ECOLE POLYTECHNIQUE

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

HIGHER ORDER LEARNING AND EVOLUTION IN GAMES

Rida LARAKI
Panayotis MERTIKOPOULOS

June 28, 2012

Cahier n° 2012-24

DEPARTEMENT D'ECONOMIE
Route de Saclay
91128 PALAISEAU CEDEX
(33) 1 69333033
http://www.economie.polytechnique.edu/
mailto:chantal.poujouly@polytechnique.edu

Higher Order Learning and Evolution in Games

Rida Laraki* Panayotis Mertikopoulos†,‡

June 28, 2012

Abstract

Continuous-time game dynamics are first order systems where payoffs are used to determine the growth rate of the players' strategy shares. In this paper, we investigate what happens beyond this first order framework by viewing payoffs as higher order forces of change, specifying e.g. the acceleration of the players' evolution instead of its velocity (a viewpoint which we show is quite natural when it comes to aggregating empirical data of past instances of play).

To that end, we derive a wide class of higher order game dynamics, generalizing all first order imitative dynamics, and, in particular, the replicator dynamics. In stark contrast to the first order setting, we show that weakly dominated strategies are eliminated in all n-th order payoff-monotonic dynamics for $n \geq 2$; moreover, strictly dominated strategies become extinct in n-th order dynamics n orders as fast as in the corresponding first order systems. Finally, we also establish a higher order analogue of the folk theorem, where, as a consequence of this higher order mode of learning, it is shown that the rate of convergence of n-th order dynamics to strict equilibria is n orders as fast as in their first order counterparts.

Keywords: Evolutionary game dynamics, higher order dynamical systems, weakly dominated strategies, folk theorem.

JEL Classification: C72, C73

*RL is with the French National Center for Scientific Research (CNRS) and the Economics Department of École Polytechnique, Paris, France. E-mail: rida.laraki@polytechnique.edu.

†PM is with the French National Center for Scientific Research (CNRS) and the Laboratoire d'Informatique de Grenoble, Grenoble, France. E-mail: panayotis.mertikopoulos@imag.fr.

‡Corresponding author.

1 Introduction

Owing to the considerable complexity of computing Nash equilibria and other rationalizable outcomes in non-cooperative games, a fundamental question that arises is whether these outcomes may be regarded as the result of a simple dynamic process where "the participants are supposed to accumulate empirical information on the relative advantages of the various pure strategies at their disposal" (Nash, 1950, p. 21). To that end, numerous classes of game dynamics have been proposed (from both a learning and an evolutionary "mass-action" perspective), each with its own distinct set of traits and characteristics – see e.g. the comprehensive survey by Sandholm (2011) for a most recent account.

Be that as it may, there are very few rationality properties that are shared by a decisive majority of classes of game dynamics, even if we focus for simplicity on the continuous-time, deterministic regime. For instance, a simple comparison between the well-known Smith dynamics (Smith, 1984) and the replicator dynamics (Taylor and Jonker, 1978) reveals that game dynamics can be innovative (Smith) or imitative (replicator); strictly dominated strategies might survive (Hofbauer and Sandholm, 2011) or, on the contrary, become extinct (Samuelson and Zhang, 1992); rest points might coincide with the Nash set of the game (Hofbauer and Sandholm, 2009) or properly contain it instead; etc. On the other hand, negative results seem to be much more ubiquitous: Hart and Mas-Colell (2003) showed that there is no class of uncoupled game dynamics that always converges to equilibrium, and even worse, weakly dominated strategies may survive in the long run, even in simple 2×2 games (Samuelson, 1993; Weibull, 1995).

As one might expect, the view is not particularly more uniform from a mathematical standpoint either: perhaps the single unifying feature of the vast majority of (deterministic, continuous-time) game dynamics is that they are first order dynamical systems that evolve over a product of simplices (the game's mixed strategy space). Beyond this basic attribute however, game dynamics hardly share any other resemblances: even continuity (an otherwise standard regularity assumption) is absent in several important classes of dynamics – e.g. as in the projection dynamics of Sandholm, Dokumacı, and Lahkar (2008).

Quite interestingly, in the closely related field of optimization (which can be likened to playing against nature), this restriction to first order dynamics is not present. To wit, as was recently shown by Alvarez, Attouch et al. (Alvarez, 2000; Alvarez, Attouch, Bolte, and Redont, 2002; Attouch, Goudou, and Redont, 2000), second order gradient ascent (the so-called "heavy ball with friction" method) has some remarkable optimization properties that first order schemes do not possess. Indeed, by interpreting the gradient of the function to be maximized as a physical, Newtonian force (and not as a first order vector field to be dutifully tracked by the system's trajectories), in many instances one can give the system enough energy to escape the basins of attraction of local maxima and converge instead to the *global* maximum of the objective function (something which is not possible in ordinary first order dynamics).

This, therefore, begs the question: *can second (or higher) order dynamics be introduced in a game theoretic setting?* And if yes, *can we by so doing obtain better convergence results and/or escape the impossibility results of first order dynamics?*

The first key challenge to overcome in this endeavor is that second order methods in optimization apply to *unconstrained* problems, whereas game dynamics must respect the (constrained) structure of the game's strategy space. To circumvent this impasse, Flåm and Morgan (2004) proposed a heavy-ball method as in Attouch et al. (2000) above, and they exogenously enforced consistence by projecting the orbits' velocity to a subspace of admissible directions when the updating would lead to inadmissible strategy profiles (say, assigning negative probability to an action). Unfortunately, as is often the case with projection-based schemes (see e.g. Sandholm et al., 2008), the resulting dynamics are not continuous, so even basic existence and uniqueness results are hard to obtain.

On the other hand, if players try to improve their performance by aggregating information on the relative payoff differences of their pure strategies, then this cumulative empirical data is *not* constrained (as mixed strategies are). Thus, a promising way to obtain a well-behaved second (or higher) order dynamical system for learning in games is to use the player's accumulated data to define an unconstrained *performance measure* for each strategy (this is where the dynamics of the process come in), and then map these "scores" to mixed strategies by means of e.g. an (inverse) logit choice model (Hofbauer, Sorin, and Viossat, 2009; Mertikopoulos and Moustakas, 2010; Rustichini, 1999; Sorin, 2009). In other words, the dynamics can first be specified on an unconstrained space, and then reflected on the game's actual strategy space via the players' choice model, which produces a mixed strategy based on each pure strategy's score.

Outline of Results

After a few preliminaries in Section 2, this approach is made precise in Section 3 where we present a higher order framework for the familiar class of *imitative dynamics* (Björnerstedt and Weibull, 1996), a class containing all payoff-monotonic dynamics, and in particular, the replicator dynamics. In fact, as a consequence of the passage from performance scores to mixed strategies, the resulting dynamics naturally inherit a game-independent "adjustment" term which slows down the orbits that approach the boundary of the game's strategy space and renders the latter invariant.

Regarding the rationality properties of the derived dynamics, we show in Section 4 that payoff-monotonic dynamics of any order eliminate strictly dominated strategies, including iteratively dominated ones: in the long run, only rationalizable strategies can survive. Qualitatively, this result is the same as its first order analogue; quantitatively however, the rate of extinction increases dramatically with the order of the dynamics considered: *dominated strategies become extinct in n-th order dynamics n orders as fast as in their first order counterparts* (Theorem 4.1). The reason for this enhanced rate of elimination is that empirical data accrues much faster if a higher order scheme is used rather than a lower order one:[1] players who look deeper into the past by using a higher order learning rule identify consistent payoff differences much faster, so they annihilate dominated strategies much faster as well.

A remarkable consequence of the above is that in all higher order ($n \geq 2$) payoff-monotonic dynamics, *even weakly dominated strategies become extinct in the long run* (Theorem 4.4). Needless to say, this comes in stark contrast to the first order setting where weakly dominated strategies survive even in simple 2×2 games such as Entry Deterrence (Weibull, 1995, Ex. 5.4). As such, the implementation of a higher order learning rule carries significant ramifications for the justification of rational behavior: the elimination of weakly dominated strategies *can* be interpreted as the outcome of a learning process, simply by considering more sophisticated players who look deeper into the past.

Extending our analysis to equilibrium play, we show in Section 5 that the folk theorem of evolutionary game theory (Hofbauer and Sigmund, 1988; Weibull, 1995) continues to hold in our setting as well (modulo certain technical modifications needed to accommodate higher order dynamics). More specifically, in all higher order payoff-monotonic dynamics, we show that: *a*) if an interior solution orbit converges, then its limit is Nash; *b*) if a point is Lyapunov stable, then it is also Nash; and *c*) if players start close enough to a strict equilibrium and with a small enough higher order learning bias, then they converge to it (Theorem 5.1). In fact, echoing our results on the rate of extinction of dominated strategies, we show that n-th order payoff-monotonic dynamics converge to strict equilibria n orders as fast as their first-order counterparts.

As a converse to (c) in first order dynamics, it is well-known that the flow of the multi-population replicator dynamics is "incompressible" (volume-preserving), so its orbits cannot coalesce to an interior point (Hofbauer, 1996; Hofbauer and Sigmund, 1988; Ritzberger and Vogelsberger, 1990); as a result, a point is asymptotically stable if and only if it is a strict equilibrium (Ritzberger and Weibull, 1995). That said, more general dynamics are not incompressible, so this

[1]For instance, $\dot{x}(t) = 1$ grows linearly in t, while the growth of $\ddot{x}(t) = 1$ is quadratic.

important equivalence ceases to hold: for instance, even the payoff-adjusted replicator dynamics of Maynard Smith exhibit interior attractors in simple 2×2 games (see e.g. Ex. 5.3 in Weibull, 1995).

On the other hand, given that payoffs do not depend on the strategies' growth rates (or other derivates), incompressibility turns out to be an intrinsic property of *all* imitative higher order dynamics for $n \geq 2$, so *non-pure points cannot be attracting in any higher order imitative dynamics* (Theorem 5.4). Higher order learning is thus seen to reinforce the link between the "epistemic" (as Ritzberger and Weibull (1995) and van Damme (1987) call it) and the dynamic instability of mixed equilibria: *if players ascribe to a higher order learning scheme, then even lossless off-equilibrium deviations become inherently unstable.*

2 Notation and Preliminaries

2.1 Notational conventions

Given a finite set $\mathcal{S} = \{s_\alpha\}_{\alpha=0}^n$, the *vector space spanned by* $\mathcal{S}$ will be the set of formal linear combinations of elements of $\mathcal{S}$ with real coefficients, i.e. the set $\mathbb{R}^{\mathcal{S}} \equiv \mathrm{Hom}(\mathcal{S}, \mathbb{R})$ of all maps $x\colon \mathcal{S} \to \mathbb{R}$. Clearly, the space $\mathbb{R}^{\mathcal{S}}$ admits a *canonical* basis $\{e_\alpha\}_{\alpha=0}^n$ consisting of the indicator functions $e_\alpha\colon \mathcal{S} \to \mathbb{R}$ which take the value $e_\alpha(s_\alpha) = 1$ on s_α and vanish otherwise. Hence, under the natural identification $s_\alpha \mapsto e_\alpha$, we will make no distinction between the elements s_α of $\mathcal{S}$ and the corresponding basis vectors e_α of $\mathbb{R}^{\mathcal{S}}$; moreover, we will frequently use the index α to refer to either s_α or e_α, and we will identify the set $\Delta(\mathcal{S})$ of probability measures on $\mathcal{S}$ with the standard n-dimensional simplex of $\mathbb{R}^{\mathcal{S}}$: $\Delta(\mathcal{S}) \equiv \{x \in \mathbb{R}^{\mathcal{S}} : \sum_\alpha x_\alpha = 1 \text{ and } x_\alpha \geq 0\}$.

In a similar vein, if $\{\mathcal{S}_k\}_{k\in\mathcal{K}}$ is a family of finite sets $\mathcal{S}_k$ indexed by $k \in \mathcal{K}$, we will also use the shorthand $\sum_\alpha^k$ for the sum $\sum_{\alpha\in\mathcal{S}_k}$. Finally, regarding players and their actions, we will follow the original convention of Nash and employ Latin indices $(j, k, \dots)$ for players, while keeping Greek ones $(\alpha, \beta, \dots)$ for their actions (pure strategies); also, unless otherwise mentioned, we will use $\alpha, \beta, \dots$, for indices that start at 0, and $\mu, \nu, \dots$, for those which start at 1.

2.2 Finite games

For our purposes, a *finite game in normal form* will consist of a finite set of *players* $\mathcal{N} = \{1, \dots, N\}$, each with a finite set of *actions* (or *pure strategies*) $\mathcal{A}_k = \{\alpha_{k,0}, \alpha_{k,1}, \dots\}$ that can be mixed by means of a probability distribution (or *mixed strategy*) $x_k = (x_{k,0}, x_{k,1}, \dots) \in \Delta(\mathcal{A}_k)$. The set $\Delta(\mathcal{A}_k)$ of a player's mixed strategies will be denoted by X_k, and aggregating over all players, the space of *strategy* *profiles* $x=(x_1,\dots,x_N)\in\prod_k \mathbb{R}^{\mathcal{A}_k}$ will be the product $X\equiv\prod_k X_k$; in this way, if $\mathcal{A}=\coprod_k \mathcal{A}_k$ denotes the (disjoint) union of the players' action sets, X can be seen as a product of simplices embedded in $\mathbb{R}^{\mathcal{A}}\cong\prod_k \mathbb{R}^{\mathcal{A}_k}$.

As is customary, when we wish to focus on the strategy of a particular (focal) player $k\in\mathcal{N}$ versus that of his *opponents* $\mathcal{N}_{-k}\equiv\mathcal{N}\setminus\{k\}$, we will employ the shorthand $(x_k;x_{-k})\equiv(x_1,\dots,x_k,\dots,x_N)\in X$ to denote the strategy profile where player k plays $x_k\in X_k$ against his opponents' strategy $x_{-k}\in X_{-k}\equiv\prod_{\ell\neq k}X_\ell$. The players' (expected) rewards are then prescribed by the game's *payoff* (or *utility*) *functions* $u_k\colon X\to\mathbb{R}$:

$$u_k(x)=\sum\nolimits_{\alpha_1}^{1}\cdots\sum\nolimits_{\alpha_N}^{N}u_k(\alpha_1,\dots,\alpha_N)\,x_{1,\alpha_1}\cdots x_{N,\alpha_N},\tag{2.1}$$

where $u_k(\alpha_1,\dots,\alpha_N)$ denotes the reward of player k in the profile $(\alpha_1,\dots,\alpha_N)\in\prod_k\mathcal{A}_k$; specifically, if player k plays $\alpha\in\mathcal{A}_k$, we will use the notation:

$$u_{k\alpha}(x)\equiv u_k(\alpha;x_{-k})=u_k(x_1,\dots,\alpha,\dots,x_N).\tag{2.2}$$

In light of the above, a *game in normal form* with *players* $k\in\mathcal{N}$, *action sets* $\mathcal{A}_k$, and *payoff functions* $u_k\colon X\to\mathbb{R}$ will be denoted by $\mathfrak{G}\equiv\mathfrak{G}(\mathcal{N},\mathcal{A},u)$. A *subgame* of $\mathfrak{G}$ will then be a game $\mathfrak{G}'\equiv\mathfrak{G}'(\mathcal{N},\mathcal{A}',u')$ played by the players of $\mathfrak{G}$, each with a subset $\mathcal{A}'_k\subseteq\mathcal{A}_k$ of their original actions, and with payoff functions $u'_k\equiv u_k|_{X'}$ suitably restricted to the reduced strategy space $X'=\prod_k\Delta(\mathcal{A}'_k)$.

Now, given a game $\mathfrak{G}\equiv\mathfrak{G}(\mathcal{N},\mathcal{A},u)$, we will say that the pure strategy $\alpha\in\mathcal{A}_k$ is (*strictly*) *dominated* by $\beta\in\mathcal{A}_k$ (and we will write $\alpha\prec\beta$) when

$$u_{k\alpha}(x)<u_{k\beta}(x)\quad\text{for all strategy profiles } x\in X.\tag{2.3}$$

More generally, we will say that $q_k\in X_k$ *is dominated by* $q'_k\in X_k$ if

$$u_k(q_k;x_{-k})<u_k(q'_k;x_{-k})\ \text{ for all strategies } x_{-k}\in X_{-k} \text{ of } k\text{'s opponents.}\tag{2.4}$$

Finally, if the above inequalities are only strict for some (but *not all*) $x\in X$, then we will employ the term *weakly dominated* and write $q_k\preccurlyeq q'_k$ instead.

By removing dominated (and, thus, rationally unjustifiable) strategies from a game, other strategies might become dominated in the resulting subgame, leading to the notion of *iteratively dominated strategies*. Specifically, given two subsets M_k and M_{-k} of X_k and X_{-k} respectively, let $\operatorname{Just}(M_k,M_{-k})\equiv\{q_k\in M_k:\forall q'_k\in M_k,\exists q_{-k}\in M_{-k}\text{ s.t. }u_k(q_k;q_{-k})\geq u_k(q'_k;q_{-k})\}$ be the set of strategies $q_k\in M_k$ that are justifiable (i.e. not dominated) with respect to any strategy $q_{-k}\in M_{-k}$. Then, starting with $X_k^0\equiv X_k$, define inductively the set of strategies that survive r elimination rounds as $X_k^r = \mathrm{Just}(X_k^{r-1}, X_{-k}^{r-1})$ where $X_{-k}^{r-1} \equiv \prod_{\ell \neq k} X_\ell^{r-1}$; similarly, the *pure* strategies that survive after r rounds will be denoted by $\mathcal{A}_k^r \equiv \mathcal{A}_k \cap X_k^r$. In this way, the sequence $\{X_k^r\}_{r=0}^\infty$ forms a descending chain $X_k^0 \supseteq X_k^1 \supseteq \ldots$ whose limit $X_k^\infty \equiv \bigcap_{r=0}^\infty X_k^r$ consists of those strategies of player k that are *rationalizable*, i.e. they survive *all* rounds of elimination. In particular, if the space $X^\infty \equiv \prod_k X_k^\infty$ of rationalizable strategies is a singleton, $\mathfrak{G}$ will be called *dominance-solvable* and the sole surviving strategy in X^∞ will be the game's *rational solution*.

Assume now that play evolves over time, say along the path $x(t) \in X$, $t \geq 0$. In that case, we will say that a pure strategy $\alpha \in \mathcal{A}_k$ *becomes extinct along* $x(t)$ if $x_{k\alpha}(t) \to 0$ as $t \to \infty$; more generally, for mixed strategies $q_k \in X_k$, we will follow Samuelson and Zhang (1992) and say that q_k becomes extinct along $x(t)$ if $\min\{x_{k\alpha}(t) : \alpha \in \mathrm{supp}(q_k)\} \to 0$, with the minimum taken over the support $\mathrm{supp}(q_k) \equiv \{\beta \in \mathcal{A}_k : q_{k\beta} > 0\}$ of q_k.

The above is equivalent to asking that the quantity $V_k(x) = \prod_{\alpha \in \mathrm{supp}(q_k)} x_{k\alpha}^{q_{k\alpha}}$ vanish as $t \to \infty$. Modulo a constant and a change of sign, the logarithm of this quantity is known as the *Kullback-Leibler divergence* $D_{\mathrm{KL}}(q_k \,\|\, x_k)$ of x_k with respect to q_k; more precisely, we have:

$$D_{\mathrm{KL}}(q_k \,\|\, x_k) = \sum_{\alpha \in \mathrm{supp}(q_k)} q_{k\alpha} \log\left(q_{k\alpha}/x_{k\alpha}\right). \tag{2.5}$$

Clearly, $D_{\mathrm{KL}}(q_k \,\|\, x_k)$ blows up to $+\infty$ whenever $\min\{x_{k\alpha} : \alpha \in \mathrm{supp}(q_k)\} \to 0$, so $q_k \in X_k$ becomes extinct along $x(t)$ iff $D_{\mathrm{KL}}(q_k \,\|\, x_k(t)) \to \infty$ as $t \to \infty$.

On the other hand, when a game cannot be solved by removing dominated strategies, one typically turns to the celebrated solution concept of a Nash equilibrium which characterizes profiles that are resilient against unilateral deviations. We will thus say that $q \in X$ is a(t) *Nash equilibrium* when

$$u_k(x_k; q_{-k}) \leq u_k(q) \quad \text{for all } x_k \in X_k \text{ and for all } k \in \mathcal{N}, \tag{2.6}$$

and if (2.6) is strict for all $x_k \in X_k \setminus \{q_k\}$, $k \in \mathcal{N}$, q will be called itself *strict*. Finally, equilibria of subgames $\mathfrak{G}'$ of $\mathfrak{G}$, will be called *restricted equilibria* of $\mathfrak{G}$.

2.3 Dynamical systems

In this section, our aim will be to give some background on dynamical systems and, in particular, to highlight the distinction between a system's *configuration space* and its *phase space*, two notions which can be used interchangeably in a first order environment, but which are very different in higher orders.

Recall first that the set of tangent vectors to the strategy space X of a normal form game at some point $x \in X$ admits the structure of a pointed convex cone, known as the *(solid) tangent cone* to X at x. As can be easily seen, this cone is naturally isomorphic to the convex cone of all rays that emanate from $x \in X$ and which intersect X in at least one other point; more precisely:

$$T_x X = \left\{ z \in \mathbb{R}^{\mathcal{A}} : \sum_\alpha^k z_{k\alpha} = 0 \text{ for all } k \in \mathcal{N} \text{ and } z_{k\alpha} \geq 0 \text{ if } x_{k\alpha} = 0 \right\}. \quad (2.7)$$

Following Lee (2003), a well-posed (first order) dynamical system on X may be regarded as a *flow* on X, i.e. as a smooth map $\Theta : X \times \mathbb{R}_+ \to X$ which satisfies the conditions *i.)* $\Theta(x, 0) = x$ for all $x \in X$, and *ii.)* $\Theta(\Theta(x,t),s) = \Theta(x, t+s)$ for all $x \in X$ and for all $s, t \geq 0$. The curve $\Theta^x : \mathbb{R}_+ \to X$, $t \mapsto \Theta(x,t)$ will be called the *orbit* (or *trajectory*) of x under Θ, and when there is no danger of confusion, $\Theta^x(t)$ will be denoted more simply as $x(t)$. Moreover, if $\dot{x}(0)$ denotes the initial velocity of the trajectory $x(t)$, we see that Θ induces a vector field V on X via the mapping $x \mapsto V(x) \equiv \dot{x}(0) \in T_x X$, so, by the fundamental theorem on flows, Θ will also be the unique flow such that $\dot{x}(t) = V(x(t))$ for all $t \geq 0$.

Given these equivalent descriptions of a dynamical system, we have the following definitions for stationarity and stability of a point $q \in X$:

- q will be called *stationary* if $V(q) = 0$ (i.e. if $\Theta^q(t) = q$ for all t).
- q will be called *Lyapunov stable* if, for every neighborhood U of q, there exists a neighborhood V of q such that $x(t) \in U$ for all $x \in V$, $t \geq 0$.
- q will be called *attracting* if $x(t) \to q$ for all x in a neighborhood U of q.
- q will be called *asymptotically stable* if it is Lyapunov stable and attracting.

The differential equation $\dot{x}(t) = V(x(t))$ (or $\dot{x} = V$ for short) represents a *first order* dynamical system on X. Higher order dynamicss of the form "$x^{(n)} = V$" can then be defined by means of the equivalent recursive formulation:

$$\begin{aligned} \dot{x}(t) &= x^1(t) \\ \dot{x}^1(t) &= x^2(t) \\ &\cdots \\ \dot{x}^{n-1}(t) &= V(x(t), x^1(t), \ldots, x^{n-1}(t)). \end{aligned} \quad (2.8)$$

In this way, an n-th order dynamical system on X may be seen as a flow on the *phase space* $\Omega \equiv \Omega(X) = \coprod_x (T_x X)^{n-1}$ whose points (n-tuples of the form

$(x, x^1, \ldots, x^{n-1})$ as above) represent all possible *states* of the system.[2] By contrast, base points $x \in X$ will keep the designation "points", and X itself will be called the *configuration space* of the system (a distinction which, as has already been noted, is redundant in first order systems).

Of course, the evolution of an n-th order dynamical system is not uniquely determined by its *initial position* $x(0)$ in the system's configuration space X, but by the system's *initial state* $(x(0), \dot{x}(0), \ldots, x^{(n-1)}(0))$ in the phase space Ω. On that account, if we wish to study how initial positions evolve over time, we will consider the corresponding *rest states* $(x(0), 0, \ldots, 0)$ which signify that the system starts *at rest*; furthermore, we will also use this natural embedding to view X as a subset of Ω and write $x \in \Omega$ instead of $(x, 0, \ldots, 0) \in \Omega$.

3 Derivation of Higher Order Dynamics

3.1 Possible approaches and their limitations

Needless to say, the most fundamental requirement for any class of game dynamics is that its solution trajectories must actually stay within the game's strategy space X (viewed as a subset of $\mathbb{R}^{\mathcal{A}}$). Accordingly, if we consider a general first order dynamical system of the form:

$$\dot{x}_{k\alpha} = F_{k\alpha}(x), \tag{3.1}$$

with Lipschitz $F_{k\alpha} \colon \mathbb{R}^{\mathcal{A}} \to \mathbb{R}$, the consistency postulate above is equivalent to the tangency requirements *a*) $\sum_{\alpha}^{k} F_{k\alpha} = 0$ for all $k \in \mathcal{N}$, and *b*) $F_{k\alpha} \geq 0$ whenever $x_{k\alpha} = 0$ – cf. the definition (2.7) of $T_x X$ and the relevant discussions in Hofbauer and Sigmund (1988), Sandholm (2011) and Weibull (1995).

On the other hand, the situation becomes much more intricate if we attempt to introduce second (or higher) order dynamics on X. To wit, if we sensibly (but, ultimately, naïvely) replace $\dot{x}$ with $\ddot{x}$ in (3.1), then the resulting dynamics will not respect the simplicial structure of X, no matter the choice of F: if players start with a sufficiently high initial growth rate vector $\dot{x}(0)$ pointing towards the exterior of X, then the corresponding solution orbits will escape X in finite time, possibly never to return.[3]

[2] More concisely, a state can be seen as an *n-jet*, with the phase space Ω being the *n-th order jet bundle* of X (Saunders, 1989).

[3] From a physical viewpoint, this behavior is entirely natural: after all, finite forces cannot contain particles of arbitrarily high energy in a bounded region. To prove this rigorously, simply use Taylor's theorem to write $x_{k\alpha}(t) = x_{k\alpha}(0) + \dot{x}_{k\alpha}(0)t + F_{k\alpha}(x(\xi_{k\alpha}))t^2$ for some $\xi_{k\alpha} \in [0, t]$, and then choose $\dot{x}_{k\alpha}(0)$ sufficiently negative (recalling that F is bounded) so that $x_{k\alpha}(t)$ becomes negative for sufficiently small $t > 0$.

One way around this obstacle was proposed by Flåm and Morgan (2004) who forced solutions to remain in X by exogenously projecting the velocity $v(t) \equiv \dot{x}(t)$ of an orbit to the tangent cone $T_x X$ of "admissible" velocity vectors – similarly to Nagurney and Zhang (1997) and Sandholm et al. (2008). In this context, one begins with the "naïve" second order version of (3.1):

$$\dot{x}_{k\alpha} = v_{k\alpha} \tag{3.2a}$$
$$\dot{v}_{k\alpha} = F_{k\alpha}, \tag{3.2b}$$

and replaces (3.2a) with the "projected" variant:

$$\dot{x} = \mathrm{proj}_{T_x X}(v), \tag{3.2a$'$}$$

where $\mathrm{proj}_{T_x X}(v)$ is the projection of v to the tangent cone $T_x X$ of X at x.[4,5]

This approach has the benefit that $x(t) \in X$ for all t for which $x(t)$ is defined, but it also carries some considerable (and, to a large extent, unavoidable) weaknesses. First, the projection operator $\mathrm{proj}_{T_x X}$ does not vary continuously with x (a consequence of the changing structure of $T_x X$ along the faces of X), so existence and (especially) uniqueness of solutions to (3.2a$'$) might (and does) fail. Second, in order for players to update (3.2a$'$), they must be able to tell *precisely* when they hit a boundary face of X in order to change the projection operator that they are using (a typical downside to non-interior methods). However, given that numerical calculations always carry unavoidable approximation errors (such as the "machine ε" of the players' calculating device (Cantrell, 2000)), this updating scheme is prone to numerical instabilities which might lead solution trajectories to escape X in practical implementations.

An alternative approach to keep solution orbits in X would be to erect an "infinite wall" at the faces of X, as in the "potential well" problems of quantum mechanics (Sakurai, 1994). These "walls" take the form of smooth barriers of infinite height that are placed at the faces of X, and which are such that solution trajectories cannot jump over them as they approach $\mathrm{bd}(X)$. Specifically, this amounts to modifying the forces (3.2b) by setting:

$$\dot{v}_{k\alpha} = F_{k\alpha} + W_{k\alpha}, \tag{3.2b$'$}$$

where the "boundary terms" $W_{k\alpha} \equiv W_{k\alpha}(x, v)$ satisfy:

[4] Flåm and Morgan actually consider only the "gradient field" $F_{k\alpha} = u_{k\alpha}$, but the same projection machinery can be applied to more general Lipschitz vector fields as well.

[5] Strictly speaking, by using (3.2a$'$) instead of (3.2a), the dynamics (3.2a$'$) are not a bona fide second order system in the sense of (2.8) except in the interior $\mathrm{int}(X)$ of X; this technicality will not concern us here, but it is nonetheless important to keep in mind.

1. $\sum_{\alpha}^{k} W_{k\alpha} = 0$, in order to be consistent with the constraint $\sum_{\alpha}^{k} x_{k\alpha} = 1$.

2. $W_{k\alpha}(x, v) \to +\infty$ as $x_{k\alpha} \to 0$, in order to ensure that $\dot{v}_{k\alpha} \to +\infty$ as $x_{k\alpha} \to 0$ (implying in turn that $x_{k\alpha}(t)$ cannot vanish if $x_{k\alpha}(0) > 0$ initially).

Nonetheless, this approach still suffers from important drawbacks. To begin with, condition (2) above implies that $x(t)$ will *rebound* at $\mathrm{bd}(X)$, meaning that (3.2b′) can never converge to a pure strategy profile (or other boundary point of X). To remedy this, we would need to impose the further condition:

3. $W_{k\alpha}(x, v) \to 0$ as $v_{k\alpha} \to 0$, in order to allow for finite accelerations $\dot{v}_{k\alpha}$ when $x(t)$ approaches $\mathrm{bd}(X)$ with *vanishing* velocity.

However, this raises new questions regarding the limit behavior of $W_{k\alpha}(x, v)$ as *both* $x_{k\alpha} \to 0$ and $v_{k\alpha} \to 0$, leading to the following, more important issue: if we simply impose an arbitrary barrier term on (3.2) satisfying the above conditions, then the adjusted dynamics (3.2a)–(3.2b′) certainly do not emerge naturally from game-theoretic considerations. Therefore, unless one can exhibit an inherent link with rationality or learning, this approach remains an artificial device with the sole, self-serving purpose of trapping solutions within X.

3.2 The second order replicator dynamics in dyadic games

An alternative approach with its roots in reinforcement learning is to overcome the restrictions imposed by the simplicial structure of X by having players update an *unconstrained* measure of their actions' performance rather than directly updating their constrained strategies (as was the case in the above approaches). In this unconstrained space of performance measures, second (or higher) order effects come about naturally when players look two (or more) steps into the past: it is the dynamics of these "scores" that induce a well-behaved dynamical system on the game's strategy space.

As an introductory example, let us consider a game where every player $k \in \mathcal{N}$ has two possible actions, say "0" and "1", which are ranked over time using the associated payoff differences $u_{k,1} - u_{k,0}$. With this regret-like information at hand, and assuming perfect information, players measure the performance of their strategies by updating the auxiliary *score variables* (or *propensities*):

$$U_k(t+1) = U_k(t) + \Delta u_k(x(t)), \tag{3.3}$$

where $x_k(t) \equiv x_{k,1}(t)$ represents the mixed strategy of player k at time t (assumed discrete for the moment). The strategies x_k are then updated themselves following the well-known *inverse logit* (or *expit*) choice model whereby actions that score better are played exponentially more often (Hofbauer et al., 2009; Mertikopoulos and Moustakas, 2010; Rustichini, 1999; Sorin, 2009):[6]

$$x_k(t+1) = \operatorname{expit}(U_k(t+1)) \equiv \frac{\exp(U_k(t+1))}{1+\exp(U_k(t+1))}. \tag{3.4}$$

This process is repeated indefinitely, so if we descend to continuous time for simplicity,[7] the system of (3.3) and (3.4) yields the coupled equations:

$$\dot{U}_k = \Delta u_k(x) \tag{3.5a}$$

$$x_k = \operatorname{expit}(U_k) = \big(1+\exp(-U_k)\big)^{-1}. \tag{3.5b}$$

Hence, by differentiating (3.5b) in order to decouple it from (3.5a), we readily obtain the 2-strategy replicator dynamics of Taylor and Jonker (1978):

$$\dot{x}_k = \frac{dx_k}{dU_k}\dot{U}_k = x_k(1-x_k)\Delta u_k(x). \tag{3.6}$$

In this well-known derivation of the replicator dynamics from the exponential reinforcement rule (3.5), the constraints $x_k \in [0,1]$, $k \in \mathcal{N}$, are automatically satisfied thanks to (3.4). On the downside however, (3.5a) itself "forgets" a lot of past (and potentially useful) information because the corresponding "discrete-time" recursion (3.3) only looks one step in the past. To remedy this, players could take (3.3) one step further by aggregating the scores U_k themselves so as to build even more momentum towards the strategies that tend to perform better. We thus obtain the second order cumulative reinforcement scheme:

$$U_k(t+1) = U_k(t) + \Delta u_k(x(t)) \tag{3.7a}$$

$$Y_k(t+1) = Y_k(t) + U_k(t+1), \tag{3.7b}$$

where, as before, the profile $x(t)$ is updated following the logistic distribution (3.4).[8] Then, by eliminating the intermediate aggregation variables U_k from (3.7), we obtain the *second order* recursion:

$$Y_k(t+1) - Y_k(t) = Y_k(t) - Y_k(t-1) + \Delta u_k(x(t)), \tag{3.8}$$

[6] We are using the term "inverse" because the expit function in (3.4) is actually the *inverse* logit function: $U_k = \log(x_k) - \log(1-x_k) = \operatorname{logit}(x_k)$, so $x_k = \operatorname{logit}^{-1}(U_k) = \operatorname{expit}(U_k)$.

[7] We should stress here that this passage to continuous time is done at a heuristic level – see Rustichini (1999) and Sorin (2009) for an exploration of some of the issues that arise in the passage from the discrete to the continuous. The discrete version of the exponential updating rule (3.5) is a very important topic to address, but since we seek to focus on the properties of the underlying continuous-time dynamics, it lies beyond the scope of this paper.

[8] Of course, players could look even deeper into the past by taking further aggregates in (3.7), but we will not deal with this issue here in order to keep our example as simple as possible.

which, in turn, leads to the continuous-time variant:

$$\ddot{Y}_k = \Delta u_k. \tag{3.9}$$

As we have already seen in the first order case, the coupled system of equations (3.4) and (3.9) automatically respects the simplicial structure of X by virtue of the logistic updating rule (3.4), so the important hurdle of actually staying in X has (finally!) been overcome. Nonetheless, for the purposes of comparison with the previous suggested approaches, it will be quite instructive to also derive the dynamics governing the evolution of the strategy profile $x(t)$ itself.

To that end, note that (3.4) gives $Y_k = \operatorname{logit}(x_k) = \log(x_k) - \log(1 - x_k)$, so a simple differentiation yields:

$$\dot{Y}_k = \frac{\dot{x}_k}{x_k} + \frac{\dot{x}_k}{1 - x_k} = \frac{\dot{x}_k}{x_k(1 - x_k)}. \tag{3.10}$$

Then, differentiating yet again, we obtain:

$$\ddot{Y}_k = \frac{\ddot{x}_k x_k(1 - x_k) - \dot{x}_k^2(1 - x_k) + \dot{x}_k^2 x_k}{x_k^2(1 - x_k)^2}, \tag{3.11}$$

and some algebra leads to the *second order replicator dynamics for dyadic games*:

$$\ddot{x}_k = x_k(1 - x_k)\Delta u_k + \frac{1 - 2x_k}{x_k(1 - x_k)}\dot{x}_k^2 \tag{3.12}$$

This derivation of a second order dynamical system on X will be the archetype for the significantly more general class of higher order dynamics of the next section, so we will not pause here to discuss (3.12) in any length. That said, it is worthy to note the following:

Remark 3.1 (Boundary terms). The second order system (3.12) is precisely of the form (3.2b′), with the last term of (3.12) playing the role of the "infinite wall" which blows up as $x_k \to 0$ and vanishes as $v_k \to 0$ (thus keeping $x(t)$ from escaping X, but allowing it to converge to $\operatorname{bd}(X)$ with zero velocity). However, whereas these same conditions were imposed artificially on (3.2b′) and were devoid of any links to learning or evolution, they now emerge naturally, as the byproduct of a logit choice model where players look deeper into the past.

Remark 3.2 (Past information). The precise sense of "looking deeper into the past" in the double aggregation scheme (3.7)–(3.9) can be made clearer if we write out the first and second order scores U_k and Y_k as explicit functions of time; in the continuous case, this gives:

$$U_k(t) = \int_0^t \Delta u_k(x(s))\, ds, \tag{3.13a}$$

$$Y_k(t) = \int_0^t U_k(s)\, ds = \int_0^t (t-s)\Delta u_k(x(s))\, ds. \tag{3.13b}$$

We thus see that the first order aggregate scores U_k assign uniform weight to all past instances of play, while the second order aggregates Y_k put (linearly) more weight on instances that are further removed into the past. This mode of weighing can be interpreted as players being reluctant to forget what has occurred, and this is precisely the reason that we describe the second order scheme (3.9) as "looking farther into the past".

It should be noted that in the theory of learning, past information is usually discounted (e.g. by an exponential factor) and ultimately discarded in favor of more recent observations (Fudenberg and Levine, 1998). As we shall see, "refreshing" observations in this way results in the players' propensity scores U_k growing at most linearly in time (see e.g. Hofbauer et al., 2009, Rustichini, 1999, and Sorin, 2009); by contrast, if players *reinforce* past observations by using (3.13b) in place of (3.13a), then their propensities may grow quadratically instead of linearly. As a result, first order learning is more conservative (leaning towards "exploring" more than "exploiting"), whereas higher order learning is less tempered and corresponds to more decisive players.

3.3 Reinforcement learning and higher order dynamics

In the general case, the reinforcement learning setup that we will be working with is as follows:

1. For every action $\alpha \in \mathcal{A}_k$, player k keeps and updates a *score* (or *propensity*) variable $y_{k\alpha} \in \mathbb{R}$ which measures the performance of α over time.

2. Players transform these scores into mixed strategies $x_k \in X_k$ by means of the Gibbs (inverse logit) choice model $G_k : \mathbb{R}^{\mathcal{A}_k} \to X_k$, $y_k \mapsto G_k(y_k)$:

$$x_{k\alpha} = G_{k\alpha}(y_k) \equiv \frac{\exp\left(\lambda_k y_{k\alpha}(t)\right)}{\sum_\beta^k \exp\left(\lambda_k y_{k\beta}(t)\right)}, \tag{GM}$$

where the "inverse temperature" $\lambda_k > 0$ controls the model's sensitivity to external stimuli (Mertikopoulos and Moustakas, 2010; Sorin, 2009).

3. The game is played and players record the payoffs $u_{k\alpha}(x)$ for each $\alpha \in \mathcal{A}_k$.

4. Players update their scores and the process is repeated ad infinitum.

Needless to say, the focal point of this learning process is the exact way in which players update the performance scores $y_{k\alpha} \in \mathbb{R}$ at each iteration of the game. We will thus take the natural extension of the aggregating framework of the previous section and consider a reinforcement scheme in which the players' scores $y_{k\alpha}$ are updated by looking n steps into the past as follows:

$$\begin{aligned} Y^{n-1}_{k\alpha}(t+1) &= Y^{n-1}_{k\alpha}(t) + u_{k\alpha}(x(t)) \\ &\cdots \\ Y^{0}_{k\alpha}(t+1) &= Y^{0}_{k\alpha}(t) + Y^{1}_{k\alpha}(t) \\ y_{k\alpha}(t+1) &= Y^{0}_{k\alpha}(t+1) \end{aligned} \tag{3.14}$$

The above scheme might appear somewhat cryptic at first, but it is quite straightforward to explain in words. In order to sharpen their performance measures as much as (consistently) possible, the players of the game accumulate data on each action's payoff; they then take an aggregate of this aggregate, and so forth up to n levels into the past; finally, they use this cumulative aggregate to update their mixed strategies via the Gibbs model (GM).

In particular, if we eliminate the intermediate aggregation variables Y^1, Y^2, etc., we obtain the straightforward n-th order recursion:

$$\Delta^n y_{k\alpha}(t) = u_{k\alpha}(x(t)), \tag{3.15}$$

where $\Delta^n y_{k\alpha}(t)$ denotes the n-th order finite difference of $y_{k\alpha}$, defined inductively as $\Delta^n y(t) = \Delta^{n-1} y(t+1) - \Delta^{n-1} y(t)$, with $\Delta^1 y(t) = y(t+1) - y(t)$. As such, if we again descend to continuous time for simplicity, we get the *n-th order exponential learning dynamics*:

$$y^{(n)}_{k\alpha}(t) = u_{k\alpha}(x(t)), \tag{LD$_\text{n}$}$$

where $x(t)$ is given by the Gibbs map (GM).

Clearly, the learning dynamics (LD$_\text{n}$) with Gibbs choice completely specify the evolution of the players' mixed strategy profile $x(t)$; however, they still fall short of our original goal to establish higher order dynamics in the players' strategy space X itself. To that end, by extending the calculations of Section 3.2 to our current general setting (see Appendix A), we obtain the *n-th order (asymmetric) replicator dynamics*:

$$x^{(n)}_{k\alpha} = \lambda_k \left(u_{k\alpha} - u_k\right) - x_{k\alpha}\left(R^{n-1}_{k\alpha} - \sum\nolimits^k_\beta x_{k\beta} R^{n-1}_{k\beta}\right), \tag{RD$_\text{n}$}$$

where the terms $R^{n-1}_{k\alpha}$ represent the effect of utilizing the higher order aggregation scheme (LD$_n$) and are given by the (game-independent!) expression:

$$R^{n-1}_{k\alpha}\left(x,\dot{x},\dots,x^{(n-1)}\right) = \sum \frac{(-1)^{m-1}n!}{m_1!\cdots m_{n-1}!}\frac{(m-1)!}{x^m_{k\alpha}}\prod_{r=1}^{n-1}\left(x^{(r)}_{k\alpha}(t)/r!\right)^{m_r}, \tag{3.16}$$

the sum being taken over all non-negative integers $m_1,\dots,m_{n-1}$ such that $m_1 + \cdots + (n-1)\cdot m_{n-1} = n$, and $m = m_1 + \cdots + m_{n-1}$.

As one would expect, for $n = 1$, we trivially obtain $R^0_{k\alpha} = 0$ for all $\alpha \in \mathcal{A}_k$, $k \in \mathcal{N}$, so (RD$_n$) is reduced to the standard replicator dynamics:

$$\dot{x}_{k\alpha} = \lambda_k x_{k\alpha}\left(u_{k\alpha}(x) - u_k(x)\right). \tag{RD$_1$}$$

As before, this derivation highlights the intimate link between the Gibbs distribution (GM) and the replicator equation: the latter is just a simple offshoot of the former combined with the learning dynamics (LD$_n$).[9]

On the other hand, for $n = 2$, the only lower order term that survives in (3.16) is for $m_1 = 2$; a bit of algebra then yields the *Newtonian replicator dynamics*:

$$\ddot{x}_{k\alpha} = x_{k\alpha}\left(u_{k\alpha}(x) - u_k(x)\right) + x_{k\alpha}\left(\dot{x}^2_{k\alpha}/x^2_{k\alpha} - \textstyle\sum^k_\beta \dot{x}^2_{k\beta}/x_{k\beta}\right). \tag{RD$_2$}$$

At first glance, the above equation seems different from the second order equation (3.12) that we derived in Section 3.2, but this is just a matter of reordering: if we restrict (RD$_2$) to two strategies, "0" and "1", and set $x_k \equiv x_{k,1} = 1 - x_{k,0}$, we will have $\dot{x}_k = \dot{x}_{k,1} = -\dot{x}_{k,0}$, and (3.12) follows immediately.

All told, the second order updating scheme $\ddot{y}_{k\alpha} = u_{k\alpha}$ that gives rise to (RD$_2$) highlights a very deep analogy between Newtonian mechanics and learning in games: in (RD$_2$), *the game's payoffs can be interpreted as the actual physical forces that determine the system's evolution, exactly as they would prescribe the orbits of a real physical system*.

Higher order imitative dynamics and payoff monotonicity. The dynamics (RD$_n$) will comprise the core of our higher order considerations, much as the replicator dynamics have become the gold standard for evolutionary and learning dynamics. More generally however, players might not base the updating (LD$_n$) of their performance scores on the payoffs $u_{k\alpha}(x)$ of the game, but on some different

[9] See also Rustichini (1999) and the more recent discussion in Hofbauer et al. (2009).

(assumed continuous) "payoff observables" $w_{k\alpha}\colon X \to \mathbb{R}$. In that case, we obtain the generalized reinforcement scheme:

$$y_{k\alpha}^{(n)} = w_{k\alpha}(x), \tag{GLD$_n$}$$

which, coupled with the Gibbs model (GM), yields the generalized dynamics:

$$x_{k\alpha}^{(n)} = \lambda_k x_{k\alpha}\left(w_{k\alpha}(x) - w_k(x)\right) - x_{k\alpha}\left(R_{k\alpha}^{n-1} - R_k^{n-1}\right), \tag{GD$_n$}$$

where w_k is the player average $w_k(x) \equiv \sum_\alpha^k x_{k\alpha} w_{k\alpha}(x)$ (and similarly for R_k).

The dynamics (GD$_n$) are characterized by the important property of "imitation", i.e. that players will never assign positive probability to a pure strategy that has become extinct in the course of play: if $x_{k\alpha}(0) = 0$ initially, then $x_{k\alpha}(t)$ remains zero for all time (cf. Remark 3.3 below). As such, the dynamics (GD$_n$) can be seen as the higher order extension of the class of imitative dynamics considered by Björnerstedt and Weibull (1996) (see also Weibull, 1995, and Sandholm, 2011): to obtain the higher order analogue of any first order dynamics of the form $\dot{x}_{k\alpha} = x_{k\alpha}(w_{k\alpha} - w_k)$, one simply needs to replace $\dot{x}_{k\alpha}$ with $x_{k\alpha}^{(n)}$ and add the (game-independent) boundary term $x_{k\alpha}(R_{k\alpha}^{n-1} - R_k^{n-1})$.

That said, if the observables $w_{k\alpha}$ are completely uncorrelated to the game's payoff functions $u_{k\alpha}$, there is little hope that the dynamics (GD$_n$) will lead to any sort of meaningful, rational play over time. It is thus natural to focus on observables w which respect the payoff ranking of a player's strategies, i.e.:

$$w_{k\alpha}(x) > w_{k\beta}(x) \text{ if and only if } u_{k\alpha}(x) > u_{k\beta}(x), \tag{PM}$$

for all $\alpha, \beta \in \mathcal{A}_k$, and for all $x \in X$. This correlation condition is usually referred to in the literature as *monotonicity* (or *payoff monotonicity*), so when the payoff observables $w_{k\alpha}$ satisfy (PM), the n-th order dynamics (GD$_n$) will be dubbed *monotonic* as well (see e.g. Hofbauer and Weibull, 1996, Samuelson and Zhang, 1992, and Weibull, 1995).

More generally, the monotonicity requirement (PM) can be broadened by replacing one (or both) of the pure strategies $\alpha, \beta \in \mathcal{A}_k$ by mixed ones (and possibly weakening the "if and only if" requirement to an "if"); in particular, (PM) can be viewed as a special case of the more stringent condition known as *aggregate monotonicity* (Samuelson and Zhang, 1992):

$$w_k(q_k'; x_{-k}) > w_k(q_k; x_{-k}) \text{ iff } u_k(q_k'; x_{-k}) > u_k(q_k; x_{-k}), \tag{AM}$$

with $x_{-k} \in X_{-k}$ and $q_k, q_k' \in X_k$. We will thus say that the dynamics (GD$_n$) are:

- *aggregate-monotonic* if (AM) holds for all $q_k, q'_k \in X_k$.
- *convex-monotonic* when the "if" direction of (AM) holds for all pure q'_k.
- *concave-monotonic* when the "if" direction of (AM) holds for all pure q_k.
- *monotonic* if (AM) holds for q_k and q'_k that are *both* pure.

(For a survey of these classes of dynamics in first order, see Hofbauer and Weibull, 1996, Weibull, 1995, or, for a most recent account, Viossat, 2011.)

In light of the above, we close this section with two book-keeping remarks on the Gibbs choice model (GM) which apply to all dynamics of the form ($\mathrm{GD_n}$):

Remark 3.3 (The faces of X). The Gibbs model (GM) maps $\mathbb{R}^{\mathcal{A}}$ to the (relative) interior $\mathrm{int}(X)$ of X, so any initial score assignment corresponds to a strategy profile that assigns positive mass to *all* actions $\alpha \in \mathcal{A}$ for all time. For the most part, we will not need to consider non-interior orbits; nonetheless, if required, we can consider initial conditions on any face X' of X simply by restricting the Gibbs map (GM) to appropriate subsets of the players' action sets (i.e. by effectively setting a non-utilized action's score to $-\infty$ and working in the associated subgame). In this manner, we see that the interior of any face of X is invariant in ($\mathrm{GD_n}$), just as in the first order case.[10]

Remark 3.4. We should also note that (GM) is not a 1-1 map between $\mathbb{R}^{\mathcal{A}}$ and $\mathrm{int}(X)$, so we may not freely pass from strategies to scores: for any $c \in \mathbb{R}$, we have $G_k(y_{k,0} + c, y_{k,1} + c, \dots) = G_k(y_{k,0}, y_{k,1}, \dots)$, so strategies can be mapped to scores only up to an additive constant (reminiscent of how the addition of a constant to a player's payoffs does not change the game). To recover a bijection, one may flag each player's "0"-th strategy and introduce the *relative scores*:

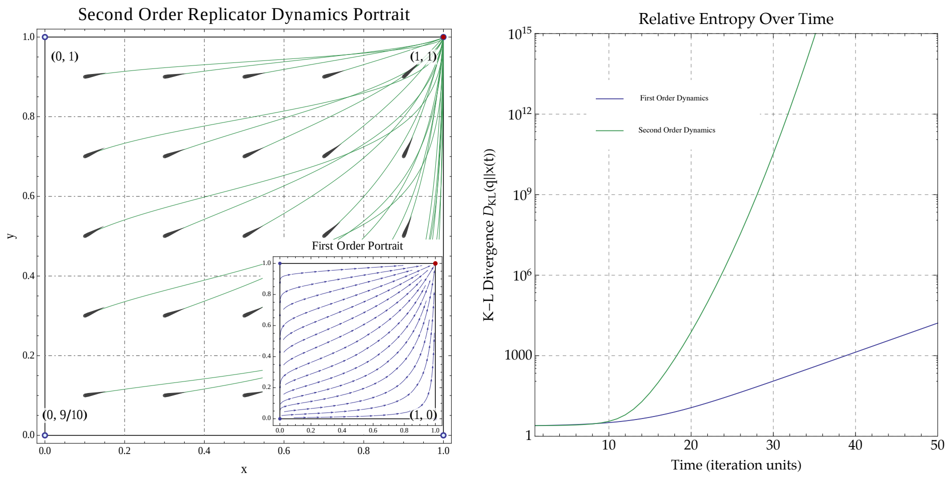
$$z_{k\mu} = y_{k\mu} - y_{k,0}, \quad \mu \in \mathcal{A}_k^* \equiv \mathcal{A}_k \setminus \{0\}, \tag{3.17}$$

which are then mapped to strategies via the *reduced* map $G_k^* \colon \mathbb{R}^{\mathcal{A}_k^*} \to X_k$:

$$G_{k,0}^*(z) = \left(1 + \textstyle\sum_\nu^k e^{z_{k\nu}}\right)^{-1}, \qquad G_{k\mu}^*(z) = e^{z_{k\mu}} \left(1 + \textstyle\sum_\nu^k e^{z_{k\nu}}\right)^{-1}. \tag{GM*}$$

In view of (GM*), the relative scores $z_{k\mu}$ may be recovered by the inverse expressions $z_{k\mu} = \log(x_{k\mu}/x_{k,0})$, $\mu \in \mathcal{A}_k^*$, $k \in \mathcal{N}$ (cf. (3.5b) in Section 3.2); furthermore, we can also identify strategy "0" of player k with the "point at negative infinity" $(-\infty, \dots, -\infty)$ in $\mathbb{R}^{\mathcal{A}_k^*}$, an observation which will be particularly useful in the asymptotic stability discussion of Section 5.

[10] This is also the reason that we need not concern ourselves with the technicalities of the fact that the dynamics ($\mathrm{GD_n}$) blow up near the boundary $\mathrm{bd}(X)$ of X.



(a) Portrait of a dominance solvable game. (b) Rate of extinction of dominated strategies.

Figure 1: Extinction of dominated strategies in the first and second order replicator dynamics. In Fig. 1(a) we plot the second order solution orbits of a dominance solvable game (see labels for the payoffs). In Fig. 1(b), we show the rate of extinction of the dominated strategy "0" by plotting the K-L divergence of a typical trajectory: the K-L distance grows exp-quadratically in second order dynamics compared to exp-linearly in first order (Theorem 4.1).

4 Elimination of dominated strategies

A fundamental rationality requirement for any class of game dynamics is that dominated strategies become extinct over time. Along these lines, our first result is that in the n-th order replicator dynamics, dominated strategies die out at a rate which is exponential in t^n:

Theorem 4.1. *Let $x(t)$ be an interior solution path of the n-th order replicator dynamics (RD_n). If $q_k \in X_k$ is iteratively dominated, we will have:*

$$D_{\mathrm{KL}}(q_k \,||\, x_k(t)) \geq \lambda_k c t^n / n! + \mathcal{O}(t^{n-1}), \tag{4.1}$$

for some constant $c > 0$; in other words, only rationalizable strategies survive. In particular, for pure strategies $\alpha \prec \beta$, we have:

$$x_{k\alpha}(t)/x_{k\beta}(t) \leq \exp\left(-\lambda_k \Delta u_{\beta\alpha} t^n / n! + \mathcal{O}(t^{n-1})\right), \tag{4.2}$$

where $\Delta u_{\beta\alpha} = \min_{x\in X}\{u_{k\beta}(x) - u_{k\alpha}(x)\} > 0$.

As an immediate corollary, we then obtain:

Corollary 4.2. *In dominance-solvable games, the n-th order replicator dynamics (RD_n) converge to the game's rational solution.*

The proof of Theorem 4.1 can be found in Appendix B; for now, we will focus on some relevant remarks:

Remark 4.1 (The asymptotic extinction rate). Even though (4.1) and (4.2) have been stated as inequalities, one can use any upper bound for the game's payoffs to show that the rate of extinction of dominated strategies in terms of the K-L divergence really *is* $\mathcal{O}(t^n)$.[11] As a result, we see that the asymptotic rate of extinction of dominated strategies in the n-th order replicator dynamics ($\mathrm{RD_n}$) is n orders as fast as in the standard first order dynamics ($\mathrm{RD_1}$), so irrational play becomes extinct much faster in higher orders.

Remark 4.2 (Irrelevance of adversarial play). Interestingly, the proof of Theorem 4.1 for dominated (but not *iteratively* dominated) strategies goes through unscathed for any (continuous) play $x_{-k}(t) \in X_{-k}$, $t \geq 0$, of k's opponents. As such, the extinction of dominated strategies is *independent* of how the focal player's opponents evolve over time – they need not even be rational.

Remark 4.3 (Payoff-monotonic dynamics). In the first order case, Samuelson and Zhang (1992) showed that in all payoff-monotonic dynamics, pure dominated strategies become extinct along interior solution orbits, while the same holds for mixed dominated strategies in the class of aggregate-monotonic dynamics. This result was extended by Hofbauer and Weibull (1996) to pure strategies which are dominated by mixed ones in convex-monotonic dynamics, whereas Viossat (2011) recently established the dual result for concave dynamics.

As it turns out, the second order landscape mirrors the first order one except for an accelerated rate of extinction (see Appendix B for the proof):

Theorem 4.3. *For any interior initial condition, we have:*

- *Payoff-monotonic n-th order dynamics eliminate all pure strategies that are dominated by pure strategies.*
- *Convex (resp. concave) monotonic n-th order dynamics eliminate all pure (resp. mixed) strategies that are dominated by mixed (resp. pure) strategies.*
- *Aggregate-monotonic n-th order dynamics eliminate all dominated strategies.*

Moreover, the rate of extinction is exponential in t^n (in the sense of (4.1)).

On the other hand, if a strategy is only *weakly* dominated, Theorem 4.1 cannot guarantee that it will be annihilated; in fact, it is well known that weakly

[11] In fact, the coefficients that make (4.1) and (4.2) into asymptotic equalities can also be determined, but we will not bother with this calculation here.

dominated strategies *do* survive even in the standard first order replicator dynamics: if the pure strategy $\alpha \in \mathcal{A}_k$ of player k is weakly dominated by $\beta \in \mathcal{A}_k$, and if all adversarial strategies $\alpha_{-k} \in \mathcal{A}_{-k}$ against which β performs better than α die out, then α may survive for an open set of initial conditions (for instance, see Example 5.4 and Proposition 5.8 in Weibull, 1995).

Quite remarkably, this can *never* be the case in a higher order setting if players start unbiased:

Theorem 4.4. *Let $x(t)$ be an interior solution orbit of the n-th order ($n \geq 2$) replicator dynamics (RD$_n$) that starts at rest: $\dot{x}(0) = \ldots = x^{(n-1)}(0) = 0$. If $q_k \in X_k$ is weakly dominated, then it becomes extinct along $x(t)$ with rate*

$$D_{\mathrm{KL}}(q_k \,||\, x_k(t)) \geq \lambda_k c\, t^{n-1} \big/ (n-1)!, \tag{4.3}$$

where λ_k is the learning rate of player k and $c > 0$ is a positive constant.

The intuition behind this surprising result (see Appendix B for the proof) can be gleaned by looking at the reinforcement learning scheme (LD$_\mathrm{n}$). If we take the case $n = 2$ for simplicity, we see that the "payoff forces" $F_{k\alpha} \equiv u_{k\alpha}$ will *never* point towards a weakly dominated strategy. As a result, solution trajectories are always accelerated away from weakly dominated strategies, and even if this acceleration vanishes in the long run, the trajectory still retains a growth rate velocity that drives it away from the dominated strategy. By comparison, this is not the case in first order dynamics; there, we only know that *growth rates* point away from weakly dominated strategies, and if these rates vanish in the long run, solution trajectories might ultimately converge to a point where weakly dominated strategies are still present (see for instance Fig. 2). In light of this, some further remarks are in order:

Remark 4.4. The assumption that solution orbits start at rest is simply there to ensure that players do not have an initial higher order "learning bias" in the form of uneven initial score derivatives $\dot{y}(0), \ddot{y}(0), \ldots \neq 0$ that might unduly skew their learning scheme towards one strategy or another.[12] As such, starting "at rest" is a very natural (and canonical) assumption to make: players may start with any mixed strategy they wish, but it is assumed that their learning process is not otherwise biased until their empirical data starts to accrue in (LD$_\mathrm{n}$).

Remark 4.5. As with strictly dominated strategies, Theorem 4.4 applies to more general imitative dynamics under the same caveats: aggregate-monotonic dynamics eliminate all weakly dominated strategies, convex (resp. concave) monotonic

[12]Indeed, $x_{k\alpha}^{(r)} = 0$ for all $r = 1, \ldots, n-1$, $\alpha \in \mathcal{A}_k$, also implies $y_{k\alpha}^{(r)} = y_{k\beta}^{(r)}$ for all $\alpha, \beta \in \mathcal{A}_k$, so being at rest is tantamount to a lack of learning bias.

dynamics eliminate pure (resp. mixed) strategies that are weakly dominated by mixed (resp. pure) strategies, and payoff-monotonic dynamics eliminate pure strategies that are weakly dominated by other pure strategies.

Remark 4.6. It is also important to note that our estimate of the rate of extinction of weakly dominated strategies is one order lower than that of strictly dominated ones; as a result, Theorem 4.4 does *not* imply the annihilation of weakly dominated strategies in first order dynamics (as well it shouldn't't!).

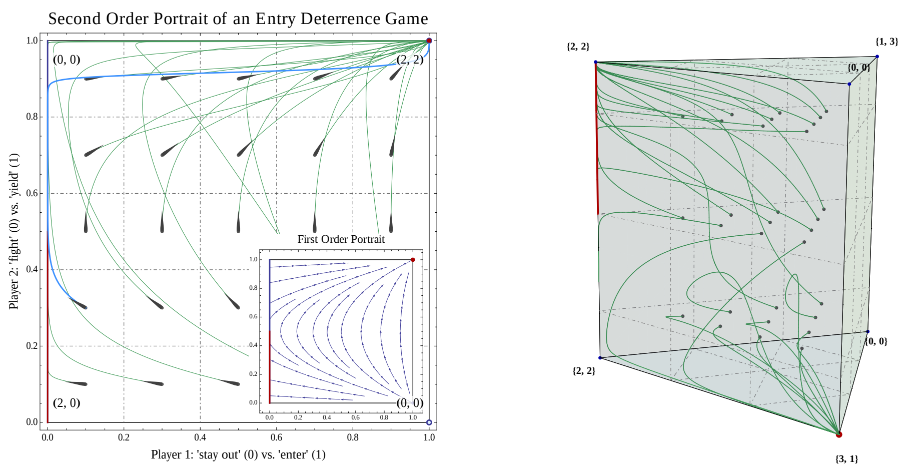
In the first order regime, it is known that if there is some adversarial strategy which constitutes evidence of domination (i.e. the weakly dominant strategy gives a greater payoff against it than the dominated one) and which always remains above a certain level, then the weakly dominated strategy becomes extinct (see e.g. Proposition 3.2 in Weibull, 1995). In our higher order setting, we show that this assumption instead implies that weakly dominated strategies become extinct as fast as *strictly* dominated ones:

Proposition 4.5. *Let $x(t)$ be an interior solution of the n-th order replicator dynamics (RD$_n$), and let $q_k \preccurlyeq q'_k$. If there exists $\alpha_{-k} \in \mathcal{A}_{-k}$ with $u_k(q_k;\alpha_{-k}) < u_k(q'_k;\alpha_{-k})$ and $x_{\alpha_{-k}}(t) \geq \varepsilon > 0$ for all $t \geq 0$, then:*

$$D_{\mathrm{KL}}(q_k \,\|\, x_k(t)) \geq \varepsilon \lambda_k \left[u_k(q'_k;\alpha_{-k}) - u_k(q_k;\alpha_{-k})\right] t^n / n! + \mathcal{O}(t^{n-1}). \tag{4.4}$$

Remark 4.7. In the first order replicator dynamics, the elimination of weakly dominated strategies when the evidence of their domination survives still requires that *all* players adhere to the same dynamics (see e.g. the proof of Proposition 3.2 in Weibull, 1995). To wit, consider a simple Entry Deterrence game where a competitor (Player 1) "enters" or "stays out" of a market controlled by a monopolist (Player 2) who can either "fight" the entrant or "share" the market, and where "fighting" is a weakly dominated strategy that yields a strictly worse payoff if the competitor "enters" (Weibull, 1995, Ex. 5.4). Under the replicator dynamics, "fight" becomes extinct if "enter" survives (cf. Figure 2); however, if Player 1 were to follow a different process under which "enter" survives but the integral of its population share over time is bounded, then "fight" does *not* become extinct. In higher orders though, the proof of Theorem 4.4 goes through for *any* continuous play $x_{-k}(t) \in X_{-k}$, $t \geq 0$, of k's opponents, so *weakly dominated strategies become extinct independently of how one's opponents evolve over time (rationally or otherwise).*

Remark 4.8. Finally, it is easy to show that Theorem 4.4 still holds if the players' initial velocities (or higher order derivates) are nonzero but small;if they are too large, weakly dominated strategies may indeed survive. This observation is important for strategies which are only *iteratively* weakly dominated because, if a



(a) Entry Deterrence

(b) Outside Option

Figure 2: Extinction of weakly dominated strategies and survival of iteratively weakly dominated ones under the second order replicator dynamics. Fig. 2(a) shows solution orbits starting at rest in an Entry Deterrence game (see labels for the payoffs): the weakly dominated strategy "fight" of Player 2 becomes extinct, in stark contrast to the first order case (compare the highlighted light blue trajectory with the first order portrait in the inlay). On the other hand, Fig. 2(b) shows an Outside Option supergame where the strategy "fight" in Fig. 2(a) is only *iteratively* weakly dominated; as it happens, this strategy pays very well against certain initial conditions, so it ends up surviving when all evidence that it is (iteratively weakly) dominated vanishes. (In both figures, Nash equilibria have been highlighted in dark red.)

strategy becomes weakly dominated after removing a strictly dominated strategy, then the system's solutions could approach the face associated with the resulting subgame with a high velocity towards the newly weakly dominated strategy (e.g. if the iteratively weakly dominated strategy pays very well against the evanescent strictly dominated one; cf. Fig. 2). Thus, although Theorem 4.4 guarantees the elimination of weakly dominated strategies, it does not extend to *iteratively* weakly dominated ones.

5 Stability of Nash Play and the Folk Theorem

In games that cannot be solved by the successive elimination of dominated strategies, one usually tries to identify the game's Nash equilibria instead. Unfortunately, given the prohibitive complexity of these solutions (Daskalakis, Goldberg, and Papadimitriou, 2006), one of the driving questions of evolutionary game theory has been to explain how Nash play might emerge over time as the byproduct of a simpler, adaptive dynamic process.

A key result along these lines is the *folk theorem of evolutionary game theory* (Hofbauer and Sigmund, 1988; Sandholm, 2011; Weibull, 1995), which, for the multi-population replicator dynamics (RD_1), can be summarized as follows:

I. Nash equilibria are stationary.

II. If an interior solution orbit converges, its limit is Nash.

III. If a point is Lyapunov stable, then it is also Nash.

IV. A point is asymptotically stable if and only if it is a strict equilibrium.

Accordingly, our aim in this section will be to derive an analogue of the folk theorem in the context of the higher order dynamics (RD_n). To that end however, it is important to recall that the higher order playing field is fundamentally different: as has been stated before, the choice of an initial strategy profile $x(0) \in X$ does not suffice to determine the evolution of (RD_n). Instead, one now needs to prescribe the full initial state $\omega(0) = (x(0), \dot{x}(0), \dots)$ in the system's phase space Ω, including the initial velocity and other higher order derivative components (up to order $n-1$).

Surprisingly, despite these differences, the folk theorem of evolutionary game theory continues to hold almost verbatim in our higher order setting:

Theorem 5.1. *Let $x(t)$ be a solution orbit of the n-th order replicator dynamics (RD_n) for a normal form game $\mathfrak{G} \equiv \mathfrak{G}(\mathcal{N}, \mathcal{A}, u)$, and let $q \in X$. Then:*

I. $x(t) = q$ *for all* $t \geq 0$ *if and only if* q *is a restricted equilibrium of* $\mathfrak{G}$.

II. *If* $x(0) \in \operatorname{int}(X)$ *and* $\lim_{t\to\infty} x(t) = q$, *then* q *is a Nash equilibrium of* $\mathfrak{G}$.

III. *If every neighborhood* U *of* q *in* X *admits an interior orbit* $x_U(t)$ *such that* $x_U(t) \in U$ *for all* $t \geq 0$, *then* q *is a Nash equilibrium of* $\mathfrak{G}$.

IV. *Let* q *be a strict equilibrium. Then, for every neighborhood* U *of* q *in* X, *there exists a neighborhood* V *of* q *in* X *and a neighborhood* W *of* $V \setminus \{q\}$ *in* Ω *such that* $x(t) \in U$ *and* $x(t) \to q$ *for all initial states* $(x(0), \dot{x}(0), \dots) \in W$; *conversely, only strict equilibria have this property.*

Moreover, as an immediate corollary of (IV), we also have:

IV$'$. *If* q *is a strict equilibrium of* $\mathfrak{G}$, *then there exists a neighborhood* U *of* q *in* X *such that* $x(t) \to q$ *whenever* $x(t)$ *starts at rest in* U *(that is,* $x(0) \in U$ *and* $\dot{x}(0) = \ldots = 0$*); conversely, only strict equilibria have this property.*

For the proof, see Appendix C; for now, the following remarks are in order:

Remark 5.1 (Points vs. states and the standard folk theorem). A natural way to discuss the stability of initial *points* $q \in X$ is via the corresponding *rest states* $(q, 0, \dots, 0) \in \Omega$ (recall also the relevant discussion in Section 2.3 and the remarks following Theorem 4.4). With this in mind, we will say that $q \in X$ is *stationary* (resp. *Lyapunov stable*, resp. *attracting*) when the associated rest state $(q, 0, \dots, 0) \in \Omega$ is itself stationary (resp. Lyapunov stable, resp. attracting).

Given this duality between points $q \in X$ and rest states $(q, 0, \dots, 0) \in \Omega$, we may draw the following parallels between the folk theorem and Theorem 5.1:

– **Parts I and II** of Theorem 5.1 are direct analogues of the corresponding first order claims; note however that (II) can now be inferred from (III).

– **Part III** is slightly stronger than the first order statement that Lyapunov stability implies Nash equilibrium. Indeed, Lyapunov stability posits robustness with respect to open neighborhoods of initial conditions (including higher order components in higher order environments), whereas Theorem 5.1 only asks that every neighborhood of the point in question admit a trajectory which is wholly contained therein. As a matter of fact, there are equilibria that, even under the standard replicator dynamics, satisfy the latter property but not the former;[13] as such, Part III of Theorem 5.1 is closer to a "bare minimum" stability characterization of Nash equilibria, especially in higher orders.

[13]For instance, the equilibrium $q = e_{11} + (e_{21} + e_{22})/2$ of the simple 2×2 game with payoff matrices $U_1 = I$ and $U_2 = 0$ is not Lyapunov stable under the replicator dynamics and is not the ω-limit of any interior trajectory, but it still satisfies the property asserted in Theorem 5.1.

– **Part IV** on the other hand is *not* tantamount to asymptotic stability – it would be if W were a neighborhood of V instead of $V\setminus\{q\}$.[14] In particular, Part IV′ shows that strict equilibria attract all nearby rest states,[15] but not all nearby non-rest states: for every nearby point $x \in X$, we can find a neighborhood $V_x \subseteq \Omega$ of initial states that converge to q, but there is no *uniform* bound on, say, initial velocities $\dot{x}(0)$ that ensures convergence to q.

This difference between first and higher orders is inextricably tied to the *sine qua non* requirement that any higher order dynamical system needs first and foremost to stay in X. Specifically, for small (but finite) $\dot{x}, \ddot{x}$, etc., the boundary term (3.16) in ($\mathrm{RD_n}$) blows up to infinity as $x_{k\alpha} \to \infty$, so the dynamics in the phase space Ω will become perpendicular to X (viewed as a subset of Ω) near $(q, 0, \ldots, 0)$, and this precludes asymptotic stability (see Fig. 3).

That said, asymptotic stability is not only too strong a requirement for higher order settings, but it is also a not very relevant one: given that players do not have direct control over their initial velocities $\dot{x}(0)$, it makes little sense to ask for robustness against different "velocity choices" by the players. What players *do* control instead is the bias $\dot{y}(0), \ddot{y}(0), \ldots$ of their learning scheme as captured by the dynamics ($\mathrm{LD_n}$); on that account, asymptotic stability should be phrased instead in terms of the players' initial *bias* in ($\mathrm{LD_n}$), and not w.r.t. their initial "velocities" in ($\mathrm{RD_n}$) – obviously, a redundant distinction for $n = 1$.

To wit, if we assume w.l.o.g. that the strict equilibrium under scrutiny corresponds to everyone playing their "0"-th strategy, the learning dynamics ($\mathrm{LD_n}$) phrased in terms of the relative scores $z_{k\mu}$ of (3.17) take the equivalent form:

$$z_{k\mu}^{(n)} = u_{k\mu}(x) - u_{k,0}(x), \tag{$\mathrm{ZD_n}$}$$

with $x = G^*(z)$ given by the reduced Gibbs map (GM*). In this formulation, the strict equilibrium $q = (0, \ldots, 0)$ corresponds to the point at negative infinity $(-\infty, \ldots, -\infty)$ and the proof of Theorem 5.1 (see Appendix C) shows that if players start at a neighborhood of $(-\infty, \ldots, -\infty)$ with learning bias $\dot{z}_{k\mu}(0), \ldots$ not exceeding some *uniform* $M > 0$, then the relative scores $z_{k\mu}$ escape to $-\infty$. In other words, *if we view the strict equilibria of $\mathfrak{G}$ as points at infinity, then they are stable and attracting in the higher order learning dynamics (ZD_n).*

Remark 5.2. As in the first order case, Theorem 5.1 applies to more general classes of imitative dynamics. In the context of payoff-monotonic dynamics in particular, our proof goes through unchanged except for the converse implication of Part IV for $n = 1$ (see also Theorem 5.4 below).

[14]We thank Josef Hofbauer for this remark.

[15]Recall that V is canonically embedded in Ω via the rest map $x \in X \mapsto (x, 0, \ldots, 0) \in \Omega$.

Now, with regards to the equilibration speed of the higher order dynamics, it can be shown that *the rate of convergence to a strict equilibrium in the n-th order dynamics (RD_n) is n orders as fast as in the first order case*. Specifically, we have:

Proposition 5.2. *Let $q = (e_{1,0}, \ldots, e_{N,0})$ be a strict Nash equilibrium of the finite game $\mathfrak{G}$, and let $x(t)$ be a solution path of the replicator dynamics (RD_n) which starts at rest and close enough to q. Then, there exists $c > 0$ such that:*

$$x_{k,0}(t) \sim 1 - \exp\left(-ct^n/n! + \mathcal{O}(t^{n-1})\right). \tag{5.1}$$

Theorem 5.1 and Proposition 5.2 (see App. C for the proof) characterize the behavior of the n-th order replicator dynamics near strict equilibria from both a qualitative and a quantitative viewpoint; on the flip side, they do little to address mixed equilibria. To study this issue, recall first that the standard asymmetric replicator dynamics preserve a certain volume form in the interior of X, so mixed equilibria cannot be attracting in first order. Ritzberger and Weibull (1995) establish this "incompressibility" property of the replicator dynamics by taking an ingenious extrinsic reparametrization which makes the replicator dynamics divergence-free in the interior of X (see also Ritzberger and Vogelsberger, 1990). On the other hand, Hofbauer and Sigmund (1988) rely implicitly on the properties of the Gibbs map (GM), and essentially show that the replicator dynamics are incompressible in the space of the score variables $y_{k\alpha}$ (see also Hofbauer, 1996). In first order, this idea can be used to show that the generalized imitative dynamics (GD_n) are volume-preserving whenever the payoff observables $w_{k\alpha}$ do not depend on $x_{k\alpha}$; in higher orders however, the dynamics are further decoupled because the $w_{k\alpha}$ are only tied to the players' mixed strategies and not their velocities. As a result, we obtain:

Proposition 5.3. *The flow of the generalized learning dynamics (GLD_n) is volume-preserving in the usual Euclidean geometry for all $n \geq 2$; the same holds for (GD_n) w.r.t. a non-Euclidean volume form on the system's phase space Ω.*

More importantly, by using Proposition 5.3, we can prove the stronger result:

Theorem 5.4. *In the generalized higher order dynamics (GD_n), $n \geq 2$, interior points cannot attract open sets of initial states; only vertices of X can be attracting. More generally, a non-pure point $q \in X$ can only attract relatively open sets of initial states whose support in X properly contains that of q.*

It is important to stress that Theorem 5.4 clashes rather strongly with the first order case $n = 1$. For instance, if we take Maynard Smith's payoff-adjusted variant of the replicator dynamics (whereby players divide (RD_1) by their average payoffs), then interior equilibria may become asymptotically stable (for instance, as in the Matching Pennies example of Weibull, 1995). In higher orders however, this is no longer the case: the learning dynamics ($\mathrm{LD_n}$) endow orbits with a tangential acceleration component, and this acceleration carries them away from interior equilibria and towards the boundary of X.

This tangential component can be tempered by including friction terms of the form $F = -\dot{y}$. In that case, the resulting dynamics cease to be incompressible and become dissipative, so trajectories may well converge to interior states; however, due to space limitations, we will not address this issue here.

6 Concluding Remarks and Future Directions

The results in the present paper suggest that higher order dynamics open the door to some intriguing new questions and directions in the study of learning and evolution in games. For one, the elimination of weakly dominated strategies is a key feature of higher order dynamics which puts them firmly apart from all their first order siblings; for another, even though the classes of higher order imitative dynamics considered here do not converge to interior equilibria, the very nature of higher order learning allows the introduction of *friction* terms which slow down trajectories and allow them to converge to interior states (a forthcoming result). In fact, we have only scratched the surface here: the higher order regime offers a extremely wide array of adjusted or altogether new classes of dynamics that simply cannot be obtained in a first order setting, and where the impossibility theorem of Hart and Mas-Colell (2003) no longer bars the way.

We should also stress here that it is hard to underestimate the role that the Gibbs choice model (GM) plays in this higher order environment. In particular, the form and properties of the (game-independent) adjustment term (3.16) in the higher order replicator dynamics ($\mathrm{RD_n}$) are direct consequences of the Gibbs model (GM), so one might ask what would happen with a different adjustment term, possibly not stemming from (GM). Surprisingly (and despite the fact that ($\mathrm{RD_n}$) retains most of its rationality properties if the replicator term is replaced by some other payoff-monotonic term), if the adjustment term (3.16) is multiplied by something as innocuous as a real number $a \in (0,1)$, then *even strictly dominated strategies may survive* (another forthcoming result). As such, the caution underlying our derivation of ($\mathrm{RD_n}$) and ($\mathrm{GD_n}$) seems to be well-justified: in the absence of solid foundations in learning, a naïve incorporation of higher order effects à la (3.2b′) may well lead to undesirable results.

That said, even in our exponential learning framework, there still remain many open (and important) questions: For instance, when is this learning pro-

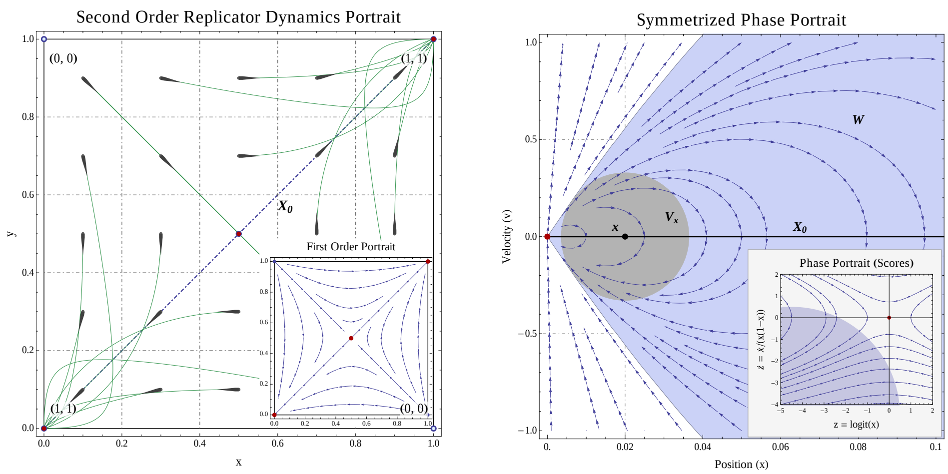


Figure 3: Second order replicator trajectories in a 2×2 coordination game with payoff matrices $U_1 = U_2 = I$. The figure to the right shows the restriction of the game's phase space to the symmetric invariant manifold X_0 which joins the game's equilibria. For every symmetric initial point x near $q = (0, 0)$, there exists a neighborhood of initial states V_x (gray) such that all orbits starting in V_x stay close and eventually converge to q. The union W of these neighborhoods (light blue) is not itself a neighborhood of q in $\Omega_0 \equiv \Omega(X_0)$, so q is not asymptotically stable in ($\mathrm{RD_n}$); however, in terms of the score variables $z = \operatorname{logit} x$, $\dot{z} = \dot{x}/x(1-x)$, the corresponding point at infinity $(-\infty, -\infty)$ *is* asymptotically stable in ($\mathrm{LD_n}$) (inlay).

cess consistent (e.g. externally, as in Sorin, 2009)? What can we expect in symmetric, single-population environments (where payoffs are no longer multilinear) or with respect to setwise solution concepts (such as sets that are closed under better replies as in Ritzberger and Weibull, 1995)? More generally, what can we expect if we move beyond a dynamical framework altogether and replace the learning process (3.13) with a more general *integral* equation of the form $Y_{k\alpha}(t) = \int_0^t \phi(t-s) u_{k\alpha}(x(s))\, ds$ where the "learning kernel" ϕ describes the weight that players assign to their past observations? The n-th order replicator dynamics can be shown to correspond to monomial learning kernels of the form $\phi(t) = t^n$, but how would learning be affected by e.g. an exponential discounting (or reinforcement) of the past?

Finally, it is important to note that our approach has been focused on continuous time with players being able to observe (or otherwise calculate) the payoffs associated to mixed strategies (the last assumption being relatively harmless in a nonatomic population setting, but potentially important from a discrete point of view). This choice has been a conscious one and it was motivated by the fact that our goal was simply to illustrate the rationality properties of the limiting continuous-time dynamics; the subtleties (and there are many!) of the descent from the continuous to the discrete and from the population to the atom would take us too far afield, so we leave this issue as a future direction (the papers by Rustichini, 1999, and Sorin, 2009, may serve as an indication of what to expect in the discrete case). Needless to say however, these are all directions that would take much more than a single paper to explore, so we chose to exhibit here only some of the most prominent features and subtleties of the higher order landscape, deferring these questions to future investigations.

Acknowledgments

This paper has benefited greatly from many discussions with Josef Hofbauer (to whom we owe the distinction between asymptotic stability in the game's strategy space and its score space), Aris Moustakas, Christina Pawlowitsch, Sylvain Sorin and Yannick Viossat; we would also like to thank the audiences of the Paris Game Theory Seminar, the Paris Working Group on Evolutionary Games, and the 2011 International Conference on the Mathematical Aspects of Game Theory and its Applications for their insightful comments and questions (Bill Sandholm and Jörgen Weibull in particular).

This work was carried out when the second author was visiting the Economics Department of École Polytechnique, Paris, France, and we would like to thank the École Polytechnique for its generous support and hospitality during 2010-11. The authors also gratefully acknowledge financial support from the Agence Nationale de la Recherche (grant ANR-10-BLAN 0112) and the GIS "Sciences de la décision".

A Derivation of Higher Order Dynamics

The starting point for the derivation of ($\mathrm{RD_n}$) is the identity

$$\log(x_{k\alpha}) - \log(x_{k\beta}) = \lambda_k(y_{k\alpha} - y_{k\beta}), \tag{A.1}$$

itself an easy consequence of (GM). Hence, by Faà di Bruno's higher order chain rule (Fraenkel, 1978), we readily obtain:

$$\frac{d^n}{dt^n}\log(x_{k\alpha}(t)) = \sum \frac{n!}{m_1!\cdots m_n!}\frac{(-1)^{m-1}(m-1)!}{x_{k\alpha}^m}\prod\nolimits_{r=1}^n \left(x_{k\alpha}^{(r)}(t)/r!\right)^{m_r}, \tag{A.2}$$

where we have set $m = m_1 + \cdots + m_n$ and the sum is taken over all non-negative integers $m_1, \ldots, m_n$ such that $\sum_{r=1}^n r\,m_r = n$. In particular, since the only term that contains $x_{k\alpha}^{(n)}$ has $m_1 = m_2 = \ldots = m_{n-1} = 0$ and $m_n = 1$, we may rewrite (A.2) as:

$$\frac{d^n}{dt^n}\log(x_{k\alpha}(t)) = \frac{x_{k\alpha}^{(n)}(t)}{x_{k\alpha}(t)} + R_{k\alpha}^{n-1}\big(x(t), \dot{x}(t), \ldots, x^{(n-1)}(t)\big), \tag{A.3}$$

where $R_{k\alpha}^{n-1}$ denotes the $(n-1)$-th order remainder of the RHS of (A.2) and is given by (3.16). Then, by taking the n-th derivative of (A.1) and substituting, we get:

$$\lambda_k\left(u_{k\alpha} - u_{k\beta}\right) = \frac{x_{k\alpha}^{(n)}}{x_{k\alpha}} - \frac{x_{k\beta}^{(n)}}{x_{k\beta}} + R_{k\alpha}^{n-1} - R_{k\beta}^{n-1}, \tag{A.4}$$

and ($\mathrm{RD_n}$) follows by applying $\sum_\beta^k x_{k\beta}(\cdot)$ on both sides (recall that $\sum_\beta^k x_{k\beta}^{(n)} = 0$).

B On the Elimination of Dominated Strategies

Proof of Theorem 4.1. We will begin by showing that if $q_k \in X_k$ is dominated by $q'_k \in X_k$, then $D_{\mathrm{KL}}(q_k \,\|\, x_k(t)) \geq ct^n/n!$ for some positive constant $c > 0$. Indeed, let $V_k(x) = D_{\mathrm{KL}}(q_k \,\|\, x_k) - D_{\mathrm{KL}}(q'_k \,\|\, x_k)$, and rewrite the Gibbs distribution (GM) as $\log x_{k\alpha} = \lambda_k y_{k\alpha} - \log(Z(y))$ where $Z(y) = \sum_\beta^k \exp(\lambda_k y_{k\beta})$ is the *partition function* of player k. Then, some algebra yields:

$$\begin{aligned} V_k(x) &= \sum_{\alpha\in\mathrm{supp}(q)} q_{k\alpha}\log\left(q_{k\alpha}/x_{k\alpha}\right) \; - \sum_{\alpha\in\mathrm{supp}(q')} q'_{k\alpha}\log\left(q'_{k\alpha}/x_{k\alpha}\right) \\ &= \sum\nolimits_\alpha^k \left(q'_{k\alpha} - q_{k\alpha}\right)\log x_{k\alpha} + h_k(q_k, q'_k) \\ &= \sum\nolimits_\alpha^k \left(q'_{k\alpha} - q_{k\alpha}\right)\lambda_k y_{k\alpha} + h_k(q_k, q'_k), \end{aligned} \tag{B.1}$$

where $h_k(q_k, q'_k)$ is a constant depending only on q_k and q'_k, and the last equality follows from the fact that $\sum_\alpha^k (q'_{k\alpha} - q_{k\alpha})\log Z = 0$ (recall that $\sum_\alpha^k q_{k\alpha} = \sum_\alpha^k q'_{k\alpha} = 1$). In this way, we readily obtain:

$$\begin{aligned} \frac{d^n}{dt^n}V_k(x(t)) &= \lambda_k \sum\nolimits_\alpha^k \left(q'_{k\alpha} - q_{k\alpha}\right) y_{k\alpha}^{(n)} = \lambda_k \sum\nolimits_\alpha^k \left(q'_{k\alpha} - q_{k\alpha}\right) u_{k\alpha}(x(t)) \\ &= \lambda_k\left[u_k(q'_k; x_{-k}(t)) - u_k(q_k; x_{-k}(t))\right] \geq \lambda_k \Delta u_k > 0, \end{aligned} \tag{B.2}$$

where the constant Δu_k is defined as $\Delta u_k = \min_{X_{-k}}\{u_k(q'_k; x_{-k}) - u_k(q_k; x_{-k})\}$ and its positivity follows from the compactness of X. Hence, if we set $c_r =$

$(r!)^{-1} \left.\frac{d^r V_k}{dt^r}\right|_{t=0}$, $r = 0\dots n-1$, Taylor's theorem with Lagrange remainder readily gives:

$$V_k(x(t)) \geq \lambda_k \Delta u_k t^n / n! + \sum_{r=0}^{n-1} c_r t^r, \tag{B.3}$$

and our assertion follows by noting that $D_{\mathrm{KL}}(q_k \| x_k(t)) \geq V_k(x(t))$. In particular, for pure strategies $\alpha \prec \beta$, we will have $V_k(x(t)) = \log\left(x_{k\beta}(t)/x_{k\alpha}(t)\right)$, so (B.3) gives:

$$\log\left(x_{k\beta}(t)/x_{k\alpha}(t)\right) \geq \lambda_k \Delta u_{\beta\alpha} t^n / n! + \mathcal{O}(t^{n-1}), \tag{B.4}$$

and (4.2) follows by exponentiating.

Now, to establish the theorem for *iteratively* dominated strategies, we will resort to induction on the rounds of elimination. To that end, assume that $D_{\mathrm{KL}}(q_k \| x_k(t)) = |\mathcal{O}(t^n)|$ for all strategies $q_k \notin X_k^r$ that do not survive r elimination rounds; in particular, if $\alpha \notin \mathcal{A}_k^r \equiv \mathcal{A}_k \cap X_k^r$, we assume that $x_{k\alpha}(t) \to 0$ as $t \to \infty$. We will show that this also holds if $q_k \in X_k^r$ survives for r deletion rounds but dies in the subsequent one.

Indeed, if $q_k \in X_k^r \setminus X_k^{r+1}$, there will be some $q_k' \in X_k^r$ with $u_k(q_k'; x_{-k}) > u_k(q_k; x_{-k})$ for all $x_{-k} \in X_{-k}^r$. With this in mind, decompose $x \in X$ as $x = x^r + z^r$ where x^r denotes the "r-rationalizable" part of x, i.e. the orthogonal projection of x on the subspace of X spanned by the surviving pure strategies $\mathcal{A}_\ell^r$, $\ell \in \mathcal{N}$. Then, if we set $\Delta u_k^r = \min\{u_k(q_k'; \alpha_{-k}) - u_k(q_k; \alpha_{-k}) : \alpha_{-k} \in \mathcal{A}_{-k}^r\}$, we will also have by (multi)linearity:

$$u_k(q_k'; x_{-k}^r) - u_k(q_k; x_{-k}^r) \geq \Delta u_k^r > 0 \quad \text{for all } x_{-k} \in X_{-k}. \tag{B.5}$$

Moreover, it is easy to see that our induction hypothesis implies $z^r(t) \to 0$ as $t \to \infty$ (recall that $x_{k\alpha}(t) \to 0$ for all $\alpha \notin \mathcal{A}_k^r$), so for large enough t, we also get:

$$|u_k(q_k'; z_{-k}^r(t)) - u_k(q_k; z_{-k}^r(t))| < \Delta u_k^r / 2. \tag{B.6}$$

Hence, by combining (B.5) and (B.6), we obtain $u_k(q_k'; x_{-k}(t)) - u_k(q_k; x_{-k}(t)) > \Delta u_k^r / 2$ for large t, and the induction is completed by plugging this last estimate into (B.2) and proceeding as in the base case $r = 0$ (our earlier assertion). □

Proof of Theorem 4.3. The crucial point in the previous proof is the n-th derivative of the entropy difference $V_k(x) = D_{\mathrm{KL}}(q_k \| x_k) - D_{\mathrm{KL}}(q_k' \| x_k)$ which determines the rate of extinction of dominated strategies. Thus, by replacing u by w in (B.2) and using the appropriate monotonicity condition for each case of dominance (pure/mixed by pure/mixed), Theorem 4.3 follows by shadowing the proof of Theorem 4.1. □

Proof of Theorem 4.4. Let $q_k \preccurlyeq q'_k$ and let $\mathcal{A}'_{-k} \equiv \{\alpha_{-k} \in \mathcal{A}_{-k} : u_k(q'_k;\alpha_{-k}) > u_k(q_k;\alpha_{-k})\}$ be the set of pure strategy profiles of k's opponents against which q'_k yields a strictly greater payoff than q_k. Then, with notation as above, we will have:

$$\frac{d^n}{dt^n} V_k(x(t)) = \lambda_k \sum\nolimits_{\alpha_{-k}\in\mathcal{A}'_{-k}} \left[u_k(q'_k;\alpha_{-k}) - u_k(q_k;\alpha_{-k})\right] x_{\alpha_{-k}}(t), \tag{B.7}$$

where $x_{\alpha_{-k}} \equiv \prod_{\ell\neq k} x_{\alpha_\ell}$ denotes the α_{-k}-th component of x. Thus, with $x(t)$ starting at rest, Faà di Bruno's formula gives $\left.\frac{d^r V_k}{dt^r}\right|_{t=0} = 0$ for all $r = 1,\dots,n-1$, and a simple integration then yields:

$$\frac{d^{n-1}}{dt^{n-1}} V_k(x(t)) = \lambda_k \sum\nolimits_{\alpha_{-k}\in\mathcal{A}'_{-k}} \left[u_k(q'_k;\alpha_{-k}) - u_k(q_k;\alpha_{-k})\right] \int_0^t x_{\alpha_{-k}}(s)\,ds, \tag{B.8}$$

However, with $x(t)$ interior, the integrals in the above equation will be positive and increasing, so for some suitably chosen $c > 0$ and t large enough, we obtain

$$\frac{d^{n-1}}{dt^{n-1}} V_k(x(t)) \geq \lambda_k c > 0, \tag{B.9}$$

and our claim follows from a $(n-1)$-fold application of the mean value theorem. □

C Convergence and Non-Convergence Results

Proof of Theorem 5.1. We will begin with stationarity of restricted equilibria. Indeed, since the payoff term of (RD$_n$) does not contain any higher order derivatives, it will vanish at $q \in X$ if and only if $u_{k\alpha}(q) = u_k(q)$ for all $\alpha \in \operatorname{supp}(q)$, implying that q is a restricted equilibrium. Conversely, let q be a Nash equilibrium in the subgame $\mathfrak{G}' \equiv \mathfrak{G}(\mathcal{N}, X', u_k|_{X'})$ with $\mathcal{A}'_k = \operatorname{supp}(q_k)$. Then, with $u_{k\alpha}(q) = u_{k\beta}(q)$ for all $\alpha,\beta \in \mathcal{A}'_k$, the updating scheme (LD$_n$) constrained to $\mathfrak{G}'$ and starting at q also gives $y^{(n)}_{k\alpha}(0) = y^{(n)}_{k\beta}(0)$ for all $\alpha,\beta \in \mathcal{A}'_k$. So, if (RD$_n$) starts at q with initial motion rates $\dot{x}(0) = \ddot{x}(0) = \cdots = 0$, we will have $y_{k\alpha}(t) - y_{k\alpha}(0) = y_{k\beta}(t) - y_{k\beta}(0)$ for all $\alpha,\beta \in \mathcal{A}'_k$, and, by the homogeneity of the Gibbs map ($G(y_1 + c, y_2 + c, \dots) = G(y_1, y_2, \dots)$ for all $c \in \mathbb{R}$), we readily obtain $x(t) = q$ for all t, i.e. q is stationary.[16]

[16] It is important to note here that Nash equilibria are *not* stationary in (LD$_n$): trajectories that are stationary in X may track a line parallel to the vector $(1,\dots,1)$ in $\mathbb{R}^{\mathcal{A}}$.

We now turn to Part (III) of the theorem – which will also prove Part (II). To that end, suppose that every neighborhood U of q in X admits an interior orbit $x(t)$ that stays in U for all $t \geq 0$; we then claim that q is Nash. Indeed, assume instead that for some $k \in \mathcal{N}$, there exists $\beta \in \mathcal{A}_k$ and $\alpha \in \operatorname{supp}(q_k)$ with $u_{k\alpha}(q) < u_{k\beta}(q)$. Then, pick $\varepsilon > 0$ and a neighborhood U of q such that $x_{k\alpha} > q_{k\alpha}/2 > 0$ and $u_{k\beta}(x) \geq u_{k\alpha}(x) + \varepsilon$ for all $x \in U$. By assumption, there exists an interior orbit $x(t)$ which stays in U for all time, so, for the associated score variables $y(t)$, we will have:

$$y_{k\beta}^{(n)}(t) - y_{k\alpha}^{(n)}(t) = u_{k\beta}(x(t)) - u_{k\alpha}(x(t)) \geq \varepsilon > 0.$$

This last inequality immediately implies that $\log\left(x_{k\beta}(t)/x_{k\alpha}(t)\right) \to +\infty$, contradicting the fact that $x_{k\alpha}(t) > q_{k\alpha}/2$ for all $t \geq 0$.

With regards to Part (IV), let $q = (e_{1,0}, \dots, e_{N,0})$ be a strict equilibrium of $\mathfrak{G}$, and consider the relative scores $z_{k\mu} = y_{k\mu} - y_{k,0}$, $\mu \in \mathcal{A}_k^* \equiv \mathcal{A}_k \setminus \{0\}$. It is then easy to see that the reduced Gibbs map $G_k^* \colon \mathbb{R}^{\mathcal{A}_k^*} \to X_k$ of (GM*) is a diffeomorphism onto its image, so the same will hold for the direct sum $G^* \equiv \bigoplus_k G_k^* \colon \mathbb{R}^{\mathcal{A}^*} \to X$ as well. Accordingly, if we take a neighborhood U_ε of q in X of the form $U_\varepsilon = \{x \in \operatorname{int}(X) : x_{k,0} \geq 1 - \varepsilon, k \in \mathcal{N}\}$, its preimage under G^* will be the set $V_h = \{z \in \mathbb{R}^{\mathcal{A}^*} : z_{k,0} \leq h, k \in \mathcal{N}\}$ where $h = (1-\varepsilon)^{-1} - 1$ ($\approx \varepsilon$ for small ε). We will show that if h is chosen small enough, then there exists $\delta > 0$ such that whenever a solution $z(t)$ of (ZD$_\text{n}$) starts at $z(0) \in V_\varepsilon$ with $\|z^{(r)}(0)\| \leq \delta$ for $r = 1, \dots, n-1$, we have $z(t) \in V_{2h}$ for all $t \geq 0$ and $z_{k\mu}(t) \to -\infty$ for all $\mu \in \mathcal{A}_k^*$, $k \in \mathcal{N}$. Since G^* is a diffeomorphism onto its image and $x \to q$ iff $z_{k\mu} \to -\infty$ for all $\mu \in \mathcal{A}_k^*$, $k \in \mathcal{N}$, this will establish the "if" direction of our claim.[17]

Indeed, let $z(t)$ be a solution of (ZD$_\text{n}$) starting in V_h and let $\tau_{2h} = \inf\{t : z(t) \notin V_{2h}\}$ be the time it takes $z(t)$ to escape from V_{2h} (with the usual convention $\inf(\varnothing) = \infty$). Then, if h is taken small enough, there will be some constant $M > 0$ such that $u_{k,0}(x) - u_{k,\mu}(x) \geq M > 0$ for all $x \in G^*(V_{2h})$ (recall that q is a strict equilibrium), so for $t \leq \tau_{2h}$, Taylor's theorem with Lagrange remainder applied to (ZD$_\text{n}$) gives:

$$z_{k\mu}(t) \leq z_{k\mu}(0) + \sum\nolimits_{r=1}^{n-1} z_{k\mu}^{(r)}(0)\, t^r / r! - M t^n / n!. \tag{C.1}$$

Hence, pick $\delta > 0$ such that the maximum of the polynomial $\sum_{r=1}^{n-1} z_{k\mu}^{(r)}(0)\, t^r / r! - M t^n / n!$ for $t \geq 0$ is strictly smaller than $\log 2$ whenever $|z_{k\mu}^{(r)}(0)| < \delta$ for $\mu \in \mathcal{A}_k^*$,

[17] Non-interior trajectories can be handled similarly by looking at the appropriate subgame.

$k \in \mathcal{N}$, and $r = 1, \dots n-1$. This readily yields $Z_{k,0}(\tau_{2h}) < \sum_{\mu}^{k} \exp\left(z_{k\mu}(0) + \log 2\right) = 2Z_{k,0}(0) < 2h$, i.e. $z(\tau_{2h}) \in V_{2h}$, a conclusion which cannot hold unless $\tau_{2h} = \infty$. We thus obtain $z(t) \in V_{2h}$ for all $t \geq 0$, so the limit of (C.1) as $t \to \infty$ gives $z_{k\mu}(t) \to -\infty$. Then, by using Taylor expansions of a lower order, one can show in a similar fashion that the same also holds for the derivatives $\dot{z}_{k\mu}(t), \dots, z_{k\mu}^{(n-1)}(t)$, as was to be shown.

For the converse implication, it is easy to show that any vertex q of X which attracts an open neighborhood of initial rest states must also be a strict Nash equilibrium: extending the reasoning of Ritzberger and Weibull (1995, Thm. 1) to our higher order setting, it suffices to consider the evolution of the dynamics in the edge which joins $q = (\alpha_k; \alpha_{-k})$ to a vertex $q' = (\alpha_k'; \alpha_{-k})$ with $u_k(q') \geq u_k(q)$. However, Theorem 5.4 shows that only a vertex $q \in X$ can attract an open set of initial states $\omega \in \Omega$ which contains a punctured neighborhood of q in X, so our assertion follows. □

Proof of Proposition 5.2. By Theorem 5.1, we know that if $x(t)$ starts close enough to q and is initially at rest, then it will always remain close to q. As such, by choosing a sufficiently small neighborhood of initial positions, the payoff differences $u_{k,0}(x(t)) - u_{k,\mu}(x(t))$ will be bounded away from 0 by some positive constant $c > 0$ for all $\mu \in \mathcal{A}_k^*$, $k \in \mathcal{N}$ and for all $t \geq 0$, so (ZD$_n$) gives $z_{k\mu}^{(n)} \leq -c < 0$ as well, and our assertion follows from an $(n-1)$-fold application of the mean value theorem. □

Proof of Proposition 5.3. Similarly to (3.14), (GLD$_n$) can be written as:

$$\begin{aligned} \dot{y}_{k\alpha}^{n-1}(t) &= w_{k\alpha}(x(t)), \\ &\dots \\ \dot{y}_{k\alpha}^{0}(t) &= y_{k\alpha}^{1}(t), \end{aligned} \tag{C.2}$$

so, given that $y_{k\alpha}^r$ does not appear in the equation for $\dot{y}_{k\alpha}^r$, it follows that the flow of (GLD$_n$) is incompressible in the standard Euclidean metric of $\mathbb{R}^{\mathcal{A}}$.[18] Using the relative scores z, the same argument applies to the dynamics (ZD$_n$) with u replaced by w, and since G^* is a local diffeomorphism, the result carries over to (GD$_n$) as well. □

Proof of Theorem 5.4. We will prove that if $q \in \text{int}(X)$, then there is no open set of initial conditions in Ω that converges to q. The result for general non-pure

[18] This remains true for $n = 1$ whenever $\partial w_{k\alpha} / \partial x_{k\alpha} = 0$, i.e. for the replicator dynamics (or, more generally, if $\operatorname{div} w = 0$).

$q \in X$ will then follow by focusing on the face X' of X which is spanned by the support of q, i.e. $X' = \prod_k \Delta(\mathcal{A}'_k)$ with $\mathcal{A}'_k = \operatorname{supp}(q_k)$; since the dynamics (GD$_n$) preserve the faces of X, the assertion follows by noting that the intersection of X' with an open set in X is open in X' by definition.

Working with the variables z of (3.17) and recalling that the map $G^*: z \mapsto x$ is a local diffeomorphism, Proposition 5.3 shows that open sets of initial states in Ω_w, the phase space of the dynamics (ZD$_n$) with u replaced by w, cannot converge to the interior state $((G^*)^{-1}(q), 0, \ldots, 0)$. Thus, to show that $z(t)$ cannot converge to the interior point $z^* \equiv (G^*)^{-1}(q)$, it suffices to show that $z(t) \to z^*$ would also imply $\lim_{t\to\infty} \dot{z}(t) = \lim_{t\to\infty} \ddot{z}(t) = \cdots = 0$.

For notational simplicity, we will only prove the case $n = 2$. To that end, assume ad absurdum that $z_{k\mu}(t) \to z^*_{k\mu}$ for some $\mu \in \mathcal{A}^*_k$, $k \in \mathcal{N}$, but that $\dot{z}_{k\mu}(t) \nrightarrow 0$. Then, without loss of generality, there exists $\varepsilon > 0$ and an increasing sequence of times $t_n \to \infty$ such that $\dot{z}_{k\mu}(t_n) \geq \varepsilon$ for all n. Thus, let J_n be the largest open interval which contains t_n and which is such that $\dot{z}_{k\mu} > \varepsilon/2$ in J_n. Then, the measure $\delta_n = m(J_n)$ of J_n must vanish as $n \to \infty$; otherwise, and by passing to a subsequence of t_n if necessary, δ_n would always exceed some positive $\delta > 0$, implying that $z_{k\mu}(t)$ grows by at least $\varepsilon\delta/2$ over J_n for all n, a contradiction (recall that $z_{k\mu}(t) \to 0$ so all subsequences of $z_{k\mu}(t)$ are Cauchy). Thus, given that $\dot{z}_{k\mu}(t_n) \geq \varepsilon$ by assumption, the mean value theorem reveals that there exists some $\xi_n \in J_n$ with $\ddot{z}_{k\mu}(\xi_n) \geq \varepsilon/2\delta_n^2$. However, since $z^*_{k\mu}$ must also be a rest point of (ZD$_n$), the dynamics (ZD$_n$) give $\ddot{z}_{k\mu}(t) \to 0$ as $t \to \infty$, a contradiction which proves our claim. □

References

Alvarez, F. (2000). On the minimizing property of a second order dissipative system in Hilbert spaces. *SIAM Journal on Control and Optimization 38*(4), 1102–1119.

Alvarez, F., H. Attouch, J. Bolte, and P. Redont (2002). A second-order gradient-like dissipative dynamical system with Hessian damping. Applications to optimization and mechanics. *Journal des Mathématiques Pures et Appliquées 81*, 774–779.

Attouch, H., X. Goudou, and P. Redont (2000). The heavy ball with friction method, I. The continuous dynamical system: global exploration of the global minima of a real-valued function by asymptotic analysis of a dissipative dynamical system. *Communications in Contemporary Mathematics 2*(1), 1–34.

Björnerstedt, J. and J. W. Weibull (1996). Nash equilibrium and evolution by imitation. In K. J. Arrow, E. Colombatto, M. Perlman, and C. Schmidt (Eds.), *The Rational Foundations of Economic Behavior*, pp. 155–181. New York, NY: St. Martin's Press.

Cantrell, C. D. (2000). *Modern mathematical methods for physicists and engineers*. Cambridge, UK: Cambridge University Press.

Daskalakis, C., P. W. Goldberg, and C. Papadimitriou (2006). The complexity of computing a Nash equilibrium. In *STOC '06: Proceedings of the 38th annual ACM symposium on the Theory of Computing*.

Flåm, S. D. and J. Morgan (2004). Newtonian mechanics and Nash play. *International Game Theory Review 6*(2), 181–194.

Fraenkel, L. E. (1978). Formulæ for higher derivatives of composite functions. *Mathematical Proceedings of the Cambridge Philosophical Society 83*(2), 159–165.

Fudenberg, D. and D. K. Levine (1998). *The Theory of Learning in Games*, Volume 2 of *Economic learning and social evolution*. Cambridge, MA: The MIT Press.

Hart, S. and A. Mas-Colell (2003). Uncoupled dynamics do not lead to Nash equilibrium. *American Economic Review 93*(5), 1830–1836.

Hofbauer, J. (1996). Evolutionary dynamics for bimatrix games: a Hamiltonian system? *Journal of Mathematical Biology 34*, 675–688.

Hofbauer, J. and W. H. Sandholm (2009). Stable games and their dynamics. *Journal of Economic Theory 144*, 1710–1725.

Hofbauer, J. and W. H. Sandholm (2011, September). Survival of dominated strategies under evolutionary dynamics. *Theoretical Economics 6*(3), 341–377.

Hofbauer, J. and K. Sigmund (1988). *The Theory of Evolution and Dynamical Systems*. Cambridge University Press.

Hofbauer, J., S. Sorin, and Y. Viossat (2009, May). Time average replicator and best reply dynamics. *Mathematics of Operations Research 34*(2), 263–269.

Hofbauer, J. and J. W. Weibull (1996). Evolutionary selection against dominated strategies. *Journal of Economic Theory 71*, 558–573.

Lee, J. M. (2003). *Introduction to Smooth Manifolds*. Number 218 in Graduate Texts in Mathematics. New York, NY: Springer-Verlag.

Mertikopoulos, P. and A. L. Moustakas (2010). The emergence of rational behavior in the presence of stochastic perturbations. *The Annals of Applied Probability 20*(4), 1359–1388.

Nagurney, A. and D. Zhang (1997). Projected dynamical systems in the formulation, stability analysis, and computation of fixed demand traffic network equilibria. *Transportation Science 31*, 147–158.

Nash, J. F. (1950). *Non-cooperative games*. Ph. D. thesis, Princeton University.

Ritzberger, K. and K. Vogelsberger (1990). The Nash field. Technical report, Institute for Advanced Studies, Vienna.

Ritzberger, K. and J. W. Weibull (1995). Evolutionary selection in normal-form games. *Econometrica 63*, 1371–99.

Rustichini, A. (1999). Optimal properties of stimulus-response learning models. *Games and Economic Behavior 29*, 230–244.

Sakurai, J. J. (1994). *Modern Quantum Mechanics*. Boston, MA: Addison Wesley.

Samuelson, L. (1993). Does evolution eliminate dominated strategies? In *Frontiers of Game Theory*. Cambridge, MA: MIT Press.

Samuelson, L. and J. Zhang (1992). Evolutionary stability in asymmetric games. *Journal of Economic Theory 57*, 363–391.

Sandholm, W. H. (2011). *Population Games and Evolutionary Dynamics*. Economic learning and social evolution. Cambridge, MA: MIT Press.

Sandholm, W. H., E. Dokumacı, and R. Lahkar (2008). The projection dynamic and the replicator dynamic. *Games and Economic Behavior 64*, 666–683.

Saunders, D. J. (1989). *The geometry of jet bundles*. Number 142 in London Mathematical Society Lecture Note Series. Cambridge, UK: Cambridge University Press.

Smith, M. J. (1984). The stability of a dynamic model of traffic assignment - an application of a method of Lyapunov. *Transportation Science 18*, 245–252.

Sorin, S. (2009). Exponential weight algorithm in continuous time. *Mathematical Programming 116*(1), 513–528.

Taylor, P. D. and L. B. Jonker (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences 40*(1-2), 145–156.

van Damme, E. (1987). *Stability and perfection of Nash equilibria*. Berlin: Springer-Verlag.

Viossat, Y. (2011). Deterministic monotone dynamics and dominated strategies. Submitted.

Weibull, J. W. (1995). *Evolutionary Game Theory*. Cambridge, MA: MIT Press.