Journal of Engineering Mathematics, Vol. 6, No. 3, July 1972
Wolters-Noordhoff Publishing–Groningen
Printed in the Netherlands

Duality in Distributed Parameter Systems

M. C. Y. KUO

Department of Electrical Engineering, Newark College of Engineering, Newark, New Jersey, U.S.A.

(Received September 8, 1971)

SUMMARY

The principle of duality for a wide class of distributed parameter systems is developed in this paper. The necessary and sufficient conditions for the primal control problem are utilized to derive the dual and converse dual theorems of the dual problem. An example of the temperature variation in a slab is given to show the application of the theory.

1. Introduction

A number of problems in the engineering field may be alternatively solved by using a duality principle. It is well known that the solution of either a primal problem or a dual problem can be solved from the solution of the other. It is sometimes computationally convenient to solve the dual problem. Principles of duality in lumped systems have been studied extensively [1, 2, 3]; however, the dual theory in distributed parameter systems has not been explored in the literature. It is the purpose of this paper to derive the dual and converse dual theorems for the optimum control of a wide class of distributed parameter systems. To complete the derivations of theorems necessary and sufficient conditions for the primal control problem are also included. The background material of relevance for the derivation can be found in [2, 4, 5].

2. Notations and Assumptions

$F(x(T, y), T)$, $f_0(t, y, x(t, y), x_{y^k}(t, y), u(t, y))$, and $H(t, y, x(t, y), x_{y^k}(t, y), u(t, y), \lambda(t, y))$ are scalar functions with continuous derivatives up to and including the second order with respect to each of its arguments. $f(t, y, x(t, y), x_{y^k}(t, y), u(t, y))$ and $h(t, y, x(t, y), u(t, y))$ are, respectively, n- and l-dimensional vector functions with continuous derivatives up to and including the second order. $x, u, \lambda, \rho, \xi, \eta$, and τ are, respectively, n-, r-, n-, l-, n-, r-, and l-dimensional functions of t and y, where y is an m-dimensional spatial coordinate vector belonging to a region Ω and $\partial\Omega$ denotes the boundary of Ω. u is required to have piecewise continuous first and second derivatives for $t \in [t_0, T]$ and $y \in \bar{\Omega}$ (the closure of Ω). x and λ are continuous in $t \in [t_0, T]$ and $y \in \bar{\Omega}$. $x_t, \lambda_t, \rho, \xi, \eta$, and τ are continuous functions of t and y except possibly for values of t and y corresponding to points of discontinuity of u. The superscript T for $(\cdot)^T$ denotes the transpose of $(\cdot)$ and subscripts denote partial derivatives. H_x and H_u are the gradient vectors of H with respect to x and u, respectively. Similar notations are applied to scalar functions of F and f_0. h_x and h_u (or f_x and f_u) are the Jacobian matrices of h (or f) with respect to x and u, respectively. The notation $x_{y^k}(t, y) = \partial^k x(t, y)/\partial y^k$ indicates all possible derivatives

$$\frac{\partial x}{\partial y_1}, \frac{\partial x}{\partial y_2}, \ldots, \frac{\partial x}{\partial y_m}, \ldots, \frac{\partial^{k_1+k_2+\ldots+k_m} x(t, y)}{\partial y_1^{k_1} \partial y_2^{k_2} \ldots \partial y_m^{k_m}},$$

where $k = k_1 + k_2 + \ldots + k_m$.

The following functions x_t, $H_{x_y^k y^k x}$, and $(h_u^T \rho)_x$ mean

$$\frac{\partial x}{\partial t}, \frac{\partial}{\partial x}\left[\frac{\partial^k}{\partial y^k}\left(\frac{\partial H}{\partial\left(\frac{\partial^k x}{\partial y^k}\right)}\right)\right] \quad \text{and} \quad \frac{\partial}{\partial x}(h_u^T \rho), \text{ respectively.}$$

The vector $\rho > 0$ means that all its components are positive. $\rho \geqq 0$ means that all the components of ρ are non-negative and at least one component is positive. Several examples of notations for $x=(x_1, x_2)$, $u=(u_1, u_2)$, $y=(y_1, y_2)$, $h=(h_1, h_2, h_3)$, $\rho=(\rho_1, \rho_2, \rho_3)$, $\zeta=(\xi_1, \xi_2)$, and $\eta=(\eta_1, \eta_2)$ are illustrated as follows:

(1) x_{y^2} gives six terms $\dfrac{\partial^2 x_1}{\partial y_1^2}, \dfrac{\partial^2 x_1}{\partial y_2^2}, \dfrac{\partial^2 x_1}{\partial y_1 \partial y_2}, \dfrac{\partial^2 x_2}{\partial y_1^2}, \dfrac{\partial^2 x_2}{\partial y_2^2}$, and $\dfrac{\partial^2 x_2}{\partial y_1 \partial y_2}$.

(2)
$$H_{x_{y^2}} = \frac{\partial H}{\partial\left(\dfrac{\partial^2 x}{\partial y^2}\right)} = \text{col.}\left(\frac{\partial H}{\partial\left(\dfrac{\partial^2 x_1}{\partial y_1^2}\right)}, \frac{\partial H}{\partial\left(\dfrac{\partial^2 x_1}{\partial y_2^2}\right)}, \frac{\partial H}{\partial\left(\dfrac{\partial^2 x_1}{\partial y_1 \partial y_2}\right)}, \frac{\partial H}{\partial\left(\dfrac{\partial^2 x_2}{\partial y_1^2}\right)}, \frac{\partial H}{\partial\left(\dfrac{\partial^2 x_2}{\partial y_2^2}\right)}, \frac{\partial H}{\partial\left(\dfrac{\partial^2 x_2}{\partial y_1 \partial y_2}\right)}\right).$$

(3)
$$H_{x_{y^2} y^2} = \frac{\partial^2}{\partial y^2}\left[\frac{\partial H}{\partial\left(\dfrac{\partial^2 x}{\partial y^2}\right)}\right] =$$
$$= \text{col}\left(\frac{\partial^2}{\partial y_1^2}\left[\frac{\partial H}{\partial\left(\dfrac{\partial^2 x_1}{\partial y_1^2}\right)}\right] + \frac{\partial^2}{\partial y_2^2}\left[\frac{\partial H}{\partial\left(\dfrac{\partial^2 x_1}{\partial y_2^2}\right)}\right] + \frac{\partial^2}{\partial y_1 \partial y_2}\left[\frac{\partial H}{\partial\left(\dfrac{\partial^2 x_1}{\partial y_1 \partial y_2}\right)}\right],\right.$$
$$\left.\frac{\partial^2}{\partial y_1^2}\left[\frac{\partial H}{\partial\left(\dfrac{\partial^2 x_2}{\partial y_1^2}\right)}\right] + \frac{\partial^2}{\partial y_2^2}\left[\frac{\partial H}{\partial\left(\dfrac{\partial^2 x_2}{\partial y_2^2}\right)}\right] + \frac{\partial^2}{\partial y_1 \partial y_2}\left[\frac{\partial H}{\partial\left(\dfrac{\partial^2 x_2}{\partial y_1 \partial y_2}\right)}\right]\right).$$

(4)
$$(H_{x_{y^2} y^2 x})^T \xi = \begin{bmatrix} \dfrac{\partial \mathscr{H}_1}{\partial x_1} & \dfrac{\partial \mathscr{H}_2}{\partial x_1} \\ \dfrac{\partial \mathscr{H}_1}{\partial x_2} & \dfrac{\partial \mathscr{H}_2}{\partial x_2} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$$
$$= \text{col.}\left(\xi_1 \frac{\partial \mathscr{H}_1}{\partial x_1} + \xi_2 \frac{\partial \mathscr{H}_2}{\partial x_1}, \ \xi_1 \frac{\partial \mathscr{H}_1}{\partial x_2} + \xi_2 \frac{\partial \mathscr{H}_2}{\partial x_2}\right),$$

where $\mathscr{H}_1$ and $\mathscr{H}_2$ are the components of the column vector in example 3.

(5)
$$\left(\frac{\partial}{\partial y}\right)^T [(\delta x) H_{x_{y^2} y}] = \left(\frac{\partial}{\partial y}\right)^T \left[(\delta x) \frac{\partial}{\partial y} \frac{\partial H}{\partial\left(\dfrac{\partial^2 x}{\partial y^2}\right)}\right] = \frac{\partial}{\partial y_1}\left[(\delta x_1) \frac{\partial}{\partial y_1} \frac{\partial H}{\partial\left(\dfrac{\partial^2 x_1}{\partial y_1^2}\right)}\right]$$
$$+ \frac{\partial}{\partial y_1}\left[(\delta x_2) \frac{\partial}{\partial y_1} \frac{\partial H}{\partial\left(\dfrac{\partial^2 x_2}{\partial y_1^2}\right)}\right] + \frac{\partial}{\partial y_1}\left[(\delta x_2) \frac{\partial}{\partial y_2} \frac{\partial H}{\partial\left(\dfrac{\partial^2 x_2}{\partial y_1 \partial y_2}\right)}\right]$$
$$+ \frac{\partial}{\partial y_2}\left[(\delta x_1) \frac{\partial}{\partial y_2} \frac{\partial H}{\partial\left(\dfrac{\partial^2 x_1}{\partial y_2^2}\right)}\right] + \frac{\partial}{\partial y_2}\left[(\delta x_2) \frac{\partial}{\partial y_2} \frac{\partial H}{\partial\left(\dfrac{\partial^2 x_2}{\partial y_2^2}\right)}\right]$$
$$+ \frac{\partial}{\partial y_2}\left[(\delta x_1) \frac{\partial}{\partial y_1} \frac{\partial H}{\partial\left(\dfrac{\partial^2 x_1}{\partial y_1 \partial y_2}\right)}\right].$$

(6) It follows from example 5 that $H_{x_{y^2}y}=0$ at the boundary of (y_1, y_2) and $t=T$ means that the following six terms

$$\frac{\partial}{\partial y_1}\frac{\partial H}{\partial\left(\frac{\partial^2 x_1}{\partial y_1^2}\right)},\ \frac{\partial}{\partial y_1}\frac{\partial H}{\partial\left(\frac{\partial^2 x_2}{\partial y_1^2}\right)},\ \frac{\partial}{\partial y_2}\frac{\partial H}{\partial\left(\frac{\partial^2 x_2}{\partial y_1 \partial y_2}\right)},\ \frac{\partial}{\partial y_2}\frac{\partial H}{\partial\left(\frac{\partial^2 x_1}{\partial y_2^2}\right)},\ \frac{\partial}{\partial y_2}\frac{\partial H}{\partial\left(\frac{\partial^2 x_2}{\partial y_2^2}\right)},$$

and

$\displaystyle\frac{\partial}{\partial y_1}\frac{\partial H}{\partial\left(\frac{\partial^2 x_1}{\partial y_1 \partial y_2}\right)}$ are equal to zero at the boundary of (y_1, y_2) and $t=T$.

$$(7)\ (h_u^T \rho)_x = \frac{\partial}{\partial x}\left\{\begin{bmatrix} \frac{\partial h_1}{\partial u_1} & \frac{\partial h_2}{\partial u_1} & \frac{\partial h_3}{\partial u_1} \\ \frac{\partial h_1}{\partial u_2} & \frac{\partial h_2}{\partial u_2} & \frac{\partial h_3}{\partial u_2} \end{bmatrix}\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{bmatrix}\right\} = \begin{bmatrix} \frac{\partial p_1}{\partial x_1} & \frac{\partial p_1}{\partial x_2} \\ \frac{\partial p_2}{\partial x_1} & \frac{\partial p_2}{\partial x_2} \end{bmatrix}$$

where

$$p_i = \frac{\partial h_1}{\partial u_i}\rho_1 + \frac{\partial h_2}{\partial u_i}\rho_2 + \frac{\partial h_3}{\partial u_i}\rho_3\,, \qquad i=1, 2\,.$$

3. Statement of Primal Problem P

Find the optimum control $u^*(t, y)$ so as to minimize the functional

$$J_p = \int_\Omega F(x(T, y), T)d\Omega + \int_{t_0}^{T}\int_\Omega f_0[t, y, x(t, y), x_{y^k}(t, y), u(t, y)]d\Omega dt \tag{3.1}$$

subject to the constraints

$$x_t(t, y) = f[t, y, x(t, y), x_{y^k}(t, y), u(t, y)] \text{ in } Q \tag{3.2}$$

$$h_i(t, y, x(t, y), u(t, y)) \leqq 0, \qquad i=1, 2, \ldots, l \tag{3.3}$$

with the initial and boundary conditions

$$x(t_0, y),\ x(t, y)|_{\partial\Omega},\quad x_y(t, y)|_{\partial\Omega}, \ldots, x_{y^{k-1}}(t, y)|_{\partial\Omega} \tag{3.4}$$

specified at t_0 and $\partial\Omega$, where $Q=(t_0, T]\times\Omega$; Ω is an open set in R^m (m-dimensional Euclidean space); the terminal time T and the spatial domain Ω are fixed.

Let the Problem P with condition (3.3) being deleted be called Problem P'. We shall first obtain the necessary optimality for Problem P'. Then, we shall see how the results are modified when the inequality constraints (3.3) are added.

4. Necessary Conditions for Problem P

Define a function H by

$$\begin{aligned} H(t, y, x(t, y), x_{y^k}(t, y), u(t, y), \lambda(t, y)) = \\ = f_0(t, y, x(t, y), x_{y^k}(t, y), u(t, y), \lambda(t, y) \\ + \lambda^T(t, y) f(t, y, x(t, y), x_{y^k}(t, y), u(t, y)) \end{aligned} \tag{4.1}$$

and consider the first order variation in δx, $\delta(x_{y^k})$, and δu. Using the standard variational technique results in the following Lemma 1.

Lemma 1 [4]: *If (x^*, u^*) is a pair of extremal solutions of Problem P', then there exists a continuous vector function $\lambda(t, y)$ such that*

$$x_t = H_\lambda = f \tag{4.2}$$

$$\lambda_t = -H_x - (-1)^k H_{x_{y^k} y^k} \tag{4.3}$$

$$F_x = \lambda(T, y) \text{ at } t = T \tag{4.4}$$

$$H_u = 0 \tag{4.5}$$

$$H_{x_{y^k} y^{k-1}} = 0 \text{ at } y = \partial\Omega \quad \text{and} \quad t = T\,. \tag{4.6}$$

In the derivation of Lemma 1 the following equality has been utilized [4].

$$(\delta x_{y^k})^T (H_{x_{y^k}}) = \Lambda + (-1)^{k-1} \left(\frac{\partial}{\partial y}\right)^T \{(\delta x)(H_{x_{y^k} y^{k-1}})\} + (-1)^k (\delta x)^T (H_{x_{y^k} y^k}) \tag{4.7}$$

where

$$\Lambda = \left(\frac{\partial}{\partial y}\right)^T \{(\delta x)_{y^{k-1}} H_{x_{y^k}}\} - \left(\frac{\partial}{\partial y}\right)^T \{(\delta x)_{y^{k-2}} H_{x_{y^k} y}\}$$
$$+ \left(\frac{\partial}{\partial y}\right)^T \{(\delta x)_{y^{k-3}} H_{x_{y^k} y^2}\} \ldots + (-1)^{k-2} \left(\frac{\partial}{\partial y}\right)^T \{(\delta x)_y H_{x_{y^k} y^{k-2}}\} \tag{4.8}$$

and

$$\int_{t_0}^{T} \int_{\Omega} \Lambda \, d\Omega \, dt = \int_{t_0}^{T} [(\delta x)_{y^{k-1}} H_{x_{y^k}} - \ldots] \, dt = 0 \tag{4.9}$$

for fixed

$$x_y(t, y)|_{\partial\Omega},\ x_{y^2}(t, y)|_{\partial\Omega},\ \ldots,\ x_{y^{k-1}}(t, y)|_{\partial\Omega}\,.$$

Now, we shall see how the results developed in Lemma 1 are modified if inequality constraints (3.3) are considered.

Let the extremal control $u^*(t, y)$ and the corresponding $x^*(t, y)$ be such that

$$\begin{aligned} &h_i(t, y, x^*, u^*) = 0 \qquad i = 1, 2, \ldots, \alpha \\ &h_i(t, y, x^*, u^*) < 0 \qquad i = \alpha+1, \alpha+2, \ldots, l\,. \end{aligned} \tag{4.10}$$

We shall assume that the matrix

$$h_u = \begin{bmatrix} \dfrac{\partial h_1}{\partial u_1} & \cdots & \dfrac{\partial h_1}{\partial u_r} \\ \vdots & & \vdots \\ \dfrac{\partial h_\alpha}{\partial u_1} & \cdots & \dfrac{\partial h_\alpha}{\partial u_r} \end{bmatrix}$$

has maximum rank at $x = x^*(t, y)$ and $u = u^*(t, y)$. Thus, for each t and y there exists a neighborhood of point $(x^*(t, y), u^*(t, y))$ in $R^n \times R^r$ such that

$$h_i(t, y, x, u) = 0, \qquad i = 1, 2, \ldots, \alpha \tag{4.11}$$

may be solved uniquely for α components of u as functions of t, y, x and the remaining $r - \alpha$ components of u. Let us define

$$u^a = \text{col.}\,(u_1, u_2, \ldots, u_\alpha) \tag{4.12}$$

$$u^b = \text{col.}\,(u_{\alpha+1}, u_{\alpha+2}, \ldots, u_r) \tag{4.13}$$

$$u = \text{col.}\,(u^a, u^b) \tag{4.14}$$

and

$$R = \text{col.}\,(h_1, h_2, \ldots, h_\alpha)\,. \tag{4.15}$$

Then, there exists a function

$$u^a = U^a(t, y, x, u^b) \tag{4.16}$$

such that

$$R(t, y, x, U^a(t, y, x, u^b), u^b) = 0 \,. \tag{4.17}$$

In view of (4.14) and (4.16) equations (3.2) and (4.1) may be, respectively, written as

$$x_t(t, y) = f(t, y, x, x_{y^k}, U^a, u^b) \tag{4.18}$$

$$H(t, y, x, x_{y^k}, U^a, u^b, \lambda) = f_0(t, y, x, x_{y^k}, U^a, u^b) + \lambda^T(t, y) f(t, y, x, x_{y^k}, U^a, u^b) \,. \tag{4.19}$$

Differentiating (4.17) with respect to x and u^b, respectively, gives

$$(R_x + R_{u^a} U^a_x)\delta x = 0 \tag{4.20}$$

$$(R_{u^b} + R_{u^a} U^a_{u^b})\delta u^b = 0 \tag{4.21}$$

from which

$$U^a_x \delta x = -(R_{u^a})^{-1} R_x \delta x \tag{4.22}$$

$$U^a_{u^b} \delta u^b = -(R_{u^a})^{-1} R_{u^b} \delta u^b \,. \tag{4.23}$$

It follows from (4.7), (4.22), and (4.23) that the first order variation of H in x, x_{y^k}, and u^b is given by

$$\delta H = [H_x^T - (H_{u^a})^T (R_{u^a})^{-1} R_x]\,\delta x + [(H_{u^b})^T - (H_{u^a})^T (R_{u^a})^{-1} R_{u^b}]\,\delta u^b + (H_{x_{y^k}})^T \delta(x_{y^k}) \tag{4.24}$$

or

$$\delta H = (\delta x)^T \{H_x - [(H_{u^a})^T (R_{u^a})^{-1} R_x]^T\} + (\delta u^b)^T \{H_{u^b} - [(H_{u^a})^T (R_{u^a})^{-1} R_{u^b}]^T\}$$
$$+ \Lambda + (-1)^{k-1} \left(\frac{\partial}{\partial y}\right)^T \{(\delta x) H_{x_{y^k} y^{k-1}}\} + (-1)^k (\delta x)^T (H_{x_{y^k} y^k}) \,. \tag{4.25}$$

Consequently, after substituting (4.25) into the first variation in J_p, expressed by

$$\delta J_p = \int_\Omega \delta F \, d\Omega + \int_{t_0}^T \int_\Omega [\delta H - \lambda^T (\delta x_t)] \, d\Omega \, dt =$$
$$= \int_\Omega (\delta x)^T [F_x - \lambda(t, y)]_{t=T} \, d\Omega +$$
$$+ \int_{t_0}^T \int_\Omega (\delta x)^T \{H_x - [(H_{u^a})^T (R_{u^a})^{-1} R_x]^T + \lambda_t + (-1)^k H_{x_{y^k} y^k}\} +$$
$$+ (\delta u^b)^T \{H_{u^b} - [(H_{u^a})^T (R_{u^a})^{-1} R_{u^b}]^T\} +$$
$$+ \int_{t_0}^T \int_\Omega (-1)^{k-1} \left(\frac{\partial}{\partial y}\right)^T \{(\delta x) H_{x_{y^k} y^{k-1}}\} \, d\Omega \, dt \,,$$

and setting δJ equal to zero, conditions (4.3) and (4.5) become

$$\lambda_t = \begin{cases} -H_x + [(H_{u^a})^T (R_{u^a})^{-1} R_x]^T - (-1)^k (H_{x_{y^k} y^k}) & \text{if} \quad h_i = 0 \\ -H_x - (-1)^k (H_{x_{y^k} y^k}) & \text{if} \quad h_j = < 0 \end{cases}$$

and

$$H_{u^b} - [(H_{u^a})^T (R_{u^a})^{-1} R_{u^b}]^T = 0 \qquad \text{if} \quad h_i = 0$$
$$H_u = 0 \qquad \text{if} \quad h_j < 0$$

where $i = 1, 2, \ldots, \alpha$ and $j = \alpha+1, \alpha+2, \ldots, l$

The necessary conditions for Problem P, thus, can be summarized in Theorem 1.

Theorem 1: *If* (x^*, u^*) *is an extremal solution of Problem P, then there exists a function*

$$
\begin{aligned}
&H(t, y, x(t, y), x_{y^k}(t, y), U^a(t, y, x, u^b), u^b(t, y), \lambda(t, y)) \\
&= f_0(t, y, x(t, y), x_{y^k}(t, y), U^a(t, y, x, u^b), u^b(t, y)) + \\
&\quad + \lambda(t, y)^T f(t, y, x(t, y), x_{y^k}(t, y), U^a(t, y, x, u^b), u^b(t, y))
\end{aligned}
\tag{4.26}
$$

such that

$$x_t = H_\lambda = f \tag{4.27}$$

$$\lambda_t = \begin{cases} -H_x + [(H_{u^a})^T (R_{u^a})^{-1} R_x]^T - (-1)^k (H_{x_{y^k} y^k}) & \text{if } h_i = 0 \\ -H_x - (-1)^k (H_{x_{y^k} y^k}) & \text{if } h_j < 0 \end{cases} \tag{4.28}$$

$$\left.\begin{aligned} &H_{u^b} - [(H_{u^a})^T (R_{u^a})^{-1} R_{u^b}]^T = 0 && \text{if } h_i = 0 \\ &H_u = 0 && \text{if } h_j < 0 \end{aligned}\right\} \tag{4.29}$$

$$F_x(T, x) = \lambda(T, y) \tag{4.30}$$

$$H_{x_{y^k} y^{k-1}} = 0 \quad \textit{for} \quad y = d\Omega \quad \textit{and} \quad t = T, \tag{4.31}$$

with specified $x(t_0, y), x(t, y)|_{\partial\Omega}, \ldots, x_{y^{k-1}}(t, y)|_{\partial\Omega}$. (4.32)

It will be demonstrated in the following that conditions (4.27–4.32) derived in this paper are equivalent to that derived from the method of Valentine [6]. Let us introduce a new variable $\rho = \text{col.}(\rho^a, \rho^b)$ such that

$$\rho^a = \text{col.}(\rho_1, \rho_2, \ldots, \rho_\alpha) = -[(H_{u^a})^T (R_{u^a})^{-1}]^T > 0 \tag{4.33}$$

and

$$\rho^b = \text{col.}(\rho_{\alpha+1}, \rho_{\alpha+2}, \ldots, \rho_l) = 0\,. \tag{4.34}$$

Then, we immediately have the following Corollary 1.

Corollary 1: If (x^*, u^*) is an extremal solution of Problem P and if $[(H_{u^b})^T (R_{u^b})^{-1}]^T < 0$, then, there exist continuous vector functions $\lambda(t, y)$ and $\rho(t, y)$ defined by (4.33) and (4.34) such that

$$x_t = H_\lambda = f \tag{4.35}$$

$$\lambda_t = -H_x - (h_x)^T \rho - (-1)^k (H_{x_{y^k} y^k}) \tag{4.36}$$

$$H_u + (h_u)^T \rho = 0 \tag{4.37}$$

$$F_x(T, y) = \lambda(T, y) \tag{4.38}$$

$$H_{x_{y^k} y^{k-1}} = 0 \quad \text{for} \quad y = \partial\Omega \text{ and } t = T \tag{4.39}$$

$$h_i^T \rho_i = 0 \qquad (i = 1, 2, \ldots, l) \quad \text{and} \quad \rho \geqq 0 \tag{4.40}$$

with specified

$$x(t_0, y),\ x(t, y)|_{\partial\Omega},\ x_y(t, y)|_{\partial\Omega}, \ldots, x_{y^{k-1}}(t, y)|_{\partial\Omega}\,. \tag{4.41}$$

5. Sufficient Conditions for Problem P

The sufficient conditions for Problem P are summarized in Theorem 2.

Theorem 2: *Let* $(x^*, u^*, \lambda^*, \rho^*)$ *be the unique solution to* (4.35–4.41). *If three conditions given below hold*

(1) *F is convex in x,* (5.1)

(2) f_0 *and* $f^T\lambda$ *are convex in* (x, x_{y^k}, u), (5.2)

(3) *h is convex in* (x, u), (5.3)

then, $u^*(t, y)$ *is optimal for Problem P.*

Proof: For simplicity we denote $f_i(t, y, x^*, x^*_{y^k}, u^*)$ by f_i^* and $f_i(t, y, x, x_{y^k}, u)$ by f_i $(i=0, 1, \ldots, n)$. Similar notations are applied to functions F and h_i $(i=1, 2, \ldots, l)$. Let $\delta x = x - x^*$, $\delta u = u - u^*$, and $\delta(x_{y^k}) = x_{y^k} - x^*_{y^k}$. Then,

$$J_p(x, x_{y^k}, u, \lambda^*, \rho^*) - J_p(x^*, x^*_{y^k}, u^*, \lambda^*, \rho^*) =$$

$$= \int_\Omega (F - F^*)d\Omega + \int_{t_0}^{T}\int_\Omega (f_0 - f_0^*)d\Omega\, dt \tag{5.4}$$

$$\geqq \int_\Omega (\delta x)^T F_x|_{t=T} d\Omega + \int_{t_0}^{T}\int_\Omega \{(\delta x)^T f_{0x} + (\delta x_{y^k})^T f_{0x_{y^k}} + (\delta u)^T f_{0u}\} d\Omega\, dt$$

(by the convexity of F and f_0)

$$= \int_\Omega (\delta x)^T F_x|_{t=T} d\Omega + \int_{t_0}^{T}\int_\Omega \{(\delta x)^T [H_x - (f_x)^T \lambda^*]$$

$$+ (\delta x_{y^k})^T [H_{x_{y^k}} - (f_{x_{y^k}})^T \lambda^*] + (\delta u)^T [H_u - (f_u)^T \lambda^*]\} d\Omega\, dt$$

(by the definition of H in (4.1))

$$= \int_\Omega (\delta x)^T F_x|_{t=T} d\Omega + \int_{t_0}^{T}\int_\Omega (\delta x)^T [-\lambda_t^* - (h_x)^T \rho^* - (-1)^k (H_{x_{y^k} y^k})$$

$$- (f_x)^T \lambda^*] d\Omega\, dt + \int_{t_0}^{T}\int_\Omega \{(-1)^{k-1} \left(\frac{\partial}{\partial y}\right)^T [(\delta x)(H_{x_{y^k} y^{k-1}})]$$

$$+ (-1)^k (\delta x)^T (H_{x_{y^k} y^k})\} d\Omega\, dt - \int_{t_0}^{T}\int_\Omega (\delta x_{y^k})^T (f_{x_{y^k}})^T \lambda^* d\Omega\, dt$$

$$+ \int_{t_0}^{T}\int_\Omega (\delta u)^T [-(h_u)^T \rho^* - (f_u)^T \lambda^*] d\Omega\, dt$$

(by equations (4.7), (4.9), (4.36), and (4.37))

$$= \int_\Omega (\delta x)^T F_x|_{t=T} d\Omega + \int_{t_0}^{T}\int_\Omega (\delta x)^T [-\lambda_t^* - (h_x)^T \rho^* - (f_x)^T \lambda^*] d\Omega\, dt$$

$$- \int_{t_0}^{T}\int_\Omega (\delta x_{y^k})^T (f_{x_{y^k}})^T \lambda^* d\Omega\, dt + \int_{t_0}^{T}\int_\Omega (\delta u)^T \{-(h_u)^T \rho^* - (f_u)^T \lambda^*\} d\Omega\, dt \tag{5.5}$$

(by equation (4.39)).

Using integration by part and noting that the incremental variation of x with respect to t at t_0 is zero (i.e. $\delta x(t_0, y) = 0$) we have

$$\int_{t_0}^{T}\int_\Omega (\delta x)^T \lambda_t^* d\Omega\, dt = \int_\Omega (\delta x)^T \lambda^*|_{t=T} d\Omega - \int_{t_0}^{T}\int_\Omega (x_t - x_t^*)^T \lambda^* d\Omega\, dt\,. \tag{5.6}$$

Substituting (5.6) into (5.5) yields

$$J_p(x, x_{y^k}, u, \lambda^*, \rho^*) - J_p(x^*, x^*_{y^k}, u^*, \lambda^*, \rho^*) \geqq$$

$$\geqq \int_\Omega [(\delta x)^T (F_x - \lambda^*)|_{t=T} d\Omega + \int_{t_0}^T \int_\Omega (x_t - x_t^*)^T \lambda^* d\Omega dt$$

$$+ \int_{t_0}^T \int_\Omega - (f_x \delta x + f_{x_{y^k}} \delta(x_{y^k}) + f_u \delta u)^T \lambda^* d\Omega dt + \int_{t_0}^T \int_\Omega - (h_x \delta x + h_u \delta u)^T \rho^* d\Omega dt$$

$$\geqq \int_{t_0}^T \int_\Omega (x_t - x_t^*)^T \lambda^* d\Omega dt + \int_{t_0}^T \int_\Omega (f^* - f)^T \lambda^* d\Omega dt$$

$$+ \int_{t_0}^T \int_\Omega (h^* - h)^T \rho^* d\Omega dt \tag{5.7}$$

(by the convexity of $f^T \lambda$ and h, and (4.38))

$$= \int_{t_0}^T \int_\Omega (h^* - h)^T \rho^* d\Omega dt \quad \text{(by (4.35))}$$

$$\geqq 0. \quad \text{(by (3.3) and (4.40))}.$$

This completes the proof.

6. Duality

Consider the following two problems:

Primal Problem P
Minimize

$$J_p = \int_\Omega F(x(T, Y), T) d\Omega + \int_{t_0}^T \int_\Omega f_0(t, y, x, x_{y^k}, u) d\Omega dt \tag{6.1}$$

subject to

$$x_t = f(t, y, x, x_{y^k}, u) \tag{6.2}$$

$$h_i(t, y, x, u) \leqq 0 \qquad i = 1, 2, \ldots, \alpha, \alpha+1, \ldots, l \tag{6.3}$$

with specified

$$x(t_0, y),\ x(t, y)|_{\partial\Omega},\ x_y(t, y)|_{\partial\Omega}, \ldots, x_{y^{k-1}}(t, y)|_{\partial\Omega}, \tag{6.4}$$

where we assume that $\alpha \leqq l$ of constraints (6.3) are equalities as given by (4.10).

Dual Problem D
Maximize

$$J_d = \int_\Omega F(x(T, y), T) d\Omega + \int_{t_0}^T \int_\Omega \{f_0(t, y, x, x_{y^k}, u) + [f(t, y, x, x_{y^k}, u) - x_t]^T \lambda$$

$$+ h^T(t, y, x, u) \rho\} d\Omega dt \tag{6.5}$$

subject to

$$\lambda_t = -H_x - h_x^T \rho - (-1)^k (H_{x_{y^k} y^k}) \tag{6.6}$$

$$H_u + h_u^T \rho = 0 \tag{6.7}$$

$$F_x(T, y) = \lambda(T, y) \tag{6.8}$$

$$H_{x_{y^k} y^{k-1}} = 0 \text{ for } y = \partial\Omega \text{ and } t = T \tag{6.9}$$

$$\rho(t, y) \geqq 0 \tag{6.10}$$

with specified

$$x(t_0, y),\ x(t, y)|_{\partial\Omega},\ x_y(t, y)|_{\partial\Omega},\ \ldots,\ x_{y^{k-1}}(t, y)|_{\partial\Omega} \tag{6.11}$$

where $\rho = \text{col.}(\rho^a, \rho^b)$ is defined by (4.33–4.34).

Equation (6.5) is the Lagrangian form of the Primal problem P subject to constraints (6.6–6.11) which are the first order necessary conditions for a minimal primal solution as stated in Corollary 1.

Theorem 3. *If F, f_0, $f^T\lambda$ and h possess the same properties as given by* (5.1–5.3), *then, the infimum of Problem P is greater than or equal to the supremum of Problem D.*

Proof: Let (x, u) satisfy (6.2–6.4) and let $(x^*, u^*, \lambda^*, \rho^*)$ satisfy (6.6–6.11). Then, it follows from (5.4) and (5.7) that

$$\begin{aligned}
&\int_\Omega [F(x(T, y), T) - F(x^*(T, y), T)]\, d\Omega + \\
&\quad + \int_{t_0}^T \int_\Omega [f_0(t, y, x, x_{y^k}, u) - f_0(t, y, x^*, x^*_{y^k}, u^*)]\, d\Omega\, dt \\
&\quad \geqq \int_{t_0}^T \int_\Omega (x_t - x_t^*)^T \lambda^*\, d\Omega + \int_{t_0}^T \int_\Omega (f^* - f)^T \lambda^*\, d\Omega\, dt + \int_{t_0}^T \int_\Omega (h^* - h)^T \rho^*\, d\Omega\, dt \\
&\quad \geqq \int_{t_0}^T \int_\Omega (f^* - x_t^*)^T \lambda^*\, d\Omega\, dt + \int_{t_0}^T \int_\Omega (h^*)^T \rho^*\, d\Omega\, dt
\end{aligned} \tag{6.12}$$

(by (6.2), (6.3), and (6.10)).

Here we note that $x_t^* \neq f^*$. Rearranging (6.12) gives

$$\begin{aligned}
&\int_\Omega F(x(T, y), T)\, d\Omega + \int_{t_0}^T \int_\Omega f_0(t, y, x, x_{y^k}, u)\, d\Omega\, dt \\
&\geqq \int_\Omega F(x^*(T, y), T)\, d\Omega + \int_{t_0}^T \int_\Omega \{f_0(t, y, x^*, x^*_{y^k}, u^*) \\
&\quad + [f(t, y, x^*, x^*_{y^k}, u^*) - x_t^*]^T \lambda^* + h^T(t, y, x^*, u^*)\rho^*\}\, d\Omega\, dt\,.
\end{aligned}$$

Therefore, the infimum of Problem P is greater than or equal to the supremum of Problem D.

Corollary 1 and Theorem 3 immediately lead to the following results.

Theorem 4: *If (x^*, u^*) is an optimal solution of Problem P, then, there exist vector functions $\lambda^*(t, y)$ and $\rho^*(t, y)$ such that $(x^*, u^*, \lambda^*, \rho^*)$ is an extremal solution of Problem D and the extremal values of Problem P and Problem D are equal.*

7. Converse Duality

In this section it is assumed that F, f_0, f and h have continuous derivatives up to and including the third order with respect to each of their arguments. We shall find conditions under which the existence of an extremal solution of Problem D implies the existence of an extremal solution to Problem P.

Let us introduce a vector $\tau = \text{col.}(\tau^a, \tau^b)$, which plays a similar role as the vector $\rho = \text{col.}(\rho^a, \rho^b)$ defined by (4.33–4.34), such that

$$\tau^a = \text{col.}(\tau_1, \tau_2, \ldots, \tau_\alpha) = 0 \tag{7.1}$$

$$\tau^b = \text{col.}(\tau_{\alpha+1}, \tau_{\alpha+2}, \ldots, \tau_l) > 0\,. \tag{7.2}$$

Then, it is clear that

$$\rho^T \tau = 0\,. \tag{7.3}$$

Define $\hat{J}_d$, H^1, H^2, and H^3, respectively, by

$$\hat{J}_d = J_d + \int_{t_0}^{T} \int_{\Omega} \{[\lambda_t + H_x + h_x^T \rho + (-1)^k H_{x_{y^k} y^k}]^T \xi + [H_u + h_u^T \rho]^T \eta + \rho^T \tau\} \, d\Omega \, dt \,,$$

$$H^1 = H_x^T \xi, \quad H^2 = [(-1)^k H_{x_{y^k} y^k}]^T \xi, \quad \text{and} \quad H^3 = H_u^T \eta \,.$$

Then, with the aid of (4.7–4.9), we have

$$\begin{aligned}\int_{t_0}^{T} \int_{\Omega} (\delta x_{y^k})^T (H^1_{x_{y^k}}) &= \int_{t_0}^{T} \int_{\Omega} \left\{(-1)^{k-1} \left(\frac{\partial}{\partial y}\right)^T [(\delta x)(H^1_{x_{y^k} y^{k-1}})] + (-1)^k (\delta x)^T (H^1_{x_{y^k} y^k})\right\} d\Omega \, dt \\ &= \int_{t_0}^{T} \int_{\Omega} \left\{(-1)^{k-1} \left(\frac{\partial}{\partial y}\right)^T [(\delta x)(H_{x x_{y^k} y^{k-1}})^T \xi] + (-1)^k [(\delta x)^T (H_{x x_{y^k} y^k})^T \xi]\right\} d\Omega \, dt\end{aligned}$$

for fixed

$$x_y(t, y)|_{\partial\Omega}, x_{y^2}(t, y)|_{\partial\Omega}, x_{y^3}(t, y)|_{\partial\Omega}, \ldots, x_{y^{k-1}}(t, y)|_{\partial\Omega} \,.$$

Similarly

$$\int_{t_0}^{T} \int_{\Omega} (\delta x_{y^k})^T (H^2_{x_{y^k}}) d\Omega \, dt = \int_{t_0}^{T} \int_{\Omega} \left\{(-1)^{2k-1} \left(\frac{\partial}{\partial y}\right)^T [(\delta x)(H_{x_{y^k} y^k x_{y^k} y^{k-1}})^T \xi] + (-1)^{2k} [(\delta x)^T (H_{x_{y^k} y^k x_{y^k} y^k})^T \xi]\right\} d\Omega \, dt$$

$$\int_{t_0}^{T} \int_{\Omega} (\delta x_{y^k})^T (H^3_{x_y}) d\Omega \, dt = \int_{t_0}^{T} \int_{\Omega} \left\{(-1)^{k-1} \left(\frac{\partial}{\partial y}\right)^T [(\delta x)(H_{u x_{y^k} y^{k-1}})^T \eta] + (-1)^k (\delta x)^T (H_{u x_{y^k} y^k})^T \eta]\right\} d\Omega \, dt \,.$$

Taking the variations of δx, δx_{y^k}, δu, $\delta\lambda$ and $\delta\rho$ about x^*, $x^*_{y^k}$, u^*, λ^* and ρ^* and noting that $\delta x(t_0, y) = 0$ we can find that the first variation in $\hat{J}_d$ has the form

$$\begin{aligned}\delta \hat{J}_d = &\int_{\Omega} (\delta x(T, y))^T (F_x - \lambda)|_{t=T} d\Omega \\ &+ \int_{t_0}^{T} \int_{\Omega} (\delta x)^T \{H_x + h_x^T \rho + (-1)^k H_{x_{y^k} y^k} + [H_{xx} + (h_x^T \rho)_x + (-1)^k H_{x_{y^k} y^k x} \\ &+ (-1)^k H_{x x_{y^k} y^k} + (-1)^{2k} H_{x_{y^k} y^k x_{y^k} y^k}]^T \xi \\ &+ [H_{ux} + (h_u^T \rho)_x + (-1)^k H_{u x_{y^k} y^k}]^T \eta + \lambda_t\} \, d\Omega \, dt \\ &+ \int_{t_0}^{T} \int_{\Omega} (\delta u)^T \{H_u + h_u^T \rho + [H_{xu} + (h_x^T \rho)_u + (-1)^k H_{x_{y^k} y^k u}]^T \xi + [H_{uu} + (h_u^T \rho)_u]^T \eta\} \, d\Omega \, dt \\ &+ \int_{t_0}^{T} \int_{\Omega} (\delta\lambda)^T \{H_\lambda - x_t + (H_{x\lambda})^T \xi + (-1)^k (H_{x_{y^k} y^k \lambda})^T \xi + (H_{u\lambda})^T \eta - \xi_t\} \, d\Omega \, dt \\ &+ \int_{\Omega} (\delta\lambda)^T \xi \Big|_{t_0}^{T} d\Omega + \int_{t_0}^{T} \int_{\Omega} (\delta\rho)^T \{h + h_x \xi + h_u \eta + \tau\} \, d\Omega \, dt \\ &+ \int_{t_0}^{T} \int_{\Omega} \left(\frac{\partial}{\partial y}\right)^T \{(\delta x)[(-1)^{k-1} H_{x_{y^k} y^{k-1}} + (-1)^{k-1} (H_{x x_{y^k} y^{k-1}})^T \xi \\ &+ (-1)^{2k-1} (H_{x_{y^k} y^k x_{y^k} y^{k-1}})^T \xi + (-1)^{k-1} (H_{u x_{y^k} y^{k-1}})^T \eta\} \, d\Omega \, dt \,.\end{aligned}$$

The necessary conditions for Problem D, obtained by setting the first variation in $\hat{J}_d$ equal to zero, are

$$F_x(x(T, y), T) = \lambda(T, y) \tag{7.4}$$

$$\begin{aligned} &H_x + h_x^T \rho + (-1)^k (H_{x_{y^k} y^k}) + [H_{xx} + (h_x^T \rho)_x + (-1)^k H_{x_{y^k} y^k x} \\ &\quad + (-1)^k H_{x x_{y^k} y^k} + (-1)^{2k} H_{x_{y^k} y^k x_{y^k} y^k}]^T \xi \\ &\quad + [H_{ux} + (h_u^T \rho)_x + (-1)^k H_{u x_{y^k} y^k}]^T \eta = -\lambda_t \end{aligned} \tag{7.5}$$

$$H_u + h_u^T \rho + [H_{xu} + (h_x^T \rho)_u + (-1)^k H_{x_{y^k} y^k u}]^T \xi + [H_{uu} + (h_u^T \rho)_u]^T \eta = 0 \tag{7.6}$$

$$H_\lambda - x_t + (H_{x\lambda})^T \xi + (-1)^k (H_{x_{y^k} y^k \lambda})^T \xi + (H_{u\lambda})^T \eta = \xi_t \tag{7.7}$$

$$\xi(t_0, y) = \xi(T, y) = 0 \tag{7.8}$$

$$h + h_x \xi + h_u \eta + \tau = 0 \tag{7.9}$$

$$\begin{aligned} &H_{x_{y^k} y^{k-1}} + (H_{x x_{y^k} y^{k-1}})^T \xi + (-1)^k (H_{x_{y^k} y^k x_{y^k} y^{k-1}})^T \xi \\ &\quad + (H_{u x_{y^k} y^{k-1}})^T \eta = 0 \ \text{ for } \ y = \partial\Omega \ \text{ and } \ t = T \end{aligned} \tag{7.10}$$

Conditions in (4.34), (7.1), (7.5–7.7), and (7.9), together with (6.6–6.7) comprise a set of $3n+2r+2l$ equations which can be utilized to solve $3n+2r+2l$ variables x, λ, ξ, u, η, ρ and τ with initial and boundary conditions given by (6.9), (6.11), (7.4), (7.8), and (7.10).

Theorem 5. *If $(x^*, u^*, \lambda^*, \rho^*)$ is an extremal solution for Problem D such that the matrix*

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{7.11}$$

is nonsingular for all $t \in [t_0, T]$ and $y \in \bar{\Omega}$, where

$$M_{11} = [H_{xx} + (h_x^T \rho)_x + (-1)^k H_{x_{y^k} y^k x} + (-1)^k H_{x x_{y^k} y} + (-1)^{2k} H_{x_{y^k} y^k x_{y^k} y^k}]^T$$

$$M_{12} = [H_{ux} + (h_u^T \rho)_x + (-1)^k H_{u x_{y^k} y^k}]^T$$

$$M_{21} = [H_{xu} + (h_x^T \rho)_u + (-1)^k H_{x_{y^k} y^k u}]^T$$

$$M_{22} = [H_{uu} + (h_u^T \rho)_u]^T ,$$

and if conditions (5.1–5.3) are held for $F, f_0, f^T \lambda$, and h, then (x^*, u^*) is an optimal solution for Problem P and the extremal values of J_p and J_d are equal.

Proof: Since $(x^*, u^*, \lambda^*, \rho^*)$ is an extremal solution for Problem D it follows from (6.6), (6.7), (7.1–7.3), and (7.5–7.6) that

$$\begin{aligned} &[H_{xx} + (h_x^T \rho)_x + (-1)^k H_{x_{y^k} y^k x} + (-1)^k H_{x x_{y^k} y^k} + (-1)^{2k} H_{x_{y^k} y^k x_{y^k} y}]^T \xi \\ &\quad + [H_{ux} + (h_u^T \rho)_x + (-1)^k H_{u x_{y^k} y^k}]^T \eta = 0 \end{aligned} \tag{7.12}$$

$$[H_{xu} + (h_x^T \rho)_u + (-1)^k H_{x_{y^k} y^k u}]^T \xi + [H_{uu} + (h_u^T \rho)_u]^T \eta = 0 \tag{7.13}$$

$$\tau(t, y) \geqq 0 \tag{7.14}$$

$$\tau_i \rho_i = 0 \qquad i = 1, 2, \ldots, l\,. \tag{7.15}$$

Since M is nonsingular we conclude that $\xi = \eta = 0$, for $t \in [t_0, T]$ and $y \in \bar{\Omega}$, is the only solution of (7.12–7.13). Equations (7.7) and (7.9) now become

$$f - x_t = 0 \tag{7.16}$$

$$h = -\tau \leqq 0 \tag{7.17}$$

in view of (7.14) and (4.1).

Hence (x^*, u^*) satisfies the constraints of Problem P. Also, equations (7.15) and (7.17) imply that

$$\rho^T h = 0 . \tag{7.18}$$

The theorem now follows from (7.16–7.18), (6.1), (6.5), and Theorem 3.

Remark: For convexity of $F, f_0, f^T\lambda$, and h the condition that the matrix M be nonsingular is equivalent to the condition that M be positive definite.

8. Example

The mathematical model for a temperature $x(t, y)$ variation in a slab, with both faces insulated at $y=0$ and $y=\pi$ and with initial temperature $x_0(y)$, may be described by

$$x_t(t, y) = cx_{y^2}(t, y) + u(t, y), \quad (t, y) \in Q = (0, T) \times (0, \pi)$$

$$x(0, y) = x_0(y), \quad y \in (0, \pi), \quad x_0 \in R^1$$

$$x_y(t, y)|_{y=0} = x_y(t, y)|_{y=\pi} = 0 ,$$

where $x_t = \partial x/\partial t$, $x_{y^2} = \partial^2 x/\partial y^2$, and $u(t, y)$ is a source of heat in Q with

$$|u(t, y)| \leqq L \quad (L \text{ is a positive constant})$$

and c is the thermal conductivity.

Problem P: Minimize

$$J_p = \tfrac{1}{2}\int_0^T \int_0^\pi [x^2(t, y) + ru^2(t, y)] dy\, dt \tag{8.1}$$

subject to

$$x_t = cx_{y^2} + u$$

$$h_1 = u - L \leqq 0$$

$$h_2 = -u - L \leqq 0$$

with specified

$$x(0, y) = x_0(y)$$

and

$$x_y(t, 0) = x_y(t, \pi) = 0$$

where r is a given constant.

Problem D: Maximize

$$J_d = \int_0^T \int_0^\pi [\tfrac{1}{2}(x^2 + ru^2) + (cx_{y^2} + u - x_t)\lambda + h_1\rho_1 + h_2\rho_2]\, dy\, dt \tag{8.2}$$

subject to

$$\lambda_t = -x - c\lambda_{y^2}$$

$$ru + \lambda + \rho_1 - \rho_2 = 0$$

$$\rho_1 \geqq 0, \ \rho_2 \geqq 0$$

with initial and boundary conditions

$$\lambda(T, y) = 0 , \qquad \lambda_y(t, y)|_0^\pi = 0 , \qquad x(0, y) = x_0(y), \quad x_y(t, y)|_0^\pi = 0 .$$

It can be shown that the matrix M in this example has the form

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & r \end{bmatrix}.$$

It is clear that M is nonsingular if $r > 0$.

Let $(x^*, u^*, \lambda^*, \rho^*)$ be the extremal solution to Problem D. Then, according to Theorem 5 (x^*, u^*) is the optimal solution to Problem P if $r > 0$. The condition that $r > 0$ to achieve the optimal solution for Problem P is similar to that of the minimum energy problem in the lumped parameter systems [7], where it is required that the matrix R in the performance index

$$J = \tfrac{1}{2}\int_{t_0}^{t_1} [x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)]\, dt$$

be positive definite for $t \in [t_0, t_1]$.

9. Conclusion

The main contribution of this paper is to develop dual and converse dual theorems for a wide class of distributed parameter systems. The dual theorem gives conditions under which an extremal solution of the primal control problem in the distributed systems yields a solution of the corresponding dual. The converse duality, on the other hand, gives conditions under which a solution of the dual problem yields a solution of the control problem. An example for the applications of the theory is illustrated.

REFERENCES

[1] R. E. Kalman, A new Approach to Linear Filtering and Prediction Problems, *Trans. ASME*, Series D, *J. Basic Engr.*, 82 (1960) 34–35.

[2] B. Mond and M. Hanson, Duality for Control Problems, *SIAM J. Control*, 6, 1, (1968) 114–120.

[3] J. D. Pearson, Duality and a Decomposition Technique, *SIAM J. Control*, 4, 1, (1966) 164–172.

[4] A. P. Sage, *Optimum System Control*, Prentice-Hall, Inc., (1968) 137–164.

[5] M. C. Y. Kuo, The Geometric Aspects of a Class of Differential Games, Part 1, Necessary Conditions, *Int. J. Control*, 10, 4, (1969) 413–432.

[6] F. A. Valentine, *The Problem of Lagrange with Differential Inequalities as Added Side Conditions*, Contributions to the Calculus of Variations, University of Chicago Press, Chicago, (1937) 407–488.

[7] M. Athans and L. P. Falb, Optimal Control, McGraw-Hill, 1966.

[8] A. G. Butkovskii, Optimum Processes in Systems with Distributed Parameters, *Automation and Remote Control*, 22 (1961) 13–21.

[9] P. K. C. Wang, Control of Distributed Parameter Systems, *Advances in Control Systems: Theory and Applications*, Academic Press, (1964) 75–172.