positive definite establishes strict convexity.

Suppose that $g(\pi)$ is chosen to be $\mathcal{O}(\pi)$. Then, the support of π is preserved under the dynamic, and $\int_{x \in \Omega} g(p(x)) dx < \infty$ whenever $\int_{x \in \omega} p(x) dx < \infty$ for all density functions p since $\int_{x \in \Omega} g(p(x)) dx < C \int_{x \in \omega} p(x) dx$ for some $C < \infty$. Define:

$$P^g(\theta) = \int_{x \in \Omega} g(\pi(x; \theta)) dx, \quad \pi^g(x; \theta) = \frac{1}{P^g(\theta)} g(\pi(x; \theta)).$$

Then, since g is nonnegative valued, $\pi^g(x; \theta)$ is a normalized density function. Then:

$$\mathbf{H}(\theta) = P^g(\theta) \mathbb{E}_{X \sim \pi^g(\cdot; \theta)}[y(X) y(X)^T] \succeq P^g(\theta) \text{Cov}_{X \sim \pi^g(\cdot; \theta)}[y(X)] = P^g(\theta) \text{Cov}_{Y \sim \pi_y^g(\cdot; \theta)}[Y].$$

It follows that \mathbf{H} is positive definite whenever the covariance of $Y \sim \pi_y^g(\cdot; \theta)$ is full rank (has rank r). The covariance is full rank whenever the convex hull of the support of π_y^g is a set of nonzero measure in \mathbb{R}^r . Since $g(p) = 0$ if and only if $p = 0$, the support of π_y^g is the support of π_y . Since g is chosen so that the dynamic is support preserving, the support of π_Y is Ψ . Suppose that the convex hull of Ψ has nonzero measure in \mathbb{R}^r . Then $Y \sim \pi_y^g(\cdot; \theta)$ must as well, so \mathbf{H} must be positive definite for all θ reachable under the dynamic. It follows that H is strictly convex on all θ reachable under the dynamic from an initial condition supported on Ω .

The last constraint on Ψ fails for transitive functions, or if the collection of embedding functions that include the constant function, as when one of the disc game components is perfectly transitive. Since the embedding functions are linearly independent, at most one is constant, so the convex hull of Ψ is, generically contained within a linear subspace of codimension at most one. The latter condition, however, is sufficient, not necessary.

The Hessian is:

$$\mathbf{H}(\theta) = P^g(\theta) \left(\bar{y}^g(\theta) \bar{y}^g(\theta) + \text{Cov}_{Y \sim \pi_y^g(\cdot; \theta)}[Y] \right) \text{ where } \bar{y}^g(\theta) = \mathbb{E}_Y$$

so $v^\top \mathbf{H}(\theta)v = 0$ requires $v \perp \bar{y}^g(\theta)$. Then, $v^\top H(\theta)v = v^\top \text{Cov}_{Y \sim \pi_y^g(\cdot; \theta)}[Y]v = \text{Var}_{Y \sim \pi_y^g(\cdot; \theta)}[v \cdot Y] = 0$ if and only if the convex hull of the support of $\pi_y^g(\cdot; \theta)$ is contained in an affine subspace of codimension one that is *perpendicular* to $\bar{y}^g(\theta)$. Since $\bar{y}^g(\theta)$ is the expected value of Y under $\pi_y^g(\cdot; \theta)$, it is contained inside the convex hull of the support of $\pi_y^g(\cdot; \theta)$. The only linear subspace of codimension one, that is perpendicular to a member of the subspace, is a subspace that passes through zero. However, Ψ cannot be contained in a linear subspace passing through the origin, since such Ψ could be rotated to include a constant function that equals zero for all x . The function $\phi_k(x) = 0$ is never a valid eigenfunction, nor is $y_k(x) = 0$ a valid basis function. It follows that Ψ cannot be contained inside a linear subspace passing through the origin, so $\mathbf{H}(\theta)$ is positive definite, so the Hamiltonian is strictly convex for all θ reachable under the dynamic from an initial condition supported on Ω . \square

6.6.6 Proof of Corollary 6.1

Corollary 6.1: [Recurrence] Suppose that the conditions of Lemma 20 hold. Then, the parameter dynamics either admit an isolated, interior equilibrium that corresponds to the choice of θ such that $\mathbb{E}_{X \sim \pi(\cdot; \theta)}[y(X)] = 0$, so correspond to a fully mixed ESS, or no such equilibrium exists. The former case requires $0 \in \text{conv}(\Psi)$ and is impossible if the convex hull of Ψ does not contain the origin. If an isolated interior equilibrium exists then it is a center and the parameter dynamics are Poincaré recurrent.

Proof: Assume that either condition of Lemma 20 hold, or, more generally, that H is strictly convex and second-differentiable. Then, since the Hessian is strictly convex it either admits a unique, isolated, interior minima, which is an isolated equilibrium for the parameter dynamics, or it does not, and all parameter orbits escape to infinity.

Suppose that there exists an isolated interior equilibrium θ_* . This requires that:

$$\mathbb{E}_{X \sim \pi(\cdot; \theta)}[y(X)] = \nabla_\theta H(\theta) = 0. \quad (144)$$

Under the conditions of Lemma 20, the total population size $P(\theta) = \int_{x \in \Omega} h^{-1}(\theta \cdot y(x) + u_0(x))dx$ is nonzero for finite θ (it only vanishes if $u(x, \theta)$ diverges in finite time, which is impossible when f is bounded, as, if f is bounded, then the rate of change in θ is bounded). It follows that, the unnormalized expectation can only equal zero if the normalized expectation equals zero. The normalized expectation lies inside the convex hull of Ψ , so, if the Hamiltonian admits an interior minima, then the convex hull of Ψ must contain the origin.

Then, since the Hamiltonian is strictly convex and second-differentiable, $H(\theta)$ is bounded from below by its quadratic Taylor approximation about θ_* :

$$H(\theta) \geq H(\theta_*) + \frac{1}{2}(\theta - \theta_*)^\top \mathbf{H}(\theta_*)(\theta - \theta_*) = H^{(2)}(\theta; \theta_*)$$

Since $\mathbf{H}(\theta)$ is positive definite, the quadratic approximation $H^{(2)}(\theta; \theta_*)$ is radially unbounded, so the Hamiltonian is radially unbounded.

It follows that the level sets of the Hamiltonian are the boundaries of a sequence of nested convex sets, and each level set is the boundary of a compact set. Then, the dynamics are both volume-preserving and bounded ($\theta(t)$ remains finite for all t if $\theta(0)$ is finite). In that case, the dynamics are Poincaré recurrent. \square

6.6.7 Proof of Corollaries 7.1 - 7.2

Corollary 7.1: [Recurrence] The parameter dynamics either admit an isolated, interior equilibrium that corresponds to a fully mixed ESS, or no such equilibrium exists. If an isolated interior equilibrium exists, then it is a center and the parameter dynamics are Poincaré recurrent. An isolated interior equilibrium exists if and only if the origin is contained inside the convex hull of Ψ .

Proof: First, Hamiltonian dynamics are volume-preserving, so the parameter dynamics do not admit any bounded attractors or sources. This rules out any isolated interior sinks or sources. So, as for homogeneous mixing, the dynamics may only admit attractors at the boundary, and all isolated interior equilibria are centers.

More strongly, since the Hamiltonian $H(\theta)$ is a sum of finitely many functions which are all strictly convex and second-differentiable, the combined Hamiltonian is strictly convex and second-differentiable. It follows that the system admits at most one isolated interior equilibrium. This equilibrium exists under the same geometric conditions on Ψ established for the homogeneous population case. If the convex hull of Ψ contains the origin, then the component Hamiltonians all admit a unique, isolated global minimizer, so their sum does as well. If the origin lies outside the convex hull of Ψ , then none of the component Hamiltonians admit an interior minimizer, so their sum does not either.

If such an equilibrium exists, then all of the component Hamiltonians are radially unbounded. It follows that the composite Hamiltonian is radially unbounded, so the parameters orbit on the boundaries of compact sets. Then, since the dynamic is volume-preserving, it must also be Poincaré recurrent. \square

Corollary 7.2: [Mixed Adaptive Dynamics] Let $\bar{y}_i(t) = \mathbb{E}_{X \sim \pi_i(\cdot, t)}[y(X)]$. Let $\bar{y}(t) = [\bar{y}_1(t), \bar{y}_2(t), \dots, \bar{y}_K(t)]$ denote the concatenation of all the centroids, $\theta(\bar{y})$ the map from the centroid back to the parameters, and \mathbf{H} the Hessian of the Hamiltonian. Let $v(y) = Uy$ denote the optimal training vector field that directs self-play. Let $m_i = \sum_{j=1}^k m_j$ denote the total per capita interaction rate for patch i . Let $\hat{M} = \text{diag}(m_1^{-1}, m_2^{-1}, \dots, m_K^{-1})M$ denote the row-stochastic version of M .

Then, the centroids obey an autonomous dynamic of the form:

$$\frac{d}{dt}\bar{y}_i(t) = m_i \mathbf{H}(\theta(\bar{y}(t)))v\left([\hat{M}\bar{y}(t)]_i\right). \quad (145)$$

Proof: Let $\bar{y}_i(t) = \mathbb{E}_{X \sim \pi_i(\cdot, t)}[y(X)]$. Then:

$$\begin{aligned} \frac{d}{dt}\bar{y}_i(t) &= \frac{d}{dt}\nabla_{\theta_i} H^{(i)}(\theta^{(i)}(t), u_0^{(i)}) = \mathbf{H}^{(i)}\left(\theta^{(i)}; u_0^{(i)}\right) \frac{d}{dt}\theta^{(i)}(t) = \mathbf{H}^{(i)}\left(\theta^{(i)}; u_0^{(i)}\right) \sum_{j=1}^K m_{ij} U \bar{y}_j(t) \\ &= \mathbf{H}^{(i)}\left(\theta^{(i)}; u_0^{(i)}\right) \sum_{j=1}^K m_{ij} v(\bar{y}_j(t)) = m_i \mathbf{H}^{(i)}\left(\theta^{(i)}; u_0^{(i)}\right) v\left(\frac{1}{m_i} \sum_{j=1}^K m_{ij} \bar{y}_j(t)\right) \end{aligned} \quad (146)$$

where $m_i = \sum_{j=1}^K m_{ij}$.

As before, this equation appears non-autonomous, since the covariance is expressed as a function of the parameters. However, since the component Hamiltonians are all strictly convex, the mapping between the parameters in each patch, and the embedding centroid of each patch, is invertible. \square

6.7 Solutions for Special Geometries

6.7.1 Decoupled Parameter Dynamics

Let $y^{(k)} = [y_{2k-1}, y_{2k}]$ denote the pair of disc game coordinates assigned to the k^{th} game. Similarly, let $\theta^{(k)} = [\theta_{2k-1}, \theta_{2k}]$ denote the parameters associated with the k^{th} game. When do the parameter dynamics, $\theta(t)$ decouple blockwise so that $\theta^{(k)}$ evolves independently of $\theta^{(k')}$ for all $k \neq k'$?

Lemma 17: [Decoupled Parameter Dynamics] Consider $Y \sim \pi_0$. If the set of component disc games $D = \{1, 2, \dots, r/2\}$ admits a partition $K = \{K_j\}$ where $K_j \subseteq D$, $\cup_j K_j = D$, and $K_j \cap K_{j'} = \emptyset$ if $j \neq j'$ such that $Y^{(k)}$ is independent of $Y^{(k')}$ when k and k' are members of different sets in the partition, then each set of parameters, $\Theta_j = \{\theta^{(k)}, k \in K_j\}$ evolves independently of each other set of parameters, and, if $Y \sim \pi(\cdot, t)$, then $\{Y^{(k)}, k \in K_j\}$ remain independent of $\{Y^{(k')}, k \in K_{j'}\}$ for all t .

Proof: Suppose that the independence assumption holds. Then, π_0 factors, $\pi_0(y) = \prod_j \pi_0^{(K_j)}(y^{(K_j)})$ and, by necessity, that Ψ can be expressed as a Cartesian product, $\Psi = \Psi^{(K_1)} \times \Psi^{(K_2)} \times \dots = \prod_j \Psi^{(K_j)}$. Then, the Hamiltonian $P(\theta)$ also factors:

$$P(\theta) = \int_{y \in \Psi} e^{\theta \cdot y} d\pi_0(y) = \prod_j \int_{y^{(K_j)} \in \Psi^{(K_j)}} e^{\theta^{(K_j)} \cdot y^{(K_j)}} d\pi_0^{(K_j)}(y^{(K_j)}) = \prod_j P^{(K_j)}(\theta^{(K_j)}). \quad (147)$$

Since the Hamiltonian is conserved under the dynamic, scaling time by the population size scales the dynamical system by $P(\theta)$, which replaces the Hamiltonian with $\log(P(\theta))$. If we can show that the time scaled dynamics decouple, then the original dynamics must also decouple.

Consider the time-scaled dynamics:

$$\frac{d}{dt}\theta(t) \propto U\nabla_\theta \log(P(\theta)) = U\nabla_\theta \sum_j \log(P^{(K_j)}(\theta^{(K_j)})) = U \sum_j \nabla_\theta \log(P^{(K_j)}(\theta^{(K_j)})) \quad (148)$$

Since $\log(P^{(K_j)}(\theta^{(K_j)}))$ depends on only the parameter blocks $k \in K_j$, its gradient is zero for all blocks $k' \notin K_j$. Since U only relates pairs of rows belonging to the same block, multiplication by U preserves this property. It follows that $\frac{d}{dt}\theta^{(K_j)}(t)$ depends only on the parameters $\theta^{(K_j)}$, and does not depend on any parameters belonging to a different component of the partition.

If $Y \sim \pi(\cdot, t)$, then separate components of the partition remain independent at all t since $\pi(y, t) = e^{\theta(t) \cdot y} \pi_0(y) = \prod_j e^{\theta^{(K_j)}(t) \cdot y^{(K_j)}} \pi_0^{(K_j)}(y^{(K_j)})$. \square

Corollary 17.1: [Fully Decoupled Parameter Dynamics] Consider $Y \sim \pi_0$. If $Y^{(k)}$ is independent of $Y^{(k')}$ for all $k \neq k'$, then each pair of parameters $\theta^{(k)}$ evolves independently of each other pair of parameters, and, if $Y \sim \pi(t)$, then $Y^{(k)}$ remains independent of $Y^{(k')}$ for all t .

Proof: The proof is a direct application of Lemma 17. \square

Corollary 17.2: [Fully Decoupled Parameter Dynamics are Quasi-Periodic] Consider $Y \sim \pi_0$. If $Y^{(k)}$ is independent of $Y^{(k')}$ for all $k \neq k'$, and the origin is contained in the convex hull of the support of π_0 , then each pair of parameters $\theta^{(k)}(t)$ evolve periodically.

Proof: The proof is immediate from previous results. If the assumptions of 18.1 hold, then each pair of parameters evolve independently. If the origin is contained in the convex hull of π_0 , then there exists a fully mixed Nashed equilibria with the same support as the initial distribution, so, by Corollary 5.3, the dynamics are recurrent. All two-dimensional, recurrent dynamics are periodic. \square

6.7.2 Rotationally Symmetric Base Measures

Lemma 18: [Rotational Symmetry] Let $\Psi \in \mathbb{R}^2$ be the domain for a single disc game, and suppose that π_0 is an initial population distribution that is an exponential tilting of a base measure ν that is symmetric under rotations. If the population evolves according to a continuous replicator dynamic, then its parameters, using base measure ν , orbit on concentric circles at a constant rate with period equal to $2\pi\|\theta(0)\|_2/\|\nabla_\theta P(\theta(0))\|_2$. In the parametrization with base measure π_0 , the parameters orbit on concentric ellipses with period $2\pi\|\theta(0)\|_2/\|\nabla_\theta P(\theta(0))\|_2$.

Proof: If π_0 is an exponential tilt of ν where ν is rotationally symmetric then, with ν as the base measure, $P(\theta)$ must be invariant to rotations of θ , so is also only a function of $\|\theta\|_2$. Then, since $P(\theta)$ is the Hamiltonian, and Hamiltonian dynamics follow level sets of their Hamiltonian function, the parameters must evolve along concentric circles. The period of each orbit is equal to the circumference of the circle divided by the rate of travel about the circle, which must also be constant, and equals $\|\nabla_\theta P(\theta)\|_2$ at θ_0 . Changing base measure corresponds to a linear transformation of the parameter space, so, using any other base measure that is absolutely continuous with respect to ν produces concentric elliptical orbits. Periods of orbits are invariant under invertible changes of coordinates. \square

6.7.3 Separable Base Measures

The continuous replicator dynamics simplify further in the case when every component of $Y \sim \pi_0$ are independent. Then, π_0 fully factors into a product of marginals:

$$\pi_0(y) = \prod_{j=1}^r \pi_0^r(y_r). \quad (149)$$

In this case, each separate pair of parameters evolves independently, but the parameters associated with the same disc game remain coupled through the dynamics.

Consider a single disc game, with a base measure that factors into products of measures on its marginals (under some rotation of the coordinates). For example, we could use π_0 equal to a uniform distribution on a rectangular region, or a Laplace distribution, or any other product of marginal measures. Then:

$$\pi_0(y) = \pi_0^1(y_1)\pi_0^2(y_2). \quad (150)$$

In this case, the Hamiltonian fully factors:

$$P(\theta_1, \theta_2) = \prod_{j=1}^2 P_j(\theta_j) \text{ where } P_j(s) = \int_{y_j} e^{sy_j} d\pi_0^j(y_j). \quad (151)$$

Notice, $P_j(s)$ is the moment generating function, or Laplace transform of, the marginal measure π_0^j for the j^{th} embedding coordinate.

Now we can write:

$$\partial_{\theta_1} P(\theta) = (\partial_{\theta_1} P_1(\theta_1)) P_2(\theta_2) = \frac{\partial_{\theta_1} P_1(\theta_1)}{P(\theta_1)} P_1(\theta_1) P_2(\theta_2) = (\partial_{\theta_1} \log(P_1(\theta_1))) P(\theta) \text{ and } \partial_{\theta_2} P(\theta) = (\partial_{\theta_2} \log(P_2(\theta_2))) P(\theta). \quad (152)$$

This form is convenient since, $\log(P_j(s)) = K_j(s)$ is the cumulant generating function for the j^{th} marginal of the base measure. Moreover, $P(\theta)$ is conserved by the dynamic, so $P(\theta(t)) = P(\theta(0)) = P$, the total population size. Then, the Hamiltonian parameter dynamics reduce to:

$$\frac{d}{dt} \theta_1(t) = Pg_2(\theta_2), \quad \frac{d}{dt} \theta_2(t) = -Pg_1(\theta_1) \quad (153)$$

where:

$$g_j(s) = \frac{d}{ds} K_j(s) = \frac{d}{ds} \log(P_j(s)) \quad (154)$$

is the slope of the cumulant generating function.

Equation (153) is especially simple when the marginals for Y_1 and Y_2 are identical. Then we can suppress the subscripts, and, scaling time by the total population size, reduce to a normal form, nonlinear oscillator equation:

$$\frac{d}{dt} \theta_1(t) \propto g(\theta_2(t)), \quad \frac{d}{dt} \theta_2(t) \propto -g(\theta_1(t)). \quad (155)$$

In the main text, we considered an example where π_0 is a Laplace distribution, so has marginals $\pi_0^j(y_j) = \frac{1}{2\lambda} \exp(-\lambda|y_j|)$. In this section, we will work out the solution presented in the main text.

The cumulant generating function for a Laplace distribution is:

$$K(s) = -\log(1 - (s/\lambda)^2) \text{ for } |s| < \lambda \quad (156)$$

so:

$$g(s) = \frac{2s/\lambda^2}{1 - (s/\lambda)^2} = 2\frac{s}{\lambda^2 - s^2} \text{ for } |s| < \lambda. \quad (157)$$

Then, the system of coupled equations is:

$$\frac{d}{dt} \theta_1(t) \propto \frac{2}{\lambda^2 - \theta_2(t)^2} \theta_2(t), \quad \frac{d}{dt} \theta_2(t) \propto -\frac{2}{\lambda^2 - \theta_1(t)^2} \theta_1(t). \quad (158)$$

We can simplify further by changing variables via $\theta = \lambda\phi$. Then:

$$\frac{d}{dt} \phi_1(t) \propto \frac{2}{(1 - \phi_2(t)^2)} \phi_2(t), \quad \frac{d}{dt} \phi_2(t) \propto -\frac{2}{(1 - \phi_1(t)^2)} \phi_1(t). \quad (159)$$

where the constant of proportionality (time-rescaling) needed to recover the original dynamic is P/λ^2 . We will from now on, assume without loss of generality, that $\lambda = 1$, and that we are working after rescaling time by P (i.e. using normalized distributions when computing expectations).

Equation (159) defines a separable elliptic oscillator, that, as expected, approaches a simple harmonic oscillator for small ϕ . It is well defined for all ϕ in the unit square $[-1, 1]^2$. The solution trajectories are the level sets of the Hamiltonian function $((1 - \phi_1^2)(1 - \phi_2^2))^{-1}$, so, by conservation of the Hamiltonian, we can solve directly for ϕ_2 as a function of ϕ_1 and visa versa. In particular,

$$\phi_2(\phi_1, P) = \pm \sqrt{1 - (P(1 - \phi_1^2))^{-1}} \quad (160)$$

where P is the value of the Hamiltonian along the level set. The level sets of the Hamiltonian converge to circles near the origin, and to squares as the parameters ϕ approach the boundary of the $[-1, 1]^2$ square.

The dynamical system admits an analytic solution for all initial conditions. To solve the system, rearrange the conservation constraint $(1 - \phi_2^2)^{-1} = P(1 - \phi_1^2)$. Then:

$$\frac{d}{dt} \phi_1(t) = 2 \frac{\phi_2(t)}{1 - \phi_2(t)^2} = 2P\phi_2(t)(1 - \phi_1(t)^2) = \pm 2\sqrt{P(1 - \phi_1^2)(P(1 - \phi_1^2) - 1)} \quad (161)$$

with sign determined by the initial conditions. In particular, $\text{sign}(\frac{d}{dt}\phi_1(t)) = \text{sign}(\phi_2(t))$.

Equation (161) is separable, so $\phi_1(t)$ must satisfy the integral equation:

$$\int^{\phi_1(t)} \frac{du}{2\sqrt{P(1-u^2)(P(1-u^2)-1)}} = t \quad (162)$$

where $P = 1/((1-\phi_1(0)^2)(1-\phi_2(0)^2))$ is the initial population size. The integral on the right hand side is an elliptic integral that can be closed using Jacobi elliptic functions. To close the integral, substitute $u = \sin(\psi)$ so that $1-u^2 = \cos(\psi)^2$ and $du = \cos(\psi)d\psi$. Then $P(1-u^2) = P\cos^2(\psi)$ and $P(1-u^2)-1 = P\cos^2(\psi)-1$. Then, the integral is:

$$\begin{aligned} \int^{\psi_1(t)} \frac{\cos(\psi)d\psi}{2\sqrt{P\cos(\psi)^2(P\cos(\psi)^2-1)}} &= \int^{\psi_1(t)} \frac{d\psi}{2\sqrt{P(P\cos(\psi)^2-1)}} = \int^{\psi_1(t)} \frac{d\psi}{2\sqrt{P(P-1)-P^2\sin(\psi)^2}} \\ &= \frac{1}{2\sqrt{P(P-1)}} \int^{\psi_1(t)} \frac{d\psi}{\sqrt{1-\frac{P}{P-1}\sin(\psi)^2}} \\ &= \frac{k}{2\sqrt{P(P-1)}} \int^{\psi_1(t)} \frac{d\psi}{\sqrt{k^2-\sin(\psi)^2}} = \frac{1}{2P} \int^{\psi_1(t)} \frac{d\psi}{\sqrt{k^2-\sin(\psi)^2}} \end{aligned} \quad (163)$$

where $k^2 = (P-1)/P$ and $\psi_1(t) = \arcsin(\phi_1(t))$.

Finally, to bring the integral into the standard elliptic form, let $\sin(\psi) = k\sin(\zeta)$ so that $\psi = \arcsin(k\sin(\zeta))$, $d\psi = (1-k^2\sin(\zeta)^2)^{-1/2}k\cos(\zeta)d\zeta$ and $(k^2-\sin(\psi)^2)^{-1/2} = (k^2(1-\sin(\zeta)^2))^{-1/2} = k^{-1}\cos(\zeta)^{-1}$. Then, Equation (161) produces the integral equation:

$$\int^{\zeta_1(t)} \frac{d\zeta}{\sqrt{1-k^2\sin(\zeta)^2}} = F(\zeta_1(t)|k^2) = 2Pt + C \text{ where } \zeta_1(t) = \arcsin\left(\frac{\phi_1(t)}{k}\right) \quad (164)$$

and where C is a constant of integration to be determined by the initial conditions and $F(\cdot|\cdot)$ is the incomplete elliptical integral of the first kind.

Inverting, and enforcing the conservation constraint, returns the solution:

$$\phi_1(t) = k\text{sn}(2Pt+C|k^2), \quad \phi_2(t) = \text{sign}(\phi_2(0))k \frac{\text{cn}(2Pt+C|k^2)}{\text{dn}(2Pt+C|k^2)} \text{ for } \begin{cases} P = ((1-\phi_1(0)^2)(1-\phi_2(0)^2))^{-1} \\ k = \sqrt{(P-1)/P} \\ C = F(\arcsin(\phi_1(0)/k|k^2)) \end{cases} \quad (165)$$

and where sn, cn, and dn are the elliptic sine, elliptic cosine, and delta amplitude functions. Each is an example of a Jacobi elliptic function [63, 119].

Since the solutions are all periodic orbits in the square $[-1, 1]^2$ we can adopt convenient initial conditions. Set $\phi_1(0) = 0$ and $\phi_2(0) = a > 0$, where a represents the amplitude of the oscillator where it crosses a coordinate axis. Then $C = 0$, $\text{sign}(\phi_2(0)) = 1$, $P = (1-\phi_2(0)^2)^{-1}$, and $k = \sqrt{1-1/P} = \sqrt{1-(1-\phi_2(0)^2)} = \phi_2(0) = a$. Then, every solution to equation (161) lies on an orbit of the form:

$$\phi_1(t) = a\text{sn}(2Pt|a^2), \quad \phi_2(t) = a \frac{\text{cn}(2Pt|a^2)}{\text{dn}(2Pt|a^2)}. \quad (166)$$

The corresponding orbits are shown in Figure 10 for amplitudes ranging from 0.1 to 0.999. Notice that, small amplitudes produce approximately circular orbits, while large amplitudes produce approximately square orbits. Also notice that, when the amplitude is small, the parameters orbit at an approximately constant velocity while, when the amplitude is large, the parameters dwell near the corners of the square, and move rapidly when near its edges. We explore this heteroclinic behavior in the main text, and in more detail in Appendix section 6.7.4.

The period of the orbits can also be computed in closed form since the Jacobi function $\text{sn}(\cdot|k^2)$ has period $4K(k)$ where $K(k) = F(\pi/2|k^2) = \int_0^{\pi/2} (1-k^2\sin(\zeta)^2)^{-1/2}d\zeta$ is the complete elliptic integral of the first kind.⁶ Therefore, the parameters orbit with amplitude dependent period:

$$T(a) = 2 \frac{K(a)}{P} = 2(1-a^2)K(a) = 2(1-a^2) \int_0^{\pi/2} \frac{d\zeta}{\sqrt{1-a^2\sin(\zeta)}}. \quad (167)$$

For small initial amplitudes $T(a) \simeq \pi$. This period is consistent with the period of the harmonic oscillator $\frac{d}{dt}\phi_1 = 2\phi_2$, $\frac{d}{dt}\phi_2 = -2\phi_1$ generated by linearizing the nonlinear oscillator about the origin. As a approaches 1 from below (large initial amplitudes), the period approaches zero, with $T(a) \sim \frac{1}{2}(1-a^2)|\ln(1-a^2)|$. In this example the period of the oscillator converges to zero as the initial amplitude a approaches 1, since, when a approaches 1, θ approaches λ , so the expected embedding coordinate \bar{y} , whose magnitude sets the rate of motion of the parameters, diverges.

⁶Not the cumulant generating function.

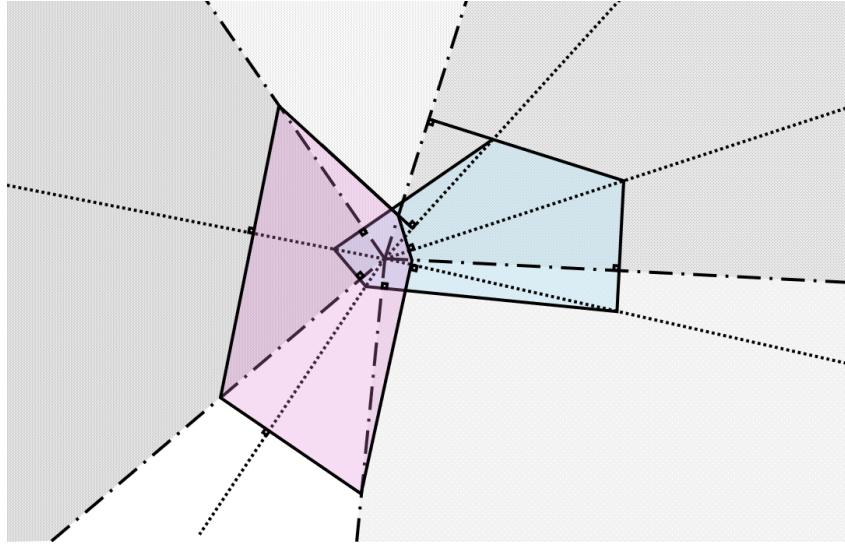


Figure 13: **Left:** Dual polygons (blue and purple) related by polar reciprocation. The vertices of each polygon lie on rays perpendicular to a side of the other polygon. **Right:** Solution trajectories for the parameter dynamics of a population evolving under the continuous replicator dynamic (colored curves, colored by value of the Hamiltonian) with initial measure equal to a uniform distribution over a polygon Ψ (shaded grey region). The Dot-dashed outer polygon is dual to Ψ . Notice that trajectories with large amplitudes (large Hamiltonian) converge to forms similar to the dual polygon.

6.7.4 Heteroclinic Dynamics Near Boundary

When initialized near the minimizer of the Hamiltonian, the parameters behave like a simple harmonic oscillator. Their orbits are elliptical, and, for the appropriate choice of base measure, circular. However, when initialized far from the minimizer, the parameter orbits are typically nonelliptical. The solutions shown in Figure 10 provide an example. The Hamiltonians plotted in the main text, show that, when the domain Ψ is polygonal, then the level sets of the Hamiltonian approach polygons, so the parameters approach polygonal orbits far from the minimizer. Since the parameter orbits, and centroid orbits, are diffeomorphic, and large exponential tilts move the centroid near to the boundary of Ψ , these limits correspond to the limiting orbits of the adaptive dynamics equation, in the limit where the centroid approaches the boundary of the convex hull of Ψ . We will show that, when this boundary is polygonal, then the corresponding parameter dynamics also produce polygonal orbits, and, the corners of the convex hull of the embedding space, Ψ , are key geometric features that generate the limiting polygonal orbit. Moreover, we will show that, as the centroid approaches the boundary, then the centroid dynamics slow down near corners of Ψ , so the centroid dynamics behave like a heteroclinic oscillator near the boundaries of Ψ .

To study these orbits, we focus on a single disc game, assume that Ψ is radially bounded so that its convex hull is compact, the convex hull contains the origin and is a polygon. To highlight the geometric relationship between the level sets of the Hamiltonian for large Hamiltonian values, and the polygonal convex hull, we will also assume that the base measure is uniform over Ψ .

The orbits of interest correspond to solutions of the parameter dynamics in a limit where the value of the Hamiltonian on the orbit diverges, or, equivalently, where $\inf_t(\|\theta(t)\|_2)$ diverges. These two conditions are equivalent when ψ is bounded, and its convex hull contains the origin since in this situation the Hamiltonian is defined for all θ , is radially unbounded, convex, and has a unique global minimizer for interior θ .

Recall that the parameters obey the Hamiltonian dynamic:

$$\frac{d}{dt}\theta(t) = U\nabla_\theta P(\theta) \propto U\nabla_\theta \log(P(\theta)) = U\bar{y}(\theta) \quad (168)$$

where $\bar{y}(\theta) = \mathbb{E}_{Y \sim p(\theta)}[Y]$ where $p(\cdot, \theta)$ is the normalized measure $\pi(\cdot, \theta)/P(\theta)$. The constant of proportionality relating both sides is the population size P . The two sides produce identical trajectories, so we will adopt the convention that all expectations are computed against the normalized population. This corresponds to an agent-based mixing model with constant per capita interaction rates. This convention simplifies the limiting analysis since P is convex, so $\|\frac{d}{dt}\theta(t)\| = \|\nabla_\theta P(\theta(t))\|_2$ may diverge as $\|\theta(t)\|_2$ diverges. In contrast, if we use constant per capita interaction rates, then, if Ψ is bounded, $\|\bar{y}(\theta)\|_2$ is bounded, so the parameter dynamics may produce orbits with finite rates as $\|\theta(0)\|_2$ diverges.

Then, following the main text argument, each corner c_j generates an angular region, bounded by the rays parallel to

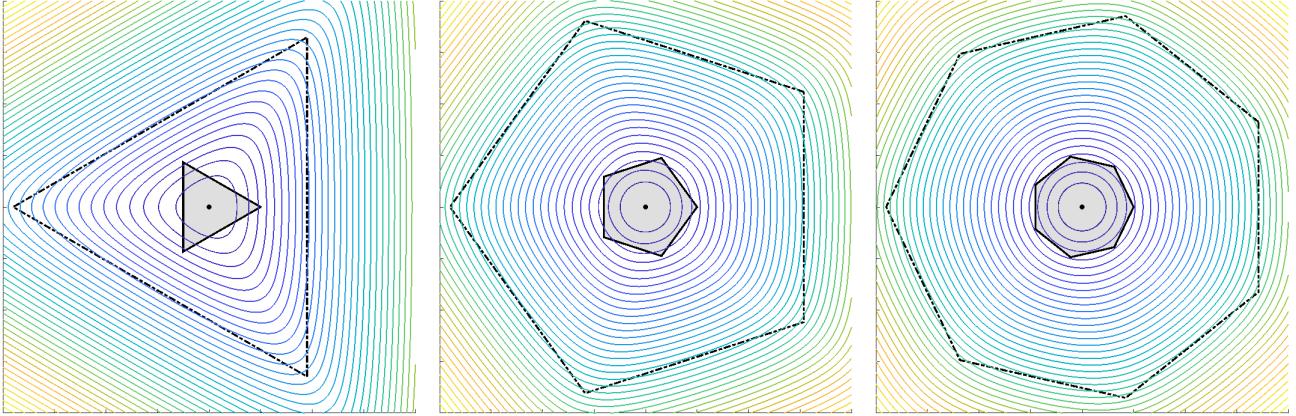


Figure 14: Parameter dynamics for a continuous replicator equation defined in a single disc game, with a base measure equal to a uniform distribution on regular polygons (the shaded grey triangle, pentagon, and septagon). For an irregular example, see Figure 13. The orbits are colored by Hamiltonian value. Orbit at large values of the Hamiltonian (large amplitude) approach a polygonal form. The outer polygon (dot-dashed outer polygon) is the dual to the domain Ψ represented in grey. Notice that, as the number of vertices increases, the polygon approaches a circle, so the base measure develops an approximate rotational symmetry. As a result, the parameter trajectories converge to concentric circles as the degree of a regular polygon increases.

\hat{n}_{j-1} and \hat{n}_j , where \hat{n}_j is the normal to the j^{th} boundary of $\text{conv}(\psi)$, such that, if $\hat{\theta}_0$ lies inside the region, then $p(y, a\hat{\theta}_0)$ approaches a delta distribution at the corner c_j . In this case, $\lim_{a \rightarrow \infty} \nabla_\theta \log(P(\theta))|_{\theta=a\hat{\theta}_0} = \lim_{a \rightarrow \infty} \bar{y}(a\hat{\theta}_0) = c_j$. Since this limit holds for all rays pointing into the j^{th} region, the level sets of the Hamiltonian, and the parameter trajectories, approach straight line segments parallel to Uc_j , bounded by their intersections with the rays $\{s\hat{n}_{j-1}\}_{s \geq 0}, \{s\hat{n}_j\}_{s \geq 0}$. Since the solutions are periodic, the collection of parameter trajectories form a series of concentric orbits. Then, since the angular regions generated by the corners partition all angles about the origin, in the limit of large initial θ , the parameter orbits must approach a series of similar polygons.

The associated polygon is dual to the polygon $\text{conv}(\Psi)$. Figure 13 shows two examples. The dual has one edge for every vertex in $\text{conv}(\Psi)$, and, as a consequence, one corner for every edge of Ψ . These corners correspond to the finite set of initial directions $\hat{\theta}_0$ for which $p(\cdot, a\hat{\theta}_0)$ does not approach a delta distribution at a corner of Ψ .

The limiting dual polygon is similar to the dual of $\text{conv}(\Psi)$ constructed by:

1. Partition \mathbb{R}^2 into angular segments bounded by rays parallel to the normal vectors of each side of $\text{conv}(\Psi)$.
2. Pick an initial point away from the origin, within the first angular segment.
3. Extend that point into a line segment, perpendicular to the vector pointing to the corner of $\text{conv}(\Psi)$ linking the two boundaries defining the angular region, extended until it intersects with the boundaries of the angular region.
4. Then, from each point where the line segment enters a new region, extend it by adding a new line segment, perpendicular to the corner of $\text{conv}(\Psi)$ assigned to the associated angular region, and extended until the line segment intersects both boundaries of the angular region.
5. Iterate until the process produces a closed orbit.

Let $P = \text{conv}(\Psi)$ denote the original polygon. Then, this procedure produces a dual polygon, P' , where:

1. The vector from the origin to the j^{th} vertex of P' is perpendicular to the j^{th} edge of P .
2. The vector from the origin to the j^{th} vertex of P is perpendicular to the j^{th} edge of P' .

If Ψ is a regular polygon, then the dual polygon associated with limiting orbits is a regular polygon of the same degree, rotated so that its edges are bisected by the vertices of Ψ , and so that its vertices bisect the edges of Ψ . For example, the square boundary observed for a Laplace base measure is the dual to the diamond shaped level sets of the Laplace density (see Figure 13). Figure 14 shows example trajectories for a uniform measure on a regular triangle, pentagon, and septagon.

For example, consider a uniform base measure on the square $[-1, 1]^2$. Then, the Hamiltonian factors, $P(\theta) = (\sinh(\theta_1)/\theta_1)(\sinh(\theta_2)/\theta_2)$ so the parameter dynamics simplify as outlined in Section 6.7.3. Then:

$$\frac{d}{dt}\theta_1(t) \propto \coth(\theta_2(t)) - \frac{1}{\theta_2(t)}, \quad \frac{d}{dt}\theta_2(t) \propto -\coth(\theta_1(t)) + \frac{1}{\theta_1(t)}. \quad (169)$$

with constant of proportionality equal to the initial population size. As usual, converting to a constant per capita interaction rate sets the constant of proportionality to one. If θ_1 and θ_2 are both large, then $\coth(\theta_j) \simeq \text{sign}(\theta_j) + \mathcal{O}(e^{-2\theta_j})$ and $1/\theta_j$ is small. So, for large θ_1 and θ_2 , the dynamics are approximated by:

$$\frac{d}{dt}\theta_1(t) = \text{sign}(\theta_2(t)) + \mathcal{O}(|\theta_2(t)|^{-1}), \quad \frac{d}{dt}\theta_2(t) = -\text{sign}(\theta_1(t)) + \mathcal{O}(|\theta_1(t)|^{-1}). \quad (170)$$