TECHNICAL RESEARCH REPORT

Stochastic Differential Games with Multiple Modes

by M.K. Ghosh, S.I. Marcus

T.R. 95-68

INSTITUTE FOR SYSTEMS RESEARCH

Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry

STOCHASTIC DIFFERENTIAL GAMES WITH MULTIPLE MODES

Mrinal K. Ghosh

Department of Mathematics
Indian Institute of Science
Bangalore 560012, India
mkg@math.iisc.ernet.in

and

Steven I. Marcus

Electrical Engineering Department and
Institute for Systems Research
University of Maryland
College Park, Maryland 20742
marcus@src.umd.edu

ABSTRACT

We have studied two person stochastic differential games with multiple modes. For the zero-sum game we have established the existence of optimal strategies for both players. For the nonzero sum case we have proved the existence of a Nash equilibrium.

Key Words. Controlled diffusions, Markov chains, Markov strategy, Optimal strategy, Isaacs equation, Nash equilibrium.

INTRODUCTION

We study a two-person stochastic differential game with multiple modes. The state of the system at time t is given by a pair $(X(t), \theta(t)) \in \mathcal{R}^d \times S, S = \{1, 2, \cdots, N\}$. The discrete component $\theta(t)$ describes the various modes of the system. The continuous component $X(t)$ is governed by a "controlled diffusion process" with a drift vector which depends on the discrete component $\theta(t)$. Thus $X(t)$ switches from one diffusion path to another at random times as the mode $\theta(t)$ changes. On the other hand, the discrete component $\theta(t)$ is a "controlled Markov chain" with a transition rate matrix depending on the continuous component. The evolution of the process $(X(t), \theta(t))$ is governed by the following equations

$$dX(t) = b(X(t), \theta(t), u_1(t), u_2(t))dt + \sigma(X(t), \theta(t))dW(t),$$

$$P(\theta(t + \delta t) = j \mid \theta(t) = i, X(s), \theta(s), s \leq t) = \lambda_{ij}(X(t))\delta t + o(\delta t), i \neq j,$$

for $t \geq 0, X(0) = x \in \mathcal{R}^d, \theta(0) = i \in S,$ where b, σ, λ are suitable functions,

$$\lambda_{ij} \geq 0, i \neq j, \sum_{j=1}^{N} \lambda_{ij} = 0,$$

$W(\cdot)$ is a standard Brownian motion, $u_1(\cdot)$ and $u_2(\cdot)$ nonanticipative processes taking values in prescribed sets U_1 and U_2, respectively, which are admissible strategies for players 1 and 2, respectively. This kind of differential game was first studied by Basar and Haurie [1] for a piecewise deterministic case ($\sigma \equiv 0$). They have studied feedback Stackelberg and Nash equilibria for a nonzero-sum game. Here we consider both zero-sum and nonzero-sum games. In a zero-sum game player 1 is trying to maximize his expected (discounted) payoff

$$E\ [\int_0^\infty\ e^{-\alpha t}\ r(X(t), \theta(t), u_1(t), u_2(t))dt]$$

over his admissible strategies, where $\alpha > 0$ is the discount factor and r is the payoff function. Player 2 is trying to minimize the same over his admissible strategies. This kind of game typically occurs in a pursuit-evasion problem where an interceptor tries to destroy a specific target. Due to quick maneuvering by the evader and the corresponding response by the interceptor the trajectories keep switching rapidly and is generally modelled as a hybrid system $(X(t), \theta(t))$ described above [12]. Though pursuit-evasion games are generally treated on a finite horizon, we have studied the infinite horizon case here. The corresponding results for the finite horizon case can be derived using analogous (almost identical) arguments. For the zero-sum game we have established existence of randomized optimal strategies for both players. We have then treated a special case where at each mode only one player controls the game (think of billiards, for example). For this special case we have shown the existence of optimal pure strategies. This may have potential applications in token ring networks [4, p. 253]. In such networks each node may be treated as a player and at any time the node having the token controls the game completely. We next consider the nonzero-sum game where each player is trying to maximize his own payoff. For player $\ell, \ell = 1, 2$, the payoff is

$$E \, [\int_0^\infty \; e^{-\alpha t} \; r_\ell(X(t), \theta(t), u_1(t), u_2(t))dt].$$

This kind of game arises in a situation where two economic agents share the same production system which is subject to random failure [10]. For a nonzero-sum game we have established the existence of (Nash) equilibrium. Our results for a nonzero-sum game extends to the several players case. We have treated two players only for notational simplicity.

Our paper is organized as follows. Section 2 describes the technical details of the problems. Zero-sum game is treated in Section 3. Section 4 deals with nonzero-sume case.

PROBLEM DESCRIPTION

Let $V_\ell = 1, 2$ be compact metric spaces and $U_\ell = \mathcal{P}(V_\ell)$ the space of probability measures on V_ℓ endowed with the topology of weak convergence. Let $V = V_1 \times V_2$ and $U = U_1 \times U_2$. Let $S = \{1, 2, \cdots, N\}$. Let

$$\bar{b} : \mathcal{R}^d \times S \times V \to \mathcal{R}^d, \quad \bar{b}(\cdot, \cdot, \cdot) = [\bar{b}_1(\cdot, \cdot, \cdot), \cdots, \bar{b}_d(\cdot, \cdot, \cdot)]'$$

$$\sigma : \mathcal{R}^d \times S \to \mathcal{R}^{d \times d}, \sigma(\cdot, \cdot) \quad [\sigma_{ij}(\cdot, \cdot)], 1 \le i, j \le d.$$

$$\lambda_{ij} : \mathcal{R}^d \to \mathcal{R}, \quad 1 \le i, j \le N, \lambda_{ij}(\cdot) \ge 0, i \ne j, \quad \sum_{j=1}^N \; \lambda_{ij}(\cdot) = 0.$$

We make the following assumptions on $\bar{b}, \sigma, \lambda$.

(A1) (i) For each $i \in S$, $\bar{b}(\cdot, i, \cdot, \cdot)$ is bounded, continuous and Lipschitz in its first argument uniformly with respect to the rest.
(ii) For each $i \in S$, $\sigma(\cdot, i)$ is bounded and Lipschitz with the least eigenvalue of $\sigma\sigma'(\cdot, i)$ uniformly bounded away from zero.
(iii) For $i, j \in S, \lambda_{ij}(\cdot)$ is bounded and Lipschitz continuous.

For $x \in \mathcal{R}^d, \lambda \in S$, $(u_1, u_2) \in U = U_1 \times U_2$, define

$$b_k(x,i,u_1,u_2) = \int_{V_1} \int_{V_2} \bar{b}_k(x,i,v_1,v_2) u_1(dv_1) u_2(dv_2), k = 1,\cdots,d$$

and

$$b(x,i,u_1,u_2) = [b_1(x,i,u_1,u_2),\cdots,b_d(x,i,u_1,u_2)]'.$$

Let $(X(\cdot),\theta(\cdot))$ be an $\mathcal{R}^d \times S$-valued process given by

$$dX(t) = b(X(t),\theta(t),u(t))dt + \sigma(X(t),\theta(t))dW(t) \tag{2.1a}$$

$$P(\theta(t+\delta t) = j \mid \theta(t) = i, \theta(s), X(s), s \leq t) = \lambda_{ij}(X(t))\delta t + 0(\delta t), i \neq j \tag{2.1b}$$

$$X(0) = x \in \mathcal{R}^d \ \theta(0) = i \in S. \tag{2.1c}$$

Here, $W(\cdot) = [W_1(\cdot),\cdots,W_d(\cdot)]'$ is a standard Wiener process, $u(\cdot) = (u_1(\cdot), u_2(\cdot))$ where $u_\ell(\cdot)$ is a U_ℓ-valued nonanticipative process (see [8] for a precise definition of nonanticipativity). The process $u_\ell(\cdot), \ell = 1,2$, as above is called an admissible strategy for player ℓ. If $u_\ell(\cdot) = v_\ell(x(\cdot),\theta(\cdot))$ for a measurable $v_\ell : \mathcal{R}^d \times S \to U_\ell$, then $u_\ell(\cdot)$ (or by an abuse of notation the map v_ℓ itself) is called a Markov strategy for the ℓth player. A strategy $u_\ell(\cdot)$ is called pure if u_ℓ is a Dirac measure, *i.e.*, $u_\ell(\cdot) = \delta_{v_\ell}(\cdot)$, where $v_\ell(\cdot)$ is a V_ℓ-valued nonanticipative process. If for each $\ell = 1,2, u_\ell(\cdot)$ is a Markov strategy then (2.1) admits a unique strong solution which is a strong Feller process [8]. Let A_ℓ, M_ℓ and D_ℓ denote the sets of arbitrary admissible, Markov and Markov pure ("deterministic") strategies, respectively, for player ℓ. For $p \geq 1$ define

$$W^{2,p}_{loc}(\mathcal{R}^d \times S) = \{f : \ \mathcal{R}^d \times S \to \mathcal{R} \ : \ \text{for each } i \in S, \ f(\cdot,i) \in W^{2,p}_{loc}(\mathcal{R}^d)\}.$$

$W^{2,p}_{loc}(\mathcal{R}^d \times S)$ is endowed with the product topology of $(W^{2,p}_{loc}(\mathcal{R}^d))^N$. Similarly, we define $\mathcal{D}(\mathcal{R}^d \times S), \mathcal{D}'(\mathcal{R}^d \times S)$, etc. For $f \in W^{2,p}_{loc}(\mathcal{R}^d \times S)$ and $v = (v_1,v_2) \in V_1 \times V_2$, we write

$$L^{v_1,v_2} f(x,i) = L_i^{v_1,v_2} f(x,i) + \sum_{j=1}^{N} \lambda_{ij} f(x,j) \tag{2.2}$$

where

$$L_i^{v_1,v_2} f(x,i) = \sum_{j=1}^{d} \bar{b}_j(x,i,v_1,v_2) \frac{\partial f(x,i)}{\partial x_j} + \frac{1}{2} \sum_{j,k=1}^{d} a_{jk}(x,i) \frac{\partial^2 f(x,i)}{\partial x_j \partial x_k} \tag{2.3}$$

$$a_{jk}(x,i) = \sum_{\ell=1}^{d} \sigma_{j\ell}(x,i)\ \sigma_{k\ell}(x,i).$$

For $(u_1, u_2) \in U_1 \times U_2$ we define

$$L^{u_1,u_2} f(x,i) = \int_{V_1} \int_{V_2} L^{v_1,v_2} f(x,i) u_1(dv_1) u_2(dv_2). \tag{2.4}$$

Zero-Sum Game.

Let $\bar{r} : \mathcal{R}^d \times S \times V_1 \times V_2 \to \mathcal{R}$ be the payoff function. We assume that

(A2) $\bar{r}$ is a bounded, continuous function, Lipschitz in its first argument uniformly with respect to the rest.

When the state of the system is $(x,i) \in \mathcal{R}^d \times S$ and the players 1, 2 choose the actions $v_1 \in V_1, v_2 \in V_2$, respectively, player 1 receives a payoff $\bar{r}(x,i,v_1,v_2)$ from player 2. The problem is to find a strategy for player 1 which will maximize his accumulated income and a strategy for player 2 which will minimize the same quantity. For performance evaluation, we consider the discounted payoff on the infinite horizon. Let $\alpha > 0$ be the discount factor. Let $(u_1, u_2) \in A_1 \times A_2$ and $(X(\cdot), \theta(\cdot))$ the solution of (2.1) corresponding to this pair of strategies. The α-discounted payoff to player 1 for the initial condition (x,i) is defined as follows:

$$R[u_1, u_2]\ (x,i) = E_{x,i}^{u_1,u_2}\ [\int_0^\infty\ e^{-\alpha t}\ r(X(t), \theta(t), u_1(t), u_2(t))dt] \tag{2.5}$$

where

$$r : \mathcal{R}^d \times S \times U_1 \times U_2 \to \mathcal{R}$$

is defined as

$$r(x,i,u_1,u_2) = \int_{V_1} \int_{V_2} \bar{r}(x,i,v_1,v_2) u_1(dv_1) u_2(dv_2). \tag{2.6}$$

A strategy $u_1^* \in A_1$ is said to be (α-discounted) optimal for player 1 if for $(x,i) \in \mathcal{R}^d \times S$

$$R[u_1^*, \tilde{u}_2](x,i) \geq \inf_{u_2 \in A_2}\ \sup_{u_1 \in A_1}\ R[u_1, u_2](x,i)\ \ := \bar{R}(x,i) \tag{2.7}$$

for any $\tilde{u}_2 \in A_2$. The function $\bar{R} : \mathcal{R}^d \times S \to \mathcal{R}$ is called upper value function of the game. Similarly, a strategy $u_2^* \in A_2$ is said to be optimal for player 2 if

$$R[\tilde{u}_1, u_2^*](x,i) \;\leq\; \sup_{u_2 \in A_2} \; \inf_{u_1 \in A_1} \; R[u_1, u_2](x,i) \;:=\; \underline{R}(x,i) \tag{2.8}$$

for any $\tilde{u}_1 \in A_1$ and $(x,i) \in \mathcal{R}^d \times S$. The function $\underline{R} : \mathcal{R}^d \times S \to \mathcal{R}$ is called the lower value function of the game. If $\bar{R} \equiv \underline{R}$, then the game is said to admit a value and the common function is denoted by R and is called the value function. Clearly the existence of a pair of optimal strategies for both players ensures that the value function exists but the converse need not hold. We will establish the existence of a value function and Markov opimal strategies for both players. Since the addition of a constant to the payoff function does not alter the optimal strategies, we may (and will) assume that $\bar{r} \geq 0$.

NonZero-Sum Game

For each $\ell = 1,2$, let $\bar{r}_\ell : \mathcal{R}^d \times S \times V_1 \times V_2 \to \mathcal{R}$ be the payoff function for player ℓ. We assume that

(**A3**) $\bar{r}_\ell$ is bounded, continuous and Lipschitz in its first argument uniformly with respect to the rest.

When the system is in state (x,i) and the action $v = (v_1, v_2) \in V_1 \times V_2$ are chosen by the players then player $\ell, \ell = 1,2$, receives payoff $\bar{r}_\ell(x,i,v_1,v_2)$. Let $r_\ell : \mathcal{R}^d \times S \times U_1 \times U_2$ be defined as follows: for $(x,i) \in \mathcal{R}^d \times S, (u_1,u_2) \in U_1 \times U_2$

$$r_\ell(x,i,u_1,u_2) = \int_{V_1} \int_{V_2} \bar{r}_\ell \; (x,i,v_1,v_2) u_1(dv_1) u_2(dv_2), \quad k = 1,2. \tag{2.9}$$

Each player wants to maximize his accumulated income. We again consider the α-discounted payoff on the infinite horizon. Let $(u_1,u_2) \in A_1 \times A_2$. Let $(X(\cdot), \theta(\cdot))$ be the solution of (2.1) corresponding to $(u_1(\cdot), u_2(\cdot))$. Then the (α-discounted) payoff to player ℓ for the initial condition (x,i) is defined as

$$R_\ell[u_1,u_2](x,i) = E_{x,i}^{u_1,u_2} \; [\int_0^\infty \; e^{-\alpha t} \; r_\ell(x(t), \theta(t), u_1(t), u_2(t)) dt]. \tag{2.10}$$

A pair of strategies $(u_1^*, u_2^*) \in A_1 \times A_2$ is said to be a (Nash) equilibrium if for $(x,i) \in \mathcal{R}^d \times S$

$$R_1[u_1^*, u_2^*](x,i) \geq R_1[u_1, u_2^*](x,i) \text{ for any } u_1 \in A_1,$$

and

$$R_2[u_1^*, u_2^*](x,i) \geq R_2[u_1^*, u_2](x,i) \text{ for any } u_2 \in A_2. \tag{2.11}$$

We will establish the existence of an equilibrium in Markov strategies. Note that the two-person nonzero-sum game can be extended to the N-person game. We are treating only the two-person case for notational simplicity.

We conclude this section by showing that both the players can confine their attention to only Markov strategies. To this end we introduce the concept of α-discounted occupation measures. Let $(u_1(\cdot), u_2(\cdot)) \in A_1 \times A_2$ and $(X(\cdot), \theta(\cdot))$ the corresponding process. The α-discounted occupation measure denoted by $\nu[u_1, u_2] \in \mathcal{P}(\mathcal{R}^d \times S \times V_1 \times V_2)$ is defined implicitly by

$$\sum_{j=1}^{N} \int_{\mathcal{R}^d \times V_1 \times V_2} f(x, i, v_1, v_2) \nu[u_1, u_2](dx, i, dv_1, dv_2)$$

$$= \alpha \, E_{x,i}^{u_1,u_2}\left[\int_0^\infty \int_{V_1} \int_{V_2} e^{-\alpha t} \, f(X(t), \theta(t), v_1, v_2) u_1(t)(dv_1) u_2(t)(dv_2) dt\right] \tag{2.12}$$

for $f \in C_b(\mathcal{R}^d \times S \times V_1 \times V_2)$. Indeed, $\nu[u_1, u_2]$ will depend on the initial condtion (x,i), but we have suppressed this dependence for notational convenience. In terms of $\nu[u_1, u_2]$, (2.5) becomes

$$R[u_1, u_2](x,i) = \alpha^{-1} \sum_{j=1}^{N} \int_{\mathcal{R}^d \times V_1 \times V_2} \bar{r}(y, j, v_1, v_2) \nu[u_1, u_2](dy, j, dv_1, dv_2) \tag{2.13}$$

and (2.10) becomes

$$R_\ell[u_1, u_2](x,i) = \alpha^{-1} \sum_{j=1}^{N} \int_{\mathcal{R}^d \times V_1 \times V_2} \bar{r}_\ell(y, j, v_1, v_2) \nu[u_1, u_2](dy, j, dv_1, dv_2). \tag{2.14}$$

Let

$$\nu[A_1, A_2] = \{\nu[u_1, u_2] \mid (u_1, u_2) \in A_1 \times A_2\}.$$

$\nu[M_1, A_2], \nu[A_1, M_2], \nu[M_1, M_2]$ etc. are defined analogously. We can closely mimic the proof of [8, Lemma 4.2], to obtain the following result.

Lemma 2.1. For any fixed initial condition

$$\nu[A_1, M_2] \;=\; \nu[M_1, M_2] \;=\; \nu[M_1, A_2].$$

It immediately follows from the above lemma that for the two person zero-sum game no player can improve his payoff by going beyond Markov strategies. For the nonzero-sum game there is no notion of value of the game. Different pairs of equilibrium strategies may yield varied payoffs to the players. However, due to the enormous complexity in implementing a non-Markov strategy both players look for equilibrium in Markov strategies only.

ZERO-SUM GAME

In this section we will establish Markov optimal strategies for both players. To this end we first study the corresponding Isaacs equation given by

$$\inf_{u_2 \in U_2} \sup_{u_1 \in U_1} [L^{u_1,u_2}\, \phi(x,i) + r(x,i,u_1,u_2)]$$

$$= \sup_{u_1 \in U_1} \inf_{u_2 \in U_2} [L^{u_1,u_2}\, \phi(x,i) + r(x,i,u_1,u_2)] \;=\; \alpha\phi(x,i). \qquad (3.1)$$

Note that (3.1) is a quasilinear system of uniformly elliptic equations with weak coupling in the sense that the coupling occurs only in the zeroth order term.

Theorem 3.1. Under (A1), (A2) the equation (3.1) has a unique solution in $C^2(\mathcal{R}^d \times S) \cap C_b(\mathcal{R}^d \times S)$.

Proof. Consider first the system of equations

$$\inf_{u_2 \in U_2} \sup_{u_1 \in U_2} [L^{u_1,u_2}\, \phi(x,i) + r(x,i,u_1,u_2)] \;=\; \alpha\phi(x,i). \qquad (3.2)$$

For $R > 0$, let $B_R = \{x \in \mathcal{R}^d \;:\; |x| < R\}$. Consider the following Dirichlet problem on $B_R \times S$

$$\inf_{u_2 \in U_1} \sup_{u_1 \in U_2} [L^{u_1,u_2}\phi(x,i) + r(x,i,u_1,u_2)] = \alpha\phi(x,i) \;\text{ in } B_R \times S \qquad (3.3a)$$

$$\phi(x,i)\;|_{\partial B_R \times S} \;=\; 0. \qquad (3.3b)$$

Under (A1), (A2) the existence of a unique solution $\phi_R(x,i)$ of (3.3) in $W^{2,p}(B_R \times S), 2 \le p < \infty$, is guaranteed by using the arguments in [11, Thm. 5.1, p. 422]. Thus to each $R > 0$ there corresponds a solution ϕ_R of (3.3) belonging to $W^{2,p}(B_R \times S), 2 \le p < \infty$. By the Sobolev imbedding theorem $\phi_R(x,i) \in C^{1,\gamma}(B_R \times S)$, for $0 < \gamma < 1, \gamma$ arbitrarily close to 1, and hence by our assumption on $\bar{b}, \lambda, \bar{r}$ $((A1),(A2))$, it is easy to see that

$$\alpha\phi(x,i) - \inf_{u_2 \in U_2} \sup_{u_1 \in U_1} [\sum_{j=1}^{d} b_j(x,i,u_1,u_2) \frac{\partial \phi_R(x,i)}{\partial x_j}$$

$$+ \sum_{j=1}^{N} \lambda_{ij}(x)\phi_R(x,j) + r(x,i,u_1,u_2)]$$

is in $C^{0,\gamma}$. By elliptic regularity [9, p. 287] applied to (3.3a) it follows that $\phi_R(x,i) \in C^{2,r}(B_R \times S)$. Standard arguments involving Ito's formula yield $\phi_R(x,i) =$

$$\inf_{u_2 \in M_2} \sup_{u_1 \in M_1} E_{x,i}^{u_1,u_2} [\int_0^{\tau_R} e^{-\alpha t} r(X(t),\theta(t),u_1(X(t),\theta(t)),u_2(X(t),\theta(t))dt], \tag{3.4}$$

where τ_R is the hitting time of ∂B_R of the process $X(\cdot)$. Since $\bar{r} \ge 0, \phi_R(x,i) \le \bar{R}(x,i)$ (the upper value of the game). Clearly $\phi_R(x,i)$ is nondecreasing in R. Let $R' > R$. Then by the interior estimates [11, pp. 398-402] $\{\phi_{R'}\}_{R'>R}$ is bounded in $B_R \times S$ uniformly in R' and $\{\nabla\phi_{R'}\}_{R'>R}$ is bounded in $W^{1,2}(B_R \times S)$ uniformly in R'. By the Sobolev imbedding theorem $W^{1,2}(B_R \times S) \hookrightarrow L^{2+\epsilon}(B_R \times S)$ for some $\epsilon > 0$. Then by suitably modifying the arguments in (4.10) in [11, p. 400], we obtain

$$|| \phi_{R'} || W^{2,2+\epsilon}_{(B_R \times S)} \quad \le \quad k_R,$$

where k_R is a constant independent of R'. (The modification is needed because of the factor $\epsilon > 0$, but it is routine). Repeating the above procedure over and over again we conclude that $\{\phi_{R'}\}_{R'>R}$ is uniformly bounded in $W^{2,p}(B_R \times S)$ for $2 \le p < \infty$. Since $W^{2,p}(B_R \times S) \hookrightarrow W^{1,p}(B_R \times S)$ and the injection is compact, it follows that $\{\phi_{R'}\}_{R'>R}$ converges strongly in $W^{1,p}(B_R \times S)$. Thus given any sequence $\{R_n\}, R_n \to \infty$ as $n \to \infty$, and for any fixed integer $N \ge 2$, we can choose a subsequence $\{R_{n_i}\}$ such that

$\{\phi_{R_{n_i}}\}$ converges strongly in $W^{1,p}(B_{N-1} \times S)$. Using a suitable diagonalization, we may assume that $\{\phi_{R_{n_i}}\}$ converges strongly in $W^{1,p}(B_{N-1} \times S)$ for each $N \geq 2$. Let ψ be a limit point of $\{\phi_{R_{n_i}}\}$. It can be shown as in [3, p. 148] (see also [11, p. 420]) that

$$\inf_{u_2 \in U_2} \sup_{u_1 \in U_1} [\sum_{j=1}^{d} b_j(x,k,u_1,u_2) \frac{\partial \phi_{R_{n_i}}(x,k)}{\partial x_j} + \sum_{j=1}^{N} \lambda_{kj} \phi_{R_{n_i}}(x,k)$$

$$+ r(x,k,u_1,u_2)]$$

$$\stackrel{n_i \to \infty}{\to} \inf_{u_2 \in U_2} \sup_{u_1 \in U_1} [\sum_{j=1}^{d} b_j(x,k,u_1,u_2) \frac{\partial \psi(x,k)}{\partial x_j} + \sum_{j=1}^{N} \lambda_{kj}(x)$$

$$+ r(x,k,u_1,u_2)]$$

strongly in $L^p(B_{N-1} \times S)$. Therefore $\psi \in W^{1,p}_{loc}(\mathcal{R}^d \times S)$ and it satisfies (3.2) in $\mathcal{D}'(\mathcal{R}^d \times S)$, $i.e.$ in the sense of distributions. By elliptic regularity $\psi \in W^{2,p}_{loc}(\mathcal{R}^d \times S)$. Then by the Sobolev imbedding theorem and elliptic regularity $\psi \in C^{2,\gamma}(\mathcal{R}^d \times S), 0 < \gamma < 1, \gamma$ arbitrarily close to 1. Using (A1), (A2), and Fan's minimax theorem [6] it follows that ψ satisfies (3.1). Clearly $\psi \in C_b(\mathcal{R}^d \times S)$. Let $v_1^*(\cdot) \in M_1$ and $v_2^*(\cdot) \in M_2$ be the outer maximizing and outer minimizing selectors respectively in (3.1). The existence of such selectors is guaranteed by a standard measurable selection theorem [2]. Then routine arguments involving Ito's formula yield

$$\psi(x,i) = E^{v_1^*,v_2^*}_{x,i}[\int_0^\infty e^{-\alpha t} r(X(t),\theta(t),v_1^*(X(t)\theta(t)),v_2^*(X(t),\theta(t)))dt]$$

$$= \inf_{v_2 \in M_1} \sup_{v_1 \in M_1} E^{v_1,v_2}_{x,i}[\int_0^\infty e^{-\alpha t} r(X(t),\theta(t),v_1(X(t)\theta(t)),v_2(X(t),\theta(t)))dt]$$

$$= \sup_{v_1 \in M_2} \inf_{v_2 \in M_2} E^{v_1,v_2}_{x,i}[\int_0^\infty e^{-\alpha t} r(X(t),\theta(t),v_1(X(t)\theta(t)),v_2(X(t),\theta(t)))dt].$$

Thus $\psi(x,i) = \underline{V}(x,i) = \bar{V}(x,i) = V(x,i)$, the value of the game. To prove uniqueness let $V'(x,i)$ be another solution of (3.1) in $C^2(\mathcal{R}^d \times S) \cap C_b(\mathcal{R}^d \times S)$. Let k be a common bound on $\mid V(\cdot,\cdot) \mid$ and $\mid V'(\cdot,\cdot) \mid$. Then it can be shown that (see [5, pp. 69-70])

$$\mid V(x,i) - V'(x,i) \mid \leq 2e^{-\alpha t}k.$$

Letting $t \to \infty$, we have $V \equiv V'$. ∎

Theorem 3.2. Assume the conditions (A1), (A2). Let $v_1^* \in M_1$ be such that

$$\inf_{v_2 \in U_2} [\sum_{j=1}^{d} b_j(x,i,v_1^*(x,i),v_2) \frac{\partial V(x,i)}{\partial x_j} + \sum_{j=1}^{N} \lambda_{ij}(x) V(x,j) + r(x,i,v_1^*(x,i),v_2)]$$

$$= \sup_{v_1 \in U_1} \inf_{v_2 \in U_2} [\sum_{j=1}^{d} b_j(x,i,v_1,v_2) \frac{\partial V(x,i)}{\partial x_j} + \sum_{j=1}^{N} \lambda_{ij}(x) V(x,j) + r(x,i,v_1,v_2)] \tag{3.5}$$

for each i and $a.e.x$. Then v_1^* is optimal for player 1.

Similarly, let $v_2^* \in M_2$ be such that

$$\sup_{v_1 \in U_1} [\sum_{j=1}^{d} b_j(x,i,v_1,v_2^*(x,i)) \frac{\partial V(x,i)}{\partial x_j} + \sum_{j=1}^{N} \lambda_{ij}(x) V(x,i) + r(x,i,v_1,v_2^*(x,i))]$$

$$= \inf_{v_2 \in U_2} \sup_{v_1 \in U_1} [\sum_{j=1}^{d} b_j(x,i,v_1,v_2) \frac{\partial V(x,i)}{\partial x_j} + \sum_{j=1}^{N} \lambda_{ij}(x) V(x,j) + r(x,i,v_1,v_2)] \tag{3.6}$$

for each i and $a.e.x$. Then v_2^* is optimal for player 2.

Proof. We prove this claim for the first player. The corresponding claim for the second player follows similarly. Let $v_1^* \in M_1$ satisfy (3.5). The existence of such v_1^* follows from a standard measurable selection theorem [2, Lemma 1]. Pick any $v_2 \in M_2$. Let $(X(\cdot), \theta(\cdot))$ be the process governed by (v_1^*, v_2) with $X(0) = x, \theta(0) = i$. Then using the same arguments as in the proof of the previous theorem, we can show that

$$V(x,i) \le R[v_1^*, v_2](x,i).$$

Hence v_1^* is optimal for player 1. ∎

We now consider a special case where in each discrete state $i \in S$, one player controls the game exclusively. In other words, we assume the following

(**A4**) Let $S_1 = \{i_1, \cdots, i_m\} \subset S, S_2 = \{j_1, \cdots, j_n\} \subset S$ be such that $S_1 \cap S_2 = \phi$ and $S_1 \cup S_2 = S$. Further assume that

$$\bar{b}(x,i,v_1,v_2) = \bar{b}_1(x,i,v_1)$$

$$\bar{r}(x,i,v_1,v_2) = \bar{r}_1(x,i,v_2)$$

for $i \in S_1$,

$$\bar{b}_1 : \mathcal{R}^d \times S_1 \times V_1 \to \mathcal{R}$$

$$\bar{r}_1 : \mathcal{R}^d \times S_1 \times V_2 \to \mathcal{R}$$

satisfying the same conditions as $\bar{b}$ and $\bar{r}_1$. Also for any $i \in S_2$

$$\bar{b}(x,i,v_1,v_2) = \bar{b}_2(x,i,v_2)$$

$$\bar{r}(x,i,v_1,v_2) = \bar{r}_2(x,i,v_2)$$

satisfying the same conditions as $\bar{b}$ and $\bar{r}$ (*cf*, (A1), (A2)).

Theorem 3.3. Under (A1), (A4), each player has Markov optimal pure strategies.

Proof. Let $i_m \in S_1$. Then under (A4) we have

$$V(x,i_m) = \sup_{u_1 \in U_1} [\sum_{k=1}^d \int_{V_1} \bar{b}_{1k}(x,i_m,v_1)u_1(dv_1) \frac{\partial V(x,i_m)}{\partial x_k}$$
$$+ \sum_{j=1}^N \lambda_{i_m}j(x)V(x,j) + \int_{V_1} \bar{r}(x,i_m,v_1)u_1(dv_1)]. \tag{3.7}$$

For each (x,i_m) the supremum in the above will be attained at an extreme point of $U_1 = \mathcal{P}(V_1)$. Thus to each $(x,i_m) \in \mathcal{R}^d \times S_1$ there exists a $v_1^*(x,i_m) \in V_1$ such that the supremum in (3.7) is obtained at $\delta_{v_1^*(x,i_m)}$. The map $(x,i_m) \to v_1^*(x,i_m)$ may be assumed to be measurable in view of the measurable selection theorem in [2]. Thus the strategy $v_1^* \in D_1$ is optimal for player 1. The claim for the second players follows similarly. ■

<u>NONZERO-SUM GAME</u>

We make the following assumption

(**A5**) $\bar{b}$ and $\bar{r}_k, k = 1,2$, are of the form

$$\bar{b}(x,i,v_1,v_2) = \bar{b}_1(x,i,v_1) + \bar{b}_2(x,i,v_2)$$

$$\bar{r}_k(x,i,v_1,v_2) = \bar{r}_{1k}(x,i,v_1) + \bar{r}_{2k}(x,i,v_2),\ k = 1,2,$$

where $\bar{b}_\ell : \mathcal{R}^d \times S \times V_\ell \to \mathcal{R}^d$, $\bar{r}_{\ell k} : \mathcal{R}^d \times S \times V_\ell \to \mathcal{R}$, satisfy the same assumptions as $\bar{b}, \bar{r}$.

Let $(v_1,v_2) \in M_1 \times M_2$. By Lemma 2.1, for any $(x,i) \in \mathcal{R}^d \times S$

$$\sup_{u_1 \in A_1} R_1[u_1, v_2](x,i) = \sup_{u_1 \in M_1} R_1[u_1, v_2](x,i)$$

$$\sup_{u_2 \in A_2} R_2[v_1, u_2](x,i) = \sup_{u_2 \in M_2} R_2[v_1, u_2](x,i).$$

In view of the results in [8] the above suprema on the right hand side can be replaced by maxima. Thus, there exist $v_1^* \in M_1, v_2^* \in M_2$ such that

$$\sup_{u_1 \in A_1} R_1[u_1, v_2](x,i) = \max_{u_1 \in A_1} R_2[u_1, v_2](x,i)$$

$$= R_1[v_1^*, v_2](x,i) := \tilde{R}_1[v_2](x,i) \tag{4.1}$$

$$\sup_{u_2 \in A_2} R_2[v_1, u_2](x,i) = \max_{u_2 \in A_2} R_2[v_1, u_2](x,i)$$

$$= R_2[v_1, v_2^*](x,i) := \tilde{R}_2[v_1](x,i). \tag{4.2}$$

Indeed, v_1^* will depend on v_2 and v_2^* will depend on v_1. v_1^* (resp. v_2^*) is called the optimal response of player 1 (resp. player 2) given player 2 (resp. player 1) is employing v_2 (resp. v_1). From [8] the following result follows.

Lemma 4.1. Fix $(v_1, v_2) \in M_1 \times M_2$. Then $\tilde{R}_1[v_2]$ is the unique solution in $W^{2,p}_{loc}(\mathcal{R}^d \times S) \cap C_b(\mathcal{R}^d \times S), 2 \le p < \infty$, of

$$\sup_{u_1 \in U_1} [L^{u_1, v_2}\phi(x,i) + r(x,i,u_1,v_2(x,i)] = \alpha\phi(x,i) \tag{4.3}$$

in $\mathcal{R}^d \times S$. A strategy $v_1^* \in M_1$ is an optimal response for player 1, given player 2 is employing v_2, if and only if

$$[\sum_{j=1}^{d} b_j(x,i,v_1^*(x,i),v_2(x,i)) \frac{\partial \tilde{R}_1[v_2](x,i)}{\partial x_j} + \sum_{j=1}^{N} \lambda_{ij}(x)\tilde{R}_1[v_2](x,j)$$

$$+ r_1(x, v_1^*(x,i), v_2(x,i))]$$

$$= \sup_{v_1 \in U_1} [\sum_{j=1}^{d} b_j(x,i,v_1,v_2(x,i)) \frac{\partial \tilde{R}_1[v_2](x,i)}{\partial x_j} + \sum_{j=1}^{N} \lambda_{ij}(x)\tilde{R}_1[v_2](x,j)$$

$$+r_1(x, v_1, v_2(x,i)]. \tag{4.4}$$

Similarly, $\tilde{R}_2[v_1]$ is the unique solution in $W^{2,p}_{loc}(\mathcal{R}^d \times S) \cap C_b(\mathcal{R}^d \times S), 2 \leq p < \infty$, of

$$\sup_{u_2 \in U_2} [L^{v_1,u_2}\psi(x,i) \ + \ r_2(x,i,v_1(x,i),u_2)] = \alpha\psi(x,i) \tag{4.5}$$

in $\mathcal{R}^d \times S$. A strategy $v_2^* \in M_2$ is an optimal response for player 2, given player 1 is employing v_1, if and only if

$$[\sum_{j=1}^{d} b_j(x,i,v_1(x,i),v_2^*(x,i)) \ \frac{\partial \tilde{R}_2[v_1](x,i)}{\partial x_j} \ + \ \sum_{j=1}^{N} \lambda_{ij}(x)\tilde{R}_2[v_1](x,j)$$

$$+r_2(x,i,v_1(x,i),v_2^*(x,i)]$$

$$= \sup_{v_2 \in U_2} [\sum_{j=1}^{d} \ b(x,i,v_1(x,i),v_2) \ \frac{\partial \tilde{R}_2[v_1](x,i)}{\partial x_j} \ + \ \sum_{j=1}^{N} \lambda(x)\tilde{R}_2[v_1](x,j)$$

$$+r_2(x,i,v_1(x,i),v_2)]. \tag{4.6}$$

Theorem 4.1. Under (A1), (A3), (A5) there exists an equilibrium $(v_1^*, v_2^*) \in M_1 \times M_2$.

Proof. Let M_1 and M_2 be endowed with the metric topology described in [8]. (See Lemma 3.2 in [8] for the convergence criterion describing the topology of M_1, M_2). Then M_1 and M_2 are compact, metric spaces. Let $M_1 \times M_2$ be endowed with the product topology. Let $(v_1, v_2) \in M_1 \times M_2$. Let $(\bar{v}_1, \bar{v}_2) \in U_1 \times U_2$. Set

$F_1(x,i,\bar{v}_1,v_2(x,i)) = \sum_{j=1}^{d} \ b_j(x,i,\bar{v}_1,v_2(x,i)) \ \frac{\partial \tilde{R}_1[v_2](x,i)}{\partial x_j}$

$$+ \ \sum_{j=1}^{N} \lambda_{ij}(x) \ \frac{\tilde{R}_1[v_2](x,i)}{\partial x_j} \ + \ r_1(x,i,\bar{v}_1,v_2(x,i))$$

$F_2(x,i,v_1(x,i),\bar{v}_2) = \sum_{j=1}^{d} \ b_j(x,i,\bar{v}_1(x,i),\bar{v}_2) \ \frac{\partial \tilde{R}_2[v_1](x,i)}{\partial x_j}$

$$+ \sum_{j=1}^{N} \lambda_{ij}(x)\tilde{R}_2[v_1](x,i) + r_2(x,i,v_1(x,i),\bar{v}_2).$$

Let

$$G_1[v_2] = \{v_1^* \in M_1 \mid F_1(x,i,v_1^*(x,i),v_2(x,i)) = \sup_{\bar{v}_1 \in U_1} F_1(x,i,\bar{v}_1,v_2(x,i))$$

$$\text{a.e.x, for each i } \}$$

$$G_2[v_1] = \{v_2^* \mid F_2(x,i,v_1(x,i),v_2^*(x,i)) = \sup_{\bar{v}_2 \in U_2} F_2(x,i,v_1(x,i),\bar{v}_2) \ a.e.x,$$

$$\text{for each } i\}.$$

Then $G_1[v_2]$ and $G_2[v_1]$ are nonempty, convex, compact subsets of M_1 and M_2, respectively. Let $G[v_1,v_2] = G_1[v_2] \times G_2[v_1]$. Then $G[v_1,v_2]$ is a noempty, convex and compact subset of $M_1 \times M_2$. Thus $(v_1,v_2) \to G[v_1,v_2]$ defines a point-to-set map from $M_1 \times M_2$ to $2^{M_1 \times M_2}$. Mimicking the arguments in [5, Thm. 5.1] this map is seen to be upper semicontinuous under the assumption (A1), (A2) and (A5). By Fan's fixed point theorem [7], there exists $(v_1^*, v_2^*) \in M_1 \times M_2$ such that $(v_1^*, v_2^*) \in G_2[v_1^*, v_2^*]$. The pair (v_1^*, v_2^*) is clearly an equilibrium. ■

Acknowledgements. We wish to thank Professor Adam Shwarz for pointing out the potential applicability of Theorem 3.3 to token ring networks.

REFERENCES

[1] Basar, T. and Haurie, A., Feedback equilibria in differential games with structural and modal uncertainties, in Advances in Large Scale Systems, vol. 1, Ed., J. B. Cruz, Jr, 1984, 163-301.

[2] Beneš, V. E., Existence of optimal strategies based on a specified information for a class of stochastic decision problems, SIAM J. Control 8 (1970), 179-188.

[3] Bensoussan, A., Stochastic Control by Functional Analysis Methods, North-Holland, Amsterdam, 1982.

[4] Bertsekas, D. and Gallager, R., Data Networks, Prentice-Hall, New Jersey, 1987.

[5] Borkar, V. S., Optimal Control of Diffusion Processes, Pitman Res. Notes Math. Ser., No. 203, Longman, Harlow, UK, 1989.

[**6**] Borkar, V. S. and Ghosh, M. K., Stochastic differential games: Occupation measure based approach, J. Optim. Theory Appl. 73 (1992), 359-385.

[**7**] Fan, K., Fixed point and minimax theorems in locally convex topological linear spaces, Proc. Natl. Acad. Sci., USA, 38 (1952), 121-126.

[**8**] Ghosh, M. K., Araposthatis, A. and Marcus, S. I., Optimal control of switching diffusions with application to flexible manufacturing systems, SIAM J. Control Optim. 31 (1993), 1183-1204.

[**9**] Grisvard, P., Elliptic Problems in Nonsmooth Domains, Pitman, Boston, 1965.

[**10**] Haurie, A., Piecewise deterministic differential games, Preprint.

[**11**] Ladyzhenskaya, O. A. and Uralćeva, N. N., Linear and Quasilinear Elliptic Equations, Academic Press, New York, 1968.

[**12**] Mariton, M., Jump Linear Systems in Automatic Control, Marcel Dekker, New York, 1990.