

On the Saddle-Point Solution of a Class of Stochastic Differential Games¹

T. BAŞAR²

Communicated by Y. C. Ho

Abstract. This paper deals with the saddle-point solution of a class of stochastic differential games described by linear state dynamics and quadratic objective functionals. The information structure of the problem is such that both players have access to a common noisy linear measurement of the state and they are permitted to utilize only this information in constructing their controls. The saddle-point solution of such differential game problems has been discussed earlier in Ref. 1, but the conclusions arrived there are incorrect, as is explicitly shown in this paper. We extensively discuss the role of information structure on the saddle-point solution of such stochastic games (specifically within the context of an illustrative discrete-time example) and then obtain the saddle-point solution of the problem originally formulated by employing an indirect approach.

Key Words. Stochastic differential games, saddle-point solutions, information structures, conjugate-point conditions, linear-quadratic games.

1. Introduction

Consider the zero-sum, two-person, stochastic differential game in which the state dynamics are described by a linear stochastic differential equation, the objective functional is quadratic, and the players have access to the same linear noisy measurements of the state. Further, assume that the noise processes in the state and measurement equations are appropriate Wiener processes, and the permissible controls of the players are chosen as causal functions of the measurement process. This stochastic differential

¹This work was done while the author was on sabbatical leave at Twente University of Technology, Department of Applied Mathematics, Enschede, Holland, from Applied Mathematics Division, Marmara Scientific and Industrial Research Institute, Gebze, Kocaeli, Turkey.

²Associate Professor of Electrical Engineering and Research and Associate Professor at Coordinated Science Laboratory, University of Illinois, Urbana, Illinois.

game has been considered earlier in Ref. 1, and it has been “proven” that the saddle-point solution satisfies the separation principle and requires the same existence conditions as the deterministic version of the problem (obtained by replacing the noisy measurement with a perfect state measurement). In this paper, we show that this conclusion is false, and that derivations of the saddle-point solution and the existence conditions require a much more involved analysis than the one given in Ref. 1.

In the next section, we provide a precise mathematical formulation of the problem and discuss inherent limitations of different direct methods that could be used in obtaining its saddle-point solution. In Section 3, we include a solution of the deterministic version of the problem, and elucidate the nature and the sources of the error committed in Ref. 1 with regard to the solution of the stochastic game. In order to have a better understanding of the role of information structures in zero-sum stochastic differential games, and of the pitfalls in the derivation of the saddle-point solution, we present in Section 4 a scalar, two-stage, discrete-time version of the original problem. In particular, we show that, while the deterministic version of this scalar dynamic game admits a pure-strategy saddle-point solution, its stochastic version admits a saddle point in only mixed strategies.

The main results of the paper are given in Section 5, and are embedded in Theorem 5.1 and Corollary 5.1. Specifically, we obtain closed-form expressions for the saddle-point strategies of the players in the linear-quadratic stochastic differential game, and prove that the required conditions of existence do not coincide with, and are in fact more stringent than, the existence conditions of the deterministic version of the problem. The paper ends with a conclusions section where we discuss possible extensions to similarly structured nonzero-sum differential games.

2. Problem Statement

The linear-quadratic, zero-sum, stochastic differential game under consideration can be defined in precise mathematical terms as follows: The evolution of the state (x) of the game is described by the linear Ito stochastic differential equation

$$\begin{aligned} dx_t &= A(t)x_t dt + B_p(t)u_t dt + B_e(t)v_t dt + F(t) dw_t^1, \\ t_0 \leq t \leq t_f, \quad x_{t_0} &= x_0, \end{aligned} \tag{1}$$

and the common observation y of the players is described by

$$dy_t = C(t)x_t dt + G(t) dw_t^2, \tag{2}$$

where

$$\begin{aligned} \dim(x) = n, & \quad \dim(u) = r_p, & \quad \dim(v) = r_e, \\ \dim(y) = m, & \quad \dim(w^1) = n, & \quad \dim(w^2) = m, \end{aligned}$$

and the matrices A, B_p, B_e, C, F, G have appropriate dimensions and are continuous on $[t_0, t_f]$. Furthermore,

$$FF' > 0, \quad GG' > 0,$$

on the interval $[t_0, t_f]$. x_0 is the value of the initial state which is assumed to be Gaussian distributed with mean zero and covariance P_0 . The functions u_t and $v_t, t \geq t_0$, represent the controls of the pursuer and the evader, respectively, and are vector stochastic processes with continuous sample paths, as to be further explained in the sequel. $x_t, t \geq t_0$, is an n -dimensional vector stochastic process with continuous sample paths; and $w_t^1, t \geq t_0$, and $w_t^2, t \geq t_0$, are statistically independent standard Wiener processes.

To delineate the information structure of this game problem, we let

$$C_m = C_m[t_0, t_f]$$

denote the space of continuous functions on $[t_0, t_f]$, with values in R^m . We further let Y_t be the sigma-field in C_m generated by the cylinder sets $\{y \in C_m, y_s \in B\}$, where B is a Borel set in R^m and $t_0 \leq s \leq t$. Then, the information gained by each player during the course of the game is completely determined by the information field $Y_t, t \geq t_0$. A permissible strategy for the pursuer will be a mapping $\gamma_p(\cdot, \cdot)$ of $[t_0, t_f] \times C_m$ into R^p with the following properties:

- (i) $\gamma_p(t, \eta)$ is continuous in t for each $\eta \in C_m$;
- (ii) $\gamma_p(t, \eta)$ is uniformly Lipschitz in η ; i.e.,

$$|\gamma_p(t, \eta) - \gamma_p(t, \xi)| \leq k \|\eta - \xi\|, \quad t \in [t_0, t_f], \quad \eta, \xi \in C_m,$$

where $\|\cdot\|$ is the standard sup norm in C_m ;

- (iii) $u_t = \gamma_p(t, \eta)$ is adapted to the information field Y_t .

We denote the class of strategies described above for the pursuer by Γ_p . Analogously, a permissible strategy $\gamma_e(\cdot, \cdot)$ for the evader will be defined as a mapping of $[t_0, t_f] \times C_m$ into R^r with properties (i) and (ii) above, and with the restriction that $v_t = \gamma_e(t, \eta)$ be adapted to the information field Y_t . Let us denote the collection of all such strategies for the evader by Γ_e . It is known that, corresponding to any pair of strategies $\{\gamma_p \in \Gamma_p, \gamma_e \in \Gamma_e\}$, the stochastic differential equation (1) admits a unique solution that is a sample-path-continuous second-order process (Refs. 2, 3). Consequently, the observation process $y_t, t \geq t_0$ will also have continuous sample paths. For any pair

of strategies $\{\gamma_p \in \Gamma_p, \gamma_e \in \Gamma_e\}$, the expected objective function for this zero-sum game is defined by

$$J(\gamma_p, \gamma_e) = E\left\{ x'_{t_f} Q_f x_{t_f} + \int_{t_0}^{t_f} (x'_t Q(t) x_t + \lambda u'_t u_t - v'_t v_t) dt \right\}, \tag{3}$$

where

$$u_t = \gamma_p(t, y_0^t), \quad v_t = \gamma_e(t, y_0^t),$$

with y_0^t denoting the stochastic process $y_s, s \geq t_0$, restricted to the time interval $[t_0, t]$, $Q_f \geq 0, Q(t) \geq 0$, for all $t \in [t_0, t_f]$, and λ is a positive scalar. It should be noted that (3) defines the most general quadratic cost function with nonnegative weights on the components of the state and positive weights on the components of the controls of the players. Actually, λ can also be absorbed in the state equation, without any loss of generality, but we prefer to keep it here as a free parameter which we will have occasion to utilize later.

The pursuer attempts to minimize J , while the evader seeks to maximize the same quantity. Then, a pair of strategies $\{\gamma_p^* \in \Gamma_p, \gamma_e^* \in \Gamma_e\}$ will be in *saddle-point equilibrium* if the pair of inequalities

$$J(\gamma_p^*, \gamma_e) \leq J(\gamma_p^*, \gamma_e^*) \leq J(\gamma_p, \gamma_e^*) \tag{4}$$

are satisfied for all $\gamma_p \in \Gamma_p, \gamma_e \in \Gamma_e$.

A direct approach to obtain the solution of this stochastic zero-sum differential game would be first to determine the security strategies of the players and then to investigate conditions under which they constitute a saddle-point equilibrium pair. Recall that a security strategy for a player (say, the pursuer) is defined as a $\gamma_p^s \in \Gamma_p$ that satisfies

$$\sup_{\gamma_e \in \Gamma_e} J(\gamma_p^s, \gamma_e) = \inf_{\gamma_p \in \Gamma_p} \sup_{\gamma_e \in \Gamma_e} J(\gamma_p, \gamma_e). \tag{5}$$

But any attempt to obtain the solution of this equality readily fails, since the optimization problem

$$\sup_{\gamma_e \in \Gamma_e} J(\gamma_p, \gamma_e)$$

structurally depends on the choice of γ_p , and thus one has to solve a different problem for each permissible choice of $\gamma_p \in \Gamma_p$. One way out of this difficulty would be to restrict γ_p to a linear class, say

$$\gamma_p(t, y_0^t) = \int_{t_0}^t N(t, \tau) dy_{\tau}, \tag{6}$$

where $N(\cdot, \cdot)$ is a matrix function of appropriate dimensions (as has been done in Ref. 4 for the case when the players have different measurements), but this approach leads to a set of coupled infinite-dimensional equations which are quite intractable as far as the explicit derivations of the saddle-point solution and the saddle-point conditions are concerned. One can even assume a finite-dimensional representation for the integral (6) and then verify that the saddle-point strategy of the pursuer indeed lies within that subset of Γ_p (as has been done in Ref. 5 for a similar problem with a somewhat different information structure), but this approach also leads to some unwieldy expressions and conditions which are not easy to manipulate. In this paper, we present an indirect way of obtaining the saddle-point solution of this problem, by relating it to a representation of the saddle-point solution of another zero-sum differential game, which is easily computable.

3. Saddle-Point Solution under Deterministic Information Pattern

Before obtaining the solution of the stochastic game problem as formulated above, let us recall the saddle-point solution of its deterministic version as far as the information structure is concerned. That is, the players now have access to the current value of the state at each instant of time and they recall past state information throughout the duration of the game. But the additive Wiener process term in the state equation is still retained in order to ensure unicity of equilibria (see Ref. 6, Section 4, for specific reasons). Then, the saddle-point solution is unique and is given by (Ref. 6)

$$\gamma_p^*(t, x_0') = -(1/\lambda)B_p'(t)S(t)x_0, \tag{7-1}$$

$$\gamma_e^*(t, x_0') = B_e'(t)S(t)x_0, \tag{7-2}$$

where $S(\cdot)$ satisfies the Riccati equation

$$\dot{S} + SA + A'S - S(B_p B_p' / \lambda - B_e B_e')S + Q = 0, \tag{8-1}$$

$$S(t_f) = Q_f. \tag{8-2}$$

The condition of existence of the saddle-point solution is precisely the existence of a symmetric matrix solution to (8), which is well known to admit conjugate points if the time interval is taken sufficiently large. However, it is also known that, for a fixed time interval $[t_0, t_f]$, and for fixed game parameters (other than λ), it is possible to find a $\lambda_0 > 0$ so that the solution to (8), and hence the saddle-point solution to the game of this section, exist for all $0 < \lambda < \lambda_0$. In particular, a λ_0 that does the job independently of the length of the time interval is that value of λ for which

$$B_p B_p' / \lambda - B_e B_e' > \epsilon I,$$

where ϵ is a sufficiently small positive parameter. However, this is only a sufficient condition, and it is always possible to find a larger λ_0 that depends on the length of the time interval.

One question that readily comes into mind now is whether one can obtain the solution of the stochastic game problem of Section 2 by making use of the solution of this deterministic problem, by paralleling an analysis that leads to the separation result of stochastic optimum control. An argument in favor of this dictum would be: "Since both players have access to the same noisy observation which is linear in the state, there is no reason why the separation theorem should not be valid for the stochastic differential game of Section 2 also." This argument was actually the motivation behind derivation of the saddle-point solution given in Ref. 1 (p. 187, Theorem 7.2) for precisely the same problem formulated in Section 2. This asserted saddle-point solution follows readily from (7) to be

$$\gamma_p^0(t, y'_0) = -(1/\lambda)B'_p(t)S(t)\hat{x}_t, \tag{9-1}$$

$$\gamma_e^0(t, y'_0) = B'_e(t)S(t)\hat{x}_t, \tag{9-2}$$

where S again satisfies (8) and where $\hat{x}_t, t \geq t_0$, is defined by

$$d\hat{x}_t = A(t)\hat{x}_t dt + B_p(t)u_t dt + B_e(t)v_t dt + K(t)[dy_t - C(t)\hat{x}_t dt], \tag{10-1}$$

$$t \geq t_0, \quad \hat{x}_{t_0} = 0, \tag{10-2}$$

and

$$K(t) = P(t)C'(t)[G(t)G'(t)]^{-1}, \tag{11-1}$$

$$\dot{P} - AP - PA' + PC'[GG']^{-1}CP - FF' = 0, \tag{11-2}$$

$$P(t_0) = P_0. \tag{11-3}$$

It is claimed in Ref. 1 that (9) indeed provides the unique solution to the stochastic differential game problem under consideration, and that it requires the same existence condition as that of the deterministic version treated above, both requiring nonexistence of a conjugate point for the Riccati equation (8) in the interval $[t_0, t_f]$. Even though the latter statement is true within a more general framework, the former statement is incorrect, since the solution (9) does not lie in the product strategy space $\Gamma_p \times \Gamma_e$. This is mainly due to the fact that \hat{x}_t given by (10) is defined as the conditional mean of x_t , given not only the past measurements $\{y_s, s \leq t\}$, but also the past control values $\{u_s, s \leq t\}$ and $\{v_s, s \leq t\}$. Hence, the proposed strategies (9-1) and (9-2) constitute a saddle-point pair, within a much larger class of strategies than $\Gamma_p \times \Gamma_e$, and moreover, within that larger class, they are not unique. This nonuniqueness is mainly due to the fact that, if the information set of a player includes both the observation and the value of control of the

other player, then a possibility of adopting different representations of the same control value emerges, which gives rise to a multiple of different saddle-point strategy pairs, each one requiring a different existence condition. Perhaps, the best illustration of this phenomenon can be provided within the context of a scalar zero-sum difference game that involves two stages, which is done in the next section. In Section 5, we turn back to the game problem of Section 2 and discuss its unique solution and a necessary existence condition, in the light of the discussion given above and the illustrative example solved in Section 4.

4. Illustrative Two-Stage Dynamic Zero-Sum Game

As an illustration of the pitfalls in the derivation of the saddle-point solutions of stochastic dynamic games, let us consider the zero-sum, two-stage, stochastic game whose state dynamics are described by the difference equation

$$x_2 = x_1 - u + w_1^1, \tag{12-1}$$

$$x_1 = x_0 + 2v + w_0^1, \tag{12-2}$$

and whose expected cost function is

$$J = E\{(x_2)^2 + \lambda u^2 - v^2\}, \tag{13}$$

where u is the scalar control of the pursuer, v is the scalar control of the evader, and $\lambda > 0$. x_0 , w_1^1 , w_0^1 are statistically independent zero-mean Gaussian random variables with unit variance.

Under the deterministic closed-loop information pattern which allows the evader to have access to the value of x_0 , i.e.,

$$v = \gamma_e(x_0),$$

and the pursuer to have access to both x_0 and x_1 , i.e.,

$$u = \gamma_p(x_1, x_0),$$

this zero-sum dynamic game admits the unique saddle-point solution

$$\gamma_p^*(x_1, x_0) = [1/(1 + \lambda)]x_1, \tag{14-1}$$

$$\gamma_e^*(x_0) = [2\lambda/(1 - 3\lambda)]x_0, \tag{14-2}$$

with the condition of existence being

$$0 < \lambda < 1/3, \tag{14-3}$$

That is, using the notation of section 3 under the right interpretation, λ_0 has the unique value $\lambda_0 = 1/3$, for this scalar problem.

If the additive noise terms w_1^1 and w_0^1 in (12) were absent, however, the solution would not have been unique. Paralleling the analysis of Ref. 7, we obtain the result that any pair

$$\gamma_p(x_1, x_0) = \psi(x_1, x_0), \quad (15-1)$$

$$\gamma_e(x_0) = [2\lambda/(1-3\lambda)]x_0, \quad (15-2)$$

with the side condition

$$\psi([1+\lambda]/(1-3\lambda)]x_0, x_0) \equiv [1/(1-3\lambda)]x_0, \quad (15-3)$$

then constitutes a saddle-point solution under the condition

$$4[1 - \bar{\psi}_{x_1}]^2 + 4\lambda(\bar{\psi}_{x_1})^2 - 1 < 0, \quad (16)$$

where

$$\bar{\psi}_{x_1} \triangleq [\nabla_{x_1}\psi(x_1, x_0)] = [(1+\lambda)/(1-3\lambda)]x_0.$$

Replacing $\bar{\psi}_{x_1}$ in (16) by

$$\bar{\psi}_{x_1} - [1/(1+\lambda)] + [1/(1+\lambda)]$$

and simplification yields the equivalent condition

$$3\lambda/(1+\lambda) + 8[\bar{\psi}_{x_1} - [1/(1+\lambda)]]^2(1+\lambda) < 0, \quad (17)$$

from which it follows that, in order to obtain the least stringent condition in λ , $\bar{\psi}_{x_1}$ must be picked so as to make the second nonnegative term in (17) zero, which in fact yields the solution (14-1), thus corroborating the conclusion arrived at in Ref. 7.

Now, let us instead endow the players with the noisy observations

$$y_0 = x_0 + w_0^2, \quad (18-1)$$

$$y_1 = x_1 + w_1^2, \quad (18-2)$$

where w_0^2 and w_1^2 are statistically independent Gaussian random variables with mean zero and variance unity. Furthermore, they are also assumed to be statistically independent of x_0 , w_0^1 , w_1^1 . The evader has access only to y_0 , i.e.,

$$v = \gamma_e(y_0),$$

whereas the pursuer has access to both y_0 and y_1 , i.e.,

$$u = \gamma_p(y_1, y_0).$$

Denote their respective strategy spaces by Γ_e and Γ_p , respectively.

Proposition 4.1. The stochastic dynamic game of this section does not admit a saddle point in $\Gamma_p \times \Gamma_e$.

Proof. Assume, to the contrary, that there exists a saddle point in $\Gamma_p \times \Gamma_e$, and denote that pair by (γ_p^*, γ_e^*) . Then, by definition of a saddle point, γ_p^* must minimize $J(\gamma_p, \gamma_e^*)$ over Γ_p . This minimization problem yields the unique relation

$$\begin{aligned} \gamma_p^*(y_1, y_0) &= [1/(1 + \lambda)]E\{x_1 | y_0, y_1, v = \gamma_e^*(\cdot)\} \\ &= [1/(1 + \lambda)][(1/3)y_1 + (1/3)y_0 + (4/3)\gamma_e^*(y_0)]. \end{aligned}$$

This implies that the structural dependence of $\gamma_p^*(y_1, y_0)$ on y_1 is linear and through the coefficient term $[1/(1 + \lambda)](1/3)$. This makes $J(\gamma_p^*(y_1, y_0), v)$ a quadratic function of v , and then the coefficient of the quadratic term in v can readily be determined to be

$$3 + [20/(1 + \lambda)9] > 0.$$

This clearly indicates that $J(\gamma_p^*(y_1, y_0), v)$ is a convex function of v , and thus the maximum of $J(\gamma_p^*, \gamma_e)$ over Γ_e does not exist, which contradicts the hypothesis that a saddle point exists. \square

We thus observe that the stochastic dynamic game under consideration in this section does not admit a saddle point, even though its deterministic version does. This does not imply, however, that the upper value of the stochastic game defined by the kernel $J(\gamma_p, \gamma_e)$ is unbounded, since there exists a well-defined security strategy for the pursuer. Letting

$$\gamma_p(y_1, y_0) = y_1,$$

the reader can easily verify for himself that the upper value is bounded for $\lambda < 1/4$, though this is not the security strategy of the pursuer. The security strategy of the pursuer is in fact in saddle-point equilibrium with an appropriate mixed strategy of the evader, which can be written as

$$\hat{\gamma}_e(y_0) = \alpha y_0 + \xi,$$

where α is a scalar and ξ is a Gaussian random variable with zero mean and an appropriate variance. This result can be verified by following an analysis parallel to the one adopted in Section 3 of Ref. 8 in solving a particular stochastic game. We will not include details of this analysis here, since it is not directly related to the main objective of this paper.

Since one can easily come up with examples of the preceding nature in the continuous-time setting, the message to be transmitted here is that the stochastic, zero-sum differential game of Section 2 might not admit a saddle-point solution, even though its deterministic version considered in

Section 3 does; this is contradictory to what is claimed in Ref. 1. Before concluding this section, we should remark that, if the pursuer has also access to the value of the evader's control v , then one pair of saddle-point strategies can readily be obtained from (14) to be

$$\gamma_p^*(y_1, y_0, v) = [1/(1+\lambda)]\hat{x}_1, \quad (19-1)$$

$$\gamma_e^*(y_0) = [2\lambda/(1-3\lambda)]\hat{x}_0, \quad (19-2)$$

where

$$\hat{x}_1 \triangleq E[x_1 | y_1, y_0, v] = \hat{x}_0 + v + \frac{1}{3}[y_1 - \hat{x}_0 - v],$$

$$\hat{x}_0 \triangleq (1/2)y_0,$$

with the existence condition being precisely (14-3). The reader can actually check this result by directly verifying that the pair (19) satisfies the relevant saddle-point inequalities. However, (19) is not the only saddle-point solution within this enlarged class. By going through an analysis parallel to the one followed in Example 1 of Ref. 9, it can be shown that any strategy of the form

$$\gamma_p(y_1, y_0, v) = [1/(1+\lambda)]\hat{x}_1 + q[v - [2\lambda/(1-3\lambda)]\hat{x}_0]$$

constitutes a saddle-point strategy for the pursuer in equilibrium with (19-2) for a wide range of values for the parameter q , the existence conditions involved (on λ) being different for each choice of q .

In the next section, we reconsider the stochastic differential game of Section 2 and obtain exact expressions for its saddle-point solution, by taking into consideration the pitfalls encountered within the context of the illustrative example of this section. We furthermore obtain a necessary condition for existence of a solution, which is more stringent than that of Ref. 1.

5. Main Results

We now obtain the saddle-point solution of the linear-quadratic stochastic differential game of Section 2, whenever it exists, and also a necessary condition of existence. To this end, let us first consider the hypothetical game in which the players are also endowed with the past values of the control vectors, and denote the corresponding spaces of permissible strategies for the pursuer and the evader by U_p and U_e , respectively. It should be noted that $U_p \supset \Gamma_p$ and $U_e \supset \Gamma_e$. We now first have the following proposition.

Proposition 5.1. The strategies (9-1) and (9-2) constitute a saddle-point pair in the product space $U_p \times U_e$.

Proof. It is, first of all, apparent that γ_p^0 given by (9-1) is an element of U_p (but not of Γ_p), and γ_e^0 given by (9-2) is an element of U_e (but not of Γ_e). Furthermore, \hat{x}_t given by (10) is the conditional mean of x_t , given y_0^t, v_0^t, u_0^t . Hence, the error term

$$\tilde{x}_t \triangleq x_t - \hat{x}_t$$

is independent of the controls of the players, and J can be written solely in the terms of $\hat{x}_t, t \geq t_0$, with some remainder terms depending only on $\tilde{x}_t, t \geq t_0$. This construction makes the problem no different from the deterministic game problem considered at the beginning of Section 3, with \hat{x}_t replacing x_t and the innovation term $dy_t - C\hat{x}_t dt$ replacing the Wiener process increment dw_t^1 . Then, the proposition follows from (7-1) and (7-2). The existence condition involved for the saddle point is clearly identical with the one of the deterministic problem. \square

Let us now assume that the stochastic differential game admits a saddle-point solution also in the space $\Gamma_p \times \Gamma_e$. Then, the following property can be proven.

Lemma 5.1. Let the stochastic differential game admit a saddle-point solution pair $\{\gamma_p^*, \gamma_e^*\}$ on the product space $\Gamma_p \times \Gamma_e$. Then, $\{\gamma_p^*, \gamma_e^*\}$ also constitutes a saddle-point pair on the product space $U_p \times U_e$.

Proof. By hypothesis, the pair of inequalities (4) is satisfied for all $\gamma_p \in \Gamma_p, \gamma_e \in \Gamma_e$. Let us first consider the left-hand-side inequality which defines a stochastic optimum control problem, i.e.,

$$(i) \quad J(\gamma_p^*, \gamma_e^*) = \max_{\Gamma_e} J(\gamma_p^*, \gamma_e) \leq \sup_{U_e} J(\gamma_p^*, \gamma_e),$$

where the last inequality follows, since $\Gamma_e \subset U_e$. Now, for each

$$\gamma_e = \gamma_e(t, y_0^t, v_0^t, u_0^t) \in U_e,$$

and with

$$u_s = \gamma_p^*(s, y_0^s),$$

there exists a unique solution

$$v_t = \hat{\gamma}_e(t, y_0^t), \quad t \geq t_0,$$

of the implicit equation

$$v_t = \gamma_e(t, y_0^t, v_0^t, \gamma_p^*(s, y_0^s)), \quad t \geq t_0,$$

because of Lipschitz continuity and causality of the strategies. Furthermore, this solution $\hat{\gamma}_e$ will be in Γ_e . Consequently, the inequality in (i) is in fact an equality. Analogously, it can be shown that

$$\min_{\Gamma_p} J(\gamma_p, \gamma_e^*) = \min_{U_p} J(\gamma_p, \gamma_e^*),$$

which completes the proof of the lemma. □

Lemma 5.2. Any two-saddle-point pairs in $U_p \times U_e$ are interchangeable.

Proof. Writing the kernel of the game as $J(\gamma_p, \gamma_e)$ suppresses all the dynamic and stochastic nature of the problem and brings it to the normal form. The result then follows as a well-known property of saddle points in zero-sum games (see, e.g., Ref. 10, p. 13). □

Theorem 5.1. Let the matrix Riccati equation (8) admit a unique symmetric solution on $[t_0, t_f]$. Then, the saddle-point solution of the stochastic differential game of Section 2 is unique, whenever it exists, and is given by

$$\gamma_p^*(t, y'_0) = -(1/\lambda)B'_p(t)S(t)z_0, \tag{20-1}$$

$$\gamma_e^*(t, y'_0) = B'_e(t)S(t)z_0, \tag{20-2}$$

where

$$dz_t = [A - (1/\lambda)B_p B'_p S + B_e B'_e S - KC]z_t dt + K dy_t, \tag{21-1}$$

$$t \geq t_0, \quad z_{t_0} = 0, \tag{21-2}$$

and $K(\cdot)$ is defined by (11).

Proof. Under the hypothesis of the theorem, and by Proposition 5.1, the strategy pair $\{\gamma_p^0, \gamma_e^0\}$ given by (9) exists as a saddle-point solution on the product space $U_p \times U_e$. Let $\{\hat{\gamma}_p \in \Gamma_p, \hat{\gamma}_e \in \Gamma_e\}$ be any solution to the stochastic differential game under consideration. Then, by Lemma 5.1, they are also in saddle-point equilibrium in the product space $U_p \times U_e$; and, by Lemma 5.2, the pairs $\{\hat{\gamma}_p, \gamma_e^0\}, \{\gamma_p^0, \hat{\gamma}_e\}$ are also saddle-point pairs on $U_p \times U_e$. Let us now consider the former pair, and in particular the right-hand side of the corresponding saddle-point inequality, which is

$$J(\hat{\gamma}_p, \gamma_e^0) \leq J(\gamma_p, \gamma_e^0), \quad \text{for all } \gamma_p \in U_p.$$

Writing out the explicit expression for $J(\gamma_p, \gamma_e^0)$, we have

$$J(\gamma_p, \gamma_e^0) = E \left\{ x_{t_f}' Q_f x_{t_f} + \int_{t_0}^{t_f} (x_t' Q(t) x_t + \lambda u_t' u_t - b_t' S B_e B_e' S b_t) dt \right\},$$

with $x_t, t \geq t_0$, and $b_t, t \geq t_0$, defined as the stochastic processes satisfying, respectively,

$$\begin{aligned} dx_t &= A x_t dt + B_p u_t dt + B_e B_e' S b_t dt + F dw_t^1, \\ x_{t_0} &= x_0, \quad t \geq t_0, \end{aligned}$$

and

$$\begin{aligned} db_t &= [A + B_e B_e' S - KC] b_t dt + B_p u_t dt + K dy_t, \\ b_{t_0} &= 0, \quad t \geq t_0, \end{aligned}$$

and with $K(\cdot)$ defined by (11). We thus observe that the problem of minimizing $J(\gamma_p, \gamma_e^0)$ over U_p is actually a standard LQG stochastic optimum control problem with state $(x_t', b_t)'$ of dimension $2n$. Since the coefficient of the quadratic term in this $2n$ -dimensional state in the cost function $J(\gamma_p, \gamma_e^0)$ is sign-indefinite, the corresponding Riccati equation could admit a conjugate point depending on the length of the time interval $[t_0, t_f]$. But we know that this is not the case here, since, by hypothesis, (γ_p^0, γ_e^0) constitutes a saddle-point pair. Hence, the kernel $J(\gamma_p, \gamma_e^0)$ is strictly convex in γ_p ; therefore, if we restrict u_t only to be a function of the observation y_0^t (i.e., if we restrict γ_p to Γ_p), the solution is clearly unique. This implies that $\hat{\gamma}_p$ is unique in Γ_p . In U_p , however, the solution of this minimization problem is not unique, since it is possible to write down different "representations" of the same control value in U_p . Nevertheless, the restriction of all these optimal representations to Γ_p is unique and is $\hat{\gamma}_p$.

Now, to obtain an expression for $\hat{\gamma}_p$, it is sufficient to note that, since $\{\gamma_p^0, \gamma_e^0\}$ constitutes a saddle-point pair, $\hat{\gamma}_p$ is in fact the unique restriction of γ_p^0 to Γ_p . Thus, $\hat{\gamma}_p$ can readily be obtained by substituting (9-1) and (9-2) into (10), and solving for the unique u_t that depends only on y_0^t , which is (20-1).

To prove validity of expression (20-2) as the unique saddle-point strategy of the evader, we start with the pair $\{\gamma_p^0, \hat{\gamma}_e\}$ considered as a saddle point, and repeat the preceding argument with some obvious modifications. The result then readily follows. □

In the light of the results obtained for the specific example of Section 4, existence of a matrix solution to the Riccati equation (8) will not be sufficient for the pair (20) to be in saddle-point equilibrium. In the following corollary, we provide a necessary condition which is in fact more stringent than the sufficiency condition of the deterministic problem.

Corollary 5.1. Assuming that the matrix Riccati equation (8) admits a well-defined solution on $[t_0, t_f]$, a necessary condition for the stochastic differential game of Section 2 to admit a saddle point is the existence of a symmetric solution $M(\cdot)$ to the $2n$ -dimensional matrix Riccati equation (22) on the interval $[t_0, t_f]$:

$$\dot{M} + M\tilde{A} + \tilde{A}'M + M\tilde{B}_e\tilde{B}'_eM + \tilde{Q} = 0, \tag{22-1}$$

$$M(t_f) = \tilde{Q}_f, \tag{22-2}$$

where

$$\tilde{B}'_e \triangleq [B'_e, 0], \tag{23-1}$$

$$\tilde{Q} \triangleq \text{diag}[Q, (1/\lambda)SB_pB'_pS], \tag{23-2}$$

$$\tilde{Q}_f \triangleq \text{diag}[Q, 0], \tag{23-3}$$

$$\tilde{A} \triangleq \left[\begin{array}{c|c} A & -(1/\lambda)B_pB'_pS \\ \hline KC & A - (1/\lambda)B_pB'_pS + B_eB'_eS - KC \end{array} \right]. \tag{23-4}$$

Proof. Under the hypothesis of the corollary, we know from Theorem 5.1 that the saddle point of the stochastic differential game of Section 2 has to be given by (20), whenever it exists. Hence, to prove the corollary, it suffices to show that (20) does not constitute a saddle-point pair if the Riccati equation (22) does not admit a well-defined solution. To this end, we consider the right-hand side inequality (4), with γ_p^* and γ_e^* given by (20-1) and (20-2), respectively. This characterizes a stochastic control problem for the evader, which is maximization of the cost function $J(\gamma_p^*, \gamma_e)$. This cost function can explicitly be written as

$$J(\gamma_p^*, v_t) = E \left\{ m'_t \tilde{Q}_f m_t + \int_{t_0}^{t_f} (m'_t \tilde{Q}(t) m_t - v'_t v_t) dt \right\}, \tag{24}$$

with $m_t, t \geq t_0$, satisfying the $2n$ -dimensional stochastic differential equation

$$dm_t = \tilde{A}m_t dt + \tilde{B}_e v_t dt + [0, K']' dy_t + [F, 0]' dw_t^1, \tag{25-1}$$

$$m_{t_0} = [x'_{t_0}, 0]', \quad t \geq t_0, \tag{25-2}$$

where \tilde{A} is defined by

$$\tilde{A} = \tilde{A} - \begin{bmatrix} 0 & 0 \\ KC & 0 \end{bmatrix},$$

and all the other relevant terms are defined by (23-1) through (23-4). Furthermore, dy_t can be written as

$$dy_t = [C, 0]m_t dt + G dw_t^2, \quad t \geq t_0.$$

This implies that the stochastic optimum control problem faced by the evader is the maximization of (24), subject to the state equation

$$dm_t = \tilde{A}m_t dt + \tilde{B}_e v_t dt + \tilde{F} dw_t, \tag{26-1}$$

$$m_{t_0} = [x'_0, 0]', \quad t \geq t_0, \tag{26-2}$$

and under the observation equation

$$dy_t = [C, 0]m_t dt + \tilde{G} dw_t, \tag{26-3}$$

where

$$w_t \triangleq (w_t^1, w_t^2)'$$

is an $(n + m)$ -dimensional standard Wiener process,

$$\tilde{G} \triangleq [G, 0],$$

and

$$\tilde{F} \triangleq \begin{bmatrix} F & 0 \\ 0 & KG \end{bmatrix}.$$

This is an LQG problem which admits a solution in Γ_e , say $\tilde{\gamma}_e$ (which is also unique), if the Riccati equation (22) admits a symmetric solution $M(\cdot)$ on the interval $[t_0, t_f]$, in which case $\tilde{\gamma}_e$ is given by

$$\tilde{\gamma}_e(t, y'_0) = \tilde{B}'_e(t)M(t)\hat{m}_t, \tag{27}$$

with \hat{m}_t satisfying the stochastic differential equation

$$d\hat{m}_t = \tilde{A}\hat{m}_t dt + \tilde{B}_e \tilde{B}'_e M \hat{m}_t dt + (\tilde{P}[C, 0]' + \tilde{F}\tilde{G}')(\tilde{G}\tilde{G}')^{-1}(dy_t - [C, 0]\hat{m}_t dt), \tag{28-1}$$

$$\hat{m}_{t_0} = 0, \quad t \geq t_0, \tag{28-2}$$

and $\tilde{P}(t)$ satisfying

$$\dot{\tilde{P}} = \tilde{A}\tilde{P} + \tilde{P}\tilde{A}' + \tilde{F}\tilde{F}' - (\tilde{P}[C, 0]' + \tilde{F}\tilde{G}')(\tilde{G}\tilde{G}')^{-1}([C, 0]'\tilde{P} + \tilde{G}\tilde{F}'), \tag{28-3}$$

$$\tilde{P}(t_0) = \begin{bmatrix} P_0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{28-4}$$

See, e.g., Ref. 1, Chapter 7. It should be noted that the required inverse in these expressions exists, since

$$\tilde{G}\tilde{G}' = [G, 0][G, 0]' = GG' > 0.$$

Therefore, this implies that, if the Riccati equation (22) has a conjugate point in the interval $[t_0, t_f]$, then there will exist a sequence of controls for the

evader which will make the value of $J(\gamma_p^*, v_t)$ sufficiently large without bound, which contradicts the right-hand side inequality (4). \square

Remark 5.1. It is now almost immediate that the necessary condition of Corollary 5.1 is more stringent than the condition of existence of a solution to (8), since, if λ is picked so that

$$(1/\lambda)B_p B_p' - B_e B_e' > \epsilon I,$$

for a sufficiently small $\epsilon > 0$, then a solution to (8) always exists (regardless of the length of the time interval), but the Riccati equation (22) will definitely have a conjugate point if the time interval is taken sufficiently long.

Remark 5.2. Even though the condition of Corollary 5.1 has been shown to be only necessary, it can be verified that it is also sufficient for the stochastic differential game of Section 2 to admit a saddle-point solution which is given by (20). A verification of this result involves some extensive manipulations in showing that:

$$(i) \quad \tilde{\gamma}_e \equiv \gamma_e^*,$$

where the former is given by (27); and

$$(ii) \quad \gamma_p^* \equiv \arg \min_{\Gamma_p} J(\gamma_p, \gamma_p^*),$$

under the conditions of Corollary 5.1.

In fact, a recent independent study (Ref. 11) proves that the same condition is sufficient for (20) to provide a saddle-point equilibrium. In other words, the author of Ref. 11 starts with the strategies (20-1)–(20-2), and establishes their saddle-point property by solving two stochastic control problems and under the condition of Corollary 5.1. Such a “verification-oriented approach”, however, is not constructive, and besides it does not lead to a complete solution of the problem; that is, the question of whether other pure-strategy saddle-point solutions exist whenever (20) ceases to be an equilibrium solution remains unanswered. The indirect derivation of this paper, however, answers this question completely; that is, the strategy pair (20) provides the only saddle-point equilibrium in pure strategies.

6. Conclusions

For a class of zero-sum stochastic differential games, we used an indirect approach to obtain the saddle-point solution in closed form, with the corresponding strategies realized as the output of an n -dimensional

system, where n is the dimension of the state. The existence conditions involved, however, are more stringent than those of the deterministic version of the problem; and, if they are not satisfied, then this implies that the stochastic differential game does not admit a saddle-point equilibrium in pure strategies. Nevertheless, the discrete-time example solved in Section 4 hints that a saddle-point value might still exist, but this time within the class of mixed strategies (for the evader). Derivation of this mixed-strategy saddle-point solution seems, at this point, quite challenging, but Refs. 8 and 12 could provide the reader with considerable insight into this difficult problem.

Yet another class of problems of interest are the stochastic, nonzero-sum differential games structured in the same way as the differential game of Section 2. But the method of approach used in this paper cannot be utilized to obtain the Nash equilibrium solutions of such differential games, since Nash solution pairs do not possess the interchangeability property (i.e., Lemma 5.2 is not valid in that case). A valid approach then would be an extension of the one used in Ref. 4, but that would require an *a priori* assumption (such as linearity) on the structure of the Nash strategies, and even then it is not conceivable that the Nash equilibrium strategies will be finite dimensional. For the discrete-time version of the same problem, however, a special version of the method of Ref. 13 can be utilized to determine the Nash equilibrium solutions (though not in recursive form). In that case, the Nash equilibrium strategies will be unique and linear, whenever they exist.

References

1. BALAKRISHNAN, A. V., *Stochastic Differential Systems, I*, Lecture Notes in Economics and Mathematical Systems, Vol. 84, Springer-Verlag, Berlin, Germany, 1973.
2. FLEMING, W. H., *Optimal Continuous-Parameter Stochastic Control*, SIAM Review, Vol. 11, 1969.
3. FLEMING, W. H., and NISIO, M., *On the Existence of Optimal Stochastic Controls*, Journal of Mathematical Mechanics, Vol. 15, No. 5, 1966.
4. BAGCHI, A., and OLSDER, G. J., *Linear-Quadratic Stochastic Pursuit-Evasion Games*, Journal of Applied Mathematics and Optimization (to appear).
5. HO, Y. C., *On the Minimax Principle and Zero-Sum Stochastic Differential Games*, Journal of Optimization Theory and Applications, Vol. 13, No. 3, 1974.
6. BAŞAR, T., *Existence of Unique Nash Equilibrium Solutions in Nonzero-sum Stochastic Differential Games*, Differential Games and Control Theory II, Edited by E. O. Roxin, P. T. Liu, and R. L. Sternberg, Marcel Dekker, New York, New York, 1977.

7. BAŞAR, T., *Some Thoughts on Saddle-Point Conditions and Information Structures in Zero-Sum Differential Games*, Journal of Optimization Theory and Applications, Vol. 18, No. 1, 1976.
8. BAŞAR, T., and MINTZ, M., *A Multistage Pursuit-Evasion Game That Admits a Gaussian Random Process as a Maximin Control Policy*, Stochastics, Vol. 1, No. 1, 1973.
9. BAŞAR, T., *Decentralized Multicriteria Optimization of Linear Stochastic Systems*, IEEE Transactions on Automatic Control, Vol. AC-23, No. 2, 1978.
10. OWEN, G., *Game Theory*, Saunders, Philadelphia, Pennsylvania, 1968.
11. KUMAR, P. R., *A Stochastic Differential Game with Jump Process Observations*, UMBC, Mathematics Research Report No. 78-7, 1978.
12. HEXNER, G., and MINTZ, M., *On the Solution of a Minimax Terminal State Estimation Problem: A Hilbert Space Approach*, International Journal of Control, Vol. 27, No. 6, 1978.
13. BAŞAR, T., *Two-Criteria LQG Decision Problems with One-Step Delay Observation Sharing Pattern*, Information and Control, Vol. 38, No. 1, 1978.