Dedicated to
Professor Czesław Olech

Control and Cybernetics
vol. **34** (2005) No. 3

Second order conditions for periodic optimal control problems

by

John Allwright and Richard Vinter

Department of Electrical and Electronic Engineering, Imperial College
London SW7 2BT, United Kingdom
e-mail: j.allwright@imperial.ac.uk, r.vinter@imperial.ac.uk

Abstract: This paper concerns second order sufficient conditions of optimality, involving the Riccati equation, for optimal control problems with periodic boundary conditions. The problems considered involve no pathwise constraints and are 'regular', in the sense that the strengthened Legendre-Clebsch condition is assumed to be satisfied. A well-known sufficient condition, which we refer to as the Riccati sufficient condition, requires the existence of a global solution to the Riccati equation whose endpoint values satisfy a certain inequality. A sharper condition, named the extended sufficient condition, takes the form of an inequality involving the solutions of a Riccati equation and two additional linear matrix equations. We highlight the superiority of the extended Riccati sufficient condition and develop a number of equivalent formulations of this condition. Not only does the extended Riccati sufficient condition supply more information about minimizers, but it is the basis of simpler numerical tests for assessing whether an extremal is a minimizer, at least in a local sense. The Riccati and also the extended Riccati sufficient conditions are applied to a variant of Speyer's 'sailboat' problem, involving parameters. It is found that the extended Riccati sufficient condition identifies a much larger set of points on parameter space for which a nominal control is optimal, in comparison to the Riccati sufficient condition.

Keywords: second order conditions, periodic optimal control, Riccati equations, dynamic programming.

1. Introduction

Consider the following optimal control problem with periodic boundary conditions

$$
(P)\begin{cases}
\text{Minimize } \int_0^T L(t,x(t),u(t))dt + g(x(T)) \\
\text{over } x \in W^{1,1}([0,T];\mathbb{R}^n) \text{ and measurable functions } u:[0,T]\to\mathbb{R}^m \\
\text{satisfying} \\
\dot{x}(t) = f(t,x(t),u(t)) \quad \text{a.e. } t\in[0,T] \\
x(0) = x(T).
\end{cases}
$$

The data for this problem comprise a number $T > 0$ and the functions $L : [0,T]\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}$, $f : [0,T]\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}$. Here, $W^{1,1}([0,T];\mathbb{R}^n)$ denotes the space of absolutely continuous $\mathbb{R}^n$-valued functions on $[0,T]$.

A *control function* (for (P)) is a measurable $\mathbb{R}^m$-valued function on $[0,T]$. Given a control function $u(.)$, a *state trajectory* $x(.)$ corresponding to $u(.)$ is a solution to the differential equation $\dot{x}(t) = f(t,x(t),u(t))$ on $[0,T]$, in the space $W^{1,1}([0,T];\mathbb{R}^n)$. A couple (x,u) comprising a control function $u(.)$ and a state trajectory $x(.)$ corresponding to $u(.)$ is referred to as a *process* (for (P)). If $x(.)$ satisfies the periodic boundary conditions, the process is called *admissible*. An *admissible process* for (P) which achieves the infimum cost over all admissible processes for (P) is referred to as a *minimizer* for (P). We shall also make use of a different, narrower, concept of optimality: an admissible process $(\bar{x}(.),\bar{u}(.))$ is said to be a weak local minimizer for (P) if there exists $\epsilon > 0$ such that

$$J(x(.),u(.)) \;\geq\; J(\bar{x}(.),\bar{u}(.))$$

for all admissible processes $(x(.),u(.))$ satisfying

$$\|x(.)-\bar{x}(.)\|_{L^\infty} \leq \epsilon \;\text{ and }\; \|u(.)-\bar{u}(.)\|_{L^\infty} \leq \epsilon.$$

Let us recall the concept of 'weak normal extremal'. This involves the Hamiltonian function

$$H(t,x,p,u) \;:=\; p^T f(t,x,u) + L(t,x,u)\,.$$

An admissible process $(\bar{x}(.),\bar{u}(.))$ is said to be a *weak normal extremal* for (P) if $(\bar{x}(.),\bar{u}(.))$ satisfies the conditions of the Pontryagin Maximum Principle, in the following special form: there exists $p(.) \in W^{1,1}([0,T];\mathbb{R}^n)$ such that,

$$
\begin{aligned}
&-\dot{p} \;=\; f_x^T(t,\bar{x}(t),\bar{u}(t))p(t) + L_x(t,\bar{x}(t),\bar{u}(t)) \\
&H_u(t,\bar{x}(t),p(t),u)|_{u=\bar{u}(t)} \;=\; 0 \quad \text{a.e. } t\in[0,T] \\
&p(T) \;=\; p(0) + g_x(\bar{x}(T))\,.
\end{aligned} \tag{1}
$$

The qualifiers 'weak' and 'normal' refer to the facts that the usual Weierstrass-type condition in the definition of the extremal is replaced by the weaker 'vanishing gradient' condition (1) and that the cost multiplier is taken to have value 1, respectively. (Notice also that the boundary condition on the costate arc $p(.)$ is the usual transversality condition specialized to the case when the endpoint constraint on state trajectories is '$x(0) = x(T)$'.) The above special form of the Pontryagin Maximum Principle is elsewhere sometimes referred to as the Euler Lagrange condition.

As is well known, the condition that an admissible process $(\bar{x}(.), \bar{u}(.))$ is a weak normal extremal (or, synonymously, satisfies the special case of the Pontryagin Maximum Principle) is a necessary condition that $(\bar{x}(.), \bar{u}(.))$ is a minimizer under unrestrictive hypotheses. However, the condition is not in general sufficient. The primary role of higher order optimality conditions in Optimal Control (and second order conditions in particular) is to provide additional tests to confirm that a given weak normal extremal is actually a weak, local minimizer. A secondary role is to provide refined necessary conditions.

The theory of second order conditions is well developed in the case of optimal control problems (such as those studied here) which do not involve pathwise control or state constraints. This centres on an analysis of the 'accessory problem', associated with a weak normal extermal $(\bar{x}(.), \bar{u}(.))$.

For each $t \in [0, T]$, define:

$$(A(t), B(t)) \ := \ (f_x(t, \bar{x}(t), \bar{u}(t)), f_u(t, \bar{x}(t), \bar{u}(t))) \tag{2}$$

and

$$\begin{pmatrix} Q(t) & D(t) \\ D^T(t) & R(t) \end{pmatrix} := \begin{pmatrix} H_{xx}(t, \bar{x}(t), p(t), \bar{u}(t)) & H_{xu}(t, \bar{x}(t), p(t), \bar{u}(t)) \\ H_{xu}^T(t, \bar{x}(t), p(t), \bar{u}(t)) & H_{uu}(t, \bar{x}(t), p(t), \bar{u}(t)) \end{pmatrix}.$$

Define also

$$J_A(y(.), v(.)) \ := \ \frac{1}{2} \int_0^T L_A(t, y(t), v(t)) dt + \frac{1}{2} y^T(T) G y(T)$$

where

$$G \ := \ g_{xx}(\bar{x}(T)) \quad \text{and} \quad L_A(t, y, v) \ = \ y^T Q(t) y + 2 y^T D(t) v + v^T R(t) v.$$

The accessory problem is

$$(A) \begin{cases} \text{Minimize } J_A(y(.), v(.)) \\ \text{over } y \in W^{1,1}([0, T]; \mathbb{R}^n) \text{ and } v \in L^2([0, T]; \mathbb{R}^m) \text{ satisfying} \\ \dot{y} = A(t) y(t) + B(t) v(t) \quad \text{a.e.} \\ y(0) = y(T). \end{cases}$$

Couples $(y(.), v(.))$ satisfying the dynamic and endpoint constraints of the accessory problem will be referred to as admissible processes for the accessory problem.

We focus on second order sufficient conditions for optimality, expressed in terms of the Riccati equation for the accessory problem:

$$\begin{cases} \dot{P} + PA + A^T P + Q - (B^T P + D^T)^T R^{-1}(B^T P + D^T) = 0 \\ P^T(.) \ = \ P(.) \end{cases} \tag{3}$$

and refinements. These conditions can be regarded as alternatives to conditions requiring that the cost of the accessory problem (A) be coercive (see Section 4 below) on the set of admissible processes (A). Sufficient conditions involving Riccati type equations have an important role in numerical optimal control. This is because they lead to simple numerical tests of whether a process, arrived at by applying some optimal control algorithm, is at least a weak local minimizer (Maurer and Pickenhain, 1995). The coercivity property, on the other hand, is not so susceptible to direct numerical verification.

A standard sufficient condition for a process $(\bar{x}(.), \bar{u}(.))$ to be a weak local minimizer (we shall refer to it as the Riccati sufficient condition) is that there exists a solution $P(.)$ to (3) satisfying

$$G + P(0) - P(T) > 0\,.$$

This condition is obtained, for example, by specializing to the periodic boundary conditions case Riccati-type sufficient conditions for problems with mixed boundary conditions given in Maurer and Pickenhain (1995) or Zeidan (2001).

An improved sufficient condition, which we call the extended Riccati sufficient condition, is reported in Stefani and Zezza (1997), as part of a broad investigation by Stefani and Zezza into optimality conditions for linear quadratic control problems with general endpoint and mixed state/control pathwise constraints. This involves three matrix differential equations, in place of the single equation of the Riccati sufficient condition. Stefani and Zezza describe the nature of the condition (it takes the form of a 'Riccati differential equation, a linear equation and an integrator') and briefly outline a proof technique, based on reduction to the separated endpoint constraints case via state augmentation.

The purpose of this paper is to highlight the superiority of the extended Riccati sufficient condition. In the interest of a self-contained treatment, we include a proof of the condition. We provide also alternative versions, expressed in terms of the transition matrix of the Hamiltonian system for the accessory problem. We further clarify the relationship between the Riccati sufficient condition and the augmented Riccati sufficient condition, by examining the special circumstances under which they coincide. Finally we illustrate the benefits of the extended Riccati sufficient condition by applying it to a variant of a periodic optimal control problem previously studied by Speyer, the 'sailboat problem'. In this problem, the extended Riccati sufficient condition can be used to identify a very much larger region in parameter space, for which a nominal process is a weak local minimizer, than the earlier Riccati sufficient condition.

We make use of two different notions of controllability for the dynamics of the accessory problems (summarized by the matrix valued functions $(A(.), B(.))$ on $[0,T]$:

> $(A(.), B(.))$ is said to be *controllable on* $[0,T]$ if, corresponding to any $\xi \in R^n$, there exists a process $(x(.), u(.)$ for (A) such that $x(0) = 0$ and $x(T) = \xi$.
>
> $(A(.), B(.))$ is said to be *controllable on* $[0,T]$ *with respect to the periodic boundary conditions* if, corresponding to any $\xi \in R^n$, there exists a process $(x(.), u(.))$ for (A) such that $x(0) = x(T) + \xi$.

Clearly, the latter condition is weaker.

Throughout, the Euclidean norm is denoted by $|.|$. For any (possibly non-symmetric) $n \times n$ matrix F, '$F > 0$' is taken to mean '$z^T F z > 0$ for all $z \neq 0$'. Likewise, '$z^T F z \geq 0$' means '$z^T F z \geq 0$ for all z'.

2. Riccati-type conditions for weak local optimality

Take $(\bar{x}(.), \bar{u}(.))$ to be the weak normal extremal of interest. We shall invoke the following hypotheses, in which $\widetilde{f}(.,.,.)$ denotes the function

$$\widetilde{f}(t,x,u) \;=\; (f(t,x,u), L(t,x,u)).$$

(H1) $\bar{u}(.)$ is essentially bounded.

(H2) $\widetilde{f}(.,x,u)$ is Lebesgue measurable for each (x,u), and $\widetilde{f} = (t,.,.)$ is of class C^2 for each $t \in [0,T]$; g is of class C^2.

(H3) $\widetilde{f}$, $\widetilde{f}_x$, $\widetilde{f}_u$, $\widetilde{f}_{xu}$, $\widetilde{f}_{xx}$ and $\widetilde{f}_{uu}$ are bounded on bounded sets.

(H4) there exists a function $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\lim_{\alpha\downarrow 0} \theta(\alpha)/\alpha = 0$ and

$$\left| \widetilde{f}(t,z) - \widetilde{f}(t,\bar{z}(t)) - \nabla\widetilde{f}(t,\bar{z}(t))(z-\bar{z}(t)) - \frac{1}{2}(z-\bar{z}(t))^T \nabla^2 \widetilde{f}(t,\bar{z}(t))(z-\bar{z}(t)) \right|$$
$$\leq \theta\left(|(z-\bar{z}(t))|^2\right)$$

for all $t \in [0,T]$ and $z \in \mathbb{R}^n \times \mathbb{R}^m$, where $\bar{z}(t) = (\bar{x}(t), \bar{x}(t))$ and $\nabla\widetilde{f}$ and $\nabla^2\widetilde{f}$ denote the gradient and Hessian of $\widetilde{f}(t)$ respectively in the (x,u) variables.

(H5) *(Strengthened Legendre-Clebsch condition.)* There exists $\epsilon > 0$ such that

$$R(t) > \epsilon I \qquad \text{for all } t \in [0,T]\ .$$

The following theorem makes reference to the the matrix equations:

$$\begin{cases} \dot{P} + PA + A^T P + Q - (B^T P + D^T)^T R^{-1}(B^T P + D^T) = 0 \\ P(.) = P^T(.) \end{cases} \tag{4}$$

$$\dot{S} + (A^T - (B^T P + D^T)^T R^{-1} B^T) S = 0 \tag{5}$$

$$\dot{M} - S^T B R^{-1} B^T S = 0. \tag{6}$$

(Here, we have written $A(t)$ as A, etc.)

THEOREM 2.1 (The Extended Riccati Sufficient Condition) *Let* $(\bar{x}(.), \bar{u}(.))$ *be a weak normal extremal for (P). Assume (H1)-(H5) and that* $(A(.), B(.))$ *is controllable on* $[0, T]$ *with respect to the periodic boundary conditions.*

Suppose that there exist solutions $(P(.), S(.), M(.))$ *to eqns. (4)–(6) on* $[0, T]$ *such that*

$$G + P(0) + 2S(0) + M(0) - (P(T) + 2S(T) + M(T)) \; > \; 0 \, . \tag{7}$$

Then $(\bar{x}(.), \bar{u}(.))$ *is a weak local minimizer.*

Recall that the matrix inequality (7), involving a possibly non-symmetric matrix is interpreted in an 'inner product' sense (see Section 1). The theorem is proved in Section 6.

Comments

1. In the above sufficient condition of optimality, $P(T)$ is chosen to be a symmetric matrix, but we do not require $S(T)$ or $M(T)$ to be symmetric matrices; in fact, allowing these matrices to be non-symmetric broadens the scope of the conditions.
2. The sufficient condition referred to earlier as the *Riccati sufficient condition*, namely that there exists a a solution $P(.)$ to (4) on $[0, T]$, satisfying $G + P(0) - P(T) \; > \; 0$, can be interpreted as a special (and therefore more restrictive) case of the sufficient condition of Theorem 2.1, in which we limit our selection of $(P(T), S(T), M(T))$ to require $S(T) = 0$ and $M(T) = 0$. (Notice that $S(T) = M(T) = 0$, then $S(.) = M(.) \equiv 0$.)

It is helpful to our understanding of the above optimality conditions at this stage to introduce some fresh notation. Define $\mathcal{D}$ to be the subspace of $\mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}$ comprising triples $(\bar{P}, \bar{S}, \bar{M})$ such that $\bar{P} = \bar{P}^T$ and there exists a solution $(P(.), S(.), M(.))$ to (4)-(6) on $[0, T]$ satisfying

$$(P(T), S(T), M(T)) \; = \; (\bar{P}, \bar{S}, \bar{M}) \, .$$

Define $V : \mathcal{D} \to \mathbb{R}$:

$$V(\bar{P}, \bar{S}, \bar{M}) = \min_{\{z \,|\, |z|=1\}} \left\{ z^T [G + P(0) + 2S(0) + M(0) - P(T) - 2S(T) - M(T)] z \right\} .$$

The extended Riccati sufficient condition can now be expressed

$$\sup \left\{ V(\bar{P}, \bar{S}, \bar{M}) \mid (\bar{P}, \bar{S}, \bar{M}) \in \mathcal{D} \right\} > 0 \, .$$

This improves on the Riccati sufficient condition

$$\sup \left\{ V(\bar{P}, \bar{S}, \bar{M}) \mid (\bar{P}, \bar{S}, \bar{M}) \in \widetilde{\mathcal{D}} \right\} \; > \; 0 \; ,$$

in which $\widetilde{\mathcal{D}}$ is the subset

$$\widetilde{\mathcal{D}} \; = \; \{ (\bar{P}, \bar{S}, \bar{M}) \in \mathcal{D} \mid \bar{S} = 0 \text{ and } \bar{M} = 0 \}.$$

3. Alternative formulations of the sufficient conditions

The test for weak local optimality implicit in Theorem 2.1, which involves a search over terminal values of solutions to the matrix differential equations (4)-(6) that satisfy the condition (7) is not straightforward to apply. The question arises then whether a set of solutions (P^*, S^*, M^*) confirming weak local optimality can more easily be identified.

In this section, we provide answers under additional hypotheses. This requires consideration of Hamilton's system of equations associated with the accessory problem:

$$\left\{ \begin{array}{l} \dot{y} = Ay - BR^{-1}(B^T p + D^T y) \\ -\dot{p} = A^T p + Qy - DR^{-1}(B^T p + D^T y) \end{array} \right. .$$

Let $\{\Psi(t,s) \,|\, 0 \le s, t \le T\}$ be the transition matrix associated with Hamilton's system of equations (H), i.e. for each $s \in [0,T]$, $\Psi(.,s)$ satisfies:

$$\left\{ \begin{array}{l} \frac{d}{dt}\Psi(t,s) \; = \; \mathcal{A}\Psi(t,s) \quad \text{a.e. } t \in [0,T] \\ \Psi(s,s) = I \end{array} \right. \tag{8}$$

where

$$\mathcal{A} = \begin{pmatrix} A - BR^{-1}D^T & -BR^{-1}B^T \\ -Q + DR^{-1}D^T & -A^T + DR^{-1}B^T \end{pmatrix}.$$

Partition the matrix $\Phi(T) := \Psi(0,T)$ into $n \times n$ matrices, thus

$$\Phi(T) \; = \; \begin{bmatrix} \Phi_{11}(T) & \Phi_{12}(T) \\ \Phi_{21}(T) & \Phi_{22}(T) \end{bmatrix}.$$

We shall impose the following hypotheses directly on the data for the accessory problem:

(H6) $A(.), Q(.) \in L^1([0,T]; R^{n\times n})$, $B(.) \in L^2([0,T]; R^{n\times m})$, $D(.) \in L^2([0,T]; R^{m\times n})$, $R(.) \in L^\infty([0,T]; R^{m\times m})$ and there exists $\epsilon > 0$ such that $R(t) > \epsilon I$ for all $t \in [0,T]$.

Theorem 3.1 *Assume (H6). Assume, furthermore,*

(a) $det[\Phi_{12}(T)] \neq 0$.

(b) There exists a solution to (4) on $[0,T]$.

Under these hypotheses there exists a solution $P(.)$ to (4) on $[0,T]$ satisfying the condition

$$det\;[\Phi_{11}(T) - \Phi_{12}(T)P(T)] \;\neq\; 0\,.$$

Define

$$(P^*, S^*, M^*) \;=\; (P(T),\, \Phi_{12}^{-1}(T)[I - \Phi_{11}(T) - \Phi_{12}(T)P(T)],\, 0)\,.$$

Then

$$sup\{V(\bar{P},\bar{S},\bar{M}) : (\bar{P},\bar{S},\bar{M}) \in \mathcal{D}\} \;=\; V(P^*,S^*,M^*)\,. \tag{9}$$

(i.e. the supremum value of $V(.)$ *is achieved at* (P^*,S^*,M^*)*). Furthermore,*

$$sup\{V(\bar{P},\bar{S},\bar{M}) : (\bar{P},\bar{S},\bar{M}) \in \mathcal{D}\} \;=\; inf_{\{z\,|\,|z|=1\}} z^T W z$$

where

$$W \;:=\; G + \Phi_{21}(T) + (\Phi_{22}(T) - I)\Phi_{12}^{-1}(T)(I - \Phi_{11}(T))\,. \tag{10}$$

A proof of Theorem 3.1 is given in Section 6. Note that there is an implicit assumption in this theorem that '$(A(.),B(.))$ is controllable on $[0,T]$'; this property follows, in fact, from condition (a). (See Lemma 6.4.)

Combining the assertions of Theorems 2.1 and 3.1, and noting (Lemma 6.4 below) that if the Riccati equation (4) has a solution on $[0,T]$, then it has a solution $P(.)$ on $[0,T]$ with boundary condition $P(T) = \alpha I$, for all α sufficiently large, we arrive at the following 'sufficiency' test for a weak local extremal to be weak local minimizer (under the hypotheses (H1)-(H6)):

(I) $det\,[\Phi_{12}(T)] \neq 0$

(II) the Riccati equation (4) with boundary condition $P(T) = \alpha I$ has a solution on $[0,T]$, for some $\alpha > 0$

(III) $W \;>\; 0$

where, as before, $\Phi_{12}(T)$, is the upper right hand block component of the partitioned matrix $\Phi_{12}(T)$ obtained from the transition matrix for the Hamiltonian system, and W is the matrix defined in (10).

An important point here is that the three conditions of this test are suitable for computations. In particular, condition (II) involves computing the maximal solution $P(.)$, backwards in time, with right endpoint condition$P(T) = \alpha I$, for increasing values of the parameter α; the condition is satisfied if, eventually, the domain includes $[0,T]$.

Sufficient conditions expressed in terms of the matrix W are a special case of sufficient conditions derived earlier by Speyer (see Speyer, 1996, and Wang and Speyer, 1990) for free time periodic optimal control problems.

Finally we make comments on circumstances when the Riccati and the extended Riccati conditions coincide or, in the terminology of Section 2, when

$$sup\left\{V(\bar{P},\bar{S},\bar{M}) \mid (\bar{P},\bar{S},\bar{M}) \in \mathcal{D}\right\} \;=\; sup\left\{V(\bar{P},\bar{S},\bar{M}) \mid (\bar{P},\bar{S},\bar{M}) \in \widetilde{\mathcal{D}}\right\}\,.$$

It is clear that, under the hypotheses of Theorem 3.1, this will be the case when the Riccati equation has a symmetric solution on $[0,T]$ with right boundary condition

$$P(T) \;=\; \Phi_{12}^{-1}(T)(I - \Phi_{11}(T))\,. \tag{11}$$

In this case, (P^*, S^*, M^*) given by (9) is simply $(P^*, S^*, M^*) = (P(T), 0, 0)$ and a 'maximizing' choice of $(P(.), S(.), M(.))$ for the extended Riccati sufficient condition can be made in which the matrix functions $S(.)$ and $M(.)$ vanish. Notice the choice (11) is only possible in the exceptional situation in which $\Phi_{12}^{-1}(T)(I - \Phi_{11}(T))$ is a symmetric matrix.

4. Coercivity of the second variation and the extended Riccati-type optimality conditions

Insights into the strength of sufficient conditions of optimality are gained by examining their relationship with necessary conditions. The following propositions, whose proofs appear in Section 6, are of interest in this regard.

Proposition 4.1 *Let $(\bar{x}(.), \bar{u}(.))$ be a minimizer for (P). Assume (H1)-(H5) and that $(A(.), B(.))$ is controllable on $[0, T]$ with respect to the periodic boundary conditions. Then*

(i) $(\bar{x}(.), \bar{u}(.))$ is a weak normal extremal, and

(ii)

$$J_A(y, v) \;\geq\; 0 \tag{12}$$

for all admissible processes (y, v) for the accessory problem.

Proposition 4.2 *Assume (H6) and that $(A(.), B(.))$ is controllable on $[0, T]$. Then the following conditions are equivalent:*

(a) the following three conditions are satisfied

(I) $det[\Phi_{12}(T)] \neq 0$

(II) the Riccati equation (4) has a solution on $[0, T]$

(III) $W \;>\; 0$.

(b) there exists a number $\gamma > 0$ such that

$$J_A(y, v) \;\geq\; \gamma \left(|y(0)|^2 + \int_0^T |v(t)|^2 dt \right) \tag{13}$$

for all admissible processes (y, v) for the accessory problem.

We remark that, under hypothesis (H6), condition (a)(I) implies '$(A(.), B(.)$ is controllable on $[0, T]$', so the *a priori* controllability hypothesis in the last proposition is required only to establish '(b) implies (a)'.

Conditions (12) and (13) are customarily referred to as 'positivity' and 'coercivity' of the second variation, respectively. The substance of the preceding propositions can be summarized as follows. Under the state hypotheses, coercivity of the second variation is equivalent to the satisfaction of the three conditions that constitute the 'sufficiency' test of the preceding section. Coercivity of the second variation itself can be regarded as a more stringent version of the 'positivity' of the second variation condition, which is a necessary condition for a weak normal extremal to be a minimizer.

The above relationships are by no means comprehensive. In the literature more general relationships linking optimality of an extremal, in some sense, coercivity/positivity of the second corresponding variation and existence of a solution to the Riccati equation, have been established, in the absence of controllability hypotheses, for a variety of possible endpoint conditions. (See Stefani and Zezza, 1997, and references therein.)

5. Speyer's sailboat problem

In this section we compare the Riccati sufficient and extended Riccati sufficient conditions, by applying them to the following example

$$\left\{\begin{array}{l} \text{Minimize } \int_0^1 (x_1^2(t) - ax_2^2(t) + dx_2^4(t) + bu^2(t))dt \\ \text{over } x \in AC^1([0,1];R^2) \text{ and } u : [0,T] \to R \text{ such that} \\ (\dot{x}_1(t), \dot{x}_2(t)) \;=\; (x_2(t), u(t)) \quad \text{a.e.} \\ (x_1(0), x_2(0)) \;=\; (x_1(T), x_2(T))\;, \end{array}\right.$$

in which a, b and d are positive parameters; $(\bar{x}(.) \equiv 0, (\bar{u}(.) \equiv 0)$ is a weak normal extremal. The corresponding accessory problem is

$$\left\{\begin{array}{l} \text{Minimize } J_A(y(.), v(.)) \\ \text{over } y \in W^{1,1}([0,T];\mathbb{R}^n) \text{ and } v \in L^2([0,T];\mathbb{R}^m) \text{ such that} \\ \dot{y} = A(t)y(t) + B(t)v(t) \quad \text{a.e.} \\ y(0) = y(T). \end{array}\right.$$

in which

$$A \;=\; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\;\; B \;=\; \begin{pmatrix} 0 \\ 1 \end{pmatrix},\;\; Q \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -a \end{pmatrix},$$
$$D \;=\; 0\;,\;\; R \;=\; b \text{ and } G \;=\; 0\;.$$

This is a fixed time version of a problem earlier studied by Speyer (1996), as part of an investigation of a phenomenon encountered in aircraft control, batch processing and other areas, where periodic operation has the potential to reduce the cost over best steady state operation. We refer to Speyer (1996) for an interpretation of the above problem, in terms of a tacking strategy for a sailboat to maximize average speed into the wind.

It is of interest to determine the values of the parameters a and b such that $(\bar{y}(.) \equiv 0, \bar{v}(.) \equiv 0)$ is a minimizer for the accessory problem because, in this case, $(\bar{x}(.) \equiv 0, \bar{u}(.) \equiv 0)$ is a weak local minimizer for the original problem, a situation in which we can expect steady state operation to be optimal.

Both the Riccati sufficient condition and the extended Riccati sufficient condition can be used to explore the region in (a, b) space such that $(\bar{y}(.) \equiv 0, \bar{v}(.) \equiv 0)$ is a minimizer for the accessory problem. To apply the Riccati sufficient condition, the positive orthant of (a, b) space was discretized, for each discrete value of (a, b) and an 'unconstrained optimization without derivatives' algorithm from the MATLAB Optimization Toolbox algorithm was used to test the condition:

$$sup \left\{ V(\bar{P}, \bar{S}, \bar{M}) \mid (\bar{P}, \bar{S}, \bar{M}) \in \widetilde{\mathcal{D}} \right\} > 0 \,,$$

in which $\widetilde{\mathcal{D}}$ is the subset

$$\widetilde{\mathcal{D}} \;=\; \{(\bar{P}, \bar{S}, \bar{M}) \in \mathcal{D} \mid \bar{S} = 0 \text{ and } \bar{M} = 0\}.$$

(According to the Riccati sufficiency test, a positive supremum value indicates $(\bar{y}(.) \equiv 0, \bar{v}(.) \equiv 0)$ is a minimizer for the accessory problem.)

Applying the extended Riccati sufficient condition, on the other hand, involved assessing the singularity of the Φ_{12} block of the the transition matrix of the Hamiltonian system, carrying out a simple line search for a solution to the Riccati equation on $[0, T]$ and testing the inequality '$W > 0$'. (See Section 3.)

The region in (a, b) space above the upper line in Fig. 1 corresponds to (a, b) values for which the Riccati sufficient condition guarantees that $(\bar{y}(.) \equiv 0, \bar{v}(.) \equiv 0)$ is a minimizer for the accessory problem. We see that the extended Riccati sufficient condition provides significantly more information, regarding situations in which $(\bar{y}(.) \equiv 0, \bar{v}(.) \equiv 0)$ is a minimizer for the accessory problem, than does the Riccati sufficient condition. It is noted also that the computational burden was considerably reduced by the use of the extended Riccati sufficient condition in place of the Riccati sufficient condition.

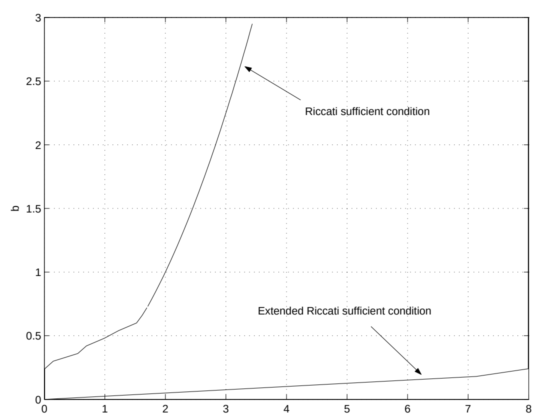

Figure 1. Regions for which the Riccati and the extended Riccati sufficient conditions predict that (y(.)=0, v(.)=0) is a minimizer for the accessory problem

The above example shows up the deficiencies of sufficient conditions, expressed simply in terms of a Riccati equation, for problems with mixed boundary conditions. An example in Stefani and Zezza (1997) illustrates the deficiencies of such conditions, for problems with separated boundary conditions, but with a cost function involving both endpoints.

6. Preliminary analysis

Throughout this section, $(\bar{x}(.), \bar{u}(.))$ is a given weak normal extremal. The following lemma summarizes well-known estimates linking problem (P) and the accessory problem. Since these estimates have a key role in the derivation of second order optimality conditions for (P), and in the interests of a self-contained treatment, a proof is given in the appendix (see also Zeidan, 2001).

LEMMA 6.1 *Assume (H1)–(H5) and that $(A(.), B(.))$ is controllable on $[0,T]$ with respect to the periodic boundary conditions. Then*

(i) Given any $\delta > 0$, there exists $\epsilon > 0$ with the following properties: given any admissible process $(x(.), u(.))$ for (P) such that

$$\|x - \bar{x}\|_{L^\infty} \leq \epsilon, \quad \|u - \bar{u}\|_{L^\infty} \leq \epsilon,$$

there exists an admissible process (y, v) for the accessory problem such that

$$J(x,u) - J(\bar{x}, \bar{u}) \geq J_A(y,v) - \delta \left[\int_0^T |v|^2 dt + |y(0)|^2 \right] .$$

(ii) Let (y, v) be an admissible process for the accessory problem. Then there exists $\bar{\alpha} > 0$, a family of admissible processes $\{(x^\alpha, u^\alpha) \mid 0 \leq \alpha \leq \bar{\alpha}\}$ for (P) and functions $\epsilon(.), \eta(.) : \mathbb{R}^+ \to \mathbb{R}^+$ such that $lim_{\alpha \downarrow 0} \epsilon(\alpha) = 0$, $lim_{\alpha \downarrow 0} \eta(\alpha) = 0$ and

$$\|x^\alpha - \bar{x}\|_{L^\infty} \leq \epsilon(\alpha), \quad \|u^\alpha - \bar{u}\|_{L^\infty} \leq \epsilon(\alpha) \tag{14}$$

$$|J(x^\alpha, u^\alpha) - J(\bar{x}, \bar{u}) - J_A(\alpha y, \alpha v)| \leq \eta(\alpha) \left[\int_0^T |\alpha v|^2 dt + |\alpha y(0)|^2 \right] \tag{15}$$

for all $\alpha \in [0, \bar{\alpha}]$.

LEMMA 6.2 *Assume (H6). Suppose that the system of equations (4)–(6) has a solution $(P(.), S(.), M(.))$ on $[0,T]$. Then for any process (y,v) for the accessory problem such that $y(0) = y(T)$ we have*

$$2J_A(y,v) \geq y^T(0) \left[G + P(0) + 2S(0) + M(0) - (P(T) + 2S(T) + M(T)) \right] y(0) \tag{16}$$

Furthermore, the above relationship is satisfied with equality if

$$v(t) = -R^{-1}[(PB+D)^T y(t) + B^T S y(0)] \quad \textit{for all } t \in [0,T]\ .$$

Proof. Take a process (y, v) for the accessory problem satisfying $y(0) = y(T)$. Define

$$s(t) = S(t)y(T) \text{ and } m(t) = y^T(T)M(t)y(T).$$

Notice that $s(.)$ and $m(.)$ satisfy the differential equations

$$\begin{cases} \dot{s}(t) + (A^T - (B^T P + D^T)^T R^{-1} B^T)s(t) = 0 \\ s(T) = S(T)y(T) \end{cases}$$

and

$$\begin{cases} \dot{m}(t) - s^T(t) B R^{-1} B^T s(t) = 0 \\ m(T) = y^T(T)M(T)y(T)\ . \end{cases}$$

The cost of (y, v) can be expressed as

$$\begin{aligned} 2J_A(y,v) &= \int_0^T y^T Q y + 2y^T D v + v^T R v dt + y^T(0) G y(0) \\ &+ \int_0^T \frac{d}{dt}[y^T P y + 2s^T y + m]dt - (y^T P y + 2s^T y + m)|_0^T. \end{aligned} \tag{17}$$

We calculate, for each $t \in [0, T]$,

$$\begin{aligned} & y^T Q y + 2y^T D v + v^T R v + \frac{d}{dt}[y^T P y + 2s^T y + m] \\ &= y^T[\dot{P} + PA + A^T P + Q]y + 2(\dot{s}^T + s^T A)y + \dot{m} + v^T R v \\ &\quad + 2(y^T D + y^T P B + s^T B)v \\ &\geq y^T[\dot{P} + PA + A^T P + Q]y + 2(\dot{s}^T + s^T A)y + \dot{m} \\ &\quad - [(PB+D)^T y + B^T s]^T R^{-1}[(PB+D)^T y + B^T s] \quad (\text{since } R > 0) \\ &= y^T[\dot{P} + PA + A^T P + Q - (PB+D)R^{-1}(PB+D)^T]y \\ &\quad + 2y^T[\dot{s} + (A^T - (PB+D)R^{-1}B^T)s] + (\dot{m} - s^T B R^{-1} B^T s) = 0. \end{aligned}$$

Furthermore, the above relations are satisfied with equality if

$$v(t) = -R^{-1}[(PB+D)^T y(t) + B^T S(t) y(T)]. \tag{18}$$

Since $y(0) = y(T)$, we deduce from equation (17) that

$$\begin{aligned} 2J_A(y,v) \ \geq\ & y^T(0)\big[G + P(0) + 2S(0) + M(0) \\ & -(P(T) + 2S(T) + M(T))\big]y(0) \end{aligned}$$

with equality when (18) is satisfied. ■

LEMMA 6.3 *Assume (H6). Assume furthermore that*

(a) det $[\Phi_{12}(T)] \neq 0$.

(b) There exists a solution $P(.)$ *on* $[0,T]$ *to (4) such that*

$$det\ [\Phi_{11}(T) - \Phi_{12}(T)P(T)] \ \neq \ 0\,.$$

Then, for all $\xi \in \mathbb{R}^n$,

$$\begin{aligned} &Min\{2J_A(y,v) \mid y(0) = y(T) = \xi\} \\ &= \ \xi^T\,(G + P(0) + 2S(0) + M(0) - (P(T) + 2S(T) + M(T)))\,\xi \\ &= \ \xi^T\,\big(G + \Phi_{21}(T) + (\Phi_{22}(T) - I)\Phi_{12}^{-1}(T)[I - \Phi_{11}(T)]\big)\,\xi\,. \end{aligned}$$

Here, $S(.), M(.))$ *are solutions to equations (5) and (6), with boundary conditions*

$$S(T) = \Phi_{12}^{-1}(0,T)[I - \Phi_{11}(0,T) - \Phi_{12}(T)P(T)] \quad and \quad M(T) = 0\,,$$

and the Φ_{ij}*'s are block components of the transition matrix for the Hamiltonian system.*

Proof. Fix $\xi \in R^n$. Let (y,v) be the process for the linearized system defined by the feedback relation

$$\left\{ \begin{array}{l} \dot{y} = Ay + Bv \\ v = -R^{-1}[(D^T + B^T P)y + B^T S(t)\xi] \\ y(0) = \xi. \end{array} \right. \tag{19}$$

Define

$$p(t) = P(t)y(t) + S(t)\xi.$$

Notice that

$$p(T) \ = \ P(T)y(T) + \Phi_{12}^{-1}(0,T)[I - \Phi_{11}(0,T) - \Phi_{12}^{-1}(0,T)P(T)]\xi.$$

and

$$v \ = \ -R^{-1}[D^T y + B^T p]\,.$$

We have

$$\begin{aligned} \dot{p}(t) &= \dot{P}(t)y(t) + \dot{S}(t)\xi + P(t)\dot{y}(t) \\ &= -[A^T P + PA + Q - (B^T P + D^T)^T R^{-1}(B^T P + D^T)]y(t) \\ &\quad + P[Ay - BR^{-1}(B^T(Py + S\xi)) + D^T y] - (A^T - (B^T P + D^T)^T R^{-1} B^T)S\xi\,. \end{aligned}$$

Gathering and cancelling terms gives

$$\begin{aligned} -\dot{p} \ &= \ A^T(Py + S\xi) + Qy - DR^{-1}B^T(Py + S\xi) - DR^{-1}D^T y \\ &= \ A^T p + Qy - DR^{-1}(B^T p + D^T y). \end{aligned}$$

Also,

$$\begin{aligned} \dot{y} &= Ay - BR^{-1}(B^T(Py + S\xi) + D^T y) \\ &= Ay - BR^{-1}(B^T p + D^T y). \end{aligned}$$

We have shown that (y,p) satisfies Hamilton's system of equations. Since $y(0) = \xi$,

$$\begin{aligned} y(0) &= \Phi_{11}(0,T)y(T) + \Phi_{12}(0,T)p(T) \\ &= \Phi_{11}(0,T)y(T) \\ &\quad + \Phi_{12}(0,T)[P(T)y(T) + \Phi_{12}^{-1}(0,T)(I - \Phi_{11}(0,T) - \Phi_{12}(T)P(T))]y(0). \end{aligned}$$

It follows that

$$[\Phi_{11}(T) + \Phi_{12}(0,T)P(T)](y(T) - y(0)) \;=\; 0\,.$$

Since, by assumption, $[\Phi_{11}(T) + \Phi_{12}(0,T)P(T)]$ is invertible, we conclude that

$$y(0) = y(1) = \xi\,. \tag{20}$$

Noting Lemma 6.2, we deduce from (19) and (20) that

$$\begin{aligned} &\mathrm{Min}\{2J_A(y',v') \mid y'(T) = y'(0) = \xi\} \\ &\quad = \; 2J_A(y,v) \\ &\quad = \; \xi^T[G + P(0) + 2S(0) + M(0) - (P(T) + 2S(T) + M(T))]\xi\,. \end{aligned}$$

Note next that, for a.e. $t \in [0,T]$, we have (writing y for $y(t)$ etc.)

$$\begin{aligned} &y^T Q y + 2y^T D v + v^T R v + \frac{d}{dt}(p^T y) \\ &\quad = \; y^T Q y + 2yDv + v^T R v + p^T(Ay + Bv) \\ &\qquad + y^T[(-Q + DR^{-1}D^T)y + (-A^T + DR^{-1}B^T)p] \\ &\quad = \; -2y^T DR^{-1}[D^T y + B^T p] - p^T BR^{-1}[D^T y + B^T p] \\ &\qquad + y^T DR^{-1}D^T y + y^T DR^{-1}B^T p \\ &\quad = \; 0\,. \end{aligned}$$

Integrating across this equation yields

$$\begin{aligned} 2J_A(y,v) \;&=\; -\int_0^T \frac{d}{dt} p^T y dt + y^T(T)Gy^T(T) \\ &=\; p^T(0)y(0) - p^T(T)y(T) + y^T(T)Gy(T)\,. \end{aligned}$$

But since $y(0) = y(T) = \xi$, we have

$$\begin{aligned} \xi &= \Phi_{11}(T)\xi + \Phi_{12}(T)p(T) \\ p(0) &= \Phi_{21}(T)\xi + \Phi_{22}(T)p(T)\,. \end{aligned}$$

Hence

$$\begin{aligned} p(T) &= \Phi_{12}^{-1}(T)[I - \Phi_{11}(T)]\xi \\ p(0) &= (\Phi_{21}(T) + \Phi_{22}(T)\Phi_{12}^{-1}(T)[I - \Phi_{11}(T)])\xi\ . \end{aligned}$$

We conclude that

$$2J_A(y,v) \;=\; \xi^T \left(\Phi_{21}(T) + (\Phi_{22}(T) - I)\Phi_{12}^{-1}[I - \Phi_{11}(T) + G]\right)\xi\ .$$

This identity completes proof of the lemma. ∎

LEMMA 6.4 *Assume (H6). Suppose that* $det\{\Phi(T)\} \neq 0$. *Then* $(A(.)B(.))$ *is controllable in* $[0,T]$.

Proof. Take any $\xi \in R^n$. Since $det\{\Phi_{12}(T)\} \neq 0$, we can choose (y,p) to be a solution to Hamilton's system of equations with right boundary condition $(y(T),p(T)) = (\xi, -\Phi_{12}^{-1}(T)\xi)$. Let v be the control $v = -R^{-1}(D^Ty + B^Tp$. It is straightforward to check that (y,v) is a process for (A) that satisfies $y(0) = 0$ and $y(T) = \xi$. Since ξ is arbitrary, we deduce that $(A(.),B(.))$ is controllable. ∎

LEMMA 6.5 *Assume (H6). Suppose that there exists a solution* $P'(.)$ *to the Riccati equation (4) on* $[0,T]$ *with right boundary condition* $P'(T) = E'$, *for some symmetric matrix* E'. *Then, (4) has a solution on* $[0,T]$ *with right boundary* $P(T) = E$ *for any symmetric matrix such that* $E \geq E'$.

Proof. Suppose that the Riccati equation (4), with right boundary condition $P(t) = E$, has a solution $P(.)$ on some sub-interval $[S,T] \subset [0,T]$. A standard argument from the theory of differential equations permits us to conclude that this solution can be extended to all of $[0,T]$, provided that we can find a positive number K, independent of S such that

$$|P(s)| \leq K\ . \tag{21}$$

(Throughout this proof, the norm on symmetric $n \times n$ matrices is the operator norm induced by the Euclidean norm on $\mathbb{R}^n$.) We now establish existence of such a K.

Fix $\xi \in \mathbb{R}^n$ and and let $y(.)$ be the solution to the linearized system equation on $[S,T]$ with boundary condition $y(S) = \xi$ generated by the linear feedback control law:

$$v(t) \;=\; -R^{-1}(B^TP + D^T)y(t)\ .$$

We have, writing $2L_A(.,.,.)$ for the cost integrand in the accessory problem,

$$\begin{aligned} \xi^TP(S)\xi &= \xi^TP(T)\xi - \int_S^T \frac{d}{dt}P(t)dt \\ &= \int_S^T L_A(t,y,v)dt + y(T)^TE(S)y(T) \end{aligned}$$

$$= \text{Min}_{(y',v')}\{\int_S^T L_A(t,y',v')dt + y'(T)^T E(S)y'(T) \,|\, y'(S)=\xi\}$$
$$\geq \text{Min}_{(y',v')}\{\int_S^T L_A(t,y',v')dt + y'(T)^T E'(S)y'(T) \,|\, y'(S)=\xi\}$$
$$= \xi^T P'(S)\xi \;\geq\; -K_1|x_i|^2\,.$$

Here, K_1 is a uniform bound on $|P'(t)|$, $0 \leq t \leq T$, that does not depend on S or ξ.

Consideration of the suboptimal process $(z(.), w(.) \equiv 0)$ on $[S,T]$ with right boundary condition $z(S) = \xi$ establishes that

$$\xi^T P(S)\xi = \text{Min}\;\{\int_S^T L_A(t,y',v')dt + y'(T)^T E(S)y'(T) \,|\, y'(S)=\xi\}$$
$$\leq \int_S^T L_A(t,z,w)dt + z(T)^T E(S)z(T) \,|\, z(S)=\xi\}$$
$$\leq K_2|\xi|^2\,,$$

for some number K_2 that does not depend on S or ξ.

The desired bound (21) follows from the above the relationships, with $K = max\{K_1, K_2\}$. ■

7. Proofs of Theorems 2.1 and 3.1 and Propositions 4.1 and 4.2

Proof of Theorem 2.1

Suppose there exists $r > 0$ and $(P(.), S(.), M(.))$ with the stated properties. Then, by continuous dependence properties of solutions to differential equations, there exist $\alpha > 0$ such that $R(t) - \alpha I$ is positive definite, for all t, and also such that there exist solutions $(\tilde{P}(.), \tilde{S}(.), \tilde{M}(.))$ to modified versions of (4)–(6), in which $R - \alpha I$ replaces R respectively, and

$$(\tilde{P}(T), \tilde{S}(T), \tilde{M}(T)) = (P(T), S(T), M(T)).$$

By reducing the size of α, if necessary, we can arrange that

$$\xi^T[G + \tilde{P}(0) + 2\tilde{S}(0) + \tilde{M}(0) - (\tilde{P}(T) + 2\tilde{S}(T) + \tilde{M}(T))]\xi > \frac{r}{2}|\xi|^2$$

for all $\xi \in \mathbb{R}^n$. Take any $\delta > 0$. According to Lemma 6.1, we can choose $\epsilon > 0$ such that, corresponding to any admissible process $(x(.), u(.))$ for (P) such that

$$\|(x,u) - (\bar{x},\bar{u})\|_{L^\infty} \leq \epsilon\,, \tag{22}$$

there exists an admissible process (y, v) for the accessory problem such that

$$J(x,u) - J(\bar{x},\bar{u}) \;\geq\; J_A(y,v) - \delta \left[\int_0^T |v|^2 dt + |y(0)|^2 \right] .$$

Write

$$\begin{aligned}\tilde{J}_A(y,v) \;:=& \frac{1}{2}\int_0^T \left(y^T(t)Q(t)y(t) + 2y^T(t)D(t)v(t) + v^T(t)[R(t) - \alpha I]v(t)\right) dt \\ &+ \frac{1}{2}y^T(T)Gy(T) .\end{aligned}$$

Choose any such (x, u) (and corresponding (y, v)). Then

$$J(y,v) \;=\; \tilde{J}_A(y,v) + \alpha \int_0^T |v^2| dt.$$

It follows now from Lemma 6.2, applied to the accessory problem with modified cost $\tilde{J}_A$, that

$$\begin{aligned}2\tilde{J}_A(y,v) &\geq y(0)^T[G + \tilde{P}(0) + 2\tilde{S}(0) + \tilde{M}(0) - (\tilde{P}(T) + 2\tilde{S}(T) + \tilde{M}(T))]y(0) \\ &\geq \frac{r}{2}|y(0)|^2.\end{aligned}$$

Now choose $\delta > 0$ such that $\delta \leq min\{r/4, \alpha/4\}$. Then

$$\begin{aligned}J(x,u) - J(\bar{x},\bar{u}) &\geq \tilde{J}_A(y,v) + (\alpha - \delta)\int_0^T |v|^2 dt - \delta|y(0)|^2 \\ &\geq (\frac{r}{2} - \delta)|y(0)|^2 + (\alpha - \delta)\int_0^T |v|^2 dt \\ &\geq (\frac{r}{4})|y(0)|^2 + (\frac{\alpha}{2})\int_0^T |v|^2 dt \;\geq\; 0.\end{aligned}$$

We have shown that $(\bar{x}, \bar{u})$ is a minimizer with respect to all admissible processes (x, u) satisfying (22). ■

Proof of Theorem 3.1

Suppose that the Riccati equation (4) has a solution $P'(.)$ on $[0, T]$. Then $\alpha I \geq P'(T)$, for some all α sufficiently large. Choose such an α. It follows from Lemma 6.5 that the Riccati equation (4) has a solution on $[0, T]$, with right boundary value αI. Notice next that, since $\Phi_{12}(T)$ is assumed to be invertible, we can write

$$\Phi_{11}(T) - \alpha\Phi_{12}(T) \;=\; \alpha\Phi_{12}^{-1}(T)[\alpha^{-1}\Phi_{12}^{-1}(T)\Phi_{11}(T) - I] .$$

It follows from this identity that, if we adjust α also to satisfy

$$\alpha \;\geq\; |\Phi_{12}^{-1}(T)||\Phi_{11}(T)| ,$$

then we retain a solution $P(,)$ to the Riccati equation (4) on $[0,T]$, but now also with the property that the matrix $(\Phi_{11}(T)-\Phi_{12}(T)P(T))$ is invertible. (In the above identity the matrix norms are operator norms generated by the Euclidean norm.) The first assertion of the theorem is proved. The remaining assertions of the theorem follow from Lemma 6.3. ■

Proof of Proposition 4.1

Suppose, in contradiction, that there exists an $r>0$ and an admissible process (y,v) for the accessory problem such that

$$J_A(y,v)<-r.$$

Take any $\epsilon>0$. Let $\{(x^\alpha,u^\alpha)|\alpha\in[0,+\bar{\alpha}]\}$, $\epsilon(.)$ and $\eta(.)$ be as in part (ii) of Lemma (6.1). Then for all $\alpha\in(0,\bar{\alpha}]$

$$J(x^\alpha,u^\alpha)-J(\bar{x},\bar{u})\le J_A(\alpha y,\alpha v)+\eta(\alpha)\left[\int_0^T|\alpha v|^2dt+|\alpha y(0)|^2\right].$$

Dividing across by α^2 and noting that $J_A(\alpha y,\alpha v)=\alpha^2 J_A(y,v)$ we have

$$\alpha^{-2}(J(x^\alpha,u^\alpha)-J(\bar{x},\bar{u}))\ \le\ -r+\eta(\alpha)K$$

where

$$K=\int\left\{|v|^2dt+|y(0)|^2\right\}.$$

It follows that, if we choose $\alpha\in(0,\alpha]$ such that $\epsilon(\alpha)<\epsilon$ and $\eta(\alpha)<\frac{r}{K}$ then

$$\|(x^\alpha,u^\alpha)-(\bar{x},\bar{u})\|_{L^\infty}\le\epsilon\quad\text{and}\quad J(x^\alpha,u^\alpha)-J(\bar{x},\bar{u})<0.$$

Since $\epsilon>0$ was arbitrary, we conclude that $(\bar{x},\bar{u})$ cannot be a weak local minimizer. ■

Proof of Proposition 4.2

Assume (a). Then there exist $r>0$ and $\alpha>0$ such that $R(t)-\alpha I$ is positive definite, for all t, and also solutions $(\tilde{P}(.),\tilde{S}(.),\tilde{M}(.))$ to modified versions of (4)–(6), in which $R-\alpha I$ replaces R respectively, and

$$(\tilde{P}(T),\tilde{S}(T),\tilde{M}(T))=(P(T),S(T),M(T))$$

and

$$\xi^TW\xi>\frac{r}{2}|\xi|^2$$

for all $\xi \in \mathbb{R}^n$. For any admissible process (y, v) for the accessory problem write

$$\tilde{J}_A \ := \frac{1}{2}\int_0^T \big(y^T(t)Q(t)y(t) + 2y^T(t)D(t)v(t) + v^T(t)[R(t) - \alpha I]v(t)\big)\, dt + \frac{1}{2}y^T(T)Gy(T)\ .$$

We have

$$\tilde{J}_A(y,v) \ = \ J_A(y,v) + \alpha \int_0^T |v|^2 dt\,.$$

From Lemma 6.2,

$$\tilde{J}_A(y,v) \ \geq \ y^T(0)^T W y(0)^T \ \geq \ \frac{r}{2}|y(0)|^2\,.$$

But then

$$J_A(y,v) \geq \gamma[|y(0)|^2 + \alpha \int_0^T |v|^2 dt]$$

where $\gamma = min\{\frac{r}{2}, \alpha\}$. (b) is confirmed.

Assume (b). We must confirm conditions (I), (II) and (III).

(I) Suppose, in contradiction, that $det\ [\Phi_{12}(T)] = 0$. Then there exists $\xi \neq 0$ such that $\Phi_{12}(T)\xi = 0$. Let (y, p) be the solution to the Hamiltonian system with boundary condition

$$(y(T), p(T)) = (0, \xi)$$

and define the control v

$$v \ = \ -R^{-1}(D^T y + B^T p)\,.$$

Since $\Phi_{12}(T)\xi = 0$, we know $y(0) = 0$. It follows that (y, v) satisfies the periodic boundary conditions. Arguing as in the proof of Lemma 6.3, we show that

$$2J_A(y,v) \ = \ p^T(0)y(0) - p^T(T)y(T) + y^T(T)Gy(T) \ = \ 0\,.$$

It follows that (y, v) is a minimizer for the accessory problem. By (b), however, $(\bar{y}, \bar{v}) \equiv (0, 0)$ is the only minimizer. Consequently, $(y, v) \equiv (0, 0)$. We deduce from Hamilton's equations that (p, v) satisfy

$$\begin{cases} -\dot{p} &= \ -A^T p \\ \ \ v &= \ B^T p \end{cases}$$

But $p(T) \neq 0$. It follows from the assumption that $(A(.), B(.))$ is controllable on $[0, T]$ that $v \neq 0$. From this contradiction, we deduce that (I) is true.

(II) $(\bar{y}, \bar{v}) \equiv (0,0)$ is, in particular, a minimizer for a variant of the accessory problem in which both endpoints of state trajectories are constrained to be zero vectors. It follows from Hestenes (1951, Thms. 13.2 and 13.3) that there exists $c > 0$ such that $(\bar{y}, \bar{v}) \equiv (0,0)$ is a minimizer for the problem *without endpoint constraints*

$$\left\{ \begin{array}{l} \text{Minimize } \int_0^T L_A(t, y(t), v(t))dt + c[|y(0)|^2 + |y(T)|^2] \\ \text{over } y \in W^{1,1} \text{ and } v \in L^2 \text{ such that} \\ \dot{x} = Ax + Bu\,. \end{array} \right.$$

Here $(1/2)L_A$ is the cost integrand of the accessory problem. The Riccati equation has a solution $P(.)$, with boundary condition $P(T) = cI$ on some interval $[S, T]$. It suffices to show that there exist constants $\alpha_1 > 0$ and $\alpha_2 > 0$, independent of S, such that

$$-\alpha_1 I \;\leq\; P(S) \;\leq\; \alpha_2 I\,,$$

for then the solution can be extended to all of $[0, T]$.

Take any $\xi \in \mathbb{R}^n$. Let $(y'(t), v'(t))$, $S \leq t \leq T$, be the process for the accessory problem on $[S, T]$ defined by the initial condition $y'(s) = \xi$ and the feedback relation

$$v' \;=\; -R^{-1}(B^T P + D^T)y'\,.$$

Let (y, v) be the process such that

$$v(t) \;=\; \left\{ \begin{array}{ll} v'(t) & \text{if } t > S \\ 0 & \text{if } t \leq S \end{array} \right.$$

and $y(S) = \xi$. It follows from a standard 'dynamic programming' analysis that

$$\xi^T P(S)\xi \;=\; \int_S^T L_A(t, y(t), v(t)) + c|y(T)|^2.$$

It is easily shown that α_1 can be chosen, independent of S and ξ, such that

$$\left| \int_0^S L_A(t, y(t), v(t))dt + c|y(0)|^2 \right| \;\leq\; \alpha_1 |\xi|^2\,.$$

But then, since $(\bar{y} \equiv 0, \bar{v} \equiv 0)$ is a minimizer for the above problem,

$$\begin{aligned} 0 \;&\leq\; \int_0^T L_A(t, y(t), v(t))dt + c(|y(0)|^2 + |y(T)|^2) \\ &=\; \int_0^S L_A(t, y(t), v(t))dt + c|y(0)|^2 + \int_S^T L_A(t, y(t), v(t))dt + c|y(T)|^2 \\ &\leq\; \alpha_1 |\xi|^2 + \xi^T P(S)\xi\,. \end{aligned}$$

It follows that $-\alpha_1 I \;\leq\; P(S)$.

For any $\xi \in \mathbb{R}^n$ choose (y, v) to be the process on $[S, T]$ defined by $v \equiv 0$ and $y(S) = \xi$. We can find $\alpha_2 > 0$, independent of S and ξ, such that

$$\xi^T P(S)\xi \ \leq \ \int_S^T L_A(t, y(t), v(t))dt + c|y(T)|^2 \ \leq \ \alpha_2|\xi|^2 \, .$$

It follows that $P(S) \ \leq \ \alpha_2 I$. (II) has been confirmed.

(III) We have shown that conditions (I) and (II) are satisfied. It follows from Lemma 6.3 that, for any $\xi \in R^n$,

$$\xi^T W\xi \ = \ \frac{1}{2} min\{J_A(y, v) \,|\, y(0) = y(T) = \xi\}.$$

It follows from (b) that $W > 0$. (III) is confirmed. ∎

Acknowledgement

The assistance of Xu Liu in this research is gratefully acknowledged.

References

Hestenes, M. R. (1951) Applications of the Theory of Quadratic Forms in Hilbert Space to Calculus of Variations. *Pacific J. Math* **1**, 525-581.

Maurer, H. and Pickenhain, S. (1995) Second order sufficient conditions for optimal control problems with mixed control-state constraints. *J. Optim. Theory and Applic.* **86**, 649-667.

Speyer, J.L. (1996) Periodic Optimal Flight. *J. Guidance, Control and Dynamics* **61**,745-754.

Stefani, G. and Zezza, P. (1997) Constrained Regular LQ-Control Problems. *SIAM J. Control and Optim.* **35**, 876-900.

Vinter, R.B. (2000) *Optimal Control*. Birkhäuser, Boston.

Wang, Q. and Speyer, J.L. (1990) Necessary and Sufficient Conditions for Local Optimality of a Periodic Process. *SIAM J. Control and Optim.* **28**, 482-497.

Zeidan, V. (2001) New Second-Order Optimality Conditions for Variational Problems with C^2-Hamiltonians. *SIAM J. Control and Optim.* **40**, 577-609.

Appendix: Proof of Lemma 6.1

(i) Fix $\delta > 0$. Let ϵ be a number, satisfying $0 < \epsilon \leq 1$, whose value will be chosen presently. Take any process (x, u) satisfying $x(0) = x(T)$ and

$$\|(\Delta\, x, \Delta\, u)\|_{L^\infty} \ \leq \ \epsilon \tag{23}$$

where

$$\Delta x := (x - \bar{x}) \quad \text{and} \quad \Delta u := (u - \bar{u}).$$

Let $z(.)$ be the solution to the linearized equation

$$\dot{z}(t) = A(t)z(t) + B(t)\Delta u$$

with boundary condition $z(0) = \Delta x(0)$. It can be deduced from the fact that the control system $(A(.), B(.))$ is $(I; -I)$ controllable on $[0, T]$ that there exists a control $v(.)$ and a state trajectory $y(.)$ for (S) such that $y(0) = y(T)$ and

$$||(y, v) - (z, \Delta u)||_{L^\infty} \ \leq \ K_1 |z(T) - z(0)| \tag{24}$$

for some constant K_1, which does not depend on our choice of (x, u). In view of (H4), there exists K_2, independent of (x, u), such that

$$\begin{aligned}
&\int |\Delta \dot{x} - A(t)\Delta x - B(t)\Delta u| dt = \\
&\int_0^T |f(t, x, u) - f(t, \bar{x}(t), \bar{u}(t)) - f_x(t, \bar{x}(t), \bar{u}(t))\Delta x - f_u(t, \bar{x}(t), \bar{u}(t))\Delta u| dt \\
&\quad \leq K_2 \int_0^T (|\Delta x|^2 + |\Delta u|^2)\, dt\,.
\end{aligned}$$

By Filippov's Existence Theorem (see, e.g. Vinter, 2000), there exists K_3, independent of (x, u), such that

$$\| \ z - \Delta x \ \|_{L^\infty} \leq \ K_3 \int_0^T \left(|\Delta x|^2 + |\Delta u|^2\right) dt.$$

By (24) and the triangle inequality,

$$\|(y, v) - (\Delta x, \Delta u)\|_{L^\infty} \ \leq \ (1 + 2K_1)K_3 \int_0^T \left(|\Delta x|^2 + |\Delta u|^2\right) dt\ . \tag{25}$$

But

$$\int_0^T \left(|\Delta x|^2 + |\Delta u|^2\right) dt \leq 2 \int_0^T (|y|^2 + |v|^2) dt + 2 \int_0^T (|y - \Delta x|^2 + |v - \Delta u|^2) dt\ . \tag{26}$$

By (23) and (25) and the triangle inequality

$$\|(y, v) - (\Delta x, \Delta u)\|_{L^\infty} \leq 2(1 + K_1)K_3 T\epsilon^2 \ \text{ and } \ \|(y, v)\|_{L^\infty} \leq \epsilon + 2(1 + K_1)K_3 T\epsilon^2\,. \tag{27}$$

From (25), then,

$$\int_0^T \left(|y-\Delta x|^2+|v-\Delta u|^2\right)\,dt \leq 4(1+K_1)^2K_3^2T^2\epsilon^2\int_0^T \left(|\Delta x|^2+|\Delta u|^2\right)dt\,.$$

It follows now from (26) that if ϵ is chosen to satisfy $8(1+K_1)^2K_3^2T^2\epsilon^2 < 1$, then

$$\int_0^T \left[|\Delta x|^2+|\Delta u|^2\right]dt \leq K_4\int_0^T \left(|y|^2+|v|^2\right)dt \tag{28}$$

where $K_4 = 2[1-8(1+K_1)^2K_3^2T^2\epsilon^2]^{-1}$.

Now consider the difference in cost between (x,u) and $(\bar{x},\bar{u})$:

$$\Delta J(x,u) := \int_0^T [L(t,x(t),u(t)) - L(t,\bar{x}(t),\bar{u}(t))]dt + g(x(0)) - g(\bar{x}(0)).$$

We can write

$$\begin{aligned}\Delta J(x,u) = &\int_0^T (L(t,x(t),u(t)) - L(t,\bar{x}(t),\bar{u}(t)))\,dt + g(x(0)) - g(\bar{x}(0))\\ &-\int_0^T p(t)\cdot[(\dot{x}-\dot{\bar{x}}) - (f(t,x(t),u(t)) - f(t,\bar{x}(t),\bar{u}(t)))]\,dt\,.\end{aligned}$$

Integrating by parts, making use of the costate arc equation and transversality conditions and noting that $x(0)=x(T)$, we deduce

$$\begin{aligned}-\int_0^T p(t)\cdot(\dot{x}-\dot{\bar{x}})\,dt = &\\ &-\nabla g(\bar{x}(0))\cdot(x(0)-\bar{x}(0))\\ &-\int_0^T (p(t)\cdot f_x(t,\bar{x}(t),\bar{u}(t)) + L_x(t,\bar{x}(t),\bar{u}(t)))\Delta x\,dt\,.\end{aligned}$$

Since, $H_u(t,\bar{x}(t),p(t),\bar{u}(t)) = 0$ a.e.,

$$\begin{aligned}\Delta J(x,u) = &\int_0^T [H(t,x(t),p(t),\bar{u}(t)) - H_u(t,\bar{x}(t),p(t),\bar{u}(t))\\ &-H_x(t,\bar{x}(t),p(t),\bar{u}(t))(x(t)-\bar{x}(t)) - H_u(t,\bar{x}(t),p(t),\bar{u}(t))(u(t)-\bar{u}(t))]\,dt\\ &+g(x(0)) - g(\bar{x}(0)) - \nabla g(\bar{x}(0))\cdot(x(0)-\bar{x}(0)).\end{aligned}$$

But then

$$\begin{aligned}\Delta J(x,u) \geq &\frac{1}{2}\int_0^T (\Delta x,\Delta u)^T\nabla^2 H(t,\bar{x}(t),p(t),\bar{u}(t))(\Delta x,\Delta u)dt\\ &+\frac{1}{2}(x(0)-\bar{x}(0))^T\nabla^2 g(\bar{x}(0))(x(0)-\bar{x}(0))\\ &-\eta(\epsilon)\left[\int_0^T (|\Delta x|^2+|\Delta u|^2)dt + |\Delta x(0)|^2\right]\end{aligned}$$

for some function $\eta : \mathbb{R}^+ \to \mathbb{R}^+$ (independent of (x,u)), such that $\eta(\epsilon) \downarrow 0$ as $\epsilon \downarrow 0$. We deduce from (23) and (25) that there exists a constant K_5 such that

$$\Delta J(x,u) \geq \frac{1}{2}\int_0^T (y,v)^T \nabla^2 \mathcal{H}(t,y(t),p(t),v(t))(y,v)dt$$
$$+\frac{1}{2}y(0)^T\nabla^2 g(\bar{x}(0))\,y(0) - \eta(\epsilon)\left[\int_0^T \left(|\Delta x|^2+|\Delta u|^2\right)dt + |\Delta x(0)|^2\right] - K_5 e$$

where

$$e = \|(y,v)-(\Delta x,\Delta u)\|_{L^\infty} \,.\, \left(\|(y,v)\|_{L^\infty} + \|(\Delta x,\Delta u)\|_{L^\infty}\right)\ .$$

But in view of (25) and (27) and since $\epsilon \leq 1$ we have

$$e \ \leq\ \epsilon K_6 \int_0^T \left(|\Delta x|^2+|\Delta u|^2\right)dt$$

for some constant K_6. We also have

$$\begin{aligned} |\Delta x(0)|^2 &\leq |y(0)|^2 + |y(0)+\Delta x(0)|.|y(0)-\Delta x(0)| \\ &\leq |y(0)|^2 + \epsilon K_7 \int_0^T \left(|\Delta x|^2+|\Delta u|^2\right)dt \end{aligned}$$

for some constant K_7. It follows that for arbitrary (x,u),

$$\Delta J(x,u) \geq \frac{1}{2}\int_0^T (y,v)^T \nabla^2 \mathcal{H}(t,y(t),p(t),v(t))(y,v)dt + \frac{1}{2}y(0)^T\nabla^2 g(\bar{x}(0))\,y(0)$$
$$-\left(\eta(\epsilon)+\epsilon K_5K_6+\eta(\epsilon)K_7\right)\left[\int_0^T (|y|^2+|v|^2)dt + |y(0)|^2\right]\ .$$

But there exists a constant $c>0$ such that, for any process for the accessory problem we have

$$\int_0^T |y|^2 dt \ \leq\ c[\int_0^T |v|^2dt + |y(0)|^2]\,.$$

It follows that, if we choose ϵ additionally to satisfy the condition

$$(1+c)\left(\eta(\epsilon)+\epsilon K_5K_6+\eta(\epsilon)K_7\epsilon\right)\ \leq\ \delta$$

then,

$$\Delta J(x,u)\ \geq\ \delta[\int_0^T |v|^2dt + |y(0)|^2]\,.$$

This is the desired inequality.

(ii) Let $\xi_1, \ldots, \xi_n$ be linearly independent vectors in $\mathbb{R}^n$. Since the linearized system is $[I, -I]$-controllable, we can choose control functions $v_1, \ldots, v_n$ and corresponding state trajectories $y_1, \ldots, y_n$ for the linearized system (S), such that

$$y_i(T) - y_i(0) \;=\; \xi_i \qquad \text{for } i = 1, \ldots, n\,.$$

For each $\alpha \in \mathbb{R}$ and $\lambda \in \mathbb{R}^n$ such that $|\alpha| + |\lambda|$ is sufficiently small, define

$$x^{\alpha,\lambda}(t) := x(t; \bar{u} + \alpha v + \sum_i \lambda_i v_i, \bar{x}(0) + \alpha y(0) + \sum_i \lambda_i y_i(0))$$

where, for any control function u' and initial state $x_0 \in \mathbb{R}^n$, $x(.; u', x_0)$ is the solution to

$$\begin{cases} \dot{x}(t) = f(t, x(t), u'(t)) \\ x(0) = x_0\,. \end{cases}$$

Now, consider the equation

$$\psi(\alpha, \lambda) = 0\,,$$

where the $\mathbb{R}^n$-valued function $\psi(.,.)$ is defined on a neighbourhood of the origin in $\mathbb{R} \times \mathbb{R}^n$ according to

$$\psi(\alpha, \lambda) \;=\; x^{\alpha,\lambda}(T) - x^{\alpha,\lambda}(0)\,.$$

By standard 'continuous dependence' results from the theory of differential equations, ψ is a C^2 function on some neighbourhood of $(0, 0)$. We see that $\psi(0, 0) = 0$. Since $\nabla_\lambda \psi(0, 0) = [\xi_1, \ldots, \xi_n]^T$ is invertible, we deduce, furthermore, from the implicit function theorem that there exists a C^2 function $\lambda(.)$ on some interval $[-\alpha', +\alpha']$ such that

$$\lambda(0) \;=\; 0 \tag{29}$$

and

$$\psi(\alpha, \lambda(\alpha)) = 0 \quad \text{for all } [-\alpha', +\alpha']. \tag{30}$$

Notice that

$$0 \;=\; \frac{d}{d\alpha}\psi(0, \lambda(0)) \;=\; \nabla_\alpha \psi(0, 0) + \nabla_\lambda \psi(0, 0)\frac{d}{d\alpha}\lambda(0)\,.$$

Hence

$$\frac{d}{d\alpha}\lambda(0) \;=\; -\nabla_\lambda^{-1}\psi(0, 0)\nabla_\alpha \psi(0, 0) = \nabla_\lambda^{-1}\psi(0, 0)(y(T) - y(0)) \;=\; 0\,. \tag{31}$$

Now define

$$x^\alpha = x^{\alpha,\lambda(\alpha)} \quad \text{and} \quad u^\alpha = \bar{u} + \alpha v + \sum \lambda_i(\alpha) v_i.$$

Then, by (30),

$$x^\alpha(T) = x^\alpha(0) \quad \text{for all } \alpha \in [0, \alpha'].$$

A routine analysis, based on Filippov's Existence Theorem and the use of (29) and (31) yields a constant $K > 0$ such that, for all sufficiently small values of α,

$$\|x^\alpha - \bar{x} - \alpha y\|_{L^\infty}, \quad \|u^\alpha - \bar{u} - \alpha v\|_{L^\infty} \leq K|\alpha|^2 \; . \tag{32}$$

Writing

$$\Delta J(x^\alpha, u^\alpha) \; = \; J(x^\alpha, u^\alpha) - J(\bar{x}, \bar{u})$$

yields

$$\Delta J(x^\alpha, u^\alpha) \; = \; \Delta J(x^\alpha, u^\alpha) + \int_0^T p(t) \cdot [\dot{x}^\alpha - f(t, x^\alpha, u^\alpha) - \dot{\bar{x}} + f(t, \bar{x}, \bar{u})] dt \tag{33}$$

for all α sufficiently small.

The second order expansion techniques employed in the proof of part (i) of this Lemma, based on (32) and (33) permit us to conclude the existence of a function $\eta : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\lim_{\alpha \downarrow 0} \eta(\alpha) = 0$ and (14) and (15) are satisfied for all $\alpha \in [0, \bar{\alpha}]$.