האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

CONTINUOUS-TIME STOCHASTIC GAMES OF FIXED DURATION

YEHUDA (JOHN) LEVY

Discussion Paper # 617 Aug 2012

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

Continuous-Time Stochastic Games of Fixed Duration*

Yehuda (John) Levy†

September 4, 2012

Abstract

We study non-zero-sum continuous-time stochastic games, also known as continuous-time Markov games, of fixed duration. We concentrate on Markovian strategies. We show by way of example that equilibria need not exist in Markovian strategies, but they always exist in Markovian public-signal correlated strategies. To do so, we develop criteria for a strategy profile to be an equilibrium via differential inclusions, both directly and also by modeling continuous-time stochastic as differential games and using the Hamilton-Jacobi-Bellman equations. We also give an interpretation of equilibria in mixed strategies in continuous-time, and show that approximate equilibria always exist.

Keywords: stochastic game, markov game, continuous-time, Markovian equilibrium

Classification: 91A15, 91A23

1 Introduction

1.1 Background

Continuous-time stochastic games, also known as *continuous-time Markov games*, are often used to model interaction in which players can take actions or events can occur at any time, not just at discrete points. Recent developments suggest that they may provide the correct framework for an array of situations in which the states change abruptly rather than smoothly. Furthermore, their study holds promise of results that can give us a better understanding of several types of dynamic games; see Neyman (2012).

*Research supported in part by Israel Science Foundation grants 1123/06 and 1596/10. Many thanks to A. Neyman and an anonymous referee for many useful comments.

†Author Address: Center for the Study of Rationality, and Department of Mathematics, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel. Tel.: (972-2)-6584135, Fax: (972-2)-6513681, E-mail: john.calculus@gmail.com

Stochastic games in discrete time were introduced by Shapley (1953). In a discrete-stochastic game, players play in stages. At each stage, the game is in one of the finitely many available states, and each player chooses an action from the finite action spaces in that state. The actions they choose then determine a probability distribution on the state space, which is used to choose the state at the next stage. There is also a payoff at each stage, which is a function of both the current state and the action profile.

Zachrisson (1964) introduced continuous-time Markov games (although his paper deals with both discrete- and continuous-time games). In the continuous-time model, the game is also always in one of the finitely many available states, and at each point in time each player chooses an action from a finite action space. Again, the players influence the transitions to the other states; however, in this case, the players only determine the *transition rate* (also referred to as the *intensity*). We again assume that at each point in time there is a payoff which is a function both of the current state and action profile.

Both Shapley (1953) and Zachrisson (1964) consider only two-player zero-sum games. Zachrisson (1964) considers games of a finite, pre-determined, and commonly known duration, and his methods of evaluating the stream of payoffs in continuous-time is simply to integrate over time. Furthermore, he considers only Markovian strategies: Players choose their action as a function of time and the current state only. These strategies do not depend on any past states or any past actions of previous players.

Indeed, whether we are studying discrete-time games of fixed duration (the payoff is just the sum or the average of the stage-by-stage payoffs), continuous-time games of fixed duration (the payoff is the integral of the stream of payoffs), zero-sum games or non-zero-sum games, Markovian strategies are a natural class to consider. Indeed, two players following two different plays reaching a state z at time t will have the same preferences over the possible continuations of play. The view that strategies should be dependent only on payoff-relevant data is highlighted by Maskin and Tirole (2001), who use it to motivate the development of the concept of *Markov Perfect Equilibria.* Harsanyi and Selten (1988) call this view the *subgame-consistency principle*, which Hellwig and Leininger (1998) describe as "the behaviour principle according to which a player's behaviour in strategically equivalent subgames should be the same, regardless of the different paths by which these subgames might be reached." Hence, the Markovian property of equilibrium strategies is, in fact, dictated by considerations of rationality and consistency.

In the discrete-time case of games of finite duration, equilibrium in Markovian strategies can be established easily by a backward induction argument; see, e.g., Zachrisson (1964) for the zero-sum case, and Rieder (1979) for the non-zero-sum case.[1] The continuous-time case is somewhat more delicate. Also for these games, Zachrisson (1964) shows that in the zero-sum case the players possess Markovian minmax strategies, and again characterizes them. However, the existence and characterization of Markovian equilibria in the non-zero-sum case have not yet been established.

Continuous-time infinite-horizon games have been studied by Guo and Hernández-Lerma, who examine both the undiscounted game in the zero-sum case under appropriate transience conditions (Guo and Hernández-Lerma, 2003) and the non-zero-sum case under discounted but unbounded payoffs (Guo and Hernández-Lerma, 2005). Neyman (2012) presents a framework for fairly general strategies, and proves the existence of uniform equilibria. Other than the works of these authors, there has been very little development in this direction.

A more common approach to continuous-time games has been via *differential games*, which were introduced by Isaacs (1951, published 1965). These models employ a continuous-state space, and the state changes continuously over time. For the zero-sum case, as well as for optimization for the single decision-maker case, if the value function is smooth,[2] it is characterized as the solution of the Hamilton–Jacobi–Bellman(–Isaacs) equation (henceforth, HJB); see Chapter 4 of Friedman (1971). Since the value function is in general not smooth, it is more generally characterized as the *viscosity solution* of the appropriate HJB equation; these generalized solutions were introduced by Crandall and Lions (1983) to deal with the non-smoothness of the value function in optimization and other differential equations, and applied to differential games by Barron et al (1984). When the game is non-zero-sum and equilibria exist in smooth Markovian strategies, the value function for each player[3] corresponding to this strategy profile also satisfies an appropriate HJB equation, which is now dependent on the other players' strategies; see Chapter 8 of Friedman (1971). However, characterization in the more general case - where equilibria are not smooth - has not been established.

Another vexing issue is the interpretation of randomized strategies in the continuous-time game. As has been pointed out - e.g., by Judd (1985) - the notion of a continuum of independent random variables is ill defined in several relevant senses. One technique employed in differential games, due to Krasovskii and Subbotin (1988), has been to discretize time and replace the original strategies with strategies constant on the induced partition, and to choose the "average" action for that interval. This technique has not yet been applied to a large class of games.

[1] Rieder (1979) works with general state spaces and compact action spaces.
[2] As a function of time and state.
[3] When the other players keep their strategies fixed.

1.2 Our Results

The main purposes of this paper are to show, by way of example, that equilibria need not exist in Markovian strategies, but they do always exist in Markovian correlated strategies. These latter strategies are correlated via a public signal, chosen randomly by Nature at each point in time.

To this end, we develop several tools. The first pair of tools consists of a differential equation describing the evolution of the payoff as a function of initial state and time, and a differential inclusion which gives a necessary and sufficient condition for equilibria. (These tools were developed by Miller (1968) in the case of a single player, using only pure strategies. Our proof of the equilibrium inclusions largely follows his. However, our proof of the evolution equation is direct, while Miller's (1968) uses the equilibrium inclusions and a duality theorem pertaining to differential equations.)

Another set of conditions similar to these, which can also be used to show that the example we present does not possess Markovian equilibrium, are the HJB equations for continuous-time stochastic games.[4] As part of the development of the HJB equations, we show that continuous-time stochastic games (in which only Markovian strategies are allowed) can be viewed as differential games in which players do not monitor others' actions; i.e., only open-loop strategies are used. The HJB equations for optimal control - in this case, applied to each player while the others hold their strategies fixed - reduce to the HJB equations for continuous-time stochastic games.

We also give an interpretation to equilibria in mixed-action strategies in continuous-time games. (As we pointed out in Section 1.1, mixed-action strategies in continuous-time are a problematic notion.) This will be done via a discretization of the continuous-time strategies similar to the one introduced by Krasovskii and Subbotin (1988) (discussed in Section 1.1). However, we will take a different approach, since they worked only in a particular class of zero-sum games. Roughly speaking, given a partition of time, we will replace Markovian strategies with mixtures over pure strategies that are constant on subintervals. Players can use randomization to choose the action that they will play in each subinterval, and the expectation of their choices should be the expectation of the original Markovian strategy in each subinterval.

As a by-product of the technical machinery that we will develop in these approximations, we will easily deduce the existence of approximate Markovian

[4]It is worth noting that in non-zero-sum differential games, the HJB equations are in general ill defined, in the sense that they don't possess unique solutions and the solutions do not depend in a continuous way on initial data; see the survey by Bressan (2011). The lack of uniqueness in our case is not surprising given the usual multiplicity of Nash equilibria; however, it is not known to what extent the solutions are "well-behaved" in terms of dependence on initial data in our class of games - which, as we mentioned, can be viewed as a very special class of differential games.

equilibria in all continuous-time stochastic games. We will also be able to state similar interpretations for the public-signal correlated equilibria, since a continuum of public signals is, for the same reasons, a problematic concept.

The construction of the example of a game without Markovian equilibria takes advantage of the structures of the manifold of Nash equilibria; in particular, it takes advantage of an example from Kohlberg and Mertens (1986), which shows that the set of equilibria in a game can be connected and yet not simply connected - in particular, it can be homeomorphic to a circle - and that even in such a case, each equilibrium can be stable, in the sense that it is the unique equilibrium of an appropriately perturbed game. This construction originally appeared in Levy (2012) in relation to discounted stochastic games which do not possess stationary equilibria. The reader is referred to there or Appendix B of this paper for details.

The intuition of the example is as follows: There is an "active state" in which the game begins, and an absorbing state with payoff 0. We will focus on a particular pair of players, C, D, out of a large set of players. If either of these two players expects a positive average payoff in the future, he will choose an action such that when the other players choose an equilibrium reply, he receives a negative payoff. And, conversely, a negative average payoff in the future will lead to a positive payoff. Hence, for both C, D, the payoff must always be 0; this is a result of the continuous time parameter. However, we take advantage of a non-simply connected structure of the equilibria of the other players - in particular, of players A, B - to have, at each point in time, at least one of the players C, D receives a non-zero payoff.

1.3 Layout of this Paper

Section 2 presents the model of continuous-time stochastic games. Section 3 gives the differential equation for the payoff function associated with a strategy profile, and Section 4 gives the optimality conditions for equilibria. Section 5 gives an interpretation of mixed strategies in continuous time and proves the existence of approximate equilibria. Section 6 presents, modulo the construction that appears in Levy (2012) or in Appendix B of this paper, the example of a continuous-time game that does not possess Markovian equilibria. Section 7 defines, characterizes, and proves the existence of correlated equilibria, and gives an interpretation of correlated strategies in continuous time. Section 8 shows how continuous-time stochastic games can be included in the framework of differential games, and gives the HJB equation. Section 9 gives some extensions of the model under which our results hold.

2 The Model

Throughout, given a bounded function $f : X \to \mathbb{R}$ for some set X, we denote

$$||f||_\infty = \sup_{x \in X} |f(x)|$$

and given a finite set X, $\Delta(X)$ denotes the set of probability distributions over X.

The general framework for a continuous-time stochastic games - also called Markov Games; see Zachrisson (1964) - with finite duration consists of the following framework:

- A finite set of states Z.
- A finite set of players $\mathcal{P}$.
- A finite set of actions I^p for each $p \in \mathcal{P}$. Denote $I^{\mathcal{P}} := \prod_{p \in \mathcal{P}} I^p$ and $\Delta^{\mathcal{P}}(I) = \prod_{p \in \mathcal{P}} \Delta(I^p)$.
- A duration $T \in \mathbb{R}$, $T > 0$.
- A payoff function $r : Z \times I^{\mathcal{P}} \to \mathbb{R}^{\mathcal{P}}$.
- A transition rate $\mu : Z \times Z \times I^{\mathcal{P}} \to \mathbb{R}$, where for all $a \in I^{\mathcal{P}}$ and $z \in Z$, $\sum_{z' \in Z} \mu(z'|z, a) = 0$ and for all $z' \neq z$, $\mu(z'|z, a) \geq 0$.
- The payoff functions and transition rates both extend multi-linearly to mixed-action profiles.

Given an initial state $z_0 \in Z$, the game - which we denote Γ_{z_0} - is played in continuous-time on the interval $[0, T]$. The states are governed by a stochastic process, in which the probability of a transition from state z to a state z' in time $[t, t + h]$, during which the players play action profile $a \in \Delta^{\mathcal{P}}(I)$, is given by $\mu(z'|z, a) \cdot h + o(h)$; we make this explicit in Proposition 1 below.

A Markovian strategy for player $p \in \mathcal{P}$ is a Lebesgue-measurable mapping $u^p : Z \times [0, T] \to \Delta(I^p)$. Let $\mathfrak{A}^p$ denote the set of Markovian strategies for Player p, and let $\mathfrak{A} = \prod_p \mathfrak{A}^p$.

Proposition 1. *For each profile of Markovian strategies $u = (u^p)_{p \in \mathcal{P}} \in \mathfrak{A}$, there exists a unique absolutely continuous Markov transition matrix function P_u defined on $[0, T] \times [0, T]$ (that is, for each s, t, $P_u(s, t)$ is a Markov transition matrix[5] and the function P_u satisfies $P_u(t, t) = I$ and $P_u(s, t) = P_u(s, r)P_u(r, t)$ for all $t, s, r \in [0, T]$) such that:*

[5] A Markov transition matrix is a square matrix with non-negative entries such that each row sums to unity.

- *For all $s \in [0,T]$, $P_u(s,\cdot)$ (which is a row vector) is the unique absolutely continuous function which satisfies a.e. the differential equation*

$$\frac{d}{dt} P_u(s,t) = P_u(s,t)\mu(u(t)) \tag{2.1}$$

and which satisfies $P(s,s) = I$ for all s, where $(\mu(u(s)))_{z,z'} = \mu(z' \mid z, u(z,s))$.

- *For each $z_0 \in Z$ there is a probability space $(\Omega_u^{z_0}, \mathbb{B}_u^{z_0}, P_u^{z_0})$ and a measurable mapping $z : \Omega_u^{z_0} \times [0,T] \to Z$ - which we will view as a random variable which assigns to each $\omega \in \Omega_u^{z_0}$ a mapping $z_\omega : [0,T] \to Z$ - such that for each ω, z_ω is right-continuous and $z_\omega(0) = z_0$, and furthermore,*

$$P_u^{z_0}(z_\omega(t) = z' \mid z_\omega(s) = z) = (P_u(s,t))_{z,z'}$$

for all $z_0, z', z \in Z$ and $s,t \in [0,T]$ with $P_u^{z_0}(z_\omega(s) = z) > 0$.

For a proof, see, e.g., Miller (1967).

Remark 2. Miller (1967) shows that under the assumptions of the above proposition, for a.e. $s \in [0,T]$, it holds for all $t \in [0,T]$ that

$$P_u(s,t) = I + \mu(u(s))(t-s) + o(t-s)$$

This gives another interpretation of μ.

The following propositions also follow:

Proposition 3. *For each $u \in \mathfrak{A}$, $z \in Z$, and each $s,t \in [0,T]$, $(P_u(s,t))_{z,z} > 0$.*

Proposition 4. *For each $u \in \mathfrak{A}$, bounded Borel-measurable mapping $\eta : Z \times \Delta^{\mathcal{P}}(I) \to \mathbb{R}$ and each $z_0 \in Z$, it holds that:*

$$E_u^{z_0}[\int_0^T \eta(z_\omega(s), u(z_\omega(s), s))ds] = \int_0^T \langle P_u^{z_0}(t), \eta(u(s))\rangle ds$$

where $P_u^{z_0}(t)$ is a vector with $(P_u^{z_0}(t))_z = (P_u^{z_0}(0,t))_{z_0,z}$, $\eta(u(s))$ is a vector with $(\eta(u(s)))_z = \eta(z, u(z,s))$, and $\langle , \rangle$ denotes the standard inner product.

We will continue to use the notations from Propositions 1 and 4, and we will write z instead of z_ω.

The payoff in the game starting in state z_0 for Player p is given by:

$$\gamma_u^p(z_0) = E_u^{z_0}\big[\int_0^T r^p(z(t), u(z(t),t))dt\big] = \int_0^T \langle P_u^{z_0}(t), r^p(u(s))\rangle ds \tag{2.2}$$

A profile $u \in \mathfrak{A}$ is a *Markovian equilibrium* if for every $z_0 \in Z$, every $p \in \mathcal{P}$ and every $\tau^p \in \mathfrak{A}^p$, we have

$$\gamma_u^p(z_0) \geq \gamma_{(\tau^p, u^{-p})}^p(z_0)$$

3 Evolution of the Payoff Function

Given a Markovian strategy profile u, for each $t \in T$, let $u_t : Z \times [0, T-t] \to \Delta^{\mathcal{P}}(I)$ be defined by $u_t(z, s) = u(z, s+t)$. Also denote, for each $t \in [0, T]$, $z \in Z$, and $p \in \mathcal{P}$,

$$\gamma_u^p(z,t) = E_u^z\big[\int_0^{T-t} r^p(z(s), u_t(z(s), s))ds\big] = \int_0^{T-t} \langle P_{u_t}^z(t), r^p(u_t(s))\rangle ds \quad (3.1)$$

where $\gamma_u = (\gamma_u^p)_p$, $\gamma_u^p(t) = (\gamma_u^p(z,t))_z$. (Note that $\gamma_u^p(z,0) = \gamma_u^p(z)$.) $\gamma_u^p(z,t)$ can be viewed as the payoff to a player p who begins playing at time t in state z under the profile u. Denote, for $z \in Z$, $t \in [0,T]$,

$$X_u^p(z,t,\cdot) = r^p(z,\cdot) + \langle \mu(z,\cdot), \gamma_u^p(t)\rangle \quad (3.2)$$

where $\mu(z,\cdot) = (\mu(z' \mid z, \cdot))_{z' \in Z}$, and $X_u = (X_u^p)_{p\in\mathcal{P}}$.

Theorem 1. *Let $(u^p)_{p\in\mathcal{P}} \in \mathfrak{A} = \prod_{p\in\mathcal{P}} \mathfrak{A}^p$.*
(a) For each $p \in \mathcal{P}$, $\gamma_u^p(\cdot) : [0,T] \to \mathbb{R}^Z$ is the unique absolutely continuous function $\psi = (\psi_z)_z$ satisfying the following differential equation for a.e. $t \in [0,T]$:

$$\frac{d\psi_z}{dt}(t) = -\big[r^p(z, u(z,t)) + \langle \mu(z, u(z,t)), \psi(t)\rangle\big] \quad (3.3)$$

with boundary condition $\psi(T) = 0$.
(b) For all $z \in Z$, and a.e. t,

$$\frac{d\gamma_u}{dt}(z,t) = -\big(X_u(z,t,u(z,t))\big)$$

Proof. (b) follows immediately from (a). The uniqueness in (a) follows from slightly generalized versions of the standard uniqueness theorems for ordinary differential equations; see, e.g., Section 2.2 of Coddington and Levinson (1972).[6]

We will show that $\gamma_u^p(z,\cdot)$ indeed satisfies (3.3) a.e.. Let $P_u(t) := P_u(0,t)$. Let $0 < t < T$ and let $|h| \ll 1$. Since the strategies are Markovian, we have $P_u(t, t+h) = P_{u_t}(h)$, and hence,

$$\begin{aligned} P_{u_{t+h}}(v-h) &= P_u(t+h, (t+h)+(v-h)) = P_u(t+h, t+v) \\ &= P_u(t+h, t) \cdot P_u(t, t+v) = P_u(t+h, t) \cdot P_{u_t}(v) \end{aligned}$$

[6]Although the theorems there are stated for dynamics of the form $\frac{dx}{dt} = f(t,x)$ when f is continuous in both parameters, it is remarked there that one can derive similar results with minor changes when f satisfies only measurability in the time coordinate.

Also, $u_{t+h}(v-h) = u_t(v)$. Hence,

$$
\begin{aligned}
&\gamma_u^p(t+h) - \gamma_u^p(t) = \\
&\quad = \int_0^{T-(t+h)} P_{u_{t+h}}(s) \cdot r^p(u_{t+h}(s))ds - \int_0^{T-t} P_{u_t}(v) \cdot r^p(u_t(v))dv \\
&\quad = \int_h^{T-t} P_{u_{t+h}}(v-h) \cdot r^p(u_{t+h}(v-h))dv - \int_0^{T-t} P_{u_t}(v) \cdot r^p(u_t(v))dv \\
&\quad = \int_0^{T-t} (P_u(t+h,t) - I) P_{u_t}(v) \cdot r^p(u_t(v))dv \\
&\quad + \int_h^0 P_u(t+h,t) P_{u_t}(v) r^p(u_t(v))dv \\
&\quad = (P_u(t+h,t) - I) \int_0^{T-t} P_{u_t}(v) \cdot r^p(u_t(v))dv \\
&\quad + P_u(t+h,t) \int_h^0 P_{u_t}(v) r^p(u_t(v))dv \\
&\quad = (P_u(t+h,t) - I)\gamma_u^p(t) + P_u(t+h,t) \int_h^0 P_u(v,t) r^p(u(t+v))dv
\end{aligned}
$$

Since $P_u(t+h,t) \underset{h\to 0}{\to} I$, we have from Remark 2 that for a.e. $t \in [0,T]$,

$$
\frac{P_u(t+h,t) - I}{h} = -P_u(t+h,t)\frac{P_u(t,t+h) - I}{h} \underset{h\to 0}{\longrightarrow} -I \cdot P_u(t,t) \cdot \mu(u(t)) = -\mu(u(t))
$$

and for a.e. $t \in [0,T]$,

$$
\begin{aligned}
&\frac{1}{h}\int_h^0 P_u(v,t) r^p(u(t+v))dv \\
&\quad = \frac{1}{h}\int_h^0 (P_u(v,t) - I) r^p(u(t+v))dv + \frac{1}{h}\int_{t+h}^t r^p(u(x))dx \\
&\quad \underset{h\to 0}{\longrightarrow} 0 - r^p(u(t)) = -r^p(u(t))
\end{aligned}
$$

so we deduce the desired result.

□

4 An Optimality Condition

Theorem 2. *A Markovian strategy $u = (u^p)_{p\in\mathcal{P}} \in \mathfrak{A} = \prod_{p\in\mathcal{P}} \mathfrak{A}^p$ is a Markovian equilibrium iff for all $z \in Z$ and a.e. $t \in [0,T]$,*

$$
u(z,t) \in NE\big(X_u(z,t,\cdot)\big) = NE\Big(\big(r^p(z,\cdot) + \langle \mu(z,\cdot), \gamma_u^p(t)\big)_{p\in\mathcal{P}}\Big)
$$

or, equivalently,

$$
\frac{d\gamma_u}{dt}(z,t) \in -NEP\big(X_u(z,t,\cdot)\big) = -NE\Big(\big(r^p(z,u(z,t)) + \langle \mu(z,u(z,t)), \gamma_u^p(t)\big)_{p\in\mathcal{P}}\Big)
$$

where NE (resp. NEP) denotes the Nash equilibria (resp. Nash equilibria payoff) correspondence, which assigns to each normal-form game its set of Nash equilibria (resp. Nash equilibria payoffs).

This theorem was established for the single-player case by Miller (1968), and our proof follows his.

Proof. Denote $P_u(t) := P_u(0,t)$. Let u be a strategy profile, fix a Player $p \in \mathcal{P}$, and let Ξ be a profile which differs only (at most) in the strategy of Player p. We first establish

$$\gamma_u^p(0) - \gamma_\Xi^p(0) = \int_0^T P_\Xi(s)\big[X_u^p(t, u(t)) - X_u^p(t, \Xi^p(t), u^{-p}(t))\big]dt$$
$$= \int_0^T P_\Xi(t)\big[r^p(u(t)) + \mu(u(t)) \cdot \gamma_u^p(t) - r^p(\Xi(t)) - \mu(\Xi(t)) \cdot \gamma_u^p(t)\big]ds \tag{4.1}$$

where we recall that $\gamma_u^p(t) = (\gamma_u^p(z,t))_{z\in Z}$, $X_u^p(t,a) = (X_u^p(z,t,a_z))_{z\in Z}$ for $a \in (\Delta^{\mathcal{P}}(I))^Z$. Since $P_u(0) = P_\Xi(0) = I$ and $\gamma_u^p(0) = 0$, we have

$$[P_u(T) - P_\Xi(T)]\gamma_u^p(T) - [P_u(0) - P_\Xi(0)]\gamma_u^p(0) = 0$$

Since the function $g : [0,T] \to \mathbb{R}$ defined by $g(t) = [P_u(t) - P_\Xi(t)]\gamma_u^p(t)$ is absolutely continuous (as P_u, P_Ξ and γ_u^p are absolutely continuous and bounded) we have

$$0 = \int_0^T g'dt = \int_0^T [\frac{d[P_u - P_\Xi]}{dt}\gamma_u^p + [P_u - P_\Xi]\frac{d\gamma_u^p}{dt}]dt$$
$$= \int_0^T [(P_u\mu(u) - P_\Xi\mu(\Xi))\gamma_u^p - (P_u - P_\Xi)(r^p(u) + \mu(u)\gamma_u^p)]dt$$
$$= \int_0^T [(P_\Xi - P_u)r^p(u) + P_\Xi(\mu(u) - \mu(\Xi))\gamma_u^p]dt$$

Combining this with

$$\gamma_u^p(0) - \gamma_\Xi^p(0) = \int_0^T [P_u r^p(u) - P_\Xi r^p(\Xi)]dt$$

gives (4.1).

(4.1) now shows that if u^p is such that for a.e. t and all z, $u_z^p(t)$ maximizes

$$X_u^p(z, t, \cdot, u_z^{-p}(t)) = r^p(z, \cdot, u_z^{-p}(t)) + \langle \mu(z, \cdot, u_z^{-p}(t)), \gamma_u^p(z,t)\rangle$$

then u^p is a best reply to u^{-p}. On the other hand, suppose u^p is a best reply to u^{-p} but that for some $\xi \in Z$ and all t in a set S of positive measure, $u_\xi^p(t)$

does not maximize $X^p_u(\xi, t, \cdot, u^{-p}_\xi(t))$. We contend that there exists a strategy Ξ^p for Player p such that for a.e. $t \in [0,T]$ and any $z \in Z$,

$$X^p_u(z, t, \Xi^p_z(t), u^{-p}(t)) = argmax[X^p_u(z, t, \cdot, u^{-p}(t))]$$

Indeed, the existence of such Ξ^p follows from standard Borel measurable selection theorems (e.g., Kuratowski and Ryll-Nardzewski (1965) or Himmelberg (1975)). In particular,

$$X^p_u(z, t, \Xi^p_z(t), u^{-p}(t)) \geq X^p_u(z, t, u(z,t))$$

with strict inequality when $z = \xi$ and $t \in S$; hence, since the diagnol elements of $P_\Xi(t)$ are positive for all t, (4.1) implies $\gamma^p_{\Xi^p, u^{-p}}(\xi, 0) > \gamma^p_u(\xi, 0)$, a contradiction. □

5 Interpretation of Mixed Strategies in Continuous Time and Existence of Approximate Equilibria

It is well known, e.g., Judd (1985), that there is no well-defined notion of a continuum of i.i.d. random variables. Hence, the notion of using mixed actions in continuous-time games is problematic, as the randomization cannot be carried out at each instant. In this section, we introduce a method for approximating strategies using mixed actions by strategies by a new type of strategies, which also employ randomization but only at finitely many points in time, resulting in a sequence of approximate equilibria which converge in the appropriate sense. This technique is based on an idea[7] introduced by Krasovskii and Subbotin (1988). In doing so, we will deduce the existence of Markovian ε-equilibria.

Let $\mathcal{T}$ be a finite partition of $[0,T]$ into intervals; we identify $\mathcal{T}$ with the set of left-most points of those intervals; i.e., $\{t_0, \ldots, t_k\}$ with $0 = t_0 < \cdots < t_k < T$, or for brevity, $\mathcal{T} = \{0 = t_0 < \cdots < t_k < T\}$. A $\mathcal{T}$-Markovian strategy (or $\mathcal{T}$-strategy for short) of Player p is a mapping $w^p : Z \times \mathcal{T} \to \Delta(I^p)$. The interpretation is that at the beginning of each interval of the partition, Player p performs his randomization of which pure action he will play at each state he may find himself in during that time interval. The collection of all such strategies for Player p will be denoted $\Sigma^p(\mathcal{T})$, which can be identified with $(\Delta(I^p))^{Z \times \mathcal{T}}$, and $\Sigma(\mathcal{T}) = \prod_{p \in \mathcal{P}} \Sigma^p(\mathcal{T})$. The collection of $\mathcal{T}$-strategies for Player p which choose only pure actions (and, hence, can be identified with $(I^p)^{Z \times \mathcal{T}}$) will be denoted $S^p(\mathcal{T})$, and $S(\mathcal{T}) = \prod_{p \in \mathcal{P}} S^p(\mathcal{T})$. An element $w^p \in \Sigma^p(\mathcal{T})$ determines a

[7]We thank an anonymous referee for pointing this reference out. The main application in Krasovskii and Subbotin (1988) is differential games of evasion and pursuit; since such games are zero-sum games, both their construction and their use of these approximations are different than ours.

Markovian strategy by

$$\hat{w}^p(z,t) = w^p(z,t_j) := \sum_{v^p \in S^p(\mathcal{T})} w^p[v] \cdot v^p(z,t_j),\ t_j \le t < t_{j+1}. \tag{5.1}$$

$\hat{w}^p(z,t)$ can be interpreted as the expectation of the action of Player p used at time t in state z if he uses the $\mathcal{T}$-Markovian strategy w^p. It is, however, the inverse operation that will be crucial for our purposes.

For $\mathcal{T} = \{0 = t_0 < \cdots < t_k < T\}$, as above, $\ell(\mathcal{T}) = \max_{0 \le j < k} |t_{j+1} - t_j|$, where $t_{k+1} = T$ for such $\mathcal{T}$. For each player p and each Markovian strategy $u^p \in \Sigma^p$, define $w^p_{\mathcal{T}} \in \Sigma^p(\mathcal{T})$ by

$$w^p(z,t_j) = \frac{1}{t_{j+1} - t_j} \int_{t_j}^{t_{j+1}} u^p(z,t)dt \tag{5.2}$$

where both sides are elements of $\Delta(I^p)$. We will call $w^p_{\mathcal{T}}$ the $\mathcal{T}$-adapted version of u^p. This terminology applies to profiles of $\mathcal{T}$-strategies as well.

Remark 5. If $w^p \in \Sigma^p(\mathcal{T})$ is the $\mathcal{T}$-adapted version of $u^p \in \mathfrak{A}^p$, then

$$\hat{w}^p(z,t) = \frac{1}{t_{j+1} - t_j} \int_{t_j}^{t_{j+1}} u^p(z,t),\ \forall t \in [t_j, t_{j+1}).$$

We define the payoff under a profile $w \in \Sigma(\mathcal{T})$ in state $z \in Z$ by

$$\gamma_w(z) := E_w[\gamma_w(z)] = \sum_{v \in S(\mathcal{T})} w[v] \cdot \gamma_v(z). \tag{5.3}$$

where on the right-hand side, each $v \in S(\mathcal{T})$ is identified with the pure Markovian strategy it induces. Indeed, the interpretation of (5.3) is that the players randomize (either at the outset of the game or at the beginning of each time interval; because of the lack of monitoring it does not matter) which action profile will be played for each state in each of the intervals.

We would also like to define a notion of payoff when one player uses (or, conceptually, deviates to) a Markovian strategy when the other players use only $\mathcal{T}$-strategies. Let $p \in \mathcal{P}$ and let $w^{-p} \in \prod_{q \ne p} \Sigma^q(\mathcal{T})$. Then we define for each Markovian strategy u^p of Player p,

$$\gamma_{u^p, w^{-p}}(z) := E_w[\gamma_{u^p, w^{-p}}(z)] = \sum_{v^{-p} \in \prod_{q \ne p} S^q(\mathcal{T})} w^{-p}[v^{-p}] \cdot \gamma_{u^p, v^{-p}}(z). \tag{5.4}$$

Hence, we say that $w \in \Sigma(\mathcal{T})$ is an ε-equilibrium of $\Gamma_z(\mathcal{T})$ if no player has a deviation, with a profit of more than ε, to a Markovian strategy in the game with initial state z.

Note that for each $z \in Z$, the game in which the players choose an action profile in $\Sigma(\mathcal{T})$, and payoffs are given by (5.3), is a finite normal-form game.[8] Let this game be denoted $\Psi_z(\mathcal{T})$. Also denote for strategy profiles $u, y \in \mathfrak{A}$ the L^1 distance between them, defined by

$$||u - v||_{L^1} = \int_0^T ||u(s) - y(s)||_\infty ds.$$

For $\varepsilon \geq 0$, a profile $u \in \mathfrak{A}$ is a *Markovian ε-equilibrium* if for every $z_0 \in Z$, every $p \in \mathcal{P}$, and every $\tau^p \in \mathfrak{A}^p$, we have

$$\gamma_u^p(z_0) + \varepsilon \geq \gamma^p_{(\tau^p, u^{-p})}(z_0).$$

Note that a Markovian 0-equilibrium is simply an equilibrium.

Theorem 3. *Every continuous-time stochastic game possesses a Markovian ε-equilibrium for every $\varepsilon > 0$. Furthermore, for each game, there is $D > 0$ such that if $\mathcal{T}$ is a partition of $[0, T]$ and w is an equilibrium of $\Psi_z(\mathcal{T})$, then $\hat{w}$ is a $D \cdot \ell(\mathcal{T})$-equilibrium of Γ.*

As we show in Section 6, a continuous-time stochastic game need not possess a Markovian equilibrium.

Theorem 4. *Let $(\mathcal{T}_n)_{n=1}^\infty$ be sequence of partitions[9] with $\ell(\mathcal{T}_n) \underset{n\to\infty}{\to} 0$. Then for any profile of Markovian strategies $p = (u^p)_{p\in\mathcal{P}}$, if w_n^p is the $\mathcal{T}_n$ version of u^p, then:*
(a) For a.e. $t \in [0, T]$ and each $z \in Z$, $\hat{w}_n^p(z, t) \to u^p(z, t)$, and $\hat{w}_n^p(z, \cdot) \to u^p(z, \cdot)$ in $L^T([0, T])$.
(b) For each $z \in Z$,

$$\gamma_{w_n}(z) \underset{n\to\infty}{\to} \gamma_u(z).$$

(c) If u is an ε-Markovian equilibrium, $\varepsilon \geq 0$, then there is a sequence of positive numbers $(\varepsilon_n)_{n\in\mathbb{N}}$ with $\limsup_{n\to\infty} \varepsilon_n \leq \varepsilon$, such that for all z, w_n is an ε_n-equilibrium of $\Psi_z(\mathcal{T}_n)$ and of $\Gamma_z(\mathcal{T}_n)$.

This justifies the use of Markovian strategies in equilibrium or ε-equilibrium. If the players wish to implement an equilibrium profile $u = (u^p)_{p\in\mathcal{P}}$ of Markovian strategies, they can discretize time arbitrarily fine via a partition $\mathcal{T}$ and use the $\mathcal{T}$-adapted version w; if $\mathcal{T}$ is fine enough, no player can profit significantly by deviating, either to another $\mathcal{T}$-strategy or even to a Markovian strategy.

The proofs of Theorems 3 and 4 rely on four lemmas. First, note that for any two action profiles $a, b \in \Delta^{\mathcal{P}}(I)$,

$$||\mu(a) - \mu(b)||_\infty \leq ||\mu||_\infty \cdot ||a - b||_\infty, \; ||r(a) - r(b)||_\infty \leq ||r||_\infty \cdot ||a - b||_\infty \quad (5.5)$$

Lemma 6 follows from standard discrete-time dynamic programming arguments. The proofs of Lemmas 7, 8, and 9 appear in Appendix A.

[8] By definition, it is multilinear over $\Sigma(\mathcal{T})$.
[9] The partitions need not be nested.

Lemma 6. *For each game and each partition $\mathcal{T}$ of $[0,T]$, there exists $w \in \Sigma(\mathcal{T})$, which is an equilibrium in $\Psi_z(\mathcal{T})$ for every $z \in Z$.*

Lemma 7. *For each game, there exists $M > 0$ such that for any two Markovian strategy profiles $u, y \in \mathfrak{A}$,*

$$||P_u(t) - P_y(t)||_\infty \le M \cdot ||u - y||_{L^1} \cdot e^{M \cdot t}, \ \forall t \in [0,T] \tag{5.6}$$

and

$$||\gamma_u - \gamma_y||_\infty \le M \cdot ||u - y||_{L^1} \tag{5.7}$$

Lemma 8. *For each game, there exists a $K > 0$ such that for every partition $\mathcal{T}$, the following statements hold:*
(a) Let $u \in \mathfrak{A}$ be any Markovian strategy profile, and let $w \in \Sigma(\mathcal{T})$ be the $\mathcal{T}$-adapted version of u. Then for all $t \in T$,

$$||P_u(t) - P_w(t)||_\infty \le K \cdot \ell(\mathcal{T}) \tag{5.8}$$

where we define $P_w(t) = \sum_{v \in S(\mathcal{T})} w[v] \cdot P_{\hat{v}}(t)$, and

$$||\gamma_u - \gamma_w||_\infty \le K \cdot \ell(\mathcal{T}) + K \cdot ||u - \hat{w}||_{L^1} \tag{5.9}$$

(b) Let $p \in \mathcal{P}$ be a player, let $u^{-p} \in \prod_{q \neq p} \mathfrak{A}^q$ be any Markovian strategy profile for the other players, and let $w^{-p} \in \Sigma(\mathcal{T})$ be the $\mathcal{T}$-adapted version of u^{-p}. Then for all $t \in T$ and all $u^p \in \mathfrak{A}^p$,

$$||P_{u^p,u^{-p}}(t) - P_{u^p,w^{-p}}(t)||_\infty \le K \cdot \ell(\mathcal{T}) \tag{5.10}$$

where we define $P_{u^p,w^{-p}}(t) = \sum_{v^{-p} \in \prod_{q \neq p} S^q(\mathcal{T})} w^{-p}[v^{-p}] \cdot P_{u^p,\hat{v}^{-p}}(t)$, and

$$||\gamma_{u^p,u^{-p}} - \gamma_{u^p,w^{-p}}||_\infty \le K \cdot \ell(\mathcal{T}) + K \cdot ||u^{-p} - \hat{w}^{-p}||_{L^1}$$

Note that for $w \in \Sigma(\mathcal{T})$, $(P_w(t))_{z,z'}$ is indeed the probability of being in state z' in time t if v is chosen according to w and $z(0) = z$.

Lemma 9. *For each game, there exists $C > 0$ such that for any partition $\mathcal{T}$, any set of Player $p \in \mathcal{P}$, and any Markovian profile $u^{-p} \in \prod_{q \neq p} \mathfrak{A}^q$ for the other players with $\mathcal{T}$-adapted versions w^{-p},*

$$\left| \sup_{w^p \in \Sigma^p(\mathcal{T})} \gamma^p_{w^p,w^{-p}} - \sup_{u^p \in \mathfrak{A}^p} \gamma^p_{u^p,w^{-p}} \right| \le C \cdot \ell(\mathcal{T}).$$

Proof. (of Theorem 3). Let $\mathcal{T}$ be a partition of $[0,T]$, and let C, K be as in Lemmas 8 and 9. Let w be an equilibrium of $\Psi_z(\mathcal{T})$ for each $z \in Z$; by Lemma 6, at least one such equilibrium exists. Take $u = \hat{w}$; we contend that u is a $(C+K)\ell(\mathcal{T})$-equilibrium of the continuous-time game.

Observe that since for each Player $p \in \mathcal{P}$, $u^{-p} = \hat{w}^{-p}$, Lemmas 8 and 9 show that

$$\begin{aligned}
\sup_{y^p \in \mathfrak{A}^p} \gamma^p_{y^p, u^{-p}}(z) &\leq \sup_{y^p \in \mathfrak{A}^p} \gamma^p_{y^p, w^{-p}}(z) + K\ell(\mathcal{T}) + ||u^{-p} - \hat{w}^{-p}||_{L^1} \\
&= \sup_{y^p \in \mathfrak{A}^p} \gamma^p_{y^p, w^{-p}}(z) + K\ell(\mathcal{T}) \\
&\leq \sup_{v^p \in \Sigma^p(\mathcal{T})} \gamma^p_{v^p, w^{-p}}(z) + (C+K)\ell(\mathcal{T}) \leq \gamma^p_w(z) + (C+K)\ell(\mathcal{T})
\end{aligned}$$

since w is an equilibrium of $\Psi_z(\mathcal{T})$. Hence, take $D = C + K$. □

Proof. (of Theorem 4). The a.e. convergence in (a) follows from the Lebesgue density theorem and Remark 5; the L^1 convergence follows from this, since a uniformly bounded a.e. convergent sequence on a finite measure space converges in L^1. (b) follows from (a) and part (a) of Lemma 8.

For (c), first we show that there exist $(\kappa_n)_{n \in \mathbb{N}}$, $\limsup_{n \to \infty} \kappa_n \leq \varepsilon$, such that w_n is a κ_n-equilibrium in $\Gamma_z(\mathcal{T}_n)$. Indeed, we will define

$$\kappa_n = \max_{p \in \mathcal{P}, z \in Z} \sup_{y^p \in \mathfrak{A}^p} \left(\gamma^p_{y^p, w_n^{-p}}(z) - \gamma^p_w(z)\right)$$

and we need to show that $\kappa_n \to 0$. It's enough to show that for each $p \in \mathcal{P}$, $z \in Z$, if

$$\delta_n(p,z) := \sup_{y^p \in \mathfrak{A}^p} \left(\gamma^p_{y^p, w_n^{-p}}(z) - \gamma^p_{w_n}(z)\right)$$

then $\limsup_{n \to \infty} \delta_n(p,z) \leq \varepsilon$. By parts (a) and (b) of Lemma 8, there is $K > 0$ such that

$$\begin{aligned}
\delta_n(p,z) &\leq 2K \cdot \ell(\mathcal{T}_n) + K \cdot ||u^{-p} - \hat{w}_n^{-p}||_{L^1} + K \cdot ||u - \hat{w}||_{L^1} \\
&\quad + \sup_{y^p \in \mathfrak{A}^p} \left(\gamma^p(y^p, u^{-p})(z) - \gamma^p_u(z)\right) \\
&\leq 2L \cdot \ell(\mathcal{T}_n) + 2K \cdot ||u - \hat{w}_n||_{L^1} \qquad (5.11)
\end{aligned}$$

since u is an equilibrium (and taking the norm over fewer coordinates only decreases the L^1 distance.) By our assumptions, and by (a), $\limsup_{n \to \infty} \delta_n(p,z) \leq \varepsilon$.

Lemma 9 now shows that if we set $\varepsilon_n = \kappa_n + C \cdot \ell(\mathcal{T}_n)$ for the appropriate $C > 0$ and all $n \in \mathbb{N}$, then since w_n is a κ_n-equilibrium of $\Gamma_z(\mathcal{T}_n)$, it also an ε_n-equilibrium of $\Psi_z(\mathcal{T}_n)$. Furthermore, $\limsup_{n \to \infty} \varepsilon_n \leq \varepsilon$ by (a). □

6 A Game without Markovian Equilibria

This section introduces some preliminary notations and results that we will need.

6.1 Some Notation

- If p is a mixed action over an action space A and $a \in A$, then $p[a]$ denotes the probability that p chooses A.
- If g is some payoff vector to some set of players $\mathcal{P}$, and $T \subseteq \mathcal{P}$, then g^T denotes the restriction of the vector to the players in T.
- Similarly, if a is an action profile of the players in $\mathcal{P}$ and $T \subseteq \mathcal{P}$, a^T is the vector of strategies of players in T.
- If Λ is a normal form game on some set of players, and α is a strategy profile of those players, then $\Lambda(\alpha)$ denotes the resulting payoff vector. If $T \subseteq \mathcal{P}$, then $\Lambda^T(\alpha)$ (resp. $\Lambda^{-T}(\alpha)$) denotes the payoff to the players in T (resp. in T^c).
- For such Λ, α, and $T \subseteq \mathcal{P}$, $\Lambda^T(\cdot, \alpha^{-T})$ denotes the expected normal-form game facing the players in T when the other players are restricted to playing α^{-T}.

6.2 Properties of The Fundamental Normal Form Games

In this section, we state the components and properties of a normal-form game (which is also dependent on a parameter) around which our example is constructed. This normal-form game was constructed in Levy (2012) to construct a discounted stochastic game with a continuum of states that does not possess a stationary equilibrium. There reader is referred to there for the construction and the proof of Proposition 10 below.

The components of the game are the following:

- The players are $\mathcal{P} = \{A, B, C, D, \theta^1, \ldots, \theta^M\}$, where $M \in \mathbb{N}$.
- The action set of Player A is $\{T, M, B\}$.
- The action set of Player B is $\{L, C, R\}$.
- The action set of each of the players $\theta^1, \ldots, \theta^M$ is $\{L, R\}$.
- The action set of each player C, D is $\{-1, 1\}$.
- The payoff r_ω (which depends on the parameter $\omega = (\omega^C, \omega^D) \in \mathbb{R}^2$) will be the sum of two payoffs, $r_\omega := r_1 + r_{2,\omega}$, where r_1 is independent of ω.
- The payoff to Players C, D under r_1 is not dependent on the actions of C, D; that is, if a, b are two action profiles with $a^{-\{C,D\}} = b^{-\{C,D\}}$, then $r_1^{C,D}(a) = r_1^{C,D}(b)$.

- The second payoff function $r_{2,\omega}$ is dependent on ω. It gives 0 payoff to all players other than C, D: That is, $r_{2,\omega}^{-\{C,D\}} \equiv 0$. To players C, D, $r_{2,\omega}$ is dependent only on $a^{C,D}$ and is given by:

$r_{2,\omega}(a) =$

$C \backslash D$	1	-1
1	ω^C, ω^D	$\frac{1}{2}\omega^C, \frac{1}{2}\omega^D$
-1	$\frac{1}{2}\omega^C, \frac{1}{2}\omega^D$	0

Let G_ω denote the game resulting from choice of parameter ω. The following proposition contains the main properties of G_ω that we will need:

Proposition 10. *Let $\omega \in \mathbb{R}^2$, and let a be an equilibrium profile in the game G_ω. The following properties hold:*

- *If $\omega^C > 0$, $r_1^C(a) \leq -\frac{1}{2}$. If $\omega^C < 0$, $r_1^C(a) \geq \frac{1}{2}$.*
- *If $\omega^D > 0$, $r_1^D(a) \leq -\frac{1}{2}$. If $\omega^D < 0$, $r_1^D(a) \geq \frac{1}{2}$.*
- *The payoffs to C, D in $r_1(a)$ are not both 0; that is, $r_1^{C,D}(a) \neq 0$. (This is regardless of ω.)*

6.3 The Example

- The players and their actions are $\mathcal{P} = \{A, B, C, D, \theta^1, \ldots, \theta^M\}$ as in Section 6.2, along with the actions sets as given there.
- The set of states is $Z = \{z_0, \overline{0}\}$, where $\overline{0}$ is an absorbing state of payoff 0. We only need to define the payoffs in/transitions from z_0, hence we will often drop reference to the state, as explicitly done below.
- The payoff function is $r(\cdot) := r(z_0, \cdot) := r_1(\cdot)$, where r_1 is as in Section 6.2.
- The transition rate $\mu : I^{\mathcal{P}} \to \mathbb{R}$, where we write $\mu(\cdot) := -\mu(z_0 \mid z_0, \cdot) \geq 0$ (the intensity of the flow *out* of z_0) is determined by the actions of Players C, D, and is given by:

$\mu(\cdot) =$

$C \backslash D$	L	R
L	0	$\frac{1}{2}$
R	$\frac{1}{2}$	1

Suppose that u is a Markovian equilibrium. Denoting $\gamma_u(\cdot) := \gamma_u(z_0, \cdot)$, $X_u(t, \cdot) := X_u(z_0, t, \cdot)$. We can now write (3.2) explicitly as:

$$X_u^p(t, \cdot) = r^p(\cdot) - \mu(\cdot)\gamma_u^p(t)$$

for all $p \in \mathcal{P}$, and we can write the conclusions of Theorem 1 and 2 in this game explicitly:

$$\frac{d\gamma_u^p}{dt}(t) = -X_u^p(t, u(t)) = -\big(r^p(u(t)) - \mu(u(t))\gamma_u^p(t)\big) \tag{6.1}$$

and for a.e. $t \in [0, 1]$,

$$u(t) \in NE(X_u(t, \cdot)) = NE\big(r(\cdot) - \mu(\cdot)\gamma_u(t)\big) \tag{6.2}$$

Remark 11. Since we are assuming that u is an equilibrium, we could have deduced (6.1) and (6.2) from Theorem 7 as well.

The following three lemmas are immediate:

Lemma 12. *For $t \in [0, 1]$,*

$$X_u(t, \cdot) = r_{\omega(t)}(\cdot) + \xi_u(t, \cdot)$$

where r_ω is of Section 6.2,

$$\omega(t) = (\omega^C(t), \omega^D(t)) := \gamma_u^{C,D}(t)$$

and ξ_u is given by

$$\xi_u^{C,D}(t) = -\gamma_u^{C,D}(t)$$

and

$$\xi_u^{-\{C,D\}}(t, a) = -\mu(a)\gamma_u^{-\{C,D\}}(t)$$

Lemma 13. *For $t \in J := (1 - \frac{1}{4||r||_\infty}, 1]$,*

$$||\xi_u(t)|| \leq ||\gamma_u(t)||_\infty = ||\omega(t)||_\infty < \frac{1}{4}$$

Lemma 14. *Let N be a finite set of players, let $I^1, \ldots, I^N$ be action spaces, and let $g_1, g_2 : N \times I^1 \times \cdots \times I^2 \to \mathbb{R}^N$ be two payoff functions, such that for any $p \in N$ and any pair of pure action profiles a, b which differ (at most) in Player p's action,*

$$g_1^p(a) - g_1^p(b) = g_2^p(a) - g_2^p(b)$$

Then the set of Nash equilibria under g_1 is the same as the set of Nash equilibria under g_2.

Note that under $\xi_u(t, \cdot)$, each player's payoff is independent of his own action. Combining this observation with Lemma 14 (where $g_1(\cdot) = X_u(t, \cdot)$ and $g_2(\cdot) = r_{\omega(t)}(\cdot)$), Lemma 12, Proposition 10, and the requirement (6.2), we deduce that for a.e. $t \in [0, 1]$,

- If $\gamma^C(t) > 0$ (resp. < 0), $r^C(u(t)) \leq -\frac{1}{2}$ (resp. $\geq \frac{1}{2}$).
- If $\gamma^D(t) > 0$ (resp. < 0), $r^D(u(t)) \leq -\frac{1}{2}$ (resp. $\geq \frac{1}{2}$).

- Regardless of the values of $\gamma^C(t), \gamma^D(t)$,

$$r^C(u(t)) \neq 0 \text{ or } r^D(u(t)) \neq 0 \tag{6.3}$$

From these, we deduce:

- By Lemma 13, in the interval $J := (1 - \frac{1}{4||r||_\infty}, 1]$, we have

$$\begin{aligned} ||r^{C,D}(\cdot) - X_u^{C,D}(t, \cdot)||_\infty &= ||(1 - \mu(\cdot))\omega(t) + \xi_u^{C,D}(t, \cdot)||_\infty \\ &\leq ||\omega(t)||_\infty + ||\xi_u^{C,D}||_\infty < \frac{1}{2} \end{aligned} \tag{6.4}$$

  Hence, the above observations and (6.1) and show that for a.e. $t \in J$, if $\gamma^C(t) > 0$ (resp. < 0), then $\frac{d\gamma^C}{dt}(t) > 0$ (resp. < 0), and similarly with $\gamma^D(t)$ and $\frac{d\gamma^D}{dt}(t)$.

- (6.3), (6.1), and Lemma 13 imply that for a.e. any $t \in J$, if $\gamma^C(t) \neq 0$ (resp. $\gamma^D(t) \neq 0$), then $\frac{d\gamma^C}{dt}(t) \neq 0$ (resp. $\frac{d\gamma^D}{dt}(t) \neq 0$).

- In particular, it follows that for a.e. $t \in [0, 1]$, it holds for at least one $j \in \{C, D\}$ that $\frac{d\gamma^j}{dt}(t) \neq 0$.

Define $G = (\gamma^C)^2 + (\gamma^D)^2$. We have deduced above that for at least one $j \in \{C, D\}$, γ^j is non-zero somewhere, and hence G is not uniformly 0. γ^C, γ^D are absolutely continuous and hence a.e. differentiable, hence, it holds a.e.,

$$G' = 2 \cdot \gamma^C \cdot \frac{d\gamma^C}{dt} + 2 \cdot \gamma^D \cdot \frac{d\gamma^D}{dt}$$

G is absolutely continuous, as both γ^C, γ^D are absolutely continuous (and bounded). Therefore, since $G' \geq 0$ a.e. and G is positive at some point, we deduce $G(d) > 0$, a contradiction.

7 Public-Signal Correlated Equilibrium

Fix a non-atomic probability space (Ω, ν). This will be the *signal space.*

7.1 Definition and Existence

A *public-signal correlated Markovian strategy* (or a *correlated strategy* for short) for Player p is a measurable[10] mapping $\sigma^p : Z \times [0, T] \times \Omega \to \prod_{p \in \mathcal{P}} \Delta(I^p)$. Intuitively, at each point in time $t \in [0, T]$, Player p observes a signal ω chosen from Ω, and plays in state $z \in Z$ the mixed action $\sigma^p(z, t, \omega)$.

[10]Where $[0, T]$ has the Lebesgue σ-algebra, Ω has the Borel σ-algebra, and $[0, T] \times \Omega$ has the induced product σ-algebra.

For a correlated strategy profile $\sigma = (\sigma^p)_{p\in\mathcal{P}}$, define for all $z \in Z$, $t \in [0,T]$,

$$u_\sigma(z,t) = \int_\Omega \sigma(z,t,\omega)d\nu(\omega).$$

u_σ is Lebesgue-measurable (by Fubini's theorem), but in general is not a Markovian strategy; that is, each $u_\sigma(z,t)$ is a distribution over $\prod_{p\in\mathcal{P}} I^p$, and it is not necessarily a product distribution. However, the transition rate function μ and the payoff function r extend linearly to $\Delta(\prod_{p\in\mathcal{P}} I^p)$, and hence P_{u_σ}, γ_{u_σ}, and X_{u_σ} can be defined as in Proposition 1, (2.2), and (3.2); we will denote them as P_σ, γ_σ, and X_σ for clarity, and indeed, the dynamics induced by u_σ is associated with σ. Note also that Proposition 4 holds for correlated strategy profiles, with u_σ replacing u. It is worth noting that if $\lambda : \Delta(\prod_{p\in\mathcal{P}} \Delta(I^p)) \to \mathbb{R}$ is any linear function (be it payoffs, transition rates, etc.), then for all $z \in Z$, $t \in [0,T]$,

$$\lambda(u_\sigma(z,t)) = \int_\Omega \lambda(\sigma(z,t,\omega))d\nu(\omega)$$

A public-signal correlated strategy σ is a *public-signal correlated Markovian equilibrium* (or a *correlated equilibrium* for short) if for every $z_0 \in Z$, every $p \in \mathcal{P}$ and every correlated strategy τ^p for Player p, we have

$$\gamma^p_\sigma(z_0) \geq \gamma^p_{(\tau^p,\sigma^{-p})}(z_0).$$

Theorem 5. *Let σ be a public-signal correlated strategy. Then σ is a public-signal correlated Markovian equilibrium iff for all $z \in Z$, a.e. $t \in [0,T]$, and ν-a.e. $\omega \in \Omega$,*

$$\sigma(z,t,\omega) \in NE\big(X_\sigma(z,t,\cdot)\big).$$

Proof. For a correlated strategy σ and a deviation τ^p for Player p, denote $\Xi = (\tau^p, \sigma^{-p})$; the parallel of (4.1) is:

$$\begin{aligned}\gamma^p_\sigma(0) - \gamma^p_\Xi(0) &= \int_0^T \Big[\int_\Omega P_\Xi(s)\big[X^p_\sigma(t,\sigma(t,\omega)) - X^p_\sigma(z,t,\Xi(t,\omega))\big]d\nu(\omega)\Big]dt \\ &= \int_0^T \Big[\int_\Omega P_\Xi(t)\big[r^p(\sigma(t,\omega)) + \mu(\sigma(t,\omega))\cdot\gamma^p_\sigma(t) \\ &\quad - r^p(\Xi(t,\omega)) - \mu(\Xi(t,\omega))\cdot\gamma^p_\sigma(t)\big]d\nu(\omega)\Big]dt \qquad (7.1)\end{aligned}$$

where $u(t,\omega) = (u(z,t,\omega))_{z\in Z}$, $X(t,a) = (X(z,t,a_z))_{z\in Z}$, and similarly for r^p, μ, as was the convention in Section 4. (7.1) can be established along the same lines as (4.1), using the fact that (3.3) now becomes

$$\frac{d\gamma^p_\sigma}{dt}(z,t) = -\int_\Omega \big[r^p(z,\sigma(z,t,\omega)) + \langle\mu(z,u(z,t,\omega)),\gamma_\sigma(t)\rangle\big]d\nu(\omega) \qquad (7.2)$$

and (2.1) becomes

$$\frac{d}{dt}P_u(s,t) = \int_\Omega P_u(s,t)\mu(u(t,\omega))d\nu(\omega) \tag{7.3}$$

Given (7.1) the proof now follows as in the proof Theorem 2.

□

Theorem 6. *Every continuous-time stochastic game of fixed duration possesses a public-signal correlated equilibrium.*

Proof. For each $z \in Z$, let $F_z : \mathbb{R}^{\mathcal{P}\times Z} \to \mathbb{R}^{\mathcal{P}}$ be the mapping which assigns to each $x = (x_z^p) \in \mathbb{R}^{\mathcal{P}\times Z}$,

$$F_z(x) = -NEP\Big(\big(r^p(z,\cdot) + \langle \mu(z,\cdot), x^p\rangle\big)_{p\in\mathcal{P}}\Big)$$

where NEP is the Nash-equilibrium payoff correspondence, let

$$G_z(x) = conv(F_z(x))$$

where $x^p = (x_z^p)_{z\in Z}$ and $conv(\cdot)$ denotes the convex hull, and let $G : \mathbb{R}^{\mathcal{P}\times Z} \to \mathbb{R}^{\mathcal{P}\times Z}$ be defined by $G(x) = \prod_{z\in Z} G_z(x)$. The correspondence G is easily seen to be upper semi-continuous (as is the Nash equilibrium payoff correspondence), and also for all $x \in \mathbb{R}^{\mathcal{P}\times Z}$,

$$||G(x)||_\infty := \sup_{y\in G(x)} ||y||_\infty \le ||r||_\infty + |Z|\cdot||\mu||_\infty \cdot ||x||_\infty$$

Under these conditions, existence theorems for solutions of differential inclusions - see, e.g., Theorem 5.1 of Deimling (1992) - guarantee that there exists an absolutely continuous mapping $\zeta : [0,T] \to \mathbb{R}^{Z\times\mathcal{P}}$ such that for a.e. $t \in [0,T]$ and each $z \in Z$,

$$\frac{d\zeta}{dt}(t) \in G(\zeta(t))$$

which satisfies $\zeta(T) = 0$. Let $\zeta = (\zeta_z)_{z\in Z}$ be such a solution, and let $S \subseteq [0,T]$ be a Borel set of full measure on which $\frac{d\zeta}{dt}(t)$ exists and the inclusion above holds. (From this point on, the proof is similar to the last part of the proof of the existence of stationary extensive-form correlated equilibrium in discrete-time stochastic games, Nowak and Raghavan (1992).) By the parametric version of Carathèeodory's theorem (see, e.g., Section 4.2 of Castaing and Valadier (1977)), there exist for each $z \in Z$ Borel measurable mappings $g_{z,0},\ldots,g_{z,n},\alpha_{z,0},\ldots,\alpha_{z,n} : S \to \mathbb{R}^n$ such that for all $t \in S$:

- $\sum_{j=0}^n \alpha_{j,z}(\omega) = 1$, and for all j, $a_{j,z} \ge 0$.
- For all $j = 0,\ldots,n$, $g_{z,j}(\omega) \in F_z(\omega)$.
- $\sum_{j=0}^n \alpha_{j,z}(\omega)g_{j,z}(\omega) = \frac{d\zeta_z}{dt}(t)$.

By the definition of F_z, we deduce from the measurable implicit function theorem (see, e.g., Section 3.6 of Castaing and Valadier (1977)) that there are Borel-measurable $f_{0,z}, \ldots, f_{n,z} : S \to \prod_{p \in \mathcal{P}} \Delta(I^p)$ such that for all $t \in S$, $j = 0, \ldots, n$, $p \in \mathcal{P}$, $z \in Z$,

$$g^p_{j,z}(t) = -r^p(z, f^p_{j,z}(t)) + \langle \mu(z, f^p_{j,z}(t)), \zeta^p(t) \rangle \tag{7.4}$$

and

$$(f^p_{j,z}(t))_{p \in \mathcal{P}} \in NE\Big(\big(r^p(z, \cdot) + \langle \mu(z, \cdot), \zeta^p(t) \rangle \big)_{p \in \mathcal{P}} \Big) \tag{7.5}$$

where recall that NE is the Nash equilibrium correspondence. Since every non-atomic probability space is isomorphic to the unit interval with Lebesgue measure, we may assume w.l.o.g. that $\Omega = [0, T]$, and ν is the Lebesgue measure. Define for $j = 0, \ldots, n$, $t \in S$,

$$\beta_j(t) = \sum_{i \leq j} \alpha_i(t)$$

and $\beta_{-1} \equiv 0$ (note that $\beta_n \equiv 1$), and then define for each[11] $t \in S$, $z \in Z$, $p \in \mathcal{P}$, $\omega \in \Omega (= [0, T])$,

$$\sigma^p(z, t, \omega) = f^p_{j,z}(t), \text{ if } \beta_{j-1}(t) < \omega \leq \beta_j(t)$$

It then follows from (7.4) that for each $z \in Z$ and a.e. $t \in [0, T]$,

$$\frac{d\zeta^p_z}{dt}(t) = -r^p(z, u_\sigma(z, t)) + \langle \mu(z, u_\sigma(z, t)), \zeta^p(t) \rangle \tag{7.6}$$

and also $\zeta(T) = 0$. Therefore, by part (a) of Theorem 1, $\zeta^p_z(t) = \gamma^p_{u_\sigma}(z, t)$, i.e., $\zeta^p_z(t) = \gamma^p_\sigma(z, t)$, for all $t \in [0, T]$, $z \in Z$, $p \in \mathcal{P}$. As such, it follows from the definition of β and σ, and from (7.5), that for a.e. $t \in [0, T]$, all $z \in Z$, $p \in \mathcal{P}$, $\omega \in \Omega$,

$$\sigma(z, t, \omega) \in NE\Big(\big(r^p(z, \cdot) + \langle \mu(z, \cdot), \gamma^p_\sigma(t) \rangle \big)_{p \in \mathcal{P}} \Big) \tag{7.7}$$

and hence, by Theorem 5, σ is a correlated equilibrium. □

7.2 Interpretation of Correlated Strategies

Given that equilibria may not exist in Markovian strategies but always exist in correlated strategies, and given that a continuum of i.i.d. random variables is a ill-defined concept (e.g., Judd (1985)), we will provide an interpretation of correlated strategies in continuous time. We proceed in a way similar to Section 5.

For any partition $\mathcal{T} = \{0 = t_0 < \cdots < t_k < T\}$ of $[0, T]$ and any finite partition[12] $\mathcal{Q}$ of Ω into sets of positive measure, let a $\mathcal{T} \times \mathcal{Q}$-correlated strategy of

[11] S is of full measure; σ can be defined arbitrarily outside of S.

[12] $\mathcal{Q}$ must cover Ω only up a ν-null set; this will be useful later when we mention an example of such a partition which satisfies additional conditions.

Player p be a mapping $w^p : Z \times \mathcal{T} \times \mathcal{Q} \to \Delta(I^p)$, where $\mathcal{Q}$ is viewed as a finite set. The interpretation is the following: At the beginning of each interval induced by the partition $\mathcal{T}$, an element of $\mathcal{Q}$ - interpreted as a "region" in Ω - is chosen by Nature by the distribution induced on it by ν - that is, $\nu(S) = \nu(\cup_{B\in S} B)$. Player p then performs his randomization on what pure action he would play at any state during that time interval, given the that the signal is in the element Nature chose, and then plays the same resulting pure action throughout that interval.[13] The collection of all such strategies for Player p will be denoted $\overline{\Sigma}^p(\mathcal{T},\mathcal{Q})$, and $\overline{\Sigma}(\mathcal{T},\mathcal{Q}) = \prod_{p\in\mathcal{P}} \overline{\Sigma}^p(\mathcal{T},\mathcal{Q})$. The collection of $\mathcal{T}\times\mathcal{Q}$-strategies for Player p which choose only pure actions (and, hence, can be identified with $(I^p)^{\mathcal{T}\times\mathcal{Q}}$) will be denoted $\overline{S}^p(\mathcal{T},\mathcal{Q})$, and $\overline{S}(\mathcal{T},\mathcal{Q}) = \prod_{p\in\mathcal{P}} \overline{S}^p(\mathcal{T},\mathcal{Q})$. An element $w^p \in \overline{\Sigma}^p(\mathcal{T},\mathcal{Q})$ determines a correlated strategy by

$$\hat{w}^p(z,t,\omega) = w^p(t_j,\omega) := \sum_{v\in\overline{S}^p(\mathcal{T})} w^p[v]\cdot v^p(t_j,z,[\omega]_{\mathcal{Q}}),\ t_j \le t < t_{j+1}. \tag{7.8}$$

where $[\omega]_{\mathcal{Q}}$ is the element of $\mathcal{Q}$ containing ω. $\hat{w}^p(z,t,\omega)$ can be interpreted as the expectation of the action of Player p used at time t in state z upon receiving signal θ if he uses the $\mathcal{T},\mathcal{Q}$-correlated strategy w^p.

For a partition $\mathcal{Q}$ of Ω, denote $\wp(\mathcal{Q}) = \max_{S\in\mathcal{Q}} \nu(S)$. For each player p and each Markovian strategy $u^p \in \Sigma^p$, define $w^p_{\mathcal{T},\mathcal{Q}} \in \overline{\Sigma}^p(\mathcal{T})$ by

$$w^p_{\mathcal{T},\mathcal{Q}}(z,t_j) = \frac{1}{(t_{j+1}-t_j)\cdot\nu(S)} \int_{t_j}^{t_{j+1}} \int_S u^p(z,t,\omega)dt\cdot d\nu(\omega), \tag{7.9}$$

where both sides are elements of $\Delta(I^p)$. We will call $w^p_{\mathcal{T},\mathcal{Q}}$ the $\mathcal{T},\mathcal{Q}$-adapted version of u^p. This terminology applies to profiles of $\mathcal{T},\mathcal{Q}$-strategies as well.

Payoffs under $\mathcal{T},\mathcal{Q}$-correlated strategy profiles, as well as when one player deviates to a general correlated strategy, are defined analogously to (5.3), (5.4), as well as to the notions of equilibria in $\Sigma_z(\mathcal{T},\mathcal{Q})$ and to the game $\Psi_z(\mathcal{T},\mathcal{Q})$.

Finally, we say that $(\mathcal{Q}_n)_{n=1}^\infty$ is a *regular sequence of partitions* of (Ω,ν) if:

- $\wp(\mathcal{Q}_n) \underset{n\to\infty}{\to} 0$.
- For ν-a.e. $\omega\in\Omega$, $\{\omega\} = \cap_n [\omega]_{\mathcal{Q}_n}$.

[13] The reason we discretize the signal space is because we will view a $\mathcal{T}\times\mathcal{Q}$-correlated strategy as a distribution over $\mathcal{T}\times\mathcal{Q}$-correlated pure strategies. If we did not discretize Ω, then the strategies for Player p in the time-discretized games would be distributions over mappings from $\mathcal{T}\times\Omega$ to I^p, and the former does not possess a "reasonable" Borel structure; see Aumann (1961). Alternatively, we could have turned the space of mappings from Ω into a standard Borel space by identifying maps that agree ν-a.e. and using the Borel structure induced by the weak-* topology.

- For any $f \in L^1(\Omega, \nu)$ it holds, for ν-a.e. $\omega \in \Omega$,

$$f(\omega) = \lim_{n\to\infty} \frac{\int_{[\omega]_{\mathcal{Q}_n}} f(z) d\nu(x)}{\nu([\omega]_{\mathcal{Q}_n})}.$$

For example, if Ω is a subset of a Euclidian space, ν is absolutely continuous w.r.t. to the Lebesgue-measure, and each $\mathcal{Q}_n$ consists of cubes (of the appropriate dimension), with volume going to 0 as $n \to \infty$. Also denote for correlated strategy profiles u, y, the $L^1([0,T], (\Omega, \nu))$-distance by

$$||u - y||_{L^1} = \int_0^T \int_{\omega\in\Omega} ||u(s,\omega) - y(s,\omega)||_\infty ds d\nu(\omega),$$

where $u(s,\omega) = (u(z,s,\omega))_{z\in Z}$.

Proposition 15. *Let $(\mathcal{T}_n)_{n=1}^\infty$ be a sequence of partitions with $\ell(\mathcal{T}_n) \underset{n\to\infty}{\to} 0$, and let $(\mathcal{Q}_n)_{n=1}^\infty$ be a regular sequence of partitions of (Ω, ν). Then for any profile of correlated strategies $p = (u^p)_{p\in\mathcal{P}}$, if w_n^p is the $\mathcal{T}_n, \mathcal{Q}_n$ version of u^p, then:*
(a) For a.e. $t \in [0,T]$, ν-a.e. $\omega \in \Omega$, and each $z \in Z$, $\hat{w}_n^p(z,t,\omega) \to u^p(z,t,\omega)$, and $\hat{w}_n^p(z,\cdot,\cdot) \to u^p(z,\cdot,\cdot)$ in $L^1([0,T], (\Omega,\nu))$.
(b) For each $z \in Z$,

$$\gamma_{w_n}(z) \underset{n\to\infty}{\to} \gamma_u(z).$$

(c) If u is a correlated equilibrium, then there is a sequence of positive numbers $(\varepsilon_n)_{n\in\mathbb{N}}$ with $\varepsilon_n \to 0$, such that for all z, w_n is a ε_n-equilibrium of $\Psi_z(\mathcal{T}_n, \mathcal{Q}_n)$ and of $\Gamma_z(\mathcal{T}_n, \mathcal{Q}_n)$.

The proof is similar to the proof of Theorem 4. Lemmas analogous to Lemmas 7, 8, and 9 hold, except that $\ell(\mathcal{T}) + \wp(\mathcal{Q})$ replaces $\ell(\mathcal{T})$, and they are proved along the same lines.

8 Continuous-Time Stochastic Games as Differential Games and the Hamilton–Jacobi–Bellman Equation

In the following sections, when we refer to differentiability on a closed set (resp. at a boundary point), we mean that the function can be extended to an open neighborhood of that set on which it is differentiable (resp. such that it is a differentiability point of the extended function.)

8.1 Differential Games

We recall a standard model of non-zero-sum continuous-time differential games, e.g., Chapter 8 of Friedman (1971), of which continuous-time stochastic games will then be a particular model of.

- There is a closed, convex set of states $X \subseteq \mathbb{R}^n$.
- A finite set of players $\mathcal{P}$.
- A compact, convex set of actions $J^p \subseteq \mathbb{R}^{n_p}$ for each $p \in \mathcal{P}$. Denote $J^{\mathcal{P}} = \prod_p I^p$.
- A duration $T > 0$.
- A payoff function $r : X \times [0, T] \times J^{\mathcal{P}} \to \mathbb{R}^{\mathcal{P}}$.
- A mapping $f : X \times J^{\mathcal{P}} \to \mathbb{R}^N$.
- An initial state $x_0 \in X$.

We assume that f and r satisfy:

- f, r are locally Lipschitz in the second coordinate, and on each compact set, the Lipschitz constant can be taken to be time-independent.
- There exists $\alpha, \beta > 0$ such that $|f(t, x, u)|, |r(t, x, u)| \leq (\alpha + \beta \cdot ||x||_\infty)$ for all t, x, u.
- For all $x \in X$, there is $\varepsilon > 0$ such that for all $u \in J^{\mathcal{P}}$ and all $0 \leq h < \varepsilon$, $x + h \cdot f(x, u) \in X$. We refer to this as the *viability condition.*

An (open-loop) strategy for Player p is a Lebesgue-measurable mapping $u^p : [0, T] \to J^p$; we do not consider more general strategies, and we denote the collection of such strategies $\mathcal{U}^p$, and $\mathcal{U} = \prod_p \mathcal{U}^p$. Given a profile $u = (u^p)_{p \in \mathcal{P}}$ of strategies, the dynamics of the game is given by:

$$\frac{dx}{dt}(t) = f(x(t), u(t)) \tag{8.1}$$

with initial state $x(0)$. By Theorem 1.1.3 in Friedman (1971), our assumptions imply that for each $u \in \mathcal{U}, t_0 \in [0, T], x_0 \in \mathbb{R}^n$, there is a unique absolutely continuous function satisfying (8.1) a.e. for which $x(t) = x$, which we denote x_{u,t_0,x_0}, or simply x_{u,x_0} when $t_0 = 0$; the conditions we have assumed also ensure that the dynamics never leave the set X.

Given a profile of strategies $u = (u^p)_{p \in \mathcal{P}}$ and an initial state x_0, the payoff to Player p is

$$\rho_u^p(x_0) = \int_0^T r^p(x_{u,x_0}(t), t, u(t))dt. \tag{8.2}$$

More generally, denote for $t \in [0, T]$ and $x \in X$,

$$\rho_u^p(x, t) = \int_t^T r^p(x_{u,t,x}(s), s, u(s))ds. \tag{8.3}$$

The *value function* associated with the profile u is given for each $p \in Pd$ by

$$V_u^p(x, t) = \max_{v^p \in \mathcal{U}^p} \rho^p_{v^p, u^{-p}}(x, t) \tag{8.4}$$

and $V_u = (V_u^p)_{p \in \mathcal{P}}$.

Remark 16. Note that V_u^p is independent of the strategy for Player p, and hence if there is only a single player, no strategy need be specified, and we will write $\mathcal{V}(x,t)$; single-player differential games are referred to as *optimal control problems.*

Proposition 17. *Let $\Upsilon(x_0)$ be an optimal control problem with associated value function $\mathcal{V}$, and assume that for almost every t, $\mathcal{V}$ is differentiable for **all** $x \in X$.*
(a) Let $W : [0,T] \times X \to \mathbb{R}$ be such that $W(T,\cdot) \equiv 0$, and for a.e. t, W is differentiable at (x,t) for all $x \in X$. Then $W = \mathcal{V}$ iff

- *W is Lipschitz.*
- *For a.e. $t \in [0,T]$ and all $x \in X$, W satisfies*

$$\frac{\partial W}{\partial t}(x,t) = -H(x,t,(\nabla_x W)(x,t)) \tag{8.5}$$

where $\nabla_x = (\frac{\partial}{x_1}, \ldots, \frac{\partial}{x_n})$, and the Hamiltonian H is defined by

$$H(z,t,q) = \max_{v \in J}[r(x,t,v) + \langle q, f(x,t,u)\rangle]$$

(b) $u \in \mathcal{U}$ is an equilibrium of $\Upsilon(x_0)$ iff for a.e. $t \in [0,T]$ and all $x \in X$,

$$u(t) \in argmax_{v \in J}[r(x,t,v) + \langle (\nabla_x W)(x,t), f(x,t,v)\rangle]. \tag{8.6}$$

It is worth noting that the strong differentiability condition we have assumed on the value function (instead of demanding differentiability for a.e.[14] point in $X \times [0,T]$, we demand that for a.e. $t \in [0,T]$, it is differentiable at (x,t) for all $x \in X$) is rarely satisfied in general differential games. Consequently, most studies of differential games make use of the viscosity solutions introduced by Crandall and Lions (1983). However, those differential games we will define to represent continuous-time stochastic games will satisfy this differentiability condition.

Proof. Part (a) follows from Theorem 5.7 of Frankowska et al (1995). Although the statement there involves notation suited to much more general optimal control problems, I will avoid introducing these notations here and simply say that under our regularity properties, the statement there indeed implies part (a). I will however remark that we do necessarily need to refer to a work such as Frankowska et al. (1995) which deals with control systems in which the dynamics need only depend on time in a measurable way, a departure from the classical case of dynamics that are jointly continuous in both time and position.

Part (b) follows along standard lines;[15] to apply these arguments, one needs to observe that since the mapping x_{u,x_0} is clearly Lipschitz and V is locally Lipschitz (this also follows from standard arguments), $t \to V(t, x_{u,x_0}(t))$ is also

[14] X is convex in a Euclidian space and hence has a natural Lebesgue measure.
[15] See, e.g., Bertsekas (2005).

Lipschitz and hence absolutely continuous. Therefore, by (8.5), it holds for a.e. $t \in [0,T]$,

$$\frac{\partial}{\partial t} V(t, x_{u,x_0}(t)) \leq -[r(t, x_{u,x_0}(t), u(t)) + \langle \nabla_x V(t, x_{u,x_0}(t)), f(t, x_{u,x_0}(t), u(t)) \rangle]$$

and we denote by S the set of t for which we have strict inequality; by the chain rule, we see that this is equivalent to that for a.e. $t \in [0,T]$,

$$\frac{d[V(t, x_{u,x_0}(t))]}{dt} \leq -r(t, x_{u,x_0}(t), u(t))$$

or, since $V(T, \cdot) \equiv 0$,

$$\int_0^T r(t, x_{u,x_0}(t), u(t)) \leq V(0,t)$$

and the last two inequalities are strict iff S has positive measure. □

8.2 Continuous-Time Stochastic Games as Differential Games

Suppose we are given a continuous-time stochastic game, as in Section 2, denoted Γ. For each $x_0 \in \Delta(Z)$, we define an auxiliary stochastic differential game Υ_{x_0} with the same set of players:

- $X = \Delta(Z)$, with initial state x_0.
- The action space J^p for Player p is $(\Delta(I^p))^Z$.
- The payoff function, which (by an abuse of notation) we will denote r, is defined for $u \in J = \prod_p J^p$ and $p \in \mathcal{P}$ by

  $$r^p(x,u) = \sum_{z \in Z} x[z] \cdot r^p(z, u_z),$$

  where $u_z = (u_z^p)_p$. Note that $r(\delta_z, u) = r(z, u_z)$, where δ_z is the Dirac measure at z (we will often identify z and δ_z).
- f is defined by

  $$f(x,u) = x' \cdot \mu(u),$$

  where $x' = (x[z])_{z \in Z}$ is viewed as a row vector.

It's easy to verify that f satisfies the conditions listed in Section 8.1, and in particular the viability condition: if $x \in \Delta(Z)$ satisfies $x_z = 0$ for some z, then $(f(x,\cdot))_z \geq 0$, while if $x_z = 1$, then $x_{z'} = 0$ for all $z' \neq z$ so $(f(x,\cdot))_z = \mu(z|z,\cdot) \leq 0$.

Note that we have a natural identification between strategies in this auxiliary game and Markovian strategies in the game of Section 2; in both cases, a strategy for Player p is a measurable mapping from $[0,T]$ to J^p - and hence we do not differentiate between them; i.e., we use the notations $\mathfrak{U}^p = \mathfrak{A}^p$, $\mathfrak{U} = \mathfrak{A}$. If $x_0 \in X$ is an initial state, u is a strategy, and $x(\cdot)$ satisfies $x(0) = x_0$ and for a.e. t,

$$\frac{dx}{dt}(t) = f(x, u(t))$$

then we have for all $t \in [0,1]$,

$$x(t) = x_0' \cdot P(0,t).$$

Furthermore, ρ_u, defined in (8.3), is affine in X.

Proposition 18. *For all $x \in X$, $t \in [0,T]$, and any strategy profile u,*

$$\rho_u(x,t) = \gamma_u(x,t),$$

where the left-hand side is defined in (8.3) and the right-hand side is the affine extension of the payoff function defined in (3.1). In particular,

$$V_u^p(x,t) = \max_{v^p} \gamma^p_{v^p,u^{-p}}(x,t) \tag{8.7}$$

where V_u^p was defined in (8.4). We now define

$$V_u^p(z,t) = \max_{v^p} \gamma^p_{v^p,u^{-p}}(z,t) \tag{8.8}$$

and if this definition is extended linearly to $\Delta(Z)$, it agrees with (8.7).

In particular:

Corollary 19. *Let $p \in \mathcal{P}$ and let u^{-p} be a profile of strategies for the other players. Then a strategy u^p of Player p is a best reply in the continuous-time stochastic game Γ iff it is a best reply in the corresponding differential game Υ_{x_0} for any initial state x_0 that has full support on Z. In particular, a profile $u = (u^p)_{p\in\mathcal{P}}$ is an equilibrium in Γ iff it is an equilibrium in Υ_{x_0} for any initial state x_0 that has full support on Z.*

8.3 The Hamilton–Jacobi–Bellman Equation

Theorem 7. *Let $u \in \mathcal{U}$ be a strategy profile.*
(a) Let $W : Z \times [0,T] \to \mathbb{R}^{\mathcal{P}}$ such that $W(\cdot,T) \equiv 0$, and for all $z \in Z$ and a.e. $t \in [0,T]$, W is differentiable at (z,t). Then $W = V_u$ iff for each $p \in \mathcal{P}$,

- *For each $z \in Z$, $p \in \mathcal{P}$, $W^p(z,\cdot)$ is Lipschitz.*
- *For a.e. $t \in [0,T]$ and each $p \in \mathcal{P}$, W^p satisfies*

$$\frac{\partial W^p}{\partial t}(z,t) = -H_u^p(z,t,(W^p(z,t))_z) \tag{8.9}$$

where the Hamiltonian H^p is defined by

$$H_u^p(z,t,q) = \max_{v\in J^p} [r^p(z,v,u^{-p}(t)) + \langle q, \mu_z(v,u^{-p})\rangle]$$

where $\mu_z(\cdot) = (\mu_{z,z'}(\cdot))_{z'\in Z}$.

(b) $u \in \mathcal{U}$ is an equilibrium of Γ iff for a.e. $t \in [0,T]$ and all $p \in \mathcal{P}$,

$$u^p(z,t) \in argmax_{u \in \Delta(I^p)}[r^p(z,u,u^{-p}(z,t)) + \langle (V_u^p(z,t))_{z\in Z}, \mu_z(u,u(t)^{-p})\rangle] \tag{8.10}$$

We mention two preliminary lemmas. We deduce the following lemma using standard techniques:

Lemma 20. *For each profile of strategies u, Player $p \in \mathcal{P}$, and state z, $V_u^p(z,\cdot)$ is Lipschitz.*

The following lemma is immediate:

Lemma 21. *Let $p \in \mathcal{P}$, and let be $\eta : X \times J^p \to \mathbb{R}$ be continuous, such that*

$$\eta(x,u) = \langle x, (\eta(z,u_z))_{z\in Z}\rangle.$$

Then for any $x_0 \in \Delta(Z)$ with full support and $u = (u_z)_{z\in Z} \in J^p$,

$$u \in argmax_{u\in J^p}\eta(x_0,u)$$

iff for all $z \in Z$,

$$u_z \in argmax_{u\in\Delta(I^p)}\eta(z,u_z).$$

Furthermore, the mapping $\psi : X \to \mathbb{R}$ defined by

$$\psi(x) = \max_{u\in J^p}\eta(x,u)$$

is affine.

Proof. (of Theorem 7). Fix some $x_0 \in \Delta(Z)$ with full support, and observe the game Υ_{x_0}. Fix a Player p and a profile of strategies u^{-p} for the other players, and observe optimal control problem $\Upsilon_{x_0}^p$ resulting from Υ_{x_0} when all players but p are required to play u^{-p}. In this case, $\mathcal{V}(x,t) = V_{u,u^{-p}}^p(x,t)$ for any strategy $u \in \mathcal{U}^p$, where $\mathcal{V}(x,t)$ was defined in Remark 16.

Since $\mathcal{V}$ is Lipschitz by Lemma 20, it is differentiable almost everywhere; since it is multilinear over $\Delta(Z)$, it follows that for a.e. t, $\mathcal{V}$ is differential at (x,t) for all $x \in \Delta(Z)$. The theorem now follows from Proposition 17, Lemma 21, Corollary 19, and the observation that if a function $G : [0,T] \times X \to \mathbb{R}$ is of the form

$$G(x,t) = \langle x, (G(z,t))_{z\in Z}\rangle$$

then

$$\nabla_x G(x,t) = (G(z,t))_{z\in Z}.$$

□

9 Extensions

We mention here several elementary ways in which the model presented in Section 2 can be generalized such that the results of this paper still apply.

Firstly, one can allow the payoffs and transition rates to be time-dependent. That is, in Section 2, we could have specified the payoffs and transitions to be of the following form:

- A bounded Lebesgue-measurable payoff function $r : Z \times [0,T] \times I^{\mathcal{P}} \to \mathbb{R}^{\mathcal{P}}$.
- A bounded[16] Lebesgue-measurable transition rate $\mu : Z \times Z \times [0,T] \times I^{\mathcal{P}} \to \mathbb{R}$, where for all $a \in I^{\mathcal{P}}$, $z \in Z$, and $t \in [0,T]$, $\sum_{z' \in Z} \mu(z'|z,t,a) = 0$, and for all $z' \neq z$, $\mu(z'|t,z,a) \geq 0$.

All the results in this paper remain correct under the appropriate notational changes.

Secondly, one can allow for a terminal payoff. That is, one can add to the components given in Section 2 a terminal payoff function $g : Z \to \mathbb{R}^{\mathcal{P}}$. The payoff for the game is then defined to be the accumulated payoff over the course of the game plus the terminal payoff in state $z(T)$; explicitly, (2.2) becomes

$$\gamma_u^p(z_0) = E_u^{z_0}\big[\int_0^T r^p(z(t), u(z(t),t))dt + g^p(z(T))\big].$$

All of the results in this paper remain correct under this change as well, except that now one must modify Theorems 1 and 7 by altering the boundary condition to $\gamma_u(z,T) = g(z)$ and $V_u(z,T) = g(z)$, respectively. (The proofs in Section 5 also become a little more involved.)

Finally, one can allow for the action spaces of the players $I^1, \ldots, I^p$ to be compact metric spaces. In this case, we require the payoff function r and the transition rate function μ to be (jointly) continuous in the actions.[17] With the exceptions of Sections 5 and 7.2, which give interpretations of randomized/correlated strategies in continuous-time, the results of this paper hold with only very minor changes in the proofs. With some care, also in the case of compact action spaces, one can provide these notions with interpretations similar to those given in the aforementioned sections by discretizing the action spaces as well as time (and, in the latter case, the signal space).

[16] It's not clear whether the boundedness of the transitions rates in this case can be replaced with some integrability condition.

[17] If we allow for r, μ to depend measurably on time, then we require a Carathéodory type of condition: for each fixed point in time, r, μ are continuous in the actions, and for each fixed action profile, r, μ are measurable in time.

10 Appendix A: Proofs from Section 5

Proof. (of Lemma 7) Note first that, $||A \cdot B||_\infty \leq C_\infty ||A||_\infty \cdot ||B||_\infty$ for an appropriate $C_\infty > 0$ and for any two $|Z| \times |Z|$ matrices A, B.[18] Further note that $||P_u||_\infty \leq 1$ for any Markovian strategy u. Using these observations and (2.1) gives:

$$
\begin{aligned}
f(t) &:= ||P_u(t) - P_y(t)||_\infty = || \int_0^t [P_u(s)\mu(u(s)) - P_y(s)\mu(y(s))]ds||_\infty \\
&\leq \int_0^t ||P_u(s)\mu(u(s)) - P_y(s)\mu(y(s))||_\infty ds \\
&\leq \int_0^t ||P_u(s)\mu(u(s)) - P_y(s)\mu(u(s))||_\infty ds \\
&\quad + \int_0^t ||P_y(s)\mu(u(s)) - P_y(s)\mu(y(s))||_\infty ds \\
&\leq C_\infty \cdot ||\mu||_\infty \cdot \int_0^t ||P_u(s) - P_y(s)||_\infty ds \\
&\quad + C_\infty ||P_y(s)||_\infty \cdot ||\mu||_\infty \cdot \int_0^t ||u(s) - y(s)||dt \\
&\leq C_\infty \int_0^t f(t) + C_\infty \cdot ||\mu||_\infty \cdot ||u - y||_{L^1}
\end{aligned}
$$

Therefore, an application of Gronwall's inequality - e.g., Hirsch and Smale, (1974) - gives (5.6) for appropriate $M > 0$. To prove (5.7), note that (3.1) shows that for each player $p \in \mathcal{P}$ and $z \in Z$,

$$
\begin{aligned}
|\gamma_u^p(z) - \gamma_y^p(z)| &= | \int_0^T [\langle P_u^z(t), r^p(u(s))\rangle - \langle P_y^z(t), r^p(y(s))\rangle]ds| \\
&\leq \int_0^T |\langle P_u^z(t), r^p(u(s))\rangle - \langle P_y^z(t), r^p(u(s))\rangle| \\
&+ \int_0^T |\langle P_y^z(t), r^p(u(s))\rangle - \langle P_y^z(t), r^p(y(s))\rangle| \\
&\leq ||r||_\infty \cdot \int_0^T ||P_u^z(s) - P_y^z(s)||_\infty ds + ||r||_\infty ||P_y||_\infty \int_0^T ||u(s) - y(s)||_\infty ds \\
&\leq ||r||_\infty \cdot M \cdot (\int_0^T e^{M \cdot s} ds) \cdot ||u - y||_{L^1} \cdot e^{M \cdot T} + ||r||_\infty \cdot ||u - y||_{L^1}
\end{aligned}
$$

and so enlarging M appropriately completes the proof of (5.7). □

Proof. (of Lemma 8) For the simplicity of notation, we will prove only (a); (b) is proved by an almost identical argument. Let $\mathcal{T} = \{0 = t_0 < \cdots < t_k < T\}$, fix

[18] For the L^2-induced operator norm, $||A \cdot B||_{L^2} \leq ||A||_{L^2} \cdot ||B||_{L^2}$, and all norms on a Euclidian space are equivalent.

τ satisfying $\ell(\mathcal{T}) < \tau < \min[T, 2\ell(\mathcal{T})]$, and let $0 = m_0 < m_1 < \cdots < m_n = k+1$ (recall $t_{k+1} = T$ be such that $\tau < t_{m_j} - t_{m_{j-1}} < 2\tau$ for all $1 \leq j < n$; such a selection is possible. Denote:

$$D_j := \sup_{t \in \{t_0, \cdots, t_{m_j}\}} ||P_u(t) - P_w(t)||_\infty$$

Clearly $D_0 = 0$. We wish to bound D_n. Fix $1 \leq j < n$, let $\mathcal{T}|_j = \{t_0, \dots, t_{m_j - 1}\}$ and $\mathcal{T}|_{j+} = \{t_{m_j}, \dots, t_{m_{j+1}-1}\}$; these can be viewed as restrictions of the partition to $[0, t_{m_j}]$ and to $[t_{m_j}, t_{m_{j+1}}]$, respectively. We will denote by $v|_j$, $v|_{j+}$ generic elements of the sets $S(\mathcal{T}|_j) \cong \prod_{p \in \mathcal{P}} (I^p)^{Z \times \mathcal{T}|_j}$, $S(\mathcal{T}|_{j+}) \cong \prod_{p \in \mathcal{P}} (I^p)^{Z \times \mathcal{T}|_{j+}}$, respectively. Note that each $w \in \Sigma(\mathcal{T})$ induces an elements of $\Sigma(\mathcal{T}|_j) \cong \prod_{p \in \mathcal{P}} (\Delta(I^p))^{Z \times \mathcal{T}|_j}$, $\Sigma(\mathcal{T}|_{j+}) \cong \prod_{p \in \mathcal{P}} (\Delta(I^p))^{Z \times \mathcal{T}|_{j+}}$ in the natural way, and this induced distribution is a product distribution on $\Sigma(\mathcal{T}|_{j+1}) = \Sigma(\mathcal{T}|_j) \otimes \Sigma(\mathcal{T}|_{j+})$. Note that for $t \leq t_{m_j}$, $P_{v|_j}(t)$ is well-defined, and for $t_{m_j} \leq t < t_{m_{j+1}}$ and any $z \in Z$, $\hat{v}|_{j+}(z, t)$ is well-defined, see (5.1). Also observe that by the definition of w, for any $t_{m_j} \leq p < q < t_{m_{j+1}}$, we have for all $z \in Z$,

$$\int_{t_p}^{t_q} \mu(u(z,s)) ds = \sum_{v|_{j+} \in \mathcal{T}|_{j+}} \int_{t_p}^{t_q} w[v|_{j+}] \cdot \mu(\hat{v}|_{j+}(z,s)) ds \tag{10.1}$$

where we recall that a pure strategy $v \in S(\mathcal{T})$ induces a pure Markovian strategy by (5.1). Finally, we note that for any Markovian strategy ϕ,

$$||P_\phi(t) - P_\phi(s)||_\infty \leq \int_s^t ||P_\phi(s) \mu(\phi(s))||_\infty ds \leq C_\infty ||\mu||_\infty (t - s) \tag{10.2}$$

where C_∞ is as in the proof of Lemma 7. Hence, letting $t_{m_j} \leq p < t_{m_{j+1}}$, and denoting $\tau_0 = t_{m_j}$, $\tau = t_p - \tau_0$, we have

$$\begin{aligned}
||P_u(t_p) - P_w(t_p)||_\infty &= || \int_0^\tau [P_u(\tau_0 + s) \mu(u(\tau_0 + s)) \\
&\quad - \sum_{v \in S(\mathcal{T}|_{j+1})} w[v] P_{\hat{v}}(\tau_0 + s) \mu(\hat{v}(\tau_0 + s))] ds ||_\infty \\
&\leq C_\infty ||\mu||_\infty \int_0^\tau ||P_u(\tau_0 + s) - P_u(\tau_0)||_\infty \\
&\quad + \sum_{v \in S(\mathcal{T}|_{j+1})} w[v] \cdot C_\infty ||\mu||_\infty \int_0^\tau ||P_{\hat{v}}(\tau_0 + s) - P_{\hat{v}}(\tau_0)||_\infty \\
&\quad + || \int_0^\tau [P_u(\tau_0) \mu(u(\tau_0 + s)) - \sum_v w[v] P_{\hat{v}}(\tau_0) \mu(\hat{v}(\tau_0 + s))] ds ||_\infty
\end{aligned} \tag{10.3}$$

By (10.2), the first two terms together are bounded by $C_\infty (||\mu||_\infty)^2 \tau^2$. As for the second term, note that for $0 \leq s \leq \tau$,

$$\sum_v w[v] P_{\hat{v}}(\tau_0) \mu(\hat{v}(\tau_0 + s)) = \Big[\sum_{v|_j \in S(\mathcal{T}|_j)} w[v|_j] \cdot P_{v|_j}(\tau_0) \Big] \cdot \Big[\sum_{v|_{j+} \in S(\mathcal{T}|_{j+})} w[v|_{j+}] \cdot \mu(\hat{v}|_{j+}(\tau_0 + s) \Big]$$

And hence, using (10.1),

$$
\begin{aligned}
&|| \int_0^\tau [P_u(\tau_0)\mu(u(\tau_0+s)) - \sum_v w[v]P_{\hat{v}}(\tau_0)\mu(\hat{v}(\tau_0+s))]ds||_\infty \\
&\quad = || \int_0^\tau P_u(\tau_0)\mu(u(\tau_0+s))ds \\
&\quad - \sum_{v|_j \in S(\mathfrak{T}|_j)} w[v|_j]\cdot P_{v|_j}(\tau_0)\cdot [\int_0^\tau \sum_{v|_{j+}} w[v|_{j+}]\cdot \mu(\hat{v}|_{j+}(\tau_0+s)ds]||_\infty \\
&\quad = || \int_0^\tau P_u(\tau_0) \int_0^\tau \mu(u(\tau_0+s))ds \\
&\quad - \sum_{v|_j \in S(\mathfrak{T}|_j)} w[v|_j]\cdot P_{v|_j}(\tau_0)\cdot \int_0^\tau \mu(u(\tau_0+s))||_\infty \\
&\quad \le \tau\cdot C_\infty||\mu||_\infty \cdot ||P_u(\tau_0) - \sum_{v|_j \in S(\mathfrak{T}|_j)} w[v|_j]\cdot P_{v|_j}(\tau_0)||_\infty \\
&\quad \le \tau\cdot C_\infty||\mu||_\infty\cdot D_j \qquad (10.4)
\end{aligned}
$$

Hence, (10.3) and (10.4) imply:

$$D_{j+1} \le \max[D_j, C_\infty(||\mu||_\infty)^2\tau^2 + D_j\cdot\tau\cdot C_\infty||\mu||_\infty]$$

with initial condition $D_0 = 0$. In particular, if $\tau < \frac{1}{C_\infty||\mu||_\infty}$, we have

$$D_{j+1} \le D_j + C_\infty(||\mu||_\infty)^2\tau^2$$

Since $\ell(\mathfrak{T})\cdot\tau < t_{m_{j+1}} - t_{m_j}$ for all $0 \le j < n$, we have $n < \frac{T}{\ell(\mathfrak{T})}$. Therefore,

$$D_n \le C_\infty(||\mu||_\infty)^2\tau^2 n \le \frac{TC_\infty(||\mu||_\infty)^2\tau^2}{\ell(\mathfrak{T})}$$

and since $\tau \le 2\ell(\mathfrak{T})$, this gives, denoting $C = 4TC_\infty(||\mu||_\infty)^2$,

$$D_n \le C\cdot\ell(\mathfrak{T}) \qquad (10.5)$$

This was under the assumption that $\tau < \frac{1}{C_\infty||\mu||_\infty}$, which can arranged as long as $\ell(\mathfrak{T}) < \frac{1}{C_\infty||\mu||_\infty}$; but C can be enlarged so that and (10.5) holds for all values of $\ell(\mathfrak{T})$. Finally, from (10.2), we have

$$||P_{v^p,u}(t) - P_{v^p,w}(t)||_\infty \le D_n + C_\infty||\mu||_\infty\cdot\ell(\mathfrak{T})$$

which completes the proof of (5.8), with $K = C + C_\infty||\mu||_\infty$. Now we prove prove (5.9). Fix $0 \le j < n$, and denote $\tau = t_{m_{j+1}} - t_{m_j}$. Using (5.8) and techniques similar to those above,

$$
\begin{aligned}
&\Big|\int_{t_{m_j}}^{t_{m_{j+1}}}[\langle P_u^z(t),r^p(u(s))\rangle - \sum_{v\in S(\mathfrak{T})} w[v]\langle P_v^z(t), r^p(\hat{v}(s))\rangle]ds\Big| \\
&\qquad\le |Z|\cdot||r^p||_\infty\cdot\tau^2 + \Big|\int_0^\tau[\langle P_u^z(t_{m_j}), r^p(u(t_{m_j}+s))\rangle \\
&\qquad - \sum_{v\in S(\mathfrak{T})} w[v]\langle P_v^z(t_{m_j}), r^p(\hat{v}(t_{m_j}+s))\rangle]ds\Big| \\
&\qquad = |Z|\cdot||r^p||_\infty\cdot\tau^2 + \Big|\int_0^\tau[\langle P_u^z(t_{m_j}), r^p(u(t_{m_j}+s))\rangle \\
&\qquad - \sum_{v|_j\in S(\mathfrak{T}|_j)} w[v|_j] \sum_{v|_{j+}\in S(\mathfrak{T}|_{j+})} w[v|_{j+}]\langle P_v^z(t_{m_j}), r^p(\hat{v}(t_{m_j}+s))\rangle]ds\Big| \\
&\qquad = |Z|\cdot||r^p||_\infty\cdot\tau^2 + \Big|\int_0^\tau[\langle P_u^z(t_{m_j}), r^p(u(t_{m_j}+s))\rangle \\
&\qquad - \langle P_w^z(t_{m_j}), r^p(\hat{w}(t_{m_j}+s))\rangle]ds\Big| \\
&\qquad \le |Z|\cdot||r^p||_\infty\cdot\tau^2 + |Z|\cdot||r^p||_\infty\cdot K\cdot\ell(\mathfrak{T})\cdot\tau \\
&\qquad + \Big|\int_0^\tau[\langle P_u^z(t_{m_j}), r^p(u(t_{m_j}+s))\rangle \\
&\qquad - \langle P_u^z(t_{m_j}), r^p(\hat{w}(t_{m_j}+s))\rangle]ds\Big| \\
&\qquad \le |Z|\cdot||r^p||_\infty\cdot\tau^2 + |Z|\cdot||r^p||_\infty\cdot K\cdot\ell(\mathfrak{T})\cdot\tau \\
&\qquad + |Z|\cdot||r||_\infty\cdot\int_{t_{m_j}}^{t_{m_{j+1}}}||u(s)-\hat{w}(s)||_\infty ds
\end{aligned}
$$

Summing over $j = 0,\dots,n-1$ and recalling $n < \frac{T}{\ell(\mathfrak{T})}$, $\tau \le 2\ell(\mathfrak{T})$, gives:

$$
\begin{aligned}
|\gamma_u^p(z) - \gamma_v^p(z)| &= \Big|\int_0^T[\langle P_u^z(t), r^p(u(s))\rangle - \langle P_v^z(t), r^p(\hat{v}(s))\rangle]ds\Big| \\
&\le \frac{T}{\ell(\mathfrak{T})}\Big(|Z|\cdot||r^p||_\infty\cdot 4(\ell(\mathfrak{T}))^2 + 2\cdot K||r^p||_\infty\cdot(\ell(\mathfrak{T}))^2\Big) \\
&\quad + |Z|\cdot||r||_\infty\cdot||u-\hat{w}||_{L^1} \\
&4T\cdot|Z|\cdot||r^p||_\infty\ell(\mathfrak{T}) + 2T\cdot K\cdot\ell(\mathfrak{T}) + |Z|\cdot||r||_\infty\cdot||u-\hat{w}||_{L^1}
\end{aligned}
$$

Enlarging K accordingly gives (5.9). □

Proof. (of Lemma 9) First, we prove

$$
\sup_{u^p\in\mathfrak{A}^p}\gamma^p_{u^p,w^{-p}} \le \sup_{w^p\in\Sigma^p(\mathfrak{T})}\gamma^p_{w^p,w^{-p}} + C\cdot\ell(\mathfrak{T})
$$

for appropriate $C > 0$. Given $w^p \in \Sigma^p(\mathfrak{T})$, set $u = \hat{w}$; by parts (a) and (b) of Lemma 8,

$$
||\gamma^p_{u^p,w^{-p}} - \gamma^p_{w^p,w^{-p}}||_\infty \le K\ell(\mathfrak{T}) + K||u-\hat{w}||_{L^1} = K\ell(\mathfrak{T})
$$

Next we prove,

$$\sup_{w^p \in \Sigma^p(\mathcal{T})} \gamma^p_{w^p, w^{-p}} \leq \sup_{u^p \in \mathfrak{A}^p} \gamma^p_{u^p, w^{-p}} + C \cdot \ell(\mathcal{T})$$

for appropriate $C > 0$. We will actually prove a bit more. We show that it suffices to take only pure Markovian strategies in the supremum on the right-hand side. Fix $w^p \in \Sigma^(\mathcal{T})$. It suffices to show that there is a pure Markovian strategy u^p satisfying

$$|\gamma^p_{w^p, w^{-p}} - \gamma^p_{u^p, w^{-p}}| \leq C \cdot \ell(\mathcal{T}) \tag{10.6}$$

for some $C > 0$ which is independent of w^p or of $\mathcal{T}$. For each interval J induced by $\mathcal{T}$, divide J into subintervals (some of which may be degenerate) $I_1, \ldots, I_{|I^p|}$, with $\ell(I_a) = \frac{\ell(I) \cdot w^p[a]}{n}$, where $\ell(\cdot)$ will be used to denote the length of an interval and a denotes the a-th element of I^p in some ordering. Note then that by the definition of u^p, and denoting $w = (w^p, w^{-p})$, we have for any interval $J = [\tau_0, \tau_1]$ in the partition induced by $\mathcal{T}$,

$$\begin{aligned} \ell(J) \cdot \sum_{a \in I^{\mathcal{P}}} w(\tau_0)[a] r^p(a) ds &= \int_J \sum_{a \in I^{\mathcal{P}}} w(\tau_0)[a] r^p(a) ds \\ &= \int_J \sum_{a \in I^{-p}} w^{-p}(\tau_0)[a^{-p}] r^p(u^p(s), a^{-p}) ds \end{aligned} \tag{10.7}$$

Note also that w^p is the $\mathcal{T}$-adapted version of u^p. Combining (5.8) and (5.10) of Lemma 8 gives:

$$\begin{aligned} |P_{w^p, w^{-p}} - P_{u^p, w^{-p}}| &\leq |P_{w^p, w^{-p}} - P_{u^p, \hat{w}^{-p}}| \\ &+ |P_{u^p, \hat{w}^{-p}} - P_{u^p, w^{-p}}| \leq 2K \cdot \ell(\mathcal{T}) \end{aligned} \tag{10.8}$$

for K as in Lemma 8. As a result, for each interval $J = [\tau_0, \tau_1]$ induced by the partition $\mathcal{T}$, we have, using (10.2) and techniques similar to those used in the proof of Lemma 8,

$$
\begin{aligned}
&\Big|\int_0^{\ell(J)} \big[\langle P^z_w(\tau_0+s), \sum_{a\in I^{\mathcal{P}}} w(\tau_0)[a]\cdot r^p(a)\rangle - \\
&\qquad \langle P^z_{u^p,w^{-p}}(\tau_0+s), \sum_{a\in I^{-p}} w(\tau_0)[a]\cdot r^p(u^p(\tau_0+s),a)\rangle\big]ds\Big| \\
&\qquad \le |Z|\cdot C_\infty||\mu||_\infty\cdot ||r||_\infty(\ell(J))^2 + \Big|\int_0^{\ell(J)} \big[\langle P^z_w(\tau_0), \sum_{a\in I^{\mathcal{P}}} w(\tau_0)[a]\cdot r^p(a)\rangle \\
&\qquad - \langle P^z_{u^p,w^{-p}}(\tau_0), \sum_{a\in I^{-p}} w(\tau_0)[a]\cdot r^p(u^p(\tau_0+s),a)\rangle\big]ds\Big| \\
&\qquad \le |Z|\cdot C_\infty||\mu||_\infty\cdot ||r||_\infty(\ell(J))^2 + \Big|\int_0^{\ell(J)} \big[\langle P^z_w(\tau_0), \sum_{a\in I^{\mathcal{P}}} w(\tau_0)[a]\cdot r^p(a)\rangle \\
&\qquad - \langle P^z_{u^p,w^{-p}}(\tau_0), \sum_{a\in I^{\mathcal{P}}} w(\tau_0)[a]\cdot r^p(a)\rangle\big]ds\Big| \\
&\qquad + \Big|\int_0^{\ell(J)} \big[\langle P^z_{u^p,w^{-p}}(\tau_0), \sum_{a\in I^{\mathcal{P}}} w(\tau_0)[a]\cdot r^p(a)\rangle \\
&\qquad - \langle P^z_{u^p,w^{-p}}(\tau_0), \sum_{a\in I^{-p}} w(\tau_0)[a]\cdot r^p(u^p(\tau_0+s),a)\rangle\big]ds\Big| \\
&\qquad \le |Z|\cdot C_\infty||\mu||_\infty\cdot ||r||_\infty(\ell(J))^2 + |Z|\cdot||r||_\infty\cdot||P_w - P^z_{u^p,w^{-p}}||\cdot\ell(J) \\
&\qquad + |Z|\cdot||P||_\infty\cdot\Big|\int_0^{\ell(J)} \big[\sum_{a\in I^{\mathcal{P}}} w(\tau_0)[a]\cdot r^p(a) \\
&\qquad - \sum_{a\in I^{-p}} w(\tau_0)[a]\cdot r^p(u^p(\tau_0+s),a)\big]\Big| \\
&\qquad \le |Z|\cdot C_\infty||\mu||_\infty\cdot ||r||_\infty(\ell(J))^2 + |Z|\cdot||r||_\infty(\ell(J))^2
\end{aligned}
$$

where the last inequality is because of (10.8) and (10.7). Denoting $L = |Z|\cdot(C_\infty||\mu||_\infty + 1)\cdot||r||_\infty$ and summing over all J induced by the partition induced by $\mathcal{T}$ gives

$$
||\gamma^p_w - \gamma^p_{u^p,w^{-p}}|| \le \sum_J L\cdot(\ell(J))^2 \le L\cdot(T\cdot\ell(\mathcal{T}) + (\ell(\mathcal{T}))^2)) \le 2L\cdot T\cdot\ell(\mathcal{T})
$$

where $\lceil\cdot\rceil$ denotes the rounding-up function, and we have the used the following elementary claim: For any $T, D > 0$,

$$
\max_{\{\sum_{i=1}^n a_i = T, \forall i, 0\le a_i\le D\}} \sum_{i=1}^n a_i^2 \le \lceil\frac{T}{D}\rceil\cdot D^2 \le T\cdot D + D^2
$$

Indeed, in our case, $\ell(J)\le\ell(\mathcal{T})$ for each J. $\square$

11 Appendix B: Construction for Section 6.2

11.1 Preliminaries

In addition to the notation introduced by 6.2, we introduce:

- For a normal-form game Λ, $NE(\Lambda)$ is the set of Nash equilibria of Λ.
- We let S denote the boundary of the square,

$$S = \{(p,q) \mid -1 \leq p, q \leq 1,\ (|p| = 1) \vee (|q| = 1)\}. \tag{11.1}$$

  We denote the four closed edges of S by $S_{\mathcal{N}}, S_{\mathcal{E}}, S_{\mathcal{S}}, S_{\mathcal{W}}$ for the north, east, south, and west edges, respectively. Note that $S_{\mathcal{N}} = -S_{\mathcal{S}}$, $S_{\mathcal{E}} = -S_{\mathcal{W}}$.

The multi-player normal game is built around a 'base' normal-form game G_0 with the following properties:

(1) The set of equilibria $NE_0 = NE(G_0)$ contains a unique hyperstable set H_0. By a hyperstable set, defined in [19], we mean a set that is minimal w.r.t. the following property: for any $\varepsilon > 0$, there is $\delta > 0$ such that the equilibria of any game G' that is in a δ-neighborhood of a game G that is equivalent[19] to G_0 are in an ε-neighborhood of H_0.

(2) H_0 is connected but not simply connected.

(3) Furthermore, there exists:

- A continuous semi-algebraic injection $\psi : S \to H_0$, which is not nulhomotopic in H_0.
- A semi-algebraic retract $\rho : NE_0 \to \psi(S)$.
- For all $\varepsilon > 0$, a semi-algebraic mapping Γ_ε from S to the ε-neighborhood of G_0, such that for each edge of S and any equilibria of any game in $\Gamma_\varepsilon(E)$ is in an ε-neighborhood of $\rho^{-1}(\psi(-E))$.

Remark 22. For later purposes, we remark that the upper-semicontinuity of the Nash equilibrium correspondence implies that for each $\varepsilon > 0$ there exists $\eta = \eta(\varepsilon)$ such that if $||H - G_0||_\infty < \eta$, then $NE(H)$ is contained in the ε-neighborhood of $NE(G_0)$.

Properties (1)-(3) would be enough for us to build a normal-form game with the desired properties. Yet, relying on these properties alone would require the some very cumbersome and technical machinery. For the sake of simplicity, we take advantage of some further properties of the game given in Appendix B of [19]:

[19] Two games are equivalent if they have the same reduced form, where the reduced form is achieved by eliminating actions that are payoff-equivalent to a convex combination of other actions.

The Game G_0			
$A\backslash B$	L	M	R
L	$1,1$	$0,-1$	$-1,1$
M	$-1,0$	$0,0$	$-1,0$
R	$1,-1$	$0,-1$	$-2,-2$

Table 11.2.a

Equilibria of G_0

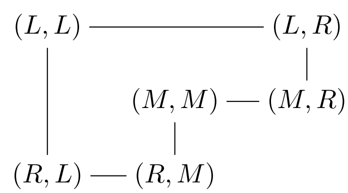


Figure 11.2.b

(11.2)

The additional properties of G_0 are the following:

(4) There are two players, A, B, with action spaces $I = \{T, M, B\}$, $J = \{L, C, R\}$, respectively.

(5) The set of Nash equilibria is hyperstable and homeomorphic to S.

(6) For each $\varepsilon > 0$, the maps[20] ψ and Γ_ε can be taken to be piecewise linear,[21] 8ε-Lipshitz, and satisfying the following property:

$$||\Gamma_\varepsilon - G_0||_\infty < 2\varepsilon, \tag{11.3}$$

such that for any edge E of S, and for any equilibrium (x, y) of any game in $\Gamma_\varepsilon(E)$ it holds that

$$||E_{x\otimes y}[\vartheta] - (-E)||_\infty < 4\varepsilon, \tag{11.4}$$

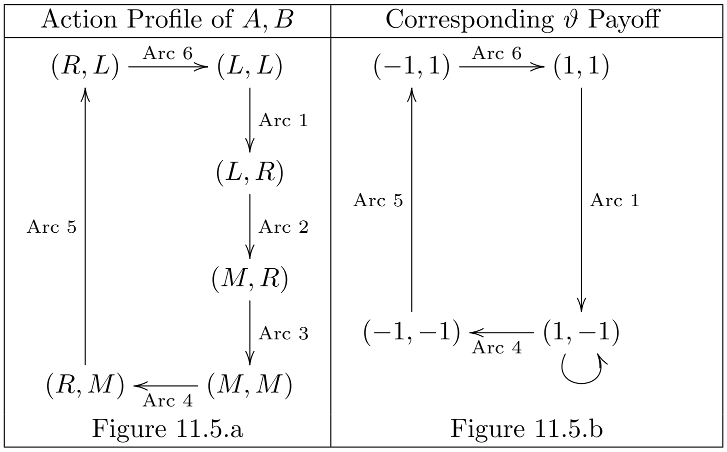
where ϑ is defined by

$\vartheta :=$

$A\backslash B$	L	M	R
L	$1,1$	$0,0$	$1,-1$
M	$0,0$	$1,-1$	$1,-1$
R	$-1,1$	$-1,-1$	$0,0$

ϑ can be understood graphically:

[20]In this case, ψ is a homeomorphism, and the retract ρ is the identity.
[21]In the sense that each edge of the square is viewed as an interval.



(11.5)

The latter part of (6) can be stated informally: For any equilibria of a game assigned to a point on E via Γ_ε, the expected payoff under ϑ is not too far from the edge opposite E. Indeed, in Section 11.3 we verify that property[22] (6) holds for the game given in (11.2). We remark that property (4) only serves to make concrete the notation in property (6) and below, and is completely irrelevant otherwise.

The following proposition is proved in Appendix B of Levy (2012):

Proposition 23. *Let I, J be finite sets*[23]*, and let $Q : S \to \mathbb{R}^{2|I\times J|}$ be a continuous and piecewise linear*[24] *map to bimatrix games on these action sets. Then for some integer M, there exist 4 normal-form games on the set of players $A, B, \theta^1, \ldots, \theta^M$, denoted $\mathfrak{K}^k$ for $k \in \{1, -1\}^2$, such that:*

(I) A, B have action spaces I, J respectively; each θ^j has an action space $\{L, R\}$. The players $\{\theta^1, \ldots, \theta^M\}$ will be called auxiliary players.

(II) The payoffs of $\theta^1, \ldots, \theta^M$ are not affected by the actions of A, B in any of the games; let $\mathfrak{K}^k_\Theta$ denote the well-defined restriction of $\mathfrak{K}^k$ to the Players $\Theta^1, \ldots, \Theta^M$.

(III) For $(p, q) \in [-1, 1]^2$, let $\mathfrak{K}(p, q)$ (resp. $\mathfrak{K}_\theta(p, q)$) denote the convex combination of the $\{\mathfrak{K}^k\}_k$ (resp. $\{\mathfrak{K}^k_\Theta\}_k$), with weights given by[25] *$(\frac{1+p}{2}, \frac{1-p}{2}) \otimes (\frac{1+q}{2}, \frac{1-q}{2})$. If $(p, q) \in S$, and a_θ is an equilibrium in the game $\mathfrak{K}_\Theta(p, q)$, then the expected payoff matrix facing A, B, given by $\mathfrak{K}^{A,B}(p, q)(\cdot, a_\theta)$, is $Q(p, q)$.*

[22] Where ψ is described by Figure 11.5.a., and ρ is the identity.

[23] The proposition also extends with almost no change in the proof to the case that Q is a map to games with any finite set of players.

[24] I.e., piecewise linear on each edge of S.

[25] $(\phi, 1-\phi)$ denotes the probability distribution choosing 1 with probability ϕ, and choosing -1 with probability $1-\phi$.

(IV) *If L denotes a Lipshitz constant of Q and if $||Q(p,q) - Q_0||_\infty \leq \kappa$ for some Q_0, some κ, and all $(p,q) \in S$, then*

$$||\mathfrak{K}^{A,B}(p,q)(\cdot, a_\theta) - Q_0||_\infty \leq L \cdot \kappa,\ \forall (p,q) \in [-1,1]^2, \forall a_\theta \in NE(\mathfrak{K}_\Theta(p,q)) \tag{11.6}$$

11.2 The Normal-Form Game

We now turn to our normal-form game, and prove the properties given in Section 6.2. Fix $\varepsilon \leq \min[\frac{1}{16}, \eta(\frac{1}{4})]$, where $\eta(\cdot)$ is defined in Remark 22. The payoff is dependent on a parameter $\omega = (\omega^C, \omega^D) \in \mathbb{R}^2$:

- The Players are $A, B, \theta^1, \ldots, \theta^M$, where M as in Proposition 23 for the function Γ_ε, as well as Players C, D.
- As in Proposition 23, Players A, B have action sets $I = \{T, M, B\}$, $J = \{L, C, R\}$, and each player θ^j has action sets $\{L, R\}$; furthermore, Players C, D have action sets $\{1, -1\}$.
- The payoff r_ω, will be the sum of two payoffs, $r_\omega := r_1 + r_{2,\omega}$, defined separately as follows:
- The first payoff function r_1 satisfies $r_1^{C,D}(a) = G^{C,D}(a^{A,B}) := \vartheta[a^{A,B}]$, where ϑ is defined in property (6) of Section 11.1, and the payoff to the other players is the same as in the game of Proposition 23 when the profile $a^{-\{C,D\}}$ is played and the choice $a^{C,D} \in \{+1, -1\}^2$ is made by Nature; namely,

$$r_1^{C,D}(a) = G^{C,D}(a^{A,B}) := \vartheta[a^{A,B}],\quad r_1^{-\{C,D\}}(a) = \mathfrak{K}^{a^{C,D}}(a^{-\{C,D\}})$$

- The second payoff function $r_{2,\omega}$ is dependent on ω. It gives a payoff of 0 to all players other than C, D: That is, $r_{2,\omega}^{-\{C,D\}} \equiv 0$. To players C, D, $r_{2,\omega}$ is dependent only on $a^{C,D}$ and is given by:

$r_{2,\omega}^{C,D}(a) =$

$C \backslash D$	1	-1
1	ω^C, ω^D	$\frac{1}{2}\omega^C, \frac{1}{2}\omega^D$
-1	$\frac{1}{2}\omega^C, \frac{1}{2}\omega^D$	0

For each $(p,q) \in S$, let $a_{p,q}$ be an equilibrium profile in the game with payoff r_1 for the players $A, B, \theta^1, \ldots, \theta^M$ when Players C, D are restricted to playing $b_{p,q} := (\frac{1+p}{2}, \frac{1-p}{2}) \otimes (\frac{1+q}{2}, \frac{1-q}{2})$; that is $a_{p,q}$ is an equilibrium in $r_1^{-\{C,D\}}(\cdot, b_{p,q})$. By applying property (III) of Proposition 23 to the mapping Γ_ε which has the properties given in property (6), we get with the help of Figure 11.5 the following relationship between p, q and the payoff in r_1 to C, D under the profile $a_{p,q}$, $r_1^{C,D}(a_{p,q}, b_{p,q})$:

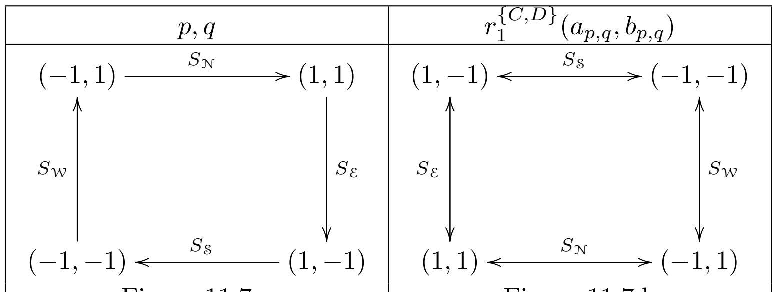


Figure 11.7.a Figure 11.7.b

(11.7)

Intuitively, as the point (p, q) goes around the square, the payoff $r_1^{C,D}(a_{p,q}, b_{p,q})$ (which is not uniquely determined) must also go 'around' the square 'close to it' - at a distance of at most 4ε from the edge on which (p, q) lies, because of (11.4).

Proposition 24. *Let $\omega \in \mathbb{R}^2$, and let a be an equilibrium profile in the game r_ω. Denote $p = 2a^C[1] - 1$, $q = 2a^D[1] - 1$. Then:*

(i) If $\omega^C > 0$, then $p = 1$; if $\omega^C < 0$, then $p = -1$. The same holds for q w.r.t. ω^D.

(ii) Hence, if $\omega^C > 0$, then $r_1^C(a) \le -\frac{1}{2}$. If $\omega^C < 0$, then $r_1^C(a) \ge \frac{1}{2}$. Similarly, if $\omega^D > 0$, then $r_1^D(a) \le -\frac{1}{2}$. If $\omega^D < 0$, then $r_1^D(a) \ge \frac{1}{2}$.

(iii) Let H be the expected matrix facing players A, B; that is, $H = r_\omega^{A,B}(\cdot, a^{-\{A,B\}})$. Then $||H - G_0||_\infty < \varepsilon$ (regardless of the values of ω^C, ω^D; this includes the case where one or both are 0), and $r_1^{C,D}(a) \neq 0$.

Proof. The first part follows from (11.4), and because C, D consider only the payoff from $r_{2,\omega}$ when making a decision. The second part follows from the first part and from (11.4), since we had chosen $\varepsilon \le \frac{1}{16}$. For the last part, first note that since $\varepsilon \le \frac{1}{16}$ and Γ_ε is 8ε-Lipshitz, Γ_ε is $\frac{1}{2}$-Lipshitz; hence, from property IV of Proposition 23 and (11.3), we see that $||H - G_0||_\infty < \varepsilon$. Finally, since $\varepsilon \le \eta(\frac{1}{4})$, with η as in Remark 22, and $\max \vartheta - \min \vartheta = 2$, we see that $r_1^{C,D}(a) = \vartheta[a^{A,B}]$ is in the $\frac{1}{2}$-neighborhood of the square S. □

11.3 Construction from Kohlberg and Mertens' Game

Let G_0 be the game defined in Figure 11.2; let $E_1, \ldots, E_6$ denote the 6 equilibria, beginning with (L, L) and proceeding clockwise, and let A_i denote the arc from E_i to $E_{i+1, mod\ 6}$. Also, for a two-player game G, the game G', defined by $G'^i(a, b)) = G^{3-i}(b, a)$, is the game where the players and action profiles are switched.

Fix $\varepsilon > 0$; we begin by defining mappings $G^1, \ldots, G^6, G^Z : [0, 1] \to \mathbb{R}^{2 \times I \times J}$, and from these we will define Γ_ε.

- $$G_1(t) :=$$

$A\backslash B$	L	M	R
L	$1+\varepsilon, 1+(1-t)\varepsilon$	$\varepsilon, -1$	$-1+\varepsilon, 1+t\cdot\varepsilon$
M	$-1, (1-t)\varepsilon$	$0, 0$	$-1, t\cdot\varepsilon$
R	$1, -1$	$0, -1$	$-2, -2$

  All equilibria in $G_1(t)$ lie on the arc A_1.

- $$G_2(t) :=$$

$A\backslash B$	L	M	R
L	$1+(1-t)\varepsilon, 1$	$(1-t)\varepsilon, -1$	$-1+(1-t)\varepsilon, 1+\varepsilon$
M	$-1, 0$	$t\cdot\varepsilon, 0$	$-1+t\cdot\varepsilon, \varepsilon$
R	$1-t\cdot\varepsilon, 0$	$-t\cdot\varepsilon, -1$	$-2, -2$

  All equilibria of $G_2(t)$ lie along A_2.

- $$G_3(t) :=$$

$A\backslash B$	L	M	R
L	$1, 1-2t\cdot\varepsilon$	$-t\cdot\varepsilon, -1$	$-1, 1-2(t-\frac{1}{2})\varepsilon$
M	$-1, -t\cdot\varepsilon$	$\varepsilon, t\cdot\varepsilon$	$-1+\varepsilon, -2(t-\frac{1}{2})\varepsilon$
R	$1-\varepsilon, -1$	$-\varepsilon, -1+t\cdot\varepsilon$	$-2, -2$

  All equilibria of $G_3(t)$ lie along A_3.

- $$G_Z(t) =$$

$A\backslash B$	L	M	R
L	$1-2t\varepsilon, 1-2(1-t)\varepsilon$	$-\varepsilon, -1$	$-1, 1-\varepsilon$
M	$-1, -\varepsilon$	ε, ε	$-1+\varepsilon, -\varepsilon$
R	$1-\varepsilon, -1$	$-\varepsilon, -1+\varepsilon$	$-2, -2$

  For $t < \frac{1}{2}$ or $t > \frac{1}{2}$, the unique equilibrium of $G_4(t)$ is (M, M).

  $$G_Z(\frac{1}{2}) =$$

$A\backslash B$	L	M	R
L	$1, 1$	$-\varepsilon, -1$	$-1, 1-\varepsilon$
M	$-1, -\varepsilon$	ε, ε	$-1+\varepsilon, -\varepsilon$
R	$1-\varepsilon, -1$	$-\varepsilon, -1+\varepsilon$	$-2, -2$

  which has pure equilibria (L, L) and (M, M), and the mixed equilibrium,

  $$(x^*, y^*) = \big((\frac{\varepsilon}{1+\varepsilon}, \frac{1}{1+\varepsilon}, 0), (\frac{\varepsilon}{1+\varepsilon}, \frac{1}{1+\varepsilon}, 0)\big) \tag{11.8}$$

  which satisfies $||(x^*, y^*) - (M, M)||_\infty = \frac{\varepsilon}{1+\varepsilon} < \varepsilon$.

- Since $G_3(0) = G'_Z(1)$, retrace our steps in the transposed games; we get

  $$G_4(t) := G'_3(1-t)$$

$$G_5(t) := G_2'(1-t)$$

$$G_6(t) := G_1'(1-t)$$

In each of these cases, all equilibria of G_j lie along A_j.

We then define

$$\Gamma_\varepsilon(p,q) = \begin{cases} G_4(\frac{1}{2}(1+p)) & \text{if } q=1 \\ G_6(\frac{1}{2}(1-q)) & \text{if } p=1 \\ G_6(\frac{1}{2}(1-p)) & \text{if } q=-1 \\ G_1(2(p+1)) & \text{if } q=-1, p \le -\frac{1}{2} \\ G_2(2(p+\frac{1}{2})) & \text{if } q=-1, -\frac{1}{2} \le p \le 0 \\ G_3(2p) & \text{if } q=-1,\ 0 \le p \le \frac{1}{2} \\ G_Z(2(p-\frac{1}{2})) & \text{if } q=-1,\ \frac{1}{2} \le p \le 1 \end{cases} \tag{11.9}$$

To see this more clearly, denote $H_j = G_j(0)$ for $j = 1, \ldots, 6, Z$. Then the map Γ_ε is the piecewise linear map given by the following diagram:

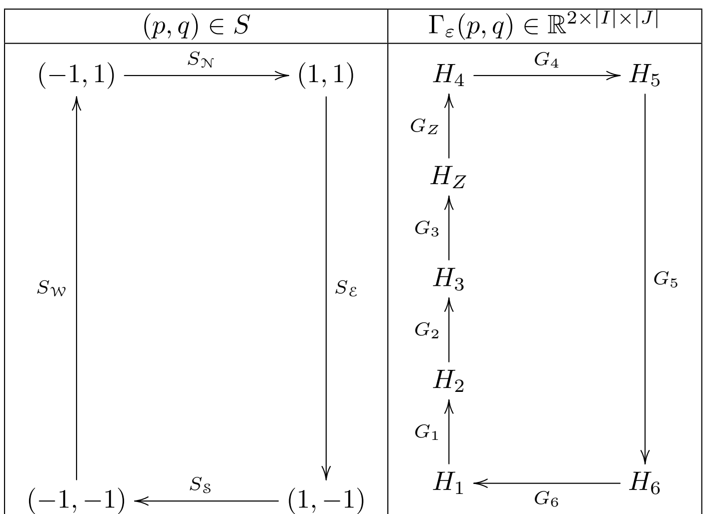


Figure 11.10.a Figure 11.10.b

(11.10)

From these figures and the explicit forms $G_1, \ldots, G_6, G_Z$ and their equilibria properties listed above, it is immediate that Γ_ε satisfies property (6) in Section 11.1.

References

[1] Aumann, R.J. (1961), Borel Structures for Function Spaces, *Illinois J. Math.*, 5, 614–630.

[2] Barron, E.N., Evans, L.C. and Jensen, R. (1984), Viscosity Solutions of Isaacs' Equation and Differential Games with Lipschitz Controls, *J. of Diff. Eq.* 53, 213–233.

[3] Bertsekas, D. (2005), *Dynamic Programming and Optimal Control*, Vol. 1, Athena Scientific, Belmont, MA.

[4] Bressan, A. (2011), Noncooperative Differential Games, *Milan J. Math*, 79, 357–427.

[5] Castaing C. and Valadier, M. (1977) , *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Mathematics 580, Springer–Verlag, New York.

[6] Coddington, E. and Levinson, N. (1972), *Theory of Ordinary Differential Equations*, McGraw-Hill Inc., NY.

[7] Crandall, M. and Lions, P. (1983), Viscosity solutions of Hamilton-Jacobi equations, *Trans. of the AMS*, 277, 1-42.

[8] Deimling, K. (1992), *Multivalued Differential Equations*, Walter de Gruyer & Co., Berlin.

[9] Frankowska, H., Plaskacz, S., and Rzezuchowski, T. (1995), Measurable Viability Theorems and the Hamilton-Jacobi-Bellman Equations, *J. of Diff. Eq.*, 116, 265–305

[10] Friedman, A. (1971), *Differential Games*, Pure and Applied Mathematics, Vol. 25, John Wiley and Sons, Inc.

[11] Guo, X. and Hernández-Lerma, O. (2003), Zero-Sum Games for Continuous-Time Markov Chains with Unbounded Transitions and Average Payoff Rates, *J. Appl. Prob.*, 40, 327–345.

[12] Guo, X. and Hernández-Lerma, O. (2005), Nonzero-Sum Games for Continuous-Time Markov Chains with Unbounded Discounted Payoffs, *J. Appl. Prob.*, 42, 303–320.

[13] Harsanyi,J.C., and Selten, R. (1988), *A General Theory of Equilibrium Selection in Games*, MIT Press.

[14] Hellwig, M. and Leininger, W. (1988), Markov-Perfect Equilibrium in Games of Perfect Information, Discussion Paper A-183, University of Bonn.

[15] Himmelberg, C.J. (1975), Measurable relations, *Fund. Math.* 87, 53-72.

[16] Hirsch, M. and Smale, S. (1974), *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press, San Diego, CA.

[17] Isaacs, R. (1965), *Differential Games: A Mathematical Theory With Applications to Warfare and Pursuit, Control and Optimization*, John Wiley and Sons, Inc.

[18] Judd, K. L. (1985), The Law of Large Numbers with a Continuum of IID Random Variables, *J. of Econ. Theory*, 35, 19–25.

[19] Kohlberg, E. and Mertens. J.F. (1986), On the Strategic Stability of Equilibria, *Econometrica*, 54, 1003–1037.

[20] Krasovskii, N.N. and Subbotin, A.I. (1988), *Game Theoretical Control Problems*, Springer-Verlag, NY.

[21] Kuratowski, K. and Ryll-Nardzewski, C. (1965), A General Theorem on Selectors, *Bull. of Pol. Acad. Sci. Math.)* 13, 379-403

[22] Levy, Y. (2012), A Discounted Stochastic Game with No Stationary Nash Equilibrium:The Case of Absolutely Continuous Transitions. DP #612, Center for the Study of Rationality, Hebrew University, Jerusalem.

[23] Maskin, E. and Tirole, J. (2001), Markov Perfect Equilibrium: I. Observable Actions, *J. Econom. Theory*, 100, 191–219.

[24] Miller, B. (1967), Finite State Continuous-Time Markov Decision Processes with Applications to a Class of Optimization Problems in Queueing Theory, Technical Report 15, Stanford University.

[25] Miller, B. (1968), Finite State Continuous Time Markov Decision Processes with a Finite Planning Horizon, SIAM J. Control, 6, 266–280.

[26] Neyman, A. (2012), Continuous–Time Stochastic Games. DP #616, Center for the Study of Rationality, Hebrew University, Jerusalem.

[27] Nowak, A.S. and Raghavan, T.E.S. (1992), Existence of Stationary Correlated Equilibria with Symmetric Information for Discounted Stochastic Games, *Math. Oper. Res.* , 17, 519-526.

[28] Rieder, U. (1979), Equilibrium Plans for Non-Zero-Sum Markov Games, in *Game Theory and Related Topics* (O. Moeschlin and D. Pallaschke, Eds.), Amsterdam, North Holland, pp. 91-102.

[29] Rudin, W. (1973), *Functional Analysis*, McGraw-Hill Inc., New York, NY.

[30] Shapley, L. (1953), Stochastic Games, *Proc. Nat. Acad.Sci. USA* 39, 1095–1100.

[31] Zachrisson, L. E. (1964), Markov Games, in M. Dresher, L. S. Shapley, and A. W. Tucker (eds.), *Advances in Game Theory*, Princeton University Press, Princeton, New Jersey, pp. 211–253.