References

[1] R. Isaacs, "Differential games I, II, III, IV," RAND Corp., Santa Monica, Calif., Research Memos. RM-1391, 1399, 1411, 1468, 1954–1956.

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

[3] L. D. Berkowitz and W. H. Fleming, "On differential games with integral payoff," in *Annals of Mathematical Studies*, no. 39. Princeton, N. J.: Princeton University Press, 1957, pp. 413–435.

[4] L. D. Berkowitz, "A variational approach to differential games," in *Annals of Mathematical Studies*, no. 52. Princeton, N. J.: Princeton University Press, 1964, pp. 127–173.

[5] Y. C. Ho, A. E. Bryson, Jr., and S. Baron, "Differential games and optimal pursuit–evasion strategies," *IEEE Trans. Automatic Control*, vol. AC-10, pp. 385–389, October 1965.

[6] Y. C. Ho, "Optimal terminal maneuver and evasion strategy," *SIAM J. Control*, vol. 4, no. 3, pp. 421–428, 1966.

[7] J. L. Speyer, "A stochastic differential game with controllable statistical parameters," *IEEE Trans. System Science and Cybernetics*, vol. SSC-3, pp. 17–20, June 1967.

[8] P. Joseph and J. Tou, "On linear control theory," *AIEE Trans. (Applications and Industry)*, vol. 80, pp. 193–196, September 1961.

[9] T. F. Gunckel and G. F. Franklin, "A general solution for linear sampled-data control systems," *Trans. ASME, J. Basic Engrg.*, vol. 85, no. 2, pp. 197–203, 1963.

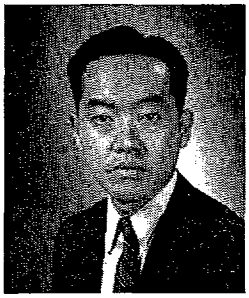
[10] R. D. Behn and Y. C. Ho, "On a class of linear stochastic differential games," Harvard University, Cambridge, Mass., Tech. Rept. 542, October 1967.

[11] R. Bellman, *Introduction to Matrix Analysis*. New York: McGraw-Hill, 1960, p. 99, exercise 7.

Robert D. Behn (S'67) was born in Washington, D. C., on September 5, 1941. He received the B.S. degree in physics from Worcester Polytechnic Institute, Worcester, Mass., in 1963, and the M.S. degree in engineering from Harvard University, Cambridge, Mass., in 1965. He is currently a Ph.D. candidate at Harvard. University.

He has been with the RAND Corp., Santa Monica, Calif., and M.I.T. Lincoln Laboratory, Lexington, Mass., where he worked on defense decision problems. He is currently a Research Assistant at Harvard University.

Mr. Behn is a member of Tau Beta Pi and Sigma Xi.



Yu-Chi Ho (S'54–M'55–SM'62) was born in China on March 1, 1934. He received the S.B.E.E. and the S.M.E.E. degrees from the Massachusetts Institute of Technology, Cambridge, in 1953 and 1955, respectively, and the Ph.D. degree in applied mathematics from Harvard University, Cambridge, Mass., in 1961.

From 1955 to 1958, he worked at the Research Laboratories Div., Bendix Corp., Southfield, Mich. Since 1961, he has been on the faculty of the Div. of Engineering and Applied Physics, Harvard University, and is now Associate Professor of Engineering and Applied Mathematics. He is a consultant to various organizations and he holds four patents in numerical control.

Dr. Ho is a member of Sigma Xi and the Society for Industrial and Applied Mathematics.

A Direct Method for a Class of Optimal Control Problems

SANG H. KYONG MEMBER, IEEE, AND ELIAS P. GYFTOPOULOS

Abstract—**A direct method is developed for the solution of a class of minimum energy control problems. The method is applicable to linear and nonlinear, stationary and time-varying systems described by input–output functional relations. It is based on the expansion of the kernels of the system and of the input, the control, in terms of a set of functions that are characteristic of the kernels. The optimality is measured by the integral of a positive definite quadratic form of the input over the control time interval. The characteristic expansions reduce the optimal control problem to that of solving a finite set of algebraic equations.**

Manuscript received August 10, 1967; revised January 16, 1968. This work was supported principally by the National Science Foundation under Grant GK-1165 and in part by the Joint Services Electronics Program under Contract DA-36-039-AMC-03200E.

S. H. Kyong was with the Dept. of Nuclear Engineering and Research Lab. of Electronics, Massachusetts Institute of Technology, Cambridge, Mass. He is now with Bell Telephone Laboratories, Inc., Whippany, N. J.

E. P. Gyftopoulos is with the Dept. of Nuclear Engineering and Research Lab. of Electronics, Massachusetts Institute of Technology, Cambridge, Mass.

I. Introduction

Optimal control problems have been studied by considering the system dynamics as described by a set of differential equations.[1],[2] Many physical systems, however, are often identified by means of input–output measurements.[3]–[5] These measurements result, in general, in input–output functional relations. It appears, therefore, that the study of optimal control methods for systems representable by input–output functional relations has merit.

The purpose of this paper is to present a direct method for solving a class of minimum energy control problems. The method is applicable to linear and nonlinear, stationary and time-varying systems representable by input–output functional relations. It is based on the expansion of the kernels of the system and of the input or control variables in terms of a set of functions that are characteristic of the kernels. Thus the optimal control problem is reduced to that of solving a finite set of simultaneous algebraic equations.

The paper is organized as follows. In Section II, the class of systems and the type of optimal control problems under consideration are presented. In Section III, the concept of the characteristic set of given functions is defined and an expansion technique in terms of the characteristic set is outlined. In Section IV, the expansion technique is applied to an optimal control problem of a linear system. Simple illustrative examples are discussed. In Section V, the same technique is applied to a nonlinear optimal control problem and an illustrative example is also included.

II. The Problem

The class of systems considered in this paper is the class of realizable multi-input–multi-output continuous nonlinear systems, which are representable by input–output functional relations of the form

$$y(t) = \sum_{i=0}^{(i)} \sum_{j=1}^{(j)} \int_{t_0}^{t} \cdots \int_{t_0}^{t} h_{ij}(t; \tau_1, \cdots, \tau_i) \cdot C_j(u; \tau_1, \cdots, \tau_i) d\tau_1, \cdots, d\tau_i \tag{1}$$

where the output $y(t)$ and the input $u(t)$ are q and r vectors, respectively, the q vector $h_{ij}(t; \tau_1, \cdots, \tau_i)$ is the ith-order kernel of the nonlinear system, the scalar $C_j(u; \tau_1, \cdots, \tau_i)$ is the jth possible product of i components of the input vector $u(t)$, each of the components in the product having $\tau_1, \tau_2, \cdots, \tau_i$ as its argument, respectively. The integer[1] (i) is the finite number of functionals that characterize the system, and the integer (j) is the number of all possible combinations of r items per i, i.e.,

$$(j) = \frac{(i + r - 1)!}{i!(r - 1)!} \cdot \tag{2}$$

The relation given by (1) is called a Volterra series and has been studied extensively by various authors in connection with the representation and analysis of nonlinear systems.[6]–[13] In general, the representation given by (1) with finite (i) is only an approximation to the real system. The actual value of (i) required for a close approximation depends upon the nature of nonlinearity in the system; and this fact constitutes a definite practical limitation of such a series representation in treating any violent nonlinearity.

Let R_q be the Euclidean q space, and let $L_{2,r}$ be the space of r-vector valued functions, each component of a member vector being square integrable on the interval $[t_0, T]$. Let an $r \times r$ matrix $Q(t)$ be defined for $t \in [t_0, T]$ such that it is symmetric and positive definite for each t, and that $q_k \in L_{2,r}$, for $k=1, 2, \cdots, r$, where q_k is the kth column of $Q(t)$. Let the transpose of a vector or a matrix be denoted by a prime.

The type of optimal control problem investigated in this paper is: given a nonlinear system described by (1) and an arbitrary initial output $y(t_0) = y_0 \in R_q$, find the input $u(t)$, $t_0 \le t \le T$ and $u \in L_{2,r}$, such that 1) the terminal output $y(T) = y_T \in R_q$, and 2) the cost functional given by

$$J = \int_{t_0}^{T} u'(t) Q(t) u(t)\, dt \tag{3}$$

is minimized.[2] The terminal time T is assumed fixed. The terminal output y_T may be fixed or given in terms of T.

Throughout the paper, it is assumed that the q vector, $h_{ij}(t; \tau_1, \cdots, \tau_i)$, belongs to $L_{2,q}$ in each of the arguments $\tau_1, \tau_2, \cdots$ and τ_i, for each value of $t \in [t_0, T]$, and for all i and j.

III. Characteristic Set of Given Functions—Definitions

Let L_2 denote the Hilbert space of square integrable functions on $[t_0, T]$. Suppose that a finite set of functions, $\{F_i(t) | F_i \in L_2;\ i=1, 2, \cdots, p\}$, is given. If there is a finite set of linearly independent functions $\{f_m(t) | f_m \in L_2;\ m=1, 2, \cdots, s\}$ such that each member of the set $\{F_i(t)\}$ is expressible as a finite linear combination of the functions of this set, i.e.,

$$F_i(t) = \sum_{m=1}^{(m)i} \alpha_m{}^i f_m(t) \quad i = 1, 2, \cdots, p \tag{4}$$

then the finite set $\{f_m(t)\}$ is called a characteristic set of the set of functions $\{F_i(t)\}$.

Assertion 1

Given an arbitrary finite set of functions $\{F_i(t) | F_i \in L_2;\ i=1, 2, \cdots, p\}$, there exists at least one characteristic set.

The proof is immediately seen by letting

$$\{f_m(t)\} = \{G_i(t)\} \tag{5}$$

where $\{G_i(t)\}$ is the set consisting of all linearly independent elements of $\{F_i(t)\}$.

Suppose that $\{f_m(t)\}$ is a characteristic set of a given set of functions $\{F_i(t) | F_i \in L_2;\ i=1, 2, \cdots, p\}$. Let the set of orthonormal functions constructed from $\{f_m(t)\}$ be denoted by $\{\tilde{\phi}_m(t)\}$. It is known[14] that the set $\{\tilde{\phi}_m(t)\}$ can always be extended to a complete orthonormal set in L_2. Let such a complete orthonormal set be denoted by $\{\phi_m(t)\}$. Then an arbitrary function in L_2, $u(t)$, can be uniquely expressed by

[1] Throughout this paper, the number of terms of a summation is denoted by placing parentheses around the running index. Similarly, if the upper limit of a summation depends on several indexes, it is denoted by placing parentheses around the running index followed by all the other indexes. These conventions are not used when the upper limit is explicitly defined.

[2] Note that the vector "output" $y(t)$, as used in this paper, can also be interpreted as representing the state vector in the usual sense.

$$u(t) = \sum_{n=1}^{\infty} b_n \phi_n(t). \tag{6}$$

The concepts defined previously provide a useful tool for the solution of the problem posed in Section II. The details for linear systems are given in Section IV and for nonlinear systems in Section V. Linear systems are included both because they provide an algebraically simple vehicle to illustrate the method developed in this paper and because some of the results are useful for the analysis of the nonlinear example presented in Section V.

It will be assumed throughout, without loss of generality, that the elements of the complete set $\{\phi_m(t)\}$ are arranged in such a way that the first s elements coincide with the elements of the set $\{\tilde{\phi}_m(t)\}$ in the same order, where s is the number of elements in $\{\tilde{\phi}_m(t)\}$.

IV. Linear Systems

A general linear system is described by the first two terms in the right-hand side of (1). Thus one has for the system (the linearity is between $\boldsymbol{u}$ and $\boldsymbol{y}-\boldsymbol{h}_0$)

$$\boldsymbol{y}(t) = \boldsymbol{h}_0(t) + \int_{t_0}^{t} \boldsymbol{H}(t;\tau)\boldsymbol{u}(\tau)\, d\tau \tag{7}$$

where $\boldsymbol{H}(t;\tau)$ is the $q\times r$ matrix function whose (kj)th entry is $h_j^k(t;\tau)$. The initial and terminal conditions are

$$\boldsymbol{y}(t_0) = \boldsymbol{y}_0 = \boldsymbol{h}_0(t_0) \in R_q \tag{8}$$

and

$$\boldsymbol{y}(T) = \boldsymbol{y}_T = \boldsymbol{h}_0(T) + \int_{t_0}^{T} \boldsymbol{H}(T;\tau)\boldsymbol{u}(\tau)\, d\tau \in R_q \tag{9}$$

respectively. Let $\boldsymbol{z}_T = \boldsymbol{y}_T - \boldsymbol{h}_0(T)$.

In what follows, the matrix $\boldsymbol{Q}(t)$ in (3) will be taken to be $\boldsymbol{I}$, the identity matrix. This does not constitute any loss of generality. For a positive definite symmetric matrix $\boldsymbol{Q}(t)$, there exists a nonsingular matrix $\boldsymbol{P}(t)$[15] such that

$$\boldsymbol{Q}(t) = \boldsymbol{P}'(t)\boldsymbol{P}(t). \tag{10}$$

Thus if

$$\boldsymbol{v}(t) = \boldsymbol{P}(t)\boldsymbol{u}(t) \tag{11}$$

and

$$\boldsymbol{K}(t;\tau) = \boldsymbol{H}(t;\tau)\boldsymbol{P}^{-1}(\tau) \tag{12}$$

then the linear system given by

$$\boldsymbol{y}(t) - \boldsymbol{h}_0(t) = \int_{t_0}^{t} \boldsymbol{H}(t;\tau)\boldsymbol{u}(\tau)\, d\tau \tag{13}$$

and the cost functional given by (3) may be transformed to, respectively,

$$\boldsymbol{y}(t) - \boldsymbol{h}_0(t) = \int_{t_0}^{t} \boldsymbol{K}(t;\tau)\boldsymbol{v}(\tau)\, d\tau \tag{14}$$

and

$$J = \int_{t_0}^{T} \boldsymbol{v}'(t)\boldsymbol{v}(t)\, dt. \tag{15}$$

The optimal input for the transformed problem $\boldsymbol{v}^*(t)$ is found by the characteristic expansion method to be developed later and the actual optimal input $\boldsymbol{u}^*(t)$ is uniquely determined by

$$\boldsymbol{u}^*(t) = \boldsymbol{P}^{-1}(t)\boldsymbol{v}^*(t). \tag{16}$$

The optimal control problem (P) may then be stated as: given an element $\boldsymbol{z}_T \in R_q$ and a bounded linear operator T from $L_{2,r}$ into R_q defined by

$$T\boldsymbol{u} = \int_{t_0}^{T} \boldsymbol{H}(T;t)\boldsymbol{u}(t)\, dt \tag{17}$$

find an element $\boldsymbol{u} \in L_{2,r}$ that minimizes

$$\|\boldsymbol{u}\|_{L_{2,r}} = \int_{t_0}^{T} \boldsymbol{u}'(t)\boldsymbol{u}(t) dt \tag{18}$$

subject to

$$T\boldsymbol{u} = \boldsymbol{z}_T. \tag{19}$$

The solution of this problem can be obtained using the notion of the characteristic set introduced in Section III.

Consider the set of functions consisting of all the entries in the jth column of $\boldsymbol{H}(T;t)$. By assumption, each element of this set is in L_2. From Assertion 1, there exists a characteristic set. Let the complete orthonormal set in L_2 constructed from the characteristic set be denoted by $\{\phi_{jm}(t)\}$. Repeating for each j, $j=1, 2, \cdots, r$, the sets $\{\phi_{1m}(t)\}$, $\{\phi_{2m}(t)\}$, $\cdots$ and $\{\phi_{rm}(t)\}$ are obtained.

Lemma 1

Given (17) and (19), there exist orthonormal sets of functions $\{\phi_{jm}(t)\}$, $j=1, 2, \cdots, r$, each complete in L_2; a set of integers $(m1)$, $(m2)$, $\cdots$ and (mr); and constants $A_{jm}^k < \infty$, for $k=1, 2, \cdots, q$, $j=1, 2, \cdots, r$, and $m=1, 2, \cdots, (mj)$; such that, for $k=1, 2, \cdots, q$,

$$z_T^k = \sum_{j=1}^{r} \sum_{m=1}^{(mj)} A_{jm}^k \int_{t_0}^{T} \phi_{jm}(t) u_j(t)\, dt. \tag{20}$$

Proof: The existence of the complete sets $\{\phi_{jm}(t)\}$ has already been shown.

Now, the kth component of (19) is, in the expanded form,

$$z_T^k = \sum_{j=1}^{r} \int_{t_0}^{T} h_j^k(T;t) u_j(t)\, dt. \tag{21}$$

But from the definition of the characteristic set, there is a finite integer (mkj) such that

$$h_j^k(T;t) = \sum_{m=1}^{(mkj)} A_{jm}^k \phi_{jm}(t) \tag{22}$$

where the constants A_{jm}^k are uniquely determined. Thus using (22) in (21),

$$z_T^k = \sum_{j=1}^{r} \sum_{m=1}^{(mkj)} A_{jm}^k \int_{t_0}^{T} \phi_{jm}(t) u_j(t)\, dt. \tag{23}$$

Equation (23), with the definitions

$$(mj) = \max_k [(mkj)] \tag{24}$$

and

$$A_{jm}^k = 0, \quad \text{for } m > (mkj) \tag{25}$$

completes the proof.

Define the matrices $\boldsymbol{A}_j = [A_{jm}^k]$, for $j=1, 2, \cdots, r$, and define a $q \times n$ matrix $\boldsymbol{A}$ by

$$\boldsymbol{A} = [\boldsymbol{A}_1 | \boldsymbol{A}_2 | \cdots | \boldsymbol{A}_r] \tag{26}$$

where

$$n = \sum_{j=1}^{r} (mj). \tag{27}$$

Since $\{\phi_{jm}(t)\}$ is complete in L_2, an arbitrary element $u_j \in L_2$ can be uniquely expressed by

$$u_j(t) = \sum_{m=1}^{\infty} B_{jm} \phi_{jm}(t). \tag{28}$$

Repeating, for $j=1, 2, \cdots, r$, for an arbitrary element $\boldsymbol{u} \in L_{2,r}$,

$$\boldsymbol{u}'(t) = \left[\sum_{m=1}^{\infty} B_{1m}\phi_{1m}(t),\ \sum_{m=1}^{\infty} B_{2m}\phi_{2m}(t), \cdots, \sum_{m=1}^{\infty} B_{rm}\phi_{rm}(t) \right] \tag{29}$$

i.e., a $\boldsymbol{u} \in L_{2,r}$ uniquely defines a set of constants B_{jm}, and hence an n vector,

$$\boldsymbol{b}' = [B_{11}, B_{12}, \cdots, B_{1(m1)} \cdots B_{r1}, B_{r2}, \cdots, B_{r(mr)}]. \tag{30}$$

Conversely, an n vector $\boldsymbol{b}$ and constants B_{jm}, for $j=1, 2, \cdots, r$, and $m=(mj)+1, (mj)+2, \cdots$, such that

$$\sum_{i=1}^{n} b_i^2 + \sum_{j=1}^{r} \sum_{m=(mj)+1}^{\infty} B_{jm}^2 < \infty$$

uniquely define an element in $L_{2,r}$.

Using (30), it is seen that (20) is reduced to

$$z_T = \boldsymbol{A}\boldsymbol{b}. \tag{31}$$

Any element in $L_{2,r}$ that satisfies (31) is a feasible input. Concerning the existence of a feasible input, the following lemma holds.

Lemma 2

Given the problem (P), there exists a feasible input in $L_{2,r}$ if rank $\boldsymbol{A}=q$.

Proof: Let rank $\boldsymbol{A}=q$, and let $\tilde{\boldsymbol{A}}$ be a nonsingular $q\times q$ submatrix of $\boldsymbol{A}$. Without loss of generality, $\tilde{\boldsymbol{A}}$ may be assumed to contain the first q columns of $\boldsymbol{A}$. There exists a q vector $\tilde{\boldsymbol{b}} = \tilde{\boldsymbol{A}}^{-1} z_T$. A $\boldsymbol{u} \in L_{2,r}$ can be found such that $b_i = \tilde{b}_i$, for $i=1, 2, \cdots, q$; $b_i = B_{jm} = 0$, for $q+1 \le i \le n$, $j=1, 2, \cdots, r$, and $m > (mj)$. Such $\boldsymbol{u}$ is clearly a feasible input, which proves the lemma.

From (29) and (30), it is seen that the cost functional $J = \|\boldsymbol{u}\|_{L_{2,r}}$ may be written as

$$J = \boldsymbol{b}'\boldsymbol{b} + \sum_{j=1}^{r} \sum_{m=(mj)+1}^{\infty} B_{jm}^2. \tag{32}$$

The following theorems contain the solution of the problem.

Theorem 1

Given the problem (P), if rank $\boldsymbol{A}=q$, then there exists an optimal input $\boldsymbol{u}^* \in L_{2,r}$, and it is unique.

Proof: Define a set X in R_n by

$$X = \{\boldsymbol{b} \mid z_T = \boldsymbol{A}\boldsymbol{b}\}. \tag{33}$$

Since rank $\boldsymbol{A}=q$, X is nonempty from Lemma 2. The mapping $\boldsymbol{A}$ is clearly continuous, and the set consisting only of a fixed element z_T in R_q is closed. Therefore, the inverse image $X \subset R_n$ is also closed.[14] Obviously, X is also convex. But a closed convex set in a Hilbert space contains a unique element of minimum norm.[14] Let $\boldsymbol{b}^* \in X$ be that element. The element $\boldsymbol{u}^* \in L_{2,r}$ defined by $\boldsymbol{b}^*$ and $B_{jm}=0$, for $j=1, 2, \cdots, r$ and $m=(mj)+1, (mj)+2, \cdots$, is obviously the unique element of minimum norm in $L_{2,r}$. Q.E.D.

Theorem 2

Given the problem (P), let $\tilde{\boldsymbol{A}}$ be a $q\times q$ nonsingular submatrix of $\boldsymbol{A}$ containing the first q columns of $\boldsymbol{A}$. Let $z_T = \tilde{\boldsymbol{A}}\tilde{\boldsymbol{b}} + \hat{\boldsymbol{A}}\hat{\boldsymbol{b}}$. Then the optimal input is given by

$$\boldsymbol{b}^* = \begin{bmatrix} \boldsymbol{c} - \boldsymbol{D}(\boldsymbol{D}'\boldsymbol{D} + \boldsymbol{I})^{-1}\boldsymbol{D}'\boldsymbol{c} \\ (\boldsymbol{D}'\boldsymbol{D} + \boldsymbol{I})^{-1}\boldsymbol{D}'\boldsymbol{c} \end{bmatrix} \tag{34}$$

$$B_{jm}^* = 0, \quad \text{for } j = 1, 2, \cdots, r, \quad m = (mj)+1, (mj)+2, \cdots$$

where $\boldsymbol{c} = \tilde{\boldsymbol{A}}^{-1} z_T$, $\boldsymbol{D} = \tilde{\boldsymbol{A}}^{-1}\hat{\boldsymbol{A}}$.

Proof: Equation (32) becomes, in terms of $\tilde{\boldsymbol{b}}$ and $\hat{\boldsymbol{b}}$,

$$J = \tilde{\boldsymbol{b}}'\tilde{\boldsymbol{b}} + \hat{\boldsymbol{b}}'\hat{\boldsymbol{b}} + \sum_j \sum_m B_{jm}^2.$$

Clearly, for minimum J, all $B_{jm}=0$. Let $\hat{J} = \tilde{\boldsymbol{b}}'\tilde{\boldsymbol{b}} + \hat{\boldsymbol{b}}'\hat{\boldsymbol{b}}$. Since $\tilde{\boldsymbol{b}} = \boldsymbol{c} - \boldsymbol{D}\hat{\boldsymbol{b}}$,

$$\hat{J} = \boldsymbol{c}'\boldsymbol{c} - 2\boldsymbol{c}'\boldsymbol{D}\hat{\boldsymbol{b}} + \hat{\boldsymbol{b}}'(\boldsymbol{D}'\boldsymbol{D} + \boldsymbol{I})\hat{\boldsymbol{b}}. \tag{35}$$

The matrix $(\boldsymbol{D}'\boldsymbol{D}+\boldsymbol{I})$ is symmetric and positive definite, and $\hat{J}$ is strictly convex and continuously differentiable in $\hat{\boldsymbol{b}}$. Thus setting the derivative to be stationary, $\hat{\boldsymbol{b}}^* = (\boldsymbol{D}'\boldsymbol{D}+\boldsymbol{I})^{-1}\boldsymbol{D}'\boldsymbol{c}$. From $\tilde{\boldsymbol{b}} = \boldsymbol{c} - \boldsymbol{D}\hat{\boldsymbol{b}}$, $\tilde{\boldsymbol{b}}^* = \boldsymbol{c} - \boldsymbol{D}\hat{\boldsymbol{b}}^*$. Q.E.D

The procedure developed previously is defined as the characteristic expansion method. It exploits the observation that, by a careful choice of bases, an infinite dimensional problem can be reduced to a finite dimensional problem. In particular, the characteristic expansion method is not to be confused with the Ritz method in the calculus of variations. Unlike the Ritz method, at no point is the construction of a minimizing sequence necessary or attempted; the reduction to a finite dimension is a natural and rigorous consequence of the choice of bases and is in no sense an approximation.

Example 1

This example is chosen from communication theory and pertains to the design of a matched filter. The description of the problem is given in Holtzman.[16] A typical analytical statement of the problem is to find $h(t)$ such that

$$y_1(T) = \int_0^T h(\tau)s_1(T-\tau)\,d\tau = 1 \tag{36}$$

$$y_2(T) = \int_0^T h(\tau)s_2(T-\tau)\,d\tau = \alpha \tag{37}$$

$$\sigma^2(T) = N\int_0^T h^2(\tau)\,d\tau = \text{minimum} \tag{38}$$

where

$$s_1(t) = 1, \qquad t \in [0, T] \tag{39}$$

$$s_2(t) = 1, \qquad t \in [0, T/2]; \qquad = 0, \qquad t \in (T/2, T] \tag{40}$$

and α and N are real constants.

Equations (36) through (38) are similar to (9) and (3). Therefore, $h(t)$ can be found by means of the characteristic expansion method. Note that the basis elements of $s_1(t)$ and $s_2(t)$ are the Haar functions[17]

$$H_1(t) = 1/\sqrt{T}, \quad H_2(t) = 1/\sqrt{T}, \qquad t \in [0, T/2]$$
$$H_2(t) = -1/\sqrt{T}, \quad t \in (T/2, T], \cdots .$$

In terms of these functions

$$s_1(T-t) = \sqrt{T}\,H_1(T-t) = \sqrt{T}\,H_1(t) \tag{41}$$

$$s_2(T-t) = \frac{\sqrt{T}}{2}\left[H_1(t) - H_2(t)\right] \tag{42}$$

and

$$h(t) = \sum_{m=1}^{\infty} B_m H_m(t). \tag{43}$$

Substitution of (41) through (43) in (36) through (38) yields

$$B_1{}^* = 1/\sqrt{T}, \quad B_2{}^* = (1-2\alpha)/\sqrt{T} \tag{44}$$

$$h^*(t) = [H_1(t) + (1-2\alpha)H_2(t)]/\sqrt{T}. \tag{45}$$

The result given by (45) is identical, as it should be to that derived by other methods.[16]

V. Nonlinear Systems

The purpose of this section is to extend the characteristic expansion method to nonlinear systems.

The kth component of the output vector of a nonlinear system described by (1) is given by the relation

$$y^k(t) = \sum_{i=0}^{(i)}\sum_{j=1}^{(j)}\int_{t_0}^{t}\cdots\int_{t_0}^{t} h_{ij}{}^k(t;\tau_1,\cdots,\tau_i) \cdot C_j(u;\tau_1,\cdots,\tau_i)\,d\tau_1\cdots d\tau_i. \tag{46}$$

To solve the optimal control problem stated in Section II, it is assumed that the kernels are degenerate, i.e.,

$$h_{ij}{}^k(T;\tau_1,\tau_2,\cdots,\tau_i) = \sum_{d=1}^{(d)}\prod_{s=1}^{i} g_w{}^p(T;\tau_s) \tag{47}$$

where the index w represents indexes k, i, j, d, and s, and the index p denotes the component of the input vector on which $g_w{}^p(T;\tau_s)$ operates. The indexes k, i, and j run from 1 to q, from 1 to (i), and from 1 to (j), respectively.

For a given p, the set of functions $g_w{}^p(T;t)$, for all w, is finite. Hence from Assertion 1, a characteristic set exists. Let the completed orthonormal set constructed from the characteristic set be $\{\phi_m{}^p(t)\}$. Then it follows from Section III that

$$g_w{}^p(T;t) = \sum_{m=1}^{(mwp)} A_{mw}{}^p\phi_m{}^p(t) \tag{48}$$

$$u_p(t) = \sum_{m=1}^{\infty} B_m{}^p\phi_m{}^p(t) \tag{49}$$

for $p = 1, 2, \cdots, r$. Then (46) becomes

$$y_T{}^k - h_0{}^k(T) = \sum_{i=1}^{(i)}\sum_{j=1}^{(j)}\sum_{d=1}^{(d)}\prod_{s=1}^{i}\left[\sum_{m=1}^{(mwp)} A_{mw}{}^pB_m{}^p\right] \tag{50}$$

for $k = 1, 2, \cdots, q$. Similarly, (3) becomes, where again, $\boldsymbol{Q} = \boldsymbol{I}$,

$$J = \sum_{p=1}^{r}\sum_{m=1}^{\infty}(B_m{}^p)^2. \tag{51}$$

Thus the optimal control problem under consideration is reduced to the solution of a nonlinear set of algebraic equations (50), subject to the minimization of J. From this point, the solution proceeds as in the case of linear systems. Note, however, that the questions of existence and uniqueness of the solutions have not been resolved for arbitrary nonlinear systems.

Admittedly, the solution of the system of (50), for an arbitrary nonlinear system, is very involved. For many practical systems, however, in which the dimensions of the input and the output, and the number of terms in the input–output functional relation are small, solutions can be found without much difficulty. For clarity, the following simple example is included to illustrate the essentials of the characteristic expansion method for nonlinear systems.

Example 2

Consider the nonlinear system described by the relations

$$y_1(t) = \int_0^t (t-\tau)u(\tau)\,d\tau + \int_0^t\int_0^t (t-\tau_1)(t-\tau_2)u(\tau_1)u(\tau_2)\,d\tau_1 d\tau_2 \quad (52a)$$

$$y_2(t) = -2\int_0^t u(\tau)\,d\tau + \int_0^t\int_0^t u(\tau_1)u(\tau_2)\,d\tau_1 d\tau_2. \quad (52b)$$

Let it be desired to find the input that brings the output to $(1, -1)$ at $t=1$ and that minimizes the cost given by the relation

$$J = \int_0^1 u^2(\tau)\,d\tau. \quad (53)$$

Note that the correct choice of the basis gives $\phi_1(t) = 1$, $\phi_2(t) = \sqrt{3} - 2\sqrt{3}t, \cdots$. Thus (52) is reduced to

$$1 = \frac{1}{2}\left(B_1 + \frac{B_2}{\sqrt{3}}\right) + \frac{1}{4}\left(B_1 + \frac{B_2}{\sqrt{3}}\right)^2 \quad (54a)$$

$$-1 = -2B_1 + B_1^2 \quad (54b)$$

which have the solutions

$$B_1 = 1, \quad B_2 = -1 \pm (1 + \sqrt{3/4})^{1/2}.$$

The minimization of J results in $B_m = 0$ for $m > 2$ and the choice of the plus sign in B_2. Thus the optimal input is

$$u^*(t) = 1 - \sqrt{3}\,[1 - (1 + \sqrt{3/4})^{1/2}] + \sqrt{12}\,[1 - (1 + \sqrt{3/4})^{1/2}]\,t. \quad (55)$$

VI. Conclusions

A direct method is presented for the solution of a class of optimal control problems of systems described by input–output functional relations. The method is based on the observation that, by a careful choice of bases in the space of inputs, an infinite dimensional problem can be transformed into a finite dimensional problem. It is shown that the method is applicable to linear and nonlinear, and stationary and time-varying systems, provided that the kernels of the system degenerate.

The method presented can be directly extended to study similar problems in distributed-parameter systems, such as might arise from a Green's function solution of a partial differential equation.

Acknowledgment

The authors thank the reviewers for several useful comments.

References

[1] L. S. Pontryagin V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes.* New York: Interscience, 1962.

[2] M. Athans and P. L. Falb, *Optimal Control, An Introduction to the Theory and Its Applications.* New York: McGraw-Hill, 1966.

[3] M. Schetzen, M.I.T. Research Lab. of Electronics, Cambridge, Mass., Quart. Progr. Rept. 60, January 15, 1961.

[4] J. Katzenelson and L. A. Gould, "The design of nonlinear filters and control systems," pt. 1, *Information and Control,* vol. 5, pp. 108–143, 1962; pt. 2, *ibid.*, vol. 7, pp. 117–145, 1964.

[5] R. Hooper, "Test signals for a class of nonlinear systems," Sc.D. dissertation, Dept. of Nuclear Engineering, Massachusetts Institute of Technology, Cambridge, Mass., 1966.

[6] V. Volterra, *Theory of Functionals and of Integral and Integro-Differential Equations.* New York: Dover, 1959.

[7] N. Wiener, *Nonlinear Problems in Random Theory.* Cambridge, Mass.: Massachusetts Institute of Technology Press, 1958.

[8] A. G. Bose, "A theory of nonlinear systems," M.I.T. Research Lab. of Electronics, Cambridge, Mass., Tech. Rept. 309, May 15, 1956.

[9] M. B. Brilliant, "Theory of the analysis of nonlinear systems," M.I.T. Research Lab. of Electronics, Cambridge, Mass., Tech. Rept. 345, March 3, 1958.

[10] D. A. George, "Continuous nonlinear systems," M.I.T. Research Lab. of Electronics, Cambridge, Mass., Tech. Rept. 355, July 24, 1959.

[11] G. Zames, "Nonlinear operators for system analysis," M.I.T. Research Lab. of Electronics, Cambridge, Mass., Tech. Rept. 370, August 25, 1960.

[12] J. F. Barrett, "The use of functionals in the analysis of nonlinear physical systems," Ministry of Supply, Great Britain, Statistical Advisory Unit Rept., January 1957.

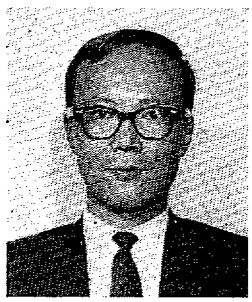
[13] R. H. Flake, "Volterra series representation of time varying nonlinear systems," *Proc. Internat'l Fed. Automatic Control Cong.* (Basle, Switzerland, 1963).

[14] A. N. Kolmogorov and S. V. Fomin, *Elements of Functional Analysis,* vols. 1 and 2. Rochester, N. Y.: Graylock, 1957.

[15] R. Bellman, *Introduction to Matrix Algebra.* New York: McGraw-Hill, 1960.

[16] J. M. Holtzman, "Signal-noise ratio maximization using the Pontryagin maximum principle," *Bell Sys. Tech. J.*, vol. 45, no. 3, pp. 473–489, 1966.

[17] A. Haar, "Zur Theorie der orthogonales Funktionsysteme," *Math. Ann.*, vol. 69, pp. 331–371, 1910.



Sang H. Kyong (S'64–M'66) was born in Seoul, Korea, on July 24, 1937. He attended Seoul University, Seoul, and he received the B.S. degree from the University of Rhode Island, Kingston, and the Ph.D. degree from the Massachusetts Institute of Technology, Cambridge, in 1961 and 1965, respectively.

From 1965 to 1966, he worked as a Reactor Physicist at the Argonne National Lab., Argonne, Ill. In 1966 he joined Bell Telephone Laboratories, Inc., Whippany, N. J., and is currently associated with the Guidance and Control Systems Group.

Dr. Kyong is a member of the Society for Industrial and Applied Mathematics, the American Nuclear Society, Sigma Xi, Tau Beta Pi, and Phi Kappa Phi.

Elias P. Gyftopoulos was born in Athens, Greece, on July 4, 1927. He received the Diploma in mechanical and electrical engineering from the Technical University of Athens, and the Sc.D. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1953 and 1958, respectively.

Since 1953, he has been with the Massachusetts Institute of Technology, where he is presently Professor of Nuclear Engineering. His areas of interest are nuclear reactor dynamics, plasma physics, and surface physics.

Dr. Gyftopoulos is a fellow of the American Academy of Arts and Sciences, a fellow of the American Nuclear Society, and a member of the Board of Directors of the American Nuclear Society.