International Journal of Control
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713393989

Existence and derivation of optimal affine incentive schemes for Stackelberg games with partial information: a geometric approach

Ying-Ping Zheng[a]; Tamer Basar[a]
[a] Decision and Control Laboratory, Coordinated Science Laboratory, University of Illinois, Urbana, Illinois 61801, U.S.A.

To cite this Article Zheng, Ying-Ping and Basar, Tamer(1982) 'Existence and derivation of optimal affine incentive schemes for Stackelberg games with partial information: a geometric approach', International Journal of Control, 35: 6, 997 — 1011
To link to this Article: DOI: 10.1080/00207178208922667
URL: http://dx.doi.org/10.1080/00207178208922667

INT. J. CONTROL, 1982, VOL. 35, NO. 6, 997–1011

Existence and derivation of optimal affine incentive schemes for Stackelberg games with partial information: a geometric approach†

YING-PING ZHENG‡§ and TAMER BAŞAR‡

Through a geometric approach, it is shown that a sufficiently large class of incentive (Stackelberg) problems with perfect or partial dynamic information admits optimal incentive schemes that are affine in the available information. As a byproduct of the analysis, explicit expressions for these affine incentive schemes are obtained, and the general results are applied to two different classes of Stackelberg game problems with partial dynamic information.

1. Introduction

The Stackelberg solution concept, first introduced by H. von Stackelberg (1934, 1952) for static games, and then extended and applied to dynamic games with open-loop information for the leader (Chen and Cruz 1972, Simaan and Cruz 1973 a, b, Cruz 1978), has recently attracted considerable attention in both the control and economics literature (Ho *et al.* 1980, Ho *et al.* 1981, Başar and Selbuz 1979, Papavassilopoulos and Cruz 1979, 1980, Tolwinski 1981, Başar 1982 a, Groves and Loeb 1979, Hurwicz and Shapiro 1978, Jennergren 1980). This recent research activity on Stackelberg games pertains to the case when the leader has access to perfect or partial dynamic information which involves the other decision maker's (follower's) past actions, in which case the problem becomes tractable only through use of some indirect methods. The main objective of the leader, in such problems, is to determine a strategy, compatible with his available information, which will force the follower to a certain behaviour which is considered to be mostly preferrable by the leader (under the several informational or structural constraints imposed on the problem). In the economics literature, such problems are known as incentive design problems, and the ' best ' policy (ies) of the leader is (are) known as optimal incentive scheme(s) (or strategies). Solutions of these incentive (Stackelberg) problems involve, in general, two steps, namely (i) derivation of an attainable bound on the leader's performance, and (ii) determination of an incentive strategy that attains this bound.

A recent reference (Başar 1982 a) was addressed primarily to the former one, and developed a general approach that led to attainable bounds for a fairly general class of Stackelberg games under both perfect and partial dynamic information. The main conclusion of Başar (1982 a) was that, even though

Received 29 December 1981.

† The research of T. Başar was supported in part by the U.S. Air Force under Grant AFOSR-78-3633 and in part by the National Science Foundation under Grant ECS-79-19396.

‡ Decision and Control Laboratory, Coordinated Science Laboratory, University of Illinois, 1101 W. Springfield Ave., Urbana, Illinois 61801, U.S.A.

§ On leave of absence from the Institute of Automation, Academia Sinica, Beijing, Peoples Republic of China.

0020-7179/82/3506 0997 $04·00

the original problem is dynamic and constitutes a highly non-trivial optimization problem if direct methods are used, the attainable performance bound can be obtained by solving a series of non-dynamic (open-loop) optimization problems.

The present paper addresses the second question raised above, namely, determination of 'optimal' incentive schemes (strategies) that achieve the bounds determined in Başar (1982 a), under both perfect and partial dynamic information; in the latter case, the observation of the leader providing the partial dynamic information is taken to be linear in the decision variable of the follower. Using a geometric approach, it is shown that for a sufficiently general class of incentive (Stackelberg) problems, and under rather mild conditions (which are precisely delineated in the paper), there exist optimal incentive schemes that are affine in the available information. Derivation of this set of 'appealing' incentive strategies are also discussed in the paper, together with two examples which illustrate the general results.

The paper is organized as follows. Section 2 is devoted to the incentive problem with perfect information for the leader. Section 3 deals with the case when the leader has only (linear) partial information on the follower's actions; a projection transformation is used to convert the problem to a new one which has perfect information for the leader, so that the results of § 2 can be used. In § 4, the results are applied to (i) linear-quadratic incentive problems defined on general Hilbert spaces and with partial information, and (ii) linear-quadratic differential game problems with sampled-data information, so as to obtain (explicitly) in each case the optimal affine incentive strategies for the leader. Finally, § 5 is devoted to some discussion on certain aspects of the general approach presented here and on possible avenues for future research.

2. Stackelberg games with perfect dynamic information for the leader

2.1. *Formulation*

Adopting the notation and terminology of Başar (1982 a), we have a two-person deterministic dynamic game problem in normal form, described by the cost functionals $J_1(\gamma_1, \gamma_2)$ and $J_2(\gamma_1, \gamma_2)$, for Player 1 (the leader) and Player 2 (the follower), respectively, where the strategies γ_1 and γ_2 belong to *a priori* determined strategy spaces Γ_1 and Γ_2, respectively. Let the decision variables of the leader and the follower be denoted by $u \in U$ and $v \in V$, respectively, which may also be considered as open-loop strategies, since the dynamic game is deterministic and the initial state (under an appropriate interpretation) is assumed to be given and fixed. Here, U (respectively, V) is the decision space of the leader (respectively, the follower), and both U and V are assumed to be appropriate Hilbert spaces. Let $J_i(u, v)$ also denote the cost functional of Player i over the product Hilbert space $U \times V$.

Now, let us further assume that:

(i) The follower's objective functional $J_2(u, v)$ is *strictly convex* on $U \times V$, i.e. for any real number c, the set $\{(u, v) | J_2(u, v) \leqslant c\}$ is strictly convex in $U \times V$.

(ii) The complete detectability condition (Assumption A of Başar (1982 a)) is satisfied, i.e. the leader can observe the follower's action(s) perfectly.

(iii) The leader is in a position to announce and enforce an *incentive scheme* (*strategy*) $\gamma_1 \in \Gamma_1$, which is a Borel-measurable mapping of V into U.

The problem faced by the leader is to find an incentive scheme which, by also taking into account rational (cost-minimizing) responses of the follower, leads to a most favourable performance for the leader. This performance may be defined as the global minimum value of J_1 (assuming that it exists)

$$J_1{}^t = \min_{(u,\, v) \in U \times V} J_1(u, v) = J_1(u^t, v^t) \tag{1}$$

which corresponds to some specific choices of $u \in U$ and $v \in V$ (in this case, $u = u^t$ and $v = v^t$) ; or, more generally, there may exist some pair (denoted again (u^t, v^t)) which is chosen according to some criterion and is considered to be most favourable to the leader. Then, the general question is : For a given pair $(u^t \in U,\ v^t \in V)$, does there exist an ' optimal ' incentive strategy $\gamma_1 \in \Gamma_1$ under which the best (J_2-minimizing) policy for the follower has the open-loop value v^t, and a corresponding decision value for the leader is u^t ? This question is addressed in the next section. Note that, in this formulation, the structure of the cost function of the leader is, in general, irrelevant, and such an incentive strategy is indeed a Stackelberg strategy for the leader under the given information structure.

2.2. *Optimal affine incentive strategies—their existence and computation*

We call an incentive strategy $\gamma_1 \in \Gamma_1$ *affine*, if it is totally characterized by an affine relation of the form

$$u = \gamma_1(v) = u_0 - Qv \tag{2}$$

where Q is a linear operator mapping V into U. If the desired ' open-loop ' solution by the leader is (u^t, v^t), this takes the form

$$u = \gamma_1(v) = u^t - Q(v - v^t) \tag{3}$$

which clearly has the open-loop value $u = u^t$ whenever the follower can be forced to the decision value $v = v^t$. Affine incentive schemes of the form (3) (for varying $Q : V \to U$—which constitute a proper subset of Γ_1) are particularly appealing because they are structurally simple, can easily be computed and implemented, and they do not have the explicit threat property that other discontinuous Stackelberg strategies possess (i.e. they provide rather soft constraints for the follower's minimization problem). Moreover, for a general class of problems, affine incentive schemes induce the desired behaviour on the follower, as will be shown below.

Towards this end, let us first note that the set

$$\Omega_t \triangleq \{(u, v) \in U \times V \,|\, J_2(u, v) \leqslant J_2(u^t, v^t)\}$$

is strictly convex, with (u^t, v^t) as a boundary point. Hence, there must exist a supporting hyperplane Π_{Ω_t} passing through (u^t, v^t), provided that Ω_t contains an interior point (Luenberger 1969, Balakrishnan 1976). The latter condition does not impose any restriction on the problem, because if an interior point does not exist, (u^t, v^t) globally minimizes J_2, and hence $\gamma_1(v) = u^t$ becomes trivially an optimal incentive strategy. If we can find an operator Q such that the submanifold defined by (3) lies on Π_{Ω_t}, then the set of admissible decision pairs (u, v), defined by the incentive scheme $u = u^t - Q(v - v^t)$, has

only the unique common point (u^t, v^t) with Ω_t. Therefore, the only solution to the follower's optimization problem is exactly v^t; in other words, the strategy $u = u^t - Q(v - v^t)$ solves the incentive problem.

In order to ensure that such an operator Q exists, we further assume that

(iv) $J_2(u, v)$ is Fréchet-differentiable on $U \times V$.

Then, the equation of the supporting hyperplane Π_{Ω_t} may be written as

$$\langle \nabla_u J_2(u^t, v^t), u - u^t \rangle + \langle \nabla_v J_2(u^t, v^t), v - v^t \rangle = 0 \tag{4}$$

where $\nabla_u J_2(u^t, v^t) \in U^*$ is the gradient of J_2 with respect to u, evaluated at (u^t, v^t), and $\nabla_v J_2(u^t, v^t) \in V^*$ is the gradient of J_2 with respect to v, evaluated also at (u^t, v^t).

Throughout this paper, when the superscript * is used in conjunction with a linear space (such as U, V) or a linear operator (such as Q), it stands for the 'adjoint'.

Lemma 1

Let X and Y be Banach spaces with given elements $x_0 \in X$, $y_0 \in Y$, $x_0 \neq 0$. Then there exists a continuous linear operator Q such that $Qx_0 = y_0$.

Proof

Define a functional $f(\cdot)$ on the subspace $\{\alpha x_0, \alpha \in R\} \subset X$

$$f(\alpha x_0) = \alpha, \quad \|f\| = 1/\|x_0\| < \infty \tag{5}$$

By the Hahn–Banach theorem (Luenberger 1969), f can be extended to the whole space X, and still retain its original norm. Denote this new functional again by $f(\cdot)$, and introduce the operator $Q(x) = f(x)y_0$, which maps X into Y. It is obvious that Q is linear and bounded (thereby continuous), with $\|Q\| = \|y_0\|/\|x_0\|$. Moreover, $Q(x_0) = f(x_0)y_0 = y_0$, which completes the proof of the lemma. □

Remark 1

When X is a Hilbert space, we may define

$$f(x) = \frac{\langle x_0, x \rangle}{\|x_0\|^2}, \quad Q(x) = \frac{\langle x_0, x \rangle}{\|x_0\|^2} y_0 \tag{6}$$

which is one possible choice for the operator Q. In fact, there exist many different possibilities, especially if we are operating in infinite dimensional spaces. □

By Lemma 1, and under the assumption that $\nabla_u J_2(u^t, v^t) \neq 0$, there exists a bounded linear operator $Q^* : U^* \to V^*$ such that

$$Q^* \nabla_u J_2(u^t, v^t) = \nabla_v J_2(u^t, v^t) \tag{7}$$

(Recall that $\nabla_u J_2(u^t, v^t) \in U^*$ and $\nabla_v J_2(u^t, v^t) \in V^*$.) Now consider the incentive strategy (3) with Q determined by (7). Taking the inner product of $u - u^t +$

$Q(v-v^t)$ with $\nabla_u J_2(u^t, v^t)$, we obtain

$$\begin{aligned} 0 &= \langle \nabla_u J_2(u^t, v^t), u-u^t+Q(v-v^t)\rangle \\ &= \langle \nabla_u J_2(u^t, v^t), u-u^t\rangle + \langle \nabla_u J_2(u^t, v^t), Q(v-v^t)\rangle \\ &= \langle \nabla_u J_2(u^t, v^t), u-u^t\rangle + \langle Q^*\nabla_u J_2(u^t, v^t), v-v^t\rangle \\ &= \langle \nabla_u J_2(u^t, v^t), u-u^t\rangle + \langle \nabla_v J_2(u^t, v^t), v-v^t\rangle \end{aligned}$$

which leads to the conclusion that (3) lies on the hyperplane (4) and passes through (u^t, v^t)—thereby providing a solution to the incentive problem. We have thus verified the following result.

Proposition 1

For the incentive problem of § 2.1, if the set $\Omega_t \triangleq \{(u, v) \in U \times V \mid J_2(u, v) \leqslant J_2(u^t, v^t)\}$ is strictly convex, and $J_2(u, v)$ is Fréchet-differentiable with $\nabla_u J_2(u^t, v^t) \neq 0$, there exists an optimal incentive strategy for the leader, in the form

$$u = \gamma_1(v) = u^t - Q(v-v^t) \tag{3}$$

where the linear operator $Q: V \to U$ is chosen according to (7). □

The requirement of strict convexity may be replaced by the following milder condition without affecting the validity of the result: Ω_t is convex in $U \times V$ and locally strictly convex at the boundary point (u^t, v^t). Furthermore, it is noteworthy that the requirement $\nabla_u J_2(u^t, v^t) \neq 0$ simply says that the cost functional of the follower has to be locally sensitive to changes in the value of the decision variable of the leader around the operating point (u^t, v^t).

Remark 2

In the finite-dimensional case, say $U = R^n$, $V = R^m$, $\nabla_u J_2(u^t, v^t)$ and $\nabla_v J_2(u^t, v^t)$ may be taken as column vectors of dimensions n and m, respectively, and the operator Q becomes an $n \times m$ matrix, satisfying the equation

$$[\nabla_u J_2(u^t, v^t)]' Q = [\nabla_v J_2(u^t, v^t)]' \tag{8}$$

where $'$ denotes a transpose. As long as $\nabla_u J_2(u^t, v^t) \neq 0$, there always exists such a matrix Q, one possible candidate being

$$Q = \nabla_u J_2(u^t, v^t)[\nabla_v J_2(u^t, v^t)]' / \|\nabla_u J_2(u^t, v^t)\|^2$$ □

Remark 3

Adopting the terminology used in Ho *et al.* (1980), Proposition 1 provides a rather mild sufficient condition for linear-incentive-controllability (l.i.c.) for a large class of Stackelberg (incentive) problems. It also proposes a general approach to determine the desired solution, which reduces the original problem to one of computation of gradients in Hilbert spaces and choosing an operator Q which satisfies a given linear relation. □

3. Stackelberg games with partial dynamic information for the leader

3.1. *Problem formulation*

Consider the Stackelberg game problem of § 2.1, with the only difference now being that the leader has access to only partial information on the follower's

actions, as in Başar (1982 a); but here this partial observation is assumed to be linear. More specifically, the leader's information space Y is a Hilbert space comprising observations of the form $y=Nv$, where N is a linear bounded operator mapping V into Y. Furthermore, N is not necessarily invertible and it has a non-trivial null space, Ker N; i.e. there exists some $\bar{v}\neq 0$ such that $N\bar{v}=0$. This implies that the leader cannot differentiate between two decisions v_1 and v_2 of the follower, if $v_1-v_2\in\text{Ker } N$. We finally assume, for the sake of simplicity in exposition, but without any loss of generality, that N has full range, that is, for any strongly positive linear operator $F>0$, the property $NFN^*>0$ holds.

The set Γ_1 of all admissible (incentive) strategies for the leader is now defined as the collection of all Borel-measurable mappings $\gamma_1: Y\to U$, i.e.

$$u=\gamma_1(y)=\gamma_1(Nv) \tag{9}$$

Later we shall see that a much smaller subset of Γ_1, comprising only affine incentive strategies of the form

$$u=\gamma_1(y)=u_0-Q(y-y_0) \tag{10}$$

constitutes a sufficiently rich class for the incentive problem under consideration.

To complete the description of the problem, we note that, for each fixed $\gamma_1\in\Gamma_1$, the follower will seek a solution $v_{\gamma_1}{}^*\in V$ to the minimization problem

$$\min_{v\in V} J_2(\gamma_1(Nv), v)=J_2(\gamma_1(Nv_{\gamma_1}{}^*), v_{\gamma_1}{}^*)$$

as a result of which the leader will incur a cost of

$$J_1{}^*(\gamma_1)=J_1(\gamma_1(Nv_{\gamma_1}{}^*), v_{\gamma_1}{}^*)$$

We will call an incentive strategy $\gamma'_1\in\Gamma_1$ *optimal*, and say that it solves the incentive (Stackelberg) problem, if it minimizes $J_1{}^*(\gamma_1)$.

This problem has been considered before in Başar (1982 a), where a tight lower bound (higher than the global minimum of $J_1(u, v)$ over $U\times V$) has been obtained for $J_1{}^*(\gamma_1{}^0)$. In the sequel, we show that this bound is achievable, under fairly general conditions, in the class of affine incentive strategies of the form (10), and that an optimum affine incentive strategy can be determined in explicit form.

3.2. *A related 'projected' problem with perfect information*

Following the discussion of Başar (1982 a), we first note that, for fixed $y\in Y$ and $u\in U$, a rational policy for the follower is to minimize $J_2(u, v)$ in the subspace determined by $y=Nv$. Since J_2 is strictly convex on this subspace, the optimizing solution for v will be unique, whenever it exists. (If such a solution does not exist for all values of $(u, y)\in U\times Y$, we may restrict attention to a subspace of $U\times Y$ for which a $v^*(u, y)$ exists. If no such subspace can be found, then the original problem does not admit a Stackelberg solution.)

Denote the solution by $v^*(u, y)$. Clearly, $y=Nv^*(u, y)$, and hence there exists a one-to-one correspondence between y and v^*, for every fixed $u\in U$. As a

result of this optimization, the leader will incur a cost of $J_1(u, v^*(u, y))$ and the follower will incur $J_2(u, v^*(u, y))$, both of which depend only on u and y. Denote them by

$$\tilde{J}_1(u, y) \triangleq J_1(u, v^*(u, y)) \tag{11 a}$$

$$\tilde{J}_2(u, y) \triangleq J_2(u, v^*(u, y)) \tag{11 b}$$

The best performance the leader can achieve is clearly the global minimum of $\tilde{J}_1$ over $U \times Y$. We now show, in the sequel, that the theory of the previous section can be used to establish the existence of an affine incentive strategy that achieves this bound. Towards this end, we first verify that $\tilde{J}_2(u, y)$ is strictly convex over $U \times Y$.

Lemma 2

The functional $\tilde{J}_2(u, y)$ defined by (11 *b*) is strictly convex over $U \times Y$, provided that $J_2(u, v)$ is strictly convex over $U \times V$.

Proof

Take two arbitrary points (u_1, y_1), $(u_2, y_2) \in U \times Y$, and an $\alpha \in (0, 1)$. Let

$$\bar{u} = \alpha u_1 + (1-\alpha) u_2 \in U$$

$$\bar{y} = \alpha y_1 + (1-\alpha) y_2 \in Y$$

Then we have the following sequence of inequalities

$$\begin{aligned}
\tilde{J}_2(\bar{u}, \bar{y}) &= J_2(\bar{u}, v^*(\bar{u}, \bar{y})) \\
&\leqslant J_2[\bar{u}, \alpha v^*(u_1, y_1) + (1-\alpha) v^*(u_2, y_2)] && (12\,a) \\
&< \alpha J_2[u_1, v^*(u_1, y_1)] + (1-\alpha) J_2[u_2, v^*(u_2, y_2)] && (12\,b) \\
&= \alpha \tilde{J}_2(u_1, y_1) + (1-\alpha) \tilde{J}_2(u_2, y_2) && (12\,c)
\end{aligned}$$

Here, (12 *a*) follows since (i) $v^*(\bar{u}, \bar{y})$ minimizes $J_2(\bar{u}, v)$ over V subject to $\bar{y} = Nv$ and for fixed $\bar{u} \in U$, and (ii) $\bar{v} \triangleq \alpha v^*(u_1, y_1) + (1-\alpha) v^*(u_2, y_2) \in V$ and satisfies the constraint $\bar{y} = N\bar{v}$. The strict inequality (12 *b*), on the other hand, follows from strict convexity of $J_2(u, v)$ over $U \times V$. Finally, (12 *c*) follows from the definition of $\tilde{J}_2$. This, then, verifies that $\tilde{J}_2$ is strictly convex. □

Therefore we now have a new ' projected ' incentive (Stackelberg) problem with perfect information (for the leader), wherein the cost functionals are $\tilde{J}_1(u, y)$, $\tilde{J}_2(u, y)$, and y is the decision variable of the follower. This incentive problem is of the type discussed in § 2, for which we let γ_1^0 denote an optimal incentive strategy. Then, we have the following equivalence between the new ' projected ' incentive problem, and the original one formulated in § 3.1.

Proposition 2

An incentive strategy $\gamma_1^0 \in \Gamma_1$ (i.e. $u = \gamma_1^0(y)$) solves the ' projected ' incentive problem with cost functionals $\tilde{J}_1(u, y)$, $\tilde{J}_2(u, y)$, and with perfect information, if and only if it solves the ' original ' incentive problem with partial information, as formulated in § 3.1.

Proof

This follows the same lines as the proof of Theorem 1 in Başar (1982 a) and is therefore omitted. □

3.3. *Solution of the incentive problem with partial information*

In § 3.2, the original incentive problem with partial information has been transformed into an equivalent ' projected ' problem with perfect information. Hence, the results of § 2.2 can readily be used to solve this problem, since we have already established (in Lemma 2) strict convexity of the new cost functional $\tilde{J}_2(u, y)$ on $U \times Y$. Accordingly, derivation of an optimal affine incentive strategy for the leader would follow the following steps.

Step 1. For fixed $u \in U$ and $y \in Y$, minimize $J_2(u, v)$ over $v \in V$ and subject to $y = Nv$. Denote the solution by $v = v^*(u, y)$. (This solution exists uniquely, since the cost functional is strictly convex and the constraint is linear.)

Step 2. Transform J_1 and J_2 into $\tilde{J}_1$ and $\tilde{J}_2$, respectively, via (11 *a*) and (11 *b*).

Step 3. Determine the global minimum of $\tilde{J}_1(u, y)$ over $(u, y) \in U \times Y$, provided that it exists, and denote the solution by (u^t, y^t). If such a solution does not exist, let (u^t, y^t) denote a pair of values which seems to be most favourable to the leader ; then Steps 4 and 5 in the sequel still determine an optimal incentive strategy.

Step 4. By Proposition 2, the set $\tilde{\Omega}_t \triangleq \{(u, y) | \tilde{J}_2(u, y) \leqslant \tilde{J}_2{}^t \triangleq \tilde{J}_2(u^t, y^t)\}$ is strictly convex, with (u^t, y^t) as a supporting point. Hence, there exists a supporting hyperplane for $\tilde{\Omega}_t$, passing through (u^t, y^t). Assume that $\tilde{J}_2$ is Fréchet-differentiable in $U \times Y$. (This condition, which places some indirect restrictions on J_2, seems to be inevitable if we seek to obtain an explicit expression for the optimal affine incentive strategy. Our subsequent analysis (in § 4) indicates that $\tilde{J}_2$ is indeed Fréchet-differentiable for two important classes of incentive problems.) Then the supporting hyperplane is uniquely defined by

$$\langle \nabla_u \tilde{J}_2(u^t, y^t), u - u^t \rangle + \langle \nabla_y \tilde{J}_2(u^t, y^t), y - y^t \rangle = 0 \tag{13}$$

Step 5. If $\nabla_u \tilde{J}_2(u^t, y^t) \neq 0$ (in space U^*), there exists an operator $Q^* : U^* \to Y^*$ such that

$$Q^*[\nabla_u \tilde{J}_2(u^t, y^t)] = \nabla_y \tilde{J}_2(u^t, y^t) \tag{14}$$

Then, the optimal affine incentive strategy is given by

$$u = \gamma_1(y) = u^t - Q(y - y^t) \tag{15}$$

which constitutes a sub-hyperplane of the supporting hyperplane (13).

Note that, since the strategy (15) has a unique point (u^t, y^t) in common with the set $\tilde{\Omega}_t$, the problem of minimizing $\tilde{J}_2(u, y)$ subject to (15) has the unique solution (u^t, y^t), and by Proposition 2 the problem of minimizing $J_2(u, v)$ subject to (15) and the linear constraint $y = Nv$ has the solution (u^t, v^t) with $Nv^t = y^t$. Therefore, the incentive strategy (15) indeed induces the desired behaviour on the follower.

We have thus established existence of an optimal affine incentive strategy for the leader, in a rather broad class of Stackelberg game problems with partial information, as summarized below.

Proposition 3

For the incentive (Stackelberg) game with partial information, as formulated in § 3.1, there exists an optimal affine incentive strategy for the leader, determined at Step 5 of the foregoing procedure, provided that

(i) $J_2(u, v)$ is strictly convex on $U \times V$, and there exists a $v^* : U \times Y \to V$ that minimizes J_2 over V and subject to $y = Nv$
(ii) $\tilde{J}_1(u, y)$ (as defined by (11 *a*)) admits a global minimum in $U \times Y$
(iii) $\tilde{J}_2(u, y)$ (as defined by (11 *b*)) is Fréchet-differentiable in $U \times Y$
(iv) $\nabla_u \tilde{J}_2(u^t, y^t)$ does not vanish. □

4. Applications

4.1. *Linear-quadratic Stackelberg games on Hilbert spaces*

As a first application of the results of the previous section, we consider the class of linear-quadratic games treated in § 5 of Başar (1982 a).

Denote the leader's decision variable by u_1 and the follower's decision variable by u_2 (instead of u and v), which belong to Hilbert spaces U_1 and U_2, respectively. The cost functional of Player i ($i = 1, 2$) is

$$J_i(u_1, u_2) = \sum_{k,\, j=1,\, 2} \langle u_k, A_{kj}{}^i u_j \rangle + \sum_{j=1,\, 2} \langle u_j, l_j{}^i \rangle \tag{16}$$

where $A_{ij}{}^k$ are linear bounded operators, $A_{ii}{}^k$ are strongly positive, $A_{ji}{}^i = 0$ for $i \neq j$, $l_j{}^i \in U_j$ are known, and

$$A_{ii}{}^i - \tfrac{1}{4} A_{ij}{}^i (A_{jj}{}^i)^{-1} A_{ij}{}^{i*} > 0, \quad i, j = 1, 2, \quad i \neq j \tag{17}$$

which make both J_1 and J_2 strictly convex on $U_1 \times U_2$. The partial information of the leader is denoted by

$$y = Nu_2 \tag{18}$$

where $y \in Y$, a Hilbert space, and the linear operator $N : U \to Y$ is bounded and has full range.

A tight attainable bound for the leader's performance in this game has been obtained in Başar (1982 a). In the sequel we show, by following Steps 1–5 of § 3.3, that there indeed exists an affine incentive strategy that achieves this bound.

Step 1

For fixed $(u, y) \in U_1 \times Y$, the unique minimum of J_2 over $u_2 \in U_2$ and subject to $y = Nu_2$ is attained by (Başar 1982 a)

$$u_2 = By + Cu_1 + Dl_2{}^2 \triangleq u_2{}^*(u_1, y) \tag{19}$$

where the bounded linear operators $B : Y \to U_2$; $C : U_1 \to U_2$; $D : U_2 \to U_2$ are defined by

$$B \triangleq (A_{22}{}^2)^{-1} N^* [N (A_{22}{}^2)^{-1} N^*]^{-1} \tag{20 a}$$

$$C \triangleq -[I - BN](A_{22}{}^2)^{-1} A_{21}{}^2 \tag{20 b}$$

$$D \triangleq -\tfrac{1}{2}[I - BN](A_{22}{}^2)^{-1} \tag{20 c}$$

Step 2

Substitution of (19) into $J_i(u_1, u_2)$ leads to

$$\begin{aligned}\tilde{J}_1(u_1, y) = {} & \langle u_1, (A_{11}{}^1 + A_{12}{}^1 C + C^* A_{22}{}^1 C)u_1\rangle + \langle u_1, (A_{12}{}^1 B + 2C^* A_{22}{}^1 B)y\rangle \\ & + \langle u_1, l_1{}^1\rangle + \langle y, B^* A_{22}{}^1 By\rangle + \langle u_1, 2(A_{12}{}^1 D + C^* A_{22}{}^1 D)l_2{}^2\rangle \\ & + \langle u_1, C^* l_2{}^1\rangle + \langle y, (2B^* A_{22}{}^1 D + B^*)l_2{}^1\rangle \\ & + \langle Dl_2{}^2, A_{22}{}^1 Dl_2{}^2\rangle + \langle Dl_2{}^2, l_2{}^1\rangle \end{aligned} \tag{21 a}$$

$$\begin{aligned}\tilde{J}_2(u_1, y) = {} & \langle u_1, (A_{11}{}^2 + C^* A_{21}{}^2 + C^* A_{22}{}^2 C)u_1\rangle + \langle u_1, (2C^* A_{22}{}^2 B + A_{21}{}^{2*} B)y\rangle \\ & + \langle u_1, l_1{}^2\rangle + \langle y, B^* A_{22}{}^2 By\rangle + \langle u_1, (A_{21}{}^{2*} D + 2C^* A_{22}{}^2 D + C^*)l_2{}^2\rangle \\ & + \langle y, (2B^* A_{22}{}^2 D + B^*)l_2{}^2\rangle + \langle l_2{}^2, (D^* A_{22}{}^2 D + D^*)l_2{}^2\rangle \end{aligned} \tag{21 b}$$

Note that both of these expressions are strictly convex and Fréchet-differentiable on $U_1 \times Y$.

Step 3

Minimization of $\tilde{J}_1(u, y)$ over $U_1 \times Y$ leads to the unique solution (Başar 1982 a)

$$y^t = K^{-1} l \tag{22 a}$$

$$u_1{}^t = -K_1{}^{-1}(K_2 y + \tilde{l}_1) \tag{22 b}$$

where

$$K_1 \triangleq 2(A_{11}{}^1 + A_{12}{}^1 C + C^* A_{22}{}^1 C) \tag{23 a}$$

$$K_2 \triangleq 2(A_{12}{}^1 B + C^* A_{22} B) \tag{23 b}$$

$$\tilde{l}_1 \triangleq l_1{}^1 + (2A_{12}{}^1 D + 2C^* A_{22}{}^1 D)l_2{}^2 + C^* l_2{}^1 \tag{23 c}$$

$$l \triangleq K_2{}^* K_1{}^{-1} \tilde{l}_1 - 2(B^* A_{22}{}^1 D)l_2{}^2 - B^* l_2{}^1 \tag{23 d}$$

$$K \triangleq 2B^* A_{22}{}^1 B - K_2{}^* K_1{}^{-1} K_2 \tag{23 e}$$

Step 4

The supporting hyperplane of the convex set

$$\tilde{\Omega}_t = \{(u_1, y) | \tilde{J}_2(u_1, y) \leqslant \tilde{J}_2(u_1{}^t, y^t)\}$$

at the point $(u_1{}^t, y^t)$ is

$$\langle \nabla_{u_1} \tilde{J}_2{}^t, u_1 - u_1{}^t\rangle + \langle \nabla_y \tilde{J}_2{}^t, y - y^t\rangle = 0 \tag{24}$$

where

$$\begin{aligned}\nabla_{u_1} \tilde{J}_2{}^t = {} & 2(A_{11}{}^2 + C^* A_{21}{}^2 + C^* A_{22}{}^2 C)u_1{}^t + (2C^* A_{22}{}^2 B + A_{21}{}^{2*} B)y^t \\ & + l_1{}^2 + (A_{21}{}^{2*} D + 2C^* A_{22}{}^2 D + C^*)l_2{}^2 \end{aligned} \tag{25 a}$$

$$\begin{aligned}\nabla_y \tilde{J}_2{}^t = {} & 2B^* A_{22}{}^2 By^t + (B^* A_{21}{}^2 + 2B^* A_{22}{}^2 C)u_1{}^t \\ & + (2B^* A_{22}{}^2 D + B^*)l_2{}^2 \end{aligned} \tag{25 b}$$

Step 5

Assume that $\nabla_{u_1} \tilde{J}_2{}^t \neq 0$. Then, an optimal affine incentive relation that achieves the bound $\tilde{J}_1(u_1{}^t, y^t)$ is

$$u_1 = \gamma_1(y) = u_1{}^t - Q(y - y^t) \tag{26}$$

where the linear bounded operator $Q: Y \to U_1$ satisfies the relation

$$Q^*[\nabla_{u_1} \tilde{J}_2{}^t] = \nabla_y \tilde{J}_2{}^t \tag{27}$$

with Q^* defined as the adjoint of Q. One possible solution to (27) is

$$Q^*(\cdot)=\langle \nabla_{u_1}\tilde{J}_2{}^t, \cdot\rangle \frac{\nabla_y \tilde{J}_2{}^t}{\langle \nabla_{u_1}\tilde{J}_2{}^t, \nabla_{u_1}\tilde{J}_2{}^t\rangle} \tag{28}$$

whose adjoint is

$$Q(\cdot)=\frac{\nabla_{u_1}\tilde{J}_2{}^t}{\langle \nabla_{u_1}\tilde{J}_2{}^t, \nabla_{u_12}\tilde{J}^t\rangle} \langle \nabla_y \tilde{J}_2{}^t, \cdot\rangle \tag{29}$$

Hence, an optimal incentive (Stackelberg) strategy for the leader is

$$\gamma_1(y)=u_1{}^t-\frac{\nabla_{u_1}\tilde{J}_2{}^t}{\langle \nabla_{u_1}\tilde{J}_2{}^t, \nabla_{u_1}\tilde{J}_2{}^t\rangle} \langle \nabla_y \tilde{J}_2{}^t, y-y^t\rangle \tag{30}$$

Of course, there are several other optimal affine incentive strategies for the leader, since (27) admits, in general, multiple solutions.

4.2. *A differential game with sampled state measurements*

Another special class of problems to which the results of § 3 are applicable is the class of linear-quadratic differential game problems in which the leader has access to sampled-data state information, as described in Başar (1981 a, § 4). The state evolves according to

$$\dot{x}=A(t)x+B_1(t)u+B_2(t)v\ ; \quad x(0)=x_0, \quad t\in[t_0, t_f]$$

and the cost functionals are

$$J_1(u, v)=\tfrac{1}{2}x'(t_f)K_{1f}x(t_f)+\tfrac{1}{2}\int_{t_0}^{t_f} [x'Q_1(t)x+u'u+v'R_1(t)v]\, dt$$

$$J_2(u, v)=\tfrac{1}{2}x'(t_f)K_{2f}x(t_f)+\tfrac{1}{2}\int_{t_0}^{t_f} [x'Q_2(t)x+u'R_2(t)u+v'v]\, dt$$

where $K_{if}\geqslant 0$, $Q_i\geqslant 0$, $R_i\geqslant 0$, $i=1, 2$, and all matrices have continuous entries.

The leader's observations are the values of state x at sampled times $t_0, t_1, \ldots, t_k=t_f$; that is, the values $x(t_0)=x_0$, $x(t_1)=x_1, \ldots, x(t_{k-1})=x_{k-1}$, $x(t_f)=x_f=x_k$.

Because of the linearity of the dynamic process

$$x_i=\phi(t_i, t_0)x_0+\int_{t_0}^{t_i} \phi(t_i, \tau)B_1u(\tau)\, d\tau+\int_{t_0}^{t_i} \phi(t_i, \tau)B_2v(\tau)\, d\tau, \quad i=1, \ldots, k$$

where $\phi(t, \tau)$ is the state transition matrix of the linear system.

Since x_0a nd $u(\tau)$ are assumed to be known to the leader, the observations are equivalent to

$$y_i=\int_{t_0}^{t_i} \phi(t_i, \tau)B_2v(\tau)\, d\tau \triangleq N_i[v], \quad i=1, \ldots, k$$

where $N_i[\cdot]$ are linear, bounded and causal operators.

Derivation of an achievable lower bound $J_1{}^*$ on the leader's cost function, and the open-loop values u^t and v^t of the leader's and follower's corresponding strategies, as well as the derivation of corresponding optimal values of the state at sampling times, have been discussed in Başar (1981 a). The results of § 3 indicate, however, that there may exist (more appealing) affine strategies for

the leader, which achieve the same bound and lead to the same optimal values of the state at sampling times. We now illustrate this via a numerical example.

Consider the first-order system $\dot{x} = -x + u + v$, $t \in [0, 2]$, $x(0) = x_0$ known and $x(2)$ free, where $v \in L_2[0, 1]$ is the follower's control action and $u \in L_2[1, 2]$ is the leader's control action. Note that they act on different time intervals, so that the causality condition is satisfied automatically.

Assume that the leader's cost functional is

$$J_1 = \tfrac{1}{2} \int_0^2 (x^2 + 2u^2 + v^2)\, dt$$

and the follower's cost functional is

$$J_2 = \tfrac{1}{2} \int_0^2 (x^2 + v^2)\, dt$$

The leader is assumed to have the information on state $x(t)$ only at $t = 1$, i.e. $x(1) = x_1$. Since $x_1 = \exp(-1)\left[x_0 + \int_0^1 \exp(t)\, v(t)\, dt\right] \triangleq \exp(-1)\,[x_0 + y]$, we may take $y = \int_0^1 \exp(t)\, v(t)\, dt = N[v(t)]$ as the observation of the leader.

Let us now apply the five steps (of derivation) of § 3.3 to this problem.

Step 1

For fixed u and y (i.e. x_1) minimize J_2 with respect to v. By standard techniques of optimal control, the minimizing solution is obtained to be

$$v^*(t, y) = (0{\cdot}03224x_0 + 0{\cdot}09506y) \exp(\sqrt{2}t) + (0{\cdot}96776x_0 - 0{\cdot}09506y) \exp(-\sqrt{2}t)$$

Note that, here, $v^*(t, y)$ is independent of u, and depends only on x_0 and y.

Step 2

Substitute $v^*(t, y)$ into J_i to obtain the expressions for $\bar{J}_i$

$$\bar{J}_1 = 0{\cdot}20249x_0^2 + 0{\cdot}08192x_0y + 0{\cdot}17542y^2 + \tfrac{1}{2}\int_1^2 \left[\exp(t-1)\, x_1 + \int_1^t \exp(t-s)\, u(s)\, ds\right]^2 dt + \tfrac{1}{2}\int_1^2 2u^2(t)\, dt$$

$$\bar{J}_2 = 0{\cdot}20249x_0^2 + 0{\cdot}08192x_0y + 0{\cdot}017542y^2 + 1{\cdot}59726x_1^2 + \frac{x_1}{2}\int_1^2 u(t)[\exp(3-t) - \exp(t-1)]\, dt + \tfrac{1}{2}\int_1^2 \left[\int_1^t \exp(t-s)\, u(s)\, ds\right]^2 dt$$

Step 3

Minimize $\bar{J}_1$ with respect to u and y. By taking the gradients of $\bar{J}_1$ with respect to $u(t)$ and y in function space and equating them to zero, we obtain

$$u^t = (0{\cdot}00469 \exp\{\sqrt{1{\cdot}5}(t-1)\} - 0{\cdot}05430 \exp\{-\sqrt{1{\cdot}5}(t-1)\})\, x_0$$

$$y^t = -0{\cdot}33754x_0$$

Step 4

Calculate the values of the gradients of $\tilde{J}_2$ at the point (u^t, y^t)

$$\nabla_y \tilde{J}_2{}^t = \nabla_y \tilde{J}_2(u^t, y^t) = 0{\cdot}23320x_0$$

$$\begin{aligned}\nabla_u \tilde{J}_2{}^t &= \nabla_u \tilde{J}_2(u^t, y^t)\\ &= \{1{\cdot}99295 \exp(-t) - 0{\cdot}03650 \exp(t)\\ &\qquad - 0{\cdot}00276 \exp(\sqrt{1{\cdot}5}t) + 0{\cdot}36957 \exp(-\sqrt{1{\cdot}5}t)\}x_0\\ &\triangleq \Psi(t)x_0\end{aligned}$$

Step 5

The supporting hyperplane of the set $\{(u, y) | \tilde{J}_2(u, y) \leqslant \tilde{J}_2(u^t, y^t)\}$ is given by

$$\nabla_y \tilde{J}_2{}^t[y - y^t] + \int_1^2 \nabla_u \tilde{J}_2{}^t[u(t) - u^t(t)]\, dt = 0 \tag{31}$$

In order to find a $u = \gamma_1(y)$ that satisfies this equation, let $u(t) = u^t(t) + \Delta(y)$, where $\Delta(y)$ is independent of t. Then Δ satisfies

$$0{\cdot}23320x_0[y - y^t] + \Delta x_0 \int_1^2 \Psi(t)\, dt = 0$$

whereby

$$\Delta = -\frac{0{\cdot}23320(y - y^t)}{\int_1^2 \Psi(t)\, dt} = -0{\cdot}69168(y - y^t)$$

Hence, an optimal incentive strategy is

$$u = \gamma_1(y) = u^t - 0{\cdot}69168(y - y^t)$$

where $u^t(t)$ and y^t were obtained at Step 3. As another alternative, let

$$Q^* = \frac{\langle \nabla_u, \cdot \rangle}{\|\nabla_u\|^2} \nabla_y, \quad Q = \frac{\nabla_u}{\|\nabla_u\|^2} \langle \nabla_y, \cdot \rangle$$

where ∇_u and ∇_y are the shorthand notations for $\nabla_u \tilde{J}_2{}^t$ and $\nabla_y \tilde{J}_2{}^t$ respectively. Then, another optimal incentive strategy would be

$$\begin{aligned}u = \gamma_1(y) &= u^t - Q[y - y^t] = u^t - \frac{\nabla_u}{\|\nabla_u\|^2} \nabla_y (y - y^t)\\ &= u^t - \frac{\Psi(t)}{\int_1^2 \Psi^2(t)\, dt} [0{\cdot}33754(y - y^t)]\\ &= u^t - 0{\cdot}33200\Psi(t)[y - y^t]\end{aligned}$$

which is of a different form. In fact, there are several other possibilities, as can be seen from (31), all of which lead to the tight lower bound $\tilde{J}_1(u^t, y^t)$. A further selection from this set of optimum incentive strategies may be made, in accordance with some additional design requirements.

5. Concluding remarks

This paper has developed a geometric approach to solve a sufficiently general class of incentive decision problems in which one of the decision makers (the leader) announces an incentive strategy in order to induce a certain behaviour on the other decision maker (the follower). The leader is assumed to have either (i) perfect information on the follower's actions or (ii) partial information provided by a linear observation. In either case, and when the cost functional of the follower is strictly convex, it is shown under rather mild conditions that there exists an affine 'optimal' incentive strategy for the leader. As a byproduct of our analysis, we have also obtained a computational scheme that yields the linear operator associated with the affine strategy. The two examples included in § 4 illustrated the applicability of this general approach and derivation of affine optimal incentive strategies.

The general form of the incentive strategy, which is either (i) $u = u^t - Q(v - v^t)$ or (ii) $u = u^t - Q(y - y^t)$, depending on whether we have perfect or partial information, may, at first glance, lead to the conclusion that our analysis and results are valid only if the leader acts after the follower does in the decision process. However, this is not totally true, because the existence of several degrees of freedom inherent in the choice of the linear operator Q (which only has to satisfy eqn. (14)) makes this analysis a viable one also for dynamic decision processes wherein the actions of the two decision makers are temporally intermingled. Hence, it is in general possible to choose this linear operator in a way that is compatible with the 'control-information dependence' requirements (such as causality) of the dynamic game under consideration, which definitely allows the decision makers to act more than once in the decision process. Two such classes of problems are the discrete-time dynamic games with linear state equation and strictly convex cost functionals (such as those considered in Başar and Selbuz (1979)), and the continuous-time differential games with linear state equation, strictly convex cost functionals and sampled-data information (such as the ones treated in Başar (1981 a), and also briefly discussed in § 4.2). Precise formulations of such 'information structural' restraints, and the derivation of causal linear operators Q compatible with these requirements, will be undertaken in future publications. We also expect extension of these results to hierarchical dynamic decision problems with more than two players, in the framework of the model (and along the lines of the analyses) of Başar (1981 b, 1982 b).

References

Balakrishnan, A. V., 1976, *Applied Functional Analysis* (New York : Springer).

Başar, T., 1981 a, *Proc. 7th IFAC Congress*, Kyoto, Japan, pp. IX-139–144 ; 1981 b, *Automatica*, **17**, 749 ; 1982 a, *Large Scale Systems*, **3**, 47. An earlier version appeared in 1980 in *Proc. 4th Int. Conf. Analysis Optim. Syst.* (New York : Springer) ; 1982 b, *J. optim. Theory Applic.* (to be published).

Başar, T., and Selbuz, H., 1979, *I.E.E.E. Trans. autom. Control*, **24**, 166.

Chen, C. I., and Cruz, J. B., Jr., 1972, *I.E.E.E. Trans. autom. Control*, **17**, 791.

Cruz, J. B., Jr., 1978, *I.E.E.E. Trans. autom. Control*, **23**, 244.

Groves, T., and Loeb, M., 1979, *Mgmt Sci.*, **25**, 221.

Ho, Y. C., Luh, P. B., and Olsder, G. J., 1980, *Proc. 4th Int. Conf. Analysis Optim. Syst.* (New York : Springer).

HO, Y. C., LUH, P. B., and MURALIDHARAN, R., 1981, *I.E.E.E. Trans. autom. Control*, **26,** 454.

HURWICZ, L., and SHAPIRO, 1978, *Bell J. Econ.*, **9,** 180.

JENNERGREN, L. P., 1980, *Mgmt Sci.*, **26,** 180.

LUENBERGER, D. G., 1969, *Optimization by Vector Space Methods* (New York : Wiley).

PAPAVASSILOPOULOS, G. P., and CRUZ, J. B., Jr., 1979, *I.E.E.E. Trans. autom. Control*, 24, 155 ; 1980, *J. optim. Theory Applic.*, **31,** 233.

SIMAAN, M., and CRUZ, J. B., Jr., 1973 a, *J. optim. Theory Applic.*, **11,** 533 ; 1973 b, *Ibid.*, **11,** 613.

STACKELBERG, H. VON, 1934, *Marktform und Gleichgewicht* (Vienna : Springer) ; 1952, *The Theory of the Market Economy* (Oxford : Oxford University Press).

TOLWINSKI, B., 1981, *J. optim. Theory Applic.*, **34,** 485.