Faculty Working Papers

ON A CERTAIN PENALTY METHOD IN OPTIMAL CONTROL AND DIFFERENTIAL GAMES

Ronald J. Stern

#74

College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

FACULTY WORKING PAPERS

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

November 21, 1972

ON A CERTAIN PENALTY METHOD IN OPTIMAL CONTROL AND DIFFERENTIAL GAMES

Ronald J. Stern

#74

To be presented at the International Management Science Conference, Tel-Aviv, Israel, June, 1973

ON A CERTAIN PENALTY METHOD IN OPTIMAL CONTROL AND DIFFERENTIAL GAMES

by

Ronald J. Stern*

ABSTRACT

The penalty technique introduced in [6] is applied to linear-quadratic optimal control problems, N-Person non-zero sum differential games, efficient point problems in linear control problems with multiple quadratic criteria, and to bicriterion optimal control problems. In all these cases the reason for applying the technique is to overcome the computational difficulty introduced by the imposition of a pointwise magnitude restraint on the feasible controls. Additional details are available in [6] - [9].

*Department of Business Administration, University of Illinois at Urbana.

ON A CERTAIN PENALTY METHOD IN OPTIMAL CONTROL AND DIFFERENTIAL GAMES

1. Introduction. In recent works a new penalty technique has been employed to derive procedures for computing optimal open loop controls for two person zero sum linear-quadratic differential games [6] and N person non-zero sum linear-quadratic differential games [7]. The method also has been applied to compute open loop solutions to bicriterion control problems and efficient points for multicriteria problems in control [8], [9]. "Open loop" solutions, it will be recalled, are solutions which are not of the feedback type; that is, they are functions of time only. In each of the problems mentioned above, the penalty method is employed to overcome computational difficulties which arise from the imposition of a pointwise magnitude restraint on the feasible controls.

In section 2 we give a detailed exposition of the technique, for the case of a general linear-quadratic optimal control problem with pointwise magnitude restraints. In subsequent sections we outline the application of the technique to the various problems mentioned in the preceeding paragraph, by referring to section 2.

2. Application to Linear-Quadratic Optimal Control. Consider the following linear system of differential equations in R^m:

$$(2.1) \quad \dot{x} = A(t)x + B(t)u \qquad (to \leqq t \leqq To)$$

with initial condition

$$(2.2) \quad x(to) = x_o$$

Here $A(t)$ and $B(t)$ are continuous $(m \times m)$ and $(m \times s)$ matrices on the compact time interval $[to, To]$. Feasible controls $u = u(t)$ are Lebesque-

measurable R^s-valued functions which almost everywhere on $[t_o,T_o]$ take values in U, the closed unit ball with center 0 in R^s. We should denote this class of feasible controls by $\mathcal{U}$.

For each $u \in \mathcal{U}$ a unique solution to (2.1) - (2.2) is determined by

$$(2.3) \quad x(t) = S(T_o,t)x_o + \int_{t_o}^{t} S(t,s)\, B(s)u(s)ds$$

where S is the fundamental solution of $\dot{x} = A(t)x$ (see e.g., [1].

We now introduce an objective function of quadratic type:

$$(2.3)\ J(u) = \left\langle x(T_o) - \xi\, ,\, W[x(T_o) - \xi\,] \right\rangle + \int_{t_o}^{T_o} \left\langle \tilde{x}(t) - C(t)x(t), Q(t)\, [\tilde{x}(t) - C(t)x(t)] \right\rangle dt - \int_{t_o}^{T_o} \left\langle u(t), u(t) \right\rangle dt$$

Here ξ is a fixed vector in R^m, W is a constant symmetric (mxm) matrix, $\tilde{x}(t)$ is a continuous R^m - valued junction on $[t_o,T_o]$, C(t) is a continuous (mxm) matrix, and Q(t) is a continuous symmetric (mxm) matrix on $[t_o,T_o]$. x(t) denotes the solution of (2.1) - (2.2) corresponding to U.

The following lemma is required. The proof, which is omitted, is of a computational nature and makes use of bounds of the various parameters in (2.1) - (2.3).

<u>Lemma 2.1.</u> There exists $N > 0$ <u>such that</u> $T_o - t_o \leqq N$ <u>implies the following:</u>

(i) J(u) is a <u>strictly concave</u> functional on $L^{2,s}(t_o,T_o)$, the space of measurable functions u satisfying $\int_{t_o}^{T_o} |u(t)|^2\, dt < \infty$, where $|\,|$ is the Euclidean norm.

(ii) $J(u)$ is bounded above over $L^{2,s}(t_o,T_o)$.

Now we introduce the following pair of problems :

P_1 maximize $J(u)$

subject to $u \in L^{2,s}(t_o,T_o)$

P_2 maximize $J(u)$

subject to $u \in \mathcal{U}$.

In view of Lemma 2.1, the condition

$$(2.4) \qquad \frac{d}{d\varepsilon} J(\bar{u} + \varepsilon v)\Big|_{\varepsilon=o} = 0 \quad \text{for all } v \in L^{2,s}(t_o,T_o)$$

is sufficient for $\bar{u}$ to be the (unique) solution of P_1. (2.4) gives rise to the following integral equation:

$$(2.5) \qquad \bar{u}(s) = B^*(s)\, S^*(T_o,s)\, W\, [\bar{x}(T_o) - \xi\,] + \int_s^{T_o} B^*(s)\, S^*(t,s)\, C^*(t)\, Q(t)\, [\tilde{x}(t) - \bar{x}(t)]\, dt.$$

Here $\bar{x}(t)$ is expressed by (2.3) in terms of $\bar{u}(t)$. We write (2.5) as follows:

$$(2.6) \qquad \bar{u} = A\bar{u}$$

where $A: C^{o,s}[t_o,T_o] \longrightarrow C^{o,s}[t_o,T_o]$ (Here we denote by $C^{o,s}[t_o,T_o]$ the space of continuous R^s-valued functions on $[t_o,T_o]$). It is easily shoen (see e.g.[2] that if T_o-t_o is sufficiently small then A is a contraction.

We therefore have

<u>Theorem 2.1</u> <u>There exists</u> $\bar{N} > 0$ <u>such that</u> $T_o-t_o < \bar{N}$ <u>implies</u> P_1 <u>has a unique solution</u>.

This solution can be identified as the uniform limit of a sequence of successive approximations using (2.5). The case for problem P_2 is not as straightforward —— we now turn our attention ot it. To this end, we first introduce a new payoff functional.

$$(2.7)\quad J^k(u) = J(u) - \int_{t_o}^{T_o} |u(t)|^{2k}\, dt$$

where k is a positive integer. The general penalty method discussed in this paper uses the (computable) optimal payoffs of an unconstrained problem with (2.7) as payoff in order to approximate the optimal payoff of P_2, as $k \to \infty$. We shall denote by P_k the control problem with payoff given by (2.7) and with the only requirement for the feasibility of u being membership in $L^{2k,s}(t_o,T_o)$.

We defer the proof of the following theorem until later in this section.

<u>Theorem 2.2.</u> <u>There exists</u> $\bar{\bar{N}} > 0$ <u>such that</u> $T_o - t_o \leqq \bar{\bar{N}}$ <u>implies</u> P_k <u>has a unique solution</u>.

We now define the following map of R^2 into R.

$$\varphi^k(v) = \begin{cases} 0 & \text{if } |v| \leqq k^{-1/2k} \\ |v|^{2k} & \text{otherwise} \end{cases}$$

We also define the following payoff functional:

$$(2.8)\quad J^{\varphi^k}(u) = J(u) - \int_{t_o}^{T_o} \varphi^k(u(t))\, dt$$

Note that

$$(2.9)\quad |J^k(u) - J^{\varphi^k}(u)| \leqq \frac{T_o - t_o}{k}$$

for any $u \in L^{2k,s}(t_o,T_o)$.

Let us denote the solution of P_k by u^k. Proof of the following result is of a computational nature (using Hölder's inequality) and will be omitted.

Lemma 2.2. $T_o - t_o \leq \bar{\bar{N}}$ implies that there is a real $Q > 0$ such that

$$\sup_k \int_{t_o}^{T_o} |u^k(t)|^{2k} \, dt \leq Q.$$

Let $\mathcal{U}^k$ be the set of vectors in R^s with Euclidean length $\leq 1 - k^{1/2k}$ and let $\rho_k(\cdot)$ denote the Euclidean distance in R^s from $\mathcal{U}^k$. Lemma 2.2 implies

$$(2.10) \quad \int_{t_o}^{T_o} \rho_k^2(U^k(t))\, dt \leq \frac{Q}{k}$$

from Theorem 2.2 and (2.9) we have, for any $u \in L^{2ks}(t_o,T_o)$, the following:

$$(2.11) \quad J(u) \geq J^{\varphi k}(u^k) - \frac{T_o - t_o}{k}$$

for each positive integer k.

Now let $\bar{P}_k$ denote the problem with objective function J(u), but where admissible controls are those Lebesgue measurable functions which almost everywhere on $[t_o,T_o]$ are valued in $\mathcal{U}^k$.

We now prove the following lemma:

Lemma 2.3. Let $T_o - t_o \leq \bar{\bar{N}}$. Then there exists a real $D > 0$ such that the following is true: for each positive integer k there is a control $\hat{u}^k$ such that:

$$(2.12) \quad J(\hat{u}^k) \geq J(u) - D(k^{-1/2} + 1 - k^{-1/2k}) \quad \frac{T_o - t_o}{k}$$

for any control u feasible for P_k.

Proof. Let v be any vector in R^s. Let

$$\hat{v} = \begin{cases} v & \text{if } v \in U^k \\ k^{-1/2k} \dfrac{v}{|v|} & \text{otherwise} \end{cases}$$

By a simple argument we have that $\hat{u}(t)$ is Lebesque measurable for any Lebesque measurable control u(t).

A routine calculation (additional details are to be found in [6]) yields (2.12). (2.10) and (2.11) are utilized here.

We now have the following:

Theorem 2.3. If $T_o - t_o \leqq \bar{\bar{N}}$ then

$$J(u^k) \longrightarrow \sup_{u \in} J(u)$$

The proof is found (subject to minor changes) in [7]. The weak topology is employed in the argument, which is similar in spirit to a result in [5], p. 209.

We turn now to the proof of Theorem 2.2.

Proof of Theorem 2.2. By arguments similar to those in the proof of Lemma 2.1 we have negativity of the second Gateoux differential of $J^k(u)$ if $T_o - t_o$ is sufficiently small. (See e.g., [2] for proof). The condition of stationarity (i.e., $J^k(u)$ have zero first Gateaux derivative) is then both a necessary and sufficient condition for optimality. The condition is given by the following integral equation (see [2] for similar equations):

$$(2.13) \qquad u(s) + 2k|u(s)|^{2k-2}u(s)$$
$$= B^*(s)\, S^*(T_o,s)\, W\, [x(T_o) - \xi\,]$$
$$+ \int_{t_o}^{s} B^*(s)\, S^*(t,s)\, C^*(t)\, Q(t)\, [\tilde{x}(t) - x(t)]\, dt.$$

(2.13) is next rewritten as follows:

$$(2.14) \quad M^k(u(s)) = T^k[M^k(u(s))]$$

Here M^k: $R^s \longrightarrow R^s$ is given by

$$M^k(v) = 2v + 2k|v|^{2k-2}v.$$

M^k has an inverse given by

$$M^{-1}(w) = \frac{w}{2+2k[r_k(|w|)]^{2k-2}}$$

where $r_k(|w|)$ is the unique real proof of the polynomial $2kx^{2k-1} + 2x - |w|$. Note that the expressability of (2.13) in the form (2.14) is dependent upon the invertibility of M^k.

Each T^k is a contraction when $T_o - t_o$ is small enough, say $\bar{\bar{N}}$; this by Theorem 4.2 in [6]. This completes the proof of the theorem.

(2.13) can be solved computationally. A procedure is given in section 5 of [6]. This procedure circumvents the problem that r_k has no explicit form when $k > 2$.

3. Application to Differential Games. The outline presented in this section summarizes results of both references [6] and [7].

The governing dynamics are given by the following system in R^m:

$$(3.1) \quad \dot{x} = A(t)x + \sum_{i=1}^{N} B_i(t)\, u_i \qquad (t_o \leq t \leq T_o)$$

with initial condition

$$(3.2) \quad x(t_o) = x_o\ .$$

We define N cost functionals of quadratic type:

$$(3.3)\quad J_i(u_1,u_2,\ldots, u_N) = \left\langle x(T_o) - \xi_i,\ W_i[x(T_o) - \xi_i]\right\rangle$$

$$+ \int_{t_o}^{T_o} \left\langle \tilde{x}_i(t) - C_i(t)x(t),\ Q_i(t)\,[\tilde{x}_i(t) - C_i(t)x(t)]\right\rangle dt$$

$$+ \int_{t_o}^{T_o} \left\langle u_i(t), u_i(t)\right\rangle dt, \qquad i = 1,2,\ldots,N$$

Assumptions on the parameters of (3.1) - (3.3) parallel those made in section 2. Admissible controls u_i for the ith player are Lebesgue measurable functions which almost everywhere take values in the unit ball of R^i.

For the above game, denoted G, we seek an open loop Nash equilibrium; that is, a vector of feasible controls $(\bar{u}_1, \bar{u}_2,\ldots,\bar{u}_N)$ such that

$$(3.4)\quad J_i(\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_N) \geqq J_i(\bar{u}_1, \bar{u}_2,\ldots, \bar{u}_{i-1}, u_i, \bar{u}_{i+1},\ldots\bar{u}_N)$$

for any admissible u_i, $i=1,2,\ldots,N$.

An approximating game, denoted G_k, is introduced, for k a positive integer. The payoff functionals for this differential game are

$$(3.5)\quad J_i^k(u_1,u_2,\ldots, u_N) = J_i(u_1,u_2,\ldots,u_N) - \int_{t_o}^{T_o} |u_i(t)|^{2k}\, dt$$

for i = 1, 2,....N.

An analog of Theorem 2.2, the proof of which may be found in [7], is the following:

<u>Theorem 3.1.</u> <u>If</u> T_o-t_o <u>is sufficiently small then</u> G^k <u>has a unique open loop Nash equilibrium</u>.

An analog of Lemma 2.3 is now given:

Lemma 3.1. If $T_o - t_o$ is sufficiently small then there exists a constant $D > 0$ such that the following is true: For each positive integer k there is a vector of controls $(\hat{u}_1^k, \hat{u}_2^k, \ldots, \hat{u}_N^k)$ feasible for G_k such that

$$(3.6) \quad J_i(\hat{u}_1^k, \hat{u}_2^k, \ldots, \hat{u}_N^k) \geqq J_i(\hat{u}_1^k, \hat{u}_2^k, \ldots \hat{u}_{i-1}^k, u_i, \hat{u}_{i+1}^k, \ldots, \hat{u}_N) - D(k^{-1/2} + 1 - k^{-1/2k}) - \frac{T_o - t_o}{k}$$

where $\bar{G}_k$ is the game with the same payoffs as G but where feasible controls u_i are Lebesque measurable and are valued almost everywhere in the ball of radius $k^{-1/2k}$ in R^i.

By properties of the weak topology it is shown in [7] that G has an open loop Nash equilibrium which can be represented as the weak limit of a subsequence of $(\hat{u}_1^k, \hat{u}_2^k, \ldots, \hat{u}_N^k)$. It is shown there that the equilibrium costs of G are computable by the method of [6], section 5.

4. Application to Multicriteria Control

$$(4.1) \quad \dot{x} = A(t) + B(t)u \quad (t_o \leq t \leq T_o)$$

with initial condition

$$(4.2) \quad x(t_o) = x_o$$

and n criterion functions

$$(4.3) \quad f_i(u) = < x(T_o) - \xi_i, W_i[x(T_o) - \xi_i] > + \int_{t_o}^{T_o} < \tilde{x}_i(t) - C_i(t)x(t), Q_i(t)[\tilde{x}(t) - C_i(t)x(t)] > dt - \int_{t_o}^{T_o} < u(t), u(t) > dt, \ (1 \leq i \leq n),$$

where assumptions on (4.1) - (4.3) parallel those made in section 1.

An <u>efficient point</u> u^o over a class of controls $\mathcal{C}$ is a control $u^o \in \mathcal{C}$ such that for no other $u \in \mathcal{C}$

$$f_i(u) \geq f_i(u^o) \qquad 1 \leq i \leq n$$

with at least one strict inequality.

The class of controls $\mathcal{C}$ in the above definition will be taken here as one of the following two classes.

$L^{2,s}(t_o T_o)$ - <u>the space of controls</u> u <u>with</u> $\int_{t_o}^{T_o} |u(t)|^2 dt < \infty$, where $| \; |$ denotes the Euclidean norm.

$\mathcal{U}$ - <u>the class of measurable controls with</u> $|u(t)| \leq 1$ <u>a.e. on</u> $[t_o, T_o]$.

The following two efficient point problems will be studied here:

E_1 <u>Find all efficient points over</u> $L^{2,s}(t_o, T_o)$.

E_2 <u>Find all efficient points over</u> $\mathcal{U}$.

For the bicriterion case n = 2 we study two other optimal control problems

M_1 <u>Maximize</u> $\min\{f_1(u), f_2(u)\}$ <u>subject to</u> $u \in L^{2,s}(t_o, T_o)$

M_2 <u>Maximize</u> $h[f_1(u), f_2(u)]$ <u>subject to</u> $u \in \mathcal{U}$,

where $h: R^2 \to R$ is continuous and non-decreasing in each of its arguments on the non-negative orthant R^2_+, and quasiconcave over the interior of R^2_+. We further assume that

$|u(t)| \leq 1$ a.e. on $[t_o, T_o] \Rightarrow f_i(u) > 0$, $i = 1,2$, which is guaranteed, for example, if W_i are positive definite and $|x_o|$ is sufficiently large.

We observe that the objective function of M_1 is a special case of the objective function $h(x,y)$ of M_2. Other examples are

$$h(x,y) = x^{\beta} y \quad ; \qquad \beta > 0$$

or

$$h(x,y) = c_1 x^{\beta_1} + c_2 y^{\beta_2};\ c_1, c_2, \beta_1, \beta_2 > 0.$$

(For other examples see e.g., [3] p. 40).

In what follows let the vector

$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ satisfy

(4.4) $\sum_{i=1}^{n} \alpha_i = 1$ and $\alpha_i \geq 0 \quad 1 \leq i \leq n.$

For each such α we define the following single criterion analog of E_1:

P_1^α Maximize $\sum_{i=1}^{n} \alpha_i f_i(u)$ subject to $u \in L^{2,s}(t_o, T_o)$

Similarly to Theorem 2.1 we have

Theorem 4.1. If $T_o - t_o$ is sufficiently small then P_1^α has a unique open loop solution.

The problems E_1 and P_1^α are related by following Theorem.

Theorem 4.2. If $T_o - t_o$ is sufficiently small

Then

(i) If the vector α satisfying (4.4) is positive (i.e., if $\alpha_i > 0$, $1 \leq i \leq n$), then the solution u^α of P_1^α is an efficient point over $L^{2,s}(t_o, T_o)$.

(ii) If u^o is an efficient point over $L^{2,s}(t_o, T_o)$ then for some α^o satisfying (4.4), u^o is the solution of $P_1^{\alpha^o}$.

Proof. (i) is obvious. (ii) is proved as in the finite dimensional case, see e.g., [4], Section 7.4.

A method for approximating a solution of P_1^α is given by Theorem 4.2 namely, u^α is the uniform limit of a sequence of successive approximations.

We now define a new problem:

P_2^α Maximize $\sum_{i=1}^{n} \alpha_i f_i(u)$ subject to $u \in \mathcal{U}$.

Although $\mathcal{U}$ is not a compact subset of $L^{2,s}(t_o,T_o)$ the consistency of $\{P_2^\alpha\}$ is guaranteed by the following Theorem.

Theorem 4.3. If T_o-t_o is sufficiently small then each problem P_2^α has a unique solution.

See [8] for proof.

The relations between E_1 and P_1^α, studied in Theorem 4.2, hold also for E_2 and P_2^α.

Theorem 4.4. If T_o-t_o is sufficiently small then (i) If the vector α satisfying (4.4) is positive, then the solution of P_2^α is an efficient point over $\mathcal{U}$.

(ii) If u^o is an efficient point over $\mathcal{U}$ then u^o is the solution of P_2^α for some α_o satisfying (4.4).

The penalty method outlined in section 2 is applicable to problem P_2^α.

The following results are proven similarly to the corresponding results in [3].

Lemma 4.1. If $T_o-t_o \leq N$, then M_1 and M_2 have optimal solutions, and at least one solution of each is efficient.

Theorem 4.5. Let $T_o-t_o < N$.
Then the following function of α

$$h[f_1(u^\alpha),\ f_2(u^\alpha)]$$

is unimodal on [0,1].

Using the customary search techniques for finding the supremum of a unimodal function, Theorems 4.1, 4.2, Lemma 4.1 and Theorem 4.5 constitute a procedure for approximately solving M_1, while Theorems 4.3, 4.4, Lemma 4.1 and Theorem 4.5 similarly constitute a procedure for approximately solving M_2.

REFERENCES

1. E. Coddington and N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill, 1955.

2. A. Friedman, Differential Games, Wiley-Interscience, 1971.

3. A. M. Geoffrion, "Solving Bicriterion Mathematical Programs," Operations Research, 15, 1967, p. 39-54.

4. S. Karlin, Mathematical Methods and Theory in Games, Programming and Economics, Vol. I, Addison-Wesley, 1959.

5. E. Lee and L. Markus, Foundations of Optimal Control Theory, Wiley, 1967.

6. R. Stern, "Open Loop Approximation of Differential Games," SIAM J. Control, 11, 1973.

7. R. Stern, "Open Loop Nash Equilibrium of N-Person Non-Zero Sum Linear-Quadratic Differential Games with Magnitude Restraints," (Forthcoming).

8. R. Stern, and A. Ben-Israel, "Multiple Criteria in Optimal Control," (Forthcoming).

9. R. Stern and A. Ben-Israel, "On Linear Optimal Control Problems with Multiple Quadratic Criteria," Proceedings of the Seminar on Multicriteria Decision Making held in Columbia, South Carolina, 1972.