KYBERNETIKA — VOLUME 16 (1980), NUMBER 1

Parameter Optimization in Nonzero-Sum Differential Games

JAROSLAV DOLEŽAL

It is assumed that in addition to the classical formulation of many-player nonzero-sum differential game one of the players, say the first, can choose the values of certain parameters to further decrease his pay-off functional with respect to its equilibrium (Nash) value. Such action causes clearly also changes in pay-off functionals of the remaining players. In contrast to the two-player zero-sum case these changes are generally both positive and negative depending on a particular pay-off functional. To solve this problem in the best possible way for the first player, a set of necessary optimality conditions is presented, which not only determine the Nash equilibrium strategies for all participating players, but also the optimal parameters. Based on these conditions an iterative numerical algorithm of gradient type is suggested. Several illustrative examples are included which were solved applying this algorithm.

1. INTRODUCTION

To begin let us summarize some facts concerning the area of parameter optimization. First of all, as it is known from control theory, the optimal behaviour of any dynamical system is usually a function of various system parameters. These parameters can be sometimes used to further improve the performance of the system in question. In optimal control theory such a class of problems is denoted as parameter optimization or optimal parameter estimation or optimal setting of constants. Here we shall show that this procedure is applicable to the case of many-players nonzero-sum differential games. The further described results can be thus regarded as the extensions of existing results in optimal control theory and two-player zero-sum differential games.

At the present time there exist several papers which contain fairly deep results dealing with parameter optimization for control systems. The fundamental results in this field are due Hofer and Sagirow [1] and Boltjanskij [2] Later Ahmed and Georganas [3] showed that the results of Boltjanskij [2] follow from the general maximum principle of Gamkrelidze [4]. Recently Georganas [5] presented imbedding techniques for optimal parameter estimation, provided that the the optimal parameter could be expressed analytically. Also Lunderstädt [6] described necessary optimality conditions for parameter optimization using the maximum principle approach.

Doležal and Černý [7] used the calculus of variations to obtain the necessary optimality conditions and proposed a first-order gradient algorithm based on the so-called influence functions for the iterative solution of parameter optimization problems. For optimal control problems this algorithm was described by Bryson and Ho [8]. Moreover, the gradient algorithm of Doležal and Černý [7] enables also treatment of control and parameter constrained problems using the projection technique, e.g., see Vasiljev [9]. An alternative gradient-restoration approach can be found in a survey paper of Miele [10].

The results of Doležal and Černý [7] were extended recently by the author [11, 12] also to the case of general two-player zero-sum differential games with nonlinear dynamics and pay-off functional and with possible control and/or final-state constraints. It was assumed that the minimizing player has the opportunity to choose the values of certain parameters before the game starts. The question was, what values of these parameters should he choose to further increase his own gain (decrease his pay-off functional) at the maximizing player's expense? In turn, the maximizing player had to solve the "worst-case" analysis problem, i.e., to determine the greatest lower bound for his expected pay-off. Otherwise speaking, saddle-point of the game in question was function of certain parameters and the minimizing player could choose the most favourable saddle-point due to his aims.

For the case of linear quadratic differential games this two-player zero-sum problem was studied by Leondes and Siu [13]. They suggested several numerical methods for iterative parameter optimization. However, their approach was limited only to system parameters (elements of the transition matrix). Moreover, in the linear quadratic differential game it was possible to use the a priori knowledge of the optimal strategies for both players.

The aim of the present contribution is to show that a similar approach is possible in many-players nonzero-sum differential games. It is assumed that one of the participating players, e.g. the first, has the above mentioned opportunity of choosing the values of certain parameters before the game starts. His aim is clearly to further decreases his pay-off functional with respect to its value in Nash equilibrium. At this place it is also assumed that the Nash equilibrium is the only attractive solution for a game in question, i.e., no coalitions, bargaining, threats, etc. are allowed. Anyhow, one must be aware of one rather important exception with respect to the two-player zero-sum case. Namely it is a priori not clear in which directions the changes of pay-off functionals of the remaining players will occur. Evidently, these changes depend on a particular form of each pay-off functional or more exactly, on the fact to what extent aims of the respective player coincide with those of the first player.

In our approach both optimal equilibrium strategies (open-loop) and optimal values of parameters are determined during the course of calculations. The developed first-order gradient algorithm can also handle problems with control and parameter constraints, and final-state constraints. This algorithm, not including parameter optimization, was used earlier in [14, 15] to obtain a numerical solution of zero-sum and nonzero-sum differential games.

In the following sections first the formulation of the problem is presented and necessary optimality conditions are discussed. Then the gradient algorithm mentioned above is described in detail. To simplify the notation only two-player nonzero-sum case is treated explicitely. Finally, the solution of concrete differential games with parameters is included which confirms the practical importance of the suggested algorithm.

2. PROBLEM FORMULATION

In this section a precise formulation of the studied problem is given. It is assumed that all vectors are the column vectors except of gradients of various functions, which are always treated as row-vectors. All further defined functions are supposed to be continuously differentiable. As E^n will be denoted the n-dimensional Euclidean space. For the sake of simplicity only the problems with fixed final time will be studied, i.e., without any loss of generality we may assume that the independent variable (time) $t \in [0, 1]$. The generalization of the next results to the case of free final time can be principally done according to the scheme indicated in [16, 17].

For the sake of notational simplicity let us consider explicitely only two-player nonzero-sum differential games. Clearly such restriction is a formal one with no loss of generality and second player represents all remaining participants. Let

$$\dot{x} = f(x, u, v, a, t), \quad x(0) = x_0, \quad t \in [0, 1], \tag{1}$$

where $x(t) \in E^n$ denotes the state and $u(t) \in E^m$ and $v(t) \in E^q$, the control variables of participating players at the time t, $a \in E^r$ the parameter and $f : E^n \times E^m \times E^q \times \times E^r \times E^1 \to E^n$. The aim of the first player (denoted I) is to choose a strategy $u(t)$ and a parameter a to minimize the cost functional (pay-off)

$$J_1(u, v, a) = [\varphi^1(x, a)]_1 + \int_0^1 L^1(x, u, v, a, t)\, \mathrm{d}t, \tag{2}$$

while the aim of the second player (denoted II) is to minimize

$$J_2(u, v, a) = [\varphi^2(x, a)]_1 + \int_0^1 L^2(x, u, v, a, t)\, \mathrm{d}t \tag{3}$$

using a strategy $v(t)$. Here $\varphi^i : E^n \times E^r \to E^{\cdot}$ and $L^i : E^n \times E^m \times E^q \times E^r \times E^1 \to \to E^1$, $i = 1, 2$. The lower indices 0 and 1 denote the evaluation of the corresponding expressions at $t = 0$ and $t = 1$.

Finally, the choices of both players must satisfy the control and parameter constraints

$$u(t) \in U \subset E^m, \quad v(t) \in V \subset E^q, \quad t \in [0, 1], \qquad a \in A \subset E^r \tag{4}$$

and the terminal (final-state) constraints

$$[\psi(x, a)]_1 = 0, \tag{5}$$

where $\psi : E^n \times E^r \to E^s$.

Let $a \in A$. We shall say that the strategy pair (u, v) is admissible if it satisfies the control constraints in (4) and the resulting trajectory according to (1) satisfies (5).

As the solution of differential game (1)–(5) let us consider the well-known Nash equilibrium, i.e., such admissible strategy pair (u^*, v^*) for which $a \in A$, and

$$\begin{aligned} J_1^*(a) &= J_1(u^*, v^*, a) \leqq J_1(u, v^*, a), \\ J_2^*(a) &= J_2(u^*, v^*, a) \leqq J_2(u^*, v, a), \end{aligned} \tag{6}$$

with (u, v^*) and (u^*, v) being any admissible strategy pairs. Usually one takes $u(t)$ and $v(t)$ as piecewise continuous functions of t.

Values $J_1^*(a)$ and $J_2^*(a)$ denote the equilibrium outcomes for respective players depending, however, on a particular choice of $a \in A$ by Player I. Assume the existence of the Nash equilibrium for each $a \in A$. Player I then clearly chooses a^* such that (if it exists)

$$J_1^* = J_1^*(a^*) = J_1(u^*, v^*, a^*) \leqq J_1(u^*, v^*, a), \quad a \in A, \tag{7}$$

i.e., $a^* = \arg\min \{J_1^*(a) \mid a \in A\}$. On the other hand, no such condition can be written for $J_2^*(a^*)$ as discussed above.

As a^* is principally known to both players when the game starts, the solution of this parametrized differential game is given by (7) provided that both players act according to (6).

3. NECESSARY OPTIMALITY CONDITIONS

It is obvious and well-known that to find an optimal parameter a^*, the augmented-state approach can be applied. Namely, the r elements of a are considered as additional state variables. For a moment let us neglect the constraints (4).

Applying now the calculus of variations it is not very difficult to show that if $u^*(t)$, $v^*(t)$, a^* form an optimal solution of the differential game with parameters given by (1)–(3), and (5), then there exist $\lambda_1(t)$, $\lambda_2(t) \in E^n$, $t \in [0, 1]$, and $v \in E^s$ such that (symbol T denotes transposition and subscripts stand for the corresponding partial derivates)

(8)
$$\dot{\lambda}_1 = -f_x^T \lambda_1 - (L_x^1)^T \,, \quad t \in [0, 1] \,,$$
$$\dot{\lambda}_2 = -f_x^T \lambda_2 - (L_x^2)^T \,, \quad t \in [0, 1] \,,$$

(9)
$$[\lambda_1 - (\varphi_x^1)^T - \psi_x^T v]_1 = 0 \,,$$
$$[\lambda_2 - (\varphi_x^2)^T - \psi_x^T v]_1 = 0 \,,$$

(10)
$$f_u^T \lambda_1 + (L_u^1)^T = 0 \,, \quad t \in [0, 1] \,,$$

(11)
$$f_v^T \lambda_2 + (L_v^2)^T = 0 \,, \quad t \in [0, 1] \,,$$

(12)
$$\int_0^1 [f_a^T \lambda_1 + (L_a^1)^T] \, dt + [(\varphi_a^1)^T + \psi_a^T v]_1 = 0 \,.$$

In these relations all functions are to be evaluated along the optimal solution. Equations (8) and (9) together with (5) define the unknown multipliers $\lambda_1(t)$, $\lambda_2(t)$ and v, while (10) and (11) give the so-called equilibrium conditions. Finally, equation (12) determines the optimal parameter a^*. Combining (1) and (5) with (8)–(12) one can easily see that, in principle, a nonlinear two-point boundary-value problem for the system of $3n$ differential equations has to be solved. In N-player case one has clearly $(N + 1)\, n$ differential equations. However, such problems cannot be generally solved in analytical way and thus iterative numerical methods must be applied.

If the constraints (4) are present, then equations (10)–(12) have the following form

(13)
$$H_1(x^*, u^*, v^*, \lambda_1, a, t) \leqq \min_{u \in U} H_1(x^*, u, v^*, \lambda_1, a, t) \,,$$
$$H_2(x^*, u^*, v^*, \lambda_2, a, t) \leqq \min_{v \in V} H_2(x^*, u^*, v, \lambda_2, a, t) \,,$$

where x^* corresponds to (u^*, v^*), $t \in [0, 1]$, $a \in A$, and where

(14)
$$H_i(x, u, v, \lambda_i, a, t) =$$
$$= L^i(x, u, v, a, t) + \lambda_i^T f(x, u, v, a, t) \,, \quad t \in [0, 1] \,, \quad i = 1, 2 \,,$$

and

(15)
$$\left\{ \int_0^1 [f_a^T \lambda_1 + (L_a^1)^T] \, dt + [(\varphi_a^1)^T + \psi_a^T v]_1 \right\}^T \delta\bar{a} \geqq 0 \,,$$

where $\delta\bar{a}$ is any feasible parameter change, i.e., $a^* + \delta\bar{a} \in A$. For further details in this respect see Ahmed and Georganas [3].

When formulating optimality conditions (8)–(12), the terminal constraints (5) were taken into the account by a single multiplier $\nu \in E^s$, being therefore common for all (in our case two) players. This fact is by far not so obvious, namely, in several pioneering works dealing with a many-player case subject to terminal constraints (5) the stated necessary optimality conditions contained the number of multipliers ν equal to that of participating players, e.g., see the paper of Sarma et all [18].

Tendency to "equip" each player with his own multiplier can be to a certain extent explained by the histoirical background going to the calculus of variations and optimal control theory. In fact, the well-known saddle-point solution of two-player zero-sum differential games is nothing else then "two-sided" maximum principle as explained by Berkovitz [19], and analogously also Nash equilibrium in the many-player case as given by Case [20]. Such interpretation of a solution differential games leads naturally also to the consideration of two-sided, resp. many-sided, optimal control problems instead of the original differential games.

Influenced by such a reasoning one "introduces" various multipliers not only with respect to the particular aims of a respective player (multipliers λ_i), but also with respect to the common terminal constraints (multipliers ν_i). However, such approach does not take into the account the "parallel" character of players' action during the course a game. Only in the case when one of the players has to play against the known strategies of remaining players the various ν_i can be allowed. This is then a basis for various gradient-type algorithms developed for iterative solution of differential game problems. An algorithm of this type is described in the next section including also parameter optimization for the first player. This discussion and explanation pertains also to the previous author's works [11, 12, 14–17] dealing with numerical solution of differential games. It is not difficult to show that in zero-sum case the multipliers ν_I and ν_{II} differ only in sign – see [14] for necessary details. In the nonzero-sum case the situation is not so obvious – see [15].

In principle, it is also possible to formulate a gradient-type algorithm having a common multiplier ν for all players. Formally one only needs to "add" the pertinent changes in terminal constraints, see (27)–(28) stated further, assuming a sole multiplier ν. Because the computed multipliers ν_i for each player are not usually the same [12, 14, 15], one has to expect certain changes in solution when taking only one ν into the account. This circumstance was also confirmed by numerical calculations and the comparison will be published elsewhere.

On the other hand, terminal constraints (5) pertain to the differential game as a whole and must be therefore adjoint by a sole multiplier ν. Only in this case also the resulting two-point boundary-value problem is meaningful, i.e., the unknown multiplier ν can be computed, at least in principle, invoking the constraints (5). Otherwise such additional unjustified multipliers will result in redundant constants which cannot be excluded or determined, and the use of indirect numerical methods, such as quasilinearization, would be prevented.

It is also worth mentioning at this place that numerical results obtained by quasilinearization and gradient algorithm (with sole multiplier v) are in a fairly good agreement and may be the use of only one multiplier v in gradient-type algorithms for the numerical solution of differential games will be more appropriate in the future.

Let us illustrate these ideas by a simple example of two-player differential game originally studied in [21]. Its various modifications were considered also by the author in the above mentioned references. Clearly the parameter optimization is unimportant from this point of view and is therefore omitted for a moment. For the sake of comparison consider also examples of differential games solved in [14] and [15]. For the remaining of this section all variables let be scalar quantities.

Consider (1) having the form

$$\begin{aligned} \dot{x}_1 &= x_2\,, & x_1(0) &= 1\,, \\ \dot{x}_2 &= -x_1 + u + v + (1 - x_1^2)\,x_2\,, & x_2(0) &= 0\,, \end{aligned} \tag{16}$$

and the pay-off functionals (2) and (3) given as

$$J_1 = -J_2 = \int_0^1 (x_1^2 + x_2^2 + 0{\cdot}25u^2 - v^2)\,\mathrm{d}t\,, \tag{17}$$

in the zero-sum case, and as

$$\begin{aligned} J_1 &= \int_0^1 (x_1^2 + x_2^2 + 0{\cdot}5u^2)\,\mathrm{d}t\,, \\ J_2 &= [x_1^2 + x_2^2]_1 + 0{\cdot}5 \int_0^1 v^2\,\mathrm{d}t\,, \end{aligned} \tag{18}$$

in the nonzero-sum case. In both cases let the terminal constraint

$$[x_1 - x_2 - 1{\cdot}5]_1 = 0 \tag{19}$$

be present. No control constraints of type (4) are assumed.

Applying the results of [14, 15] or directly (8)–(11) it is possible to show the final conditions of the type (9) have the form (λ has clearly two components)

$$[\lambda_1 - v]_1 = 0\,, \quad [\lambda_2 + v]_1 = 0\,, \tag{20}$$

in the zero-sum, and

$$\begin{aligned} &[\lambda_1^1 - v]_1 = 0\,, & &[\lambda_1^2 + v]_1 = 0, \\ &[\lambda_2^1 - 2x_1 - v]_1 = 0\,, & &[\lambda_2^2 - 2x_2 + v]_1 = 0\,, \end{aligned} \tag{21}$$

in the nonzero-sum case. Eliminating v one obtains the missing terminal condition with former case, resp. three missing conditions in the latter case. Finally, on substituting for u and v according (9) into the corresponding equations (8) and (16) the desired two-point boundary-value problems

$$
(22)\qquad \begin{aligned}
\dot{x}_1 &= x_2\,, & x_1(0) &= 1\,,\\
\dot{x}_2 &= -x_1 + (1 - x_1^2)\,x_2 - 1{\cdot}5\lambda_1\,, & x_2(0) &= 0\,,\\
\dot{\lambda}_1 &= (1 + 2x_1x_2)\,\lambda_2 - 2x_1\,, & \lambda_1(1) &= -\lambda_2(1)\,,\\
\dot{\lambda}_2 &= -\lambda_1 - (1 - x_1^2)\,\lambda_2 - 2x_2\,, & x_1(1) - x_2(1) &= 1{\cdot}5\,,
\end{aligned}
$$

and

$$
(23)\qquad \begin{aligned}
\dot{x}_1 &= x_2\,, & x_1(0) &= 1\,,\\
\dot{x}_2 &= -x_1 + (1 - x_1^2)\,x_2 - \lambda_2^1 - \lambda_2^2\,, & x_2(0) &= 0\,,\\
\dot{\lambda}_1^1 &= (1 + 2x_1x_2)\,\lambda_1^2 - 2x_1\,, & \lambda_1^1(1) &= -\lambda_1^2(1)\,,\\
\dot{\lambda}_1^2 &= -\lambda_1^1 - (1 - x_1^2)\,\lambda_1^2 - 2x_2\,, & 2(x_1(1) + x_2(1)) &= \lambda_2^1(1) + \lambda_2^2(1)\,,\\
\dot{\lambda}_2^1 &= (1 + 2x_1x_2)\,\lambda_2^2\,, & 2x_1(1) &= \lambda_1^1(1) + \lambda_2^2(1)\,,\\
\dot{\lambda}_2^2 &= -\lambda_2^1 - (1 - x_1^2)\,\lambda_2^2\,, & x_1(1) - x_2(1) &= 1{\cdot}5\,,
\end{aligned}
$$

are obtained in the zero-sum and nonzero-sum cases, respectively.

Both problems can be solved applying the quasilinearization method [22]. Satisfactory solution (quadratic change in two consecutive iterations less than 10^{-20}) was achieved in 5, resp. in 4 iterations. The obtained solutions are identical with those determined by a gradient algorithm, however, using a sole multiplier v. Analogical discussion pertains also to the case with parameters – see Examples 2 and 4 presented further and [23].

5. FIRST-ORDER GRADIENT ALGORITHM

The numerical approach to the studied problem is based on the first-order gradient algorithm originally described by Bryson and Ho [8] for optimal control problems. Its applicability to zero-sum and nonzero-sum differential games was demonstrated by the author [11, 12, 14–17]. Let us only point out the fact that control and parameter constraints (4) are treated applying the idea of a projection [9].

The derivation of the algorithm is omitted, because it can be done rather easily having in mind the just mentioned references, e.g., see [14, 15]. Recall the fact mentioned in the last section, that a "two-sided" optimization problem will be solved, i e., two distinct parameters v_I and v_{II} will be used. The resulting algorithm then consists of the following steps.

STEP 1. Select the feasible initial solution estimate, i.e., strategies $u(t)$, $v(t)$, $t \in [0, 1]$, and the value of parameter a not violating the constraints (4).

STEP 2. Integrate the system (1) in the sense of the increasing time (forward run) using the given initial condition, on applying the values estimated in Step 1. Record the histories $x(t)$, $u(t)$ and $v(t)$, $t \in [0, 1]$, and the values $[\varphi_x^1]_1$, $[\varphi_x^2]_1$, $[\varphi_a^1]_1$, $[\psi]_1$, $[\psi_a]_1$ and $[\psi_x]_1$.

STEP 3. Integrate in the sense of the decreasing time (backward run) to obtain n-dimensional influence functions $p_i(t)$, $i = 1, 2$, and $(n \times s)$-dimensional influence function $R(t)$, $t \in [0, 1]$ according to formulas

$$\dot{p}_i = -f_x^T p_i - (L_x^i)^T, \quad p_i(1) = [\varphi_x^i]_1^T, \quad i = 1, 2, \tag{24}$$

$$\dot{R} = -f_x^T R, \qquad R(1) = [\psi_x^T]_1. \tag{25}$$

STEP 4. Compute the following expressions (dimensions are obvious from the preceding considerations)

$$\begin{aligned} I_{\psi\psi}^I &= \int_0^1 R^T f_u W_I f_u^T R \,\mathrm{d}t, \\ I_{\psi\psi}^{II} &= \int_0^1 R^T f_v W_{II} f_v^T R \,\mathrm{d}t, \\ I_{J\psi}^I &= \int_0^1 (p_1^T f_u + L_u^1)\, W_I f_u^T R \,\mathrm{d}t, \\ I_{J\psi}^{II} &= \int_0^1 (p_2^T f_v + L_v^2)\, W_{II} f_v^T R \,\mathrm{d}t, \\ I_{\psi a} &= [\psi_a]_1 + \int_0^1 R^T f_a \,\mathrm{d}t, \\ I_{Ja} &= [\varphi_a^1]_1 + \int_0^1 (p_1^T f_a + L_a^1) \,\mathrm{d}t. \end{aligned} \tag{26}$$

Here $W_I(t)$ and $W_{II}(t)$, $t \in [0, 1]$, are positive definite matrix functions having the dimensions $(m \times m)$ and $(q \times q)$, respectively.

STEP 5. Select

$$\begin{aligned} \delta\psi_I &= -\varepsilon_I[\psi]_1, \quad 0 \leqq \varepsilon_I \leqq 1, \\ \delta\psi_{II} &= -\varepsilon_{II}[\psi]_1, \quad 0 \leqq \varepsilon_{II} \leqq 1, \end{aligned} \tag{27}$$

to achieve better satisfaction of (5). Then compute s-vectors

$$\begin{aligned} v_I &= -A_{\psi\psi}^{-1}(\delta\psi_I + A_{J\psi}^T), \\ v_{II} &= -(I_{\psi\psi}^{II})^{-1}(\delta\psi_{II} + (I_{J\psi}^{II})^T), \end{aligned} \tag{28}$$

where

(29)
$$A_{\psi\psi} = I^I_{\psi\psi} + I_{\psi a} W_a I^T_{\psi a}\,,$$
$$A_{J\psi} = I^I_{J\psi} + I_{Ja} W_a I^T_{\psi a}\,,$$

with W_a being a positive definite $(r \times r)$-matrix. It is assumed that all indicated inversions exist.

STEP 6. The existing solution estimates $u(t)$, $v(t)$, and a are updated by adding the corrections

(30)
$$\delta u = -W_I[L^1_u + (p_1 + Rv_I)^T f_u]^T\,, \quad t \in [0, 1]\,,$$
$$\delta v = -W_{II}[L^2_v + (p_2 + Rv_{II})^T f_v]^T\,, \quad t \in [0, 1]\,,$$
$$\delta a = -W_a[I_{Ja} + v^T_I I_{\psi a}]^T\,.$$

STEP 7. Check, if the resulting new solution estimates

(31)
$$u(t) \triangleq u(t) + \delta u(t)\,, \quad t \in [0, 1]\,,$$
$$v(t) \triangleq v(t) + \delta v(t)\,, \quad t \in [0, 1]\,,$$
$$a \triangleq a + \delta a\,,$$

satisfy the constraints (4). If this is not the case, perform projection according to formulas

(32)
$$u(t) \triangleq \operatorname{proj}\,[u(t) \mid U]\,, \quad t \in [0, 1]\,,$$
$$v(t) \triangleq \operatorname{proj}\,[v(t) \mid V]\,, \quad t \in [0, 1]\,,$$
$$a \triangleq \operatorname{proj}\,[a \mid A]\,,$$

where $\gamma_0 \in E^n$ and $Q \subset E^n$ we define

(33)
$$\operatorname{proj}\,[\gamma_0 \mid Q] = \arg\min\,\{\|\gamma - \gamma_0\| \mid \gamma \in Q\}\,,$$

i.e., under the projection of the point γ_0 we understand its nearest point $\tilde{\gamma} \in Q$.

STEP 8. Using the projected values compute the corresponding feasible changes $\delta\bar{u}(t)$, $\delta\bar{v}(t)$, $t \in [0, 1]$, and $\delta\bar{a}$, and evaluate the relations

(34)
$$\mathscr{E}_1 = \delta\bar{a}^T W_a^{-1} \delta\bar{a} + \int_0^1 (\delta\bar{u}^T W_I^{-1} \delta\bar{u})\, \mathrm{d}t\,,$$
$$\mathscr{E}_2 = \int_0^1 (\delta\bar{v}^T W_{II}^{-1} \delta\bar{v})\, \mathrm{d}t\,.$$

If $\mathscr{E}_i < \varepsilon$, $i = 1, 2$, and $|[\psi]_1| < \delta$ (ε, δ are the permitted errors in optimality conditions and terminal constraints violation, respectively), then stop the computations; else go to Step 2.

The weighting matrices can be roughly determined using the comparison with the so-called predicted values, e.g., see [8], [14, 17]. Anyhow, it is usually sufficient to adjust these values only at the beginning of computations performing several trail runs of the algorithm. As in optimal control theory it will be always necessary to try various initial solution estimates in order to avoid the possibility of obtaining only a local optimum. The stopping condition (34) can be alternatively derived from the resulting changes of the pay-off functionals J_1, J_2 in two successive iterations.

It is also evident that the required projection (32) can cause certain difficulties, because we are not always able to compute it in analytical way. However, in a number of practically important cases, where the constraining sets are given as parallelepipeds, spheres, balls, etc., the desired projection is easily determined. Thus this approach represents an interesting numerical tool for the studied parameter optimization problem.

6. ILLUSTRATIVE EXAMPLES

To illustrate the practical importance of the developed algorithm let us solve in this section several illustrative examples. All examples were solved using the SIMFOR simulation program of Černý [24] in the connection with EAI PACER 600 computer (digital part). This interactive simulation program for the solution of two-point boundary-value problems simplifies the realization of many numerical methods for dynamic optimization. Such an approach saves a lot of routine programmer's work and enables a direct use of the whole EAI PACER 600 computer system installed at the Institute of Information Theory and Automation.

In this section all variables are scalars. The obvious transformation, if necessary, to the normalized time interval is not explicitely mentioned in the sequel. However, all indications concerning the number of iterations for convergence, weighting constants, etc., pertain to this normalized form. As stopping conditions the values $\varepsilon = 10^{-10}$ and $\delta = 10^{-6}$, if necessary, are used. As above, a parameter a is used by Player I (control u), to decrease his pay-off in the studied differential game. Let us also note that the all figures are the direct prints of computer display using the Hard Copy Unit. All integrations were done using the 3rd order variable step Runga-Kutta method with overall permitted error $e_{\max} = 10^{-4}$. Definite integrals were evaluated using Simpson's rule.

Example 1. Consider a modification of the problem studied in the previous section with first component of the initial state as a parameter. First, to make comparison possible, the zero-sum case is investigated, which originally appeared in [12]. The system equations are then as follows

$$(35)\qquad \begin{aligned} \dot{x}_1 &= x_2\,, & x_1(0) &= a\,,\\ \dot{x}_2 &= -x_1 + u - v + (1 - x_1^2)\,x_2\,, & x_2(0) &= 1\,. \end{aligned}$$

The pay-off functional

$$J_1 = -J_2 = \frac{1}{2}[x_1^2]_{1.5} + \frac{1}{2}\int_0^{1.5}(0.5u^2 - 2.0v^2)\,dt, \tag{36}$$

and constraint

$$|u(t)| \leqq 0.8, \quad t \in [0, 1]. \tag{37}$$

To obtain the problem with fixed initial state, as required in (1), let us perform the substitution

$$\begin{aligned} y_1 &= x_1 - a, \\ y_2 &= x_2. \end{aligned} \tag{38}$$

In this way we obtain the following differential game with parameter

$$\begin{aligned} \dot{y}_1 &= y_2, & y_1(0) &= 0, \\ \dot{y}_2 &= -(y_1 + a) + u - v + y_2 - (y_1 + a)^2 y_2, & y_2(0) &= 1, \end{aligned} \tag{39}$$

$$J_1 = -J_2 = \frac{1}{2}[(y_1 + a)^2]_{1.5} + \frac{1}{2}\int_0^{1.5}(0.5u^2 - 2.0v^2)\,dt. \tag{40}$$

The nominal solution estimates were $u(t) = v(t) = 0$, $t \in [0, 1]$ and $a = 0$. The desired accuracy was achieved in 18 iterations using the weighting constants $W_I =$

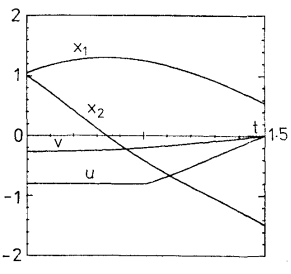

Fig. 1. Optimal solution of Example 1.

$= W_{II} = 0.5$ and $W_a = 1.0$. The obtained optimal values are $a^* = 1.0465$ and $J_1^* = -J_2^* = 0.25453$. Time-histories of all variables are shown in Fig. 1. This problem without the constraint (37) was solved in [11] with $a^* = 1.0285$ and $J_1^* = -J_2^* = 0.24941$, i.e., as it was possible to expect the constraint (37) results in a less favourable situation for the minimizing player. Also in this example the choice of initial solution estimates is not crucial and the convergence is easily achieved for other initial estimates.

Example 2. Consider again the system (35) with the pay-off functional

$$J_1 = -J_2 = \frac{1}{2}\int_0^{1.5} (x_1^2 + x_2^2 + 0{\cdot}5u^2 - 2{\cdot}0v^2)\,\mathrm{d}t \tag{41}$$

and the terminal (final-state) constraint

$$[x_1 - x_2 - 2]_{1.5} = 0\,, \tag{42}$$

i.e., after the above transformation,

$$J_1 = -J_2 = \frac{1}{2}\int_0^{1.5} [(y_1 + a)^2 + y_2^2 + 0{\cdot}5u^2 - 2{\cdot}0v^2]\,\mathrm{d}t \tag{43}$$

and

$$[y_1 + a - y_2 - 2]_{1.5} = 0\,. \tag{44}$$

With the same initial estimate as in Example 1 and with weighting constants $W_I = W_{II} = W_a = 0{\cdot}5$, and $\varepsilon_I = \varepsilon_{II} = 0{\cdot}7$, the desired accuracy was reached in 22 iterations with the optimal values $a^* = 0{\cdot}65536$ and $J_1^* = -J_2^* = 1{\cdot}2226$. The cor-

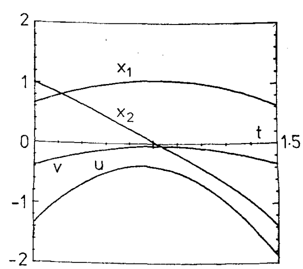

Fig. 2. Optimal solution of Example 2.

responding final state was $x_1 = 0{\cdot}62089$ and $x_2 = -1{\cdot}3791$, and terminal multipliers $v_I = -0{\cdot}93768$ and $v_{II} = 0{\cdot}68596$. The optimal solution is depicted in Fig. 2. If the terminal constraint (42) are not present, the algorithm converges to $a^* = 0{\cdot}15323$ and $J_1^* = -J_2^* = 0{\cdot}73674$, i.e., the satisfaction of (42) is rather convenient for Player II.

As also recently observed in [12], in this example the optimal solution can slightly vary with the concrete numerical procedure. For example, changing only $W_a = 0{\cdot}3$, the obtained optimal values are $a^* = 0{\cdot}65971$ and $J_1^* = -J_2^* = 1{\cdot}2195$. This circumstance can be roughly explained by the rather flat optimum and the existing "equilibrium" in satisfaction the optimality conditions and terminal constraint in the studied case. The necessary optimality conditions (10)–(12) cannot distinguish the exact optimum within the given accuracy in this constrained case. This explanation is further supported by the fact that the optimal value of pay-off functionals J_1^*, J_2^* was generally reached in less then one half of the iterations needed for convergence. Maybe the use of more sophisticated numerical methods (quasilinearization, conjugate gradients) can clear up this matter.

Example 3. Also now let the system (35) be given, however, let us study the non-zero-sum case, where

$$J_1 = \int_0^1 (x_1^2 + x_2^2 + u^2)\,\mathrm{d}t\,, \tag{45}$$

$$J_2 = [x_1^2 + x_2^2]_1 + \int_0^1 v^2\,\mathrm{d}t\,.$$

The obvious transformation analogical to (40) is not indicated here. Using same initial estimates as before the optimal solution was reached in 17 iterations having $W_I = W_{II} = 0{\cdot}5$ and $W_a = 0{\cdot}12$. It is shown in Fig. 3. Optimal parameter $a^* =$

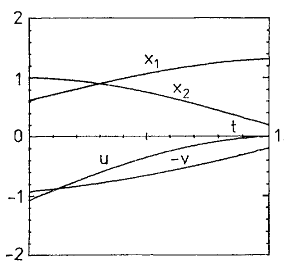

Fig. 3. Optimal solution of Example 3.

$= 0{\cdot}61448$, while $J_1^* = 1{\cdot}9428$ and $J_2^* = 2{\cdot}2111$. When solving this game with fixed initial parameter estimate $a = 0$, then, applying the algorithm of [15], the values $J_1^0 = 3{\cdot}1699$ and $J_2^0 = 5{\cdot}6285$ are obtained. Thus the pay-off functionals of both players decrease considerably as the result of optimal parameter selection performed by Player I. Other choice of (45) can result, on the other hand, in the increased pay-off functional J_2^*.

Example 4. In addition to the formulation of Example 3 let us assume the terminal constraint of the form (42), i.e.,

$$[x_1 - x_2 - 1]_1 = 0 . \tag{46}$$

With the same initial estimate as in the preceding examples and with $W_I = W_{II} = 0{\cdot}4$, $W_a = 0{\cdot}12$ and $\varepsilon_I = \varepsilon_{II} = 0{\cdot}7$, the prescribed accuracy was reached in 18 iterations with the optimal values $a^* = 0{\cdot}23080$, and $J_1^* = 0{\cdot}88603$ and $J_2^* = 1{\cdot}2194$.

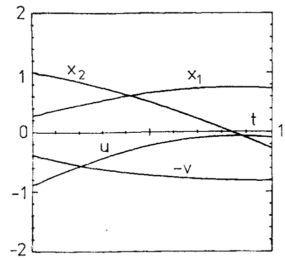

Fig. 4. Optimal solution of Example 4.

For the optimal solution see Fig. 4. The corresponding final state was $x_1 = 0{\cdot}70191$ and $x_2 = -0{\cdot}29809$, and terminal multipliers $v_I = -0{\cdot}15796$, $v_{II} = -2{\cdot}6171$. In comparison with Example 3, i.e. the case without the terminal constraint (46), one can conclude that the satisfaction of constraints is more convenient for Player I. Similar discussion as to the zero-sum case (Examples 1 and 2) pertains also to the studied nonzero-sum case (Examples 3 and 4).

7. CONCLUSIONS

For a general class of two-player nonzero-sum games, where the first player has the additional choice of certain parameters, the necessary optimality conditions were briefly derived using the calculus of variations. Based on these conditions a previously published gradient algorithm [14–17] was extended to handle also the problems of this type. In this way it was possible to treat nonlinear differential games with parameters numerically to obtain both, equilibrium strategies and parameters. Moreover, the formulation considered enables to treat in a simple way various constraints on controls and parameters (included using a projection technique) and on final state (included using multipliers).

A question of a number of terminal multipliers was discussed in detail. It shows that terminal constraints are included by a sole multiplier, which is common for all players. On the other hand, in certain situation it is possible to have distinct terminal multipliers for each player, e.g. when solving a sequence of successive optimal control problems to determine a numerical solution of a particular game — see the described gradient algorithm.

Practical experience with the suggested algorithm was reported and illustrative examples of zero-sum and nonzero-sum games were solved in detail. The obtained results confirm the applicability of the developed algorithm to the numerical solution of practical problems.

(Received June 27, 1979.)

REFERENCES

[1] E. Hofer, P. Sagirow: Optimal Parameter Selection. Regelungstechnik *15* (1957), 12, 537 to 540. In German.

[2] V. G. Boltjanskij: Mathematical Methods of Optimal Control. Nauka, Moscow 1969. In Russian.

[3] N. U. Ahmed, N. D. Georganas: On Optimal Parameter Selections. IEEE Trans. on Automatic Control *AC-18* (1973), 3, 313—314.

[4] R. V. Gamkrelidze: On Some Extremal Problems in the Theory of Differential Equations with Applications to the Theory of Optimal Control. SIAM Journal of Control *3* (1965), 1, 106—128.

[5] N. D. Georganas: Optimal Parameter Selection by Imbedding Techniques. IEEE Trans. on Automatic Control *AC-20* (1975), 1, 166—167.

[6] R. Lunderstädt: Transportations Systems with Optimal Parameters. Problems of Control and Information Theory *5* (1976), 2, 109—116.

[7] J. Doležal, P. Černý: The Application of Optimal Control Methods to the Determination of Multifunctional Catalyst. Automatizace *21* (1978), 1, 3—8. In Czech.

[8] A. E. Bryson, Y. C. Ho: Applied Optimal Control. Ginn and Company, Waltham, Massachusetts 1969.

[9] F. P. Vasiljev: Lectures on Solution Methods for Extremum Problems. Moscow University Publishers, Moscow 1974. In Russian.

[10] A. Miele: Recent Advances in Gradient Algorithms for Optimal Control Problems. Journal of Optimization Theory and Applications *17* (1975), 5—6, 361—430.

[11] J. Doležal: Optimal Parameter Estimation in Two-Player Zero-Sum Differential Games. In "Trans. of the 8-th Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes", M. Driml (ed.), Vol. A. Academia, Prague 1978, 143—156.

[12] J. Doležal: Parameter Optimization for Two-Player Zero-Sum Differential Games. Trans. of the ASME, Journal of Dynamic Systems, Measurement, and Control *101* (1979), Ser. G, 4.

[13] C. T. Leondes, T. K. Siu: Parameter Optimization for Linear Quadratic Differential Games. Trans. of the ASME, Journal of Dynamic Systems, Measurement, and Control *99* (1977), Ser. G, 1, 58—62.

[14] J. Doležal: A Gradient-Type Algorithm for the Numerical Solution of Two-Player Zero-Sum Differential Game Problems. Kybernetika *14* (1978), 6, 429—446.

[15] J. Doležal, M. Thoma: On the Numerical Solution of N-Player Nonzero-Sum Differential Games Via Nash Solution Concept. Problems of Control and Information Theory *8* (1979), 6.

[16] J. Doležal: A Numerical Approach to the Solution of Differential Games. Research Report No. 709, Institute of Information Theory and Automation, Prague 1976.

[17] J. Doležal: A SIMFOR Realization of the First-Order Gradient Algorithm. Research Report No. 848, Institute of Information Theory and Automation, Prague 1977.

[18] U. R. Prasad, I. G. Sarma: Theory of N-Person Differential Games. In "Proc. of the First International Conference on the Theory and Application of Differential Games", Y. C. Ho, G. Leitmann (eds.), Amherst, Mass. 1969. VIII 12—17.

[19] L. Berkovitz: Lectures on Differential Games. In "Differential Games and Related Topics", H. W. Kuhn, G. P. Szegö (eds.), North-Holland, Amsterdam 1971, 3—45.

[20] J. H. Case: Toward a Theory of Many Player Differential Games. SIAM Journal on Control *7* (1969), 2, 179—197.

[21] D. H. Quintana, E. J. Davison: Two Numerical Techniques to Solve Differential Game Problems. International Journal of Control *16* (1972), 3, 465—474.

[22] J. Doležal: On the Modified Quasilinearization Method for Discrete Two-Point Boundary-Value Problems. Research Report No. 788, Institute of Information Theory and Automation, Prague 1977.

[23] J. Doležal: Differential Games with Parameters. 9-th IFIP Conference on Optimization Techniques, Warsaw, September 3—9, 1979. To appear.

[24] P. Černý: Digital Simulation Program SIMFOR for the Solution of Two-Point Boundary-Value Problems. Research Report No. 639, Institute of Information Theory and Automation, Prague 1975. In Czech.

Ing. Jaroslav Doležal, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8, Czechoslovakia.