SIAM J. APPL. MATH.
Vol. 76, No. 3, pp. 976–999

SUFFICIENT CONDITIONS FOR A PATH-CONNECTED SET OF LOCAL SOLUTIONS TO AN OPTIMAL CONTROL PROBLEM*

ANDY BORUM† AND TIMOTHY BRETL†

Abstract. Consider a fixed-endpoint, fixed time optimal control problem with state and input taking values in Euclidean space. We show that if the Hamiltonian function associated with this problem satisfies a scale invariance property, then the set of all local solutions to this problem—over all possible terminal state constraints—is path-connected. We also show that this result extends to problems with additional constraints such as a bound on total cost. Finally, we use this result to show that the set of all curves that can be realized by a planar elastica is path-connected, and we describe how this result can be applied to the problem of robotic manipulation for nonrigid objects.

Key words. optimal control, scale invariance, path-connectedness, conjugate points

AMS subject classifications. 34H05, 49K15, 49K30, 49N90, 93C15

DOI. 10.1137/15M1025608

1. Introduction.

Consider an optimal control problem with state $q(t) \in \mathbb{R}$ on the fixed time interval $t \in [0, T]$ with $q(0) = 0$. Suppose the upper curve in Figure 1 is a local solution of this optimal control problem for the terminal state constraint $q(T) = 1$, and the lower curve is a local solution for the terminal state constraint $q(T) = -1$. In this paper, we ask whether it is possible to continuously deform the upper curve into the lower curve in such a way that each trajectory along the deformation is a local solution of the optimal control problem for some choice of $q(T)$. (One possible continuous deformation is shown in Figure 1 by the dashed curves.) In other words, we ask whether the set of all local solutions of the optimal control problem over all possible choices of the terminal state constraint is path-connected.

Our main contribution in this paper is a sufficient condition for the set described above to be path-connected. This result relies on the assumption that the Hamiltonian function associated with the optimal control problem satisfies a scale invariance property. When this scale invariance property is satisfied, the necessary and sufficient conditions for optimality are scale-invariant with respect to time. This allows us to construct a path in the set of local solutions of the optimal control problem connecting two given solutions. This path corresponds to a continuous deformation of one given solution into the other, such as the deformation shown in Figure 1.

After establishing conditions for the set of all local solutions to be path-connected, we ask whether this set remains path-connected if we place additional constraints on the solutions of the optimal control problem. The constraints we consider include bounds on the state and the costate of the optimal control problem. We also pay particular attention to the constraint that the total cost must be less than a given bound.

We then consider an optimal control problem whose local solutions are stable

*Received by the editors June 16, 2015; accepted for publication (in revised form) February 26, 2016; published electronically May 26, 2016. This work was supported by NSF under grant IIS-1320519. This paper generalizes and extends results that appeared in a paper by the authors presented at the 2015 International Conference on Robotics and Automation (ICRA) [6].

http://www.siam.org/journals/siap/76-3/M102560.html

†Department of Aerospace Engineering, University of Illinois, Urbana, IL 61801 (borum2@illinois.edu, tbretl@illinois.edu). The work of the first author was supported by NSF-GRFP under grant DGE-1144245.

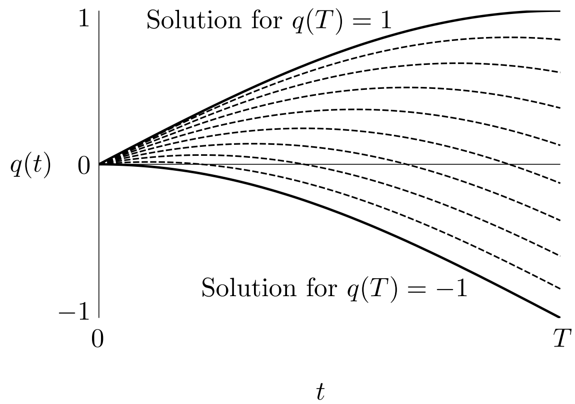

Fig. 1. *Two local solutions of an optimal control problem, one having the terminal state constraint $q(T) = 1$, and the other having the terminal state constraint $q(T) = -1$. In this paper, we ask whether either of these trajectories can be continuously deformed into the other trajectory, all the while remaining a local solution of the optimal control problem for some terminal state constraint. The dashed curves show a few trajectories along such a continuous deformation.*

equilibrium configurations of a planar elastica, i.e., a planar curve with given endpoints and slopes at the endpoints that minimizes curvature. Using the results in this paper, we show that the set of all such curves over all possible terminal state constraints is path-connected. Finally, we describe how this result can be applied to the problem of robotic manipulation for a canonical type of deformable object.

The motivating application and related work are discussed in section 2. In section 3, we review necessary and sufficient conditions for optimal control problems. In section 4, we describe optimal control problems with path-connected sets of local solutions, and we consider five example problems that have topologically distinct sets of local solutions. In section 5, we state our main result which is a sufficient condition for the set of all local solutions to be path-connected. In section 6, we consider the set of all local solutions that satisfy additional constraints. In section 7, we apply these results to an optimal control problem whose local solutions are stable equilibrium configurations of a planar elastica. Concluding remarks are given in section 8.

The results in this paper both generalize and extend our previous work, presented at a conference, in which we showed that the free configuration space of a Kirchhoff elastic rod is path-connected [6]. Here, we consider a much wider class of optimal control problems, and we state conditions for the set of all local solutions to be path-connected in terms of the Hamiltonian function rather than the differential equations that govern the associated Hamiltonian system.

2. The motivating application and related work. In this section, we discuss the motivation for this paper and the previous literature related to our results. The main application, which is robotic manipulation of nonrigid objects, is discussed in section 2.1. Previous work is discussed in section 2.2.

2.1. The motivating application. The main application that motivates this work is robotic manipulation of thin deformable objects such as an elastic wire, cable, or rod [7, 12, 14, 16, 17, 26]. When manipulating an elastic rod, the objective is to find a path of the robotic grippers holding the ends of the rod that causes the rod to move from an initial shape to a goal shape. Two types of models can be used to predict the rod's response to a change in gripper placement. First, a time-dependent dynamic model of the rod can be used. Such a model would capture vibrations and other dynamic phenomena that develop as the rod is manipulated. Alternatively, a time-independent quasi-static model can be used. This model assumes that the rod is being manipulated slowly enough so that dynamic effects are negligible, and at each point in time, the rod is in static equilibrium.

In this paper, we use a quasi-static model of the elastic rod. Under this quasi-static assumption, for given placements of the robotic grippers, the shape of the rod locally minimizes elastic potential energy. Therefore, the shape of the rod is a local solution of an optimal control problem whose total cost is the elastic potential energy stored in the rod. Quasi-statically deforming the rod using the robotic grippers is equivalent to planning a path of the rod through the set of all local solutions of this optimal control problem over all possible choices of the endpoint constraints. This set can be thought of as the feasible configuration space of the rod.

It is important to note that when using a quasi-static model of the elastic rod, an optimal control problem is solved to find stable equilibrium shapes of the rod. Optimal control is not used to find optimal paths of the robotic grippers holding the ends of the rod. This approach to manipulation has been applied to both planar [16] and spatially deformed rods [7]. As we will see in section 7, the results in this paper allow us to show that the feasible configuration space of the planar rod is path-connected. Furthermore, we are able to construct paths in this configuration space connecting any two shapes of the elastic rod. These paths in the feasible configuration space provide us with paths of the robotic grippers that cause the rod to move to the desired shape.

2.2. Related work. The sets that we consider in this paper (i.e., sets of all local solutions of an optimal control problem) have been previously analyzed for a few specific optimal control systems and calculus of variations problems. Bretl and McCarthy analyzed an optimal control problem whose local solutions are stable equilibrium configurations of a Kirchhoff elastic rod [7]. They showed that the set of all local solutions of this optimal control problem over all terminal state constraints is a subset of a smooth six-dimensional manifold. The same was done for a planar elastica by Matthews and Bretl [16]. The authors of this paper showed in [6] that this subset is path-connected for the Kirchhoff elastic rod. Neukirch and Henderson considered the related problem of characterizing the set of extremals of the Kirchhoff elastic rod [19] and used numerical continuation to compute the set [11]. Ivey and Singer considered the set of Kirchhoff elastic rods that have quasi-periodic centerlines and showed that this set is parameterized by a two-dimensional disc [13].

On a sub-Riemannian manifold, the set of all extremals (not necessarily local optima) originating from a point on the manifold corresponds to the preimage of the exponential map on the sub-Riemannian manifold. This set has been analyzed for a variety of problems [3, 8, 18, 22, 23, 24, 25]. Of these papers, Sachkov and Sachkova's characterization of the exponential map for a planar elastica is the one most related to our work [25]. After decomposing the preimage of the exponential map into subsets, Sachkov showed that some of these subsets are path-connected. However, it was not established that the set of all local optima is path-connected. Finally, we note that for linear quadratic optimal control problems, the set of all local optima is always path-connected [2]. We will discuss this in more detail in section 4 after exploring some examples.

When using continuation and homotopy methods to solve an optimal control problem, one is often interested in showing that the set of all solutions of the problem over all choices of the continuation parameter is path-connected [5, 9, 29]. When applying continuation methods to systems of algebraic equations, Smale gave conditions that guarantee the existence of a continuous path connecting the solutions as the continuation parameter changes [27]. However, in the case of optimal control systems, such conditions have been difficult to find [9]. When using a continuation method to find geodesics on a Riemannian manifold, Bonnard, Shcherbakova, and Sugny gave a sufficient condition for the existence of a continuous deformation connecting the geodesics as the continuation parameter changes [5]. This condition requires the time interval of the optimal control problem to be shorter than the injectivity radius, where the injectivity radius is the minimum time at which trajectories cease to be optimal [10].

To prove our main result in section 5, we rely on a symmetry property of differential equations known as scale invariance. This invariance property can be formulated as a one-dimensional Lie group symmetry [20, 28], and this property can be used to reduce the dimension of a system of differential equations by one [4]. We are particularly interested in applying this scale invariance property to the Hamiltonian system that results from applying Pontryagin's maximum principle to an optimal control problem [21]. In previous work, scale invariance was used by Ardentov to bound conjugate times for a sub-Riemannian problem on the Engel group [3]. The scale invariance property of a planar elastica that we consider in section 7 is mentioned briefly in [22] and [25], but the implications of this property were not analyzed.

3. Necessary and sufficient conditions for optimal control problems.

In this section, we state necessary and sufficient conditions for a trajectory to be a local solution of an optimal control problem. The Pontryagin maximum principle [21], which associates a Hamiltonian system to the optimal control problem, provides necessary conditions for optimality and is stated in section 3.1. The sufficient conditions for optimality in section 3.2 are based on the theory of conjugate points in optimal control problems [1]. These necessary and sufficient optimality conditions provide a set of ordinary differential equations that must be satisfied by local optima of the optimal control problem. In section 5, we will describe a scaling property that is sometimes satisfied by these differential equations, and we will show that this property can be completely characterized by the Hamiltonian function defined in section 3.1.

3.1. Necessary conditions for optimality.

Let $U = \mathbb{R}^m$ for some $m > 0$, and let $f\colon \mathbb{R}^n \times U \to \mathbb{R}^n$ and $g\colon \mathbb{R}^n \times U \to \mathbb{R}$ be smooth functions for some $n > 0$. Fix $T > 0$, and consider the optimal control problem

$$\begin{aligned} &\underset{q,u}{\text{minimize}} && C(q,u) = \int_0^T g(q(t),u(t))\ dt \\ &\text{subject to} && \dot{q}(t) = f(q(t),u(t)), \\ &&& q(0)=0,\ q(T)=b \end{aligned} \tag{3.1}$$

for some $b \in \mathbb{R}^n$, where $(q,u) : [0,T] \to \mathbb{R}^n \times U$. Necessary conditions for a trajectory (q,u) to be a local optimum of (3.1) are provided by Pontryagin's maximum principle [21]. To apply the maximum principle, we define the parameterized Hamiltonian function $\widehat{H}\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times U \to \mathbb{R}$ by

$$\widehat{H}(p,q,k,u) = p^T f(q,u) - kg(q,u), \tag{3.2}$$

where $p \in \mathbb{R}^n$ is called the costate.

THEOREM 3.1 (necessary conditions for optimality). *Suppose* $(q^*, u^*)\colon [0,T] \to \mathbb{R}^n \times U$ *is a local optimum of* (3.1). *Then, there exists* $k \geq 0$ *and a solution* $(p,q)\colon [0,T] \to \mathbb{R}^n \times \mathbb{R}^n$ *of the time-varying Hamiltonian system*

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}, \qquad \frac{dp}{dt} = -\frac{\partial H}{\partial q}, \tag{3.3}$$

where $H\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ *is given by* $H(p,q,t) = \widehat{H}(p,q,k,u^*(t))$*, that satisfies* $q(t) = q^*(t)$ *and*

$$H(p(t),q(t),t) = \max_{u \in U} \widehat{H}(p(t),q(t),k,u) \tag{3.4}$$

for all $t \in [0,T]$*. If* $k = 0$*, then* $p(t) \neq 0$ *for all* $t \in [0,T]$.

Proof. See Theorem 12.10 of Agrachev and Sachkov [1]. ∎

We call the solution (p,q) in Theorem 3.1 an abnormal extremal when $k = 0$ and a normal extremal otherwise. When $k \neq 0$, we may assume that $k = 1$. We call (q,u) abnormal if it is the projection of an abnormal extremal. We call (q,u) normal if it is the projection of a normal extremal and is not abnormal.

3.2. Sufficient conditions for optimality. Let $\pi\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ denote the projection map $\pi(p,q) = q$. The following assumptions (A1)–(A3) are needed in order to apply the sufficient conditions in Theorem 3.2. We will assume these conditions hold throughout the paper.

Suppose $(p,q)\colon [0,T] \to \mathbb{R}^n \times \mathbb{R}^n$ is a normal extremal of (3.1). Define the Hamiltonian function H by

$$H(p,q) = \max_{u \in U} \widehat{H}(p,q,1,u). \tag{3.5}$$

(A1) The Hamiltonian function H defined in (3.5) is smooth, and solutions of the Hamiltonian system (3.3) with H given by (3.5) exist for all time $t \in [0,T]$.

(A2) There is no other solution (p',q') of (3.3) satisfying $q(t) = q'(t)$ for all $t \in [0,T]$.

(A3) The maximum in (3.5) exists and $\partial^2 \widehat{H}/\partial u^2 < 0$.

Assumptions (A1)–(A3) place some restrictions on the type of optimal control problems we consider. In particular, the smoothness of the Hamiltonian function H typically requires that there be no constraints on the control input u, and this is why we have chosen the control set $U = \mathbb{R}^m$. A wide class of optimal control problems that satisfy these assumptions are problems with control-affine dynamic constraints and cost functions that are quadratic in the control input. We will see in section 4.2 that even these systems, which satisfy assumptions (A1)–(A3), can have sets of local optima with distinct topological properties.

Under condition (A1), solutions of (3.3) starting at $q(0) = 0$ are uniquely defined on the interval $[0,T]$ by the choice of initial costate $p(0) \in \mathbb{R}^n$, and the map from $p(0)$ to (p,q) is smooth and bijective. Denote this map by Γ, so that $(p,q) = \Gamma(p(0))$. Also let $\mathcal{A} \subset \mathbb{R}^n$ denote the set of all initial costates $p(0) \in \mathbb{R}^n$ that map to normal extremals of (3.1), i.e.,

$$\mathcal{A} = \{a \in \mathbb{R}^n \colon (p,q) = \Gamma(a) \text{ is a normal extremal of } (3.1)\}. \tag{3.6}$$

When (A2) holds, each $a \in \mathcal{A}$ is mapped smoothly to a unique $q = \pi(p,q) = \pi(\Gamma(a))$. When (A1)–(A3) are all true, each $a \in \mathcal{A}$ is mapped bijectively to a normal (q,u), and this map is smooth by assumption (A3) (due to the implicit function theorem). Denote this smooth map by $(q,u) = \Psi(a)$. Also define

$$\mathcal{C} = \Psi(\mathcal{A}) \subset C^\infty([0,T], \mathbb{R}^n \times U). \tag{3.7}$$

Note that for each $a \in \mathcal{A}$, $(q,u) = \Psi(a)$ satisfies Theorem 3.1 for some choice of terminal state constraint $b \in \mathbb{R}^n$. Therefore, $\mathcal{C}$ is the set of all normal (q,u) of (3.1) over all possible choices of the terminal state constraint $b \in \mathbb{R}^n$.

Theorem 3.2 provides a sufficient condition to test which normal (q,u) from Theorem 3.1 are local optima of (3.1).

THEOREM 3.2 (sufficient conditions for optimality). *Suppose* $(p,q)\colon [0,T] \to \mathbb{R}^n \times \mathbb{R}^n$ *is a normal extremal of* (3.1). *Define* $H \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$ *as in* (3.5), *and define* $u\colon [0,T] \to U$ *so that* $u(t)$ *is the unique maximizer of* (3.5) *at* $(p(t), q(t))$. *Let* $\varphi\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ *be the flow of* (3.3), *and define the endpoint map* $\phi_t\colon \mathbb{R}^n \to \mathbb{R}^n$ *by* $\phi_t(a) = \pi \circ \varphi(t,a,q(0))$. *Then* (q,u) *is a local optimum of* (3.1) *if and only if there exists no* $t \in (0,T]$ *for which* ϕ_t *is degenerate at* $p(0)$.

Proof. See Theorem 21.8 of Agrachev and Sachkov [1]. ∎

A point at which ϕ_t is degenerate is called a conjugate point, and the endpoint map ϕ_t is degenerate when its Jacobian matrix is singular. To compute this Jacobian matrix, extend the normal extremal (p,q) in Theorem 3.2 to $(p,q)\colon [0,T] \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ so that $(p(t,a), q(t,a))$ is the normal extremal with initial costate $p(0) = a$. Let $\mathbf{J}(t)$ be the Jacobian matrix of ϕ_t, and let $\mathbf{M}(t)$ be the Jacobian matrix of the costate $p(t,a)$ with respect to a, i.e.,

$$[\mathbf{J}(t)]_{ij} = \frac{\partial q_i(t,a)}{\partial a_j}, \qquad [\mathbf{M}(t)]_{ij} = \frac{\partial p_i(t,a)}{\partial a_j}.$$

Then the entries of $\mathbf{M}$ and $\mathbf{J}$ can be found by solving the differential equations

$$\frac{d}{dt}[\mathbf{M}(t)]_{ij} = \frac{\partial}{\partial a_j}\frac{\partial p_i(t,a)}{\partial t} = -\frac{\partial}{\partial a_j}\frac{\partial H}{\partial q_i} = -\sum_{k=1}^{n}\left(\frac{\partial^2 H}{\partial q_k \partial q_i}\frac{\partial q_k}{\partial a_j} + \frac{\partial^2 H}{\partial p_k \partial q_i}\frac{\partial p_k}{\partial a_j}\right), \tag{3.8}$$

$$\frac{d}{dt}[\mathbf{J}(t)]_{ij} = \frac{\partial}{\partial a_j}\frac{\partial q_i(t,a)}{\partial t} = \frac{\partial}{\partial a_j}\frac{\partial H}{\partial p_i} = \sum_{k=1}^{n}\left(\frac{\partial^2 H}{\partial q_k \partial p_i}\frac{\partial q_k}{\partial a_j} + \frac{\partial^2 H}{\partial p_k \partial p_i}\frac{\partial p_k}{\partial a_j}\right), \tag{3.9}$$

where H is the Hamiltonian function defined in (3.5). These equations can be written in matrix form as

$$\frac{d}{dt}\begin{bmatrix}\mathbf{M}\\ \mathbf{J}\end{bmatrix} = \begin{bmatrix}-H_{pq} & -H_{qq}\\ H_{pp} & H_{qp}\end{bmatrix}\begin{bmatrix}\mathbf{M}\\ \mathbf{J}\end{bmatrix}, \tag{3.10}$$

where subscripts denote derivatives with respect to the variables p and q. From the definitions of $\mathbf{M}$ and $\mathbf{J}$, it is clear that

$$\mathbf{M}(0) = \mathbf{I}_{n\times n}, \qquad \mathbf{J}(0) = \mathbf{0}_{n\times n}. \tag{3.11}$$

The system (3.10) along with the initial conditions (3.11) give a set of linear time-varying matrix differential equations that, when solved, provide the Jacobian of the endpoint map ϕ_t. A conjugate point occurs when $\det(\mathbf{J}(t)) = 0$ for some $t \in (0,T]$.

4. The set of local optima may or may not be path-connected.

In the previous section, we constructed a set $\mathcal{C}$ in (3.7) that contains all normal (q, u) of (3.1) for all possible choices of the terminal state constraint b. The set of all normal (q, u) that are local optima of (3.1) is some subset of $\mathcal{C}$. In this section, we consider the topological properties of this subset. First, in section 4.1 we describe the problem of showing that this subset is path-connected. Then, in section 4.2 we analyze this subset for five example optimal control problems. Through these examples, we will see that this subset either can be path-connected or can consist of two or more disjoint components.

4.1. A formal statement of the question about local optima.

We previously defined $\mathcal{A}$ in (3.6) to be the set of initial costate values $p(0) \in \mathbb{R}^n$ that map to normal extremals of the system (3.3). Theorem 3.2 gave a test to determine which normal extremals correspond to local minima of (3.1). Now define $\mathcal{A}_{\min} \subset \mathcal{A}$ by

$$\mathcal{A}_{\min} = \{a \in \mathcal{A} \colon (p, q) = \Gamma(a) \text{ satisfies the conditions in Theorem 3.2}\}. \tag{4.1}$$

$\mathcal{A}_{\min}$ is the set of $a \in \mathcal{A}$ corresponding to normal extremals $(p, q) = \Gamma(a)$ that do not have conjugate points on the interval $(0, T]$. The set of all local optima of (3.1) over all possible choices of the terminal state constraint b is then

$$\mathcal{C}_{\min} = \Psi(\mathcal{A}_{\min}) \subset \mathcal{C}. \tag{4.2}$$

The main question that we address in this paper is the following:

When is the set $\mathcal{C}_{\min}$ path-connected? That is, assume (q_0, u_0) and $(q_1, u_1) \colon [0, T] \to \mathbb{R}^n \times U$ are local optima of (3.1) *for $b = q_0(T)$ and $b = q_1(T)$, respectively. When are these two solutions connected by a path in the set of all local optima of* (3.1) *over all possible choices of b?*

To show that (q_0, u_0) and (q_1, u_1) are connected by a path in $\mathcal{C}_{\min}$, we must find a continuous function $\gamma \colon [0, T] \times [0, 1] \to \mathbb{R}^n \times U$ that satisfies the following properties:

1. $\gamma(t, 0) = (q_0(t), u_0(t))$ for all $t \in [0, T]$;
2. $\gamma(t, 1) = (q_1(t), u_1(t))$ for all $t \in [0, T]$;
3. for all $s \in [0, 1]$, $(q_s, u_s) = \gamma(\cdot, s)$ is a local optimum of (3.1) for $b = \gamma(T, s)$.

The following lemma simplifies this problem by allowing us to search for a continuous function $\alpha \colon [0, 1] \to \mathcal{A}_{\min}$.

LEMMA 4.1. *If $\mathcal{A}_{\min}$ is path-connected, then $\mathcal{C}_{\min}$ is path-connected.*

Proof. $\mathcal{C}_{\min} = \Psi(\mathcal{A}_{\min})$ and Ψ is continuous. Thus, if $\mathcal{A}_{\min}$ is path-connected, then $\mathcal{C}_{\min}$ is also path-connected. ∎

In the remainder of the paper, we will be concerned with showing that the set $\mathcal{A}_{\min}$ is path-connected.

4.2. Five examples with distinct sets of local optima.

We will now compute $\mathcal{A}_{\min}$ for five one-dimensional optimal control problems. Whereas the application that will be considered in section 7 has physical relevance, these example problems are included simply to show that the properties of the set of local optima can vary for different optimal control problems. In some of these examples, we will be able to explicitly compute or bound $\partial q(t, a)/\partial a$ instead of using the differential equations (3.10) to compute the Jacobian of the endpoint map ϕ_t in Theorem 3.2. The set $\mathcal{A}_{\min}$

will exhibit distinct properties in each of the five examples. In particular, we will see that the set $\mathcal{A}_{\min}$ can be empty, the entire set $\mathbb{R}^n$, a path-connected subset of $\mathbb{R}^n$, or a disconnected subset of $\mathbb{R}^n$ (made up of two or more connected components).

The five example optimal control problems that we consider all have the form

$$\begin{aligned} &\underset{q,u}{\text{minimize}} && \int_0^{2\pi} \frac{1}{2}u^2 + F_i(q)\ dt \\ &\text{subject to} && \dot{q}(t) = u(t), \\ & && q(0) = 0,\ q(2\pi) = b, \end{aligned} \tag{4.3}$$

where $(q,u)\colon [0,2\pi] \to \mathbb{R}\times\mathbb{R}$, and F_i is a smooth function. In each case, the optimal control, found from (3.4) is $u(t) = p(t)$, and it is easy to show that assumptions (A1)–(A3) hold. Also, there are no abnormal extremals, since setting $k = 0$ in (3.2) implies $p(t) = 0$. Thus $\mathcal{A} = \mathbb{R}$ in all five problems. Throughout this section, we will use a to denote the initial condition for $p(t)$ in the Hamiltonian system (3.3).

Example 1. First, consider the linear quadratic problem with

$$F_1(q) = 0. \tag{4.4}$$

Applying Theorem 3.1 gives the Hamiltonian system

$$\dot{q} = p, \qquad \dot{p} = 0.$$

The normal extremal beginning at $p(0) = a$ is

$$q(t) = at, \qquad p(t) = a.$$

It is clear that $\partial q(t,a)/\partial a$ is positive for $t \in (0, 2\pi]$, so every normal (q,u) is a local optimum by Theorem 3.2. Thus $\mathcal{A}_{\min} = \mathbb{R}$ and $\mathcal{C}_{\min}$ is path-connected by Lemma 4.1.

Example 2. Now consider the linear quadratic problem with

$$F_2(q) = -\frac{1}{2}q^2. \tag{4.5}$$

Applying Theorem 3.1 gives the Hamiltonian system

$$\dot{q} = p, \qquad \dot{p} = -q.$$

The normal extremal beginning at $p(0) = a$ is

$$q(t) = a\sin(t), \qquad p(t) = a\cos(t).$$

It is clear that $\partial q(t,a)/\partial a$ is zero at $t = \pi \in (0, 2\pi]$. Therefore, $\mathcal{A}_{\min} = \emptyset$ and there are no local optima. In Figure 2a, we plot the first time t_c at which $\partial q(t,a)/\partial a = 0$ (i.e., the first conjugate time) as a function of the choice of $a = p(0)$. In this case, $t_c(a)$ is a horizontal line at π. (Note that if we had chosen the final time to be less than π, then every extremal would be a local optimum and $\mathcal{A}_{\min} = \mathbb{R}$.)

In both of the previous examples, we found that the Jacobian $\partial q(t,a)/\partial a$ is independent of the choice of a, i.e., $q(t)$ depends linearly on a. This is true of all linear quadratic optimal control problems [2]. The existence of local optima in linear quadratic problems depends only on the dynamics $f(q,u)$, the cost function $g(q,u)$, and the final time T, and does not depend on the choice of a (and therefore does not depend on the choice of $b = q(T)$).

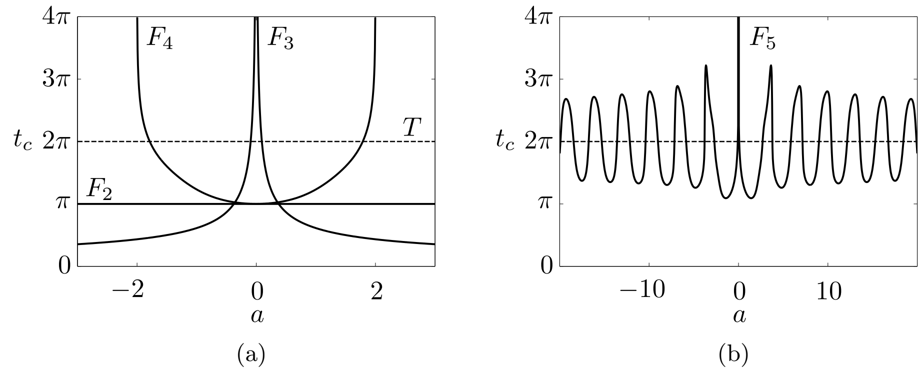

FIG. 2. *The first conjugate time t_c as a function of $a \in \mathcal{A}$ for four optimal control problems of the form* (4.3) *with* (a) *F_i given by* (4.5)–(4.7) *and* (b) *F_i given by* (4.8). *Local solutions of* (4.3) *correspond to $a \in \mathcal{A}$ for which $t_c(a) > T = 2\pi$. The sets of local solutions for these four example problems exhibit distinct topological properties.*

Example 3. As a more interesting nonlinear example, now consider

$$F_3(q) = -\frac{1}{4}q^4. \tag{4.6}$$

Applying Theorem 3.1 gives the Hamiltonian system

$$\dot{q} = p, \qquad \dot{p} = -q^3,$$

and $\partial q(t,a)/\partial a$ is found by solving (3.8) and (3.9), which are given by

$$\dot{\mathbf{J}} = \mathbf{M}, \qquad \dot{\mathbf{M}} = -3q(t)^2\mathbf{J}.$$

If we choose $a = 0$, then $\mathbf{J}(t) = t$, so $\partial q(t,a)/\partial a > 0$ for all $t \in (0, 2\pi]$. Thus $\mathcal{A}_{\min}$ is not empty, since $0 \in \mathcal{A}_{\min}$. The first conjugate time t_c was computed numerically and is shown in Figure 2a for a range of values of $a \in \mathcal{A}$. The set $\mathcal{A}_{\min}$ corresponds to points a for which $t_c(a) > 2\pi$. We see that in this example, $t_c(a)$ decreases as the magnitude of a increases, and $\mathcal{A}_{\min}$ consists of one path-connected component. Therefore, by Lemma 4.1, $\mathcal{C}_{\min}$ is path-connected.

Example 4. Now consider a second nonlinear problem with

$$F_4(q) = \cos(q). \tag{4.7}$$

We claim that for this problem, $\mathcal{A}_{\min}$ is disconnected. Applying Theorem 3.1 gives the Hamiltonian system

$$\dot{q} = p, \qquad \dot{p} = -\sin(q).$$

We can now use (3.8) and (3.9) to find $\partial q(t,a)/\partial a$ by solving

$$\dot{\mathbf{J}} = \mathbf{M}, \qquad \dot{\mathbf{M}} = -\cos(q(t))\mathbf{J}.$$

If we choose $a = 0$, then $\mathbf{J}(\pi) = 0$. Thus $\partial q(t,a)/\partial a = 0$ at $t = \pi \in (0, 2\pi]$ and $0 \notin \mathcal{A}_{\min}$. In Appendix A, we show that if $|a| > 2$, then $a \in \mathcal{A}_{\min}$. Thus $\mathcal{A}_{\min}$ is not path-connected. The first conjugate time t_c (computed numerically) is shown in Figure 2a for a range of $a \in \mathcal{A}$. Since $a \in \mathcal{A}_{\min}$ if and only if $t_c(a) > 2\pi$, we see that $\mathcal{A}_{\min}$ has two path-connected components.

Example 5. As a final example, consider

$$F_5(q) = -\cos(q) - \frac{1}{8}q^2. \tag{4.8}$$

Theorem 3.1 gives the Hamiltonian system

$$\dot{q} = p, \qquad \dot{p} = \sin(q) - \frac{1}{4}q,$$

and $\partial q(t,a)/\partial a$ is found by solving (3.8) and (3.9), which are given by

$$\dot{\mathbf{J}} = \mathbf{M}, \qquad \dot{\mathbf{M}} = \left(\cos(q(t)) - \frac{1}{4}\right)\mathbf{J}.$$

If we choose $a = 0$, then

$$\mathbf{J}(t) = \frac{1}{\sqrt{3}}\left(e^{\frac{\sqrt{3}}{2}t} - e^{-\frac{\sqrt{3}}{2}t}\right),$$

and $\partial q(t,a)/\partial a > 0$ for all $t \in (0, 2\pi]$. Therefore, $0 \in \mathcal{A}_{\min}$. The first conjugate time t_c was computed numerically and is shown in Figure 2b. As a varies, $t_c(a)$ oscillates around 2π, and therefore $\mathcal{A}_{\min}$ consists of many path-connected components.

Based on these five examples, we have seen that the set $\mathcal{A}_{\min}$ can have distinct topological properties for different optimal control problems. If the optimal control problem cannot be solved analytically (as was the case in Examples 3, 4, and 5), then showing that $\mathcal{A}_{\min}$ is path-connected without resorting to numerical computations is a nontrivial task.

5. Sufficient conditions for a path-connected set of local optima.

In section 4.2, we saw that the topological properties of $\mathcal{A}_{\min}$, defined in (4.1), vary for different optimal control problems, and there was no obvious way to determine whether $\mathcal{A}_{\min}$ is path-connected. In this section, we give a sufficient condition for $\mathcal{A}_{\min}$ to be path-connected. Our main result is stated in section 5.1 and proved in section 5.2.

5.1. The main result.

In Theorem 5.1, we state a sufficient condition for the set $\mathcal{A}_{\min}$ to be path-connected, which sometimes allows us to determine the topological properties of the set of local optima. Since these are sufficient conditions, there are optimal control problems for which these conditions are not satisfied, but $\mathcal{A}_{\min}$ is path-connected. We will see examples of this when we apply Theorem 5.1 to the examples considered in section 4.2. Before stating the theorem, we need one piece of notation. For $L > 0$ and $\mu \in \mathbb{R}^n$, define the matrix $S_\mu^L \in \mathbb{R}^{n\times n}$ by $S_\mu^L = \operatorname{diag}(L^{\mu_1}, \ldots, L^{\mu_n})$.

THEOREM 5.1 (sufficient conditions for a path-connected set of local optima). *Suppose $\mathcal{A}$ is path-connected. If there exist $\mu \in \mathbb{R}^n$ and $\mu_0 \in \mathbb{R}$ such that the Hamiltonian function in* (3.5) *satisfies*

$$H(S_\nu^L p, S_\mu^L q) = L^{\mu_0+1}H(p,q) \tag{5.1}$$

for all p and $q \in \mathbb{R}^n$ and positive real numbers $L > 0$, where $\nu \in \mathbb{R}^n$ satisfies $\nu_i = \mu_0 - \mu_i$, then $\mathcal{A}_{\min}$ is path-connected.

This sufficient condition allows us to check whether $\mathcal{A}_{\min}$ is path-connected by only checking that the Hamiltonian function H satisfies a certain scaling property. It is somewhat surprising that the set of all local optima can be characterized in this way without explicitly computing the sufficient conditions for optimality described in Theorem 3.2. As we will show in the proof of Theorem 5.1, this scaling property allows us to construct paths in the set $\mathcal{A}_{\min}$ connecting any two local solutions.

Looking back to the example problems considered in section 4.2 and applying Theorem 5.1 to the problem with F_i given by (4.4), the Hamiltonian H is given by

$$H(p, q) = \frac{1}{2}p^2.$$

Choosing $\mu = 0$ and $\mu_0 = 1$ satisfies the condition in (5.1). Therefore, we could have determined that $\mathcal{C}_{\min}$ is path-connected by using Theorem 5.1.

For the example problem with F_i given by (4.5), the Hamiltonian function is

$$H(p, q) = \frac{1}{2}\left(p^2 + q^2\right).$$

No choice of μ and μ_0 satisfies (5.1). In this example, if we had set the final time to be less than π, then $\mathcal{A}_{\min} = \mathbb{R}$, which is path-connected. This example demonstrates that Theorem 5.1 is not necessary for a path-connected set of local optima.

With F_i given by (4.6), the Hamiltonian is given by

$$H(p, q) = \frac{1}{2}p^2 + \frac{1}{4}q^4.$$

Choosing $\mu = 1$ and $\mu_0 = 3$ satisfies the condition in (5.1), so we could have concluded that $\mathcal{C}_{\min}$ is path-connected using Theorem 5.1.

With F_i given by (4.7), the Hamiltonian function is

$$H(p, q) = \frac{1}{2}p^2 - \cos(q).$$

Since $\mathcal{A}_{\min}$ is not path-connected for this problem, the converse of Theorem 5.1 tells us that no choice of μ and μ_0 satisfies (5.1). This fact can easily be checked. The same result is true with F_i given by (4.8).

5.2. Proof of the main result. In this section, we prove Theorem 5.1. We first show that the condition (5.1) in Theorem 5.1 implies that the necessary and sufficient conditions in Theorems 3.1 and 3.2 are scale-invariant with respect to time. This property allows us to define scaling relationships between certain solutions of (3.3) and (3.10). These scaling relationships will provide us with canonical paths in the set $\mathcal{A}_{\min}$, i.e., canonical ways of deforming local optima of (3.1). We then use these paths to show that any two points in $\mathcal{A}_{\min}$ can be connected by a path contained in $\mathcal{A}_{\min}$. This result shows that $\mathcal{A}_{\min}$ (and therefore $\mathcal{C}_{\min}$, defined in (4.2)) is path-connected.

Scale-invariant Hamiltonian systems. In section 2, we briefly discussed a property satisfied by some systems of differential equations called scale invariance. This property is often defined for a general system of differential equations [20, 28, 4]. The following definition specializes this general property to Hamiltonian systems.

DEFINITION 5.2. *Consider the $2n$-dimensional Hamiltonian system*

$$\frac{dQ}{dt} = \frac{\partial H(P,Q)}{\partial P}, \qquad \frac{dP}{dt} = -\frac{\partial H(P,Q)}{\partial Q}, \tag{5.2}$$

where H is smooth. If there exist μ and $\nu \in \mathbb{R}^n$ such that for any $L > 0$ and any solution (P, Q) of (5.2)*, the functions $q_i(t) = L^{\mu_i} Q_i(Lt)$ and $p_i(t) = L^{\nu_i} P_i(Lt)$ satisfy*

$$\frac{dq}{dt} = \frac{\partial H(p,q)}{\partial p}, \qquad \frac{dp}{dt} = -\frac{\partial H(p,q)}{\partial q},$$

then we say that the system (5.2) *is scale-invariant.*

When analyzing a scale-invariant Hamiltonian system, knowing a single solution (P, Q) of (5.2) provides a family of solutions defined by

$$(p(t), q(t)) = (S_\nu^L P(Lt), S_\mu^L Q(Lt)), \tag{5.3}$$

with $L \in (0, \infty)$. Two solutions of (5.2) in this family can then be continuously deformed into one another within the set of solutions of (5.2) by continuously varying the scaling parameter L.

Checking that the conditions in Definition 5.2 hold for all solutions (P, Q) of (5.2) requires us to compute these solutions, which is often not possible. The following lemma provides a condition in terms of the Hamiltonian function for the system (5.2) to be scale-invariant. This is easier to check than the conditions in Definition 5.2.

LEMMA 5.3. *If there exist $\mu \in \mathbb{R}^n$ and $\mu_0 \in \mathbb{R}$ such that the Hamiltonian function H in* (5.2) *satisfies the condition* (5.1)*, then the system* (5.2) *is scale-invariant.*

Proof. Let $L > 0$, let (P, Q) be a solution of (5.2), define $\nu \in \mathbb{R}^n$ as in Theorem 5.1, and define (p, q) as in (5.3). Using (5.1), (5.2), (5.3), and the fact that $\mu_i + \nu_i = \mu_0$ for each $i = 1, \ldots, n$, we have

$$\begin{aligned}
\frac{d}{dt}(q_i(t)) &= \frac{d}{dt}(L^{\mu_i} Q_i(Lt)) \\
&= L^{1+\mu_i} \frac{d}{d(Lt)}(Q_i(Lt)) \\
&= L^{1+\mu_i} \frac{\partial}{\partial P_i}(H(P(Lt), Q(Lt))) \\
&= L^{1+\mu_i+\nu_i} \frac{\partial}{\partial p_i}(H(P(Lt), Q(Lt))) \\
&= \frac{\partial}{\partial p_i}(L^{1+\mu_0} H(P(Lt), Q(Lt))) \\
&= \frac{\partial}{\partial p_i}(H(S_\nu^L P(Lt), S_\mu^L Q(Lt))) \\
&= \frac{\partial}{\partial p_i} H(p(t), q(t)).
\end{aligned}$$

A similar calculation shows that

$$\frac{d}{dt}(p_i(t)) = -\frac{\partial}{\partial q_i} H(p(t), q(t)). \tag{5.4}$$

∎

For given $\mu \in \mathbb{R}^n$ and $\mu_0 \in \mathbb{R}$, if the Hamiltonian function H satisfies the conditions in Lemma 5.3, then we say that H is invariant under the scaling (μ, μ_0).

Thus far, we have only considered solutions of the Hamiltonian system (5.2). The following lemma shows that if the Hamiltonian function H is invariant under the scaling (μ, μ_0), then solutions of (3.10) also satisfy scaling relationships.

LEMMA 5.4. *Assume H is invariant under the scaling (μ, μ_0), and define $\nu \in \mathbb{R}^n$ as in Theorem* 5.1. *Let (P, Q) be a solution of* (5.2), *let $A = P(0)$, and let*

$$[\mathbf{J}]_{ij}(t) = \frac{\partial Q_i(t, A)}{\partial A_j},$$

where we have extended Q to depend explicitly on A, as was done in section 3.2. *Let $L > 0$, define (p, q) as in* (5.3), *and let $a = p(0)$. Then the matrix $\widehat{\mathbf{J}}(t)$ with entries*

$$[\widehat{\mathbf{J}}]_{ij}(t) = \frac{\partial q_i(t, a)}{\partial a_j}$$

satisfies

$$\widehat{\mathbf{J}}(t) = \left(S_\mu^L\right) \mathbf{J}(Lt) \left(S_\nu^L\right)^{-1}. \tag{5.5}$$

Proof. For each component of $\widehat{\mathbf{J}}(t)$, we have

$$[\widehat{\mathbf{J}}]_{ij}(t) = \frac{\partial q_i(t, a)}{\partial a_j} = \frac{\partial L^{\mu_i} Q_i(Lt, A)}{\partial L^{\nu_j} A_j} = \left[S_\mu^L\right]_{ii} \Big([\mathbf{J}]_{ij}(Lt)\Big) \left[S_\nu^L\right]_{jj}^{-1}.$$

We conclude that (5.5) holds. ∎

Constructing paths in the set of local solutions. We have shown that the symmetry admitted by a scale-invariant Hamiltonian system partitions the solutions of the Hamiltonian system into families, with solutions in each family related by (5.3). This symmetry also provides a canonical way of continuously deforming a solution into any other solution in the same family. In section 3, Theorems 3.1 and 3.2 relate local solutions of the optimal control problem (3.1) to solutions of the Hamiltonian system (3.3). Thus, if the Hamiltonian system (3.3) is scale-invariant, the local solutions of (3.1) might be partitioned in a similar way. This result is not immediately obvious since local solutions of (3.1) do not correspond to all solutions of (3.3), but only to normal solutions of (3.3) that satisfy the sufficient conditions in Theorem 3.2.

The following two lemmas show that, even with these two additional constraints (i.e., normality and the absence of conjugate points), the set of local solutions of (3.1) can be partitioned using the relation (5.3). First, recall from (3.6) that $\mathcal{A}$ is the set of initial costate values $p(0)$ that map to normal extremals (p, q). The following lemma shows that for any $a \in \mathcal{A}$, we can scale a according to (5.3) and remain in the set $\mathcal{A}$.

LEMMA 5.5. *Assume H is invariant under the scaling (μ, μ_0) and define $\nu \in \mathbb{R}^n$ as in Theorem* 5.1. *If $a \in \mathcal{A}$ and $L > 0$, then $S_\nu^L a \in \mathcal{A}$.*

Proof. Let $a \in \mathcal{A}$, $L > 0$, $(P, Q) = \Gamma(a)$, and $(p, q) = \Gamma(S_\nu^L a)$. By Lemma 5.3, the Hamiltonian system (3.3) is scale-invariant. Note that

$$(p(0), q(0)) = (S_\nu^L P(0), S_\mu^L Q(0)).$$

By existence and uniqueness of solutions of (3.3) (from assumption (A1)), (P, Q) and (p, q) must satisfy (5.3). Thus, by Definition 5.2, both (P, Q) and (p, q) solve (3.3) for the same Hamiltonian function. If (p, q) is abnormal, this implies (P, Q) is abnormal, which contradicts $a \in \mathcal{A}$. We conclude that $S_\nu^L a \in \mathcal{A}$. ∎

The previous lemma shows that the set $\mathcal{A}$ can be partitioned using the scaling $a \mapsto S_\nu^L a$ with $L > 0$. The next lemma shows that the same is true for $\mathcal{A}_{\min}$ (recall from (4.1) that $\mathcal{A}_{\min}$ is the set of all points in $\mathcal{A}$ corresponding to normal extremals without conjugate points). Before stating this lemma, recall the result that conjugate times along normal extremals are positive and discrete [1]. Therefore, if the normal extremal $(p, q) = \Gamma(a)$ for some $a \in \mathcal{A}$ has conjugate times, there exists a smallest conjugate time. Denote this smallest conjugate time by $t_c(a)$ as was done in section 4. If the extremal $\Gamma(a)$ does not have conjugate times, let $t_c(a) = \infty$.

Lemma 5.6. *Assume H is invariant under the scaling (μ, μ_0), and define $\nu \in \mathbb{R}^n$ as in Theorem* 5.1. *If $a \in \mathcal{A}$ and $0 < L < t_c(a)/T$, then $S_\nu^L a \in \mathcal{A}_{\min}$.*

Proof. Let $a \in \mathcal{A}$, $0 < L < t_c(a)/T$, $(P, Q) = \Gamma(a)$, and $(p, q) = \Gamma(S_\nu^L a)$. The Hamiltonian system (3.3) is scale-invariant by Lemma 5.3, and (p, q) is a normal extremal of (3.1) by Lemma 5.5.

Let $\mathbf{J}(t)$ be the Jacobian matrix computed in (3.9) of the endpoint map ϕ_t along the normal extremal (P, Q), and let $\widehat{\mathbf{J}}(t)$ be the Jacobian along the normal extremal (p, q). By Lemma 5.4, $\mathbf{J}(t)$ and $\widehat{\mathbf{J}}(t)$ are related by (5.5). Since $L > 0$, S_μ^L and S_ν^L are nonsingular. Thus $\widehat{\mathbf{J}}(t)$ is singular if and only if $\mathbf{J}(Lt)$ is singular.

By definition, $t_c(a)$ is the first time at which $\mathbf{J}(t)$ is singular. For $t \in [0, T]$, we have $Lt < t_c(a)$ by assumption. Therefore, $\det(\widehat{\mathbf{J}}(t)) \neq 0$ for all $t \in [0, T]$. By Theorem 3.2, we conclude that $S_\nu^L a \in \mathcal{A}_{\min}$. ∎

The previous lemma shows that, given any $a \in \mathcal{A}$, the curve $S_\nu^L a$, with $0 < L < t_c(a)/T$, is completely contained in $\mathcal{A}_{\min}$. To show that $\mathcal{A}_{\min}$ is path-connected, we will show that any two such curves can be connected using a third curve completely contained in $\mathcal{A}_{\min}$. This will establish Theorem 5.1. Before proving this result, we need the following lemma, which shows that conjugate points are bounded from below along continuous paths in $\mathcal{A}$.

Lemma 5.7. *Let $\alpha \colon [0, 1] \to \mathcal{A}$ be continuous. Then there exists $\epsilon > 0$ such that*

$$\epsilon < \inf_{s \in [0,1]} t_c(\alpha(s)).$$

Proof. See Proposition 2.6 of Sachkov [23]. ∎

Proof of Theorem 5.1. We can now prove our main result. In the proof, we will only consider the case when $\mathcal{A}_{\min}$ is nonempty. If $\mathcal{A}_{\min} = \emptyset$, we will consider it to be path-connected.

Let $a_0, a_1 \in \mathcal{A}_{\min}$. Let $\alpha \colon [0, 1] \to \mathcal{A}$ be a continuous path such that $\alpha(0) = a_0$ and $\alpha(1) = a_1$. Such a path exists since $\mathcal{A}$ is assumed to be path-connected. Choose $L_{\min}$ such that

$$0 < L_{\min} < \min \left\{ \frac{1}{T} \left(\inf_{s \in [0,1]} t_c(\alpha(s)) \right), 1 \right\},$$

which is possible by Lemma 5.7. Now consider the path $\alpha_1 \colon [L_{\min}, 1] \to \mathcal{A}$ given by

$$\alpha_1(s) = S_\nu^s a_0. \tag{5.6}$$

Since $a_0 \in \mathcal{A}_{\min}$, we have $T < t_c(a_0)$. This implies that $s < t_c(a_0)/T$ for all $s \in [L_{\min}, 1]$. Therefore, by Lemma 5.6, we have $\alpha_1(s) \in \mathcal{A}_{\min}$ for all $s \in [L_{\min}, 1]$. Next, consider the path $\alpha_2 \colon [0, 1] \to \mathcal{A}$ given by

$$\alpha_2(s) = S_\nu^{L_{\min}} \alpha(s). \tag{5.7}$$

Since $L_{\min} < t_c(\alpha(s))/T$ for all $s \in [0,1]$, we have $\alpha_2(s) \in \mathcal{A}_{\min}$ for all $s \in [0,1]$ by Lemma 5.6. Finally, consider the path $\alpha_3 \colon [L_{\min}, 1] \to \mathcal{A}$ given by

$$\alpha_3(s) = S_\nu^s a_1. \tag{5.8}$$

Since $a_1 \in \mathcal{A}_{\min}$, we have $T < t_c(a_1)$. This implies that $s < t_c(a_1)/T$ for all $s \in [L_{\min}, 1]$. By Lemma 5.6, we have $\alpha_3(s) \in \mathcal{A}_{\min}$ for all $s \in [L_{\min}, 1]$.

The union $\beta = \alpha_1 \cup \alpha_2 \cup \alpha_3$ is a continuous path in $\mathcal{A}_{\min}$ connecting a_0 and a_1. We conclude that $\mathcal{A}_{\min}$ (and therefore $\mathcal{C}_{\min}$) is path-connected. ∎

6. Sets of local optima under additional constraints.

Thus far, we have considered the set of all local optima for all possible choices of the terminal state. In some situations, it may be of interest to consider the set of all local optima that satisfy some additional constraints. The scale invariance property used in the previous section to prove Theorem 5.1 can sometimes be used to show that the set of constrained local optima is path-connected. As an example, suppose we want to show that the set

$$\left\{ (q,u) \in \mathcal{C}_{\min} \colon \|q_j(t)\|_\infty \equiv \max_{t \in [0,T]} |q_j(t)| \leq q_{\max} \right\} \tag{6.1}$$

is path-connected for some $j \in \{1, 2, \ldots, n\}$ and some $q_{\max} > 0$. Showing that this set is path-connected is equivalent to showing that some subset of $\mathcal{A}_{\min}$ is path-connected.

Suppose that $(q,u) = \Psi(a)$ for some $a \in \mathcal{A}_{\min}$ satisfies the constraint on $q_j(t)$ and the Hamiltonian function (3.5) is invariant under the scaling (μ, μ_0). Then, if $(q', u') = \Psi(S_\nu^L a)$ for some $0 < L \leq 1$, we have from (5.3) that $q'_j(t) = L^{\mu_j} q_j(Lt)$ and

$$\|q'_j(t)\|_\infty = L^{\mu_j} \|q_j(Lt)\|_\infty \leq L^{\mu_j} \|q_j(t)\|_\infty \leq L^{\mu_j} q_{\max}.$$

Therefore, if $\mu_j > 0$, $\|q'_j(t)\|_\infty$ decreases as L decreases. To show that the constrained set of local optima (6.1) is path-connected, the proof of Theorem 5.1 can be modified by choosing $L_{\min}$ to be small enough so that the constraint on $q_j(t)$ is satisfied at each point along the path β.

We conclude that if $\mu_j > 0$, then the set (6.1) is path-connected. To enforce a bound on $\|p_j(t)\|_\infty$, we require that $\nu_j > 0$. Bounds on multiple states and costates are handled by requiring the corresponding components of μ and ν to be positive.

Arguments based on scale invariance similar to those just used can be applied to other types of constraints. In section 6.1 below, we give particular consideration to a constraint that bounds the total cost associated to each locally optimal trajectory.

6.1. Trajectories with bounded total cost.

In this section, we consider the set of all optima with a bounded total cost. If we assume that the cost function is nonnegative and scales in a similar way to the Hamiltonian in Theorem 5.1, then the set of all local optima with a bounded total cost is path-connected. To prove this result, first recall that by using assumptions (A1)–(A3) and the implicit function theorem, we can locally write the optimal control u found from condition (3.4) as a function of the costate p and the state q. Define this function by $u(t) = \Lambda(p(t), q(t))$. Furthermore, define the functions $f_\Lambda \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and $g_\Lambda \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by

$$f_\Lambda(p,q) = f(q(t), \Lambda(p,q)), \qquad g_\Lambda(p,q) = g(q(t), \Lambda(p,q)).$$

The cost function and dynamic constraints can now be expressed as functions of the extremal (p,q). The following lemma gives conditions on f_Λ and g_Λ for the Hamiltonian system (3.3) to be scale-invariant.

LEMMA 6.1. *If there exist $\mu \in \mathbb{R}^n$ and $\mu_0 \in \mathbb{R}$ such that the cost function and dynamic constraints in* (3.1) *satisfy*

$$g_\Lambda(S_\nu^L p, S_\mu^L q) = L^{\mu_0+1} g_\Lambda(p,q), \qquad f_\Lambda(S_\nu^L p, S_\mu^L q) = L S_\mu^L f_\Lambda(p,q) \tag{6.2}$$

for all p and $q \in \mathbb{R}^n$ and positive real numbers $L > 0$, where $\nu \in \mathbb{R}^n$ satisfies $\nu_i = \mu_0 - \mu_i$, then the Hamiltonian system (3.3) *is scale-invariant.*

Proof. Using (3.2), we have

$$\begin{aligned} H(S_\nu^L p, S_\mu^L q) &= \left(S_\nu^L p\right)^T f_\Lambda(S_\nu^L p, S_\mu^L q) - k g_\Lambda(S_\nu^L p, S_\mu^L q) \\ &= p^T \left(S_\nu^L\right)^T L S_\mu^L f_\Lambda(p,q) - k L^{\mu_0+1} g_\Lambda(p,q) \\ &= L^{\mu_0+1}\left(p^T f_\Lambda(p,q) - k g_\Lambda(p,q)\right) \\ &= L^{\mu_0+1} H(p,q). \end{aligned}$$

By Lemma 5.3, the Hamiltonian function in (3.4) is invariant under the scaling (μ, μ_0), and the Hamiltonian system (3.3) is therefore scale-invariant. ∎

Recall from (3.1) that $C(q,u)$ denotes the total cost associated with the trajectory (q,u). The following lemma shows that the total cost along paths in $\mathcal{A}$ of the form $S_\nu^L a$ with $0 < L \leq 1$ and $a \in \mathcal{A}$ can be bounded.

LEMMA 6.2. *Suppose the conditions in Theorem* 5.1 *are satisfied and g_Λ satisfies the conditions in Lemma* 6.1. *Furthermore, suppose g_Λ is nonnegative. If $a \in \mathcal{A}$ and $0 < L \leq 1$, then*

$$C(\Psi(S_\nu^L a)) \leq L^{\mu_0} C(\Psi(a)).$$

Proof. Let $(P,Q) = \Gamma(a)$ and $(p,q) = \Gamma(S_\nu^L a)$, which are related by (5.3). Using (6.2), we have

$$\begin{aligned} C(\Psi(S_\nu^L a)) &= \int_0^T g_\Lambda(p(t), q(t))\, dt \\ &= \int_0^T g_\Lambda(S_\nu^L P(Lt), S_\mu^L Q(Lt))\, dt \\ &= \int_0^T L^{\mu_0+1} g_\Lambda(P(Lt), Q(Lt))\, dt \\ &= L^{\mu_0} \int_0^{LT} g_\Lambda(P(\tau), Q(\tau))\, d\tau \\ &\leq L^{\mu_0} \int_0^{T} g_\Lambda(P(\tau), Q(\tau))\, d\tau \\ &= L^{\mu_0} C(\Psi(a)). \end{aligned}$$

The inequality holds since g_Λ is nonnegative and $L \leq 1$. ∎

We can now state Theorem 6.3, which provides a sufficient condition for the set of local optima with a bounded total cost to be path-connected.

THEOREM 6.3. *Suppose the conditions in Theorem* 5.1 *are satisfied with $\mu_0 > 0$, and let $C_0 > 0$. If g_Λ is nonnegative and satisfies the conditions in Lemma* 6.1, *then the set $\{(q,u) \in \mathcal{C}_{min} \colon C(q,u) \leq C_0\}$ is path-connected.*

Proof. If the set $\{a \in \mathcal{A}_{\min} : C(\Psi(a)) \leq C_0\}$ is empty, we will consider it and $\{(q, u) \in \mathcal{C}_{\min} : C(q, u) \leq C_0\}$ to be path-connected.

Now assume the set $\{a \in \mathcal{A}_{\min} : C(\Psi(a)) \leq C_0\}$ is nonempty, and let a_0 and $a_1 \in \mathcal{A}_{\min}$ be such that $C(\Psi(a_0))$ and $C(\Psi(a_1)) \leq C_0$. Let $\alpha\colon [0,1] \to \mathcal{A}$ be a continuous path such that $\alpha(0) = a_0$ and $\alpha(1) = a_1$. Such a path exists since $\mathcal{A}$ is assumed to be path-connected. Choose L_1 such that

$$0 < L_1 < \frac{1}{T}\left(\inf_{s\in[0,1]} t_c(\alpha(s))\right),$$

which is possible by Lemma 5.7.

Since $g(q, u)$ is smooth, $C(\Psi(\alpha(s)))$ is bounded from above for all $s \in [0, 1]$. Let $C_{\max} > 0$ be such that

$$\sup_{s\in[0,1]} C(\Psi(\alpha(s))) < C_{\max}.$$

Now choose L_2 such that

$$0 < L_2 < \left(\frac{C_0}{C_{\max}}\right)^{\frac{1}{\mu_0}}.$$

Now let $L_{\min} = \min\{L_1, L_2, 1\}$ and consider the path $\beta = \alpha_1 \cup \alpha_2 \cup \alpha_3$ constructed in the proof of Theorem 5.1 for this choice of $L_{\min}$. Since $L_{\min} \leq L_1$, this path is completely contained within $\mathcal{A}_{\min}$. Also, since $\mu_0 > 0$, Lemma 6.2 gives

$$\begin{aligned} C(\Psi(\alpha_1(s))) &\leq s^{\mu_0} C(\Psi(a_0)) \leq C(\Psi(a_0)) \leq C_0, \\ C(\Psi(\alpha_3(s))) &\leq s^{\mu_0} C(\Psi(a_1)) \leq C(\Psi(a_1)) \leq C_0 \end{aligned}$$

for all $s \in [L_{\min}, 1]$. Finally, using Lemma 6.2, we have

$$C(\Psi(\alpha_2(s))) = C(\Psi(S_\nu^{L_{\min}}\alpha(s))) \leq L_{\min}^{\mu_0} C(\Psi(\alpha(s))) \leq L_2^{\mu_0} C_{\max} < C_0$$

for all $s \in [0, 1]$. Therefore, the path $\beta = \alpha_1 \cup \alpha_2 \cup \alpha_3$ is contained within the set $\{a \in \mathcal{A}_{\min} : C(\Psi(a)) \leq C_0\}$, and we conclude that this set is path-connected. Since Ψ is continuous, the image of this set, which is $\{(q, u) \in \mathcal{C}_{\min} : C(q, u) \leq C_0\}$, is path-connected. ∎

7. Application to the planar elastica. In this section, we apply the results in Theorems 5.1 and 6.3 to an optimal control problem whose associated Hamiltonian function is scale-invariant. Consider a thin inextensible wire or rod whose ends are held by robotic grippers, and assume the rod is confined to deform in a plane. We will model the rod as a planar elastica [22, 23, 25]. Without loss of generality, we may assume that the rod has length 1. A point along the rod at arc-length $t \in [0, 1]$ is described by the vector $q(t) \in \mathbb{R}^3$, where $(q_1(t), q_2(t))$ gives the position in the plane and $q_3(t)$ is the angle between the tangent to the rod and the q_1-axis. We will assume that the base of the rod is held fixed at the origin, so that $q(0) = 0$.

As discussed in section 2.1, we use a quasi-static model of the rod. This model ignores the dynamics of the rod and assumes that the grippers holding the ends of the rod move slow enough so that the rod is in static equilibrium at each point in time. Under this quasi-static assumption, the configuration of the rod for a given set of boundary conditions (i.e., robot gripper poses) is a shape that locally minimizes the stored elastic energy in the rod. We call a shape that locally minimizes the elastic energy a stable equilibrium configuration. Under the assumptions of linear elasticity, the elastic energy is proportional to the integral of the squared curvature along the length of the rod. The shape of the rod is a local solution of the optimal control problem

$$\begin{aligned} &\underset{q,u}{\text{minimize}} \quad C(q,u) = \frac{1}{2}\int_0^1 u(t)^2\, dt \\ &\text{subject to} \quad \dot{q}(t) = \begin{bmatrix} \cos(q_3(t)) \\ \sin(q_3(t)) \\ u(t) \end{bmatrix}, \\ &\qquad\qquad\quad q(0) = 0,\ q(1) = b \end{aligned} \tag{7.1}$$

for some gripper placement $b \in \mathbb{R}^3$. In the optimal control problem (7.1), the control input u is the bending strain along the rod, the cost function is the elastic energy stored in the rod, and the dynamic constraints ensure that the rod is inextensible. Although we assume that the rod behaves according to the laws of linear elasticity, we are using a geometrically exact (i.e., geometrically nonlinear) model of the rod. Therefore, while linear elasticity assumes that the bending strain u stays small, changes in the rod's configuration (q, u) do not need to be small.

Due to the geometric nonlinearity of the elastic rod model (7.1), specifying the boundary condition b in (7.1) does not uniquely determine the configuration (q, u) of the rod. This ambiguity is one reason why finding a path of the robotic gripper holding the end of the rod that causes the rod to move from an initial shape to a goal shape appears to be challenging. However, we have shown that points in the space $\mathcal{A}$ are in one-to-one correspondence with normal (q, u) of (7.1), i.e., equilibrium configurations of the rod. Furthermore, we have shown that points in $\mathcal{A}_{\min}$ are in one-to-one correspondence with local optima of (7.1), i.e., stable equilibrium configurations of the rod. As a consequence, we can use Theorem 5.1 to construct a path in $\mathcal{A}_{\min}$ that corresponds to a continuous deformation of the rod through the set of stable equilibrium configurations, and that can be implemented without ambiguity by a path of the robotic gripper. Note the importance of keeping the rod in static equilibrium while it is being deformed—otherwise, the quasi-static assumption we made in deriving a model of the rod would be violated.

We begin by applying Theorem 3.1 to the optimal control problem (7.1). The parameterized Hamiltonian is given by

$$\widehat{H}(p, q, k, u) = p_1 \cos(q_3) + p_2 \sin(q_3) + p_3 u - \frac{k}{2} u^2.$$

First, consider the abnormal case with $k = 0$. The parameterized Hamiltonian $\widehat{H}$ is extremized in u when $p_3 = 0$. Now consider the normal case with $k = 1$. The optimal control, found from (3.4), is $u = p_3$. The Hamiltonian system (3.3) is given by

$$\begin{aligned} &\dot{q}_1 = \cos(q_3), \qquad \dot{q}_2 = \sin(q_3), \qquad \dot{q}_3 = p_3, \\ &\dot{p}_1 = 0, \qquad\qquad\ \ \dot{p}_2 = 0, \qquad\qquad\ \ \dot{p}_3 = p_1 \sin(q_3) - p_2 \cos(q_3). \end{aligned} \tag{7.2}$$

In the Hamiltonian system (7.2), the costate trajectory p has a physical interpretation. The costates p_1 and p_2 describe the forces acting on the rod, and the costate p_3 describes the torque acting on the rod [7].

Abnormal extremals are achieved by choosing the initial condition $p_2(0) = p_3(0) = 0$. Physically, the abnormal extremals correspond to the straight configuration of the rod. Therefore, using our earlier notation,

$$\mathcal{A} = \{(a_1, a_2, a_3) \in \mathbb{R}^3 \colon (a_2, a_3) \neq (0, 0)\},$$

which is path-connected. The existence and uniqueness assumptions in (A1) are satisfied by (7.2) [22]. To establish assumption (A2), it suffices to show that (p, q) can be uniquely determined from a normal (q, u) and its derivatives. Taking derivatives of the differential equation for p_3 in (7.2) and recalling that $u = p_3$, we have

$$p_3(0) = u(0), \quad -p_2(0) = \dot{u}(0), \quad p_1(0)p_3(0) = \ddot{u}(0), \quad p_2(0)(p_3(0)^2 - p_1(0)) = \dddot{u}(0).$$

Since (q, u) is normal, $p_2(0)$ and $p_3(0)$ cannot both be zero. Therefore, these four equations allow us to uniquely solve for $p(0)$, which allows us to find (p, q) by solving (7.2). Finally, note that assumption (A3) is satisfied since $\partial^2 \widehat{H}/\partial u^2 = -1$. Theorem 3.2 can therefore be used as a test for local optimality.

We now claim that the Hamiltonian function

$$H(p, q) = p_1 \cos(q_3) + p_2 \sin(q_3) + \frac{1}{2} p_3^2 \tag{7.3}$$

satisfies the conditions in Theorem 5.1. Let $\mu = (-1, -1, 0)$ and $\mu_0 = 1$. Then, from Theorem 5.1, we have $\nu = (2, 2, 1)$ and

$$H(S_\nu^L p, S_\mu^L q) = L^2 \left(p_1 \cos(q_3) + p_2 \sin(q_3) + \frac{1}{2} p_3^2 \right) = L^{\mu_0 + 1} H(p, q)$$

for all $L > 0$. Thus the Hamiltonian function (7.3) is invariant under the scaling (μ, μ_0). We can therefore conclude the following result using Theorem 5.1.

Theorem 7.1. *The set of all stable equilibrium configurations of the planar elastica (i.e., $\mathcal{C}_{\min}$) is path-connected.*

In [25], a decomposition of $\mathcal{A}_{\min}$ into disjoint sets is defined, and it is shown that some of these sets are path-connected. This result is proved by analyzing the properties of the exponential map, which sends initial costate values (i.e., $a \in \mathcal{A}$) to state trajectories $q(t)$. The scale invariance property that we use in Theorem 5.1 allows us to show that the entire set $\mathcal{A}_{\min}$ is path-connected.

Furthermore, it is clear that $g_\Lambda(p, q) = \frac{1}{2} p_3^2$ is nonnegative and satisfies the conditions in Lemma 6.1. Therefore, since $\mu_0 > 0$, we can conclude the following result using Theorem 6.3.

Theorem 7.2. *The set of all stable equilibrium configurations of the planar elastica with an elastic potential energy less than a given bound is path-connected.*

To see how these results can be applied to the problem of manipulation planning, assume that initial and goal stable equilibrium configurations of the rod are given. This is equivalent to providing two points a_0 and $a_1 \in \mathcal{A}_{\min}$. Using the construction in the proof of Theorem 5.1, we can find a path $\beta \colon [0, 1] \to \mathcal{A}_{\min}$ connecting these two points. If we require the stored elastic energy in the rod to be bounded, the construction in the proof of Theorem 6.3 can be used.

Once the path $\beta(s)$ is found, the map $\Psi \circ \beta$ gives a continuous sequence of local solutions of (7.1), given by $(q_s(t), u_s(t)) \equiv \Psi \circ \beta(s)$ with $s \in [0, 1]$. To execute this

FIG. 3. *A sequence of configurations of a planar elastica corresponding to the path α in Figure* 4a. *The leftmost and rightmost configurations are local solutions of* (7.1)*, i.e., they are stable equilibrium configurations. The six intermediate configurations are not local minima of* (7.1) *and are therefore unstable equilibrium configurations.*

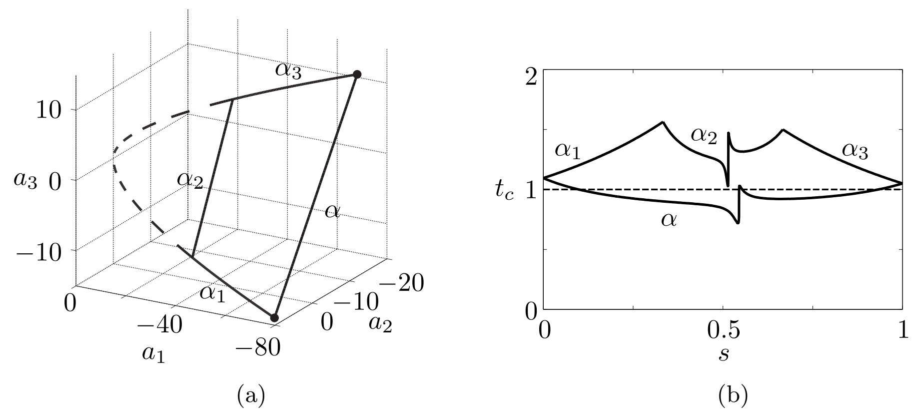

FIG. 4. (a) *The space $\mathcal{A} \subset \mathbb{R}^3$ corresponding to all initial costate values that lead to normal extremals of* (7.2)*, and* (b) *the first conjugate time along the paths α, α_1, α_2, and α_3. The path α (which is not contained in $\mathcal{A}_{min}$) corresponds to the sequence of configurations in Figure* 3*. The paths α_1, α_2, and α_3 from* (5.6)–(5.8) *are contained in $\mathcal{A}_{min}$ and correspond to the sequence of rod configurations in Figure* 5.

sequence of stable equilibrium configurations, the robotic gripper holding the rod at $T = 1$ should follow the path $q_s(1)$ for $s \in [0, 1]$.

An example of this procedure is shown in Figures 3–5. Consider the problem of manipulating the leftmost rod in Figure 3 into the rightmost rod in Figure 3. Both of these rod configurations are local solutions of (7.1). The three-dimensional space $\mathcal{A}$ is shown in Figure 4a, and the leftmost and rightmost rod configurations in Figure 3 correspond to the black circles in Figure 4a. These two points are connected by a straight line, which we denote by $\alpha\colon [0, 1] \to \mathcal{A}$. The sequence of rod configurations in Figure 3 corresponds to moving along the path α.

The lower line in Figure 4b shows the first conjugate time at each point along the path α in $\mathcal{A}$. We see that at the endpoints of the path, the first conjugate time occurs after the final time $T = 1$, verifying that the leftmost and rightmost rod configurations in Figure 3 are local solutions of (7.1). In between the endpoints, conjugate times

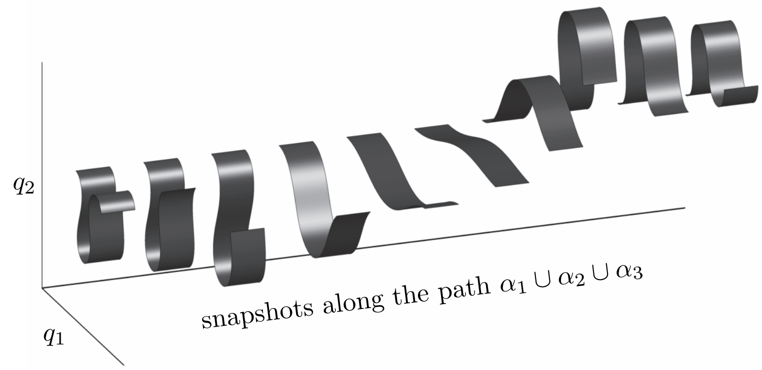

FIG. 5. *A sequence of configurations of a planar elastica corresponding to the paths α_1, α_2, and α_3 in Figure* 4a. *All of these configurations are local minima of* (7.1) *and are therefore stable equilibrium configurations of the elastic rod. Note that the leftmost and rightmost configurations in this figure are, respectively, identical to the leftmost and rightmost configurations in Figure* 3.

occur before the final time $T = 1$. Therefore, the path α is not contained in $\mathcal{A}_{\min}$. In fact, the intermediate rod configurations in Figure 3 do not locally minimize the total cost in (7.1) and are therefore unstable equilibrium configurations.

Using the procedure described in Theorem 5.1, we can deform the path α into a path in $\mathcal{A}_{\min}$. We see from Figure 4b that along α, the smallest conjugate time is approximately 0.75. We can choose $L_{\min}$ in Theorem 5.1 to be 0.7. The three paths α_1, α_2, and α_3, defined in (5.6)–(5.8), are shown in the space $\mathcal{A}$ in Figure 4a. The dashed curves in Figure 4a show the paths α_1 and α_3 extended to the origin, i.e., the dashed curves are $\alpha_1(s)$ and $\alpha_3(s)$ with $s \in [0, L_{\min}]$.

The upper line in Figure 4b shows the first conjugate time along the curves α_1, α_2, and α_3 (where the length of the curve $\alpha_1 \cup \alpha_2 \cup \alpha_3$ has been normalized to 1). As expected, the first conjugate time occurs after $T = 1$ along the entire path $\alpha_1 \cup \alpha_2 \cup \alpha_3$, so this path is contained in $\mathcal{A}_{\min}$. Figure 5 shows a sequence of rod configurations along this path in $\mathcal{A}_{\min}$, with the first three configurations corresponding to α_1, the next four corresponding to α_2, and the final three configurations corresponding to α_3.

8. Conclusion.

We have derived a sufficient condition for the set of all local optima of an optimal control problem over all possible terminal state constraints to be path-connected. This condition relies on the assumption that the necessary and sufficient conditions for optimality are scale-invariant. This scale invariance property can easily be checked by showing that the Hamiltonian function provided by the Pontryagin maximum principle satisfies the scaling property in Theorem 5.1. We also showed that under a similar scaling condition on the cost function, the set of all local optima with a total cost less than a given bound is path-connected.

We then applied these results to an optimal control problem whose local solutions are stable equilibrium configurations of a planar elastic rod. We showed that the set of all stable equilibrium configurations is path-connected, and we provided an example of how to construct a path in this set. This procedure can be used to plan paths of robotic grippers holding the ends of the rod that cause the rod to move between any two stable equilibrium configurations.

Many extensions of the results in this paper could be explored in future work. The conditions in Theorems 5.1 and 6.3 are sufficient, but not necessary for a path-connected set of local optima. More complete characterizations of optimal control problems with path-connected sets of local optima, in terms of either less stringent sufficient conditions or conditions that are both necessary and sufficient, would be interesting. Depending upon the applications being considered, constraints other than those analyzed in section 6 could be considered. For example, scale invariance is used in [6] to show that the set of all Kirchhoff elastic rods (i.e., nonplanar elasticas) that do not contain self-intersections is path-connected. This constraint corresponds to an injectivity condition on $q(t)$.

Another line of future work is to consider perturbations in elements of the optimal control problem other than the terminal state constraint. If the optimal control problem (3.1) is formulated in a geometric setting as an optimal control problem on a Lie group, invariance of the optimal control problem with respect to the initial value of the state $q(0)$ can be related to invariance of the Hamiltonian function (3.4) under the left-action of the Lie group [15]. Therefore, a geometric statement of Theorem 5.1 combined with a left-invariance assumption may be sufficient to show that the set of all local optima over all possible initial and terminal state constraints is path-connected.

Finally, extensions of the application considered in section 7 could be investigated. Conservative external forces, such as gravity, break the scale invariance property we used to establish path-connectedness of the set of local optima. However, the set of local optima may remain path-connected in the presence of these forces. The model used in section 7, although geometrically nonlinear, assumes the rod behaves according to linear elasticity. A cost function derived from nonlinear elasticity could be used in (7.1) to more accurately model an elastic rod experiencing large deformations. If this nonlinear elastic model could be shown to satisfy the scale invariance properties discussed in this paper, we could prove that the set of local optima remains path-connected under this more general model of elastic deformation.

Appendix A. Conjugate times in the optimal control problem (4.6). In this appendix, we analyze the optimal control problem (4.3) with $F_i(q)$ given by (4.6). We show that if $|a| > 2$, then $a \in \mathcal{A}_{\min}$. This result, along with the fact that $0 \notin \mathcal{A}_{\min}$ (which was shown in section 4.2), shows that $\mathcal{A}_{\min}$ is not path-connected.

Since the Hamiltonian function in Theorem 3.1 is conserved, we have

$$\frac{1}{2}p(t)^2 - \cos(q(t)) = \frac{1}{2}p(0)^2 - \cos(q(0)) = \frac{1}{2}a^2 - 1.$$

Since $\dot{q}(t) = p(t)$ and $|a| > 2$, this gives $\dot{q}(t)^2 > 0$ and

$$\dot{q}(t,a) = \operatorname{sign}(a)\sqrt{a^2 + 2\cos(q(t)) - 2}.$$

Assume that $a > 2$. We then have

$$\begin{aligned}
\frac{\partial}{\partial t}\left(\frac{\partial q(t,a)}{\partial a}\right) &= \frac{\partial}{\partial a}\left(\frac{\partial q(t,a)}{\partial t}\right) \\
&= \frac{1}{\sqrt{a^2 + 2\cos(q(t)) - 2}}\left(a - \sin(q(t))\frac{\partial q(t,a)}{\partial a}\right) \\
&> 1 - \frac{1}{\sqrt{a^2 - 4}}\left|\frac{\partial q(t,a)}{\partial a}\right|.
\end{aligned}$$

Let $\hat{a} = (a^2 - 4)^{-1/2}$, and let y be the solution of

$$\dot{y} = 1 - \hat{a}|y|, \qquad y(0) = 0.$$

Since $\partial q(0,a)/\partial a = 0$, we have $\partial q(t,a)/\partial a \geq y(t)$. The solution of this equation is

$$y(t) = \frac{1}{\hat{a}}\left(1 - e^{-\hat{a}t}\right),$$

which is positive for all $t > 0$. Therefore $\partial q(t,a)/\partial a > 0$ for all $t > 0$. A similar argument can be used for the case $a < -2$. Therefore $a \in \mathcal{A}_{\min}$ if $|a| > 2$. We conclude that $\mathcal{A}_{\min}$ is not path-connected.

REFERENCES

[1] A. A. Agrachev and Y. L. Sachkov, *Control Theory from the Geometric Viewpoint*, Springer, Berlin, 2004. doi:10.1007/978-3-662-06404-7.

[2] A. Agrachev, L. Rizzi, and P. Silveira, *On conjugate times of LQ optimal control problems*, J. Dyn. Control Syst., 21 (2015), pp. 625–641. doi:10.1007/s10883-014-9251-6.

[3] A. A. Ardentov and Yu. L. Sachkov, *Conjugate points in nilpotent sub-Riemannian problem on the Engel group*, J. Math. Sci., 195 (2013), pp. 369–390. doi:10.1007/s10958-013-1584-2.

[4] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* I. *Asymptotic Methods and Perturbation Theory*, McGraw Hill Book Company, New York, 1978; Springer Science+Business Media, New York, 1999. doi:10.1007/978-1-4757-3069-2.

[5] B. Bonnard, N. Shcherbakova, and D. Sugny, *The smooth continuation method in optimal control with an application to quantum systems*, ESAIM Control Optim. Calc. Var., 17 (2011), pp. 267–292. doi:10.1051/cocv/2010004.

[6] A. Borum and T. Bretl, *The free configuration space of a Kirchhoff elastic rod is path-connected*, in Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, 2015, pp. 2958–2964. doi:10.1109/ICRA.2015.7139604.

[7] T. Bretl and Z. McCarthy, *Quasi-static manipulation of a Kirchhoff elastic rod based on a geometric analysis of equilibrium configurations*, Int. J. Robot. Res., 33 (2014), pp. 48–68. doi:10.1177/0278364912473169.

[8] Y. A. Butt, Yu. L. Sachkov, and A. I. Bhatti, *Extremal trajectories and Maxwell strata in sub-Riemannian problem on group of motions of pseudo-Euclidean plane*, J. Dyn. Control Syst., 20 (2014), pp. 341–364. doi:10.1007/s10883-014-9239-2.

[9] J.-B. Caillau, O. Cots, and J. Gergaud, *Differential continuation for regular optimal control problems*, Optim. Methods Softw., 27 (2012), pp. 177–196. doi:10.1080/10556788.2011.593625.

[10] M. P. do Carmo, *Riemannian Geometry*, Translated from the second Portuguese edition by F. Flaherty, Math. Theory Appl., Birkhäuser Boston, Inc., Boston, MA, 1992.

[11] M. E. Henderson and S. Neukirch, *Classification of the spatial equilibria of the clamped elastica: Numerical continuation of the solution set*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 14 (2004), pp. 1223–1239. doi:10.1142/S0218127404009971.

[12] J. E. Hopcroft, J. K. Kearney, and D. B. Krafft, *A case study of flexible object manipulation*, Int. J. Robot. Res., 10 (1991), pp. 41–50. doi:10.1177/027836499101000105.

[13] T. A. Ivey and D. A. Singer, *Knot types, homotopies and stability of closed elastic rods*, Proc. London Math. Soc., 79 (1999), pp. 429–450. doi:10.1112/S0024611599011983.

[14] F. Lamiraux and L. E. Kavraki, *Planning paths for elastic objects under manipulation constraints*, Int. J. Robot. Res., 20 (2001), pp. 188–208. doi:10.1177/02783640122067354.

[15] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*, Texts Appl. Math. 17, Springer, New York, 1999. doi:10.1007/978-0-387-21792-5.

[16] D. Matthews and T. Bretl, *Experiments in quasi-static manipulation of a planar elastic rod*, in Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, Portugal, 2012, pp. 5420–5427. doi:10.1109/IROS.2012.6385876.

[17] M. Moll and L. E. Kavraki, *Path planning for deformable linear objects*, IEEE Trans. Robot., 22 (2006), pp. 625–636. doi:10.1109/TRO.2006.878933.

[18] F. Monroy-Pérez and A. Anzaldo-Meneses, *Optimal control on the Heisenberg group*, J. Dyn. Control Syst., 5 (1999), pp. 473–499. doi:10.1023/A:1021787121457.

[19] S. Neukirch and M. E. Henderson, *Classification of the spatial equilibria of the clamped elastica: Symmetries and zoology of solutions*, J. Elasticity, 68 (2002), pp. 95–121. doi:10.1023/A:1026064603932.

[20] P. J. Olver, *Applications of Lie Groups to Differential Equations*, Grad. Texts Math. 107, Springer-Verlag, New York, 1986. doi:10.1007/978-1-4684-0274-2.

[21] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, John Wiley, New York, 1962.
[22] Yu. L. Sachkov, *Maxwell strata in the Euler elastic problem*, J. Dyn. Control Syst., 14 (2008), pp. 169–234. doi:10.1007/s10883-008-9039-7 .
[23] Yu. L. Sachkov, *Conjugate points in the Euler elastic problem*, J. Dyn. Control Syst., 14 (2008), pp. 409–439. doi:10.1007/s10883-008-9044-x.
[24] Yu. L. Sachkov, *Cut locus and optimal synthesis in the sub-Riemannian problem on the group of motions of a plane*, ESAIM Control Optim. Calc. Var., 17 (2011), pp. 293–321. doi:10.1051/cocv/2010005.
[25] Yu. L. Sachkov and E. F. Sachkova, *Exponential mapping in Euler's elastic problem*, J. Dyn. Control Syst., 20 (2014), pp. 443–464. doi:10.1007/s10883-014-9211-1.
[26] M. Saha and P. Isto, *Manipulation planning for deformable linear objects*, IEEE Trans. Robot., 23 (2007), pp. 1141–1150. doi:10.1109/TRO.2007.907486.
[27] S. Smale, *A convergent process of price adjustment and global Newton methods*, J. Math. Econom., 3 (1976), pp. 107–120. doi:10.1016/0304-4068(76)90019-7.
[28] H. Stephani, *Differential Equations: Their Solution Using Symmetries*, Cambridge University Press, Cambridge, UK, 1990. doi:10.1017/CBO9780511599941.
[29] E. Trélat, *Optimal control and applications to aerospace: Some results and challenges*, J. Optim. Theory Appl., 154 (2012), pp. 713–758. doi:10.1007/s10957-012-0050-5.