J. Indian Inst. Sci., Sept.-Oct. 1987, 67, pp. 361–368
© Indian Institute of Science.

On suboptimal feedback Nash controls for nonzero-sum differential games*

P. K. BHARATHAN†
Department of Electrical and Computer Engineering, University of West Indies, St. Augustine, Trinidad.

Abstract

Nonzero-sum differential games are primarily concerned with control situations evolving over time and are essentially competitive in nature characterized by differential equations, where there are N players, each one extremizing his individual payoff function, the dynamic situation being common to all. The Nash equilibrium strategy offers the most secure solution, because, any player is bound to lose by unilaterally deviating from his Nash control, assuming that other players hold fast to their respective Nash controls.

Key words: Optimal control, differential games.

1. Introduction

Pursuit-evasion problems usually encountered in military applications fall in the category of two-person zero-sum differential game, since one player gains only at the expense of the other, their interests being exactly conflicting. This situation may be considered as a particular case of a more general differential game formulation, where there are N players, each one minimising his own individual cost, the dynamic system being common to all. Clearly, there are as many pay-off functions as there are players and the cost associated with any one player, may, in general, be influenced by not only his individual control and the common state vector but also by the controls of other players as well. Moreover, the sum of all the players' costs is neither zero, nor a constant. The control situation is otherwise known as a nonzero-sum differential game[1,2].

A general, nonzero-sum, N-player differential game may be described as follows: Given a dynamic system characterized by a vector differential equation,

$$\dot{x} = f(x, u_1, \ldots, u_N, t); x(t_0) = x_0. \tag{1.1}$$

*First presented at the Platinum Jubilee Conference on Systems and Signal Processing held at the Indian Institute of Science, Bangalore, India, during December 11–13, 1986.

†Present Address: Department of Electrical Engineering, Regional Engineering College, Calicut 673 601, Kerala State.

For $i = 1,\ldots,N$, player i wishes to choose his control u_i to minimise the cost

$$J_i = K_i(x(t_f)) + \int_{t_0}^{t_f} L_i(x, u_1, \ldots, u_N, t)\,\mathrm{d}t \tag{1.2}$$

where x is the n-vector state. Here it is also assumed that the initial state x_0 and the terminal time t_f are fixed and known, $x(t_f)$ is free and there are no constraints on the control vectors $u_1,\ldots u_N$.

In order to solve the above control situation, it is necessary for one to demand that the solution has some attribute such as minimax, Nash equilibrium, stability against coalition, Pareto optimality and so on. Also, one must specify what information is available to each player during the course of the game. Here, only the Nash equilibrium strategy is considered, assuming that each player knows the current value of the state vector as well as all system parameters and cost functions, but he does not know the strategies of the other players. If $J_1(u_1,\ldots u_N),\ldots,J_N(u_1\ldots,u_N)$ are cost functions for players $1,\ldots,N$, then the strategy set $(u_1^*,\ldots,u_N^*)$ is a Nash equilibrium strategy set if, for $i = 1,\ldots,N$,

$$J_i(u_1^*,\ldots,u_N^*) \leq J_i(u_1^*,\ldots,u_{i-1}^*,\ u_i,\ u_{i+1}^*,\ldots,\ u_N^*). \tag{1.3}$$

Thus, Nash equilibrium strategy offers the most secure solution since no player can achieve a better result by unilaterally deviating from his Nash equilibrium control provided other players are assumed to hold fast to their respective Nash controls.

It is easily seen[3] that the optimal control problem and the two-person zero-sum differential game are only particular cases of the general problem posed above. When there is only one player (N = 1), it is the optimal control problem. The number of players being limited to two (N = 2), the problem becomes a two-person zero-sum differential game, if the players have entirely conflicting interests ($J_1 = -J_2$). Also, in this case, the familiar saddle point solution is none other than the Nash equilibrium strategy set.

In this paper, the specific optimal control approach using time-invariant-feedback gains and integral-state feedback[4–6], is extended to study the Nash equilibrium strategies for nonzero-sum differential games. The open-loop, closed-loop and suboptimal strategies are considered for the general case. A scalar example[7] is worked out for numerical results.

2. Open-loop and closed-loop Nash control

The open-loop and closed-loop Nash controls are, in general, different for a nonzero-sum differential game even under deterministic conditions[2]. This is clearly against the common notion in respect of optimal control problems and zero-sum differential games that the two solutions are just different ways of describing the same outcome. This is not true for nonzero-sum differential games, the actual solution itself depending on which type of solution is sought at the outset. Whichever type of Nash solution is required, one could, in principle, solve for the Nash strategies in respect of all the players in advance if there are no unpredictable inputs to the system. The principle of optimality holds, and the open-loop and closed-loop strategies for any one player are the same, provided the nature of strategies of the other players are fixed and, in particular whether they are going to be open-loop or closed loop.

The necessary conditions in respect of open-loop, closed-loop and suboptimal feedback solutions are given below.

2.1 *Open-loop solution*

Consider the dynamic system (1.1) and the cost (1.2). If the ith player is to choose his open-loop control $u_i(t)$ satisfying the Nash equilibrium condition (1.3), a variational approach may be used to obtain the following necessary conditions, which hold only if all the controls are essentially open loop.
For $i = 1, \ldots, N$,

$$\dot{x} = f(x, u_1, \ldots, u_N, t); \; x(t_0) = x_0 \tag{2.1}$$

$$\dot{\lambda}_i = -\frac{\partial H_i}{\partial x} \; ; \lambda_i(t_f) = \frac{\partial K_i(x(t_f))}{\partial x(t_f)} \tag{2.2}$$

$$\frac{\partial H_i}{\partial u_i} = 0 \tag{2.3}$$

where the Hamiltonian H_i is given as

$$H_i(x, u_1, \ldots, u_N, t, \lambda_i) = L_i(x, u_1, \ldots, u_N, t) + \lambda_i' f(x, u_1, \ldots, u_N, t). \tag{2.4}$$

One may note that the ith player actually solves an optimal-control problem. Hence, the principle of optimality holds and open-loop and closed-loop solutions are the same. But, this is made possible only because, at the outset it is known that other players are going to use only open-loop Nash controls.

2.2 *Closed-loop solution*

If a closed-loop Nash control $u_i(x, t)$ is sought at the outset for the player i, when the other players are assumed to use their respective closed-loop Nash controls, it is known that equation (2.2) is replaced by:

$$\dot{\lambda}_i = \frac{-\partial H_i}{\partial x} - \sum_{\substack{j=1 \\ j \neq i}}^{N} \left(\frac{\partial u_j}{\partial x}\right) \frac{\partial H_i}{\partial u_j} \; ; \lambda_i(t_f) = \frac{\partial K_i(x(t_f))}{\partial x(t_f)} \; . \tag{2.5}$$

The second term in (2.5) makes all the difference which is crucial. One may note that this 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 the open-loop case for nonzero-sum game (because $\partial u_j/\partial x = 0$, since u_j is a function of time only). Thus the term is responsible for making the open-loop and closed-loop Nash controls different, in general for the ith player, provided the controls of the other players are not specified at the beginning itself.

2.3 *Suboptimal feedback solution*

Consider again the dynamic system (1.1) and the cost (1.2). Let the suboptimal feedback control for the ith player be specified in the beginning itself as

$$u_i(t) = g_i(x, \int_{t_0}^{t} x\, \mathrm{d}t, \rho_i, t) \tag{2.6}$$

where only the constant feedback gain vector ρ_i is to be chosen in an optimal way. The Nash equilibrium condition (1.3) may be rewritten as, for $i = 1,\ldots,N$,

$$J_i(\rho_1^*,\ldots,\rho_N^*) \leq J_i(\rho_1^*,\ldots,\rho_{i-1}^*,\rho_i,\rho_{i+1}^*,\ldots,\rho_N^*). \tag{2.7}$$

The ith player solves his optimal control problem resulting in a feedback solution. Following the basic results[5], the necessary conditions may be easily given as, for $i = 1,\ldots N$,

$$\dot{x} = f[x, g_1(x,y,\rho_1,t),\ldots,g_N(x,y,\rho_N,t),t];\ x(t_0) = x_0\ ; \tag{2.8}$$

$$\dot{y} = x;\ y(t_0) = 0; \tag{2.9}$$

$$\dot{\lambda}_i = \frac{-\partial H_i}{\partial x}\ ;\ \lambda_i(t_f) = \frac{\partial K_i(x(t_f))}{\partial x(t_f)}; \tag{2.10}$$

$$\dot{\xi}_i = \frac{-\partial H_i}{\partial y};\ \xi_i(t_f) = 0\ ; \tag{2.11}$$

$$\int_{t_0}^{t_f} \frac{\partial H_i}{\partial \rho_i} \mathrm{d}t = 0; \tag{2.12}$$

where the Hamiltonian H_i is given by:

$$\begin{aligned} H_i(x,y,\lambda_i,\rho_i,\xi_i,t) = {} & L_i[x,g_1(x,y,\rho_1,t),\ldots,g_N(x,y,\rho_N,t),t] \\ & + \lambda_i' f[x,g_1(x,y,\rho_1,t),\ldots,g_N(x,y,\rho_N,t),t] + \xi_i' x. \end{aligned}$$

Thus, for a nonzero-sum differential game, when the players are assumed to use suboptimal strategies (2.6), the open-loop and closed-loop solutions are the same, the suboptimal solution being essentially a feedback solution.

The general results of Section 2 may be particularised for the simple and common case of the two-player, nonzero-sum, linear quadratic differential game. The extension, being obvious and straightforward, is not repeated here.

3. Numerical example

A scalar two-player, linear quadratic, nonzero-sum differential game is worked to study the open-loop, closed-loop and the suboptimal strategies. The example is taken from Simaan and Cruz[7] where the open-loop Nash solution has been obtained as a function of time explicitly.

Consider the scalar example of a linear time-invariant dynamic system

$$\dot{x} = x + u_1 - u_2;\ x(0) = 1,\ t \ \varepsilon\ (0, 1)$$

and the quadratic costs for the two players

$$J_1 = x^2(1) + \frac{1}{2}\int_0^1 \left(\frac{4}{3}x^2 + \frac{1}{3}u_1^2\right) dt;$$

$$J_2 = \frac{1}{2}x^2(1) + \frac{1}{2}\int_0^1 \left(4x^2 + u_2^2\right) dt.$$

The following truly optimal and suboptimal Nash equilibrium strategies are considered for illustration.

(i) Open-loop strategies
(ii) Closed-loop strategies
(iii) Proportional suboptimal strategies

$$u_1 = a_1 x, u_2 = a_2 x$$

(iv) Proportional plus integral suboptimal strategies

$$u_1 = a_1 x + b_1 \int_0^t x\, \mathrm{d}t;$$

$$u_2 = a_2 x + b_2 \int_0^t x\, \mathrm{d}t.$$

In this context, it may be mentioned that the above specific suboptimal control structures only have been considered since they are known[4-6] to be better linear control policies both from the point of view of performance degradation and feasibility of implementation.

Following the developments of section 2.1 through 2.3, necessary conditions are obtained in respect of all the four pairs of strategies given above and these equations are solved using a digital computer. The numerical results are given in Table I. For all the four cases, u_1, u_2 and x are shown in figs 1–3 respectively.

It is obvious that case (i), being an open-loop situation, is distinctly different from the rest of the feedback control situations. The addition of integral feedback to the proportional control generally brings the solutions closer to the truly optimal closed-loop solution. The off-line computation of the time-invariant suboptimal feedback control parameters and the ease of implementation of the suboptimal strategies on-line are, of course, added advantages. Cases (ii)–(iv), being all feedback solutions, the corresponding minimum costs in respect of players u_1 and u_2 may be compared in Table I for the three cases. One may note that the minimum costs are less for player u_1 for suboptimal control compared to the unconstrained case. Thus, it appears that a suboptimal feedback scheme is advantageous to player u_1 compared to the optimal case provided the other player also is adopting a similar suboptimal scheme. Naturally, this can lead to persuasion to his advantage on the part of u_1 to have constrained controls for both players instead of a free choice of controls. The condition under which a suboptimal

Table I
Computer results for the numerical example

Case	Control	Optimal parameters	Minimum costs
(i)	Open-loop	..	$J_1 = 0.2634$ $J_2 = 0.5646$
(ii)	Closed-loop	..	$J_1 = 0.3373$ $J_2 = 0.5638$
(iii)	Proportional		
	$u_1 = a_1 x$ $u_2 = a_2 x$	$a_1 = -2.0293$ $a_2 = 1.2087$	$J_1 = 0.3102$ $J_2 = 0.6087$
(iv)	Proportional plus integral	$a_1 = -1.8308$ $b_1 = -0.5353$	
	$u_1 = a_1 x + b_1 \int_0^t x \, dt$	$a_2 = 1.2065$	$J_1 = 0.3147$
	$u_2 = a_2 x + b_2 \int_0^t x \, dt$	$b_2 = -0.1210$	$J_2 = 0.6132$

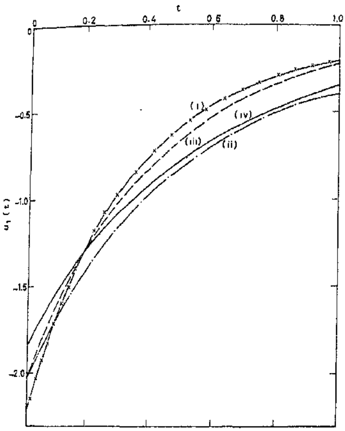

FIG. 1. Controls $u_1(t)$.

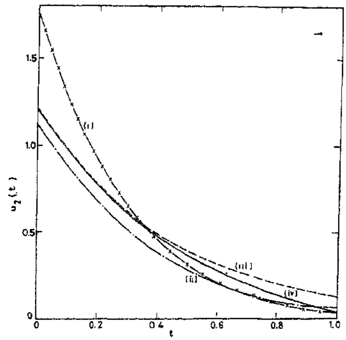

FIG. 2. Controls $u_2(t)$.

policy can be advantageous compared to the unconstrained one for any one player may be obtained at least for the two-player linear quadratic case in terms of the sign definiteness of an 'advantageous matrix'. However, this is not discussed here.

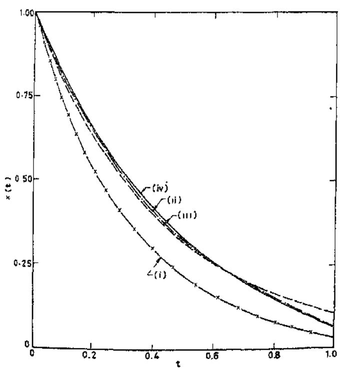

FIG. 3. State trajectories.

4. Conclusion

Open-loop and closed-loop and suboptimal feedback Nash equilibrium strategies ar examined for nonzero-sum differential games. Open-loop and closed-loop strategies ar quite different and give rise to different response trajectories.

The introduction of integral state feedback generally brings the suboptimal feedbac solution closer to the unconstrained case. Also, it is possible that suboptimal constraine strategies may be advantageous to certain players compared to pure strategies. Othe aspects such as mixed strategies[6], cooperation, etc., are not examined in the presen study.

References

1. Starr, A. W. and Ho, Y. C — Non zero-sum differential games, *J. Optimisation Theory Applic.*, 1969, 3 184–206.
2. Starr, A. W. — Further properties of non zero-sum differential games, *J. Optimisatio Theory Applic.*, 1969, **4**, 207–219.
3. Ho, Y. C. — Differential games, dynamic optimisation and generalised control theory, *J Optimisation Theory Applic.*, 1970, **9**, 179–209.
4. Bharathan, P. K., Natarajan, T. and Chandran, S. R. — Class of suboptimal feedback policies for linear and nonlinear systems *Proc. IEE*, 1975, **122**, 444–448.
5. Bharathan, P. K., Natarajan, T. and Seshadri, V. — On proportional plus proportional-integral suboptimal control, *IEEI Trans.*, 1975, **AC–20**, 561–563.
6. Bharathan, P. K. — Suboptimal strategies for zero-sum differential games, *Proc. Natn. Systems Conf.*, 1977, P.S.G. College of Technology, Coimbatore, India.
7. Simaan, M. and Cruz, J. B. Jr. — On the solution of the open loop Nash Riccati equations in linear quadratic differential games, *Int. J. Control*, 1973, **18**, 57–63.