A necessary and sufficient condition for non-smooth linear-state optimal control problems

Martimort, David and Stole, Lars

October 2009

Working paper

NECESSARY AND SUFFICIENT CONDITIONS FOR NON-SMOOTH LINEAR-STATE OPTIMAL CONTROL PROBLEMS[1]

October 2009

David Martimort and Lars Stole

We present a set of necessary and sufficient conditions for a class of optimal control problems with pure state constraints for which the objective function is linear in the state variable but the objective function is otherwise only restricted to be upper semicontinuous in the control variable.

Keywords: Optimal control, non-smooth optimization, convex analysis.

1. THE THEOREM

We are interested in the following pure-state control program ($\mathcal{P}$):

$$\text{Maximize } \Lambda(x) \equiv \int_0^1 \left(S(\theta, u(\theta)) - x(\theta)f(\theta)\right) d\theta$$
$$\text{subject to } x \in AC(\Theta, \mathbb{R}),\ \dot{x}(\theta) = u(\theta),\ x(\theta) \geq 0 \text{ for all } \theta \in \Theta \equiv [0, 1].$$

The constraints require that the state variable x is a non-negative, absolutely continuous function, $x \in AC(\Theta, \mathbb{R})$. x is said *admissible* if it satisfies these constraints. Note that the integrand $L(\theta, x, u) = (S(\theta, u) - x)f(\theta)$ is linear in x and that the state constraint, $x \geq 0$, is independent of θ. These two restrictions within the class of state-constrained, non-smooth optimal control problems are the source of many sharp results in the analysis that follows.

We assume that $S(\theta, \cdot)$ is an upper-semi continuous function bounded from above and that $f(\theta)$ is a positive and bounded from above function so that $F(\theta) \equiv \int_{[\underline{\theta},\theta]} f(\theta)$ is absolutely continuous. Without loss of generality, we normalize f such that $F(1) = 1$ and interpret F as a continuous probability distribution. Lastly, we assume that $S(\cdot, \cdot)$ is $\mathcal{L} \times \mathcal{B}$-measurable, where $\mathcal{L}$ denotes the set of Lebesgue measurable subsets of Θ and $\mathcal{B}$ is the set of Borel measurable subsets of $\mathbb{R}$. Importantly, we do not assume *a priori* that $S(\theta, \cdot)$ is a continuous function. We present our main result for this class of problems.

Theorem 1 *$\overline{x}$ is a solution to program ($\mathcal{P}$) if and only if $\overline{x}$ is admissible and there exists a probability measure μ defined over the Borel subsets of Θ with an associated adjoint function, $\overline{M} : \Theta \to [0, 1]$, defined by $\overline{M}(\underline{\theta}) = 0$ and for $\theta > \underline{\theta}$,*

$$\overline{M}(\theta) \equiv \int_{[\underline{\theta},\theta)} \mu(ds),$$

such that the following two conditions are satisfied:

$$\text{(1.1)} \qquad \operatorname{supp}\{\mu\} \subseteq \{\theta \,|\, \overline{x}(\theta) = 0\},$$

[1]We are especially thankful to John Birge for many helpful discussions.
[a]Paris School of Economics-EHESS, david.martimort@parisschoolofeconomics.eu
[b]University of Chicago Booth School of Business,lars.stole@chicagobooth.edu

$$\dot{\overline{x}}(\theta) \in \arg\max_{v \in \mathbb{R}} \; S(\theta, v) + (F(\theta) - \overline{M}(\theta))v, \; \textit{for a.e. } \theta \in \Theta. \tag{1.2}$$

Furthermore, if

$$y(\theta, \sigma) \equiv \arg\max_{v \in \mathbb{R}} \; S(\theta, v) + (F(\theta) - \sigma)v$$

is single-valued and continuous over the domain $(\theta, \sigma) \in \Theta \times [0, 1]$*, then the solution* $\overline{x}$ *to* $(\mathcal{P})$ *is continuously differentiable.*

Remarks:

• Theorem 1 is very similar to Theorem 1 in Jullien (2000). In both theorems, necessary and sufficient conditions are stated in terms of a probability measure which serves to express a "complementary slackness condition" (1.1) and an optimality condition (1.2). Moreover, both theorems use a similar condition to establish the continuity of $\dot{\overline{x}}(\theta)$ in the solution to (P). Jullien's Theorem, however, uses the stronger hypothesis that $S(\cdot)$ is twice continuously differentiable. Our technical contribution is to weaken these hypotheses to requirements of upper semi-continuity. This generalization allows us to apply the necessary and sufficient conditions above to our class of common agency games with upper-semi continuous contract menus.

• The condition that $y(\theta, \sigma)$ is single-valued and continuous is implied by the strict concavity of $S(\theta, \cdot)$. It is also implied by the weaker condition in Jullien (2000, Assumption 2) that $S(\theta, v) - (\sigma - F(\theta))v$ is strictly quasi-concave in v for any $\sigma \in [0, 1]$.

• The adjoint function $\overline{M}(\theta)$. Note in particular that the function $\overline{M}$ is constructed to be *left*-continuous rather than right-continuous.

2. PROOF OF NECESSITY

We prove necessity by specializing Theorem 3 from Vinter and Zheng (1998), exploiting fact that our integrand in Λ is a linear function of x and that the state constraint $x(\theta) \geq 0$ is linear and independent of θ.

Preliminaries for Non-Smooth Analysis. We first introduce some additional notation. We draw heavily from Vinter and Zheng (1998) in the following presentation.[1]

Take a closed set $A \subseteq \mathbb{R}^k$ and a point $x \in A$. A vector $p \in \mathbb{R}^k$ is a *limiting normal* to A at x if there exists a sequence $(x_i, p_i) \to (x, p)$ and a $K \geq 0$ such that for each i in the sequence $p_i \cdot |x_i - x| \leq K|x_i - x|^2$. The cone of limiting normal vectors to A at x is denoted $N_A(x)$. Given a lower semi-continuous function $g : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ and a point $x \in \mathbb{R}^k$ such that $g(x) < +\infty$, the *limiting subdifferential* of g at x is defined as

$$\partial g(x) \equiv \{\xi \,|\, (\xi, -1) \in \; N_{\text{epi}\{g\}}(x, g(x))\},$$

where epi$\{g\}$ is the *epigraph* of the function g defined as

$$\text{epi}\{g\} \equiv \{(x, \alpha) \in \mathbb{R}^k \times \mathbb{R} \,|\, \alpha \geq g(x)\}.$$

[1] A complete treatment can be found in the monograph of Vinter (2000). Theorem 3 from Vinter and Zheng (1998) appears as Theorem 10.2.1 in Vinter (2000).

The *asymptotic limiting subdifferential* of g at x, written $\partial^\infty g(x)$, is defined as

$$\partial^\infty g(x) \equiv \{\xi \,|\, (\xi, 0) \in N_{\text{epi}\{g\}}(x, g(x))\}.$$

Two results from nonsmooth analysis (e.g., Vinter (2000), Propositions 4.3.3 and 4.3.4) that we use are (1) $\partial^\infty g(x) = \{0\}$ if g is Lipschitz continuous and (2) for any x such that $g(x)$ is finite,

$$N_{\text{epi}\{g\}}(x, g(x)) = \{(\xi d, -\xi) \,|\, \xi > 0\, d \in \partial g(x)\} \cup \{\partial^\infty g(x) \times \{0\}\}.$$

We denote the Euclidean norm in $\mathbb{R}^k$ by $|\cdot|$, and denote the norm on the space of absolutely continuous functions by

$$||x|| \equiv |x(\underline{\theta})| + \int_\Theta |\dot{x}(\theta)| d\theta.$$

A *local maximizer* of $\Lambda(x)$ is a feasible arc, $\overline{x}$, which maximizes $\Lambda(x)$ over all feasible arcs $x \in AC(\Theta, \mathbb{R}_+)$ within an ε neighborhood of $\overline{x}$,

$$||\overline{x} - x|| \leq \varepsilon.$$

A *local minimizer* is defined analogously.

For completeness, we reproduce here Theorem 3 of Vinter and Zheng (1998) which provides necessary conditions for solutions to the following minimization program:

$$(\mathcal{P}') : \qquad \text{Minimize } J(x) \equiv \textstyle\int_{\underline{\theta}}^{\overline{\theta}} L(\theta, x(\theta), \dot{x}(\theta)) d\theta$$
$$\text{subject to } x \in AC(\Theta, \mathbb{R}) \text{ and } h(\theta, x(\theta)) \leq 0 \text{ for all } \theta \in \Theta \equiv [\underline{\theta}, \overline{\theta}].$$ [2]

Theorem 2 *(Vinter and Zheng (1998), Theorem 3) Let $\overline{x}$ be a AC local minimizer for $(\mathcal{P}')$ such that $J(\overline{x}) < +\infty$. Assume that the following hypotheses are satisfied:*

H1. $L(\cdot, x, \cdot)$ is $\mathcal{L} \times \mathcal{B}$ measureable for each x and $L(\theta, \cdot, \cdot)$ is lower semi-continuous for a.e. $\theta \in \Theta$.

H2. For every $K > 0$ there exists $\delta > 0$ and $k \in L^1$ such that

$$|L(\theta, x', v) - L(\theta, x, v)| \leq k(\theta)|x' - x|, \quad L(\theta, \overline{x}(\theta), v) \geq -k(\theta)$$

for a.e. $\theta \in \Theta$, for all $x, x' \in \overline{x}(\theta) + \delta B$ and $v \in \dot{\overline{x}}(\theta) + KB$, where B is a unit Euclidean ball.

H3. h is upper semi-continuous near $(\theta, \overline{x}(\theta))$ for all $\theta \in \Theta$, and there exists a constant k_h such that

$$|h(\theta, x') - h(\theta, x)| \leq k_h |x' - x|$$

for all $\theta \in \Theta$ and all $x', x \in \overline{x}(\theta) + \delta B$.

Then there exist an arc $p \in AC$, a constant $\lambda \geq 0$, a non-negative measure μ on the Borel subsets of Θ and a μ-integrable function $\gamma : \Theta \to \mathbb{R}$, such that

[2] We specialize their theorem to our present problem in which the range of $x(\theta)$ is one-dimensional and there is no endpoint cost function.

(i). $\lambda + \max_{\theta \in \Theta} |p(\theta)| + \int_\Theta \mu(ds) = K > 0$ *(where K is an arbitrary normalization constant),*[3]

(ii).

$$\dot{p}(\theta) \in co\left\{\eta \,|\, (\eta, p(\theta) + \int_{[\underline{\theta},\theta)} \gamma(s)\mu(ds), -\lambda) \in N_{\text{epi}\{L(\theta,\cdot,\cdot)\}}(\overline{x}(\theta), \dot{\overline{x}}(\theta), L(\theta, \overline{x}(\theta), \dot{\overline{x}}(\theta)))\right\} \text{ a.e.,}$$

(iii).

$$p(\underline{\theta}) = p(\overline{\theta}) - \int_\Theta \gamma(s)\mu(ds) = 0,$$

(iv).

$$\left(p(\theta) + \int_{[\underline{\theta},\theta)} \gamma(s)\mu(ds)\right) \cdot \dot{\overline{x}}(\theta) - \lambda L(\theta, \overline{x}(\theta), \dot{\overline{x}}(\theta)) \geq \left(p(\theta) + \int_{[\underline{\theta},\theta)} \gamma(s)\mu(ds)\right) \cdot v - \lambda L(\theta, \overline{x}(\theta), v)$$

for all $v \in \mathbb{R}$ a.e.,

(v). $\gamma(\theta) \in \partial_x^> h(\theta, \overline{x}(\theta))$ *μ-a.e. and* $\text{supp}\{\mu\} \subseteq \{t \,|\, h(\theta, \overline{x}(\theta)) = 0\}$*, where*

$$\partial_x^> h(\theta, x) \equiv \text{co}\{\lim_i \xi_i \,|\, \exists t_i \to t, x_i \to x \textit{ such that } h(\theta, x_i) > 0 \textit{ and } \xi_i \in \partial_x h(t_i, x_i) \textit{ for all } i\}.$$

We apply this result to our setting by substituting $xf(\theta) - S(\theta, v)$ in program $(\mathcal{P})$ in place of $L(\theta, x, v)$ and thereby converting the maximization functional Λ in program $(\mathcal{P})$ to the minimization functional J in program $(\mathcal{P}')$. We complete the transformation by requiring that $h(\theta, x) = -x$, and that $L(\theta, x, v)$ is a linear function of x for any (θ, v).

First, we verify that hypotheses H1-H3 are satisfied for our program $(\mathcal{P})$. Because $S(\theta, \cdot)$ is upper semi-continuous and $\mathcal{B}$-measurable, and because $L(\theta, x, v)$ is linear in x, H1 is satisfied. H2 requires that $L(\theta, \cdot, v)$ is Lipschitz continuous, which is trivial given that L is linear in x with coefficient $f(\theta)$. Because the transformed program has $h(\theta, x) = -x$, h is a continuous linear functional of x and thus H3 is also satisfied.

Next, we specialize the conclusions of Vinter and Zheng (1998) by making use of the additional restrictions on $L(\cdot)$ and $h(\cdot)$. We present this in the following Lemma.

Lemma 1 *Suppose that $L(\theta, x, v)$ is a linear function of x and that $h(\theta, x) = -x$. Then the conclusions (i)-(v) of Theorem 2 imply*

(a). $\lambda + \max_{\theta \in \Theta} |p(\theta)| + \int_\Theta \mu(ds) = K > 0,$

(b). $\dot{p}(\theta) = \lambda f(\theta)$ *a.e.,*

(c). $p(\underline{\theta}) = p(\overline{\theta}) + \int_\Theta \gamma(s)\mu(ds) = 0$

[3] We choose to state the Theorem using $K > 0$ as an arbitrary normalization rather than $K = 1$, which is the normalization chosen in Vinter and Zheng (1998). Later, by setting $K = 3$, we will succeed in normalizing μ to a probability measure which is a more familiar object.

(d). $\dot{\overline{x}}(\theta) \in \arg\max_{v\in\mathbb{R}} \left(p(\theta) + \int_{[\underline{\theta},\theta)} \gamma(s)\mu(ds)\right) \cdot v + \lambda S(\theta, v),$ *a.e.,*
(e). $\gamma(\theta) = -1$ *μ-a.e. and* $\operatorname{supp}\{\mu\} \subseteq \{t \mid h(\theta, \overline{x}(\theta)) = 0\}$.

Proof of Lemma 1: Implications (i) and (a) are identical. Implication (ii) requires almost everywhere that

$$\dot{p}(\theta) \in \operatorname{co}\left\{\eta \mid \left(\eta, p(\theta) + \int_{[S,t)} \gamma(s)\mu(ds), -\lambda\right) \in N_{\operatorname{epi}(L(\theta,\cdot,\cdot))}\left(\overline{x}, \dot{\overline{x}}, L(\theta, \overline{x}, \dot{\overline{x}})\right)\right\}.$$

Because $L(\theta, \overline{x}(\theta), \dot{\overline{x}}(\theta)) = f(\theta)\overline{x}(\theta) - S(\theta, \dot{\overline{x}}(\theta))$ is finite, the limiting normal cone in the above expression can be written as

$$\begin{aligned} &N_{\operatorname{epi}(L(\theta,\cdot,\cdot))}\left(\overline{x}, \dot{\overline{x}}, \overline{L}\right) \\ &\qquad = \left\{(\xi d_1, \xi d_2, -\xi) \mid \xi > 0,\ (d_1, d_2) \in \partial\left(f(\theta)\cdot\overline{x}(\theta) - S(\theta, \dot{\overline{x}}(\theta))\right)\right\} \\ &\qquad\qquad \bigcup \left\{\partial^\infty\left(f(\theta)\cdot\overline{x}(\theta) - S(\theta, \dot{\overline{x}}(\theta))\right) \times \{0\}\right\}. \end{aligned}$$

Using the fact that $L(\cdot)$ is additively separable in x and $\dot{x}$, a basic chain rule for lower semi-continuous functions (RW, Proposition 10.5) yields

$$\begin{aligned} \partial\left(f(\theta)\overline{x}(\theta) - S(\theta, \dot{\overline{x}}(\theta))\right) &= \partial\left(f(\theta)\overline{x}(\theta)\right) \times \partial\left(-S(\theta, \dot{\overline{x}}(\theta))\right) \\ &= \left\{f(\theta) \times \partial\left(-S(\theta, \dot{\overline{x}}(\theta))\right)\right\}, \end{aligned}$$

and

$$\begin{aligned} \partial^\infty\left(f(\theta)\overline{x}(\theta) - S(\theta, \dot{\overline{x}}(\theta))\right) &\subseteq \partial^\infty\left(f(\theta)\overline{x}(\theta)\right) \times \partial^\infty\left(-S(\theta, \dot{\overline{x}}(\theta))\right) \\ &= \left\{\{0\} \times \partial^\infty\left(-S(\theta, \dot{\overline{x}}(\theta))\right)\right\}, \end{aligned}$$

where the last equality uses the fact that a linear function is Lipschitz continuous and hence $\partial^\infty(f(\theta)\overline{x}(\theta)) = \{0\}$. Substituting these subdifferentials into the expression for the limiting normal cone, we have a simple inclusion:

$$\begin{aligned} N_{\operatorname{epi}(L(\theta,\cdot,\cdot))}\left(\overline{x}, \dot{\overline{x}}, \overline{L}\right) &\subseteq \left\{(\xi f(\theta), \xi d_2, -\xi) \mid \xi > 0,\ d_2 \in \partial\left(-S(\theta, \dot{\overline{x}}(\theta))\right)\right\} \\ &\qquad \bigcup \left\{\{0\} \times \partial^\infty\left(-S(\theta, \dot{\overline{x}}(\theta))\right) \times \{0\}\right\}. \end{aligned}$$

This simplifies yet again to the inclusion

$$\begin{aligned} &N_{\operatorname{epi}(L(\theta,\cdot,\cdot))}\left(\overline{x}, \dot{\overline{x}}, \overline{L}\right) \\ &\qquad \subseteq \left\{(\xi f(\theta), \xi d_2, -\xi) \mid \xi \geq 0,\ d_2 \in \partial\left(-S(\theta, \dot{\overline{x}}(\theta))\right) \cup \partial^\infty\left(-S(\theta, \dot{\overline{x}}(\theta))\right)\right\}. \end{aligned}$$

The key point to note is that any vector in the limiting normal cone must point in the same direction in the $(\overline{x}, \overline{L})$ plane, regardless of d_2. Returning to implication (ii), we see that any point η in the given convex hull must satisfy $(\eta, \cdot, -\lambda) = (\xi f(\theta), \cdot, -\xi)$ for some $\xi \geq 0$, and hence the convex hull reduces to $\{\lambda f(\theta)\}$. We conclude that implication (ii) simplifies to implication (b) given that $L(\cdot)$ is both additively separable and linear in x.

Implication (iii) is identical to implication (c).

Using the transformation $L(\theta, x, v) = xf(\theta) - S(\theta, v)$, implication (iv) simplifies to implication (d). Lastly, the fact that $h(\theta, x) = -x$ yields $\partial_x h(\theta, \overline{x}(\theta)) = \partial_x^{>} h(\theta, \overline{x}(\theta)) = \{-1\}$. Thus, implication (v) simplifies to $\gamma(\theta) = -1$ μ-a.e. and $supp\{\mu\} \subseteq \{t \,|\, \overline{x}(\theta) = 0\}$. This is implication (e). *Q.E.D.*

An immediate inspection of conditions (a)-(e) suggest further simplifications by combining these conditions. Conditions (b) and (c) jointly yield

$$p(\theta) = \lambda F(\theta).$$

Because $p(\overline{\theta}) = \lambda$ and $\gamma(\theta) = -1$ a.e. with respect to μ, condition (c) also implies

$$\int_\Theta \mu(ds) = \lambda.$$

Because we also have $\max_{\theta\in\Theta} |p(\theta)| = \lambda$, condition (a) implies $\lambda > 0$ and in particular $\lambda = \frac{K}{3}$. Because the choice of K is arbitrary, we choose $K = 3$ as a normalization, yielding $\lambda = 1$ and $\int_\Theta \mu(ds) = 1$. Thus, the normalization makes μ a probability measure on Θ. Defining $\overline{M}(\theta) = \int_{[\underline{\theta},\theta)} \mu(ds)$, the implication in (d) is therefore

$$\dot{\overline{x}}(\theta) \in \arg\max_{v\in\mathbb{R}} \; S(\theta, v) + \big(F(\theta) - \overline{M}(\theta)\big)\, v, \; a.e.,$$

which is condition (1.2) of Theorem 1. Lastly, the implication of (e) delivers the complementary slackness condition (1.1). We have therefore proven the necessity of the conditions in Theorem 1.

3. PROOF OF SUFFICIENCY

Sufficiency is proven by generalizing Arrow's Sufficiency Theorem to non-smooth optimal control problems and specializing the theorem to the case in which the objective integrand is a linear function of x. We adapt the argument of Arrow's Sufficiency Theorem using the basic approach of Seierstad and Sydsaeter (1987) but relaxing their continuity and smoothness assumptions. The regularity of the optimal solution follows from arguments involving the necessary conditions.

Let x be any admissible arc: $x \in AC(\Theta, \mathbb{R})$ and $x(\theta) \geq 0$ for all $\theta \in \Theta$. Define

$$\Delta = \int_\Theta \big\{\big(S(\theta, \dot{\overline{x}}(\theta)) - \overline{x}(\theta) f(\theta)\big) - \big(S(\theta, \dot{x}(\theta)) - x(\theta) f(\theta)\big)\big\}\, d\theta.$$

We will demonstrate that, under conditions (1.1) and (1.2) of Theorem 1, $\Delta \geq 0$.

To this end, it is useful to define the Hamiltonian for program (P) using $M(\theta) - F(\theta)$ as the adjoint equation which satisfies conditions(1.1) and (1.2):

$$H(\theta, x, v) \equiv S(\theta, v) - x \cdot f(\theta) - (\overline{M}(\theta) - F(\theta)) \cdot v.$$

Note that $\overline{M}(\theta)$ is defined for $\theta \in (\underline{\theta}, \overline{\theta}]$ and thus $H(\cdot)$ inherits the same domain. Nonetheless, because μ is not part of expression of Δ and F is absolutely continuous, we can ignore the point $\underline{\theta}$ in the integral and conclude that

$$\Delta = \int_{(\underline{\theta},\overline{\theta}]} \left(H(\theta,\overline{x}(\theta),\dot{\overline{x}}(\theta)) - H(\theta,x(\theta),\dot{x}(\theta))\right) d\theta + \int_{\Theta} (F(\theta) - \overline{M}(\theta)) \left(\dot{x}(\theta) - \dot{\overline{x}}(\theta)\right) d\theta.$$

Define the optimized Hamiltonian as

$$\hat{H}(\theta, x) \equiv \sup_{v \in \mathbb{R}} H(\theta, x, v).$$

Because $\overline{M}(\theta) - F(\theta)$ is bounded on $(\underline{\theta}, \overline{\theta}]$ and $S(\theta, \cdot)$ is bounded from above by assumption, we note that $\hat{H}(\cdot)$ must be finite. Condition (1.2) implies that

$$\hat{H}(\theta, \overline{x}(\theta)) = H(\theta, \overline{x}(\theta), \dot{\overline{x}}(\theta))$$

and for any admissible $x \in AC(\Theta; \mathbb{R}_+)$,

$$\hat{H}(\theta, x(\theta)) \geq H(\theta, x(\theta), \dot{x}(\theta)).$$

Combining these facts, we obtain

$$\begin{aligned} H(\theta,\overline{x}(\theta),\dot{\overline{x}}(\theta)) - H(\theta,x(\theta),\dot{x}(\theta)) &\geq \hat{H}(\theta,\overline{x}(\theta)) - \hat{H}(\theta,x(\theta)) \\ &= f(\theta)(x(\theta) - \overline{x}(\theta)). \end{aligned}$$

The last statement relies fundamentally on the linearity of $H(\cdot)$ in x. Substituting into the previous statement for Δ, we have

$$\begin{aligned} \Delta &\geq \int_{(\underline{\theta},\overline{\theta}]} f(\theta)(x(\theta) - \overline{x}(\theta))d\theta + \int_{\Theta} (F(\theta) - \overline{M}(\theta)) \left(\dot{x}(\theta) - \dot{\overline{x}}(\theta)\right) d\theta \\ &= \int_{\Theta} \left(f(\theta)(x(\theta) - \overline{x}(\theta)) + F(\theta)\left(\dot{x}(\theta) - \dot{\overline{x}}(\theta)\right)\right) d\theta - \int_{(\underline{\theta},\overline{\theta}]} \overline{M}(\theta)\left(\dot{x}(\theta) - \dot{\overline{x}}(\theta)\right) d\theta \\ &= \int_{\Theta} \frac{d}{d\theta}[F(\theta)(x(\theta) - \overline{x}(\theta))]d\theta - \int_{(\underline{\theta},\overline{\theta}]} \overline{M}(\theta)\left(\dot{x}(\theta) - \dot{\overline{x}}(\theta)\right) d\theta \\ &= (x(1) - \overline{x}(1)) - \int_{(\underline{\theta},\overline{\theta}]} \overline{M}(\theta)\left(\dot{x}(\theta) - \dot{\overline{x}}(\theta)\right) d\theta. \end{aligned}$$

It follows that $\Delta \geq 0$ if

$$(x(1) - \overline{x}(1)) - \int_{(\underline{\theta},\overline{\theta}]} M(\theta)\left(\dot{x}(\theta) - \dot{\overline{x}}(\theta)\right) d\theta \geq 0.$$

If $\overline{M}$ were absolutely continuous, we would be able to integrate the second term by parts and reach such a conclusion. Because $\overline{M}$ is possibly discontinuous, we must proceed more carefully. Note that $\overline{M}$ is non-decreasing on $(\underline{\theta}, \overline{\theta}]$ with at most a countable number of upward jump discontinuities. Furthermore, $\overline{M}$ is absolutely continuous elsewhere, allowing us to integrate by parts between any pair of discontinuities. Also note that at any such upward jump point, τ, $\overline{M}$ is left and right continuous with $\overline{M}(\tau) < \overline{M}(\tau^+)$ and (by condition (1.1)) we have $\overline{x}(\tau^+) = 0$.

Denote the set of jump discontinuities by $\{\tau_1, \tau_2, \ldots\}$, a possibly infinite set. Let $\mathcal{I}$ be the index set of τ_i. Between any two points τ_i and τ_{i+1}, we know

$$\int_{(\tau_i,\tau_{i+1}]} \overline{M}(\theta)\left(\dot{x}(\theta) - \dot{\overline{x}}(\theta)\right) d\theta$$

$$
\begin{aligned}
&= \overline{M}(\theta)(x(\theta)-\overline{x}(\theta))\big|_{t=\tau_i^+}^{\tau_{i+1}} - \int_{(\tau_i,\tau_{i+1})}(x(\theta)-\overline{x}(\theta))\mu(\theta)d\theta \\
&= \overline{M}(\tau_{i+1})(x(\tau_{i+1})-\overline{x}(\tau_{i+1})) - \overline{M}(\tau_i^+)(x(\tau_i)-\overline{x}(\tau_i)) \\
&\quad -\int_{(\tau_i,\tau_{i+1})}(x(\theta)-\overline{x}(\theta))\mu(\theta)d\theta.
\end{aligned}
$$

The second equality above uses the fact that x and $\overline{x}$ are continuous on Θ.

Define the size of the jump discontinuity at τ by $d(\tau)=\overline{M}(\tau^+)-\overline{M}(\tau)>0$. Then we may write

$$
\begin{aligned}
&\int_{(\underline{\theta},\overline{\theta}]}\overline{M}(\theta)\left(\dot{x}(\theta)-\dot{\overline{x}}(\theta)\right)d\theta \\
&=\sum_{i\in\mathcal{I}}\overline{M}(\tau_{i+1})(x(\tau_{i+1})-\overline{x}(\tau_{i+1}))-(d(\tau_i)+\overline{M}(\tau_i))(x(\tau_i)-\overline{x}(\tau_i)) \\
&-\int_{(\tau_i,\tau_{i+1})}(x(\theta)-\overline{x}(\theta))\mu(\theta)d\theta \\
&\quad =(x(1)-\overline{x}(1))-\sum_{i\in\mathcal{I}}d(\tau_i)(x(\tau_i)-\overline{x}(\tau_i))-\int_{(\tau_i,\tau_{i+1})}(x(\theta)-\overline{x}(\theta))\mu(\theta)d\theta.
\end{aligned}
$$

By complementary slackness in condition (1.1), we know $\overline{x}(\theta)\mu(\theta)=0$ and at any jump point τ we must have $\overline{x}(\tau)=0$. Thus,

$$
\int_{(\underline{\theta},\overline{\theta}]}\overline{M}(\theta)\left(\dot{x}(\theta)-\dot{\overline{x}}(\theta)\right)d\theta=(x(1)-\overline{x}(1))-\sum_{i\in\mathcal{I}}d(\tau_i)x(\tau_i)-\int_{(\tau_i,\tau_{i+1})}x(\theta)\mu(\theta)d\theta.
$$

We deduce

$$
\begin{aligned}
\Delta &\geq (x(1)-\overline{x}(1))-\int_{(\underline{\theta},\overline{\theta}]}M(\theta)\left(\dot{x}(\theta)-\dot{\overline{x}}(\theta)\right)d\theta \\
&= \sum_{i\in\mathcal{I}}d(\tau_i)x(\tau_i)+\int_{(\tau_i,\tau_{i+1})}x(\theta)\mu(\theta)d\theta.
\end{aligned}
$$

Because $x(\theta)\geq 0$, μ is a non-negative measure, and jump discontinuities $d(\tau_i)$ are positive, we conclude $\Delta\geq 0$ as claimed. We have proven that conditions (1.1) and (1.2) are sufficient for a solution.

Smoothness of the solution, $\overline{x}$: We add the hypothesis that

$$
y(\theta,\sigma)\equiv\arg\max_{v\in\mathbb{R}}\ S(\theta,v)+(F(\theta)-\sigma)v
$$

is single-valued and continuous for $(\theta,\sigma)\in\Theta\times[0,1]$. It follows that $y(\theta,\sigma)$ is non-increasing in σ and from condition (1.2), that $\dot{\overline{x}}(\theta)=q(\theta,\overline{M}(\theta))$ a.e.

Suppose to the contrary that $\dot{\overline{x}}$ is discontinuous at some point $\tau\in\Theta$. Initially, suppose that Condition (1.2) is extended to hold for all $\theta\in\Theta$ rather than for a.e. $\theta\in(\underline{\theta},\overline{\theta}]$; call this Condition (1.2'). Condition (1.2') and the additional hypothesis that $y(\theta,\sigma)$ is continuous in (θ,σ) jointly imply that $\dot{\overline{x}}(\theta)$ is discontinuous at τ only if $\overline{M}$ is also discontinuous at τ. Any discontinuity in $\overline{M}$, however, must be an upward jump, $d(\tau)=\overline{M}(\tau^+)-\overline{M}(\tau)>0$, implying that $\dot{\overline{x}}(\theta)$ must jump downwards. Complementary slackness (condition (1.1), however, imposes that $\overline{x}(\tau)=0$, with the implication that a downward discontinuity at τ would violate the state constraint $x(\theta) \geq 0$ in the neighborhood to the immediate right of τ. Hence, continuity must hold for all points $\theta \in [\underline{\theta}, \theta)$ under Condition (1.2'). Furthermore, because $\overline{M}$ is left continuous at $t = 1$, no jump in $\dot{\overline{x}}(\theta)$ is possible at this endpoint. We conclude that Condition (1.2') implies that $\dot{\overline{x}}(\theta)$ is continuous for all $\theta \in \Theta$. The weaker Condition (1.2) allows $\dot{\overline{x}}(\theta)$ to violate the maximization condition on sets of measure zero, including at $\theta = \underline{\theta}$. But such violations have no effect on the solution $\overline{x}$ which is absolutely continuous. Thus, $\overline{x}$ is smooth as posited.

REFERENCES

ALIPRANTIS, C. AND K. BORDER (1990). *Infinite Dimensional Analysis*, 2nd ed. Berlin: Springer.

GALBRAITH, G AND R. VINTER (2004). Regularity of Optimal Controls for State-Constrained Problems. *Journal of Global Optimization* **28** 305–317.

JULLIEN, B. (2000). Participation Constraints in Adverse Selection Models. *Journal of Economic Theory* **93** 1–47.

LOEWEN, P. AND T. ROCKAFELLAR (1995). New Necessary Conditions for the Generalized Bolza Problem. *SIAM Journal of Control and Oprimization* **49** 294–313.

ROCKAFELLAR, T. AND R. WETS (2004). *Variational Analysis*. Berlin: Springer.

ROYDEN, H.L. (1988). *Real Analysis*. New York: Prentice Hall.

SEIERSTAD, A. AND K. SYDSAETER (1987). *Optimal Control Theory with Economic Applications*. Amsterdam: North Holland.

VINTER, R. (2000). *Optimal Control*. Boston: Birkhauser.

VINTER, R. AND H. ZHENG (1995). Necessary Conditions for Optimal Control Problem with State Constraints. *Transactions of the American Mathematical Society* **350** 1181–1204.