Probabilistic Interpretations of Integrability for Game Dynamics*

William H. Sandholm†

April 7, 2013

Abstract

In models of evolution and learning in games, a variety of proofs of convergence rely on the assumption that the players' choice functions are integrable. This assumption does not have an obvious game-theoretic interpretation. We address this question by introducing probability models defined in terms of piecewise smooth closed curves through $\mathbb{R}^n$; these curves describe cycles in the performances of the available actions. We establish that a choice function is integrable if and only if in the probability model induced by each such curve, the rate at which players switch to a randomly drawn action is uncorrelated with a certain binary signal. The binary signal specifies whether the performance of the randomly drawn action is improving or worsening, and can also be interpreted as a signal about the performances of actions other than the one randomly drawn.

1. Introduction

Models of evolution and learning in games commonly specify a player's choice probabilities as a function of the performances of the n available actions, which are represented by a point in $\mathbb{R}^n$. For instance, consider the problem of finding consistent strategies for repeated play of a normal form game. During each period, each player chooses an action; he then receives a payoff that depends on his own action and the actions of the others. The performance of action i as of period t is measured by the player's *regret* for (not having chosen) this action, which is defined as the difference between two terms: the average payoff he would have obtained had he chosen i in periods 1 through t, and the average

*I thank George Mailath for asking the question that motivated this paper, and Larry Samuelson and two referees for helpful comments. Financial support from the National Science Foundation under Grants SES-0092145 and SES-1155135 is gratefully acknowledged.

†Department of Economics, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706, USA. e-mail: whs@ssc.wisc.edu; website: http://www.ssc.wisc.edu/~whs.

payoff he actually obtained during those periods. A repeated game strategy is *consistent* (or satisfies *no regret*) if it ensures that for any sequence of payoff realizations, the player's regrets for each of his actions become nonpositive as t approaches infinity. If each player follows a no-regret strategy, then the time average of play must converge to the set of *coarse correlated equilibria*: this concept is the generalization of correlated equilibrium obtained when players' decisions about whether to follow the proposed correlated strategy are made at the ex ante stage.

Hannan (1957) and Blackwell (1956) were the first to construct consistent repeated game strategies, and a variety of consistent and ε-consistent strategies have been found since that time.[1] These strategies typically are expressed in terms of a choice function that maps the current regrets to the mixed action to be played next period. One property used to ensure that this map generates a consistent repeated game strategy is that it be *integrable*: in other words, that it be expressible as the gradient of a scalar-valued function called a *potential function*. Fudenberg and Levine (1998) use integrable choice functions to construct ε-consistent repeated game strategies. Hart and Mas-Colell (2001) propose a general class of consistent repeated game strategies; the conditions on choice functions they use to ensure consistency are continuity, monotonicity, and integrability.

Integrability is also central to the analysis of another basic process of heuristic learning in games: stochastic fictitious play. In the classic fictitious play process of Brown (1949, 1951), each player plays optimally given the performances of the n actions, here measured by the payoffs these actions obtain against the time average of his opponents' past choices. In order to ensure that convergence of these time averages implies convergence of actual play, Fudenberg and Kreps (1993) introduced stochastic fictitious play, in which the players' optimizations are performed after the payoffs to each of his pure actions are subject to random perturbations. Hofbauer and Sandholm (2002) prove that stochastic fictitious play converges to approximate Nash equilibrium in zero-sum games, potential games (Monderer and Shapley (1996)), and games with an interior evolutionarily stable strategy (Maynard Smith and Price (1973)). Their analysis relies heavily on the fact that the map from base payoffs to choice probabilities generated by the perturbed optimization problem is integrable.

Integrability has played a similar role in recent work in evolutionary game theory. Hofbauer and Sandholm (2009) consider the evolution of aggregate behavior in a class of population games called *contractive games* (or *stable games*), a class which includes zero-sum games, games with an interior ESS, and concave potential games (Sandholm (2001)) as special cases. They define population dynamics in terms of choice functions called *revision*

[1]For overviews of the heuristic learning literature, see Young (2004) and Hart (2005).

protocols, which describe the rates at which agents switch between actions. Hofbauer and Sandholm (2009) consider a variety of protocols in which the switching rates to each action are a function of the performances of the n actions. Here performance is measured using *excess payoffs*, which are the differences between each action's payoff and the the average payoff in the population. To guarantee convergence to Nash equilibrium, they require their revision protocols to satisfy both a form of monotonicity and integrability.

Monotonicity conditions, which require better performing actions to be assigned higher probabilities, at least in some average sense, are natural requirements to impose in models of learning and evolution of games. In contrast, the rationale for imposing integrability is not as clear. Integrability is useful from a technical point of view, because the potential function of an integrable choice function can be used to construct a Lyapunov function—a scalar-valued function whose value decreases over time under the learning process, at least in expectation. However, despite the usefulness of integrability for proving convergence results, the game-theoretic interpretation of the integrability of choice functions is obscure.

The aim of this paper is to provide such interpretations. We set the stage in Section 2, where we consider the case of excess payoff dynamics in contractive games in greater detail. What is known in this environment suggests that integrability should be viewed as a restriction on how the switching rate to each action depends on the performances of other actions.

We substantiate this idea by providing a probabilistic characterization of integrability. In Section 3, we introduce a class of probability models, each of which is defined in terms of a closed curve through $\mathbb{R}^n$. Points on the curve represent the performances of the n actions, and the curve itself describes a possible path for the n strategies' performances if players' behavior were to enter a cycle. The probability model specifies a random draw of a performance vector from this curve and a random draw of an action, where the probabilities that define these draws are generated from a parameterization of the curve that traverses it at constant ℓ^1 speed.

In Section 4, we establish that a choice function is integrable if in the probability model induced by every piecewise smooth closed curve, the switching rate to the randomly drawn action is uncorrelated with a certain binary signal. This signal specifies whether the performance of the randomly drawn action is improving or worsening as the curve is traversed. Thus, roughly speaking, the integrability of a choice function is equivalent to the requirement that along all performance cycles, the switching rate to the drawn action does not depend systematically on whether this action is becoming more or less appealing.[2]

To obtain another interpretation of integrability, we establish that this same binary signal can be understood as a bit of information concerning the performances of actions other than the one randomly drawn. Thus, integrability allows the rate of switching to the randomly drawn action to depend on the performances of other actions, but requires that this dependence not be too systematic in nature along performance cycles. To provide context for this interpretation of integrability, we characterize the separable choice functions using a stronger independence requirement for the same class of probability models induced by closed curves.

2. Motivation: Excess Payoff Dynamics in Contractive Games

To motivate the analysis to follow, we review the role of integrability of revision protocols in establishing convergence results for contractive games.[3]

To begin, we define population games played by a unit mass of agents. In such a game, each agent chooses a pure action from the set $S = \{1, \dots, n\}$. The aggregate behavior of these agents is described by a *population state*; this is an element of the simplex $X = \{x \in \mathbb{R}^n_+ \colon \sum_{j \in S} x_j = 1\}$, with x_j representing the proportion of agents choosing action j. We identify a *population game* with a Lipschitz continuous payoff function $F \colon X \to \mathbb{R}^n$. The scalar $F_i(x)$ represents the payoff to action i when the population state is x.

We let $\bar{F}(x) = \sum_{j \in S} x_j F_j(x)$ denote the *average payoff* obtained in the population at state x, and let $\hat{F}_i(x) = F_i(x) - \bar{F}(x)$ denote the *excess payoff* to action i relative to the population average. Finally, we call state x a *Nash equilibrium* of F if $x_i > 0$ implies that $F_i(x) \geq F_j(x)$ for all $j \in S$, so that no agent can improve his payoff by unilaterally switching actions.

An *evolutionary dynamic* for population game F is a differential equation $\dot{x} = V^F(x)$ on the simplex. This differential equation describes the aggregate consequences of individual agents' adjusting their action choices over time by means of simple myopic choice rules.[4]

Sandholm (2005), building on work by Brown and von Neumann (1950), Weibull (1996), and Hofbauer (2000), introduces the class of excess payoff dynamics. These dynamics are

[2] This requirement on switching rates is not unreasonable in the models described above, in which agents are assumed to act in a completely myopic way. It would be much less reasonable in alternative models in which agents make simple forecasts about the likely directions of change in the performances of their actions before deciding which action to play. The use of such forecasts can generate dynamics with excellent convergence properties—see Shamma and Arslan (2005) and Arslan and Shamma (2006).

[3] For background on population games and evolutionary dynamics, see Sandholm (2010).

[4] For formal analyses linking the stochastic decisions of individual agents to a deterministic description of aggregate behavior via a differential equation, see Benaïm and Weibull (2003) and Roth and Sandholm (2013).

defined in terms of a Lipschitz continuous function $\rho\colon \mathbb{R}^n \to \mathbb{R}^n_+$, called a *revision protocol*, that maps vectors of excess payoffs π to vectors of switching rates $\rho(\pi)$. Specifically, $\rho_i(\pi)$ is the rate at which a revising agent switches to action i, and this rate does not depend on the action the agent was playing when the revision opportunity arose.

The revision protocol is assumed to satisfy a payoff monotonicity condition called *acuteness*:

(1) $$\pi \cdot \rho(\pi) > 0 \text{ whenever } \pi \notin \mathbb{R}^n_-.$$

Intuitively, acuteness requires that away from Nash equilibrium, actions with high payoffs are switched to more often than actions with low payoffs, at least in an average sense.[5]

Suppose that during recurrent play of game F, all agents update their choices using revision protocol ρ, switching from their current action to action j at rate $\rho_j(\hat{F}(x))$. The resulting dynamics of aggregate behavior are described by the *excess payoff dynamic*

(2) $$\dot{x}_i = \rho_i(\hat{F}(x)) - x_i \sum_{j\in S} \rho_j(\hat{F}(x)).$$

The original dynamic of form (2), the *Brown-von Neumann-Nash* dynamic (Brown and von Neumann (1950)), is derived from the separable, semilinear protocol $\rho_j(\pi) = [\pi_j]_+$, yielding the differential equation

(3) $$\dot{x}_i = [\hat{F}_i(x)]_+ - x_i \sum_{j\in S} [\hat{F}_j(x)]_+.$$

An appealing property of the BNN dynamic is that its rest points, the points where $V^F(x) = \mathbf{0}$, are precisely the Nash equilibria of the underlying game F, a property called *Nash stationarity*. This property provides an interpretation of Nash equilibrium as a description of the aggregate behavior of agents who update their action choices using a simple myopic rules. In fact, the specific functional form of the BNN dynamic is unimportant for this result: Sandholm (2005) shows that every excess payoff dynamic satisfies Nash stationarity. This paper also shows that out of equilibrium, excess payoff dynamics satisfy a monotonicity condition called *positive correlation*:

(4) $$F(x) \cdot V^F(x) > 0 \text{ whenever } V^F(x) \neq \mathbf{0}.$$

[5]The restriction of the excess payoff vector π to the complement of $\mathbb{R}^n_-$ in condition (1) reflects two facts. First, the excess payoff vector cannot lie in $\text{int}(\mathbb{R}^n_-)$, since this would mean that all actions generate a lower than average payoff. Second, the excess payoff vector $\hat{F}(x)$ lies on $\text{bd}(\mathbb{R}^n_-)$ if and only if x is a Nash equilibrium of F: see Proposition 3.4 of Sandholm (2005).

This condition is enough to ensure that in *potential games*, that is, games satisfying $F \equiv \nabla f$ for some potential function $f\colon X \to \mathbb{R}$, excess payoff dynamics ascend the potential function and converge to Nash equilibrium from all initial conditions.[6]

Hofbauer and Sandholm (2009) (see also Hofbauer (2000)) consider the case in which F is a *contractive game* (or *stable game*): that is, a game that satisfies

$$(y - x) \cdot (F(y) - F(x)) \leq 0 \text{ for all } x, y \in X.$$

They prove that in contractive games, excess payoff dynamics (2) converge to Nash equilibrium from all initial conditions whenever the revision protocol ρ satisfies not only acuteness (1), but also *integrability*

$$\text{there exists a } C^1 \text{ function } \gamma\colon \mathbb{R}^n \to \mathbb{R} \text{ such that } \nabla\gamma(\pi) = \rho(\pi) \text{ for all } \pi \in \mathbb{R}^n. \tag{5}$$

To do so, they show that when F is a contractive game, the function $\Gamma\colon X \to \mathbb{R}_+$, defined by $\Gamma(x) = \gamma(\hat{F}(x))$, is a strict Lyapunov function for the dynamic (2): its value decreases along solution trajectories of (2), strictly so unless x is a Nash equilibrium of F.

The role of intergrability in the proof of convergence is clear: the function γ, called the *revision potential*, is used to construct the Lyapunov function Γ, the existence of which implies global convergence. However, the game-theoretic interpretation of this condition is not obvious.

We can get some sense of the meaning of integrability by considering a stronger requirement, *separability*:

$$\rho_i(\pi) \text{ is independent of } \pi_{-i} \text{ for all } \pi \in \mathbb{R}^n \text{ and } i \in S.$$

Clearly, any separable revision protocol is integrable: if we write $\rho_i(\pi) = \phi_i(\pi_i)$, then the function

$$\gamma(\pi) = \sum_{j \in S} \int_0^{\pi_j} \phi_j(p_j)\, \mathrm{d}p_j.$$

is a potential function for ρ. Moreover, separability, which is satisfied by the protocol for the BNN dynamic (3), has a clear game-theoretic interpretation: the rate of switching to action j depends only on the excess payoff to action j.

For further insight into the meaning of integrability, we consider an example in which convergence fails to occur. The setting for the example is the population game

[6] See Sandholm (2001, 2009).

$$F(x) = Ax = \begin{pmatrix} 0 & -2 & 3 \\ 3 & 0 & -2 \\ -2 & 3 & 0 \end{pmatrix} \begin{pmatrix} x_R \\ x_P \\ x_S \end{pmatrix},$$

called *good Rock-Paper-Scissors*. It is easy to verify that this game is contractive. To specify the revision protocol, fix $\varepsilon > 0$, and let $g^\varepsilon : \mathbb{R} \to \mathbb{R}$ be a continuous decreasing function that equals 1 on $(-\infty, 0]$, equals ε^2 on $[\varepsilon, \infty)$, and is linear on $[0, \varepsilon]$; then define the revision protocol ρ by

$$\begin{pmatrix} \rho_R(\pi) \\ \rho_P(\pi) \\ \rho_S(\pi) \end{pmatrix} = \begin{pmatrix} [\pi_R]_+ g^\varepsilon(\pi_S) \\ [\pi_P]_+ g^\varepsilon(\pi_R) \\ [\pi_S]_+ g^\varepsilon(\pi_P) \end{pmatrix}. \tag{6}$$

Under this protocol, the weight placed on an action is proportional to positive part of the action's excess payoff, as in the protocol for the BNN dynamic; however, this weight is only of order ε^2 if the action it beats in good RPS has an excess payoff greater than ε.

It is easy to verify that protocol (6) satisfies acuteness (1), which implies that the corresponding excess payoff dynamic satisfies Nash stationarity and positive correlation. Nevertheless, Hofbauer and Sandholm (2009) show that when $\varepsilon \leq \frac{1}{10}$, many solution trajectories of the resulting excess payoff dynamic (2) approach a limit cycle in good RPS.

The source of the cycling in this example is that under protocol (6), the rates at which agents switch to each action depend systematically on the payoffs of the *next* action in the best response cycle. This suggests that integrability of the revision protocol, which would guarantee convergence to Nash equilibrium in good RPS, is a restriction on the way that the switch rates to each action depend on the performances of other actions. Our aim in what follows is to substantiate this intuition.

3. The Model

With this background in hand, we now develop our interpretations of integrability. We do so using a general model that encompasses the various environments described in the introduction.

3.1 Integrability and closed curves

As above, we take $S = \{1, \ldots, n\}$ to be a set of actions. We define a *choice function* to be a continuous function $\rho : \mathbb{R}^n \to \mathbb{R}^n$. The input to a choice function is a vector $\pi \in \mathbb{R}^n$

describing the performances of the n actions (e.g., a vector of excess payoffs), and the output $\rho(\pi) \in \mathbb{R}^n$ describes the switching rates to the n actions.[7]

Choice function ρ is *integrable* if there exists a C^1 function $\gamma \colon \mathbb{R}^n \to \mathbb{R}$ satisfying $\nabla\gamma(\pi) = \rho(\pi)$ for all $\pi \in \mathbb{R}^n$. While integrability is commonly used to prove convergence results for game dynamics, its game-theoretic meaning is unclear. In what follows, we provide game-theoretic interpretations of integrability by introducing probability models defined in terms of closed curves through the space $\mathbb{R}^n$ of performance vectors. These curves represent paths that the actions' performances might follow if the population's behavior were to enter a cycle.

The set $C \subset \mathbb{R}^n$ is a *piecewise smooth closed curve* if it is the image of a piecewise C^1 function $q \colon [0,1] \to C$ with $q(0) = q(1)$. This q is called a *parameterization* of C. Endowing $\mathbb{R}^n$ with the ℓ^1 norm, $\|\pi\| = \sum_{i=1}^n |\pi_i|$, the *length* of curve C is

$$L = \int_0^1 \left\|\dot{q}(t)\right\| \mathrm{d}t.$$

Because C is piecewise smooth, L is finite. We henceforth assume that q is a *natural parameterization* of C: $\left\|\dot{q}(t)\right\| = L$ at all points t at which q is differentiable, so that q traverses C at a constant rate with respect to the ℓ^1 norm.

It is well-known that ρ is integrable if and only if for every piecewise smooth closed curve C, the line integral of ρ over C,

$$\oint_C \rho(\pi) \cdot \mathrm{d}\pi \equiv \int_0^1 \rho(q(t)) \cdot \dot{q}(t)\, \mathrm{d}t,$$

evaluates to 0. This characterization is the basis for our probabilistic interpretation of integrability.

3.2 Probability models from closed curves

We now introduce a probability model in which we define a random draw of a performance vector π from the closed curve C and an action i from $S = \{1, \ldots, n\}$. The vector π from the curve is drawn uniformly at random with respect to the natural parameterization. Conditional on the draw of this vector, the action i is drawn with probability proportional to its contribution to the tangent vector to C at π, so that actions whose

[7]Here and in what follows, we call π a "performance vector" to emphasize that it is an element of $\mathbb{R}^n$, but despite this terminology, π should be viewed as a point in space rather than a velocity through space. Also, the interpretation of $\rho(\pi) \in \mathbb{R}^n$ suggests that its components should be nonnegative, but this property is not needed in our analysis.

performances are changing more quickly are assigned higher probabilities. As we explain below, this ensures that the overall probability with which action i is drawn is equal to action i's contribution to the length of the curve.[8]

The definitions of all of the random variables introduced below depend on the choice of the curve C. While we suppress this dependence in our notation, it is important to keep this dependence in mind to understand the characterization theorems.

To define the probability model formally, we introduce the probability triple $(\Omega, \mathcal{F}, \mathbb{P})$, where $\Omega = [0,1] \times (0,1]$, $\mathcal{F}$ is the Borel σ-algebra, and $\mathbb{P}$ is Lebesgue measure. If we let (t,u) denote a typical element of Ω, then with some abuse of notation, we can regard the parameterization q as a random vector—that is, as a function from Ω to $\mathbb{R}^n$—with the understanding that $q(t,u) \equiv q(t)$ only depends on the first component of (t,u). Thus q represents our random draw of a performance vector.

The random draw of an action is represented by the random variable $I\colon \Omega \to S$. If $t \in [0,1]$ is a point at which q is differentiable, we define $I(t,u)$ by

$$I(t,u) = i \;\text{ when }\; u \in \left(\frac{1}{L} \sum_{j=1}^{i-1} \left|\dot{q}_j(t)\right|, \frac{1}{L} \sum_{j=1}^{i} \left|\dot{q}_j(t)\right| \right].$$

If not, $I(t,u)$ is defined in an arbitrary way. Notice that the probability with which action i is drawn is proportional to the total variation in π_i along curve C, which is action i's contribution to the length of the curve:

$$\mathbb{P}(I = i) = \mathbb{P}\left(\left\{ (t,u)\colon u \in \left(\frac{1}{L} \sum_{j=1}^{i-1} \left|\dot{q}_j(t)\right|, \frac{1}{L} \sum_{j=1}^{i} \left|\dot{q}_j(t)\right| \right] \right\} \right) = \frac{1}{L} \int_0^1 \left|\dot{q}_i(t)\right| \mathrm{d}t.$$

We call action $i \in S$ *trivial* (for a given choice of C) if q_i is constant, so that the value of π_i is fixed on C. An action is trivial if and only if $\dot{q}_i$ is identically zero; thus, the realization of I is nontrivial with probability one.

3.3 Performance levels, signals, and choice rates

Recall that the elements of $S = \{1, \dots, n\}$ are actions, each vector $\pi \in C \subset \mathbb{R}^n$ is a possible evaluation of the actions' performances, and the vector $\rho(\pi) \in \mathbb{R}^n$ describes the rates at which agents switch to each action in S given the vector of performances. We

[8]Our results depend on our defining our probability model using the parameterization of C that moves at a constant ℓ^1 rate. In particular, if C is a closed curve of performance vectors induced by a closed trajectory through the set of population states, it can be endowed with this trajectory's parameterization, but we do not use this parameterization to define our probability model.

therefore call the random vector q the *performance vector draw* and the random variable I the *action draw*.

We then introduce the random variables

$$Q = Q(t,u) = q_{I(t,u)}(t), \text{ and}$$
$$Z = Z(t,u) = \operatorname{sgn}(\dot{q}_{I(t,u)}(t)).$$

Here $\operatorname{sgn}(\cdot)$ is the signum function: $\operatorname{sgn}(x)$ is 1 if x is positive, -1 if x is negative, and 0 if x is 0. As with $I(t,u)$, $Z(t,u)$ is defined in an arbitrary way when t is not a point of differentiability of q.

The random variable Q, the *performance level*, is defined to be the performance of the drawn action under the drawn performance vector.

The random variable Z is called the *signal*. Evidently, Z specifies whether the performance of the drawn action is improving or worsening as one proceeds from the drawn performance vector along curve C, moving in the direction specified by the parameterization q. We argue below that Z can also be viewed as a binary signal about the performances of actions other than the drawn action under the drawn performance vector.

Finally, given a choice function ρ, we define the random variable

$$R = R(t,u) = \rho_{I(t,u)}(q(t)).$$

The random variable R, the *choice rate*, is the rate at which agents switch to the drawn action under the drawn performance vector.

4. Analysis and Characterizations

To begin the analysis, we establish that the signal Z has the properties mentioned above. Lemma 1 shows that it is binary.

Lemma 1. $\mathbb{P}(Z \in \{-1,1\}) = 1$.

Proof. Let $t \in [0,1]$ be a point at which q is differentiable. If $\dot{q}_i(t) = 0$, then by construction, $I(t,u) \neq i$ for all $u \in (0,1]$. It follows that $\dot{q}_{I(t,u)}(t) \neq 0$ for all $u \in (0,1]$. But since $Z(t,u) = 0$ if and only if $\dot{q}_{I(t,u)}(t) = 0$, and since q is differentiable at all but finitely many $t \in [0,1]$, we conclude that $\mathbb{P}(Z = 0) = 0$. ■

Proposition 1 justifies our interpretation of Z as a signal about the performances of actions other than the drawn action. It does so by showing that Z is independent of the performance of the drawn action, both conditional on the identity of that action and unconditionally.

Proposition 1. *For every piecewise smooth closed curve C, the performance level Q and the signal Z are independent conditional on the action draw I, and are also (unconditionally) independent.*

The proof of Proposition 1 requires the following auxiliary result.

Lemma 2. *The signal Z and the action draw I are independent. In particular,* $\mathbb{P}(Z = 1 \mid I = i) = \mathbb{P}(Z = -1 \mid I = i) = \frac{1}{2}$ *for all nontrivial $i \in S$.*

Proof. Let $i \in S$ be nontrivial. In light of Lemma 1, it is enough to show that $\mathbb{E}(Z \mid I = i) = 0$, or equivalently, that $\mathbb{E}Z1_{\{I=i\}} = 0$. But

$$\begin{aligned}
\mathbb{E}Z1_{\{I=i\}} &= \mathbb{E}\,\mathrm{sgn}(\dot{q}_i)1_{\{I=i\}} \\
&= \int_0^1 \int_0^1 \mathrm{sgn}(\dot{q}_i(t))1_{\{I(t,u)=i\}}\,\mathrm{d}u\,\mathrm{d}t \\
&= \int_0^1 \mathrm{sgn}(\dot{q}_i(t)) \times \frac{1}{L}\left|\dot{q}_i(t)\right|\mathrm{d}t \\
&= \frac{1}{L}\int_0^1 \dot{q}_i(t)\,\mathrm{d}t \\
&= \frac{1}{L}\left(q_i(1) - q_i(0)\right) \\
&= 0,
\end{aligned}$$

where the final equality holds because $q(1) = q(0)$. ■

Proof of Proposition 1. Because Z and I are independent by Lemma 2, the second claim is an easy implication of the first.[9] To prove the first claim, it is enough to show that for all $c \in \mathbb{R}$, all $z \in \{-1, 1\}$ (by Lemma 1), and all nontrivial i,

$$\mathbb{P}(Q \le c, Z = z \mid I = i) = \mathbb{P}(Q \le c \mid I = i)\,\mathbb{P}(Z = z \mid I = i).$$

Since $\mathbb{P}(Z = 1 \mid I = i) = \mathbb{P}(Z = -1 \mid I = i) = \frac{1}{2}$ by Lemma 2, it is sufficient to show that

$$\mathbb{P}(Q \le c, Z = 1 \mid I = i) = \mathbb{P}(Q \le c, Z = -1 \mid I = i),$$

[9]Without the independence of Z and I, this implication would not hold. The situation is analogous to a basic one in Bayesian statistics. There, two observations that are independent conditional on an unknown (i.e., random) parameter are not unconditionally independent, because the value of the first observation provides information about the parameter, which in turn provides information about the second observation.

or equivalently, that $\mathbb{E}Z1_{\{Q\leq c, I=i\}} = 0$.

To do so, we compute as follows:

$$
\begin{aligned}
\mathbb{E}Z1_{\{Q\leq c, I=i\}} &= \int_0^1 \int_0^1 \operatorname{sgn}(\dot{q}_i(t))\, 1_{\{q_i(t)\leq c\}} 1_{\{I(t,u)=i\}}\,\mathrm{d}u\,\mathrm{d}t \\
&= \int_0^1 \operatorname{sgn}(\dot{q}_i(t)) 1_{\{q_i(t)\leq c\}} \times \frac{1}{L}\left|\dot{q}_i(t)\right| \mathrm{d}t \\
&= \frac{1}{L}\int_0^1 \dot{q}_i(t) 1_{\{q_i(t)\leq c\}}\,\mathrm{d}t.
\end{aligned}
$$

To evaluate the integral in the last expression, define $p\colon [0,1] \to \mathbb{R}$ by $p(t) = \min\{q_i(t), c\}$. Then Danskin's envelope theorem (see Theorem A.4 of Hofbauer and Sandholm (2009)) implies that $\dot{p}(t) = \dot{q}_i(t)1_{\{q_i(t)\leq c\}}$ for almost every $t \in [0,1]$. Thus since $q(1) = q(0)$, we have that

$$
\mathbb{E}Z1_{\{Q\leq c, I=i\}} = \frac{1}{L}\int_0^1 \dot{p}(t)\,\mathrm{d}t = \frac{1}{L}\left(p(1) - p(0)\right) = 0,
$$

establishing the proposition. ■

To provide context for our interpretation of integrability, Proposition 2 characterizes the separable choice functions in terms of probability models generated from piecewise smooth closed curves. It shows that ρ is separable if and only if for every such curve, conditional on any realization of the drawn action, the choice rate of this action is independent of the binary signal about the performances of the other actions. Furthermore, it shows that separability implies that the choice rate and the binary signal are unconditionally independent.

Proposition 2. *Choice function ρ is separable if and only if for every piecewise smooth closed curve C, the choice rate R and the signal Z are independent conditional on the action draw I. If these statements are true, then for each such curve C, R and Z are also (unconditionally) independent.*

Proof. As in the proof of Proposition 1, the second statement follows from the first and Lemma 2. Moreover, since separability means that $\rho_i(\pi)$ is measurable with respect to π_i, the "only if" direction of the first statement follows directly from Proposition 1.

To prove the "if" direction, we establish the contrapositive. Suppose that ρ is not separable: for some action i, there exist π_i, π_{-i}, and $\tilde{\pi}_{-i}$ such that $\rho_i(\pi_i, \pi_{-i}) \neq \rho_i(\pi_i, \tilde{\pi}_{-i})$. Without loss of generality, we can suppose that $\rho_i(\pi_i, \pi_{-i}) > r > \rho_i(\pi_i, \tilde{\pi}_{-i})$. Since ρ is continuous, these inequalities remain true if we replace π_i with any $\tilde{\pi}_i \in [\pi_i, \pi_i + h]$, where $h > 0$ is sufficiently small.

Now let C be the rectangle with vertices $a = (\pi_i, \pi_{-i}), b = (\pi_i + h, \pi_{-i}), c = (\pi_i + h, \tilde{\pi}_{-i})$, and $d = (\pi_i, \tilde{\pi}_{-i})$, and suppose that q traverses these points in this order. Then, up to a set of measure zero,

$$\{(t,u) \colon I(t,u) = i\} = \left\{(t,u) \colon q(t) \in \overline{ab}\right\} \cup \left\{(t,u) \colon q(t) \in \overline{cd}\right\},$$

and the two events in the union have equal positive measure. On the first event, $\rho_i(q) > r$ and $\operatorname{sgn}(\dot{q}_i) = 1$, and on the second, $\rho_i(q) < r$ and $\operatorname{sgn}(\dot{q}_i) = -1$. Thus $\rho_i(q)$ and $\operatorname{sgn}(\dot{q}_i)$ are not independent conditional on $I = i$. ■

The construction in this proof connects the two interpretations of the binary signal Z in the simplest possible setting. Conditional on an action draw of $I = i$, the value of the binary signal is $Z \equiv \operatorname{sgn}(\dot{q}_i) = 1$ if and only if the performances of the strategies other than i are $Q_{-i} = \pi_{-i}$, and $Z \equiv \operatorname{sgn}(\dot{q}_i) = -1$ if and only if $Q_{-i} = \tilde{\pi}_{-i}$. Furthermore, the value of Z is independent of Q_i conditional on $I = i$, as required by Proposition 1.

The final result, Proposition 3, shows that a choice function is integrable if and only if for every piecewise smooth closed curve, the choice rate R and the binary signal Z are uncorrelated. Taking the primitive view of Z as the direction of change of the drawn action's performance, Proposition 3 allows us to interpret integrability as a requirement that along performance cycles, actions' choice rates not depend systematically on whether their performances are improving or worsening.

Alternatively, viewing Z as a signal about the performances of actions other than the drawn action, Proposition 3 verifies our intuitive view of integrability as a restriction on the dependence between the rate at which agents switch to an action and the quality of other actions. While separability requires choice rates and signals about other actions' performances to be independent, integrability allows for dependence, but demands that this dependence not be too systematic in nature.

Proposition 3. *Choice function ρ is integrable if and only if for every piecewise smooth closed curve C, the choice rate R and the signal Z are uncorrelated.*

Proof. Lemma 2 implies that $\mathbb{P}(Z = 1) = \mathbb{P}(Z = -1) = \frac{1}{2}$, and so that $\mathbb{E}Z = 0$. Therefore, R and Z are uncorrelated if and only if $\mathbb{E}RZ = 0$. But

$$\begin{aligned}
\mathbb{E}RZ &= \mathbb{E}\rho_I(q)\operatorname{sgn}(\dot{q}_I) \\
&= \int_0^1 \int_0^1 \rho_{I(t,u)}(q(t))\operatorname{sgn}(\dot{q}_{I(t,u)}(t))\,\mathrm{d}u\,\mathrm{d}t \\
&= \int_0^1 \left(\sum_{i=1}^n \rho_i(q(t))\operatorname{sgn}(\dot{q}_i(t)) \times \frac{1}{L}\left|\dot{q}_i(t)\right| \right) \mathrm{d}t
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{L} \int_0^1 \left(\sum_{i=1}^n \rho_i(q(t))\, \dot{q}_i(t) \right) \mathrm{d}t \\
&= \frac{1}{L} \int_0^1 \rho(q(t)) \cdot \dot{q}(t)\, \mathrm{d}t \\
&= \frac{1}{L} \oint_C \rho(\pi) \cdot \mathrm{d}\pi.
\end{aligned}
$$

The final expression equals 0 for every piecewise smooth closed curve C if and only if ρ is integrable, completing the proof of the proposition. ■

References

Arslan, G. and Shamma, J. S. (2006). Anticipatory learning in general evolutionary games. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 6289–6294.

Benaïm, M. and Weibull, J. W. (2003). Deterministic approximation of stochastic evolution in games. *Econometrica*, 71:873–903.

Blackwell, D. (1956). Controlled random walks. In Gerretsen, J. C. H. and Groot, J. D., editors, *Proceedings of the International Conference of Mathematicians 1954*, volume 3, pages 336–338. North Holland, Amsterdam.

Brown, G. W. (1949). Some notes on computation of games solutions. Report P-78, The Rand Corporation.

Brown, G. W. (1951). Iterative solutions of games by fictitious play. In Koopmans, T. C. et al., editors, *Activity Analysis of Production and Allocation*, pages 374–376. Wiley, New York.

Brown, G. W. and von Neumann, J. (1950). Solutions of games by differential equations. In Kuhn, H. W. and Tucker, A. W., editors, *Contributions to the Theory of Games I*, volume 24 of *Annals of Mathematics Studies*, pages 73–79. Princeton University Press, Princeton.

Fudenberg, D. and Kreps, D. M. (1993). Learning mixed equilibria. *Games and Economic Behavior*, 5:320–367.

Fudenberg, D. and Levine, D. K. (1998). *The Theory of Learning in Games*. MIT Press, Cambridge.

Hannan, J. (1957). Approximation to Bayes risk in repeated play. In Dresher, M., Tucker, A. W., and Wolfe, P., editors, *Contributions to the Theory of Games III*, volume 39 of *Annals of Mathematics Studies*, pages 97–139. Princeton University Press, Princeton.

Hart, S. (2005). Adaptive heuristics. *Econometrica*, 73:1401–1430.

Hart, S. and Mas-Colell, A. (2001). A general class of adaptive strategies. *Journal of Economic Theory*, 98:26–54.

Hofbauer, J. (2000). From Nash and Brown to Maynard Smith: Equilibria, dynamics, and ESS. *Selection*, 1:81–88.

Hofbauer, J. and Sandholm, W. H. (2002). On the global convergence of stochastic fictitious play. *Econometrica*, 70:2265–2294.

Hofbauer, J. and Sandholm, W. H. (2009). Stable games and their dynamics. *Journal of Economic Theory*, 144:1665–1693.

Maynard Smith, J. and Price, G. R. (1973). The logic of animal conflict. *Nature*, 246:15–18.

Monderer, D. and Shapley, L. S. (1996). Potential games. *Games and Economic Behavior*, 14:124–143.

Roth, G. and Sandholm, W. H. (2013). Stochastic approximations with constant step size and differential inclusions. *SIAM Journal on Control and Optimization*, 51:525–555.

Sandholm, W. H. (2001). Potential games with continuous player sets. *Journal of Economic Theory*, 97:81–108.

Sandholm, W. H. (2005). Excess payoff dynamics and other well-behaved evolutionary dynamics. *Journal of Economic Theory*, 124:149–170.

Sandholm, W. H. (2009). Large population potential games. *Journal of Economic Theory*, 144:1710–1725.

Sandholm, W. H. (2010). *Population Games and Evolutionary Dynamics*. MIT Press, Cambridge.

Shamma, J. S. and Arslan, G. (2005). Dynamic fictitious play, dynamic gradient play, and distributed convergence to Nash equilibria. *IEEE Transactions on Automatic Control*, 50:312–327.

Weibull, J. W. (1996). The mass action interpretation. Excerpt from 'The work of John Nash in game theory: Nobel Seminar, December 8, 1994'. *Journal of Economic Theory*, 69:165–171.

Young, H. P. (2004). *Strategic Learning and Its Limits*. Oxford University Press, Oxford.