SIAM J. CONTROL AND OPTIMIZATION
Vol. 19, No. 5, September 1981

0363-0129/81/1905-0007 $01.00/0

SOLUTION OF SOME STOCHASTIC QUADRATIC NASH AND LEADER-FOLLOWER GAMES*

G. P. PAPAVASSILOPOULOS†

Abstract. The linear quadratic Gaussian static Nash and Stackelberg two-player games are considered and completely solved. Necessary and sufficient conditions for existence and uniqueness of the solutions are presented as well as the procedure for finding all the solutions. For the Nash game, in particular, it is shown that if there exists a solution there will exist a solution affine in the information, and that the solution will be nonunique if (intuitively) the coupling of the information of the two players equals some power of the inverse of the coupling of their costs. Many interesting dynamic cases with nested information structures can be reduced to static ones and are essentially covered by the analysis presented.

1. Introduction. It has been recognized that the single objective optimization problem cannot capture all the aims of a decision procedure. Usually there are many conflicting objectives that a decision maker has to meet, and the formulation of a single objective as a weighted sum of the several objectives is not necessarily the only way to go. Also, there might exist many decision makers with conflicting objectives who do not agree on an overall average objective. On the other hand, an existing hierarchy among the several decision makers in a certain organization should not be ignored when one creates the mathematical model. Such considerations make game theory a natural vehicle for studying multiobjective hierarchical decision procedures. In particular, the so-called Nash and leader-follower (or Stalkelberg) games offer themselves for studying such situations. For definitions and some properties of these games see [2], [3]. (See also [13] for some recent results concerning leader-follower games and their relation to the theory of incentives in economics.)

There are several results concerning Nash and leader-follower games, but there are still many open problems. In this paper we study, and solve completely the static Nash and leader-follower games, where the players have quadratic costs and linear measurements of a random variable which enters linearly into the costs, see (1)–(7). Several dynamic cases (where there is time evolution), are included in the static formulation, as long as appropriate nestedness conditions [4] are imposed on the information of the players. For example, the stochastic linear quadratic discrete time Gaussian Nash game, where the players share at each stage all the past information with one step delay, belongs to the class of dynamic games that can be reduced to static ones. Another example is the stochastic linear quadratic, discrete time Gaussian leader-follower game, where the leader has information only at the first stage, whereas the follower in addition to having his own information acquires the information of the leader with one step delay. Although such dynamic problems can be handled by the methods developed here, we will focus on the static case only. We will nonetheless provide in § 5 the procedure, which reduces dynamic problems of this type to static ones.

The only existing results concerning such types of stochastic games are in [6], [7]. There, sufficient conditions for existence and uniqueness of solutions are found by imposing a contraction assumption. As a consequence, the results of [6], [7], in addition

* Received by the editors March 17, 1980, and in revised form January 20, 1981. This work was supported in part by the U.S. Air Force Office of Scientific Research under grant AFOSR-80-0171.

† Dept. of Electrical Engineering, University of Southern California, University Park, Los Angeles, CA 90007.

to being extremely conservative, cannot answer the important question of how the interplay among the information and the costs of the players affects the solution.

The structure of the present paper is the following. In § 2 we pose the problems and in § 3 we give the complete solution of the Nash game. It is shown that the solution will be nonunique if some numbers, which are products of even powers of the canonical correlation coefficients of the information of the two players, are inverses of eigenvalues of a matrix (R_1R_2), which represents the strength of the coupling of the costs. Intuitively, the solution will be nonunique if the coupling of the information is equal to some power of the inverse of the coupling of the costs. It is also shown that if there exists a solution there will exist a solution linear in the information. The way to construct all the solutions is also given. In § 4 we solve the leader-follower game. In § 5 we derive a sufficient condition for the existence and uniqueness of a solution of an equation which is a generalization of an equation playing a central role in § 3. A special case of this condition was presented in [5], but our result, in addition to being more powerful, is proven in a much easier fashion. This condition can be used to guarantee existence and uniqueness of solutions of Nash and leader-follower games, if there are many players and one is not willing to generalize the exact results of §§ 3 and 4 to the many player case. In this section we also sketch a way to reduce dynamic problems with nested information to static ones. Finally, § 6 presents the solution of a simple one-dimensional Nash static stochastic game which can serve to illustrate some aspects of the whole analysis. § 7 is a conclusions section. The proofs of two lemmas used in § 3 are given in Appendix A and B. Appendix C contains some results concerning an operator which turns out to be of importance when solving the Nash or leader-follower games.

In our analysis, we will consider two players only, but generalization to the many player case is possible.

2. Statement of the problems. Let $x:\Omega\to R^n$ be a Gaussian random variable with respect to a probability space $(\Omega, \mathcal{F}, P)$, which, without loss of generality, is assumed to have zero mean and unit covariance matrix. Let

$$y_i = C_i x, \qquad i=1,2, \tag{1}$$

where the C_i's are real matrices of dimension $n_i\times n$. The random variable $y_i:\Omega\to R^{n_i}$ generates a minimal sub σ-field $\mathcal{F}_i$ of $\mathcal{F}$ in Ω. Let U_i, $i=1, 2$ denote the space of functions $u_i:\Omega\to R^{m_i}$, which are $\mathcal{F}_i$ measurable and for which $\|u_i\|^2=\langle u_i, u_i\rangle<+\infty$, where the inner product in U_i is defined by

$$\langle u_i, v_j\rangle=\int_\Omega u_i' v_j \, dP, \qquad u_i, v_i\in U_i. \tag{2}$$

U_i is a separable Hilbert space and u_i can be considered as a function of y_i; see [8]. For a given pair $(u_1, u_2)\in U_1\times U_2$ consider

$$J_1(u_1, u_2)=E[\tfrac{1}{2}u_1'u_1+u_1'R_1u_2+u_1'S_1x+u_2'Q_{12}u_2+u_2'F_1x], \tag{3}$$

$$J_2(u_1, u_2)=E[\tfrac{1}{2}u_2'u_2+u_2'R_2u_1+u_2'S_2x+u_1'Q_{21}u_1+u_1'F_2x], \tag{4}$$

where E denotes total expectation, and $Q_{ij}=Q_{ij}'$, R_i, S_i, F_i are real matrices of appropriate dimensions. We want to solve the problems N and S.

Problem N. Find the pairs $(u_1^*, u_2^*)\in U_1\times U_2$ for which

$$J_1(u_1^*, u_2^*)\leqq J_1(u_1, u_2^*) \quad \forall u_1\in U_1, \tag{5}$$

$$J_2(u_1^*, u_2^*)\leqq J_2(u_1^*, u_2) \quad \forall u_2\in U_2. \tag{6}$$

Problem S. Find the pairs $(u_1^*, u_2^*) \in U_1 \times U_2$ which solve:

$$\begin{aligned} &\text{minimize} && J_1(u_1, u_2) \\ &\text{subject to} && (u_1, u_2) \in U_1 \times U_2 \quad \text{and} \quad J_2(u_1, u_2) \leqq J_2(u_1, \bar{u}_2) \quad \forall \bar{u}_2 \in U_2. \end{aligned} \tag{7}$$

The formalism (1)–(6) describes a two-player nonzero sum, static, stochastic, Nash game, where player i has information y_i, chooses u_i and wants to minimize J_i. The formalism (1)–(4), (7) describes a two-player, nonzero sum, static, stochastic, leader-follower game, where player i has information y_i, chooses u_i and wants to minimize J_i, and, in addition, player 1 (leader) decides and announces his decision u_1, first, before player 2 (follower) decides on u_2.

Without loss of generality, we make the following assumption, which is assumed to hold throughout the present paper:

Assumption.

$$\operatorname{rank} C_i = n_i, \qquad i = 1, 2.$$

The formula $E[x \mid y_i] = C_i'(C_iC_i')^{-1}C_ix$, will be used repeatedly in the later sections.

3. Solution of the problem N. In this section we solve problem N. For fixed $u_2 \in U_2$, the problem

$$\text{minimize} \qquad J_1(u_1, u_2), \qquad u_1 \in U_1, \tag{8}$$

is a quadratic minimization problem in the Hilbert space $U_{m_1} = \{u : \Omega \to R^{m_1}, u \text{ is } \mathcal{F}$ measurable and $\|u\| < +\infty\}$, which has U_1 as a closed subspace. Use of the projection theorem yields (9) as a necessary and sufficient condition for u_1 to solve (8):

$$u_1 + E[R_1u_2 \mid y_1] + E[S_1x \mid y_1] = 0; \tag{9}$$

see [1] for details. $E[\cdot \mid y_i]$ denotes conditional expectation. Similarly for fixed $u_1 \in U_1$, the problem

$$\text{minimize} \qquad J_2(u_1, u_2), \qquad u_2 \in U_2, \tag{10}$$

has u_2 as a solution if and only if

$$u_2 + E[R_2u_1 \mid y_2] + E[S_2x \mid y_2] = 0. \tag{11}$$

Substituting u_2 from (11) into (9), we conclude that the study of Problem N is equivalent to the study of the equation

$$u_1 - R_1R_2E[E[u_1 \mid y_2] \mid y_1] = R_1S_2E[E[x \mid y_2] \mid y_1] - S_1E[x \mid y_1], \tag{12}$$

on which we will concentrate from now on.

We will need the following lemma.

LEMMA 1. *There exist nonsingular square matrices T_1, T_2 so that the matrices*

$$\bar{C}_1 = T_1C_1, \qquad \bar{C}_2 = T_2C_2,$$

have the following properties:

(1) $\bar{C}_1\bar{C}_1' = I$, $\bar{C}_2\bar{C}_2' = I$.

(2)

$$\bar{C}_1 = \begin{bmatrix} \bar{C}_{111} \\ \bar{C}_{112} \\ \bar{C}_{12} \end{bmatrix}, \qquad \bar{C}_2 = \begin{bmatrix} \bar{C}_{222} \\ \bar{C}_{221} \\ \bar{C}_{21} \end{bmatrix}, \qquad \bar{C}_{12} = \bar{C}_{21}, \qquad \bar{C}_1\bar{C}_2' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \bar{C}_{112}\bar{C}_{221}' & 0 \\ 0 & 0 & I \end{bmatrix}.$$

(3) *$\bar{C}_{112}$ and $\bar{C}_{221}$ have the same dimensions $k \times n$ and*

$$\bar{C}_{112}\bar{C}'_{221} = \begin{bmatrix} \sqrt{\mu_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\mu_k} \end{bmatrix} \quad \textit{where } 0 \lneqq \sqrt{\mu_i} \lneqq 1.$$

(4) *The dimensions of all the component matrices are uniquely determined.*

The proof of this lemma can be found in standard books on statistics [12], where the elements of $\bar{C}_1\bar{C}'_2$ are called the canonical correlation coefficients of y_1, y_2 and algorithms for finding T_1, T_2 are described. For the sake of completeness we present a proof of Lemma 1 in Appendix A.

The importance of Lemma 1 for our problem is that since T_i is nonsingular, the minimal σ-fields generated by y_i and T_iy_i are the same, and thus, we can consider equivalently that u_i is a function of T_iy_i. *In the following we will assume that C_1, C_2 have been brought into the form suggested by Lemma* 1 (and drop the bars from $\bar{C}_1$, $\bar{C}_2$).

In terms of the information structure of the game, Lemma 1 allows us to consider y_1 and y_2 as normal Gaussian vectors which can be decomposed into independent components as $y_1 = (y_{112}, y_{111}, y_{12})$, $y_2 = (y_{221}, y_{222}, y_{21})$, where $y_{12} = y_{21}$ is the common information, y_{111} is known only to player 1, y_{222} is known only to player 2 and y_{112}, y_{221} represent the nontrivial coupling of the information.

Let us introduce the following notation:

$$\begin{gathered} y_i = C_ix, \quad y_{ij} = C_{ij}x, \quad y_{ijl} = C_{ijl}x, \\ P_i = E[\cdot \mid y_i], \quad P_{ij} = E[\cdot \mid y_{ij}], \quad P_{ijl} = E[\cdot \mid y_{ijl}], \qquad i, j, l = 1, 2. \end{gathered} \tag{13}$$

P_i, P_{ij}, P_{ijl} and E are projections in the Hilbert space U_{m_1}. An equivalent interpretation of Lemma 1 is that, if without loss of generality, we impose $E[u] = 0$ in U_{m_1}, then P_1 and P_2 can be decomposed into sums of orthogonal projections; i.e., $P_1 = P_{111} + P_{112} + P_{12}$, $P_2 = P_{222} + P_{221} + P_{21}$, where $P_{12} = P_{21}$, $P_{111}P_2 = 0$, $P_{222}P_1 = 0$, $P_{112}P_{221} \neq P_{221}P_{112}$ and $\|P_{221}P_{112}\| = \|P_{112}P_{221}\| < 1$ (see Lemma 2). One can verify that $P_{12} = \lim (P_1P_2)^n$ as $n \to +\infty$, $P_{111} = \lim (P_1(I - P_2))^n$, $P_{222} = \lim (P_2(I - P_2))^n$, $P_{112} = P_1 - P_{12} - P_{111}$, $P_{221} = P_2 - P_{21} - P_{222}$ (see [10, problem 96]).

We can write (12) as

$$u_1 - R_1R_2P_1P_2u_1 = Sy_1, \tag{14}$$

where

$$S = R_1S_2C'_2C_2C'_1 - S_1C'_1, \qquad S = [s_1|\cdots|s_{n_1}]. \tag{15}$$

We will construct an orthonormal complete set for U_1. Let

$$p_n(z) = \frac{(-1)^n}{\sqrt{n!}} e^{\frac{1}{2}z^2} \frac{d^n}{dz^n} e^{-\frac{1}{2}z^2}, \qquad n = 0, 1, 2, \cdots, \tag{16}$$

be the Hermite polynomials which constitute an orthonormal complete set with respect to the Gaussian measure $\mu(z)$; i.e.,

$$\int_{-\infty}^{+\infty} p_n(z)p_l(z)\, d\mu(z) = \begin{cases} 1 & \text{if } n = l, \\ 0 & \text{if } n \neq l, \end{cases} \tag{17}$$

where

$$\mu(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}w^2}\, dw \tag{18}$$

(see [9, p. 217]). Let $y_1=(z_1, \cdots, z_{n_1})'$. Since y_1 is a normal Gaussian random vector,

$$\{p_{k_1}(z_1)p_{k_2}(z_2)\cdots p_{k_{n_1}}(z_{n_1})\}, \qquad k_1, \cdots, k_{n_1}\in\{0, 1, 2, \cdots\} \tag{19}$$

constitute an orthonormal complete set with which we can express every component of u_1, see [11, p. 56]. Let us enumerate this set and denote it by

$$\{\bar{p}_0(y_1), \bar{p}_1(y_1), \cdots\}. \tag{20}$$

We will assume, without loss of generality, that $\bar{p}_0=1$, $(\bar{p}_1(y_1), \cdots, \bar{p}_{n_1}(y_1))'=y_1$, and that as n increases, the power of each component of y_1 in $\bar{p}_n(y_1)$ goes to infinity. Each $u_1\in U_1$ can be expressed as

$$u_1(y_1)=\sum_{n=0}^{\infty} c_n\bar{p}_n(y_1), \tag{21}$$

where $c_n\in R^{m_1}$ and

$$\sum_{n=0}^{\infty} \|c_n\|^2<+\infty. \tag{22}$$

Equation (14) can be written equivalently as

$$\sum_{n=0}^{\infty} c_n\bar{p}_n(y_1)-R_1R_2\sum_{n=0}^{\infty} c_nP_1P_2\bar{p}_n(y_1)=Sy_1. \tag{23}$$

We will need the following lemma, the proof of which is given in Appendix B.

LEMMA 2. *Let*

$$a_{nl}=E[\bar{p}_n(y_1)P_1P_2\bar{p}_l(y_1)], \qquad a_n=a_{nn}.$$

Then the following hold:

(1) $a_{nl}=a_{ln}$.

(2) $a_{nl}=0$ *if* $n\neq l$.

(3) $a_n=0$ *if* $\bar{p}_n$ *depends on* y_{111}.

(4) $a_n=1$ *if* $\bar{p}_n$ *depends only on* y_{12}, *or if* $n=0$.

(5) *If* $\bar{p}_n(y_1)=\bar{p}_{n_3}(y_{12})\cdot p_{m_1}(z_1)\cdots p_{m_k}(z_k)$, *where* $y_{112}=(z_1, \cdots, z_k)'$ *then* $a_n=\mu_1^{m_1}\cdots\mu_k^{m_k}$, a_n *is independent of* n_3, $0<a_n<1$ *and these* a_n*'s constitute a sequence decreasing strictly to zero.*

(6) *The operators* $P_{112}P_{221}$, $P_{221}P_{112}$, *restricted on the domain of the* u*'s with* $E[u]=0$, *have norm equal to* $\max\{\mu_1, \cdots, \mu_k\}<1$.

Multiplying both sides of (23) by $\bar{p}_n(y_1)$, taking expectation and using Lemma 2 (2) yields

$$c_n-R_1R_2c_na_n=\begin{cases} s_n & \text{if } n=1, \cdots, n_1,\\ 0 & \text{otherwise,}\end{cases} \tag{24}$$

or more concisely

$$[c_0, c_1, \cdots]-R_1R_2[c_0, c_1, \cdots]\begin{bmatrix} a_0 & & 0\\ & a_1 & \\ 0 & & \ddots\end{bmatrix}=[\underbrace{0}_{m_1\times 1}|S|00\cdots]. \tag{25}$$

The conditions for solvability of (25) are apparent.

Let us now state formally all the previous analysis, in the form of a theorem.

THEOREM 1. *Consider equation* (12):

(1) *It has a solution if and only if there exist* $c_0, c_1, \cdots\in\mathbb{R}^{m_1}$ *which satisfy* (25).

(2) *If there exists at least one solution, then there exists a solution linear in* (*the information*) y_1.

(3) *The general solution, if it exists, has the form*

$$(26)\qquad u_1=c_0+[c_1,\cdots,c_{n_1}]y_1+\sum_{k=1}^{q} c_{\bar n_k}\bar p_{\bar n_k}(y_{112})\phi_k(y_{12})+[c_{\bar m_1},\cdots,c_{\bar m_r}]\phi(y_{12}),$$

where $c_0, c_1,\cdots,c_{n_1}, c_{\bar n_1},\cdots,c_{\bar n_q}$ *satisfy* (24), $\bar n_1,\cdots,\bar n_q>n_1$, $c_{\bar m_1},\cdots,c_{\bar m_r}$ *constitute a basis for the null space of* $(I-R_1R_2)$, ϕ *and* ϕ_k *are arbitrary measurable functions of* y_{12} *taking values in* R^r *and* R *respectively,* $E[\phi_k^2]=1$, $E[\phi'\phi]<+\infty$, *and* ϕ *contains no affine term in* y_{12} (*i.e.,* $E[\phi]=0$, $E[y_{12}\cdot\phi']=0$).

Proof. The proof is immediate from the previous analysis. It need only be pointed out that the $c_{\bar n_k}$'s will be finite in number since as n increases the a_n's which correspond to $\bar p_n(y_{112})$ decrease to zero, and R_1R_2 has a finite number of eigenvalues. c_0, $c_{\bar m_1},\cdots,c_{\bar m_r}$ are all eigenvectors of R_1R_2 corresponding to the eigenvalue 1. The appearance of the product $\bar p_{\bar n_k}(y_{112})\phi_k(y_{12})$ is a consequence of Lemma 2(5). □

The procedure suggested by Theorem 1 for solving (12) is the following.

Step 1. Try to find $u_L=c_0+L_1y_1$, which solves (12). ($c_0\in R^{m_1}$ and L_1 will equal $[c_1,\cdots,c_{n_1}]$ see (26).) This is equivalent to solving the equations

$$(I-R_1R_2)c_0=0,$$

$$L_1-R_1R_2L_1C_1C_2'(C_2C_2')^{-1}C_2C_1'(C_1C_1')^{-1}$$
$$=R_1S_2C_2'(C_2C_2')^{-1}C_2C_1'(C_1C_1')^{-1}-S_1C_1'(C_1C_1')^{-1}.$$

(Notice that we do not need to find $\bar C_i=T_iC_i$ in order to carry out this step. If this has been done and we use $\bar C_i$ in place of C_i then it is easily seen that the two equations above are equivalent to (25) with $n=0, 1,\cdots,n_1$.) If there exists no such L_1, stop and conclude that there is no solution. Otherwise go to step 2.

Step 2. Solve for L_2 the equation

$$(I-R_1R_2)L_2=0,$$

where L_2 is an $m_1\times r$ matrix and where r is the dimension of the null space of $I-R_1R_2$. Set $u_0=L_2\varphi(y_{12})$, where φ is any $r\times 1$ vector function of y_{12}, which satisfies $E[\varphi'\varphi]<+\infty$, $E[\varphi]=0$, $E[y_{12}\varphi']=0$ (L_2 will equal $[c_{\bar m_1},\cdots,c_{\bar m_r}]$, see (26)).

Step 3. Calculate the eigenvalues of R_1R_2, the μ_i's of Lemma 1, and consider T_iy_i in place of y_i in accordance with Lemma 1. Check whether for some nonnegative integers $m_1,\cdots,m_k$, not all zero $\mu_1^{m_1},\cdots,\mu_k^{m_k}$ is the inverse of some eigenvalue of R_1R_2. This check will stop in a finite number of steps, since $\mu_1^{m_1}\cdots\mu_k^{m_k}$ goes to zero because at least one of the m_i's increases. Let $\mu_1^{m_1}\cdots\mu_k^{m_k}$ be an inverse eigenvalue of R_1R_2 with corresponding eigenvectors $c_{\bar n_1},\cdots,c_{\bar n_f}$. If μ_i corresponds to the ith component of y_{112}, z_i (see Lemma 2 (5)) consider

$$u_d=\sum_{k=1}^{f} c_{\bar n_k}p_{m_1}(z_1)\cdots p_{m_k}(z_k)\varphi_{\bar n_k}(y_{12}),$$

and set

$$u_c=\sum_d u_d \quad \text{(finite sum)},$$

(u_c corresponds to the third term of (26)). Then the solution of (12) is $u_1=u_L+u_0+u_c$.

It should be noticed that the part of u_1 which depends on y_{111} is determined uniquely and is linear in y_{111}, because every $\bar p_n$ which depends only on y_{111} has $a_n=0$, and thus the corresponding c_n is either equal to s_n or 0.

Let us now consider the impact of the possible nonunique solutions for (u_1, u_2) to the costs J_1, J_2 of the Nash game. Let $Q_{ij}=0, F_i=0, i\neq j$. If (u_1, u_2) is a solution, then using (9), we obtain

$$E[u_1'R_1u_2]=E[P_1(u_1'R_1u_2)]=E[u_1'R_1P_1u_2]$$
$$=E[u_1'(-u_1-S_1P_1x)]=-E[u_1'u_1]-E[u_1'S_1x].$$

Thus,

$$J_1^*(u_1, u_2)=-\tfrac{1}{2}E[u_1'u_1],$$

and from (26)

$$J_1^*(u_1, u_2)=-\tfrac{1}{2}\{\|c_0\|^2+\|c_1\|^2+\cdots+\|c_{n_1}\|^2$$
$$+\|c_{\bar{n}_1}\|^2+\cdots+\|c_{\bar{n}_q}\|^2+E[\phi'C'\cdot C\phi]\},\quad \text{where } C=[c_{\bar{m}_1}\cdots c_{\bar{m}_r}].$$

If Q_{12}, F_1 are not zero, we can again calculate J_1^*. J_1^* will have some more quadratic terms in the c_i's, but it will also include the term $E[\phi'C'R_2'Q_{12}R_2C\phi]$. It is clear that in the case of nonuniqueness of solutions, by choosing the c_n's, and ϕ appropriately, we can vary J_1^*, J_2^*. If Q_{12} and F_1 are not set equal to zero, but are chosen so as to convexify J_1^* as a function of the c_i's and ϕ, then the possibility of arbitrarily small J_1^* will be ruled out. It will then necessarily hold that $C'(-I+2R_2'Q_{12}R_2)C\geqq 0$ (i.e., $-I+2R_2'Q_{12}R_2\geqq 0$ on the null space of $I-R_1R_2$) and thus player 1 will choose $\phi=0$. Thus, if $C'(-I+2R_2'Q_{12}R_2)C\geqq 0$, a second level deterministic problem can be introduced where player 1 will determine the c_i's to find the best (for himself) out of the many Nash solutions. Of course additional restrictions will have to be imposed in order to guarantee the convexity of the second level deterministic problem.

It should also be noticed that Theorem 1 reveals the dependence of the solution of the Nash game upon the relation between the matrices which determine the information and the matrices which determine the cost. Obviously only C_1, C_2, R_1R_2 have an impact as far as it concerns the existence and uniqueness of the solution. Nonetheless, the matrices Q_{ij}, F_i have an influence in the choice of the best out of the many Nash solutions when the second level optimization problem is solved and this influence can be very drastic (see § 6).

In Appendix C we present some useful results concerning the operator $I-R_1R_2P_1P_2$, which is obviously of central importance in our analysis.

4. Solution of problem S. The purpose of this section is to solve Problem S. Our development will be brief in view of the analysis of § 3.

For a given u_1 the follower solves problem (10), finds u_2 in terms of u_1 (see (11)), and u_2 is substituted in J_1 so that the leader has to solve

$$\begin{aligned}\underset{u_1\in U_1}{\text{minimize}}\, J(u_1)=E[&\tfrac{1}{2}u_1'(I-(R_1R_2+R_2'R_1'+R_2'Q_{12}R_2)P_2)u_1\\ &+u_1'(S_1+P_2(-R_1S_2+R_2'Q_{12}S_2-R_2F_1))x\\ &+x'(\tfrac{1}{2}S_2'Q_{12}S_2-S_2'F_1)P_2x].\end{aligned}\tag{27}$$

To guarantee $\inf_{u_1} J(u_1)>-\infty$, we assume that J_1 is convex in u_1 for $u_1=P_1u_1$, i.e., we assume that

$$P_1\geqq RP_1P_2P_1,\tag{28}$$

where

$$R=R_1R_2+R_2'R_1'+R_2'Q_{12}R_2.\tag{29}$$

If (28) holds, then u_1 is a solution if and only if

$$u_1 - RP_1P_2u_1 + P_1(S_1 + P_2(-R_1S_2 + R_2'Q_{12}S_2 - R_2F_1))x = 0, \tag{30}$$

which is of exactly the same type as (12) and thus all the analysis of § 3 carries over. The only question pertaining to Problem S in particular is present in assumption (28). Using (21) and Lemma 2 we obtain the following relation equivalent to (28):

$$\sum_{n=0}^{\infty} \|c_n\|^2 \geqq \sum_{n=0}^{\infty} c_n'(a_nR)c_n, \quad \forall c_0, c_1, \cdots, \quad \text{with } \sum_{n=0}^{\infty} \|c_n\|^2 < +\infty,$$

or

$$a_nR \leqq I, \qquad n = 0, 1, 2, \cdots, \tag{31}$$

and since $a_0 = 1$, we conclude that (28) is equivalent to

$$R \leqq I. \tag{32}$$

(32) could have been deduced directly from (28) by allowing u = constant, but (31) can be useful if we decide to restrict u. For example, if u is restricted to being a nonlinear function of y_{112}, only, then using (B-5) and (37), a weaker condition which will involve the μ_i's, can be substituted for (32).

In light of the discussion above and the analysis of § 3 we can easily conclude the following concerning the leader-follower game:

(i) If $R < I$, then there is a unique solution and it is linear in the information (since if $R < I$ no inverse eigenvalue of R can be equal to some $\mu_1^{m_1} \cdots \mu_k^{m_k}$).

(ii) If there exists a solution there will exist a solution affine in the information.

(iii) If R has some eigenvalue equal to 1 then the general solution of (30), if it exists, will be of the form $l_0 + L_1y_1 + L_2\varphi(y_{12})$, where $l_0 + L_1y_1$ is a solution and the columns of L_2 constitute a basis for the null space of $I - R$.

(iv) If (32) does not hold then $\inf J_1 = -\infty$ (since if $u = c \in R^{m_1}$, where $(I-R)c = \lambda c$, $\lambda < 0$, then $J_1 \to -\infty$ as $\|c\| \to +\infty$).

Solving the leader-follower game is less demanding than solving the Nash game since the calculation of $\mu_1, \cdots, \mu_k$ is not necessary. Of course, one needs to find y_{12} if $I - R$ is singular.

It is clear that the discussion in § 3 about different values of the cost induced by different solutions, convexification and a second level game, carry over to the leader-follower game as well.

5. A sufficient condition, extensions to dynamic cases. It is an immediate consequence of the analysis of § 3 that, if R_1R_2 has no eigenvalues in $[1, +\infty)$, then the Nash problem admits a unique solution which will be linear in the information. This can be proved independently as a consequence of Theorem 2 below. Theorem 2 is related to [5, Thm. 1], which is a special case of our Theorem 2. In addition our proof of Theorem 2 is much simpler than the one given in [5].

THEOREM 2. *Let H be a Hilbert space over the complex numbers and P an orthogonal projection in H. Let $Q: H \to H$ be a continuous linear operator (P and Q do not necessarily commute) and v an element of H. Then, a sufficient condition that the equation*

$$PQu + Pv = 0, \qquad Pu = u \tag{33}$$

have a unique solution $u \in H$ is that there exist a continuous linear invertible operator

$E: H \to H$ which commutes with P, and that the following holds:

$$QE^* + EQ^* \geqq I \quad \text{on } PH. \tag{34}$$

If (34) *holds, then the solution of* (33) *is given by*

$$u = P \sum_{n=0}^{\infty} [(I - \bar{E}^{-1}Q)P]^n \bar{E}^{-1} v, \tag{35}$$

where $\bar{E} = \delta E$ for any $\delta > \|QQ^*\|$.

Proof. The requirement that u solve (33) is equivalent to the requirement that u solve

$$PQu - PEu + Eu + Pv = 0 \tag{36}$$

for some E, as in the statement of the theorem. Obviously, (33) implies (36). Conversely, if (36) holds, applying P to both sides of (36) yields

$$PQu + Pv = 0, \tag{37}$$

and (36) together with (37) implies $-PEu + Eu = 0$, i.e., $Pu = u$. (36) can be written as

$$[I - P[I - E^{-1}Q]] + u + PE^{-1}v = 0. \tag{38}$$

A sufficient condition that (38) have a unique solution given by (35) is that

$$\|P[I - E^{-1}Q]\| < 1,$$

or equivalently

$$P[E^{-1}Q - I][Q^*E^{-1^*} - I]P \leqq (1 - \varepsilon)I \quad \text{on } H, \tag{39}$$

for some ε, $0 < \varepsilon < 1$. Taking into consideration the fact that we can multiply E by any $\delta > 0$, we can easily conclude that (39) is equivalent to (34). □

Condition (34) holds if $Q = Q' > 0$, is a real matrix, if we choose $E = \varepsilon I$, where ε is sufficiently large and positive. This special case was proved in [5] by more complicated arguments.

To apply Theorem 2 to the Nash game we first bring R_1R_2 into its Jordan form, $TR_1R_2T^{-1} = J$ and let $u = Tu_1$, $v = TSy_1$. If u_i is a component of u and ρ_i an eigenvalue of R_1R_2, it suffices to be able to invert the operator $1 - \rho_i P_1 P_2$. The role of Q in Theorem 2 will be now played by $1 - \rho_i P_2$. Taking E to be any complex number $\neq 0$ and using (34), we conclude that if $\rho_i \notin [1, +\infty)$ then the solution of (12) exists, is unique and linear in the information. (It should be pointed out that J does not need to be diagonal: if a 2×2 block of J involves the eigenvalue ρ and has a 1 in the upper right corner, then we first invert the $1 - \rho P_1 P_2$ associated with the component of u, u_i corresponding to the bottom row of this block and move to the above row in order to solve for the other component of u, u_{i-1}; the 1 of the Jordan block multiplies u_i which is already known, and so we have to invert $1 - \rho P_1 P_2$ again.)

Another application of Theorem 2 is in the study of equations of the form

$$PQu + Pv = 0, \qquad Pu = u, \tag{40}$$

where $H = H_1 \oplus H_2 \oplus \cdots \oplus H_n$, $P = \operatorname{diag}[P_1, \cdots, P_n]$, Q is a real matrix, $v \in H$ and P_i is the projection of H onto H_i. Such an equation will appear if we consider the n-player Nash game instead of the 2-player game of § 2. It will also appear in the study of dynamic linear quadratic Nash games with noisy linear state measurements, a discrete time evolution equation and appropriate nestedness conditions (see [4]) on the information of the players. Application of Theorem 2 to (40) yields that if there exists a real matrix $E=\operatorname{diag}[E_1,\cdots,E_n]$, $E_i:H_i\to H_i$, E_i, with

$$EQ^*+QE^*>0, \tag{41}$$

then (40) admits a unique solution. Of course, if the H_i's admit Hermite polynomials as a complete set of orthonormal eigenvectors, one can follow a procedure identical to the one of § 3, but the formulae derived will be quite complicated. Finally notice that if $n=2$ and

$$Q=\begin{bmatrix} I & R_1 \\ R_2 & I \end{bmatrix},$$

then (40) represents another way of writing (9) and (11). Application of the condition (41) is possible, but the result will be weaker than the one derived by first transforming (12) into its Jordan form and then applying Theorem 2.

In the rest of this section we will show how some dynamic problems can be reduced to static ones (see also [4]), and how one can solve them. Let

$$x_{k+1}=A_kx_k+B_k^1u_k^1+B_k^2u_k^2+w_k, \qquad k=0,1,\cdots,N,$$

$$y_k^i=C_k^ix_k+v_k^i, \qquad i=1,2$$

$$J_i(u_1,u_2)=E\left[x_{N+1}'Q_{N+1}^ix_{N+1}+\sum_{k=0}^N x_k'Q_k^ix_k+u_k^{i'}u_k^i+u_k^{j'}R_k^{ij}u_k^j\right], \qquad i,j=1,2, \quad i\neq j.$$

x_0, w_k, v_k^i are independent Gaussian random variables with nonsingular covariance matrices. The real matrices $Q_k^i=Q_k^{i'}\geqq 0$, R_k^{ij}, A_k, B_k^i,C_k^i have appropriate dimensions, $x_k\in R^n$, $u_k^i\in R^{n_i}$, $y_k^i\in R^{m_i}$ and $u_i=(u_0^i,u_1^i,\cdots,u_k^i)'$. Player 1 chooses u_k^1 as a function of $(y_0^1,y_1^1,\cdots,y_k^1,y_0^2,\cdots,y_{k-1}^2)$, and player 2 chooses u_k^2 as a function of $(y_0^1,\cdots,y_{k-1}^1,y_0^2,\cdots,y_k^2)$. Using the evolution equation, we can express the J_i's as quadratic functions of x_0, w_k, v_k^i, u_k^1, u_k^2 and the y_k^i's as functions of x_0, $w_0,\cdots,w_{k-1}$, $v_0^1,\cdots,v_k^1$, $v_0^2,\cdots,v_{k-1}^2$ and $u_0^1,\cdots,u_{k-1}^1$, $u_0^2,\cdots,u_{k-1}^2$. Because of the nestedness of the information we can do away with the presence of the u_l^i's $0\leqq l\leqq k-1$, in the expression for y_k^1, and similarly for y_k^2. Let $x=(x_0,w_0,\cdots,w_N,v_0^1,\cdots,v_N^1,v_0^2,\cdots,v_N^2)$. We have thus transformed our problem into the following:

$$J_1(u_1,u_2)=E[u_1'Q_{11}u_1+u_1'R_{12}u_2+u_1'S_1x+u_2'Q_{12}u_2+u_2'F_1x+x'L_1x],$$

$$J_2(u_1,u_2)=E[u_2'Q_{22}u_2+u_2'R_{21}u_1+u_2'S_2x+u_1'Q_{21}u_1+u_1'F_2x+x'L_2x],$$

$$y_{10}=C_{10}x, \qquad y_{11}=C_{11}x,\cdots,y_{1N}=C_{1N}x,$$

$$y_{20}=C_{20}x, \qquad y_{21}=C_{21}x,\cdots,y_{2N}=C_{2N}x.$$

Let

$$P_{1k}=E[\cdot\,|\,y_0^1,\cdots,y_k^1,y_0^2,\cdots,y_{(k-1)}^2],$$

$$P_{2k}=E[\cdot\,|\,y_0^1,\cdots,y_{(k-1)}^1,y_0^2,\cdots,y_k^2],$$

$$P_1=\operatorname{diag}[P_{10},\cdots,P_{1N}], \qquad P_2=\operatorname{diag}[P_{20},\cdots,P_{2N}].$$

($P_{1k}P_{2l}=P_{2l}$, $P_{2k}P_{1l}=P_{1l}$ if $l\leqq k-1$ and $P_{1k}P_{1l}=P_{1l}$, $P_{2k}P_{2l}=P_{2l}$ if $l\leqq k$.) If we are interested in the Nash solution we can write down the analogues of (9) and (11), which can be viewed together, as an equation of the type (40). We can thus either apply a generalization of the analysis of § 3 or settle for less and use Theorem 2. (If we are interested in the leader-follower solution, we have to assume that $y_k^1=0$ for $k\geqq 1$.)

6. Example of a Nash problem. In this section we will solve a one-dimensional Nash problem. Let

$$y_1 = x_1, \qquad y_2 = x_1 + ax_2.$$

x_1, x_2 are normally distributed, independent Gaussian random variables and $a \neq 0$. Let θ be the absolute value of the correlation coefficient of y_1, y_2, $\theta = (1+a^2)^{-1}$, $0<\theta<1$,

$$J_1(u_1, u_2) = E[\tfrac{1}{2}u_1^2 + r_1u_1u_2 + u_1(s_{11}x_1 + s_{12}x_2) + \tfrac{1}{2}q_1u_2^2 + u_2(t_{11}x_1 + t_{12}x_2)],$$

$$J_2(u_1, u_2) = E[\tfrac{1}{2}u_2^2 + r_2u_1u_2 + u_2(s_{21}x_1 + s_{22}x_2) + \tfrac{1}{2}q_2u_1^2 + u_2(t_{21}x_1 + t_{22}x_2)],$$

where r_i, s_{ij}, q_i are reals. (12) assumes the form

$$u_1 - \rho E[E[u_1 | x_1 + ax_2] | x_1] = sy_1, \tag{42}$$

where

$$\rho = r_1r_2, \qquad s = r_1(s_{21} + as_{22}) - s_{11}.$$

Let $u = \sum_{n=0}^{\infty} c_np_n(y_1)$, where the p_n are the one-dimensional Hermite polynomials (see [9]), and $\sum c_n^2 < +\infty$. A straightforward application of Theorem 1 yields the following. Consider the equations for the c_n's,

$$c_n(1-\rho\theta^n) = \begin{cases} s & \text{if } n=1, \\ 0 & \text{if } n = 0, 2, 3, \cdots. \end{cases}$$

If:

(i) $1 \neq \rho\theta^n$, $n = 0, 1, 2, \cdots$, then the solution exists, is unique and is given by $u_1(y_1) = s(1-\rho\theta)^{-1}y_1$.

(ii) $1 \neq \rho\theta$, but $\rho\theta^n = 1$ for some $n = 0, 2, 3, \cdots$, then the solution is $u_1(y_1) = s(1-\rho\theta)^{-1}y_1 + cp_n(y_1)$, c arbitrary and real.

(iii) $1 = \rho\theta$ and $s = 0$, then the solution is $u_1 = (y_1) = ly_1$, l is arbitrary and real.

(iv) $1 = \rho\theta$ and $s \neq 0$, then there is no solution.

If $1 = \rho\theta^n$ for some $n \geqq 2$, then case (ii) holds and an easy calculation shows that

$$J_1 = \tfrac{1}{2}c^2[-1 + q_1\theta^n r_2^2] + \text{constant},$$

since $\theta^n = 1/(r_1r_2) > 0$. We conclude that if $r_1/r_2 > q_1$ player 1 can make his cost arbitrarily small for sufficiently large c. If $r_1/r_2 < q_1$ he will do well to choose $c = 0$. If both $r_1/r_2 < q_1$ and $r_2/r_1 < q_2$, hold then both players will agree on $c = 0$ (or on c sufficiently small if $r_1/r_2 > q_1$ and $r_2/r_1 > q_2$). If $r_1/r_2 = q_1$, then player 1 does not care about c. Conflict will arise about the choice of c if $r_1/r_2 > q_1$ and $r_2/r_1 < q_2$, in which case player 1 will want c as big as possible whereas player 2 will want $c = 0$. If player 1 is faster than 2, he calculates his u_1 through (42) first, realizes the possibility of choosing c arbitrarily and by declaring his decision he forces player 2 to use (11) to find his decision and thus player 1 imposes his choice of c. Therefore, the case of nonunique Nash solutions carries hidden in it the concept of the leader-follower game. Finally, notice that if J_1 is convex in u_1, u_2, i.e., $q_i \geqq r_1^2$, since $r_1r_2 = 1/\theta^n > 1$, then we obtain $1/r_2 < r_1$ and thus $r_1/r_2 < r_1^2 \leqq q_1$, i.e., $q_1 \geqq r_1/r_2$; therefore player 1 will prefer $c = 0$, in agreement with the fact that the convexity of J_1 cannot permit it to go to $-\infty$. Nonetheless, it might very well be that $r_1/r_2 < q_1 < r_1^2$ in which case player 1 will again prefer $c = 0$, although J_1 is not convex in u_1 and u_2, i.e., he cannot make J_1 arbitrarily small although J_1 is not convex in (u_1, u_2). This situation is due to the fact that what matters is the convexity of J_1 in $u_1 = P_1u_1$, $u_2 = P_2u_2$, i.e., convexity on some subspace and this convexity is guaranteed by $q_1 > r_1/r_2$.

The condition $1=\rho\theta^n$, i.e., $\theta=$ correlation coefficient of y_1, $y_2=(r_1r_2)^{-n}$, is critical. r_1r_2 can be interpreted as the coupling of J_1, J_2, whereas θ is the coupling of the information. We can thus interpret the condition $1=\rho\theta^n$, as saying that if the coupling of the information equals the inverse of some power $n\neq 1$ of the coupling of the costs, the solution will be nonunique.

7. Conclusions. Here we will point out several directions in which the analysis presented can be generalized, or problems which suggest themselves for study within our framework.

The second level problems that have to be solved in the case of nonunique solutions, as discussed at the end of §§ 3 and 4 are of definite importance. In the Nash case, J_1^* is a quadratic function of the c_i's (we set $\phi=0$) and the constraints on the c_i's are finite, since the c_i's involved are finite in number. Thus player 1 is faced with a classical quadratic deterministic optimization problem subject to linear constraints. Although it is an easy problem, it merits special attention because it will provide the best Nash decision to player 1.

To generalize our analysis to the many player case one needs to extend Lemma 1 and Lemma 2. One can go one step further and allow different components of u_1 to have different information or even more, one can study equations of the form $PQu+Pv=0$, where $P=\operatorname{diag}[P_1,\cdots,P_n]$, $\bar{P}u=u$, $\bar{P}=\operatorname{diag}[\bar{P}_1,\cdots,\bar{P}_k]$, where P_i, $\bar{P}_i$ are projections. Such extensions are important in order to be able to handle dynamic games with nested information structures (although conceptually they are covered by the methods presented here).

Another interesting problem whose study lies within the capabilities of the methods presented, is the one where one leader is followed by two followers, which followers play Nash.

Appendix A: Proof of Lemma 1. Let $R(C)$ denote the range of a matrix C. All bases to be mentioned are orthonormal. Let the rows of $\bar{C}_{12}=\bar{C}_{21}$ be a basis for $R(C_1')\cap R(C_2')$. Let the rows of $\bar{C}_{111}$ be a basis for $R(C_1')\cap R(C_2')^{\perp}$. Let the rows of $\bar{C}_{222}$ be a basis for $R(C_1')^{\perp}\cap R(C_1')$. Choose $\bar{C}_{112}$ so that its rows together with those of $\bar{C}_{111}$ and $\bar{C}_{12}$ constitute a basis for $R(C_1')$. Choose $\bar{C}_{221}$ so that its rows together with those of $\bar{C}_{222}$ and $\bar{C}_{21}$ constitute a basis for $R(C_2')$. This construction proves (1), (2) and (4). Let us concentrate on $\bar{C}_{112}$, $\bar{C}_{221}$, which we will denote by D_1, D_2. If D_1 is $k_1\times n$, D_2 is $k_2\times n$ and $k_1\neq k_2$, let $k_1>k_2$, without loss of generality. Then there are nonsingular square matrices L_1, L_2 so that $L_1D_1D_2'L_2'$ will have its last row equal to zero, which means that the last row of L_1D_1 is an element of $R(D_1')$ perpendicular to $R(D_2')$. Such elements, nonetheless, were put in $R(\bar{C}_{111}')$ and thus k_1 cannot be strictly greater than k_2. Reversing the roles of k_1 and k_2 we conclude that $k_1=k_2=k$ and that D_1D_2' is a square nonsingular matrix. Let Λ be the diagonal Jordan equivalent of D_1D_1', i.e.,

$$(D_1D_1')U=U\Lambda,$$

where U is the matrix of the orthogonal eigenvectors. Let M be the diagonal Jordan equivalent for which

$$(\Lambda^{-1/2}U'D_1D_2'(D_2D_2')^{-1}D_2D_1'U\Lambda^{-1/2})V=VM,$$

where V is the matrix of the orthogonal eigenvectors. Let

$$\bar{D}_1=V'\Lambda^{-1/2}U'D_1,$$
$$\bar{D}_2=M^{-1/2}V'\Lambda^{-1/2}U'D_1D_2'(D_2D_2')^{-1}D_2.$$

It can be verified that

$$\bar{D}_1\bar{D}_1' = I, \qquad \bar{D}_2\bar{D}_2' = I, \qquad \bar{D}_1\bar{D}_2' = M^{1/2},$$

so we can use $\bar{D}_1, \bar{D}_2$ for $\bar{C}_{112}, \bar{C}_{221}$. $R(\bar{D}_1)$ and $R(\bar{D}_2)$ have no common elements, since if they had one, it should have been placed in $R(\bar{C}_{12}')$ from the beginning. M is diagonal has positive elements, and each $\sqrt{\mu_{ii}}$ is the product of two nonidentical unit length vectors (rows of $\bar{D}_1$, $\bar{D}_2$). Thus $0<\sqrt{\mu_{ii}}<1$. □

Appendix B: Proof of Lemma 2.

$$a_{nl} = E[\bar{p}_n P_1 P_2 \bar{p}_l] = E[\bar{p}_n P_2 \bar{p}_l] = E[(P_2\bar{p}_n)\cdot \bar{p}_l] = a_{ln},$$

since P_i is self-adjoint and $P_1\bar{p}_n = \bar{p}_n$. $\bar{p}_n$ and $\bar{p}_l$ have the form

$$\bar{p}_n(y_1) = \bar{p}_{n_1}(y_{111})\bar{p}_{n_2}(y_{112})\bar{p}_{n_3}(y_{12}),$$
$$\bar{p}_l(y_1) = \bar{p}_{l_1}(y_{111})\bar{p}_{l_2}(y_{112})\bar{p}_{l_3}(y_{12}).$$

Using the independence of some of the components of y_1, y_2, we have

$$\begin{aligned}
a_{nl} &= E[\bar{p}_n(y_1)P_1P_2\bar{p}_l(y_1)]\\
&= E[\bar{p}_n(y_1)P_2\bar{p}_l(y_1)]\\
&= E[\bar{p}_n(y_1)E[\bar{p}_{l_1}(y_{111})\bar{p}_{l_2}(y_{112})\bar{p}_{l_3}(y_{12})\,|\,y_{12}, y_{221}, y_{222}]]\\
&= E[\bar{p}_n(y_1)\bar{p}_{l_3}(y_{12})E[\bar{p}_{l_1}(y_{111})\bar{p}_{l_2}(y_{112})\,|\,y_{12}, y_{221}, y_{222}]]\\
&= E[\bar{p}_n(y_1)\bar{p}_{l_3}(y_{12})E[\bar{p}_{l_1}(y_{111})\bar{p}_{l_2}(y_{112})\,|\,y_{221}]]\\
\text{(B.1)}\quad &= E[\bar{p}_n(y_1)\bar{p}_{l_3}(y_{12})E[E[\bar{p}_{l_1}(y_{111})\bar{p}_{l_2}(y_{112})\,|\,y_{221}, y_{112}]\,|\,y_{221}]]\\
&= E[\bar{p}_n(y_1)\bar{p}_{l_3}(y_{12})E[\bar{p}_{l_2}(y_{112})E[\bar{p}_{l_1}(y_{111})\,|\,y_{221}, y_{112}]\,|\,y_{221}]]\\
&= E[\bar{p}_n(y_1)\bar{p}_{l_3}(y_{12})E[\bar{p}_{l_2}(y_{112})E[\bar{p}_{l_1}(y_{111})]\,|\,y_{221}]]\\
&= E[\bar{p}_n(y_1)\cdot \bar{p}_{l_3}(y_{12})E[p_{l_1}(y_{111})]\cdot E[\bar{p}_{l_2}(y_{112})\,|\,y_{221}]]\\
&= E[\bar{p}_{n_1}(y_{111})\bar{p}_{n_2}(y_{112})\bar{p}_{n_3}(y_{12})\bar{p}_{l_3}(y_{12})E[\bar{p}_{l_1}(y_{111})]\cdot E[\bar{p}_{l_2}(y_{112})\,|\,y_{221}]]\\
&= E[\bar{p}_{n_1}(y_{111})]\cdot E[\bar{p}_{l_2}(y_{111})]\cdot E[\bar{p}_{n_3}(y_{12})\bar{p}_{l_3}(y_{12})]\cdot E[\bar{p}_{n_2}(y_{112})E[\bar{p}_{l_2}(y_{112})\,|\,y_{221}]].
\end{aligned}$$

It holds that

$$\text{(B.2)}\qquad \begin{aligned}
E[\bar{p}_{n_1}(y_{111})] &= \begin{cases} 0 & \text{if } n_1 \neq 0,\\ 1 & \text{if } n_1 = 0,\end{cases}\\
E[\bar{p}_{n_3}(y_{12})\bar{p}_{l_3}(y_{12})] &= \begin{cases} 0 & \text{if } n_3 \neq l_3,\\ 1 & \text{if } n_3 = l_3.\end{cases}
\end{aligned}$$

Let

$$\bar{p}_{n_2}(y_{112}) = p_{m_1}(z_1)\cdots p_{m_k}(z_k),$$
$$\bar{p}_{l_2}(y_{112}) = p_{s_1}(z_1)\cdots p_{s_k}(z_k),$$

where $y_{112} = (z_1, \cdots, z_k)'$. Let $y_{221} = (w_1, \cdots, w_k)'$.

Since z_i depends only on w_i and vice versa (Lemma 1, (3)) we have

$$\begin{aligned}
&E[p_{s_1}(z_1)\cdots p_{s_k}(z_k)\,|\,w_1,\cdots,w_k]\\
&\quad=E[E[p_{s_1}(z_1)\cdots p_{s_k}(z_k)\,|\,z_1,w_1,\cdots,w_k]\,|\,w_1,\cdots,w_k]\\
&\quad=E[p_{s_1}(z_1)E[p_{s_2}(z_2)\cdots p_{s_k}(z_k)\,|\,z_1,w_1,\cdots,w_k]\,|\,w_1,\cdots,w_k]\\
&\quad=E[p_{s_1}(z_1)\,|\,w_1]\cdot E[p_{s_2}(z_2)\cdots p_{s_k}(z_k)\,|\,w_2,\cdots,w_k]\\
&\quad=\cdots=E[p_{s_1}(z_1)\,|\,w_1]\cdot E[p_{s_2}(z_2)\,|\,w_2]\cdots E[p_{s_k}(z_k)\,|\,w_k].
\end{aligned}$$

Therefore,

$$\begin{aligned}
&E[\bar p_{n_2}(y_{112})E[\bar p_{l_2}(y_{112})\,|\,y_{221}]]\\
\text{(B.3)}\qquad&\quad=E[p_{m_1}(z_1)\cdots p_{m_k}(z_k)E[p_{s_1}(z_1)\,|\,w_1]\cdots E[p_{s_k}(z_k)\,|\,w_k]]\\
&\quad=E[p_{m_1}(z_1)E[p_{s_1}(z_1)|w_1]]\cdots E[p_{m_k}(z_k)E[p_{s_k}(z_k)|w_k]].
\end{aligned}$$

Since $E[z_1w_1]=\sqrt{\mu_1}$, $E[p_{s_1}(z_1)\,|\,w_1]$ will be a polynomial of order s_1 in w_1 and the leading coefficient of this polynomial will be the leading coefficient of $p_{s_1}(z_1)$ multiplied by $(\sqrt{\mu_1})^{s_1}$. Similarly, $E[E[p_{s_1}(z_1)\,|\,w_1]\,|\,z_1]$ will be a polynomial in z_1 of order s_1, with the leading coefficient of p_{s_1} multiplied by $(\sqrt{\mu_1})^{2s_1}$. Thus $E[E[p_{s_1}(z_1)\,|\,w_1]\,|\,z_1]=\mu_1^{s_1}p_{s_1}(z_1)+$ Hermite polynomials in z_1 of order strictly less than s_1. Thus we conclude that

$$\begin{aligned}
E[p_{m_1}(z_1)\cdot E[p_{s_1}(z_1)\,|\,w_1]]&=E[p_{m_1}(z_1)E[E[p_{s_1}(z_1)\,|\,w_1]\,|\,z_1]]\\
&=\begin{cases}0 & \text{if } s_1<m_1,\\ \mu^{m_1} & \text{if } m_1=s_1.\end{cases}
\end{aligned}$$

Since $E[p_{m_1}(z_1)E[p_{s_1}(z_1)\,|\,w_1]]=E[p_{s_1}(z_1)E[p_{m_1}(z_1)\,|\,w_1]]$, we conclude that

$$\text{(B.4)}\qquad E[p_{m_1}(z_1)E[p_{s_1}(z_1)\,|\,w_1]]=\begin{cases}0 & \text{if } m_1\neq s_1,\\ \mu_1 & \text{if } m_1=s_1.\end{cases}$$

From (B.3), (B.4) we now obtain

$$\text{(B.5)}\qquad E[p_{n_2}(y_{112})E[p_{l_2}(y_{112})\,|\,y_{221}]]=\begin{cases}0 & \text{if } n_2\neq l_2,\\ \mu_1^{m_1}\cdots\mu_k^{m_k} & \text{if } n_2=l_2 \text{ and}\\ & p_{n_2}(y_{112})=p_{m_1}(z_1)\cdots p_{m_k}(z_k).\end{cases}$$

Equations (B.1), (B.2) and (B.5) prove (2)–(5).

Let us now prove (6). To find $\|P_{112}P_{221}\|$ we will calculate $P_{112}P_{221}u$. u can be restricted to depend only on y_{112} and thus $u=\sum_{n=1}^{\infty}c_n\bar p_n(y_{112})$, where $\sum_{n=1}^{\infty}\|c_n\|^2<\infty$, ($c_0=0$ so that $E[u]=0$).

$$\begin{aligned}
\|P_{112}P_{221}u\|&=E[(P_{112}P_{221}u)'\cdot P_{112}P_{221}u]\\
&=E[u'\cdot P_{221}P_{112}P_{221}u]\\
&=\sum_{n,l\geqq 1}^{\infty}c_n'c_lE[\bar p_n(y_{112})P_{221}P_{112}P_{221}\bar p_l(y_{112})]\\
&=\sum_{n,l\geqq 1}c_n'c_l\bar a_{nl}.
\end{aligned}$$

An argument similar to the one used before shows that $\bar a_{nl}=0$ if $n\neq l$, and if $n=l$,

$$\bar a_{nn}=\bar a_n\doteq\mu_1^{2m_1}\cdots\mu_k^{2m_k}=a_n^2.$$

Thus, $\|P_{112}P_{221}u\|^2=\sum_{n=1}^{\infty}\|c_n\|^2a_n^2$. Also,

$$\|u\|^2=\sum_{n,l\geqq 1} c'_n c_l E[\bar{p}_n\cdot p_l]=\sum_{n=1}^{\infty}\|c_n\|^2,$$

and therefore

$$\frac{\|P_{112}P_{221}u\|^2}{\|u\|^2}=\frac{\sum_{n=1}^{\infty}\|c_n\|^2a_n^2}{\sum_{n=1}^{\infty}\|c_n\|^2}.$$

Obviously $\|P_{112}P_{221}\|=\sup a_n$, and since a_n decreases, because $0<\mu_i<1$, we conclude $\|P_{112}P_{221}\|=\max\{\mu_1,\cdots,\mu_k\}$ 1. For reasons of symmetry, $\|P_{221}P_{112}\|=\max\{\mu_1,\cdots,\mu_k\}=\|P_{112}P_{221}\|$. □

Appendix C: The operator $I-RP_1P_2$. From the analysis of § 3 it is obvious that if instead of having to solve (14) we had to solve

$$(I-RP_1P_2)u=v, \tag{C.1}$$

where $v\in U_1$ (and thus $v=\sum_{j=0}^{\infty}d_j\bar{p}_j$, $d_j\in R^{m_1}$, $\sum\|d_j\|^2<+\infty$), we would end up with the equivalent system of linear equations

$$(I-a_iR)c_i=d_i,\qquad i=0,1,\cdots. \tag{C.2}$$

If (C.1) has a solution $c_0, c_1,\cdots$, with $c_j\in\mathbb{R}^{m_1}$ then

$$u=\sum_{j=1}^{\infty}c_jp_j,\qquad \sum_{j=1}^{\infty}\|c_j\|^2<+\infty,$$

is a solution of (C.2). Therefore the R's for which $I-RP_1P_2$ is invertible are those which do not have any of the $1/a_n$'s (for $a_n\neq 0$) as eigenvalues.

To find $\|I-RP_1P_2\|$, let $u=\sum_{n=0}^{\infty}c_n\bar{p}_n$. Thus $\|u\|^2=\sum_{n=0}^{\infty}c_n^2<\infty$,

$$\begin{aligned}\|(I-RP_1P_2)u\|^2&=E[u'u+u'R'RP_1P_2P_1u-2u'RP_1P_2u]\\&=\sum_{n=0}^{\infty}(\|c_n\|^2+a_n^2c'_nR'Rc_n-2a_nc'_nRc_n)\\&=\sum_{n=0}^{\infty}c'_n(I-a_nR)'(I-a_nR)c_n\leqq\left(\sum_{n=0}^{\infty}\|c_n\|^2\right)\sup_n\|I-a_nR\|^2,\end{aligned}$$

and obviously $\|I-RP_1P_2\|=\sup_n\|I-a_nR\|$. If $(I-RP_1P_2)^{-1}$ exists, to find $\|(I-RP_1P_2)^{-1}\|$, let $v=\sum_{n=0}^{\infty}d_n\bar{p}_n$, $\|v\|^2=\sum_{n=0}^{\infty}\|d_n\|^2<+\infty$. Then $(I-RP_1P_2)^{-1}v=\sum_{n=0}^{\infty}(I-a_nR)^{-1}d_p\bar{p}_n=\sum_{n=0}^{\infty}c_n\bar{p}_n$, $\|(I-RP_1P_2)^{-1}v\|^2=\sum_{n=0}^{\infty}\|C_n\|^2=\sum_{n=0}^{\infty}d'_n(I-a_nR)^{-1'}\cdot(I-a_nR)^{-1}d_n$. Thus $\|(I-RP_1P_2)^{-1}\|=\sup_n\|(I-a_nR)^{-1}\|$. (It is easy to see that if $(I-a_nR)^{-1}$ exists for all a_n, then $\sup_n(\|(I-a_nR)^{-1}\|)<+\infty$.)

Let us formalize this discussion into a proposition.

PROPOSITION 1.

(1) spectrum $(RP_1P_2)=\{a_nr;\ r=\textit{eigenvalue of } R,\ n=0,1,2,\cdots\}$.

(2) $\|I-RP_1P_2\|=\sup_n\|I-a_nR\|$.

(3) $(I-RP_1P_2)^{-1}$ *exists, if and only if* $1\neq a_nr$ *for all* n *and then* $\|(I-RP_1P_2)^{-1}\|=\sup_n\|(I-a_nR)^{-1}\|$.

We can use (3) in the case where we have to solve for u the equation $(I-RP_1P_2)u+f(u)=v$, where $\|f(u)-f(\bar{u})\|\leqq L\|u-\bar{u}\|$ and $f:U_1\to U_1$, $v\in U_1$. If $L<\inf_n[\|(I-a_nR)^{-1}\|^{-1}]$ the contraction mapping theorem is applicable and yields existence and uniqueness of a solution. Equations of this form can arise when the cost J_1 is nonlinear in u_1, u_2.

REFERENCES

[1] R. RADNER, *Team decision problems*, Ann. Math. Statis., 33 (1962), pp. 857–881.

[2] A. W. STARR AND Y. C. HO, *Nonzero-sum differential games*, J. Optim. Theory Appl., 3 (1969), pp. 184–206.

[3] J. B. CRUZ Jr., *Leader-follower strategies for multilevel systems*, IEEE Trans. Automat. Control, AC-23 (1978), pp. 244–255.

[4] Y. C. HO AND K. C. CHU, *Team decision theory and information structures in optimal control problems—part I*, IEEE Trans. Automat. Control, AC-17 (1972), pp. 15–22.

[5] J. F. RUDGE, *Series solutions to static team control problems*, Math. Oper. Res., 1 (1976), pp. 67–81.

[6] T. BASAR, *Equilibrium solutions in two-person quadratic decision problems with static information structures*, IEEE Trans. Automat. Control, AC-20 (1975), pp. 320–328.

[7] ———, *Decentralized multicriteria optimization of linear stochastic systems*, IEEE Trans. Automat. Control, AC-23 (1978), pp. 233–243.

[8] R. B. ASH, *Real Analysis and Probability*, Academic Press, New York, 1972.

[9] K. R. PARTHASARATHY, *Introduction to Probability and Measure*, Springer-Verlag, New York, 1977.

[10] P. R. HALMOS, *A Hilbert Space Problem Book*, Van Nostrand, Princeton, NJ, 1967.

[11] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics*, vol. I, Interscience Publishers, New York, 1953.

[12] T. W. ANDERSON, *An Introduction to Multivariate Statistical Analysis*, John Wiley, New York, 1958.

[13] Y. C. HO, P. B. LUH AND R. MURALIDHARAN, *Information structure: Stackelberg games and incentive controllability*, IEEE Trans. Automat. Control, (1981), to appear.