

Staticization, Its Dynamic Program and Solution Propagation

William M. McEneaney^a, Peter M. Dower^b

^a*Dept. of Mechanical & Aerospace Eng., University of California San Diego, La Jolla, CA 92093, USA.*

^b*Department of Electrical & Electronic Engineering, University of Melbourne, Victoria 3010, Australia.*

Abstract

Stationary-action formulations of dynamical systems are considered. Use of stationary-action formulations allow one to generate fundamental solutions for classes of two-point boundary-value problems (TPBVPs). One solves for stationary points of the payoff as a function of inputs rather than minimization/maximization, a task which is significantly different from that in optimal control problems. Both a dynamic programming principle (DPP) and a Hamilton-Jacobi partial differential equation (HJ PDE) are obtained for a class of problems subsuming the stationary-action formulation. Although convexity (or concavity) of the payoff may be lost as one propagates forward, stationary points continue to exist, and one must be able to use the DPP and/or HJ PDE to solve forward to such time horizons. In linear/quadratic models, this leads to a requirement for propagation of solutions of differential Riccati equations past finite escape times. Such propagation is also required in (nonlinear) n -body problem formulations where the potential is represented via semiconvex duality. The dynamic programming tools developed here are applicable.

Key words. dynamic programming, stationary action, two-point boundary value problem, Hamilton-Jacobi, optimal control.

1 Introduction

The classical approach to solution of energy-conserving dynamical systems is integration of Newton's second law. An alternative viewpoint is that a system evolves along a path which makes the action functional stationary, i.e., such that the first-order differential around the path is the zero element. This latter viewpoint appears particularly useful in some applications in modern physics, including gravitational systems where relativistic effects are non-negligible and systems in the quantum domain (cf. [9–11,17]). Our interests are more pedestrian; the stationary-action formulation has recently been found to be quite useful for generation of fundamental solutions to two-point boundary-value problems (TPBVPs) for conservative dynamical systems. For sufficiently short time horizons, stationarity of the action typically corresponds to minimization of the action. That is, the stationary point is a global minimum of that action (cf., [6,14,15]). For longer time horizons, the stationary point is more typically a saddle.

As our motivating interest is in solution of TPBVPs for conservative dynamical systems, we note that this specifically includes mass-spring, wave equation and n -body problems [6,14,15]. By appending a min-plus delta function terminal cost to the action functional, we obtain a fundamental solution object for such TPBVPs. Min-plus convolutions of this object with functionals associated to specific terminal conditions yield the solutions of the specific TPBVPs. As a change in the boundary data only requires convolution with a different functional, our object may best be termed a fundamental solution for TPBVPs, corresponding to the given time horizon. It is worth remarking that, further, one can populate the fundamental solution semigroup by convolving the fundamental solution with itself, enabling solution of the TPBVP for all strictly positive horizons.

As noted above, for sufficiently short time horizons, one may obtain the stationary action solution by minimization of the action functional, in which case it is obvious that the fundamental solution is derived from the value function for an optimal control formulation. However, for longer horizons, we must find the stationary point, and this requires a new set of tools. We define stationarity and value for such problems. Surprisingly, for a specific class of terminal costs, one may obtain a dynamic programming principle (DPP) for stationarity, where this is

* Research partially supported by NSF, AFOSR and the Australian Research Council.

Email addresses: wmceneaney@ucsd.edu (William M. McEneaney), pdower@unimelb.edu.au (Peter M. Dower).

directly analogous to standard DPPs (for optimization). We do not look for the absolute weakest assumptions, but only a reasonable first-foray set. We also formally write the corresponding Hamilton-Jacobi partial differential equation (HJ PDE), and then obtain a verification result, which is also quite similar to that found for classical optimal control problems. We remark that a verification result implies that any solution of the HJ PDE in the specified class must be the value function. This validates the approach of solving a stationarity problem (and hence the related TPBVP if the stationarity problem is generated by such) by solving the associated HJ PDE problem.

In the mass-spring case (which appears in Section 3 as a motivating example), the stationary-action problem is linear-quadratic, and the HJ PDE reduces to a differential Riccati equation (DRE). By the above-noted verification result, we see that solution of the DRE yields solution of the stationarity problem, and any corresponding TPBVP. The wave equation [6] also yields a DRE, albeit infinite dimensional. The n -body problem may be reduced to a parameterized set of time-dependent DREs [14,15]. We see that in all cases, solutions of DREs form a critical building block. Of course, DREs can exhibit finite escape times, and do so in these cases. In classical optimal control, one is not interested in propagation of the solution past such escape times. However, in stationarity problems, these may correspond to points where one loses convexity [concavity] of the payoff. Although the minimum [maximum] may go to $-\infty$ [$+\infty$], the stationary value may be well-defined and finite past such asymptotes, and one must propagate the solution beyond them. The DPP yields a means for propagation through escape times, and this will be indicated.

Although stationary action is the motivating problem class, the theory developed below is applicable to wider classes of problems, where one is seeking a stationary point. An obvious example is that of certain differential games. Extensions to stochastic cases appear possible as well, but are not considered here.

Section 2 contains relevant definitions. Section 3 presents a simple mass-spring TPBVP motivating example. Section 4 contains the main results – the DPP and HJ PDE verification theorem. Section 5 reduces to the linear/quadratic case, and indicates a means for propagation of DREs past escape times. Section 6 very briefly indicates some application areas.

2 Stationarity definitions

Recall that we are seeking stationary points of payoffs, which is unusual in comparison to the standard classes of problems in optimization. In analogy with the language for minimization and maximization, we will refer to the search for stationary points as staticization, with these

points being statica (in analogy with minima/maxima) and a single such point being a staticum (in analogy with minimum/maximum). Prior to the development, we make the following definitions. Suppose $\mathcal{Y}$ is a generic normed vector space with $\mathcal{G}_{\mathcal{Y}} \subseteq \mathcal{Y}$, and suppose $F : \mathcal{G}_{\mathcal{Y}} \rightarrow \mathbb{R}$. We say $\bar{y} \in \text{argstat}\{F(y) \mid y \in \mathcal{G}_{\mathcal{Y}}\}$ if $\bar{y} \in \mathcal{G}_{\mathcal{Y}}$ and either

$$\limsup_{y \rightarrow \bar{y}, y \in \mathcal{G}_{\mathcal{Y}} \setminus \{\bar{y}\}} \frac{|F(y) - F(\bar{y})|}{|y - \bar{y}|} = 0, \quad (1)$$

or there exists $\delta > 0$ such that $\mathcal{G}_{\mathcal{Y}} \cap B_{\delta}(\bar{y}) = \{\bar{y}\}$ (where $B_{\delta}(\bar{y})$ denotes the ball of radius δ around $\bar{y}$). If $\text{argstat}\{F(y) \mid y \in \mathcal{G}_{\mathcal{Y}}\} \neq \emptyset$, we define

$$\begin{aligned} \text{stat}_{y \in \mathcal{G}_{\mathcal{Y}}} F(y) &\doteq \text{stat}\{F(y) \mid y \in \mathcal{G}_{\mathcal{Y}}\} \\ &\doteq \{F(\bar{y}) \mid \bar{y} \in \text{argstat}\{F(y) \mid y \in \mathcal{G}_{\mathcal{Y}}\}\}. \end{aligned} \quad (2)$$

If $\text{argstat}\{F(y) \mid y \in \mathcal{G}_{\mathcal{Y}}\} = \emptyset$, $\text{stat}_{y \in \mathcal{G}_{\mathcal{Y}}} F(y)$ is undefined. Throughout, we will abuse notation by writing $\bar{y} = \text{argstat}\{F(y) \mid y \in \mathcal{G}_{\mathcal{Y}}\}$ in the event that the argstat is the single point, $\{\bar{y}\}$, and similarly for stat .

In the case where $\mathcal{Y}$ is a Hilbert space, and $\mathcal{G}_{\mathcal{Y}} \subseteq \mathcal{Y}$ is an open set, $F : \mathcal{G}_{\mathcal{Y}} \rightarrow \mathbb{R}$ is Fréchet differentiable at $\bar{y} \in \mathcal{G}_{\mathcal{Y}}$ with Fréchet derivative $F_y(\bar{y}) \in \mathcal{Y}$ if

$$\lim_{v \rightarrow 0, \bar{y} + v \in \mathcal{G}_{\mathcal{Y}} \setminus \{\bar{y}\}} \frac{|F(\bar{y} + v) - F(\bar{y}) - \langle F_y(\bar{y}), v \rangle|}{|v|} = 0. \quad (3)$$

The following is immediate from the above definitions.

Lemma 1 *Suppose $\mathcal{Y}$ is a Hilbert space, with open set $\mathcal{G}_{\mathcal{Y}} \subseteq \mathcal{Y}$ and $\bar{y} \in \mathcal{G}_{\mathcal{Y}}$. Then, $\bar{y} \in \text{argstat}\{F(y) \mid y \in \mathcal{G}_{\mathcal{Y}}\}$ if and only if $F_y(\bar{y}) = 0$.*

3 Motivational examples

As indicated in the introduction, an important problem class which motivates this effort is that of TPBVPs for conservative systems.

3.1 A simple mass-spring problem

We first examine the classic one-dimensional mass-spring example in a substantial detail in order to provide motivation and insight. Although the problem is essentially trivial, it provides a nice means for obtaining a sense of the stationary action principle as a tool for understanding system dynamics and TPBVPs. Further, as remarked above, the stationary action viewpoint is the accepted viewpoint in modern physics (cf., [9–11,17]), and as such, it will be ultimately necessary for advanced applications. It also provides exceptional computational advantages for difficult classes of problems, such as TPBVPs in the gravitational n -body case [14,15].

Remark 2 *Although the mass-spring model has an analytically solvable form due to the quadratic potential, this potential is not physically reasonable (the potential approaches $+\infty$ as $|x| \rightarrow \infty$), and induces degeneracies, particularly at half-period times. Nonetheless, it is useful for building intuition.*

Consider the mass-spring problem with mass, m , and spring-constant, K (typically given as $\dot{\xi} = -(K/m)\xi$). The associated stationary action TPBVP payoff, $J^\infty : \hat{T} \times \mathbb{R} \times \mathcal{U}_\infty \times \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}$ with $\hat{T} \doteq \{(s, t) \in \mathbb{R}^2 \mid 0 \leq s \leq t < \infty\}$ and $\mathcal{U}_\infty \doteq \mathcal{L}_2^{loc}(0, \infty)$, is given by

$$J^\infty(s, t, x, u, z) = \int_s^t \frac{m}{2} u^2(r) - \frac{K}{2} \xi^2(r) dr + \psi^\infty(\xi(t), z), \quad (4)$$

$$\text{where } \dot{\xi}(r) = u(r), \quad r \in (s, t), \quad \xi(s) = x, \quad (5)$$

$$\psi^\infty(x, z) \doteq \begin{cases} 0 & \text{if } x = z, \\ +\infty & \text{otherwise.} \end{cases} \quad (6)$$

Solution of this stationary-action problem will yield solution of the TPBVP given by dynamics $m\dot{\xi}(r) = -\nabla V(\xi(r))$ with initial position $x \in \mathbb{R}^n$ and terminal position $z \in \mathbb{R}^n$ for the given duration t and given potential function [14,15], where in this example the potential is $V(x) = (K/2)x^2$.

The stationary action solution, u^* , is such that $J_u^\infty(s, t, x, u^*, z) = 0$, where J_u^∞ denotes the Fréchet derivative of J^∞ with respect to u as per (3). Here, we take $K = m = 1$. In [14,15], one notes that if $t - s < \pi/2$, then J^∞ is strictly convex in u , and defining $W^\infty(t - s, x, z) \doteq \inf_{u \in \mathcal{U}_\infty} J^\infty(s, t, x, u, z)$, one finds

$$W^\infty(t - s, x, z) = \frac{1}{2} [P(t)x^2 + 2Q(t)xz + R(t)z^2], \quad (7)$$

where, formally, $P(s) = R(s) = -Q(s) = +\infty$, $\dot{P} = -1 - P^2$, $\dot{Q} = -PQ$, and $\dot{R} = -Q^2$. Letting

$$\psi^c(x, z) \doteq \frac{c}{2} |x - z|^2 \quad (8)$$

for $c \in (0, \infty)$, and noting that $\psi^c \rightarrow \psi^\infty$ as $c \rightarrow \infty$, one may show that the solution does, in fact, have this form, and further that the solution is given by [14,15]

$$P(t) = R(t) = \cot(t - s), \quad Q(t) = -1/\sin(t - s), \quad (9)$$

which is guaranteed by the aforementioned convexity to be valid on at least $t - s \in (0, \pi/2)$. From here, one may obtain the control and state trajectories solving the TPBVP, which are given, in the case $s = 0$ as $u^*(r) = P(t - r)\xi^*(r) + Q(t - r)z = \frac{\cos(t - r)}{\sin(t - r)}\xi^*(r) - \frac{1}{\sin(t - r)}z$, and $\xi^*(r) = x \cos(r) + \frac{z - x \cos(t)}{\sin(t)} \sin(r)$. However, at $t = \pi/2$, one loses convexity of J^∞ in u , and one must seek a staticum rather than a minimum.

One method for extending past $\pi/2$ to the stationary-over- u case is to break the interval into multiple segments of duration less than $\pi/2$, and then concatenate these. Suppose we wish to find the solution to TPBVPs for this mass-spring example with duration $3\pi/4$. As an illustration of this approach, let us break the interval up into two halves, where the payoff is then convex on each half-interval. Suppose we introduce an intermediate point, $\zeta \in \mathbb{R}$. Then, the stationary action problem with $s = 0$ is given by

$$\begin{aligned} W^\infty(3\pi/4, x, z) &\doteq \text{stat}_{u \in \mathcal{L}_2(0, \frac{3\pi}{4})} J^\infty(0, 3\pi/4, x, u, z) \\ &= \text{stat}_{\zeta \in \mathbb{R}} \left\{ \text{stat}_{u \in \mathcal{L}_2(0, \frac{3\pi}{8})} J^\infty(0, 3\pi/8, x, u, \zeta) \right. \\ &\quad \left. + \text{stat}_{u \in \mathcal{L}_2(\frac{3\pi}{8}, \frac{3\pi}{4})} J^\infty(3\pi/8, 3\pi/4, \zeta, u, z) \right\} \\ &= \text{stat}_{\zeta \in \mathbb{R}} \left\{ \text{stat}_{u \in \mathcal{L}_2(0, \frac{3\pi}{8})} J^\infty(0, 3\pi/8, x, u, \zeta) \right. \\ &\quad \left. + \text{stat}_{u \in \mathcal{L}_2(0, \frac{3\pi}{8})} J^\infty(0, 3\pi/8, \zeta, u, z) \right\}, \end{aligned}$$

where the second equality follows from time-invariance. Next, noting the convexity of each subsegment,

$$\begin{aligned} W^\infty(3\pi/4, x, z) &= \text{stat}_{\zeta \in \mathbb{R}} \left\{ W^\infty(3\pi/8, x, \zeta) \right. \\ &\quad \left. + W^\infty(3\pi/8, \zeta, z) \right\} \\ &= \frac{1}{2} \text{stat}_{\zeta \in \mathbb{R}} \left\{ P(3\pi/8)x^2 + 2Q(3\pi/8)x\zeta + R(3\pi/8)\zeta^2 \right. \\ &\quad \left. + P(3\pi/8)\zeta^2 + 2Q(3\pi/8)\zeta z + R(3\pi/8)z^2 \right\}, \end{aligned}$$

and setting the derivative with respect to ζ equal to zero, we find this is

$$\begin{aligned} &= \frac{1}{2} \left\{ [P(3\pi/8) - Q^2(3\pi/8)/(P(3\pi/8) + R(3\pi/8))] \right. \\ &\quad \cdot (x^2 + z^2) \\ &\quad \left. - 2Q^2(3\pi/8)/(P(3\pi/8) + R(3\pi/8))xz \right\}, \end{aligned}$$

which by (9) and trigonometric identities,

$$\begin{aligned} &= \frac{1}{2} \left\{ \cot(3\pi/4)(x^2 + z^2) - (2/\sin(3\pi/4))xz \right\} \\ &= \frac{1}{2} \left\{ P(3\pi/4)x^2 + 2Q(3\pi/4)xz + R(3\pi/4)z^2 \right\}. \quad (10) \end{aligned}$$

That is, surprisingly, one finds that the solution is identical to that which one would obtain if one had naively propagated the analytical solution of the DRE forward. In fact, introducing additional intermediary points for increasingly long time durations, one continues to find that the solution is identical to that found by naïve application of the above analytical solution (7),(9) for arbitrarily long periods. We will see that this is not a coincidence, and solution of stationary action problems may be obtained in the linear-quadratic case by proper propagation of the DRE past escape times.

3.2 Other example problem classes

Although the motivational mass-spring example is sufficiently simple to be presented in near entirety, staticization tools are applicable to a wide variety of problems, specifically including our main motivating class: TPBVPs for conservative systems. As indicated in the introduction, one may solve such problems by *minimizing* the action functional if the time horizon is sufficiently short. For general system durations, it is necessary to develop the staticization method so as to find stationary points of the action functional. Major problem subclasses include the wave equation and the n -body problem. In both cases, least action has been applied to obtain fundamental solutions of TPBVPs, and this is being extended to arbitrarily long-duration problems with the aid of the theory developed here.

Least action for the wave equation is discussed in [4–6]. One TPBVP in this class, with underlying domain $[0, L]$, $L > 0$, and known initial/terminal position data, is expressed as

$$\ddot{x} = -\mathcal{A}x, \quad (11)$$

$$x(0, \cdot) = y(\cdot), \quad x(t, \cdot) = z(\cdot), \quad t > 0, \quad (12)$$

where $\mathcal{A}$ is a positive, self-adjoint, unbounded linear operator densely defined on $\mathcal{X} = \mathcal{L}_2[0, L]$ by

$$\mathcal{A}y = -\partial^2 y, \quad y \in \text{dom}(\mathcal{A}), \quad (13)$$

$$\text{dom}(\mathcal{A}) = \mathcal{X}_0 = \left\{ y \in \mathcal{X} \left| \begin{array}{l} y, \partial y \text{ abs. cts.}, \partial^2 y \in \mathcal{X}, \\ y(0) = 0 = y(L) \end{array} \right. \right\},$$

in which ∂ is the (spatial) differentiation operation.

A fundamental solution for TPBVP (11), (12) is constructed in [4–6]. For short time horizons $t > 0$, this construction proceeds via the optimal control problem

$$W^{\infty, \mu}(t, y, z) = \inf_{u \in \mathcal{U}} \left\{ \int_0^t T^\mu(u(r)) - V(\xi(r)) dr + \psi^\infty(\xi(t), z) \right\}, \quad (14)$$

$$\dot{\xi}(r) = u(r), \quad \xi(0) = y, \quad y, z \in \mathcal{X}_{\frac{1}{2}}, \quad \mu \in \mathbb{R},$$

$$T^\mu(u) = \frac{1}{2} \|u\|^2 + \frac{\mu^2}{2} \|u\|_{\frac{1}{2}}^2, \quad V(y) = \frac{1}{2} \|y\|_{\frac{1}{2}}^2,$$

$$\langle y, z \rangle_{\frac{1}{2}} = \langle \mathcal{A}^{\frac{1}{2}} y, \mathcal{A}^{\frac{1}{2}} z \rangle, \quad \|y\|_{\frac{1}{2}}^2 = \langle y, y \rangle_{\frac{1}{2}},$$

in which the controlled state $\xi(r)$ evolves in a Hilbert space $\mathcal{X}_{\frac{1}{2}} = \text{dom}(\mathcal{A}^{\frac{1}{2}})$ as a consequence of an applied input $u \in \mathcal{U} = \mathcal{L}_2([0, t]; \mathcal{X}_{\frac{1}{2}})$. Here, $\mathcal{X}_{\frac{1}{2}}$ is equipped with an inner product $\langle \cdot, \cdot \rangle_{\frac{1}{2}}$ that is defined with respect to the unique, positive, self-adjoint, unbounded square-root $\mathcal{A}^{\frac{1}{2}}$ of $\mathcal{A}$, and the standard inner product $\langle \cdot, \cdot \rangle$ on $\mathcal{X}$, as indicated above.

The Hamiltonian associated with the optimal control problem (14) defines a pair of characteristic equations that together take the abstract Cauchy form [2]

$$\begin{pmatrix} \dot{\xi}(r) \\ \dot{\pi}(r) \end{pmatrix} = \mathcal{A}_\mu \begin{pmatrix} \xi(r) \\ \pi(r) \end{pmatrix}, \quad r \in [0, t], \quad (15)$$

in which $\mathcal{A}_\mu \in \mathcal{L}(\mathcal{X}_{\frac{1}{2}} \oplus \mathcal{X})$ is known. The uniformly continuous semigroup of bounded linear operators generated by $\mathcal{A}_\mu$ defines approximations for all solutions of the wave dynamics (11). In particular, given any $r \in [0, t]$, the corresponding element $\mathcal{T}_\mu(r) \in \mathcal{L}(\mathcal{X}_{\frac{1}{2}} \oplus \mathcal{X})$ of this semigroup converges strongly as $\mu \rightarrow 0$ to an operator $\mathcal{T}_0(r) \in \mathcal{L}(\mathcal{X}_{\frac{1}{2}} \oplus \mathcal{X})$ that describes precisely all solutions of (11) on $[0, r]$. Consequently, as any solution of TPBVP (11), (12) can be represented in terms of $\mathcal{T}_0(t)$, the optimal control problem (14) can be used to construct any TPBVP solution via the corresponding characteristics (15), and the aforementioned strong convergence property of the associated semigroup. That is, this construction yields a fundamental solution semigroup for TPBVP (11), (12), see [4–6].

For longer time horizons, convexity of the cost in (14) is lost, rendering the optimal control problem undefined. However, the fundamental solution semigroup, constructed as indicated above for a sufficiently short horizon, may be propagated to arbitrarily long time horizons without modification. This propagation is enabled by a weakening of the least-action approach described to one of stationary-action, again see [4–6].

TPBVPs in the class of n -body problems are discussed in [12, 14, 15]. In [12], the problem of one small body (e.g., a spacecraft or asteroid) moving among $n - 1$ large bodies on predetermined orbits is considered, while in [14, 15], the general case is discussed. A typical problem in this class is given as follows. Suppose the set of positions of the n bodies at any time $r \in [0, t]$ is denoted by $\bar{\xi}(r) = (\xi^1(r), \xi^2(r), \dots, \xi^n(r)) \in \mathbb{R}^{3n}$. Suppose the initial and terminal positions are specified as $\bar{\xi}(0) = x \in \mathbb{R}^{3n}$ and $\bar{\xi}(t) = z \in \mathbb{R}^{3n}$. Classically, the problem is given as

$$\ddot{\xi}^i(r) = \sum_{j \neq i} -Gm_j \frac{\xi^j(r) - \xi^i(r)}{|\xi^j(r) - \xi^i(r)|^3}, \quad i \in \{1, 2, \dots, n\}$$

$$\bar{\xi}(0) = x, \quad \bar{\xi}(t) = z.$$

Alternatively, the least-action formulation for the fundamental solution is given as [14, 15]

$$W^\infty(t, x, z) = \max_{\alpha \in \mathcal{A}} \min_{u \in \mathcal{U}} \left\{ \int_0^t T(u(r)) - V^\alpha(r, \bar{\xi}(r)) dr + \psi^\infty(\bar{\xi}(t), z) \right\},$$

$$\dot{\bar{\xi}}(r) = u(r), \quad \bar{\xi}(0) = x,$$

$$V^\alpha(r, x) = - \sum_{j>i} \mu_{i,j} \alpha_{i,j}(r) \left[1 - \frac{(\alpha_{i,j}(r)|x^i - x^j|)^2}{2} \right],$$

where $\mu_{i,j} = (3/2)^{(3/2)} G m_i m_j$, $T(v) = \sum_{i=1}^n m_i |v^i|^2$, and $\mathcal{U}$ and $\mathcal{A}$ are appropriately dimensioned L_2 and L_∞ spaces, respectively. Applying a Riccati differential equation method as an aid for solving the inner (minimizing) player subproblem, one finds that there exists a convex, finite-dimensional set, Ω , such that for any x, z , the value, $W^\infty(t, x, z)$ is given by the maximization of an affine functional over Ω . That is, the set Ω functions as a fundamental solution. A numerical example in the case of a small body moving among a set of large bodies appears in [12]. There, the staticization was only performed over intermediary points on the trajectories, where the subsegments were sufficiently short to allow for least-action to be used in place of stationary action. The theory developed here will allow for a stationary-action approach, and hence simpler analysis of arbitrary-duration TPBVPs in astrodynamics.

4 Dynamic programming for staticization

In this section, we will obtain a dynamic programming principle as well as a verification result for the appropriate Hamilton-Jacobi PDE.

Remark 3 For clarity, we provide several definitions. Let $\mathcal{Y}, \mathcal{Z}$ be Hilbert spaces, and let $F : \mathcal{Y} \times \mathcal{Z} \rightarrow \mathbb{R}$. We say F is strictly uniformly convex on $\mathcal{Y}$ with respect to $z \in \mathcal{Z}$, if there exists $C_F > 0$ such that for all $y, v \in \mathcal{Y}$, $z \in \mathcal{Z}$, $|F(y - v, z) - 2F(y, z) + F(y + v, z)| \geq C_F |v|^2$. We say F is coercive on $\mathcal{Y}$ if given $\bar{R}, M < \infty$, there exists $\hat{R} < \infty$ such that $F(y, z) \geq F(0, z) + M$ for all $|y| \geq \hat{R}$, $|z| \leq \bar{R}$. We say F has bounded second-order differences, if there exists $M_F < \infty$ such that

$$|F(y + \delta_1^y + \delta_2^y, z + \delta_1^z + \delta_2^z) - F(y + \delta_1^y, z + \delta_1^z) - F(y + \delta_2^y, z + \delta_2^z) + F(y, z)| \leq M_F [|\delta_1|^2 + |\delta_2|^2]$$

for all $z, \delta_1^z, \delta_2^z \in \mathcal{Z}$, and all $y, \delta_1^y, \delta_2^y \in \mathcal{Y}$ where $\delta_k \doteq (\delta_k^y, \delta_k^z)$ for $k = 1, 2$.

We assume the time-invariant dynamics are given as

$$\dot{\xi}(r) = f(\xi(r), u(r)), \quad \xi(s) = x \in \mathbb{R}^n,$$

where $f \in C^2(\mathbb{R}^n \times \mathbb{R}^m; \mathbb{R}^n)$ is globally Lipschitz. For $(s, t) \in \hat{T}$, let $\mathcal{U}_{s,t} \doteq \mathcal{L}_2([s, t]; \mathbb{R}^m)$. The inclusion of the left endpoint in the domain of elements of $\mathcal{U}_{s,t}$, albeit a set of measure zero, will be helpful in definitions and discussions below. Then, for $x, z \in \mathbb{R}^n$ and $u \in \mathcal{U}_{s,t}$, we consider payoff $J : \hat{T} \times \mathbb{R}^n \times \mathcal{U}_{s,t} \times \mathbb{R}^n \rightarrow \mathbb{R}$ given by

$$J(s, t, x, u, z) \doteq \int_s^t L(\xi(r), u(r)) dr + \psi(\xi(t), z), \quad (16)$$

where $L \in C^2(\mathbb{R}^n \times \mathbb{R}^m)$, $\psi \in C^2(\mathbb{R}^n \times \mathbb{R}^n)$ with ψ convex and coercive in x . We also allow $\psi = \psi^\infty$ given by (6), generalized to domain $\mathbb{R}^n$. Define

$$W(t - s, x, z) = \operatorname{stat}_{u \in \mathcal{U}_{s,t}} J(s, t, x, u, z). \quad (17)$$

Conditions guaranteeing the existence of W for specific classes of problems, particularly n -body problems, are given in [14,15]. (These conditions were *not* satisfied by the mass-spring example.) The next step will be to obtain the DPP for W .

We make the following assumptions.

For any $(s, t) \in \hat{T}$ and $x, z \in \mathbb{R}^n$, there exists a unique staticum of $J(s, t, x, \cdot, z)$, $u^* = u^*(s, t, x, z)$, that is, $\{u^*\} = \operatorname{argstat}_{u \in \mathcal{U}_{s,t}} J(s, t, x, u, z)$. (A.1)

Given $x, z \in \mathbb{R}^n$, there exists $\delta^c > 0$ such that if $0 \leq s \leq t < s + \delta^c$, then for any $\zeta \in \mathbb{R}^n$, $J(s, t, \zeta, \cdot, z)$ is strictly uniformly convex on $\mathcal{U}_{s,t}$ with respect to $\zeta \in \mathbb{R}^n$ and coercive on $\mathcal{U}_{s,t}$. (A.2)

Remark 4 We note that Assumption (A.2) is motivated by the fact that for the standard stationary-action problems (cf., [14,15] and the references therein), the action is strictly uniformly convex and coercive over sufficiently short time-horizons. Moreover, the stationary-action model is

$$\dot{\xi}(r) = u(r), \quad \xi(s) = x \in \mathbb{R}^n,$$

$$\psi(x, z) = \psi^\infty(x, z) \text{ and } L(x, v) = T(v) - V(x),$$

where $T(u(r)) = T(\dot{\xi}(r))$ denotes the kinetic energy at time $r \in [s, t]$, and $V(\xi(r))$ denotes the potential energy.

Let $(s, t) \in \hat{T}$ and $x, z \in \mathbb{R}^n$. For $u^0 \in \mathcal{U}_{0,s}$ and $u^1 \in \mathcal{U}_{s,t}$, let

$$J_d(s, t, x, u^0, u^1, z) \doteq \int_0^s L(\xi^0(r), u^0(r)) dr + \int_s^t L(\xi^1(r), u^1(r)) dr + \psi(\xi^1(t), z), \quad (18)$$

where $\xi^0(r) = u^0(r)$ on $(0, s)$ with $\xi^0(0) = x$ and $\xi^1(r) = u^1(r)$ on (s, t) with $\xi^1(s) = \xi^0(s^-)$. Define the concatenation $\mathcal{C} : \mathcal{U}_{0,s} \times \mathcal{U}_{s,t} \rightarrow \mathcal{U}_{0,t}$ by

$$[\mathcal{C}(u^0, u^1)](r) \doteq \begin{cases} u^0(r) & \text{if } r \in [0, s), \\ u^1(r) & \text{if } r \in [s, t), \end{cases}$$

where we suppress the detail that these are equivalence classes of functions equal almost everywhere. One immediately sees that, letting $u^0 \in \mathcal{U}_{0,s}$, $u^1 \in \mathcal{U}_{s,t}$ and

$\bar{u} \doteq \mathcal{C}(u^0, u^1)$, one has $\bar{u} \in \mathcal{U}_{0,t}$ and

$$J(0, t, x, \bar{u}, z) = J_d(s, t, x, u^0, u^1, z). \quad (19)$$

Analogously, letting $\bar{u} \in \mathcal{U}_{0,t}$, $u^0(r) = \bar{u}(r)$ for $r \in [0, s]$ and $u^1(r) = \bar{u}(r)$ for $r \in [s, t]$, one again has (19). We take the norm on $\mathcal{U}_{0,s} \times \mathcal{U}_{s,t}$ to be $|(u^0, u^1)| = \{|u^0|^2 + |u^1|^2\}^{1/2}$, where the norms on the right-hand side are the respective $\mathcal{L}_2$ norms on $[0, s]$ and $[s, t]$. The following is easily obtained from the above.

Lemma 5 For $(s, t) \in \hat{T}$, $\mathcal{C} : \mathcal{U}_{0,s} \times \mathcal{U}_{s,t} \rightarrow \mathcal{U}_{0,t}$ is a linear bijection. Further, the induced operator norm satisfies $|\mathcal{C}| = |\mathcal{C}^{-1}| = 1$.

Lemma 6 Let $(s, t) \in \hat{T}$ and $x, z \in \mathbb{R}^n$. Then,

$$\begin{aligned} W(t, x, z) &= \operatorname{stat}_{\bar{u} \in \mathcal{U}_{0,t}} J(0, t, x, \bar{u}, z) \\ &= \operatorname{stat}_{(u^0, u^1) \in \mathcal{U}_{0,s} \times \mathcal{U}_{s,t}} J_d(s, t, x, u^0, u^1, z). \end{aligned}$$

Proof: The left-hand equality is simply (17) with $s = 0$, where existence and uniqueness are guaranteed by Assumption (A.1). In particular,

$$W(t, x, z) = J(0, t, x, \bar{u}^*, z), \quad (20)$$

where $\bar{u}^* = \operatorname{argstat}_{\bar{u} \in \mathcal{U}_{0,t}} J(0, t, x, \bar{u}, z)$. Let $(s, t) \in \hat{T}$ and $x, z \in \mathbb{R}^n$. By (1),

$$0 = \lim_{\bar{u} \rightarrow \bar{u}^*} \frac{|J(0, t, x, \bar{u}, z) - J(0, t, x, \bar{u}^*, z)|}{|\bar{u} - \bar{u}^*|}. \quad (21)$$

For any $\bar{u} \in \mathcal{U}_{0,t}$, let $(u^0, u^1) = \mathcal{C}^{-1}(\bar{u})$, and in particular, let $(u^{0,*}, u^{1,*}) = \mathcal{C}^{-1}(\bar{u}^*)$. By (19), (21) and Lemma 5,

$$0 = \lim_{(u^0, u^1) \rightarrow (u^{0,*}, u^{1,*})} \frac{|J_d(s, t, x, u^0, u^1, z) - J_d(s, t, x, u^{0,*}, u^{1,*}, z)|}{|(u^0, u^1) - (u^{0,*}, u^{1,*})|},$$

which implies

$$(u^{0,*}, u^{1,*}) \in \operatorname{argstat}_{(u^0, u^1) \in \mathcal{U}_{0,s} \times \mathcal{U}_{s,t}} J_d(s, t, x, u^0, u^1, z). \quad (22)$$

Similarly, $(\hat{u}^0, \hat{u}^1) \in \operatorname{argstat}_{J_d(s, t, x, u^0, u^1, z)}$ implies that $\mathcal{C}(\hat{u}^0, \hat{u}^1) \in \operatorname{argstat}_{\bar{u} \in \mathcal{U}_{0,t}} J(0, t, x, \bar{u}, z) = \bar{u}^*$. Consequently, $(u^{0,*}, u^{1,*})$ is the unique staticizer, and

$$J_d(s, t, x, u^{0,*}, u^{1,*}, z) = \operatorname{stat}_{(u^0, u^1) \in \mathcal{U}_{0,s} \times \mathcal{U}_{s,t}} J_d(s, t, x, u^0, u^1, z).$$

Also, by (19), $J_d(s, t, x, u^{0,*}, u^{1,*}, z) = J(0, t, x, \bar{u}^*, z)$. Combining these last two equalities with (20) completes the proof. $\square$

The DPP is obtained as follows:

Theorem 7 Let $(s, t) \in \hat{T}$, $t - s < \delta^c$ (where δ^c is given in Assumption (A.2)) and $x, z \in \mathbb{R}^n$. Suppose that for any $\zeta \in \mathbb{R}^n$, $J(0, t - s, \zeta, \cdot, z) \in C^2(\mathcal{U}_{0,t-s})$. Suppose $J(0, t, x, \mathcal{C}(\cdot, \cdot), z) \in C^2(\mathcal{U}_{0,s} \times \mathcal{U}_{s,t})$, and that it has bounded second-order differences. Then,

$$\begin{aligned} W(t, x, z) &= \operatorname{stat}_{u^0 \in \mathcal{U}_{0,s}} \left\{ \int_0^s L(\xi^0(r), u^0(r)) dr \right. \\ &\quad \left. + W(t - s, \xi^0(s), z) \right\}, \end{aligned}$$

where $\xi^0(r) = f(\xi^0(r), u^0(r))$ for $r \in (0, s)$, with $\xi^0(0) = x$.

Proof: Let $(s, t) \in \hat{T}$ and $x, z \in \mathbb{R}^n$. By Lemma 6,

$$W(t, x, z) = \operatorname{stat}_{(u^0, u^1) \in \mathcal{U}_{0,s} \times \mathcal{U}_{s,t}} J_d(s, t, x, u^0, u^1, z),$$

which by (18),

$$\begin{aligned} &= \operatorname{stat}_{(u^0, u^1) \in \mathcal{U}_{0,s} \times \mathcal{U}_{s,t}} \left\{ \int_0^s L(\xi^0(r), u^0(r)) dr \right. \\ &\quad \left. + \int_s^t L(\xi^1(r), u^1(r)) dr + \psi(\xi^1(t), z) \right\} \\ &= \operatorname{stat}_{(u^0, u^1) \in \mathcal{U}_{0,s} \times \mathcal{U}_{s,t}} \left\{ \int_0^s L(\xi^0(r), u^0(r)) dr \right. \\ &\quad \left. + J(s, t, \xi^0(s), u^1, z) \right\} \\ &= \operatorname{stat}_{(u^0, \hat{u}^1) \in \mathcal{U}_{0,s} \times \mathcal{U}_{0,t-s}} \left\{ \int_0^s L(\xi^0(r), u^0(r)) dr \right. \\ &\quad \left. + J(0, t - s, \xi^0(s), \hat{u}^1, z) \right\}. \quad (23) \end{aligned}$$

By the assumptions, the term in curly brackets on the right-hand side of (23), as a function of $u^0, \hat{u}^1$ over $\mathcal{U}_{0,s} \times \mathcal{U}_{0,t-s}$ satisfies the conditions of Lemma 17 (with $\hat{u}^1$ taking the role of y , and u^0 taking the role of z there). Consequently,

$$\begin{aligned} &\operatorname{stat}_{(u^0, \hat{u}^1) \in \mathcal{U}_{0,s} \times \mathcal{U}_{0,t-s}} \left\{ \int_0^s L(\xi^0(r), u^0(r)) dr \right. \\ &\quad \left. + J(0, t - s, \xi^0(s), \hat{u}^1, z) \right\} \\ &= \operatorname{stat}_{u^0 \in \mathcal{U}_{0,s}} \min_{\hat{u}^1 \in \mathcal{U}_{0,t-s}} \left\{ \int_0^s L(\xi^0(r), u^0(r)) dr \right. \\ &\quad \left. + J(0, t - s, \xi^0(s), \hat{u}^1, z) \right\} \\ &= \operatorname{stat}_{u^0 \in \mathcal{U}_{0,s}} \left\{ \int_0^s L(\xi^0(r), u^0(r)) dr \right. \end{aligned}$$

$$+ \min_{\hat{u}^1 \in \mathcal{U}_{0,t-s}} J(0, t-s, \xi^0(s), \hat{u}^1, z) \Big\},$$

which by the assumed differentiability of $J(0, t-s, \xi^0(s), \cdot, z)$ and the definition of W ,

$$= \text{stat}_{u^0 \in \mathcal{U}_{0,s}} \left\{ \int_0^s L(\xi^0(r), u^0(r)) dr + W(t-s, \xi^0(s), z) \right\}. \quad (24)$$

Combining (23) and (24) yields the desired result. $\square$

In addition to the above DPP, we now obtain a verification theorem for the associated HJ PDE. The HJ PDE problem, for each $z \in \mathbb{R}^n$, is

$$0 = \text{stat}_{v \in \mathbb{R}^m} \{ L(x, v) - W_r(r, x, z) \quad (25)$$

$$+ W_x(r, x, z) \cdot f(x, v) \}, \quad (r, x) \in (0, t) \times \mathbb{R}^n, \\ W(0, x, z) = \psi(x, z), \quad x \in \mathbb{R}^n. \quad (26)$$

Theorem 8 *Let $t \in (0, \infty)$. Suppose $\bar{W} \in C^3((0, t) \times \mathbb{R}^{2n}) \cap C([0, t] \times \mathbb{R}^{2n})$ satisfies (25), (26). Suppose there exist $K_L, K_f < \infty$ such that*

$$|L_{xx}(x, v)|, |L_{xv}(x, v)|, |L_{vv}(x, v)| \leq K_L, \\ |f_x(x, v)|, |f_v(x, v)| \leq K_f, \\ |f_{xx}(x, v)|, |f_{xv}(x, v)|, |f_{vv}(x, v)| \leq K_f,$$

for all $(x, v) \in \mathbb{R}^n \times \mathbb{R}^m$. Suppose there exists $\tilde{u} : [0, t] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfying $\tilde{u}(r, x) \in \text{argstat}_{v \in \mathbb{R}^m} [\bar{W}_x(t-r, x, z) \cdot f(x, v) + L(x, v)]$, such that $\tilde{u}$ is bounded on bounded sets, and such that there exists $\tilde{K}_f < \infty$ such that $|f(x, \tilde{u}(r, x)) - f(y, \tilde{u}(r, y))| \leq \tilde{K}_f |x - y|$ for all $x, y \in \mathbb{R}^n$ and $r \in [0, t]$. Then, $\bar{W}(t, x, z) = W(t, x, z) = \text{stat}_{u \in \mathcal{U}_{0,t}} J(0, t, x, u, z)$.

Proof: Let $t, x, z, \bar{W}, \tilde{u}$ be as assumed. By standard results, there exists unique absolutely continuous $\tilde{\xi} : [0, t] \rightarrow \mathbb{R}^n$ satisfying $\dot{\tilde{\xi}}(r) = f(\tilde{\xi}(r), \tilde{u}(r, \tilde{\xi}(r)))$, $\tilde{\xi}(0) = x$ such that $\tilde{u}(\cdot, \tilde{\xi}(\cdot)) \in \mathcal{U}_{0,t}$. By the Fundamental Theorem of Calculus,

$$\bar{W}(t, x, z) = \bar{W}(0, \tilde{\xi}(t), z) + \int_0^t \bar{W}_r(r, \tilde{\xi}(t-r), z) \\ - \bar{W}_x(r, \tilde{\xi}(t-r), z) \cdot f(\tilde{u}(t-r, \tilde{\xi}(t-r))) dr,$$

which by (25), (26) and the choice of $\tilde{u}$,

$$= \psi(\tilde{\xi}(t), z) + \int_0^t L(\tilde{\xi}(t-r), \tilde{u}(t-r, \tilde{\xi}(t-r))) dr \\ = \int_0^t L(\tilde{\xi}(r), \tilde{u}(r, \tilde{\xi}(r))) dr + \psi(\tilde{\xi}(t), z). \quad (27)$$

Let $u^*(r) \doteq \tilde{u}(r, \tilde{\xi}(r))$, $\xi^*(r) = \tilde{\xi}(r)$ for all $r \in [0, t]$. Then, (27) may be rewritten as

$$\bar{W}(t, x, z) \quad (28) \\ = \int_0^t L(\xi^*(r), u^*(r)) dr + \psi(\xi^*(t), z) = J(0, t, x, u^*, z).$$

Now we must demonstrate that u^* is the *argstat* of $J(0, t, x, \cdot, z)$. Let $\hat{u} \in \mathcal{U}_{0,t}$, $\delta \doteq \hat{u} - u^* \in \mathcal{L}_2((0, t); \mathbb{R}^n)$ with $|\delta| \leq 1$, and let $\hat{\xi}$ be the corresponding trajectory (i.e., $\dot{\hat{\xi}}(r) = x + \int_0^r f(\hat{\xi}(r), \hat{u}(r)) dr$). By (28) and the Fundamental Theorem of Calculus again,

$$J(0, t, x, u^*, z) = \bar{W}(t, x, z) \\ = \bar{W}(0, \hat{\xi}(t), z) + [\bar{W}(t, x, z) - \bar{W}(0, \hat{\xi}(t), z)] \\ = \bar{W}(0, \hat{\xi}(t), z) + \int_0^t \bar{W}_r(r, \hat{\xi}(t-r), z) \\ - \bar{W}_x(r, \hat{\xi}(t-r), z) \cdot f(\hat{\xi}(t-r), \hat{u}(t-r)) dr,$$

which by (25), (26) (implying that the last term in brackets below is zero),

$$= \psi(\hat{\xi}(t), z) + \int_0^t L(\hat{\xi}(t-r), \hat{u}(t-r)) dr \\ + \int_0^t [-L(\hat{\xi}(t-r), \hat{u}(t-r)) + \bar{W}_r(r, \hat{\xi}(t-r), z) \\ - \bar{W}_x(r, \hat{\xi}(t-r), z) \cdot f(\hat{\xi}(t-r), \hat{u}(t-r))] \\ - [-L(\xi^*(t-r), u^*(t-r)) + \bar{W}_r(r, \xi^*(t-r), z) \\ - \bar{W}_x(r, \xi^*(t-r), z) \cdot f(\xi^*(t-r), u^*(t-r))] dr,$$

and noting that the first and second terms comprise the payoff for $\hat{u}$,

$$= J(0, t, x, \hat{u}, z) \\ + \int_0^t [-L(\hat{\xi}(t-r), \hat{u}(t-r)) + \bar{W}_r(r, \hat{\xi}(t-r), z) \\ - \bar{W}_x(r, \hat{\xi}(t-r), z) \cdot f(\hat{\xi}(t-r), \hat{u}(t-r))] \\ - [-L(\xi^*(t-r), u^*(t-r)) + \bar{W}_r(r, \xi^*(t-r), z) \\ - \bar{W}_x(r, \xi^*(t-r), z) \cdot f(\xi^*(t-r), u^*(t-r))] dr.$$

That is,

$$|J(0, t, x, \hat{u}, z) - J(0, t, x, u^*, z)| \quad (29) \\ \leq \int_0^t \left| [-L(\hat{\xi}(t-r), \hat{u}(t-r)) + \bar{W}_r(r, \hat{\xi}(t-r), z) \\ - \bar{W}_x(r, \hat{\xi}(t-r), z) \cdot f(\hat{\xi}(t-r), \hat{u}(t-r))] \right. \\ \left. - [-L(\xi^*(t-r), u^*(t-r)) + \bar{W}_r(r, \xi^*(t-r), z) \\ - \bar{W}_x(r, \xi^*(t-r), z) \cdot f(\xi^*(t-r), u^*(t-r))] \right| dr.$$

For compactness of presentation, let $|\Xi|$ denote the integrand in (29), and let

$$G(r, x, v) = G(r, x, v; z) \doteq \overline{W}_x(r, x, z) \cdot f(x, v), \quad (30)$$

for $x, z \in \mathbb{R}^n$, $v \in \mathbb{R}^m$ and $r \in (0, t)$. Using the assumed smoothness, at any $r \in (0, t)$, and where also for compactness of presentation, we drop the z argument of $\overline{W}$ where confusion will not occur, the integrand in (29) is

$$\begin{aligned} |\Xi| = & \left| -L_x(\xi^*(t-r), u^*(t-r))\Delta_{t-r} \right. \\ & - L_v(\xi^*(t-r), u^*(t-r))\delta_{t-r} \\ & - \frac{1}{2}[\Delta'_{t-r}L_{xx}(\tilde{x}_{t-r}^0, \tilde{u}_{t-r}^0)\Delta_{t-r} \\ & + 2\Delta'_{t-r}L_{xv}(\tilde{x}_{t-r}^0, \tilde{u}_{t-r}^0)\delta_{t-r} + \delta'_{t-r}L_{vv}(\tilde{x}_{t-r}^0, \tilde{u}_{t-r}^0)\delta_{t-r}] \\ & + \overline{W}_{rx}(r, \xi^*(t-r))\Delta_{t-r} + \frac{1}{2}\Delta'_{t-r}\overline{W}_{rxx}(r, \tilde{x}^1)\Delta_{t-r} \\ & - G_x(r, \xi^*(t-r), u^*(t-r))\Delta_{t-r} \\ & - G_v(r, \xi^*(t-r), u^*(t-r))\delta_{t-r} \\ & - \frac{1}{2}[\Delta'_{t-r}G_{xx}(r, \tilde{x}_{t-r}^2, \tilde{v}_{t-r}^2)\Delta_{t-r} \\ & + 2\Delta'_{t-r}G_{xv}(r, \tilde{x}_{t-r}^2, \tilde{v}_{t-r}^2)\delta_{t-r} \\ & + \delta'_{t-r}G_{vv}(r, \tilde{x}_{t-r}^2, \tilde{v}_{t-r}^2)\delta_{t-r}] \Big|, \end{aligned} \quad (31)$$

where $\delta_{t-r} \doteq \delta(t-r)$, $\Delta_{t-r} \doteq \hat{\xi}(t-r) - \xi^*(t-r)$, and for all $r \in (0, t)$, $(\tilde{x}_{t-r}^0, \tilde{u}_{t-r}^0)$, $(\tilde{x}_{t-r}^1, \tilde{u}_{t-r}^1)$, $(\tilde{x}_{t-r}^2, \tilde{u}_{t-r}^2)$ lie on the line segment from $(\xi^*(t-r), u^*(t-r))$ to $(\hat{\xi}(t-r), \hat{u}(t-r))$ (where the use of subscript $t-r$ for time-dependence is helpful in saving space). Using the assumed bounds on the second derivatives of L , (31) yields

$$\begin{aligned} |\Xi| \leq & \left| -L_x(\xi^*(t-r), u^*(t-r))\Delta_{t-r} \right. \\ & - L_v(\xi^*(t-r), u^*(t-r))\delta_{t-r} \\ & + \overline{W}_{rx}(r, \xi^*(t-r))\Delta_{t-r} + \frac{1}{2}\Delta'_{t-r}\overline{W}_{rxx}(r, \tilde{x}^1)\Delta_{t-r} \\ & - G_x(r, \xi^*(t-r), u^*(t-r))\Delta_{t-r} \\ & - G_v(r, \xi^*(t-r), u^*(t-r))\delta_{t-r} \\ & - \frac{1}{2}[\Delta'_{t-r}G_{xx}(r, \tilde{x}_{t-r}^2, \tilde{v}_{t-r}^2)\Delta_{t-r} \\ & + 2\Delta'_{t-r}G_{xv}(r, \tilde{x}_{t-r}^2, \tilde{v}_{t-r}^2)\delta_{t-r} \\ & + \delta'_{t-r}G_{vv}(r, \tilde{x}_{t-r}^2, \tilde{v}_{t-r}^2)\delta_{t-r}] \Big| \\ & + \bar{C}_1(|\Delta_{t-r}|^2 + |\delta_{t-r}|^2) \quad \forall r \in (0, t) \end{aligned} \quad (32)$$

for an appropriate choice of $\bar{C}_1 < \infty$.

Now, by the choice of u^* , Lemma 1 and (25),

$$\begin{aligned} & L_v(\xi^*(t-r), u^*(t-r)) + G_v(r, \xi^*(t-r), u^*(t-r)) \\ & = \frac{\partial}{\partial v} \{L(\xi^*(t-r), v) \\ & \quad + \overline{W}_x(r, \xi^*(t-r)) \cdot f(\xi^*(t-r), v)\} \Big|_{v=u^*(t-r)} = 0. \end{aligned} \quad (33)$$

Next, note that

$$\begin{aligned} & L_x(\xi^*(t-r), u^*(t-r)) - \overline{W}_{rx}(r, \xi^*(t-r)) \\ & \quad + G_x(r, \xi^*(t-r), u^*(t-r)) \\ & = \frac{\partial}{\partial x} \{L(x, u^*(t-r)) - \overline{W}_r(r, x) \\ & \quad + \overline{W}_x(r, x) \cdot f(x, u^*(t-r))\} \Big|_{x=\xi^*(t-r)} \\ & = \frac{d}{dx} \{L(x, \tilde{u}(t-r, x)) - \overline{W}_r(r, x) \\ & \quad + \overline{W}_x(r, x) \cdot f(x, \tilde{u}(t-r, x))\} \Big|_{x=\xi^*(t-r)} \\ & \quad - \frac{\partial}{\partial v} \{L(x, v) - \overline{W}_r(r, x) \\ & \quad + \overline{W}_x(r, x) \cdot f(x, v)\} \Big|_{\substack{x=\xi^*(t-r) \\ v=\tilde{u}(t-r, \xi^*(t-r))}} \cdot \frac{d\tilde{u}}{dx}(r, \xi^*(t-r)), \end{aligned}$$

where the total derivative notation $\frac{d}{dx}$ outside the first curly bracket is abused to denote the derivative with respect to all x explicitly appearing in the bracketed expression. By (33), this is

$$\begin{aligned} & = \frac{d}{dx} \{L(x, \tilde{u}(t-r, x)) - \overline{W}_r(r, x) \\ & \quad + \overline{W}_x(r, x) \cdot f(x, \tilde{u}(t-r, x))\} \Big|_{x=\xi^*(t-r)} \\ & = 0, \end{aligned} \quad (34)$$

where the last equality follows by (25) as the previous right-hand side is the derivative of zero. Substituting (33) and (34) in (32), we find

$$\begin{aligned} |\Xi| \leq & \left| \frac{1}{2}\Delta'_{t-r}\overline{W}_{rxx}(r, \tilde{x}^1)\Delta_{t-r} \right. \\ & - \frac{1}{2}[\Delta'_{t-r}G_{xx}(r, \tilde{x}_{t-r}^2, \tilde{v}_{t-r}^2)\Delta_{t-r} \\ & + 2\Delta'_{t-r}G_{xv}(r, \tilde{x}_{t-r}^2, \tilde{v}_{t-r}^2)\delta_{t-r} \\ & + \delta'_{t-r}G_{vv}(r, \tilde{x}_{t-r}^2, \tilde{v}_{t-r}^2)\delta_{t-r}] \Big| \\ & + \bar{C}_1(|\Delta_{t-r}|^2 + |\delta_{t-r}|^2) \quad \forall r \in (0, t). \end{aligned} \quad (35)$$

In order to proceed, we need L_∞ bounds on the trajectories, ξ^* and $\hat{\xi}$. Note that for $r \in (0, t)$,

$$\begin{aligned} |\xi^*(r) - \hat{\xi}(r)| \leq & \int_0^r |f(\hat{\xi}(\rho), \hat{u}(\rho)) - f(\xi^*(\rho), \hat{u}(\rho))| \\ & + |f(\xi^*(\rho), \hat{u}(\rho)) - f(\xi^*(\rho), u^*(\rho))| d\rho, \end{aligned}$$

which by the assumed bounds on the derivatives of f ,

$$\leq K_f \int_0^r |\hat{\xi}(\rho) - \xi^*(\rho)| + |\hat{u}(\rho) - u^*(\rho)| d\rho,$$

which upon application of the Hölder's inequality,

$$\leq K_f \int_0^r |\hat{\xi}(\rho) - \xi^*(\rho)| d\rho + K_f |\delta| \sqrt{r}.$$

Applying Gronwall's inequality to this, one finds that there exists $\bar{C}_2 = \bar{C}_2(t) < \infty$ such that

$$|\xi^*(r) - \hat{\xi}(r)| \leq \bar{C}_2 |\delta| \leq \bar{C}_2 \quad \forall r \in (0, t). \quad (36)$$

Similarly, using

$$|\xi^*(r) - x| \leq \int_0^r |f(\xi^*(\rho), \tilde{u}(\rho, \xi^*(\rho))) - f(x, \tilde{u}(\rho, x))| d\rho + \int_0^r |f(x, \tilde{u}(\rho, x))| d\rho,$$

one may show there exists $\bar{C}_3 = \bar{C}_3(t, |x|) < \infty$ such that

$$|\xi^*(r)| \leq \bar{C}_3 \quad \forall r \in (0, t), \quad (37)$$

where we do not include the details. Using (36) and (37), one finds there exists $\bar{C}_4 = \bar{C}_4(t, |x|) < \infty$

$$|\hat{\xi}(r)|, |\xi^*(r)| \leq \bar{C}_4 \quad \forall r \in (0, t). \quad (38)$$

By (38), the assumed smoothness of $\bar{W}$ and f , and the assumed bounds on derivatives of f , there exists $\bar{C}_5 = \bar{C}_5(t, |x|) < \infty$ such that

$$|\bar{W}_{rxx}(r, \tilde{x}^1)|, |G_{xx}(r, \tilde{x}_{t-r}^2, \tilde{v}_{t-r}^2)|, |G_{xv}(r, \tilde{x}_{t-r}^2, \tilde{v}_{t-r}^2)|, |G_{vv}(r, \tilde{x}_{t-r}^2, \tilde{v}_{t-r}^2)| \leq \bar{C}_5 \quad \forall r \in (0, t). \quad (39)$$

Finally, combining (35) and (39), there exists $\bar{C}_6 = \bar{C}_6(t, |x|) < \infty$ such that

$$|\Xi| \leq \bar{C}_6 \left(|\Delta_{t-r}|^2 + |\delta_{t-r}|^2 \right) \quad \forall r \in (0, t). \quad (40)$$

Applying (40) in (29), we have

$$|J(0, t, x, \hat{u}, z) - J(0, t, x, u^*, z)| \leq \bar{C}_6 \int_0^t \left(|\Delta_{t-r}|^2 + |\delta_{t-r}|^2 \right) dr = \bar{C}_6 \left(|\Delta|^2 + |\delta|^2 \right)$$

which by (36),

$$\leq \bar{C}_7 |\delta|^2,$$

for appropriate choice of $\bar{C}_7 = \bar{C}_7(t, |x|) < \infty$. Recalling (1), we see that this implies that u^* is the argstat of $J(0, t, x, \cdot, z)$. $\square$

5 Linear-quadratic example

We consider the linear-quadratic problem given by

$$L(x, v) = \frac{1}{2} v' D v - \frac{1}{2} x' B x, \quad f(x, v) = v, \quad (41)$$

$$\psi(x, z) = \psi^c(x, z), \quad (42)$$

for all $x, v, z \in \mathbb{R}^n$, where $D \succ dI$ (where we write $A \succ B$ if $A - B$ is positive definite), $d > 0$, B and D are symmetric, and $c \in (0, \infty)$. We look for $\bar{W}$ of the form

$$\bar{W}(t, x, z) = \frac{1}{2} [x' P(t) x + 2x' Q(t) z + z' R(t) z]. \quad (43)$$

With the above quadratic cost and given dynamics, the HJ PDE (25) takes the form

$$\begin{aligned} 0 &= \text{stat}_{v \in \mathbb{R}^n} \left[\frac{1}{2} v' D v - \frac{1}{2} x' B x - \bar{W}_r(r, x, z) \right. \\ &\quad \left. + \bar{W}_x(r, x, z) \cdot v \right] \\ &= \min_{v \in \mathbb{R}^n} \left[\frac{1}{2} v' D v - \frac{1}{2} x' B x - \bar{W}_r(r, x, z) \right. \\ &\quad \left. + \bar{W}_x(r, x, z) \cdot v \right] \end{aligned} \quad (44)$$

$$\begin{aligned} &= -\frac{1}{2} x' B x - \bar{W}_r(r, x, z) \\ &\quad - \frac{1}{2} \bar{W}_x(r, x, z) D^{-1} \bar{W}_x(r, x, z). \end{aligned} \quad (45)$$

Remark 9 That (44) is a minimum, in spite of the fact that, for sufficiently long duration problems, the value is obtained as a staticum, may appear at first glance to be contradictory. The consistent minimum in the HJ PDE is due to the infinitesimal limit implicit there. If one examines the DPP (Theorem 7), we see that there are two terms inside the outer staticization, where for sufficiently short durations, $t - s$, the payoff underlying the second term is convex with respect to the input in $\mathcal{U}_{s,t}$. It is also worth noting that at the outset of the proof of the DPP, one is already working with a value, W , which is defined as a staticum rather than a minimum.

Substituting form (43) in (45), one obtains

$$\begin{aligned} 0 &= -\frac{1}{2} x' B x - \frac{1}{2} [x' \dot{P}(t) x + 2x' \dot{Q}(t) z + z' \dot{R}(t) z] \\ &\quad - \frac{1}{2} (P(t)x + Q(t)z)' D^{-1} (P(t)x + Q(t)z). \end{aligned}$$

Equating like terms yields

$$\dot{P}(t) = -B - P(t) D^{-1} P(t), \quad (46)$$

$$\dot{Q}(t) = -P(t) D^{-1} Q(t), \quad (47)$$

$$\dot{R}(t) = -Q'(t) D^{-1} Q(t), \quad (48)$$

and the initial condition (i.e., (42), (26) and (8)) imply

$$P(0) = R(0) = cI = -Q(0). \quad (49)$$

Note that if P, Q, R are well-defined on $(0, t)$, then the optimal control is given by $u^*(r) = P(t-r)x + Q(t-r)z$ for $r \in (0, t)$, and one has $u^* \in \mathcal{U}_{0,t}$. Further, one can verify that the assumptions of verification Theorem 8 are valid on this interval. Consequently, $\bar{W}$ is the value function on this interval.

We are now faced with the prospect that P, Q and R may exhibit finite escape times, while there may exist

an argstat remaining finite (a.e.) indefinitely. In fact, one finds that $\bar{W}$ continues to have form (43) past such escape times, and one must obtain the means to correctly propagate the solution of (46)–(48) past such times. The means for this was indicated in the specific mass-spring example above, and we proceed to obtain this in a more general context.

Suppose we have successfully propagated forward to time $s \in (0, \infty)$, and wish to propagate to $t > s$, where $t - s < \delta^c$ (defined in (A.2)), and one might have an escape in (s, t) . We will use the staticization DPP (i.e., Theorem 7) to propagate from s up to t . For $y \in \mathbb{R}^n$, define

$$\mathcal{U}_{0,s}^y \doteq \left\{ u \in \mathcal{U}_{0,s} \mid \int_0^s u(r) dr = y \right\}. \quad (50)$$

We need to verify that the conditions of Theorem 7 hold for sufficiently small $\delta^c > 0$ (see [14,15] for similar computations). We have:

Lemma 10 *Suppose $\delta^c < \sqrt{2d/\max\{1, \hat{\lambda}\}}$ where $\hat{\lambda}$ is the maximal eigenvalue of B in (41). Let $\tau \in (0, \delta^c)$ and $\zeta, z \in \mathbb{R}^n$. Let J be given by (16), where L, f are given by (41). $J(0, \tau, \zeta, \cdot, z)$ is a convex quadratic function on $\mathcal{U}_{0,\tau}$.*

Proof: We prove only the convexity. Let $u, \delta \in \mathcal{U}_{0,\tau}$, $u^+ \doteq u + \delta$, $u^- \doteq u - \delta$; $\xi(0) = \xi^+(0) = \xi^-(0) = \zeta$; $\dot{\xi}(r) = u(r)$, $\dot{\xi}^+(r) = u^+(r)$, $\dot{\xi}^-(r) = u^-(r)$ on $(0, \tau)$. One easily sees that

$$\begin{aligned} & J(0, \tau, \zeta, u^+, z) - 2J(0, \tau, \zeta, u, z) + J(0, \tau, \zeta, u^-, z) \\ &= \frac{1}{2} \int_0^\tau \delta(r)' D \delta(r) - \left(\int_0^\tau \delta(\rho) d\rho \right)' B \left(\int_0^\tau \delta(\rho) d\rho \right) \\ &\quad + \frac{c}{2} \left| \int_0^\tau \delta(r) dr \right|^2 \\ &\geq \frac{d}{2} |\delta|_{\mathcal{L}_2(0,\tau)}^2 - \frac{\hat{\lambda}}{2} \int_0^\tau \left| \int_0^r \delta(\rho) d\rho \right|^2 dr \\ &\geq \frac{d}{2} |\delta|_{\mathcal{L}_2(0,\tau)}^2 - \frac{\hat{\lambda}}{2} \int_0^\tau r \int_0^r |\delta(\rho)|^2 d\rho dr \\ &\geq \left[\frac{d}{2} - \frac{\hat{\lambda}\tau^2}{4} \right] |\delta|_{\mathcal{L}_2(0,\tau)}^2. \quad \square \end{aligned}$$

The assumptions of bounded second-order differences and C^2 behavior are not difficult to verify, and we do not include verifications of these. Uniqueness assumption (A.1) does not always hold for purely quadratic problems, and we simply assume it here. (One may note the nonuniqueness of trajectories for the scalar mass-spring system when $x = z = 0$ and the duration is a

half-period, as an example.) Applying Theorem 7, and recalling notation (50),

$$\begin{aligned} W(t, x, z) &= \text{stat}_{u^0 \in \mathcal{U}_{0,s}} \left\{ \int_0^s L(\xi^0(r), u^0(r)) dr \right. \\ &\quad \left. + W(t - s, \xi^0(s), z) \right\} \\ &= \text{stat}_{\zeta \in \mathbb{R}^n} \text{stat}_{u^0 \in \mathcal{U}_{0,s}^{\zeta-x}} \left\{ \int_0^s L(\xi^0(r), u^0(r)) dr \right. \\ &\quad \left. + W(t - s, \zeta, z) \right\} \\ &= \text{stat}_{\zeta \in \mathbb{R}^n} \left\{ \text{stat}_{u^0 \in \mathcal{U}_{0,s}^{\zeta-x}} \left[\int_0^s L(\xi^0(r), u^0(r)) dr + \psi^\infty(\xi^0(s), \zeta) \right] \right. \\ &\quad \left. + W(t - s, \zeta, z) \right\} \\ &= \text{stat}_{\zeta \in \mathbb{R}^n} \left\{ W(s, x, \zeta) + W(t - s, \zeta, z) \right\}. \quad (51) \end{aligned}$$

As we have already propagated forward to s , we have

$$W(s, x, \zeta) = \frac{1}{2} [x' P(s)x + 2x' Q(s)\zeta + \zeta' R(s)\zeta]. \quad (52)$$

Further, as $t - s < \delta^c$, we have strict convexity on the $t - s$ duration segment, and consequently we have

$$W(t - s, \zeta, z) = \frac{1}{2} [\zeta' P(t - s)\zeta + 2\zeta' Q(t - s)z + z' R(t - s)z]. \quad (53)$$

Combining (51)–(53), one finds staticizing point $\zeta^* = -[R(s) + P(t - s)]^{-1}[Q'(s)x + Q(t - s)z]$, and consequently,

$$W(t, x, z) = \frac{1}{2} [x' P(t)x + x' Q(t)z + z' Q(t)x + z' R(t)z], \quad (54)$$

where $P(t), Q(t), R(t)$ are given by

$$P(t) = P(s) - Q(s)[R(s) + P(t - s)]^{-1}Q'(s), \quad (55)$$

$$Q(t) = Q(s)[R(s) + P(t - s)]^{-1}Q(t - s), \quad (56)$$

$$R(t) = R(t - s) - Q'(t - s)[R(s) + P(t - s)]^{-1}Q(t - s). \quad (57)$$

That is, (55)–(57) allows us to propagate past the finite escape time occurring in interval (s, t) . We remark that forward propagation by repeated application of updates similar to (55)–(57) for general DREs is discussed in [16] for the finite-dimensional case, and in [6–8] for the infinite-dimensional case.

6 Applications

The initial motivation for consideration of stationarity problems is the stationary action principle for dynamical systems. However, from the above, one sees that the main results, specifically the dynamic programming

principle and the relation to the associated HJ PDE, are valid on a much larger domain. One may seek the staticum in a wide variety of problems. One obvious additional application is in the area of zero-sum games. In particular, when value exists, this value is a minimax solution over the control pair, say $(u(\cdot), w(\cdot))$ where u is the minimizing player control and w is the maximizing player control. If the payoff is continuously differentiable, the minimax value function is a staticum over the combined control, $\bar{u} \doteq (u, w)$. Consequently the results here may be applied for solution. Of course, there is more structure in a game than what is assumed or discussed here.

Returning to the original motivation domain, recall that the stationary action may be the least action over a short time-interval. In the least-action case, one may apply standard techniques from optimal control to solve the associated control-problem formulation. This is discussed in the context of the n -body problem in [12,14,15], and in the context of the wave equation in [4–6]. The results obtained in this paper allow such results to be extended to arbitrary duration problems.

References

- [1] F.L. Baccelli, G. Cohen, G.J. Olsder and J.-P. Quadrat, *Synchronization and Linearity*, John Wiley, New York, 1992.
- [2] R. Curtain and H. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*, Vol. 21 of Texts in Applied Mathematics, Springer-Verlag, New York, 1995.
- [3] E.J. Davison and M.C. Maki, “The numerical solution of the matrix Riccati differential equation”, IEEE Trans. Auto. Control, 18 (1973), 71–73.
- [4] P.M. Dower and W.M. McEneaney, “A max-plus fundamental solution semigroup for a class of lossless wave equations”, Proc. SIAM Conf. on Control and Its Applics. (2015), 400–407.
- [5] P.M. Dower and W.M. McEneaney, “Solving two-point boundary value problems for a wave equation via the principle of stationary action and optimal control”, Preprint arXiv:1501.02006 (2015), in review.
- [6] P.M. Dower and W.M. McEneaney, “A fundamental solution for an infinite dimensional two-point boundary value problem via the principle of stationary action,” Proc. Australian Control Conf. (2013), 270–275.
- [7] P.M. Dower and W.M. McEneaney, “A max-plus based fundamental solution for a class of infinite dimensional Riccati equations”, Proc. IEEE Conf. Dec. Control (2011), 615–620.
- [8] P.M. Dower and W.M. McEneaney, “A max-plus dual space fundamental solution for a class of operator differential Riccati equations”, SIAM J. Control and Optim., 53 (2015), 969–1002.
- [9] R.P. Feynman, “Space-time approach to non-relativistic quantum mechanics”, Rev. of Mod. Phys., 20 (1948) 367–387.
- [10] R.P. Feynman, *The Feynman Lectures on Physics*, Vol. 2, Basic Books, (1964) 19-1–19-14.
- [11] C.G. Gray and E.F. Taylor, “When action is not least”, Am. J. Phys. 75, (2007), 434–458.
- [12] S.H. Han and W.M. McEneaney, “The principle of least action and two-point boundary value problems in orbital mechanics”, Proc. Amer. Control Conf. (2014), 1939–1944
- [13] Tsoy-Wo Ma, *Classical Analysis on Normed Spaces*, World Scientific, Singapore, 1995.
- [14] W.M. McEneaney and P.M. Dower, “The principle of least action and fundamental solutions of mass-spring and n -body two-point boundary value problems”, SIAM J. Control and Optim. (2015), 53(5):2898–2933.
- [15] W.M. McEneaney and P.M. Dower, “The principle of least action and solution of two-point boundary value problems on a limited time horizon”, Proc. SIAM Conf. on Control and Its Applics. (2013), 199–206.
- [16] W.M. McEneaney, “A new fundamental solution for differential Riccati equations Arising in Control”, Automatica (2008), 44:920–936.
- [17] T. Padmanabhan, *Gravitation: Foundations and Frontiers*, Cambridge Univ. Press, 2010.
- [18] R.T. Rockafellar and R.J. Wets, *Variational Analysis*, Springer-Verlag, New York, 1997.
- [19] H. Zhang and P.M. Dower, “A max-plus primal space fundamental solution for a class of difference Riccati equations”, Proc. Math. Theory Networks and Systems (2014), 1446–1452.

7 Appendix

Some technical results related to staticization follow. Suppose $\mathcal{Y}, \mathcal{Z}$ are Hilbert spaces, and let the inner product on $\mathcal{Y} \times \mathcal{Z}$ be given by $\langle (y^1, z^1), (y^2, z^2) \rangle \doteq \langle y^1, y^2 \rangle_{\mathcal{Y}} + \langle z^1, z^2 \rangle_{\mathcal{Z}}$, where $\langle \cdot, \cdot \rangle_{\mathcal{Y}}$ and $\langle \cdot, \cdot \rangle_{\mathcal{Z}}$ denote the inner products on $\mathcal{Y}$ and $\mathcal{Z}$, respectively. Suppose $\mathcal{G} \subseteq \mathcal{Y}$ is open and convex. Suppose $F : \mathcal{G} \times \mathcal{Z} \rightarrow \mathbb{R}$. We present results under subsets of the following assumptions.

$$F \in C^2(\mathcal{G} \times \mathcal{Z}; \mathbb{R}). \quad (\text{A.F1})$$

$$F \text{ is strictly uniformly convex on } \mathcal{G}. \text{ (See Remark 3.)} \quad (\text{A.F2})$$

$$\text{Either: a) } \mathcal{G} = \mathcal{Y} \text{ and } F \text{ is coercive on } \mathcal{Y} \text{ as defined in Remark 3; or b) } \mathcal{G} \text{ is bounded, and for every } z \in \mathcal{Z}, \text{ there exists } \bar{y} = \bar{y}(z) \in \mathcal{G} \text{ such that } F(\bar{y}, z) < F(y, z) \text{ for all } y \in \mathcal{G} \setminus \{\bar{y}\}. \quad (\text{A.F3})$$

The following result is standard, and we do not include a proof. See also [18]. Note that for $r > 0$, we let $B_r \doteq \{y \in \mathcal{Y} \mid |y| < r\}$, where throughout, the relevant underlying space will be implicitly clear.

Lemma 11 *Suppose F satisfies (A.F2) and (A.F3). For each $z \in \mathcal{Z}$, there exists a unique $\bar{y}(z) \in \mathcal{G}$ such that $F(\bar{y}(z), z) < F(y, z)$ for all $y \in \mathcal{G} \setminus \{\bar{y}(z)\}$. Further, given $R_z < \infty$, there exists $R_y = R_y(R_z) < \infty$ such that $\bar{y}(z) \in B_{R_y}$ for all $z \in B_{R_z}$.*

Lemma 12 Suppose F satisfies (A.F2) and (A.F3). Suppose that for each $z \in \mathcal{Z}$, $F(\cdot, z) \in C^2(\mathcal{G}; \mathbb{R})$. Suppose that for any $R_y, R_z < \infty$, there exists $\bar{K}_0 < \infty$ such that $|F(y, z^1) - F(y, z^2)| \leq \bar{K}_0 |z^1 - z^2|$ for all $y \in B_{R_y}$, $z^1, z^2 \in B_{R_z}$. Let $\bar{y} : \mathcal{Z} \rightarrow \mathcal{G}$ be as defined in Lemma 11. There exists $\bar{K}_H = \bar{K}_H(R_z) < \infty$ such that $\bar{y}$ is Hölder continuous, with $|\bar{y}(z^1) - \bar{y}(z^2)| \leq \bar{K}_H |z^1 - z^2|^{1/2}$ for all $z^1, z^2 \in B_{R_z}(0) \subset \mathcal{Z}$.

Proof: Let $R_z < \infty$, $z^1, z^2 \in B_{R_z}$ and $\bar{y} : \mathcal{Z} \rightarrow \mathcal{G}$ be as given. Let $\hat{v} \doteq \bar{y}(z^2) - \bar{y}(z^1)$. Suppose $\hat{v} \neq 0$; otherwise there is nothing more to prove. By (A.F2), for all $\bar{\delta} \in (-1, 1)$, $\bar{\delta} \in (-1, 0) \cup (0, 1)$,

$$[F(\bar{y}(z^1) + \bar{\delta}\hat{v} - \delta\hat{v}, z^1) - 2F(\bar{y}(z^1) + \bar{\delta}\hat{v}, z^1) + F(\bar{y}(z^1) + \bar{\delta}\hat{v} + \delta\hat{v}, z^1)]/\delta^2 \geq C_F. \quad (58)$$

Define $G(\bar{\delta}) \doteq F(\bar{y}(z^1) + \bar{\delta}\hat{v}, z^1)$ for all $\bar{\delta} \in [-1, 1]$, where we note that by Assumption (A.F1), $G \in C^2$. Further, by (58),

$$\frac{d^2 G}{d\bar{\delta}^2}(\bar{\delta}) \geq C_F \quad \forall \bar{\delta} \in (-1, 1). \quad (59)$$

Also, by Lemma 11, $G(\bar{\delta}), G(-\bar{\delta}) \geq G(0)$ for all $\bar{\delta} \in (-1, 1)$, and consequently, $\frac{dG}{d\bar{\delta}}(0) = 0$. Combining this with (59), we see that $G(\bar{\delta}) - G(0) \geq (C_F/2)\bar{\delta}^2$ for all $\bar{\delta} \in [-1, 1]$. Taking $\bar{\delta} = 1$, and recalling the definition of G ,

$$F(\bar{y}(z^2), z^1) - F(\bar{y}(z^1), z^1) \geq (C_F/2)|\bar{y}(z^2) - \bar{y}(z^1)|^2,$$

or, equivalently,

$$|\bar{y}(z^2) - \bar{y}(z^1)|^2 \leq (2/C_F)[F(\bar{y}(z^2), z^1) - F(\bar{y}(z^1), z^1)]. \quad (60)$$

Proceeding similarly, but at z^2 instead of z^1 , one finds

$$|\bar{y}(z^1) - \bar{y}(z^2)|^2 \leq (2/C_F)[F(\bar{y}(z^1), z^2) - F(\bar{y}(z^2), z^2)]. \quad (61)$$

Now, by Lemma 11, there exists $R_y = R_y(R_z) < \infty$ such that $\bar{y}(z) \in B_{R_y}$ for all $z \in B_{R_z}$, and in particular, $\bar{y}(z^1), \bar{y}(z^2) \in B_{R_y}$. Then, using the assumed local Lipschitz nature, there exists $\bar{K}_0(R_z) < \infty$ such that

$$|F(\bar{y}(z^1), z^2) - F(\bar{y}(z^1), z^1)| \leq \bar{K}_0 |z^2 - z^1|, \quad (62)$$

$$|F(\bar{y}(z^2), z^1) - F(\bar{y}(z^2), z^2)| \leq \bar{K}_0 |z^2 - z^1|. \quad (63)$$

Summing (60) and (61), we have

$$2|\bar{y}(z^2) - \bar{y}(z^1)|^2 \leq (2/C_F)[F(\bar{y}(z^2), z^1) - F(\bar{y}(z^1), z^1) + F(\bar{y}(z^1), z^2) - F(\bar{y}(z^2), z^2)], \quad (64)$$

which by (62), (63),

$$\leq \frac{4\bar{K}_0}{C_F} |z^2 - z^1|. \quad \square$$

Lemma 13 Suppose F satisfies (A.F1), (A.F2) and (A.F3). Let $\bar{y} : \mathcal{Z} \rightarrow \mathcal{G}$ be as defined in Lemma 11. Let $R_z < \infty$. There exists $\bar{K}_L = \bar{K}_L(R_z) < \infty$ such that $|\bar{y}(z^1) - \bar{y}(z^2)| \leq \bar{K}_L |z^1 - z^2|$ for all $z^1, z^2 \in B_{R_z}$.

Proof: The conditions assumed in Lemma 13 are sufficient to guarantee the conditions of Lemma 12. Consequently, by (64), we have

$$|\bar{y}(z^2) - \bar{y}(z^1)|^2 \leq (2/C_F)[F(\bar{y}(z^2), z^1) - F(\bar{y}(z^1), z^1) + F(\bar{y}(z^1), z^2) - F(\bar{y}(z^2), z^2)]. \quad (65)$$

For $y^a, y^b \in \mathcal{G}$, we will let $\overline{y^a, y^b}$ denote the line segment from y^a to y^b , and similarly on $\mathcal{Z}$. By (A.F1), there exists $z^3, z^4 \in \overline{z^1, z^2}$ such that

$$F(\bar{y}(z^2), z^2) - F(\bar{y}(z^2), z^1) = \langle F_z(\bar{y}(z^2), z^3), z^2 - z^1 \rangle, \quad (66)$$

and

$$F(\bar{y}(z^1), z^2) - F(\bar{y}(z^1), z^1) = \langle F_z(\bar{y}(z^1), z^4), z^2 - z^1 \rangle, \quad (67)$$

where partial $F_z(y, z) \in \mathcal{Z}$ for each pair $(y, z) \in \mathcal{G} \times \mathcal{Z}$. Combining (65)–(67), we have

$$\begin{aligned} |\bar{y}(z^2) - \bar{y}(z^1)|^2 &\leq (2/C_F) \langle F_z(\bar{y}(z^1), z^4) - F_z(\bar{y}(z^2), z^3), z^2 - z^1 \rangle \\ &\leq (2/C_F) |F_z(\bar{y}(z^1), z^4) - F_z(\bar{y}(z^2), z^3)| |z^2 - z^1|. \end{aligned} \quad (68)$$

As $\mathcal{G}$ is open, there exists $\epsilon > 0$ such that $\zeta : (-\epsilon, 1+\epsilon) \rightarrow \mathcal{G} \times \mathcal{Z}$, where

$$\zeta(\lambda) \doteq (1 - \lambda) (\bar{y}(z^1), z^4) + \lambda (\bar{y}(z^2), z^3).$$

Then, letting $G(\lambda) \doteq F_z(\zeta(\lambda))$, we have $G \in C^1((-\epsilon, 1+\epsilon); \mathcal{Z})$. With a slight abuse of notation,

$$\begin{aligned} |F_z(\bar{y}(z^1), z^4) - F_z(\bar{y}(z^2), z^3)| &= |F_z(\zeta(0)) - F_z(\zeta(1))| = |G(0) - G(1)|. \end{aligned} \quad (69)$$

Defining $M_{0,1} \doteq \max_{\lambda \in [0,1]} \left| \frac{dG}{d\lambda}(\lambda) \right| < \infty$, and using the Integral Mean Value Theorem (cf., [13]), (69) implies

$$\begin{aligned} |F_z(\bar{y}(z^1), z^4) - F_z(\bar{y}(z^2), z^3)| &\leq M_{0,1} |\zeta(1) - \zeta(0)| \\ &= M_{0,1} |(\bar{y}(z^1), z^4) - (\bar{y}(z^2), z^3)| \\ &\leq M_{0,1} [|\bar{y}(z^1) - \bar{y}(z^2)| + |z^1 - z^2|]. \end{aligned} \quad (70)$$

Combining (68) and (70), we have

$$\begin{aligned} |\bar{y}(z^2) - \bar{y}(z^1)|^2 &\leq (2/C_F)M_{0,1} [|\bar{y}(z^1) - \bar{y}(z^2)| |z^1 - z^2| \\ &\quad + |z^1 - z^2|^2] \\ &\leq \frac{1}{2} |\bar{y}(z^1) - \bar{y}(z^2)|^2 + \frac{1}{2} \left[(2/C_F)^2 M_{0,1}^2 + (2/C_F)M_{0,1} \right] \\ &\quad \cdot |z^1 - z^2|^2, \end{aligned}$$

which yields the desired result. $\square$

Lemma 14 Suppose F satisfies (A.F1), (A.F2) and (A.F3). Let $\bar{y} : \mathcal{Z} \rightarrow \mathcal{G}$ be as defined in Lemma 11. Then, $F(\bar{y}(\cdot), \cdot)$ is Gateaux differentiable, with Gateaux derivative $\partial F((\bar{y}(z), z); v) = \langle F_z(\bar{y}(z), z), v \rangle$ for all $z, v \in \mathcal{Z}$, where $F_z(\hat{y}, z)$ denotes the partial (Fréchet) derivative of F in z at $\hat{y} \in \mathcal{G}$.

Proof: Let $z^0, v \in \mathcal{Z}$, $|v| = 1$, $\omega \in \mathbb{R}$, $y^0 \doteq \bar{y}(z^0)$. Let $\hat{z}(\omega) \doteq z^0 + \omega v$. By (A.F1), given $\omega \in [0, 1]$, there exists $\tilde{y} \in y^0, \bar{y}(\hat{z}(\omega))$ such that

$$\begin{aligned} F(\bar{y}(\hat{z}(\omega)), \hat{z}(\omega)) &= F(y^0, z^0) \\ &+ \langle F_y(y^0, z^0), \bar{y}(\hat{z}(\omega)) - y^0 \rangle_y \\ &+ \frac{1}{2} \langle \bar{y}(\hat{z}(\omega)) - y^0, F_{yy}(\tilde{y}, z^0)(\bar{y}(\hat{z}(\omega)) - y^0) \rangle_y \\ &+ F(\bar{y}(\hat{z}(\omega)), \hat{z}(\omega)) - F(\bar{y}(\hat{z}(\omega)), z^0), \end{aligned}$$

which by (A.F1) and the definition of y^0 ,

$$\begin{aligned} &= \frac{1}{2} \langle \bar{y}(\hat{z}(\omega)) - y^0, F_{yy}(\tilde{y}, z^0)(\bar{y}(\hat{z}(\omega)) - y^0) \rangle_y \\ &+ F(y^0, z^0) + F(\bar{y}(\hat{z}(\omega)), \hat{z}(\omega)) - F(\bar{y}(\hat{z}(\omega)), z^0), \end{aligned}$$

which, using (A.F1) again, for appropriate choice of $\tilde{z} \in \overline{z^0, \hat{z}(\omega)}$,

$$\begin{aligned} &= F(y^0, z^0) + \langle F_z(\bar{y}(\hat{z}(\omega)), z^0), \hat{z}(\omega) - z^0 \rangle_z \\ &+ \frac{1}{2} \langle \hat{z}(\omega) - z^0, F_{zz}(\bar{y}(\hat{z}(\omega)), \tilde{z})(\hat{z}(\omega) - z^0) \rangle_z \\ &+ \frac{1}{2} \langle \bar{y}(\hat{z}(\omega)) - y^0, F_{yy}(\tilde{y}, z^0)(\bar{y}(\hat{z}(\omega)) - y^0) \rangle_y. \end{aligned}$$

Equivalently,

$$\begin{aligned} &F(\bar{y}(\hat{z}(\omega)), \hat{z}(\omega)) - F(y^0, z^0) \\ &= \langle F_z(\bar{y}(\hat{z}(\omega)), z^0), \hat{z}(\omega) - z^0 \rangle_z \\ &+ \frac{1}{2} \langle \hat{z}(\omega) - z^0, F_{zz}(\bar{y}(\hat{z}(\omega)), \tilde{z})(\hat{z}(\omega) - z^0) \rangle_z \\ &+ \frac{1}{2} \langle \bar{y}(\hat{z}(\omega)) - y^0, F_{yy}(\tilde{y}, z^0)(\bar{y}(\hat{z}(\omega)) - y^0) \rangle_y, \end{aligned}$$

which implies

$$\begin{aligned} &|F(\bar{y}(\hat{z}(\omega)), \hat{z}(\omega)) - F(y^0, z^0) \\ &\quad - \langle F_z(y^0, z^0), \hat{z}(\omega) - z^0 \rangle_z| \\ &\leq |F_z(\bar{y}(\hat{z}(\omega)), z^0) - F_z(y^0, z^0)| |\hat{z}(\omega) - z^0| \\ &\quad + \frac{1}{2} |F_{zz}(\bar{y}(\hat{z}(\omega)), \tilde{z})| |\hat{z}(\omega) - z^0|^2 \end{aligned}$$

$$+ \frac{1}{2} |F_{yy}(\tilde{y}, z^0)| |\bar{y}(\hat{z}(\omega)) - y^0|^2. \quad (71)$$

Employing the Integral Mean Value Theorem (cf., [13]), one finds that there exists $\hat{y} \in \overline{\bar{y}(\hat{z}(\omega)), y^0}$ such that

$$|F_z(\bar{y}(\hat{z}(\omega)), z^0) - F_z(y^0, z^0)| \leq |F_{yz}(\hat{y}, z^0)| |\bar{y}(\hat{z}(\omega)) - y^0|. \quad (72)$$

Substituting (72) in (71), we have

$$\begin{aligned} &|F(\bar{y}(\hat{z}(\omega)), \hat{z}(\omega)) - F(y^0, z^0) \\ &\quad - \langle F_z(y^0, z^0), \hat{z}(\omega) - z^0 \rangle_z| \\ &\leq |F_{yz}(\hat{y}, z^0)| |\bar{y}(\hat{z}(\omega)) - y^0| |\hat{z}(\omega) - z^0| \\ &\quad + \frac{1}{2} |F_{zz}(\bar{y}(\hat{z}(\omega)), \tilde{z})| |\hat{z}(\omega) - z^0|^2 \\ &\quad + \frac{1}{2} |F_{yy}(\tilde{y}, z^0)| |\bar{y}(\hat{z}(\omega)) - y^0|^2, \end{aligned}$$

which by Lemma 13,

$$\begin{aligned} &\leq \left[\bar{K}_L |F_{yz}(\hat{y}, z^0)| + \frac{1}{2} |F_{zz}(\bar{y}(\hat{z}(\omega)), \tilde{z})| \right. \\ &\quad \left. + \frac{1}{2} \bar{K}_L^2 |F_{yy}(\tilde{y}, z^0)| \right] \omega^2. \end{aligned} \quad (73)$$

We will obtain bounds on the second-derivative terms, independent of $\omega \in [0, 1]$. Consider first, the $|F_{yy}(\tilde{y}, z^0)|$ term. For each $\omega \in [0, 1]$, there exists $\lambda \in [0, 1]$ such that

$$(\tilde{y}, z^0) = (\lambda y^0 + (1 - \lambda)\bar{y}(z^0 + \omega v), z^0). \quad (74)$$

Consider

$$\begin{aligned} \mathcal{C} &\doteq \{(y, z^0) \in \mathcal{G} \times \mathcal{Z} \mid \exists \lambda, \omega \in [0, 1] \text{ s.t.} \\ &\quad y = \lambda y^0 + (1 - \lambda)\bar{y}(z^0 + \omega v)\}, \end{aligned}$$

and let $\bar{G} : [0, 1]^2 \rightarrow \mathcal{C}$ be given by $\bar{G}(\lambda, \omega) \doteq (\lambda y^0 + (1 - \lambda)\bar{y}(z^0 + \omega v), z^0)$. Let (λ, ω) , $(\hat{\lambda}, \hat{\omega}) \in [0, 1]^2$. Then,

$$\begin{aligned} |\bar{G}(\lambda, \omega) - \bar{G}(\hat{\lambda}, \hat{\omega})| &\leq |\lambda - \hat{\lambda}| |y^0| + |\lambda - \hat{\lambda}| |\bar{y}(z^0 + \omega v)| \\ &\quad + |1 - \hat{\lambda}| |\bar{y}(z^0 + \hat{\omega} v) - \bar{y}(z^0 + \omega v)|, \end{aligned}$$

which, noting $|1 - \hat{\lambda}| \leq 1$, and using Lemma 13,

$$\leq (2|y^0| + \bar{K}_L) |\lambda - \hat{\lambda}| + \bar{K}_L |\omega - \hat{\omega}|,$$

which implies that $\bar{G}$ is Lipschitz. As $[0, 1]^2$ is compact and $\bar{G}$ is Lipschitz, we see that $\mathcal{C}$ is compact. Then, by (A.F1), the restriction of $|F_{yy}(\cdot, \cdot)|$ to $\mathcal{C}$ is a continuous, functional on a compact set, and consequently, there exists $D_{yy} = D_{yy}(v) < \infty$ such that

$$|F_{yy}(\tilde{y}, z^0)| \leq D_{yy} \quad \forall \omega \in [0, 1]. \quad (75)$$

By a completely analogous proof, there exists $D_{yz} = D_{yz}(v) < \infty$ such that

$$|F_{yz}(\hat{y}, z^0)| \leq D_{yz} \quad \forall \omega \in [0, 1]. \quad (76)$$

Only $|F_{zz}(\bar{y}(\hat{z}(\omega)), \tilde{z})|$ remains. In a similar manner to above, note that for each $\omega \in [0, 1]$, there exists $\lambda \in [0, 1]$ such that

$$(\bar{y}(\hat{z}(\omega)), \tilde{z}) = (\bar{y}(z^0 + \omega v), z^0 + (1 - \lambda)\omega v).$$

Letting

$$\mathcal{C}_{zz} \doteq \{(y, z) \in \mathcal{G} \times \mathcal{Z} \mid \exists \lambda, \omega \in [0, 1] \text{ s.t. } y = \bar{y}(z^0 + \omega v) \text{ and } z = z^0 + (1 - \lambda)\omega v\},$$

and $\widehat{G} : [0, 1]^2 \rightarrow \mathcal{C}_{zz}$ be given by $\widehat{G}(\lambda, \omega) \doteq (\bar{y}(z^0 + \omega v), z^0 + (1 - \lambda)\omega v)$, we find that $\widehat{G}$ is Lipschitz on a compact domain, $[0, 1]^2$, and consequently, $\mathcal{C}_{zz}$ is compact. Then, again similarly to the above, there exists $D_{zz} = D_{zz}(v) < \infty$ such that

$$|F_{zz}(\bar{y}(\hat{z}(\omega)), \tilde{z})| \leq D_{zz} \quad \forall \omega \in [0, 1]. \quad (77)$$

By (73)–(77),

$$\begin{aligned} & |F(\bar{y}(\hat{z}(\omega)), \hat{z}(\omega)) - F(y^0, z^0)| \\ & - \langle F_z(y^0, z^0), \hat{z}(\omega) - z^0 \rangle_z \\ & \leq [\bar{K}_L D_{yz} + \tfrac{1}{2} D_{zz} + \tfrac{1}{2} \bar{K}_L^2 D_{yz}] \omega^2, \end{aligned}$$

which yields the assertions.

Lemma 15 *Suppose F satisfies (A.F1), (A.F2) and (A.F3). Let $\bar{y} : \mathcal{Z} \rightarrow \mathcal{G}$ be as defined in Lemma 11. Suppose F has bounded second-order differences. Then, $F(\bar{y}(\cdot), \cdot)$ is Fréchet differentiable with derivative $DF(\bar{y}(z), z) = F_z(\hat{y}, z)$ for all $z, v \in \mathcal{Z}$, where $\hat{y} = \bar{y}(z)$ and $F_z(\hat{y}, z)$ denotes the partial (Fréchet) derivative of F in z .*

Proof: Fix $\tilde{z} \in \mathcal{Z}$. For $z \in \mathcal{Z}$,

$$\begin{aligned} & |F(\bar{y}(z), z) - F(\bar{y}(\tilde{z}), \tilde{z})| \\ & \leq |F(\bar{y}(z), z) - F(\bar{y}(\tilde{z}), z)| + |F(\bar{y}(\tilde{z}), z) - F(\bar{y}(\tilde{z}), \tilde{z})| \\ & \leq |F(\bar{y}(z), z) - F(\bar{y}(\tilde{z}), z) - F(\bar{y}(z), \tilde{z}) + F(\bar{y}(\tilde{z}), \tilde{z})| \\ & + |F(\bar{y}(z), \tilde{z}) - F(\bar{y}(\tilde{z}), \tilde{z})| + |F(\bar{y}(\tilde{z}), z) - F(\bar{y}(\tilde{z}), \tilde{z})|, \end{aligned}$$

and because F has bounded second-order differences, there exists $M_1 < \infty$ such that this is

$$\begin{aligned} & \leq M_1 [| \bar{y}(z) - \bar{y}(\tilde{z}) |^2 + |z - \tilde{z}|^2] \\ & + |F(\bar{y}(z), \tilde{z}) - F(\bar{y}(\tilde{z}), \tilde{z})| + |F(\bar{y}(\tilde{z}), z) - F(\bar{y}(\tilde{z}), \tilde{z})|, \end{aligned}$$

and by (A.F1) and the choice of $\bar{y}$, there exists $M_2 < \infty$ such that this is

$$\begin{aligned} & \leq M_1 [| \bar{y}(z) - \bar{y}(\tilde{z}) |^2 + |z - \tilde{z}|^2] + M_2 [| \bar{y}(z) - \bar{y}(\tilde{z}) |^2] \\ & + |F(\bar{y}(\tilde{z}), z) - F(\bar{y}(\tilde{z}), \tilde{z})|, \end{aligned}$$

which by Lemma 13,

$$\begin{aligned} & \leq [M_1(1 + \bar{K}_L^2) + M_2 \bar{K}_L^2] |z - \tilde{z}|^2 \\ & + |F(\bar{y}(\tilde{z}), z) - F(\bar{y}(\tilde{z}), \tilde{z})|. \end{aligned}$$

This implies

$$\begin{aligned} & |F(\bar{y}(z), z) - F(\bar{y}(\tilde{z}), \tilde{z})| / |z - \tilde{z}| \\ & \leq [M_1(1 + \bar{K}_L^2) + M_2 \bar{K}_L^2] |z - \tilde{z}| \\ & + |F(\bar{y}(\tilde{z}), z) - F(\bar{y}(\tilde{z}), \tilde{z})| / |z - \tilde{z}|. \end{aligned}$$

Consequently, using (A.F1), the Fréchet derivative of $G(z) \doteq F(\bar{y}(z), z)$ exists and has the asserted form.

Lemma 16 *Suppose F satisfies (A.F1), (A.F2) and (A.F3), and has bounded second-order differences. Let $\bar{y} : \mathcal{Z} \rightarrow \mathcal{G}$ be as defined in Lemma 11. Suppose $(y^*, z^*) = \text{argstat}_{(y,z) \in \mathcal{G} \times \mathcal{Z}} F(y, z)$ (where uniqueness is implicit). Then, $y^* = \bar{y}(z^*)$ and $z^* = \text{argstat}_{z \in \mathcal{Z}} F(\bar{y}(z), z)$.*

Proof: By Lemma 1, $DF(y, z)|_{(y,z)=(y^*,z^*)} = 0$, where 0 denotes the zero element of $L(\mathcal{Y} \times \mathcal{Z}; \mathcal{Y} \times \mathcal{Z})$. By (A.F1) and standard results, this implies

$$D_y F(y, z^*)|_{y=y^*} = 0 \text{ and } D_z F(y^*, z)|_{z=z^*} = 0, \quad (78)$$

where we continue to use 0 to generically denote the zero element of the appropriate space, and D_y, D_z denote partial Fréchet derivatives with respect to the indicated variables. Next, note that by (A.F1) and the definition of $\bar{y}$,

$$D_y F(y, z)|_{y=\bar{y}(z)} = 0 \quad \forall z \in \mathcal{Z}. \quad (79)$$

By definition and (79),

$$F(\bar{y}(z), z) = \min_{y \in \mathcal{G}} F(y, z) = \text{stat}_{y \in \mathcal{G}} F(y, z). \quad (80)$$

In particular, $\bar{y}(z^*) = \text{argstat}_{y \in \mathcal{G}} F(y, z^*)$. Combining this with (78), we see

$$y^* = \bar{y}(z^*), \quad (81)$$

which is the first assertion. Now, by (81) and Lemma 15, and then applying (78),

$$DF(\bar{y}(z), z)|_{z=z^*} = D_z(y^*, z)|_{z=z^*} = 0. \quad (82)$$

Then, by (82) and Lemma 1,

$$z^* \in \text{argstat}_{z \in \mathcal{Z}} F(\bar{y}(z), z). \quad (83)$$

Now suppose $z^\dagger \in \text{argstat}_{z \in \mathcal{Z}} F(\bar{y}(z), z)$, which by Lemma 1, is equivalent to

$$DF(\bar{y}(z), z)|_{z=z^\dagger} = 0. \quad (84)$$

Let $y^\dagger \doteq \bar{y}(z^\dagger)$. By (84) and Lemma 11,

$$D_z F(y^\dagger, z)|_{z=z^\dagger} = 0. \quad (85)$$

Also, by the definition of $y^\dagger$ and (A.F1),

$$D_y F(y, z^\dagger)|_{y=y^\dagger} = 0. \quad (86)$$

By (85), (86) and (A.F1), $DF(y, z)|_{(y,z)=(y^\dagger,z^\dagger)} = 0$, which implies, using Lemma 1, that $(y^\dagger, z^\dagger) \in \text{argstat}_{(y,z) \in \mathcal{G} \times \mathcal{Z}} F(y, z)$, which, by the lemma assumptions, implies $(y^\dagger, z^\dagger) = (y^*, z^*)$, and of course, this implies $z^\dagger = z^*$. Combining this with (83) yields the desired result. $\square$

Lemma 17 *Under the same conditions as Lemma 16, one has*

$$\begin{aligned} \text{stat}_{(y,z) \in \mathcal{G} \times \mathcal{Z}} F(y, z) &= F(y^*, z^*) = \text{stat}_{z \in \mathcal{Z}} F(\bar{y}(z), z) \\ &= \text{stat}_{z \in \mathcal{Z}} \min_{y \in \mathcal{G}} F(y, z) = \text{stat}_{z \in \mathcal{Z}} \text{stat}_{y \in \mathcal{G}} F(y, z). \end{aligned}$$

Proof: The first equality is simply a restatement of one of the assumptions. Recalling (80), $\text{stat}_{y \in \mathcal{G}} F(y, z) = \min_{y \in \mathcal{G}} F(y, z) = F(\bar{y}(z), z)$ for all $z \in \mathcal{Z}$. Combining this with the existence implied by the second assertion of Lemma 16 yields the last two equalities. Also by the last assertion of Lemma 16, and then the first assertion of Lemma 16,

$$\text{stat}_{z \in \mathcal{Z}} F(\bar{y}(z), z) = F(\bar{y}(z^*), z^*) = F(y^*, z^*). \quad \square$$