JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 55, 418–433 (1976)

Conjugate Point Properties for Linear Quadratic Problems*

P. M. MEREAU AND W. F. POWERS

Department of Aerospace Engineering, The University of Michigan, Ann Arbor, Michigan 48105

Submitted by George Leitmann

Analogs of certain conjugate point properties in the calculus of variations are developed for optimal control problems. The main result in this direction is concerned with the characterization of a parameterized family of extremals going through the first backward conjugate point, t_c. A corollary of this result is that for the linear quadratic problem (LQP) there exists at least a one-parameter family of extremals going though the conjugate point which gives the same cost as the candidate extremal, i.e., the extremal control is optimal but nonunique on $[t_c, t_f]$. An analysis of the effect on the conjugate point of employing penalty functions for terminal equality constraints in the LQP is presented, also. It is shown that the sequence of approximate conjugate points is always conservative, and it converges to the conjugate point of the constrained problem. Furthermore, it is proved that the addition of terminal constraints has the effect of causing the conjugate point to move backward (or remain the same).

1. INTRODUCTION AND PRELIMINARIES

Since an optimal linear feedback control does not exist on an interval containing a conjugate point, the conjugate point in control problems has received a good deal of attention [1–5]. The trend in the literature has been either to make adequate assumptions to insure that no conjugate point occurs [6–9] or to concentrate on methods to test for conjugate points [10–15] and several aspects of the conjugate point widely studied in the calculus of variations have not been developed for control problems. In this paper, several properties of the conjugate point in linear quadratic problems (LQP) are presented.

In Section 2, analogs of results well known in the calculus of variations such as envelope contacts and the envelope theorem are given for the LQP and control oriented proofs are provided. The main result in this direction is that extremals of the LQP are nonproper optima if the end points are

* This research was supported by the National Science Foundation under Grant GK 30115.

conjugate. In Section 3, the behavior of the conjugate point is examined when the method of penalty functions is employed to approximate the terminal constraints and a conservative, economical test for conjugate points is proposed.

We shall consider the following problem:

$$
\text{(LQP)}\begin{cases}
\text{minimize } J = \frac{1}{2}x_f^T F x_f + \frac{1}{2}\int_{t_0}^{t_f} [x^T C(t)\, x + u^T E(t)\, u]\, dt & (1)\\
\text{subject to } \dot{x} = A(t)\, x + B(t)\, u \qquad t \in [t_0, t_f] & (2)\\
\qquad x(t_0) = x_0 \qquad\qquad t_0, t_f \text{ prescribed} & (3)\\
\qquad T x_f + \psi_0 = 0 & (4)
\end{cases}
$$

where $x \in R^n$, $u \in R^m$, $m \leqslant n$. F, $C(t)$, $E(t)$ are respectively symmetric $(n \times n)$, $(n \times n)$, $(m \times m)$ matrices, and $A(t)$, $B(t)$, T are respectively $(n \times n)$, $(n \times m)$, $(p \times n)$ matrices, $p \leqslant n$. $A(t)$, $B(t)$, $C(t)$ and $E(t)$ are continuous on $[t_0, t_f]$. It is assumed furthermore that,

A.1. The problem is normal [16].

A.2. $E(t)$ is positive definite for all $t \in [t_0, t_f]$ (generalized Legendre–Clebsch condition which implies nonsingularity).

A.3. T has maximum rank.

Assumption A.3 implies that the terminal constraints (4) are independent, which can always be achieved by a proper change of state variables. Also, because of assumption A.2, there is no loss of generality in assuming no mixed term of the form $x^T D(t) u$ in the integrand of the cost functional J since such a problem would be equivalenct to a problem of the above form with $C - DE^{-1}D^T$ instead of C and $A - BE^{-1}D^T$ instead of A [12, p. 156].

The necessary conditions for an admissible pair (u, x) [i.e., constraints (2), (3), (4) are satisfied] to be optimal are that there exist an absolutely continuous function $\lambda(t) \in R^n$, $t \in [t_0, t_f]$, and a constant $\nu \in R^p$ such that,

$$\dot{\lambda}(t) = -C(t)\, x(t) - A^T(t)\, \lambda(t), \qquad t \in [t_0, t_f] \tag{5}$$

$$E(t)\, u(t) + B^T(t)\, \lambda(t) = 0, \qquad t \in [t_0, t_f] \tag{6}$$

$$E(t) \geqslant 0, \qquad t \in [t_0, t_f] \tag{6'}$$

$$\lambda(t_f) = F x(t_f) + T^T \nu \tag{7}$$

Using (6) and assumption A.2 to solve for u, one obtains the following two point boundary value problem (TPBVP):

$$\begin{aligned} \dot{x} &= A(t)\,x - B(t)\,E^{-1}(t)\,B^T(t)\,\lambda \\ \dot{\lambda} &= -C(t)\,x - A^T(t)\,\lambda \end{aligned} \tag{8}$$

$$\begin{aligned} x(t_0) &= x_0 \\ Tx(t_f) + \psi_0 &= 0, \\ \lambda(t_f) &= Fx(t_f) + T^T\nu. \end{aligned} \tag{9}$$

Suppose that we have an extremal pair $(u^*(t), x^*(t))$, $t \in [t_0, t_f]$, with associated multipliers $\lambda^*(t)$, $t \in [t_0, t_f]$, and ν^*, then a conjugate point on the trajectory x^* is defined as follows (see Ref. [1]),

DEFINITION 1. If there exists a solution (x, λ) of Eqs. (8), distinct from (x^*, λ^*), satisfying (9) and $x(t_c) = x^*(t_c)$ for some $t_c < t_f$, then t_c is a conjugate point to t_f on the trajectory x^*.

A numerically oriented method to solve the TPBVP is the "backward sweep method" due to McReynolds and Bryson [11]. This method is based upon the fact that there exists a linear relationship between λ, x, and ν, say,

$$\lambda(t) = S(t)\,x(t) + R(t)\nu, \tag{10}$$

where the matrices S and R are solutions of [12]

$$\begin{aligned} \dot{S} &= -SA(t) - A^T(t)\,S - C(t) + SB(t)\,E^{-1}(t)\,B^T(t)\,S, \\ S(t_f) &= F, \end{aligned} \tag{11}$$

$$\begin{aligned} \dot{R} &= -[A^T(t) - S(t)\,B(t)\,E^{-1}(t)\,B^T(t)]\,R, \\ R(t_f) &= T^T, \end{aligned} \tag{12}$$

and the constant vector ν satisfies

$$R^T(t)\,x(t) + Q(t)\,\nu + \psi_0 = 0 \tag{13}$$

where the matrix Q is a solution of

$$\begin{aligned} \dot{Q} &= R^T(t)\,B(t)\,E^{-1}(t)\,B^T(t)\,R(t), \\ Q(t_f) &= 0. \end{aligned} \tag{14}$$

When $Q^{-1}(t)$ exists, (13) and (10) give

$$\lambda(t) = M(t)\,x(t) - R(t)\,Q^{-1}(t)\psi_0 \tag{15}$$

and the matrix

$$M(t) \triangleq S(t) - R(t) Q^{-1}(t) R^T(t) \tag{16}$$

provides a conjugate point test according to:

LEMMA 1. *$t_c < t_f$ is a conjugate point if and only if the matrix $M(t)$ is infinite*[1] *at t_c.*

A proof of this lemma is given in Ref. [4].

The matrix M is obtained by integrating backward equations (11), (12) and (14). However, it may happen that the matrix S becomes infinite before M does and thus computational difficulties are created. The procedure used to handle this problem is to integrate backward Eq. (11), (12), and (14) until $Q^{-1}(t)$ is defined,[2] say at $\bar{t}$, then $M(\bar{t})$ is defined and since M satisfies the same differential equation as S[13], it may be obtained directly for $t < \bar{t}$ by integrating (11). Note that M cannot be obtained directly from t_f since $M(t_f)$ is not defined. When there is no terminal constraint (4), $R \equiv 0$ and the conjugate point test involves only the matrix S, which is defined at t_f. We shall return to this question in Section 3 and propose a conservative test based on a matrix defined at t_f.

2. Optimality of Extremals when the End Points are Conjugate

Another way to solve the TPBVP is to consider the transition matrix $\Omega(t, t_0)$ of Eq. (8). Let $\Omega(t, t_0)$ be partitioned into 4 $n \times n$ blocks,

$$\Omega(t, t_0) = \begin{bmatrix} \Omega_{11}(t, t_0) & \Omega_{12}(t, t_0) \\ \Omega_{21}(t, t_0) & \Omega_{22}(t, t_0) \end{bmatrix}. \tag{17}$$

It can be shown that (see [17, p. 156])

$$S(t) = [\Omega_{22}(\tau, t) - S(\tau)\, \Omega_{12}(\tau, t)]^{-1} [S(\tau)\, \Omega_{11}(\tau, t) - \Omega_{21}(\tau, t)], \tag{18}$$

$$R(t) = [\Omega_{22}(\tau, t) - S(\tau) \Omega_{12}(\tau, t)]^{-1} R(\tau), \tag{19}$$

$$M(t) = [\Omega_{22}(\tau, t) - M(\tau)\, \Omega_{12}(\tau, t)]^{-1} [M(\tau)\, \Omega_{11}(\tau, t) - \Omega_{21}(\tau, t)], \tag{20}$$

for any $t \leqslant \tau \leqslant t_f$ such that $S(\tau)$, $R(\tau)$ and $M(\tau)$ are defined. It follows that $M(t)$ is infinite if and only if the matrix

$$\overline{M}(t, \tau) = \Omega_{22}(\tau, t) - M(\tau)\, \Omega_{12}(\tau, t) \tag{21}$$

[1] A matrix is infinite if at least one of its elements is infinite.

[2] Q^{-1} is defined at $t_f - \epsilon$, $\epsilon > 0$ arbitrarily small, if the system (2) is completely controllable or if the system (2)–(4) is completely controllable in the reduced sense [4].

is singular for some $\tau \in (t, t_f)$ (it is proved in Ref. [18] that if $\overline{M}(t, \tau)$ is singular for some $\tau \in (t, t_f)$, then it is singular and has the same null space for all $\tau \in (t, t_f)$).

Let t_c be a conjugate point on the extremal trajectory $x^*(t)$, $t \in [t_0, t_f]$, and d be the dimension of the null space of $\overline{M}(t_c, \tau)$, $d \geqslant 1$, we have

THEOREM 1. *There exists a d-parameter family of extremals going through the conjugate point (i.e., $x = x^*(t_c)$ at $t = t_c$).*

Proof. Let e be a unit vector in the null space of $\overline{M}(t_c, \tau)$ and integrate Eqs. (8) forward with boundary conditions

$$\begin{aligned} x(t_c) &= x^*(t_c) \\ \lambda(t_c) &= \lambda^*(t_c) + \alpha e, \end{aligned} \tag{22}$$

where α is a scalar parameter.

Using (17), the solution can be expressed as

$$\begin{aligned} x_\alpha(t) &= \Omega_{11}(t, t_c)\, x^*(t_c) + \Omega_{12}(t, t_c)\, \lambda(t_c) = x^*(t) + \alpha\Omega_{12}(t, t_c)\, e, \\ \lambda_\alpha(t) &= \Omega_{21}(t, t_c)\, x^*(t_c) + \Omega_{22}(t, t_c)\, \lambda(t_c) = \lambda^*(t) + \alpha\Omega_{22}(t, t_c)\, e. \end{aligned} \tag{23}$$

We wish to verify that the necessary conditions for optimality given in Section 1 are satisfied for any α. Obviously, conditions (2), (5), and (6) are satisfied with

$$u_\alpha(t) = -E^{-1}(t)\, B^T(t)\, \lambda_\alpha(t)$$

and it remains to check (4) and (7). We have, for some $\nu_\alpha \in R^p$,

$$\begin{aligned} \lambda_\alpha(t_f) - Fx_\alpha(t_f) - T^T\nu_\alpha = \lambda^*(t_f) &- Fx^*(t_f) - T^T\nu_\alpha \\ &+ \alpha[\Omega_{22}(t_f, t_c) - F\Omega_{12}(t_f, t_c)]e \end{aligned}$$

and, using the multiplier ν^*,

$$\lambda_\alpha(t_f) - Fx_\alpha(t_f) - T^T\nu_\alpha = T^T(\nu^* - v_\alpha) + \alpha[\Omega_{22}(t_f, t_c) - F\Omega_{12}(t_f, t_c)]e. \tag{24}$$

But, from the choice of e

$$\overline{M}(t_c, \tau)e = [\Omega_{22}(\tau, t_c) - M(\tau)\, \Omega_{12}(\tau, t_c)]e = 0 \ \text{ for all } \tau \in (t_c, t_f) \tag{25}$$

then, using (16) and multiplying on the left by $[\Omega_{22}(t_f, \tau) - F\Omega_{12}(t_f, \tau])$ gives,

$$\begin{aligned} [\Omega_{22}(t_f, \tau) &- F\Omega_{12}(t_f, \tau)][\Omega_{22}(\tau, t_c) - S(\tau)\, \Omega_{12}(\tau, t_c)]e \\ &= -[\Omega_{22}(t_f, \tau) - F\Omega_{12}(t_f, \tau)]\, R(\tau)\, Q^{-1}(\tau)\, R^T(\tau)\, \Omega_{12}(\tau, t_c)e \end{aligned}$$

and from (18) and (19) with $\tau = t_f$, $t = \tau$,

$$[\Omega_{22}(t_f, \tau) - F\Omega_{12}(t_f, \tau)]\, \Omega_{22}(\tau, t_c)\, e - [F\Omega_{11}(t_f, \tau) - \Omega_{21}(t_f, \tau)]\, \Omega_{12}(\tau, t_c)e$$
$$= -T^T Q^{-1}(\tau)\, R^T(\tau)\, \Omega_{12}(\tau, t_c)e$$

or,

$$\begin{aligned}\{\Omega_{22}(t_f, \tau)\, \Omega_{22}(\tau, t_c) &+ \Omega_{21}(t_f, \tau)\, \Omega_{12}(\tau, t_c) \\ &- F[\Omega_{12}(t_f, \tau)\, \Omega_{22}(\tau, t_c) + \Omega_{11}(t_f, \tau)\, \Omega_{12}(\tau, t_c)]\}e \\ &= -T^T Q^{-1}(\tau)\, R^T(\tau)\, \Omega_{12}(\tau, t_c)e\end{aligned}$$

It can be shown from (17) and $\Omega(t_f, t_c) = \Omega(t_f, \tau)\, \Omega(\tau, t_c)$ that

$$\Omega_{22}(t_f, \tau)\, \Omega_{22}(\tau, t_c) + \Omega_{21}(t_f, \tau)\, \Omega_{12}(\tau, t_c) = \Omega_{22}(t_f, t_c),$$
$$\Omega_{12}(t_f, \tau)\, \Omega_{22}(\tau, t_c) + \Omega_{11}(t_f, \tau)\, \Omega_{12}(\tau, t_c) = \Omega_{12}(t_f, t_c),$$

and we have

$$[\Omega_{22}(t_f, t_c) - F\Omega_{12}(t_f, t_c)]e = -T^T Q^{-1}(\tau)\, R^T(\tau)\, \Omega_{12}(\tau, t_c)e.$$

Then (24) becomes

$$\lambda_\alpha(t_f) - Fx_\alpha(t_f) - T^T\nu_\alpha = T^T[\nu^* - \nu_\alpha - Q^{-1}(\tau)\, R^T(\tau)\, \Delta x(\tau)] \tag{26}$$

where $\Delta x(\tau) \triangleq \alpha\Omega_{12}(\tau, t_c)e = x_\alpha(\tau) - x^*(\tau)$. Now, we show that $K(\tau) \triangleq Q^{-1}(\tau)\, R^T(\tau)\, \Delta x(\tau)$ is constant. Using (8), (12), (14) and $(d/d\tau)\, Q^{-1}(\tau) = Q^{-1}(\tau)(d/d\tau)\, Q(\tau)\, Q^{-1}(\tau)$. We have,

$$\begin{aligned}(d/d\tau)\, K(\tau) = &-Q^{-1}R^TBE^{-1}B^TRQ^{-1}R^T\Delta x - Q^{-1}R^T(A - BE^{-1}B^TS)\, \Delta x \\ &+ Q^{-1}R^T(A\Delta x - BE^{-1}B^T\Delta\lambda)\end{aligned}$$

where

$$\Delta\lambda(\tau) \triangleq \lambda_\alpha(\tau) - \lambda^*(\tau),$$
$$(d/d\tau)\, K(\tau) = -Q^{-1}R^TBE^{-1}B^T[\Delta\lambda - (S - RQ^{-1}R^T)\, \Delta x] = 0.$$

Since, from (23) and (25), $\Delta\lambda(\tau) = \alpha\Omega_{22}(\tau, t_c)e = \alpha M(\tau)\, \Omega_{12}(\tau, t_c)e = M(\tau)\, \Delta x(\tau)$. $K(\tau)$ being constant, we can choose $\nu_\alpha = \nu^* - K(\tau)$ and it follows that (26) implies (7) for all α.

In order to check (4), multiply $\nu_\alpha - \nu^* = -Q^{-1}(\tau)\, R^T(\tau)\, \Delta x(\tau)$ on the left by $Q(\tau)$ and let τ go to t_f, then (12) and (14) imply

$$\begin{aligned}0 = \lim_{\tau\to t_f} Q(\tau)\, (\nu_\alpha - \nu^*) &= -\lim_{\tau\to t_f} R^T(\tau)\, \Delta x(\tau) = -T(x_\alpha(t_f) - x^*(t_f)) \\ &= -Tx_\alpha(t_f) - \psi_0\,.\end{aligned}$$

Thus the proof that $(x_\alpha(t), \lambda_\alpha(t))$ is extremal for any α is complete. If $d > 1$, one can choose $d - 1$ vectors e_i in the null space of $\overline{M}(t_c, \tau)$ such that the vectors $e, e_1, \ldots, e_{d-1}$ are linearly independent and construct, as above, a one-parameter family of extremals associated with each e_i. ▮

The property given in Theorem 1 is an extension to control problems of the geometrical point of view in the calculus of variations in which the conjugate point is defined as the contact point of an extremal curve with the envelope of a family of extremal curves [16, p. 26]. Indeed, since the problem is quadratic, each member of a family of extremals depends linearly upon the parameter (see Eqs. (23)) and the envelope of the family degenerates into a single point, which is a conjugate point on all the extremals. The analog of the envelope theorem [16, p. 26] in control problems is,

THEOREM 2. *The value of the cost functional* (1) *is the same between* t_c *and* t_f *along each extremal of the family of Theorem* 1.

A proof of this theorem is given in Ref. [1]. Theorem 2 indicates that the extremals of the LQP are still optimal when the initial time t_0 is conjugate to the final time t_f. This property is not necessarily true for nonlinear problems where nonzero third or higher variations may give the cost functional a value smaller than the nominal value, as shown in the example below.

EXAMPLE 1. minimize

$$J = \tfrac{1}{2} \int_{t_0}^{\pi} (u^2 + u^3 - x^2)\, dt,$$

subject to

$$\dot{x} = u; \qquad x(t_0) = 0.$$

It can be shown that $u^*(t) \equiv x^*(t) \equiv 0$ is an extremal solution and that $t_c = \pi/2$ is a conjugate point. When $u =$ constant $= \alpha$, the value of the cost functional with $t_0 = \pi/2$ is $J_\alpha = (\pi/4)\, \alpha^2(1 + \alpha - \pi^2/12)$ and α can be chosen less than $(\pi^2/12) - 1$ so that $J_\alpha < 0 = J^*$.

We shall close this section with the following property well known in the calculus of variations [16, p. 30].

THEOREM 3. *A conjugate point is isolated*; *that is, there exists* $\epsilon > 0$ *such that if* t_1 *is not a conjugate point, there is no conjugate point on* $[t_1 - \epsilon, t_1]$ *and if* t_c *is a conjugate point, there is no conjugate point on* $(t_c, t_c + \epsilon]$.

Proof. This property follows directly from formula (20). Indeed, if t_1 is not a conjugate point, $M(t_1)$ is finite and by choosing $\tau = t_1$ and $t = t_1 - \epsilon$, (21) becomes

$$\overline{M}(t_1 - \epsilon, t_1) = \Omega_{22}(t_1, t_1 - \epsilon) - M(t_1)\,\Omega_{12}(t_1, t_1 - \epsilon)$$

and since $\Omega_{22}(t_1, t_1) = I$ (unit matrix) and $\Omega_{12}(t_1, t_1) = 0$, we can find $\bar{\epsilon}$ small enough such that $\overline{M}(t_1 - \epsilon, t_1)$ is nonsingular for all $\epsilon \leqslant \bar{\epsilon}$. Then it follows that there is no conjugate point on $[t_1 - \bar{\epsilon}, t_1]$. If t_c is a conjugate point, there is no conjugate point on $(t_c, t_c + \epsilon)$ for ϵ small enough since the contrary would imply that the solution (x, λ) of Eq. (8) in Definition 1 is such that $x(t_c) = x^*(t_c)$ and $x(t_c)$ and $x(t_c + \eta) = x^*(t_c + \eta)$ for all $\eta > 0$ less than some $\bar{\epsilon}$ which would imply $(x, \lambda) \equiv (x^*, \lambda^*)$, a contradiction.

3. Penalty Function Approximation to Conjugate Points

An approximate way to handle the terminal constraints (4) is to adjoin them to the cost functional through the use of penalty functions, which are specified so that the minimization of the augmented cost will force the constraints to be satisfied to a desired degree of accuracy. By eliminating terminal constraints, a penalty function approach allows a conjugate point test with the single Riccati equation (11) as opposed to the M matrix test discussed in Section 1. However, the conjugate point obtained is only an approximation to the conjugate point of the original problem and it is of interest to know the worth of this approximation.

Let k be a positive integer and consider the following problem without terminal constraints,

$$(\mathrm{ULQP})_k \begin{cases} \text{minimize} \\ J_k = \tfrac{1}{2}x_f^T F x_f + \tfrac{1}{2}x_f^T T^T P(\lambda_k)\, T x_f + \tfrac{1}{2}\displaystyle\int_{t_0}^{t_f} [x^T C(t)\, x + u^T E(t)\, u]\, dt, \\ \text{subject to} \\ \quad \dot{x} = A(t)\, x + B(t)\, u, \qquad t \in [t_0, t_f], \\ x(t_0) = x_0, \qquad\qquad\quad t_0, t_f \text{ prescribed} \end{cases}$$

where λ_k is a scalar increasing with k ($k = 0, 1, 2,\ldots$) and $P(\lambda_k)$ is a $(p \times p)$ penalty matrix such that,

(i) $P(\lambda_0) = 0$

(ii) $P(\lambda_k) > 0$. (i.e., positive definite) for all $k \geqslant 1$

(iii) $P(\lambda_{k+1}) - P(\lambda_k) > 0$ for all $k \geqslant 0$.

(iv) $\lim_{k \to +\infty} P(\lambda_k) = +\infty$ (i.e., $v^T P(\lambda_k)\, v \to +\infty$ for all $v \in R^p$, $v \neq 0$).

An example is $P(\lambda_k) = k \times I$, where I is the $(p \times p)$ unit matrix. Let $S^k(t)$ be the matrix solution of (11), but with the boundary condition

$$S^k(t_f) = F + T^T P(\lambda_k) T \tag{27}$$

and designate by t_s, t_c and $t_s{}^k$ the first times (going backward) when the matrices $S(t)$, $M(t)$ (associated with LQP)) and $S^k(t)$, respectively, become infinite. (Note that t_s, t_c and/or $t_s{}^k$ may be equal to $-\infty$). Consider the following Riccati differential equation

$$\dot{X} = -XU(t) - U^T(t)X - V(t) + XW(t)X \tag{28}$$

where the matrices X, U, V, W have dimension $(n \times n)$ and V, W are symmetric. Assume that $X_1(t)$ and $X_2(t)$ are two solutions of (28) with boundary condition $X_1(t_1) = L_1$, $X_2(t_1) = L_2$ for some $t_1(L_1$ and L_2 are symmetric $(n \times n)$ matrices, which implies that $X_1(t)$ and $X_2(t)$ are symmetric) and let t_{x_1} and t_{x_2} be the first times when $X_1(t)$ and $X_2(t)$ become infinite when going backward from t_1. We shall need the following property,

LEMMA 2. *If $W(t) > 0$ (positive definite) for all t and $L_1 - L_2 \geqslant 0$, then*

(i) $t_{x_1} \leqslant t_{x_2}$

(ii) $X_1(t) - X_2(t) \geqslant 0$ *for all* $t \in (t_{x_2}, t_1]$.

Part (ii) of Lemma 2 is proved in Ref. [17, Theorem 3, p. 160] where it is also stated, without proof, that $t_{x_2} \leqslant t_{x_1}$, which is incorrect. The correct result (i.e., part (i) of Lemma 2) is proved in Ref. [18] (Appendix B).

THEOREM 4. *For all integers k, $l \geqslant 0$, we have*

(i) $t_c \leqslant t_s^{k+l} \leqslant t_s{}^k \leqslant t_s{}^0 = t_s < t_f$

(ii) *The existence of t_c (i.e., $t_c > -\infty$) implies the existence of $t_s{}^k$, and the existence of $t_s{}^k$ implies the existence of $t_s{}^l$, $l = 0, 1,..., k-1$.*

(iii) *$S^{k+1}(t) - S^k(t)$ is positive semidefinite for all $t \in (t_s{}^k, t_f]$.*

Proof. From the definition of $P(\lambda_k)$ and (27) we have, for all k, $l \geqslant 0$

$$S^{k+l}(t) - S^k(t) = T^T[P(\lambda_{k+l}) - P(\lambda_k)]\, T \geqslant 0$$

and Lemma 2 implies $t_s^{k+l} \leqslant t_s{}^k$ and $S^{k+l}(t) - S^k(t) \geqslant 0$ for all $t \in (t_s{}^k, t_f]$. Also, from (16),

$$\begin{aligned} &M(t_f - \epsilon) - S^k(t_f - \epsilon) \\ &\quad = S(t_f - \epsilon) - S^k(t_f - \epsilon) - R(t_f - \epsilon)\, Q^{-1}(t_f - \epsilon)\, R^T(t_f - \epsilon) \end{aligned}$$

When $\epsilon > 0$ is small enough, $Q^{-1}(t_f - \epsilon)$ exists and is negative definite (follows from (14) and $E(t) > 0$). When ϵ goes to zero, $Q^{-1}(t_f - \epsilon)$ becomes infinite, $R(t_f - \epsilon)$ tends to T^T and $S(t_f - \epsilon) - S^k(t_f - \epsilon)$ tends to $-T^T P(\lambda_k) T$ (finite when k is finite). Thus, ϵ can be chosen small enough so that the matrix $RQ^{-1}R^T |_{t_f - \epsilon} \leqslant 0$ dominates the matrix $S - S^k |_{t_f - \epsilon}$ and it follows that $M(t_f - \epsilon) - S^k(t_f - \epsilon) \geqslant 0$. Also ϵ can be chosen small enough such that t_c, $t_s^k < t_f - \epsilon$, and Lemma 2 implies $t_c \leqslant t_s^k$ for all $k \geqslant 0$. Condition (ii) follows by contradiction since $t_s^k = -\infty$ (i.e., nonexistence of t_s^k) and (i) would imply $t_c = -\infty$, a contradiction with the existence of t_c. ▮

Now assume that $t_0 > t_c$, which insures the existence of a solution $(u^*(t), x^*(t))$, $t \in [t_0, t_f]$, to (LQP) and let k go to infinity.

THEOREM 5. *Suppose $t_0 > t_c$. Then,*

(i) *There exists $K > 0$ such that for each integer $k > K$, problem* $(\text{ULQP})_k$ *has a solution $(u^{*k}(t), x^{*k}(t))$, $t \in [t_0, t_f]$.*

(ii) $\lim_{k \to +\infty} [x^{*k}(t_f)]^T T^T P(\lambda_k) T x^{*k}(t_f) = 0$

(iii) $\lim_{k \to +\infty} J_k^* = J^*$

(iv) $\lim_{k \to +\infty} u^{*k}(t) = u^*(t)$, $t \in [t_0, t_f]$

(v) $\lim_{k \to +\infty} t_s^k = t_c$

where J_k^* and J^* are the values of the cost functionals of problems $(\text{ULQP})_k$ *and* (LQP) *due respectively to (u^{*k}, x^{*k}) and (u^*, x^*).*

The proof of Theorem 4 is lengthy and is therefore given in Appendix A.

After being insured of the convergence of t_s^k, towards t_c, the next point of interest is the rate of convergence. We shall give two simple examples which exhibit a rapid rate of convergence for small values of k.

EXAMPLE 2.

Minimize

$$J = \tfrac{1}{2} \int_{t_0}^{\pi} (u^2 - x^2)\, dt$$

Subject to

$$\dot{x} = u; \qquad x(t_0) = x_0; \qquad x(\pi) = 0.$$

It can be proved that $t_s = \pi/2$, $t_c = 0$, and for $P(\lambda_k) = k$,

$$t_s^k = \tan^{-1}(1/k).$$

The variation of t_s^k when k increases is shown in Fig. 1.

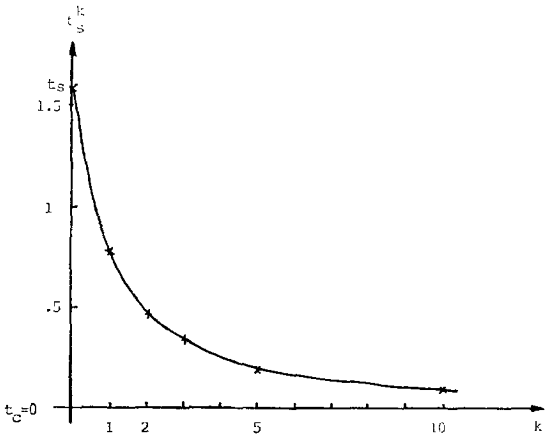

FIG. 1. Convergence of t_s^k toward t_c for Example 2.

EXAMPLE 3.

Minimize

$$J = -\tfrac{3}{4}x_{2f}^2 + \tfrac{1}{2}\int_{t_0}^{1} (u_1^2 + u_2^2)\, dt,$$

Subject to

$$\dot{x}_1 = u_1, \qquad x_1(t_0) = x_{10},$$

$$\dot{x}_2 = u_2, \qquad x_2(t_0) = x_{20}.$$

It can be proved that $t_s = 1/3$, $t_c = -1/3$ and, for

$$P(\lambda_k) = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix},$$

$$t_s^k = \frac{1}{3} + \frac{1}{4k} - \left[\frac{1}{3k} + \left(\frac{2}{3} - \frac{1}{4k}\right)^2\right]^{1/2}.$$

The variation of t_s^k when k increases is shown in Fig. 2.

It can be seen, for the examples above, that the gap between t_s and t_c is almost covered with "reasonable" values of k (i.e., values such that the penalty term in J_k has the same order of magnitude as the term $\frac{1}{2}x_f^TFx_f$). This suggests the following economical, conservative test for conjugate points: instead of going through the procedure explained at the end of Section1, one chooses a "reasonable" value of k (or $P(\lambda_k)$), calculates the $S^k(t)$-matrix, and considers t_s^k as a conservative approximation to the conjugate point t_c.

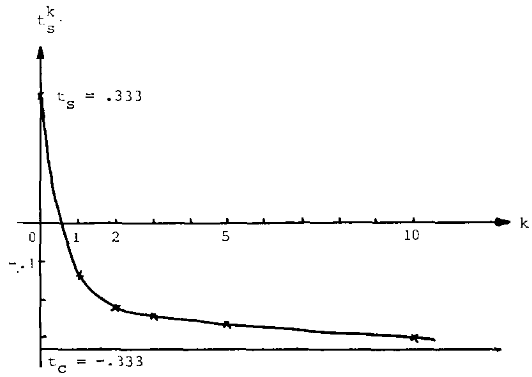

FIG. 2. Convergence of t_s^k toward t_c for Example 3.

We shall close this section by presenting a property of a conjugate point with respect to terminal constraints.

THEOREM 6. *Any addition of constraints on the final value of the state has the effect of making the conjugate point go backward (or remain the same).*

Proof. Let $(\overline{\text{LQP}})$ and $\overline{\text{ULQP}})_k$ be respectively the same problems as (LQP) with extra terminal constraints $Vx_f = 0$, where V is a $(q \times n)$ matrix of maximal rank, $1 \leqslant q \leqslant n - p$, and $(\text{ULQP})_k$ with the extra penalty term $\frac{1}{2}x_f^T V^T P_1(\lambda_k) Vx_f$ in the cost functional, where the $(q \times q)$ penalty matrix $P_1(\lambda_k)$ has the same properties as $P(\lambda_k)$. Designate by $\bar{t}_c$ and $\bar{t}_s^k$ the first backward conjugate points of $(\overline{\text{LQP}})$ and $(\overline{\text{ULQP}})_k$ and define the matrix $\bar{S}^k(t)$ as $S^k(t)$ with the boundary condition

$$\bar{S}^k(t_f) = F + T^T P(\lambda_k)\, T + V^T P_1(\lambda_k) V$$

We have, $\bar{S}^k(t_f) - S^k(t_f) = V^T P_1(\lambda_k) V \geqslant 0$ for all k and Lemma 2 implies $\bar{t}_s^k \leqslant t_s^k$ for all k. Then

$$\lim_{k\to\infty} \bar{t}_s^k \leqslant \lim_{k\to\infty} t_s^k$$

and $\bar{t}_c \leqslant t_c$ follows from Theorem 5(v). ∎

Therefore, the more constrained the final value of the state, the greater the chance for an extremal solution to be locally optimal. This agrees with intuition since by constraining the terminal state, one reduces the class of admissible solutions and thus lessens the chance to find a solution giving a smaller value to the functional. Also, Theorem 6 implies that one should be careful when using a penalty method to handle terminal constraints since, by eliminating all constraints, a penalty method "moves forward" the conjugate point (see also Theorem 4(i) and one may end up rejecting solutions (because of the presence of a conjugate point) that are true local optima.

4. Conclusions

A number of properties of conjugate points for the LQP are developed. In particular, envelope theorems from the classical calculus of variations are generalized to the LQP, and the results are employed to give a characterization of the optimal control on the interval $[t_c, t_f]$. In addition, the behavior of the conjugate point in the LQP with penalty functions for terminal equality constraints is studied. It is found that a conservative approximation of the conjugate point is always obtained, and that the addition of terminal constraints causes the conjugate point to move backward (or remain the same). Such results are useful in the synthesis of neighboring optimal feedback controls since the occurrence of a conjugate point implies the nonexistence of finite feedback gains at t_c.

Appendix A: Proof of Theorem 5

The solution x of Eqs. (2) and (3) can be considered as an implicit functional of u and problems (LQP) and $(\text{ULQP})_k$ can be expressed implicitly in terms of u as follows,

$$(\text{LQP}) \begin{cases} \text{minimize} & J = I[u] \\ \text{subject to} & \Psi[u] = 0 \end{cases}$$

$$(\text{ULQP})_k \left\{ \underset{u}{\text{minimize}}\ J_k = I[u, \lambda_k] = I[u] + \Phi[u, \lambda_k] \right.$$

where $\Phi[u, \lambda_k] = \frac{1}{2}\Psi^T[u]\, P(\lambda_k)\, \Psi[u]$ and u varies in the space of continuous[3] real functions defined on $[t_0, t_f]$. It follows from the structures of (LQP) and $(\text{ULQP})_k$ that the functionals I, Φ, and Ψ are continuous on $[t_0, t_f]$. Define, for any $\epsilon > 0$ and $\alpha \in R^1$,

$$\begin{aligned} U_\epsilon &= \{u : \Psi^T[u]\, \Psi[u] \leqslant \epsilon\}, \\ U(\alpha) &= \{u : I[u] \leqslant \alpha\}, \\ U^k(\alpha) &= \{u : I[u, \lambda_k] \leqslant \alpha\}. \end{aligned}$$

[3] It is obvious from (6) and $E(t) > 0$ that the optimal control, when it exists, is continuous and the search for an optimal u can be restricted to continuous controls.

Note that the set of admissible controls of (LQP) is $U_0 \triangleq U_{\epsilon=0}$ and that $\Phi[u, \lambda_k] \geqslant \Phi[u, \lambda_K] \geqslant 0$ implies $U^k(\alpha) \subset U^K(\alpha) \subset U(\alpha)$ for all α and $k \geqslant K$.

The assumption $t_0 > t_c$ implies that problem (LQP) has a unique optimal solution $u^*(t)$, $t \in [t_0, t_f]$. We shall now make another assumption typical of penalty function type proofs.

Assumption A.1. There exists α_1 and $\epsilon_1 > 0$ such that the set $U(\alpha_1) \cap U_{\epsilon_1}$ is nonempty and compact.

Since $U(\alpha) \subset U(\alpha_1)$ for all $\alpha \leqslant \alpha_1$ and $U_\epsilon \subset U_{\epsilon_1}$ for all $\epsilon \leqslant \epsilon_1$, Assumption A.1 and the fact that the sets are closed imply that the sets $U(\alpha) \cap U_\epsilon$ are compact (possibly empty) for all $\alpha \leqslant \alpha_1$ and $\epsilon < \epsilon_1$.

Proof of (i). Choose $\bar{u} \not\equiv u^*$ in U_0 such that $0 \leqslant I[\bar{u}] - I[u^*] < \eta$, where η is small enough so that $\bar{\alpha} \triangleq I[\bar{u}] < I[u^*] + \eta < \alpha_1$ (the existence of such a $\bar{u}$ is implied by the normality of (LQP)). We wish to show that the set $U^k(\bar{\alpha})$ is compact when k is large enough. $U^k(\bar{\alpha})$ is not empty since it contains at least $\bar{u}$ and u^*. A control $\tilde{u}$ in $U^k(\bar{\alpha})$ is such that

$$I[\tilde{u}, \lambda_k] = I[\tilde{u}] + \Psi^T[\tilde{u}]\, P(\lambda_k)\, \Psi[\tilde{u}] \leqslant \bar{\alpha}. \tag{A.1}$$

Since $P(\lambda_k)$ increases and goes to $+\infty$ as k increases and goes to $+\infty$, (A.1) implies that $\Psi^T[\tilde{u}]\, \Psi[\tilde{u}]$ must be small when k is large and it follows that for any $\epsilon > 0$ given, there exists K_ϵ such that, for all $k \geqslant K_\epsilon$, any control $\tilde{u}$ in $U^k(\bar{\alpha})$ is also in U_ϵ. Therefore, there exists K_{ϵ_1} such that $U^k(\bar{\alpha}) \subset U_{\epsilon_1}$, for all $k \geqslant K_\epsilon$, and the compactness of the sets $U^k(\bar{\alpha}) \subset U(\bar{\alpha}) \cap U_{\epsilon_1}$, $k \geqslant K_{\epsilon_1}$, follows from Assumption A.1. Then, for a given $k \geqslant K_{\epsilon_1}$, the continuous functional $I[u, \lambda_k]$ attains its minimum over $U^k(\bar{\alpha})$, that is, there exists a control $u_k^*(t)$, $t \in [t_0, t_f]$, such that

$$I[u_k^*, \lambda_k] \leqslant I[u, \lambda_k] \qquad \text{for all} \quad u \in U^k(\bar{\alpha}).$$

When $u \notin U^k(\bar{\alpha})$, $I[u, \lambda_k] > \bar{\alpha} = I[\bar{u}] = I[\bar{u}, \lambda_k] \geqslant I[u_k^*, \lambda_k]$ since $\bar{u} \in U^k(\bar{\alpha})$, and it follows that

$$I[u_k^*, \lambda_k] \leqslant I[u, \lambda_k] \qquad \text{for all} \quad u \tag{A.2}$$

which completes the proof of part (i).

Proof of (ii)–(iv). Consider the sequence $\{u_k^*\}$ consisting of the solutions of each $(\mathrm{ULQP})_k$, $k \geqslant k_{\epsilon_1}$. For each k, (A.2) and $\Phi(u, \lambda_{k+1}) \geqslant \Phi(u, \lambda_k)$ imply

$$I[u_k^*, \lambda_k] \leqslant I[u^*, \lambda_k] = I[u^*]$$

and,

$$I[u_k^*, \lambda_k] \leqslant I[u_{k+1}^*, \lambda_k] \leqslant I[u_{k+1}^*, \lambda_{k+1}].$$

The sequence $\{I[u_k^*, \lambda_k]\}$ is thus monotonically increasing, bounded above, and it converges toward a limit $\tilde{I}$ as k goes to infinity.

For each $k \geqslant K_{\epsilon_1}$, $u_k^* \in U^k(\bar{\alpha}) \cap U_{\epsilon_1} \subset U^{K_{\epsilon_1}}(\bar{\alpha}) \cap U_{\epsilon_1}$ and the sequence $\{u_k^*\}$ has a limit u_l as k goes infinity since the set $U^{K_{\epsilon_1}}(\bar{\alpha}) \cap U_{\epsilon_1} \subset U(\bar{\alpha}) \cap U_{\epsilon_1}$ is compact. We have,

$$\lim_{k\to+\infty} \Phi[u_k^*, \lambda_k] = \Psi^T[u_l] \, [\lim_{k\to+\infty} P(\lambda_k)] \, \Psi[u_l].$$

Then, the boundedness of the right-hand term (due to the existence of $\tilde{I} \leqslant I[u^*]$) and $\lim_{k\to+\infty} P(\lambda_k) = +\infty$ imply that $u_l \in U_0$. Now, assume that (ii) does not hold, i.e., $\lim_{k\to+\infty} \Phi[u_k^*, \lambda_k] > 0$, then

$$\lim_{k\to+\infty} I[u_k^*, \lambda_k] = I[u_l] + \lim_{k\to+\infty} \Phi[u_k^*, \lambda_k] \leqslant I[u^*] \tag{A.3}$$

implies that $I[u_l] < I[u^*]$, which contradicts the definition of u^* and $u_l \in U_0$. Thus part (ii) must hold. Then, (A.3) implies $\tilde{I} = I[u_l] \leqslant I[u^*]$ and part (iii) follows from $I[u^*] \leqslant I[u_l]$ (since $u_l \in U_0$).

Now if $u^l \not\equiv u^*$, the uniqueness of the optimal control u^* of (LQP) and $u_l \in U_0$ imply $I[u^l] > I[u^*]$, which is a contradiction with (A.3) and (ii). Thus we have $u^l \equiv u^*$ and part (iv) is proved.

Proof of (v). Since, from Theorem 4, $t_c \leqslant t_s^k$ for all k, (v) will follow if for any given $\epsilon > 0$ small enough, there exists k_ϵ such that $t_s^k - t_c < \epsilon$ for all $k > k_\epsilon$. Note that the arguments used above are valid for any $t_0 > t_c$ and we can choose $t_0 = t_c + \eta$, where $\eta > 0$ can be chosen arbitrarily small. We know from part (i) that each problem $(\text{ULQP})_k$ has a solution $u_k^*(t)$, $t \in [t_0, t_f]$, when $k > k_{\epsilon_1}$, which implies $t_s^k \leqslant t_0$ for all $k > k_{\epsilon_1}$. Therefore given $\epsilon > 0$, we can find K_ϵ such that (choosing $\eta = \epsilon/2$),

$$t_c \leqslant t_s^k \leqslant t_0 = t_c + \epsilon/2 \qquad \text{for all} \quad k > K_\epsilon$$

then

$$t_s^k - t_c \leqslant \epsilon/2 < \epsilon \qquad \text{for all} \quad k > K_\epsilon$$

which completes the proof of Theorem 5. ∎

References

1. J. V. Breakwell and Y. C. Ho, On the conjugate point condition for the control problem, *Int. J. Engineering Science* **2** (1965), 565–579.
2. I. Lee, Optimal trajectory, guidance, and conjugate points, *Information and Control* **8** (1965), 589–606.
3. S. E. Dreyfus, Control problems with linear dynamics, quadratic criterion, and linear terminal constraints, *IEEE Transactions on Automatic Control* **AC-12** (1967), 323–324.

4. W. E. SCHMITENDORF AND S. J. CITRON, On the applicability of the sweep method to optimal control problems, *IEEE Transactions on Automatic Control* **AC-14** (1969), 69–72.
5. H. G. MOYER AND H. J. KELLEY, Conjugate points on extremal rocket paths, *in* "Proceedings, XIXth International Astronautical Federation Congress," Pergamon Press, Oxford, 1970.
6. R. E. KALMAN, Contributions to the theory of optimal control, *Bol. Sociedad Matematica Mexicana* **5** (1960), 102–119.
7. J. V. BREAKWELL, J. L. SPEYER, AND A. E. BRYSON, Optimization and Control of nonlinear systems using the second variation, *SIAM J. on Control* **1** (1963), 193–223.
8. S. K. MITTER, Successive approximation methods for the solution of optimal control problems, *Automatica* **3** (1966), 135–149.
9. M. ATHANS AND P. L. FALB, "Optimal Control," McGraw–Hill, New York, 1966.
10. A. Y. VASILEV, Necessary and sufficient conditions for the optimality of controlled systems, *Automatica and Remote Control* **25** (1965), 1262–1270.
11. S. R. MCREYNOLDS AND A. E. BRYSON, A successive sweep method for solving optimal programming problems, *in* "Proceedings, 1965 JACC," Troy, New York, 1965, pp. 551–555.
12. A. E. BRYSON AND Y. C. HO, "Applied Optimal Control," Blaisdell, Waltham, Massachusetts, 1969.
13. S. J. CITRON, "Elements of Optimal Control," Holt, Rinehart and Winston, New York, 1969.
14. H. G. MOYER, Optimal control problems that test for envelope contacts, *J. Optimization Theory. Appl.* **6** (1970), 287–298.
15. L. J. WOOD AND A. E. BRYSON, Second-order optimality conditions for variable end time terminal control problems, *AIAA Journal* **11** (1973), 1241–1246.
16. G. A. BLISS, "Lectures on the Calculus of Variations," University of Chicago Press, Chicago, Ill., 1946.
17. R. W. BROCKETT, "Finite Dimensional Linear Systems," Wiley, New York, 1970.
18. P. MEREAU, "Global Optimality Conditions and the Darboux Point," The University of Michigan, Ph.D. Thesis, 1974.