Reinforcement Learning in Large Population Models: A Continuity Equation Approach[*]

Ratul Lahkar[†] and Robert M. Seymour[‡]

September 22, 2009

Abstract

We study an evolutionary model in which strategy revision protocols are based on agent specific characteristics rather than wider social characteristics. We assume that agents are primed to play mixed strategies. At any time, the distribution of mixed strategies over agents in a population is described by a probability measure. In each round, a pair of randomly chosen agents play a game, after which they update their mixed strategies using certain reinforcement driven rules based on payoff information. The distribution over mixed strategies thus changes. In a continuous-time limit, this change is described by non-linear continuity equations. We provide a general solution to these equations, which we use to analyze some simple evolutionary scenarios: negative definite symmetric games, doubly symmetric games, generic 2×2 symmetric games, and 2×2 asymmetric games. A key finding is that, when agents carry mixed strategies, distributional considerations cannot be subsumed under a classical approach such as the deterministic replicator dynamics.

Keywords: Reinforcement Learning; Continuity Equation; Replicator Dynamics.

1 Introduction

It has long been recognized that the strong rationality assumptions on which modern economic theory is founded are unrealistic as a representation of actual human behavior (Simon, 1956, 1957), and the increasingly urgent challenge to rework these foundations has prompted much research towards shifting the emphasis from substantive to procedural rationality; i.e to the way people actually make economic decisions. This research has in common the downgrading of hyper-rationality, with its associated superhuman cognitive demands, in exchange for the upgrading of certain kinds of psychological complexities that are claimed to better characterize the behavior of real people in decision making contexts. One strand of this response is the 'behavioral economics' movement - the attempt to account for empirically observed deviations from standard rationality assumptions by building into utility functions various inferred psychological dispositions (Camerer, 2003). Another strand is 'bounded rationality', both in the original 'satisficing' sense of Herbert Simon (1956, 1957, 1983)[1], and in the more recent 'adaptive toolbox' perspective on human decision making, which rejects optimization in favor of the 'rules of thumb' said to be part of the human species-specific evolutionary heritage (Gigerenzer and Selten, 2001).[2] Within bounded rationality, there has emerged a substantial literature on learning in a game theoretic context.[3] This is the attempt to explain convergence to equilibrium by boundedly rational players through the use of simple behavior-modifying mechanisms such as reinforcement learning (Borgers and Sarin, 1997, 2000; Erev and Roth, 1998), regret matching (Hart and Mas-Colell, 2000) and stochastic fictitious play (Fudenberg and Levine, 1995).

A related response to the rationality challenge emphasizes learning in a social context through repeated experience, but by boundedly rational agents using simple procedures of trial-and-error or imitation through which players learn that some behaviors perform better than others. This is the approach of 'evolutionary game theory', which studies the processes through which populations of interacting agents find their way (or not) to an equilibrium.[4] The assumption that agents do not have hyper-rational cognitive abilities is reflected in the simple strategy adjustment procedures that are designed to reward better performing behavior; analogously to the way a competitive market rewards behavior leading to higher profits (Samuelson, 1997, Chapter 1). Thus, most recent evolutionary models assume that each agent is primed to play, unreflectively, a pure strategy, which is retained until a revision opportunity becomes available. Any cognitive sophistication is subsumed under the revision protocol that agents use.

Evolutionary game theory potentially has great application in studying the impact of individual level boundedly rational behavior on aggregate social outcomes. However, there has also been a sharp dichotomy between evolutionary game theory and the larger game theoretic literature on

---

[1]See also Rubinstein (1998) for a more 'rationalistic' approach to bounded rationality.

[2]This approach has been dubbed 'ecological rationality', through its emphasis on fast and frugal heuristics that are matched to specific environmental characteristics (Gigerenzer et al., 1999).

[3]See Young (2005) for a review of this literature.

[4]See Weibull (1995), Hofbauer and Sigmund (1988, 1998), Samuelson (1997), and Sandholm (2009) for book level studies of evolutionary game theory.

learning through boundedly rational rules that has prevented the realization of this potential. Our primary aim in this paper is to construct a theoretical framework that can resolve this dichotomy and allow the study of boundedly rational learning in the context of large population models. The source of this dichotomy lies in the nature of the rules that individual agents use to revise their actions in the two contexts. In the learning literature, a finite set of individuals repeatedly interact in a game by conditioning their current strategic behaviour on the experience they have accumulated in previous rounds of the game. The strategy revision rules in this context are therefore history dependent, allowing the players to 'learn' from their previous experience. For example, in reinforcement learning, actions which have been more successful in the past are employed with higher probability in the present. In contrast, the revision protocols in evolutionary game theory are generally conditioned on the current social state, which is the proportions of agents in a population (or populations) playing different pure actions. For example, in imitative revision protocols (Börnerstedt and Weibull, 1996; Schlag, 1998) that generate the replicator dynamic (Taylor and Jonker, 1978), an agent may randomly select and imitate the strategy of a rival member of the population.[5] Revision protocols of this form, which emphasize social imitation, therefore fail to take into account a very significant part of the cognitive abilities of human agents—the ability to learn from *personal* experience (albeit in a social context). There is however considerable evidence that such learning does take place and that past experience has a significant bearing on future course of action; for example, Erev and Roth (1998) document empirical evidence in support of reinforcement learning. In this paper, we resolve this dichotomy by developing a general theoretical framework that allows agents to apply history dependent learning procedures in the setting of large population games and analyze one particular such procedure, reinforcement learning, in detail.

The conditioning of revision protocols in evolutionary game theory on the social state is a legacy of the origins in biology. The evolutionary process in biology is modeled as an automatic process, driven by births and deaths, working to increase the frequency of better performing strategies. While the motivation of an evolutionary process in economics is the same, it requires to be micro-founded on decision behavior by agents, given the higher (relative to animals) cognitive abilities of humans. This has led to a variety of revision protocols conditioned on the current social state (see footnote 5). However, it is unlikely that in a decentralized environment agents would be privy to such detailed information about the social state, thereby rendering the implementation of such revision protocols essentially unfeasible. Some characterizations of revision protocols that generate the classical replicator dynamic can, of course, be much more parsimonious in their level of information requirement. For example, agents are required to observe only their immediate

---

[5]Other examples of the difference between revision procedures in the two fields can be cited. In learning under regret matching, actions which would have reduced aggregate past regret become more probable. Thus, if r_i is the aggregate average regret of not always having played strategy i, then under the Hart and Mas-Colell (2000) regret matching rule, the current probability of playing i is proportional to $\max\{0, r_i\}$. In contrast, the revision protocol to generate the 'Brown–Nash–von Neumann' (BNN) dynamic (Sandholm, 2005) assigns probabilities proportional to the excess of the payoffs over the mean population payoff, with payoffs calculated as expectations over the current social state. Note that the functional forms of the two revision procedures are similar but the state variable each is conditioned on differ. Similarly, in evolution under perturbed best response agents play a near best response to the current social state (Hofbauer and Sandholm, 2005).

payoff against a randomly matched opponent, rather than the expected payoff against the current social state[6], as long as they can also imitate the current (pure) strategy of some other randomly selected agent. However, in some contexts, this assumption of the observability of a potential rival's strategy may be a very strong informational assumption; for example, traders in a stock market would normally take care to conceal their planned strategies from rivals. Nevertheless, in most scenarios in evolutionary game theory it remains the case that, for behaviors more sophisticated than random imitation, the informational requirements are significantly more onerous, thus severely compromising their feasibility.

In contrast, the evolutionary framework that we construct is free from such onerous informational requirements, being based on strategy revision protocols that require only knowledge of *agent-specific characteristics* rather than of wider social characteristics. This allows a greater range of behaviors – whether resulting from conscious deliberation or from essentially subconsciously processes – to be feasibly implemented, thereby doing justice to a greater range of the cognitive abilities of human agents. In particular, we use revision protocols based on *reinforcement learning*, since these are both simple and extremely parsimonious in the information required for their implementation.

We focus on reinforcement learning since it is the simplest and most widely studied class of learning rules, and also has significant empirical support (Erev and Roth, 1998). However, our general theoretical framework can be adapted to other learning mechanisms like regret matching or stochastic fictitious play.[7] To allow the application of procedures like reinforcement learning, we postulate that each agent in a population game is primed with a mixed strategy, which we interpret as the agent's *behavioral disposition*, expressing her intrinsic uncertainty about what pure action to take in the game when called upon to play. In this approach, the population state is specified by a probability measure over the set of mixed strategies, defining the 'mass' of agents using a particular mixed strategy. We note that this introduces a radical form of agent *heterogeneity* into the population, extending the classical setting in which all agents use a fixed mixed strategy or, equivalently, a fixed mixture of pure strategies. As players revise their individual mixed strategies in the light of experience using some learning protocol such as reinforcement learning, the population state changes. We track the changes in population states using a generalization to a probability measure setting of a first-order partial differential equation system akin to the *continuity equations* used in physics in the study of conserved quantities, such as bulk fluids.[8] These evolutionary

---

[6]If agents are being randomly matched to play a symmetric normal form game with payoff matrix U, they are required only to observe the realized payoffs u_{ij} rather than the expected payoffs of the form $\sum_j u_{ij}x_j$, where x_j is the proportion of agents playing pure action j.

[7]Among the well known learning rules, reinforcement learning and regret matching require an agent to remember only the actions he has played in the past. Stochastic fictitious play requires agents to remember the actions of the opponents they encountered in the earlier rounds of the game. The application of these rules in large population models would therefore require agents to know only their own actions or the actions of those opponents they personally encountered in the past. It is of course not necessary that in applying a rule such as reinforcement, agents are consciously increasing the probability on some action. It may be equally true that they are responding to some subconscious cues that incline them towards adopting some actions with greater likelihood. We discuss this point in greater detail in Section 12. Our technical framework is general enough to accommodate both interpretations.

[8]In physics, the continuity equation is a linear partial differential equation that describes the rate of change in the

3

dynamics can also track the change in the mean of the population state which is the proportion of agents playing different pure actions thereby permitting comparison with the results from the classical approach to evolutionary game theory. We use this approach to study the impact of heterogeneity on the long run population state in Sections 10 and 11.

To illustrate the application of these evolutionary ideas in an economic context, consider a simple pricing game in which a population of homogeneous sellers interact with a population of homogeneous buyers. The sellers can quote one of three prices; p_0, p_1, p_2 with $0 < p_0 < p_1 < p_2$. A consumer randomly matches two suppliers and asks them to submit quotations. Consumers are of two types: a proportion y_1 simply pick one of the two quotations at random and pays whatever price it asks to the supplier, and the remaining proportion y_2 compares both quotations and buys at the lower price, breaking ties with a random choice. This situation defines a 2-player, 3×3 symmetric game between suppliers.[9] Under certain conditions, it is easy to show that the game has a unique mixed strategy equilibrium. A conventional equilibrium-focused approach would therefore predict price dispersion as an equilibrium in this model.

However, evolutionary models studied in Hopkins and Seymour (2002) (using the replicator dynamic) and Lahkar (2009) (using the logit dynamic – Fudenberg and Levine, 1995) show that under standard evolutionary adjustment procedures, the mixed equilibrium is unstable; a conclusion that is validated by evidence both from the field (Lach, 2002) and the laboratory Cason et. al. (2005). These evolutionary models are based on sellers conditioning their behaviour on the current social state. We have already expressed our reservations about how feasible it is to implement these protocols (see footnote 9); particularly one as onerous as the logit best response where sellers would actually need to observe the social state in order to update their actions. In contrast, to apply reinforcement learning in our framework, sellers are construed as (heterogeneous) carriers of mixed strategies over possible prices, and simply need to update their probability of charging a particular price on the basis of whether or not the presently quoted price results in a sale. Note that in this strategy revision process, a seller need not even observe the price charged by the opponent with whom he is currently matched. This is certainly a far less onerous requirement than that imposed by even the simplest imitation revision protocol, where the seller would need to observe the price that a randomly selected rival seller would charge were that seller invited to submit a quotation.

We derive our evolution equations for two models of population games, asymmetric and sym-

---

mass of fluid in any part of the medium through which the fluid is flowing. See, for example, Margenau and Murphy (1962). However, our continuity equations differ from classical versions encountered in physics in that they contain non-linearities. See Ramsza and Seymour (2009) for an application of continuity equation techniques to track the evolution of fictitious play updating weights in a population game. Our paper provides a more general method of constructing continuity equations that can be used for a variety of learning algorithms.

[9]This is a simplified case of the finite dimensional Burdett and Judd (1983) model studied in Lahkar (2009). The payoff matrix of the game is as follows. Matched with a seller charging price p_j, the seller charging p_i obtains expected payoff

$$\begin{array}{ll} \frac{1}{2}p_i y_1, & \text{if } p_i > p_j, \\ \frac{1}{2}p_i(y_1 + y_2) = \frac{1}{2}p_i, & \text{if } p_i = p_j, \\ \frac{1}{2}p_i(y_1 + 2y_2), & \text{if } p_i < p_j. \end{array}$$

Let x_i be the proportion of sellers charging price p_i. The expected payoff from charging p_i is then $\pi_i = \frac{1}{2}p_i\left(y_1 + 2y_2\left(\frac{x_i}{2} + \sum_{j>i} x_j\right)\right)$.

4

metric. In the asymmetric case, there are two populations of agents with one player from each population randomly matched in each round to play a two-player asymmetric normal form game. If we let the time interval between each round of play go to zero, we obtain a coupled pair of non-linear continuity equations, one for each population. In symmetric games, agents are matched with members of their own population leading to a single continuity equation tracking the change in the probability measure of that population.

From the general continuity equation, we generate one particular form—the *replicator continuity equation*—using two reinforcement learning based updating rules, one from Börgers and Sarin (1997) and a second that we propose as an alternative. When applied in the context of learning, the expected change in mixed strategy of an agent under either of these rules is given by the classical replicator dynamic. Such a relationship between reinforcement rules and the replicator dynamic has been established in Börgers and Sarin (1997) and Hopkins (2002). However, by applying learning algorithms explicitly to large population models, our work provides a more general perspective on the link between learning and evolution.

We solve the replicator continuity equation using standard methods based on Liouville's formula.[10] To characterize solutions explicitly requires us to derive an associated ODE system whose solutions describe trajectories of certain aggregate quantities associated to the population means. We call this ODE system the *distributional replicator dynamics.* We show that the continuity replicator dynamics has many stationary solutions, in particular any probability distribution over mixed strategies having mean that is a Nash equilibrium. Thus, equilibrium populations can be very heterogeneous, with individuals playing any mixed strategy with positive probability, but with population mean a Nash equilibrium. However, in contrast to the classical case, the population mean itself is construed as a mixed strategy, not as a mixture of pure strategies. We then analyze some simple evolutionary scenarios: negative definite symmetric games, doubly symmetric games, generic 2×2 symmetric games, and 2×2 asymmetric games, focusing on the convergence properties of the mean of the population state since this is the observable *aggregate social state.* For negative definite symmetric games (for example, games with an interior evolutionary stable state, or ESS) and doubly symmetric games (for example, coordination games), we show global convergence of the mean of the population state to a Nash equilibrium, although, importantly, this need not imply that the asymptotic population state itself is a point mass on the globally attracting mixed Nash equilibrium. In 2×2 symmetric games, we find that the aggregate social state is the same pure strategy Nash equilibrium that would result under the classical replicator dynamic, provided the initial point in the latter case is identical to the initial aggregate state in the former case. However, this conclusion does not hold either for $n \times n$ symmetric games with $n > 2$, or for 2×2 asymmetric games, where we provide counter-examples in which the two dynamics converge to different pure

---

[10]This formula expresses the time evolution of the probability density function as a function of the initial probability density and the deterministic trajectories of the underlying characteristic ODE system, which describes the motion of individual agents in the population– see, section 4 below. The classical Liouville formula describes the change in volume along flow lines of an underlying dynamical system see, for example, Hartman (1964). Related versions are discussed in Weibull (1995) and Hofbauer and Sigmund (1998).

equilibria. Hence, expanding the behavioural flexibility of agents to allow use of mixed strategies in evolutionary contexts has real consequences, in that it can lead to radically different conclusions about the equilibrium social state.

In some ways, our paper revives some of the early work in biology (Hines, 1980; Zeeman, 1981) concerning the evolution of mixed strategies. These papers use straightforward adaptations of different versions of the replicator dynamic (see footnote 24) to study the evolution of mixed strategies using the standard biological motivation that the growth of the population share using a strategy is proportional to the advantage of that strategy over the mean strategy. The analogous motivation in the context of human interaction and learning, that agents imitate successful strategies is, however, not applicable to mixed strategies since such strategies are not directly observable. We therefore cannot directly apply the functional form of the replicator dynamic to mixed strategies. Hence the necessity to use the continuity equation approach for this purpose. Nevertheless, our result on the convergence of the mean of the population to the Nash equilibria in negative definite games is comparable to the results of Hines (1980) and Zeeman (1981) on convergence of the mean to an ESS. This paper is also related, though in a more peripheral manner, to the literature on evolutionary dynamics in games with continuous pure action space.[11] However, while these papers focus on the evolution of probability measures over pure actions, our concern is the evolution of probability measures over mixed strategies in finite games.

The remainder of this paper is organized as follows. In section 2, we derive the general, nonlinear continuity equations for 2 player asymmetric and symmetric games. Section 3 presents two updating rules that generate continuity equations based on the classical replicator dynamic, and we consider equilibrium conditions for distributions under these dynamics. In section 4, we introduce Liouville's formula in a general context, and use this formula to solve a generalized form of the replicator continuity equation in section 5. In section 6, we introduce the distributional replicator dynamics, a system of autonomous ODEs whose solutions determine the distributional solutions of the replicator continuity equations. Section 7 shows stability in the mean for negative definite and semi-definite games, and section 8 proves that expected payoff increases along non-equilibrium trajectories for doubly-symmetric games. Sections 9 and 11 contain analyses of generic 2×2 symmetric and asymmetric games respectively, while section 10 presents a 3×3 symmetric game for which classical and distributional trajectories convergence to different equilibria. Section 12 contains a discussion of the paper and concludes. Certain proofs and additional technical material are presented in the appendix.

2 The General Continuity Equation for Population Games

We derive the continuity equations in the setting of population games. First, we consider the case in which two players, each chosen from a separate population, are randomly matched to play an

---

[11]Some papers in this field are Cressman (2005), Cressman and Hofbauer (2005), Cressman, Hofbauer and Riedel (2006), Hofbauer, Oechssler and Riedel (2008) and Oechssler and Riedel (2001, 2002).

asymmetric normal form game.[12] Next, we look separately at the case where two players chosen from the same population are randomly matched to play a symmetric normal form game.

2.1 Two-population Asymmetric Games

Consider a society consisting of the set of populations $\mathcal{P} = \{1, 2\}$. We assume both populations are of fixed probability mass 1. Let S_l be the strategy set and n_l be the number of strategies of population $l \in \mathcal{P}$. We denote by Δ_l the simplex corresponding to population l. Thus,

$$\Delta_l = \left\{ x \in I\!R_+^{n_l} : \sum_{i \in S_l} x_i = 1 \right\}. \tag{1}$$

A mixed strategy used by a player in population l belongs to Δ_l. We will use x and y to denote a typical mixed strategy of a player in populations 1 and 2 respectively. Then $\Delta = \Delta_1 \times \Delta_2$ is the set of mixed strategies of pairs of players, one from each population.

Let P be a Borel probability measure defined on the space of mixed strategies for population 1, Δ_1.[13] Thus, if $B \subseteq \Delta_1$ is a Borel set, then $P(B)$ can be regarded as the proportion of agents in population 1 playing mixed strategies in B. The mean strategy in population 1, $\langle x \mid P \rangle \in \Delta_1$, is given by

$$\langle x \mid P \rangle = \int_{\Delta_1} x P\,(dx). \tag{2}$$

More generally, given a real vector-valued continuous function $\phi(x)$ on Δ_1, we define its expectation with respect to P by

$$\langle \phi \mid P \rangle = \int_{\Delta_1} \phi(x) P(dx). \tag{3}$$

Similarly, we denote by Q the Borel probability measure over the space of mixed strategies for population 2 characterizing proportions of agents using these strategies. The mean strategy $\langle y \mid Q \rangle \in \Delta_2$ for population 2 is defined analogously to (2), and the expectation, $\langle \psi \mid Q \rangle$, of a continuous function $\psi(y)$ on Δ_2, is defined analogously to (3).

We interpret the means $(\langle x \mid P \rangle, \langle y \mid Q \rangle)$ as the *aggregate social state* generated by the measures P and Q. Even though agents are actually playing mixed strategies, the observable aggregate social state is the proportion of agents playing different pure actions. By the law of large numbers, this distribution over pure actions is equal to $(\langle x \mid P \rangle, \langle y \mid Q \rangle)$. We make use of this concept of the aggregate social state later in our applications of the continuity equation in 2×2 games.

Two players, one from each population, are randomly matched to play an asymmetric normal form game. We denote by P_t and Q_t the measures characterizing the two population states at time $t \geq 0$. Our objective is to track the evolution of the two measures P_t and Q_t over time. We derive

---

[12]We confine ourselves to two-player asymmetric games merely for notational convenience. All the ideas involved can be extended easily to multipopulation asymmetric games at the cost of more cumbersome notation.

[13]That is, P is a non-negative measure of total mass 1, defined on the σ-field of Borel sets in Δ_1, the smallest σ-field containing the closed sets of Δ_1 – see, for example, Dunford and Schwatz (1964), p 137.

the continuity equations for this purpose as follows.

Suppose the two chosen players use the mixed strategy profile $(x, y) \in \Delta$. The probability that they play the action profile $(i, j) \in S = S_1 \times S_2$ is given by

$$\pi_{ij}(x, y) = x_i y_j. \tag{4}$$

Of course, $\sum_{i,j} \pi_{ij}(x, y) = 1$ for all (x, y). After a play of the game, a player updates his mixed strategy according to some updating protocol of the following general form. Given that the action profile (i, j) has been played, the row player updates his strategy $x \in \Delta_1$ to x' given by an updating rule of the form:

$$x' = x + \tau f_{ij}(x), \tag{5}$$

where τ is a small time parameter representing the length of a round in which the game is played. Similarly, the column player updates her strategy $y \in \Delta_2$ to y' given by an updating rule of the form:

$$y' = y + \tau g_{ij}(y). \tag{6}$$

Thus, f_{ij} and g_{ij} are functions, $f_{ij} : \Delta_1 \to I\!\!R_0^{n_1}$ and $g_{ij} : \Delta_2 \to I\!\!R_0^{n_2}$, where $I\!\!R_0^n = \{z \in I\!\!R^n : \sum_i z_i = 0\}$. We call these the *forward state change* functions: they specify how the players' states change going forward in time, and therefore are rules to update the mixed strategies x and y respectively.

The associated *backward state change* functions specify where current states came from, going backward in time. Thus the backward state changes are functions $b_{ij} : \Delta_1 \to I\!\!R_0^{n_1}$ and $c_{ij} : \Delta_2 \to I\!\!R_0^{n_2}$ which satisfy:

$$(x, y) = (u + \tau f_{ij}(u), v + \tau g_{ij}(v)) \iff (u, v) = (x - \tau b_{ij}(x), y - \tau c_{ij}(y)). \tag{7}$$

Between times t and $t + \tau$, the two probability measures make the transition from P_t and Q_t to $P_{t+\tau}$ and $Q_{t+\tau}$ respectively. Using (4), the relationships between the probability measures at the two time periods are given by

$$P_{t+\tau}(dx) = \sum_{i,j \in S} \int_{y \in \Delta_2} (x - \tau b_{ij}(x))_i \, y_j P_t \left(d[x - \tau b_{ij}(x)]\right) Q_t(dy), \tag{8}$$

$$Q_{t+\tau}(dy) = \sum_{i,j \in S} \int_{x \in \Delta_1} x_i \, (y - \tau c_{ij}(y))_j \, Q_t \left(d[y - \tau c_{ij}(y)]\right) P_t(dx). \tag{9}$$

In order to derive the continuity equations, we multiply (8) and (9) by smooth, real-valued, but otherwise arbitrary 'test functions' $\phi(x)$ and $\psi(y)$ respectively, and then integrate. We obtain

$$\langle \phi \mid P_{t+\tau} \rangle = \sum_{i,j \in S} \int_{x \in \Delta_1} \phi(x) \, (x - \tau b_{ij}(x))_i \, P_t \left(d[x - \tau b_{ij}(x)]\right) \langle y_j \mid Q_t \rangle \tag{10}$$

$$\langle \psi \mid Q_{t+\tau} \rangle = \sum_{i,j \in S} \int_{y \in \Delta_2} \psi(y) \, (y - \tau c_{ij}(y))_j \, Q_t \left(d[y - \tau c_{ij}(y)]\right) \langle x_i \mid P_t \rangle. \tag{11}$$

8

We now focus on (10) to obtain the continuity equation for population 1. We show in Appendix A.1 that the continuous-time limit $\tau \to 0$ leads to the following *weak form of the continuity equation*[14]:

$$\frac{d}{dt}\langle \phi \mid P_t \rangle = \int_{x \in \Delta_1} \nabla\phi(x) \cdot [\mathcal{F}(x)\langle y \mid Q_t \rangle]\, P_t(dx), \tag{12}$$

where $\mathcal{F}(x)$ is the $n_1 \times n_2$ matrix whose (i,j)-th entry is:

$$\mathcal{F}_{ij}(x) = \sum_{r \in S_1} x_r f_{rj,i}(x). \tag{13}$$

A similar derivation gives the weak form of the continuity equation for Q_t analogous to (12):

$$\frac{d}{dt}\langle \psi \mid Q_t \rangle = \int_{y \in \Delta_2} \nabla\psi(y) \cdot [\mathcal{G}(y)\langle x \mid P_t \rangle]\, Q_t(dy), \tag{14}$$

where $\mathcal{G}(y)$ is the $n_2 \times n_1$ matrix defined analogously to (13):

$$\mathcal{G}_{ji}(y) = \sum_{s \in S_2} y_s g_{is,j}(y). \tag{15}$$

The weak continuity equations (12) and (14) provide the dynamical equations that describe the evolution of the probability measures P_t and Q_t from specified initial measures P_0 and Q_0. Note that, given Q_t, the form (12) is linear in P_t, and given P_t, the form (14) is linear in Q_t. However, taken together, this pair of equations is a coupled non-linear system.

2.1.1   Absolute Continuity and Strong Forms of the Continuity Equations

If a probability measure P is absolutely continuous with respect to Lebesgue measure, then we can write $P(dx) = p(x)dx$ for some L_1 probability density function $p(x)$. We show in Appendix A.2 that if P_t is a solution to (12) with absolutely continuous initial measure P_0, then P_t is absolutely continuous for all $t \geq 0$. Thus, if Q_0 is also absolutely continuous, then the weak continuity equations (12) and (14) can be expressed in terms of probability densities $p_t(x)$ and $q_t(y)$.

If we assume in addition that the density function $p_t(x)$ is differentiable in both x and t, then we may obtain the *strong form of the continuity equation*. Thus, taking $\phi(x) = 0$ for $x \in \partial\Delta_1$,

---

[14]For a differentiable scalar function $\phi : \mathbb{R}^n \to \mathbb{R}$, the *gradient* of ϕ, $\nabla\phi$, is the vector field on $\mathbb{R}^n$ defined by $\nabla\phi(x) = (\partial\phi(x)/\partial x_1, \dots, \partial\phi(x)/\partial x_n)$. For later use, we also note that the *divergence* of a vector field $f : \mathbb{R}^n \to \mathbb{R}^n$ is the scalar function on $\mathbb{R}^n$ defined by $\nabla \cdot f(x) = \sum_{i=1}^n \partial f_i(x)/\partial x_i$. See, for example, Margenau and Murphy (1962) for a general discussion of these operators.

using (3) and integrating by parts on the right-hand side in (12), we obtain[15]:

$$\int_{\Delta_1} \phi(x) \frac{\partial p_t(x)}{\partial t} dx = - \int_{\Delta_1} \phi(x) \nabla \cdot [\mathcal{F}(x)\langle y \mid q_t \rangle p_t(x)] dx$$

That is:

$$\int_{\Delta_1} \phi(x) \left\{ \frac{\partial p_t(x)}{\partial t} + \nabla \cdot [\mathcal{F}(x)\langle y \mid q_t \rangle p_t(x)] \right\} dx = 0.$$

Since this holds for all differentiable test functions $\phi(x)$ which vanish on $\partial \Delta_1$, we obtain the differential form of the continuity equation:

$$\frac{\partial p_t(x)}{\partial t} + \nabla \cdot [\mathcal{F}(x)\langle y \mid q_t \rangle p_t(x)] = 0, \qquad x \in \text{int } \Delta_1, \, t > 0. \qquad (16)$$

This is the strong form of the continuity equation, which applies to differentiable density functions.

When $q(y, t)$ is also differentiable in y and t, we can obtain the strong form of the continuity equation analogous to (16):

$$\frac{\partial q_t(y)}{\partial t} + \nabla \cdot [\mathcal{G}(y)\langle x \mid p_t \rangle q_t(y)] = 0, \qquad y \in \text{int } \Delta_2, \, t > 0. \qquad (17)$$

The strong continuity equations for smooth densities, (16) and (17), give the dynamical equations that describe the evolution of the probability densities p_t and q_t. Intuitively, $\mathcal{F}(x)\langle y \mid q_t \rangle$ represents the adaptation 'velocity' of mixed strategy x.[16] That is, $[\mathcal{F}(x)\langle y \mid q_t \rangle]\tau$ is the expected change in mixed strategy x in the small time interval τ in response to a play of the game. Since the mass of x is represented by $p_t(x)$, $[\mathcal{F}(x)\langle y \mid q_t \rangle] p_t(x)$ gives the probability mass flow at x. The divergence of this vector field therefore gives the rate at which the probability mass in a small neighbourhood of x is expanding or contracting. Since $\frac{\partial p_t(x)}{\partial t}$ is precisely the rate of change of the probability mass of x, we are led to the continuity equation (16).

2.2 Symmetric Games

We now consider a symmetric normal form game with players chosen from a single population.[17] We denote by $S = \{1, 2, \cdots, n\}$ the set of actions in the game. The set of mixed strategies is the

---

[15]The formal argument has the following form. For X a vector field on the domain Δ, we use the identity $\nabla \cdot [\phi X] = \phi \nabla \cdot X + \nabla \phi \cdot X$ to obtain

$$\int_\Delta \nabla \phi \cdot X dV = \int_\Delta \nabla \cdot [\phi X] dV - \int_\Delta \phi \nabla \cdot X dV$$

Now use the divergence theorem (Margenau and Murphy, 1962) together with the assumption that $\phi = 0$ on $\partial \Delta$ to obtain:

$$\int_\Delta \nabla \cdot [\phi X] dV = \int_{\partial \Delta} (u \cdot X) \phi dA = 0.$$

[16]In the next section, we provide two strategy updating rules in which this velocity is given by the replicator dynamic.

[17]To describe the symmetric case, we adopt the notation of population 1 of the asymmetric case but drop the population subscript.

n-simplex $\Delta = \{x \in I\!\!R^n_+ : \sum_{i \in S} x_i = 1\}$. The microfoundations and the derivation of the continuity equation for a one population symmetric game now proceeds analogously to the asymmetric case. The only difference is that instead of two probability measures, we need only track the evolution of a single measure. Players from the same population are matched in pairs at each time interval to play the game. The event that player 1 uses mixed strategy x and player 2 uses mixed strategy y occurs with probability $P_t(dx)P_t(dy)$, where P_t is the probability measure for the population at time t. Thus, given that the chosen players use strategies $x, y \in \Delta$, the probability, $\pi_{ij}(x, y)$, that they play the pair of pure strategies $i, j \in S$ is given by (4). As in the asymmetric case, players update their mixed strategies using a rule of the form (5).

The updating equation corresponding to (8) is:

$$P_{t+\tau}(dx) = \sum_{i,j \in S} \int_{y \in \Delta} (x - \tau b_{ij}(x))_i \, y_j P_t \left(d\left[x - \tau b_{ij}(x) \right] \right) P_t(dy). \tag{18}$$

The difference between (8) and (18) is that the $Q_t(dy)$ term in (8) is replaced by $P_t(dy)$. The derivation now proceeds as in the asymmetric case to obtain the weak symmetric form corresponding to (12):

$$\frac{d}{dt} \langle \phi \mid P_t \rangle = \int_{x \in \Delta} \nabla \phi(x) \cdot [\mathcal{F}(x) \langle x \mid P_t \rangle] \, P_t(dx), \tag{19}$$

with $\phi(x)$ a smooth test function, and $\mathcal{F}(x)$ given by (13) with $n_2 = n_1$. From this, when P_t is absolutely continuous with differentiable density function $p_t(x)$ we derive the symmetric analogue of the strong form (16) :

$$\frac{\partial p_t(x)}{\partial t} + \nabla \cdot [\mathcal{F}(x) \langle x \mid p_t \rangle p_t(x)] = 0, \qquad x \in \text{int}\, \Delta, \, t > 0. \tag{20}$$

Note that the form (20) is non-linear in p_t.

3 Replicator Continuity Equations

Equations (16) and (17) give the general form of the continuity equations for 2-population, asymmetric games for differentiable density functions. In this section, we derive a particular form of the continuity equations—the *replicator continuity equations*. We first introduce two alternative forward state change rules $f_{ij}(x)$ and $g_{ij}(y)$. Both these rules are based on the idea of reinforcement. We then show that these updating rules lead to the replicator continuity equations. These updating rules therefore provide the microfoundations to the replicator continuity dynamic.

Reinforcement models have been widely studied in the learning literature. A group of players, one in each role in the game, employ mixed strategies in each round of a game. Reinforcement models are based on the idea that if the action currently employed obtains a high payoff, then the probability assigned to it increases in the next round of play. Reinforcement models are therefore extremely naive models of learning. Agents mechanically respond to stimuli from their environment

without seeking to create any model of the situation or strategically evaluate how they are doing. Hence, they do not seek to exploit the pattern of opponents' past play and predict the future behaviour of their opponents.[18] In this sense, agents are boundedly rational.

The two forward state change rules we consider are described below. We assume that players from population 1 play the role of the row player while those from population 2 are column players. If the row player plays action $i \in S_1$ and the column player plays action $j \in S_2$, the payoff to the row player is u_{ij} and to the column player is v_{ji}. The expected payoff to $i \in S_1$ against mixed strategy $y \in \Delta_2$ is $(Uy)_i$, where U is the $n_1 \times n_2$ payoff matrix (u_{ij}). Similarly, the payoff to $j \in S_2$ against mixed strategy $x \in \Delta_1$ is $(Vx)_j$, where V is the $n_2 \times n_1$ payoff matrix (v_{ji}). Thus

$$(Uy)_i = \sum_{j \in S_2} u_{ij} y_j, \tag{21}$$

$$(Vx)_j = \sum_{i \in S_1} v_{ji} x_i. \tag{22}$$

We consider a player in a 2-player game who employs strategy $x \in \Delta$, uses action i and encounters an opponent who uses action j in the current round. The player then updates her strategy to x' according to an updating rule $f_{ij}(x)$, as in (5). For brevity, we present only the rules for population 1. For population 2, the updated strategy y' and the updating vector $g_{ij}(y)$ take analogous forms, as in (6).

In enumerating the two rules, we need to assume that all payoffs are positive for Rule 1 and negative for Rule 2 in order to ensure that all probabilities x'_r are less than 1.[19] Since it is always possible to rescale payoffs to make them all positive or negative without affecting incentives, we do not consider this a severe restriction.

1. This rule is from Börgers and Sarin (1997) and is a special case of a general class of reinforcement rules introduced in Börgers and Sarin (2000). Under this rule, the mixed strategy x' and the forward state change vector take the form

$$x'_r = \delta_{ir} u_{ij} \tau + (1 - u_{ij}\tau)x_r, \tag{23}$$

$$f_{ij,r}(x) = (\delta_{ir} - x_r)u_{ij}. \tag{24}$$

For τ small enough, a sufficient condition for (23) to represent an updating rule is $u_{ij} > 0$, for all $i, j \in S$.

The general class of rules in Börgers and Sarin (2000) is based on the idea of *aspiration*. To explain this rule, let us momentarily set $\tau = 1$. Suppose that at round t of play, a player aspires

---

[18]Börgers and Sarin (1997) provide some justification of why agents respond to very limited information in these models–only their own payoffs. They argue that the acquisition or processing of new information may be too costly relative to benefits. Hence, they say, reinforcement models may be more plausible if agents' behaviour is habitual rather than the result of careful reflection.

[19]For large τ, we would also need to assume that the payoffs are less than 1 (more than -1) for Rule 1 (Rule 2) to ensure that the probabilities are positive. Since we are primarily concerned with the case where τ is arbitrarily small, we dispense with this restriction.

to a payoff of a_t. The probability of playing a strategy $r \neq i$ is then $x'_r = x_r + (a_t - u_{ij})x_r$[20]. Hence, if $u_{ij} > a_t$, then action i gets reinforced. By setting a_t identically equal to zero, we obtain (23). Note that in this case, the current action i is always reinforced.

2. We now consider a revision rule which applies when all payoffs u_{ij} are negative. The updated strategy and the state change rule we consider is

$$
\begin{aligned}
x'_r &= x_r + \tau u_{rj} x_r, \quad \text{for } r \neq i, & (25) \\
f_{ij,r}(x) &= u_{rj} x_r, \quad\quad\quad \text{for } r \neq i, & (26)
\end{aligned}
$$

with the residual probability being alloted to i. For small τ, it is sufficient to assume that u_{rj} is negative for x' to be a probability distribution.

Revision rule (25) has a similar interpretation to (23). We interpret the negative payoffs as costs that the consumer incurs. Suppose a_t is the maximum (non-negative) cost that the consumer is willing to incur in period t. The probability of playing $r \neq i$ in the next round is given by[21] $x'_r = x_r + (u_{rj} - a_t)x_r$. Action i is therefore reinforced if $u_{rj} < a_t$, for all $r \neq i$. By setting a_t identically equal to zero, we obtain (25) and ensure that the current action i is always reinforced when all payoffs are negative.

Recalling the notation of (21) and (22), we introduce the following operators

$$
R^1_i(x)\, y = x_i \left\{ (Uy)_i - x \cdot Uy \right\}, \tag{27}
$$

$$
R^2_j(y)\, x = y_j \left\{ (Vx)_j - y \cdot Vx \right\}. \tag{28}
$$

Clearly, the vector field generated by the two operators in (27) and (28) is identical to the vector field generated by the bimatrix replicator dynamic (Taylor, 1979) on $\Delta = \Delta_1 \times \Delta_2$. Hence, we call the $n_1 \times n_2$ matrix operator $R^1(x)$ and the $n_2 \times n_1$ matrix operator $R^2(y)$, the *replicator operators* for the two populations.

We now establish that the two updating rules described above generate the replicator operators for the two populations which arise as $\mathcal{F}(x)$ and $\mathcal{G}(y)$ in the associated continuity equations, as defined in (13) and (15) .

Lemma 3.1 *For each of the updating protocols enumerated earlier in this section, $\mathcal{F}(x) = R^1(x)$ and $\mathcal{G}(y) = R^2(y)$.*

Proof. We prove the result only for Rule 1 for population 1. The proof for Rule 2 is similar.

---

[20]For the moment, we are ignoring the requirement of imposing restrictions on a_t and u_{ij} such that the probability x'_r actually makes sense

[21]We once again momentarily set $\tau = 1$ and ignore any restriction we need to put on a_t for x' to be a probability distribution. We also temporarily drop the assumption that the u_{rj} are negative.

We show that for $f_{ij}(x)$ given by (24), $[\mathcal{F}(x)y]_r = R_r^1(x)y$, for $r \in S_1$. ¿From (13) we have

$$
\begin{aligned}
[\mathcal{F}(x)\,y]_r &= \sum_{i,j \in S} x_i f_{ij,r}(x) y_j \\
&= \sum_{i,j \in S} x_i y_j \left(\delta_{ir} - x_r \right) u_{ij} \\
&= x_r \left(\sum_{j \in S_2} u_{rj} y_j - \sum_{i,j \in S} x_i u_{ij} y_j \right) \\
&= x_r \left\{ [Uy]_r - x \cdot Uy \right\} \\
&= R_r^1(x)y.
\end{aligned}
$$

The proof for population 2 and $\mathcal{G}(y)$ is similar. ∎

The following proposition is now immediate.

Proposition 3.2 *Under the forward state change rules (24) and (26), the continuity equations (12) and (14) are given by*

$$
\frac{d}{dt} \langle \phi \mid P_t \rangle = \int_{x \in \Delta_1} \nabla \phi(x) \cdot \left[R^1(x) \langle y \mid Q_t \rangle \right] P_t(dx), \tag{29}
$$

$$
\frac{d}{dt} \langle \psi \mid Q_t \rangle = \int_{y \in \Delta_2} \nabla \psi(y) \cdot \left[R^2(y) \langle x \mid P_t \rangle \right] Q_t(dy). \tag{30}
$$

We call (29) and (30) the *replicator continuity equations*.[22]

In a similar way, we obtain the replicator continuity equation for a single population, symmetric game.[23] Let $f_{ij}(x)$ be the mixed strategy rule in a symmetric game where $f_{ij}(x)$ can take the form in (24) or (26). We write $R = R^1$, as in (27), for the replicator operator in the symmetric case.[24]

---

[22]These are the weak forms of the replicator continuity equations. There are obvious strong forms corresponding to (16) and (17) for measures characterized by differentiable density functions.

[23]For the theoretical study of continuity equations, we use simple normal form games where realized payoff is identical to expected payoff. In the price dispersion game discussed in the Introduction, there is, however, a difference. If the two matched sellers charge prices (p_1, p_2), then, for the firm charging p_1, realized payoff is either p_1 or 0, depending upon whether a sale materializes or not. However, the expected payoffs are as given in footnote 9. It is, however, not difficult to extend the theoretical framework we develop in this section to allow the sellers to exercise reinforcement learning on the basis of their realized payoffs only. The consumer can be treated as 'Nature' which makes an initial random move that determines the game between sellers: there are two choices of consumer strategy chosen with probabilities y_1 and y_2. Once such a choice is made, the seller payoff matrix is determined. If we then apply, for example, the Börgers and Sarin (1997) reinforcement rule (23), the expected change in a mixed strategy continues to be given by the replicator dynamic. Hence, the change in the population state—the probability measure—also continues to be given by the replicator continuity equation.

[24]Consider a symmetric game with payoff matrix U and let $\mu_t = \langle x | p_t \rangle$ be the mean. The strong form of the replicator continuity equation is $\frac{\partial p_t(x)}{\partial t} + \nabla \cdot [R(x)\mu_t p_t(x)] = 0$. We note that this dynamic is very different from those used in the early biology literature concerning the evolution of mixed strategies. For example, Zeeman (1981) uses a straightforward adaptation of the replicator dynamic having the form $\frac{\partial p_t(x)}{\partial t} = p_t(x) \left(x \cdot U\mu - \mu \cdot U\mu \right)$, and Hines (1980) uses the mean payoff adjusted replicator dynamic, $\frac{\partial p_t(x)}{\partial t} = \frac{p_t(x)}{\mu \cdot U\mu} \left(x \cdot U\mu - \mu \cdot U\mu \right)$ introduced in Maynard Smith (1982) for pure strategies.

Proposition 3.3 *Let P_t be the probability measure over mixed strategies in a symmetric game. Then, under each of the updating protocols (24)-(26), the weak continuity equation (19) is given by*

$$\frac{d}{dt}\langle \phi \mid P_t \rangle = \int_{x \in \Delta} \nabla \phi(x) \cdot [R(x)\langle x \mid P_t \rangle] \, P_t(dx). \tag{31}$$

3.1 Example: Mean replicator dynamics

As an example of (31), it is instructive to derive a more explicit form of the dynamics for the mean $\mu_t = \langle x \mid P_t \rangle$. We focus on the symmetric case since it simplifies notation even as the ideas involved can be readily adapted to the asymmetric case. Thus, taking $\phi(x) = x_i$ in (31), we have:

$$\begin{aligned}
\frac{d}{dt}\mu_i(t) &= \int_\Delta R_i(x)\mu(t)P_t(dx) \\
&= \int_\Delta x_i\,(e_i - x) \cdot U\mu(t)P_t(dx) \\
&= (e_i \cdot U\mu(t)) \int_\Delta x_i P_t(dx) - \int_\Delta x_i x \cdot U\mu(t)P_t(dx) \\
&= (e_i \cdot U\mu(t))\,\mu_i(t) - (\mu(t) \cdot U\mu(t))\,\mu_i(t) - \mathrm{Cov}_t\,(x_i, x) \cdot U\mu(t) \\
&= R_i\,(\mu(t))\,\mu(t) - \mathrm{Cov}_t\,(x_i, x) \cdot U\mu(t),
\end{aligned}$$

where $\mathrm{Cov}(x, x)$ is the covariance matrix

$$\mathrm{Cov}(x, x)_{ij} = \mathrm{Cov}(x_i, x_j) = \int_\Delta (x_i - \mu_i)(x_j - \mu_j)P(dx). \tag{32}$$

That is, the continuity replicator dynamics for means can be written in the form

$$\frac{d\mu}{dt} = R(\mu)\mu - \mathrm{Cov}(x, x)U\mu. \tag{33}$$

Equation (33) makes clear that the continuity replicator dynamics of means differs from the classical replicator dynamics applied to the population mean through a covariance term, which cannot be reduced to a function of the mean.[25]

3.2 Rest points

If x^* is a rest point of the symmetric replicator operator, i.e. $R(x^*)x^* = 0$, and we take $P_0 = \delta_{x^*}$,[26] the mass-point distribution at x^*, then the initial mean is $\langle x \mid P_0 \rangle = x^*$, and hence from (31),

---

[25]We note that Hines (1980) derives an equation for mean dynamics from the mean-payoff adjusted replicator dynamic (see footnote 24). His equation has the form

$$\dot{\mu} = \frac{1}{\mu \cdot U\mu}\mathrm{Cov}\,(x, x)\,U\mu.$$

[26]The mass-point, or Dirac measure at $x^* \in \Delta$ is defined by: $\langle \phi \mid \delta_{x^*} \rangle = \phi(x^*)$ for any differentiable function ϕ on Δ. By convention, this distribution is represented by the Dirac probability 'density': $\delta_{x^*}(dx) = \delta(x - x^*)dx$. We sometimes adopt this convention.

$\frac{d}{dt}\langle \phi \mid P_t\rangle|_{t=0} = 0$. Thus, δ_{x^*} defines a rest point of the replicator continuity dynamics (31).

In particular, the pure strategy distributions, $\delta_{e_i}, 1 \leq i \leq n$, are all rest points of the continuity dynamics. More generally, it also follows easily from (31) that any distribution which is a mixture of pure strategies:

$$P = \sum_{i=1}^{n} x_i \delta_{e_i}, \qquad x \in \Delta, \tag{34}$$

is a rest point of the continuity dynamics. In this case, the mean is $\mu = x$. Intuitively, this is because reinforcement learning (23) always reinforces the pure strategies that are played, and hence agents who consistently play a pure strategy never have an incentive to modify their strategy. This result is in complete contrast to the usual interpretation of the situation modelled by the classical replicator dynamics, in which the population is always a mixture of pure strategy players. We note however, that distributions of the form (34) are often unstable under the continuity replicator dynamics; they are vulnerable to invasion by mixed-strategy players.

Also, any Nash equilibrium x^* is a rest point of the classical replicator dynamics, and hence δ_{x^*} is a rest point of the continuity dynamics (31). More generally, we show that *any* initial distribution P_0 with mean a Nash equilibrium is a stationary solution of the continuity dynamics, with a similar statement in the asymmetric case.

Proposition 3.4 A: Let (x^*, y^*) be a Nash equilibrium of an asymmetric game with payoff matrix (U, V). Let P_0 and Q_0 be two measures satisfying $\langle x \mid P_0\rangle = x^*$ and $\langle y \mid Q_0\rangle = y^*$. Then $(P_t, Q_t) = (P_0, Q_0)$ is a stationary solution of the continuity replicator dynamics.

B: Let x^* be a Nash equilibrium of a symmetric game with payoff matrix U. Let P_0 be a measure satisfying $\langle x \mid P_0\rangle = x^*$. Then $P_t = P_0$ is a stationary solution of the continuity replicator dynamics.

Proof. A. First suppose that (x^*, y^*) has full support. Then $Uy^* = \pi_1^* e^1$ and $V\pi_2^* e^2$, where $\pi_1^* = x^* \cdot Uy^*$ and $\pi_2^* = y^* \cdot Vx^*$ are the equilibrium payoffs to players 1 and 2, respectively, and $e^k \in \mathbb{R}^{n_k}$ is the vector all of whose entries are 1. Observe from (27)-(28) that $R^1(x)y^* = 0$ and $R^2(y)x^* = 0$ for any $(x, y) \in \Delta$. By assumption, at $t = 0$, $\langle x \mid P_0\rangle = x^*$ and $\langle y \mid Q_0\rangle = y^*$. Hence, from equations (29)-(30), $\frac{d}{dt}\langle \phi \mid P_t\rangle|_{t=0} = \frac{d}{dt}\langle \psi \mid Q_t\rangle|_{t=0} = 0$. That is, (P_0, Q_0) is a rest point of equations (29)-(30).

Now suppose that (x^*, y^*) does not have full support. Let $\partial\Delta_1(x^*)$ be the lowest dimensional face of Δ_1 containing x^*. Then $\langle x \mid P_0\rangle = x^*$ implies that P_0 is supported on $\partial\Delta_1(x^*)$. That is $\Delta_1 \setminus \partial\Delta_1(x^*)$ has P_0-measure 0. Similarly, define $\partial\Delta_2(y^*)$. Then (x^*, y^*) has full support on $\partial\Delta_1(x^*) \times \partial\Delta_2(y^*)$, and the argument above shows that (P_0, Q_0) is a rest point of (29)-(30).

B. The proof in the symmetric case is similar. ∎

4 Solution of the General Continuity Equation: Liouville's Formula

Our approach to solving the non-linear continuity equations we have constructed is to begin by solving a different, but related problem. Thus, instead of confronting the non-linearities directly, we first consider a linear continuity equation, but one defined by an explicitly time-dependent vector field. We will later show how a solution of the non-linear continuity equations of interest can be constructed from explicit solutions of linear continuity equations of this type.

4.1 Liouville's Formula

Let $X = X(x,t) \in \mathbb{R}^n$ be a (possibly time-dependent) smooth vector field defined for x in a neighbourhood of the state space $\Omega \subset \mathbb{R}^n$, where Ω is a compact, connected domain with non-empty interior and piecewise smooth boundary. We assume that Ω is invariant under the flow determined by $X(x,t)$. Let P_t be a probability measure on Ω satisfying the linear weak continuity equation

$$\frac{d}{dt}\langle \phi \mid P_t \rangle = \int_{\Omega} \nabla\phi(x) \cdot X(x,t) P_t(dx), \tag{35}$$

for all smooth test functions $\phi(x)$, and for given initial measure P_0. The solution to this initial-value problem may be described as follows.

We first introduce some notation to describe the solution trajectories to the (non-autonomous) differential equations defined by X,

$$\frac{dx}{dt} = X(x,t). \tag{36}$$

Let $x_{t_0,t}(x), t \in \mathbb{R}$, denote the solution trajectory to (36) that passes through the point $x \in \Omega$ at time t_0. Thus, the trajectory that passes through x at time t starts at the point $x_{t,0}(x)$ when $t = 0$.[27] After time $s \geq 0$, this trajectory has reached the point $x_{t,s}(x) = x_{0,s}(x_{t,0}(x))$. In particular, $x_{t,t}(x) = x_{0,t}(x_{t,0}(x)) = x$, and by definition $x_{t,0}(x_{0,t}(x)) = x$.

We can now write down the solution to the initial value problem (35):

$$P_t(B) = P_0(x_{t,0}(B)). \tag{37}$$

This is Liouville's formula for measures. A proof is given in Appendix A.2.

In the case in which P_0 is absolutely continuous, so that $P_0(dx) = p_0(x)dx$ for an initial density function $p_0(x)$, then it is also shown in Appendix A.2 that the solution (37) is described by a more classical form of Liouville's formula which determines the density function $p_t(x)$ associated to P_t:

$$p_t(x) = p_0(x_{t,0}(x)) \exp\left\{ -\int_0^t [\nabla \cdot X](x_{t,s}(x), s)\, ds \right\}. \tag{38}$$

---

[27]Note that the situation for a non-autonomous vector field is more complicated than for the more familiar autonomous case. This is because the explicit time dependence of $X(x,t)$ imposes an absolute, rather than a relative, time-scale on the dynamics. In particular, the initial time $t = 0$ is exogenously determined.

4.2 Expected Values

Liouville's formula (37) allows us to calculate expected values of associated variables in terms of the initial measure P_0 and solutions of the characteristic system (36). Thus, for a smooth function $\phi(x)$, define its expected value with respect to the probability density P_t satisfying (35) by:

$$\langle \phi \mid P_t \rangle = \int_\Omega \phi(x) P_t(dx). \tag{39}$$

Then we have:

Proposition 4.1 *The expected value $\langle \phi \mid P_t \rangle$ may be expressed in the form:*

$$\langle \phi \mid P_t \rangle = \int_\Omega \phi\left(x_{0,t}(x)\right) P_0(dx). \tag{40}$$

A proof is given in Appendix A.2.

As an example of the use of (40), the following Corollary shows that the trajectories of the underlying characteristic dynamics (36) may be recovered as solutions of the continuity equation (35) for initial conditions which are mass points.

Corollary 4.2 *Suppose $p_0(x) = \delta(x - x_0)$ for some $x_0 \in \Omega$ is an initial mass point density. Then $p_t(x) = \delta\left(x - x_{0,t}(x_0)\right)$ for all $t \geq 0$, where $x_{0,t}(x_0)$ is the solution trajectory of the characteristic equations $\dot{x} = X(x,t)$ with initial condition x_0.*

Proof. If $p_0(x) = \delta(x - x_0)$ in (35), it follows from (40) that $\langle \phi \mid p_t \rangle = \phi\left(x_{0,t}(x_0)\right)$. Since this is true for any continuous function $\phi(x)$, it follows from (39) that $p_t(x) = \delta\left(x - x_{0,t}(x_0)\right)$. ∎

5 Application of Liouville's Formula to the Replicator Continuity Equation

In this section, we use Liouville's formula (37) to lay the foundations for a solution to the pair of coupled continuity equations (29) and (30). To do this, we first "freeze" population 2 in the following sense. Suppose that mixed strategies are distributed over agents in population 2 by a fixed, time dependent probability measure Q_t that is independent of any process in population 1. This measure determines a mean history, $y(t) = \langle y \mid Q_t \rangle \in \Delta_2$, which in turn determines the evolution of the probability measure P_t for population 1 via the continuity equation (29). In effect, this scenario replaces the population 2 of responsive agents by a non-stationary environment with which agents in population 1 interact, and whose behaviour is determined by the fixed, but now non-stationary process Q_t.

The outcome of this "freezing" process is that we can consider the population 1 continuity equation (29) as decoupled from (30). In the next section, we shall recover this coupling by considering a simultaneous "freezing" procedure for both populations.

We have defined the replicator operators $R^1(x) : \Delta_2 \to I\!\!R^{n_1}$ in (27). Suppose given a specified history $y(t) \in \Delta_2$, as described above. We associate a *pseudo replicator dynamic* to this trajectory, whose solutions specify the time-development of row-player responses to this history. This takes the form of the explicitly time-dependent dynamical system

$$\dot{x}_i = R_i^1(x)y(t) = x_i(e_i^1 - x) \cdot Uy(t), \tag{41}$$

where $e_i^1 \in I\!\!R^{n_1}$ is the i-th standard basis vector. This is an explicitly time-dependent dynamical system, which we consider as the characteristic ODE (36) for a general continuity equation (35). To solve this continuity equation, we begin by solving the characteristic system (41). We can then find the solution of any associated initial value problem of the form (35) by means of Liouville's formula (37).

5.1 Solution of the pseudo Replicator Dynamics

Write $c(t) = Uy(t) \in I\!\!R^{n_1}$, a time-dependent vector-payoff stream to row players. Then the pseudo-replicator equations (41) can be written as:

$$\frac{dx_i}{dt} = x_i \left(e_i^1 - x\right) \cdot c(t). \tag{42}$$

Write

$$C(t) = \int_0^t c(s)ds. \tag{43}$$

Then we can express the solutions of (42) as follows.

Proposition 5.1 *The solution trajectory of the pseudo-replicator dynamics (42) passing through $x \in \Delta_1$ at time $t = t_0$ is:*

$$x_{t_0,t}(x)_i = \frac{x_i e^{C_i(t) - C_i(t_0)}}{x \cdot e^{C(t) - C_i(t_0)}}. \tag{44}$$

In particular:

$$x_{0,t}(x)_i = \frac{x_i e^{C_i(t)}}{x \cdot e^{C(t)}}, \qquad and \qquad x_{t,0}(x)_i = \frac{x_i e^{-C_i(t)}}{x \cdot e^{-C(t)}}. \tag{45}$$

Proof. With $x_{t_0,t}(x)$ given by (44), a direct calculation gives

$$\frac{d}{dt} \left[x_{t_0,t}(x)_i \right] = x_{t_0,t}(x)_i \left\{ e_i - x_{t_0,t}(x) \right\} \cdot c(t),$$

which shows that $x_{t_0,t}(x)$ is a solution of (42). It also follows from (44) that $x_{t_0,t_0}(x) = x$, as required. ∎

5.2 Solution of the Replicator Continuity Equation

In the case in which P_0 is absolutely continuous, we may use Liouville's formula (38), together with Proposition 5.1, to compute the solution to the replicator continuity equation associated with a

pseudo-replicator vector field of the form (42). This is given in the following proposition, proved in Appendix A.3.

Proposition 5.2 *The solution of the initial value problem (35) with initial density $p_0(x)$, associated to the characteristic vector field (42) is:*

$$p_t(x) = p_0 \left(\frac{xe^{-C(t)}}{x \cdot e^{-C(t)}} \right) \left(\frac{1}{x \cdot e^{-C(t)}} \right)^{n_1} \exp \left\{ -e^1 \cdot C(t) \right\}, \tag{46}$$

where $C(t) \in \mathbb{R}^{n_1}$ is given by (43) and $e^1 \in \mathbb{R}^{n_1}$ is the vector all of whose entries are 1.

More generally, we may obtain the expected value of a continuous function $\phi(x)$ from (40) and (45):

$$\langle \phi \mid P_t \rangle = \int_{\Delta_1} \phi \left(\frac{\xi e^{C(t)}}{\xi \cdot e^{C(t)}} \right) P_0(d\xi). \tag{47}$$

We now present two immediate implications of (47) giving conditions under which a row-player pure strategy is eventually eliminated.

Proposition 5.3 *Suppose there exists an i such that $[C_i(t) - C_j(t)] \to \infty$ as $t \to \infty$ for some $j \neq i$, and the i-th face, $\partial\Delta_1^{(i)} = \{x \in \Delta_1 : x_i = 0\}$, has P_0-measure zero. Then $\langle x_j \mid P_t \rangle \to 0$ as $t \to \infty$.*

Proof. For $\xi \in \Delta_1 \setminus \partial\Delta_1^{(i)}$, we have $\xi_i > 0$. Thus:

$$\frac{\xi_j e^{C_j(t)}}{\xi \cdot e^{C(t)}} = \frac{\xi_j e^{C_j(t)}}{\xi_i e^{C_i(t)} + \sum_{k \neq i} \xi_k e^{C_k(t)}} = \frac{\xi_j e^{-[C_i(t) - C_j(t)]}}{\xi_i + \sum_{k \neq i} \xi_k e^{-[C_i(t) - C_k(t)]}} \to 0 \qquad \text{as } t \to \infty,$$

since the denominator is never zero. Hence, from (47),

$$\langle x_j \mid P_t \rangle = \int_{\Delta_1 \setminus \partial\Delta_1^{(i)}} \left(\frac{\xi_j e^{C_j(t)}}{\xi \cdot e^{C(t)}} \right) P_0(d\xi) \to 0 \qquad \text{as } t \to \infty. \quad \blacksquare$$

We may translate Proposition 5.3 into a game-theoretic context as follows. Suppose a population of row players plays a game with $n_1 \times n_2$ payoff matrix U against 'Nature', characterized by a column-player strategy profile $y : \mathbb{R}_+ \to \Delta_2$. If s is a column-player pure strategy, we say that s is *persistent* if $Y_s(T) = \int_0^T y_s(t)dt \to \infty$ as $T \to \infty$. Then we have:

Corollary 5.4 *Suppose the row-player strategy j is weakly dominated by the row-player strategy i, and that the i-th face $\partial\Delta_1^{(i)}$ has P_0-measure zero. Suppose also that there is a persistent column-player strategy s for which $u_{is} > u_{js}$. Then $\langle x_j \mid P_t \rangle \to 0$ as $t \to \infty$. That is, j is eventually eliminated with probability 1.*

Proof. That j is weakly dominated by i implies that $u_{ir} \geq u_{jr}$ for all column-player strategies r. Thus, $c_i(t) = [Uy(t)]_i = \sum_r u_{ir} y_r(t) \geq \sum_r u_{jr} y_r(t) = c_j(t)$ for all $t \geq 0$, with this inequality strict for any t for which $y_s(t) \neq 0$. Thus, $[C_i(t) - C_j(t)] = \sum_r (u_{ir} - u_{jr}) Y_r(t) \geq (u_{is} - u_{js}) Y_s(t) \to \infty$ as $t \to \infty$, since s is persistent. The result therefore follows from Proposition 5.3. ■

6 Distributional Replicator Dynamics

In this section we show how, in the asymmetric case, a solution to the pair of coupled continuity equations (29) and (30), or, in the symmetric case, to the corresponding single continuity equation (31), can be obtained from the "frozen" solution (47) for population 1, and an analogous frozen solution for population 2. The coupling of these solutions is then tracked by solutions of an associated ODE system, which we term the *distributional replicator dynamics.*

6.1 Asymmetric Games

Consider solutions to equations (29) and (30). From (42) and (43), these equations can be construed in the first instance as independent ("frozen") continuity equations associated with the time-dependent vectors (one for each population):

$$c(t) = \frac{dC(t)}{dt} = U\langle y \mid Q_t \rangle \in \mathbb{R}^{n_1}, \tag{48}$$

$$d(t) = \frac{dD(t)}{dt} = V\langle x \mid P_t \rangle \in \mathbb{R}^{n_2}. \tag{49}$$

Thus, (47) gives:

$$\langle x \mid P_t \rangle = \int_{\Delta_1} \left(\frac{\xi e^{C(t)}}{\xi \cdot e^{C(t)}} \right) P_0(d\xi), \tag{50}$$

$$\langle y \mid Q_t \rangle = \int_{\Delta_2} \left(\frac{\zeta e^{D(t)}}{\zeta \cdot e^{D(t)}} \right) Q_0(d\zeta). \tag{51}$$

We therefore obtain the system of $n_1 + n_2$ differential equations in the variables $C_1, \ldots, C_{n_1}$ and $D_1, \ldots, D_{n_2}$:

$$\frac{dC_i}{dt} = \sum_{k=1}^{m} u_{ik} \int_{\Delta_2} \left(\frac{\zeta_k e^{D_k}}{\zeta \cdot e^D} \right) Q_0(d\zeta), \qquad C_i(0) = 0, \quad 1 \leq i \leq n_1, \tag{52}$$

$$\frac{dD_j}{dt} = \sum_{l=1}^{n} v_{jl} \int_{\Delta_1} \left(\frac{\xi_l e^{C_l}}{\xi \cdot e^C} \right) P_0(d\xi), \qquad D_j(0) = 0, \quad 1 \leq j \leq n_2. \tag{53}$$

We call these equations the asymmetric *distributional replicator Dynamics* associated with the pair of initial probability measures P_0 and Q_0. The solutions of these equations with the given initial conditions define trajectories $C(t)$ and $D(t)$, in terms of which the continuity dynamics can

be completely specified as in (47), or (46) if measures are absolutely continuous, with analogous formulae for population 2.

Note that at most $n_1 - 1$ of the C_i's and at most $n_2 - 1$ of the D_j's are independent.[28] For example, setting $A_i = C_i - C_{n_1}$ and $B_j = D_j - D_{n_2}$, equations (52) and (53) can be reduced to:

$$\frac{dA_i}{dt} = \sum_{k=1}^{m} (u_{ik} - u_{n_1 k}) \int_{\Omega_2} \left(\frac{\zeta_k e^{B_k}}{\zeta \cdot e^B} \right) Q_0(d\zeta), \qquad A_i(0) = 0, \quad 1 \le i \le n_1 - 1, \tag{54}$$

$$\frac{dB_j}{dt} = \sum_{l=1}^{n} (v_{jl} - v_{n_2 l}) \int_{\Omega_1} \left(\frac{\xi_l e^{A_l}}{\xi \cdot e^A} \right) P_0(d\xi), \qquad B_j(0) = 0, \quad 1 \le j \le n_2 - 1, \tag{55}$$

where Ω_1 and Ω_2 are the projections of Δ_1 and Δ_2 onto $I\!\!R^{n_1-1}$ and $I\!\!R^{n_2-1}$, respectively, given by $x_{n_1} = 1 - \sum_{i=1}^{n_1-1} x_i$ and $y_{n_2} = 1 - \sum_{j=1}^{n_2-1} y_j$ (see definition (96) of Appendix A.3). Of course, $A_{n_1} = B_{n_2} = 0$.

6.2 Symmetric Games

We consider the continuity equation (31) associated with a 2-player, n-strategy symmetric game having $n \times n$ payoff matrix U. In terms of the theory of section 5, this is the continuity equation associated to the time-dependent mixed strategy profile $y : I\!\!R \to \Delta$ given by $y(t) = \langle x \mid P_t \rangle$. That is, $c(t) = U \langle x \mid P_t \rangle$. Thus, from (43) we have

$$c(t) = \frac{dC(t)}{dt} = U \langle x \mid P_t \rangle, \tag{56}$$

and using (47) we therefore obtain a system of n differential equations in the variables $C_1, \ldots, C_n$:

$$\frac{dC_i}{dt} = \sum_{j=1}^{n} u_{ij} \int_{\Delta} \left(\frac{\xi_j e^{C_j}}{\xi \cdot e^C} \right) P_0(d\xi), \qquad C_i(0) = 0, \quad 1 \le i \le n. \tag{57}$$

Following section 6.1, we call equations (57) the symmetric *distributional replicator dynamics* associated with the initial measure P_0. The solutions of these equations with the given initial conditions define trajectories $C(t)$, in terms of which the continuity dynamics can be completely specified as in (47), or (46) in the absolutely continuous case.

Again, at most $n - 1$ of equations (57) are independent. For example, setting $A_i = C_i - C_n$, equations (57) can be reduced to

$$\frac{dA_i}{dt} = \sum_{j=1}^{n} (u_{ij} - u_{nj}) \int_{\Omega} \left(\frac{\xi_j e^{A_j}}{\xi \cdot e^A} \right) P_0(d\xi), \qquad A_i(0) = 0, \quad 1 \le i \le n - 1, \tag{58}$$

where $\Omega \subset I\!\!R^{n-1}$ is the projection of Δ onto $I\!\!R^{n-1}$ obtained by setting $x_n = 1 - \sum_{i=1}^{n-1} x_i$. Of course $A_n = 0$. Note that the formulae (46) and (47) can be expressed in terms of the A_i's.

---

[28]Because of the constraints $\sum_j \langle y_j \mid Q_t \rangle = \sum_i \langle x_i \mid P_t \rangle = 1$.

6.3 The modified replicator formulation

We now write the distributional dynamics (52-53) for the asymmetric game with payoff matrix (U, V) in the form

$$\frac{dC}{dt} = UF\left(e^D \mid Q_0\right), \qquad C(0) = 0, \tag{59}$$

$$\frac{dD}{dt} = VF\left(e^C \mid P_0\right), \qquad D(0) = 0, \tag{60}$$

where $F(\cdot \mid P_0) : \mathbb{R}_+^{n_1} \to \Delta_1$ is the function

$$F_i(\xi \mid P_0) = \int_{\Delta_1} \left(\frac{\xi_i x_i}{\xi \cdot x}\right) P_0(dx). \tag{61}$$

Clearly $0 \leq F_i \leq 1$ and $e^1 \cdot F = 1$, and hence $F \in \Delta_1$. Note also that $F(\alpha\xi \mid P_0) = F(\xi \mid P_0)$ for any positive scaler α. In addition, if int Δ_1 has non-zero P_0-measure, then $F_i(\xi \mid P_0)$ is monotonically increasing in ξ_i, with $\lim_{\xi_i \to 0} F_i(\xi \mid P_0) = 0$ and $\lim_{\xi_i \to \infty} F_i(\xi \mid P_0) = 1$. On the other hand, $F_i(\xi \mid P_0)$ is monotonically decreasing in ξ_j for $j \neq i$, with $\lim_{\xi_j \to \infty} F_i(\xi \mid P_0) = 0$. Also $F(e^1 \mid P_0) = \langle x \mid P_0 \rangle$. Clearly, $F(\cdot \mid Q_0) : \mathbb{R}_+^{n_2} \to \Delta_2$ can be defined analogously.

Now define new variables $\xi = e^C/(e^1 \cdot e^C) \in \Delta_1$ and $\zeta = e^D/(e^2 \cdot e^D) \in \Delta_2$. Then a straightforward calculation shows that the distributional replicator equations (59)-(60) can be written in the form

$$\frac{d\xi}{dt} = R^1(\xi)F(\zeta \mid Q_0), \qquad \xi(0) = \frac{1}{n_1}e^1, \tag{62}$$

$$\frac{d\zeta}{dt} = R^2(\zeta)F(\xi \mid P_0), \qquad \zeta(0) = \frac{1}{n_2}e^2, \tag{63}$$

where $R^1(\xi)$ and $R^2(\zeta)$ are the replicator operators (27) and (28).

Similarly in the symmetric case, using the variables $\xi = e^C/(e \cdot e^C) \in \Delta$, we have

$$\frac{d\xi}{dt} = R(\xi)F(\xi \mid P_0), \qquad \xi(0) = \frac{1}{n}e. \tag{64}$$

These versions express the distributional replicator dynamics as dynamical systems which are restricted to simplex phase spaces.

7 Stability Analysis for Negative Definite Games

We have observed in Proposition 3.4 that any probability measure whose mean is a Nash equilibrium is a stationary solution of the replicator continuity equation. In this section, we study the convergence of trajectories of this dynamic to such stationary points in the class of negative definite and negative semi-definite games. These results are of interest since this class of games encompass a wide variety of well known games. For example, games with an interior ESS are negative definite

games whereas two player zero-sum games are negative semi-definite games The stability properties of Nash equilibria in such games have been established under a wide range of evolutionary dynamics.[29] Our interest is in seeing whether such stability results can be extended to our case in which agents are assumed to operate in a far more parsimonious information structure than is conventionally assumed.

We first consider the convergence of the mean of the social state to a Nash equilibrium. We say a symmetric game with $n \times n$ payoff matrix U is negative semi-definite on $I\!\!R_0^n$ if

$$z \cdot Uz \leq 0 \qquad \text{for all } z \in I\!\!R_0^n,$$

and is negative definite if this inequality is strict. An attractive feature of negative definite games is that they have a unique Nash equilibrium (Sandholm, 2009; Theorem 3.3.16.) For a negative semi-definite game, the set of Nash equilibria is convex. Zero-sum games are the most prominent examples of negative semi-definite games.

We can now prove our stability result for negative definite games. We show that in negative definite games, the mean social state always converges towards the Nash equilibrium under the replicator continuity dynamics.

First we state the following key lemmas, which are proved in Appendix A.5. To do this we require some notation. Let $S \subseteq \{1, 2, \ldots, n\}$ be a non-empty set of pure strategies. Define $e_S = \sum_{i \in S} e_i \in I\!\!R^n$ and the subspace $I\!\!R_S^n = \{x \in I\!\!R^n \mid x_j = 0 \text{ for } j \notin S\}$. Then we note that $I\!\!R_S^n$ may be decomposed as $I\!\!R_S^n = [e_S] \bigoplus I\!\!R_{S0}^n$, where $[e_S]$ is the 1-dimensional subspace generated by e_S, and $I\!\!R_{S0}^n = \{x \in I\!\!R_S^n \mid \sum_i x_i = 0\}$.

Lemma 7.1 *Suppose* $\text{int} \Delta$ *has positive* P*-measure. Then the Jacobian derivative* $DF(\xi \mid P) = \left(\partial F_i(\xi \mid P) / \partial C_j \right)$ *is symmetric. Let* $S = \text{supp}(\xi)$*. Then the* j*-th row and* j*-th column of* $DF(\xi)$ *are zero for* $j \notin S$*, and* $e_S \cdot DF(\xi) = DF(\xi)e_S = 0$*. Further,* $DF(\xi)$ *is positive-definite on* $I\!\!R_{S0}^n$*.*

Lemma 7.2 *Suppose* $\text{int} \Delta$ *has positive* P*-measure, and let* $u \in \Delta$*. Then there is a unique* $\xi^* = \xi^*(u) \in \Delta$ *such that* $F(\xi^* \mid P) = u$*. Further,* $\text{supp}(\xi^*) = \text{supp}(u)$*.*

Theorem 7.3 *Consider a symmetric* $n \times n$ *game with payoff matrix* U*, and suppose that* U *is negative definite on* $I\!\!R_0^n$*, with unique Nash equilibrium* $x^* \in \Delta$*. Let* P_0 *be a probability distribution on* Δ *for which* $\text{int} \Delta$ *has positive* P_0*-measure, and let* P_t *be the solution of the continuity replicator dynamics with initial condition* P_0*. Then the mean population state* $\mu_t = \langle x \mid P_t \rangle \to x^*$*. In particular, the Nash equilibrium is globally asymptotically stable under the mean replicator dynamics.*

If the game U *is negative semi-definite, then the convex set of Nash equilibria is Lyapunov stable.*

Proof. Let $\xi^* = \xi^*(x^*) \in \Delta$ be the unique point satisfying $F(\xi^* \mid P_0) = x^*$ (Lemma 7.2). We show the global dynamic stability of the equilibrium ξ^* under the dynamics (64) using the Lyapunov

[29]See Sandholm (2009) for a discussion of these results. Sandholm (2009) refers to negative semi- definite games as stable games and negative definite games as strictly stable games.

function

$$K\left(\xi \mid P_0\right) = -x^* \cdot \ln \xi + \int_\Delta \ln\left(\xi \cdot x\right) P_0(dx). \tag{65}$$

This is well-defined on the subset $\{\xi \in \Delta \mid S(x^*) \subseteq S(\xi)\}$. One checks that x^* is a global minimum of K on Δ (cf. Appendix A.5, proof of lemma 7.2). Then, for $\xi \neq \xi^*$,

$$
\begin{aligned}
\frac{dK}{dt} &= \sum_{i=1}^n \frac{1}{\xi_i}\left\{-x_i^* + F_i(\xi \mid P_0)\right\} \dot{\xi}_i \\
&= \sum_{i=1}^n \left\{-x_i^* + F_i(\xi \mid P_0)\right\}\left(e_i^1 - \xi\right)\cdot U F(\xi) \quad \text{using (64)} \\
&= -\left(x^* - F(\xi)\right)\cdot U F(\xi) \\
&= \left(x^* - F(\xi)\right)\cdot U\left(x^* - F(\xi)\right) - \left(x^* - F(\xi)\right)\cdot U x^*.
\end{aligned}
$$

The second term is non-negative since x^* is a Nash equilibrium, and the first term is negative if U is negative-definite, since $x^* - F(\xi) \in \mathbb{R}_0^n$. Clearly $\dot{K}(\xi^*) = 0$. Thus, $K(\xi) - K(\xi^*)$ is a global Lyapunov function, and the result follows.

If U is negative semi-definite, then $\dot{K}(\xi) \leq 0$, and the Lypunov stability of any Nash equilibrium x^* follows. ∎

7.0.1 Asymmetric games

Asymmetric games cannot be negative-definite, but can be negative semi-definite in the following sense. An asymmetric game (U, V) is negative semi-definite if

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \cdot \begin{pmatrix} 0 & U \\ V & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = 0 \qquad \text{for all } (z_1, z_2) \in \mathbb{R}_0^{n_1} \times \mathbb{R}_0^{n_2}.$$

For example, asymmetric zero-sum games are negative semi-definite. The set of Nash equilibria (x^*, y^*) for such a game forms a convex subset of $\Delta = \Delta_1 \times \Delta_2$. A similar proof to that given for Theorem 7.3, using the Lyapunov function $K(\xi \mid P_0) + K(\zeta \mid Q_0)$, shows that this set is Lyapunov stable under analogous assumptions on (P_0, Q_0).

7.1 Limiting distributions

Theorem 7.3 shows that the asymptotics of the mean are essentially independent of the initial distribution. However, this does not mean that the asymptotic distribution itself is independent of the initial distribution. As is evident from Proposition 3.4, there are many such stationary distributions, and in particular P_t need not converge to a mass-point distribution at x^*.

This is very important. Although on average, the equilibrium population plays the Nash equilibrium, at the individual level the population can be very heterogeneous.

To see this, we consider a $n \times n$ symmetric negative definite game with interior equilibrium x^*, as in Theorem 7.3. Consider the solution of the distributional replicator equations given by (58). From the global stability of the equilibrium A^*,[30] we have

$$
\begin{aligned}
x_{0,t}(x)_i \quad &= \quad \frac{x_i e^{C_i(t)}}{\sum_{j=1}^{n-1} x_j e^{C_j(t)} + x_n e^{C_n(t)}} = \frac{x_i e^{A_i(t)}}{\sum_{j=1}^{n-1} x_j e^{A_j(t)} + x_n} \\
&\longrightarrow \quad \frac{x_i e^{A_i^*}}{\sum_{j=1}^{n-1} x_j e^{A_j^*} + x_n} = x_i^*(x \mid P_0) \qquad \text{as } t \to \infty.
\end{aligned} \tag{66}
$$

Hence,

$$
\langle \phi \mid P_t \rangle \longrightarrow \langle \phi \mid P_\infty \rangle = \lim_{t \to \infty} \int_\Delta \phi\left(x_{0,t}(x)\right) P_0(dx) = \int_\Delta \phi\left(x^*(x \mid P_0)\right) P_0(dx). \tag{67}
$$

This defines the limiting probability distribution P_∞.

7.1.1   Example

Consider a generic 2×2 symmetric game with invertible payoff matrix U. Then U has eigenvalue λ on $I\!\!R_0^2 = \{(\xi, -\xi) \in I\!\!R^2 \mid \xi \in I\!\!R\}$ and equilibrium $(x^*, 1 - x^*)$ where

$$
\lambda = (u_{11} - u_{21}) + (u_{22} - u_{12}), \qquad \text{and} \qquad x^* = \frac{u_{22} - u_{12}}{(u_{11} - u_{21}) + (u_{22} - u_{12})}. \tag{68}
$$

Then $x^* \in (0, 1)$ and U is negative-definite on $I\!\!R_0^2$ if and only if the payoff differences $u_{11} - u_{21}$ and $u_{22} - u_{12}$ are both negative.

Suppose P_0 is represented by a probability density function on $[0, 1]$, $P_0(dx) = p_0(x)dx$. Then we can use the formula (46) in Proposition 5.2 to obtain the limiting density:

$$
p_\infty(x) = p_0 \left(\frac{x e^{-A^*}}{1 - x + x e^{-A^*}} \right) \frac{e^{-A^*}}{\left(1 - x + x e^{-A^*}\right)^2}, \tag{69}
$$

where A^* is the unique solution of $\int_0^1 \frac{x e^A}{1 - x + x e^A} p_0(x)dx = x^*$.

For example, for the uniform distribution, $P_0(dx) = dx$, A^* is the solution of

$$
\int_0^1 \frac{x e^A}{1 - x + x e^A} dx = \frac{e^A \left(e^A - A - 1\right)}{\left(e^A - 1\right)^2} = x^*. \tag{70}
$$

This is illustrated in Figure 1.

---

[30]Recall that $A_i = C_i - C_n$, and hence $e^{A_i} = \xi_i/\xi_n$. Thus, A^* is defined by $A_i^* = \ln(\xi_i^*) - \ln(\xi_n^*)$, where $\xi^* = \xi^*(x^*) \in \text{int } \Delta$, is as in Lemma 7.2.
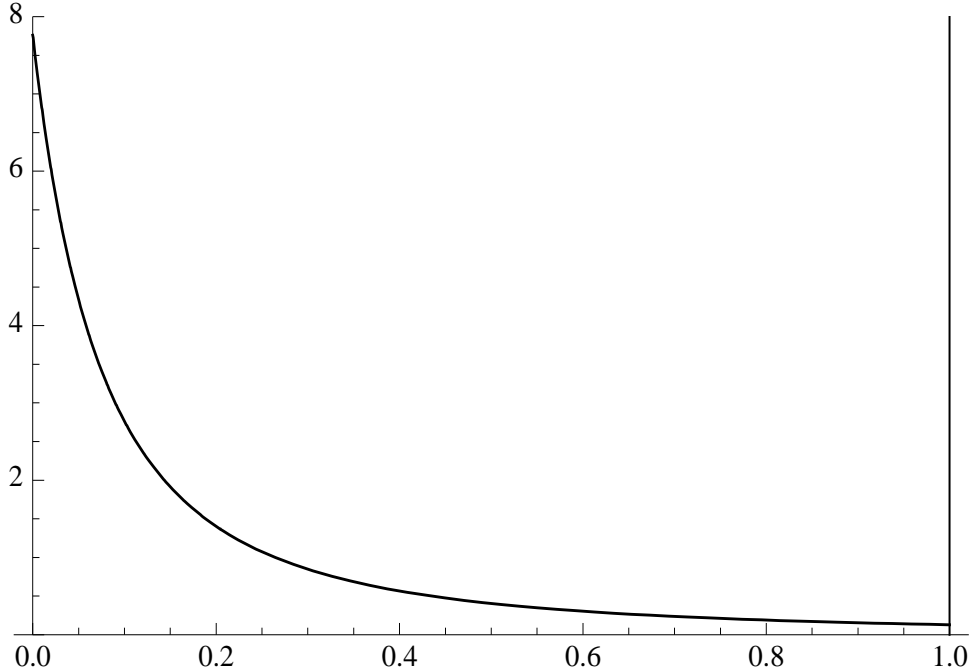
Figure 1: The limiting density $p_\infty(x)$ defined by equation (69) for the uniform distribution $p_0(x) = 1$. In this example, $x^* = 0.2$, and $A^* = -2.0491$. This density has mean x^* and variance 0.04875.

8 Doubly symmetric games and mean payoff

A symmetric population game is called doubly symmetric if the payoff matrix U itself is symmetric, so that $u_{ij} = u_{ji}$ (e.g. a coordination game with positive diagonal elements and zero non diagonal elements). Define the *mean payoff* with respect to the distribution P to be $\bar{w}(\mu) = \mu \cdot U\mu$, where $\mu = \langle x \mid P \rangle$ is the population mean. We first show that $\bar{w}(\mu)$ increases along non-equilibrium trajectories.[31] We then use this result to establish convergence of the population mean to the set of Nash equilibria.

Theorem 8.1 *Let U be the payoff matrix of a symmetric game, and suppose that U is a symmetric matrix. Suppose that $int\,\Delta$ has positive P_0-measure. Then $\bar{w}(\mu_t)$ increases along non-equilibrium trajectories of the continuity replicator dynamics, and the mean μ_t converges to a level set (with respect to $\bar{w}$) of Nash equilibria.*

---

[31]The biological interpretation of this result is the well known Fundamental Theorem of Natural Selection in classical population genetics, in which the entries of U are genotype fitnesses. We also note that the alternative definition of mean payoff given by $\bar{w} = \langle x \cdot Ux \rangle$ need not increase along non-equlibrium trajectories.

Proof. Since U is symmetric we have

$$
\begin{aligned}
\frac{1}{2}\dot{w}(\mu) &= \mu \cdot U\dot{\mu} \\
&= F(e^C \mid P_0) \cdot U\dot{F}(e^C \mid P_0) \\
&= F(e^C) \cdot UDF(e^C)\dot{C} \\
&= F(e^C) \cdot UDF(e^C)UF(e^C) \qquad \text{from the symmetric form of (59)} \\
&= [(UF) \cdot DF\,(UF)]\,(\xi).
\end{aligned}
$$

This is positive by Lemma 7.1, provided $UF(\xi)$ has a non-zero component in $I\!R^n_{S0}$, where $S = \text{supp}(\xi)$. This is not the case if and only if $UF(\xi) = \pi^* e_S + v$ for some constant π^* and $v \in I\!R^n$ with $v_i = 0$ for $i \in S$, in which case $\dot{w}(\mu) = 0$. If ξ^* is such a point, then $R(\xi^*)F(\xi^*) = 0$, and hence ξ^* is an equilibrium of the distributional dynamics (64). If $x^* = F(\xi^*)$, this implies that x^* is a rest point of the standard replicator dynamic, $R(x^*)x^* = 0$, since $\text{supp}(x^*) = \text{supp}(\xi^*)$.

Hence, $\dot{w}(\mu) \geq 0$, for all μ with the equality holding only if μ is a rest point of the classical replicator dynamic. However, it is known that any local maximum of the mean payoff function of a doubly symmetric game is a Nash equilibrium (see Sandholm, 2009; Theorem 3.1.7). Hence, the mean population state converges to a level set of Nash equilibria along non-equilibrium trajectories.
∎

9   Generic 2×2 Symmetric Games

For a generic symmetric 2×2 game with payoff matrix U, the classical replicator dynamics can be expressed in terms of the two characteristic parameters λ and x^* defined in (68). Thus, for $x \in [0, 1]$ there is a single replicator dynamic equation

$$
\dot{x} = \lambda x(1-x)(x-x^*). \tag{71}
$$

This may be compared to the continuity replicator dynamics for means given by equation (33), which in this case reduces to the single equation

$$
\dot{\mu} = \lambda \left\{ \mu(1-\mu) - V \right\} (\mu - x^*), \tag{72}
$$

where V is the variance:

$$
V_t = \int_0^1 (x - \mu_t)^2 P_t(dx). \tag{73}
$$

As observed in section 3.2, any distribution of the form $P_0 = (1-\alpha)\delta_0 + \alpha\delta_1$ with $\alpha \in [0, 1]$ defines a rest point of the symmetric continuity equation (31), with mean $\mu = \alpha$ and variance $V = \alpha(1-\alpha)$. In particular δ_0 and δ_1 are rest points. However, in contrast to the classical case (71), the sense in which these are (if they are) locally stable rest points of the continuity dynamics

is not immediately clear.

To address this question, we consider the distributional replicator dynamics in the form (58) which, for the 2×2 case, reduces to the single equation

$$\dot{A} = \lambda \left\{ -x^* + F(e^A \mid P_0) \right\}, \qquad F(e^A \mid P_0) = \int_0^1 \frac{x e^A}{1 - x + x e^A} P_0(dx), \tag{74}$$

with initial condition $A(0) = 0$. The case of most interest is when x^* defines an interior equilibrium, $0 < x^* < 1$, and $\lambda > 0$; i.e. both payoff differences $u_{11} - u_{21}$ and $u_{22} - u_{12}$ are positive. For the replicator dynamic (71), the equilibria $x = 0$ and $x = 1$ are both locally asymptotically stable, with basins of attraction $0 \le x < x^*$ and $x^* < x \le 1$, respectively. For the distributional dynamic (74), the following lemma relates the asymptotic behaviour of $A(t)$ to the initial density function.

Lemma 9.1 *Suppose $\lambda > 0$ and $x^* \in (0, 1)$, and that the initial distribution P_0 has no mass point at $x = 0$ or $x = 1$. Let $\mu_0 = \langle x \mid P_0 \rangle$ be the associated mean mixed strategy.*

1. If $\mu_0 < x^*$, then $A(t)$ is monotonically decreasing in t, and $A(t) \to -\infty$ as $t \to \infty$.

2. If $\mu_0 > x^*$, then $A(t)$ is monotonically increasing in t, and $A(t) \to \infty$ as $t \to \infty$.

Proof. From (74), we have $\dot{A}(0) = \lambda (-x^* + \mu_0)$. Since $\lambda > 0$, $\dot{A}(0) > 0$ if $\mu_0 > x^*$ and $\dot{A}(0) < 0$ if $\mu_0 < x^*$. Moreover, the monotonicity properties of $F(e^A \mid P_0)$ imply that the initial conditions are self-reinforcing as t increases. Hence, if $\mu_0 > x^*$, then $\dot{A}(t) > 0$, and if $\mu_0 < x^*$, then $\dot{A}(t) < 0$, for all $t \ge 0$. ∎

We now use Proposition 5.3 and Lemma 9.1 to derive the following proposition.

Proposition 9.2 *Consider a generic 2×2 symmetric game with $\lambda > 0$ and $0 < x^* < 1$. Let $\mu_t = \langle x \mid P_t \rangle$ be the mean with respect to the solution measure P_t. If P_0 has no mass point at $x = 0$ or $x = 1$, then $P_t \to \delta_1$ if $\mu_0 > x^*$, and hence $\mu_t \to 1$, and $P_t \to \delta_0$ if $\mu_0 < x^*$, and hence $\mu_t \to 0$ as $t \to \infty$.*

9.1 Further remarks

Earlier, we interpreted the mean $\mu_t = \langle x \mid P_t \rangle$ as the aggregate social state generated by P_t. Proposition 9.2 implies in the type of 2×2 symmetric games we are considering, and for suitable initial distributions, there is no difference in the long-run aggregate social state under the replicator continuity equation and the classical replicator dynamic. In the long run either all agents play action 1 or all play action 2. Nevertheless, the time-course trajectories of the aggregate state under the two dynamics generally differ, as we show with an example in Figure 2A. In principle, therefore, it would be possible to distinguish whether agents are playing pure or mixed strategies by observing the solution trajectories.

This coincident asymptotic behavior is not, however, a general result, and does not hold for $n \times n$ symmetric games with $n > 2$. We give an example for $n = 3$ below. In section 11.3 we present

an example of a 2×2 *asymmetric* game in which the long run social state differs radically under the classical replicator dynamic and the replicator continuity equations.

10   A 3×3 symmetric example

Consider a symmetric 3×3 game with diagonal payoff matrix $U = \text{diag}\{\lambda_1, \lambda_2, \lambda_3\}$. The classical replicator dynamics are

$$\dot{x}_1 = x_1 \left\{\lambda_1 x_1 (1 - x_1) - \lambda_2 x_2^2 - \lambda_3 x_3^2\right\}, \tag{75}$$

$$\dot{x}_2 = x_2 \left\{-\lambda_1 x_1^2 + \lambda_2 x_2 (1 - x_2) - \lambda_3 x_3^2\right\}, \tag{76}$$

with $x_3 = 1 - x_1 - x_2$. If $\lambda_1, \lambda_2, \lambda_3$ are positive, then the pure strategy equilibria e_1, e_2, e_3 are all asymptotically stable, and there is an interior equilibrium x^* with $x_i^* \propto \lambda_i^{-1}$, which is unstable.

Consider an initial probability distribution of the form $P_0 = \alpha_1 \delta_{a_1} + \alpha_2 \delta_{a_2} + \alpha_3 \delta_{a_3}$, with $\alpha = (\alpha_1, \alpha_2, \alpha_3), a_1, a_2, a_3 \in \text{int } \Delta$. That is, the population initially consists of a mixture three types of player, playing mixed strategies a_1, a_2 and a_3, respectively. The mean for this distribution is $\mu_0 = \alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3$. The associated distributional replicator dynamics (58) are

$$\frac{dA_1}{dt} = \sum_{k=1}^{3} \alpha_k \frac{\lambda_1 a_{k1} e^{A_1} - \lambda_3 a_{k3}}{a_{k1} e^{A_1} + a_{k2} e^{A_2} + a_{k3}}, \tag{77}$$

$$\frac{dA_2}{dt} = \sum_{k=1}^{3} \alpha_k \frac{\lambda_2 a_{k2} e^{A_2} - \lambda_3 a_{k3}}{a_{k1} e^{A_1} + a_{k2} e^{A_2} + a_{k3}}, \tag{78}$$

with initial conditions $A_1(0) = A_2(0) = 0$. If $(A_1(t), A_2(t))$ is the solution trajectory of these equations, then the associated trajectory of the mean $\mu_t = \langle x \mid P_t \rangle$ is

$$\mu_{t,i} = \sum_{k=1}^{3} \alpha_k \frac{a_{ki} e^{A_i(t)}}{a_{k1} e^{A_1(t)} + a_{k2} e^{A_2(t)} + a_{k3}}, \qquad i = 1, 2. \tag{79}$$

We compare this trajectory with the trajectory of the replicator dynamics (75)-(76) with initial condition $x_0 = \mu_0$, and show that parameters can be chosen so that these two trajectories converge to different pure-strategy equilibria.

An example is illustrated in Figure 2B. The basin of attraction of the equilibrium $e_2 = (0, 1)$ for classical replicator trajectories is the region satisfying $e_2 \cdot Ux > x \cdot Ux$. This is shaded in Figure 2B. However, the distributional replicator trajector does not stay within this region, but moves into the basin of attraction of the equilibrium $e_3 = (0, 0)$. [32]

---

[32] A standard proof of stability of ESS equilibria for the classical replicator dynamics uses the entropy function (e.g. Hofbauer and Sigmund, 1998, Chapter 7). For a Nash equilibrium of a symmetric game, x^*, this is defined on the subset $\{x \in \Delta \mid S(x^*) \subseteq S(x)\}$ by

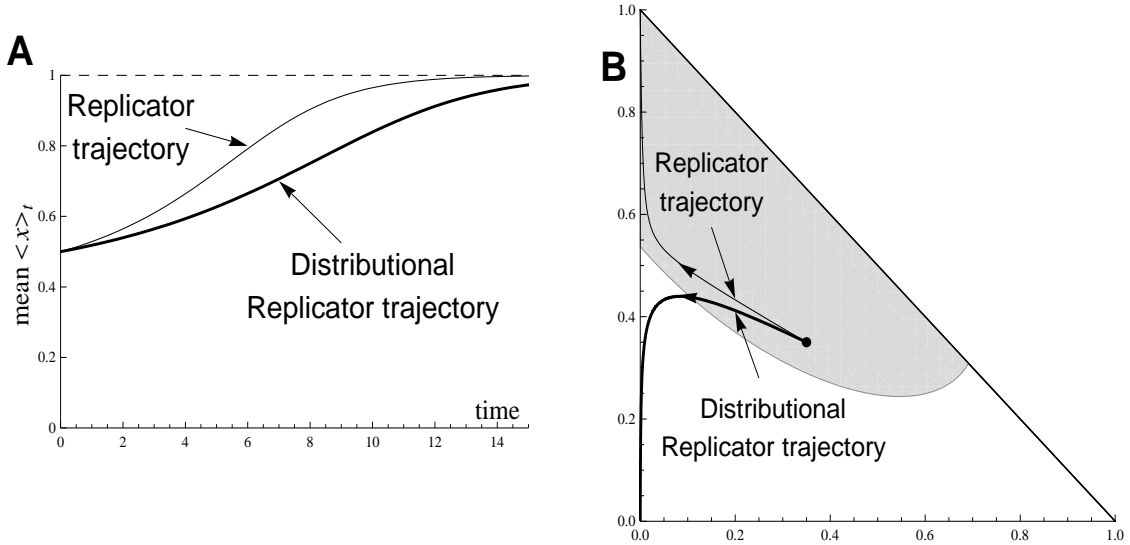$$L(x) = -\sum_i x_i^* \ln(x_i).$$

Figure 2: **A:** Trajectories of the classical replicator dynamic (71) (thin curve) starting from initial condition x_0, and of the mean $\mu_t = \langle x \rangle_t = \langle x \mid p_t \rangle$ under the dynamic (74) (thick curve), starting from an initial condition with mean, $\mu_0 = \langle x \rangle_0 = x_0$. These trajectories converge to the Nash equilibrium $x = 1$, but with different time lines. In this example, the distributional replicator dynamics has uniform initial density $p_0(x) = 1$. Other parameters are $x^* = 0.4$ and $\lambda = 1$. **B:** Trajectories of the replicator dynamics (75)-(76) (thin curve) and the mean (79) associated with the distributional replicator dynamics (77)-(78). The shaded region is the basin of attraction of $e_2 = (0, 1)$ for classical replicator trajectories. Initial conditions for both trajectories are the same: $(x_1, x_2) = (\mu_{1,0}, \mu_{2,0})$. The parameters are: $(\lambda_1, \lambda_2, \lambda_3) = (0.5, 1.12, 1.3)$; $(\alpha_1, \alpha_2, \alpha_3) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$; $(a_{11}, a_{12}, a_{13}) = (0.9, 0.05, 0.05)$; $(a_{21}, a_{22}, a_{23}) = (0.05, 0.9, 0.05)$; $(a_{31}, a_{32}, a_{33}) = (0.1, 0.1, 0.8)$.

31

11   Generic 2×2 Asymmetric Games

11.1   The Replicator Dynamic

We use the distributional replicator dynamics for asymmetric games to analyze the dynamics of the probability measures over mixed strategies for 2×2 asymmetric games. The payoff matrices to the row and column players are U and V.

Since each player has only two strategies, we denote a mixed strategy for population 1 by $(x, 1-x) \in \Delta_1$, and a mixed strategy for population 2 by $(y, 1-y) \in \Delta_2$. The standard replicator dynamics is then

$$\dot{x} = \lambda_U x(1-x)(y - y^*), \tag{80}$$

$$\dot{y} = \lambda_V y(1-y)(x - x^*), \tag{81}$$

where

$$\lambda_U = (u_{11} - u_{21}) + (u_{22} - u_{12}), \qquad y^* = \frac{u_{22} - u_{12}}{(u_{11} - u_{21}) + (u_{22} - u_{12})}, \tag{82}$$

$$\lambda_V = (v_{11} - v_{21}) + (v_{22} - v_{12}), \qquad x^* = \frac{v_{22} - v_{12}}{(v_{11} - v_{21}) + (v_{22} - v_{12})}. \tag{83}$$

The dynamics (80), (81) have equilibria at $(x, y) = (0,0), (0,1), (1,1), (1,0)$ and (x^*, y^*). The latter lies in the interior the state space $0 \leq x, y \leq 1$ provided the payoff differences $(u_{11} - u_{21})$ and $(u_{22} - u_{12})$ are non-zero and have the same sign, and similarly for $(v_{11} - v_{21})$ and $(v_{22} - v_{12})$. In particular, if these signs are all positive, then λ_U and λ_V are both positive, and in this case $(0,0)$ and $(1,1)$ are locally asymptotically stable Nash Equilibria, with all other equilibria unstable. There is therefore an equilibrium selection problem in this case. Which of the two stable Nash equilibria is the asymptotic outcome of a replicator dynamic trajectory depends on the initial condition.

In the distributional case, consider the mean entropy: $\bar{L}_t = \langle L \mid P_t \rangle$. Then using (47), we obtain

$$\bar{L}_t = -\sum_i x_i^* \int_\Delta \ln \left(\frac{e^{C_i} x_i}{e^C \cdot x} \right) P_0(dx) = -\sum_i x_i^* \int_\Delta \ln \left(\frac{\xi_i x_i}{\xi \cdot x} \right) P_0(dx)$$

$$= \langle L \mid P_0 \rangle - x^* \cdot \ln(\xi) + \int_\Delta \ln(\xi \cdot x) P_0(dx).$$

This provides a relationship between mean entropy and the Lyapunov function (65) used in the proof of theorem 7.3, namely:

$$\bar{L}_t = \bar{L}_0 + K(\xi \mid P_0).$$

In particular, it follows from the proof of Theorem 7.3 that if U is negative definite on $\mathbb{R}_0^n$, with unique equilibrium x^*, then mean entropy decreases along trajectories of the distributional replicator dynamics (64). However, for the 3×3 symmetric game discussed here, with equilibrium $x^* = e_2$, entropy $L(x)$ decreases along a trajectory of the classical replicator dynamic, but mean entropy $\bar{L}$ does not decrease along the corresponding trajectory of the distributional replicator dynamic. The latter decreases only as long as the distributional replicator trajectory stays within the shaded region in Figure 2B, but then increases after leaving this region.

We note also that neither is it the case that entropy of the mean, $L\left(\langle x \mid P_t \rangle \right)$, necessarily decreases along trajectories.

11.2 The Distributional Replicator Dynamics

The reduced distributional replicator dynamics (55) and (55) is 2-dimensional, with variables $A = A_1$ and $B = A_2$. We can therefore write these dynamics as

$$\dot{A} = \lambda_U \left\{ -y^* + F\left(e^B \mid Q_0\right) \right\}, \qquad F\left(e^B \mid Q_0\right) = \int_0^1 \left(\frac{\zeta e^B}{1 - \zeta + \zeta e^B} \right) Q_0\left(d\zeta\right), \qquad (84)$$

$$\dot{B} = \lambda_V \left\{ -x^* + F\left(e^A \mid P_0\right) \right\}, \qquad F\left(e^A \mid P_0\right) = \int_0^1 \left(\frac{\xi e^A}{1 - \xi + \xi e^A} \right) P_0\left(d\xi\right), \qquad (85)$$

with initial condition $A(0) = B(0) = 0$.

Equations (84) and (85) therefore constitute an ODE system in $\mathbb{R}^2$. We claim that, as $t \to \infty$, either $|A(t)|, |B(t)| \to \infty$ or the solution trajectories exhibit periodic orbits. These properties are summarized in the following proposition. We omit a detailed proof.

Proposition 11.1 *Let P_0 and Q_0 be the initial distributions for populations 1 and 2 respectively, which have no mass points at $x = 0$ or 1 and $y = 0$ or 1. Then,*

1. *Either: Both P_t and Q_t converge to mass points on 0 or 1 as $t \to \infty$. That is, $P_t \to \delta_0$ or δ_1, and $Q_t \to \delta_0$ or δ_1.*

2. *Or: The trajectories of P_t and Q_t are periodic.*

Proof. We show that under the dynamics (84), (85), $(|A(t)|, |B(t)|) \to (\infty, \infty)$, or the trajectories $A(t)$ and $B(t)$ describe a closed orbit in the (A, B)-plane[33]. By Proposition 5.3, if $A(t) \to \infty$, $P_t \to \delta_1$ and if $A(t) \to -\infty$, $P_t \to \delta_0$. Similarly, $Q_t \to \delta_1$ if $B(t) \to \infty$ or δ_0 if $B(t) \to -\infty$.

On the other hand, if $A(t)$ and $B(t)$ exhibit periodic motion, it follows from (47) that the trajectories of P_t and Q_t are periodic. ∎

Proposition 11.1 implies that P_t and Q_t never converge to probability measures whose aggregate state is the mixed strategy Nash equilibrium. This conclusion evokes the well known result that in 2×2 asymmetric games, a mixed strategy Nash equilibrium is never asymptotically stable under the replicator dynamic (Selten, 1980; Schuster and Sigmund, 1981; Hofbauer and Sigmund, 1981). However, unlike in 2×2 symmetric games, the convergence behaviour of the aggregate state under the replicator continuity equation does not necessarily replicate the convergence behaviour of the state variable under the replicator dynamic (80), (81). We present an example that establishes this fact. Thus, we shall construct a game in which the replicator dynamic converges to the Nash equilibrium $(1, 1)$ from given initial conditions (x_0, y_0). However, under the replicator continuity equation, and with appropriate initial distributions satisfying $(\langle x \mid P_0 \rangle, \langle y \mid Q_0 \rangle) = (x_0, y_0)$, the density functions over mixed strategies, (P_t, Q_t), converge to (δ_0, δ_0). Hence, for the aggregate state, $(\langle x \mid P_t \rangle, \langle y \mid Q_t \rangle) \to (0, 0)$.

---

[33] These assertions are proved by exploiting the Hamiltonian structure of the distributional dynamics (84), (85). In particular, this allows us to find linear asymptotes for $(A(t), B(t))$ in the case where these are not periodic. The details are lengthy but essentially straightforward.

11.3 A 2×2 asymmetric example

We aim to show that there are situations in which the means of the distributions determined by the classical replicator dynamics and the distributional replicator dynamics exhibit radically different asymptotic behaviours (cf section 10). As discussed in section 11.1, we assume that all payoff differences are positive, so that $0 < x^*, y^* < 1$ and λ_U and λ_V are both positive.

First consider the replicator equations (80) and (81) with initial conditions (x_0, y_0) satisfying $0 < x_0 < x^*$ and $y^* < y_0 < 1$. Make the time change $t \to t' = \lambda_U t$. Then (80) and (81) can be written as

$$\dot{x} = x(1-x)(y - y^*), \tag{86}$$
$$\dot{y} = \lambda y(1-y)(x - x^*). \tag{87}$$

Regarded as a distributional dynamics, these equations are associated with the initial probability distributions $P_0 = \delta_{x_0}$ and $Q_0 = \delta_{y_0}$ (Corolloary 4.2).

Now consider a distributional dynamics as follows. Assume that population 2 is initially homogeneous, in the sense that all agents use a common mixed strategy y_0. However, population 1 consists initially of two types, agents who use a mixed strategy a_0, with $0 < a_0 < x^*$, and agents who use a mixed strategy a_1, with $x^* < a_1 < 1$. The population proportions of these agents are $1 - \alpha$ and α, with $0 < \alpha < 1$. Thus, the initial distributions of the two populations are $P_0 = (1-\alpha)\delta_{a_0} + \alpha\delta_{a_1}$ for population 1, and $Q_0 = \delta_{y_0}$ for population 2. We assume that the initial means of the two populations are

$$x_0 = \langle x \mid P_0 \rangle = (1-\alpha)a_0 + \alpha a_1, \qquad y_0 = \langle y \mid Q_0 \rangle. \tag{88}$$

The distributional replicator dynamics (84)-(85) associated to these initial densities are:

$$\dot{A} = \left\{ -y^* + \frac{y_0 e^B}{1 - y_0 + y_0 e^B} \right\}, \qquad A(0) = 0, \tag{89}$$
$$\dot{B} = \lambda \left\{ -x^* + (1-\alpha)\frac{a_0 e^A}{1 - a_0 + a_0 e^A} + \alpha\frac{a_1 e^A}{1 - a_1 + a_1 e^A} \right\}, \qquad B(0) = 0. \tag{90}$$

We wish to compare these dynamics to those associated with the replicator dynamics (86)-(87) having initial conditions (x_0, y_0). We show that parameters can be chosen so that the replicator trajectory satisfies $(x_t, y_t) \to (1, 1)$, whereas the trajectory of means with the same initial conditions (88) satisfies $(\langle x \mid P_t \rangle, \langle y \mid Q_t \rangle) \to (0, 0)$ as $t \to \infty$. A numerical example illustrated in Figure 3.

In fact, it can be shown that this example belongs to a larger class of examples. Thus, for *any* initial distributions of the form $P_0 = (1-\alpha)\delta_{a_0} + \alpha\delta_{a_1}$ and $Q_0 = \delta_{y_0}$ with fixed $(x^*, y^*) \in (0, 1)$, $0 < a_0 < x^* < a_1 < 1$, $y_0 \in (y^*, 1)$ and any $\alpha > 0$ for which $x_0 < x^*$, a positive constant $\lambda = \lambda_V/\lambda_U$ can be chosen so that the means of the trajectories of the two dynamics converge to different Nash equilibria. In particular, a_0, x_0 and a_1 can be arbitrarily close together.
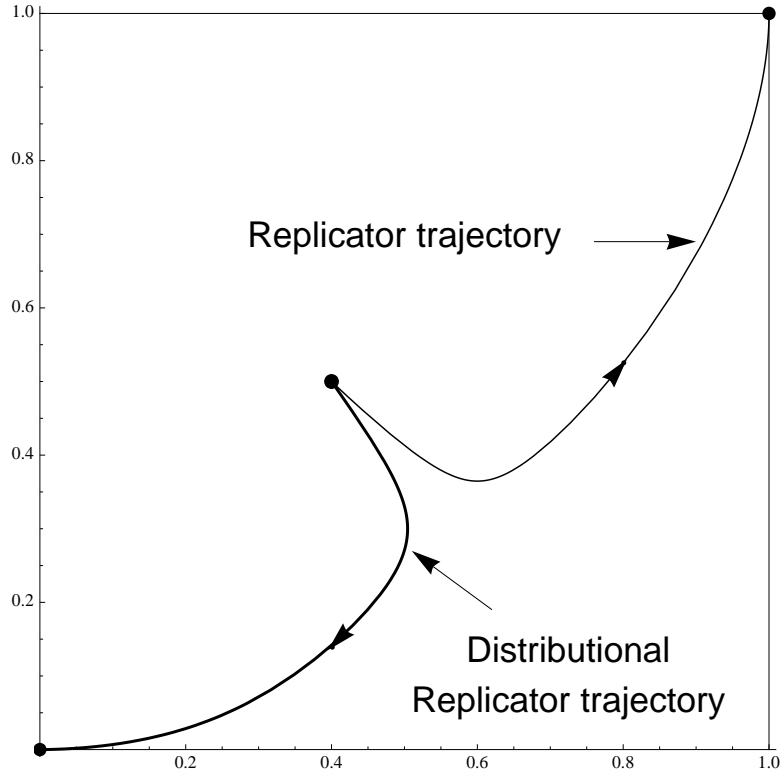
Figure 3: Trajectories of the means $(\langle x \mid P_t\rangle, \langle y \mid Q_t\rangle)$, starting from a common initial condition, $(\langle x \mid P_0\rangle, \langle y \mid Q_0\rangle) = (x_0, y_0)$, for the two dynamics (89)-(90) (thick curve) and (86)-(87) (thin curve). These trajectories converge to the Nash equilibria $(0,0)$ and $(1,1)$, respectively. Parameters are: $(x^*, y^*) = (0.6, 0.3)$, $a_0 = 0.2$, $a_1 = 0.8$, $y_0 = 0.5$, $\alpha = \frac{1}{2}\alpha^* = 0.33$, with $x_0 = (1-\alpha)a_0 + \alpha a_1 = 0.4$, and $\lambda = 5/6$.

12 Discussion and Conclusion

The motivation behind this paper was to provide more meaningful economic foundations to evolutionary game theory by allowing agents in large populations to practice a wider range of behaviors based upon their individual histories and experience. We have constructed a general theoretical framework that allows the incorporation of behavioral rules from the boundedly rational learning literature. A byproduct of this approach is that it allows a wide range of possible agent behaviors even in an environment where they may not have any knowledge of wider social characteristics such as the current social state. This is in contrast to the classical approach to evolutionary game theory where imitation of a random agent is the least informationally demanding behavior (that generates the replicator dynamics). To apply such learning procedures, we have assumed that individual agents are primed to play mixed strategies. The evolution of the probability measure over mixed strategies – the social state in our model – is determined by the general *continuity equation*, derived in Section 2, and defined for a large class of mixed strategy updating rules. We have analyzed one specific behavioral procedure – reinforcement learning – in great detail and derived the associated *replicator continuity equation* in Section 3. Although this equation cannot be solved explicitly (any more than can the classical replicator dynamic), we have proposed a general solution method using Liouville's formula and an associated finite-dimensional, autonomous ODE system that we call 'distributional dynamics', which can be applied to any finite normal form game.

We have shown that the replicator continuity dynamics admits a large class of stationary solutions; in particular, any probability distribution (or pair of distributions in the asymmetric case) whose mean over the space of mixed strategies is a Nash equilibrium. This implies that equilibrium populations can be very heterogeneous, in that different players can play different mixed strategies. However, the population is a 'mixture' of mixed strategy players, rather than of pure strategy players, as in the classical case. This constitutes a much richer behavioral structure than is usually considered. In particular, 'rationality' exists only at the aggregate mean level, with individual agents possibly exhibiting inconsistent choices even at equilibrium. We have shown that such a rich equilibrium arises for negative-definite symmetric games (Section 7), and that the unique Nash equilibrium for such a game is globally attracting for the population mean. However, although the equilibrium population mean is fixed, the equilibrium *distribution* depends on the initial distribution, and thus is a function of 'history' – in particular, on the initial distribution of agents' behavioral dispositions (mixed strategies). We have also shown that the population mean converges globally to a set of Nash equilibria in doubly symmetric games. Such results on convergence in negative definite games and doubly symmetric games are, of course, standard in both learning and evolutionary game theory. It is, however, significant that such results are obtained even when we have integrated the two approaches here. This raises the interesting possibility that other well known results from the learning literature may also be obtained in the context of population games. For example, it would be interesting to examine whether, under the Hart and Mas-Colell (2000) regret matching rule, regrets would be eliminated for all agents in the population as they are in learning in finite player games

By allowing agents to employ different mixed strategies, we have also been able to analyze the effect of heterogeneity of agent behavior in the population. In the classical approach to evolutionary game theory, agent behavior is homogeneous since all agents play the same mixed strategy at any given time (equivalently, agents play only pure strategies so that there is a given mixture of pure strategies in the population). We have focused on situations in which results from our distributional theory differ markedly from corresponding results for the standard replicator dynamics. We have constructed examples of 3×3 symmetric, and 2×2 asymmetric games in which the replicator continuity equations lead to very different predictions about the observed social state from those of the classical replicator dynamic (Sections 10 and 11.3). This shows that local stability properties of (in these cases, pure) ESS equilibria can be quite different in the distributional context.

These examples illustrate the impact of agent heterogeneity on the long run social state. It is possible to imagine an economic scenario which can be interpreted in this context. For example, consider the 2×2 asymmetric game discussed in Section 11.3. Suppose the heterogeneous population (population 1) consists of two types of consumer: either 'socially responsive' (say, more environmentally conscious), or 'indifferent' (concerned mostly with value-for-money). The homogeneous population (population 2) consists of producers having the choice of being responsive to the consumers' concerns (at additional cost) or not being so responsive. Suppose payoffs are structured so that there are two alternative pure equilibria: a 'value/profit maximizing' equilibrium, and a 'socially preferred' equilibrium (that mitigates harmful externalities). Figure 3 illustrates a situation where homogeneous behavior by indifferent consumers leads to the 'value/profit maximizing' equilibrium (the classical replicator trajectory). However, by replacing a fraction of indifferent consumers by more socially responsive consumers (leaving the initial population mean strategy unchanged), the very different 'socially preferred' equilibrium is attained.

These examples therefore illustrate that the behavioral flexibility inherent in allowing agents to play mixed strategies does have observational consequences. Even when agents play mixed strategies, it may be argued that it is only the mean of the distribution that would be observed, and by the law of large numbers, this mean would be identical to the proportion of agents playing different pure strategies. Despite this, the final observed social state can be quite different depending upon whether agents are playing mixed strategies or pure strategies. In order, therefore, to decide which particular approach—pure or mixed strategy— would be more relevant to model any particular situation, it is necessary to make appropriate assumptions about the nature of behavioural flexibility that agents may exhibit in that situation. The conventional practice of confining agents to playing pure strategies has its origins in biological models of evolution. Such an assumption may be perfectly justified in biological contexts, but seems excessively naive in models of human interaction.

Nevertheless, that agents play mixed strategies does not necessarily assume a high degree of cognitive sophistication. In particular, we do not suppose that agents consciously use randomizing devices as part of a rational calculation. We can assume instead that agents make their decisions within a largely subconscious, but stochastic internal environment (which is modulated by infor-

mation from the external environment) that offers them 'cues' that they use to condition their choice of action in the game. How this conditioning takes place depends on the agent's *behavioural disposition*, conceived simply as a function that converts the subconscious cues into actions. Thus, it is the internal stochastic environment that acts as a randomizing device, and this, together with the agent's disposition, generates a (pre-play) mixed strategy that characterizes her response when she is called upon to play. It is this disposition that is updated by reinforcement in response to payoff information (see Section 3). Though in this interpretation, as in the classical case, agents are of very limited cognitive capacity, with many (probably most) of the cues that condition the agent's action processed subconsciously, she may of course tell herself elaborate stories about why her action is the 'rational' response to the situation with which she is faced – and even convince herself that she has chosen her action on this basis.[34]

It should be possible to use the continuity equation approach to analyze mixed strategy evolution in other types of player-matching schemes than the simple pairwise-matching scheme discussed here. In this paper, a player interacts with a potentially different partner in each round of the game. However, the theory has a straightforward extension to the case in which some fixed proportion of agents are matched in each round. Alternatively, one may fix the population into matched pairs of players at the beginning, and allow these pairs to interact repeatedly using some learning protocol. The change in the distribution of mixed strategies in the populations can then be studied using a continuity equation.[35] Or one can consider a more realistic scenario of a combination of the two matching schemes—where players play with a fixed partner for a certain number of periods and then change partners. Such problems can form a substantial research agenda for the future.

Finally, we return to the pricing game used as a motivational example in the Introduction. It is easy to verify that the game is a positive definite game.[36] If we consider the case with a unique interior equilibrium and apply the classical replicator dynamic, we can argue using the intuition of the bad Rock-Paper-Scissor game that solution trajectories in the pricing game will diverge away from the equilibrium and become asymptotic to the boundary of the simplex.[37] An evolutionary analysis of the price dispersion model is credible because (a) there is a large number of sellers in such a model so that it can be analyzed as a population game; and (b) sellers can be expected to behave in a myopic fashion since in the presence of a large number of competitors, it would be unrealistic to assume that they would possess the level of rationality and knowledge required to coordinate on the exact mixed equilibrium prediction. It is therefore reasonable to expect dispersed price equilibria to be unstable. However, even this calculation is predicated on the caveat that sellers would have the knowledge of the social state; an assumption which, we have argued, is highly improbable. Our

---

[34]Of course, the analysis and conclusions of this paper are not predicated on any particular view of the mechanics of decision making. All that is required is that, somehow, such decision making can be represented by mixed strategies.

[35]In fact, for a two population, 2×2 asymmetric game in which one member of each pair is chosen from each population, the associated continuity equation is of classical, linear form.

[36]Analogous to negative definite games, a positive-definite game is defined as follows. Let U be the payoff matrix of a two player symmetric $n \times n$ game. The game is positive definite if $z \cdot Uz > 0$, for all $z \in \mathbb{R}_0^n, z \neq 0$.

[37]Lahkar (2009) analyzes this model using the logit dynamic. Under this dynamic, trajectories converge to a limit cycle in the interior of the simplex. As an explanation of observed price dispersion, the limit cycle prediction has greater empirical (Lach, 2002) and experimental support (Cason et. al., 2005).

analysis in this paper allows us to apply the continuity dynamics driven by reinforcement learning to this problem. Although we do not give a rigorous proof here, the analysis of Section 7 suggests that the mean solution trajectories would diverge from the mixed equilibrium under the replicator continuity dynamic. The instability of dispersed equilibria therefore holds in this new scenario even if sellers cannot observe the behaviour of rivals and needs to rely to his personal history of pricing strategies.

A   Appendix

A.1   The weak form of continuity equations

We work with probability measures defined on the Borel sets in Δ. Let P_t be a probability measure at time $t \geq 0$ for population 1, and Q_t be the corresponding probability measure for population 2. In the notation of section 2, set $\beta_{ij}(x) = x - \tau b_{ij}(x) : \Delta \to \Delta$. Then the updated probability measure at time $t + \tau$ for population 1 is given by equation (8), namely:

$$P_{t+\tau}(B) = \sum_{i,j \in S} \beta_{ij}(x)_i P_t \left(\beta_{ij}(B) \right) \langle y_j \mid Q_t \rangle,$$

for any Borel set B. Thus, if $\phi(x)$ is a smooth, real-valued test function, then:

$$\langle \phi \mid P_{t+\tau} \rangle = \sum_{i,j \in S} \int_{x \in \Delta_1} \phi(x) \beta_{ij}(x)_i P_t \left(d\beta_{ij}(x) \right) \langle y_j \mid Q_t \rangle.$$

Making the change of notation $\beta_{ij}(x) \to x$ (for each i and j) and using (7), we obtain

$$\langle \phi \mid P_{t+\tau} \rangle = \sum_{i,j \in S} \int_{x \in \Delta_1} \phi(x + \tau f_{ij}(x)) x_i P_t \left(dx \right) \langle y_j \mid Q_t \rangle.$$

Now Taylor expand the $\phi(\cdot)$ term up to terms of order τ:

$$\langle \phi \mid P_{t+\tau} \rangle = \sum_{i,j \in S} \int_{x \in \Delta_1} \{ \phi(x) + \tau \nabla \phi(x) \cdot f_{ij}(x)) \} x_i P_t \left(dx \right) \langle y_j \mid Q_t \rangle.$$

Noting that $\sum_{i,j} x_i \langle y_j \mid Q_t \rangle = 1$, and using (4), this can be written in the form:

$$\frac{1}{\tau} \left\{ \int_{x \in \Delta_1} \phi(x) P_{t+\tau}(dx) - \int_{x \in \Delta_1} \phi(x) P_t(dx) \right\}$$

$$= \sum_{i,j \in S} \int_{x \in \Delta_1} \nabla \phi(x) \cdot f_{ij}(x) x_i P_t(dx) \langle y_j \mid Q_t \rangle$$

$$= \int_{x \in \Delta_1} \nabla \phi(x) \cdot \left\{ \sum_{i,j \in S} x_i f_{ij}(x) \langle y_j \mid Q_t \rangle \right\} P_t(dx)$$

$$= \int_{x \in \Delta_1} \nabla \phi(x) \cdot [\mathcal{F}(x) \langle y \mid Q_t \rangle] P_t(dx).$$

Taking the limit as $\tau \to 0$ therefore gives:

$$\frac{d}{dt} \langle \phi \mid P_t \rangle = \int_{x \in \Delta_1} \nabla \phi(x) \cdot [\mathcal{F}(x) \langle y \mid Q_t \rangle] P_t(dx). \tag{91}$$

This is the *weak form of the continuity equation* for Borel probability measures, which exists provided the integral on the right exists for all $t \geq 0$. This is the case if, for example, the forward state change vectors, $f_{ij}(x)$, are continuous in x, since then $\mathcal{F}(x)$, given by (13), is also continuous, and hence bounded on Δ_1. Since $\nabla \phi(x)$ is continuous, and hence bounded, and P_t is a probability measure, it follows that the integral always exists. This shows that $\langle \phi \mid P_t \rangle$ is differentiable in t, with time-derivative given by (91).

Clearly, an analogous equation holds for population 2.

A.2 Proof of Liouville's formula

We are required to solve a weak continuity equation of the form

$$\frac{d}{dt} \langle \phi \mid P_t \rangle = \int_{x \in \Omega} \nabla \phi(x) \cdot X(x, t) P_t(dx), \tag{92}$$

where $X(x, t)$ is a smooth, time-dependent vector field on a compact, regular domain $\Omega \subset \mathbb{R}^n$, a domain which is invariant under the flow defined by X, and we are give an initial probability measure P_0 at time $t = 0$.

Consider the generalized function γ_t defined by:

$$\gamma_t(\phi) = \int_\Omega \phi(x_{0,t}(x)) P_0(dx).$$

Then

$$\frac{d\gamma_t}{dt}(\phi) = \frac{d}{dt} \int_\Omega \phi\left(x_{0,t}(x)\right) P_0\left(dx\right)$$

$$= \int_\Omega \nabla\phi(x_{0,t}(x)) \cdot \dot{x}_{0,t}(x) P_0\left(dx\right)$$

$$= \int_\Omega \nabla\phi(x_{0,t}(x)) \cdot X\left(x_{0,t}(x), t\right) P_0\left(dx\right).$$

Now apply the smooth change of variables $\xi = x_{0,t}(x)$, which has inverse $x = x_{t,0}(\xi)$. Then:

$$\frac{d\gamma_t}{dt}(\phi) = \int_\Omega \nabla\phi(\xi) \cdot X\left(\xi, t\right) P_t\left(d\xi\right),$$

where P_t is the measure defined by[38]

$$P_t(B) = P_0\left(x_{t,0}(B)\right). \tag{93}$$

We also have:

$$\gamma_t(\phi) = \int_\Omega \phi(\xi) P_t(d\xi) = \langle\phi \mid P_t\rangle.$$

This shows that P_t given by (93) is the solution of the weak form of the continuity equation (92) with the given initial measure P_0. Equation (93) is a measure-theoretic form of Liouville's formula. We also obtain expected values of smooth test functions:

$$\langle\phi \mid P_t\rangle = \int_\Omega \phi(\xi) P_t(d\xi) = \int_\Omega \phi\left(x_{0,t}(x)\right) P_0\left(dx\right). \tag{94}$$

This yields the formula (40), and hence proves Proposition 4.1.

A.2.1 Absolute continuity

Now suppose that P_0 is absolutely continuous with respect to Lebesgue measure. That is, there is a Lebesgue-integrable density function $p_0(x)$ such that $P_0(dx) = p_0(x)dx$. Then it follows from (93) that $P_t(d\xi) = P_0\left(x_{t,0}(d\xi)\right) = p_0\left(x_{t,0}(\xi)\right) dx_{t,0}(\xi)$. We also have $dx = dx_{t,0}(\xi) = |J_t(x;\xi)|d\xi$, where $J_t(x;\xi)$ is the Jacobian matrix:

$$J_t(x;\xi) = \det\left(\frac{\partial x_i}{\partial \xi_j}\right).$$

To compute this Jacobian, consider the generalized Jacobian

$$J_{t,s}(\xi) = \det\left(\frac{\partial x_{t,s}(\xi)_i}{\partial \xi_j}\right).$$

Then $J_t(x;\xi) = J_{t,0}(\xi)$, and $J_{t,t}(\xi) = 1$. Next, observe that, by definition of the trajectories $x_{t,s}(\xi)$,

---

[38]See Dunford and Schwartz (1964), Lemma 8, p 182.

we have

$$\frac{d}{ds}\left[\frac{\partial x_{t,s}(\xi)_i}{\partial \xi_j}\right] = \frac{\partial}{\partial \xi_j}\left[\frac{dx_{t,s}(\xi)_i}{ds}\right] = \frac{\partial}{\partial \xi_j}\left[X_i\left(x_{t,s}(\xi),s\right)\right] = \sum_{k=1}^{n}\frac{\partial X_i}{\partial x_k}\left(x_{t,s}(\xi),s\right)\frac{\partial x_{t,s}(\xi)_k}{\partial \xi_j}. \qquad (95)$$

Let $J_{t,s}^{(i)}(\xi)$ be the determinant of the matrix obtained from $J_{t,s}(\xi)$ by taking the time derivatives with respect to s of the entries in the i-th row, as in (95), but leaving the other rows unchanged. Let $[J_{t,s}(\xi)]_{i,j}$ be the ij-th minor of $J_{t,s}(\xi)$.[39] Then:

$$\begin{aligned}
\frac{dJ_{t,s}(\xi)}{ds} &= \sum_{i=1}^{n} J_{t,s}^{(i)}(\xi) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}(-1)^{i+j}\frac{d}{ds}\left[\frac{\partial x_{t,s}(\xi)_i}{\partial \xi_j}\right][J_{t,s}(\xi)]_{i,j} \qquad \text{expanding } J_{t,s}^{(i)}(\xi) \text{ by the } i\text{-th row} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}(-1)^{i+j}\frac{\partial X_i}{\partial x_k}\left(x_{t,s}(\xi),s\right)\frac{\partial x_{t,s}(\xi)_k}{\partial \xi_j}[J_{t,s}(\xi)]_{i,j} \qquad \text{using (95)} \\
&= \sum_{i=1}^{n}\sum_{k=1}^{n}(-1)^{i+k}\frac{\partial X_i}{\partial x_k}\left(x_{t,s}(\xi),s\right)\left\{\sum_{j=1}^{n}(-1)^{k+j}\frac{\partial x_{t,s}(\xi)_k}{\partial \xi_j}[J_{t,s}(\xi)]_{i,j}\right\} \\
&= \sum_{i=1}^{n}\sum_{k=1}^{n}(-1)^{i+k}\frac{\partial X_i}{\partial x_k}\left(x_{t,s}(\xi),s\right)\delta_{ik}J_{t,s}(\xi).
\end{aligned}$$

The last equality holds because, for $k \neq i$, the expression in $\{\}$ is the determinant of an $n \times n$ matrix whose i-th and k-th rows are identical, and hence this determinant is zero. We therefore have:

$$\frac{dJ_{t,s}(\xi)}{ds} = J_{t,s}(\xi)\sum_{i=1}^{n}\frac{\partial X_i}{\partial x_i}\left(x_{t,s}(\xi),s\right) = J_{t,s}(\xi)\left[\nabla \cdot X\right]\left(x_{t,s}(\xi),s\right).$$

Integrating this from $s = 0$ to $s = t$ and recalling that $J_{t,t}(\xi) = 1$ and $J_{t,0}(\xi) = J_t(x;\xi)$, gives:

$$|J_t(x;\xi)| = \exp\left\{-\int_0^t\left[\nabla \cdot X\right]\left(x_{t,s}(\xi),s\right)ds\right\}.$$

It now follows that P_t is absolutely continuous with respect to Lebesgue measure, with associated density function $p_t(\xi) = p_0\left(x_{t,0}(\xi)\right)|J_t(x;\xi)|$. That is:

$$p_t(\xi) = p_0\left(x_{t,0}(\xi)\right)\exp\left\{-\int_0^t\left[\nabla \cdot X\right]\left(x_{t,s}(\xi),s\right)ds\right\}.$$

This yields the probability-density function form of Liouville's formula (38).

---

[39]That is, the determinant of the $(n-1) \times (n-1)$-matrix obtained from $J_{t,s}(\xi)$ by deleting the i-th row and the j-th column.

A.3 Proof of Proposition 5.2

For the pseudo-replicator vector field $X(x,t) = R^1(x)y(t)$ on the simplex $\Delta_1 \subset \mathbb{R}^{n_1}$, we have $\sum_{i=1}^{n_1} x_i = 1$ and $\sum_{i=1}^{n_1} X_i = 0$. Hence, the independent components are x_i and X_i for $1 \leq i \leq n_1 - 1$. We therefore take the state space to be the projection of Δ_1 into $\mathbb{R}^{n_1-1}$ defined by:

$$\Omega_1 = \left\{ (x_1, \ldots, x_{n_1-1}) \in \mathbb{R}^{n_1-1} : 0 \leq x_i \leq \sum_{i=1}^{n_1-1} x_i \leq 1 \right\}. \tag{96}$$

Then, if $(x_1, \ldots, x_{n_1-1}) \in \Omega_1$, the associated point $x \in \Delta_1$ is $x = (x_1, \ldots, x_{n_1-1}, x_{n_1})$ with $x_{n_1} = 1 - \sum_{i=1}^{n_1-1} x_i$. Generally x denotes a point in Δ_1, but relevant operations often involve only the independent components, i.e. the associated point in Ω_1.

Let $L_{ij}(x) = x_i(\delta_{ij} - x_j)$. Then, from (42) we can write the divergence of X on Ω_1 as:

$$\nabla \cdot X(x,t) = \sum_{i=1}^{n_1-1} \left\{ \frac{\partial}{\partial x_i} - \frac{\partial}{\partial x_{n_1}} \right\} X_i(x,t) = \sum_{i=1}^{n_1-1} \sum_{j=1}^{n_1} \left\{ \frac{\partial}{\partial x_i} - \frac{\partial}{\partial x_{n_1}} \right\} L_{ij}(x)c_j(t),$$

for $x \in \Delta_1$. Also, for $1 \leq i, j \leq n_1 - 1$:

$$\frac{\partial}{\partial x_i} [L_{ij}(x)] = (1 - x_i)\delta_{ij} - x_j, \qquad \frac{\partial}{\partial x_i} [L_{in_1}(x)] = -x_{n_1},$$

$$\frac{\partial}{\partial x_{n_1}} [L_{ij}(x)] = 0, \qquad \frac{\partial}{\partial x_{n_1}} [L_{in_1}(x)] = -x_i.$$

Hence,

$$\begin{aligned} \nabla \cdot X(x,t) &= \sum_{i,j=1}^{n_1-1} \left\{ (1 - x_i)\delta_{ij} - x_j \right\} c_j(t) + \sum_{i=1}^{n_1-1} (x_i - x_{n_1})c_{n_1}(t) \\ &= \sum_{i=1}^{n_1} c_i(t) - n_1 \sum_{i=1}^{n_1} x_i c_i(t) \\ &= \left\{ e^1 - n_1 x \right\} \cdot c(t), \end{aligned}$$

where $e^1 = \sum_{i=1}^{n_1} e_i^1 \in \mathbb{R}^{n_1}$ is the vector all of whose entries are 1.

It now follows that, if $x_{t,s}(x)$ are the solution trajectories of the pseudo-replicator equations (44), then we obtain

$$[\nabla \cdot X](x_{t,s}(x), s) = \left\{ e^1 - n_1 x_{t,s}(x) \right\} \cdot c(s) = e^1 \cdot c(s) - n_1 \sum_{i=1}^{n_1} \frac{x_i c_i(s) e^{C_i(s,t)}}{x \cdot e^{C(s,t)}},$$

43

where $C(s, t) = C(s) - C(t)$. Thus

$$
\begin{aligned}
\int_0^t \left[\nabla \cdot X\right](x_{t,s}(x), s)\, ds &= e^1 \cdot \int_0^t c(s)ds - n_1 \sum_{i=1}^{n_1} \int_0^t \frac{x_i e^{C_i(s,t)}}{x \cdot e^{C(s,t)}} c_i(s) ds \\
&= e^1 \cdot C(t) - n_1 \int_0^t \frac{d}{ds}\left[\ln\left(x \cdot e^{C(s,t)}\right)\right] ds \\
&= e^1 \cdot C(t) + n_1 \ln\left[x \cdot e^{-C(t)}\right],
\end{aligned}
$$

because $C(t, t) = 0$, $C(0, t) = -C(t)$ and $e^1 \cdot x = 1$. We therefore have:

$$
\exp\left\{-\int_0^t \left[\nabla \cdot X\right](x_{t,s}(x), s)\, ds\right\} = \left(\frac{1}{x \cdot e^{-C(t)}}\right)^{n_1} \exp\left\{-e^1 \cdot C(t)\right\}.
$$

Substituting in Liouville's formula (38), it now follows that the solution of the weak continuity equation for density functions associated to a pseudo-replicator vector field (42) is given by (46). ∎

A.4 Proof of Lemma 7.1

By definition, $\xi = e^C/(e \cdot e^C)$, and $DF(\xi) = \left(\frac{\partial F_i(\xi)}{\partial C_j}\right)$. A calculation from (61) shows that

$$
\frac{\partial F_i(\xi)}{\partial C_j} = \int_\Delta \left(\frac{\xi_i x_i}{\xi \cdot x}\right)\left\{\delta_{ij} - \left(\frac{\xi_j x_j}{\xi \cdot x}\right)\right\} P(dx). \tag{97}
$$

Clearly DF is symmetric. Also, if $S = \mathrm{supp}(\xi)$, then $F_j(\xi) = 0$ and the j-th row and j-th column of DF are zero for $j \notin S$. However, since $\mathrm{int}\,\Delta$ has positive P-measure, DF has positive diagonal entries and negative off-diagonal entries for row and column indices $i, j \in S$. Further, from (97), we have $DF e_S = e_S \cdot DF = 0$. Hence, DF maps $\mathbb{R}^n$ onto $\mathbb{R}^n_{S0}$. A straightforward calculation from (97) now shows that, for $z \in \mathbb{R}^n_{S0}$,

$$
z \cdot DF(\xi) z = \sum_{i \in S} F_i(\xi)(z_i - \bar{z})^2, \qquad \text{where } \bar{z} = z \cdot F(\xi),
$$

and hence $DF(\xi)$ is positive-definite on $\mathbb{R}^n_{S0}$. ∎

A.5 Proof of Lemma 7.2

For $\xi \in \Delta$, consider the dynamical system

$$
\dot{\xi} = u - F(\xi \mid P) \in \mathbb{R}^n_0, \tag{98}
$$

because $u, F \in \Delta$. Also,

$$
\dot{\xi}_i\big|_{\xi_i=0} = u_i - F_i(\xi \mid P)\big|_{\xi_i=0} = u_i \geq 0,
$$

from which it follows that Δ is forward-invariant under the flow of the system (98). It now follows from standard results that Δ contains at least one equilibrium $\xi^* = \xi^*(u)$ of (98).[40] Further, it is clear from the definition of F in (61) that $\operatorname{supp}(\xi^*) = \operatorname{supp}(u)$.

It remains to show that ξ^* is unique. Suppose that u, and hence ξ^* has full support. For $\xi \in I\!\!R_+^n$, let $\zeta = \ln \xi$, and consider the potential function

$$K(\zeta \mid P) = -u \cdot \zeta + \int_\Delta \ln(e^\zeta \cdot x) P(dx).$$

Then $\nabla K = -u + F$, and hence $\nabla K(\zeta^*) = 0$, where $\zeta^* = \ln \xi^*$. Further

$$[\nabla^2 K]_{ij} = \frac{\partial^2 K}{\partial \zeta_i \partial \zeta_j} = \frac{\partial F_i}{\partial \zeta_j}. \tag{99}$$

That is $\nabla^2 K = DF$, which is positive definite on $I\!\!R_0^n$ by Lemma 7.1. Hence, ζ^* is the unique global minimum of K subject to the constraint $e \cdot e^\zeta = e \cdot \xi = 1$. Since (98) can be written as $\dot\zeta = -e^{-\zeta} \nabla K(\zeta)$, it follows that any equilibria must satisfy $\nabla K(\zeta) = 0$, and hence $\xi^* = e^{\zeta^*}$ is the unique equilibrium satisfying the constraint $\xi^* \in \Delta$.

Now suppose that u does not have full support. If $u = e_i$, then it is clear from the definition (61) that $\xi^* = e_i$ is the unique solution of $F = e_i$. So, we may suppose that the support of u contains at least two elements. If $S = \operatorname{supp}(u)$, then (99) defines an $|S| \times |S|$ symmetric matrix, $\nabla^2 K_S$, by taking $i, j \in S$. The argument of lemma 7.1 (see Appendix A.4), shows that this matrix is positive definite on $I\!\!R_{S0}^n$ (which has dimension at least 1), and hence $\xi_S^* = e^{\zeta_S^*} \in \Delta_S \subset I\!\!R_S^n$ is the unique solution of $F = u$. $\blacksquare$

References

[1] Binmore, K., Samuelson, L., 1997. Muddling Through: Noisy Equilibrium Selection. *Journal of Economic Theory*, 74, 235-265.

[2] Binmore, K., Gale, J., Samuelson, L., 1995. Learning To Be Imperfect: The Ultimatum Game. *Games and Economic Behavior*, 8, 56-90.

[3] Björnerstedt, J., Weibull, J.W., 1996. Nash equilibrium and evolution by imitation. In: Arrow, K.J., et al. (Eds.), *The Rational Foundations of Economic Behavior*. St. Martins, New York, pp. 155181.

[4] Börgers, T., Sarin, R., 1997. Learning Through Reinforcement and Replicator Dynamics. *Journal of Economic Theory*, 77, 1-14.

[5] Börgers, T., Sarin, R., 2000. Naive reinforcement learning with endogenous aspirations. *International Economic Review*, 41, 921-950.

---

[40]See, for example, Spanier (1966), Theorem 12, p197.

[6] Burdett, K., Judd, K., 1983. Equilibrium Price Dispersion. *Econometrica* 51, 955-969.

[7] Cabrales, A., 2000. Stochastic Replicator Dynamics. *International Economic Review* 41, 451-481.

[8] Camerer, C, 2003. *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton University Press, Princeton NJ.

[9] Cason, T., Friedman, D., Wagener, F., 2005. The Dynamics of Price Dispersion, or Edgeworth Variations. *Journal of Economic Dynamics & Control* 29, 801-822.

[10] Cressman, R., 2005. Stability of the replicator equation with continuous strategy space. *Math. Soc. Sci.* 50, 127-147.

[11] Cressman, R., Hofbauer, J., 2005. Measure dynamics on a one-dimensional continuous trait space: Theoretical foundations for adaptive dynamics. *Theoretical Population Biology* 67, 47-59.

[12] Cressman, R., Hofbauer, J., Riedel F., 2006. Stability of the replicator equation for a single-species with a multi-dimensional continuous trait space. *Journal of Theoretical Biology* 239, 273-288

[13] Dunford, N., Schwartz, J.T., 1964. *Linear Operators. Part I: General Theory*, 2nd ed. Interscience Publishers, Inc. New York.

[14] Erev, I, Roth, A.E., 1998. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *Am. Econ. Rev.* 88(4), 848-881.

[15] Fudenberg, D., Levine, D., 1998. *Theory of Learning in Games.* MIT Press.

[16] Gigerenzer, G., Selten, R., 2001. *Bounded Rationality. The Adaptive Toolbox.* MIT Press, Cambridge Mass.

[17] Gigerenzer, G., Todd, P.M. and the ABC Research Group, 1999. *Simple Heuristics That Make Us Smart.* Oxford University Press, New York.

[18] Hart, S., Mas-Colell, A., 2000. A Simple Adaptive Procedure leading to Correlated Equilibrium. *Econometrica* 68, 1127-1150.

[19] Hartman, P., 1964. *Ordinary Differential Equations.* John Wiley & Sons. New York, London, Sydney.

[20] Hines, W.G.S., 1980. Strategy stability in complex populations. *Journal of Applied Probability* 17, 600-610

[21] Hofbauer, J., Oechssler, J., Riedel F., 2008. Brown-von Neumann-Nash Dynamics: The Continuous Strategy Case. *Journal of Economic Theory* In Press.

[22] Hofbauer, J., Sandholm, W., 2005. Evolution in Games with Randomly Disturbed Payoffs. *Journal of Economic Theory* 132, 47-69.

[23] Hofbauer, J., Sigmund, K., 1988. *Theory of Evolution and Dynamical Systems.* Cambridge: Cambridge University Press.

[24] Hofbauer, J., Sigmund, K., 1998. *Evolutionary Games and Population Dynamics.* Cambridge: Cambridge University Press.

[25] Hopkins, E., 1998. Two Competing Models of How People Learn in Games. *Econometrica* 70, 2141-2166.

[26] Hopkins, E., Seymour, R., 2002. The Stability of Price Dispersion under Seller and Consumer Learning. *International Economic Review* 43, 1157-1190.

[27] Lahkar, R., 2009. The Dynamic Instability of Dispersed Price Equilibria. Unpublished Manuscript, IFMR, Chennai. http://ifmr.ac.in/facultyportal/ratullahkar/wp-content/uploads/2009/08/price-dispersion-draft1.pdf

[28] Lach, S., 2002. Existence and Persistence of Price Dispersion: An Empirical Analysis. *Review of Economics and Statistics* 84, 433–444.

[29] Margenau, H., Murphy, GM., 1962. *The Mathematics of Physics and Chemistry*, 2nd edition. Van Nostrand. Princeton, NJ.

[30] Maynard Smith, J., Price,G.R., 1973. The Logic of Animal Conflict; *Nature* 246, 15 – 18

[31] Maynard Smith, J. 1982. *Evolution and the Theory of Games.* Cambridge University Press, Cambridge.

[32] Oechssler, J., Riedel F., 2001. Evolutionary dynamics on infinite strategy spaces, *Economic Theory* 17, 141–162.

[33] Oechssler, J., Riedel F., 2002. On the Dynamics Foundations of Evolutionary Stability in Continuous Models. *Journal of Economic Theory* 107, 223-252.

[34] Ramsza, M., Seymour, R.M., 2009. Fictitious Play in an Evolutionary Environment. *Games and Economic Behavior.* Published online: doi:10.1016/j.geb.2009.05.003

[35] Rubinstein, A., 1998. *Modeling Bounded Rationality.* MIT Press, Cambridge Mass.

[36] Samuelson, L., 1997. *Evolutionary Games and Equilibrium Selection.* MIT Press, Cambridge Mass.

[37] Sandholm, W.H., 2005. Excess Payoff Dynamics and Other Well-Behaved Evolutionary Dynamics. *Journal of Economic Theory*, 124, 149-170.

[38] Sandholm, W.H., 2009. Population Games and Evolutionary Dynamics. University of Wisconsin-Madison. To be published by MIT Press. http://www.ssc.wisc.edu/ whs/book/pged.pdf

[39] Selten, R., 1980. A note on evolutionarily stable strategies in asymmetrical animal conflicts. *Journal of Theoretical Biology* 84, 93-101.

[40] Selten, R., 2001. What is Bounded Rationality? In: G. Gigerenzer and R. Selten (eds): *Bounded Rationality. The Adaptive Toolbox.* MIT Press, Cambridge Mass.

[41] Schlag, K., 1998. Why imitate, and if so, how? A Boundedly Rational Approach to Multi-armed Bandits. *Journal of Economic Theory* 78, 130-156.

[42] Simon, H., 1956. Rational choice and the structure of the environment. *Psychological Review* 63, 129-138.

[43] Simon, H., 1957. *Models of Man.* Wiley, New York.

[44] Simon, H., 1983. *Models of bounded rationality*, Vols 1 and 2. MIT Press, Cambridge, Mass.

[45] Spanier, E.H., 1966. *Algebraic Topology.* McGraw-Hill Book Company, New York

[46] Schuster, P., Sigmund, K., 1983. Replicator dynamics. *Journal of Theoretical Biology*, 100, 533538.

[47] Taylor, P.D., Jonker, L., 1978. Evolutionarily stable strategies and game dynamics. *Math. Biosci.* 40, 145156.

[48] Taylor, P.D., 1979. Evolutionarily stable strategies with two types of players. *Journal of Applied Probability* 16, 7683.

[49] Weibull, J. W., 1995. *Evolutionary Game Theory.* MIT Press, Cambridge.

[50] Yosida, K., 1968. *Functional Analysis*, 2nd ed. Springer-Verlag, Berlin, Heidelberg, New York.

[51] Young, H. P., 2005. *Strategic Learning and Its Limits.* Arne Ryde Memorial Lectures Series. Oxford University Press, USA.

[52] Zeeman, E.C., 1981. Dynamics of the evolution of animal conflicts; *Journal of Theoretical Biology*; 89, 249-270.