Necessary and sufficient conditions for non-smooth, linear-state optimal control problems

David Martimort (Paris School of Economics)
Lars A. Stole (University of Chicago)

January 2014

We are interested in the solution to the following pure-state control program (P):

$$\text{Maximize } \Lambda(x) \equiv \int_0^1 \left(S(t, \dot{x}(t)) - x(t) \cdot f(t)\right) dt$$
$$\text{subject to } x \in W^{1,1}([0,1], \mathbb{R}),\ x(t) \geq 0 \text{ for all } t \in [0,1].$$

The constraints require that the state variable x is a non-negative, absolutely continuous real function. We say that x is **admissible** if it satisfies these constraints. Note that the integrand in $\Lambda(x)$ is a linear function in x and that the state constraint, $x(t) \geq 0$, is independent of t. These two restrictions within the class of state-constrained, non-smooth optimal control problems are the source of many of the sharp results in the analysis that follows.

We maintain the following assumptions on the data of the problem. We assume that $S(t, \cdot)$ is an upper-semi continuous function bounded from above and that $f(t)$ is a positive function that is also bounded from above with a corresponding absolutely continuous distribution function, $F(t) \equiv \int_{[0,t]} f(t)$. Without loss of generality, we normalize S and f such that $F(1) = 1$ and so we can think of F as a probability distribution. Lastly, we assume that $S(\cdot, \cdot)$ is $\mathcal{L} \times \mathcal{B}$-measurable, where $\mathcal{L}$ is the set of Lebesgue measurable subsets of $[0,1]$ and $\mathcal{B}$ is the set of Borel measurable subsets of $\mathbb{R}$.

Importantly, we do not assume *a priori* that $S(t, \cdot)$ is a continuous function. We present our main result for this class of problems.

Theorem .1. *$\bar{x}$ is a solution to program (P) if and only if $\bar{x}$ is admissible and there exists a probability measure μ defined over the Borel subsets of $[0,1]$ with distribution function*

$$M(t) \equiv \int_{[0,t)} \mu(ds)$$

such that the following two conditions are satisfied:

1. $\int_{[0,1]} \bar{x}(s) \mu(ds) = 0$,
2. $\dot{\bar{x}}(t) \in \arg\max_{v \in \mathbb{R}}\ S(t, v) + (F(t) - M(t)) \cdot v$, *for a.e.* $t \in [0,1]$.

Furthermore, if

$$y(t, \alpha) \equiv \arg\max_{v \in \mathbb{R}}\ S(t, v) + (F(t) - \alpha) \cdot v$$

is single-valued and continuous over the domain $(t, \alpha) \in [0,1]^2$, then the solution $\bar{x}$ to (P) is continuously differentiable.

Remarks:

- The statement in Theorem .1 is very similar to Theorem 1 in Jullien (2000). In both theorems, the necessary and sufficient conditions are stated in terms of a probability measure which serves as the multiplier equation for the state constraint and an optimization condition constructed from this probability measure. Moreover, both theorems use a similar condition to establish the continuity of $\dot{\bar{x}}(t)$ in the solution to (P). Jullien's Theorem, however, uses the stronger hypothesis that S is twice continuously differentiable. Our technical contribution is to weaken these hypotheses to requirements of upper semi-continuity and measurability. This generalization allows us to apply the necessary and sufficient conditions above to our class of common agency games with upper semicontinuous contract menus.

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

- The distribution $M(t)$ defined with respect to μ is nondecreasing on $(0, 1]$ and it is undefined at $t = 0$. Note in particular that the distribution M is constructed to be *left*-continuous rather than right-continuous.

- Program (P) allows for an unrestricted choice of $\dot{x}(t) \in \mathbb{R}$. With some minor technical conditions, we can generalize program (P) to require $\dot{x}(t) \in \mathcal{Q}(t)$, where $\mathcal{Q}(t)$ is a t-dependent, convex and closed "velocity" set.[1]

Proof of Theorem .1:

Overview. 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 $x(t) \geq 0$ is linear and independent of t. 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. The regularity of the optimal solution follows from arguments involving the necessary conditions. While the proof seems straightforward when viewed from this broad vantage point, a considerable investment in concepts and notation from non-smooth, non-convex analysis is required along the way.

Non-smooth analysis. We first introduce some notation from non-smooth analysis. We draw heavily from Vinter and Zheng (1998) in the following presentation. 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).

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

[1]Proposition 8 in Vinter and Zheng (1998) (reproduced as Proposition 10.4.2 in Vinter (2000)) indicates how one can easily incorporate a non-state-dependent velocity constraint set. The statement of Theorem .1 would need only be modified with this domain restriction. Indeed, the necessary conditions of the Theorem .1 apply even to a program in which the velocity set $\mathcal{Q}$ depends upon both t and $x(t)$. With state dependence, however, our sufficiency proof cannot be used and so this generalization is limited.

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

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

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 the norm on $W^{1,1}$ (the space of absolutely continuous functions) by

$$||x||_{W^{1,1}} \equiv |x(0)| + \int_{[0,1]} |\dot{x}(t)| dt.$$

A **local maximizer** of $\Lambda(x)$ is a feasible arc, $\bar{x}$, which maximizes $\Lambda(x)$ over all feasible arcs $x \in W^{1,1}([0,1], \mathbb{R}_+)$ within an ε neighborhood of $\bar{x}$,

$$||\bar{x} - x||_{W^{1,1}} \leq \varepsilon.$$

A **local minimizer** is defined analogously.

Necessity. For completeness, we state Theorem 3 of Vinter and Zheng (1998) which provides necessary conditions for solutions to the following minimization program (P'):

$$\text{Minimize } J(x) \equiv \ell(x(0), x(1)) + \textstyle\int_0^1 L(t, x(t), \dot{x}(t)) dt$$
$$\text{subject to } x \in W^{1,1}([0,1], \mathbb{R}^k) \text{ and } h(t, x(t)) \leq 0 \text{ for all } t \in [0,1].$$

We further specialize their theorem to our present problem in which the range of $x(t)$ is one-dimensional and there is no endpoint cost function, $\ell(x(0), x(1)) \equiv 0$. The modified statement of their result for $\ell = 0$ and $x(t) \in \mathbb{R}$ follows.

Theorem .2. (Vinter and Zheng (1998), Theorem 3) *Let $\bar{x}$ be a $W^{1,1}$ local minimizer for (P') such that $J(\bar{x}) < +\infty$. Assume that the following hypotheses are satisfied:*

(H1) $L(\cdot, x, \cdot)$ is $\mathcal{L} \times \mathcal{B}$ measureable for each x and $L(t, \cdot, \cdot)$ is lower semi-continuous for a.e. $t \in [0,1]$.

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

$$|L(t, x', v) - L(t, x, v)| \leq k(t)|x' - x|, \quad L(t, \bar{x}(t), v) \geq -k(t)$$

for a.e. $t \in [0,1]$, for all $x, x' \in \bar{x}(t) + \delta B$ and $v \in \dot{\bar{x}}(t) + KB$, where B is a unit Euclidean ball.

(H3) h *is upper semi-continuous near* $(t, \bar{x}(t))$ *for all* $t \in [0,1]$*, and there exists a constant* k_h *such that*

$$|h(t,x') - h(t,x)| \le k_h |x' - x|$$

for all $t \in [0,1]$ *and all* $x', x \in \bar{x}(t) + \delta B$.

Then there exist an arc $p \in W^{1,1}$*, a constant* $\lambda \ge 0$*, a non-negative measure* μ *on the Borel subsets of* $[0,1]$*, and a* μ*-integrable function* $\gamma : [0,1] \to \mathbb{R}$ *such that*

(i). $\lambda + \max_{t\in[0,1]} |p(t)| + \int_{[0,1]} \mu(ds) = 1,$

(ii).

$$\dot{p}(t) \in co\left\{\eta \,|\, (\eta, p(t) + \int_{[0,t)} \gamma(s)\mu(ds), -\lambda) \in N_{\text{epi}\{L(t,\cdot,\cdot)\}}(\bar{x}(t), \dot{\bar{x}}(t), L(t, \bar{x}(t), \dot{\bar{x}}(t)))\right\} \text{ a.e.,}$$

(iii).

$$p(0) = p(1) - \int_{[0,1]} \gamma(s)\mu(ds) = 0,$$

(iv).

$$\left(p(t) + \int_{[0,t)} \gamma(s)\mu(ds)\right) \cdot \dot{\bar{x}}(t) - \lambda L(t, \bar{x}(t), \dot{\bar{x}}(t)) \ge \left(p(t) + \int_{[0,t)} \gamma(s)\mu(ds)\right) \cdot v - \lambda L(t, \bar{x}(t), v)$$

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

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

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

We apply this result to our setting by substituting $x \cdot f(t) - S(t,v)$ in program (P) in place of $L(t,x,v)$ and thereby converting the maximization functional Λ in program (P) to the minimization functional J in program (P'). We complete the transformation by requiring that $h(t,x) = -x$, and that L is a continuous linear functional in x.

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

Next, we specialize the conclusions of Vinter and Zheng's (1998) result my making use of the additional restrictions on L and h. We present this in the following Lemma.

Lemma 1. *Suppose that* $L(t,x,v)$ *is a linear function of* x *and that* $h(t,x) = -x$*. Then the conclusions (i)-(v) of Theorem .2 imply* *(a).* $\lambda + \max_{t\in[0,1]} |p(t)| + \int_{[0,1]} \mu(ds) = 1$,

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

(c). $p(0) = p(1) + \int_{[0,1]} \gamma(s)\mu(ds) = 0$

(d). $\dot{\bar{x}}(t) \in \arg\max_{v\in\mathbb{R}} \left(p(t) + \int_{[0,t)} \gamma(s)\mu(ds)\right) \cdot v + \lambda S(t,v)$, *a.e.*,

(e). $\gamma(t) = -1$ *μ-a.e. and* $\text{supp}\{\mu\} \subseteq \{t \,|\, h(t,\bar{x}(t)) = 0\}$.

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

$$\dot{p}(t) \;\in\; \text{co}\left\{\eta \mid \left(\eta, p(t) + \int_{[S,t)} \gamma(s)\mu(ds), -\lambda\right) \in N_{\text{epi}(L(t,\cdot,\cdot))}\left(\bar{x}, \dot{\bar{x}}, L(t,\bar{x},\dot{\bar{x}})\right)\right\}, \quad a.e.$$

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

$$\begin{aligned} N_{\text{epi}(L(t,\cdot,\cdot))}\left(\bar{x},\dot{\bar{x}},\bar{L}\right) &= \{(\xi d_1, \xi d_2, -\xi) \,|\, \xi > 0,\, (d_1,d_2) \in \partial\left(f(t)\cdot\bar{x}(t) - S(t,\dot{\bar{x}}(t))\right)\} \\ &\quad \bigcup \{\partial^\infty\left(f(t)\cdot\bar{x}(t) - S(t,\dot{\bar{x}}(t))\right) \times \{0\}\}. \end{aligned}$$

Using the fact that L is additively separable in x and $\dot{x}$, a basic chain rule for lower semi-continuous functions (Rockafellar and Wets (2004), Proposition 10.5) yields

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

and

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

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

$$\begin{aligned} N_{\text{epi}(L(t,\cdot,\cdot))}\left(\bar{x},\dot{\bar{x}},\bar{L}\right) &\subseteq \{(\xi f(t), \xi d_2, -\xi) \,|\, \xi > 0,\, d_2 \in \partial\left(-S(t,\dot{\bar{x}}(t))\right)\} \\ &\quad \bigcup \{\{0\} \times \partial^\infty\left(-S(t,\dot{\bar{x}}(t))\right) \times \{0\}\}. \end{aligned}$$

This simplifies yet again to the inclusion

$$N_{\text{epi}(L(t,\cdot,\cdot))}\left(\bar{x},\dot{\bar{x}},\bar{L}\right) \subseteq \{(\xi f(t), \xi d_2, -\xi) \,|\, \xi \geq 0,\, d_2 \in \partial\left(-S(t,\dot{\bar{x}}(t))\right) \cup \partial^\infty\left(-S(t,\dot{\bar{x}}(t))\right)\}.$$

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

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

Using the transformation $L(t,x,v) = x \cdot f(t) - S(t,v)$, implication (iv) simplifies to implication (d). Lastly, the fact that $h(t,x) = -x$ yields $\partial_x h(t,\bar{x}(t)) = \partial_x^{>} h(t,\bar{x}(t)) = \{-1\}$. Thus, implication (v) simplifies to $\gamma(t) = -1$ μ-a.e. and $supp\{\mu\} \subseteq \{t \mid \bar{x}(t) = 0\}$. This is implication (e). $\square$

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

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

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

$$\int_{[0,1]} \mu(ds) = \lambda.$$

Because we also have $\max_{t\in[0,1]} |p(t)| = \lambda$, condition (a) implies $\lambda > 0$ and in particular $\lambda = \frac{1}{3}$. Because $\lambda > 0$ and the conditions (a)-(e) are linear homogenous in λ, we can arbitrarily choose $\lambda = 1$ and in the process normalize μ so that $\int_{[0,1]} \mu(ds) = 1$ and thus μ is a probability measure on $[0,1]$. Defining $M(t) = \int_{[0,t)]} \mu(ds)$, the implication in (d) is therefore

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

which is condition (2) of Theorem .1. Lastly, the implication of (e) can be summarized as the complementary slackness condition $\mu(t) \cdot \bar{x}(t) = 0$, which is equivalent to condition (1). We have therefore proven the necessity of the conditions in Theorem .1.

Sufficiency: We adapt the argument of Arrow's sufficiency theorem using the basic approach of Seirestad and Sydsaeter (1987) but relaxing their continuity and smoothness assumptions.

Let x be any admissible arc: $x \in W^{1,1}([S,T],\mathbb{R})$ and $x(t) \geq 0$ for all $t \in [0,1]$. Define

$$\Delta = \int_{[0,1]} \{(S(t,\dot{\bar{x}}(t)) - \bar{x}(t)f(t)) - (S(t,\dot{x}(t)) - x(t)f(t))\} \, dt.$$

We will demonstrate that under the conditions (1) and (2) of Theorem .1, it follows that $\Delta \geq 0$.

To this end, it is useful to define the Hamiltonian for program (P) using $M(t) - F(t)$ as the adjoint equation which satisfies conditions (1) and (2):

$$H(t,x,v) \equiv S(t,v) - x \cdot f(t) - (M(t) - F(t)) \cdot v.$$

Note that $M(t)$ is defined for $t \in (0,1]$ and thus H inherits the same domain. Nonetheless, because μ is not part of expression of Δ and F is absolutely continuous, we can ignore the point $t = 0$ in the integral and conclude that

$$\Delta = \int_{(0,1]} (H(t,\bar{x}(t),\dot{\bar{x}}(t)) - H(t,x(t),\dot{x}(t))) \, dt + \int_{[0,1]} (F(t) - M(t)) \, (\dot{x}(t) - \dot{\bar{x}}(t)) \, dt.$$

Define the optimized Hamiltonian as

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

Because $M(t)-F(t)$ is bounded on $(0,1]$ and $S(t,\cdot)$ is bounded from above by assumption, we note that $\hat{H}$ must be finite. Condition (2) implies that

$$\hat{H}(t,\bar{x}(t)) = H(t,\bar{x}(t),\dot{\bar{x}}(t))$$

and for any admissible $x \in W^{1,1}([0,1];\mathbb{R}_+)$,

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

Combining these facts, we obtain

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

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

$$\begin{aligned} \Delta &\geq \int_{(0,1]} f(t)\cdot(x(t)-\bar{x}(t))dt + \int_{[0,1]} (F(t)-M(t))\,(\dot{x}(t)-\dot{\bar{x}}(t))\,dt \\ &= \int_{[0,1]} (f(t)\cdot(x(t)-\bar{x}(t)) + F(t)\,(\dot{x}(t)-\dot{\bar{x}}(t)))\,dt - \int_{(0,1]} M(t)\,(\dot{x}(t)-\dot{\bar{x}}(t))\,dt \\ &= \int_{[0,1]} \frac{d}{dt}[F(t)(x(t)-\bar{x}(t))]dt - \int_{(0,1]} M(t)\,(\dot{x}(t)-\dot{\bar{x}}(t))\,dt \\ &= (x(1)-\bar{x}(1)) - \int_{(0,1]} M(t)\,(\dot{x}(t)-\dot{\bar{x}}(t))\,dt. \end{aligned}$$

It follows that $\Delta \geq 0$ if

$$(x(1)-\bar{x}(1)) - \int_{(0,1]} M(t)\,(\dot{x}(t)-\dot{\bar{x}}(t))\,dt \geq 0.$$

If M were absolutely continuous, we would be able to integrate the second term by parts and reach such a conclusion. Because M is possibly discontinuous, we must proceed more carefully. Note that M is nondecreasing on $(0,1]$ with at most a countable number of upward jump discontinuities. Furthermore, 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, M is left and right continuous with $M(\tau) < M(\tau^+)$ and (by condition (1)) we have $\bar{x}(\tau^+) = 0$.

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

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

The second equality above makes use of the fact that x and $\bar{x}$ are continuous on $[0, 1]$.

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

$$\begin{aligned}
&\int_{(0,1]} M(t) \left(\dot{x}(t) - \dot{\bar{x}}(t)\right) dt \\
&\qquad = \sum_{i \in \mathcal{I}} M(\tau_{i+1})(x(\tau_{i+1}) - \bar{x}(\tau_{i+1})) - (d(\tau_i) + M(\tau_i))(x(\tau_i) - \bar{x}(\tau_i)) \\
&\qquad\qquad - \int_{(\tau_i, \tau_{i+1})} (x(t) - \bar{x}(t))\mu(t) dt \\
&\qquad = (x(1) - \bar{x}(1)) - \sum_{i \in \mathcal{I}} d(\tau_i)(x(\tau_i) - \bar{x}(\tau_i)) - \int_{(\tau_i, \tau_{i+1})} (x(t) - \bar{x}(t))\mu(t) dt.
\end{aligned}$$

By complementary slackness in condition (1), we know $\bar{x}(t)\mu(t) = 0$ and at any jump point τ we must have $\bar{x}(\tau) = 0$. Thus,

$$\int_{(0,1]} M(t) \left(\dot{x}(t) - \dot{\bar{x}}(t)\right) dt = (x(1) - \bar{x}(1)) - \sum_{i \in \mathcal{I}} d(\tau_i) x(\tau_i) - \int_{(\tau_i, \tau_{i+1})} x(t)\mu(t) dt.$$

We deduce

$$\begin{aligned}
\Delta &\geq (x(1) - \bar{x}(1)) - \int_{(0,1]} M(t) \left(\dot{x}(t) - \dot{\bar{x}}(t)\right) dt \\
&= \sum_{i \in \mathcal{I}} d(\tau_i) x(\tau_i) + \int_{(\tau_i, \tau_{i+1})} x(t)\mu(t) dt.
\end{aligned}$$

Because $x(t) \geq 0$, 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) and (2) are sufficient for a solution.

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

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

is single-valued and continuous for $(t, \alpha) \in [0, 1]^2$. It follows that $y(t, \alpha)$ is nonincreasing in α and from condition (2), that $\dot{\bar{x}}(t) = q(t, M(t))$ a.e.

Suppose to the contrary that $\dot{\bar{x}}$ is discontinuous at some point $\tau \in [0, 1]$. Initially, suppose that Condition (2) is extended to hold for all $t \in [0, 1]$ rather than for a.e. $t \in (0, 1]$; call this Condition (2'). Condition (2') and the additional hypothesis that $y(t, \alpha)$ is continuous in (t, α) jointly imply that $\dot{\bar{x}}(t)$ is discontinuous at τ only if M is also discontinuous at τ. Any discontinuity in M, however, must be an upward jump, $d(\tau) = M(\tau^+) - M(\tau) > 0$, implying that $\dot{\bar{x}}(t)$ must jump downwards. Complementary slackness (condition (3)), however, imposes that $\bar{x}(\tau) = 0$, with the implication that a downward discontinuity at τ would violate the state constraint $x(t) \geq 0$ in the neighborhood to the immediate right of τ. Hence, continuity must hold for all points $t \in [0, 1)$ under Condition (2'). Furthermore, because M is left continuous at $t = 1$, no jump in $\dot{\bar{x}}(t)$ is possible at this endpoint. We conclude that Condition (2') implies that $\dot{\bar{x}}(t)$ is continuous for all $t \in [0, 1]$. The weaker Condition (2) allows $\dot{\bar{x}}(t)$ to violate the maximization condition on sets of measure zero, including at $t = 0$. But such violations have no effect on the solution $\bar{x}$ which is absolutely continuous. Thus, $\bar{x}$ is smooth as posited. $\square$