Jaroslav Doležal

A gradient-type algorithm for the numerical solution of two-player zero-sum differential game problems

*Kybernetika*, Vol. 14 (1978), No. 6, (429)--446

Persistent URL: http://dml.cz/dmlcz/124273

KYBERNETIKA — VOLUME 14 (1978), NUMBER 6

# A Gradient-Type Algorithm for the Numerical Solution of Two-Player Zero-Sum Differential Game Problems

JAROSLAV DOLEŽAL

A computational method for the determination of open-loop saddle-point strategies for a broad class of two-player zero-sum differential games is suggested. In this way the often used numerical approach to the numerical treatment of optimal control problems is extended to deal to with deal differential game problems involving control and/or final-state constraints. To illustrate the practical importance of the algorithm several concrete examples of differential games are solved in detail. The existing results in this area are also briefly reviewed.

## 1. INTRODUCTION

The aim of this paper is to show that the gradient-type algorithm of the author [1] for optimal control problems can be successfully applied when seeking for open-loop saddle-point strategies of two-player zero-sum differential games. Such possibility was briefly reported about in [2] and some additional results are to be found in [3]. As the algorithm of [1] is of rather general nature, it was in the meantime also successfully applied to the numerical solution of two-player zero-sum differential games with parameters [4–5] and general $N$-player nonzero-sum differential games [6].

The structure of this paper is as follows. In the next section a brief historical review of some of the existing results concerning two-player zero-sum differential games is presented. Similar review for the nonzero-sum case is given in [6]. In this way the reader can obtain brief information about developments in the field of differential games, one of the most important and attractive area of applied mathematics in the last decade. Special attention will be paid to the numerical solution of differential game problems, because this is in a very close connection with practical applications of the differential game theory.

The following section is then devoted to the precise formulation of a differential game problem containing the constraints on controls and final state. The concept of a saddle-point solution is introduced and necessary optimality conditions are briefly discussed.

The fourth section contains the derivation of a gradient method based upon the perturbations of linearized form of the original problem. The alternative derivation of necessary optimality conditions is obtained in this way. Then the numerical algorithm is summarized and some aspects of its computer implementation are discussed.

Finally, several concrete examples of two-player zero-sum differential games are solved in detail and the computational experience with the suggested algorithm is reported. The obtained results confirm the practical applicability of the algorithm when solving differential game problems.

## 2. A BRIEF HISTORICAL SURVEY

The subject of differential games is relatively new area of applied mathematics. It is closely related to the theory of optimal control, which is, in fact, a trivial one-player game. The overall interest in this subject is dated by the appearence of the classical Isaacs monograph [7] in 1965, although there has been some activity in this field before. However, not on a such large scale as after the Isaac's book.

Differential game is mathematical denotation of a conflicting situation which evolves continuously in time. If in the conflict participate only two individuals with strictly antagonistic aims, one speaks about the so-called two-player zero-sum differential game. The first period of the development of differential game theory is represented by the names of Pontrjagin, Berkovitz, Pšeničnyj, e.g., see [8 – 11]. The effort of many other contributors culminated in [12], where various aspects of differential game theory are studied. In the same time several books have appeared giving rigorous and comprehensive treatment of this subject. Let us mention at least the geometric approach of Blaquière et al. [13], the extremum construction of Krasovskij [14] or the approximate construction of Friedman [15] based on the so-called upper and lower game.

At this time also various special problems connected with differential games were studied. The existence of saddle-point solutions is treated by Varaiya and Lin [16], Friedman [15] and later by Elliott et al. [17] and Parthasarathy and Raghavan [18], where also the case of the so-called relaxed (mixed) strategies was considered. Mixed strategy solutions for quadratic games are explicitely given by Wilson [19].

It is not possible to mention all important results and contributors in this area due to the limited space. Therefore let us close this part of the survey by mentioning several books [20 – 23], mostly conference proceedings, where the interested reader can find a lot of supplementary material and further references. Here let us only point out the papers [24 – 26], where concrete cases of simple pursuit-evasion games are studied and analytical solutions are given. As up-to-date the most general treatment of "feedback" differential games it is necessary to mention the book of Krasovskij and Subbotin [27], where various problems of differential games are studied in detail. 

Now let us finally discuss some papers dealing with a numerical solution of two-player zero-sum differential games. As we shall see in the following section, the necessary optimality conditions for a saddle-point solution result, in general, into a nonlinear two-point boundary-value problem, and thus iterative numerical techniques must be applied. Only in the case of linear dynamics and quadratic cost functional it is possible, similarly as in optimal control, to deduce the analytical feedback control law for both participating players [28–29].

A brief survey of some numerical techniques for differential games is due to Tabac [30], however several important works are not included. Seemingly the first conceptual algorithm for the numerical iterative solution of general differential games is due Starr [31]. It is based upon the perturbations of necessary optimality conditions. In [32] the Newton-Raphson and gradient algorithms are constructed to solve zero-sum differential games with nonlinear dynamics and quadratic cost functional. On the other hand, in [33] the general complex cost functional is decomposed into simpler functionals together with a special search technique applied to determine saddle--points.

Alternative approach has been explored also by the author applying the idea of differential dynamic programming and a first-order and second-order algorithms has been worked out in [34]. For the possible application of other techniques, such as quasilinearization, penalty methods, or gradient-restoration approach, etc., see [35–39]. The practically important problem of computing the so-called closed-loop strategies (functions of time and state) is studied by Anderson [40–41] applying the method of neighbouring extremals and Järmark [42] using the differential dynamic programming approach.

The first-order gradient algorithm presented in this paper can be regarded as the extension of the previous author's work [2]. Conceptually is this simple algorithm related to [31, 32, 39]. It enables also the additional treatment of various constraints similarly as the algorithms in [35, 38], which are of more complicated nature.

In principal, any numerical method for optimal control problems can be extended to handle also differential game problems. Similarly as in optimal control it is not possible to develop a universal method due to fairly large variety of problems leading to differential games. Usually it is advisable to try more methods in the case of difficulty to exclude a possible dependence of solutions on the particular method.

## 3. STATEMENT OF THE PROBLEM

In this section let us give the precise formulation of a two-player zero-sum differential game. For the sake of simplicity the matrix notation will be used throughout the paper. It is assumed that all considered vectors are 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 an $n$-dimensional Euclidean space. To simplify further considerations only the problems with fixed final time will be studied, i.e., without any loss of generality it is then assumed, that the independent variable (time) $t \in [0, 1]$. The case of free final time, e.g., the pursuit-evasion differential games, can be treated analogously as in optimal control. Some details in this respect can be found in [2–3].

As two-player zero-sum differential game is denoted a dynamic situation evolving in time, which is influenced by two individuals with strictly conflicting aims. More exactly, consider a dynamic system described by a vector differential equation and initial conditions

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

subject to the final-state constraints

$$[\psi(x)]_1 = 0 \tag{2}$$

and the control constraints

$$u(t) \in U \subset E^{m_1}\,, \quad v(t) \in V \subset E^{m_2}\,, \quad t \in [0, 1]\,. \tag{3}$$

The cost functional is given as

$$J = [\varphi(x)]_1 + \int_0^1 L(x, u, v, t)\,\mathrm{d}t\,. \tag{4}$$

In these equations, $x(t) \in E^n$ denotes the state vector, $u(t) \in E^{m_1}$ is the control vector of the minimizing player (denoted $P$), and $v(t) \in E^{m_2}$ is the control vector of the maximizing player (denoted $E$). The various functions are defined as follows:

$$\begin{aligned} &f : E^n \times E^{m_1} \times E^{m_2} \times E^1 \to E^n\,, \quad \psi : E^n \to E^q\,, \\ &L : E^n \times E^{m_1} \times E^{m_2} \times E^1 \to E^1\,, \quad \varphi : E^n \to E^1\,, \end{aligned} \tag{5}$$

and index 1 denotes the evaluation of the corresponding expressions at $t = 1$.

The aim of player $P$ is to choose a control $u$ to minimize $J$, whereas the aim of player $E$ is strictly antagonistic, i.e., to maximize $J$. As an admissible strategy pair $(u, v)$ is denoted a pair of piecewise-continuous vector functions $u(t)$, $v(t)$, $t \in [0, 1]$, not violating the constraints (3), and such that the resulting state trajectory according to (1) is a unique absolutely continuous function satisfying the final-state constraints (2) at $t = 1$. In this respect one can speak about the open-loop strategies, because the controls of both players are only the functions of $t$.

The solution of the differential game (1)–(4) is the well-known saddle-point, i.e., such admissible strategy pair $(u^\circ, v^\circ)$ for which

$$J(u^\circ, v) \leqq J(u^\circ, v^\circ) \leqq J(u, v^\circ) \tag{6}$$

for any $u$ and $v$ constituting the admissible pairs $(u, v^\circ)$ and $(u^\circ, v)$, respectively. Note that the inequalities (6) state that if $E$ plays non-optimaly, $P$ can reduce the cost functional $J$ still playing $u^\circ$, while if $P$ plays non-optimaly, $E$ can increase $J$ by playing $v^\circ$.

Now let us assume the existence of a saddle-point, the so-called "solution in small", i.e. no barriers or other possible singular surfaces are encountered. For the sufficient existence conditions see [15–18]. Then the necessary optimality conditions for a saddle-point solution can be applied, see [9, 13]. First, let us neglect for a moment the constraints (3).

Applying the calculus of variations it is not very hard to show that if $(u^\circ, v^\circ)$ is a saddle-point of the differential game given by (1)–(2) and (4), then there exist multipliers $\lambda(t) \in E^n$, $t \in [0, 1]$, and $\nu \in E^q$ such that (symbol $T$ denotes transposition and subscripts stand for the corresponding partial derivatives)

$$\dot{\lambda} = -f_x^T \lambda - L_x^T, \quad t \in [0, 1], \tag{7}$$

$$[\lambda - \varphi_x^T - \psi_x^T \nu]_1 = 0, \tag{8}$$

$$f_u^T \lambda + L_u^T = 0, \quad t \in [0, 1], \tag{9}$$

$$f_v^T \lambda + L_v^T = 0, \quad t \in [0, 1]. \tag{10}$$

In these relations all functions (arguments omitted for the sake of simplicity) are to be evaluated along the saddle-point solution $(u^\circ, v^\circ)$ and the corresponding saddle-point trajectory $x^\circ$. Observe that (7)–(8) together with (1)–(2) define the unknown multipliers $\lambda(t)$ and $\nu$, while (9)–(10) give the so-called saddle-point conditions.

On combining (1)–(2) and (7)–(10) one easily concludes that, in principle, a non-linear two-point boundary-value problem for the system of $2n$ differential equations has to be solved. As it is well-known, such problem cannot be generally solved in an analytical way, and therefore some iterative numerical techniques must be applied, e.g. see [43] in this respect. This approach is then denoted as an indirect one, based upon the application of necessary optimality conditions and excluding $u$ and $v$ according to (9) and (10), respectively.

If the constraints (3) are present the equations (9)–(10) have the form

$$H(u^\circ, v^\circ) = \min_{u \in U} H(u, v^\circ) = \max_{u \in V} H(u^\circ, v), \quad t \in [0, 1], \tag{11}$$

where

$$H(u, v) = L(x, u, v, t) + \lambda^T f(x, u, v, t), \quad t \in [0, 1]. \tag{12}$$

Constraints (3) will be incorporated into the numerical procedure using the idea of projection (clipping-off technique) [44]. It should be also noted that the constraints (3) can be generally time-dependent, i.e. to be of the form $U(t)$, $V(t)$, $t \in [0, 1]$, where $U$ and $V$ are now the set-valued functions in $E^{m_1}$ and $E^{m_2}$, respectively.

## 4. DERIVATION OF THE ALGORITHM

As in optimal control the basis for the further described gradient algorithm is the selection of a suitable nominal strategy pair $(u, v)$ together with a procedure for the appropriate modification of this nominal solution estimate to obtain an improved value of the cost functional $J$, and in the same time, to achieve a better satisfaction of constraints (2)–(3).

As the constraints (3) will be treated directly applying the projection, let us neglect them when deriving the basic algorithm. Here presented derivation is based upon the perturbations of linearized necessary saddle-point conditions. In optimal control a similar approach is described in [45]. Alternatively also the approach based upon the application of variational methods [46] is possible, e.g. see [2–3].

To simplify further considerations let us exclude the integral part of the original cost functional (4) introducing the additional state variable $x^{n+1}$. In an obvious way the following equations are obtained:

$$\tilde{x} = \begin{pmatrix} x \\ x^{n+1} \end{pmatrix}, \quad \tilde{f} = \begin{pmatrix} f \\ L \end{pmatrix}, \quad \tilde{x}_0 = \begin{pmatrix} x \\ 0 \end{pmatrix}, \tag{13}$$

$$\tilde{\psi} = \psi, \quad \tilde{\varphi} = \varphi + x^{n+1}.$$

Then (1), (2) and (4) change to

$$\dot{\tilde{x}} = \tilde{f}(\tilde{x}, u, v, t), \quad \tilde{x}(0) = \tilde{x}_0, \quad t \in [0, 1], \tag{14}$$

$$[\tilde{\psi}(\tilde{x})]_1 = 0, \tag{15}$$

$$\tilde{J} = [\tilde{\varphi}(\tilde{x})]_1. \tag{16}$$

Now assume nominal strategies $u$ and $v$ constituting a nominal saddle-point solution estimate $(u, v)$ and the corresponding trajectory $\tilde{x}$ according to (14). Further let $\Delta u$ and $\Delta v$ be strategy changes and suppose that the updated strategies $u + \Delta u$ and $v + \Delta v$ produce an updated trajectory $\tilde{x} + \Delta\tilde{x}$ such that the following linearization of (14) is satisfied:

$$\Delta\dot{\tilde{x}} = \tilde{f}_{\tilde{x}}\, \Delta\tilde{x} + \tilde{f}_u\, \Delta u + \tilde{f}_v\, \Delta v, \quad \Delta\tilde{x}(0) = 0, \quad t \in [0, 1]. \tag{17}$$

All here and further introduced expressions are to be evaluated at $\tilde{x}, u, v$. The resulting first order change of (16) due to $\Delta u$ and/or $\Delta v$ can be expressed as

$$\Delta\tilde{J} = [\tilde{\varphi}(\tilde{x} + \Delta\tilde{x})]_1 - [\tilde{\varphi}(\tilde{x})]_1 = [\tilde{\varphi}_{\tilde{x}}\, \Delta\tilde{x}]. \tag{18}$$

First let us study only the effect of the strategy change $\Delta u$ of player $P$, i.e. assume $\Delta v = 0$. It is the well-known fact that the solution of (17) is then given by the formula

$$[\Delta\tilde{x}]_1 = \int_0^1 \tilde{X}(t)\, \tilde{f}_u\, \Delta u(t)\, \mathrm{d}t, \tag{19}$$

where the $(n + 1) \times (n + 1)$-dimensional matrix $\tilde{X}(t)$ satisfies the matrix differential equation

(20) $$\dot{\tilde{X}} = -\tilde{X}\tilde{f}_{\tilde{x}}, \quad \tilde{X}(1) = I, \quad t \in [0, 1],$$

with $I$ the appropriate unit matrix. Then

(21) $$\Delta \tilde{J}_P = \int_0^1 \tilde{J}_u(t)\, \Delta u(t)\, \mathrm{d}t,$$

where

(22) $$\tilde{J}_u(t) = [\tilde{\varphi}_{\tilde{x}}]_1\, \tilde{X}(t)\, \tilde{f}_u = \tilde{p}_P^T(t)\, \tilde{f}_u, \quad t \in [0, 1].$$

Clearly $p_P(t)$ satisfies the $(n + 1)$-dimensional differential equation

(23) $$\dot{\tilde{p}}_P = -\tilde{f}_{\tilde{x}}^T \tilde{p}_P, \quad \tilde{p}_P(1) = [\tilde{\varphi}_{\tilde{x}}^T], \quad t \in [0, 1].$$

Now let us derive the expression for the change $[\Delta\tilde{\psi}_P]_1$ of $[\tilde{\psi}(\tilde{x})]_1$ due to $\Delta u$. Using the same arguments as above, it can be shown that

(24) $$[\Delta\tilde{\psi}_P]_1 = \int_0^1 \tilde{\psi}_u(t)\, \Delta u(t)\, \mathrm{d}t,$$

where

(25) $$\tilde{\psi}_u(t) = \tilde{R}_P^T \tilde{f}_u, \quad t \in [0, 1],$$

and $\tilde{R}_P(t)$ satisfies the $((n + 1) \times q)$-dimensional differential equation

(26) $$\dot{\tilde{R}}_P = -\tilde{f}_{\tilde{x}}^T \tilde{R}_P, \quad \tilde{R}_P(1) = [\tilde{\psi}_{\tilde{x}}^T]_1, \quad t \in [0, 1].$$

On the other hand, if the effect of $\Delta v$ for $\Delta u = 0$ is studied, one can easily realize, having in mind the maximization of (16), that only the final condition in (23) will have the reversed sign, i.e. the following expressions are valid for player $E$:

(27) $$\Delta \tilde{J}_E = \int_0^1 \tilde{J}_v(t)\, \Delta v(t)\, \mathrm{d}t = \int_0^1 \tilde{p}_E^T(t)\, \tilde{f}_v\, \Delta v(t)\, \mathrm{d}t,$$

where

(28) $$\dot{\tilde{p}}_E = -f_{\tilde{x}}^T \tilde{p}_E, \quad \tilde{p}_E(1) = -[\tilde{\varphi}_{\tilde{x}}^T]_1, \quad t \in [0, 1],$$

and

(29) $$[\Delta\tilde{\psi}_E]_1 = \int_0^1 \tilde{\psi}_v(t)\, \Delta v(t)\, \mathrm{d}t = \int_0^1 \tilde{R}_E(t)\, \tilde{f}_v\, \Delta v(t)\, \mathrm{d}t,$$

where

(30) $$\dot{\tilde{R}}_E = -\tilde{f}_{\tilde{x}}^T \tilde{R}_E, \quad \tilde{R}_E(1) = [\tilde{\psi}_{\tilde{x}}^T]_1, \quad t \in [0, 1].$$

Composing now (23) with (28) and (26) with (30), it is obtained that

$$\tilde{p}_P(t) = -\tilde{p}_E(t) = \tilde{p}(t)\,, \quad t \in [0, 1]\,, \tag{31}$$
$$\tilde{R}_P(t) = \tilde{R}_E(t) = R(t)\,, \quad t \in [0, 1]\,.$$

Having in mind the definition (13) one easily obtains that the last component of $\tilde{p}(t)$ is equal to 1 and the last row of $\tilde{R}(t)$ is zero. Then in terms of the original formulation (4) we have

$$\Delta J_P = \int_0^1 J_u(t)\, \Delta u(t)\, \mathrm{d}t = \int_0^1 (p^T(t) f_u + L_u)\, \Delta u(t)\, \mathrm{d}t\,, \tag{32}$$
$$\Delta J_E = \int_0^1 J_v(t)\, \Delta v(t)\, \mathrm{d}t = \int_0^1 (-p^T(t) f_v - L_v)\, \Delta v(t)\, \mathrm{d}t\,,$$

$$[\Delta \psi_P]_1 = \int_0^1 \psi_u(t)\, \Delta u(t)\, \mathrm{d}t = \int_0^1 R^T(t) f_u\, \Delta u(t)\, \mathrm{d}t\,, \tag{33}$$
$$[\Delta \psi_E]_1 = \int_0^1 \psi_v(t)\, \Delta v(t)\, \mathrm{d}t = \int_0^1 R^T(t) f_v\, \Delta v(t)\, \mathrm{d}t\,,$$

where the $n$-vector $p(t)$ and $(n \times q)$-matrix $R(t)$ are given by the relations

$$\dot{p} = -f_x^T p - L_x^T\,, \quad p(1) = [\varphi_x^T]_1\,, \quad t \in [0, 1]\,, \tag{34}$$
$$\dot{R} = -f_x^T R\,, \quad R(1) = [\psi_x^T]_1\,, \quad t \in [0, 1]\,.$$

Alternatively, $p(t)$ and $R(t)$ are denoted as the influence functions.

Consider now the augmented cost functional

$$\hat{J} = J + \nu^T[\psi(x)]_1\,, \tag{35}$$

where $\nu \in E^q$ is a constant vector of multipliers. It follows from (32)–(33) that

$$\hat{J}_u(t) = J_u(t) + \nu_P^T\, \psi_u(t)\,, \quad t \in [0, 1]\,, \tag{36}$$
$$\hat{J}_v(t) = J_v(t) + \nu_E^T\, \psi_v(t)\,, \quad t \in [0, 1]\,.$$

Then clearly the choices

$$\Delta u(t) = -W_P(t)\, \hat{J}_u^T(t) = -W_P(t)\, (L_u + (p + R\nu_P)^T f_u)^T\,, \quad t \in [0, 1]\,, \tag{37}$$
$$\Delta v(t) = -W_E(t)\, \hat{J}_v^T(t) = W_E(t)\, (L_v + (p - R\nu_E)^T f_v)^T\,, \quad t \in [0, 1]\,,$$

will guarantee the desired changes of $\hat{J}$ for both players. Here $W_P(t)$ and $W_E(t)$ are positive definite matrices having the dimensions $(m_1 \times m_1)$ and $(m_2 \times m_2)$, respectively. Then we have , in fact,

$$\Delta \hat{J}_P < 0\,, \quad \Delta \hat{J}_E > 0\,. \tag{38}$$

This is of course valid for sufficiently small changes $\Delta u(t)$ and $\Delta v(t)$.

Substituting $\Delta u(t)$, $\Delta v(t)$ from (37) into (32) and (33), we find that

$$\begin{aligned} v_P &= -(I^P_{\psi\psi})^{-1}\left([\Delta\psi_P]_1 + (I^P_{J\psi})^T\right), \\ v_E &= -(I^E_{\psi\psi})^{-1}\left([\Delta\psi_E]_1 + (I^E_{J\psi})^T\right), \end{aligned} \tag{39}$$

and the predicted changes finally are

$$\begin{aligned} \Delta J_P &= -(I^P_{JJ} - I^P_{J\psi} v_P)\,, \\ \Delta J_E &= -(I^E_{JJ} - I^E_{J\psi} v_E)\,. \end{aligned} \tag{40}$$

In these relations the following shorthand notation is used (dimensions are obvious from the above considerations)

$$\begin{aligned} I^P_{\psi\psi} &= \int_0^1 R^T f_u W_P f_u^T R \,\mathrm{d}t\,, \\ I^E_{\psi\psi} &= \int_0^1 R^T f_v W_E f_v^T R \,\mathrm{d}t\,, \\ I^P_{J\psi} &= \int_0^1 (p^T f_u + L_u)\, W_P f_u^T R \,\mathrm{d}t\,, \\ I^E_{J\psi} &= -\int_0^1 (p^T f_v + L_v)\, W_E f_v^T R \,\mathrm{d}t\,, \\ I^P_{JJ} &= \int_0^1 (p^T f_u + L_u)\, W_P (f_u^T p + L_u^T) \,\mathrm{d}t\,, \\ I^E_{JJ} &= \int_0^1 (p^T f_v + L_v)\, W_E (f_v^T p + L_v^T) \,\mathrm{d}t\,. \end{aligned} \tag{41}$$

It is assumed that the inversions indicated in (39) exist.

If the optimum is approached, it follows from (37) that

$$\begin{aligned} L_u + (p + R v_P)^T f_u &\to 0\,, \quad t \in [0, 1]\,, \\ L_v + (p - R v_E)^T f_v &\to 0\,, \quad t \in [0, 1]\,, \end{aligned} \tag{42}$$

which is in agreement with the optimality conditions (9)–(10).

The weighting matrices $W_P(t)$ and $W_E(t)$ in (37) should be chosen in such manner to obtain the approximate agreement of actual and predicted changes in $J$ and $\psi$

according to (32) and (33). In fact, the values $[\Delta\psi_P]_1$ and $[\Delta\psi_E]_1$ in (33) are to be selected to achieve the better final-state constraints satisfaction, e.g., one can put

$$(43)\qquad \begin{aligned} [\Delta\psi_P]_1 &= -\varepsilon_P[\psi(x)]_1\,, \quad 0 \leqq \varepsilon_P \leqq 1\,,\\ [\Delta\psi_E]_1 &= -\varepsilon_E[\psi(x)]_1\,, \quad 0 \leqq \varepsilon_E \leqq 1\,. \end{aligned}$$

It should be also noted, that letting $\varepsilon_P = \varepsilon_E = 0$ in (43), the relations (39) then correspond to those of [45]. However, in the further summarized algorithm such choice usually leads to a slower final-state constraints satisfaction, as it was confirmed by practical computations.

As stopping conditions we can require (see (39)–(40) with $[\Delta\psi_P]_1 = [\Delta\psi_E]_1 = 0$)

$$(44)\qquad \begin{aligned} &[\psi(x)]_1 = 0\,,\\ &I^P_{JJ} - I^P_{J\psi}(I^P_{\psi\psi})^{-1}\,(I^P_{J\psi})^T = 0\,,\\ &I^E_{JJ} - I^E_{J\psi}(I^E_{\psi\psi})^{-1}\,(I^E_{J\psi})^T = 0 \end{aligned}$$

with the desired accuracy.

If the strategy constraints (3) are present, the new solution estimates

$$(45)\qquad \begin{aligned} u(t) &\triangleq u(t) + \Delta u(t)\,, \quad t \in [0, 1]\,,\\ v(t) &\triangleq v(t) + \Delta v(t)\,, \quad t \in [0, 1]\,, \end{aligned}$$

are always checked if they meet these constraints. This not being the case, the following projection is performed, e.g., see [44, p. 263], according to formulas

$$(46)\qquad \begin{aligned} u(t) &\triangleq \operatorname{proj}\,[u(t) \mid U]\,, \quad t \in [0, 1]\,,\\ v(t) &\triangleq \operatorname{proj}\,[v(t) \mid V]\,, \quad t \in [0, 1]\,, \end{aligned}$$

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

$$(47)\qquad \operatorname{proj}\,[\gamma_0 \mid Q] = \arg\min\,[\|\gamma - \gamma_0\| \mid \gamma \in Q]\,,$$

i.e., under the projection of a point $\gamma_0$ we understand its nearest point $\tilde{\gamma} \in Q$.

It should be noted at this place that such projection is sometimes troublesome to perform analytically in spite of its simple geometric interpretation. The existence of this projection is guarranted if the set $Q$ is compact. If $Q$ is moreover strictly convex, the projection will be unique. However, in a number of practical cases, where the constraining sets in (3) are given as parallelepipeds, spheres, elipsoids, etc., the desired projection is easily performed. Observe that this reasoning will be valid if the sets $U(t)$ and $V(t)$ vary with $t$ and have a piecewise-continuous boundary.

In this case the stopping conditions (44) must be modified in the following way. We consider only the "feasible" changes $\Delta\bar{u}(t)$ and $\Delta\bar{v}(t)$, i.e., such portions of $\Delta u(t)$

and $\Delta v(t)$ given by (37), that the new solution estimates (45) satisfy the constraints (3). Then instead of (44) we use

(48)
$$\mathscr{E}_P = \int_0^1 \Delta\bar{u}^{\mathrm{T}}(t)\ W_P^{-1}(t)\ \Delta\bar{u}(t)\ \mathrm{d}t\ ,$$
$$\mathscr{E}_E = \int_0^1 \Delta\bar{v}^{T}(t)\ W_E^{-1}(t)\ \Delta\bar{v}(t)\ \mathrm{d}t\ .$$

It can be shown that if no constraints (3) are present the conditions (48) are identical with those given by (44).

## 5. SUMMARY OF THE ALGORITHM

In the section let us summarize the computational steps of the just derived first-order gradient algorithm for differential game problems (1)–(4). The resulting algorithm consists of the following steps.

**Step 1.** Select a nominal feasible saddle-point estimate $(u, v)$, i.e., functions $u(t)$ and $v(t)$, $t \in [0, 1]$, not violating the constraints (3).

**Step 2.** Using this estimate integrate the system equations (1) in the sense of the increasing time (forward run) with the specified initial condition $x_0$. Record the time-histories $x(t)$, $u(t)$ and $v(t)$, $t \in [0, 1]$, and the values $[\varphi_x]_1$, $[\psi]_1$, $[\psi_x]_1$.

**Step 3.** Integrating in the sense of the decreasing time (backward run) determine $n$-dimensional function $p(t)$ and $(n \times q)$-dimensional function $R(t)$, $t \in [0, 1]$ according to formulas (34).

**Step 4.** Compute the expressions given by the relations (41).

**Step 5.** Select $[\Delta\psi_P]_1$ and $[\Delta\psi_E]_1$ according to (43) by an appropriate choice of $\varepsilon_P$ and $\varepsilon_E$, and compute $\nu_P$ and $\nu_E$ from the relations (39).

**Step 6.** Update the existing solution estimates $u(t)$ and $v(t)$ by adding the corrections $\Delta u(t)$ and $\Delta v(t)$, according to (37) so that (45) is obtained.

**Step 7.** Check, if the resulting new solution estimates (45) satisfy the constraints (3). This not being the case, perform the projection as indicated in relations (46).

**Step 8.** Using the projected values compute the feasible changes $\Delta\bar{u}(t)$ and $\Delta\bar{v}(t)$, $t \in [0, 1]$, and evaluate relations $\mathscr{E}_P$ and $\mathscr{E}_E$ given by (48). If $\mathscr{E}_E < \varepsilon$, $\mathscr{E}_P < \varepsilon$, and $[\psi]_1 < \delta$, where $\varepsilon$ and $\delta$ are the permitted errors in optimality conditions and final-state constraints, respectively, then stop the computations; else go to **Step 2**.

In practical realization of the algorithm the appropriate weighting matrices $W_P(t)$ and $W_E(t)$, $t \in [0, 1]$, and the constants $\varepsilon_P$ and $\varepsilon_E$ are usually selected only in the first iteration of the algorithm and these values are used throughout the remaining computational process. Practical experience has further shown that the mentioned weighting matrices can be chosen constant.

Alternatively, the stopping condition (48) can be replaced on considering the actual changes in the cost functional $J$. If its changes produced by the algorithm are satisfactorily small, then the computations are stopped. Let us also remark that, in general, it is advisable, similarly as in optimal control theory to try various nominal solution estimates to decrease the possibility of obtaining only a "local" saddle-point.

## 6. ILLUSTRATIVE EXAMPLES

The practical importance of the presented algorithm is illustrated by the solution of several concrete examples of two-player zero-sum differential games. All examples were solved using the SIMFOR simulation program of Černý [47] in the connection with EAI PACER 600 hybrid computer (digital part) installed at the Institute of Information Theory and Automation. This interactive simulation program for the solution of two-point boundary-value problems simplifies to a great extent the realization of many numerical methods for dynamic optimization. Moreover, this program saves a lot of routine programmer's work and enables a direct use of the whole EAI PACER 600 computer system.

Through this section all variables are scalars. As stopping criteria the values $\varepsilon = 10^{-10}$ and $\delta = 10^{-6}$ (if necessary) are used. Let us also note that all figures are the direct prints of the computer display using the Hard Copy Unit. All integrations were performed using the 3rd order variable step Runge-Kutta method with the overall permitted error $e_{\max} = 10^{-4}$. Definite integrals were evaluated using the Simpson's rule.

The further reported convergence results (weights $W_P$, $W_E$, constants $\varepsilon_P$, $\varepsilon_E$) pertain to the normalized formulation (1)–(4) having $t \in [0, 1]$. If the problem has $t \in$ $t \in [t_0, t_f]$, one performs the evident time-scale transformation

$$t' = \frac{t - t_0}{t_f - t_0} \tag{49}$$

to have $t' \in [0, 1]$. This transformation is not explicitely mentioned in the sequel. The given convergence results are those obtained for the nominal solution estimate $u(t) = v(t) = 0$, $t \in [t_0, t_f]$. The corresponding weighting matrices $W_P(t)$ and $E_E(t)$ are assumed to be constant for $t \in [t_0, t_f]$.

For all examples the most important data are collected in Table 1. The meaning of all symbols is obvious, only $N$ stands for the number of iterations needed for convergence, $t_f$ denotes the final time and $J^\circ$ are the optimal costs.

**Example 1.** Consider a simple example of a linear-quadratic differential game. Given the system equations

$$\begin{aligned} \dot{x}_1 &= -x_1 + ax_2 + v, & x_1(0) &= 2\cdot 0, \\ \dot{x}_2 &= 0\cdot 5x_1 + 0\cdot 5x_2 + 1\cdot 5u, & x_2(0) &= 2\cdot 0, \end{aligned} \tag{50}$$

and the cost functional

$$J = \tfrac{1}{2}[x_1^2 + x_2^2]_{0\cdot 9} + \tfrac{1}{2}\int_0^{0\cdot 9} (u^2 - v^2)\,\mathrm{d}t, \tag{51}$$

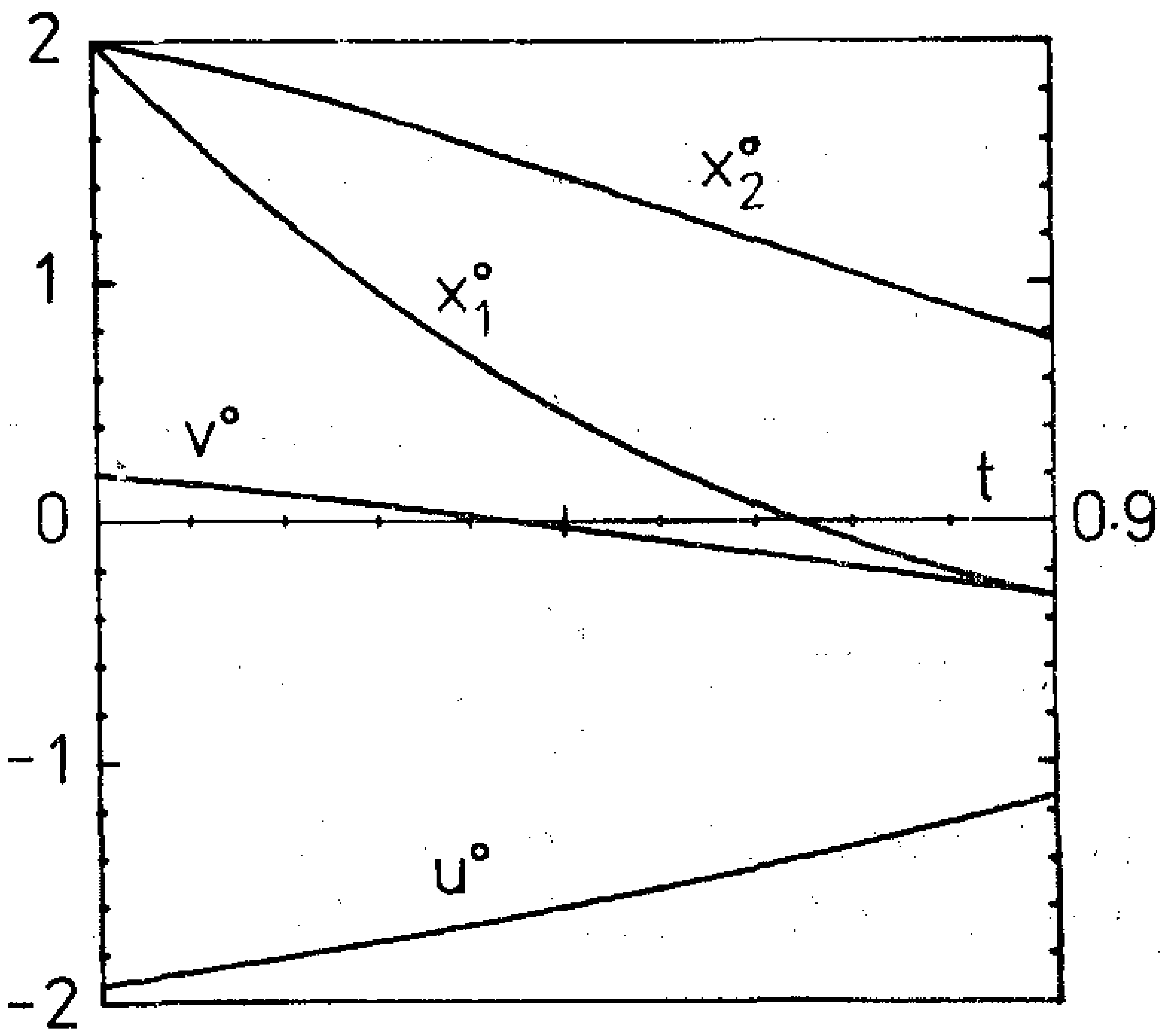


**Fig. 1.** Optimal solution of Example 1.

i.e., $t \in [0, 0\cdot 9]$. Further let $a = -1\cdot 3657$ as in [4–5], where the problem of the additional parameter optimization by the minimizing player was studied. For this value of parameter $a$, which is optimal for the minimizing player, the data concerning the convergence are summarized in Table 1 and the optimal solution (saddle-point strategies $u^\circ(t)$, $v^\circ(t)$ and the corresponding state trajectory $x^0(t)$) is shown in Fig. 1.

In this simple case it is possible to obtain the saddle-point strategies also in a "feedback" form when solving the corresponding Riccati-like equations by backward integration, e.g., see [28–29]. In fact, a linear two-point boundary-value problem is to be solved. Then clearly a considerably less computational effort will be needed in comparison with the iterative procedure of this first-order algorithm. The example was included here to illustrate the principal possibility to solve this class of problems.

However, the main importance of the algorithm is in its applicability for nonlinear problem or problems with various constraints. For example, if in this example certain control constraints are assumed, the mentioned direct solution via the Riccati equation is no more possible.

**Table 1.** Summary of the convergence results for the illustrative examples.

| | $J^{\circ}$ | $x_1^{\circ}(t_f)$ | $x_2^{\circ}(t_f)$ | $N$ | $W_P$ | $W_E$ |
|---|---|---|---|---|---|---|
| Example 1 | 1·4838 | −0·31143 | 0·76331 | 26 | 0·3 | 0·5 |
| Example 2 | 0·54896 | 0·57774 | −0·30865 | 36 | 0·47 | 0·45 |
| Example 3 | 0·96897 | 0·65257 | −0·67400 | 29 | 0·5 | 0·5 |
| Example 4 | 0·97166 | 0·63882 | −0·71527 | 30 | 0·5 | 0·5 |
| Example 5 | 0·99885 | 0·60374 | −0·89626 | 13 | 1·0 | 0·5 |

**Example 2.** Consider the system dynamics representing the Van der Pol oscillator

$$\begin{aligned} &\dot{x}_1 = x_2\,, && x_1(0) = 0{\cdot}0\,, \\ &\dot{x}_2 = -x_1 + u - v + (1 - x_1^2)\,x_2\,, && x_2(0) = 1{\cdot}0\,, \end{aligned} \tag{52}$$

and the cost functional

$$J = \tfrac{1}{2}[x_1^2]_{1.5} + \tfrac{1}{2}\int_0^{1.5} (0{\cdot}5u^2 - 2{\cdot}0v^2)\,\mathrm{d}t\,, \tag{53}$$

i.e., $t \in [0, 1{\cdot}5]$. The obtained results are given in Table 1 and Fig. 2. It can be seen that a very good agreement with the results of [32] is reached.

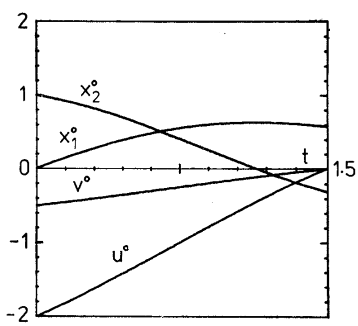


**Fig. 2.** Optimal solution of Example 2.

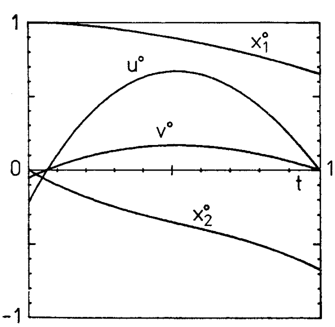


**Fig. 3.** Optimal solution of Example 3.

**Example 3.** Consider the system (52) with the reversed initial condition and $t \in [0, 1]$:

$$(54) \qquad \dot{x}_1 = x_2\,, \qquad x_1(0) = 1{\cdot}0\,,$$
$$\dot{x}_2 = -x_1 + u - v + (1 - x_1^2)\,x_2\,, \quad x_2(0) = 0{\cdot}0\,,$$

and the integral cost functional

$$(55) \qquad J = \int_0^1 (x_1^2 + x_2^2 + 0{\cdot}25u^2 - v^2)\,\mathrm{d}t\,.$$

The optimal solution is depicted in Fig. 3. However, see also Table 1 for further details.

**Example 4.** In addition to (54)–(55) assume the strategy constraints

$$(56) \qquad |u(t)| \leqq 0{\cdot}5\,, \quad |v(t)| \leqq 0{\cdot}15\,, \quad t \in [0, 1]\,.$$

These constraints were treated applying the described projection technique. For the optimal solution in this case see Fig. 4 and Table 1.

**Example 5.** Consider once more (54)–(55) and the additional final-state constraint

$$(57) \qquad [x_1 - x_2 - 1{\cdot}5]_1 = 0\,.$$

Also this problem was solved rather easily using the described algorithm with $\varepsilon_P = \varepsilon_E = 0{\cdot}5$ to achieve faster constraint satisfaction. The obtained optimal solution is shown in Fig. 5 and described again in Table 1. For optimal solution we had $v_P = -0{\cdot}41126$ and $v_E = 0{\cdot}53544$. Let us note, that in this example the optimal

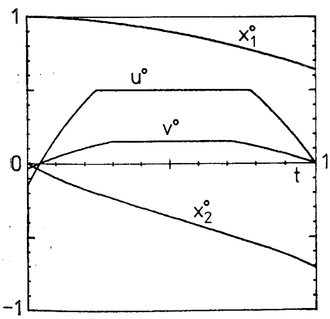


Fig. 4. Optimal solution of Example 4.

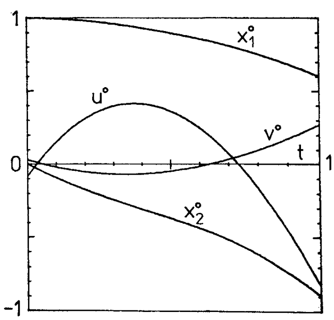


Fig. 5. Optimal solution of Example 5.

solution slightly varied with a concrete computational process, i.e., practically unimportant dependence on the choice of $W_P$, $W_E$, $\varepsilon_P$ and $\varepsilon_E$ was observed. This circumstance can be probably explained by the "equilibrium" in satisfaction of the saddle-point conditions (9)–(10) and the final-state constraints (2) included using the multiplier $v$. Within the given accuracy it is not further possible to distinguish the exact optimum.

In all examples the various nominal solution estimates were tested, all leading to the same final results. Also the additional tests were perform to be sure that the converged solutions are really saddle-point, e.g., by small perturbations of the converged strategies $u^\circ(t)$, and $v^\circ(t)$ in the connection with condition (6).

To the overall characteristics of the presented examples let us remark that the optimal costs $J^\circ$ in Table 1 were usually reached in a smaller number of iterations than needed for convergence. This shows that the saddle-point solutions are rather flat and that the prescribed accuracy is sufficient to obtain meaningful results.

## 7. CONCLUSIONS

A first-order gradient algorithm was developed for a broad class of nonlinear two-player zero-sum differential game problems involving strategy and final-state constraints. The derivation of this algorithm based on a linearization of the original differential game and introduction of the so-called influence functions. The strategy constraints are treated directly applying the idea of projection, while the final-state constraints are incorporated indirectly using multipliers.

The practical experience with the suggested algorithm is reported and several concrete examples of two-player nonzero-sum differential games are solved in detail to illustrate various computational aspects of this numerical procedure. The main advantage is its simple form, which is easy to realize especially in the connection with the existing interactive simulation program SIMFOR described in [47].

(Received June 29, 1978.)

REFERENCES

[1] J. Doležal, P. Černý: The application of optimal control methods to the determination of multifunctional catalysts. 23rd CHISA Conference, Mariánské Lázně 1976. See also: Automatizace *21* (1978), 1, 3–8. In Czech.

[2] J. Doležal, P. Černý: Numerical solution of differential games. In "Proceedings of Cybernetics Conference", Czechoslovak Cybernetical Society, Prague 1976, 39–51. In Czech.

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

[4] J. Doležal: Optimal parameter estimation in two-player zero-sum differential games. In "Transactions of the 8th Prague Conference on Information Theory, Statistical Decision Functions, Random Processes", Academia, Prague 1978, 143–156.

[5] J. Doležal: Parameter optimization for two-player zero-sum differential games. Transactions ASME, Ser. G. J. Dynamic Systems, Measurement, and Control. Submitted for publication.

[6] J. Doležal, M. Thoma: On the numerical solution of $N$-player nonzero-sum differential games. Problems of Control and Information Theory. Submitted for publication.

[7] R. Isaacs: Differential Games. Wiley, New York 1965.

[8] L. S. Pontrjagin: On differential games theory. Uspechi Mat. Nauk *XXI* (1966), 4, 219—274. In Russian.

[9] L. D. Berkovitz: Necessary conditions for optimal strategies in a class of differential games and control problems. SIAM J. Control *5* (1967), 1, 1—24.

[10] B. N. Pšeničnyj: Structure of differential games. Dokl. Acad. Nauk SSSR *184* (1969), 2, 285—287. In Russian.

[11] A. V. Balakrishnan, L. W. Neustadt (eds.). Mathematical Theory of Control. Academic Press, New York 1967.

[12] Y. C. Ho, G. Leitmann (eds.): Proceedings of the First International Conference on the Theory and Applications of Differential Games. Amherst, Mass. 1969.

[13] A. Blaquière, F. Gérard, G. Leitmann: Quantitative and Qualitative Games. Academic Press, New York 1969.

[14] N. N. Krasovskij: Pursuit-Evasion Game Problems. Nauka, Moscow 1970. In Russian.

[15] A. Friedman: Differential Games. Wiley-Interscience, New York 1971.

[16] P. Varaiya, J. Lin: Existence of saddle points in differential Games. SIAM J. Control *7* (1969), 1, 141—157.

[17] R. J. Elliott, N. J. Kalton, L. Marcus: Saddle points for linear differential games. SIAM J. Control *11* (1973), 1, 100—112.

[18] T. Parthasarathy, T. E. S. Raghavan: Existence of saddle points and Nash equilibrium points for differential games. SIAM J. Control *11* (1975), 5, 977—980.

[19] D. J. Wilson: Mixed strategy solutions for quadratic games. J. Optimization Theory Appl. *13* (1974), 3, 319—333.

[20] H. W. Kuhn, G. P. Szegö (eds.): Differential Games and Related Topics. North-Holland, Amsterdam 1971.

[21] A. Blaquière (ed.): Topics in Differential Games. North-Holland, Amsterdam 1973.

[22] J. D. Grote (ed.): Theory and Applications of Differential Games. Reidel, Dordrecht 1975.

[23] P. Hagedorn, H. W. Knobloch, G. J. Olsder (eds.): Differential Games and Applications. Springer-Verlag, Berlin 1977.

[24] E. Cockayne: Plane pursuit with curvature constraints. SIAM J. Appl. Math. *15* (1967), 6, 1511—1517.

[25] A. W. Merz: The game of two identical cars. J. Optimization Theory Appl. *9* (1972), 5, 324—343.

[26] J. P. Marec, Nguyen Van Nhan: Two-dimensional pursuit-evasion game with penalty on turning rates. J. Optimization Theory Appl. *23* (1977), 2, 305—345.

[27] N. N. Krasovskij, A. I. Subbotin: Positional Differential Games. Nauka, Moscow 1974. In Russian.

[28] Y. C. Ho, A. E. Bryson, Jr., S. Baron: Differential games and optimal pursuit-evasion strategies. IEEE Trans. Automatic Control *AC-10* (1965), 4, 385—389.

[29] Y. C. Ho: Differential games, dynamic optimization, and generalized control theory. Survey paper. J. Optimization Theory Appl. *6* (1970), 3, 179—209.

[30] D. Tabac: Numerical solutions of differential game problems. Int. J. Systems Sci. *6* (1975), 6, 591—599.

[31] A. W. Starr: Computation of Nash equilibria for nonlinear nonzero-sum differential games. See [12], IV 13—18.

[32] V. H. Quintana, E. J. Davison: Two numerical techniques to solve differential game problems. Int. J. Control *16* (1972), 3, 465—474.

[33] V. Gourishankar, A. Salama: A technique for solving a class of differential games. Int. J. Control *15* (1972), 3, 529—539.

[34] J. Doležal: Differential Dynamic Programming Approach to the Numerical Solution of Differential Games. Research Report No. 372, Institute of Information Theory and Automation, Prague 1971. In Czech.

[35] R. G. Graham, C. T. Leondes: Solution of differential games by quasilinearization. Int. J. Systems Sci. *5* (1974), 6, 523—544.

[36] J. Hartung: Penalty-methods for optimal control problem and differential games. In "Optimization and Optimal Control Conference", Lecture Notes in Mathematics No. 477, Springer-Verlag, Berlin 1975. In German.

[37] L. W. Taylor, Jr.: Application of the epsilon technique to a realistic optimal pursuit-evasion problem. J. Optimization Theory Appl. *15* (1975), 6, 685—702.

[38] N. Leffler, C. T. Leondes: An algorithm for the solution of differential games with bounds on the state and control variables — I, II. Comput. Elect. Engng. *2* (1975), 1, 67—81, 83—94.

[39] B. N. Grevcov, I. O. Mel'c: Computation of improving variations for minimax problems without discontinuities in control. Avtomat. i Telemech. (1978), 3, 5—13. (In Russian.)

[40] G. M. Anderson: A near-optimal closed-loop solution method for nonsingular zero-sum differential games. J. Optimization Theory Appl. *13* (1974), 4, 303—318.

[41] G. M. Anderson: A transition matrix method for generating near-optimal closed-loop solutions to non-linear zero-sum differential games. Int. J. Systems Sci. *7* (1976), 5, 529—543.

[42] B. Järmark: Near-Optimal Closed-Loop Strategy for Aerial Combat Games. Research Report TRITA-REG-7602, The Royal Institute of Technology, Stockholm 1976.

[43] S. M. Roberts, J. S. Shipman: Two-Point Boundary-Value Problems: Shooting Methods. American Elsevier, New York 1972.

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

[45] M. A. Connor, M. Vlach: A new augmented penalty function technique for optimal control problems. J. Optimization Theory Appl. *21* (1977), 1, 39—49.

[46] A. E. Bryson, Y. C. Ho: Applied Optimal Control, Blaisdell, Waltham, Mass. 1969.

[47] 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. Czechoslovakia.*