

Further Properties of Nonzero-Sum Differential Games¹

A. W. STARR² AND Y. C. HO³

Abstract. The general nonzero-sum differential game has N players, each controlling a different set of inputs to a single nonlinear dynamic system and each trying to minimize a different performance criterion. These general games have several interesting features which are absent in the two best-known special cases (the optimal control problem and the two-person, zero-sum differential game). This paper considers some of the difficulties which arise in attempting to generalize ideas which are well known in optimal control theory, such as the *principle of optimality* and the relation between *open-loop* and *closed-loop* controls. Two types of solutions are discussed: the *Nash equilibrium* and the *noninferior set*. Some simple multistage discrete games are used to illustrate phenomena which also arise in the continuous formulation.

1. Introduction

In the general N -player, nonzero-sum differential game, the i th player chooses u_i trying to minimize

$$J_i = \int_{t_0}^{t_f} L_i(x, t, u_1, \dots, u_N) dt + K_i(x(t_f)) \quad (1)$$

subject to the n -dimensional state equation, common to all players,

$$\dot{x} = f(x, t, u_1, \dots, u_N), \quad x(t_0) = x_0 \quad (2)$$

and possibly subject to various inequality or equality constraints on the state and/or control variables; these are omitted here for simplicity.

¹ Paper received July 8, 1968. This research was supported by Joint Services Electronics Contracts Nos. N00014-67-A-0298-0006, 0005, 0008 and by NASA Grant No. NGR 22-007-068.

² Graduate Student, Division of Engineering and Applied Physics, Harvard University, Cambridge, Massachusetts.

³ Gordon McKay Professor of Engineering and Applied Mathematics, Division of Engineering and Applied Physics, Harvard University, Cambridge, Massachusetts.

This problem, which includes the optimal control problem ($N = 1$) and the two-person, zero-sum differential game ($N = 2, J_1 = -J_2$) as special cases, is of interest in analyzing a dynamic system with inputs controlled by several *players* with not entirely conflicting goals.

One naturally expects that methods for computing solutions to these problems can be obtained by generalizing well-known methods of optimal control theory. While this is true to some extent, several difficulties arise which are absent in control problems and two-person, zero-sum differential games. In this paper, we consider generalizations of two ideas which are very useful in solving optimal control problems: (a) the relation between *open-loop* and *closed-loop* optimal controls and (b) the *principle of optimality*.

Nonzero-sum differential games were discussed by Starr and Ho (Ref. 1), who concluded that there is no single satisfactory definition of *optimality* for these problems. Depending upon the application, various types of solutions are relevant.

One interesting type of solution is the *Nash equilibrium*. It is *optimal* in the sense that no player can achieve a better result by deviating from his *Nash controls* as long as the other players continue to use *their* Nash controls. If the control strategy and the cost for the i th player are denoted by u_i and J_i , respectively, the Nash equilibrium strategy set $\{u_1^*, \dots, u_N^*\}$ has the property that, for $i = 1, \dots, N$,

$$J_i(u_1^*, \dots, u_N^*) = \min_{u_i} J_i(u_1^*, \dots, u_{i-1}^*, u_i, u_{i+1}^*, \dots, u_N^*)$$

Letting $u^* = \{u_1^*, \dots, u_N^*\}$ and $J = \{J_1, \dots, J_N\}$, we sometimes refer to u^* as a *Nash saddle point* of $J(u)$.

Depending on the formulation of the problem, u_i may be one of a finite set of controls (static bimatrix game), a function of time (open-loop differential game), a function of the state vector and time (closed-loop differential game), and so on.

In the analysis of competitive dynamic systems (for example, several rival firms in an imperfectly competitive market), the restriction that no binding agreements can be made among the players leads naturally to the *secure* Nash solutions. Then, one would like to know what has been sacrificed to obtain this security, i.e., do solutions exist which reduce the costs of *all* players below their Nash costs? This leads us to a second type of interesting solution, the *set of noninferior strategies*. If \hat{u} is noninferior, then there exists no control u such that

$$J_i(u) \leq J_i(\hat{u}), \quad i = 1, \dots, N$$

with strict inequality for at least one i . Any *negotiated* solutions with all players cooperating but no transfer payments allowed should be chosen from this class. In most differential games, there is a single Nash solution but an $(N - 1)$ -parameter family of noninferior, or *undominated*, solutions.

2. Relationship Between Open-Loop and Feedback Nash Solutions

In *optimal control problems*, one often distinguishes between *open-loop* solutions, where the optimal control for a trajectory through a specified initial state x_0 is given as a function of time, and *closed-loop* or *feedback* solutions, which give the optimal control as a function of the state x and time t everywhere in an appropriate region of the state-time space. It is well known that, in deterministic problems,⁴ the open-loop solution $u^0(t)$, $t_0 \leq t \leq t_f$, can be generated from the feedback solution $u^0(x, t)$ simply by integrating the state equation forward from the initial point (x_0, t_0) . This would be a reasonable way to find the open-loop control if an algorithm based on a dynamic-programming approach were available for computing the closed-loop optimal controls in a region containing the given initial point.

Alternately, if a successful open-loop algorithm based on a variational approach is available for calculating $u^0(t)$ for a trajectory through (x_0, t_0) , then the closed-loop control law can, at least in principle, be generated by successively solving the open-loop problem for each initial point (x_0, t_0) .

In what appears to be the most interesting class of differential games, all players know the current state vector, so that a *closed-loop* Nash solution is required.⁵ There may also be interesting *open-loop* problems where the entire sequence of controls for each player must be chosen prior to the initial time.

Whichever type of Nash solution is required, one could in principle solve for the Nash strategies for all the players in advance, since there are no *unpredictable* inputs to the system. Therefore, one is tempted to conclude that the same relation exists between the open-loop and closed-loop strategies as exists in the optimal control problem, i.e., they are just different ways of describing the same outcome. The purpose of this section is to demonstrate that such a conclusion is false. Although our real interest is continuous

⁴ These are problems where all the parameters and all the inputs to the system over the time interval under consideration are known at the initial time.

⁵ More realistically, they might have imperfect, noisy measurements of the state vector; but here we assume exact knowledge of the state vector as well as all the system parameters including the cost functions for the other players.

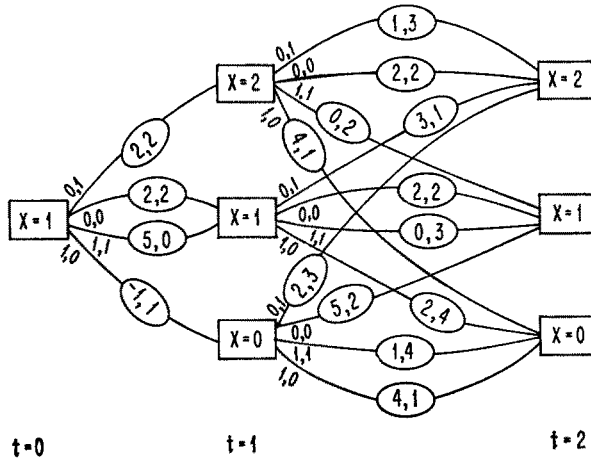


Fig. 1

differential games, we first illustrate the basic idea by considering a very simple discrete, finite-state multistage game.

In the two-player game in Fig. 1, each player has two possible controls, labeled 0 and 1. At each stage t , both players simultaneously choose a control. The resulting control pair determines the transition to the next stage. There are four possible transitions, leading to three possible stages x , and associated with each transition are costs c_1, c_2 (encircled) for the two players. Each player wants to minimize his total cost in reaching $t = 2$, the terminal stage.

Let us try to find the *closed-loop* Nash solution by following the *dynamic programming* approach. At stage $t = 1$ and state $x = 2$, the situation for the two players is represented by the bimatrix game in Fig. 2a. Clearly, the controls 0, 0 are the only pair with the Nash property, since Player 1 would increase his cost from 2 to 4 by playing 1, while Player 2 would increase his cost from 2 to 3 by playing 1. (As far as the Nash equilibrium is concerned, it does not matter what would happen should *both* players play a non-Nash control.) The Nash costs are $c_1 = c_2 = 2$. Similarly, we see from Fig. 2b that the Nash controls at $x = 1, t = 1$ are 1, 1 with costs 0, 3; and, from Fig. 2c, we see that, at $x = 0, t = 1$, the Nash control pair 1, 0 gives costs 4, 1. Moving back to the initial stage $t = 0$, we assume that the players play their Nash controls at $t = 1$; thus we add the Nash costs 2, 2 associated with state 2 to the costs of the transition leading to state 2, and so on. The resulting situation is given in Fig. 2d. The Nash control pair is then 0, 1 with costs 4, 4 for the entire game. The *trajectory* is $x(1) = 2, x(2) = 2$.

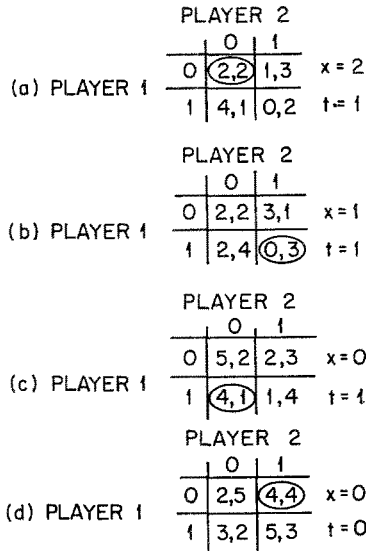


Fig. 2

Then, can we conclude that this trajectory with its associated control sequences 00, 10 is also the open-loop Nash trajectory? In Fig. 3, the costs are tabulated for each pair of open-loop control sequences. Inspection of this bimatrix game shows that only the control sequence pair 11, 00 has the Nash property, giving costs 3, 2. The closed-loop Nash solution 00, 10 does *not* have the Nash property in the open-loop table. The open-loop Nash trajectory is $x(1) = 0, x(2) = 0$.

One reason for this difference between the open-loop and closed-loop solutions is the fact that several control sequences were eliminated from consideration at $t = 0$ by the assumption that the player would choose Nash

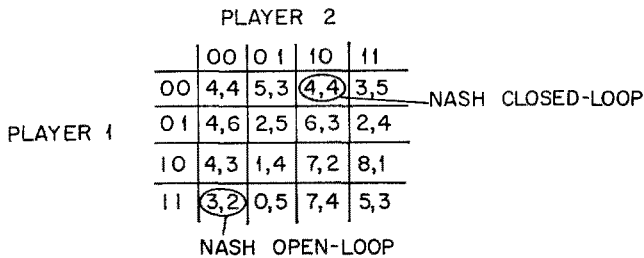


Fig. 3

controls only at $t = 1$, based on knowledge of state at $t = 1$. This assumption that the players always attempt to *optimize* the remaining part of the trajectory based on the current state regardless of previous actions is the natural extension of the basic principle of optimality found in all dynamic-programming type of calculations. Yet, it is not always safe to employ such assumptions in the nonzero-sum case. Another interesting point to note is that the Nash open-loop costs (3, 2) in Fig. 3 are strictly superior to the closed loop costs (4, 4) calculated via *dynamic programming*. This casts further doubt on the applicability of the principle of optimality. We have more to say on this in Section 3.

It should be pointed out that the two-stage game with closed-loop control in Fig. 1 can also be represented as a single bimatrix game, but not the same game as that obtained in Fig. 3 for open-loop controls. Since each player has eight possible feedback strategies, the closed-loop bimatrix game is an 8×8 table. In this array, only the closed loop strategy pair

$$\begin{aligned} u_1(0, 0) &= 0, & u_2(0, 0) &= 1 \\ u_1(1, 1) &= 1, & u_2(1, 1) &= 1 \\ u_1(2, 1) &= 0, & u_2(2, 1) &= 0 \end{aligned}$$

has the Nash property. Obviously, this would be a cumbersome way to find closed-loop Nash strategies, especially with a larger number of states, stages, controls, or players.

2.1. Continuous Differential Games. A general conceptual method for finding the *closed-loop* Nash equilibrium control $u_1^*(x, t), \dots, u_n^*(x, t)$ was presented in Ref. 1. One finds the *remaining cost functions* $V_i(x, t)$, $i = 1, \dots, N$, by solving a set of coupled partial differential equations

$$-\partial V_i / \partial t = \min_{u_i} H_i(x, t, u_1, \dots, u_N, \partial V_i / \partial x), \quad i = 1, \dots, N \quad (3)$$

where the Hamiltonian for the i th player is

$$H_i(x, t, u_1, \dots, u_N, \partial V_i / \partial x) = L_i(x, t, u_1, \dots, u_N) + (\partial V_i / \partial x) f(x, t, u_1, \dots, u_N) \quad (4)$$

On the terminal surface,

$$V_i(x(t_f), t_f) = K_i(x(t_f)) \quad (5)$$

The Nash controls are the controls u_i which achieve the required minima. If the functions L_i and f are continuously differentiable in u_i and if the

minimum is in the interior of the set of admissible controls, then u_i can be found by solving the relation

$$\partial H_i / \partial u_i = 0, \quad i = 1, \dots, N \tag{6}$$

to obtain u_i explicitly as a function of $x, t, \partial V / \partial x$. Then, one must solve the set of partial differential equations for $V_i(x, t)$, from which one finally obtains $u_i^*(x, t)$.

To find the open-loop Nash solutions, one first uses a variational method to derive the necessary conditions. Case (Ref. 2) obtained the following conditions, which hold only if the controls are all open-loop:

$$\dot{x} = f(x, t, u_1, \dots, u_N) \tag{7}$$

$$\dot{\lambda}_i^T = -\partial H_i / \partial x \tag{8}$$

$$\lambda_i^T(t_f) = (\partial / \partial x(t_f)) K_i(x(t_f)) \tag{9}$$

$$H_i(x, t, u_1, \dots, u_N, \lambda_i^T) \equiv \min \text{ w.r.t. } u_i \tag{10}$$

where

$$H_i(x, t, u_1, \dots, u_N, \lambda_i^T) = L_i(x, t, u_1, \dots, u_N) + \lambda_i^T f(x, t, u_1, \dots, u_N) \tag{11}$$

Computational algorithms can be obtained from these necessary conditions. Necessary conditions for the *closed-loop* Nash controls $\Psi_1(x, t), \dots, \Psi_N(x, t)$ were obtained by Starr and Ho (Ref. 1) by replacing (8) with

$$\dot{\lambda}_i^T = -\partial H_i / \partial x - \sum_{j=1, j \neq i}^N (\partial H_i / \partial u_j) (\partial \Psi_j(x, t) / \partial x) \tag{12}$$

The presence of the summation term in (12) makes the necessary conditions (7), (12), (9), (10) virtually useless for deriving computational algorithms. Note that this troublesome term is absent in the optimal control problem (because $N = 1$), in the two-person zero-sum game (because $H_1 = -H_2$, so that $\partial H_1 / \partial u_2 = -\partial H_2 / \partial u_2 = 0$), and in the open-loop, nonzero-sum problem (because $\partial \Psi_j / \partial x = 0$). One certainly expects the open-loop and closed-loop solutions to be different whenever this term is nonzero.

Using reasoning familiar from optimal control theory, one can interpret (12) as follows: λ_i is the *influence function* for the i th player, i.e., the sensitivity of his cost to a perturbation in the state vector. If the other players are using feedback strategies, any perturbation δx of the state vector causes them to change their controls by the amount $(\partial \Psi_j / \partial x) \delta x$. If the i th Hamiltonian were already extremized with respect to the control $u_j, j \neq i$, this would

not affect the i th player's cost; but, since generally $\partial H_i/\partial u_j \neq 0$ for $i \neq j$, the reactions of the other players to the perturbation influence the i th player's cost, and the i th player must account for this effect in considering variations of the trajectory.

In fact, a rather peculiar situation arises when the i th player makes a small change δu_i in his control in the vicinity of the Nash trajectory. Since $\partial H_i/\partial u_i = 0$, the effect of δu_i on the i th player's cost is only of the second order in δu_i , but the effects on all the other players' costs are of the first order because $\partial H_j/\partial u_i \neq 0$ for $i \neq j$. In making fine adjustments to reach his minimum cost, the i th player may cause wild fluctuations, either beneficial or harmful, in his rivals' costs. If they are able to react to this change (i.e., they have closed-loop control), they in turn cause first-order changes in the i th player's cost, so that another second-order term in δu_i , due to the reactions of the rivals, must be added to the *direct* second-order effect of δu_i on the i th cost. Thus, it is easy to see that the equilibrium conditions (and, consequently, the trajectories which satisfy them) are not the same in the open-loop and closed-loop problems. Even for the simplest nonzero-sum differential game, the *linear-quadratic* case, entirely different Nash solutions have been obtained by the authors for the open-loop and closed-loop formulations.

3. The Optimality Principle

The well-known *principle of optimality* has been of great use in providing a conceptual framework for solving optimal control problems. The same principle, which Isaacs called the *tenet of transition*, is the basis of a general method for finding optimal strategies in zero-sum, two-person differential games. Thus, it is naturally interesting to inquire what principle of optimality, if any, holds for a more general, N -person, nonzero-sum differential game. In this section, we discuss the relation between the noninferior solutions, the Nash solution, and the optimality principle.

In a static, nonzero-sum game, we speak of a *prisoners' dilemma* situation⁶ whenever the Nash solution does not belong to the noninferior set. For example, in Fig. 2, the *prisoners' dilemma* occurs in bimatrix games a and d, but not in b or c. It should also be clear what is meant by the statement that the vector Hamiltonian

$$H = \{H_1, \dots, H_N\}$$

⁶ See footnote in the introduction of Ref. 1.

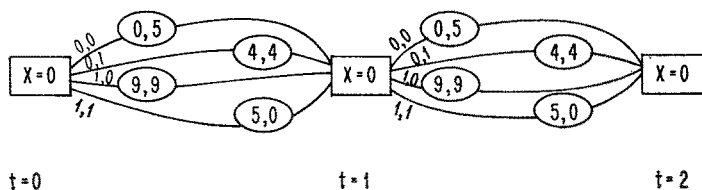


Fig. 4

(with H_i defined in (11)) has a *prisoners' dilemma* for some particular values of $x, t, \lambda_1, \dots, \lambda_N$.

Now, consider a dynamic game, either a differential game or a multistage game, whose closed-loop Nash solution is obtained via the *dynamic-programming* approach used in Section 2. One is tempted to guess that, if no *prisoners' dilemma* occurs at any stage or state during the computation of the Nash equilibrium, then the Nash solution is noninferior. But this conjecture is false, as we show below. Again, we start with a discrete multistage game. The game in Fig. 4 is almost trivial; it is really a single static, bimatrix game played twice. Since there is only one state, there is no difference between *open-loop* and *closed-loop*.⁷ One can see by inspection that the *prisoners' dilemma* does not occur at either stage in the Nash solution. The pair of control sequences 00, 11 gives the Nash solutions. At no stage does the *prisoners dilemma* situation occur; i.e., the Nash solution at each stage is noninferior. Can we conclude from this that the Nash solution is noninferior globally over two stages? In other words, is there no *cooperative* solution by which *both* players can reduce their costs? To answer this question, we tabulate the costs for all possible pairs of control sequences in Fig. 5.

Inspection of Fig. 5 shows that there are *eight* noninferior solutions (marked with an asterisk), but the Nash solution is not among them. By

		PLAYER 2				
		00	01	10	11	
PLAYER 1	00	*0,10	*4,9	*4,9	(8,8) — NASH	
	01	9,14	*5,5	13,13	*9,4	
	10	9,14	13,13	*5,5	*9,4	
	11	18,18	14,9	14,9	*10,0	

Fig. 5

⁷ A more complicated counterexample where *state* is important can also be constructed, but the game in Fig. 4 is adequate for our purposes.

playing either 01 against 01 or 10 against 10, we see that the costs are 5, 5, compared to the Nash costs 8, 8 obtained by playing 00 against 11. But, to obtain the costs 5, 5 by the sequence 01, 01, Player 2 must trust Player 1 not to try to optimize, by playing control 0, at $t = 1$. Similarly, if the costs 5, 5 are to be obtained by the sequences 10, 10, then Player 1 must trust Player 2.

This very simple game illustrates two basic points about nonzero-sum multistage games: (a) the absence of a *prisoners' dilemma* situation at every stage in solving for the Nash controls does not guarantee that the Nash solution is noninferior over all stages; and (b) noninferior solutions generally require trusting the rivals to play nonoptimal controls, not only at the present stage but at all future stages as well.

More basically, the principle of optimality, which is obvious in control problems, also applies to zero-sum differential games, because it is reasonable to base the choice of action at one time on an assumed mode of behavior of the players at later times (they seek a minimum or a saddle point). In nonzero-sum games, since the meaning of *optimality* is nonunique, it is natural but not necessarily realistic to assume that the rivals continuously seek one particular form of solution, in this case the Nash equilibrium. Cooperation should thus be considered not only at any given stage, but over several stages.

The noninferior solutions to the general differential game were also presented in Ref. 1. They could be obtained by solving the $(N - 1)$ -parameter set of scalar optimization problems

$$\min_{u_1, \dots, u_N} \sum_{i=1}^N \mu_i J_i, \quad \sum_{i=1}^N \mu_i = 1, \quad \mu_i > 0 \quad (13)$$

provided that certain convexity conditions are satisfied.⁸

For a given time-invariant weighting vector μ , the associated noninferior trajectory can be found by solving the Hamilton–Jacobi equation

$$-\partial \hat{V}(x, t, \mu) / \partial t = \min_{u_1, \dots, u_N} \hat{H}(x, t, u_1, \dots, u_N, \partial \hat{V} / \partial x, \mu) \quad (14)$$

where

$$\hat{H} = \sum_{i=1}^N \mu_i L_i(x, t, u_1, \dots, u_N) + (\partial \hat{V} / \partial x) f(x, t, u_1, \dots, u_N) \quad (15)$$

⁸ It is sufficient that the set of cost vectors $[J_1, \dots, J_N]$ generated by all the admissible controls is convex. Weaker sufficient conditions, involving *directional convexity*, can be obtained (see Ref. 4). In any case, the solutions to (13) are always noninferior.

and

$$\hat{V}(x(t_f), t_f) = \sum_{i=1}^N \mu_i K_i(x(t_f), t_f) \tag{16}$$

Now, assume that the closed-loop Nash solution has been found by solving Eq. (3). Generally, the Nash solution does not belong to the noninferior set. But suppose that our game has the special property that the controls for the Nash trajectory through any initial point are also the controls for the noninferior trajectory for some time-invariant weighting vector μ^* . Then, the remaining noninferior cost must be related to the remaining Nash costs by

$$\hat{V}(x, t, \mu^*) = \sum_{i=1}^N \mu_i V_i(x, t) \tag{17}$$

If (17) is substituted into (15) with $\mu = \mu^*$, the Nash Hamiltonians H_i are related to the noninferior Hamiltonian \hat{H} by

$$\hat{H}(\mu^*) = \sum_{i=1}^N \mu_i^* H_i \tag{18}$$

Thus, the assumption that the Nash solution is noninferior implies that, at each time t on the trajectory, the set of controls which satisfies the Nash condition for the static vector function $[H_1, \dots, H_N]$ also minimizes some time-invariant, positive-weighted linear combination of H_i , $i = 1, \dots, N$. In other words, as we solve the infinite sequence of *static Nash saddle-point problems* to obtain the Nash trajectory, we never encounter the *static prisoners' dilemma situation*.

This is a necessary condition for the Nash solution to be noninferior. In effect, it says that it is impossible for all players to gain by playing *cooperative* controls in the time interval $[t, t + dt]$ and then reverting to the local Nash controls in the interval $[t + dt, t_f]$. Without the requirement that μ^* be time invariant, it is not sufficient that the static prisoners' dilemma situation never occurs along the Nash trajectory.

Suppose that the Nash solution has already been obtained for a given game. We wish to determine whether or not this solution is noninferior. A simple way to check this is to start at the terminal time and compute the controls which minimize at time t_f the linear combination

$$\sum_{i=1}^N \mu_i H_i$$

for some arbitrary positive weighting μ . By iteration, we then attempt to find a μ^* satisfying

$$\sum_{i=1}^N \mu_i^* = 1, \quad \mu_i^* > 0, \quad i = 1, \dots, N$$

which gives controls coinciding at time t_f with the Nash controls. Three results are possible: (a) no such μ^* exists, in which case the Nash solution is *not* noninferior; (b) a unique μ^* is obtained; or (c) μ^* is not uniquely determined, in which case more conditions are obtained by repeating this procedure at earlier times.

If a unique μ^* is found, one can solve the optimal control problem with the scalar cost criterion

$$\hat{J} = \sum_{i=1}^N \mu_i^* J_i$$

starting at the terminal point of the Nash trajectory. The resulting noninferior trajectory, holding μ^* constant, coincides with the Nash trajectory if, and only if, the latter is noninferior.

4. Conclusions

The previous two sections have illustrated some of the interesting phenomena which arise when the optimal control problem (alternately, the *strictly competitive* zero-sum differential game) is generalized by allowing several controllers with different cost criteria. If one seeks a Nash equilibrium trajectory, one must specify whether or not the controllers have instantaneous access to the state vector, since the *open-loop* and *closed-loop* formulations lead to entirely different solutions. If one wonders whether a different solution exists which produces a better result for all *players* than the *secure* closed-loop Nash set of control strategies, it is not sufficient to examine the set of Hamiltonians at each point on the Nash trajectory. This *vector Hamiltonian* contains the information necessary for computing the closed-loop Nash controls at time t , provided the problem has already been solved for the remaining time interval, but it does not contain information about noninferior solutions, open-loop Nash solutions, or any other solutions which may be of interest.

Also central to the discussion in Sections 2 and 3 was the fact that, on a Nash trajectory, each player's cost is minimized with respect to his

own control but not with respect to the other players' controls. Generally, there is no set of controls which simultaneously minimizes *all* the players' costs. Should such a set of controls exist, the problem would degenerate into N uncoupled optimal control problems, with *each* player controlling all the N controls. All players would arrive at the same set of N optimal controls, and the Nash solution would thus be noninferior for *every* positive weighting vector μ .

Because his cost is not minimized with respect to the j th player's control (that is, $\partial H_i / \partial u_j \neq 0$) the i th player is very sensitive to changes in his rivals' controls. This fact is the cause of considerable difficulty in developing algorithms for computing Nash controls for nonlinear problems.

References

1. STARR, A. W., and HO, Y. C., *Nonzero-Sum Differential Games*, Journal of Optimization Theory and Applications, Vol. 3, No. 3, 1969.
2. CASE, J. H., *Equilibrium Points of N-Person Differential Games*, University of Michigan, Department of Industrial Engineering, TR No. 1967-1, 1967.
3. ISAACS, R., *Differential Games*, John Wiley and Sons, New York, 1965.
4. DACUNHA, N. O., and POLAK, E., *Constrained Minimization Under Vector-Valued Criteria in Finite-Dimensional Spaces*, University of California at Berkeley, Electronics Research Laboratory, Memorandum No. ERL-M188, 1966.