Automatica, Vol. 16, pp. 409–414
Pergamon Press Ltd. 1980. Printed in Great Britain

0005-1098/80/0701-0409 $02.00/0

Brief Paper

Team-Optimal Closed-Loop Stackelberg Strategies in Hierarchical Control Problems*

TAMER BAŞAR† and GEERT JAN OLSDER‡

Key Words—Differential games; game theory; hierarchical systems; large-scale systems; linear systems; multicriteria optimization; optimal control; optimal systems; Stackelberg strategies; team theory.

Abstract—This paper is concerned with the derivation of closed-loop Stackelberg (CLS) solutions of a class of continuous-time two-player nonzero-sum differential games characterized by linear state dynamics and quadratic cost functionals. Explicit conditions are obtained for both the finite and infinite horizon problems under which the CLS solution is a representation of the optimal feedback solution of a related team problem which is defined as the joint minimization of the leader's cost function. First, a specific class of representations is considered which depend linearly on the current and initial values of the state, and then the results are extended to encompass a more general class of linear strategies that also incorporate the whole past trajectory. The conditions obtained all involve solutions of linear matrix equations and are amenable to computational analysis for explicit determination of CLS strategies.

1. *Introduction*

AN APPROPRIATE solution concept for hierarchical multicriteria decision problems is the Stackelberg solution concept which was first introduced in economics within the context of static economic competition (von Stackelberg, 1934). Its dynamic version later entered the control literature through the works of Chen and Cruz (1972) and Simaan and Cruz (1973a, b), and found applications in nonzero-sum differential games where one player has enough ability or power to enforce his strategy on the other player(s). Within the context of two-player differential games, the more powerful player is called the leader and the other one is called the follower. An extension is, of course, possible to 'one leader'–'many followers' and even to 'many leaders'–'many followers' situations.

The initial application of the Stackelberg solution concept to differential games has been within the framework of the open-loop information structure, and the extension to closed-loop information structures has remained as a challenge for a long time. The main difficulty arises from the fact that under closed-loop information, the reaction set of the follower cannot be determined explicitly, even in the case of linear quadratic (LQ) problems, thus resulting in a nonclassical control problem faced by the leader. One way to circumvent this difficulty is to assume specific parametric structures for the strategy of the leader, in which case the follower's reaction can be explicitly determined, and then the problem faced by the leader is to optimize on those parameters subject to the constraint imposed by the reaction set of the follower. Such an approach has, in fact, been adopted by Medanic (1977), within the context of LQ nonzero-sum two-person differential games and under a linear feedback structure for the strategies of both players, and the author arrives at a set of intercoupled complicated equations in terms of the coefficient matrices sought and depending on the probability distribution of the initial state.

Quite recently, a new indirect approach has been proposed to obtain the closed-loop Stackelberg (CLS) solution of nonzero-sum differential games, which, in essence, relates the CLS strategy of the leader to a representation of optimal feedback solution of a particular team problem (Başar and Selbuz, 1979b). This approach has been carried out by Başar and Selbuz (1979b) to obtain the CLS solution of a class of two-person nonzero-sum dynamic games described by linear difference equations and quadratic cost functionals, yielding linear CLS strategies for both the leader and the follower, with the former's being of the one-step memory type. Besides, the optimal strategies can be obtained recursively. A specific example solved in detail by Basar and Selbuz (1979a) clearly displays several important features of the new approach taken, and properties of the CLS strategies obtained within that context. The essentials of this new technique are as follows: first the leader's cost function is minimized over the controls of the leader and the follower, yielding optimal (team) strategies for both players in feedback form, together with an optimal trajectory. But there exist different CL representations of the leader's optimal team strategy on this optimal trajectory. Then the question is whether there exists one particular representation for the leader to which the optimal response of the follower (obtained by minimization of his own cost function) coincides with his (the follower's) team strategy, thus yielding the optimal team trajectory. If it exists, such a strategy would clearly be the leader's CLS strategy, since it provides the leader with an optimal team cost which is the lower bound on the CLS cost. It has been shown in Başar and Selbuz (1979b) that, within the context of the specific class of dynamic game problems considered, such an approach can effectively be used to obtain the CLS strategy of the leader.

In the present paper, we extend this approach to two-person continuous-time linear differential game problems under quadratic cost functions for both players, and we treat both the finite-horizon and the infinite-horizon cases. The specific techniques used here to obtain the CLS solution are inherently different from those used by Basar and Selbuz (1979b) for the discrete-time version. In the main bulk of the paper the class of admissible representations for the leader are assumed to depend only on the current value and the initial value of the state vector, whereas in the last section an extension to more general memory representations is discussed. Almost concurrently with this work is the recent paper by Papavassilopoulos and Cruz (1979), who have also utilized the approach of Basar and Selbuz (1979b) in the continuous-time setting, to obtain a set of sufficiency conditions for the leader to be able to enforce the optimal trajectory within the class of linear memory strategies described by Lebesgue-Stieltjes measures. Within this general setting, the sufficiency conditions obtained by the authors are in the form of integro-differential equations, which are difficult to manipulate even in the scalar case. We should also mention that the specific techniques used by Papavassilopoulos and Cruz (1979) to arrive at those

*Received 8 February 1979; revised 5 February 1980. The original version of this paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor D. Tabak. Research reported in this paper was performed while the first author was on sabbatical leave at Twente University of Technology, Department of Applied Mathematics, Enschede, The Netherlands.

†Applied Mathematics Division, Marmara Research Institute, P.K. 21, Gebze, Kocaeli, Turkey.

‡Department of Applied Mathematics, Twente University of Technology, P.O. Box 217, 7500 AE Enschede, The Netherlands.

sufficiency conditions are inherently different from those employed in this paper.

In the next section, we provide a formulation of the continuous-time Stackelberg problem to be considered in the paper, and give the optimal feedback solution of the related team problem. In Section 3, we obtain conditions (in the form of linear matrix differential equations) for a particular representation of the leader's team strategy to force the follower to the team solution. The obtained CLS strategy of the leader depends explicitly on the current and the initial values of the state, while that of the follower is in feedback form. Section 4 treats the infinite-horizon version of the general problem, and obtains explicit and computationally attractive expressions. It also includes the Stackelberg solution of a nonscalar differential game problem. Section 5 contains an extension of results of Section 3 to more general representations that involve continuous memory, so that the leader can have flexibility in his choices. A conclusion section ends the paper.

2. *General problem statement and solution of the related team problem*

Using the conventional notation for continuous-time systems, we let the evolution of the state be described by the differential equation

$$\frac{dx}{dt}=A(t)x+B_1(t)u+B_2(t)v,$$
$$x(0)=x_0,\quad 0\leq t\leq T, \tag{2.1}$$

where $u(\cdot)$ is an r-vector function controlled by Player 1 (the leader), $v(\cdot)$ is an r'-vector function controlled by Player 2 (the follower), $x(\cdot)$ is the state vector of dimension n, $A(\cdot)$, $B_1(\cdot)$ and $B_2(\cdot)$ are matrices of appropriate dimensions and with entries continuous on $[0,T]$. All this information and the value of the initial state x_0 are known to both players. The terminal time T will be assumed, for the moment, to be a finite number. In Section 4, however, we will also consider the infinite-horizon problem.

The cost functions for Players 1 and 2 are given by J_1 and J_2, respectively, where

$$J_i=\tfrac{1}{2}x'(T)Q_{if}x(T)+\tfrac{1}{2}\int_0^T\{x'Q_i(t)x$$
$$+u'R_{i1}(t)u+v'R_{i2}(t)v\}\,dt \tag{2.2}$$

with $R_{11}(\cdot)>0$, $R_{22}(\cdot)>0$, $R_{12}(\cdot)>0$, all other weighting matrices being nonnegative definite, and all having entries continuous on $[0,T]$. The symbol $'$ denotes the transpose operation.

Each player has access to closed-loop perfect state information and can utilize that in the choice of his control. We denote typical strategies for the leader and the follower by γ_1 and γ_2, respectively, whose realizations are denoted above by u and v, respectively; that is $u(t)=\gamma_1(t,x(s),s\leq t)$ and $v(t)=\gamma_2(t,x(s),s\leq t)$. We further assume that γ_1 and γ_2 satisfy the standard measurability, continuity and growth conditions so that the differential equation admits a unique solution on $[0,T]$, for each pair (γ_1,γ_2). We denote the space of such admissible strategies for Player i by Γ_i, $i=1,2$. Finally, we denote the value of J_i given by (2.2) for an admissible pair of strategies (γ_1,γ_2) by $J_i(\gamma_1,\gamma_2)$.

Within the framework of this differential game problem, the closed-loop Stackelberg (CLS) solution concept can be introduced as follows: The leader determines a strategy $\gamma_1\in\Gamma_1$ and announces it ahead of time. Then, knowing what the leader's strategy (but *not* the control value) is, the follower chooses a $\gamma_2\in\Gamma_2$ so as to minimize $J_2(\gamma_1,\gamma_2)$. Since this choice affects the cost incurred to the leader, the leader has to take this into account before announcing his strategy. This then leads to the following definition of a CLS strategy for the leader, where we also take into consideration the possibility that the follower's reaction to every strategy of the leader might not be unique.

Definition 1 For the differential game formulated above, a strategy $\gamma_1^*\in\Gamma_1$ constitutes a *closed-loop Stackelberg* (CLS) *strategy* for the leader if

$$\sup_{R(\gamma_1^*)}J_1(\gamma_1^*,\gamma_2)\leq\sup_{R(\gamma_1)}J_1(\gamma_1,\gamma_2),\quad\forall\gamma_1\in\Gamma_1, \tag{2.3}$$

where $R(\gamma_1)$ denotes the rational reaction set of the follower defined by

$$R(\gamma_1)=\{\gamma_2^0\in\Gamma_2:J_2(\gamma_1,\gamma_2^0)\leq J_2(\gamma_1,\gamma_2),\quad\forall\gamma_2\in\Gamma_2\}. \tag{2.4}$$

The CLS cost of the leader is then

$$J_1^*=\sup_{R(\gamma_1^*)}J_1(\gamma_1^*,\gamma_2)=\inf_{\Gamma_1}\sup_{R(\gamma_1)}J_1(\gamma_1,\gamma_2). \tag{2.5}$$

Furthermore, any strategy $\gamma_2^*\in R(\gamma_1^*)$ for which $J_1(\gamma_1^*,\gamma_2^*)=J_1^*$ is known as a CLS strategy for the follower. □

Now, as it has been discussed in Section 1 (and more extensively by Başar and Selbuz (1979b) for the discrete-time problem), an indirect approach to obtain the CLS solution is first to determine an easily computable lower bound on the CLS cost J_1^* and then to seek existence of a strategy for the leader, which will force the follower to such a behaviour that the eventual cost incurred to the leader is that lower bound. To this end, we first note the natural bound

$$J_1^*\geq\min_{\Gamma_1}\min_{\Gamma_2}J_1(\gamma_1,\gamma_2)=J_1(\tilde\gamma_1,\tilde\gamma_2)\overset{\Delta}{=}\tilde J_1 \tag{2.6}$$

obtained by jointly minimizing $J_1(\gamma_1,\gamma_2)$ over Γ_1 and Γ_2, which is, in fact, the best the leader can hope to accomplish. The RHS of (2.6) defines a team problem that admits a well-defined solution, since $J_1(\gamma_1,\gamma_2)$ is a strictly convex functional of its arguments. Furthermore, using standard results of LQ control theory, its unique solution in feedback form can readily be determined to be

$$\tilde\gamma_i(t,x(t))=-R_{1i}^{-1}(t)B_i'(t)S(t)x(t),\quad i=1,2 \tag{2.7}$$

with $S(\cdot)$ defined as the unique nonnegative-definite matrix solution of

$$\dot S=-Q_1-A'S-SA+S[B_1R_{11}^{-1}B_1'+B_2R_{12}^{-1}B_2']S;$$
$$S(T)=Q_{1f}, \tag{2.8}$$

where we have suppressed the time dependence of the matrices, for ease in notation. The unique trajectory resulting from application of strategies (2.7) is

$$\bar x(t)=\Phi(t,0)x_0 \tag{2.9}$$

where $\Phi(t,\tau)$ is the state transition matrix of the system

$$\dot x=[A-(B_1R_{11}^{-1}B_1'+B_2R_{12}^{-1}B_2')S]x. \tag{2.10}$$

Let us now consider a subset of Γ_1 consisting of all strategies for the leader with open-loop representation identical to $\tilde\gamma_1(t,\bar x(t))$, that is, introduce a set $\tilde\Gamma_1\subset\Gamma_1$ defined as

$$\tilde\Gamma_1=\{\gamma_1\in\Gamma_1:\gamma_1(t,\bar x(s),s\leq t)$$
$$=-R_{11}^{-1}(t)B_1'(t)S(t)\Phi(t,0)x_0\}. \tag{2.11}$$

Analogously, define a set $\tilde\Gamma_2\subset\Gamma_2$ for the follower:

$$\tilde\Gamma_2=\{\gamma_2\in\Gamma_2:\gamma_2(t,\bar x(s),s\leq t)$$
$$=-R_{12}^{-1}(t)B_2'(t)S(t)\Phi(t,0)x_0\}. \tag{2.12}$$

Our interest in these sets stems from the following property:

Proposition 1 If there exists a strategy $\gamma_1^* \in \tilde{\Gamma}_1$ such that all solutions of the minimization problem

$$\inf_{\Gamma_2} J_2(\gamma_1^*, \gamma_2) \tag{2.13}$$

lie in $\tilde{\Gamma}_2$, then γ_1^* constitutes a CLS strategy for the leader and $\tilde{\gamma}_2$ is a CLS strategy for the follower.

Proof This follows readily, since then the lower bound in (2.6) is attained. The reason why we do not allow any solution to (2.13) outside $\tilde{\Gamma}_2$, is because then, if the follower adopts that particular strategy, the team cost might not be attained. □

In the remaining sections of the paper we will investigate to what extent the hypothesis of Proposition 1 is satisfied for the class of LQ differential game problems formulated in this section. First, we will restrict ourselves to an even smaller subset of $\tilde{\Gamma}_1$, comprised of strategies of the form

$$\gamma_1(t, x(t), x_0) = -R_{11}^{-1}(t)B_1'(t)S(t)x(t) + P'(t)[x(t) - \bar{x}(t)] \tag{2.14}$$

which depend at time t only on $x(t)$ and x_0. Later, in Section 5, we will extend our approach to encompass a broader class of strategies. Throughout the analysis we will suppress the t-dependences of all the vectors and matrices involved, whenever there is no ambiguity from the context.

3. *Sufficient conditions for the leader to enforce his team solution within a specific representation class*

Assuming that the leader chooses a strategy in the form (2.14), where $P'(\cdot)$ is an $(r \times n)$-dimensional matrix whose entries are continuous functions on $[0, T]$, then the follower's reaction will be determined from the solution of the following LQ optimal control problem:

$$\begin{aligned}\min_{\gamma_2 \in \Gamma_2} \Big[\tfrac{1}{2}x'(T)Q_{2f}x(T) &+ \tfrac{1}{2}\int_0^T [x'\{Q_2 + (P' - R_{11}^{-1}B_1'S)'R_{21}(P' - R_{11}^{-1}B_1'S)\}x \\ &- 2x'(P' - R_{11}^{-1}B_1'S)'R_{21}P'\bar{x} + u_2'R_{22}u_2]\,dt\Big]\end{aligned} \tag{3.1}$$

subject to

$$\frac{dx}{dt} = (A - B_1R_{11}^{-1}B_1'S + B_1P')x - B_1P'\bar{x} + B_2u_2, \qquad x(0) = x_0 \tag{3.2}$$

and with

$$u_2(t) = \gamma_2(t, x(s), s \leqq t). \tag{3.3}$$

Now, since (3.1) defines a strictly convex function in $u_2(\cdot)$, the preceding optimization problem admits a unique open-loop solution in Γ_2 for each fixed matrix function $P(\cdot)$. Hence, for the class of representations (2.14), we can replace $\tilde{\Gamma}_2$ in the statement of Proposition 1, with the larger set Γ_2, without any loss of generality.

A straightforward approach to test the hypothesis of Proposition 1, in this case, would be first to obtain the solution of the above minimization problem (using any one of the available standard techniques), and then to investigate existence of a matrix $P(\cdot)$ for which the unique open-loop representation of the optimal strategy found coincides with the unique open-loop strategy in $\tilde{\Gamma}_2$. But such an approach leads to rather cumbersome expressions which are difficult to manipulate.* An alternative to this approach is to make use of the sought property of equivalence between the optimal trajectory of (3.1) and (3.2) and the optimal trajectory of the team problem defined in Section 2, throughout the derivation. To this end, we adopt the Hamiltonian approach (whose first order conditions are also sufficient since we are dealing with a strictly convex LQ problem), and define the Hamiltonian H corresponding to (3.1) and (3.2) as

$$\begin{aligned}H = \tfrac{1}{2}x'\{Q_2 + (P' - R_{11}^{-1}B_1'S)'R_{21}(P' - R_{11}^{-1}B_1'S)\}x \\ - x'(P' - R_{11}^{-1}B_1'S)'R_{21}P'\bar{x} + \tfrac{1}{2}u_2'R_{22}u_2 \\ + \mu'(A - B_1R_{11}^{-1}B_1'S + B_1P')x - \mu'B_1P'\bar{x} + \mu'B_2u_2,\end{aligned} \tag{3.4}$$

where $\mu(\cdot)$ denotes the n-dimensional costate vector, satisfying the differential equation

$$\begin{aligned}\dot{\mu} = -\frac{\partial H}{\partial x} &= -\{Q_2 + (P' - R_{11}^{-1}B_1'S)'R_{21}(P' - R_{11}^{-1}B_1'S)\}x \\ &\quad + (P' - R_{11}^{-1}B_1'S)'R_{21}P'\bar{x} \\ &\quad - (A - B_1R_{11}^{-1}B_1'S + B_1P')'\mu; \\ \mu(T) &= Q_{2f}x(T).\end{aligned} \tag{3.5}$$

Maximization of H with respect to u_2 yields

$$u_2(t) = -R_{22}^{-1}B_2'\mu(t), t \in [0, T], \tag{3.6}$$

and this has to be matched, in view of (3.5), with any element of $\tilde{\Gamma}_2$, in particular with the open-loop strategy.

Let us now assume, for the moment, that there exists a matrix function $P(\cdot)$ such that the hypothesis of Proposition 1 is valid. Then, equating (3.6) with $\tilde{\gamma}_2(t, \bar{x}(t))$, we obtain the relation

$$R_{22}^{-1}(t)B_2'(t)\mu(t) = R_{12}^{-1}(t)B_2'(t)S(t)\bar{x}(t), \quad t \in [0, T], \tag{3.7}$$

and furthermore, since the optimal trajectories of the team problem and the present optimal control problem will coincide, we can let $x(\cdot) \equiv \bar{x}(\cdot)$ in (3.5) to yield the differential equation

$$\begin{aligned}\dot{\mu} = \{Q_2 - (P' - R_{11}^{-1}B_1'S)'R_{21}R_{11}^{-1}B_1'S\}\bar{x} \\ - (A - B_1R_{11}^{-1}B_1'S + B_1P')'\mu; \\ \mu(T) = Q_{2f}\bar{x}(T),\end{aligned} \tag{3.8}$$

whose solution depends linearly on $\bar{x}(\cdot)$, which in turn is related to x_0 through an invertible transformation. Therefore, letting $\mu(t) = M(t)\bar{x}(t)$ in (3.8), where $M(\cdot)$ is an $(n \times n)$ matrix function whose entries are continuously differentiable on $[0, T]$, we readily obtain the following linear matrix differential equation for M to satisfy

$$\begin{aligned}\dot{M} + MF + F'M + PB_1'M - PR_{21}R_{11}^{-1}B_1'S + Q_2 \\ + SB_1R_{11}^{-1}R_{21}R_{11}^{-1}B_1'S + SB_2R_{12}^{-1}R_{22}R_{12}^{-1}B_2'S = 0; \\ M(T) = Q_{2f},\end{aligned} \tag{3.9}$$

where

$$F \triangleq A - (B_1R_{11}^{-1}B_1' + B_2R_{12}^{-1}B_2')S. \tag{3.10}$$

An equivalent relation replacing (3.7) is now

$$R_{22}^{-1}(t)B_2'(t)M(t) = R_{12}^{-1}(t)B_2'(t)S(t), \tag{3.11}$$

and hence, if there exists a $P(\cdot)$ such that the solution of (3.9) satisfies (3.11), the strategy (2.14) with that particular $P(\cdot)$ constitutes a CLS strategy for the leader since it then satisfies the hypothesis of Proposition 1. This conclusion is made precise in the following Theorem:

*This has actually been the approach taken by Papavassilopoulos and Cruz (1979) to obtain a set of sufficient conditions for the desired property to hold true.

Condition 1 There exists an $(n\times r)$ matrix function $P(\cdot)$ with continuous entries so that the unique solution of the linear matrix differential equation (3.9) satisfies the relation (3.11), where $S\geq 0$ is defined as the unique solution of (2.8) and the team trajectory $\bar{x}(\cdot)$ is given by (2.9). □

Theorem 1 Let Condition 1 be satisfied for a matrix function $P^*(\cdot)$. Then, there exists a CLS solution for the differential game of Section 2, which is given by

$$\gamma_1^*(t,x(t),x_0)=[P^{*\prime}(t) \\ -R_{11}^{-1}(t)B_1'(t)S(t)]x(t)-P^{*\prime}(t)\bar{x}(t), \quad (3.12a)$$

$$\gamma_2^*(t,x(t))=-R_{12}^{-1}(t)B_2'(t)S(t)x(t), \quad (3.12b)$$

and the CLS costs for the leader and the follower are given respectively by

$$J_1^*=\tfrac{1}{2}x_0'S(0)x_0, \quad (3.13a)$$

$$J_2^*=\tfrac{1}{2}x_0'M(0)x_0, \quad (3.13b)$$

Proof As discussed prior to the statement of the Theorem, this result follows readily since (3.12a) satisfies the hypothesis of Proposition 1 under Condition 1, in other words, if (3.12a) is adopted as a policy by the leader, the follower's unique optimal reaction in feedback strategies is the one given by (3.12b) which is also the team strategy obtained in Section 2. Furthermore, since (3.12a) is a representation of (2.7), the trajectory determined by (3.12) is precisely the team trajectory $\bar{x}(\cdot)$, and hence the team cost will be equal to CLS cost for the leader, which is (3.13a). To verify (3.13b) as the CLS cost for the follower, it is sufficient to note that the minimum value of the LQ optimal control problem faced by the follower is given by $\frac{1}{2}x'\mu(0)$, where μ is defined by (3.5), from which expression (3.13b) follows by inspection under Condition 1. □

4. *The infinite-horizon problem*

We now consider the closed-loop Stackelberg solution of the differential game problem for the case when $T\to\infty$. For a meaningful formulation, we take all the parameters of the problem as constants, and further let $Q_{1f}=Q_{2f}=0$. Since this infinite-horizon problem can be considered as the limiting case of the finite-horizon differential game, versions of Condition 1 and Theorem 1 can easily be obtained by letting the derivative terms vanish in the differential equations for S and M, but provided that certain stabilizability conditions are satisfied. To this end, we first note that the related team problem admits the unique feedback solution

$$\bar{\gamma}_i(x(t))=-R_{1i}^{-1}B_i'Sx(t), \quad i=1,2 \quad (4.1)$$

with $S>0$ defined as the unique solution of

$$S[B_1R_{11}^{-1}B_1'+B_2R_{12}^{-1}B_2']S-SA-A'S-Q_1=0, \quad (4.2)$$

provided that the pair

$$\{A,(B_1,B_2)\} \quad (4.3)$$

is stabilizable,* a condition which we assume to hold *a priori*. Then, the feedback matrix

$$F=A-(B_1R_{11}^{-1}B_1'+B_2R_{12}^{-1}B_2')S \quad (4.4)$$

is a stable matrix, and the optimal team trajectory satisfies the differential equation

$$\dot{\bar{x}}=F\bar{x}, \quad x(0)=x_0. \quad (4.5)$$

*Together with R_{11}, R_{12}, $Q_1>0$, this is a necessary and sufficient condition for existence of a unique solution to the team problem. If, however, $Q_1\geq 0$, instead of $Q_1>0$, the said condition is only sufficient.

The statement of Proposition 1 is equally valid in this case with some obvious notational modifications, as well as the approach taken in Section 3 to obtain a candidate Stackelberg solution for the leader. Hence, adopting the representation

$$\gamma_1(t,x(t),x_0)=-R_{11}^{-1}B_1'Sx(t)+P'[x(t)-\bar{x}(t)], \quad (4.6)$$

for the leader, where P is an $(n\times r)$ matrix yet to be determined, we seek to solve the infinite-horizon optimization problem faced by the follower. This optimization problem is as defined by (3.1) and (3.2) with only $Q_{2f}=0$ and $T=\infty$. Since $\bar{x}(t)$ approaches zero exponentially as $t\to\infty$, a sufficient condition for the optimization problem to be meaningful is to assume that P can be chosen such that the pair

$$\{A-B_1R_{11}^{-1}B_1'S+B_1P',B_2\} \quad (4.7)$$

is stabilizable. Let us denote the class of all $(n\times r)$ matrices P that lead to satisfaction of this condition by $\mathscr{P}$. Then the problem is to investigate existence of an element in $\mathscr{P}$ so that the follower's optimal reaction coincides with (4.1) with $i=2$. By going through similar steps as in Section 3, it is now not difficult to visualize the following versions of Condition 1 and Theorem 1, which require no further justification:

Condition 2 There exists an element P of $\mathscr{P}$ such that (i) if λ_j denotes any eigenvalue of F and μ_i denotes any eigenvalue of $F+B_1P'$, $\lambda_j+\mu_i\neq 0$, (ii) the unique solution of the linear matrix equation

$$MF+F'M+PB_1'M-PR_{21}R_{11}^{-1}B_1'S+Q_2 \\ +SB_1R_{11}^{-1}R_{21}R_{11}^{-1}B_1'S \\ +SB_2R_{12}^{-1}R_{22}R_{12}^{-1}B_2'S=0 \quad (4.8)$$

satisfies the constraint

$$R_{22}^{-1}B_2'M=R_{12}^{-1}B_2'S. \quad \square \quad (4.9)$$

Theorem 2 Let Condition 2 be satisfied for a specific matrix P^*. Then, there exists a CLS solution to the infinite-horizon LQ differential game, which is given by

$$\gamma_1^*(t,x(t),x_0)=[P^*-R_{11}^{-1}B_1'S]x(t)-P^{*\prime}\bar{x}(t) \quad (4.10a)$$

$$\gamma_2^*(x(t))=-R_{12}^{-1}B_2'Sx(t). \quad \square \quad (4.10b)$$

A special class of problems of particular interest are those with controls of dimension one, in which case considerable simplifications result with regard to Condition 2. To investigate this somewhat further, let us first note that when the control inputs are scalar, by appropriate rescaling of the cost functions and by redefining B_1, it is always possible to take $R_{11}=R_{12}=R_{22}=1$ (recall that these quantities were required to be positive). Then, the only weighting term for the controls, left as a variable, is R_{21} which we denote, in this case, by d. Under this set-up, Condition 2 becomes equivalent to existence of a vector $P'=(p_1,\ldots,p_n)$ so that the solution of

$$MF+F'M+PB_1'M-dPB_1'S+Q_2 \\ +dSB_1B_1'S+SB_2B_2'S=0 \quad (4.11a)$$

satisfies the relation

$$B_2'[M-S]=0. \quad (4.11b)$$

Defining a new $(n\times n)$ matrix Z as $M-S$, we now note that Z should satisfy the linear matrix equation

$$ZF+F'Z+PB_1'Z+(1-d) \\ \times[P-SB_1]B_1'S+Q_2-Q_1=0 \quad (4.12a)$$

together with the relation

$$B_2'Z=0, \quad (4.12b)$$

and this condition is equivalent to Condition 2(ii) for this special version of the problem. It should be noted that Condition 2(i) still ensures existence of a unique solution to (5.12a). Let us now consider a specific example to illustrate the results of this section.

Example 1 Consider the infinite-horizon differential game of state space dimension 2 and with the following values for the parameters used in the general formulation: $B_1' = (1, 1)$, $B_2' = (2, 1)$, $R_{11} = R_{12} = R_{22} = 1$, $R_{21} = 2$,

$$A = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \quad Q_1 = \begin{bmatrix} 1 & -1 \\ -1 & 5 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

To obtain the CL Stackelberg solution of this differential game with Player 1 as the leader, we now make direct use of Theorem 2 after verifying satisfaction of the stabilizability requirement and of Condition 2. To this end, let us first note that the pair $(A, (B_1, B_2))$ is stabilizable and hence the first requirement is satisfied. The matrix S characterizing the team solution turns out to be

$$S = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix},$$

and by (4.12b), Z should have the form

$$Z = \begin{pmatrix} z_1 & z_2 \\ -2z_1 & -2z_2 \end{pmatrix}.$$

Thus, using all these in (4.12a) with $d = 2$, we obtain the vector equation

$$\underbrace{\begin{pmatrix} -2-p_1 & -1 \\ -1 & -2-p_1 \\ 7-p_2 & 2 \\ 2 & 7-p_2 \end{pmatrix}}_{G} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 \\ -1+p_1 \\ -1 \\ 3+p_2 \end{pmatrix}}_{g}$$

which should admit a consistent solution by a specific choice for the pair $\{p_1, p_2\}$. Condition 2(i) now requires, in the present context, the matrix G to be full of full rank, which is true if, for example, p_1 is not allowed the values -1 and -3. Now, by adding the second column of G on g, we obtain the equivalent condition

$$\operatorname{rank}[G, \tilde{g}] = 2$$

where

$$\tilde{g} = (-1, -3, 1, 10).$$

This condition is satisfied if, and only if, the following two equations admit a solution (p_1, p_2), with $p_1 \neq -1, -3$:

$$-p_1^2 + 9p_1 + (1-p_1)p_2 = 0$$
$$-10p_1^2 - 17p_1 + 3 - (5+3p_1)p_2 = 0.$$

If we now solve for p_2 from the first equation in terms of p_1, and then substitute it into the second one, the result is a cubic polynomial equation in terms of p_1,

$$13p_1^3 - 15p_1^2 - 65p_1 + 3 = 0,$$

which admits three real roots $p_1^* = \pm\sqrt{5}$, $3/13$. The corresponding p_2 values are $p_2^* = -7.26$, -5.26, -0.26, respectively, to the nearest three decimal places. All these three sets of solutions satisfy the required conditions, and thus $p^* = (p_1^*, p_2^*)$, when substituted in (4.10a), provides a CLS strategy (though nonunique) for the leader. Theorem 2, therefore, yields for this differential game problem, the CLS solutions

$$\gamma_1^*(t, x, x_0) = (p_1^*, p_2^* - 1)x(t) - (p_1^*, p^*)\exp(Ft)x_0$$
$$\gamma_2^*(x(t)) = -(1, 0)x(t)$$

where

$$F = \begin{bmatrix} -2 & -1 \\ -1 & -2 \end{bmatrix}$$

and p_1^*, p_2^* are as given above.

5. *A more general representation*

Heretofore, we have restricted the leader's strategies to representations of the functional form (2.14), and have obtained sufficient conditions under which he can enforce his team solution on the follower. It is possible, however, to provide more flexibility for the leader by allowing him a representation of his team strategy more general than (2.14). In particular, if we replace (2.14) with

$$\gamma_1(t, x(s), s \leq t) = -R_{11}^{-1}B_1'Sx(t) + P'(t)[x(t) - \bar{x}(t)] + W'(t)[y(t) - \bar{y}(t)], \qquad (5.1)$$

where $y(\cdot)$ is an n-dimensional vector function satisfying the differential equation

$$\dot{y} = Cy + Dx, \quad y(0) = 0 \qquad (5.2)$$

and $\bar{y}(\cdot)$ denotes the solution of (5.2) with $x(\cdot)$ replaced by the optimal team trajectory $\bar{x}(\cdot)$, then the leader has the freedom of choosing the matrix functions $W(\cdot)$, $C(\cdot)$, $D(\cdot)$, in addition to $P(\cdot)$, which clearly provides him with a greater degree of flexibility.

Now, assuming the leader announces the strategy (5.1) for one appropriate dimensional quadruple $\{P, W, C, D\}$ with continuous entries, the follower's reaction to that can easily be determined, and conditions for a new version of Proposition 1 (with obvious modifications) to be satisfied can be derived by basically following the steps taken in Section 3, assuming of course that we are dealing with a finite-horizon problem. The Hamiltonian that would replace (3.4), in the present context, is

$$\begin{aligned} H = {} & \tfrac{1}{2}x'Q_2x + \tfrac{1}{2}[R_{11}^{-1}B_1'Sx + W'(\bar{y} - y) \\ & + P'(\bar{x} - x)]'R_{21}[R_{11}^{-1}B_1'Sx + W'(\bar{y} - y) \\ & + P'(\bar{x} - x)] + \tfrac{1}{2}u_2'R_{22}u_2 + \lambda_1' \\ & \times [(A - B_1R_{11}^{-1}B_1'S)x + B_1P'(x - \bar{x}) \\ & + B_1W'(y - \bar{y}) + B_2u_2] + \lambda_2'[Cy + Dx], \end{aligned} \qquad (5.3)$$

where $\lambda_1(\cdot)$, $\lambda_2(\cdot)$ denote n-dimensional costate vectors, satisfying the differential equations

$$\begin{aligned} \dot{\lambda}_1 = {} & -\{Q_2 - (P - SB_1R_{11}^{-1})R_{21}R_{11}^{-1}B_1'S\}\bar{x} \\ & -(A - B_1R_{11}^{-1}B_1'S + B_1P')'\lambda_1 \\ & -D'\lambda_2; \quad \lambda_1(T) = Q_{2f}\bar{x}(T) \end{aligned} \qquad (5.4a)$$

$$\dot{\lambda}_2 = +WR_{21}R_{11}^{-1}B_1'S\bar{x} - WB_1'\lambda_1 - C'\lambda_2; \quad \lambda_2(T) = 0 \qquad (5.4b)$$

where we have also made use of the sought property that the optimal trajectory of this optimization problem coincides with the optimal team trajectory, i.e., $x(\cdot) \equiv \bar{x}(\cdot)$, $y(\cdot) \equiv \bar{y}(\cdot)$. Hence, the set of differential equations (5.4) now replace (3.8) within the context of the present more general version. Maximization of H with respect to u_2 yields

$$u_2(t) = -R_{22}^{-1}B_2'\lambda_1(t)$$

which is sought to be equal to $\bar{\gamma}_2(t, x(t))$, i.e.

$$R_{22}^{-1}B_2'\lambda_1(\cdot) \equiv R_{12}^{-1}B_2'S(\cdot)\bar{x}(\cdot). \qquad (5.5)$$

Since the solution set of (5.4) depends linearly on x_0, and $\bar{x}(\cdot)$ is dependent on x_0 through a linear nonsingular transformation, we can easily let λ_1 and λ_2 to be related to $\bar{x}(\cdot)$ through the linear matrix transformations $\lambda_1(t)$

$= \Lambda_1(t)\bar{x}(t)$, $\lambda_2(t) = \Lambda_2(t)\bar{x}(t)$, where $\Lambda_1(\cdot)$ and $\Lambda_2(\cdot)$ satisfy matrix differential equations

$$\dot{\Lambda}_1 + \Lambda_1 F + F'\Lambda_1 + PB_1'\Lambda_1 - PR_{21}R_{11}^{-1}B_1'S + Q_2 + SB_1R_{11}^{-1}R_{21}R_{11}^{-1}B_1'S + SB_2R_{12}^{-1}R_{22}R_{12}^{-1}B_2'S + D'\Lambda_2 = 0; \quad \Lambda_1(T) = Q_{2f} \tag{5.6a}$$

$$\dot{\Lambda}_2 + \Lambda_2 F + C'\Lambda_2 + WB_1'\Lambda_1 - WR_{21}R_{11}^{-1}B_1'S = 0; \quad \Lambda_2(T) = 0. \tag{5.6b}$$

Furthermore, in terms of this new notation, (5.5) becomes

$$R_{22}^{-1}B_2'\Lambda_1(t) \equiv R_{12}^{-1}B_2'S(t), \quad t \in [0, T]. \tag{5.7}$$

Hence, we immediately have the following versions of Condition 1 and Theorem 1, in the present context, which we give without proof.

Condition 3 There exist appropriate dimensional matrix functions $P(\cdot)$, $W(\cdot)$, $C(\cdot)$ and $D(\cdot)$ with continuous entries so that the unique solution of the coupled linear matrix differential equations (5.6) satisfies the relation (5.7). □

Theorem 3 Let Condition 3 be satisfied for matrix functions $P^*(\cdot)$, $W^*(\cdot)$, $C^*(\cdot)$ and $D^*(\cdot)$. Then, there exists a CLS solution for the differential game of Section 2, which is given by

$$\gamma_1^*(t, x(s), s \leq t) = [P^{*\prime} - R_{11}^{-1}B_1'S]x(t) - P^{*\prime}\bar{x}(t) + W^{*\prime}[y(t) - \bar{y}(t)] \tag{5.8a}$$

$$\gamma_2^*(t, x(t)) = -R_{12}^{-1}B_2'Sx(t), \tag{5.8b}$$

where y and $\bar{y}$ respectively satisfy

$$\dot{y} = C^*y + D^*x(t), \quad y(0) = 0 \tag{5.9a}$$

$$\dot{\bar{y}} = C^*\bar{y} + D^*\bar{x}(t), \quad \bar{y}(0) = 0. \tag{5.9b}$$

Furthermore, the CLS costs for the leader and the follower are given, respectively, by

$$J_1^* = \tfrac{1}{2}x_0'S(0)x_0, \tag{5.10a}$$

$$J_2^* = \tfrac{1}{2}x_0'\Lambda_1(0)x_0. \quad \square \tag{5.10b}$$

We should note at this point that the only difference between (3.9) and (5.6a) is the additional term $D'\Lambda_2$ appearing in the latter, which actually provides the leader with more freedom in ensuring satisfaction of relation (5.7).

6. *Conclusion*

We have obtained, in this paper, the closed-loop Stackelberg (CLS) solution of an important class of continuous-time two player nonzero-sum differential games described by linear state equations and quadratic cost functionals, and under certain (not totally restrictive) conditions on the system parameters. Both the finite- and infinite-horizon problems are considered, and in both cases, the derived CLS strategies have the important property that they generate the same state trajectory as that of a related optimal team problem. The objective function of this team problem is precisely the cost function of the leader, so that the derived CLS solution can also be considered as an extreme Pareto-optimal solution to the leader's advantage.

In the derivation of the CLS solutions that inherit the above property, two different classes of representations have been adopted for the leader: (i) linear strategies that depend only on the current and initial values of the state, (ii) linear strategies that depend on the whole past trajectory. For the former class of representations, both the finite and the infinite-horizon problems are solved, with conditions of existence related to solutions of linear matrix equations. For the class of strategies that depend on the whole past trajectory, a set of sufficiency conditions that are related to the solutions of two linear matrix differential equations are obtained, under which the finite-horizon problem admits a CLS solution with the above cited feature. The scalar version of the differential game problem of this paper has been studied thoroughly in Başar and Olsder (1979) and it has been shown, in that context, that one of the following two situations can arise: (a) the differential game admits a CLS solution within the prescribed linear class, or (b) there exist only ε-Stackelberg strategies for the leader, with the corresponding sequence of costs converging to the minimum value of his cost function. Hence, to seek the CLS solution of differential games within the class of representations of related team solutions is a reasonable approach, which is clearly also corroborated by Example 1 of Section 4.

REFERENCES

Başar, T. and G. J. Olsder (1979). Team-optimal closed-loop Stackelberg strategies in linear-quadratic differential games. *Proc. 17th Allerton Conference on Communication, Control and Computing*, Urbana, Illinois.

Başar, T. and H. Selbuz (1979a). A new approach for derivation of closed-loop Stackelberg strategies. *Proc. IEEE Conf. on Decision and Control*, San Diego, January, pp. 1113–1118.

Başar, T. and H. Selbuz (1979b). Closed-loop Stackelberg strategies with applications in the optimal control of multi-level systems. *IEEE Trans. Aut. Control* **AC-24,** 166–179.

Chen, C. I. and J. B. Cruz, Jr. (1972). Stackelberg solution for two-person games with biased information patterns. *IEEE Trans. Aut. Control* **AC-17,** 791–798.

Medanic, J. (1977). Closed-loop Stackelberg strategies in linear-quadratic problems. *Proc. 1977 JACC*, pp. 1324–1329, San Francisco.

Papavassilopoulos, G. P. and J. B. Cruz, Jr. (1979). Sufficient conditions for Stackelberg and Nash strategies with memory. *Proc. Conference on Systems Engineering for Power: Organizational Forms for Large Scale Systems*, Vol. II, U.S. Dept. of Energy, Washington, DC, October, pp. 2.59–2.69.

Simaan, M. and J. B. Cruz, Jr. (1973a). On the Stackelberg strategy in nonzero-sum games, *J. Optimiz. Theory Appl.* **11,** 535–555.

Simaan, M. and J. B. Cruz, Jr. (1973b). Additional aspects of the Stackelberg strategy in nonzero-sum games. *J. Optimiz. Theory Appl.* **11,** 613–620.

von Stackelberg, H. (1934). *Marktform und Gleichgewicht*. Springer, Berlin.