Journal of Mathematical Economics 46 (2010) 691–707

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

Gradient dynamics in population games: Some basic results

Daniel Friedman[a,*], Daniel N. Ostrov[b]

[a] Economics Department, University of California, Santa Cruz, CA 95064, United States
[b] Department of Mathematics and Computer Science, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053-0290, United States

ARTICLE INFO

Article history:
Received 29 October 2009
Received in revised form 2 August 2010
Accepted 4 August 2010
Available online 11 August 2010

JEL classification:
C73
D70

Keywords:
Population games
Gradient dynamics
Potential games

ABSTRACT

When each player in a population game continuously adjusts her action to move up the payoff gradient, then the state variable (the action distribution) obeys a nonlinear partial differential equation. We find conditions that render gradient adjustment myopically optimal and analyze two broad classes of population games. For one class, we use known results to establish the existence and uniqueness of solutions to the PDE. In some cases, these solutions exhibit shock waves or rarefaction waves. For a second class, we use a local form of Nash equilibrium to characterize the steady state solutions of the PDE and find sufficient conditions for asymptotic convergence.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

In financial markets, participants continuously adjust leverage. Their actions alter returns, which provoke further leverage adjustments and alterations of returns, sometimes leading to bubbles and crashes. In order to attract voters, politicians continuously adjust their stances on issues. These political dynamics can produce polarization or a tweedledum–tweedledee uniformity. In order to attract females, the males of many avian species evolve special plumages and dances, giving rise to evolutionary dynamics that can produce mudhens or peacocks.

These examples share a common structure: they are population games in which participants continuously adjust their strategies by moving up the payoff (or fitness) gradient. In this paper we propose a fairly general approach to such games. We show that the dynamics can exhibit shock waves and rarefaction waves and that behavior can converge to a dispersed distribution or to a clumped (pure atomic) distribution. The approach brings together previously disparate literatures in fluid dynamics and evolutionary game theory, and opens the way to new applications in biology and the social sciences.

The basic premise is that each player incrementally adjusts her strategy so as to increase her own payoff. From the player's perspective, the payoff function looks like a landscape in which she seeks to go uphill.

Nontrivial dynamics arise from the interplay between the distribution and the landscape. All players continuously adjust their strategies to increase their own fitness, so the population distribution changes. The change in the distribution alters the landscape, provoking further adjustments by the players and further shifts in the landscape. This process may or may not converge to an equilibrium. When it does converge, we shall see for a broad class of games that the resulting landscape is dominated by peaks and mesas, reminiscent of the American Southwest.

* Corresponding author.
E-mail addresses: dan@ucsc.edu (D. Friedman), dostrov@scu.edu (D.N. Ostrov).

0304-4068/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jmateco.2010.08.006

Some of our work builds on the conservation law and *Hamilton–Jacobi* literatures in partial differential equations; see Smoller (1994), Evans (1998), and Dafermos (2005) for recent textbook treatments. It also builds on evolutionary games literature, whose roots in theoretical biology include Wright (1949) and Kauffman (1993), as well as Maynard Smith and Price (1973). See Weibull (1995) for an early textbook treatment of evolutionary game dynamics and Sandholm (2011) for the most recent.

We depart from the evolutionary games literature in two major respects. First, most of that literature deals with one or more populations, each with an action set that consists of a finite number $n \geq 2$ of distinct alternative actions. With one population the state space is the $n-1$ dimensional simplex, and with several populations it is the Cartesian product of such simplices. By contrast, we consider a continuum of alternative actions, and the state space is the infinite-dimensional set of distribution functions.

Second, and perhaps more important, our dynamics are tied to the structure of the action space. A handful of recent articles (e.g., Cressman and Hofbauer, 2005; Oechssler and Riedel, 2002; Hofbauer et al., 2009) generalize standard dynamics—replicator and BNN—to continuous action spaces. However, at the level of individual players, these dynamics involve arbitrarily large jumps.[1] The ordering and topology of the action space are irrelevant.[2]

By contrast, but consistent with Darwin's dictum *Natura non facit saltum*, we assume that individual players do not jump around but rather move continuously over the landscape. They respect the ordering and topology of the real number line. As we will see, the underlying reason might be prohibitively large costs of adjustment for jumps, or simple physical or biological constraints on the evolutionary process.

Local adjustment processes for fixed landscapes have been considered by social scientists at least since Simon (1957) and Cyert and March (1963). Selten and Buchta (1998) propose non-parametric sign-preserving adjustment dynamics in a discrete-time setting. Arrow and Hurwicz (1960) explain gradient dynamics as "cautious" Cournot adjustment, argue that "it is of considerable behavioral as well as computational importance" to study their properties, and find sufficient conditions for convergence to Nash equilibrium in n-player (not population) games. Hart and Mas-Colell (2003) obtain convergence to correlated equilibrium in certain types of 2-player games in which each player responds to regret (payoff shortfalls from the ex-post best response) by adjusting mixed strategies in the simplex via a gradient process. Anderson et al. (2004) obtain convergence to logit equilibrium in n-player potential games with noisy gradient adjustment within continuous strategy spaces.

Three papers are even more closely related to ours. Sonnenschein (1982) treats gradient dynamics in the context of a specific population game, in which firms adjust product attributes as well as prices. Friedman and Ostrov (2008) analyze gradient dynamics in populations of consumers who seek to outrank others via conspicuous consumption. We will see that both of these population games are special cases of the present model, and that some of their results can be generalized. Friedman and Ostrov (2009) is a long working paper from which much of the material presented below is extracted. It also analyzes a variety of applications, including some related to the financial, political, and biological examples mentioned above.

Our presentation begins in the next section by laying out elements of the general model. For simplicity, we restrict attention to a single large population of strategically identical players with a continuous action set. Our first result justifies gradient dynamics as a rational response to quadratic adjustment costs. We then briefly discuss the effect of noise in the chosen action and derive the corresponding equation governing the evolution of all players actions. We focus on the limit of this equation as the noise goes to zero, which yields a very general first order partial differential equation that governs gradient dynamics.

The general form of the governing PDE is very broad. There is no unified theory for this general form, so we work separately with two broad classes of population games within this general form. Section 3 focuses on games whose payoff functions depend on the current distribution only through the *local* value of the cumulative distribution function or the density. Here we are able to draw on existing results for nonlinear first order PDEs to ensure the existence and uniqueness of solutions to the PDE governing gradient dynamics. To obtain uniqueness, it is crucial to recall that our first order PDE is derived from the limit as noise disappears from the second order PDEs governing actions with noise. Even some rather simple examples within this class of games have solutions that include shock waves, in which a behavior distribution that is initially continuous becomes discontinuous in finite time, and rarefaction fans, where initial discontinuities disappear immediately.

Section 4 focuses on *non-local* games whose payoffs depend on the entire distribution of behavior, but not specifically on the local value. We begin by connecting a variant on the standard equilibrium concept, Nash equilibrium, to the dynamic equilibria of our model and show that, for a broad class of non-local payoff functions, the equilibrium states, static or

[1] Replicator dynamics are explained in biology in terms of differing birth and survival rates for different strategies. The social science explanation is imitation: strategy switches are more likely to a recently observed strategy with a higher payoff. BNN dynamics have a similar explanation: only a small fraction of individuals get the opportunity to adjust during any short time interval, and the adjustments are jumps to the better replies to the current distribution.

[2] The literature on supermodular games, e.g., Vives (1990), deals with ordered action spaces, but it uses rather different mathematical techniques than ours to study the set of Nash equilibria, not adjustment processes. A strand of the theoretical biology literature called "adaptive dynamics" uses ordinary differential equations to model evolutionary dynamics in continuous action spaces under a very strong homogeneity assumption— at each point of time, almost all players employ exactly the same action; see Hofbauer and Sigmund (2003, Section 4.2).

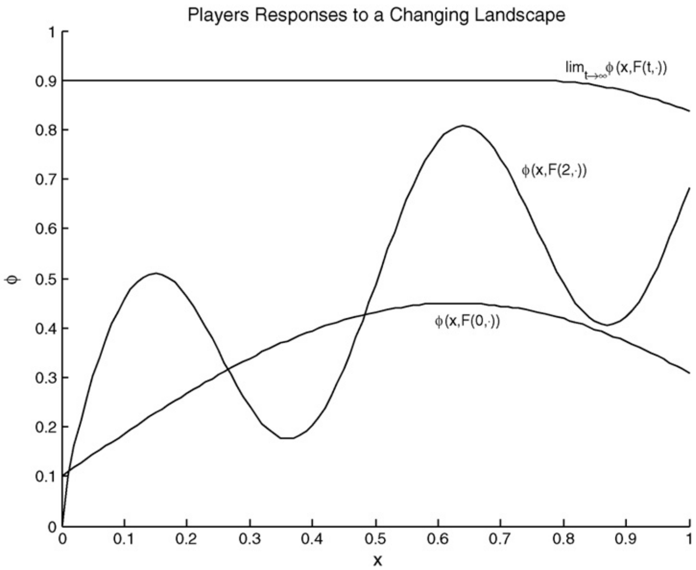

Fig. 1. Three landscapes. From initial distribution $F(0,\cdot)$, players' responses to the changing landscape could lead by time $t=2$ to a distribution $F(2,\cdot)$, which corresponds to a three peaked landscape, and ultimately to a distribution $\lim_{t\to\infty} F(t,\cdot)$, where the landscape is dominated by a mesa (flat local maximum) and a peak at the lower boundary, $x=0$.

dynamic, correspond to Southwestern landscapes. We then show how an arbitrary symmetric 2-player payoff function can be transformed into the payoff function for a non-local population game. In contrast to the case of local dynamics in Section 3, solutions in a wide class of non-local games are known to remain smooth on the interior of the action space. The rest of the section presents convergence results for a subset of such games. Extending the work of Monderer and Shapley (1996) and of Anderson et al. (2004) on *n*-player *potential* games, we show how to construct a Lyapunov function for population games obtained from symmetric 2-player potential games, and we use it to prove convergence to equilibrium sets. We sprinkle examples throughout the presentation.

2. The model

Each player has a closed strategy (or action) space $A \subset \mathbb{R} = (-\infty, \infty)$. Two leading examples are $A = [0, 1]$ and $A = \mathbb{R}$. Time is continuous, denoted by $t \in [0, \infty)$. At any particular time, the distribution of action choices is represented by a cumulative distribution function $F(t,\cdot)$, where $F(t, x)$ denotes the fraction of the population choosing strategies $y \in A$ that are less than or equal to x. The distribution $F(t,\cdot)$ encapsulates the present state of the system. It may or may not have a density $f = F_x$.

Let $\phi(x, F(t,\cdot)) \in \mathbb{R}$ be the payoff to a player using action $x \in A$ at time $t \in [0, \infty)$ when the population distribution is $F(t,\cdot)$. The "landscape" at time t is the graph of the function $x \mapsto \phi(x, F(t,\cdot))$. It changes as the distribution $F(t,\cdot)$ evolves (see Fig. 1).

That each player faces the same landscape reflects two maintained assumptions. First, we are dealing with only a single population confronting the same strategic situation. This assumption is just for simplicity; although we will not do so in this paper, there is no conceptual problem in extending the analysis to several distinct populations that interact strategically. The second maintained assumption is that each player is "small" relative to the entire population. All players face the same distribution of other players' actions, hence the same landscape, because no individual player can appreciably affect the overall distribution.

The key behavioral assumption is that each player adjusts her action x with velocity $V = \phi_x$. [3] That is, each player moves towards higher payoffs at a rate proportional to the slope of the landscape, and, through scaling, this proportionality constant is set equal to 1. Thus the state F (or f if it exists) evolves as the entire population of players moves up the payoff gradient.

2.1. Gradient dynamics as optimal adjustment

Biologists since Wright (1949) have noted that gradient dynamics arise naturally in gene frequency models and, since Lande (1976), in evolutionary models for continuous traits. Gradient dynamics are also natural in differential games in which players can only choose rates of change in their positions (see, for example, Isaacs, 1999).[4]

[3] If the gradient points outward at a boundary point of the action set A, population mass accumulates at that boundary point, as discussed further below.

[4] Of course, the differential games literature deals mainly with 2-person zero-sum games of pursuit and evasion, and not with population games.

It is less clear how to justify gradient dynamics in social science applications where players can choose to maintain the current action even when the gradient is positive or negative. Indeed, players might choose to move against the gradient, or might move with it at a disproportionate rate, or might even take a discontinuous jump. In this subsection we argue that gradient dynamics are optimal, or at least sensible, when players face common sorts of adjustment cost.

A player who adjusts her action $x \in A$ with velocity $\dot{x}(s) = V(s, x)$ over the time interval $s \in [t, t + \Delta t]$ faces *quadratic adjustment cost* if, for some constant $a > 0$, her gross payoff over that time interval is reduced by $a \int_t^{t+\Delta t} [V(s, x)]^2 ds$. Quadratic adjustment cost is quite natural in many applications. For example, in financial markets, the per-share trading cost increases approximately linearly in the net amount traded in a given short time interval (e.g., Hasbrouck, 1991), so the overall adjustment cost (net trade in shares times per-share trading cost) is quadratic.

We will say that a player facing quadratic adjustment cost is *myopically rational* if, over any short time horizon $\Delta t > 0$, she chooses an action adjustment velocity $V(s, x)$ to maximize $\phi(x, F(t + \Delta t, \cdot)) - a \int_t^{t+\Delta t} [V(s, x)]^2 ds$ under the standard population game assumption that $F(s, \cdot)$ is unaffected by the choices of any single player.

Theorem 1. At every point (t, x) in the interior of $[0, \infty) \times A$ at which $[\phi(x, F(t, \cdot))]_x$ is continuous in t and x, a myopically rational player facing quadratic adjustment costs chooses adjustment rate $V(t, x) = [1/(2a)][\phi(x, F(t, \cdot))]_x$ as the time horizon, Δt, shrinks to zero.

The proof appears in the Appendix. Of course, we can rescale time so that $[1/(2a)] = 1$. Thus myopically rational players with quadratic adjustment costs always obey gradient dynamics at interior points of A.

What happens at boundary points of a finite interval like $A = [0, 1]$? Here we impose the physical constraint $V \geq 0$ at $x = 0$ and $V \leq 0$ at $x = 1$. When the constraint does not bind, the gradient dynamics are the same as for an interior point. When the constraint binds because the gradient points outwards from the boundary point, then we enforce $V = 0$ and so mass accumulates at that boundary point.

An alternative definition of myopic rationality is that the player seeks to maximize the net return $\dot{\phi} - aV^2$, as in Sonnenschein (1982). This also yields the desired first order condition, $V(t, x) = [1/(2a)]\phi_x(x, F(t, \cdot))$, and again justifies gradient dynamics.

Myopic rationality is about as far as evolutionary games and population games usually go, since a primary concern of most of the literature (including the present paper) is how and when full optimality or static equilibrium might be attained for adjustment processes where predicting others' behavior much into the future is not possible.

There are some works that consider what happens when knowledge of future behavior is possible. For example, there is a strand of literature beginning with Matsui and Matsuyama (1995) that studies population games in order to refine the set of Nash equilibria of the corresponding stage game (usually a 2×2 bimatrix), using dynamics arising from rational, foresighted players who know that they will only be able to switch strategies at discrete random (Poisson) times.

Another example is Lasry and Lions (2007), who investigate Nash equilibria for populations of players who seek to maximize average payoff (or to minimize average expected cost) over an infinite time horizon. Their analysis requires periodic strategy spaces (so the A in their paper is the topological equivalent of a circle) and generally, to ensure ergodic behavior, requires a non-zero random noise component in players' action adjustments (corresponding to $\sigma > 0$ in Section 2.2).

In the context of our model, Friedman and Ostrov (2009) considers the case where each (fully rational) player knows the nature of the distribution into the indefinite future and chooses an adjustment path $\{V(s) : s \geq 0\}$ so as to maximize the expected present value of the net payoff stream. The paper shows that these optimal paths approximate gradient adjustment and shows how this approximation gets better as the discount rate increases.

2.2. Vanishing viscosity

In addition to the systematic effect of the gradient $V = \phi_x$, players' adjustments might also be affected by a small random noise component. To capture this possibility, one can describe the adjustment of x by the stochastic differential equation

$$dx = Vdt + \sigma dB,$$

where dB is the change in the Brownian motion, scaled by the constant volatility parameter $\sigma > 0$. The Fokker–Planck–Kolmogorov equation tells us that, for this evolution equation, the density $f(t, x)$ must be governed by the following parabolic PDE:

$$f_t = -(Vf)_x + \frac{1}{2}\sigma^2 f_{xx}. \tag{1}$$

This equation just represents conservation of population mass. It is exactly parallel to conservation of momentum in fluid dynamics, where $1/2\,\sigma^2$ is the viscosity. See the Appendix of Friedman and Ostrov (2009) for a derivation of this equation and further discussion.

The limit of equation (1) as the volatility σ goes to zero is

$$f_t = -(Vf)_x. \tag{2}$$

In fluid mechanics, this conservation of mass equation is called the "continuity equation," where "continuity" refers to the fluid *continuum*. In this context, the quantity Vf on the right hand side of (2) is called the *flux*, as it represents the flux of mass in the fluid continuum. Integrating (2) from $-\infty$ to x we obtain

$$F_t = -VF_x, \tag{3}$$

which is valid even when the density $f = F_x$ is discontinuous.

Henceforth we will refer to either (2) or (3) as *the gradient dynamics PDE*. The left hand side of (3) evaluated at a point $x \in A$ is the rate of change in the mass of players that choose actions $y \leq x$. The right hand side is the adjustment speed times population density, and represents the mass of players adjusting their actions downward (hence the minus sign) from above x to below.

Thus the gradient dynamics PDE (in either form) says that the distribution changes via continuous adjustment by individual players—no population mass is gained or lost overall, and nobody takes instantaneous jumps in their action. The fact that these equations result from letting the volatility go to zero will be crucial for establishing uniqueness of the solutions to (2) and (3) for the payoff functions examined in the next section.

3. Local dependence on the distribution

Perhaps surprisingly, the gradient dynamics PDE (2) can have solutions that become discontinuous in finite time.[5] But the existence and uniqueness (and even the meaning) of discontinuous solutions to a partial *differential* equation is unclear. In this section, we explicate these and other related issues via well-studied special cases, known to PDE researchers as *conservation laws* and *Hamilton–Jacobi equations*. For these equations $\phi_x = [\phi(x, F(t, \cdot))]_x = V$, where V takes the form $V(t, f(t, x))$ or $V(t, F(t, x))$. Clearly, for these cases, the dependence of the velocity on the distribution is local.

In a conservation law PDE, V generally takes the form $V(x, f(t, x))$, which we will focus on until Section 3.2. A simple economic example is a pricing game for a symmetrically differentiated good, where a higher price $x \in A = [0, 1]$ tends to be more profitable, but the mass of players choosing lower prices reduces payoff, as captured in the payoff function $\phi(x, F(t, \cdot)) = ax - bF(t, x)$ where $a > b > 0$. The same payoff function could represent farmers' location along a stream (now x is the distance from the source) in a dry location with good bottom land or, in a biological application, $1 - x \in A = [0, 1]$ could represent height in a jungle forest, where sunlight is a scarce resource. For this payoff function, the gradient is $[\phi(x, F(t, x))]_x = V(x, f(t, x)) = a - bf(t, x)$.

Using the *flux function* $H(x, f(t, x)) = V(x, f(t, x))f(t, x)$, the gradient dynamics PDE (2) can be written[6]

$$f_t(t, x) + [H(x, f(t, x))]_x = 0. \tag{4}$$

The solution to (4) can be discontinuous even if $H(\cdot, \cdot)$ and the initial condition are smooth. The problem can be seen clearly in the special case where $V(x, F(t, \cdot)) = f(t, x)$. Then (4) is called Burgers' equation, and it reads

$$f_t + [f^2]_x = 0. \tag{5}$$

To complete the specification, set $A = \mathbb{R}$ and consider the following initial "density"[7]:

$$f_0(x) = \begin{cases} \dfrac{1}{2} & \text{if } x \leq 0 \\ \dfrac{1-x}{2} & \text{if } 0 < x \leq 1 \\ 0 & \text{if } x > 1. \end{cases} \tag{6}$$

One common tool for seeing how solutions to (4) evolve from their initial condition is the *method of characteristics*. Friedman and Ostrov (2009, Sections 2 and 3) explain this method in detail; it is also covered in the PDE textbooks cited earlier. To use this method, one considers a curve $x = \xi(t)$ in the (t, x) plane (called a *characteristic*) that starts at a generic location on the $t = 0$ line and evolves according to $d\xi/dt = H_f$. From the chain rule, we have that $[H(x, f(t, x))]_x = H_x + H_f f_x$ and also that the evolution of f along the characteristic curve is given by $df/dt = f_t + f_x d\xi/dt$. Combining these facts with (4), we see that $df/dt = -H_x$ along the characteristic. Solving for $\xi(t)$ and $f(t, \xi(t))$ from each point on the initial condition, we get an ODE flow that forms a surface in (t, x, f) space, thereby yielding solutions to our PDE. It often aids intuition to project this flow onto the (t, x) plane, and to sketch families of characteristic curves.

A key property is that the value of f along a characteristic curve depends on the initial condition, f_0, only at $x = \xi(0)$ and nowhere else. Equivalently, information from the initial condition at any point only travels along the characteristic curve, so evolution of the solution from any specific point strictly follows the one dimensional characteristic curve—as opposed to spreading out in all directions as in the heat equation.

[5] This is ironic given that the name of this equation in fluid mechanics is the "continuity equation"!

[6] We use the letter H for the flux since, in the context of Section 3.2 below, this function is commonly known as the *Hamiltonian*.

[7] The constraint that the integral of the density must equal one clutters the example unnecessarily, so we drop it for the moment.

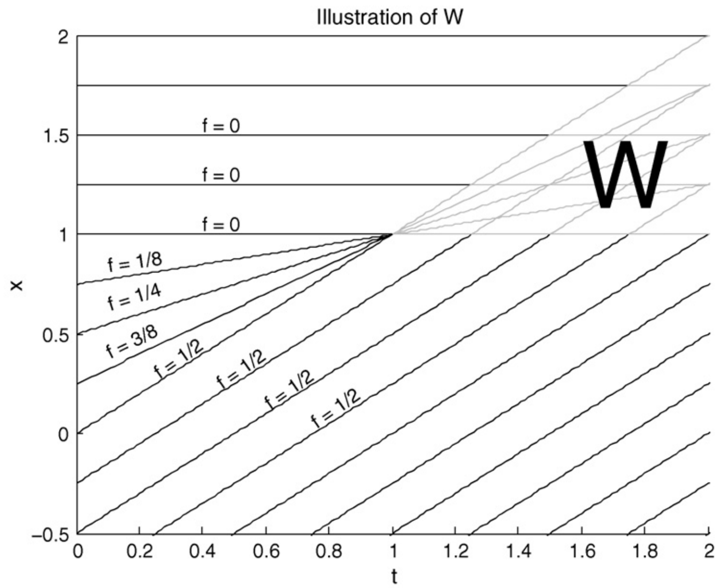

Fig. 2. Inside the wedge W, there are many different characteristics that might be followed. This generates an infinite number of candidate solutions.

For Burgers' equation, (5), the characteristic equations are

$$\frac{d\xi}{dt} = 2f(t, \xi(t)), \quad \xi(0) = x_0 \tag{7}$$

$$\frac{df(t, \xi(t))}{dt} = 0, \quad f(0, \xi(0)) = f_0(x_0). \tag{8}$$

Thus, for Burgers' equation, f stays constant on a characteristic. Therefore, the characteristics' slopes are constant, meaning the characteristics are straight lines in the (t, x) plane. For our initial condition, (6), the characteristics that emanate from the initial condition between $x = 0$ and $x = 1$ have values of f that decrease from 1/2 to 0, so their slopes decrease from 1 to 0 in such a way that all of these characteristics collide when $t = 1$ at the point $x = 1$—suggesting in some sense that $f(1, 1)$ should take all values between 0 and 1/2!

Clearly uniqueness becomes problematic inside the region labelled W in Fig. 2, where characteristics overlap. Further, all the conceivable candidate solutions are discontinuous even though our initial condition and flux function are continuous.

To resolve this non-uniqueness problem, recall that the gradient dynamics PDE (2) is the limit of (1) as the volatility parameter $\sigma \to 0$. Fortunately, initial value problems for (1) are known to have unique smooth solutions, and these solutions have a well defined limit as $\sigma \to 0$. Therefore, we focus on that limit, known as the *vanishing viscosity solution*.[8]

For our example of Burgers' equation, (1) becomes

$$f_t^\sigma + [(f^\sigma)^2]_x = \frac{1}{2}\sigma^2 f_{xx}. \tag{9}$$

For the moment, we ignore the initial condition (6). One can verify that

$$f^\sigma(x, t) = \frac{1}{2\left(1 + e^{\frac{x - \frac{1}{2}t}{\sigma^2}}\right)}$$

is a solution to (9). Clearly, f^σ is essentially equal to 1/2 when x is significantly below the $x = t/2$ line and essentially equal to 0 when x is significantly above the $x = t/2$ line. As we cross the $x = t/2$ line, f^σ smoothly transitions from 1/2 to 0, and this transition is faster for smaller $\sigma > 0$. At $t = 0$, we have the initial condition

$$f_0^\sigma(x) = \frac{1}{2(1 + e^{\frac{x}{2\sigma}})}.$$

[8] This solution is also called the entropy solution because it can also be derived by enforcing entropy, that is, by requiring that order cannot come from disorder, in a manner explained in Friedman and Ostrov (2009).

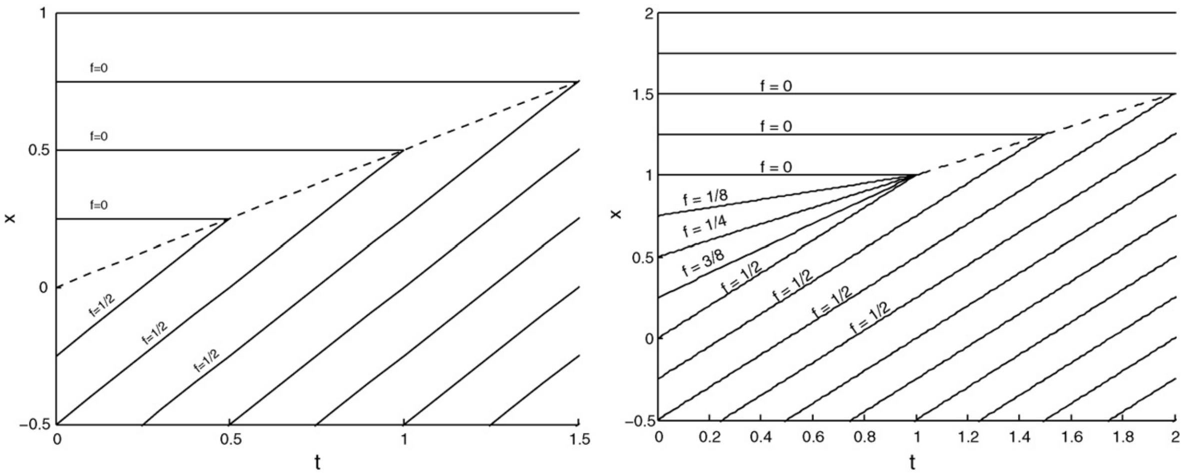

Fig. 3. Vanishing viscosity solution to Burgers' equation with initial condition (10) in Panel a, and with initial condition (6) in Panel b.

Now let $\sigma \to 0$. The initial condition clearly converges[9] to

$$f_0(x) = \begin{cases} \dfrac{1}{2} & \text{if } x < 0 \\ 0 & \text{if } x > 0, \end{cases} \tag{10}$$

and the solution f^σ clearly converges to the following discontinuous viscosity solution shown in Panel a of Fig. 3:

$$f(x,t) = \begin{cases} \dfrac{1}{2} & \text{if } x < \dfrac{t}{2} \\ 0 & \text{if } x > \dfrac{t}{2}. \end{cases} \tag{11}$$

The initial condition (10) just determined corresponds to the state of the solution at $t = 1$ for our original example with the initial condition (6)—we only need translate the condition up one unit in x. Therefore, it becomes clear how to resolve the nonuniqueness issue inside the wedge W in our original example, and we obtain the unique vanishing viscosity solution shown in Panel b of Fig. 3.

3.1. Shock discontinuities and rarefaction waves

When $\sigma > 0$, the solutions to (1) are smooth, but, as we just have seen, in the limit as $\sigma \to 0$, the solutions to (2) can fail to be smooth along curves in the (t, x) plane. These curves are called *shocks*, and we denote a generic shock curve by $x = s(t)$. Note that characteristics collide and terminate on shocks. Since the solution $f(t, x)$ is discontinuous at shocks, the meaning of the derivatives in the PDE (4) become unclear at shocks.

The answer is that the PDE is satisfied in a weak sense, called the sense of distributions, explained in our working paper and standard references. Briefly, one proceeds as follows: multiply the PDE by a smooth test function $\psi(t, x)$ with compact support and integrate, and then use integration by parts to transfer the derivatives to the test function. Then simply impose the resulting equation, which contains no derivatives of f. That is, f is a (weak or distributional) solution of the PDE (4) provided that, for any smooth test function $\psi(t, x)$ with compact support,

$$\int_0^\infty \int_A f\psi_t + H\psi_x dx\, dt + \int_A f_0 \psi(0, x) dx = 0. \tag{12}$$

It turns out that applying (12) along a shock curve $x = s(t)$ results in the Rankine–Hugoniot jump condition, which requires that the slope of the shock curve satisfy

$$\frac{ds}{dt} = \lim_{\varepsilon \to 0^+} \frac{H(s(t), f(t, s(t)+\varepsilon)) - H(s(t), f(t, s(t)-\varepsilon))}{f(t, s(t)+\varepsilon) - f(t, s(t)-\varepsilon)}. \tag{13}$$

[9] Ideally in this example we would use the same discontinuous inital condition regardless of the value of σ to generate all the (still smooth) solutions to the non-zero viscous equations. To simplify the calculations, we instead used changing inital conditions that converge to the discontinuous initial condition. See Ostrov (2007) for a rigorous demonstration that this simplification is harmless.

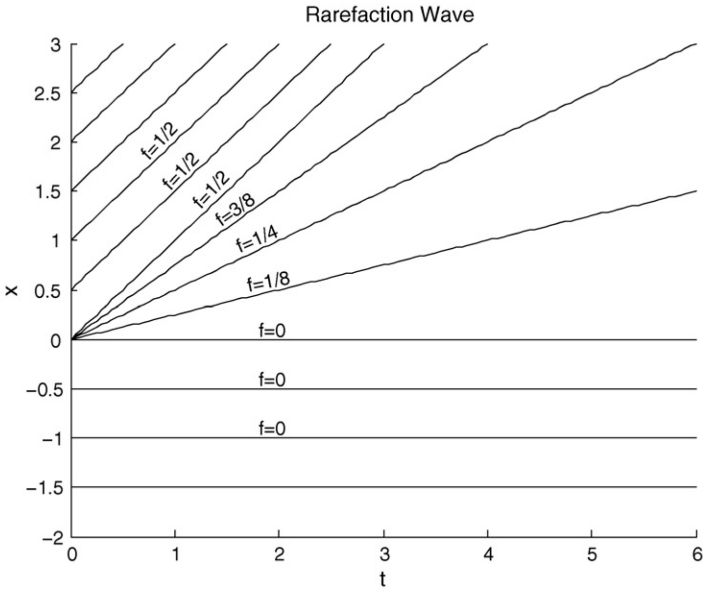

Fig. 4. Rarefaction wave for Burger's equation with initial condition (14).

For example, if we consider Burgers' equation, then $H(x, p) = p^2$, and so, for initial condition (10), the shock curve satisfies the Rankine–Hugoniot condition $ds/dt = \frac{(1/2)^2 - 0^2}{1/2 - 0} = 1/2$. Thus the shock curve starting at the point (0, 0) must be $x = s(t) = t/2$ as we see in Panel a of Fig. 3.

There is one more complication to consider before proceeding further. Locations where the initial condition is discontinuous lead immediately to either a shock or, as we now describe, a *rarefaction wave*. If the characteristics have slopes that cause them to cross, we have a shock; if the characteristics have slopes that cause them to point away from each other, we have a rarefaction wave. For example, the initial condition (10) causes an immediate shock, but if we switch the values in the initial condition so that

$$f_0(x) = \begin{cases} 0 & \text{if } x \le 0 \\ \dfrac{1}{2} & \text{if } x > 0, \end{cases} \tag{14}$$

we have characteristics with slope 0 covering the region of the (t, x) plane where $x \le 0$, so $f = 0$ in this region, and we have characteristics with slope 1 covering the region where $x > t$, so $f = 1/2$ in this region. Since these characteristics point away from each other, we have the obvious question: what happens to the solution in the seeming vacuum of information between these regions; that is, in the wedge of the (t, x) plane where $0 < x/t < 1$?

Within the wedge we look for a solution to the conservation law of the form $f(t, x) = \hat{f}(w)$ where $w = x/t$ in this wedge region. We find that by making this assumption we end up with a solution that is continuous when $t > 0$. Because the solution is continuous, it has no shocks and therefore is a vanishing viscosity solution. Burgers' equation applied to $\hat{f}(w)$ reduces to an ODE where $\hat{f}'(w)$ can be factored out. Simple algebra then yields $\hat{f}(w) = w/2$, so the full solution to Burgers' equation with initial condition (14) is

$$f(x, t) = \begin{cases} 0 & \text{if } \dfrac{x}{t} \le 0 \\ \dfrac{x}{2t} & \text{if } 0 < \dfrac{x}{t} \le 1 \\ \dfrac{1}{2} & \text{if } 1 < \dfrac{x}{t}. \end{cases}$$

The solution, illustrated in Fig. 4, is continuous when $t > 0$; there are no shocks. Within the wedge, we have a set of straight characteristics that fan out from the origin, which corresponds to information expanding out from the initial condition's discontinuity. The fan is called a *rarefaction wave* since the density of the characteristic lines decreases as time increases and the fan spreads out.

3.2. Equations for the cumulative distribution function

We now extend the analysis to two separate cases in which the gradient dynamics PDE is written in terms of the cumulative distribution function F, rather than the density f.

In the first case, F just plays the role that f played in the conservation law PDEs examined above. For example, Friedman and Ostrov (2008) analyze an application to conspicuous consumption games in the spirit of Veblen (1899/1994). A special case of their model (in which envy of other players with greater conspicuous consumption is the dominant motive) yields Burgers' equation for F instead of f:

$$F_t + \left[F^2\right]_x = 0.$$

Of course, the earlier results for f now apply to F. For example, before we had that f stays constant along a characteristic for Burgers' equation; now we have that F stays constant along a characteristic. The interpretation in the application is that, although individuals generally differ in the amount, $1 - x$, of conspicuous consumption, their rank of conspicuous consumption within the population doesn't change, except in some sense when characteristics collide and form a shock. Here a shock wave represents a positive and increasing fraction of the population choosing exactly the same amount of conspicuous consumption. The paper interprets this sort of shock as a homogeneous and growing middle class.

In the second case, we obtain the equation for F by integrating the conservation law form for f. Specifically, just as we changed the form of the gradient dynamics PDE from (2) to (3) by integrating from $-\infty$ to x, we apply this same integration to change the conservation law form $f_t(t,x) + [H(x, f(t,x))]_x = 0$ in (4) to what is called a *Hamilton–Jacobi* form $F_t(t,x) + [H(x, F_x(t,x))] = 0$, where the flux function, H, is now called the *Hamiltonian* function. This equation is essentially unchanged: information travels as before, so the characteristic curves $x = \xi(t)$ are the same and we have the same patterns of shocks and rarefaction waves. Now, however, since $F_x = f$ we have that the *derivative* of the solution, F_x, is discontinuous at the shock, as opposed to the solution being discontinuous. F remains continuous at the shock; indeed, it can be shown that the Rankine-Hugoniot jump condition in this context exactly corresponds to this requirement of continuity for F at shocks.

3.3. Finite action intervals

When the action space is a finite interval $A = [x_l, x_u]$, we need to specify boundary conditions at the two endpoints over time. Characteristics may start or may terminate at each boundary point (t, x_l) or (t, x_u). That is, a characteristic curve can originate from the point and go forward in time, or a previously existing characteristic may enter into the point on the boundary. If a characteristic starts from the point, we require continuity of the solution with the boundary condition at that point. But if characteristics only terminate at the point, it is impossible to enforce that the value of f specified by the characteristic will agree with the value given by the boundary condition. Therefore, at such a point, we do not require that the limit of the solution as we approach the boundary will equal the value specified by the boundary condition.

When we have an equation for the density, f, we must be careful to keep track of the mass accumulated at each point on the boundary. When our equation is for F, this is much clearer and easier. Since the mass is trapped between x_l and x_u, our boundary conditions are $F = 0$ at x_l and $F = 1$ at x_u.[10] Therefore, at any given time, the difference between the limit of the solution as a boundary is approached and the boundary's value, 0 or 1, indicates the amount of mass built up at that boundary.

3.4. An existence and uniqueness theorem

Let us consider the gradient dynamics PDE of the conservation form $f_t(t,x) + [H(x, f(t,x))]_x = 0$ in (4), along with an initial condition $f(0,x) = f_0(x)$. If $A = \mathbb{R}$, then from Kruskov (1970), we know there exists a unique *BV* (i.e., bounded variation) solution if the initial condition f_0 is *BV*. If A is a finite interval, then we specify boundary conditions at the two endpoints, x_l and x_u; specifically, $f(t, x_l) = f_{x_l}(t)$ and $f(t, x_u) = f_{x_u}(t)$. In this finite domain case, uniqueness and existence of a *BV* solution is known from Bardos et al. (1979), who use the following condition.

Condition 1: For any solution, $f(t,x)$, we require at points (t, x_l) on the lower boundary that

$$\min_{k \in I_l} \left[\left(H(x_l, f(t, x_l^+)) - H(x_l, k) \right) sgn(f_{x_l}(t) - f(t, x_l^+)) \right] = 0,$$

where $f(t, x_l^+) = \lim_{x \to x_l^+} f(t,x)$ and I_l is the interval $\left[\min\{0, f(t, x_l^+)\}, \max\{0, f(t, x_l^+)\}\right]$. At points (t, x_u) on the upper boundary, we require that

$$\min_{k \in I_u} \left[\left(H(x_u, k) - H(x_u, f(t, x_u^-)) \right) sgn(f_{x_u}(t) - f(t, x_u^-)) \right] = 0,$$

where $f(t, x_u^-) = \lim_{x \to x_u^-} f(t,x)$ and I_u is the interval $\left[\min\{0, f(t, x_u^-)\}, \max\{0, f(t, x_u^-)\}\right]$.

Condition 1 uses the sign function $sgn(y) = -1, 0, 1$ as $y < 0$, $= 0$, > 0. As discussed in the previous subsection, this condition enforces the boundary condition at boundary points where characteristics originate, but not where they terminate.

[10] Technically, since $F(t,x)$ is defined to be the fraction of the population at time t choosing strategies y that are less than *or equal to* x, a positive fraction of the population choosing strategy x_l implies that $F(t, x_l) \neq 0$. However, since $F(t,x) = 0$ for all $x < x_l$, this technicality is irrelevant in the PDE analysis. Therefore, we should set $F(t, x_l) = 0$ when solving the PDE, and then, after solving, we can enforce continuity from the right both at $x = x_l$ and any other points of discontinuity, should they exist, to obtain the technically correct F distribution.

We are now ready to state the existence and uniqueness theorem of Kruskov and Bardos et al. for conservation laws. The theorem imposes a condition on H_x that prevents f from blowing up in finite time as it evolves along a characteristic. The notation C^n means that the first n derivatives exist and are continuous.

Theorem 2. (Kruskov, Bardos et al.). Assume $H \in C^2$, $f_0 \in BV$, and, when A is a finite interval, assume f_{x_l} and $f_{x_u} \in C^2$. Also, assume that H_x is Lipschitz continuous with respect to f, uniformly over the region $(0, T] \times A$ in the (t, x) plane, where T is any positive time. Finally, when A is a finite interval, assume Condition 1 holds. Then the gradient dynamics PDE (4) has a unique BV vanishing viscosity solution.

The Sonnenschein (1982) model provides a non-trivial economic example. Here the state $f(t, x)$ represents the density of firms' physical locations, $x \in A = [0, 1]$, at time t. (Alternatively, x could be the products' attributes.) Assuming inelastic unit supply by each firm, the payoff function $\phi(x, f(x, t))$ is the market clearing price at location x (minus fixed cost) given a specified demand function and the current supply density f. Sonnenschein imposes the "circle" constraint (relaxed in followup work) that f and ϕ have the same values at $x = 0$ as at $x = 1$, apparently in order to avoid dealing with boundary points or with mass leaking to $\pm\infty$. Theorem 2 assures us that the gradient dynamics PDE for Sonnenschein's model has a unique BV vanishing viscosity solution. Indeed, as noted in Friedman and Ostrov (2009), the forces resisting congestion dominate behavior in Sonnenschein's model. Specifically, from an arbitrary smooth initial state, the density evolves monotonically towards the uniform distribution on A.

Theorem 2 also applies with minor modifications to the *Hamilton–Jacobi* case $F_t(t, x) + [H(x, F_x(t, x))] = 0$. Since the solution F is the integral of f, the solution space is now BUC (bounded, uniformly continuous functions) instead of BV. Extensions of these existence and uniqueness results to a much wider class of *Hamilton–Jacobi* equations are contained in Crandall and Lions (1983). Their paper introduces the notion of *viscosity solutions*, which have been used in many applications. See the text by Evans (1998) for an excellent introduction.

We close this section on a cautionary note. Although known variants of Theorem 2 cover conservation laws and *Hamilton–Jacobi* equations, they do not cover all possible payoff functions with local dependence on the distribution. For example, Friedman and Ostrov (2009) develop a population game for Bertrand pricing (or preemption) with payoff function $\phi(x, f(t, \cdot)) = (x - c)\int_x^1 f(t, y)dy$. The gradient is

$$V = [\phi(x, F(t, \cdot))]_x = 1 - F(t, x) - (x - c)f(t, x). \qquad (15)$$

The corresponding gradient dynamics PDE is a hybrid between conservation laws and *Hamilton–Jacobi* equations, and little is known about its properties, including the existence and uniqueness of solutions.

4. Non-local dependence on the distribution

We now consider the wide class of cases where $\phi(x, F(t, \cdot))$ has only non-local dependence on the distribution. For example, ϕ might depend on $F(t, \cdot)$ via its mean $\bar{x}(t) = \int_A y dF(t, y)$ or its variance or higher moments, or its order statistics, or (as in Section 4.2) the expectation of a given function with respect to the $dF(t, \cdot)$ measure.

The distinction from the class of payoff functions considered in the previous section is that here the specific choice of x does not affect ϕ through its dependence on F. In particular, it is non-local in the sense that $\phi_x = [\phi(x, F(t, \cdot))]_x = [\phi_x(x, F(t, \cdot))]$.

4.1. Equilibrium landscapes

There are two very different notions of equilibrium in dynamic economic models. The first is a steady state or, in our model, a distribution that is invariant under the gradient dynamics PDE. The other notion is a static consistency condition, such as Nash equilibrium. We shall now connect the two notions, and characterize the geometry of equilibrium landscapes. We begin with some general definitions that apply to all population games on continuous action spaces A.

Supp(F), the *support* of distribution function F, is the closure of $\{x \in A : \forall \varepsilon > 0, F(x - \varepsilon) < F(x + \varepsilon)\}$.

A *Nash equilibrium* (NE) of a population game $\phi(x, F)$ is a distribution F^* such that $\phi(x, F^*) \geq \phi(y, F^*)$ for all $y \in A$ and all $x \in Supp(F^*)$. The interpretation is that all actions actually used in NE are best responses; nobody has an incentive to try something different.

A distribution F^* is a *local Nash equilibrium* (LNE) if there is an $\varepsilon > 0$ such that, for each $x \in Supp(F^*)$, the NE condition $\phi(x, F^*) \geq \phi(y, F^*)$ holds for all $y \in A \cap (x - \varepsilon, x + \varepsilon)$. In LNE, nobody has an incentive to try something slightly different.

A distribution F^* is *a steady state* (or *invariant*) under gradient dynamics for ϕ if the solution to (3) with $V = \phi_x$ and initial condition $F_0 = F^*$ satisfies $F(t, x) = F^*(x)$ for all $t > 0$ and $x \in A$.

When the solution to (3) is not classical, we must extend the definition of steady state to deal with weak solutions as in (12). Note first that any candidate steady state $F^*(x)$ is a monotone function, so the subset of A in which it fails to be differentiable (or continuous) has measure zero. Assume ϕ_{xx} exists[11] for all $x \in A$. Then, as in Section 3, a weak solution,

[11] The results in this subsection also apply to the local dependence class of payoff functions studied in Section 3 if we assume that the gradient is a differentiable function of x. However, that is a strong assumption when $F(x)$ can contain discontinuities.

$F(t,x)$, must satisfy

$$\int_0^T \int_A F(t,x)\psi_t(t,x)dxdt + \int_A F_0(x)\psi(0,x)dx = -\int_0^T \int_A F(t,x)\left[\phi_x(t,F(t,\cdot))\psi(t,x)\right]_x dxdt + \sum_{x_0 \in \partial A} \int_0^T F(t,x_0)\phi_x(x_0,F(t,\cdot))\psi(t,x_0)\nu(x_0)dt \tag{16}$$

for any smooth test function $\psi(t,x)$ with compact support, where $\nu(x_0) = -1$ if x_0 is a lower boundary point in ∂A and $\nu(x_0) = 1$ if x_0 is an upper boundary point in ∂A, e.g., for $A = [0, 1]$, we set $\nu(0) = -1$ and $\nu(1) = 1$. Hence our general definition of a steady state is a distribution F^* such that, if $F_0 = F^*$, then each side of equation (16) equals 0 for every time $T > 0$.

The phrase *$\phi(\cdot, F)$ is maximal on $X \subset A$* means that there is an $\epsilon > 0$ such that for any $x \in X$, the NE condition $\phi(x, F(\cdot)) \geq \phi(y, F(\cdot))$ holds for all $y \in A \cap (x - \varepsilon, x + \varepsilon)$. Note that if ϕ_x exists, being maximal implies that $\phi_x = 0$ on X. Intuitively, maximality is a stability condition—population mass should not flow away from X—although we will not attempt to formalize that intuition in this paper.

A distribution function $F(x)$ defined on A is *piecewise C^1* if its derivative is continuous except at a finite number of points $x \in A$. At these points, if any, F may fail to have a continuous derivative, fail to be differentiable or fail to be continuous. Of course, since F is a distribution, it is BV, and therefore at each $x \in A$, F has a limit from the left and a limit from the right.

Theorem 3. Assume that $F^*(x)$ is piecewise C^1 and that $\phi_{xx}(x, F^*(\cdot))$ exists at all $x \in A$. Then the following three statements are equivalent:

1. F^* is a local Nash equilibrium for ϕ.
2. $\phi(\cdot, F^*)$ is constant on every connected component of $Supp(F^*)$ and is maximal on the boundary of $Supp(F^*)$.
3. F^* is a steady state under gradient dynamics for ϕ, and $\phi(\cdot, F^*)$ is maximal on the boundary of $Supp(F^*)$.

The proof, again, is in the Appendix. The theorem assures us that LNE (hence NE) distributions are always steady states, and that they locally provide an equal payoff for every action actually employed. This generalizes the equal expected payoff property of mixed NE in standard games and, in the context of population games, generalizes the Bishop-Cannings theorem (Maynard Smith, 1982, p. 182) to continuous action spaces.

Geometrically, the result is that equilibrium landscapes consists of mesas over the support components, with peaks at isolated points of the support. That is, the landscape is flat on each support component and elsewhere drops off into canyonlands (or at least is not any higher nearby). In a NE landscape, the mesas and peaks all have the same altitude.

4.2. Potential games

Section 3 shows that local dependence in the payoff function ϕ can lead to nonlinear equations with discontinuous solutions. We now consider a class of ϕ with non-local dependence that leads to classical solutions; discontinuities can only occur on the boundary of A. These games are derived from symmetric 2-player games as follows.

Let $g : A \times A \to \mathbb{R}$ be the normal form for any symmetric 2-player game. Thus $g(x, y)$ is the payoff to a player choosing strategy $x \in A$ when the other player chooses $y \in A$. The corresponding population game payoff function ϕ is the expectation of g given the current distribution $F(t, \cdot)$ of choices of other players. That is,

$$\phi(x, F(t,\cdot)) = \int_A g(x,y)dF(t,y),$$

and consequently, for interior points of A, we have that $V = \phi_x$ is

$$V(x, F(t,\cdot)) = \int_A g_x(x,y)dF(t,y). \tag{17}$$

An interpretation is that ϕ is the average actual payoff for simultaneous play against the entire field of other players. Another interpretation is that ϕ is the expected payoff against a randomly drawn opponent from the current population distribution.

We shall assume for the rest of this section that the initial distribution and the underlying 2-player payoff function g satisfy

Assumption A (regularity): Assume that g_{xxx} exists and is bounded for all (x, y). Assume that the initial distribution $F_0 \in C^2$ is also bounded.

Friedman and Ostrov (2009) show that Assumption A ensures that the solution F to the PDE (3) using (17) is *classical*. That is, $F_t(t, x)$ and $F_x(t, x)$ are defined at all points in the interior of $A \times [0, \infty)$. (Of course, F may be discontinuous at the boundary of A if characteristics evolve into it.) Friedman and Ostrov (2009) also show that when $g(x, y)$ is discontinuous then shocks can arise from smooth initial conditions, yielding a non-classical solution.

Lemma 4.4 of Monderer and Shapley (1996) motivates the following definitions. A (symmetric 2-player) payoff function $g(x, y)$ represents a *potential game* if there is a smooth function $P : A \times A \to \mathbb{R}$ such that, for all $(a, b) \in A \times A$, we have

$$g_x(a, b) = P_x(a, b) \quad \text{and} \quad g_x(b, a) = P_y(a, b). \tag{18}$$

The function P is called a *potential function* for g. The intuition is that P provides the same local incentives for both players as g, so relative maxima of P correspond to LNE of g.

Theorem 4.5 of Monderer and Shapley (1996) tells us that the following condition is necessary and sufficient for a smooth g to represent a potential game: for all $(a, b) \in A \times A$,

$$g_{xy}(a, b) = g_{xy}(b, a). \tag{19}$$

The same theorem tells us that in this case a potential function for g is

$$P(a, b) = \int_0^1 [a g_x(sa, sb) + b g_x(sb, sa)] ds. \tag{20}$$

For example, the classic Cournot model (with constant marginal cost $m < 1$ and linear demand with unit slope and intercept) has $g(x, y) = x(1 - x - y - m)$. Here $g_{xy}(a, b) = -1 = g_{xy}(b, a)$, so (19) holds, and (20) yields $P(x, y) = (x + y)(1 - m) - (x^2 + y^2 + xy)$. It is easy to check that P satisfies (18), so this Cournot model is one example of a potential game.

When g admits a potential function P, the corresponding population game $\phi(a, F) = \int_A g(a, b) dF(b)$ has a gradient that, in view of (17) and (18), can be written

$$V(a) = \phi_x(a, F) = \int_A P_x(a, b) dF(b). \tag{21}$$

We are able to leverage this fact to show that, if $F(t, \cdot)$ is a classical solution to the gradient dynamics PDE, then

$$L(t) = \int_A \int_A P(a, b) dF(t, a) dF(t, b) \tag{22}$$

is a Lyapunov function that strictly increases with time, except when $F(t, \cdot)$ is in a set Ω, which includes all steady state distributions for P. Specifically, Ω is the set of all distributions $F(\cdot)$ such that for x in the interior of A, we have $V(x) = \int_A P_x(x, b) dF(b) = 0$ almost everywhere in the $dF(\cdot)$ measure, for x that are lower boundary points of ∂A with positive $dF(\cdot)$ measure, we have $\int_A P_x(x, b) dF(b) \leq 0$, and for x that are upper boundary points of ∂A with positive $dF(\cdot)$ measure, we have $\int_A P_x(x, b) dF(b) \geq 0$. Note from this definition that for any distribution $F \in \Omega$, the boundary conditions that we enforce on ∂A imply that $V = 0$ at all boundary points with positive $dF(\cdot)$ measure.

More formally, we have the following

Theorem 4. Given Assumption A, let the population game ϕ be derived from a symmetric two-player game g with a potential function P. Then the function $L(t)$ defined in (22) is a Lyapunov function for ϕ: we have $\dot{L}(t) > 0$ for all t where $F(t, \cdot) \notin \Omega$, and $\dot{L}(t) = 0$ for all $t > t_0$ if $F(t_0, \cdot) \in \Omega$.

The proof is in the Appendix. When, contrary to Assumption A, all the population mass is always concentrated at a single point whose location may change over time, this Lyapunov result is already known; e.g., see Eq. (6) in Anderson et al. (2004).

Lyapunov functions help in establishing asymptotic convergence. The goal is to find conditions under which the *trajectory of distributions* $\mathcal{T} = \{F(t, \cdot) : t \geq 0\}$ defined by the solution to the gradient dynamics PDE (3) must converge (in a sense still to be defined) as $t \to \infty$ to a subset of Ω.

Convergence when $A = \mathbb{R}$ can be problematic because population mass may "leak out" to $x = \pm\infty$ as $t \to \infty$. An example is the uniform distribution on $[-t, t]$; the limit as $t \to \infty$ is the constant function $F(x) = 0$, which is not in the space of cumulative distribution functions. Of course mass cannot leak from $A = [0, 1]$ or any other finite interval.

When A is unbounded, we want to establish that the family of distributions $\mathcal{T}$ is *tight*, which informally means there is a uniform bound for the entire family that prevents leakage of mass to $x = \pm\infty$. Formally, a family of distributions is tight if, for any $\varepsilon > 0$, there is a finite interval $[-R_\varepsilon, R_\varepsilon]$ such that $F(R_\varepsilon) - F(-R_\varepsilon) \geq 1 - \varepsilon$ for every distribution in the family. (See, for example, Billingsley, 1999, for more details.) To establish that $\mathcal{T}$ is tight, we will make use the following notion:

Definition: The potential function $P(\cdot, \cdot)$ defined on the domain $A \times A \subset \mathbb{R}^2$ is *proper* if A is bounded or, if A is unbounded, we have that $P(x, y) \to -\infty$ as $x^2 + y^2 \to \infty$.

With this definition in place, we have the following theorem establishing the tightness of $\mathcal{T}$:

Theorem 5. Given Assumption A, let the population game ϕ be derived from a symmetric two-player game g with a proper potential function P. Then the trajectory of distributions $\mathcal{T}$ is tight.

The proof is in the Appendix. From Prokhorov's theorem (see, for example, Billingsley, 1999), we know that when $\mathcal{T}$ is tight, it is also sequentially compact in the weak convergence topology. Specifically, every sequence in $\mathcal{T}$ contains a subsequence that converges in the Lévy–Prokhorov metric to a distribution function. By Theorem 4, every sequence in $\mathcal{T}$ with t values that go to infinity must contain a subsequence that converges in the Lévy–Prokhorov metric to a distribution function in

Ω. Therefore, since $t \mapsto F(t, \cdot)$ is continuous in the Lévy–Prokhorov metric, we have the following corollary as an immediate consequence of Theorem 4, Theorem 5, and Prokhorov's theorem:

Corollary 6. Given Assumption A, let the population game ϕ be derived from a symmetric two-player game g with a proper potential function P. Then

1. Every sequence of distributions $F(t_n, \cdot)$ with $\lim_{n\to\infty} t_n = \infty$ that comes from the solution of the gradient dynamics PDE (3) has a subsequence that weakly converges to a distribution in Ω.
2. The ω-limit set of $\mathcal{T}$ is a connected subset (in the Lévy–Prokhorov metric) of Ω.
3. If there is sequence in $\mathcal{T}$ that weakly converges to a distribution F^* that is an isolated point (in the Lévy–Prokhorov metric) within the set Ω, then $F(t, \cdot)$ weakly converges to $F^*(\cdot)$.
4. If Ω consists of a single distribution, $F^*(\cdot)$, then $F(t, \cdot)$ weakly converges to $F^*(\cdot)$.

To illustrate, suppose $A = [0, 1]$ and consider the population game ϕ arising from $g(x, y) = -a(x-y)^2$, where $a > 0$. This ϕ captures so-called agglomeration effects—it is advantageous to have a physical location x closer to others' locations y, or more generally, to move in a continuous fashion towards other players' choices. Equation (18) shows that g is its own potential function. The gradient is

$$\phi_x = V(t,x) = \int_0^1 g_x(x,y)dF(t,y) = \int_0^1 -2a(x-y)dF(t,y) = -2a[x - \bar{x}(t)], \tag{23}$$

where, by integration by parts,

$$\bar{x}(t) = \int_0^1 ydF(t,y) = 1 - \int_0^1 F(t,y)dy \tag{24}$$

is the mean of the current distribution $F(t, \cdot)$. Assuming $F_0 \in C^2$ and there is no initial mass at either endpoint (so $\lim_{x\to 0^+} F_0(x) = 0$ and $\lim_{x\to 1^-} F_0(x) = 1$) we can differentiate (24) with respect to t, apply the gradient dynamics PDE (3) and then substitute the integral of (23) to see that the mean does not change over time: $\bar{x}_t = \int_0^1 V(t,y)dF(t,y) = -2a[\bar{x}(t) - \bar{x}(t)] = 0$. The last expression in (23) shows that the gradient is positive for $x < \bar{x}$ and negative for $x > \bar{x}$. Hence population mass slides towards the initial mean $\bar{x}(0)$ from both sides. The ω-limit set, Ω, consists of distributions for which $V = 0$ almost everywhere in the dF measure, so again by (23) these are distributions that concentrate the full measure of mass at a single point $x = \bar{x} = x^* \in A$. But on a trajectory with initial mean $\bar{x}(0)$, the only distribution in Ω has all of its probability mass at $\bar{x}(0)$. Hence, by Corollary 6, as $t \to \infty$, the trajectory converges weakly to F^*, where $F^*(x) = 1$ if $x \geq \bar{x}(0)$ and is 0 otherwise.

On the other hand, let us replace g by $-g$, so our new $g = a(x-y)^2$ where $a > 0$. The economic interpretation is now a form of diversity-seeking—one's own payoff increases the farther one's location is from others. In this case, a straightforward adaptation of the preceding analysis coupled with Corollary 6 shows that the ω-limit set of $\mathcal{T}$ is in Ω, which consists of distributions whose support rests solely on three points: the two endpoints of A and the center of mass, $\bar{x}$. While not covered by the corollary, the Lyapunov function analysis can be extended to establish that convergence must be to a single distribution in Ω.

These examples come from the wider class where $g(x, y) = g(y, x)$. These are known as *games of common interest* or *team games* because all players have identical incentives to make the distribution more favorable. Games of common interest are automatically potential games since (19) holds with $P = g$. For such games, the Lyapunov function $L(t)$ is the mean payoff since

$$L(t) = \int_A \int_A g(x,y)dF(t,x)dF(t,y) = \int_A \phi(x,F)dF(t,x) = \bar{\phi}(t).$$

Theorem 4 tells us that games of common interest are *progressive* in the sense that, apart from steady states, the mean payoff increases under the gradient dynamics PDE. In progressive games, the indirect effect of changes in the landscape never overshadows the direct effect of players moving up the gradient, so the mean payoff never decreases.

As a final illustration, consider a 2-player game with potential function P in which there is a quadratic loss when the player deviates from a preferred action $x_0 \in A = \mathbb{R}$, specifically, $P(x,y) = -a\left[(x-x_0)^2 + (y-x_0)^2\right]$, where $a > 0$. Then from (21) we have $V(x) = \int_A P_x(x, y)dF(y) = -2a(x-x_0)\int_A dF(y) = -2a(x-x_0)$. By definition, Ω consists of all cdf's in x that set $V(x) = 0$ with probability 1, so Ω is the singleton $\{F^*\}$, where $F^*(x) = 1$ if $x \geq x_0$ and otherwise is 0. Obviously P is proper, and, from $P_x(x, y) = g_x(x, y)$, we have that $g_{xxx} = 0$, which is obviously bounded. Therefore, for any bounded, C^2 initial distribution, F_0, we have from the last part of Corollary 6 that as $t \to \infty$, the trajectory of distributions must converge weakly to F^*.

5. Discussion

To summarize, behind a few scattered examples in the existing literature lies a wide and underexplored set of population games. In these games, each player in a large population continuously adjusts her action within a continuum of alternatives, seeking to climb the payoff gradient. Under gradient dynamics, the current population distribution of actions changes over time, and the payoff landscape shifts until, perhaps, a steady state is achieved.

We showed that gradient dynamics arise naturally, not only from physical constraints, but also from adjustment costs. We noted that known conservation law and *Hamilton–Jacobi* results guarantee that the gradient dynamics PDE has a unique solution for a wide class of such population games. We identified the "Southwestern" payoff landscapes that characterise equilibrium for another wide class. We also obtained Lyapunov functions and convergence results for potential games.

Much work remains. On the theoretical side, one would like to characterize the broadest class of population games for which solutions to the gradient dynamics PDE exist and are unique. Otherwise put, what sorts of payoff dependence on the current action distribution might preclude a solution to the gradient dynamics PDE, e.g., might cause a blow-up in finite time? Also, within the set of games with unique solutions, for which subset does the solution converge to an asymptotically stable steady state distribution? The results so far are fragmentary.

Equally important is to extend current results to action spaces $A \subset \mathbb{R}^n$ for $n > 1$. It is encouraging that the existence and uniqueness results of Kruskov (1970) and Bardos et al. (1979) for conservation laws extend to $\mathbb{R}^n$. Despite the fact that there is no longer a one-to-one correspondence between conservation laws and *Hamilton–Jacobi* equations when $n > 1$, the viscosity solution theory of Crandall and Lions (1983) provides extensions of existence and uniqueness results to $\mathbb{R}^n$ for *Hamilton–Jacobi* equations. Results outside these classes of the gradient dynamics PDE are less explored. Another important extension is to two or more strategically distinct populations rather than just one. The current results should provide helpful guidance for such future work.

The value of such theoretical advances lies mainly in the substantive applications they allow. As hinted in Section 1, we see promising applications to financial markets, politics and biology as well as to economics, and we hope to see them further developed by researchers in these fields.

Acknowledgement

This paper was written with the support of National Science Foundation grant SES-0436509. Eric Edem provided excellent research assistance; Ralph Abraham, William Sandholm, and Joel Yellin offered helpful advice at earlier stages of the project; and Andreu Mas-Colell first got us to think about gradient dynamics. The final version of the paper benefitted considerably from the suggestions of two anonymous referees. Our heartfelt thanks to all.

Appendix A. Proofs of theorems

Theorem 1. At every point (t, x) in the interior of $[0, \infty)\times A$ at which $[\phi(x, F(t, \cdot))]_x$ is continuous in t and x, a myopically rational player facing quadratic adjustment costs chooses adjustment rate $V(t, x) = [1/(2a)][\phi(x, F(t, \cdot))]_x$ as the time horizon, Δt, shrinks to zero.

Proof. Given a velocity function $V(t)$ (where the dependence of V on the action is suppressed, since the action is also a function of t) and the assumption of myopic rationality, the net payoff for a player at time $t + \Delta t$ given that she takes action x at time t is

$$\phi\left(x + \int_t^{t+\Delta t} V(s)ds, F(t+\Delta t, \cdot)\right) - a\int_t^{t+\Delta t} [V(s)]^2 ds. \tag{25}$$

We look for a function $V(s)$ that maximizes this net payoff. First, we establish that the maximizing $V(s)$ is a constant function. Since the distribution F is unaffected by the actions of any single player, the payoff function ϕ is only affected by the choice of $V(s)$ through $\int_t^{t+\Delta t} V(s)ds$ in its first argument. For a fixed value of this integral, we consider what function $V(t)$ minimizes the running cost $\int_t^{t+\Delta t}[V(t)]^2 dt$. By Jensen's inequality we know that this running cost $\int_t^{t+\Delta t}[V(s)]^2 ds \geq (1/\Delta t)\left(\int_t^{t+\Delta t} V(s)ds\right)^2$. Since we are only considering $V(s)$ where the right hand side of this inequality is fixed, we minimize the running cost by choosing a function $V(s)$ that gives us equality in Jensen's expression, which is the constant function, $V(s) = V$.

Now we look for the constant V that maximizes

$$\phi(x + V\Delta t, F(t+\Delta t, \cdot)) - aV^2\Delta t, \tag{26}$$

the net payoff (25) for constant velocities. Differentiating the net payoff with respect to V yields the optimal value $V = \frac{1}{2a}\phi_x(x + V\Delta t, F(t+\Delta t, \cdot))$. In the limit as the time horizon, Δt, shrinks to zero, we obtain from the continuity of ϕ_x in t and x that

$$V(t, x) = \frac{1}{2a}\phi_x(x, F(t, \cdot)). \quad \square \tag{27}$$

Theorem 3. Assume that $F^*(x)$ is piecewise C^1 and that $\phi_{xx}(x, F^*(\cdot))$ exists at all $x \in A$. Then the following three statements are equivalent:

1. F^* is a local Nash equilibrium for ϕ.
2. $\phi(\cdot, F^*)$ is constant on every connected component of $Supp(F^*)$ and is maximal on the boundary of $Supp(F^*)$.
3. F^* is a steady state under gradient dynamics for ϕ, and $\phi(\cdot, F^*)$ is maximal on the boundary of $Supp(F^*)$.

Proof. Statement 1 ⇒ Statement 2: Recall that the support of an arbitrary distribution function, F, is the closure of $\{x \in A : \forall \varepsilon > 0, F(x-\varepsilon) < F(x+\varepsilon)\}$. Since $Supp(F^*)$ is closed, each connected component of $Supp(F^*)$ is a closed interval, which we generically denote $[a, b]$. By the definition of the LNE, we have for some fixed $\varepsilon > 0$ that $\phi(a, F^*) \geq \phi(x, F^*)$ for all $x \in [a, \min\{a+\varepsilon, b\}]$. Since each of these x values are in $Supp(F^*)$, the definition of the LNE also tells us that that $\phi(x, F^*) \geq \phi(a, F^*)$. Therefore ϕ is constant for all these x values. If $b < a+\varepsilon$, then we are done. Otherwise, replace a with $a+\varepsilon$, repeat the argument, and conclude that ϕ is constant on the interval $[a, \min\{a+2\varepsilon, b\}]$. Further repetition of the argument tells us that ϕ is constant on $[a, \min\{a+n\varepsilon, b\}]$, where n is an arbitrarily large positive integer, hence on all of $[a, b]$. Simple extension of this argument covers the semi-infinite and infinite interval cases $(-\infty, b]$, $[a, \infty)$ and $(-\infty, \infty)$. Finally, the definition of LNE immediately implies that ϕ is maximal on the boundary of $Supp(F^*)$.

Statement 2 ⇒ Statement 3: We first show statement 2 implies that for each $x \in A$ either $\phi_x(x, F^*(\cdot)) = 0$ or $(F^*)'(x) = 0$. By the definition of the support of F^*, we know that the complement of $Supp(F^*)$ is an open set where $(F^*)'(x) = 0$. Since $\phi(\cdot, F^*)$ is constant on every connected component of $Supp(F^*)$, we have that $\phi_x(x, F^*(\cdot)) = 0$ for any interior point in $Supp(F^*)$. Since ϕ is maximal on the boundary of $Supp(F^*)$ and ϕ_x is assumed to exist, we must also have that $\phi_x(x, F^*(\cdot)) = 0$ at all points on the boundary of $Supp(F^*)$. Since all points in A are either in the complement, boundary or interior of $Supp(F^*)$, this establishes that either $\phi_x(x, F^*(\cdot)) = 0$ or $(F^*)'(x) = 0$ for all points in A.

Now consider the weak solution in (16). Recall we have a steady state if the expression on either side of (16) can be shown to be zero. To the right hand side of this expression, we apply integration by parts to the x integral over each connected region within A where $(F^*)'$ is defined. Given that either $\phi_x(x, F^*(\cdot)) = 0$ or $(F^*)'(x) = 0$, the resulting integral over each of these regions is zero. The contribution from the boundaries that correspond to the finite locations where $(F^*)'$ is not defined must also be shown to be zero. Since $(F^*)'$ is undefined, we must have that $\phi_x = 0$ at these locations and further, since ϕ_{xx} is defined for all $x \in A$, ϕ_x must be continuous at these locations. But this implies for any test function ψ and any generic location, x_0, in the interior of A where F^* is not differentiable, that

$$\lim_{x \to x_0^-} \left[F^*(x)\phi_x(x, F^*(\cdot))\psi(x,t)\right] = \lim_{x \to x_0^+} \left[F^*(x)\phi_x(x, F^*(\cdot))\psi(x,t)\right] = 0,$$

which implies the contribution from these interior boundaries is zero. (Note also, by the Rankine–Hugoniot condition, this implies the speed of all shocks are zero, as expected.) A similar calculation establishes there is no contribution from locations on the boundary of A where $(F^*)'$ is undefined. Since the right hand side of (16) is zero, we have a steady state.

Statement 3 ⇒ Statement 1: For a steady state, F^*, eqn (16) implies that within the interior of each of the pieces of A where $(F^*)'$ is continuous, we have that either $(F^*)' = 0$ or $\phi_x = 0$ at almost every point. Since the set of points within these interiors where neither $(F^*)'$ nor ϕ_x is zero has measure zero, each point in this set must have a sequence of points that converge to it where either $(F^*)' = 0$ for all the points in the sequence or $\phi_x = 0$ for all the points in the sequence. But both $(F^*)'$ and ϕ_x are continuous in x within these interiors, so we must have that either $(F^*)' = 0$ or $\phi_x = 0$ at each point, and therefore, $(F^*)' = 0$ or $\phi_x = 0$ at all points where $(F^*)'$ is continuous.

Now consider each of the finite number of x values where $(F^*)'$ fails to be continuous. If F^* isn't continuous at one of these points, $x = x_0$, then, since the Rankine–Hugoniot condition for Eq. (16) implies that $\left[\lim_{x \to x_0^-} F^*(x) - \lim_{x \to x_0^+} F^*(x)\right] \phi_x(x_0, F^*(\cdot))\psi(x_0, t) = 0$ holds for all test functions ψ, we must have that $\phi_x = 0$ at x_0. If F^* is continuous, but not continuously differentiable at one of these points, $x = x_0$, then x_0 cannot have a neighborhood around it where $(F^*)' = 0$ everywhere other than at $x = x_0$ since that would force $(F^*)'$ to equal zero and be continuous at x_0. Therefore, there must be a sequence of points that converges to x_0 where $\phi_x = 0$ for all the points in the sequence, and therefore, by the continuity of ϕ_x, $\phi_x = 0$ at x_0.

We have now established that at each $x \in A$ we either have that $\phi_x(x, F^*(\cdot)) = 0$ or $(F^*)'(x) = 0$. Now consider $Supp(F^*)$. By definition, an interior point, x_0, of $Supp(F^*)$ cannot be surrounded by a neighborhood of points where $(F^*)'(x) = 0$. Therefore, there is a sequence of points that converges to x_0 where, for each point in the sequence, $(F^*)'(x)$ does not exist or is not equal to 0. But then $\phi_x(x, F^*(\cdot)) = 0$ on this sequence, and so, by the continuity of ϕ_x, we have that $\phi_x = 0$ at x_0, and we can conclude that $\phi_x = 0$ at all interior points in $Supp(F^*)$. Since ϕ is maximal on the boundary of $Supp(F^*)$, there is an $\epsilon > 0$ such that for any x on this boundary, $\phi(x, F^*) \geq \phi(y, F^*)$ holds for all $y \in A \cap (x-\varepsilon, x+\varepsilon)$. Since $\phi_x = 0$ at all interior points of $Supp(F^*)$, using the same ϵ as on the boundary, we have for every x in the interior of $Supp(F^*)$ that $\phi(x, F^*) \geq \phi(y, F^*)$ if $y \in A \cap (x-\varepsilon, x+\varepsilon)$. Since $Supp(F^*)$ is the union of its interior and its boundary, we have, by definition, that F^* is a LNE for ϕ. □

Theorem 4. Given Assumption A, let the population game ϕ be derived from a symmetric two-player game g with a potential function P. Then the function $L(t)$ defined in (22) is a Lyapunov function for ϕ: we have $\dot{L}(t) > 0$ for all t where $F(t, \cdot) \notin \Omega$, and $\dot{L}(t) = 0$ for all $t > t_0$ if $F(t_0, \cdot) \in \Omega$.

Proof. We establish the theorem by showing that $\dot{L} = 2\overline{V^2}$, two times the average of the square of V, where V is the velocity from Eq. (2) or (3) and "average" means the integral in the $dF(t, \cdot)$ measure.

Since it is more straightforward, we begin with the case where $A = \mathbb{R}$ where we know the density $f(t, x)$ exists. (Further, the methods in Friedman and Ostrov (2009) are easy to extend to show that if we strengthen Assumption A to assume g_{xxxx} exists and is bounded and the initial condition, $F_0 \in C^3$, is also bounded, we have that $f(t, x)$ is differentiable in t and x.) In this case

$$L(t) = \int_A \int_A P(a, b) f(t, a) f(t, b) da db.$$

Differentiating $L(t)$ with respect to time, we have

$$\dot{L} = \int_A \int_A P(a, b)[f_t(t, a) f(t, b) + f(t, a) f_t(t, b)] da db.$$

Now we substitute $f_t = -(Vf)_x$ from Eq. (2),

$$\dot{L} = -\int_A \int_A P(a, b) f(t, b)[V(a, f) f(t, a)]_x da db - \int_A \int_A P(a, b) f(t, a)[V(b, f) f(t, b)]_x da db$$

Note by our subscript convention that the first term's x derivative corresponds to changes in a while the second term's x derivative corresponds to changes in b. Applying integration by parts, we obtain

$$\dot{L} = \int_A \int_A P_x(a, b) f(t, b) V(a, f) f(t, a) da db + \int_A \int_A P_y(a, b) f(t, a) V(b, f) f(t, b) da db. \quad (28)$$

From the definition of the potential function, we substitute $P_x(a, b) = g_x(a, b)$ and $P_y(a, b) = g_x(b, a)$, which shows the two integrals in each of the two lines of (28) are identical. Recalling our integral for V in Eq. (17), we have that (28) reduces to

$$\dot{L} = 2\int_A V^2(a, f) f(t, a) da = 2\overline{V^2} \geq 0. \quad (29)$$

Of course, in general, if A is a finite interval, there may be mass accumulating at the endpoints. We now tackle this more general case. For simplicity, we will restrict ourselves to $A = [0, 1]$. Since mass at the endpoints corresponds to f not existing, we must use forms of our equations that involve Stieltjes integrals using F. So, our Lyapunov function again takes the form given in (22):

$$L(t) = \int_0^1 \int_0^1 P(a, b) dF(t, a) dF(t, b).$$

Applying integration by parts twice and using the boundary conditions $F(t, 0) = 0$ and $F(t, 1) = 1$ yields

$$L(t) = P(1, 1) - \int_0^1 P_y(1, b) F(t, b) db - \int_0^1 P_x(a, 1) F(t, a) da - \int_0^1 \int_0^1 P_{xy}(a, b) F(t, a) F(t, b) da db.$$

Now we differentiate with respect to t and substitute $F_t = -VF_x$ from Eq. (3). Using, as before, the (now Stieltjes) integral for $V(x, F) = \int_0^1 g_x(x, y) dF(t, y)$ in (17) along with the facts that $P_x(a, b) = g_x(a, b)$ and $P_y(a, b) = g_x(b, a)$, which imply $P_{xy}(a, b) = g_{xy}(a, b) = g_{xy}(b, a)$, yields

$$\dot{L} = 2\int_0^1 g_x(a, 1) V(a, F) dF(t, a) - 2\int_0^1 \int_0^1 g_{xy}(a, b) V(a, F) F(t, b) db \; dF(t, a).$$

Applying integration by parts to the y derivative gives our expected result:

$$\dot{L} = 2\int_0^1 V(a, F) \left(\int_0^1 g_x(a, b) dF(t, b) \right) dF(t, a) = 2\overline{V^2} \geq 0. \quad \square$$

Theorem 5. Given Assumption A, let the population game ϕ be derived from a symmetric two-player game g with a proper potential function P. Then the trajectory of distributions $\mathcal{T}$ is tight.

Proof. By definition, the trajectory set of distributions $\mathcal{T} = \{F(t, \cdot) : t \geq 0\}$ is tight if, for any $\varepsilon > 0$, there is a finite interval $[-R_\varepsilon, R_\varepsilon]$ such that for all t, $F(t, R_\varepsilon) - F(t, -R_\varepsilon) \geq 1 - \varepsilon$. Informally, this just means the probabilistic mass does not escape to $\pm\infty$ as t increases.

For the case where $A = [0, 1]$, $\mathcal{T}$ is clearly tight by choosing $R_\varepsilon = 1$ for all ε.

For the case where $A = \mathbb{R}$, we need to use the fact that P is proper more carefully. Since P is proper, there exists a point (a_0, b_0) where P is maximized. Let P_{max} be this maximum value of P. Also, define $P(R)$ be the maximum value that P takes over the region $\{(x, y) | \max\{|x|, |y|\} \geq R\}$. This region is the full (x, y) plane with a square of size $2R \times 2R$ centered at $(0, 0)$ removed. Since P is proper, it follows that $P(R) \to -\infty$ as R grows.

If $\mathcal{T}$ is not tight, there must be a value $\epsilon_0 > 0$ such that, for any fixed R, there is a time, $t(R)$, where $F(t(R), R) - F(t(R), -R) < 1 - \epsilon_0$. This implies the probability of being greater than R or less than $-R$ is greater than or equal to ϵ_0. Since $\dot{L} \geq 0$, we know that $L(0) \leq L(t)$ for all times. Therefore, using the definition of L in (22) it follows that for all R large enough that $P(R) < 0$, we must have that $L(0) \leq L(t(R)) \leq \max\{0, P_{max}\} + P(R)(2 - \epsilon_0)\epsilon_0$. But this is in contradiction with the fact that $P(R) \to -\infty$ as R gets big. Therefore, when P is proper, $\mathcal{T}$ must be tight. □

References

Anderson, S., Goeree, J., Holt, C., 2004. Noisy Directional learning and the logit equilibrium. Scandinavian Journal of Economics 106 (3), 581–602.

Arrow, K., Hurwicz, L., 1960. Stability of the gradient process in *n*-person games. Journal of the Society of Industrial and Applied Mathematics 8, 280–294.

Bardos, C., le Roux, A.Y., Nédélec, J.-C., 1979. First order quasilinear equations with boundary conditions. Communications in Partial Differential Equations 4 (9), 1017–1034.

Billingsley, P., 1999. Convergence of Probability Measures. John Wiley & Sons, Inc., New York, NY.

Crandall, M.G., Lions, P.-L., 1983. Viscosity solutions of Hamilton–Jacobi equations. Transactions of the American Mathematical Society 277 (1), 1–42.

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.

Cyert, R., March, J., 1963. A Behavioral Theory of the Firm. Prentice-Hall, Englewood Cliffs, NJ.

Dafermos, C.M., 2005. Hyperbolic conservation laws in continuum physics. second edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 325. Springer–Verlag, Berlin, xx+626 pp. ISBN: 978-3-540-25452-2; 3-540-25452-8.

Evans, L.C., 1998. Partial differential equations. Graduate Studies in Mathematics, vol. 19. American Mathematical Society, Providence, RI, xviii+662 pp. ISBN: 0-8218-0772-2.

Friedman, D., Ostrov, D., 2008. Conspicuous consumption dynamics. Games and Economic Behavior 64 (1), 121–145.

Friedman, D., Ostrov, D., 2009. Gradient Dynamics and Evolving Landscapes in Population Games. Manuscript, UCSC Economics Department.

Hart, S., Mas-Colell, A., 2003. Regret-based continuous-time dynamics. Games and Economic Behavior 45, 375–394.

Hasbrouck, J., 1991. Measuring the information content of stock trades. Journal of Finance 46, 179–207.

Hofbauer, J., Sigmund, K., 2003. Evolutionary game dynamics. Bulletin of the American Mathematical Society 40 (4), 479–519.

Hofbauer, J., Oechssler, J., Riedel, F., 2009. Brown–von Neumann–Nash dynamics: the continuous strategy case. Games and Economic Behaviour 65 (March), 406–429.

Isaacs, R., 1999. Differential Games. Dover.

Kauffman, S.A., 1993. The Origins of Order: Self-organization and Selection in Evolution. Oxford University Press, New York.

Kruskov, S.N., 1970. First order quasilinear equations with several independent variables. (Russian) Mat. Sb. (N. S.) 81 (123), 228–255.

Lande, R., 1976. Natural selection and random genetic drift in pheontypic evolution. Evolution 30 (2), 314–334.

Lasry, J.M., Lions, P.L., 2007. Mean field games. Japanese Journal of Mathematics 2 (1), 229–260.

Matsui, A., Matsuyama, K., 1995. An approach to equilibrium selection. Journal of Economic Theory 65, 415–434.

Maynard Smith, J., 1982. Evolution and The Theory of Games. Cambridge University Press, New York.

Maynard Smith, J., Price, G.R., 1973. The logic of animal conflict. Nature 246, 15–18.

Monderer, D., Shapley, L., 1996. Potential games. Games and Economic Behavior 14, 124–143.

Oechssler, J., Riedel, F., 2002. On the dynamic foundation of evolutionary stability in continuous models. Journal of Economic Theory 107, 223–252.

Ostrov, D.N., 2007. Nonuniqueness for the vanishing viscosity solution with fixed initial condition in a nonstrictly hyperbolic system of conservation laws. Journal of Mathematical Analysis and Applications 335 (2), 996–1012.

Sandholm, W., 2011. Population Games and Evolutionary Dynamics. MIT Press.

Selten, R., Buchta, J., 1998. Experimental sealed bid first price auctions with directly observed bid functions. In: Budescu, I., Erev, I., Zwick, R. (Eds.), Games and Human Behavior: Essays in Honor of Amnon Rapoport. Lawrence Erlbaum Associates, Mahwah, NJ.

Simon, H., 1957. Models of Man. John Wiley, NY.

Smoller, J. Shock waves and reaction-diffusion equations. second edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 258. Springer–Verlag, New York, xxiv+632 pp. ISBN: 0-387-94259-9.

Sonnenschein, H., 1982. Price dynamics based on the adjustment of firms. American Economic Review 72 (5), 1088–1096.

Veblen, T., 1899/1994. The Theory of the Leisure Class. Penguin Books, New York, NY, USA.

Vives, X., 1990. Nash equilibrium with strategic complementarities. Journal of Mathematical Economics 19, 305–321.

Weibull, J., 1995. Evolutionary Game Theory. MIT Press.

Wright, S., 1949. Adaptation and selection. In: Jepson, G.L., Simpson, G.G., Mayr, E. (Eds.), Genetics, Paleontology and Evolution, pp. 356–389. Princeton University Press, Princeton, NJ.