Applied Mathematics & Optimization (2021) 84 (Suppl 1):S733–S765
https://doi.org/10.1007/s00245-021-09784-6

Exploiting Characteristics in Stationary Action Problems

Vincenzo Basco[1] · Peter M. Dower[1] · William M. McEneaney[2] · Ivan Yegorov[3]

Accepted: 24 April 2021 / Published online: 6 July 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
Connections between the principle of least action and optimal control are explored with a view to describing the trajectories of energy conserving systems, subject to temporal boundary conditions, as solutions of corresponding systems of characteristics equations on arbitrary time horizons. Motivated by the relaxation of least action to stationary action for longer time horizons, due to loss of convexity of the action functional, a corresponding relaxation of optimal control problems to stationary control problems is considered. In characterizing the attendant stationary controls, corresponding to generalized velocity trajectories, an auxiliary stationary control problem is posed with respect to the characteristic system of interest. Using this auxiliary problem, it is shown that the controls rendering the action functional stationary on arbitrary time horizons have a state feedback representation, via a verification theorem, that is consistent with the optimal control on short time horizons. An example is provided to illustrate application via a simple mass-spring system.

Keywords Stationary action · Optimal control · Characteristics · Hamilton–Jacobi–Bellman partial differential equations

Supported by AFOSR Grant FA2386-16-1-4066. A preliminary version of this work appeared in [8].

✉ Peter M. Dower
pdower@unimelb.edu.au

Vincenzo Basco
vincenzo.basco@unimelb.edu.au; vincenzobasco@gmail.com

William M. McEneaney
wmceneaney@eng.ucsd.edu

Ivan Yegorov
ivan.egorov@ndsu.edu; ivanyegorov@gmail.com

1 Department of Electrical & Electronic Eng., University of Melbourne, Victoria 3010, Australia

2 Department of Mech. & Aero. Eng., UC San Diego, La Jolla, CA 92093, USA

3 Department of Mathematics, North Dakota State University, Fargo, ND 58102, USA

Mathematics Subject Classification Primary 65M25 · 49J15 · 49S05 · Secondary 35F21

1 Introduction

In recent investigations [6,10], connections between Hamilton's action principle and optimal control have been exploited to synthesize fundamental solutions for conservative systems of differential equations, in finite and infinite dimensions, and their related two point boundary value problems (TPBVPs). In each case, an optimal control problem is identified whose cost is representative of the desired *action*, leading to a characteristic system corresponding to the desired conservative system. The tools of optimal control, including dynamic programming, semigroup theory, idempotent algebra, and convex analysis, subsequently provide a pathway for construction of its fundamental solution, for large classes of boundary conditions, see for example [10].

For short time horizons, convexity of the action functional with respect to the generalized velocity trajectory is typically guaranteed for finite dimensional dynamics. This ensures that an associated optimal control problem is well-defined, c.f. [10]. Consequently, stationary action is achieved as *least* action [9], as characterised by a corresponding value function, while the associated equations of motion are described by the characteristic system corresponding to a standard Hamilton–Jacobi–Bellman (HJB) partial differential equation (PDE).

For longer or infinite time horizons, or for configurations with infinite dimensional dynamics, the equivalence of stationary action and optimal control breaks down, typically due to a loss of convexity of the action [7,10]. This leads to finite escape phenomena exhibited by the value function, and hence an inability to propagate solutions beyond these times. This limitation is particularly severe in the infinite dimensional setting [6,7], and motivates exploration of *stationary* control problems, as opposed to optimal control problems, whose value can propagate through these finite escape phenomena to longer horizons [5,11–13].

An optimal control problem can be relaxed to a stationary control problem by formally replacing the infimum (or supremum) operation in the definition of the attendant value function with a *stat* operation [11–13]. As indicated, this stat operation requires only stationarity of its cost function argument, rather than optimality. In the stationary action problems considered to date, see for example [5–7,11–14], this has involved the characterization of open loop controls that render the cost stationary. However, motivated by the notion of minimax solutions or minimal selections considered in [4,16,17], it is also reasonable to consider initial adjoint or generalized momentum variables that render an associated characteristics based cost stationary. An investigation in this direction forms the basis of this work, as initiated in [8]. The main results provide an equivalence between two stationary control problems, subject to uniqueness of solutions of a TPBVP [5], and a verification result for stationary trajectories posed with respect to a suitable HJB PDE. An illustrative example is included.

In terms of organization, Sect. 2 reviews the connection between least action and optimal control, and states the main assumptions used throughout. Section 3 relaxes optimality to stationarity in the minimax solution/minimal selection representation [4, 16,17] of the optimal control problem encapsulating least action, yielding a stationary control problem and value function compatible with stationary action. This relaxed problem is used to characterize the stationary trajectories of interest, both via TPBVPs as per earlier work [5], and via a verification theorem involving an HJB PDE. The paper concludes with a simple example in Sect. 4 followed by a brief conclusion in Sect. 5. An appendix is included, containing proofs diverted from the main body of the paper.

Throughout, $\mathbb{R}$, $\mathbb{Z}$, $\mathbb{N}$ denote the real, integer, and natural numbers respectively, with extended reals defined as $\overline{\mathbb{R}} \doteq \mathbb{R} \cup \{\pm\infty\}$. $a \vee b$ denotes the maximum of $a, b \in \overline{\mathbb{R}}$. $|\cdot|$ and $\langle\cdot,\cdot\rangle$ denote the Euclidean norm and scalar product, respectively. The space of continuous mappings between Banach spaces $\mathscr{X}$ and $\mathscr{Y}$ is denoted by $C(\mathscr{X};\mathscr{Y})$. The set of bounded linear operators between the two spaces is denoted by $\mathcal{L}(\mathscr{X};\mathscr{Y})$, or $\mathcal{L}(\mathscr{X})$ if $\mathscr{X}$ and $\mathscr{Y}$ coincide. The norm on $\mathscr{X}$ is denoted by $\|\cdot\|_{\mathscr{X}}$, or simply $\|\cdot\|$ if contextually clear. If $\mathscr{X}$ is a Hilbert space, the inner product is denoted by $\langle\cdot,\cdot\rangle_{\mathscr{X}}$, or $\langle\cdot,\cdot\rangle$ again if contextually clear. Given a closed interval $I \subset \mathbb{R}$ and Hilbert space $\mathscr{X}$, the space of square summable measurable functions on I, endowed with the standard inner product, is denoted by $\mathscr{L}^2(I;\mathscr{X})$.

Given two Banach spaces $\mathscr{X}$, $\mathscr{Y}$, a function $f \in C(\mathscr{X};\mathscr{Y})$ is *Fréchet differentiable at* $x \in \mathscr{X}$, with derivative $Df(x) = D_x f(x) \in \mathcal{L}(\mathscr{X};\mathscr{Y})$, if the function $df_x : \mathscr{X} \to \mathscr{Y}$ defined by

$$df_x(h) \doteq \begin{cases} 0 & \|h\|_{\mathscr{X}} = 0, \\ \frac{f(x+h)-f(x)-Df(x)\,h}{\|h\|_{\mathscr{X}}} & \|h\|_{\mathscr{X}} > 0, \end{cases} \tag{1}$$

satisfies $\lim_{\|h\|_{\mathscr{X}}\to 0} \|df_x(h)\|_{\mathscr{Y}} = 0$; i.e. f is Fréchet differentiable at x if $h \mapsto df_x(h)$ is continuous at 0. The function f is *Fréchet differentiable* (everywhere) if it is Fréchet differentiable at every $x \in \mathscr{X}$. A function f is continuously Fréchet differentiable, denoted $f \in C^1(\mathscr{X};\mathscr{Y})$, if $Df : \mathscr{X} \to \mathcal{L}(\mathscr{X};\mathscr{Y})$ is continuous. Higher order Fréchet derivatives are similarly defined, with $f \in C^k(\mathscr{X};\mathscr{Y})$ if f is k-times continuously Fréchet differentiable, $k \in \mathbb{N}$. The Fréchet derivative Df of a real-valued function $f : \mathscr{X} \to \mathbb{R}$, where it exists, has the Riesz representation $Df(x)\,h = \langle\nabla f(x),\, h\rangle$ for $x, h \in \mathscr{X}$, in which $\nabla f = \nabla_x f : \mathscr{X} \to \mathscr{X}$. Where f is twice Fréchet differentiable, $D^2 f(x)\,h\,\tilde{h} = \langle D\nabla f(x)\,\tilde{h}, h\rangle = \langle\nabla^2 f(x)\,\tilde{h},\, h\rangle = \langle\nabla^2 f(x)\,h,\, \tilde{h}\rangle$ for all $h, \tilde{h} \in \mathscr{X}$, in which $\nabla^2 f = D\nabla f : \mathscr{X} \to \mathcal{L}(\mathscr{X})$ denotes the Fréchet derivative of ∇f.

2 Least Action and Optimal Control

The *action* associated with a desired energy conserving motion is classically defined as the time integral of the Lagrangian formed by the difference of the generalized kinetic and potential energies involved. Where the action is uniformly bounded below (or above) with respect to the generalized velocity trajectory, it may be interpreted as the integrated running cost associated with an optimal control problem, to which a terminal cost may be added so as to encode desired terminal conditions of the motion.

Formalizing, let $\mathscr{X}$ denote a real Hilbert space of instantaneous generalized positions of the motion, and let $T \in \mathbb{R}_{\geq 0}$ and $t \in [0, T]$ denote the final and initial times of that motion. Let $\mathscr{U}[t, T] \doteq \mathscr{L}^2([t, T]; \mathscr{X})$ define the space of associated admissible generalized velocity trajectories. Given an initial generalized position $x \in \mathscr{X}$, a coercive self-adjoint inertia operator $\mathcal{M} \in \mathcal{L}(\mathscr{X})$, a potential field $V : \mathscr{X} \to \mathbb{R}$, and an artificial terminal cost $\psi : \mathscr{X} \to \mathbb{R}$, the cost function $J_T(t, x, \cdot) : \mathscr{U}[t, T] \to \mathbb{R}$ encapsulating the action is defined by

$$J_T(t, x, u) \doteq \int_t^T \tfrac{1}{2} \langle u_s, \mathcal{M} u_s \rangle - V(\bar{x}_s)\, ds + \psi(\bar{x}_T), \tag{2}$$

in which $u \in \mathscr{U}[t, T]$ is the generalized velocity trajectory, and $\bar{x} \in C([t, T]; \mathscr{X})$ is the corresponding unique generalized position trajectory, satisfying

$$\bar{x}_s \doteq x + \int_t^s u_\sigma\, d\sigma, \tag{3}$$

for all $s \in [t, T]$. The inertial operator, potential field, and terminal cost are assumed to satisfy the following throughout:

$$\begin{aligned} &\exists\ m \in \mathbb{R}_{>0},\ K \in R_{\geq 0} \quad s.t. \quad \forall\, x, h \in \mathscr{X}, \\ &\begin{cases} m\, \|h\|^2 - \langle h, \mathcal{M} h \rangle \leq 0, \\ \|\nabla^2 V(x)\|_{\mathcal{L}(\mathscr{X})} \vee \|D\nabla^2 V(x)\|_{\mathcal{L}(\mathscr{X};\mathcal{L}(\mathscr{X}))} \vee \|D^2\nabla^2 V(x)\|_{\mathcal{L}(\mathscr{X}\times\mathscr{X};\mathcal{L}(\mathscr{X}))} \leq \frac{K}{2}, \\ \|\nabla^2 \psi(x)\|_{\mathcal{L}(\mathscr{X})} \leq \frac{K}{2}; \end{cases} \end{aligned} \tag{4}$$

i.e., the inertia operator $\mathcal{M}$ is coercive (and hence boundedly invertible), while the third and fourth derivatives of the potential field, along with the second derivative of the terminal cost, are uniformly bounded.

Under assumption (4) and for sufficiently short time horizons $T - t$, the cost $u \mapsto J_T(t, x, u)$ is strongly convex, see Lemma 1 below. Consequently, an optimal control problem defined via (2) encapsulates stationary action as *least* action, with the value function $W_T : [0, T] \times \mathscr{X} \to \mathbb{R}$ involved defined for all $t \in [0, T]$, $x \in \mathscr{X}$ by

$$W_T(t, x) \doteq \inf_{u \in \mathscr{U}[t,T]} J_T(t, x, u). \tag{5}$$

Theorem 1 *Given $m, K \in \mathbb{R}_{\geq 0}$ as per (4), $T \in \mathbb{R}_{\geq 0}$, $t \in [0, T]$ such that $((T - t) \vee 1)\,(T - t) < \frac{m}{K}$, and any $x \in \mathscr{X}$, there exists a unique $\bar{u} \in \mathscr{U}[t, T]$ such that $W_T(t, x) = J_T(t, x, \bar{u}) \in \mathbb{R}$; i.e. the value function (5) is well-defined and real-valued.*

The proof of Theorem 1 uses the following lemma.

Lemma 1 *Given arbitrary $T \in \mathbb{R}_{\geq 0}$, $t \in [0, T]$, $x \in \mathscr{X}$, and $u \in \mathscr{U}[t, T]$, cost $J_T(t, x, \cdot) : \mathscr{U}[t, T] \to \mathbb{R}$ is twice Fréchet differentiable, with second derivative $D_u \nabla_u J_T(t, x, u) \in \mathcal{L}(\mathscr{U}[t, T])$ given by*

$$D_u \nabla_u J_T(t,x,u)\,\delta_u = (\mathcal{M} - \Delta_T(t,x,u))\,\delta_u \tag{6}$$

for all $\delta_u \in \mathscr{U}[t,T]$, in which $\nabla_u J_T(t,x,u) \in \mathscr{U}[t,T]$ denotes the Riesz representation of the first Fréchet derivative at u, with $\Delta_T(t,x,u) \in \mathcal{L}(\mathscr{U}[t,T])$ given by

$$[\Delta_T(t,x,u)\,\delta_u]_r \doteq \int_t^T \left[\int_{r\vee\rho}^T \nabla^2 V(\bar{x}_\sigma)\,d\sigma - \nabla^2\psi(\bar{x}_T)\right][\delta_u]_\rho\,d\rho$$

for all $r \in [t,T]$, $\delta_u \in \mathscr{U}[t,T]$. Moreover, given $m, K \in \mathbb{R}_{\geq 0}$ as per (4),

$$\langle \delta_u,\ D_u\nabla_u J_T(t,x,u)\,\delta_u\rangle_{\mathscr{U}[t,T]} \geq K\,(\tfrac{m}{K} - ((T-t)\vee 1)\,(T-t))\,\|\delta_u\|^2_{\mathscr{U}[t,T]} \tag{7}$$

for all $\delta_u \in \mathscr{U}[t,T]$, so that $D_u\nabla_u J_T(t,x,u)$ is coercive and $J_T(t,x,\cdot)$ is strongly convex, provided that $T-t \in \mathbb{R}_{\geq 0}$ is sufficiently small.

Proof Fix $T \in \mathbb{R}_{\geq 0}$, $t \in [0,T]$, $x \in \mathscr{X}$, and $u \in \mathscr{U}[t,T]$. Hölder's inequality and the second inequality in (4) yield $\Delta_T(t,x,u) \in \mathcal{L}(\mathscr{U}[t,T])$, with

$$\|\Delta_T(t,x,u)\|_{\mathcal{L}(\mathscr{X})} \leq K((T-t)\vee 1)\,(T-t). \tag{8}$$

Twice Fréchet differentiability of $J_T(t,x,\cdot) : \mathscr{U}[t,T] \to \mathbb{R}$, boundedness of the second derivative involved, i.e. $D_u\nabla_u J_T(t,x,u) \in \mathcal{L}(\mathscr{U}[t,T])$, and (6), subsequently follow by a minor generalization of [5, Theorem 3.6]. Combining (8) with the first inequality in (4) via Cauchy–Schwartz and (6) yields

$$\begin{aligned}\langle \delta_u,\ D_u\nabla_u J_T(t,x,u)\,\delta_u\rangle_{\mathscr{U}[t,T]} &= \langle \delta_u,\ (\mathcal{M} - \Delta_T(t,x,u))\,\delta_u\rangle_{\mathscr{U}[t,T]}\\ &\geq \langle \delta_u, (m - K((T-t)\vee 1)\,(T-t))\,\delta_u\rangle_{\mathscr{U}[t,T]}\\ &= K\,(\tfrac{m}{K} - ((T-t)\vee 1)\,(T-t))\,\|\delta_u\|^2_{\mathscr{U}[t,T]},\end{aligned}$$

for all $\delta_u \in \mathscr{U}[t,T]$, which is (7). By inspection, for sufficiently short time horizons, i.e. $((T-t)\vee 1)\,(T-t) < \frac{m}{K}$, it follows that $D_u\nabla_u J_T(t,x,u) \in \mathcal{L}(\mathscr{U}[s,t])$ is coercive. Given an arbitrary $\hat{u} \in \mathscr{U}[t,T]$, and $\tilde{u} \doteq \hat{u} - u \in \mathscr{U}[t,T]$, Taylor's theorem further implies that

$$\begin{aligned}J_T(t,x,\hat{u}) &= J_T(t,x,u) + \langle \tilde{u}, \nabla_u J_T(t,x,u)\rangle_{\mathscr{U}[t,T]}\\ &\quad + \left\langle \tilde{u},\ \left(\int_0^1 (1-\sigma)\,D_u\nabla_u J_T(t,x,u+\sigma\,\tilde{u})\,d\sigma\right)\tilde{u}\right\rangle_{\mathscr{U}[t,T]}\\ &\geq J_T(t,x,u) + \langle \tilde{u}, \nabla_u J_T(t,x,u)\rangle_{\mathscr{U}[t,T]}\\ &\quad + \tfrac{1}{2}\,K\,(\tfrac{m}{K} - ((T-t)\vee 1)\,(T-t))\,\|\tilde{u}\|^2_{\mathscr{U}[t,T]}.\end{aligned}$$

Hence, $J_T(t,x,\cdot)$ is strongly convex by (4). □

Proof (Theorem 1) Fix $T \in \mathbb{R}_{\geq 0}$, $t \in [0, T]$ as per the hypothesis, and any $x \in \mathscr{X}$. Lemma 1 implies that the cost $J_T(t, x, \cdot) : \mathscr{U}[t, T] \to \mathbb{R}$ of (2) is strongly convex. Hence, the infimum in (5) is achieved at a unique minimizer $\bar{u} \in \mathscr{U}[t, T]$, thereby yielding a well-defined and real-valued optimal cost $J_T(t, x, \bar{u}) = W_T(t, x)$. □

Theorem 1 ensures that for sufficiently short time horizons the principle of stationary action can be formulated as a principle of *least* action [5,9,10], via the optimal control problem defined by the value function (5). Applying standard tools from optimal control [2–4] for finite dimensional $\mathscr{X}$, this value function may subsequently be characterized as the viscosity solution of a non-stationary HJB PDE constrained by a terminal condition. Indeed, by strengthening by (4) to include boundedness of V and ψ, [4, Theorems 5.2.12, 7.4.14] implies that the value function W_T of (5) is the unique viscosity solution W of

$$\begin{cases} 0 = -\dfrac{\partial W}{\partial t}(t, x) + H(x, \nabla_x W(t, x)), & (t, x) \in [0, T] \times \mathscr{X}, \\ W(T, x) = \psi(x), & x \in \mathscr{X}, \end{cases} \tag{9}$$

in which the Hamiltonian $H : \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ is

$$H(x, p) \doteq V(x) + \tfrac{1}{2}\langle p, \mathcal{M}^{-1} p\rangle = V(x) + \sup_{u \in \mathscr{X}} \left\{ -\langle p, u\rangle - \tfrac{1}{2}\langle u, \mathcal{M} u\rangle \right\} \tag{10}$$

for all $x, p \in \mathscr{X}$. (Alternatively, boundedness of V and ψ may be replaced with a restriction on the instantaneous generalized velocities, i.e. $u_s \in U \subset \mathscr{X}$ for all $s \in [t, T]$, in which U is compact; see [4, Theorem 7.4.14].)

In establishing the equations of motion imparted by the least action principle, the characteristic system for (9) is given by the final value problem (FVP)

$$\begin{cases} \dot{\bar{x}}_s = -\nabla_p H(\bar{x}_s, \bar{p}_s) = -\mathcal{M}^{-1} \bar{p}_s, \quad s \in [t, T], & \text{(11a)} \\ \dot{\bar{p}}_s = \nabla_x H(\bar{x}_s, \bar{p}_s) = \nabla V(\bar{x}_s), & \text{(11b)} \\ \dot{\bar{z}}_s = -\langle \bar{p}_s, \nabla_p H(\bar{x}_s, \bar{p}_s)\rangle + H(\bar{x}_s, \bar{p}_s) = V(\bar{x}_s) - \tfrac{1}{2}\langle \bar{p}_s, \mathcal{M}^{-1} \bar{p}_s\rangle, & \text{(11c)} \\ \bar{x}_T = y, \quad \bar{p}_T = \nabla\psi(y), \quad \bar{z}_T = \psi(y), \quad y \in \mathscr{X}, & \text{(11d)} \end{cases}$$

in which ∇_x and ∇_p refer to the Riesz representations of the respective Fréchet derivatives, and terminal generalized position $y \in \mathscr{X}$ is used to parameterize the solutions. The first two equations (11a), (11b) describe the motion imparted by the principle of least action, corresponding respectively to the generalized position and additive inverse of the momentum, while the third equation (11c) describes the temporal evolution of the cost, i.e. action. As formalized later in Lemma 2, equations (11a), (11b) coupled with initial data, or terminal data as per (11d), exhibit a unique classical solution, so that the second derivative $\ddot{\bar{x}}_s$ is well defined. In particular,

$$\ddot{\bar{x}}_s = -\mathcal{M}^{-1} \dot{\bar{p}}_s = -\mathcal{M}^{-1} \nabla V(\bar{x}_s)$$

for all $s \in [t, T]$, which describes Newton's second law. Observe also that the Hamiltonian H of (10) corresponds to the total energy, i.e. the sum of potential and kinetic energies. As expected, the chain rule implies that

$$\begin{aligned}
\tfrac{d}{ds} H(\bar{x}_s, \bar{p}_s) &= \langle \nabla_x H(\bar{x}_s, \bar{p}_s), \dot{\bar{x}}_s \rangle + \langle \nabla_p H(\bar{x}_s, \bar{p}_s), \dot{\bar{p}}_s \rangle \\
&= -\langle \nabla_x H(\bar{x}_s, \bar{p}_s), \nabla_p H(\bar{x}_s, \bar{p}_s) \rangle + \langle \nabla_p H(\bar{x}_s, \bar{p}_s), \nabla_x H(\bar{x}_s, \bar{p}_s) \rangle \\
&= 0,
\end{aligned}$$

for all $s \in [t, T]$, i.e. the total energy is conserved.

3 Stationary Action and Stationary Control

The connection between least action and optimal control is known to break down for longer time horizons, due typically to a loss of convexity of the action encapsulated by the cost (2), see for example [7, Lemma 1]. This can be glimpsed in Lemma 1, where the sufficient condition for convexity of the cost (2) is no longer valid, rendering the optimal control interpretation of Theorem 1 inapplicable. In practice, as the horizon length increases and convexity of the cost is lost, the value function (5) involved experiences finite escape phenomena.

However, on longer time horizons, it is well known that the principle of stationary action (rather than least action) continues to describe the motion of energy conserving systems [9]. In order to encapsulate this description in a framework that is analogous to optimal control, the infimum operation appearing in (5) is relaxed to a *stat* operation [5,10–12].

Definition 1 The *stat* operation, along with the corresponding *argstat* operation, is defined with respect to a function $F \in C^1(\mathscr{Z}; \mathbb{R})$ by

$$\begin{aligned}
\operatorname*{stat}_{\zeta \in \mathscr{Z}} F(\zeta) &\doteq \left\{ F(\bar{\zeta}) \,\middle|\, \bar{\zeta} \in \operatorname*{arg\,stat}_{\zeta \in \mathscr{Z}} F(\zeta) \right\}, \\
\operatorname*{arg\,stat}_{\zeta \in \mathscr{Z}} F(\zeta) &\doteq \left\{ \zeta \in \mathscr{Z} \,\middle|\, 0 = \lim_{y \to \zeta} \frac{|F(y) - F(\zeta)|}{\|y - \zeta\|} \right\},
\end{aligned} \tag{12}$$

in which $\mathscr{Z}$ is a Banach space. The elements in $\operatorname{arg\,stat}_{\zeta \in \mathscr{Z}} F(\zeta)$ are called *stationary points* for F.

Relaxing the *inf* appearing in (5) to a *stat* gives rise to the notion of a stationary control problem.

3.1 Stationary Control Problem

With $\mathscr{Z} \doteq \mathscr{U}[t, T]$ and $F \doteq J_T(t, x, \cdot)$ in (12), define the *stationary control problem* [5,6,11,12] corresponding to (5) via the relaxed value function $\widetilde{W}_T : [0, T] \times \mathscr{X} \to \mathbb{R}$ given by

$$\widetilde{W}_T(t,x) \doteq \operatorname*{stat}_{u\in\mathscr{U}[t,T]} J_T(t,x,u), \tag{13}$$

for all $t \in [0,T]$, $x \in \mathscr{X}$, in which J_T is the same cost (2), and the stat operation is as per (12). The utility of (13), relative to (5), in recovering the desired dynamics on arbitrary time horizons is illustrated via the following standard calculus of variations result [5].

Theorem 2 *Given $T \in \mathbb{R}_{\geq 0}$, $t \in [0,T]$, $x \in \mathscr{X}$, the following statements are equivalent:*

(i) *there exists $\bar{u} \in \mathscr{U}[t,T]$ such that the cost $J_T(t,x,\cdot)$ of (2) is stationary, i.e.*

$$\bar{u} \in \operatorname*{arg\,stat}_{u\in\mathscr{U}[t,T]} J_T(t,x,u); \tag{14}$$

(ii) *there exist $y \in \mathscr{X}$ and $(\bar{x},\bar{p}) \in (\mathscr{U}[t,T])^2$ that the TPBVP defined by FVP (11) and $\bar{x}_t = x$ is satisfied.*

Moreover, the maps $s \mapsto \bar{u}_s$ and $s \mapsto \bar{p}_s$ satisfy

$$\bar{u}_s = -\mathcal{M}^{-1}\,\bar{p}_s \quad a.e. \quad s \in [t,T]. \tag{15}$$

Proof See for example [5, Theorem 3.9]. □

Consistent with optimal control, dynamic programming is applicable to stationary control problems, albeit currently in a restricted setting, see [11]. The dynamic programming principle obtained yields an HJB PDE analogous to (9), which may be used for verification of stationarity given an explicit candidate for the value function (13). However, rather than adopt that approach here, the aim is instead to use an alternative to cost (2) that explicitly encapsulates the characteristic system (11a), (11b) describing the motion, and to use the ensuing analysis to explore how TPBVPs involving the motion might otherwise be solved.

To this end, let $(s,y) \mapsto (X^-_{s,T}(y), P^-_{s,T}(y), Z^-_{s,T}(y)) \in \mathscr{X}^2 \times \mathbb{R}$ denote the solution map for the FVP (11) integrated backwards in time. With $\bar{x}_s = X^-_{s,T}(y)$, $\bar{p}_s = P^-_{s,T}(y)$, and $\bar{u}_s \doteq -\mathcal{M}^{-1} P^-_{s,T}(y) = -\mathcal{M}^{-1}\,\bar{p}_s$ for $s \in [t,T]$, note by (11c) that

$$\begin{aligned} Z^-_{s,T}(y) = \bar{z}_s &= \psi(y) + \int_s^T \tfrac{1}{2}\,\langle P^-_{r,T}(y), \mathcal{M}^{-1}\,P^-_{r,T}(y)\rangle - V(X^-_{r,T}(y))\,dr \\ &= \psi(y) + \int_s^T \tfrac{1}{2}\,\langle \bar{p}_r, \mathcal{M}^{-1}\,\bar{p}_r\rangle - V(\bar{x}_r)\,dr \\ &= \psi(y) + \int_s^T \tfrac{1}{2}\,\langle \bar{u}_r, \mathcal{M}\,\bar{u}_r\rangle - V(\bar{x}_r)\,dr. \end{aligned} \tag{16}$$

Hence, $Z^-_{t,T}(y) = J_T(t,x,\bar{u})$, c.f. (2), when $X^-_{t,T}(y) = x$. For short horizons, this motivates an equivalent characterization of the value function W_T of (5) as the *minimax solution* [16] or *minimal selection* [4], i.e.

$$W_T(t,x) = \inf_{y\in\mathscr{X}} \left\{ Z^-_{t,T}(y) \;\middle|\; X^-_{t,T}(y) = x \right\}, \tag{17}$$

for all $t \in [0, T]$, $x \in \mathscr{X}$. That is, the value function $W_T(t, x)$ describes the minimal action accumulated by the family of solutions of the characteristic system (11) as parameterized by terminal generalized positions $\bar{x}_T = y$ that is compatible with the fixed initial generalized position x and terminal generalized momentum constraint $\bar{p}_T = \nabla\psi(y)$ imposed by the terminal cost ψ. As per [17] and in view of (16), (17), it is reasonable to reparameterize these solutions with respect to the initial adjoint variable $\bar{p}_t \doteq p \in \mathscr{X}$, rather than the terminal generalized position variable $\bar{x}_T = y$. For short horizons, the alternative value function $\widehat{W}_T : [0, T] \times \mathscr{X} \to \mathbb{R}$ of interest is defined by [8]

$$\widehat{W}_T(t,x) \doteq \inf_{p\in\mathscr{X}} \bar{J}_T(t,x,p) \tag{18}$$

for all $t \in [0, T]$, $x \in \mathscr{X}$, in which the associated cost $\bar{J}_T(t, x, \cdot) : \mathscr{X} \to \mathbb{R}$ from (16) is

$$\bar{J}_T(t,x,p) \doteq \int_t^T \tfrac{1}{2}\langle \bar{p}_s,\, \mathcal{M}^{-1}\,\bar{p}_s\rangle - V(\bar{x}_s)\,ds + \psi(\bar{x}_T), \tag{19}$$

and $s \mapsto \bar{x}_s$ and $s \mapsto \bar{p}_s$ satisfy the corresponding *initial value problem* (IVP) involving (11a), (11b), i.e.

$$\begin{cases} \dot{\bar{x}}_s = -\mathcal{M}^{-1}\,\bar{p}_s, \quad s \in [t, T], \\ \dot{\bar{p}}_s = \nabla V(\bar{x}_s), \\ \quad \bar{x}_t = x, \quad \bar{p}_t = p, \quad x, p \in \mathscr{X}. \end{cases} \tag{20}$$

Consistent with the relaxation of the optimal control value function W_T of (5) for short horizons to $\widetilde{W}_T$ of (13) for arbitrary horizons, the infimum in the definition (18) of $\widehat{W}_T$ may also be relaxed to stat (12), yielding the corresponding value function $\overline{W}_T : [0, T] \times \mathscr{X} \to \mathbb{R}$ defined by

$$\overline{W}_T(t,x) \doteq \operatorname*{stat}_{p\in\mathscr{X}} \bar{J}_T(t,x,p) \tag{21}$$

for all $T \in \mathbb{R}_{\geq 0}$, $t \in [0, T]$, $x \in \mathscr{X}$, with cost $\bar{J}_T$ as per (19).

Remark 1 For short horizons as per Theorem 1, the value function W_T of (5), the minimax solution/minimal selection (17), and the stationary control problem value function $\widetilde{W}_T$ of (13) all coincide. In particular, there exists a unique optimal control $\bar{u} \in \mathscr{U}[t, T]$ as per Theorem 1, so that the argstat in (13), (14) is the singleton $\{\bar{u}\}$, i.e. $\widetilde{W}_T(t, x)$ is single-valued and real. Moreover, Theorem 2 implies the existence of $y \in \mathscr{X}$ and a solution $s \mapsto (\bar{x}_s, \bar{p}_s)$ of the characteristic system (11) satisfying $\bar{x}_t = x$, $\bar{x}_T = y$, $\bar{p}_T = \nabla\psi(y)$, with $\bar{u}_s = -\mathcal{M}^{-1}\,\bar{p}_s$ for all $s \in [t, T]$. Hence, the set $\{y \in \mathscr{X} \mid X^-_{t,T}(y) = x\}$ in (17) is non-empty, i.e. $W_T(t, x)$ is single-valued and real.

For arbitrary horizons, convexity of the argstat (if known) is sufficient to guarantee that the value functions $\widetilde{W}_T$, $\overline{W}_T$ of (13), (21) are single-valued and real. □

The subsequent analysis is concerned the relationship between the argstats associated with (13) and (21), i.e. characterizing the former via the latter. With this analysis in mind, it is convenient for brevity of notation to define $f : \mathscr{X}^2 \to \mathscr{X}^2$, $l : \mathscr{X}^2 \to \mathbb{R}$, and $\Psi : \mathscr{X}^2 \to \mathbb{R}$ by

$$\begin{aligned} f(X) &\doteq \begin{pmatrix} -\mathcal{M}^{-1} p \\ \nabla V(x) \end{pmatrix}, \quad X \doteq Y_p(x) \doteq \begin{pmatrix} x \\ p \end{pmatrix} \in \mathscr{X}^2, \\ l(X) &\doteq \tfrac{1}{2} \langle p, \mathcal{M}^{-1} p \rangle - V(x), \quad \Psi(X) \doteq \psi(x). \end{aligned} \tag{22}$$

Using this notation, observe that (19), (20), (21) correspond to

$$\widetilde{J}_T(t, Y) = \bar{J}_T(t, x, p) = \int_t^T l(X_s)\, ds + \Psi(X_T), \tag{23}$$

$$\dot{X}_s = f(X_s), \; s \in [t, T], \; X_t = Y \doteq Y_p(x), \tag{24}$$

$$\overline{W}_T(t, x) = \operatorname*{stat}_{p \in \mathscr{X}} \widetilde{J}_T\left(t, Y_p(x)\right). \tag{25}$$

3.2 Fréchet Differentiation of the Cost

The objective now is to characterize the argstat in (21) via differentiation of (19), (23). With this in mind, some intermediate lemmas are useful. The proofs involved rely on classical arguments, such as those in [15, Chap. 5], and are delayed to Appendix A.

Lemma 2 *Given any $T \in \mathbb{R}_{\geq 0}$, $t \in [0, T]$, $Y \in \mathscr{X}^2$, the initial value problem (24) has a unique classical solution $\overline{X}(Y) \in C([t, T]; \mathscr{X}^2) \cap C^1((t, T); \mathscr{X}^2)$.*

Lemma 3 *The map $Y \mapsto \overline{X}(Y)$ defined via the unique classical solution of Lemma 2 is continuous, i.e. $\overline{X} \in C(\mathscr{X}^2; C([t, T]; \mathscr{X}^2))$. In particular, there exists an $\alpha \in \mathbb{R}_{\geq 0}$ such that*

$$\|\overline{X}(Y + h) - \overline{X}(Y)\|_\infty \leq \|h\| \exp(\alpha\,(T - t)) \tag{26}$$

for all $Y, h \in \mathscr{X}^2$.

Lemma 4 *The map $Y \mapsto \overline{X}(Y) \in C(\mathscr{X}^2; C([t, T]; \mathscr{X}^2))$ of Lemma 3 is Fréchet differentiable with derivative given by*

$$D\overline{X}(Y) \in \mathcal{L}(\mathscr{X}^2; C([t, T]; \mathscr{X}^2)), \qquad [D\overline{X}(Y)\, h]_s = U_{s,t}(Y)\, h, \quad s \in [t, T], \tag{27}$$

for all $Y, h \in \mathscr{X}^2$, in which $U_{s,r}(Y) \in \mathcal{L}(\mathscr{X}^2)$, $r, s \in [t, T]$, is an element of the two-parameter family of evolution operators generated by $A(Y)_s \in \mathcal{L}(\mathscr{X}^2)$, i.e.

$$U_{s,r}\, h = U_{s,r}(Y)\, h = h + \int_r^s A(Y)_\sigma\, U_{\sigma,r}(Y)\, h\, d\sigma, \qquad \forall\, r, s \in [t, T],\ h \in \mathscr{X}, \tag{28}$$

in which $s \mapsto A(Y)_s$ is defined uniquely, given Y, by

$$A(Y)_s \doteq \Lambda(\overline{X}(Y)_s) \doteq \begin{pmatrix} 0 & -\mathcal{M}^{-1} \\ \nabla^2 V(\bar{x}(Y)_s) & 0 \end{pmatrix}, \tag{29}$$

for all $s \in [t, T]$, and $\overline{X}(Y) = \begin{pmatrix} \bar{x}(Y) \\ \bar{p}(Y) \end{pmatrix}$.

Lemma 5 *Given $T \in \mathbb{R}_{>0}$, $t \in [0, T)$, the map $Y \mapsto U_{s,r}(Y)$ of (28) is twice Fréchet differentiable, uniformly in $r, s, \in [t, T]$.*

Remark 2 Regularity of the map $Y \mapsto U_{s,r}(Y)$, $s, r \in [t, T]$, of (28) is ultimately determined by regularity of V. If V is k-times Fréchet differentiable with $k \geq 2$, then $Y \mapsto U_{s,r}(Y)$ is $k - 2$ times Fréchet differentiable, uniformly in $r, s, \in [t, T]$. Note by (4) that $k \doteq 4$ is assumed throughout, so that $Y \mapsto U_{s,r}(Y)$ must be twice differentiable, as stated in Lemma 5. For further details, see the proof of Lemma 5 in Appendix A. □

By applying these lemmas to (23), Fréchet regularity of the cost $(t, Y) \mapsto \widetilde{J}_T(t, Y)$ may be demonstrated.

Proposition 1 *Given $T \in \mathbb{R}_{\geq 0}$, the map $(t, Y) \mapsto \widetilde{J}_T(t, Y)$ of (23) is continuously Fréchet differentiable with derivative $D\widetilde{J}_T$ given by*

$$D\widetilde{J}_T(t, Y)\,(\delta, h) = \big(D_t \widetilde{J}_T(t, Y)\,\delta,\ D_Y \widetilde{J}_T(t, Y)\, h\big), \tag{30}$$

where $D_t \widetilde{J}_T(t, Y)\,\delta = -l(\overline{X}(Y)_t)\,\delta$ and $D_Y \widetilde{J}_T(t, Y)\, h = \langle \nabla_Y \widetilde{J}_T(t, Y), h\rangle_{\mathscr{X}^2}$ for all $t \in [0, T)$, $Y, h \in \mathscr{X}^2$, $\delta \in (-t, T - t)$, in which $\nabla_Y \widetilde{J}_T(t, Y) \in \mathscr{X}^2$ is the corresponding Riesz representation of $D_Y \widetilde{J}_T(t, Y)$, given by

$$\nabla_Y \widetilde{J}_T(t, Y) = U_{T,t}(Y)'\, \nabla\Psi(\overline{X}(Y)_T) + \int_t^T U_{s,t}(Y)'\, \nabla l(\overline{X}(Y)_s)\, ds. \tag{31}$$

Moreover, the map $(t, Y) \mapsto D\widetilde{J}_T(t, Y)$ is also continuously Fréchet differentiable.

Proof Fix $T \in \mathbb{R}_{\geq 0}$, $t \in [0, T]$, $Y, h \in \mathscr{X}^2$, and $\delta \in (-t, T - t)$. Note immediately that $(t, Y) \mapsto \widetilde{J}_T(t, Y)$ is Fréchet differentiable if and only if $t \mapsto \widetilde{J}_T(t, Y)$ and $Y \mapsto \widetilde{J}_T(t, Y)$ are Fréchet differentiable, using for example the norm $\|(t, Y)\|^2 \doteq |t|^2 + \|Y\|^2_{\mathscr{X}^2}$. By inspection of (23), $t \mapsto \widetilde{J}_T(t, Y)$ is Fréchet differentiable, with the derivative indicated in the left-hand equality in (30).

In order to demonstrate that the map $Y \mapsto \widetilde{J}_T(t, Y)$ is Fréchet differentiable, with derivative as per the right-hand equality in (30), the chain rule for Fréchet differentiation [1] may be applied. To this end, in view of (23), define $\widetilde{I} : C([t, T]; \mathscr{X}^2) \to \mathbb{R}$ and $\widetilde{\iota} : C([t, T]; \mathscr{X}^2) \to \mathcal{L}(C([t, T]; \mathscr{X}^2); \mathbb{R})$ by

$$\begin{aligned} \widetilde{I}(Z) &\doteq \int_t^T l(Z_s)\, ds + \Psi(Z_T), \\ \iota(Z)\, \delta &\doteq \int_t^T \langle \nabla l(Z_s), \delta_s \rangle_{\mathscr{X}^2}\, ds + \langle \nabla \Psi(Z_T), \delta_T \rangle_{\mathscr{X}^2} \end{aligned} \tag{32}$$

for all $Z, \delta \in C([t, T]; \mathscr{X}^2)$, in which $\langle (v, x), (w, y) \rangle_{\mathscr{X}^2} \doteq \langle v, w \rangle_{\mathscr{X}} + \langle x, y \rangle_{\mathscr{X}}$ for all $v, w, x, y \in \mathscr{X}$. Note in particular that $\widetilde{J}_T(t, Y) = \widetilde{I} \circ \overline{X}(Y)$, with $\overline{X} \in C(\mathscr{X}^2; C([t, T]; \mathscr{X}^2))$ Fréchet differentiable by Lemma 4, and the candidate derivative of $z \mapsto D\widetilde{I}(Z)$ is $\iota(Z)$ in (32). Fix an arbitrary such $Z, \delta \in C([t, T]; \mathscr{X}^2)$. By inspection,

$$\begin{aligned} |\widetilde{I}(Z + \delta) - \widetilde{I}(Z) - \widetilde{\iota}(Z)\, \delta| \leq \int_t^T & |l(Z_s + \delta_s) - l(Z_s) - \langle \nabla l(Z_s), \delta_s \rangle_{\mathscr{X}^2}|\, ds \\ & + |\Psi(Z_T + \delta_T) - \Psi(Z_T) - \langle \nabla \Psi(Z_T), \delta_T \rangle_{\mathscr{X}^2}|. \end{aligned}$$

As $l, \Psi \in C^3(\mathscr{X}^2; \mathbb{R})$ by (4), (22), and $Dl(Y)\, h = \langle \nabla l(Y), h \rangle$, the mean value theorem implies that

$$\begin{aligned} |\widetilde{I}(Z + \delta) - \widetilde{I}(Z) - \widetilde{\iota}(Z)\, \delta| &\leq \int_t^T \left| \int_0^1 (1 - \eta)\, \langle \delta_s, \nabla^2 l(Z_s + \eta\, \delta_s)\, \delta_s \rangle_{\mathscr{X}^2}\, d\eta \right| ds \\ &\quad + \left| \int_0^1 (1 - \eta)\, \langle \delta_T, \nabla^2 \Psi(Z_T + \eta\, \delta_T)\, \delta_T \rangle_{\mathscr{X}^2}\, d\eta \right| \\ &\leq C \int_t^T \|\delta_s\|_{\mathscr{X}^2}^2\, ds + C\, \|\delta_T\|_{\mathscr{X}^2}^2 \leq C\, (T - t) \vee 1\, \|\delta\|_\infty^2 \end{aligned}$$

in which $C < \infty$ is given by

$$C \doteq \frac{1}{2} \sup_{Y \in \mathscr{X}^2} \max(\|\nabla^2 l(Y)\|_{\mathcal{L}(\mathscr{X}^2)}, \|\nabla^2 \Psi(Y)\|_{\mathcal{L}(\mathscr{X}^2)}),$$

and finiteness follows by (4), (22). Recalling (1) yields $|d\widetilde{I}_Z(\delta)| \leq C \max(T - t, 1)\, \|\delta\|_\infty$, i.e. $\widetilde{I}$ is Fréchet differentiable with derivative $D\widetilde{I} = \widetilde{\iota}$. Hence, $\widetilde{J}_T(t, Y) = \widetilde{I} \circ \overline{X}(Y)$, in which $\widetilde{I} : C([t, T]; \mathscr{X}^2) \to C([t, T]; \mathscr{X}^2)$ is Fréchet differentiable, as demonstrated above, and $\overline{X} \in C(\mathscr{X}^2; C([t, T]; \mathscr{X}^2)$ is Fréchet differentiable by Lemma 4. The chain rule, along with (27), (32), thus yield

$$\begin{aligned} D_Y \widetilde{J}_T(t, Y)\, h &= D\widetilde{I}(\overline{X}(Y))\, D\overline{X}(Y)\, h = \widetilde{\iota}(\overline{X}(Y))\, U_{\cdot,t}(Y)\, h \\ &= \int_t^T \langle \nabla l(\overline{X}(Y)_s), U_{s,t}(Y)\, h \rangle\, ds + \langle \nabla \Psi(\overline{X}(Y)_T), U_{T,t}(Y)\, h \rangle_{\mathscr{X}^2} \end{aligned}$$

$$= \left\langle \int_t^T U_{s,t}(Y)' \, \nabla l(\overline{X}(Y)_s) \, ds + U_{T,t}(Y)' \, \nabla \Psi(\overline{X}(Y)_T), \; h \right\rangle_{\mathscr{X}^2}$$
$$= \langle \nabla_Y \widetilde{J}_T(t, Y), \; h \rangle_{\mathscr{X}^2},$$

in which $\nabla_Y \widetilde{J}_T(t, Y)$ is as per the lemma statement. Hence, the right-hand equality in (30) holds.

It may be verified that $(t, Y) \mapsto D\widetilde{J}_T(t, Y)$ is continuous. In particular, by inspection of (30) and Lemma 3 that $(t, Y) \mapsto D_t \widetilde{J}_T(t, Y)$ is continuous, i.e. $l \in C^4(\mathscr{X}^2; \mathbb{R})$ by (4), (22), and $t \mapsto \overline{X}(Y)_t$ and $Y \mapsto \overline{X}(Y)_t$ are both continuous. Similarly, by inspection of (30), (31), $(t, Y) \mapsto D_Y \widetilde{J}_T(t, Y)$ is continuous as $Y \mapsto U_{s,t}(Y)$ is continuous, uniformly in $s \in [t, T]$, by Lemma 5.

Twice continuous Fréchet differentiability follows similarly, via (4), (27), (30), and Lemma 5. □

Proposition 1 and (23) may be combined directly to obtain corresponding Riesz representations for the Fréchet derivatives of the cost function $\bar{J}_T$ of (19), with respect to the generalized position and adjoint variable.

Proposition 2 *Given $T \in \mathbb{R}_{>0}$, $t \in [0, T)$, the maps $x \mapsto \bar{J}_T(t, x, p)$ and $p \mapsto \bar{J}_T(t, x, p)$ of (19) are Fréchet differentiable with derivatives given by*

$$\begin{aligned} D_x \bar{J}_T(t, x, p) &\in \mathcal{L}(\mathscr{X}; \mathbb{R}), \quad D_x \bar{J}_T(t, x, p) \, h = \langle \nabla_x \bar{J}_T(t, x, p), \; h \rangle, \\ D_p \bar{J}_T(t, x, p) &\in \mathcal{L}(\mathscr{X}; \mathbb{R}), \quad D_p \bar{J}_T(t, x, p) \, h = \langle \nabla_p \bar{J}_T(t, x, p), \; h \rangle, \end{aligned} \tag{33}$$

for all $x, p, h \in \mathscr{X}$, in which the Riesz representations are $\nabla_x \bar{J}_T(t, x, p) = \begin{pmatrix} \mathcal{I} & 0 \end{pmatrix} \nabla_Y \widetilde{J}_T \left(t, Y_p(x)\right)$ and $\nabla_p \bar{J}_T(t, x, p) = \begin{pmatrix} 0 & \mathcal{I} \end{pmatrix} \nabla_Y \widetilde{J}_T \left(t, Y_p(x)\right)$ respectively, with $\nabla \widetilde{J}_T(t, \cdot)$, $Y_p(x)$ as per (31), (22). Moreover, given $(x, p) \in \mathscr{X}^2$ and

$$\begin{pmatrix} \bar{x}_s \\ \bar{p}_s \end{pmatrix} \doteq \overline{X}(Y_p(x))_s \in \mathscr{X}^2, \quad \zeta_s \doteq \begin{pmatrix} \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s) \\ \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s) \end{pmatrix} \in \mathscr{X}^2, \tag{34}$$

the map $s \mapsto \zeta_s$ satisfies

$$\zeta_s = U_{T,s}(Y_p(x))' \, \nabla \Psi(\overline{X}(Y_p(x))_T) + \int_s^T U_{\sigma,s}(Y_p(x))' \, \nabla l(\overline{X}(Y_p(x))_\sigma) \, d\sigma, \tag{35}$$

for all $s \in [t, T]$. Equivalently, $\zeta_s = \begin{pmatrix} \bar{p}_s - \pi_s \\ \xi_s \end{pmatrix}$ for all $s \in [t, T]$, where $s \mapsto \begin{pmatrix} \pi_s \\ \xi_s \end{pmatrix}$ is the unique solution of the FVP

$$\begin{cases} \dot{\xi}_s = -\mathcal{M} \, \pi_s, \quad s \in [t, T], & \text{(36a)} \\ \dot{\pi}_s = \nabla^2 V(\bar{x}_s) \, \xi_s, & \text{(36b)} \\ \xi_T = 0, \quad \pi_T = \bar{p}_T - \nabla \psi(\bar{x}_T), & \text{(36c)} \end{cases}$$

c.f. (29).

Proof Fix arbitrary $T \in \mathbb{R}_{\geq 0}$, $t \in [0, T)$, $x, p, h \in \mathscr{X}$. By (19), (23), $\bar{J}_T(t, x, p) = \widetilde{J}_T\big(t, Y_p(x)\big)$, and note that the maps $x \mapsto Y_p(x)$ and $p \mapsto Y_p(x)$ are Fréchet differentiable with respective derivatives given by $D_x Y_p(x)\, h = (\mathcal{I}\ 0)'\, h$ and $D_p Y_p(x)\, h = (0\ \mathcal{I})'\, h$, with $0, \mathcal{I} \in \mathcal{L}(\mathscr{X})$ denoting the zero and identity maps. Applying the chain rule, and Proposition 1,

$$\begin{aligned} D_x \bar{J}_T(t, x, p)\, h = D_Y \widetilde{J}_T(t, Y_p(x))\, D_x Y_p(x)\, h &= \langle \nabla_Y \widetilde{J}_T(t, Y_p(x)),\, D_x Y_p(x)\, h \rangle_{\mathscr{X}^2} \\ &= \langle \big(\mathcal{I}\ 0\big)\, \nabla_Y \widetilde{J}_T(t, Y_p(x)),\ h \rangle, \end{aligned}$$

yielding the first asserted Riesz representation, with the other asserted representation following similarly.

For the remaining assertions (35), (36a), (36b), (36c), given (34), note that

$$\zeta_s = \nabla_Y \widetilde{J}_T(s, Y_{\bar{p}_s}(\bar{x}_s)) = \nabla_Y \widetilde{J}_T(s, \overline{X}(Y_p(x))_s),$$

so that (35) follows by Proposition 1. By inspection, $s \mapsto \zeta_s$ of (35) is differentiable, with the Leibniz integral rule yielding

$$\begin{aligned} \dot{\zeta}_s &= (\tfrac{\partial}{\partial s} U_{T,s}(Y_p(x)))'\, \nabla \Psi(\overline{X}(Y_p(x))_T) - \nabla l(\overline{X}(Y_p(x))_s) \\ &\quad + \int_s^T (\tfrac{\partial}{\partial s} U_{\sigma,s}(Y_p(x)))'\, \nabla l(\overline{X}(Y_p(x))_\sigma)\, d\sigma \\ &= -A(Y_p(x))_s'\, \zeta_s - \nabla l(\overline{X}(Y_p(x))_s) \\ &= \begin{pmatrix} 0 & -\nabla^2 V(\bar{x}_s) \\ \mathcal{M}^{-1} & 0 \end{pmatrix} \zeta_s + \begin{pmatrix} \nabla V(\bar{x}_s) \\ -\mathcal{M}^{-1}\, \bar{p}_s \end{pmatrix}, \end{aligned}$$

for all $s \in (t, T)$. Defining $s \mapsto (\pi_s,\ \xi_s)$ via $\begin{pmatrix} \bar{p}_s - \pi_s \\ \xi_s \end{pmatrix} = \zeta_s$ as per the hypothesis, it follows that

$$\begin{aligned} \dot{\xi}_s &= \mathcal{M}^{-1}(\bar{p}_s - \pi_s) - \mathcal{M}^{-1}\, \bar{p}_s = -\mathcal{M}^{-1}\, \pi_s, \\ \dot{\pi}_s &= \dot{\bar{p}}_s - [\dot{\bar{p}}_s - \dot{\pi}_s] = \nabla V(\bar{x}_s) - [-\nabla^2 V(\bar{x}_s)\, \xi_s + \nabla V(\bar{x}_s)] = \nabla^2 V(\bar{x}_s)\, \xi_s, \end{aligned}$$

for all $s \in [t, T]$. That is, (36a), (36b) hold.

Moreover, as $\mathcal{U}_{T,T}(Y_p(x)) = \mathcal{I}$, and Ψ is as per (22), note by (35) that

$$\zeta_T = \nabla \Psi(\overline{X}(Y_p(x))_T) = \begin{pmatrix} \nabla \psi(\bar{x}_T) \\ 0 \end{pmatrix},$$

i.e. the terminal conditions (36c) hold. Uniqueness of this solution follows via arguments as per Lemma 2. □

3.3 Characterization of Stationary Trajectories

With the Fréchet derivative of the cost $\bar{J}_T$ of (19) given by Proposition 2, the argstat in (21) may be further investigated with a view to characterizing the argstat (14) underlying the stationary control problem (13), via Theorem 2. To this end, given $T \in \mathbb{R}_{>0}$, $t \in [0, T]$, $x \in \mathscr{X}$, observe by Proposition 2 that

$$p \in \operatorname*{arg\,stat}_{q \in \mathscr{X}} \bar{J}_T(t, x, q) \iff \begin{cases} \textit{the unique solution } s \mapsto (\xi_s, \pi_s) \textit{ of FVP } (36) \textit{ satisfies} \\ \qquad 0 = \nabla_p \bar{J}_T(t, x, p) = \nabla_p \bar{J}_T(t, \bar{x}_t, \bar{p}_t) = \xi_t. \end{cases}$$

Note that if the left-hand argstat condition holds, the implication is that the unique solution $s \mapsto (\xi_s, \pi_s)$ of FVP (36) satisfies its terminal conditions (36c) and an *additional* boundary condition $\xi_t = 0$. However, the totality of these conditions, i.e. $\xi_t = 0 = \xi_T$ and $\pi_T = \bar{p}_T - \nabla\psi(\bar{x}_T)$, is insufficient to guarantee that $0 = \pi_T = \bar{p}_T - \nabla\psi(\bar{x}_T)$, as required by Theorem 2 in order to characterize the argstat (14). In particular, the TPBVP defined by (36a), (36b), and the boundary conditions $\xi_t = 0 = \xi_T$, does not necessarily have the trivial solution as its only solution. However, this is the case if $0 = \pi_t = p - \nabla_x \bar{J}_T(t, x, p)$ is also imposed.

This motivates an additional argstat condition involving the cost $\bar{J}_T$ of (19), allowing the argstat in (13), (14) to be characterized via the argstat in (21), as formalized by the following lemma and theorem.

Lemma 6 *Given $T \in \mathbb{R}_{>0}$, $t \in [0, T)$, $x, p \in \mathscr{X}$, and $(\bar{x}_s, \bar{p}_s) \doteq \overline{X}(Y_p(x))_s$ for all $s \in [t, T]$, the following statements are equivalent:*

(i) $0 = \nabla_p \bar{J}_T(t, x, p)$ *and* $p = \nabla_x \bar{J}_T(t, x, p)$*;*
(ii) $0 = \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$ *for all* $s \in [t, T]$*;*
(iii) $\bar{p}_s = \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$ *for all* $s \in [t, T]$.

Proof Fix $T \in \mathbb{R}_{>0}$, $t \in [0, T)$, $x, p \in \mathscr{X}$. By Proposition 2, recall that

$$\begin{pmatrix} \xi_s \\ \pi_s \end{pmatrix} \doteq \begin{pmatrix} \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s) \\ \bar{p}_s - \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s) \end{pmatrix} \tag{37}$$

satisfies (36a), (36b) for all $s \in (t, T)$.

(i) $\Longrightarrow$ (ii), (iii): Suppose *(i)* holds, i.e. $0 = \nabla_p \bar{J}_T(t, x, p)$, $p = \nabla_x \bar{J}_T(t, x, p)$. Consequently, selecting $s = t$ in (37), $\xi_t = 0 = \pi_t$. Hence, by (36a), (36b), it follows that $\xi_s = 0 = \pi_s$ for all $s \in [t, T]$, so that by (37), $0 = \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$ and $\bar{p}_s = \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$ for all $s \in [t, T]$. That is, both *(ii)* and *(iii)* hold.

(iii) $\Longrightarrow$ (ii): Suppose *(iii)* holds, i.e. $0 = \bar{p}_s - \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s) \doteq \pi_s$ for all $s \in [t, T]$. Then, by (36a), $\dot{\xi}_s = -\mathcal{M}^{-1} \pi_s = 0$ for all $s \in [t, T]$. Moreover, selecting $s = T$ in (37) yields $\xi_T = \nabla_p \bar{J}_T(T, \bar{x}_T, \bar{p}_T) = \nabla_p \psi(\bar{x}_T) = 0$, see (19). As $\dot{\xi}_s = 0$ for all $s \in [t, T]$ and $\xi_T = 0$, integration yields $\xi_s = 0$ for all $s \in [t, T]$. That is, by (37), $0 = \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$, and so *(ii)* holds.

(ii) $\Longrightarrow$ *(iii):* Suppose *(ii)* holds, i.e. $0 = \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s) \doteq \xi_s$ for all $s \in [t, T]$. Then, by (36a), $-\mathcal{M}^{-1} \pi_s = \dot{\xi}_s = 0$, so that $0 = \pi_s = \bar{p}_s - \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$ for all $s \in [t, T]$. That is, *(iii)* holds.

(ii) $\Longrightarrow$ *(i):* Suppose *(ii)* holds. By the preceding implication, *(iii)* also holds. Selecting $s = t$ in both *(ii)* and *(iii)* yields $0 = \nabla_p \bar{J}_T(t, \bar{x}_t, \bar{p}_t) = \nabla_p \bar{J}_T(t, x, p)$ and $p = \bar{p}_t = \nabla_x \bar{J}_T(t, \bar{x}_t, \bar{p}_t) = \nabla_x \bar{J}_T(t, x, p)$ respectively. That is, *(i)* holds. □

Theorem 3 *Given $T \in \mathbb{R}_{\geq 0}$, $t \in [0, T]$, and $x \in \mathscr{X}$, the following statements are equivalent:*

(i) there exists $\bar{u} \in \arg\operatorname{stat}_{u \in \mathscr{U}[t,T]} J_T(t, x, u)$, as per (14);
(ii) there exists $p \in \arg\operatorname{stat}_{q \in \mathscr{X}} \bar{J}_T(t, x, q)$ such that $(\bar{x}_s, \bar{p}_s) \doteq \overline{X}(Y_p(x))_s$ satisfies

$$\bar{p}_s \in \underset{q \in \mathscr{X}}{\arg\operatorname{stat}}\, \bar{J}_T(s, \bar{x}_s, q) \quad \forall\, s \in [t, T].$$

Moreover, $\bar{u}$ of (i) and $(\bar{x}, \bar{p}) \doteq \overline{X}(Y_p(x))$ satisfy (15), i.e. $\bar{u}_s = -\mathcal{M}^{-1} \bar{p}_s$ a.e. $s \in [t, T]$.

Proof Fix arbitrary $T \in \mathbb{R}_{\geq 0}$, $t \in [0, T]$, and $x \in \mathscr{X}$.

(ii) $\Rightarrow$ *(i):* Suppose *(ii)* holds, i.e. there exists $p \in \arg\operatorname{stat}_{q \in \mathscr{X}} \bar{J}_T(t, x, q)$ such that $(\bar{x}_s, \bar{p}_s) \doteq \overline{X}(Y_p(x))_s$ satisfies $0 = \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$ for all $s \in [t, T]$. That is, assertion *(ii)* of Lemma 6 holds, so that the equivalent assertion *(iii)* also holds, i.e. $\bar{p}_s = \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$ for all $s \in [t, T]$. In particular, $\bar{p}_T = \nabla_x \bar{J}_T(T, \bar{x}_T, \bar{p}_T) = \nabla\psi(\bar{p}_T)$. Hence, $s \mapsto (\overline{X}(Y_p(x))_s, \bar{z}_s)$, $s \in [t, T]$, c.f. (16), solves the TPBVP defined by FVP (11) with $y \in \bar{x}_T$ and $\bar{x}_t = x$, and so Theorem 2 implies the existence of $\bar{u} \in \arg\operatorname{stat}_{u \in \mathscr{U}[t,T]} J_T(t, x, u)$, with its explicit form (15).

(i) $\Rightarrow$ *(ii):* Suppose *(i)* holds, i.e. there exists $\bar{u} \in \arg\operatorname{stat}_{u \in \mathscr{U}[t,T]} J_T(t, x, u)$. By Theorem 2 and (16), there exists a solution $s \mapsto (\bar{x}_s, \bar{p}_s, \bar{z}_s)$ to the TPBVP defined by FVP (11) with $y = \bar{x}_T$ and $\bar{x}_t = x$. Let $p \doteq \bar{p}_t$, and note that $(\bar{x}_s, \bar{p}_s) \doteq \overline{X}(Y_p(x))_s$, $s \in [t, T]$. As $\bar{p}_T = \nabla\psi(\bar{x}_T)$ by (11b), (11d), FVP (36) has the trivial solution as its unique solution, i.e. $\xi_s = 0 = \pi_s$ for all $s \in [t, T]$. Consequently, by Proposition 2, i.e. (34),

$$\begin{pmatrix} \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s) \\ \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s) \end{pmatrix} = \zeta_s = \begin{pmatrix} \bar{p}_s - \pi_s \\ \xi_s \end{pmatrix} = \begin{pmatrix} \bar{p}_s \\ 0 \end{pmatrix}, \qquad s \in [t, T], \tag{38}$$

so that $0 = \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$, i.e. $\bar{p}_s \in \arg\operatorname{stat}_{q \in \mathscr{X}} \bar{J}_T(s, \bar{x}_s, q)$, for all $s \in [t, T]$. □

Theorem 4 *Given $T \in \mathbb{R}_{>0}$, $t \in [0, T)$, and $x \in \mathscr{X}$, the following statements are equivalent:*

(i) there exists $\bar{u} \in \arg\operatorname{stat}_{u \in \mathscr{U}} J_T(t, x, u)$, as per (14);
(ii) there exists $p \in \mathscr{X}$ such that

$$p \in \underset{q \in \mathscr{X}}{\arg\operatorname{stat}}\, \bar{J}_T(t, x, q), \qquad x \in \underset{y \in \mathscr{X}}{\arg\operatorname{stat}} \left\{ \langle y, p \rangle - \bar{J}_T(t, y, p) \right\}. \tag{39}$$

Moreover, $\bar{u}$ *of (i) and* $(\bar{x}, \bar{p}) \doteq \overline{X}(Y_p(x))$ *satisfy* (15)*, i.e.* $\bar{u}_s = -\mathcal{M}^{-1}\,\bar{p}_s$ *a.e.* $s \in [t, T]$.

Proof Fix $T \in \mathbb{R}_{>0}$, $t \in [0, T)$, $x \in \mathscr{X}$.

(i) ⇒ *(ii):* Suppose that *(i)* holds, i.e. there exists $\bar{u} \in \arg\operatorname{stat}_{u\in\mathscr{U}} J_T(t, x, u)$. That is, assertion *(i)* of Theorem 3 holds, so that the equivalent assertion *(ii)* of Theorem 3 also holds, i.e. there exists $p \in \mathscr{X}$ such that $(\bar{x}_s, \bar{p}_s) \doteq \overline{X}(Y_p(x))_s$ satisfies $\bar{p}_s \in \arg\operatorname{stat}_{q\in\mathscr{X}} \bar{J}_T(t, \bar{x}_s, q)$ for all $s \in [t, T]$. Hence, $s \mapsto (\bar{x}_s, \bar{p}_s)$ satisfies assertion *(iii)* of Lemma 6, so that the equivalent assertion *(i)* of Lemma 6 also holds, yielding (39). That is, *(ii)* holds.

(ii) ⇒ *(i):* Suppose that *(ii)* holds. Reversing the sequence of implications provided by the equivalences in Lemma 6 and Theorem 3 in the above argument yields *(i)*.

□

3.4 Verification via an HJB PDE

A verification theorem is provided for the cost $\bar{J}_T$ of (19), (21), formulated with respect to the extended Hamiltonian $\bar{H} : \mathscr{X}^2 \times \mathscr{X}^2 \to \mathbb{R}$ defined by

$$\bar{H}(x, p, \hat{\pi}, \xi) \doteq -\tfrac{1}{2}\,\langle p,\, \mathcal{M}^{-1}\, p\rangle + V(x) + \langle \hat{\pi},\, \mathcal{M}^{-1}\, p\rangle - \langle \xi,\, \nabla V(x)\rangle \tag{40}$$

for all $x, p, \hat{\pi}, \xi \in \mathscr{X}$.

Theorem 5 *Given* $T \in \mathbb{R}_{>0}$ *suppose there exists a* $W \in C([0, T] \times \mathscr{X}^2; \mathbb{R}) \cap C^1((0, T) \times \mathscr{X}^2; \mathbb{R})$ *satisfying the HJB PDE and terminal condition*

$$\begin{cases} -\dfrac{\partial W}{\partial t}(t, x, p) + \bar{H}(x, p, \nabla_x W(t, x, p), \nabla_p W(t, x, p)) = 0, \\ \quad (t, x, p) \in (0, T) \times \mathscr{X}^2, \\ W(T, x, p) = \psi(x), \quad (x, p) \in \mathscr{X}^2, \end{cases} \tag{41}$$

in which $\bar{H}$ *is as per* (40)*, and* ψ *is the terminal cost appearing in* (2)*,* (19)*. Then,* $\bar{J}_T(t, x, p) = W(t, x, p)$ *for all* $t \in (0, T)$*,* $x, p \in \mathscr{X}$*, where* $\bar{J}_T$ *is as per* (19)*.*

Conversely, $\bar{J}_T$ *of* (19) *always satisfies* (41)*, and is consequently its unique solution.*

Proof Fix $T \in \mathbb{R}_{>0}$, $t \in (0, T)$, and let W be as per the theorem statement. Fix any $x, p \in \mathscr{X}$. With $\overline{X}$ as per Lemmas 2, 3, 4, let $(\bar{x}_s, \bar{p}_s) \doteq \overline{X}(Y_p(x))_s$ for all $s \in [t, T]$. Note in particular that $s \mapsto (\bar{x}_s, \bar{p}_s)$ is a classical solution of the IVP (20), (24). Hence, by the asserted regularity of W, $s \mapsto W(s, \bar{x}_s, \bar{p}_s)$ is differentiable, so that the chain rule and (41) yield

$$\begin{aligned} \tfrac{d}{ds} W(s, \bar{x}_s, \bar{p}_s) &= \tfrac{\partial}{\partial s} W(s, \bar{x}_s, \bar{p}_s) + \langle \nabla_x W(s, \bar{x}_s, \bar{p}_s),\ \dot{\bar{x}}_s\rangle + \langle \nabla_p W(s, \bar{x}_s, \bar{p}_s),\ \dot{\bar{p}}_s\rangle \\ &= -[-\tfrac{\partial}{\partial s} W(s, \bar{x}_s, \bar{p}_s) + \bar{H}(x, p, \nabla_x W(s, \bar{x}_s, \bar{p}_s), \nabla_p W(s, \bar{x}_s, \bar{p}_s))] \\ &\quad + \bar{H}(x, p, \nabla_x W(s, \bar{x}_s, \bar{p}_s), \nabla_p W(s, \bar{x}_s, \bar{p}_s)) \\ &\quad + \langle \nabla_x W(s, \bar{x}_s, \bar{p}_s),\ -\mathcal{M}^{-1}\,\bar{p}_s\rangle + \langle \nabla_p W(s, \bar{x}_s, \bar{p}_s),\ \nabla V(\bar{x}_s)\rangle \end{aligned}$$

$$
\begin{aligned}
&= -\tfrac{1}{2}\,\langle \bar{p}_s,\, \mathcal{M}^{-1}\,\bar{p}_s\rangle + V(\bar{x}_s) + \langle \nabla_x W(s, \bar{x}_s, \bar{p}_s),\, \mathcal{M}^{-1}\,\bar{p}_s\rangle \\
&\quad - \langle \nabla_p W(s, \bar{x}_s, \bar{p}_s),\, \nabla V(\bar{x}_s)\rangle \\
&\quad + \langle \nabla_x W(s, \bar{x}_s, \bar{p}_s),\, -\mathcal{M}^{-1}\,\bar{p}_s\rangle + \langle \nabla_p W(s, \bar{x}_s, \bar{p}_s),\, \nabla V(\bar{x}_s)\rangle \\
&= -\tfrac{1}{2}\,\langle \bar{p}_s,\, \mathcal{M}^{-1}\,\bar{p}_s\rangle + V(\bar{x}_s),
\end{aligned}
$$

for all $s \in (t, T)$. Integrating with respect to $s \in (t, T)$, and recalling the boundary condition in (41), subsequently yields

$$
\begin{aligned}
\psi(\bar{x}_T) - W(t, x, p) &= W(T, \bar{x}_T, \bar{p}_T) - W(t, x, p) \\
&= \int_t^T -\tfrac{1}{2}\,\langle \bar{p}_s,\, \mathcal{M}^{-1}\,\bar{p}_s\rangle + V(\bar{x}_s)\, ds.
\end{aligned}
$$

Rearranging, and recalling (19), yields the asserted equality $\bar{J}_T(t, x, p) = W(t, x, p)$. Recalling that $t \in (0, T)$, $x, p \in \mathscr{X}$ are arbitrary yields the first assertion.

For the converse, note by Proposition 1 that $\bar{J}_T \in C([0, T] \times \mathscr{X}^2; \mathbb{R}) \cap C^1((0, T) \times \mathscr{X}^2; \mathbb{R})$. Fix $x, p \in \mathscr{X}$. Note by (19) that $\bar{J}_T(T, x, p) = \psi(x)$, so that the terminal condition in (41) trivially holds. Fix $t \in (0, T)$, and let $(\bar{x}_s, \bar{p}_s) \doteq \overline{X}(Y_p(x))_s$ for all $s \in [t, T]$. Fix $r \in (t, T]$. By (19),

$$
\begin{aligned}
\bar{J}_T(t, x, p) &= \int_t^r +\tfrac{1}{2}\,\langle \bar{p}_s,\, \mathcal{M}^{-1}\,\bar{p}_s\rangle - V(\bar{x}_s)\, ds \\
&\quad + \int_r^T +\tfrac{1}{2}\,\langle \bar{p}_s,\, \mathcal{M}^{-1}\,\bar{p}_s\rangle - V(\bar{x}_s)\, ds - \psi(\bar{x}_T) \\
&= \int_t^r +\tfrac{1}{2}\,\langle \bar{p}_s,\, \mathcal{M}^{-1}\,\bar{p}_s\rangle - V(\bar{x}_s)\, ds + \bar{J}_T(r, \bar{x}_r, \bar{p}_r).
\end{aligned}
$$

Dividing through by $r - t$ and sending $r \to t^+$, Proposition 1 implies that

$$
\begin{aligned}
&-\tfrac{1}{2}\,\langle p,\, \mathcal{M}^{-1}\,p\rangle + V(x) = \tfrac{d}{dt}\bar{J}_T(t, \bar{x}_t, \bar{p}_t) \\
&= \tfrac{\partial}{\partial t}\bar{J}_T(t, x, p) + \langle \nabla_x \bar{J}_T(t, x, p),\, -\mathcal{M}^{-1}\,p\rangle + \langle \nabla_p \bar{J}_T(t, x, p),\, \nabla V(x)\rangle,
\end{aligned}
$$

i.e. $\bar{J}_T$ satisfies (41), and uniqueness follows by the first assertion. □

Verification Theorem 5 may be used to restate the characterization of stationary controls provided by Theorem 4. Note in particular that a form of state feedback characterization of the stationary control is inherited from the earlier theorem, which mirrors the corresponding characterization provided by a standard verification theorem for optimal control.

Theorem 6 *Given $T \in \mathbb{R}_{>0}$, suppose there exists a $W \in C([0, T] \times \mathscr{X}^2; \mathbb{R}) \cap C^1((0, T) \times \mathscr{X}^2; \mathbb{R})$ satisfying the HJB PDE and terminal condition (41). Suppose further that, given $t \in [0, T]$ and $x \in \mathscr{X}$, there exists a $p \in \mathscr{X}$ such that*

$$p \in \underset{q \in \mathscr{X}}{\arg\operatorname{stat}}\, W(t,x,q), \qquad x \in \underset{y \in \mathscr{X}}{\arg\operatorname{stat}} \{\langle y, p\rangle - W(t,y,p)\}. \tag{42}$$

Then, there exists $\bar{u} \in \arg\operatorname{stat}_{u \in \mathscr{U}[t,T]} J_T(t,x,u)$ *such that*

$$\begin{aligned} \bar{x}_s &= x + \int_t^s \bar{u}_\sigma \, d\sigma, \quad \bar{p}_s \in \underset{q \in \mathscr{X}}{\arg\operatorname{stat}}\, W(s,\bar{x}_s,q), \\ \bar{u}_s &= -\mathcal{M}^{-1} \nabla_x W(s,\bar{x}_s,\bar{p}_s) = -\mathcal{M}^{-1}\,\bar{p}_s, \end{aligned} \tag{43}$$

for all $s \in [t,T]$.

Proof Fix $T \in \mathbb{R}_{>0}$, $W \in C([0,T]\times\mathscr{X}^2;\mathbb{R}) \cap C^1((0,T)\times\mathscr{X}^2;\mathbb{R})$, $t \in [0,T]$, $x \in \mathscr{X}$, as per the theorem statement. Suppose that $p \in \mathscr{X}$ exists such that (42) holds. Observe by Theorem 5 that $\bar{J}_T \equiv W$. Hence, by (42) and Theorem 4, there exists $\bar{u} \in \arg\operatorname{stat}_{u\in\mathscr{U}[t,T]} J_T(t,x,u)$ satisfying $\bar{u}_s = -\mathcal{M}^{-1}\,\bar{p}_s$ for all $s \in [t,T]$. Moreover, by Lemma 6, $\bar{p}_s = \nabla_x \bar{J}_T(s,\bar{x}_s,\bar{p}_s) = \nabla_x W(s,\bar{x}_s,\bar{p}_s)$, and $0 = \nabla_p \bar{J}_T(s,\bar{x}_s,\bar{p}_s) = \nabla_p W(t,\bar{x}_s,\bar{p}_s)$ for all $s \in [t,T]$, so that (43) holds. □

Remark 3 The characteristic system for (41) is the FVP given by

$$\begin{cases} \dot{\hat{x}}_s = -\nabla_{\hat{\pi}} \bar{H}(\hat{x}_s,\hat{p}_s,\hat{\pi}_s,\hat{\xi}_s) = -\mathcal{M}^{-1}\,\hat{p}_s, \quad s\in[t,T], & \text{(44a)}\\ \dot{\hat{p}}_s = -\nabla_{\hat{\xi}} \bar{H}(\hat{x}_s,\hat{p}_s,\hat{\pi}_s,\hat{\xi}_s) = \nabla V(\hat{x}_s), & \text{(44b)}\\ \dot{\hat{\pi}}_s = \nabla_{\hat{x}} \bar{H}(\hat{x}_s,\hat{p}_s,\hat{\pi}_s,\hat{\xi}_s) = \nabla V(\hat{x}_s) - \nabla^2 V(\hat{x}_s)\,\hat{\xi}_s, & \text{(44c)}\\ \dot{\hat{\xi}}_s = \nabla_{\hat{p}} \bar{H}(\hat{x}_s,\hat{p}_s,\hat{\pi}_s,\hat{\xi}_s) = -\mathcal{M}^{-1}(\hat{p}_s - \hat{\pi}_s), & \text{(44d)}\\ \dot{\hat{z}}_s = -\langle (\hat{\pi}_s,\hat{\xi}_s),\, (\nabla_{\hat{\pi}} \bar{H}(\hat{x}_s,\hat{p}_s,\hat{\pi}_s,\hat{\xi}_s), \nabla_{\hat{\xi}} \bar{H}(\hat{x}_s,\hat{p}_s,\hat{\pi}_s,\hat{\xi}_s)\rangle_{\mathscr{X}^2} & \\ \qquad\quad + \bar{H}(\hat{x}_s,\hat{p}_s,\hat{\pi}_s,\hat{\xi}_s) & \\ \quad = V(\hat{x}_s) - \tfrac{1}{2}\langle \hat{p}_s,\, \mathcal{M}^{-1}\,\hat{p}_s\rangle, & \text{(44e)}\\ \hat{x}_T = y, \quad \hat{p}_T = q, \quad \hat{\pi}_T = \nabla\psi(y), \quad \hat{\xi}_T = 0, \quad \hat{z}_T = \psi(y), & \\ \qquad y, q \in \mathscr{X}. & \text{(44f)} \end{cases}$$

Fix $y \in \mathscr{X}$ and $q \doteq \nabla\psi(y) \in \mathscr{X}$, and let the respective solutions of FVP (11) and FVP (44a), (44b), (44e), (44f) be denoted by $s \mapsto (\bar{x}_s,\bar{p}_s,\bar{z}_s)$ and $s \mapsto (\hat{x}_s,\hat{p}_s,\hat{z}_s)$ for all $s \in [t,T]$. Note that they are identical, by choice of q. Note further that the remaining equations in (44) can be written as the FVP

$$\begin{cases} \dot{\hat{p}}_s - \dot{\hat{\pi}}_s = \nabla^2 V(\bar{x}_s)\,\hat{\xi}_s, \\ \dot{\hat{\xi}}_s = -\mathcal{M}^{-1}(\hat{p}_s - \hat{\pi}_s), \\ \quad \hat{p}_T - \hat{\pi}_T = 0, \quad \hat{\xi}_T = 0, \end{cases}$$

which has the trivial solution as its unique solution, i.e. $\hat{p}_s - \hat{\pi}_s = 0 = \hat{\xi}_s$ for all $s \in [t,T]$. Moreover, by inspection of (10), (40), $H(\bar{x}_s,\bar{p}_s) = \bar{H}(\hat{x}_s,\hat{p}_s,\hat{\pi}_s,\hat{\xi}_s)$ for all $s \in [t, T]$. That is, the characteristic systems and Hamiltonians for the optimal and stationary control problems coincide. Recall by Theorem 5 that the HJB PDE (41) has a unique classical solution $W = \bar{J}_T$. Interpreting the associated characteristics as satisfying $\hat{\pi}_s = \nabla_x W(s, \hat{x}_s, \hat{p}_s)$ and $\hat{\xi}_s = \nabla_p W(s, \hat{x}_s, \hat{p}_s)$, note that $\bar{p}_s = \hat{\pi}_s = \nabla_x W(s, \bar{x}_s, \bar{p}_s) = \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$ and $0 = \hat{\xi}_s = \nabla_p W(s, \bar{x}_s, \bar{p}_s) = \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$ for all $s \in [t, T]$, which is equivalent to (39), (42) by Lemma 6. □

Remark 4 Theorem 5 provides a characterization of the cost $\bar{J}_T$ of (19) via the HJB PDE (41). With this characterization available, it is possible to provide an auxiliary statement of Proposition 2 that yields corresponding assertions, and in particular that (36a), (36b) hold. This auxiliary statement appears in Appendix B. □

4 A One-Dimensional Example

A one-dimensional linear mass-spring system consists of a mass $\mathcal{M} \doteq m \in \mathbb{R}_{>0}$ located at position $x \in \mathscr{X} \doteq \mathbb{R}$ whose motion is a consequence of a quadratic potential field $V : \mathbb{R} \to \mathbb{R}_{\geq 0}$, $V(x) \doteq \frac{1}{2} \kappa x^2$, $x \in \mathbb{R}$. Suppose that an initial velocity $\dot{x}_t \in \mathbb{R}$ of this mass is sought, at a fixed initial time t, so as to achieve a particular terminal velocity $\dot{x}_T = v \in \mathbb{R}$ later, at a fixed final time T. The corresponding terminal cost $\psi : \mathbb{R} \to \mathbb{R}$ in (19), (23) encapsulating this requirement is defined by $\psi(x) \doteq -m\, v\, x$ for all $x \in \mathbb{R}$, i.e. so that $\bar{p}_T = \nabla\psi(\bar{x}_T) = -m\, v$ in (11), (20), (44). Observe by inspection that assumption (4) holds, with $K \doteq 2\kappa$. With a view to determining the initial velocity required, an explicit solution to the HJB PDE (41) is constructed, and Theorem 6 subsequently applied. To this end, fix $t \in [0, T]$ and define

$$\Sigma \doteq \begin{pmatrix} \kappa & 0 \\ 0 & -\frac{1}{m} \end{pmatrix}, \qquad \Gamma \doteq \begin{pmatrix} 0 & -\frac{1}{m} \\ \kappa & 0 \end{pmatrix}. \tag{45}$$

Recalling (40), (45), note that the HJB PDE (41) may be written as

$$\begin{aligned} 0 &= -\frac{\partial W}{\partial s}(s, x, p) + \tfrac{1}{2}\left\langle \begin{pmatrix} x \\ p \end{pmatrix}, \Sigma \begin{pmatrix} x \\ p \end{pmatrix} \right\rangle - \left\langle \begin{pmatrix} \nabla_x W(s, x, p) \\ \nabla_p W(s, x, p) \end{pmatrix}, \Gamma \begin{pmatrix} x \\ p \end{pmatrix} \right\rangle \\ &= -\frac{\partial W}{\partial s}(s, Y) + \tfrac{1}{2}\langle Y, \Sigma Y\rangle - \langle \nabla_Y W(s, Y), \Gamma Y\rangle, \end{aligned} \tag{46}$$

for all $s \in (t, T)$, $Y \doteq Y_p(x) = (x, p) \in \mathbb{R}^2$. Define a solution candidate $\breve{W} : [t, T] \times \mathbb{R}^2 \to \mathbb{R}$ by

$$\breve{W}(s, Y) \doteq \tfrac{1}{2}\langle Y, P_s Y\rangle + \langle Q_s, Y\rangle, \tag{47}$$

in which

$$
\begin{aligned}
P_s &\doteq - \int_s^T \exp(\Gamma' (\sigma - s))\, \Sigma\, \exp(\Gamma (\sigma - s))\, d\sigma, \\
Q_s &\doteq \exp(\Gamma' (T - s)) \begin{pmatrix} -m\, v \\ 0 \end{pmatrix},
\end{aligned} \tag{48}
$$

for all $s \in (t, T)$, $Y \in \mathbb{R}^2$. Applying Leibniz, note that

$$
\dot{P}_s = \Sigma - \Gamma' P_s - P_s \Gamma, \qquad \dot{Q}_s = -\Gamma' Q_s, \tag{49}
$$

for all $s \in (t, T)$. Differentiating (47) yields

$$
\tfrac{\partial}{\partial s} \breve{W}(s, Y) = \tfrac{1}{2} \langle Y,\, \dot{P}_s Y \rangle + \langle \dot{Q}_s,\, Y \rangle, \qquad \nabla_Y \breve{W}(s, Y) = P_s Y + Q_s.
$$

Substituting these derivatives in the right-hand side of (46), and applying (49), subsequently yields

$$
\begin{aligned}
&- \frac{\partial \breve{W}}{\partial s}(s, Y) + \tfrac{1}{2} \langle Y,\, \Sigma Y \rangle - \langle \nabla_Y \breve{W}(s, Y),\, \Gamma Y \rangle \\
&= -\tfrac{1}{2} \langle Y,\, \dot{P}_s Y \rangle - \langle \dot{Q}_s,\, Y \rangle + \tfrac{1}{2} \langle Y,\, \Sigma Y \rangle - \langle P_s Y + Q_s,\, \Gamma Y \rangle \\
&= \tfrac{1}{2} \langle Y,\, [-\dot{P}_s + \Sigma - P_s \Gamma - \Gamma' P_s]\, Y \rangle + \langle -\dot{Q}_s - \Gamma' Q_s,\, Y \rangle \\
&= 0,
\end{aligned}
$$

for all $s \in (t, T)$, $Y \in \mathbb{R}^2$. Note further that $\breve{W}(T, x, p) = \breve{W}(T, Y) = \langle Q_T,\, Y \rangle = -m\, v\, x = \psi(x)$. That is $\breve{W}$ of (47) is a solution of (41). Hence, by Theorem 5, the cost $\bar{J}_T(s, x, p)$ of (19) is given explicitly by $\bar{J}_T(s, x, p) = \breve{W}(s, Y_p(x))$ for all $s \in [t, T]$. Diagonalizing Γ and integrating (48) yields

$$
\begin{aligned}
P_s &= \tfrac{1}{2} \begin{pmatrix} -\frac{\kappa}{\omega} \sin(2\,\omega\,(T - s)) & 1 - \cos(2\,\omega (T - s)) \\ 1 - \cos(2\,\omega\,(T - s)) & \frac{1}{m\,\omega} \sin(2\,\omega\,(T - s)) \end{pmatrix}, \qquad \omega \doteq \sqrt{\frac{\kappa}{m}}, \\
Q_s &= -m\, v \begin{pmatrix} \cos(\omega\,(T - s)) \\ -\frac{\omega}{\kappa} \sin(\omega\,(T - s)) \end{pmatrix}.
\end{aligned}
$$

With a view to illustrating Theorem 6, fix $x \in \mathbb{R}$, and note that

$$
\begin{aligned}
\nabla_x \breve{W}(t, x, p) &= (\,1\;\;0\,)\, \nabla_Y \breve{W}(t, Y_p(x)) = (\,1\;\;0\,) \left(P_t \begin{pmatrix} x \\ p \end{pmatrix} + Q_t \right) \\
&= -\tfrac{\kappa}{2\omega} \sin(2\,\omega\,(T - t))\, x + \tfrac{1}{2} [1 - \cos(2\,\omega (T - t))]\, p - m\, v\, \cos(\omega\,(T - t)), \\
\nabla_p \breve{W}(t, x, p) &= (\,0\;\;1\,)\, \nabla_Y \breve{W}(t, Y_p(x)) = (\,0\;\;1\,) \left(P_t \begin{pmatrix} x \\ p \end{pmatrix} + Q_t \right) \\
&= \tfrac{1}{2} [1 - \cos(2\,\omega\,(T - t))]\, x + \tfrac{1}{2\,m\,\omega} \sin(2\,\omega\,(T - t))\, p + m\, v\, (\tfrac{\omega}{\kappa}) \sin(\omega\,(T - t)),
\end{aligned} \tag{50}
$$

for all $p \in \mathbb{R}$. Note that $m\,\omega = \sqrt{\kappa\, m} = \frac{\kappa}{\omega}$. Motivated by (42), let $p \in \mathbb{R}$ be such that $0 = \nabla_p \breve{W}(t, x, p)$ and $p = \nabla_x \breve{W}(t, x, p)$ via (50). Collecting and simplifying these two equations via double angle formulae subsequently yields

$$\Omega_t \begin{pmatrix} x\sqrt{\kappa\, m} \\ p \end{pmatrix} = \Theta_t, \tag{51}$$

in which

$$\Omega_t \doteq \begin{pmatrix} \sin^2(\omega\,(T-t)) & \sin(\omega\,(T-t))\,\cos(\omega\,(T-t)) \\ \sin(\omega\,(T-t))\,\cos(\omega\,(T-t)) & \cos^2(\omega\,(T-t)) \end{pmatrix},$$
$$\Theta_t \doteq -m\,v \begin{pmatrix} \sin(\omega\,(T-t)) \\ \cos(\omega\,(T-t)) \end{pmatrix},$$

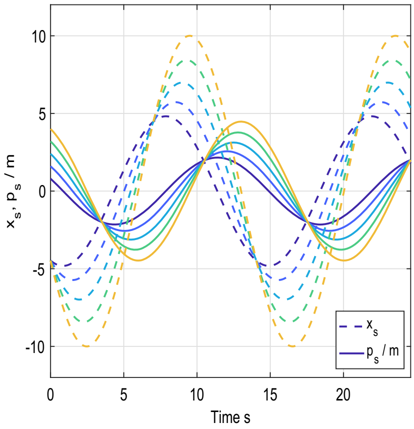
By inspection, the matrix $(\,\Omega_t \big| \Theta_t\,) \in \mathbb{R}^{2\times 3}$ is rank one, i.e. the two equations in (51) are linearly dependent. Some minor manipulations yield four possible cases for selecting p, given x, t, T, and v, namely,

$$\begin{cases} p = -x\,\sqrt{\kappa\, m}\,\tan\,(\omega\,(T-t)) - m\,v\,\sec(\omega\,(T-t)), \\ \qquad\qquad\qquad\qquad \omega\,(T-t) \notin \left\{n\,\pi,\ (n+\tfrac{1}{2})\,\pi\,:\, n\in\mathbb{Z}\right\}, \\ p = (-1)^{n+1}\,m\,v, \ \omega\,(T-t) \in \{n\,\pi\,:\,n\in\mathbb{Z}\},\ x\in\mathbb{R}, \\ p \ \text{ arbitrary}, \qquad \omega\,(T-t)\in\{(n+\tfrac{1}{2})\,\pi\,:\,n\in\mathbb{Z}\},\ x = (-1)^{n+1}\,(\tfrac{v}{\omega}), \\ p \ \text{ does not exist}, \ \ \omega\,(T-t)\in\{(n+\tfrac{1}{2})\,\pi\,:\,n\in\mathbb{Z}\},\ x \neq (-1)^{n+1}\,(\tfrac{v}{\omega}). \end{cases}$$

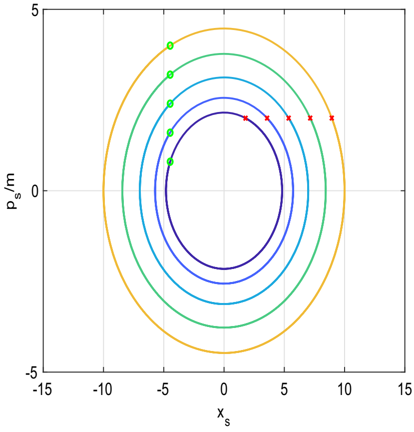
Note in the second case that p must correspond to the desired terminal generalized momentum, with sign determined by whether $T-t$ is a period or half-period of the mass-spring oscillation. In the third and fourth cases, $T-t$ corresponds to a quarter or three quarter period of the mass-spring oscillation, and p is either arbitrary, or does not exist, depending on the specific choice of x. Where p exists, note that the initial velocity that achieves the desired terminal velocity $\dot{x}_T = v$ is given by $\dot{x}_t = -p/m$. An example of the third case, where p and hence $\dot{x}_t$ is arbitrary, is illustrated in Fig. 1, for $v \doteq -2$ and $x = (-1)^4\,(\frac{v}{\omega}) \approx -4.47$. Note in particular that the desired terminal $\dot{x}_T = v$ is achieved for every trajectory, irrespective of its initial velocity, as expected.

5 Conclusions

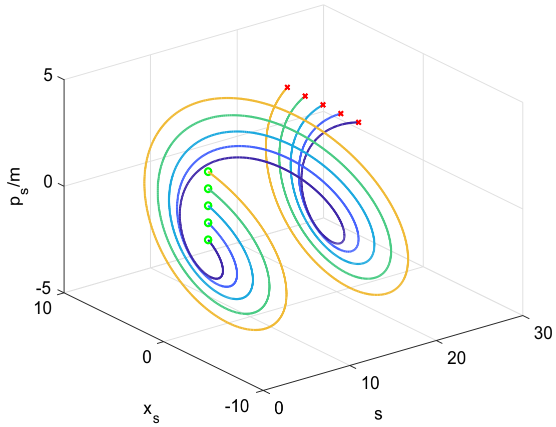
Connections between stationary action and stationary control are explored with a view to characterizing trajectories of energy conserving systems with temporal boundary conditions, evolving on arbitrary time horizons. An auxiliary stationary control problem is defined with respect to the characteristic system associated with the energy conserving dynamics of interest, and a verification theorem developed. This verification theorem provides a characterization of generalized velocity trajectories that render the associated action functional stationary for arbitrary time horizons, in an analogous way to existing verification results available for optimal control problems encapsulating least action on short time horizons. Application of this verification theorem is illustrated via a simple mass-spring example.



(a) Time evolution.



(b) Phase portrait.



(c) Time evolution of phase portrait.

Fig. 1 State and costate trajectories. **a** Time evolution. **b** Phase portrait. **c** Time evolution of phase portrait

Funding This research was supported by the US Air Force Office of Scientific Research.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

A Proofs of Lemmas 2, 3, 4, and 5

Proof (Lemma 2) The proof employs a standard fixed point argument, exploiting global Lipschitz continuity of f of (22), see for example [15, Theorem 5.1, p. 127]. Note that global Lipschitz continuity of $\nabla V(x)$ in (22) follows directly from the second bound assumed in (4). □

Proof (Lemma 3) Fix $T \in \mathbb{R}_{\geq 0}$, $t \in [0, T]$, and $Y, h \in \mathscr{X}^2$. Applying Lemma 2, there exist unique classical solutions $\overline{X}(Y)$ and $\overline{X}(Y+h)$ to (24) satisfying respectively $\overline{X}(Y)_t = Y$ and $\overline{X}(Y+h)_t = Y + h$. In integral form,

$$\begin{aligned} \overline{X}(Y)_s &= Y + \int_t^s f(\overline{X}(Y)_\sigma)\, d\sigma, \\ \overline{X}(Y+h)_s &= Y + h + \int_t^s f(\overline{X}(Y+h)_\sigma)\, d\sigma, \end{aligned} \tag{52}$$

so that

$$\overline{X}(Y+h)_s - \overline{X}(Y)_s = h + \int_t^s f(\overline{X}(Y+h)_\sigma) - f(\overline{X}(Y)_\sigma)\, d\sigma$$

for all $s \in [t, T]$. Consequently, as f is globally Lipschitz by inspection of (22),

$$\begin{aligned} \|\overline{X}(Y+h)_s - \overline{X}(Y)_s\| &\leq \|h\| + \int_t^s \|f(\overline{X}(Y+h)_\sigma) - f(\overline{X}(Y)_\sigma)\|\, d\sigma \\ &\leq \|h\| + \alpha \int_t^s \|\overline{X}(Y+h)_\sigma - \overline{X}(Y)_\sigma\|\, d\sigma \end{aligned}$$

in which $\alpha \in \mathbb{R}_{\geq 0}$ is the associated Lipschitz constant. Applying Gronwall's inequality, and recalling the definition of $\|\cdot\|_\infty$, yields

$$\|\overline{X}(Y+h) - \overline{X}(Y)\|_\infty \leq \|h\| \exp(\alpha\,(T-t)),$$

so that (26) holds. As $Y, h \in \mathscr{X}^2$ are arbitrary, the asserted continuity follows. □

Proof (Lemma 4) Fix $T \in \mathbb{R}_{\geq 0}$, $t \in [0, T]$, and $\overline{X} \in C(\mathscr{X}^2; C([t, T]; \mathscr{X}^2))$ as per Lemma 3. Fix $Y \in \mathscr{X}^2$ and $s \mapsto A(Y)_s$ as per (29), and note that (28) follows by [15, Theorem 5.2, p. 128]. Fix any $h \in \mathscr{X}^2$, $s \in [t, T]$, and note by inspection of (22) that $A(Y)_s = Df(\overline{X}(Y)_s)$. Hence, recalling (52),

$$\begin{aligned} \overline{X}(Y+h)_s - \overline{X}(Y)_s - U_{s,t}(Y)\, h &= \int_t^s f(\overline{X}(Y+h)_\sigma) - f(\overline{X}(Y)_\sigma) - A(Y)_\sigma\, U_{\sigma,t}(Y)\, h\, d\sigma \\ &= \int_t^s f(\overline{X}(Y)_\sigma + [\overline{X}(Y+h)_\sigma - \overline{X}(Y)_\sigma]) - f(\overline{X}(Y)_\sigma) - Df(\overline{X}(Y)_\sigma)\, U_{\sigma,t}(Y)\, h\, d\sigma \\ &= \int_t^s f(\overline{X}(Y)_\sigma + [\overline{X}(Y+h)_\sigma - \overline{X}(Y)_\sigma]) - f(\overline{X}(Y)_\sigma) - Df(\overline{X}(Y)_\sigma)\, [\overline{X}(Y+h)_\sigma - \overline{X}(Y)_\sigma] \\ &\quad + Df(\overline{X}(Y)_\sigma)\, [\overline{X}(Y+h)_\sigma - \overline{X}(Y)_\sigma - U_{\sigma,t}(Y)\, h]\, d\sigma. \end{aligned} \tag{53}$$

Define $\bar{I}_f : C([t, T]; \mathscr{X}^2) \to C([t, T]; \mathscr{X}^2)$ by

$$\bar{I}_f(X)_s \doteq \int_t^s f(X_\sigma)\, d\sigma \tag{54}$$

for all $X \in C([t,T]; \mathscr{X}^2)$. Note that $Y \mapsto f(Y)$ is twice Fréchet differentiable by (4), with $D^2 f(Y) \in \mathcal{L}(\mathscr{X}^2; \mathcal{L}(\mathscr{X}^2)) = \mathcal{L}(\mathscr{X}^2 \times \mathscr{X}^2; \mathscr{X}^2)$ for all $Y \in \mathscr{X}^2$. Again by (4), there exists an $M \in \mathbb{R}_{>0}$ such that

$$\sup_{Y \in \mathscr{X}^2} \|D^2 f(Y)\|_{\mathcal{L}(\mathscr{X}^2 \times \mathscr{X}^2; \mathscr{X}^2)} \le M < \infty.$$

Hence, by the mean value theorem, given $X, \delta \in C([t,T]; \mathscr{X}^2)$,

$$\begin{aligned}
&\left\| \bar{I}_f(X+\delta)_s - \bar{I}_f(X)_s - \int_t^s Df(X_\sigma)\,\delta_\sigma\, d\sigma \right\| \le \int_t^s \|f(X_\sigma + \delta_\sigma) - f(X_\sigma) - Df(X_\sigma)\,\delta_\sigma\|\, d\sigma \\
&= \int_t^s \left\| \left(\int_0^1 \int_0^1 D^2 f(X_\sigma + \hat{\eta}\,\eta\,\delta_\sigma)\, d\hat{\eta}\,\eta\, d\eta \right) (\delta_\sigma, \delta_\sigma) \right\| d\sigma \\
&\le \int_t^s \int_0^1 \int_0^1 \|D^2 f(X_\sigma + \eta\,\delta_\sigma)\|_{\mathcal{L}(\mathscr{X}^2 \times \mathscr{X}^2; \mathscr{X}^2)} d\hat{\eta}\,\eta\, d\eta\, \|\delta_\sigma\|^2\, d\sigma \\
&\le \tfrac{M}{2} \int_t^s \|\delta_\sigma\|^2\, d\sigma \le \tfrac{M}{2}\,(s-t)\,\|\delta\|_\infty^2.
\end{aligned}$$

That is,

$$\left\| \bar{I}_f(X+\delta) - \bar{I}_f(X) - \int_t^{(\cdot)} Df(X_\sigma)\,\delta_\sigma\, d\sigma \right\|_\infty \le \tfrac{M}{2}\,(T-t)\,\|\delta\|_\infty^2,$$

so that $\bar{I}_f$ is Fréchet differentiable with derivative

$$[D\bar{I}_f(X)\,\delta]_s = \int_t^s Df(X_\sigma)\,\delta_\sigma\, d\sigma \tag{55}$$

for all $X, \delta \in C([t,T]; \mathscr{X}^2)$, $s \in [t,T]$. So, recalling (53), and (1),

$$\begin{aligned}
\overline{X}(Y+h)_s - \overline{X}(Y)_s - U_{s,t}(Y)\,h &= [d[\bar{I}_f]_{\overline{X}(Y)}(\overline{X}(Y+h) - \overline{X}(Y))]_s\, \|\overline{X}(Y+h) - \overline{X}(Y)\|_\infty \\
&\quad + \int_t^s Df(\overline{X}(Y)_\sigma)\,[\overline{X}(Y+h)_\sigma - \overline{X}(Y)_\sigma - U_{\sigma,t}(Y)\,h]\, d\sigma.
\end{aligned}$$

Noting that $L \doteq \sup_{\sigma \in [t,T]} \|Df(\overline{X}(Y)_\sigma)\|_{\mathcal{L}(\mathscr{X}^2)} < \infty$, taking the norm of both sides yields

$$\begin{aligned}
&\|\overline{X}(Y+h)_s - \overline{X}(Y)_s - U_{s,t}(Y)\,h\| \\
&\quad \le \|d[\bar{I}_f]_{\overline{X}(Y)}(\overline{X}(Y+h) - \overline{X}(Y))\|_\infty\, \|\overline{X}(Y+h) - \overline{X}(Y)\|_\infty \\
&\qquad + \int_t^s L\, \|\overline{X}(Y+h)_\sigma - \overline{X}(Y)_\sigma - U_{\sigma,t}(Y)\,h\|\, d\sigma.
\end{aligned}$$

Hence, by Gronwall's inequality,

$$\|\overline{X}(Y+h)_s - \overline{X}(Y)_s - U_{s,t}(Y)\,h\|$$

$$\leq \|d[\bar{I}_f]_{\overline{X}(Y)}(\overline{X}(Y+h) - \overline{X}(Y))\|_\infty \, \|\overline{X}(Y+h) - \overline{X}(Y)\|_\infty \, \exp(L\,(T-t)),$$

or, with $\theta_Y(h) \doteq \|d[\bar{I}_f]_{\overline{X}(Y)}(\overline{X}(Y+h) - \overline{X}(Y))\|_\infty$,

$$\begin{aligned}
&\|\overline{X}(Y+h) - \overline{X}(Y) - U_{\cdot,t}(Y)\,h\|_\infty \\
&\quad \leq \theta_Y(h)\,\|\overline{X}(Y+h) - \overline{X}(Y)\|_\infty \, \exp(L\,(T-t)) \\
&\quad \leq \theta_Y(h)\,\|\overline{X}(Y+h) - \overline{X}(Y) - U_{\cdot,t}(Y)\,h\|_\infty \, \exp(L\,(T-t)) \\
&\qquad + \theta_Y(h) \sup_{s\in[t,T]} \|U_{s,t}(Y)\|_{\mathcal{L}(\mathscr{X}^2)}\,\|h\| \, \exp(L\,(T-t)).
\end{aligned}$$

As θ_Y is continuous at 0, there exists an $r > 0$ sufficiently small such that $\|h\| < r$ implies that $\theta_Y(h)\exp(L\,(T-t)) < \frac{1}{2}$. Hence, with $\|h\| < r$,

$$\begin{aligned}
\|\overline{X}(Y+h) - \overline{X}(Y) - U_{\cdot,t}(Y)\,h\|_\infty &< 2\,\theta_Y(h) \sup_{s\in[t,T]} \|U_{s,t}(Y)\|_{\mathcal{L}(\mathscr{X}^2)}\,\|h\| \, \exp(L\,(T-t)) \\
&= Q\,\theta_Y(h)\,\|h\|,
\end{aligned}$$

in which $Q \doteq 2\,\sup_{s\in[t,T]} \|U_{s,t}(Y)\|_{\mathcal{L}(\mathscr{X}^2)} \, \exp(L\,(T-t))$. Consequently, taking a limit,

$$\lim_{\|h\|\to 0} \frac{\|\overline{X}(Y+h) - \overline{X}(Y) - U_{\cdot,t}(Y)\,h\|_\infty}{\|h\|} \leq \lim_{\|h\|\to 0} Q\,\theta_Y(h) = 0.$$

That is, $Y \mapsto \overline{X}(Y)$ is Fréchet differentiable, with the indicated derivative. □

Proof (Lemma 5) Fix $T \in \mathbb{R}_{>0}$, $t \in [0,T]$ as per the lemma statement. It is first demonstrated that $Y \mapsto U_{s,r}(Y)$ is continuous, uniformly in $r, s \in [t,T]$, as this motivates the subsequent proof of continuous differentiability. Fix $r, s \in [t,T]$, $h, \hat{h} \in \mathscr{X}^2$. As $U_{s,r}(Y) \in \mathcal{L}(\mathscr{X}^2)$ is an element of the two-parameter family of evolution operators generated by $A(Y)_s \in \mathcal{L}(\mathscr{X}^2)$, see (29),

$$\begin{aligned}
U_{s,r}(Y)\,h &= h + \int_r^s A(Y)_\sigma\, U_{\sigma,r}(Y)\,h\,d\sigma, \\
U_{s,r}(Y+\hat{h})\,h &= h + \int_r^s A(Y+\hat{h})_\sigma\, U_{\sigma,r}(Y+\hat{h})\,h\,d\sigma,
\end{aligned}$$

so that

$$\begin{aligned}
[U_{s,r}(Y+\hat{h}) - U_{s,r}(Y)]\,h &= \int_r^s [A(Y+\hat{h})_\sigma\, U_{\sigma,r}(Y+\hat{h}) - A(Y)_\sigma\, U_{\sigma,r}(Y)]\,h\,d\sigma \\
&= \int_r^s [A(Y+\hat{h})_\sigma - A(Y)_\sigma]\,[U_{\sigma,r}(Y+\hat{h}) - U_{\sigma,r}(Y)]\,h\,d\sigma \\
&\quad + \int_r^s A(Y)_\sigma\,[U_{\sigma,r}(Y+\hat{h}) - U_{\sigma,r}(Y)]\,h\,d\sigma + \int_r^s [A(Y+\hat{h})_\sigma - A(Y)_\sigma]\,U_{\sigma,r}(Y)\,h\,d\sigma.
\end{aligned} \tag{56}$$

Hence, by the triangle inequality,

$$\begin{aligned}
&\|[U_{s,r}(Y+\hat{h}) - U_{s,r}(Y)]\,h\| \\
&\quad \le \int_r^s \|A(Y+\hat{h})_\sigma - A(Y)_\sigma\|_{\mathcal{L}(\mathscr{X}^2)}\, \|[U_{\sigma,r}(Y+\hat{h}) - U_{\sigma,r}(Y)]\,h\|\, d\sigma \\
&\qquad + \int_r^s \|A(Y)_\sigma\|_{\mathcal{L}(\mathscr{X}^2)}\, \|[U_{\sigma,r}(Y+\hat{h}) - U_{\sigma,r}(Y)]\,h\|\, d\sigma \\
&\qquad + \int_r^s \|A(Y+\hat{h})_\sigma - A(Y)_\sigma\|_{\mathcal{L}(\mathscr{X}^2)}\, \|U_{\sigma,r}(Y)\,h\|\, d\sigma. \qquad (57)
\end{aligned}$$

Recalling (4), and in particular the uniform bound on $x \mapsto D\nabla^2 V(x)$, given $x, \bar{x} \in \mathscr{X}^2$, the mean value theorem implies that $\nabla^2 V(x) - \nabla^2 V(\bar{x}) = (\int_0^1 D\nabla^2 V(\bar{x} + \eta\,(x - \bar{x}))\, d\eta)(x - \bar{x})$, so that $\|\nabla^2 V(x) - \nabla^2 V(\bar{x})\|_{\mathcal{L}(\mathscr{X})} \le \frac{K}{2}\, \|x - \bar{x}\|$. Hence, by (29), there exists an $\alpha_1 \in \mathbb{R}_{\ge 0}$ such that $\Lambda : \mathscr{X}^2 \to \mathcal{L}(\mathscr{X}^2)$ satisfies $\|\Lambda(Z) - \Lambda(\bar{Z})\|_{\mathcal{L}(\mathscr{X}^2)} \le \alpha_1 \|Z - \bar{Z}\|$ for all $Z, \bar{Z} \in \mathscr{X}^2$. So, applying Lemma 3, there exists an $\alpha \in \mathbb{R}_{\ge 0}$, $L_0 \doteq \sup_{\sigma \in [t,T]} \|A(0)_\sigma\|_{\mathcal{L}(\mathscr{X}^2)} < \infty$, $L_1 \doteq \alpha_1\, \exp(\alpha\,(T - t)) < \infty$, such that

$$\begin{aligned}
\sup_{\sigma \in [t,T]} \|A(Y+\hat{h})_\sigma - A(Y)_\sigma\|_{\mathcal{L}(\mathscr{X}^2)} &\le \alpha_1 \sup_{\sigma \in [t,T]} \|\overline{X}(Y+\hat{h})_\sigma - \overline{X}(Y)_\sigma\| \le L_1\, \|\hat{h}\|, \\
\sup_{\sigma \in [t,T]} \|A(Y)_\sigma\|_{\mathcal{L}(\mathscr{X}^2)} &\le L_0 + L_1\, \|\hat{h}\|, \qquad (58)
\end{aligned}$$

in which the second inequality follows from the first, via the triangle inequality, by selecting $\hat{h} = -Y$. Note further that as $\sigma \mapsto A(Y)_\sigma$ is continuous, $L_2 \doteq \sup_{\sigma \in [t,T]} \|U_{\sigma,t}(Y)\|_{\mathcal{L}(\mathscr{X}^2)} < \infty$, see [15, Theorem 5.2, p.128]. Hence, substituting these inequalities in (57) yields

$$\begin{aligned}
\|[U_{s,r}(Y+\hat{h}) - U_{s,r}(Y)]\,h\| &\le (L_0 + 2\,L_1\, \|\hat{h}\|) \int_r^s \|[U_{\sigma,r}(Y+\hat{h}) - U_{\sigma,r}(Y)]\,h\|\, d\sigma \\
&\quad + (T - t)\, L_1\, L_2\, \|\hat{h}\|\, \|h\|.
\end{aligned}$$

Gronwall's inequality subsequently implies that

$$\sup_{r,s \in [t,T]} \|U_{s,r}(Y+\hat{h}) - U_{s,r}(Y)\|_{\mathcal{L}(\mathscr{X}^2)} \le (T - t)\, L_1\, L_2\, \|\hat{h}\|\, \exp((L_0 + 2\,L_1\, \|\hat{h}\|)(T - t)). \qquad (59)$$

Continuity of $Y \mapsto U_{s,r}(Y)$, uniformly in $r, s \in [t, T]$, thus follows.

Next, $Y \mapsto U_{s,r}(Y)$ is shown to be Fréchet differentiable, uniformly in $r, s \in [t, T]$. Appealing to the contraction theorem and Picard's principle, for any $t \le r < s \le T$ and $Y \in \mathscr{X}$, consider the two-parameter family of operators $V_{s,r}(Y) \in \mathcal{L}(\mathscr{X}^2; \mathcal{L}(\mathscr{X}^2))$ solving

$$V_{s,r}(Y)\,\hat{h}\,h = \int_r^s A(Y)_\sigma\, V_{\sigma,r}(Y)\,\hat{h}\,h\, d\sigma + \int_r^s D_Y A(Y)_\sigma\, \hat{h}\, U_{\sigma,r}(Y)\,h\, d\sigma \qquad (60)$$

for all $h, \hat{h} \in \mathscr{X}^2$, $r, s \in [t, T]$, in which $D_Y A(Y)_\sigma = D\Lambda(\overline{X}(Y)_\sigma)\, U_{\sigma,t}(Y) \in \mathcal{L}(\mathscr{X}^2; \mathcal{L}(\mathscr{X}^2))$ by the chain rule and Lemma 4. Note in particular by (4), (29), and Lemma 3 that

$$L_3 \doteq \sup_{\sigma \in [t,T]} \|D\Lambda(\overline{X}(Y)_\sigma)\|_{\mathcal{L}(\mathscr{X}^2; \mathcal{L}(\mathscr{X}^2))} < \infty.$$

Applying the triangle inequality to (60), and recalling the definitions of L_0, L_1, L_2, yields

$$\begin{aligned}\|V_{s,r}(Y)\,\hat{h}\,h\| &\le \int_r^s (L_0 + L_1\,\|\hat{h}\|)\,\|V_{s,r}(Y)\,\hat{h}\,h\|\,d\sigma + \int_r^s L_3\,\|\hat{h}\|\,L_2\,\|h\|\,d\sigma \\ &\le (T-t)\,L_2\,L_3\,\|\hat{h}\|\,\|h\| + (L_0 + L_1\,\|\hat{h}\|) \int_r^s \|V_{s,t}(Y)\,\hat{h}\,h\|\,d\sigma,\end{aligned}$$

so that by Gronwall's inequality,

$$\|V_{s,r}(Y)\,\hat{h}\,h\| \le (T-t)\,L_2\,L_3\,\|\hat{h}\|\,\|h\|\,\exp\Big((L_0 + L_1\,\|\hat{h}\|)\,(T-t)\Big).$$

As $\hat{h}, h \in \mathscr{X}^2$ are arbitrary, it follows immediately that $V_{s,r}(Y) \in \mathcal{L}(\mathscr{X}^2; \mathcal{L}(\mathscr{X}^2))$ for all $r, s \in [t, T]$. Recalling (56), observe by adding and subtracting terms that

$$\begin{aligned}&[U_{s,r}(Y+\hat{h}) - U_{s,r}(Y) - V_{s,r}(Y)\,\hat{h}]\,h \\ &= \int_r^s [A(Y+\hat{h})_\sigma\, U_{\sigma,r}(Y+\hat{h}) - A(Y)_\sigma\, U_{\sigma,r}(Y)]\,h\,d\sigma - V_{s,r}(Y)\,\hat{h}\,h \\ &= \int_r^s A(Y)_\sigma\,[U_{\sigma,r}(Y+\hat{h}) - U_{\sigma,r}(Y) - V_{\sigma,r}(Y)\,\hat{h}]\,h\,d\sigma \\ &\quad + \int_r^s [A(Y+\hat{h})_\sigma - A(Y)_\sigma]\,[U_{\sigma,r}(Y+\hat{h}) - U_{\sigma,r}(Y)]\,h\,d\sigma \\ &\quad + \int_r^s [A(Y+\hat{h})_\sigma - A(Y)_\sigma - D_Y A(Y)_\sigma\,\hat{h}]\,U_{\sigma,r}(Y)\,h\,d\sigma \\ &\quad - \left[V_{s,r}(Y)\,\hat{h}\,h - \int_r^s A(Y)_\sigma\, V_{\sigma,r}(Y)\,\hat{h}\,h\,d\sigma - \int_r^s D_Y A(Y)_\sigma\,\hat{h}\,U_{\sigma,r}(Y)\,h\,d\sigma\right],\end{aligned} \tag{61}$$

and the last term in square brackets is zero by definition (60) of $V_{s,r}(Y)$. Define $\hat{A} : C([t,T]; \mathscr{X}^2) \to C([t,T]; \mathcal{L}(\mathscr{X}^2))$ by $\hat{A}(X)_\sigma \doteq A(X_\sigma)$ for all $X \in C([t,T]; \mathscr{X}^2)$, and note that the range of $\hat{A}$ follows by (4), (29). Fix $X, \delta \in C([t,T]; \mathscr{X}^2)$, and (for convenience) write $X_\sigma = ([X_1]_\sigma, [X_2]_\sigma) \in \mathscr{X}^2$, $\delta_\sigma = ([\delta_1]_\sigma, [\delta_2]_\sigma) \in \mathscr{X}^2$ for all $\sigma \in [t,T]$, with $X_1, X_2, \delta_1, \delta_2 \in C([t,T]; \mathscr{X})$. Combining (4), (29) with the mean value theorem, there exists $\hat{\alpha} \in \mathbb{R}_{\ge 0}$ such that

$$\begin{aligned}&\|\hat{A}(X+\delta) - \hat{A}(X) - DA(X)\,\delta\|_{C([t,T];\mathcal{L}(\mathscr{X}^2))} \\ &= \sup_{\sigma\in[t,T]} \|A(X_\sigma + \delta_\sigma) - A(X_\sigma) - DA(X_\sigma)\,\delta_\sigma\|_{\mathcal{L}(\mathscr{X}^2)}\end{aligned}$$

$$
\begin{aligned}
&= \sup_{\sigma\in[t,T]} \left\| \begin{pmatrix} 0 & 0 \\ \nabla^2 V([X_1]_\sigma + [\delta_1]_\sigma) - \nabla^2 V([X_1]_\sigma]) - D\nabla^2 V([X_1]_\sigma)\,[\delta_1]_\sigma & 0 \end{pmatrix} \right\|_{\mathcal{L}(\mathscr{X}^2)} \\
&= \hat{\alpha} \sup_{\sigma\in[t,T]} \|\nabla^2 V([X_1]_\sigma + [\delta_1]_\sigma) - \nabla^2 V([X_1]_\sigma) - D\nabla^2 V([X_1]_\sigma)\,[\delta_1]_\sigma\|_{\mathcal{L}(\mathscr{X})} \\
&= \hat{\alpha} \sup_{\sigma\in[t,T]} \left\| \left(\int_0^1 \int_0^1 D^2\nabla^2 V([X_1]_\sigma + \hat{\eta}\,\eta\,[\delta_1]_\sigma)\, d\hat{\eta}\,\eta\, d\eta \right) ([\delta_1]_\sigma, [\delta_1]_\sigma) \right\|_{\mathcal{L}(\mathscr{X})} \\
&\le \hat{\alpha} \sup_{\sigma\in[t,T]} \sup_{\hat{\eta},\eta\in[0,1]} \left\| D^2\nabla^2 V([X_1]_\sigma + \hat{\eta}\,\eta\,[\delta_1]_\sigma) \right\|_{\mathcal{L}(\mathscr{X}\times\mathscr{X};\mathcal{L}(\mathscr{X}))} \sup_{\sigma\in[t,T]} \|[\delta_1]_\sigma\|_{\mathscr{X}}^2 \\
&\le \hat{\alpha}\,(\tfrac{K}{2})\,\|\delta\|^2_{C([t,T];\mathscr{X}^2)}
\end{aligned} \tag{62}
$$

for all $X, \delta \in C([t,T]; \mathscr{X}^2)$. Dividing both sides by $\|\delta\|_{C([t,T];\mathscr{X}^2)}$ and taking the limit as $\|\delta\|_{C([t,T];\mathscr{X}^2)} \to 0$ subsequently yields that $\hat{A}$ is Fréchet differentiable with derivative $D\hat{A}(X) \in \mathcal{L}(C([t,T]; \mathscr{X}^2); C([t,T]; \mathcal{L}(\mathscr{X}^2)$. Hence, taking the norm of both sides of (61), applying the triangle inequality, (59), (62), and recalling the definitions of L_1, L_2, L_3,

$$
\begin{aligned}
&\|[U_{s,r}(Y+\hat{h}) - U_{s,r}(Y) - V_{s,r}(Y)\,\hat{h}]\,h\| \\
&\quad \le (L_0 + L_1\,\|\hat{h}\|) \int_r^s \|[U_{s,r}(Y+\hat{h}) - U_{s,r}(Y) - V_{s,r}(Y)\,\hat{h}]\,h\|\, d\sigma \\
&\qquad + (T-t)^2\, L_1^2\, L_2\, \|\hat{h}\|^2 \exp((L_0 + 2\,L_1\,\|\hat{h}\|)(T-t))\,\|h\| \\
&\qquad + (T-t)\, L_2\, \|\hat{A}\circ\overline{X}(Y+\hat{h}) - \hat{A}\circ\overline{X}(Y) - D\hat{A}(\overline{X}(Y))\, D\overline{X}(Y)\,\hat{h}\|_{C([t,T];\mathcal{L}(\mathscr{X}^2))}\,\|h\| \\
&\quad = (L_0 + L_1\,\|\hat{h}\|) \int_r^s \|[U_{s,r}(Y+\hat{h}) - U_{s,r}(Y) - V_{s,r}(Y)\,\hat{h}]\,h\|\, d\sigma \\
&\qquad + (T-t)^2\, L_1^2\, L_2\, \|\hat{h}\|^2 \exp((L_0 + 2\,L_1\,\|\hat{h}\|)(T-t))\,\|h\| \\
&\qquad + (T-t)\, L_2\, \|d(\hat{A}\circ\overline{X})_Y(\hat{h})\|_{C([t,T];\mathcal{L}(\mathscr{X}^2))}\,\|\hat{h}\|\,\|h\|,
\end{aligned}
$$

in which $d(\hat{A}\circ\overline{X})_Y(\cdot)$ is defined via (1). Hence, by Gronwall's inequality,

$$
\begin{aligned}
&\|[U_{s,r}(Y+\hat{h}) - U_{s,r}(Y) - V_{s,r}(Y)\,\hat{h}]\,h\| \\
&\quad \le (T-t)\, L_2 \left[(T-t)\, L_1^2\, \|\hat{h}\| \exp((L_0 + 2\,L_1\,\|\hat{h}\|)(T-t)) \right. \\
&\qquad \left. + \|d(\hat{A}\circ\overline{X})_Y(\hat{h})\|_{C([t,T];\mathcal{L}(\mathscr{X}^2))} \right] \|\hat{h}\|\,\|h\| \\
&\qquad \times \exp((L_0 + L_1\,\|\hat{h}\|)\,(T-t)).
\end{aligned}
$$

As $\hat{h}, h \in \mathscr{X}^2$ are arbitrary,

$$
\begin{aligned}
&\lim_{\|\hat{h}\|\to 0} \frac{\sup_{r,s\in[t,T]} \|U_{s,r}(Y+\hat{h}) - U_{s,r}(Y) - V_{s,r}(Y)\,\hat{h}\|_{\mathcal{L}(\mathscr{X}^2)}}{\|\hat{h}\|} \\
&\quad \le \lim_{\|\hat{h}\|\to 0} \|d(\hat{A}\circ\overline{X})_Y(\hat{h})\|_{C([t,T];\mathcal{L}(\mathscr{X}^2))} = 0.
\end{aligned} \tag{63}
$$

Hence, $Y \mapsto U_{s,r}(Y)$ is Fréchet differentiable, uniformly in $r, s \in [t, T]$, with derivative $V_{s,r}(Y)$.

It remains to be shown that $Y \mapsto U_{s,r}(Y)$ is twice Fréchet differentiable via (60). To this end, define $\upsilon_s \doteq V_{s,r}(Y)\,\hat{h} \in \mathcal{L}(\mathscr{X}^2)$ and $w_s \doteq D_Y A(Y)_s\,\hat{h}\,U_{s,r}(Y) \in \mathcal{L}(\mathscr{X}^2)$ for all $s \in [t, T]$, and note by (60) that

$$\upsilon_s = \int_r^s A(Y)_\sigma\,\upsilon_\sigma + w_\sigma\,d\sigma,$$

for all $s \in [t, T]$, recalling that $h \in \mathscr{X}$ in (60) is arbitrary. Equivalently, $s \mapsto \upsilon_s$ is the unique solution of the IVP $\dot{\upsilon}_s = A(Y)_s\,\upsilon_s + w_s$ for all $s \in (t, T)$, subject to $\upsilon_r = 0 \in \mathcal{L}(\mathscr{X}^2)$. By definition, $s \mapsto A(Y)_s$ generates the two-parameter family $U_{s,r}(Y)$, $r, s \in [t, T]$, so that $s \mapsto \upsilon_s = V_{s,r}(Y)\,\hat{h}$ satisfies

$$\begin{aligned} V_{s,r}(Y)\,\hat{h} = \upsilon_s &= U_{s,r}(Y)\,\upsilon_r + \int_r^s U_{s,\sigma}(Y)\,w_\sigma\,d\sigma = \int_r^s U_{s,\sigma}(Y)\,w_\sigma\,d\sigma \\ &= \int_r^s U_{s,\sigma}(Y)\,D_Y A(Y)_\sigma\,\hat{h}\,U_{\sigma,r}(Y)\,d\sigma \end{aligned} \tag{64}$$

for all $r, s \in [t, T]$, in which the third equality follows as $\upsilon_r = V_{r,r}(Y)\,\hat{h} = 0 \in \mathcal{L}(\mathscr{X}^2)$, either by (60) or directly as $V_{r,r}(Y) \doteq D_Y U_{r,r}(Y) = D_Y I = 0$. Hence, by inspection of (64), the map $Y \mapsto V_{s,r}(Y) \doteq D_Y U_{s,r}(Y)$ is also Fréchet differentiable, with

$$\begin{aligned} D_Y V_{s,r}(Y)\,h\,\hat{h} = \int_r^s & V_{s,\sigma}(Y)\,h\,D_Y A(Y)_\sigma\,\hat{h}\,U_{\sigma,r}(Y) \\ & + U_{s,\sigma}(Y)\,D_Y^2 A(Y)_\sigma\,h\,\hat{h}\,U_{\sigma,r}(Y) + U_{s,\sigma}(Y)\,D_Y A(Y)_\sigma\,\hat{h}\,V_{\sigma,r}(Y)\,h\,d\sigma, \end{aligned}$$

in which $D_Y^2 A(Y)_\sigma \in \mathcal{L}(\mathscr{X}^2 \times \mathscr{X}^2; \mathcal{L}(\mathscr{X}^2))$, $\sigma \in [t, T]$, exists by (29) and (4). □

B An Auxiliary Statement of Proposition 2

Proposition 3 *Given $T \in \mathbb{R}_{>0}$, $t \in [0, T)$, $x, p \in \mathscr{X}$, and $(\bar{x}_s, \bar{p}_s) \doteq \overline{X}(Y_p(x))_s$ for all $s \in [t, T]$, the maps $s \mapsto \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$ and $s \mapsto \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$ are continuously differentiable, with derivatives given by*

$$\tfrac{d}{ds}\left[\nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)\right] = -\mathcal{M}^{-1}\left(\bar{p}_s - \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)\right), \tag{65}$$

$$\tfrac{d}{ds}\left[\nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)\right] = \nabla V(\bar{x}_s) - \nabla^2 V(\bar{x}_s)\,\nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s), \tag{66}$$

for all $s \in (t, T)$*. Moreover,* $s \mapsto \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)$ *is twice continuously differentiable, and satisfies*

$$0 = \tfrac{d^2}{ds^2}\left[\nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)\right] + \mathcal{M}^{-1}\,\nabla^2 V(\bar{x}_s)\,\nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s), \tag{67}$$

for all $s \in (t, T)$*.*

Proof Fix $T \in \mathbb{R}_{>0}$, $x, p \in \mathscr{X}$, and let $(\bar{x}_s, \bar{p}_s) \in \mathscr{X}^2$, $s \in [t, T]$, be as per the lemma statement. Fix $h \in \mathscr{X}$. Applying Proposition 1, $(s, x, p) \mapsto \bar{J}_T(s, x, p)$ is twice continuously differentiable, and the order of differentiation may be swapped. In particular,

$$\begin{aligned}
&\tfrac{d}{ds}\,[D_p\,\bar{J}_T(s, \bar{x}_s, \bar{p}_s)\,h] \\
&= \tfrac{\partial}{\partial s}\,[D_p\,\bar{J}_T(s, \bar{x}_s, \bar{p}_s)\,h] + D_x\,[D_p\,\bar{J}_T(s, \bar{x}_s, \bar{p}_s)\,h]\,\dot{\bar{x}}_s + D_p\,[D_p\,\bar{J}_T(s, \bar{x}_s, \bar{p}_s)\,h]\,\dot{\bar{p}}_s \\
&= \left(D_p\,\tfrac{\partial}{\partial s}\,\bar{J}_T(s, \bar{x}_s, \bar{p}_s) + D_x\,D_p\,\bar{J}_T(s, \bar{x}_s, \bar{p}_s)\,\dot{\bar{x}}_s + D_p\,D_p\,\bar{J}_T(s, \bar{x}_s, \bar{p}_s)\,\dot{\bar{p}}_s\right) h.
\end{aligned} \tag{68}$$

Meanwhile, $\bar{J}_T$ satisfies (41) by Theorem 5, i.e.

$$0 = -\tfrac{\partial}{\partial s}\,\bar{J}_T(s, x, p) - \tfrac{1}{2}\,\langle p,\, \mathcal{M}^{-1}\,p\rangle + V(x) + D_x \bar{J}_T(s, x, p)\,\mathcal{M}^{-1}\,p - D_p \bar{J}_T(s, x, p)\,\nabla V(x), \tag{69}$$

for all $s \in (t, T)$, $x, p \in \mathscr{X}$. Differentiating (69) with respect to p,

$$\begin{aligned}
0 &= -D_p(\tfrac{\partial}{\partial s}\,\bar{J}_T(s, x, p))\,h - \langle \mathcal{M}^{-1}\,p,\, h\rangle + D_p\,(D_x\,\bar{J}_T(s, x, p))\,h\,\mathcal{M}^{-1}\,p \\
&\quad + D_x\,\bar{J}_T(s, x, p)\,\mathcal{M}^{-1}\,h - D_p\,(D_p\,\bar{J}_T(s, x, p)\,\nabla V(x))\,h \\
&= -\langle \mathcal{M}^{-1}\,(p - \nabla_x \bar{J}_T(s, x, p)),\, h\rangle \\
&\quad - \left(D_p\,\tfrac{\partial}{\partial s}\,\bar{J}_T(s, x, p) - D_x\,D_p\,\bar{J}_T(s, x, p)\,\mathcal{M}^{-1}\,p + D_p\,D_p\,\bar{J}_T(s, x, p)\,\nabla V(x)\right) h.
\end{aligned}$$

Evaluating along the trajectory $s \mapsto (\bar{x}_s, \bar{p}_s)$ corresponding to $\overline{X}(Y_p(x))$, i.e. as per (20), yields

$$\begin{aligned}
&\left(D_p\,\tfrac{\partial}{\partial s}\,\bar{J}_T(s, \bar{x}_s, \bar{p}_s) + D_x\,D_p\,\bar{J}_T(s, \bar{x}_s, \bar{p}_s)\,\dot{\bar{x}}_s + D_p\,D_p\,\bar{J}_T(s, \bar{x}_s, \bar{p}_s)\,\dot{\bar{p}}_s\right) h \\
&\quad = -\langle \mathcal{M}^{-1}\,(\bar{p}_s - \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)),\, h\rangle.
\end{aligned}$$

Substitution in (68) subsequently yields

$$\langle \tfrac{d}{ds}[\nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)],\, h\rangle = \tfrac{d}{ds}\,[D_p\,\bar{J}_T(s, \bar{x}_s, \bar{p}_s)\,h] = -\langle \mathcal{M}^{-1}\,(\bar{p}_s - \nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)),\, h\rangle. \tag{70}$$

Recalling that $h \in \mathscr{X}$ is arbitrary immediately yields (65).

Similarly, for (66), observe that

$$\begin{aligned}
&\tfrac{d}{ds}[D_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)\, h] = \tfrac{\partial}{\partial s}[D_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)\, h] + D_x\, [D_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)\, h]\, \dot{\bar{x}}_s \\
&\quad + D_p\, [D_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)\, h]\, \dot{\bar{p}}_s \\
&= \left(D_x\, \tfrac{\partial}{\partial s} \bar{J}_T(s, \bar{x}_s, \bar{p}_s) + D_x\, D_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)\, \dot{\bar{x}}_s + D_p\, D_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)\, \dot{\bar{p}}_s\right) h.
\end{aligned} \tag{71}$$

Differentiating (69) with respect to x,

$$\begin{aligned}
0 &= -D_x(\tfrac{\partial}{\partial s} \bar{J}_T(s, x, p))\, h + D_x V(x)\, h + D_x\, (D_x \bar{J}_T(s, x, p))\, h\, \mathcal{M}^{-1}\, p \\
&\quad - D_x\, (D_p \bar{J}_T(s, x, p))\, h\, \nabla V(x) - D_p \bar{J}_T(s, x, p)\, D_x \nabla V(x)\, h \\
&= -\left(D_x\, \tfrac{\partial}{\partial s} \bar{J}_T(s, x, p) + D_x\, D_x \bar{J}_T(s, x, p)\, (-\mathcal{M}^{-1}\, p) + D_p\, D_x \bar{J}_T(s, x, p)\, \nabla V(x)\right) h \\
&\quad + \left(D_x V(x) - D_p \bar{J}_T(s, x, p)\, D_x \nabla V(x)\right) h
\end{aligned}$$

Evaluating along the trajectory $s \mapsto (\bar{x}_s, \bar{p}_s)$ corresponding to $\overline{X}(Y_p(x))$, i.e. as per (20), yields

$$\begin{aligned}
&\left(D_x \tfrac{\partial}{\partial s} \bar{J}_T(s, \bar{x}_s, \bar{p}_s) + D_x\, D_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)\, \dot{\bar{x}}_s) + D_p\, D_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)\, \dot{\bar{p}}_s\right) h \\
&\quad = \langle \nabla V(\bar{x}_s) - \nabla^2 V(\bar{x}_s)\, \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s),\ h\rangle.
\end{aligned}$$

Substitution in (71) subsequently yields

$$\langle \tfrac{d}{ds}[\nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)],\ h\rangle = \langle \nabla V(\bar{x}_s) - \nabla^2 V(\bar{x}_s)\, \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s),\ h\rangle.$$

Recalling that $h \in \mathscr{X}$ is arbitrary immediately yields (66).

The remaining assertion regarding twice differentiability is immediate by inspection of (65), (66), with

$$\begin{aligned}
\tfrac{d^2}{ds^2}[\nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)] &= -\mathcal{M}^{-1}\left(\dot{\bar{p}}_s - \tfrac{d}{ds}[\nabla_x \bar{J}_T(s, \bar{x}_s, \bar{p}_s)]\right) \\
&= -\mathcal{M}^{-1}\left(\nabla V(\bar{x}_s) - [\nabla V(\bar{x}_s) - \nabla^2 V(\bar{x}_s)\, \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s)]\right) \\
&= -\mathcal{M}^{-1}\, \nabla^2 V(\bar{x}_s)\, \nabla_p \bar{J}_T(s, \bar{x}_s, \bar{p}_s),
\end{aligned}$$

as required. □

References

1. Baggett, L.W.: Functional Analysis. Dekker, London (1991)
2. Bardi, M., Capuzzo-Dolcetta, I.: Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations. Systems & Control: Foundations & Application. Birkhauser, Basel (1997)
3. Basco, V., Frankowska, H.: Lipschitz continuity of the value function for the infinite horizon optimal control problem under state constraints. In: Trends in Control Theory and Partial Differential Equations, Springer INdAM Series, pp. 17–38 (2019)

4. Cannarsa, P., Sinestrari, C.: Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control. Progress in Nonlinear DIfferential Equations and Their Applications. Birkhaüser, Basel (2004)
5. Dower, P.M., McEneaney, W.M.: On existence and uniqueness of stationary action trajectories. In: Proceedings 22nd International Symposium on Mathematical Theory of Networks and Systems (Minneapolis MN, USA), pp. 624–631 (2016)
6. Dower, P.M., McEneaney, W.M.: Solving two-point boundary value problems for a wave equation via the principle of stationary action and optimal control. SIAM J. Control Optim. **55**(4), 2151–2205 (2017)
7. Dower, P.M., McEneaney, W.M.: Verifying fundamental solution groups for lossless wave equations via stationary action and optimal control. Appl. Math. Optim. (2020). https://doi.org/10.1007/s00245-020-09700-4
8. Dower, P.M., McEneaney, W.M., Yegorov, I.: Exploiting characteristics in stationary action problems. In: Proceedings of SIAM Conference on Control & Its Applications (Chengdu, China), pp. 75–82 (2019)
9. Gray, C.G., Taylor, E.F.: When action is not least. Am. J. Phys. **75**(5), 434–458 (2007)
10. McEneaney, W.M., Dower, P.M.: The principle of least action and fundamental solutions of mass-spring and n-body two-point boundary value problems. SIAM J. Control Optim. **53**(5), 2898–2933 (2015)
11. McEneaney, W.M., Dower, P.M.: Staticization, its dynamic program, and solution propagation. Automatica **81**, 56–67 (2017)
12. McEneaney, W.M., Dower, P.M.: Static duality and a stationary action application. J. Differ. Equ. **264**(2), 525–549 (2018)
13. McEneaney, W.M., Dower, P.M.: Staticization and associated Hamilton-Jacobi and Riccati equations. SIAM Conf. Control Theory Appl. **2015**, 376–383 (2015)
14. McEneaney, W.M., Dower, P.M.: Staticization-based representations for Schrödinger equations driven by Coulomb potentials. In: 3^{rd} IFAC Workshop on Thermodynamic Foundations for a Mathematical Systems Theory (Louvain-la-Neuve, Belgium). https://doi.org/10.1016/j.ifacol.2019.07.005 (2019)
15. Pazy, A.: Semigroups of Linear Operators and Applications to Partial Differential Equations. Applied Mathematical Sciences, vol. 44. Springer, New York (1983)
16. Subbotin, A.I.: Generalized Solutions of First-Order PDEs: The Dynamical Optimization Perspective. Systems & Control: Foundations & Application. Springer, New York (1995)
17. Yegorov, I., Dower, P.M.: Perspectives on characteristics based curse-of-dimensionality-free numerical approaches to solving Hamilton Jacobi equations. Appl. Math. Optim. (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.