CONTROL AND OPTIMIZATION OF HYBRID SYSTEMS ON RIEMANNIAN MANIFOLDS

Farzin Taringoo

Department of Electrical and Computer Engineering
McGill University, Montréal

9 October 2012

A thesis submitted to McGill University
in partial fulfilment of the requirements of the degree of
Doctor of Philosophy

ABSTRACT

The fundamental motivation for the work in this thesis is the analysis of the optimal control of hybrid systems on Riemannian manifolds using the language of differential geometry. Hybrid systems theory constitutes one of the major frameworks within which one may model and analyze the behaviour of large and complex systems; in particular, the optimal control of hybrid systems has been a focus of research over the last decades resulting in the important generalization of Minimum (Maximum) Principle of classic optimal control to hybrid systems.

In the work of Shaikh and Caines (2007) and their predecessors, a formulation for a class of optimal control problems for general hybrid systems with nonlinear dynamics and autonomous or controlled switchings at switching states and times is proposed. In this thesis we extend the framework of Shaikh and Caines (2007) to a general class of hybrid systems defined on Riemannian manifolds. Due to the formulation generality, this class of hybrid systems covers a vast range of practical examples arising in such different areas as mechanical systems, chemical processes, air traffic control systems and cooperative robotic manipulator systems.

In this thesis, a formulation for general hybrid systems on differentiable Riemannian manifolds is first presented. In the case of autonomous switchings, switching manifolds are modelled by embedded orientable submanifolds of the ambient state manifold and consequently hybrid optimal control problems are defined for hybrid systems in this general setting.

Second, the classic Minimum Principle is extended to the Hybrid Minimum Principle (HMP) yielding the optimality necessary conditions for hybrid systems at the optimal switching states and times. The HMP statement in this thesis is obtained by employing the so-called needle control variation in the control value space. This class of control variations results in state trajectory variations along the nominal state trajectory in the ambient state manifold where the optimality conditions are derived by analyzing the cost function variation with respect to state variations.

Third, in order to optimize switching states and times, numerical optimization algorithms (Gradient Geodesic-HMP, Newton Geodesic-HMP) are formulated by employing the HMP equations on general Riemannian state manifolds. The convergence analysis for the proposed algorithms is based upon the LaSalle Invariance Theorem. Technically these algorithms generalize the standard steepest descent and Newton methods in Euclidean spaces to Reimannian manifolds by employing the notion of Levi-Civita connections.

Fourth, the derivation of the HMP results for hybrid systems on Riemannian manifolds is carried out for hybrid systems on Lie groups. The group structure of the ambient state manifold gives rise to a special form for the adjoint processes and Hamiltonian functions as the solutions for the optimality equations. In this thesis hybrid optimal control problems on Lie groups are only considered for the class of left invariant systems, however, the analysis can be easily modified to right invariant systems. In the setting of left invariant hybrid systems on Lie groups, the Gradient Geodesic-HMP and Newton Geodesic-HMP algorithm are modified into algorithms called the Gradient Exponential-HMP and Newton Exponential-HMP algorithms.

The fifth and last part of the thesis focuses on the problem of optimization of autonomous hybrid optimal control problems with respect to the geometrical features of switching manifolds. Such features include first order and second order information on the switching manifolds such as curvature tensors and normal differential forms.

RÉSUMÉ

La motivation première du travail accompli dans cette thèse est l'analyse du contrôle optimal de systèmes hybrides sur les variétés riemanniennes en utilisant le language de la géométrie differentielle. La théorie des systèmes hybrides constitue un des cadres majeurs dans lequel on peut modeler et analyser le comportement de systèmes grands et complexes; en particulier, le contrôle optimal de systèmes hybrides a été le centre d'intérêt des recherches dans les décennies précédentes ayant comme résultat une importante généralisation du Principe Minimum (Maximum) du contrôle optimal classique aux systèmes hybrides.

Le travail de Shaikh et Caines (2007) et leurs prédécesseurs propose une formule pour une classe de problèmes de contrôle optimal pour les systèmes hybrides généraux avec des dynamiques non linéaires et autonomes ou des commutations contrôlées aux états et temps de commutation. Cette thèse élargit le cadre de Shaikh et Caines (2007) à une classe générale de systèmes hybrides définis sur les variétés riemanniennes. En raison de la nature générale de la formulation, cette classe de systèmes hybrides couvre un vaste éventail d'exemples pratiques survenant dans différents domaines tels que les systèmes mécaniques, les procédés chimiques, le contrôle des systèmes de navigation aérienne, ainsi que les systèmes de manipulation de la robotique coopérative.

Premièrement, cette thèse présente une formulation pour le cas des systèmes hybrides généraux sur les variétés riemaniennes différentielles. Dans le cas des commutations autonomes, les variétés de commutation sont modélisées par les sous-variétés prolongées et orientables de la variété d'état ambiante et conséquemment, les problèmes de contrôle optimal hybrides sont définis pour les systèmes hybrides dans ce contexte général.

Deuxièmement, le Principe Minimum classique est étendu au Principe Minimum Hybride (HMP), produisant les conditions nécessaires d'optimalité pour les systèmes hybrides aux états et temps optimaux de commutation. L'énoncé du Principe Minimum Hybride (HMP) dans cette thèse est obtenu en utilisant la commande de variation d'aiguille, ainsi nommée, dans l'espace de valeur de contrôle. Cette classe de variation donne des variations de trajectoire au long de la trajectoire d'état nominale dans la variété d'état ambiante. Les conditions d'optimalité sont obtenues en analysant la variation de la fonction de coût en respectant les variations d'état.

Troisièmement, dans le but d'optimiser les états et temps de commutation, des algorithmes numériques d'optimisation (Géodésique Gradient-HMP, Géodésique Newton-HMP) sont formulés en utilisant les équations du Principe Minimum Hybride (HMP) sur les variétés d'état riemanniennes générales. L'analyse de convergence pour les algorithmes proposés est basée sur le théorème d'invariance de LaSalle. Techniquement, ces algorithmes généralisent l'algorithme standard de la plus profonde descente et les méthodes de Newton dans les espaces euclidien aux variétés riemanniennes en utilisant la notion des connexions Levi-Civita.

Quatrièmement, la dérivation des résultats du Principe Minimum Hybride (HMP) sur les variétés riemanniennes est appliqué aux systèmes hybrides sur les groupes de Lie. La structure du groupe de la variété d'état ambiante engendre une forme spéciale des processus adjoints et des fonctions hamiltoniennes comme solutions pour les équations d'optimalité. Dans cette thèse, les problèmes de contrôle optimal sur les groupes de Lie sont seulement considérés pour la classe des systèmes invariants de gauche. Par contre, l'analyse peut facilement être modifiée pour les systèmes invariants de droite. Dans ce contexte, les algorithmes du Géodésique Gradient-HMP et du Géodésique Newton-HMP sont développés aux algorithmes du Gradient Exponentiel-HMP et à l'algorithme Newton Exponentiel-HMP pour les systèmes hybrides invariants de gauche sur les groupes de Lie.

Finalement, la dernière partie de la thèse met l'accent sur la question d'optimisation des problèmes de contrôle optimal autonomes hybrides en ce qui concerne les caractéristiques géométriques des variétés de commutation. De telles caractéristiques comprennent des informations de premier et de second ordre sur les variétés de commutation telles que les tenseurs de courbures et les formes différentielles normales.

Claims of Originality and Published Work

Claims of Originality

The following original contributions are presented in this thesis:

- An extension of the Hybrid Minimum Principle (HMP) derived for hybrid systems defined on Euclidean spaces to that of impulsive hybrid systems defined on Riemannian manifolds.
- The design and development of numerical algorithms for the optimization of switching states and times for a general class of hybrid systems defined on Riemanian manifolds and Lie groups.
- The convergence analyses of the proposed numerical algorithms based upon the optimality equations and the geometry of ambient state manifolds.
- The derivation and extension of the HMP for left invariant hybrid systems defined on Lie groups and modifying the numerical optimization algorithms in order to take into account the group properties of ambient state manifolds.
- The formulation and derivation of necessary conditions for optimality of hybrid cost functions with respect to the geometry of switching manifolds in the case of autonomous hybrid systems.

Publications

Publications in Journals and Conferences:

- *Papers published or to appear*

[1.] F. Taringoo and P. E. Caines, “The Exponential Gradient HMP Algorithm for the Optimization of Hybrid Systems On Lie Groups.” *4th IFAC Conference on Analysis and Design of Hybrid Systems,* Eindhoven, The Netherlands, pages:33-38, June 2012.

[2.] F. Taringoo and P. E. Caines, “On the Extension of the Hybrid Minimum Principle to Hybrid Systems On Lie Groups.” *to appear in Proc. American Control Conference,* Montreal, Canada, June 2012, https://css.paperplaza.net/conferences/conferences/2012ACC/program.

[3.] F. Taringoo and P. E. Caines, “Gradient Geodesic and Newton Geodesic HMP Algorithms for the Optimization of Hybrid Systems.” *IFAC Annual Reviews in Control*, vol. 35, pages:187-198, 2011.

[4.] F. Taringoo and P. E. Caines, “On the Extension of the Hybrid Maximum Principle to Riemannian Manifolds.” *50th IEEE Conference on Decision and Control*, Orlando, USA, pages: 3301-3306, December 2011.

[5.] F. Taringoo and P. E. Caines, “Gradient-Geodesic HMP Algorithms for the Optimization of Hybrid Systems based on the Geometry of Switching Manifolds.” *49th IEEE Conference on Decision and Control*, Georgia, USA, pages: 1534-1539, December, 2010.

[6.] F. Taringoo and P. E. Caines, “On the Geometry of Switching Manifolds for Autonomous Hybrid Systems.” *10th International Workshop on Discrete Event Systems*, Berlin, Germany, pages: 45-50, August, 2010.

[7.] F. Taringoo and P. E. Caines, “Geometry and Deformation of Switching Manifolds for Autonomous Hybrid Systems.” *19th International Symposium on Mathematical Theory of Networks and Systems*, Budapest, Hungary, pages: 761-766, July, 2010.

[8.] F. Taringoo and P. E. Caines, "The Sensitivity of Hybrid Systems Optimal Cost Functions with Respect to Switching Manifold Parameters." *Hybrid Systems: Computation and Control, Springer Verlag*, San Fransisco, Eds: R. Majumdar and P. Tabuada pages: 475-479, April, 2009.

[9.] F. Taringoo and P. E. Caines, "Geometrical Properties of Optimal Hybrid System Trajectories and the Optimization of Switching Manifolds." *3rd IFAC Conference on Analysis and Design of Hybrid Systems*, Zaragoza, Spain, September, 2009.

- *Papers submitted for publication*

[1.] F. Taringoo and P. E. Caines, "On the Optimal Control of Impulsive Hybrid Systems on Riemannian Manifolds." *Submitted to SIAM Journal on Control and Optimization.*

[2.] F. Taringoo and P. E. Caines, "Newton-Geodesic HMP Algorithms for the Optimization of Hybrid Systems and the Geometric Properties of Hybrid Value Functions." *Submitted to 51st IEEE Conference on Decision and Control.*

- *Papers in preparation*

[1.] F. Taringoo and P. E. Caines, "On the Extension of Hybrid Minimum Principle to Lie Groups and the Exponential Gradient HMP Algorithm." *to be submitted to IEEE transaction on Automatic Control.*

[2.] F. Taringoo and P. E. Caines, "On the Geometry and Deformation of Switching Manifolds for Autonomous Hybrid Systems." *to be submitted to IEEE Journal of Selected Topics in Signal Processing: Special Issue on Differential Geometry in Signal Processing.*

Contribution of Co-authors

For all the papers cited above, Peter E. Caines' contributions consisted of work on the problem formulations and their analyses. These contributions amounted to 25% of those papers.

Acknowledgements

First, I want to express my appreciation and respect to my Ph.D. supervisor, Professor Peter E. Caines, especially in directing me towards a topic which proved to be stimulating and promising. Throughout my study and research in McGill University, I have benefited a lot from his keen scientific foresight, erudite mathematical guidance, painstaking academic attitude, and solid financial support. I also appreciate the opportunities he has provided to attend conferences and workshops and to interact with leading researchers around the world.

Second, many thanks go to my friends and colleagues for their help and support. Specially I would like to thank Mojtaba Nourian, Sayed Rahi Modirnia, Peng Jia, Arman Kizilkale, and Dmitry Gromov for listening to and providing suggestions for all my research problems along the way.

Thanks also go to Benoit Boulet, Cynthia Davidson, Jan Binder, and Patrick McLean for all their consistent support and professional maintenance of the CIM research environment. I am grateful for the financial support throughout my studies provided by National Science and Engineering Research Council.

Finally, I am extremely thankful to my father, Hossein Taringoo, my mother, Narges Habibi, my sister Farzaneh Taringoo and my girlfriend Sibel Abaci for all their love, support, and encouragement through the years, which have constituted a crucial part of my life. I present this thesis as a gift to them.

Farzin Taringoo

McGill University, Montréal, Quebéc

October 2012

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

Introduction

The key notion of hybrid systems is the continuous and discrete nature of their state space and dynamics. This fundamental concept has been characterized and crystallized over the last few years (see [**15**, **16**, **51**, **58**, **66**, **83**, **84**]).

Examples of hybrid systems can be found in a vast area of engineering and industrial applications including telecommunication and transportation networks, mechanical systems, chemical processes and water systems (see [**36**, **47**, **54**]). One important problem arising in the context of hybrid systems is optimization and optimal control of hybrid systems. This problem is addressed in [**7**–**9**, **22**, **26**, **27**, **31**–**34**, **44**, **45**, **55**, **56**, **58**, **66**, **68**, **69**, **84**–**86**, **88**]. Among these, and [**58**, **66**] and [**69**] present versions of the Hybrid Minimum Principle (HMP).

Technically the HMP results are extensions of Pontryagin's Minimum Principle (often called Pontryagin's Maximum Principle) for hybrid systems. In particular, Shaikh and Caines proposed a derivation for the Hybrid Minimum Principle based upon the control needle variation for autonomous and controlled hybrid systems (see [**66**]). Their methods generalize the approaches of [**25**,**87**] where optimality conditions are obtained by analyzing the cost variation propagation along the optimal trajectory. Employing the HMP results, they also proposed a class of numerical algorithms (HMP Algorithms) for the optimization of switching states and switching times for both autonomous and controlled hybrid systems. The efficacy of their HMP algorithm has been illustrated via numerical comparisons with a gradient algorithm proposed by Xu and Antsaklis in [**84**].

In the first chapter of this thesis, following [**66**], we generalize the formulation of hybrid systems to those on Riemannian manifolds and then we obtain the Hybrid Minimum Principle for a class of impulsive hybrid systems where the state trajectories are discontinuous at switching instants. Our development is based on a geometric version of Pontryagin's Minimum Principle for a general class of state manifolds which is given in [**3**, **6**]. It is shown that under appropriate hypotheses on the differentiability of the hybrid value function, the discontinuity of the adjoint variable at the optimal switching state and switching time is proportional to a differential form of the hybrid value function defined on the cotangent bundle of the state manifold. In the case of open control sets and Euclidean state spaces, these results appeared in [**57**] without using the language of differential geometry. The results obtained are also extended to the cases of time varying switching manifolds, time varying impulse jumps and multiple switching hybrid systems. The discontinuity equations of the adjoint variables are also derived in the presence of discrete switching costs at switching states.

Chapter 3 presents numerical algorithms for the optimization of switching states and switching times for autonomous hybrid systems. However, the analysis can be carried out for controlled hybrid systems as well. The central core of the proposed algorithms relies on the extension of the well-known Gradient Descent and Newton algorithms for cost functions defined on Riemannian manifolds. In a natural way, the notion of straight lines in Euclidean spaces are generalized to geodesic lines on Riemannian manifolds, see [**37**, **42**]. In [**37**], it is shown that geodesics are locally length minimizers where the neighbourhoods for which geodesics are length minimizers are determined by the geometric structure of Riemannian manifolds (see [**42**] Chapter 10). In general, for a Riemannian manifold $(\mathcal{M}, g_{\mathcal{M}})$, geodesics are locally computed as solutions of a second order differential equation involving the Christoffel symbols (see [**37**, **42**]). We introduce the so-called Gradient Geodesic HMP algorithm (GG-HMP) which is an extension of the HMP algorithm in [**66**] where switching states are updated along the geodesics in switching manifolds. The convergence analysis is performed using the Lasalle Invariance Principle (see [**60**]).

In order to further improve the convergence rate of the GG-HMP, the so-called Newton Geodesic version of the GG-HMP is formulated in the local coordinate system of the switching state. In the general case of Reimannian manifolds, it is always not possible to define a Hessian matrix as can be done in Euclidean spaces, (see e.g. [**29**], [**67**]). However, employing the notion of *Levi-Civita* connection ∇ on Riemannian manifolds, the Hessian may be defined as a bilinear symmetric form, [**37**]. Again the Lasalle Invariance Principle provides a proof of convergence for the Newton Geodesic method.

In Chapter 4 we further extend the results of Chapter 2 to hybrid systems defined on Lie groups. By definition, Lie groups are differentiable manifolds associated with group structures which are multiplication and inversion (see e.g. [**19**,**38**]). Examples of dynamical systems on Lie groups can be found in many mechanical systems where state manifolds constitute group properties. Such examples include rigid body motion and rotational systems (see [**10**, **18**, **19**]). Optimal control of dynamical systems on Lie groups is presented in [**38**], Chapter 12, for general left and right invariant control systems. The theory of Hamiltonian systems on Lie groups is based upon a special realization of the cotangent bundle of Lie groups. This realization enables us to obtain the Hamiltonian equations which correspond to the optimality of control systems, in a special form of differential equations on Lie algebras of state Lie groups. For a hybrid system defined on a Lie group, it is shown that the difference of the adjoint variables gives a projection of the hybrid value function differential form on the Lie algebra of the state manifold.

The numerical algorithms in Chapter 3 are then modified to Gradient Exponential HMP algorithms to optimize switching states and times for hybrid systems defined on Lie groups. The group properties of state manifolds generate different possible directions for updating equations in descent algorithms performing on Lie groups. One of these possibilities is given by the exponential map on Lie groups.

Note that depending upon the metric associated to a Lie group, the exponential flow may coincide with the geodesic flow on the group, see e.g. [**49**].

Chapter 5 treats the problem of optimization of autonomous hybrid systems with respect to the geometry of switching manifolds. The problem presented there considers the variation of the switching manifold configurations which determine the autonomous (uncontrolled) discrete state switchings. The optimal cost variation (i.e. derivative) as a function of the switching manifold parameters is described by the solution of a set of differential equations generating the state and costate sensitivity functions. This problem is addressed in [**11**, **12**, **61**, **62**] from different points of views. In [**62**] the maximization of a measure of the stability of systems with periodic behaviour is analyzed in terms of the adjustment of the switching surfaces. In [**11**, **12**] a method for deriving the cost variation induced by shifting the switching manifold position is proposed. One direct extension of the optimal control problem for autonomous hybrid systems concerns the notion of switching manifold geometry, where this is interpreted as the shaping and displacement of switching manifolds in order to optimize system performance. In this chapter, we analyze HSOC sensitivity with respect to the parameters determining systems switching manifolds. Similar to [**11**, **12**], we employ a method for deriving the optimal cost variation as a function of the switching manifold parameters, but we consider a general class of hybrid systems for which there are continuous controls in each distinct discrete state.

Finally, in Chapter 6, suggestions for possible future research and extensions are presented.

CHAPTER 2

Optimal Control of Hybrid Systems On Riemannian Manifolds

In the following definition the standard hybrid systems framework (see e.g. [**15**, **66**, **75**]) is generalized to the case where the continuous state space is a smooth manifold, where henceforth in this paper smooth means C^∞.

Definition 2.1. *A hybrid system with autonomous discrete transitions is a five-tuple*

$$\mathbf{H} := \{H = Q \times \mathcal{M}, U, F, \mathcal{S}, \mathcal{J}\} \tag{2.1}$$

where:
$Q = \{1, 2, 3, ..., |Q|\}$ *is a finite set of discrete (valued) states (components) and* $\mathcal{M}$ *is a smooth* n *dimensional Riemannian continuous (valued) state (component) manifold with associated metric* $g_{\mathcal{M}}$.
H *is called the hybrid state space of* $\mathbf{H}$.
$U \subset \mathbb{R}^u$ *is a set of admissible input control values, where* U *is a compact set in* $\mathbb{R}^u$. *The set of admissible input control functions is* $\mathcal{I} := (L_\infty[t_0, t_f), U)$*, the set of all bounded measurable functions on some interval* $[t_0, t_f), t_f < \infty$*, taking values in* U.
F *is an indexed collection of smooth, i.e.* C^∞*, vector fields* $\{f_{q_i}\}_{q_i \in Q}$*, where* f_{q_i} :

$\mathcal{M} \times U \to T\mathcal{M}$ is a controlled vector field assigned to each discrete state; hence each f_{q_i} is continuous on $\mathcal{M} \times U$ and continuously differentiable on $\mathcal{M}$ for all $u \in U$.
$\mathcal{S} := \{n_\gamma^k : \gamma \in Q \times Q, 1 \leq k \leq K < \infty, n_\gamma^k \subset \mathcal{M}\}$ is a collection of embedded time independent pairwise disjoint switching manifolds (except in the case where $\gamma = (p,q)$ is identified with $\gamma^{'} = (q,p)$) such that for any ordered pair $\gamma = (p,q), n_\gamma^k$ is an open smooth, oriented codimension 1 submanifold of $\mathcal{M}$, possibly with boundary ∂n_γ^k. By abuse of notation, we describe the manifolds locally by $n_\gamma^k = \{x : n_\gamma^k(x) = 0, x \in \mathbb{R}^n\}$.
$\mathcal{J}$ shall denote the family of the state jump functions on the manifold $\mathcal{M}$. For an autonomous switching event from $p \in Q$ to $q \in Q$, the corresponding jump function is given by a smooth map $\zeta_{p,q} : \mathcal{M} \to \mathcal{M}$: if $x(t^-) \in \mathcal{S}$ the state trajectory jumps to $x(t) = \zeta_{p,q}(x(t^-)) \in \mathcal{M}$, $\zeta_{p,q} \in \mathcal{J}$. The non-jump special case is given by $x(t) = x(t^-)$. We use the term impulsive hybrid systems for those hybrid systems where the continuous part of the state trajectory may have discontinuous transitions (i.e. jump) at controlled or autonomous discrete state switching times.

We assume:
A1: The *initial state* $h_0 := (x(t_0), q_0) \in H$ is such that $x_0 = x(t_0) \notin \mathcal{S}$ for all $q_i \in Q$. A *(hybrid) input function* u is defined on a half open interval $[t_0, t_f), t_f \leq \infty$, where further $u \in \mathcal{I}$. A *(hybrid) state trajectory* with initial state h_0 and (hybrid) input function u is a triple (τ, q, x) consisting of a finite strictly increasing sequence of times (boundary and switching times) $\tau = (t_0, t_1, t_2, \dots)$, an associated sequence of discrete states $q = (q_0, q_1, q_2, \dots)$, and a sequence $x(\cdot) = (x_{q_0}(\cdot), x_{q_1}(\cdot), x_{q_2}(\cdot), \dots)$ of absolutely continuous functions $x_{q_i} : [t_i, t_{i+1}) \to \mathcal{M}$ satisfying the continuous and discrete dynamics given by the following definition.

Definition 2.2. *The continuous dynamics of a hybrid system* **H** *with initial condition $h_0 = (x_0, q_0)$, input control function $u \in \mathcal{I}$ and hybrid state trajectory (τ, q, x) are specified piecewise in time via the mappings*

$$(x_{q_i}, u) : [t_i, t_{i+1}) \to \mathcal{M} \times U, \quad i = 0, ..., L, \quad 0 < L < \infty, \tag{2.2}$$

where $x_{q_i}(.)$ *is an integral curve of* $f_{q_i}(.,u(.)) : \mathcal{M} \times [t_i, t_{i+1}) \to T\mathcal{M}$ *satisfying*

$$\dot{x}_{q_i}(t) = f_{q_i}(x_{q_i}(t), u(t)), \quad a.e.\, t \in [t_i, t_{i+1}),$$

where $x_{q_{i+1}}(t_{i+1})$ *is given recursively by*

$$x_{q_{i+1}}(t_{i+1}) = lim_{t \uparrow t_{i+1}^-} \zeta_{q_i,q_{i+1}}(x_{q_i}(t)), \quad h_0 = (q_0, x_0), t < t_f. \tag{2.3}$$

The discrete autonomous switching dynamics are defined as follows:
For all p, q*, whenever an admissible hybrid system trajectory governed by the controlled vector field* f_p *meets any given switching manifold* $n_{p,q}$ *transversally, i.e.* $f_p(x(t_s^-), t_s^-) \notin T_{x(t_s^-)}\mathcal{S}$*, there is an autonomous switching to the controlled vector field* f_q*, equivalently, discrete state transition* $p \to q$*. Conversely, any autonomous discrete state transition corresponds to a transversal intersection.*

A system trajectory is not continued after a non-transversal intersection with a switching manifold. Given the definitions and assumptions above, standard arguments give the existence and uniqueness of a hybrid state trajectory (τ, q, x)*, with initial state* $h_0 \in H$ *and input function* $u \in \mathcal{I}$*, up to* T*, defined to be the least of an explosion time or an instant of non-transversal intersection with a switching manifold.*

We adopt:

A2: (Controllability) For any $q \in Q$, all pairs of states (x_1, x_2) are mutually accessible in any given time period $[0,t], 0 < t < t_f$, via the controlled vector field $\dot{x}_q(t) = f_q(x_q(t), u(t))$, for some $u \in \mathcal{I} = (L_\infty[0,t_f), U)$.

A3: $\{l_{q_i}\}_{q_i \in Q}$, is a family of *loss functions* such that $l_{q_i} \in C^k(\mathcal{M} \times U; \mathbb{R}^+), k \geq 1$, and h is a *terminal cost function* such that $h \in C^k(\mathcal{M}; \mathbb{R}^+), k \geq 1$.

Henceforth, Hypotheses ***A1-A3*** will be in force unless otherwise stated. Let L be the number of switchings and $u \in \mathcal{I}$ then we define the *hybrid cost function* as

$$J(t_0, t_f, h_0; L, u) := \sum_{i=0}^{L} \int_{t_i}^{t_{i+1}} l_{q_i}(x_{q_i}(s), u(s))ds + h(x_{q_L}(t_f)),$$
$$t_{L+1} = t_f < T, u \in \mathcal{I}, \tag{2.4}$$

where we observe the conditions above yield $J(t_0, t_f, h_0; L, u) < \infty$.

Definition 2.3. *For a hybrid system* $\mathbf{H}$*, given the data* $(t_0, t_f, h_0; L)$*, the Bolza Hybrid Optimal Control Problem (BHOCP) is defined as the infimization of the hybrid cost function* $J(t_0, t_f, h_0; L, u)$ *over the hybrid input functions* $u \in \mathcal{I}$*, i.e.*

$$J^o(t_0, t_f, h_0; L) = inf_{u \in \mathcal{I}} J(t_0, t_f, h_0; L, u).$$

Definition 2.4. *A Mayer Hybrid Optimal Control Problem (MHOCP) is defined as the special case of the BHOCP where the cost function given in (2.4) is evaluated only on the terminal state of the system, i.e.* $l_{q_i} = 0, \ \ i = 1, ..., L$.

In general, different control inputs result in different sequences of discrete states of different cardinality. However, in this chapter, we shall restrict the infimization to be over the class of control functions, generically denoted $\mathcal{U} \subset \mathcal{I}$, which generates an a priori given sequence of discrete transition events.

We adopt the following standard notation and terminology, see [**19**, **75**, **79**]. The time dependent flow associated to a differentiable time independent vector field f_{q_i} is a map $\Phi_{f^u_{q_i}}$ satisfying ($f^u_{q_i}(.)$ is used here for brevity instead of $f_{q_i}(., u(t))$ since the calculations are performed with respect to a given control u):

$$\Phi_{f^u_{q_i}} : [t_i, t_{i+1}) \times [t_i, t_{i+1}) \times \mathcal{M} \to \mathcal{M}, \quad (t, s, x) \to \Phi^{(t,s)}_{f^u_{q_i}}(x) := \Phi_{f^u_{q_i}}((t, s), x) \in \mathcal{M}, \tag{2.5}$$

where

$$\Phi^{(t,s)}_{f^u_{q_i}} : \mathcal{M} \to \mathcal{M}, \quad \Phi^{(s,s)}_{f^u_{q_i}}(x) = x, \tag{2.6}$$

$$\frac{d}{dt}\Phi^{(t,s)}_{f^u_{q_i}}(x)|_t = f_{q_i}\big(\Phi^{(t,s)}_{f^u_{q_i}}(x(s))\big), \;\; t,s \in [t_i, t_{i+1}). \tag{2.7}$$

We associate $T\Phi^{(t,s)}_{f^u_{q_i}}(.)$ to $\Phi^{(t,s)}_{f^u_{q_i}} : \mathcal{M} \to \mathcal{M}$ via the push-forward of $\Phi^{(t,s)}_{f^u_{q_i}}$.

$$T\Phi^{(t,s)}_{f^u_{q_i}} : T_x\mathcal{M} \to T_{\Phi^{(t,s)}_{f^u_{q_i}}(x)}\mathcal{M}. \tag{2.8}$$

Following [**19**], the corresponding *tangent lift* of $f^u_{q_i}(.)$ is the time dependent vector field $f^{T,u}_{q_i}(.) \in TT\mathcal{M}$ on $T\mathcal{M}$

$$f^{T,u}_{q_i}(v_x) := \frac{d}{dt}|_{t=s} T\Phi^{(t,s)}_{f^u_{q_i}}(v_x), \quad v_x \in T_x\mathcal{M}, \tag{2.9}$$

which is given locally as

$$f^{T,u}_{q_i}(x, v_x) = \left[f^{u,i}_{q_i}(x)\frac{\partial}{\partial x^i} + (\frac{\partial f^{u,i}_{q_i}}{\partial x^j} v^j)\frac{\partial}{\partial v^i} \right]^n_{i,j=1}, \tag{2.10}$$

and $T\Phi^{(t,s)}_{f^u_{q_i}}(.)$ is evaluated on $v_x \in T_x\mathcal{M}$, see [**19**]. The following lemma gives the relation between the push-forward of $\Phi^{(t,s)}_{f_{q_i}}$ and the tangent lift introduced in (2.10). For simplicity and uniformity of notation, we use f_{q_i} instead of $f^u_{q_i}$.

Lemma 2.1 ([**6**]). *Consider $f_{q_i}(., u(.)) : \mathcal{M} \times I \to T\mathcal{M}, I = [t_i, t_{i+1})$ as a time dependent vector field on $\mathcal{M}$ and $\Phi^{(t,s)}_{f_{q_i}}$ as its corresponding flow. The flow of $f^{T,u}_{q_i}$, denoted by $\Psi : I \times I \times T\mathcal{M} \to T\mathcal{M}$, satisfies:*

$$\Psi(t, s, (x, v)) = (\Phi^{(t,s)}_{f_{q_i}}(x), T\Phi^{(t,s)}_{f_{q_i}}(v)) \in T\mathcal{M}, \quad (x, v) \in T\mathcal{M}. \tag{2.11}$$

□

2.1. The Pontryagin Minimum Principle for Standard Optimal Control Problems

In this section we focus on the *Pontryagin Minimum Principle* (PMP) for standard (non-hybrid) optimal control problems defined on a Riemannian manifold $\mathcal{M}$. A standard optimal control problem (OCP) can be obtained from a BHOCP, see (2.4), by fixing the discrete states q_i to q, and hence L to the value 0. The resulting optimal control problem in Bolza form becomes that of the infimization of the cost (2.4) with respect to state dynamics which by suppressing notation of q may be written $\dot{x} = f(x(t), u(t)),\ x(t) \in \mathcal{M}, u(t) \in \mathcal{U}, t \in [t_0, t_f]$.

2.1.1. The Relationship between Bolza and Mayer Problems. In Section 2 both the BHOCP and the MHOCP were introduced; since the results in this chapter are only stated for the Mayer problem we now briefly explain the relationship between them.

In general (see [**6**]), a Bolza problem can be converted to a Mayer problem with state variable $\hat{x} := (x, x_{n+1})$ by adjoining an auxiliary state x_{n+1} to the state x, one then defines the dynamics to be given by

$$\dot{\hat{x}}(t) = \begin{bmatrix} \dot{x}(t) \\ \dot{x}_{n+1}(t) \end{bmatrix} = \begin{bmatrix} f(x(t), u(t)) \\ l(x(t), u(t)) \end{bmatrix}, \tag{2.12}$$

where f and l are respectively the dynamics and the running cost of the Bolza problem. Then the equivalent Mayer problem is obtained by the infimization of the penalty function $\hat{h}(.)$ defined as follows:

$$\hat{h}(\hat{x}(t_f)) \equiv \hat{h}(x(t_f), x_{n+1}(t_f)) := x_{n+1}(t_f) + h(x(t_f)) = J(t_0, t_f, x_0, u), \tag{2.13}$$

where h is the terminal cost function of the Bolza problem. Note that after such a transformation from a Bolza problem the state space of the resulting Mayer problem is $\mathcal{M} \times R$, where $\mathcal{M}$ is the state manifold of the Bolza problem.

2.1.2. Elementary Control and Tangent Perturbations. We now present some results from [**3**], [**6**] and [**41**]. It is essential to note that henceforth in this paper we treat the general Mayer problem with state space manifold denoted by $\mathcal{M}$. In the special case where Mayer OCP is derived from a Bolza problem, $\mathcal{M}$ takes the product form given in the previous section.

Consider the nominal control input $u(.)$ and define the associated perturbed control as

$$u_{\pi(t^1,u_1)}(t,\epsilon) = \begin{cases} u_1 & t^1-\epsilon \le t \le t^1, \\ u(t) & \text{elsewhere,} \end{cases} \tag{2.14}$$

where $0 \le \epsilon < \infty, u_1 \in U$. For brevity in notation $u_{\pi(t^1,u_1)}(t,\epsilon)$ shall be written $u_\pi(t,\epsilon)$.

Associated to $u_\pi(.,.)$ we have the corresponding state trajectory $x_\pi(.,.)$ on $\mathcal{M}$. It may be shown under suitable hypotheses, $lim_{\epsilon\to 0} x_\pi(t,\epsilon) = x(t)$ uniformly for $t_0 \le t \le t_f$, see [**30**] and [**41**]. Following (2.6), the flow resulting from the perturbed control is defined as:

$$\Phi^{(t,s),x}_{\pi,f}(.) : [0,\tau] \to \mathcal{M}, \quad x \in \mathcal{M}, t,s \in [t_0,t_f], \tau \in \mathbb{R}^+, \Phi^{(t,s),x}_{\pi,f}(\epsilon) \in \mathcal{M},$$

where $\Phi^{(t,s),x}_{\pi,f}(.)$ is the flow corresponding to the perturbed control $u_\pi(t^1,\epsilon)$, i.e. $\Phi^{(t,s),x}_{\pi,f}(\epsilon) := \Phi^{(t,s)}_{f^{u_\pi(t^1,\epsilon)}}(x(s))$. The following lemma gives the formula of the variation of $\Phi^{(t,s),x}_{\pi,f}(.)$ at the limit from the right $0^+ := lim_{\epsilon\downarrow 0}\epsilon$. We recall that the point $t^1 \in (t_0,t_f)$ is called a Lebesgue point of $u(.)$ if, ([**3**]):

$$\lim_{s_1 \downarrow t^1} \frac{1}{|s_1 - t^1|} \int_{t^1}^{s_1} |u(\tau) - u(t^1)| d\tau = 0. \tag{2.15}$$

For any $u \in L_\infty([t_0,t_f],U)$, u may be modified on a set of measure zero so that all points are Lebesgue points (see [**59**], page 158, and [**63**]) in which case, necessarily, the value of any cost function is unchanged.

Lemma 2.2 ([**6**]). *For a Lebesgue time t^1, the curve $\Phi^{(t^1,s),x}_{\pi,f}(.) := \Phi^{(t^1,s)}_{f^{u_\pi(t^1,\epsilon)}}(x(s))$: $[0,\tau] \to \mathcal{M}$ is differentiable from the right at $\epsilon = 0$ and the corresponding tangent* *vector* $\frac{d}{d\epsilon}\Phi_{\pi,f}^{(t^1,s),x}|_{\epsilon=0}$ *is given by*

$$\frac{d}{d\epsilon}\Phi_{\pi,f}^{(t^1,s),x}|_{\epsilon=0} = f(x(t^1), u_1) - f_q(x(t^1), u(t^1)) \in T_{x(t^1)}\mathcal{M}. \tag{2.16}$$

□

The tangent vector $f(x(t^1), u_1) - f(x(t^1), u(t^1))$ is called the *elementary perturbation vector* associated to the perturbed control $u_\pi(.,.)$ at $(x(t), t)$. The displacement of the tangent vectors at $x \in \mathcal{M}$ is given by the push-forward of the vector field f_q, see sections below.

2.1.3. Adjoint Processes and the Hamiltonian.

In this section we present the definitions of the adjoint process and the Hamiltonian function which appear in the statement of the Minimum Principle. In the case $\mathcal{M} = \mathbb{R}^n$, by the smoothness of f we may define the following system of differential equations:

$$\dot{\lambda}^T(t) = -\lambda^T(t)\frac{\partial f}{\partial x}(x(t), u(t)),\ t \in [t_0, t_f],\ x(t_0) \in \mathbb{R}^n. \tag{2.17}$$

The matrix solution φ of $\dot{\varphi}(t) = \frac{\partial f}{\partial x}(x(t), u(t))\varphi(t)$, where $\varphi(0) = I$, gives the transformation between tangent vectors on the state trajectory $x(t)$ from time t^1 to t^2 (see [**41**]), in other words, considering v_1 as a tangent vector at $x(t^1)$, the push-forward of v_1 under $\Phi_f^{(t^2,t^1)}$ is

$$v_2 = T\Phi_f^{(t^2,t^1)}(v_1) = \varphi(t^2 - t^1)v_1,\ \ v_1 \in T_{x(t^1)}\mathbb{R}^n \simeq \mathbb{R}^n, t^1, t^2 \in [t_0, t_f]. \tag{2.18}$$

Evidently the vector $v(t) = \phi(t)v(0)$ is the solution of the following differential equation:

$$\dot{v}(t) = \frac{\partial f}{\partial x}(x(t), u(t))v(t), \quad v(0) = v_0, v(t) \in T_{x(t)}\mathbb{R}^n \simeq \mathbb{R}^n. \tag{2.19}$$

A key feature of the solution of (2.17) is that along $x(.)$, $\lambda^T(.)v(.)$ remains constant since

$$\frac{d}{dt}(\lambda^T(t)v(t)) = \dot{\lambda}^T(t)v(t) + \lambda^T(t)\dot{v}(t) = -\lambda^T(t)\frac{\partial f}{\partial x}v(t) + \lambda^T(t)\frac{\partial f}{\partial x}v(t) = 0. \tag{2.20}$$

For a general Riemannian manifold $\mathcal{M}$, the role of the adjoint process λ is played by a trajectory in the cotangent bundle of $\mathcal{M}$, i.e. $\lambda(t) \in T^*_{x(t)}\mathcal{M}$. As in the definition of the tangent lift, we define the *cotangent lift* which corresponds to the variation of a differential form $\alpha \in T^*\mathcal{M}$ (see [**81**]):

$$f^{T^*,u}(\alpha_x) := \frac{d}{dt}|_{t=-s} T^*\Phi_{f^u}^{(-t,s)}(\alpha_x), \quad \alpha_x \in T^*_x\mathcal{M}, \tag{2.21}$$

where $x = x(t) = \Phi_{f^u}^{(t,s)}(x(s))$. As in (2.10), in the local coordinates, (x, p), of $T^*\mathcal{M}$, we have

$$f^{T^*,u}(x,p) = \left[f^{u,i}(x)\frac{\partial}{\partial x^i} - (\frac{\partial f_t^{u,i}}{\partial x^j}p^j)\frac{\partial}{\partial p^i} \right]_{i,j=1}^n, \tag{2.22}$$

where $T^*\Phi_{f^u}^{(-t,s)}(.)$ is the pull back of $\Phi_{f^u}^{(-t,s)}$ applied to differential forms $\alpha_x \in T^*_x\mathcal{M}$. The minus sign in front of t in (2.21) is due to the fact that pull backs act in the opposite sense to push forwards, therefore the variation of a covector α_x at $x = x(s)$ depends upon Φ^{-1} which notationally corresponds to $-t$, see [**81**]. The following lemma gives the connection between the cotangent lift defined in (2.21) and its corresponding flow on $T^*\mathcal{M}$. Let $(T^*\Phi_f^{(t,s)})^{-1} = T^*\Phi_f^{(-t,s)}$, the pull back of Φ^{-1}, whose existence is guaranteed since $\Phi : \mathcal{M} \to \mathcal{M}$ is a diffeomorphism, see [**81**].

Lemma 2.3 ([**6**]). *Consider $f(x(t), u(t))$ as a time dependent vector field on $\mathcal{M}$, then the flow $\Gamma : I \times I \times T^*\mathcal{M} \to T^*\mathcal{M}$, $I = [t_0, t_f]$, satisfies*

$$\Gamma(t, s, (x, p)) = (\Phi_f^{(t,s)}(x), (T^*\Phi_f^{(t,s)})^{-1}(p)), \ (x, p) \in T^*_x\mathcal{M}, \tag{2.23}$$

and Γ is the corresponding integral flow of $f^{T^*,u}$.

□

We now generalize (2.17) and (2.19) to differentiable manifolds. Along a given trajectory $\lambda(.) \in T^*\mathcal{M}$, the variation with respect to time, $\dot{\lambda}(t)$, is an element of $TT^*\mathcal{M}$. The vector field defined in (2.21) is thus the mapping $f^{T^*,u} : T^*\mathcal{M} \to TT^*\mathcal{M}$, which generalizes (2.17) to a mapping from $\lambda(t) \in T^*\mathcal{M}$ to $\dot{\lambda}(t) \in TT^*\mathcal{M}$. The generalization of (2.20) to $\mathcal{M}$ is given in the following proposition.

PROPOSITION 2.1 ([**6**]). *Let $f_q(., u(.)) : \mathcal{M} \times I \to T\mathcal{M}, \quad I = [t_0, t_f]$, be a time dependent vector field giving rise to the associated pair $f^{T,u}, f^{T^*,u}$; then along an integral curve of $f(., u)$ on $\mathcal{M}$*

$$\langle \Gamma, \Psi \rangle : I \to \mathbb{R}, \tag{2.24}$$

is a constant map, where Γ is an integral curve of $f^{T^*,u}$ in $T^*\mathcal{M}$ and Ψ is an integral curve of $f^{T,u}$ in $T\mathcal{M}$. □

The integral curves Γ and Ψ are the generalizations of $\lambda(.)$ and $v(.)$ in (2.19) and (2.20) in R^n to the case of a differentiable manifold $\mathcal{M}$. The corresponding variation of the elementary tangent perturbation in Lemma 2.2 is given in the following proposition.

PROPOSITION 2.2 ([**6**]). *Let $\Psi : [t^1, t_f] \to T\mathcal{M}$ be the integral curve of $f_q^{T,u}$ with the initial condition $\Psi(t^1) = [f(x(t^1), u_1) - f(x(t^1), u(t^1))] \in T_{x(t^1)}\mathcal{M}$, then*

$$\frac{d}{d\epsilon}\Phi_{\pi,f}^{(t,t^1),x}|_{\epsilon=0} = \Psi(t), \quad t \in [t^1, t_f]. \tag{2.25}$$

□

By the result above and Lemma 2.1 we have

$$\frac{d}{d\epsilon}\Phi_{\pi,f}^{(t,t^1),x}|_{\epsilon=0} = T\Phi_f^{(t,t^1)}([f(x(t^1), u_1) - f(x(t^1), u(t^1))]) \in T_{x(t)}\mathcal{M}. \tag{2.26}$$

2.1.4. Hamiltonian Functions and Vector Fields. Here we recall the notions of Hamiltonian vector fields (see e.g. [**5**]), which were employed in [**3**] to obtain a Minimum Principle for optimal control problems in a geometrical framework.

For an optimal (non-hybrid) control problem defined on the state manifold $\mathcal{M}$ (q fixed, $L = 1$), with controlled vector field $f_q(x(t), u(t)) \in T_{x(t)}\mathcal{M}$, the Hamiltonian function for the Mayer problem is defined as:

$$H : T^*\mathcal{M} \times U \to \mathbb{R}, \tag{2.27}$$

$$H(p, x, u) = \langle p, f_q(x, u) \rangle, \quad p \in T_x^*\mathcal{M}, \;\; f_q(x, u) \in T_x\mathcal{M}. \tag{2.28}$$

In general, the Hamiltonian is a smooth function $H \in C^\infty(T^*\mathcal{M})$ with an associated Hamiltonian vector field $\overrightarrow{H} \in \mathfrak{X}(T^*\mathcal{M})$ defined by (see [**3**])

$$\omega_\lambda(., \overrightarrow{H}) = dH, \quad \lambda \in T^*\mathcal{M}, \tag{2.29}$$

where $\omega_\lambda \in \Omega^2(T^*\mathcal{M})$ is the symplectic form (see e.g. [**28**], [**43**]) defined on $T^*\mathcal{M}$ (see [**3, 38**]) and $\mathfrak{X}(T^*\mathcal{M})$ is the space of smooth vector fields defined on $T^*\mathcal{M}$). The Hamiltonian vector field satisfies $i_{\overrightarrow{H}}\omega_\lambda = -dH$, (see [**3**]) where $i_{\overrightarrow{H}}$ is the contraction mapping (see [**37, 43**]) along the vector field $\overrightarrow{H}$. In the local coordinates (x, p) of $T^*\mathcal{M}$, we have:

$$dH = \sum_{i=1}^{n} \frac{\partial H}{\partial p^i} dp^i + \frac{\partial H}{\partial x^i} dx^i, \quad \overrightarrow{H} = \sum_{i=1}^{n} \frac{\partial H}{\partial p^i}\frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i}\frac{\partial}{\partial p^i}. \tag{2.30}$$

So the Hamiltonain system $\dot{\lambda}(t) = \overrightarrow{H}(\lambda)$, $\lambda \in T^*\mathcal{M}$ is locally written as:

$$\dot{x}(t) = \frac{\partial H}{\partial p^i}, \qquad \dot{p}(t) = -\frac{\partial H}{\partial x^i},$$
$$\text{where} \quad \lambda(t) = (x(t), p(t)) \in T^*\mathcal{M}, \;\; x(t_0) = x_0, \lambda(t_f) = dh(x(t_f)) \in T^*_{x(t_f)}\mathcal{M}, \tag{2.31}$$

where $dh = \sum_{i=1}^{n} \frac{\partial h}{\partial x^i} dx^i \in \Omega^1(\mathcal{M})$.

2.1.5. Pontryagin Minimum Principle. For standard (non-hybrid) optimal control problems defined on a Riemannian manifold $\mathcal{M}$ we have the following result known as Pontryganin Minimum Principle.

THEOREM 2.1 ([**41**]). *Consider an OCP satisfying hypotheses **A1-A3** ($L = 0, q_i = q$) defined on a Riemannian manifold $\mathcal{M}$. Then corresponding to an optimal control and optimal state trajectory pair, (u^o, x^o) there exists a nontrivial adjoint trajectory $\lambda^o(.) = (x^o(.), p^o(.)) \in T^*\mathcal{M}$, defined along the optimal state trajectory, such that:*

$$H(x^o(t), p^o(t), u^o(t)) \leq H(x^o(t), p^o(t), v), \;\; \forall v \in U, t \in [t_0, t_f],$$

and the corresponding optimal adjoint trajectory $\lambda^o(.) \in T^*\mathcal{M}$ satisfies:

$$\dot{\lambda^o}(t) = \overrightarrow{H}(\lambda^o(t)), \quad t \in [t_0, t_f].$$

The Minimum Principle gives necessary conditions for optimality; the conditions under which the Minimum Principle is sufficient for optimality are discussed in [**13**] and [**24**].

2.2. The Hybrid Minimum Principle for Autonomous Impulsive Hybrid Systems

Here we consider a simple impulsive autonomous hybrid system consisting of one switching manifold. Consider a hybrid system with a single switching from the discrete state q_0 to the discrete state q_1 at the unique switching time t_s on the optimal trajectory $x^o(.)$ associated with the dynamics:

$$\dot{x}_{q_0}(t) = f_{q_0}(x(t), u(t)), \quad a.e.\, t \in [t_0, t_s), \tag{2.32}$$

$$\dot{x}_{q_1}(t) = f_{q_1}(x(t), u(t)), \quad a.e.\, t \in [t_s, t_f], \tag{2.33}$$

where $x(t_0) = x_0$ and

$$f_{q_i}(., u(.)) : \mathcal{M} \times [t_i, t_{i+1}) \to T\mathcal{M}, \quad i = 0, 1, \tag{2.34}$$

together with a smooth state jump $\zeta := \zeta_{q_0,q_1} : \mathcal{M} \to \mathcal{M}$ with the following action:

$$x^o(t_s) = \zeta(x^o(t_s^-)) = lim_{t \to t_s^-} \zeta(x(t)), \quad x^o(t_s^-) \in \mathcal{S} \subset \mathcal{M}. \tag{2.35}$$

We shall assume the switching manifold $\mathcal{S}$ is an embedded $n-1$ dimensional submanifold $\mathcal{S} := \mathcal{S}_{q_0,q_1}$ which consists of a single switching manifold (see Section 2). Following [**66**], the control needle variation analysis is performed in two distinct cases. In the first case, the variation is applied after the optimal switching time, therefore there is no state variation propagation through the state trajectory before the switching manifold, while in the second case, the control needle variation is applied before the optimal switching time. In this case there exists a state variation propagation along the state trajectory which passes through the switching manifold, see [**66**] and Figure 2.1. Recalling assumption ***A2*** in the Bolza problem and assuming the existence of optimal controls for each pair of given switching state and switching time, let us define $v(x, t)$ for a hybrid system with one autonomous switching, i.e. $L = 1$, as follows:

$$v(x, t) = inf_{u \in \mathcal{U}} J(t_0, t_f, h_0, u), \quad x \in \mathcal{M}, t \in (t_0, t_f), \tag{2.36}$$

where

$$x = \Phi_{f_{q_0}}^{(t^-, t_0)}(x_0) \in \mathcal{S} \subset \mathcal{M}. \tag{2.37}$$

2.2.1. Non-Interior Optimal Switching States. In this subsection, we show that the optimal switching state for an MHOCP derived from a BHOCP (see (2.12)) cannot be an interior point of the *attainable switching set* $\mathcal{A}(x_0, t_s) \subset \mathcal{S}$, $t_0 < t_s < t_f$, for an MHOCP which is defined as

$$\mathcal{A}(x_0, t_s) = \left\{ x \in \mathcal{S} \ \ s.t. \ \exists u \in \mathcal{U}, \Phi_{f_{q_0}^u}^{(t_s^-, t_0)}(x_0) = x \right\}.$$

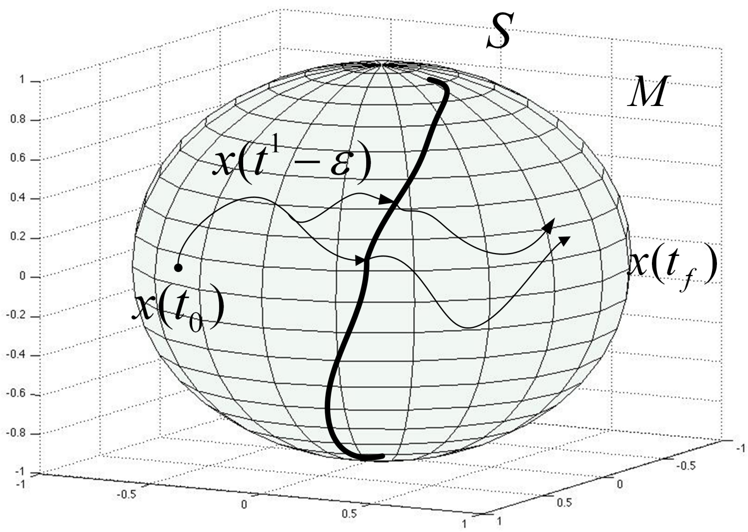

FIGURE 2.1. Hybrid State Trajectory On the Sphere

Note that the state manifold of a Mayer problem derived from a Bolza problem is $\mathcal{M}_B \times \mathbb{R}$ where $\mathcal{M}_B$ is the state manifold of the Bolza problem. In this paper, for simplicity and uniformity of notation, the state manifold and the switching manifold of a Mayer problem shall also be denoted by $\mathcal{M}$ and $\mathcal{S}$ respectively.

LEMMA 2.4. *Consider an MHOCP derived from a BHOCP as in (2.12), (2.13) with a single switching from the discrete state* q_0 *to the discrete state* q_1 *at the unique switching time* t_s *on the optimal trajectory* $(x^o(.), u^o(.))$ *and an* n *dimensional switching manifold* $\mathcal{S} = \mathcal{S}_B \times \mathbb{R} := n_{q_0,q_1}$ *defined in an* $n+1$ *dimensional manifold* $\mathcal{M} = \mathcal{M}_B \times \mathbb{R}$*, where* $\mathcal{S}_B \subset \mathcal{M}_B$ *is the switching manifold of the BHOCP. Then an optimal switching state* $x^o(t_s^-) \in \mathcal{S}$ *at the optimal switching time* t_s *cannot be an interior point of* $\mathcal{A}(x_0, t_s)$ *in the induced topology of* $\mathcal{S}$ *from* $\mathcal{M}$*.*

PROOF. If $\mathcal{A}(x_0, t_s)$ has empty interior in the topology induced on $\mathcal{S}$ from $\mathcal{M}$ the result is immediate. Assume $x^o(t_s^-)$ is an interior point of $\mathcal{A}(x_0, t_s)$, i.e. there exists an open neighbourhood $B_{x^o(t_s^-)} \subset \mathcal{A}(x_0, t_s)$ of $x^o(t_s^-) \in \mathcal{S}$. Let us denote a coordinate system around $x^o(t_s^-)$ by $(x_1^o, ..., x_{n+1}^o)$, where x_{n+1}^o corresponds to the running cost of the Bolza problem, see (2.12). Since the switching manifold of the MHOCP is defined by $\mathcal{S} = \mathcal{S}_B \times \mathbb{R}$, we may choose a neighbourhood $B_{x^o(t_s^-)}$ of $x^o(t_s^-)$ in the induced topology of $\mathcal{S}$ with the last coordinate x_{n+1} free to vary in an open set in $\mathbb{R}$. Hence fixing $x_1^o(t_s^-), ..., x_n^o(t_s^-)$, there exists $y \in B_{x^o(t_s^-)}$ such that

$$y_i = x_i^o(t_s^-), \;\; i = 1, ..., n, \quad y_{n+1} < x_{n+1}^o(t_s^-),$$

which is accessible by f_{q_0} subject to a new control $\hat{u}(t),\;\; t_0 \leq t < t_s$, where $\hat{u}$ is not necessarily equal to u^o. Set the control $u(t) = u^o(t),\;\; t_s \leq t \leq t_f$; then $u(.)$ results in an identical state trajectory on $[t_s, t_f]$ for the Bolza problem (since the variables $x_1, ..., x_n$ do not change). However, the final hybrid cost corresponding to the new switching state y is

$$J(t_0, t_f, (x_0, q_0); 1, (\hat{u}, u^o)) = y_{n+1} + \int_{t_s}^{t_f} l_1(x^o(t), u^o(t))dt + h(x^o(t_f)) < v(x^o(t_s^-), t_s),$$

where $y_{n+1} = \int_{t_0}^{t_s} l_0(\hat{x}(t), \hat{u}(t))dt < x_{n+1}^o = \int_{t_0}^{t_s} l_0(x^o(t), u^o(t))dt$, contradicting the optimality of $x^o(t_s^-)$; we conclude $x^o(t_s^-)$ lies on the boundary of $\mathcal{A}(x_0, t_s)$. □

However the lemma above implies that the hybrid value function defined by (2.36) cannot be differentiated in all directions at the optimal switching state for MHOCPs derived from BHOCPs. Hence the main HMP Theorem 2.2 for MHOCPs below applies in potential to all MHOCPs derived from BHOCPs. The general HMP statement given below employs a differential form dN_x corresponding to the normal vector to the switching manifold $\mathcal{S} \subset M$ at the optimal switching state $x^o(t_s)$. Now in the special case where the value function can be differentiated in all directions at $x^o(t_s) \in \mathcal{S}$, it may be shown that $dN_{x^o(t_s)} = \mu dv(x^o(t_s), t_s)$ for some scalar μ, see [**70**], Lemma A.1; this fact has significant implications for the theory of HMP as is shown in [**71**,**72**,**80**].

2.2.2. Preliminary Lemmas. In order to use the methods introduced in [**3**, **6**, **41**], we establish Lemma 2.5 using the perturbed control $u_\pi(.,.)$ and the associated state variation at the final state $x^o(t_f)$. Denote by $t_s(\epsilon)$ the switching time corresponding to $u_\pi(t,\epsilon)$. Note that, in general, $\Phi^{(t,t_0)}_{\pi,f_q}(x_0)$ does not necessarily intersect the switching manifold at t_s. Hence, we introduce the following perturbed control to guarantee that eventually the state trajectory meets the switching manifold.

$$u_\pi(t,\epsilon) = \begin{cases} u^o(t) & t \leq t^1 - \epsilon \\ u_1 & t^1 - \epsilon \leq t \leq t^1 \\ u^o(t) & t^1 < t \leq t_s \\ u^o(t_s) & t_s \leq t < t_s(\epsilon) \end{cases}, \tag{2.38}$$

The following lemma shows that under the control above, the hybrid state trajectory always intersects the switching manifold for sufficiently small $\epsilon \in \mathbb{R}^+$.

Lemma 2.5. *For an MHOCP satisfying **A1-A3** with a single switching from the discrete state q_0 to the discrete state q_1 at the unique switching time t_s on the optimal trajectory $x^o(.)$, the state trajectory associated to the control needle variation $u_\pi(t^1,\epsilon)$ in (2.38) intersects the $n-1$ dimensional switching manifold $\mathcal{S} \subset \mathcal{M}$ for all sufficiently small $\epsilon \in R^+$ and the corresponding switching time $t_s(\epsilon)$ is differentiable with respect to ϵ.*

Proof. The proof is given in Appendix A.1. □

Lemma 2.6. *For an MHOCP satisfying hypotheses **A1-A3** with a single switching from the discrete state q_0 to the discrete state q_1 at the unique switching time t_s on the optimal trajectory $x^o(.)$, the state variation at the switching time t_s, i.e. $\frac{d\Phi^{(t_s(\epsilon),t^1),x}_{\pi,f_{q_1}}}{d\epsilon}|_{\epsilon=0}$, is given as follows:*

$$\begin{aligned} \frac{d\Phi^{(t_s(\epsilon),t^1),x}_{\pi,f_{q_1}}}{d\epsilon}|_{\epsilon=0} &= T\zeta \circ T\Phi^{(t_s^-,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))] \\ &\quad +(\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}).\Big(T\zeta\big[f_{q_0}(x^o(t_s^-),u^o(t_s^-))\big] - f_{q_1}(x^o(t_s),u^o(t_s))\Big), \\ &\qquad t^1 \in [t_0,t_s). \end{aligned} \tag{2.39}$$

Proof. The proof is obtained by the differentiation of the state flow combination; it is given in Appendix A.2. □

The following lemma gives a variational inequality as a necessary condition for the minimality of the Mayer hybrid cost function $h(x(t_f)) = J(t_0, t_f, x_0, u)$ defined by (2.13) where for the simplicity in notation the Mayer cost function $\hat{h}(.)$ is replaced by $h(.)$. This inequality enables us to construct an adjoint curve $\lambda(.) \in T^*\mathcal{M}$ which satisfies the HMP equations.

In order to prove the following lemma we use the Taylor expansion of a smooth function defined on a Riemannian manifold, see [**4**] and [**67**]. For a given smooth function $h : \mathcal{M} \to \mathbb{R}$ and a vector field $X \in \mathfrak{X}(\mathcal{M})$, where $\mathfrak{X}(\mathcal{M})$ defines the space of all smooth vector fields on $\mathcal{M}$, the Taylor expansion of h around $p \in \mathcal{M}$ along a tangent vector $X_p \in T_p\mathcal{M}$ is given by (see [**67**]):

$$\begin{aligned} h(exp_p\theta X_p) &= h(p) + \theta(\nabla_X h)(p) + ... + \frac{\theta^{n-1}}{(n-1)!} \times (\nabla_X^{n-1} h)(p) \\ &\quad + \frac{\theta^n}{(n-1)!}\int_0^1 (1-t)^{n-1}(\nabla_X^n h)(exp_p t\theta X)dt, \quad 0 < \theta < \theta^*, \end{aligned} \tag{2.40}$$

where $exp_p\theta X_p$ is the geodesic emanating from $p \in \mathcal{M}$ with the velocity $X_p \in T_p\mathcal{M}$, $X(p) = X_p$ and θ^* is the upper bound of the existence of geodesics on the Riemannian manifold $\mathcal{M}$. The existence of θ^* is guaranteed by the fundamental theorem of existence and uniqueness of geodesics (see [**37**]). In (2.40), $\nabla : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$ is the *Levi-Civita* connection on $\mathcal{M}$ which satisfies the following characteristic relations:

$$Xg_{\mathcal{M}}(Y,Z) = g_{\mathcal{M}}(\nabla_X Y, Z) + g_{\mathcal{M}}(Y, \nabla_X Z), \quad \forall X, Y, Z \in \mathfrak{X}(\mathcal{M}),$$

$$(i) : \nabla_X Y - \nabla_Y X = [X,Y], \quad (ii) : \nabla_X f = X(f) \quad \forall X, Y \in \mathfrak{X}(\mathcal{M}), f \in C^\infty(\mathcal{M}). \tag{2.41}$$

Based on the fundamental theorem of existence of geodesics on $\mathcal{M}$ (see [**37**]), for each $v_\pi(t_f) \in T_{x(t_f)}\mathcal{M}$ there exists a geodesic emanating from $x(t_f)$ with the velocity $v_\pi(t_f)$.

LEMMA 2.7. *For an MHOCP satisfying* ***A1-A3*** *with a single switching from the discrete state* q_0 *to the discrete state* q_1 *at the unique switching time* t_s *on the optimal trajectory* $(x^o(.), u^o(.))$*,*

$$\langle dh(x^o(t_f)), v_\pi(t_f)\rangle \geq 0, \quad \forall v_\pi(t_f) \in K_{t_f}, \tag{2.42}$$

where

$$K_{t_f} = K^1_{t_f} \cup K^2_{t_f}, \tag{2.43}$$

and where

$$K^1_{t_f} = \bigcup_{t_s \leq t \leq t_f} \bigcup_{u_1 \in U} T\Phi^{(t_f,t)}_{f_{q_1}}[f_{q_0}(x^o(t), u_1) - f_{q_0}(x^o(t), u^o(t))] \subset T_{x^o(t_f)}\mathcal{M}, \tag{2.44}$$

and

$$\begin{aligned} K^2_{t_f} = &\bigcup_{t_0 \leq t < t_s} \bigcup_{u_1 \in U} T\Phi^{(t_f,t_s)}_{f_{q_1}} \circ T\zeta \circ T\Phi^{(t_s^-,t)}_{f_{q_0}}[f_{q_0}(x^o(t), u_1) - f_{q_0}(x^o(t), u^o(t))] \\ &+ (\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}) T\Phi^{(t_f,t_s)}_{f_{q_1}}\Big(T\zeta\big[f_{q_0}(x^o(t_s^-), u^o(t_s^-))\big] - f_{q_1}(x^o(t_s), u^o(t_s))\Big) \\ &\subset T_{x^o(t_f)}\mathcal{M}. \end{aligned} \tag{2.45}$$

PROOF. To apply (2.40) to h, one needs to extend $v_\pi(t_f) \in T_{x(t_f)}\mathcal{M}$ to a smooth vector field denoted by $\tilde{\mathcal{V}}_\pi \in \mathcal{X}(\mathcal{M})$ such that $\tilde{\mathcal{V}}_\pi(x(t_f)) = v_\pi(t_f)$. It is shown in [**43**] that this extension always exists.

Employing (2.40) on h along $v_\pi(t_f)$ and using the extended smooth vector field $\tilde{\mathcal{V}}_\pi \in \mathcal{X}(\mathcal{M})$, we have

$$h(exp_{x^o(t_f)}\theta v_\pi(t_f)) = h(x^o(t_f)) + \theta(\nabla_{\tilde{\mathcal{V}}_\pi(x(t_f))}h)(x^o(t_f)) + o(\theta), 0 < \theta < \theta^*. \tag{2.46}$$

Here we show that K_{t_f}, defined in Lemma 2.7, contains all the state perturbations at t_f. Lemma 2.2 and Proposition 2.2 together imply that $K^1_{t_f} = \bigcup_{t_s \leq t \leq t_f} \bigcup_{u\in U} T\Phi^{(t_f,t)}_{f_{q_1}}[f_{q_0}(x^o(t,u)) - f_{q_0}(x^o(t), u^o(t))]$ contains all the state perturbations at $x(t_f)$ for all the elementary control perturbations inserted after t_s. For all the control perturbations applied at $t_0 < t < t_s$, either $t_s(\epsilon) < t_s$ or $t_s \leq t_s(\epsilon)$, where $t_s(\epsilon)$ is the switching time corresponding to $u_\pi(t,\epsilon)$.

Following Lemma 2.6, in a local chart around $x(t_s)$, the differentiability of $t_s(\epsilon)$ with respect to ϵ implies

$$\begin{aligned}\frac{d\Phi^{(t_s(\epsilon),t^1),x}_{\pi,f_{q_1}}}{d\epsilon}|_{\epsilon=0} &= T\zeta \circ T\Phi^{(t_s^-,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))] \\ &+(\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0})\Big(T\zeta\big[f_{q_0}(x^o(t_s^-),u^o(t_s^-))\big] - f_{q_1}(x^o(t_s),u^o(t_s))\Big) \\ &\in T_{x^o(t_s)}\mathcal{M},\end{aligned} \tag{2.47}$$

therefore

$$\begin{aligned}K^2_{t_f} = \bigcup_{t_0<t<t_s}\bigcup_{u\in U}\big\{&T\Phi^{(t_f,t_s)}_{f_{q_1}} \circ T\zeta \circ T\Phi^{(t_s^-,t)}_{f_{q_0}}[f_{q_0}(x^o(t,u_1)) - f_{q_0}(x^o(t),u^o(t))] \\ +(\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}).&T\Phi^{(t_f,t_s)}_{f_{q_1}}\Big(T\zeta\big[f_{q_0}(x^o(t_s^-),u^o(t_s^-))\big] - f_{q_1}(x^o(t_s),u^o(t_s))\Big)\big\} \\ &\subset T_{x^o(t_f)}\mathcal{M}, t \in (t_0,t_s),\end{aligned} \tag{2.48}$$

contains all the state variations at $x^o(t_f)$ corresponding to all elementary control perturbations at $t \in (t_0, t_s)$.

Since K_{t_f} contains all the state perturbations at $x^o(t_f)$, choosing $v_\pi(t_f) \in K_{t_f} \subset T_{x(t_f)}\mathcal{M}$ implies that at least at one particular time, one particular elementary control variation $\big(u_\pi(t(v_\pi), \epsilon), u_1(v_\pi)\big)$ results in the final state variation $v_\pi(t_f) \in K_{t_f}$.

Note that choosing $\epsilon = \theta$, $h(exp_{x^o(t_f)}\theta v_\pi(t_f))$ and $h(x_\epsilon(t_f))$, where $x_\epsilon(t_f)$ is the final state curve obtained with respect to ϵ variation, are equal to first order since they have the same first order derivative with respect to ϵ. By the construction of $u_\pi(t, \epsilon) \in \mathcal{U}$, $x_\epsilon(t_f)$ is a curve in the reachable set of the hybrid system at t_f. The minimality of $x^o(t_f)$ consequently implies that $h(x^o(t_f)) \leq h(x_\epsilon(t_f))$; then $h(x_\epsilon(t_f)) - h(exp_{x^o(t_f)}\epsilon v_\pi(t_f)) = o(\epsilon)$ together with (2.46) implies

$$0 \leq (\nabla_{\tilde{\mathcal{V}}_\pi(x(t_f))} h)(x^o(t_f)), \quad \tilde{\mathcal{V}}_\pi(x^o(t_f)) = v_\pi(t_f), \quad \forall v_\pi(t_f) \in K_{t_f}. \tag{2.49}$$

For the smooth function $h : \mathcal{M} \to \mathbb{R}$, (2.41) (ii) implies

$$\tilde{\mathcal{V}}_\pi(h) = \nabla_{\tilde{\mathcal{V}}_\pi} h(x_{t_f}) = \sum_{i=1}^{n} v_\pi^i(x(t_f)) \frac{\partial h}{\partial x_i}, \tag{2.50}$$

where the second equality uses local coordinates. Therefore by the definition of dh we have

$$\nabla_{\tilde{\mathcal{V}}_\pi} h(x_{t_f}) = \langle dh(x(t_f)), v_\pi(x(t_f)) \rangle, \tag{2.51}$$

which implies

$$\langle dh(x(t_f)), v_\pi(x(t_f)) \rangle \geq 0, \quad \forall v_\pi(t_f) \in K_{t_f}, \tag{2.52}$$

and completes the proof. □

2.2.3. Statement of the Hybrid Minimum Principle. Generalizing the results for $\mathcal{M} = \mathbb{R}^n$ in [**66**], we have the following theorem which gives the HMP for autonomous hybrid systems with only one autonomous switching which occurs on the switching manifold $\mathcal{S} \subset \mathcal{M}$.

For an MHOCP with a single switching from the discrete state q_0 to the discrete state q_1 at the unique switching time t_s on the optimal trajectory $(x^o(.), u^o(.))$, where the switching manifold is an $n-1$ dimensional oriented submanifold of $\mathcal{M}$, we have

$$\forall X \in T_x\mathcal{S},\ \ g_{\mathcal{M}}(N_x, X) = 0, \tag{2.53}$$

where $N_x \in T_x^{\perp}\mathcal{S} \subset T_x\mathcal{M}$ is the normal vector at $x^o(t_s^-)$ (the metric $g_{\mathcal{M}}$ is positive definite). For use below we define a one form dN_x, corresponding to N_x by

$$dN_x := g_{\mathcal{M}}(N_x, .) \in T_x^*\mathcal{M}, \tag{2.54}$$

where the linearity of dN_x follows from the bi-linearity of $g_{\mathcal{M}}$.

THEOREM 2.2. *Consider an impulsive MHOCP satisfying hypotheses **A1-A3**. Then corresponding to an optimal control and optimal state trajectory, u^o and x^o with a single switching state at $(x^o(t_s), t_s)$, there exists a nontrivial adjoint trajectory $\lambda^o(.) = (x^o(.), p^o(.)) \in T^*\mathcal{M}$, defined along the optimal state trajectory, such that:*

$$H_{q_i}(x^o(t), p^o(t), u^o(t)) \leq H_{q_i}(x^o(t), p^o(t), u_1),\ \ \forall u_1 \in U, t \in [t_0, t_f],\ \ i = 0, 1, \tag{2.55}$$

and the corresponding optimal adjoint trajectory $\lambda^o(.) \in T^*\mathcal{M}$ satisfies:

$$\dot{\lambda}^o(t) = \overrightarrow{H}_{q_i}(\lambda^o(t)), \quad t \in [t_0, t_f], i = 0, 1, \tag{2.56}$$

for optimal switching state and switching time $(x^o(t_s), t_s)$, there exists $dN_x \in T_x^*\mathcal{M}$ such that

$$\begin{aligned} &p^o(t_s^-) = T^*\zeta(p^o(t_s)) + \mu dN_{x^o(t_s^-)}, \\ &p^o(t_s^-) \in T^*_{x^o(t_s^-)}\mathcal{M}, \quad p^o(t_s) \in T^*_{x^o(t_s)}\mathcal{M}, \\ &x^o(t_s) = \zeta(x^o(t_s^-)), \end{aligned} \tag{2.57}$$

$$x^o(t_0) = x_0,\ \ p^o(t_f) = dh(x^o(t_f)) \in T^*_{x^o(t_f)}\mathcal{M},\ \ dh = \sum_{i=1}^{n} \frac{\partial h}{\partial x^i} dx^i \in T_x^*\mathcal{M}, \tag{2.58}$$

where $\mu \in \mathbb{R}$ *and* $T^*\zeta : T^*\mathcal{M} \to T^*\mathcal{M}$. *The continuity of the Hamiltonian at* $(x^o(t_s), t_s)$ *is given as follows*

$$H_{q_0}(x^o(t_s^-), p^o(t_s^-), u^o(t_s^-)) = H_{q_1}(x^o(t_s), p^o(t_s), u^o(t_s)). \tag{2.59}$$

PROOF. The proof is based on the control needle variation along the optimal state trajectory and employs the results of Lemma 2.7; it is given in Appendix A.3. □

In the case where $dim(\mathcal{S}) < n - 1$, the normal vector at the optimal switching state is not uniquely defined and (2.57) becomes

$$p^o(t_s^-) - T^*\zeta(p^o(t_s)) \in T^{*^\perp}_{x^0(t_s^-)}\mathcal{S} \quad p^o(t_s^-) \in T^*_{x^o(t_s^-)}\mathcal{M}, \quad p^o(t_s) \in T^*_{x^o(t_s)}\mathcal{M}, \tag{2.60}$$

where $T^{*^\perp}_x\mathcal{S} := \{\alpha \in T^*_x\mathcal{M}, \quad s.t. \ \forall X \in T_x\mathcal{S}, \langle \alpha, X \rangle = 0\}$.

2.2.4. Interior Optimal Switching States. Here we specify a hypothesis for MHOCP which expresses the HMP statement based on a differential form of the hybrid value function.

A4: For an MHOCP, the value function $v(x,t), \ x \in \mathcal{M}, t \in (t_0, t_f)$, is assumed to be differentiable at the optimal switching state $x^o(t_s^-)$ in the switching manifold $\mathcal{S}$, where the optimal switching state is an interior point of the attainable switching states on the switching manifold.

We note that ***A4*** rules out MHOCPs derived from BHOCPs (see Lemma 2.3). The following theorem gives the HMP statement for an accessible MHOCP satisfying ***A4***.

THEOREM 2.3. *Consider an impulsive MHOCP satisfying* ***A1-A4***. *Then corresponding to the optimal control and optimal state trajectory* u^o, x^o *with a single switching state at* $(x^o(t_s), t_s)$, *there exists a nontrivial adjoint trajectory* $\lambda^o(.) =$

$(x^o(.), p^o(.)) \in T^*\mathcal{M}$ defined along the optimal state trajectory such that:

$$H_{q_i}(x^o(t), p^o(t), u^o(t)) \leq H_{q_i}(x^o(t), p^o(t), u_1), \forall u_1 \in U, t \in [t_0, t_f], i = 0, 1, \quad (2.61)$$

and the corresponding optimal adjoint variable $\lambda^o(.) \in T^*\mathcal{M}$ satisfies:

$$\dot{\lambda}^o(t) = \overrightarrow{H}_{q_i}(\lambda^o(t)), \quad t \in [t_0, t_f], i = 0, 1. \quad (2.62)$$

At the optimal switching state and switching time $x^o(t_s), t_s$, we have

$$\begin{aligned} &p^o(t_s^-) = T^*\zeta(p^o(t_s)) + \mu dv(x^o(t_s^-), t_s), \\ &p^o(t_s^-) \in T^*_{x^o(t_s^-)}\mathcal{M}, \quad p^o(t_s) \in T^*_{x^o(t_s)}\mathcal{M}, \\ &x^o(t_s) = \zeta(x^o(t_s^-)), \end{aligned} \quad (2.63)$$

where $\mu \in \mathbb{R}, T^*\zeta : T^*\mathcal{M} \to T^*\mathcal{M}$, and

$$dv(x^o(t_s^-), t_s) = \sum_{j=1}^{n} \frac{\partial v(x^o(t_s^-), t_s)}{\partial x^j} dx^j \in T^*_{x^o(t_s)}\mathcal{M}. \quad (2.64)$$

The continuity of the Hamiltonian at $(x^o(t_s), t_s)$ is given as follows

$$H_{q_0}(x^o(t_s^-), p^o(t_s^-), u^o(t_s^-)) = H_{q_1}(x^o(t_s), p^o(t_s), u^o(t_s)). \quad (2.65)$$

PROOF. The proof closely parallels the proof of Theorem 2.2 with the role of $dN_{x(t)}$ being replaced by $dv(x, t)$ whose existence is guaranteed by ***A4***; this is presented in Appendix A.4. □

2.3. Time Varying Switching Manifolds and Discontinuity of the Hamiltonian

In this section we extend the results obtained in the previous section to impulsive autonomous hybrid systems with time varying switching manifolds. The HMP proof parallels the proof of time invariant cases with a modification in the variation of the value function $v(x, t)$ with respect to the switching time. Since $\mathcal{S}$ is time varying, we decompose the metric of $\mathcal{M} \times \mathbb{R}$ as

$$g_{\mathcal{M}\times\mathbb{R}} = g_{\mathcal{M}} \oplus g_{\mathbb{R}}, \tag{2.66}$$

where $g_{\mathbb{R}}$ is the Euclidean metric of $\mathbb{R}$. Now the one form corresponding to the normal vector $N_{(x,t)}$ at $(x,t) \in \mathcal{S} \subset \mathcal{M} \times \mathbb{R}$ is defined as

$$dN_{(x,t)} := g_{\mathcal{M}\times\mathbb{R}}(N_{(x,t)}, .) \in T^*(\mathcal{M} \times \mathbb{R}) = T^*\mathcal{M} \oplus T^*\mathbb{R}. \tag{2.67}$$

Based on the special form of $g_{\mathcal{M}\times\mathbb{R}}$, we can decompose $dN_{(x,t)}$ as

$$dN_{(x,t)} = dN_x \oplus dN_t, \;\; dN_x \in T^*\mathcal{M}, dN_t \in T^*_t\mathbb{R} \simeq \mathbb{R}. \tag{2.68}$$

Theorem 2.4. *Consider an impulsive MHOCP satisfying hypotheses* ***A1-A3*** *where the switching manifold is an n dimensional embedded time varying switching submanifold $\mathcal{S} \subset \mathcal{M} \times \mathbb{R}$ and where the switching state jump is given by a smooth function $\zeta : \mathcal{M} \rightarrow \mathcal{M}$ whenever $(x(t^-), t) \in \mathcal{S}$. Then corresponding to the optimal control and optimal trajectory u^o, x^o with a single switching state at $(x^o(t_s), t_s)$, there exists a nontrivial adjoint trajectory $\lambda^o(.) = (x^o(.), p^o(.)) \in T^*\mathcal{M}$ defined along the optimal state trajectory such that:*

$$H_{q_i}(x(t), p^o(t), u^o(t)) \leq H_{q_i}(x(t), p^o(t), u_1), \quad \forall u_1 \in U, t \in [t_0, t_f], i = 0, 1, \tag{2.69}$$

and the corresponding optimal adjoint variable $\lambda^o(.) \in T^*\mathcal{M}$, (locally given by $\lambda^o(.) = (x^o(.), p^o(.))$) satisfies:

$$\dot{\lambda}^o(t) = \overrightarrow{H}_{q_i}(\lambda^o(t)), \quad t \in [t_0, t_f], i = 0, 1. \tag{2.70}$$

At the optimal switching state and switching time $(x^o(t_s), t_s)$, there exists $dN_x \in T^*_x\mathcal{M}$ such that

$$p^o(t_s^-) = T^*\zeta(p^o(t_s)) + \mu dN_{x^o(t_s^-)},$$
$$p^o(t_s^-) \in T^*_{x^o(t_s^-)}\mathcal{M}, \quad p^o(t_s) \in T^*_{x^o(t_s)}\mathcal{M},$$
$$x^o(t_s) = \zeta(x^o(t_s^-)), \tag{2.71}$$

$$x(0) = x_0^o, \;\; p^o(t_f) = dh(x^o(t_f)) \in T^*_{x^o(t_f)}\mathcal{M}, \;\; dh = \sum_{i=1}^{n} \frac{\partial h}{\partial x^i} dx^i \in T^*_x\mathcal{M}, \tag{2.72}$$

where $\mu \in \mathbb{R}$ *and* $T^*\zeta : T^*\mathcal{M} \to T^*\mathcal{M}$. *The discontinuity of the Hamiltonian at* $(x^o(t_s), t_s)$ *is given by*

$$H_{q_0}(x^o(t_s^-), p^o(t_s^-), u^o(t_s^-)) = H_{q_1}(x^o(t_s), p^o(t_s), u^o(t_s)) - \mu\langle dN_{t_s}, \frac{\partial}{\partial t}\rangle, \tag{2.73}$$

where dN_{t_s} *is the differential form corresponding to the time component of the normal vector at* $(x(t_s^-), t_s)$ *on the switching manifold* $\mathcal{S}$.

Proof. The proof is given in Appendix A.5. □

2.3.1. Time Varying Impulsive Jumps. In this section we investigate the HMP equations in the case of time varying impulsive jumps. For a HOCP with two discrete states, consider the state jump function as a smooth time varying map $\hat{\zeta} : \mathcal{M} \times \mathbb{R} \to \mathcal{M}$. Therefore $T\hat{\zeta} : T\mathcal{M} \oplus T\mathbb{R} \to T\mathcal{M}$ and $T^*\hat{\zeta} : T^*\mathcal{M} \to T^*\mathcal{M} \oplus T\mathbb{R}$. we denote $T\hat{\zeta} = T\zeta \oplus D_t\zeta$, where

$$T\zeta : T\mathcal{M} \to T\mathcal{M}, \quad D_t\zeta : T\mathbb{R} \to T\mathcal{M}, \tag{2.74}$$

where $T\zeta$ and $D_t\zeta$ are the pushforwards of $\hat{\zeta}$ with respect to $t \in \mathbb{R}$ and $x \in \mathcal{M}$ respectively. The following theorem gives the HMP for hybrid impulsive systems in the case of time varying impulse jumps which is consistent with the results presented in [**57**].

THEOREM 2.5. *Consider an impulsive MHOCP satisfying hypotheses* ***A1-A3****. The switching manifold is assumed to be an n dimensional embedded time varying submanifold $\mathcal{S} \subset \mathcal{M} \times \mathbb{R}$ and the switching state jump is given by a time varying smooth function $\hat{\zeta} : \mathcal{M} \times \mathbb{R} \to \mathcal{M}$ which is enabled whenever $(x(t^-), t) \in \mathcal{S}$; then corresponding to the optimal control and optimal trajectory u^o, x^o, with a single switching state at $(x^o(t_s), t_s)$, there exists a nontrivial adjoint trajectory $\lambda^o(.) = (x^o(.), p^o(.)) \in T^*\mathcal{M}$ defined along the optimal state trajectory such that:*

$$H_{q_i}(x^o(t), p^o(t), u^o(t)) \leq H_{q_i}(x^o(t), p^o(t), u_1), \quad \forall u_1 \in U, t \in [t_0, t_f], i = 0, 1, \tag{2.75}$$

and the corresponding optimal adjoint trajectory $\lambda^o(.) \in T^*\mathcal{M}$, locally given by $\lambda^o(.) = (x^o(.), p^o(.))$, satisfies

$$\dot{\lambda}^o(t) = \overrightarrow{H}_{q_i}(\lambda^o(t)), \quad t \in [t_0, t_f], i = 0, 1. \tag{2.76}$$

At the optimal switching state $x^o(t_s)$ and switching time t_s, there exists $dN_x \in T^*_x\mathcal{M}$ such that

$$\begin{aligned} &p^o(t_s^-) = T^*\zeta(p^o(t_s)) + \mu dN_{x^o(t_s^-)}, \\ &p^o(t_s^-) \in T^*_{x^o(t_s^-)}\mathcal{M}, \quad p^o(t_s) \in T^*_{x^o(t_s)}\mathcal{M}, \\ &x^o(t_s) = \zeta(x^o(t_s^-)), \end{aligned} \tag{2.77}$$

$$x^o(t_0) = x_0, \;\; p^o(t_f) = dh(x^o(t_f)) \in T^*_{x^o(t_f)}\mathcal{M}, \;\; dh = \sum_{i=1}^{n} \frac{\partial h}{\partial x^i} dx^i \in T^*_x\mathcal{M}, \tag{2.78}$$

where $\mu \in \mathbb{R}$,

$$T^*\hat{\zeta} = T^*\zeta \oplus D^*_t\zeta : T^*\mathcal{M} \to T^*\mathcal{M} \oplus T^*\mathbb{R}, \tag{2.79}$$

and

$$T^*\zeta : T^*\mathcal{M} \to T^*\mathcal{M}, \quad D_t^*\zeta : T^*\mathcal{M} \to T^*\mathbb{R}. \tag{2.80}$$

The discontinuity of the Hamiltonian at $(x^o(t_s), t_s)$ *is given by*

$$H_{q_0}(x^o(t_s^-), p^o(t_s^-), u^o(t_s^-)) =$$
$$H_{q_1}(x^o(t_s), p^o(t_s), u^o(t_s)) - D_t^*\zeta(p^o(t_s)) - \mu\langle dN_{t_s}, \frac{\partial}{\partial t}\rangle. \tag{2.81}$$

Proof. The proof is given in Appendix A.6. □

2.4. Extension to Multiple Switchings Cases

In this section we obtain the HMP theorem statement for multiple switching hybrid systems where switching manifolds are time invariant. The standing assumption in this section is that $x^o(.)$ is an optimal trajectory under the optimal control $u^o(.)$ for a given MHOCP; it is further assumed that this is a sequence of autonomous transitions along $x^o(.)$ at the distinct time instants $t_0, t_1, ..., t_L$ and $\mathcal{S}_i$ is a time invariant switching manifold subcomponent of $\mathcal{M}$.

Lemma 2.8. *Without loss of generality, assume that for all sufficiently small* $0 \le \epsilon$ *the needle variation* $u_\pi(t, .)$ *applied at* t^1, $t_{j-1} < t^1 < t_j$, *the resulting perturbed trajectories intersect only* $\mathcal{S}_i, i = 0, ..., L$ *and assume further that switching times are greater than the optimal switching times, i.e.* $t_i \le t_i(\epsilon), \quad i = j, ..., L$. *Then the state variation at* t_f *is given as*

$$\frac{d}{d\epsilon}\Phi_\pi^{(t_f,t^1),x}|_{\epsilon=0} = \Big(\prod_{i=0}^{L-j} T\Phi_{f_{q_{i+j}}}^{(t_{i+j+1},t_{i+j})} \circ T\zeta_{i+j}\Big) \circ T\Phi_{f_{q_j}}^{(t_j,t^1)}$$
$$\times\big(f_{q_j}(x(t^1), u_1) - f_{q_j}(x(t^1), u(t^1))\big) + \sum_{i=0}^{L-j}\big(\prod_{l=i}^{L-j} T\Phi_{f_{q_{l+j}}}^{(t_{l+j+1},t_{l+j})} \circ T\zeta_{l+i}\big)$$
$$\times\big(\frac{dt_{i+j}(\epsilon)}{d\epsilon}|_{\epsilon=0}(f_{q_{i+j+1}} - T\zeta_{i+j} f_{q_{i+j}})\big) \in T_{x(t_f)}\mathcal{M}, \tag{2.82}$$

where $T\zeta_{L+1} = I$ *and for simplicity we use* ζ_i *instead of* $\zeta_{q_i,q_{i+1}}$ *for* $i = 0, ..., L$.

PROOF. The proof is based on the results of Lemma 2.2 and an extension of (A.24) and (A.25) to the case where $t_{i+j}(\epsilon)$ is the $(i+j)$th switching time corresponding to $u_\pi(t,\epsilon)$. □

Employing the Lemma above, Lemma 2.7 can be generalized to multiple switching hybrid systems as follows:

LEMMA 2.9. *For a HOCP corresponding to a given sequence of event transitions* $i=0,...,L$, *we have*

$$\langle dh(x^o(t_f)), v_\pi(t_f)\rangle \geq 0, \quad \forall v_\pi(t_f) \in K_{t_f}, \tag{2.83}$$

where

$$K_{t_f} = \bigcup_{r=1}^{L} K_{t_f}^r, \tag{2.84}$$

and

$$\begin{aligned} K_{t_f}^r =& \bigcup_{t_{r-1}\leq t<t_r} \bigcup_{u\in U} \Big(\prod_{i=0}^{L-r} T\Phi_{f_{q_{i+r}}}^{(t_{i+r+1},t_{i+r})} \circ T\zeta_{i+r}\Big) \circ \big\{T\Phi_{f_{q_r}}^{(t_r,t)}\big(f_{q_r}(x(t),u_1) \\ & - f_{q_r}(x(t),u(t))\big)\big\} \\ & + \bigcup_{t_{r-1}\leq t<t_r} \bigcup_{u\in U} \sum_{i=0}^{L-r} \big(\prod_{l=i}^{L-r} T\Phi_{f_{q_{l+r}}}^{(t_{l+r+1},t_{l+r})} \circ T\zeta_{l+i}\big)\big(\frac{dt_{i+r}(\epsilon)}{d\epsilon}|_{\epsilon=0}(f_{q_{i+r+1}} - T\zeta_{i+r} f_{q_{i+r}})\big), \end{aligned} \tag{2.85}$$

PROOF. The proof is parallel to that of Lemma 2.7 and employs the results of Lemma 2.8. □

The following theorem gives the HMP statement for the case of multiple switchings impulsive hybrid systems.

THEOREM 2.6. *Consider a multiple switching impulsive MHOCP satisfying hypotheses* ***A1-A3***. *Then corresponding to the optimal control and optimal state trajectory* u^o, x^o, *there exists a nontrivial* $\lambda^o(.) \in T^*\mathcal{M}$ *along the optimal state trajectory* *such that:*

$$H_{q_i}(x^o(t), p^o(t), u^o(t)) \leq H_{q_i}(x^o(t), p^o(t), u_1), \ \ \forall u_1 \in U, t \in [t_0, t_f], \tag{2.86}$$

and the corresponding optimal adjoint trajectory $\lambda^o(.) \in T^*\mathcal{M}$*, locally given by* $\lambda^o(.) = (x^o(.), p^o(.))$*, satisfies:*

$$\dot{\lambda^o}(t) = \overrightarrow{H}_{q_i}(\lambda^o(t)), \quad t \in [t_0, t_f], \ i = 0, ..., L. \tag{2.87}$$

At the optimal switching state and switching time $(x^o(t_i), t_i)$*, there exists* $dN_x^i \in T_x^*\mathcal{S}_i$ *such that*

$$\begin{aligned}
&p^o(t_i^-) = T^*\zeta_i(p^o(t_i)) + \mu_i dN^i_{x^o(t_i^-)}, \\
&p^o(t_i^-) \in T^*_{x^o(t_i^-)}\mathcal{M}, \quad p^o(t_i) \in T^*_{x^o(t_i)}\mathcal{M}, \\
&x^o(t_i) = \zeta_i(x^o(t_i^-)),
\end{aligned} \tag{2.88}$$

where $\mu_i \in R$ *and* $T^*\zeta_i : T^*\mathcal{M} \to T^*\mathcal{M}$*. The continuity of the Hamiltonian at the switching instants* $(x^o(t_i^-), t_i), i = 0, ..., L$*, is given by*

$$H_{q_i}(x^o(t_i^-), p^o(t_i^-), u^o(t_i^-)) = H_{q_{i+1}}(x^o(t_i), p^o(t_i), u^o(t_i)), \ \ i = 0, ..., L. \tag{2.89}$$

Proof. The proof parallels the proof of Theorem 2.2 employing the results of Lemma 2.9. □

2.5. Simulation Results

To illustrate the results above we consider an HOCP and employ the *Gradient Geodesic-HMP* (GG-HMP) algorithm (see [**77**]).

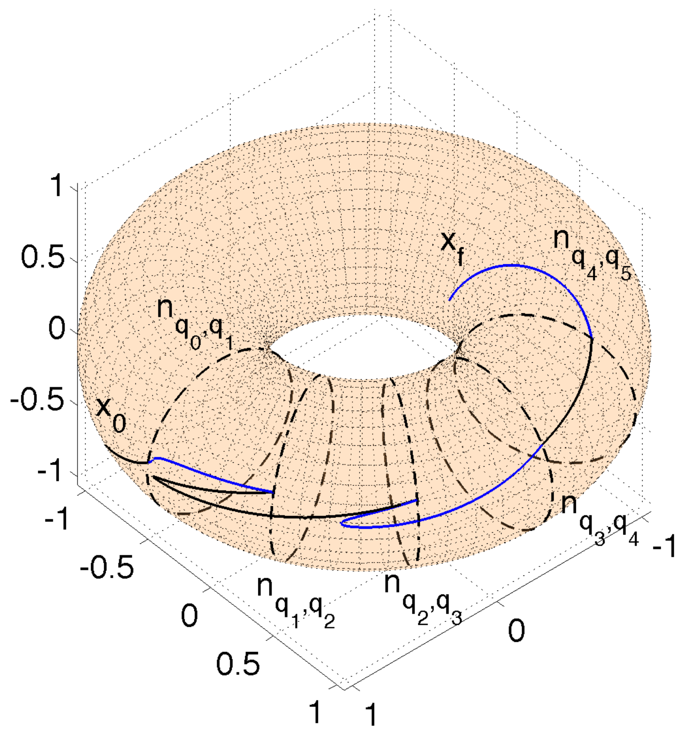

FIGURE 2.2. Hybrid State Trajectory On the Torus

The HOCP is defined on a torus with the following parametrization:

$$x(\zeta, w) = (R + rcos(w))cos(\zeta),$$
$$y(\zeta, w) = (R + rcos(w))sin(\zeta),$$
$$z(\zeta, w) = rsin(w), w, \zeta \in [0, 2\pi). \tag{2.90}$$

where $R = 1, r = 0.5$. The induced Riemannian metric is given by

$$g_{T^2}(\zeta, w) = (R + rcos(w))^2 d\zeta \otimes d\zeta + r^2 dw \otimes dw. \tag{2.91}$$

The hybrid system trajectory goes through each discrete state in numerical order and the dynamics are given in the local parametrization space of the torus T^2 as

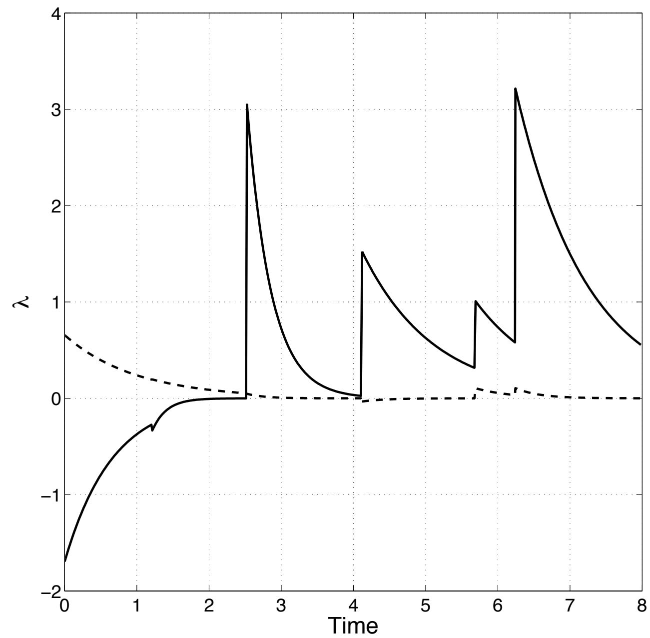

FIGURE 2.3. Adjoint Process

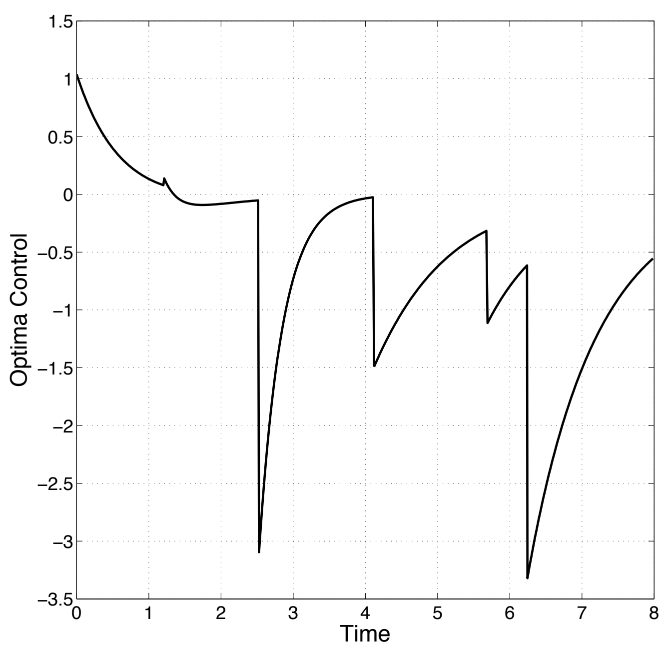

FIGURE 2.4. Control Function

follows:

$$q_0 \quad \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 1.5 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \zeta \\ w \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u, \tag{2.92}$$

$$q_1 \quad \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \zeta \\ w \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u, \tag{2.93}$$

$$q_2 \quad \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} \zeta \\ w \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u, \tag{2.94}$$

$$q_3 \quad \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} \zeta \\ w \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u, \tag{2.95}$$

$$q_4 \quad \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \zeta \\ w \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u, \tag{2.96}$$

$$q_5 \quad \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} \zeta \\ w \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u. \tag{2.97}$$

The switching submanifolds and the cost function are defined as follows:

$$n_{q_0,q_1} = \{0 \le w < 2\pi, \zeta = 0\}, \quad n_{q_1,q_2} = \{0 \le w < 2\pi, \zeta = \frac{\pi}{6}\}, \tag{2.98}$$

$$n_{q_2,q_3} = \{0 \le w < 2\pi, \zeta = \frac{\pi}{3}\}, \quad n_{q_3,q_4} = \{0 \le w < 2\pi, \zeta = \frac{\pi}{2}\}, \tag{2.99}$$

$$n_{q_4,q_5} = \{0 \le w < 2\pi, \zeta = \frac{2\pi}{3}\}, \quad J = \frac{1}{2}\int_0^8 u^2(t)dt, \tag{2.100}$$

and the boundary conditions are given as:

$$x_0 = (1.4117, -0.4367, -0.1478) \in R^3, \tag{2.101}$$
$$x_f = (-0.1478, -0.49980, 0.10130) \in R^3.$$

The hamiltonian functions are given as

$$H_{q_i}(\begin{pmatrix} \zeta \\ w \end{pmatrix}, p(t), u(t)) = (p_1(t), p_2(t)) \begin{pmatrix} \dot{\zeta} \\ \dot{w} \end{pmatrix} + \frac{1}{2}u^2(t), \;\; i = 0, ..., 5, \;\; t \in [t_i, t_{i+1}). \tag{2.102}$$

The GG-HMP algorithm is an extension to Riemannian manifolds of the HMP algorithm introduced in [**66**]; this is done by introducing a geodesic gradient flow algorithm on $\mathcal{S}$ and constructing an HMP algorithm along geodesics on $\mathcal{S}$. Figure 2.2 shows the state trajectory on the torus and Figure 2.3 depicts the adjoint variable with the discontinuity at the optimal switching times
$\overline{t_s} = [1.2137, 2.6250, 4.0145, 5.2821, 6.6382]$.

CHAPTER 3

Gradient Geodesic and Newton Geodesic HMP Algorithms

In this chapter we consider hybrid systems defined on $\mathcal{M} = R^{n+1}$ and switching manifolds are given by $\mathcal{N} = M := \{m_{i,j} = 0, \ i,j \in Q\}$. The following theorem gives the Hybrid Maximum Principle in an extended class of the cases treated in [**66**], specifically the autonomous switchings case is extended to the time varying switching manifold case. It is shown that the discontinuity of the Hamiltonian functions and adjoint variables at an optimal switching state and switching time gives important information about the geometry of the switching manifold M at switching states.

Theorem 3.1 ([**66**]). *Consider a hybrid system satisfying the assumptions **A1-A3** in Chapter 2 and define*

$$H_q(x,\sigma,u,\lambda) = \lambda^T f_{\sigma(q)}(x,u) + l_{\sigma(q)}(x,u), \ \ \lambda \in R^{n+1}, u \in U, q \in Q.$$

Assume that the hybrid system contains only autonomous switchings and let

$$J^o(t_0,t_f,h_0,L) = inf_{u\in\mathcal{U}} J(t_0,t_f,h_0,L,u)$$

be the infimized cost function with infimizing control u^o and trajectory (x^o,q^o) which are both assumed to exist. Let us assume we have L autonomous switchings and let

$t_1, t_2, ..., t_L$, *denote the switching times along the optimal trajectory.*
Finally, assume that almost everywhere along an optimal trajectory, the continuous state x*, satisfies the controllability condition given in [**66**]. Then:*
(i) There exists a piecewise absolutely continuous adjoint process satisfying

$$\dot{\lambda}_j^o = -\frac{\partial H_j}{\partial x}(x^o, \sigma^o, \lambda, u^o), \quad u_t^o \in U \ \ a.e., t \in (t_j, t_{j+1}).$$

(ii) At the switching times the adjoint process and Hamiltonain function satisfy

$$\lambda_j(t_j^-) = \lambda_{j+1}(t_j) + p_j \nabla_x m_{j,j+1}(x(t_j), t_j), \ 1 \leq j \leq L, \tag{3.1}$$

$$H_j(t_j^-) = H_{j+1}(t_j) - p_j \nabla_t m_{j,j+1}(x(t_j), t_j), \ 1 \leq j \leq L. \tag{3.2}$$

(iii) Along the optimal trajectory the Hamiltonian minimization condition is satisfied

$$H_j(x^o(t), \sigma^o(t), \lambda_j(t), u^o(t)) \leq H_j(x^o(t), \sigma^o(t), \lambda_j(t), v), \\ \forall v \in U, \quad t \in [t_j, t_{j+1}), j \in [0, 1, , ..., L]. \tag{3.3}$$

i.e.

$$\lambda^{oT}(t) f_{\sigma^o(q^o(t-))}(x_t^o, u_t^o) + l_{\sigma^o(q^o(t-))}(x_t^o, u_t^o) \leq \\ \lambda^{oT}(t) f_{\sigma^o(q^o(t-))}(x_t^o, v) + l_{\sigma^o(q^o(t-))}(x_t^o, v), \quad \forall v \in U.$$

□

For simplicity of the analysis we only consider the case of one autonomous switching from $q_0 \in Q$ to $q_1 \in Q$, and so $u^o \equiv (u_1^o, u_2^o)$; the extension to the general case is straightforward but engenders significant complexity, see [**66**].

3.1. The HMP Algorithm

In this section we review the HMP algorithm presented in detail in [**66**].

By the assumption of the existence of solutions for the Hybrid Optimal Control Problem (HOCP), there exists a switching time t_s and a switching state $x_s \in M$ which locally minimize the hybrid cost function. As stated in Theorem 3.1, the boundary conditions for the Hamiltonian and the adjoint variables are as follows: $\lambda_j(t_j^-) = \lambda_{j+1}(t_j) + p_j \nabla_x m_{j,j+1}(x(t_j), t_j), \quad H_j(t_j^-) = H_{j+1}(t_j) - p_j \nabla_t m_{j,j+1}(x(t_j), t_j), \quad p_j \in R$, where, by the Pontryagin Maximum Principle (PMP), $\dot{\lambda} = -H_x, \dot{x} = H_\lambda$, together with the appropriate boundary conditions ([**66**]).

In the case of an autonomous hybrid system consisting of two phases, separated by a switching manifold M, the optimal hybrid cost is

$$J^* = inf_{t_s,x_s}[inf_{u\in U} J(t_0, t_f, h_0, L, u, t_s, x_s)], \tag{3.4}$$

where x_s, t_s are the switching state and switching time on the switching manifold M. The HMP algorithm proceeds with initializing (x_s, t_s) and taking ϵ_f such that $0 < \epsilon_f << 1$. Setting the iteration number k to zero, the optimal control $u_1^k(t), 0 < t < t_s$ and $u_2^k(t), t_s \le t < t_f$, are evaluated. Incrementing k by 1, $\nabla_x M(x_s^{k-1}, t_s^{k-1})$ and $\nabla_t M(x_s^{k-1}, t_s^{k-1})$ are evaluated at the previous switching state and switching time respectively. The updating procedure is given as follows:

$$\begin{aligned} t_s^k &= t_s^{k-1} - r_k\big(H_1^k(t_s^{k-1}) - H_2^k(t_s^{k-1}) \\ &\quad - \frac{\partial M}{\partial t}(t_s^{k-1}, x_s^{k-1})p^k\big) - r_k \frac{\partial M}{\partial t}(t_s^{k-1}, x_s^{k-1}) M(t_s^{k-1}, x_s^{k-1}), \end{aligned} \tag{3.5}$$

$$\begin{aligned} x_s^k &= x_s^{k-1} - r_k\big(\lambda_2^k(t_s^{k-1}) - \lambda_1^k(t_s^{k-1}) \\ &\quad - \frac{\partial M}{\partial x}(t_s^{k-1}, x_s^{k-1})p^k\big) - r_k \frac{\partial M}{\partial x}(t_s^{k-1}, x_s^{k-1}) M(t_s^{k-1}, x_s^{k-1}), \end{aligned} \tag{3.6}$$

where

$$p^k = (Q_k^T Q_k)^{-1} Q_k^T \begin{pmatrix} H_1^k(t_s^{k-1}) - H_2^k(t_s^{k-1}) \\ \lambda_2^k(t_s^{k-1}) - \lambda_1^k(t_s^{k-1}) \end{pmatrix}, \tag{3.7}$$

and

$$Q_k = \begin{pmatrix} H_1^k(t_s^{k-1}) - H_2^k(t_s^{k-1}) \\ \lambda_2^k(t_s^{k-1}) - \lambda_1^k(t_s^{k-1}) \end{pmatrix}. \tag{3.8}$$

By the algorithm above, the necessary conditions for the optimality are satisfied whereby the switching state and switching time lie on the switching manifold, [**66**]. A disadvantage of the HMP algorithm is that it does not guarantee $x_s^k \in M$ at each step k. The last terms in (3.5, 3.6) are added in order to enforce the approach of x_s^k, t_s^k to the switching manifold in the limit. In the next section the general geodesic gradient flow algorithm is defined on M, and then is used to construct an HMP algorithm on the hybrid switching surface M.

3.2. Geodesic-Gradient Flow Algorithm

3.2.1. Formulation and Analysis of the GG-HMP Algorithm. In the modified version of the HMP algorithm the initial x_s^0 is chosen such that $x_s^0 \in M$ and this constraint is maintained by moving along geodesics on M by recursively choosing the hybrid value function gradient to initialize the geodesic search directions (see [**74**]).

For a hybrid control problem which consists of two distinct phases with a time invariant switching manifold M, the hybrid value function $v(.,.)$ is defined as:

$$v(x,t) = inf_{u\in\mathcal{U}} J(t_0, t_f, h_0; x_{t_s}, t_s, u)|_{t_s=t, x_{t_s}=x}, \;\; x_{t_s} \in M. \tag{3.9}$$

In this chapter we shall always assume $v(.,.) \in C^{k,l}(M,R)$ for sufficiently large k,l to make the analysis under discussion valid. The following lemmas, respectively, give the relations between the sensitivity of the hybrid value function and the discontinuity of the Hamiltonian function and adjoint variables at the switching state and switching time.

Lemma 3.1 ([**80**]). *Let $(x(t_s), t_s) = (x(t_s^o), t_s^o)$ be the optimal switching state and switching time subject to the hypotheses of the HSOC problem and of Theorem 3.1, then*

$$(i) \quad \frac{\partial v(x,t)}{\partial t}|_{(x(t_s^o),t_s^o)} = 0, \tag{3.10}$$

$$(ii) \quad \nabla_x v(x,t)|_{(x(t_s^o),t_s^o)} \perp T_{x(t_s^o)}M, \tag{3.11}$$

where $T_{x(t_s)}M$ is the tangent space at the switching state $x(t_s)$.

Proof. The proof is given in Chapter 5 (Lemma 5.1). □

Lemma 3.2 ([**1**,**80**]). *For the HSOC problem defined in Theorem 3.1 the following relations hold for all $(x(t_s), t_s) \in R^{n+1} \times R$,*

$$\frac{\partial v(x,t)}{\partial t}|_{(x(t_s),t_s)} = H_1(t_s^-) - H_2(t_s), \tag{3.12}$$

$$\nabla_x v(x(t_s), t_s) = \lambda_2(t_s) - \lambda_1(t_s^-). \tag{3.13}$$

□

The basic notions of Riemannian manifolds needed for the rest of the chapter are as follows: For a given n dimensional switching manifold M, a Riemannian metric $g(x) \in T_x^*M \otimes T_x^*M$ is defined as (T^*M is the cotangent bundle of the switching manifold M):

$$g(x) = \sum_{i,j=1}^{n} g_{ij}(x) dx_i \otimes dx_j, \quad i,j = 1, ..., n. \tag{3.14}$$

For a given curve $\gamma : [a,b] \to M$ which is locally described as $\gamma(t) = (x_1(t), ..., x_n(t))$, the associated length with respect to the Riemannian metric g is defined as

$$L(\gamma) := \int_a^b (\sum_{i,j=1}^{n} g_{ij}(x(s)) \dot{x}_i(s) \dot{x}_j(s))^{\frac{1}{2}} ds. \tag{3.15}$$

Fixing the initial and final points $\gamma(a), \gamma(b) \in M$, a minimal length curve which connects $\gamma(a)$ to $\gamma(b)$ is called *geodesic*, [**37**]. The solution of the Euler-Lagrange variational problem associated with (3.15) shows that all the geodesics on M connecting $\gamma(a), \gamma(b)$ must satisfy the following system of ordinary differential equations:

$$\ddot{x}_i(s) + \sum_{j,k=1}^{n} \Gamma^i_{j,k} \dot{x}_j(s) \dot{x}_k(s) = 0, \quad i = 1, ..., n, \tag{3.16}$$

where

$$\Gamma^i_{j,k} = \frac{1}{2} \sum_{l=1}^{n} g^{il}(g_{jl,k} + g_{kl,j} - g_{jk,l}), \quad g_{jl,k} = \frac{\partial}{\partial x_k} g_{jl}. \tag{3.17}$$

All the indices i, j, k here run from 1 up to n and $[g^{ij}] = [g_{ij}]^{-1}$. The existence of the geodesics on the switching manifold M is given by the following theorem.

Theorem 3.2 ([**37**]). *Let M be a Riemannian manifold, choose $p \in M$ and $v \in T_pM$, then there exist, $\epsilon > 0$ and precisely one geodesic γ such that $\gamma : [0, \epsilon] \to M$ and $\gamma(0) = p, \dot{\gamma}(0) = v$.* □

The assumption that the Riemannian manifold is complete as a metric space implies that it is *geodesically* complete [**37**], that is to say the interval $[0, \epsilon]$ can be extended to R for all $x \in M$ and all $v \in T_pM$. For the sake of simplicity in notation we consider x alone as the optimization variable. However all the results can be modified by extending the state variable by t_s^k. Here we formulate the projection of the value function gradient on the tangent space of the switching manifold. Following [**29**] and [**46**] we consider a time invariant switching manifold M as an embedded n dimensional surface in R^{n+1}, $x(.) \in R^{n+1}$. Similar to the HMP algorithm, the analysis and the algorithms in this chapter can be extended to the case where the switching manifold M is time varying, i.e. locally described by $M(x, t) = 0$.

We assume the switching manifold M to be an n dimensional Riemannian embedded connected submanifold of R^{n+1} with metric $g(.,.)$ induced by the Euclidean metric of R^{n+1}, i.e

$$g(\varphi_x, \omega_x) :=< \varphi_x, \omega_x >, \quad \forall \varphi_x, \omega_x \in T_x M, \tag{3.18}$$

where $<,>$ is the corresponding Euclidean inner product in R^{n+1}. Then the projected gradient of the value function $v(x)$ is defined as the unique vector $\nabla_M^\gamma v(x) \in T_x M$, such that (see [**29**])

$$g(\nabla_M^\gamma v(x), \omega_x) =< \nabla_x v(x), \omega_x >, \quad \forall \omega_x \in T_x M, \tag{3.19}$$

where $\nabla_x v(x)$ is the Euclidean gradient of $v(.)$ in R^{n+1} given by Lemma 3.2. It is important to notice that $\nabla_M^\gamma v(x)$ satisfying (3.19) can be defined for an arbitrary g, which is not necessarily induced by the Euclidean metric

Definition 3.1. *Let M be a geodesically complete Riemannian manifold, $p \in M$ and $V_p = \{w \in T_p M; \gamma_w$ is defined on $[0, \theta]\}$, then $exp_\theta^p : V_p \to M$ is defined by $exp_\theta^p(w) = \gamma_w(\theta)$, where γ_w is the geodesic emanating from p and $\dot{\gamma}_w = w$.* □

Here we give a continuous version of the HMP algorithm along geodesics (GG-HMP) and prove the convergence by using the Lasalle Theory. By initializing the switching state x_s^k on the switching manifold M, we solve two boundary value problems to obtain optimal controls for (x_0, x_s^k) and (x_s^k, x_f) and compute $\nabla_M^\gamma v(x_s^k)$ via Lemma 3.2.

The Geodesic-Gradient flow is defined as follows:

Definition 3.2. *(Geodesic-Gradient Flow) Let $\theta^0 = 0$, and $x(\theta^0) = x^0 \in M$, then for all $0 \le k$ and all x^k such that $\nabla_M^\gamma v(x^k) \neq 0$, define*

$$\begin{aligned} x(\theta) &= \gamma_{x^k}(\theta) = exp_{x^k}(-\theta \nabla_M^\gamma v(x^k(\theta^k))), \\ &\qquad \theta \in [\theta^k, \theta^{k+1}), \quad x(\theta) \in M, \end{aligned} \tag{3.20}$$

where

$$\theta^{k+1} = \sup_s \{s; \frac{dv(x(t))}{dt} \leq 0, t \in [\theta^k, s), s \in [\theta^k, \theta^k + 1)\}, \tag{3.21}$$

where $\gamma_{x^k}(s)$ is the geodesic emanating from x^k with velocity $-\nabla^\gamma_M v(x^k)$ given by (3.19). □

The next switching state is specified as

$$x^{k+1} = x(\theta^{k+1}), \quad 0 \leq k < \infty, \tag{3.22}$$

and is the origin of the next step geodesic unless $\nabla^\gamma_M v(x^{k+1}) = 0$, in which case $x^{k+1} = x(\theta^{k+1})$ is defined to be a finite iteration equilibrium point of the flow.
In other words, using (3.20) and (3.21) we construct a flow φ on the switching manifold M which is differentiable from the right at all points $x \in \varphi(\theta, x^0)$. The derivative from the right of φ at the switching state x^i is given as

$$\begin{aligned} \frac{dv(x^i(s))}{d\theta}|_{s=0} &= \mathcal{L}_{-\nabla^\gamma_M v(x^i)} v(x^i(s))|_{s=0} \\ &= < \nabla_x v(x^i), -\nabla^\gamma_M v(x^i) > \\ &= -g(\nabla^\gamma_M v(x^i), \nabla^\gamma_M v(x^i)) \\ &\equiv -||\nabla^\gamma_M v(x^i)||^2, \quad x(0) = x^i, \end{aligned} \tag{3.23}$$

where the third equality is given by (3.19). The equation above turns into the following formula for the derivative from the right of the value function at the point x which is not a switching state.

$$\frac{dv(x(s))}{d\theta}|_{s=0} = \mathcal{L}_X v(x(s))|_{s=0}, \tag{3.24}$$

where X is the tangential vector field to the geodesic curve γ at x as:

$$X = \frac{d\gamma(s)}{ds}|_{s=0}, \quad \gamma(0) = x, s \in [\theta^k, \theta^{k+1}), k \geq 0. \tag{3.25}$$

Over the interval of existence $[0, \omega)$ we denote the total flow induced by (3.20)-(3.23) as

$$\varphi(\theta, x^0) = \Pi_{i=1}^{n} \psi_i(\theta^{i-1}, \theta^i, x^{i-1}) \circ \psi_n(\theta^n, \theta, x^n), \tag{3.26}$$

where

$$\psi_i(\theta^{i-1}, \theta^i, x^{i-1}) = \gamma_{x^{i-1}}(\theta^i - \theta^{i-1})), \tag{3.27}$$

θ^i is the switching time to the next iteration and n is the index number of the last switching before θ. By the continuity of the geodesic flows $\{\psi_i, \quad 1 \leq i\}$, φ is a continuous map on $[0, \omega)$. In the notation of topological dynamics, and in particular Lasalle Theory (see e.g. [**20**, **60**]), the limit set of the initial state x^0 is denoted as $\Omega(x^0)$, where

$$y \in \Omega(x^0) \Rightarrow \exists \theta_n, n \geq 1, \quad s.t. \quad lim_{n \to \infty} x(\theta_n) = y, \tag{3.28}$$

when $lim_{n \to \infty}(\theta_n) = \omega$. Note the sequence $\{\theta_n\}$ is in general distinct from $\{\theta^n\}$.

H1: There exists $0 < b < \infty$ such that the associated sublevel set $\mathcal{N}_b = \{x \in M; \quad v(x) < b\}$ is (i) open (ii) connected, (iii) contains a strict local minimum x_* which is the only local minimum in $\mathcal{N}_b$, and (iv) $\mathcal{N}_b$ has compact closure.

Now choose $x^0 \in \mathcal{N}_{b-\epsilon}$ for $0 < \epsilon < b$. By the construction of φ, for all $0 < \theta < \theta'$ we have

$$v(\varphi(\theta', x^0)) \leq v(\varphi(\theta, x^0)) \leq v(x^0) < b - \epsilon < b, \tag{3.29}$$

and hence for $\Phi^+ := \{\varphi(\theta, x^0); 0 \leq \theta < \omega\}$

$$\overline{\Phi^+} \subset \overline{\mathcal{N}}_{b-\epsilon} \subset \mathcal{N}_b. \tag{3.30}$$

Hence the flow φ is defined everywhere in $\overline{\mathcal{N}}_{b-\epsilon} \subset \mathcal{N}_b$, where $\overline{\mathcal{N}}_{b-\epsilon}$ is compact since $\overline{\mathcal{N}}_b$ is compact. So for all $x \in \mathcal{N}_{b-\epsilon}$ we have an extension of $\varphi(., x^0)$ in $\mathcal{N}_b$, therefore the maximum interval of existence of $\varphi(., x^0)$ in $\mathcal{N}_b$ is infinite.

THEOREM 3.3. *Subject to the hypothesis* ***H1*** *on* $\mathcal{N}_b$ *and with an initial state* x^0 *such that* $x^0 \in \mathcal{N}_{b-\epsilon} \subset M$, $0 < \epsilon < b$, *either the Geodesic-Gradient flow,* φ, *reaches an equilibrium after a finite number of switchings, or it satisfies*

$$\varphi(\theta, x^0) \to \Omega(x^0) \subset v^{-1}(c), \quad c \in R, \tag{3.31}$$

as $\theta \to \infty$, *for some* $c \in R$, *where*

$$\forall y \in \Omega(x^0), \quad \frac{dv(y)}{d\theta}|_{\theta=0} = 0, \tag{3.32}$$

and, furthermore, the switching sequence $\{x\}_0^\infty = \{x^0, x^1, \cdots, \}$ *converges to the limit point* $x_* \in \Omega(x^0) \subset \mathcal{N}_b$, *where* x_* *is the unique element of* $\mathcal{N}_b$ *such that* $\nabla_M^\gamma v(x_*) = 0$.

PROOF. The first statement of the theorem is immediate by Definition 3.2. To prove the second statement, similar to the proof of the Lasalle Theorem, we proceed by showing that $v(.)$ is constant on the set $\Omega(x^0)$. The precompactness of Φ^+ ((i): $\overline{\Phi^+} \subset \overline{\mathcal{N}}_b$ (ii): there does not exist $\theta_i \to \omega, i \to \infty$, such that $\varphi(\theta_i, x^0) \to \partial\mathcal{N}_b$, i.e. $\overline{\Phi^+} \bigcap \partial\mathcal{N}_b = \oslash$), imply $\Omega(x^0) \neq \oslash$, see [**20**]. By the definition of $\Omega(x^0)$ we have

$$\forall y \in \Omega(x^0) \Rightarrow \exists \theta_n, n \geq 1, \quad s.t. \quad \varphi(\theta_n, x^0) \to y, \theta_n \to \infty, \tag{3.33}$$

and since $v(.) \in C^1$,

$$lim_{n\to\infty} v(x(\theta_n)) = lim_{n\to\infty} v(\varphi(\theta_n, x^0)) = v(y) =: c. \tag{3.34}$$

Now choose $y' \in \Omega(x^0), y' \neq y$, then by the existence of a convergent sequence $x(\theta_n')$ to y' we have

$$\begin{aligned} &\forall \epsilon > 0 \Rightarrow \exists n, n_i, k \quad s.t. \quad \theta_n < \theta_{n_i}' < \theta_{n+k} \\ &c - \epsilon < v(x(\theta_{n+k})) \leq v(x(\theta_{n_i}')) \leq v(x(\theta_n)) < c + \epsilon, \end{aligned} \tag{3.35}$$

i.e. $v(y^{'}) = c$, hence $\Omega(x^0) \subset v^{-1}(c)$. To prove stationarity, i.e (3.32), we observe that $\Omega(x^0)$ is positive invariant under the flow φ, i.e

$$\varphi(\theta, \Omega(x^0)) \subset \Omega(x^0), \quad \theta > 0. \tag{3.36}$$

This follows from the continuity of $\varphi(.,.)$, see [**20**]. Differentiability from the right for all $x \in \varphi(\theta, x^0), 0 < \theta$, implies

$$\begin{aligned} \frac{dv}{d\theta}|_{\theta=0} &= lim_{\theta \to 0^+} \frac{v(\varphi(\theta, y)) - v(\varphi(0, y))}{\theta} \\ &= lim_{\theta \to 0^+} \frac{c - c}{\theta} = 0, \quad y \in \Omega(x^0), \end{aligned} \tag{3.37}$$

since $\varphi(\theta, y) \in \Omega(x^0)$ by (3.36) and $v(\Omega(x^0)) = c$ by (3.35).

It remains to prove the statement for the sequence of the switching states $\{x\}_0^\infty = \{x^0, x^1, \cdots\}$. The switching sequence $\{x\}_0^\infty$ consists of the switching points on $\varphi(\theta, x^0)$ which by (3.21) is an infinite sequence.

The precompactness of Φ^+ with respect to $\mathcal{N}_b$ implies the existence of a convergent subsequence of $\{x\}_0^\infty$ such that

$$lim_{i \to \infty} \varphi(\theta_i^n, x^0) = x^* \in \Omega(x^0), \quad \Omega(x^0) \subset \overline{\Phi^+} \subset \overline{\mathcal{N}_{b-\epsilon}}. \tag{3.38}$$

Since $v \in C^\infty(\mathcal{N}_b)$

$$lim_{i \to \infty} \nabla_M^\gamma v(\varphi(\theta_i^n, x^0)) = \nabla_M^\gamma v(x^*), \tag{3.39}$$

and

$$lim_{i \to \infty} \frac{dv(\varphi(\theta_i^n, x^0))}{d\theta}|_{\theta=0} = \frac{dv(x^*)}{d\theta}|_{\theta=0}. \tag{3.40}$$

But since the state $\varphi(\theta_i^n, x^0))$ is a switching state chosen from the switching sequence $\{x\}_0^\infty$,

$$\begin{aligned} \frac{dv(\varphi(\theta_i^n, x^0))}{d\theta}|_{\theta=0} &= < \nabla_{\varphi(\theta_i^n, x^0)} v(x), \nabla_M^\gamma v(\varphi(\theta_i^n, x^0)) > \\ &= -||\nabla_M^\gamma v(\varphi(\theta_i^n, x^0))||^2, \end{aligned} \tag{3.41}$$

As is stated in (3.38), the limit point x^* is an element of the limit set $\Omega(x^0)$, therefore by (3.37) we have

$$\frac{dv(x^*)}{d\theta}|_{\theta=0} = 0. \tag{3.42}$$

From (3.39)-(3.41) we have

$$\begin{aligned} 0 &= \frac{dv(x^*)}{d\theta}|_{\theta=0} = lim_{i\to\infty} \frac{dv(\varphi(\theta_i^n, x^0))}{d\theta}|_{\theta=0} \\ &= lim_{i\to\infty}(-||\nabla_M^\gamma v(\varphi(\theta_i^n, x^0))||^2) = -||\nabla_M^\gamma v(x^*)||^2. \end{aligned} \tag{3.43}$$

Hence

$$\nabla_M^\gamma v(x^*) = 0. \tag{3.44}$$

But by ***H1***, x_* is the unique point in $\mathcal{N}_{b-\epsilon} \subset \mathcal{N}_b$ for which this holds, hence all subsequences of $\{x\}_0^\infty$ converge to $x_* = x^*$ an hence so does the sequence. □

DEFINITION 3.3. *(Conceptual GG-HMP Algorithm)*
Consider the hybrid system with two phases separated by the switching manifold M, and the performance function $v(.)$.

Generate the Geodesic-Gradient flow, (3.20)-(3.22), on M with $\nabla_M^\gamma v(x)$, $x \in M$, evaluated by (3.12), (3.13), (3.18), (3.19).

Stopping rule: for a given $0 < \beta$, if $||\nabla_M^\gamma v(x)|| < \beta$ stop. □

THEOREM 3.4. *Assume **H1** holds for $\mathcal{N}_b \subset M$, for the HOCP with the switching manifold M and the performance function $v(.)$, then the GG-HMP with data (M, v, β)*

halts at $x^{k(\beta)}(\beta)$*, where either* $x^{k(\beta)}(\beta)$ *is a finite equilibrium point of the Geodesic-Gradient flow, and hence* $\nabla^{\gamma}_{M} v(x^{k(\beta)}(\beta)) = 0$ *and* $x^{k(\beta)}(\beta) = x_*$*, where* x_* *is the unique point of* $\mathcal{N}_b \subset M$ *such that* $||\nabla^{\gamma}_{M} v(x_*)|| = 0$*, or* $x^{k(\beta)}(\beta)$ *is such that*

$$x^{k(\beta)}(\beta) \to x_*, \quad k(\beta) \to \infty, \quad as \quad \beta \to 0. \tag{3.45}$$

Proof. The first statement is immediate by Definition 3.3. The second holds since $v(.)$ has a unique local minimum at x_*, and $v(.) \in C^1(\mathcal{N}_b)$ with $||\nabla^{\gamma}_{M} v(x_*)|| = 0$; hence

$$\rho_\beta(x_*) := \sup\{d_M(x, x_*); ||\nabla^{\gamma}_{M} v(x)|| < \beta, x \in M\}, \tag{3.46}$$

where $d(.,.)$, the geodesic distance on M, is such that $\rho_\beta(x_*) \to 0$ as $\beta \to 0$. Hence by Theorem 3.3, $x^{k(\beta)}(\beta) \to x_* \quad as \quad \beta \to 0$. □

3.2.2. Simulation Results. Here we simulate the GG-HMP for a simple example and compare the results with the HMP algorithm in [**66**]. Consider the following Hybrid system which consists of two different phases

$$S_1 \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} u, \tag{3.47}$$

$$S_2 \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} u, \tag{3.48}$$

where

$$x_0 = (2, 1, 4), \quad x_f = (4, 1, 3), \quad J = \frac{1}{2}\int_0^{10} u^2(t)dt, \tag{3.49}$$

and the switching manifold M is considered to be the time invariant plane $m(x, y, z) = x + y - z = 0$ and the geodesic curves on M are straight lines. Figure 3.1 shows the convergence rate of the HMP and GG-HMP algorithms. As is evident, in this

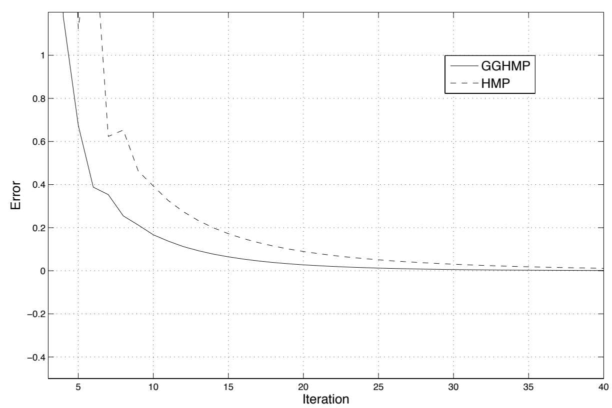

FIGURE 3.1. HMP and GG-HMP Convergence Rates

example the GG-HMP converges faster than the HMP to the optimal switching state and switching time. It is entirely likely that this is because the switching and gradient evaluation points are all on the switching surface and they are connected through the shortest path by the GG-HMP. Figure 3.2 shows the state space evolution of the example above derived by the optimal hybrid control; the optimal switching time is 5.5678s and the optimal switching state is $(-0.3801, -0.7268, -1.1069)$. Both GG-HMP and HMP start from an initial switching point $(0, 2, 2)$.

3.3. GG-HMP Algorithm Along Local Parameterizations

In this section we present a simplified version of the GG-HMP algorithm in order to reduce the computational load. As stated in the second step of the GG-HMP algorithm, the updating equations require a solution of (3.16) on the given manifold M. In general solving (3.16) imposes a significant computational load and slows down the computational rate. In order to reduce this complexity, we propose an algorithm which searches for the critical values of the value function in the local coordinates of the switching manifold [**71**].

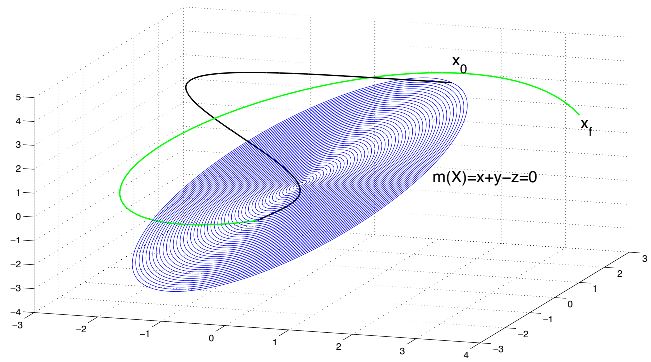

FIGURE 3.2. A Switching Manifold and the Corresponding Hybrid State Trajectory

Consider an n dimensional smooth differentiable manifold M and $x \in M$ for which the local coordinate chart and the associated neighbourhood are denoted as (x^{-1}, V); by their definitions x and x^{-1} are continuous maps which are furthermore assumed to be smooth:

$$x^{-1} : V \subset M \to R^n, \quad x : W \subset R^n \to M \subset R^{n+1}, \tag{3.50}$$

where V and W are open sets in M and R^n respectively. The mapping x is called the local parametrization of the switching manifold M. By the results of elementary differential geometry, the tangent space at the point $x \in M$ is spanned by $(\frac{\partial x}{\partial x_1}, ..., \frac{\partial x}{\partial x_n})$ or equivalently $(\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n})$. The hybrid value function minimization problem in the local coordinates of the switching manifold converts to an unconstrained optimization problem and we have the following lemma:

LEMMA 3.3. *Consider the hybrid system defined in Theorem 3.1 with the associated time invariant switching manifold M, where by assumption the optimal switching state and time (x^o, t^o) is an interior point of $M \times R$. Then, in the local coordinates of the optimal switching state and switching time,*

$$\frac{\partial v(x,t)}{\partial x_i}|_{x=x^o} = 0, \ \frac{\partial v(x,t)}{\partial t}|_{t=t^o} = 0, \ \ i = 1, ..., n. \tag{3.51}$$

Proof. Applying the chain rule in the ordinary differential equations implies

$$\frac{\partial v(x,t)}{\partial x_i}|_{x=x^o} = \frac{\partial v(x,t)}{\partial x}|_{x=x^o}\frac{\partial x}{\partial x_i}.$$

By the results of Lemma 3.1, $\frac{\partial v(x,t)}{\partial x}|_{x=x^o}$ is normal to the tangent space of the switching manifold at x^o, which proves the first statement. The second statement holds since the optimization with respect to the switching time is unconstrained. □

In this chapter we assume that the optimization process is performed in an open subset of the switching manifold which is covered by a single chart. Hence the parametrization is unchanged during the optimization process. The resulting *GG-HMP algorithm along local parameterizations*, (denoted as the GGAP-HMP Algorithm), for a time varying switching manifold which is considered locally as $m(x,t) = 0$, has the following specifications:

Definition 3.4. *GGAP-HMP (Multiple Autonomous Switchings) Algorithm*

For a hybrid system with one switching manifold:

(1) Initialize the switching state x_s^k *and switching time* t_s^k *on the time varying switching manifold* $m(x,t)$ *and compute* $g(\nabla_M^\gamma v(x_s^k,t_s^k), \frac{\partial}{\partial x_i^k}) =< \nabla_M^\gamma v(x_s^k,t_s^k), \frac{\partial}{\partial x_i^k} >= \nabla_{x_i} v(x^k), \quad i = 1,...,n+1,$ *where* $(x_1^k,...,x_n^k)$ *are local coordinates for* x_s^k, $x_{n+1}^k = t_s^k$ *and* $\nabla_x v(x^k) = [\nabla_{x_1} v(x^k),...,\nabla_{x_{n+1}} v(x^k)]^T$.

(2) Update the local coordinates of (x_s^k,t_s^k) *by the following equation:*

$$x_i^{k+1} = x_i^k - \tau^k g(\nabla_M^\gamma v(x_s^k,t_s^k), \frac{\partial}{\partial x_i^k}), i = 1,...,n+1, \tag{3.52}$$

which is equivalent to

$$x_i^{k+1} = x_i^k - \tau^k(\lambda^+(t_s^k) - \lambda^-(t_s^k)).\frac{\partial x}{\partial x_i}|_{x=x^k}, i = 1,...n, \tag{3.53}$$

$$t_s^{k+1} = t_s^k - \tau^k(H_1(x_s^k, \lambda^-(t_s^k)) - H_2(x_s^k, \lambda^+(t_s^k))), \tag{3.54}$$

and

$$\tau^{k+1} = \sup_s\{\frac{dv(x(t))}{dt} \leq 0\}, t \in [\tau^k, s), s \in [\tau^k, \tau^k + 1). \tag{3.55}$$

(3): If $||\nabla_x v(x^k)|| < \beta$, where β is a predefined bound then stop, otherwise go to step (1) with the next initial state (x_s^{k+1}, t_s^{k+1}). □

The proof for the convergence of the continuous version of the GGAP-HMP is same as the proof of Theorem 3.3. Since the optimization problem is unconstrained in the local coordinates of x, by Lemma 3.3 the updating step for (x_s^k, t_s^k) are given by (3.53),(3.54). It should be noted that the geodesic curves are straight lines in R^n ([**37**]).

3.3.1. Simulation Results. We simulate the GGAP-HMP algorithm for the given hybrid system in the first example with the switching manifold $m(x, y, z) = x^2 + y^2 - z = 0$. As is obvious from Figure 3.3, the convergence rate of the GGAP-

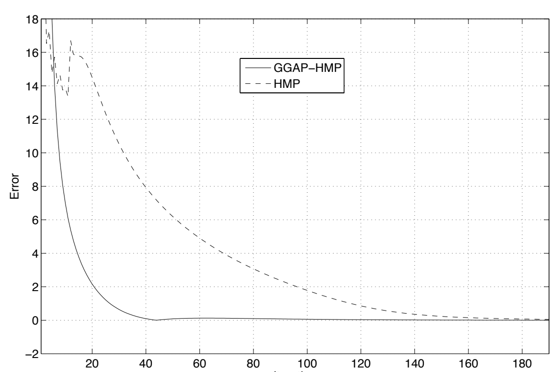

FIGURE 3.3. HMP and GGAP-HMP Convergence Rates

HMP algorithm is significantly faster than the HMP algorithm for the given example. The state trajectory is shown in Figure 3.4. The optimal switching state and switching time are $(0.0886, 0.1428, 0.0282)$ and 4.88s respectively. Both the HMP and GGAP-HMP start with the initial switching state $(2, 2, 8)$ on the switching manifold.

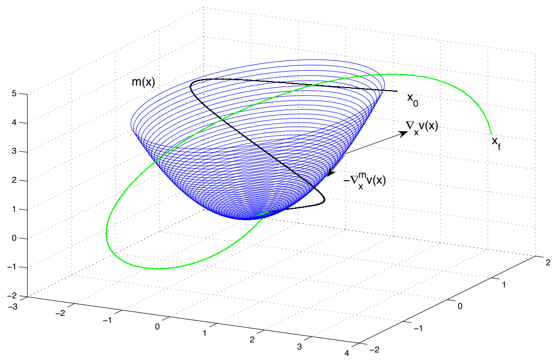

FIGURE 3.4. A Switching Manifold and the Corresponding Hybrid State Trajectory

Remark: Following [**37**], we see that geodesic curves in M are given in a special coordinates system (*normal coordinates*) for which they satisfy the differential equations $\ddot{x}_i(0) = 0, i = 1, ..., n$.

3.4. NG-HMP Algorithm

3.4.1. Formulation and Analysis of the NG-HMP Algorithm. In this section we define a version of Newton's method along geodesics (*NG-HMP*) for hybrid systems as a search algorithm to find the critical points of the hybrid value function v. The update equation based on the standard Newton method for a function $h(.) \in C^2(R^n)$ is as follows:

$$x^{k+1} = x^k - (\tilde{H}_k^{-1})\nabla_x h(x^k), \tag{3.56}$$

where $\tilde{H}_k$ denotes the (assumed nonsingular) Hessian matrix of h.

In the general case of Reimannian manifolds, it is not possible to define a Hessian matrix as can be done in Euclidean spaces, (see e.g. [**29**], [**67**]). However, employing the *Levi-Civita* connection ∇ on a Riemannian manifold, the Hessian $\tilde{H}$ may be defined as a bilinear symmetric form, [**37**]. We recall for all $f, l \in C^\infty(M)$ and

$X, Y, Z \in \mathcal{X}(M)$, where $\mathcal{X}(M)$ is the space of smooth vector fields on M, the *Levi-Civita* connection on M with respect to the Riemannian metric g is uniquely specified by the following axioms:

$$\nabla_X f = X(f), \quad f \in C^\infty(M), \tag{3.57}$$

$$\nabla_X fY = X(f)Y + f\nabla_X Y, \tag{3.58}$$

$$\nabla_{fX+lZ} Y = f\nabla_X Y + l\nabla_Z Y, \tag{3.59}$$

$$\nabla_X Y - \nabla_Y X = [X, Y], \tag{3.60}$$

$$Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z). \tag{3.61}$$

For a vector field $X \in \mathcal{X}(M)$ the covariant derivative is defined as $\nabla_X : TM \to TM$, where

$$\nabla_X(Y)|_x = \nabla_{X(x)} Y, \quad Y \in \mathcal{X}(M). \tag{3.62}$$

The Taylor expansion on a Riemannian manifold for the value function v is then given as follows ([**67**]):

$$\begin{aligned} & v(exp_x \theta X) = v(x) + \theta(\nabla_X v)(x) + ... + \frac{\theta^{n-1}}{(n-1)!} \times \\ & (\nabla_X^{n-1} v)(x) + \frac{\theta^n}{(n-1)!} \int_0^1 (1-t)^{n-1} \times \\ & (\nabla_X^n v)(x)(exp_x t\theta X)dt, \quad 0 < \theta < \theta^*, \end{aligned} \tag{3.63}$$

which is equivalent to

$$v(exp_x\theta X) = v(x) + \theta(dv(X))|_x + ... + \frac{\theta^{n-1}}{(n-1)!} \times$$
$$(\nabla_X^{n-2} dv)(X)|_x + \frac{\theta^n}{(n-1)!}\int_0^1 (1-t)^{n-1} \times$$
$$(\nabla_X^{n-1} dv)(X)|_x(exp_x t\theta X)dt, \quad 0 < \theta < \theta^*, \tag{3.64}$$

where dv is the differential one form of v. As before we assume that x_* is a strict local minimum of v on M, then (3.63) implies

$$(\nabla_X v)|_{x_*} = 0, \quad 0 < (\nabla_X^2 v)|_{x_*}, \quad \forall X \in M, \tag{3.65}$$

where

$$\nabla_X \nabla_Y v = \nabla_X dv(Y) =: \tilde{H}_v(X, Y), \quad \forall X, Y \in T_x M, \tag{3.66}$$

and $\tilde{H}_v(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})$ may be obtained from (3.68) below. The following lemma gives the covariant derivative of the one form dv:

Lemma 3.4 ([**37**]). *On a smooth Riemannian manifold M, for all C^∞ one forms $dx_i, i = 1, ..., n$, the covariant derivative $\nabla_{\frac{\partial}{\partial x_j}} dx_i$ is given as follows:*

$$\nabla_{\frac{\partial}{\partial x_j}} dx_i = -\sum_{i,j=1}^{n} \Gamma_{i,j}^k dx_k, \tag{3.67}$$

where $\Gamma_{i,j}^k$ are introduced in (3.17). □

Then Lemma 3.4 and (3.66) together imply

$$\tilde{H}_v = \sum_{i,j=1}^{n} \Big((\frac{\partial^2 v}{\partial x_i \partial x_j}) - \sum_{k=1}^{n} \Gamma_{j,i}^k \frac{\partial v}{\partial x_k} \Big) dx_i \otimes dx_j \in T_x^* M \otimes T_x^* M. \tag{3.68}$$

The following lemma is essential for the proof of the local convergence of the version of Newton's method presented below.

LEMMA 3.5. *For a strict local minimum x_* of the value function v, there exists a neighbourhood $\mathcal{N}_{x_*}$ such that $0 < \tilde{H}_v(X,X)|_{x\in\mathcal{N}_{x_*}}$ for all $X \in \mathcal{X}(M), X \neq 0$.*

PROOF. As in (3.65, 3.66), x_* a strict local minimum implies that $0 < \tilde{H}_v(X,X)|_{x_*}$; the result then follows from the smoothness of both v and the Riemannian metric g on M. □

For any given $X, Y \in \mathcal{X}(M)$, by (3.65) and (3.66), $\nabla dv(X) = \tilde{H}_v(X,.)$ which induces an isomorphism between T_xM and T_x^*M, where $\nabla dv : \mathcal{X}(M) \to \mathcal{X}^*(M)$, and $\mathcal{X}^*(M)$ is the space of all smooth covectors defined on M. (The functionality, one to one and onto properties of ∇dv can be verified from (3.65) and (3.66).) Therefore $\tilde{H}^{-1} : T_x^*M \to T_xM$. The Newton recursion along geodesics in M is given by the following update equation:

$$x^{k+1} = exp_{x^k}(-\theta \tilde{H}_k^{-1} dv|_{x^k}),\ x^1 = x_0, 1 \leq k < \infty. \tag{3.69}$$

Similar to the flow defined for the GG-HMP algorithm, we define a flow for the Newton HMP algorithm on the switching manifold M as follows:

DEFINITION 3.5. *(The Newton-Geodesic Flow) Let $\theta^0 = 0$, and $x(\theta^0) = x^0 \in M$, then for all $0 \leq k < \infty$ and all x^k such that $\tilde{H}_k^{-1}$ is nonsingular, define*

$$x(\theta) = \gamma_{x^k}(\theta) = exp_{x^k}(-\theta \tilde{H}_k^{-1} dv|_{x^k}),$$
$$\theta \in [\theta^k, \theta^{k+1}), \quad x(\theta) \in M, \quad 1 \leq k < \infty \tag{3.70}$$

where

$$\theta^{k+1} = \sup_s \{s; \frac{dv(x(t))}{dt} \leq 0, t \in [\theta^k, s), s \in [\theta^k, \theta^k + 1)\}, \tag{3.71}$$

where $\gamma_{x^k}(s)$ is the geodesic emanating from x^k with velocity $-\tilde{H}_k^{-1} dv|_{x^k}$. □

The switching state in (3.70) is defined as

$$x^{k+1} = x(\theta^{k+1}), \quad 0 \leq k < \infty, \tag{3.72}$$

and is the origin of the subsequent geodesic segment unless $\tilde{H}_k^{-1} dv|_{x^k} = 0$, in which case $x^{k+1} = x(\theta^{k+1})$ is defined to be a finite iteration equilibrium point of the flow.

Over the interval of existence $[0, \omega)$ we denote the total flow induced by (3.69)) as

$$\varphi(\theta, x^0) = \Pi_{i=1}^{n} \psi_i(\theta^{i-1}, \theta^i, x^{i-1}) \circ \psi_n(\theta^n, \theta, x^n), \tag{3.73}$$

where

$$\psi_i(\theta^{i-1}, \theta^i, x^{i-1}) = \gamma_{x^{i-1}}(\theta^i - \theta^{i-1}), \;\; \gamma_{x^0}(\theta^1 - \theta^0) = \gamma_{x^0}(\theta^1), \theta^0 = 0, \tag{3.74}$$

$\theta^i - \theta^{i-1}$ is the elapsed time between the switching times θ^i, θ^{i-1} to the next iteration and n is the index number of the last switching before the instant θ. By the continuity of the geodesic flows $\{\psi_i, 1 \leq i < \infty\}$, φ is a continuous map on $[0, \omega)$.

H2: There exists $0 < b < \infty$ such that the associated sublevel set $\mathcal{N}_b = \{x \in M; \quad v(x) < b\}$ is (i) open (ii) connected, (iii) contains a strict local minimum x_* which is the only local minimum in $\mathcal{N}_b$, (iv) $\mathcal{N}_b$ has compact closure and (v) $\mathcal{N}_{x_*} \subset \mathcal{N}_b$.

Without loss of generality, we assume $\mathcal{N}_{x_*} \subset \mathcal{N}_{b-\epsilon}$ for some $\epsilon > 0$, then by selecting $x^0 \in \mathcal{N}_{x_*} \subset \mathcal{N}_{b-\epsilon} \subset \mathcal{N}_b$ we prove $\omega = \infty$ by the following lemma:

LEMMA 3.6. *For an initial state $x^0 \in \mathcal{N}_{x_*}$, the existence interval of the flow defined in (3.73) is unbounded.*

PROOF. By ***H2*** we have $\mathcal{N}_{x_*} \subset \mathcal{N}_{b-\epsilon}$. Choose $0 < \theta < \theta'$ then if θ is not a switching time by the construction of ϕ, i.e. (3.71)

$$v(\varphi(\theta', x^0)) \leq v(\varphi(\theta, x^0)) \leq v(x^0) < b - \epsilon < b. \tag{3.75}$$

We need to prove the statement above when θ is a switching time. The derivative from the right of the flow φ, which is the combination of the flows defined in (3.70) at the switching state x^k, is given by

$$\frac{dv(x^k(\theta))}{d\theta}|_{\theta=0} = dv(-\tilde{H}_k^{-1}dv|_{x^k}) = -dv(\tilde{H}_k^{-1}dv|_{x^k}). \tag{3.76}$$

By the definition of $\tilde{H}$ we have

$$\nabla_X dv = \tilde{H}(.,X), \quad X \in \mathcal{X}(M), \tag{3.77}$$

therefore $\nabla_{-\tilde{H}^{-1}dv}dv = -\nabla_{\tilde{H}^{-1}dv}dv = \tilde{H}(.,-\tilde{H}^{-1}dv) = -dv$, as may be seen by evaluating the expression on any $Y \in T_xM$. Hence

$$-dv(\tilde{H}_k^{-1}dv|_{x^k}) = -\nabla_{\tilde{H}_k^{-1}dv}dv(\tilde{H}_k^{-1}dv)|_{x^k} = -\tilde{H}_k(\tilde{H}_k^{-1}dv, \tilde{H}_k^{-1}dv)|_{x^k}, \tag{3.78}$$

where the last equality holds by (3.77). By ***H2*** and Lemma 3.5 we have

$$\frac{dv(x^k(\theta))}{d\theta}|_{\theta=0} = -\tilde{H}_k(\tilde{H}_k^{-1}dv, \tilde{H}_k^{-1}dv)|_{x^k} \leq 0. \tag{3.79}$$

It follows by the construction of φ in (3.73), that for all $0 < \theta < \theta'$,

$$v(\varphi(\theta', x^0)) \leq v(\varphi(\theta, x^0)) \leq v(x^0) < b - \epsilon < b, \tag{3.80}$$

and hence for $\Phi^+ := \{\varphi(\theta, x^0); 0 \leq \theta < \omega\}$

$$\overline{\Phi^+} \subset \overline{\mathcal{N}}_{b-\epsilon} \subset \mathcal{N}_b. \tag{3.81}$$

So the flow φ is defined everywhere in $\overline{\mathcal{N}_{b-\epsilon}}$, where $\mathcal{N}_b$ has compact closure. Hence for all $x \in \mathcal{N}_{b-\epsilon}$ we have an extension of φ in $\mathcal{N}_b$, therefore the maximum interval of existence of $\varphi(., x^0)$ in $\mathcal{N}_b$ is infinite.

□

THEOREM 3.5. *Subject to the hypothesis **H2** on $\mathcal{N}_b$ and with an initial state x^0 such that $x^0 \in \mathcal{N}_{x_*} \subset M$, either the Newton geodesic flow, φ, reaches an equilibrium after a finite number of switchings, or it satisfies*

$$\varphi(\theta, x^0) \to \Omega(x^0) \subset v^{-1}(c), \quad c \in R, \tag{3.82}$$

as $\theta \to \infty$, for some $c \in R$, where

$$\forall y \in \Omega(x^0), \quad \frac{dv(y)}{d\theta}|_{\theta=0} = 0, \tag{3.83}$$

and, furthermore, the switching sequence $\{x\}_0^\infty = \{x^0, x^1, \cdots, \}$ converges to the limit point $x_* \in \Omega(x^0) \subset \mathcal{N}_b$, where x_* is the unique element of $\mathcal{N}_{x_*}$ such that $\nabla_M^\gamma v(x_*) = 0$.

Proof. This follows by the same argument as in the proof of Theorem 3.3. □

3.4.2. Convergence Rate of the NG-HMP Algorithm. In this section we discuss the convergence rate of the NG-HMP algorithm based upon the analysis given in [**67**]. As is shown in [**43**], for a given curve $\gamma : I \to M$ and a tangent vector $V_0 \in T_{\gamma(t_0)}M$, there exists a unique parallel vector field V along γ such that $V(t_0) = V_0$. This parallel translation defines a linear isomorphism $P_{t_0 t_1} : T_{\gamma(t_0)}M \to T_{\gamma(t_1)}M$. The corresponding dual map is denoted by $P^*_{t_0 t_1} : T^*_{\gamma(t_1)}M \to T^*_{\gamma(t_0)}M$.

Consider the map $P^*_{t_0(.)}dv : R \to T^*_{\gamma(t_0)M}$, where $dv \in T^*_{\gamma(t)M}$, then we can apply the Taylor expansion on $P^*_{t_0(.)}dv(V)$ where $V \in T_{\gamma(t_0)}M$. For a given $X \in T_{x_0}M$ we set $x(\theta) = x(\theta, x_0) = exp_{x_0}(\theta X)$ then for a differential form $dv_{x_\theta} \in T^*_{exp_x(\theta X)}M$ we have

$$\begin{aligned} P^*_\theta dv_{x_\theta} &= dv_x + \theta(\nabla_X dv)_x + ... + \frac{\theta^{n-1}}{(n-1)!}(\nabla_X^{n-1} dv)_x \\ &+ \frac{\theta^n}{(n)!}(\nabla_X^n dv)_{x_\alpha} \circ P_\alpha, \end{aligned} \tag{3.84}$$

where $\alpha \in [0, \theta]$ and without loss of generality we assume $t_0 = 0$. Let us consider x^k, x^{k+1} for a sufficiently large k. As proved in Theorem 3.5, x^k converges to x_*, the unique local minimum of $v(.)$ in $\mathcal{N}_b$, hence $\lim_{k\to\infty} d(x^k, x^{k+1}) = 0$, where d is the minimum length of the curves connecting x^k and x^{k+1}. Employing the *Normal Neighbourhood Lemma* in [**43**], we may select k large enough such that $x^k, x^{k+1}, x_* \subset \mathcal{U}_k$, where $\mathcal{U}_k$ denotes a normal neighbourhood corresponding to x^k, see [**37**]. Let $\theta_k = \theta^{k+1} - \theta^k$, the running time of the geodesic flow from x^k to x^{k+1}, where $x^{k+1} =$

$exp_{x^k}(-\theta_k \tilde{H}_k^{-1} dv|_{x^k})$. Since $x^{k+1}, x_* \subset \mathcal{U}_k$

$$\exists X_k \in T_{x^k}M, X_{k+1} \in T_{x^{k+1}}M, \quad s.t \quad x_* = exp_{x^k}(X_k),$$
$$x_* = exp_{x^{k+1}}(X_{k+1}), \tag{3.85}$$

where the existence of X_{k+1} is guaranteed since x^{k+1} is close enough to x_*.

As is shown in [**67**], we have

$$X_k = -\theta_k \tilde{H}_k^{-1} dv|_{x^k} + P_{\theta_k}^{-1} X_{k+1} + \mathcal{E}, \tag{3.86}$$

where $\mathcal{E} \in T_{x^k}M$ is the additional third order displacement term ([**67**]). By applying $\tilde{H}$ on both sides of (3.86) we have

$$\tilde{H}_k(P_{\theta_k}^{-1} X_{k+1}) = \theta_k dv|_{x^k} + \tilde{H}_k(X_k) - \tilde{H}_k(\mathcal{E}). \tag{3.87}$$

The Taylor expansion of dv_x at x_* gives

$$dv_{x^k} = -\tilde{H}_k(X_k) - \frac{1}{2}\nabla^2_{\tilde{X}_k} dv_{x^k} \circ P_\alpha, \quad \alpha \in [0,1], \tag{3.88}$$

where inserting (3.88) into (3.87) yields

$$\tilde{H}_k(P_{\theta_k}^{-1} X_{k+1}) = \tilde{H}_k(X_k)(1-\theta_k)\frac{1}{2}\nabla^2_{\tilde{X}_k} dv_{x^k} \circ P_\alpha - \tilde{H}_k(\mathcal{E}). \tag{3.89}$$

As is shown in [**67**], the smoothness of v and g gives

$$\delta_1 ||X_k|| \leq ||\tilde{H}_k(X_k)|| \leq \delta_2 ||X_k||,$$
$$||\nabla^2 dv_{x^k}(X_k, X_k)|| \leq \hat{\delta} ||X_k||^2, \quad 0 < \delta_1, \delta_2, \hat{\delta}, \tag{3.90}$$

for some $\delta_1, \delta_2, \hat{\delta}$. Therefore in the case where $\theta_k \neq 1$ we have a linear convergence rate as follows:

$$\exists \delta_k \in R_+ \quad s.t \quad d(x^{k+1}, x_*) \leq \delta_k d(x^k, x_*). \tag{3.91}$$

In the case $\theta_k = 1$, a quadratic convergence rate can be obtained from (3.88) and (3.90).

3.4.3. NG-HMP Algorithm for Embedded Surfaces in R^{n+1}. In order to implement the Newton Geodesic algorithm, the second derivative terms in (3.68) must be computed. In this subsection a system of differential equations is introduced in order to compute these quantities. Let us define the state and the adjoint variations in a local coordinate system of the switching state for a hybrid system consisting of two distinct phases as follows:

$$y_i^x(t) = lim_{\delta x_i \to 0} \frac{\delta x(t)}{\delta x_i}, \qquad i = 1, ..., n, t \in [0, t_f], \tag{3.92}$$

$$z_i^x(t) = lim_{\delta x_i \to 0} \frac{\delta \lambda(t)}{\delta x_i}, \qquad i = 1, ..., n, t \in [0, t_f], \tag{3.93}$$

where $\delta x, \delta\lambda$ are state and adjoint variations with respect to the variation of the ith component of the local coordinate of $x(t_s)$. In [**71**] it is shown that $y_i^x(t), z_i^x(t)$ satisfy the following differential equations (see Appendix B):

$$\dot{y}_i^x(t) = \frac{\partial f_1}{\partial x} y_i^x(t) + \frac{\partial f_1}{\partial \lambda} z_i^x(t), \qquad t \in [t_0, t_s), \tag{3.94}$$

$$\dot{z}_i^x(t) = -\frac{\partial^2 H_1}{\partial x^2} y_i^x(t) - \frac{\partial^2 H_1}{\partial \lambda \partial x} z_i^x(t), \quad t \in [t_0, t_s), \tag{3.95}$$

$$\dot{y}_i^x(t) = \frac{\partial f_2}{\partial x} y_i^x(t) + \frac{\partial f_2}{\partial \lambda} z_i^x(t), \qquad t \in [t_s, t_f] \tag{3.96}$$

$$\dot{z}_i^x(t) = -\frac{\partial^2 H_2}{\partial x^2} y_i^x(t) - \frac{\partial^2 H_2}{\partial \lambda \partial x} z_i^x(t), \quad t \in [t_s, t_f], \tag{3.97}$$

where $H_i, i = 1, 2$ is the associated Hamiltonian function corresponding to f_1 and f_2. For a fixed end point optimal hybrid trajectory the boundary conditions for (3.92) and (3.93) are given as follows:

$$y_i^x(0) = y_i^x(t_f) = 0, \quad y_i^x(t_s) = \frac{\partial x}{\partial x_i}|_{x(t_s)}. \tag{3.98}$$

The following theorem gives the relations between the second order variation of the hybrid value function v and geometrical properties of the switching manifold involving the second fundamental forms.

THEOREM 3.6. *At the optimal switching state x_s^o on the switching manifold M, and at the switching time t_s^o the following holds:*

$$\begin{aligned} -H_{ik} &= \mu\frac{\partial x^T}{\partial x_i}\frac{\partial^2 v(x_s^o,t_s^o)}{\partial x^2}\frac{\partial x}{\partial x_k}+T_i\frac{\partial x}{\partial x_k} \\ &= \mu\frac{\partial x^T}{\partial x_k}\frac{\partial^2 v(x_s^o,t_s^o)}{\partial x^2}\frac{\partial x}{\partial x_i}+T_k\frac{\partial x}{\partial x_i}, i,k=1,...,n, \end{aligned} \tag{3.99}$$

*where $T_i, T_k \in T_{x^o}M$, μ is the discontinuity parameter appearing in the adjoint process boundary condition at the switching time (see [**58, 66**]) and H_{ik} is the second fundamental form of the switching manifold at x^o, see [**40**].*

PROOF. A proof is given in Chapter 5 (Theorem 5.3). □

Based upon the relations given in Theorem 3.6, the following corollary is established:

COROLLARY 3.1. *In the local coordinates of the optimal switching state x_s^o we have*

$$\frac{\partial x^T}{\partial x_i}\frac{\partial^2 v(x_s^o,t_s^o)}{\partial x^2}\frac{\partial x}{\partial x_k}=\frac{\partial^2 v(x_s^o,t_s^o)}{\partial x_i\partial x_k}-\mu^{-1}H_{ik}, \;\; i,k=1,...,n. \tag{3.100}$$

PROOF. A proof is given in Chapter 5 (Corollary 1). □

The second variation of the hybrid value function v at the optimal switching state x_s^o and switching time t_s^o is given in the following lemma by using the results derived in Theorem 3.6.

LEMMA 3.7. *The local Hessian matrix components of the value function of the hybrid system at the optimal switching state x_s^o satisfy the following equations for*

$i,k=1,...,n$:

$$\begin{aligned}\frac{\partial^2 v(x_s^o,t_s^o)}{\partial x_i \partial x_k} &= y_i^x(t)(z_k^x(t_s^{o^+})-z_k^x(t_s^{o^-}))+\mu^{-1}H_{ik} \\ &= y_k^x(t^o)(z_i^x(t_s^{o^+})-z_i^x(t_s^{o^-}))+\mu^{-1}H_{ki}\end{aligned} \tag{3.101}$$

where $(z_k^x(t_s^{o^+})-z_k^x(t_s^{o^-}))$ is the discontinuity of the z solution of (3.95),(3.97).

Proof. A proof is given in Chapter 5 (Lemma 5.4). □

For an arbitrary given switching state x which is not necessarily optimal, the local Hessian matrix of the hybrid value function is given as follows:

Lemma 3.8 ([**73**]). *The local Hessian matrix of the value function of the hybrid system at a non-optimal state x_s and the switching time t_s satisfies the following equations for $i,k=1,...,n$:*

$$\frac{\partial^2 v(x_s,t_s)}{\partial x_i \partial x_k} = y_i^x(t)(z_k^x(t_s^+)-z_k^x(t_s^-))+(\lambda^+(t_s)-\lambda^-(t_s))\frac{\partial^2 x}{\partial x_i \partial x_k}, \tag{3.102}$$

where $(z_k^x(t_s^+)-z_k^x(t_s^-))$ is the discontinuity of the z solution of (3.95),(3.97).

Proof. The proof parallels the proof of Lemma 3.7 where the analysis is performed at a generic switching state x_s. □

The following lemma gives the second order variation of the hybrid value function v with respect to the local coordinates of the switching state $x^i, i=1,...,n$ and the switching time t_s.

LEMMA 3.9 ([**73**]). *The components* $\frac{\partial^2 v(x,t)}{\partial x_i \partial t_s}$, $\quad i = 1, ..., n$ *of the Hessian matrix* $H(v(x_s, t_s)) = \begin{pmatrix} \frac{\partial^2 v(x_s,t_s)}{\partial x_i^2} & \frac{\partial^2 v(x_s,t_s)}{\partial x_i \partial t_s} \\ \frac{\partial^2 v(x_s,t_s)}{\partial t_s \partial x_i} & \frac{\partial^2 v(x_s,t_s)}{\partial t_s^2} \end{pmatrix}$ *is computed as*

$$\begin{aligned} \frac{\partial^2 v(x_s, t_s)}{\partial x_i \partial t_s} = & \frac{\partial H_1(t_s^-)}{\partial x} y_i^x(t_s^-) - \frac{\partial H_2(t_s^+)}{\partial x} y_i^x(t_s^+) + \\ & \frac{\partial H_1(t_s^-)}{\partial \lambda} z_i^x(t_s^-) - \frac{\partial H_2(t_s^+)}{\partial \lambda} z_i^x(t_s^+), \end{aligned} \tag{3.103}$$

where H_1 *and* H_2 *are the corresponding Hamiltonians of the hybrid phases before and after switching time respectively.*

PROOF. The proof parallels the proof of Lemma 3.8 where the analysis is performed at a generic pair of switching state x_s and switching time t_s. □

Similar to the variations defined in (3.94) and (3.96), the state and adjoint variations with respect to t_s are defined as follows:

$$y_s^t(t) := lim_{\delta t_s \to 0} \frac{\delta x(t)}{\delta t_s}, \quad t \in [t_0, t_f], \tag{3.104}$$

$$z_s^t(t) := lim_{\delta t_s \to 0} \frac{\delta \lambda(t)}{\delta t_s}, \quad t \in [t_0, t_f]. \tag{3.105}$$

The following lemma gives $\frac{\partial^2 v(x_s,t_s)}{\partial t_s^2}$ which appears in the Hessian of the hybrid value function.

LEMMA 3.10 ([**73**]). *The component* $\frac{\partial^2 v(x_s,t_s)}{\partial t_s^2}$ *of the Hessian matrix* $H(v(x_s, t_s))$ *satisfies*

$$\begin{aligned} \frac{\partial^2 v(x_s, t_s)}{\partial t_s^2} = & \frac{\partial H_1(t_s^-)}{\partial x} y_i^t(t_s^-) - \frac{\partial H_2(t_s^+)}{\partial x} y_i^t(t_s^+) + \\ & \frac{\partial H_1(t_s^-)}{\partial \lambda} z_i^t(t_s^-) - \frac{\partial H_2(t_s^+)}{\partial \lambda} z_i^t(t_s^+), \end{aligned} \tag{3.106}$$

where $z_s^t(t)$ *is the solution of*

$$\dot{y}_s^t(t) = \frac{\partial^2 H_i(x,\lambda)}{\partial x \partial \lambda} y_s^t(t) + \frac{\partial^2 H_i(x,\lambda)}{\partial^2 \lambda} z_s^t(t), \tag{3.107}$$

$$\dot{z}_s^t(t) = -\frac{\partial^2 H_i(x,\lambda)}{\partial^2 x} y_s^t(t) - \frac{\partial^2 H_i(x,\lambda)}{\partial x \partial \lambda} z_s^t(t), \tag{3.108}$$

where $y_s^t(t_0) = y_s^t(t_f) = 0, \quad y_s^t(t_s^-) - y_s^t(t_s^+) = f_2(u(t_s), x(t_s)) - f_1(u(t_s), x(t_s)).$

Proof. The proof parallels the proof of Lemma 3.8 where the analysis is performed at a generic switching time t_s. □

Definition 3.6. *NG-HMP (Newton-Geodesic-HMP) Algorithm*

For a hybrid system with one switching manifold:

(1) Initialize the switching state x_s^k *on the switching manifold* M *and the switching time* t_s^k *then compute* $\nabla_M^\gamma v(x_s^k, t_s^k)$, *where* $(x_1^k, ..., x_n^k)$ *are local coordinates for* x_s^k, $x_{n+1}^k = t_s^k$.

(2) Compute y_i^x, z_i^x *as the solution of (3.94) and (3.95) and* y_i^t *and* z_i^t *as the solution of (3.107) and (3.108). Compute* $H^k = \begin{pmatrix} \frac{\partial^2 v(x_s^k,t_s^k)}{\partial x^{i2}} & \frac{\partial^2 v(x_s^k,t_s^k)}{\partial x^i \partial t_s} \\ \frac{\partial^2 v(x_s^k,t_s^k)}{\partial t_s \partial x^i} & \frac{\partial^2 v(x_s^k,t_s^k)}{\partial t_s^2} \end{pmatrix}$ *using Lemmas 3.8-3.10.*

(3) Update the local coordinates of (x_s^k, t_s^k) *by the following equation:*

$$(x_s^{k+1}, t_s^{k+1}) = exp_{(x_s^k, t_s^k)}(-\tilde{H}_k^{-1} dv|_{x_s^k, t_s^k}), \tag{3.109}$$

where $\tilde{H}_k$ *is computed by (3.68) together with* H^k *which is the second order variation of the value function in (3.68).*

(4) If $||\nabla_x v(x^k)|| < \beta$, *where* β *is a predefined bound then stop, otherwise go to step (1) with the next initial state* (x_s^{k+1}, t_s^{k+1}). □

3.4.4. Simulation Results. Consider the following hybrid system which possesses the two phases:

$$S_1 : \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} -.01 & 1 & 1 \\ 0 & -.01 & 1 \\ 0 & 0 & -.01 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} u, \tag{3.110}$$

$$S_2 : \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} -.01 & 0 & 0 \\ 1 & -.01 & 0 \\ 1 & 1 & -.01 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} u, \tag{3.111}$$

with the initial and terminal conditions and the cost function given by

$$x_0 = (2,1,4), \quad x_f = (4,1,3), J = \frac{1}{2}\int_0^{10} u^2(t)dt, \tag{3.112}$$

The switching manifold M is taken to be $m(x,y,z) = x^2 + y^2 - z = 0$. Figure 3.6 shows the convergence rate of the NG-HMP and GG-HMP algorithms.

Simulations of the GG-HMP and NG-HMP algorithms for the given hybrid system with the switching manifold $m(x,y,z) = x^2 + y^2 - z = 0$ resulted in the state and cost trajectories shown in Figures 3.5 and 3.6 respectively. As is obvious from Figure

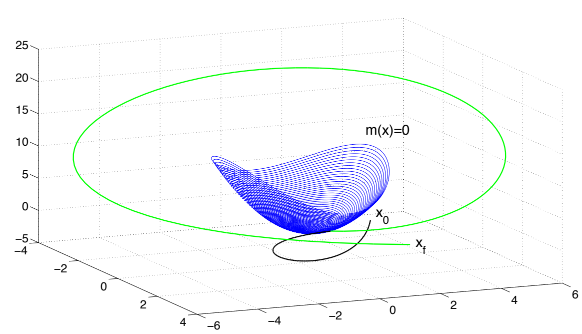

FIGURE 3.5. The Switching Manifold and the Corresponding Hybrid State Trajectory

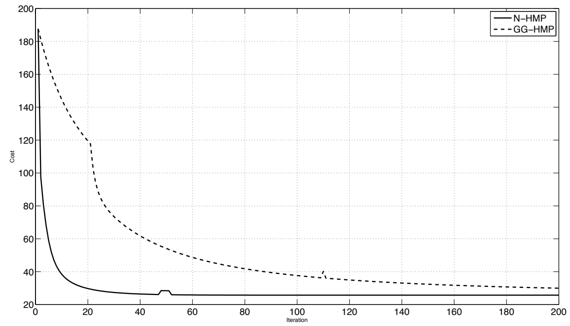

FIGURE 3.6. NG-HMP and GG-HMP Convergence

3.6, the convergence rate for the NG-HMP algorithm is significantly faster than the GGAP-HMP algorithm for the given example.

The optimal switching state and switching time generated by the algorithms were respectively

$(1.1787, 0.1837, 1.4232)$ and 2.5353s. Here both the NG-HMP and GGAP-HMP start with the same initial switching state $(2, 2, 8)$ on the switching manifold.

CHAPTER 4

The Hybrid Minimum Principle On Lie Groups

4.1. Control Systems on Lie Groups

In this section we introduce control systems on Lie groups and then extend the definition of hybrid systems to that of hybrid control systems defined on Lie groups (see [**76**, **78**]).

4.1.1. Lie Groups and Lie Algebras.

Definition 4.1. *A group $(G, \star)$ is called a Lie Group if, (see [**82**]):*
(1): G is a smooth manifold,
(2): The group operations are smooth.
(The group operations are associative multiplication and inversion, i.e. $\forall g_1, g_2, g_3, g \in G, \quad g_1 \star (g_2 \star g_3) = (g_1 \star g_2) \star g_3 \in G, g^{-1} \in G, g^{-1} \star g = g \star g^{-1} = e$). □

In this chapter it is assumed that the continuous part of the hybrid system evolves on a Lie group G.

Definition 4.2. *A Lie Algebra V is a real vector space endowed with a bilinear operation $[.,.] : V \times V \to V$ such that (see [**19**, **82**]):*

(1): $\forall \zeta, \eta \in V, \quad [\zeta, \eta] = -[\eta, \zeta]$

(2):$[\zeta, [\eta, \gamma]] + [\eta, [\gamma, \zeta]] + [\gamma, [\zeta, \eta]] = 0 \quad \forall \zeta, \eta, \gamma \in V$ □

The Lie algebra $\mathcal{L}$ of a Lie group G may be identified with the tangent space at the identity element e with the associated Lie bracket defined on the tangent space of G, i.e. $\mathcal{L} = T_e G$. A vector field X on G is called *left invariant* if

$$\forall g_1, g_2 \in G, \quad X(g_1 \star g_2) = TL_{g_1} X(g_2), \tag{4.1}$$

which immediately implies $X(g \star e) = X(g) = TL_g X(e)$ where $L_g : G \to G, \quad L_g(h) = g \star h$ where $TL_g : T_h G \to T_{g \star h} G$.

Definition 4.3. *Corresponding to a left invariant vector field X, we define the exponential map as follows:*

$$exp : \mathcal{L} \to G, \quad exp(tX(e)) := \Phi(t, X), t \in R, \tag{4.2}$$

□

where $\Phi(t, X)$ is the solution of $\dot{g}(t) = X(g(t))$ with the boundary condition $g(0) = e$. The following theorem gives the flow of a left invariant vector field with an arbitrary initial state $g \in G$.

Theorem 4.1 ([**82**]). *Let G be a Lie group with the corresponding Lie algebra $\mathcal{L}$, then for a left invariant vector field X*

$$\Phi(t, X, g) = L_g \circ exp(tX(e)), t \in R, \tag{4.3}$$

where $\Phi(t, X, g)$ is the flow of X starting at $g \in G$. □

A left invariant control system defined on a given Lie group G is defined as follows: (see [**19**, **21**, **38**])

$$\dot{g}(t) = f(g(t), u) = TL_{g(t)} f(e, u), \quad g(t) \in G, u \in R^u, \tag{4.4}$$

where $f(g(t), u)$ is a left invariant vector field on G. Similar to left invariant systems, right invariant systems are defined. In this chapter we only consider hybrid systems where the associated vector fields are left invarian, however the analysis can also be applied to right invariant hybrid systems.

4.1.2. Left Invariant Optimal Control Systems. A *left invariant optimal control system* is an optimal control problem where the ambient state manifold $\mathcal{M}$ is replaced by a Lie group G.

The corresponding vector field f_q is a left invariant vector field defined on G for any given $u \in \mathcal{U}$ such that

$$f_q(., u(.)) : G \times [t_0, t_f] \to TG, \tag{4.5}$$

and the cost function is defined as

$$J := \int_{t_0}^{t_f} l_q(g(s), u(s))ds, \quad u \in \mathcal{U}, \tag{4.6}$$

where $l_q(g(s), u(s))$ is assumed to be left invariant i.e. $l_q(L_h g(s), u(s)) = l_q(g(s), u(s))$. In general, a Bolza problem can be converted to a Mayer problem using an auxiliary state variable in the dynamics, see [**66**] and [**6**]. The following lemma shows the equivalence of a Bolza problem defined on a Lie group G and its Mayer extension.

LEMMA 4.1. *Consider a left invariant Optimal Control Problem (OCP) defined on a Lie group G with the following dynamics and cost function:*

$$\dot{g}(t) = f(g(t), u), \quad g(t) \in G, u \in R^u, \tag{4.7}$$

$$J = \int_{t_0}^{t_f} l(g(s), u(s))ds. \tag{4.8}$$

Then the Mayer problem associated to the optimal control problem above is defined on the Lie group $G \times R$ and the corresponding dynamics are left invariant.

PROOF. The state space equation of the Mayer problem concerning the Bolza problem is given as follows:

$$\begin{pmatrix} \dot{g} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} f(g(t), u(t)) \\ l(g(t), u(t)) \end{pmatrix} = F(\bar{g}(t), u), \tag{4.9}$$

where $\bar{g} = (g, z), \quad g \in G, z \in R$. The group action defined on $G \times R$ is given as follows:

$$(g_1, z_1)\bar{\star}(g_2, z_2) = (g_1 \star g_2, z_1 + z_2), \tag{4.10}$$

where $\star$ corresponds to the group action of G and $\bar{\star}$ is the group action of $G \times R$. Since G is a Lie group it follows that $(G \times R, \bar{\star})$ is also a Lie group. It only remains to show $F(\bar{g}, u)$ is left invariant. The left translation on $G \times R$ is defined by

$$L_{\bar{g}}(\bar{h}) = (L_g h, z_g + z_h), \quad \bar{g} = (g, z_g), \bar{h} = (h, z_h), \tag{4.11}$$

therefore

$$TL_{\bar{g}}F(\bar{h}, u) = TL_g f(h, u) \oplus l(L_g h, u) = f(g \star h, u) \oplus l(g \star h, u), \tag{4.12}$$

which yields $F(\bar{h}, u)$ is left invariant since $f(g, u)$ and $l(g, u)$ are both left invariant. □

4.2. Optimal Control Problems On Lie Groups

The optimal control problem on Lie groups has been addressed in [**17**,**18**,**38**,**39**]. In this section we give the Minimum Principle results presented in [**39**] for optimal control problems defined on a Lie group G. As shown in [**38**], the left translation gives an isomorphism between TG and $G \times \mathcal{L}$. Since $L_{g^{-1}}$ maps g to e, then $TL_{g^{-1}} : T_g G \to T_e G = \mathcal{L}$ is the corresponding isomorphism. This statement also holds between T^*G and $G \times \mathcal{L}^*$ where $\mathcal{L}^*$ is the dual space of the Lie algebra $\mathcal{L}$. The corresponding isomorphism is given by $T^*L_g : T_g^*G \to T_e^*G = \mathcal{L}^*$. We use the equivalence $T^*G \approx$

$G \times \mathcal{L}^*$ associated to the isomorphism above to construct Hamiltonian functions on Lie groups.

4.2.1. Hamiltonian Systems on $T^*\mathcal{M}$ and T^*G. By definition, for an optimal control problem defined on an n dimensional differentiable manifold $\mathcal{M}$, a Hamiltonian function is defined as a smooth function $H : T^*\mathcal{M} \times U \to R$, see [**3**,**38**]. The associated Hamiltonian vector field $\overrightarrow{H}$ is defined as follows (see [**3**]):

$$\sigma_\lambda(., \overrightarrow{H}) = dH, \quad \lambda \in T^*\mathcal{M}, \tag{4.13}$$

where σ is the symplectic form defined on $T^*\mathcal{M}$ which is locally written as follows:

$$\sigma = \sum_{i=1}^{n} d\zeta_i \wedge dx_i, \tag{4.14}$$

and (ζ, x) is the local coordinate representation of λ in $T^*\mathcal{M}$.

The Hamiltonian system of the ODE corresponding to H is

$$\dot{\lambda} = \overrightarrow{H}(\lambda), \tag{4.15}$$

where locally we have

$$\begin{cases} \dot{x}_i = \frac{\partial H}{\partial \zeta_i}, \quad i = 1, ..., n, \\ \dot{\zeta}_i = -\frac{\partial H}{\partial x_i}, \quad i = 1, ..., n. \end{cases} \tag{4.16}$$

Similar to Hamiltonian systems on smooth manifolds we can define Hamiltonian functions for left invariant vector fields on the cotangent bundle of a Lie group G. This is done by using the isomorphism between $\mathcal{L}^* \times G$ and T^*G introduced above which is denoted by $\mathcal{I}$:

$$(1) \quad \mathcal{I}(X, g) \in T^*G, \quad X \in \mathcal{L}^*, g \in G, \tag{4.17}$$

$$(2) \quad \mathcal{I}(., g) : \mathcal{L}^* \to T_g^*G \;\text{ is a linear isomorphism.} \tag{4.18}$$

A Hamiltonian function for a left invariant vector field X on G is defined as

$$H_X(g, \lambda) := \langle \lambda, X(e) \rangle = \langle \lambda, TL_{g^{-1}} X(g) \rangle, \quad \lambda \in \mathcal{L}^*. \tag{4.19}$$

The preceding identification realizes that $TT^*G \simeq T(G \times \mathcal{L}^*) = (G \times \mathcal{L}) \times (\mathcal{L}^* \times \mathcal{L}^*)$, therefore the tangent vector at $(g, \lambda) \in G \times \mathcal{L}^*$ is an element of $\mathcal{L} \times \mathcal{L}^*$ denoted by $T = (X, \gamma)$. The symplectic form σ along a given curve $\Gamma(t) \in T^*G$ satisfies the following equation, see [**3**, **38**]:

$$\sigma_\Gamma(T_1(\Gamma), T_2(\Gamma)) = \langle \gamma_2(t), X_1(t) \rangle - \langle \gamma_1(t), X_2(t) \rangle - \langle \lambda(t), [X_1(t), X_2(t)] \rangle, \tag{4.20}$$

where $T_i = (X_i, \gamma_i)$. Similar to Hamiltonian systems on $T^*\mathcal{M}$, the Hamiltonian vector field $\overrightarrow{H}$ on $G \times \mathcal{L}^*$ satisfies the following equation

$$dH = \sigma_{\Gamma(t)}(., \overrightarrow{H}). \tag{4.21}$$

The following theorem gives the Minimum Principle for optimal control problems defined on Lie groups.

THEOREM 4.2 ([**38**]). *For a left invariant optimal control problem defined by (4.5) and (4.6), along the optimal state and optimal control $g^o(t), u^o(t)$, there exists a nontrivial adjoint curve $\lambda^o(t) \in \mathcal{L}^*$ such that the following equations hold:*

$$H(g^o(t), \lambda^o(t), u^o(t)) \leq H(g^o(t), \lambda^o(t), u), \quad \forall u \in U, \tag{4.22}$$

and locally

$$\frac{dg^o}{dt} = TL_{g^o(t)}\left(\frac{\partial H}{\partial \lambda}\right), \tag{4.23}$$

$$\frac{d\lambda^o}{dt} = -(ad)^*_{\frac{\partial H}{\partial \lambda}}(\lambda^o(t)), \tag{4.24}$$

where $H(g, \lambda, u) := \langle \lambda, TL_{g^{-1}} f(g, u) \rangle$.

PROOF. The proof is by employing (4.21) and applying the symplectic form σ given by (4.20). A complete proof can be found in [**38**], Chapter 12, Theorem 1. □

For each $\eta, \zeta \in \mathcal{L}$ we have the following definition

$$ad_{\zeta} : \mathcal{L} \to \mathcal{L}, \ \ ad_{\zeta}(\eta) = [\zeta, \eta]. \tag{4.25}$$

For each $\zeta, \eta \in \mathcal{L}, \gamma \in \mathcal{L}^*$, ad^* is defined as

$$\langle ad^*_{\zeta}(\gamma), \eta\rangle := \langle \gamma, ad_{\zeta}(\eta)\rangle. \tag{4.26}$$

For more information about the definition above see [**2, 82**]. It should be noted that, in general, for a Hamiltonian function defined on $G \times \mathcal{L}^*$, the integral curve of the Hamiltonian vector field, i.e. (4.23) and (4.24), satisfies the following equations (see [**38**]):

$$\frac{dg}{dt} = TL_{g(t)}(\frac{\partial H}{\partial \lambda}), \tag{4.27}$$

$$\frac{d\lambda}{dt} = -T^*L_{g(t)}(\frac{\partial H}{\partial g}) - (ad)^*_{\frac{\partial H}{\partial \lambda}}(\lambda(t)). \tag{4.28}$$

In our framework since the Hamiltonian function is g invariant then $T^*L_{g(t)}(\frac{\partial H}{\partial g})$ does not appear in the statement of Theorem 4.2. Since the tangent space of T^*G is identified with $\mathcal{L} \times \mathcal{L}^*$, by the definition of the Hamiltonian $H : G \times \mathcal{L}^* \to R$, it is noted that $\frac{\partial H}{\partial \lambda} \in \mathcal{L}^{*^*} = \mathcal{L}$ and $\frac{\partial H}{\partial g} \in T^*_g G$.

4.3. Hybrid Systems on Lie Groups

The definition of hybrid systems on Lie groups is similar to that of hybrid systems given in Definition 2.1 for which the ambient manifold $\mathcal{M}$ is replaced by a Lie group G. Here we only consider a hybrid system consisting of two different phases with the associated left invariant vector fields f_1, f_2 as follows:

$$\dot{g}(t) = f_1(g(t), u(t)), \ \dot{g}(t) = f_2(g(t), u(t)), \ u(t) \in \mathcal{U}. \tag{4.29}$$

The switching manifold $\mathcal{N}$ associated to the autonomous phase change is considered to be a submanifold of G which is by definition a regular Lie subgroup. The hybrid cost function is defined as

$$J = \sum_{i=0}^{1} \int_{t_i}^{t_{i+1}} l_{q_i}(g_{q_{i+1}}(s), u(s))ds + h(g_{q_L}(t_f)),\, u \in \mathcal{U}, \tag{4.30}$$

where $l_i, i = 1, 2$ are left invariant smooth functions on G. The hybrid optimal control problem is to find the optimal switching state $g_s^o \in \mathcal{N}$, optimal switching time $t_s \in R$ and the associated optimal controls u_1^o and u_2^o in order to minimize the hybrid cost defined by (4.30). Here we assume the state variable for both dynamics evolve on the same Lie group G. Similar to the proof in [**75**], we apply the needle control variation in two different parts. First, the control needle variation is applied after the optimal switching time so there is no state propagation along the state trajectory through the switching manifold. Second, the control needle variation is applied before the optimal switching time. In this case there exists a state variation propagation through the switching manifold, see [**66**], [**73**]. With the assumption of accessibility of $\dot{g}(t) = f_{q_1}(g(t), u(t))$, for a hybrid system with one autonomous switching, we define the hybrid value function $v(g, t)$ same as the definition given in Chapter 2.

$$v(g, t) = inf_{u \in \mathcal{U}} J(t_0, t_f, h_0, u), \quad g \in G, t \in R, \tag{4.31}$$

where $g(t, t_0, g_0) = g \in G$. We use the value function v to explain the discontinuity of the adjoint process appears in the statement of the Hybrid Minimum Principle in the next sections.

4.4. Non-Interior Optimal Switching States

In general the hybrid value function for a Mayer type problem attains its minimum on the boundary of the attainable switching states on the switching manifold hence is not differentiable. In this case the discontinuity of the adjoint process in the HMP statement is given based on a normal vector at the switching time on the switching manifold. In order to have a normal vector N the switching manifold we need to define a Riemannian metric on G. A left invariant Riemannian metric $\mathbf{G}$ on $(G, \star)$ is defined as follows:

$$\mathbf{G}(g)(X, Y) = \mathbf{G}(h \star g)(TL_h(X), TL_h(Y)), \tag{4.32}$$

where $X, Y \in T_gG$. An inner product $\mathbf{I}$ on $\mathcal{L}$ is given by $\mathbf{I} : \mathcal{L} \times \mathcal{L} \to R$. Then the following theorem gives a Riemannian metric on G with respect to the inner product $\mathbf{I}$ defined on $\mathcal{L}$.

LEMMA 4.2. *([**19**]) The inner product $\boldsymbol{I}$ on $\mathcal{L}$ determines a smooth left invariant Riemannian metric $\boldsymbol{G}$ on G as follows:*

$$\boldsymbol{G}(g)(X, Y) = \boldsymbol{I}(TL_{g^{-1}}X, TL_{g^{-1}}Y), \tag{4.33}$$

where $X, Y \in T_gG$. □

It is also shown that a left invariant Riemannian metric $\mathbf{G}$ on G determines an inner product $\mathbf{I}$ via left translation operation, see [**19**], Theorem 5.38. A normal vector N at the switching state $g(t_s)$ on $\mathcal{N}$ satisfies

$$\mathbf{G}(g)(N, Y) = 0, \;\; \forall Y \in T_{g(t_s)}\mathcal{N} \subset T_{g(t_s)}G, \tag{4.34}$$

where by Lemma 4.2 we have $\mathbf{I}(TL_{g^{-1}}N, TL_{g^{-1}}Y) = 0$. By the linear property of the inner product $\mathbf{I}$ on the vector space $\mathcal{L}$ we can defined the following one form

$$D_gN : \mathcal{L} \to R, \;\; D_gN = \mathbf{I}(TL_{g^{-1}}N, .) \in \mathcal{L}^*. \tag{4.35}$$

The following lemma shows that the one form $\mathbf{G}_g(N, .)$ is the pullback of $\mathbf{I}(TL_{g^{-1}}N, .)$ under the map $TL_{g^{-1}}$.

LEMMA 4.3. *For a Lie group $(G, \star)$ associated with an inner product $\boldsymbol{I}$ on $\mathcal{L}$ we have*

$$\forall g \in G, \;\; \boldsymbol{G}(g)(N, .) = T^*L_{g^{-1}}D_gN \in T^*_gG, \tag{4.36}$$

Proof. We show $\forall X \in T_gG$, $\mathbf{G}(g)(N,X) = \langle T^*L_{g^{-1}}D_gN, X\rangle$. As is obvious $TL_{g^{-1}} : T_gG \to \mathcal{L}$ therefore

$$\langle T^*L_{g^{-1}}D_gN, X\rangle = \langle T^*L_{g^{-1}}\mathbf{I}(TL_{g^{-1}}N, .), X\rangle, \tag{4.37}$$

By the definition of pullbacks, see [**41**], we have

$$\begin{aligned}&\langle T^*L_{g^{-1}}\mathbf{I}(TL_{g^{-1}}N, .), X\rangle = \\ &\langle \mathbf{I}(TL_{g^{-1}}N, .), TL_{g^{-1}}X\rangle = \\ &\mathbf{I}(TL_{g^{-1}}N, TL_{g^{-1}}X) = \mathbf{G}(g)(N,X),\end{aligned} \tag{4.38}$$

where the second equality comes from the definition of $\mathbf{I}$. □

The following theorem gives the HMP statement for hybrid systems defined on Lie groups in the case of non-differentiability of the value function.

Theorem 4.3. *Consider a hybrid system satisfying the hypotheses presented in* ***A1, A2, A3*** *(presented in Chapter 2) on a Lie group G and an embedded switching submanifold $\mathcal{N} \subset G$ with an associated inner product $\boldsymbol{I} : \mathcal{L} \times \mathcal{L} \to R$. Then corresponding to the optimal control and optimal state trajectory $u^o(t), g^o(t)$, there exists a nontrivial $\lambda^o \in \mathcal{L}^*$ along the optimal state trajectory such that:*

$$H_{q_i}(g^o(t), \lambda^o(t), u^o(t)) \le H_{q_i}(g^o(t), \lambda^o(t), u_1), \forall u_1 \in U, t \in [t_0, t_f], i = 1, 2, \tag{4.39}$$

and at the optimal switching state and switching time $g^o(t_s), t_s$ we have

$$\lambda(t_s^{-o}) = \lambda(t_s^{+o}) + \mu\boldsymbol{I}(TL_{{g^o}^{-1}(t_s)}N, .) \in \mathcal{L}^*. \tag{4.40}$$

and the continuity of the Hamiltonian is given as follows:

$$H_{q_1}(g^o(t_s^-), \lambda^o(t_s^-), u^o(t_s^-)) = H_{q_2}(g^o(t_s), \lambda^o(t_s), u^o(t_s)). \tag{4.41}$$

The optimal adjoint variable λ^o *satisfies*

$$\frac{dg^o}{dt} = TL_{g^o(t)}\big(\frac{\partial H_{q_i}}{\partial \lambda}\big), \quad \frac{d\lambda^o}{dt} = -(ad)^*_{\frac{\partial H_{q_i}}{\partial \lambda}}(\lambda^o(t)), \;\; t \in [t_i, t_{i+1}), q_i \in Q, \tag{4.42}$$

where

$$H_{q_i}(g, \lambda, u) := \langle \lambda, TL_{g^{-1}} f_{q_i}(g, u)\rangle. \tag{4.43}$$

□

It should be noted that in the case which the normal vector is not uniquely given, the discontinuity of the adjoint process is given by

$$\lambda^o(t_s^-) - \lambda^o(t_s)) \in T^* L_{g^o(t_s)}(T^{*^\perp}_{g^0(t_s^-)} \mathcal{N}), \tag{4.44}$$

where

$$T^{*^\perp}_g \mathcal{N} := \{\alpha \in T^*_g G, \; s.t. \; \forall X \in T_g \mathcal{N}, \langle \alpha, X\rangle = 0\}. \tag{4.45}$$

In order to prove Theorem 4.5, we employ the notion of control needle variation which has been used in the optimal control literature, see [**3**, **6**, **41**].

4.4.1. Control Needle Variation. Similar to the control needle variation introduced in the proof of the Hybrid Maximum Principle in [**66**], we introduce the following control needle variation for a left invariant control system.

$$u_\pi(t^1, \epsilon) = \begin{cases} u_1 & t^1 - \epsilon \le t \le t^1 \\ u^o(t) & \text{elsewhere} \end{cases}, \tag{4.46}$$

where $u_1 \in U$. Let us denote the state flow of the left invariant control system $\dot{g}(t) = f(g, u)$ as $g(t) = g(t, s, g_0)$ where s is the initial time and g_0 is the initial state.

Due to the needle variation, the perturbed control system is given by

$$\dot{g}_{(\pi,\epsilon)}(t) = f(g_{(\pi,\epsilon)}(t), u_\pi(t)), \quad t \in [t_0, t_f]. \tag{4.47}$$

The following theorem gives the state variation of a left invariant control system with respect to a control needle variation.

LEMMA 4.4. *For a Lebesgue time* t^1*, the curve* $g_{\pi,f}^{(t,s),g}(\epsilon):[0,\tau]\to G$ *is differentiable at* $\epsilon=0$ *and the corresponding tangent vector* $\frac{d}{d\epsilon}g_{\pi,f}^{(t^1,s),g}|_{\epsilon=0}$ *is*

$$\frac{d}{d\epsilon}g_{\pi,f}^{(t^1,s),g}|_{\epsilon=0}=TL_{g(t^1)}\big(f(e,u_1)-f(e,u^o(t^1))\big). \tag{4.48}$$

PROOF. The proof is based on the left invariance property of f. As is shown by Lemma 2.3 the state variation with respect to the control needle variation is given by

$$f(g(t^1),u_1)-f(g(t^1),u^o(t^1))=TL_{g(t^1)}\big(f(e,u_1)-f(e,u^o(t^1))\big), \tag{4.49}$$

which completes the proof □

The following lemma gives the state variation at an arbitrary time t, where $t^1<t$, for a non-hybrid left invariant control system.

LEMMA 4.5. *Let* $g_{(\pi,\epsilon)}(t):[t_0,t_f]\to G$ *be a solution of* $\dot{g}_{(\pi,\epsilon)}(t)=f(g_{(\pi,\epsilon)}(t),u_\pi(t))$ *then for* $t^1<t\leq t_f$

$$\begin{aligned}\frac{d}{d\epsilon}g_{\pi,f}^{(t,t^1),x}|_{\epsilon=0}=TR_{exp((t-t^1)f(e,u^o))}\circ\\ TL_{g(t^1)}\big(f(e,u_1)-f(e,u^o(t^1)\big)\in T_{g(t)}G,\end{aligned} \tag{4.50}$$

where $TR_{exp((t-t^1)f(e,u^o))}$ *is the push forward of the right translation* $R_{exp((t-t^1)f(e,u^o))}$ *at* $g(t^1)$*.*

PROOF. As is shown in [**6**] for a given control system on a differentiable manifold $\mathcal{M}$, the state variation at time t where $t^1<t$ is given as follows:

$$\frac{d}{d\epsilon}\Phi_{\pi,f_q}^{(t,t^1),x}|_{\epsilon=0}=T\Phi_{f_q}^{(t,t^1)}([f_q(x(t^1),u_1)-f_q(x(t^1),u(t^1))])\in T_{x(t)}\mathcal{M}, \tag{4.51}$$

where $\Phi_{\pi,f_q}^{(t,t^1),x}$ is the flow initiates form x and corresponds to the control u_π, see [**6**]. The push-forward of $\Phi_{f_q}^{(t,t^1),x}$, i.e. $T\Phi_{f_q}^{(t,t^1)}$ is computed along the nominal control $u(t)$

and is evaluated at $x(t^1)$. For a left invariant control system evolving on G, based on Definition 4.5 and Theorem 4.1 we have

$$g_{(\pi,\epsilon)}(t) = g_0 \circ exp(tf(e,u_\pi)) = g_0 \circ exp((t^1)f(e,u_\pi) + (t-t^1)f(e,u_\pi)). \quad (4.52)$$

Since $u_\pi(t) = u^o(t), \quad t \in [t^1, t_f]$, by the one parameter subgroup property of exp (see [**82**]) we have

$$g_{(\pi,\epsilon)}(t) = g_{(\pi,\epsilon)}(t^1) \circ exp((t-t^1)f(e,u^o)), \quad t^1 < t \leq t_f. \quad (4.53)$$

Therefore, by evaluating the push forward of composition maps, we have

$$\frac{d}{d\epsilon} g_{\pi,f}^{(t,t^1),x}|_{\epsilon=0} = TR_{exp((t-t^1)f(e,u^o))}\big(\frac{d}{d\epsilon} g_{\pi,f}^{(t^1,s),g}|_{\epsilon=0}\big), \quad (4.54)$$

which together with Lemma 4.4 and (4.49) yields the statement. □

We analyze the HOCP with the cost defined in (4.29) and (4.30) by defining a differential form of the penalty function $h(.)$ which is differentiable by the hypotheses. Let us denote

$$dh := \frac{\partial h}{\partial g} \in T_g^* G. \quad (4.55)$$

In order to use the method introduced in [**3, 6, 41**], we prove the following lemma using the optimal control $u^o(.)$ and the associated final state $g^o(t_f)$. We denote $t_s(\epsilon)$ as the associated switching time corresponding to $u_\pi(t,\epsilon)$ which is assumed to be differentiable with respect to ϵ for all $u \in U$.

LEMMA 4.6. *For a Hybrid Optimal Control Problem (HOCP) defined on a Lie group G, at the optimal final state of the trajectory $g^o(t)$ we have*

$$\langle \mathcal{I}^{-1}_{g^o(t_f)}\big(dh(g^o(t_f))\big), TL_{{g^o}^{-1}(t_f)}\big(v_\pi(t_f)\big)\rangle \geq 0, \forall v_\pi(t_f) \in K_{t_f}, \quad (4.56)$$

where

$$K^1_{t_f} = \bigcup_{t_s \le t < t_f} \bigcup_{u_1 \in U} TR_{exp((t_f - t)f(e,u^o))}$$
$$\circ TL_{g(t)}\big(f(e,u_1) - f(e,u^o(t))\big) \subset T_{g(t_f)}G,\ t \in [t_s, t_f], \tag{4.57}$$

and

$$K^2_{t_f} = \bigcup_{t_0 \le t < t_s} \bigcup_{u_1 \in U} TR_{exp((t_f - t_s)f_2(e,u^o))} \circ$$
$$TR_{exp((t_s - t)f_1(e,u^o))} \circ TL_{g(t)}\big(f(e,u_1) - f(e,u^o(t))\big)$$
$$+ \frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0} TR_{exp((t_f - t_s)f_2(e,u^o))} \circ$$
$$TL_{g(t_s)}\big(f_{q_2}(e,u^o(t_s)) - f_{q_1}(e,u^o(t_s))\big)$$
$$\subset T_{g(t_f)}G,\ t \in [t_0, t_s), \tag{4.58}$$

and

$$K_{t_f} = K^1_{t_f} \bigcup K^2_{t_f}. \tag{4.59}$$

PROOF. Based on the definition of pull backs (see [**2**, **19**]), we have

$$\langle \mathcal{I}^{-1}_{g^o(t_f)}\big(dh(g^o(t_f))\big), TL_{{g^o}^{-1}(t_f)}\big(v_\pi(t_f)\big)\rangle =$$
$$\langle T^*L_{g^o(t_f)} \circ \mathcal{I}^{-1}_{g^o(t_f)}\big(dh(g^o(t_f))\big), v_\pi(t_f)\rangle, \tag{4.60}$$

and since by the definition $\mathcal{I}_g = T^*L_{g^{-1}}$, then

$$\langle \mathcal{I}^{-1}_{g^o(t_f)}\big(dh(g^o(t_f))\big), TL_{{g^o}^{-1}(t_f)}\big(v_\pi(t_f)\big)\rangle = \langle dh(g^o(t_f)), v_\pi(t_f)\rangle. \tag{4.61}$$

As shown in [**73**, **75**], $\langle dh(g^o(t_f)), v_\pi(t_f)\rangle \ge 0$ for all $v_\pi(t_f) \in K_{t_f}$. The set K_{t_f}, as is constructed above, contains all the possible final state variation of $g^o(t_f) \in G$, therefore the statement follows. □

The following lemma gives the relation between $\mathbf{G}(g^o(t_s))(N,.)) \in T^*_{g^o(t_s)}G$ and any tangent vector $X \in T_{g^o(t_s)}\mathcal{N} \subset T_{g^o(t_s)}G$.

Lemma 4.7. *Consider an autonomous HOCP consisting of two different regimes separated by a k dimensional embedded switching manifold $\mathcal{N} \subset G$; then at the optimal switching state $g^o(t_s) \in \mathcal{N}$ and switching time t_s we have*

$$\langle \mathcal{I}^{-1}_{g^o(t_s)}\big(\boldsymbol{G}(g^o(t_s))(N,.)\big), TL_{{g^o}^{-1}(t_s)}X\rangle = 0, \forall X \in T_{g^o(t_s)}\mathcal{N}. \tag{4.62}$$

Proof. The proof is same as the proof given in [**75**] since

$$\langle \mathcal{I}^{-1}_{g^o(t_s)}\Big(\mathbf{G}(g^o(t_s))(N,.)\Big), TL_{{g^o}^{-1}(t_s)}X\rangle = G(g^o(t_s))(N,X) = 0. \tag{4.63}$$

□

Here we give the proof for the HMP theorem on G.

Proof. *Step 1*: First consider $t_s < t^1$ where the needle variation is applied at time t^1. As shown in [**75**], we have

$$0 \leq \langle dh, v_\pi(t_f)\rangle, \quad \forall v_\pi \in K_{t_f}, \tag{4.64}$$

where $dh \in T^*_{g(t_f)}G$. As mentioned before the cotangent bundle of the Lie group is identified by $G \times \mathcal{L}$ therefore

$$\mathcal{I}^{-1}_{g(t_f)}(dh) \in \mathcal{L}^*. \tag{4.65}$$

By employing (4.64), we have

$$\begin{aligned} &0 \leq \langle \mathcal{I}^{-1}_{g(t_f)}(dh), TL_{g^{-1}(t_f)} \circ TR_{exp((t_f-t^1)f(e,u^o))} \circ \\ &TL_{g(t^1)}\big(f(e,u_1) - f(e,u^o(t^1))\big)\rangle. \end{aligned} \tag{4.66}$$

The flow of the left invariant system on G implies

$$g(t_f) = L_{g(t)}exp((t_f - t)f(e,u)), \tag{4.67}$$

then by the vector space properties of $\mathcal{L}$ and one parameter subgroups property of exp we have

$$g(t) = L_{g(t_f)} exp(-(t_f - t) f(e,u)), \tag{4.68}$$

which finally gives

$$0 \leq \langle T^* L_{g(t^1)} \circ T^* R_{exp((t_f - t^1) f(e,u))}(dh), f(e,u_1) - f(e,u^o(t^1)) \rangle. \tag{4.69}$$

Therefore $\forall u \in U$

$$\begin{aligned} &\langle T^* L_{g(t^1)} \circ T^* R_{exp((t_f - t^1) f(e,u))}(dh), f(e,u^o(t^1)) \rangle \leq \\ &\langle T^* L_{g(t^1)} \circ T^* R_{exp((t_f - t^1) f(e,u))}(dh), f(e,u_1) \rangle, \end{aligned} \tag{4.70}$$

and

$$T^* L_{g(t^1)} \circ T^* R_{exp((t_f - t^1) f(e,u))}(dh) \in \mathcal{L}^*. \tag{4.71}$$

The adjoint variable is then defined as

$$\lambda(t) = T^* L_{g(t)} \circ T^* R_{exp((t_f - t) f(e,u))}(dh) \in \mathcal{L}^*, \;\; t_s \leq t \leq t_f. \tag{4.72}$$

Step 2: Second consider $t_0 \leq t^1 < t_s$ where t^1 is the needle variation time. Similar to the approach in [**75**] we introduce the value $v(g_s, t_s)$ function with respect to the switching state and switching time (g_s, t_s), $g_s \in G, t_s \in R$. For a given switching time t, the differential form of the value function is then given by $dv(g,t) \in T^*_g G$. In the case for which $t_s(\epsilon) < t_s = t^o_s$ we have (see Lemma 2.6)

$$\begin{aligned} &\frac{dg_{\pi,f_1}^{(t_s(\epsilon),t^1),g(t^1))}}{d\epsilon}|_{\epsilon=0} = (\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}) \\ &\times T L_{g(t_s)}\big(f_1(e,u^o(t_s))\big) \\ &+ T R_{exp(t_s - t^1) f_1(e,u^o)} \circ T L_{g(t^1)}\big(f_1(e,u) - f_1(e,u^o(t^1))\big) \subset T_{g(t_s)} G. \end{aligned} \tag{4.73}$$

Note that the differentiability of $t_s(\epsilon)$ with respect to ϵ is shown in Lemma 2.5 for hybrid systems on Riemannian manifolds.

And for the case in which $t_s < t_s(\epsilon)$ we have

$$\frac{dg^{(t_s(\epsilon),t^1),g(t^1))}_{\pi,f_1}}{d\epsilon}|_{\epsilon=0} = -(\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0})TL_{g(t_s)}\big(f_1(e,u^o(t_s))\big)$$
$$+TR_{exp(t_s-t^1)f_1(e,u^o)} \circ TL_{g(t^1)}\big(f_1(e,u) - f_1(e,u^o(t^1))\big) \subset T_{g(t_s)}G. \tag{4.74}$$

Equation (4.62) together with Lemma 4.5 implies that

$$\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0} = -\langle \mathcal{I}^{-1}_{g(t_s)}\Big(\mathbf{G}(g(t_s))(N,.)\Big), f_1(e,u^o(t_s))\rangle^{-1}$$
$$+\langle T^*L_{g(t^1)} \circ T^*R_{exp((t_s-t^1)f(e,u))}\Big(\mathbf{G}(g(t_s))(N,.)\Big),$$
$$f_1(e,u_1) - f_1(e,u^o(t^1))\rangle, \tag{4.75}$$

since $g^{(t_s(.),t^1,g(t^1))}_{\pi,f_1} : [0,\epsilon] \to \mathcal{N}$ and $\frac{dg^{(t_s(\epsilon),t^1),g(t^1))}_{\pi,f_1}}{d\epsilon}|_{\epsilon=0} \in T_{g(t_s)}\mathcal{N} \subset T_{g(t_s)}G$. In the second case

$$\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0} = \langle \mathcal{I}^{-1}_{g(t_s)}\big(\mathbf{G}(g(t_s))(N,.)\big), f_1(e,u^o(t_s))\rangle^{-1} \times$$
$$\langle T^*L_{g(t^1)} \circ T^*R_{exp((t_s-t^1)f(e,u))}\big(\mathbf{G}(g(t_s))(N,.)\big), f_1(e,u_1) - f_1(e,u^o(t^1))\rangle. \tag{4.76}$$

In order to obtain the state variation at t_s in the case (ii) we use the push-forward of the combination of the flows before and after t_s as follows:

$$\frac{dg^{(t_s,t_s(\epsilon))}_{\pi,f_2} \circ g^{(t_s(\epsilon),t^1,g(t^1))}_{\pi,f_1}}{d\epsilon}|_{\epsilon=0} = TR_{exp(t_s-t^1)f_1(e,u^o)}$$
$$\circ TL_{g(t^1)}\big(f_1(e,u_1) - f_1(e,u^o(t^1))\big)$$
$$+\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}TL_{g(t_s)}\big(f_1(e,u^o(t^1)) - f_2(e,u^o(t^1))\big) \in T_{g(t_s)}G, \tag{4.77}$$

and for case (i)

$$\frac{dg_{\pi,f_2}^{(t_s(\epsilon),t_s)}\circ g_{\pi,f_1}^{(t_s,t^1,g(t^1))}}{d\epsilon}|_{\epsilon=0}=TR_{exp(t_s-t^1)f_1(e,u^o)}$$
$$\circ TL_{g(t^1)}\big(f_1(e,u_1)-f_1(e,u^o(t^1))\big)+\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}$$
$$\times TL_{g(t_s)}\big(f_2(e,u^o(t^1))-f_1(e,u^o(t^1))\big)\in T_{g(t_s)}G. \tag{4.78}$$

The final state variation at the final time t_f is now given as follows:

$$\frac{dg_{\pi,f_2}^{(t_f,t^1,g(t^1))}(\epsilon)}{d\epsilon}|_{\epsilon=0}=TR_{exp(t_f-t_s)f_2(e,u^o)}\frac{dg_{\pi,f_2}^{(t_s(\epsilon),t_s)}\circ g_{\pi,f_1}^{(t_s,t^1,g(t^1))}}{d\epsilon}|_{\epsilon=0}. \tag{4.79}$$

Therefore

$$0\leq\langle dh(g^o(t_f)),TR_{exp(t_f-t_s)f_2(e,u^o)}$$
$$\times\big[\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}TL_{g(t_s)}\big(f_2(e,u^o(t_s))-f_1(e,u^o(t_s))\big)$$
$$+TR_{exp(t_s-t^1)f_1(e,u^o)}\circ TL_{g(t^1)}\big(f_1(e,u_1)-f_1(e,u^o(t^1))\big)\big]\rangle\ , \tag{4.80}$$

Hence

$$0\leq\langle dh(g^o(t_f)),TR_{exp(t_f-t_s)f_2(e,u^o)}$$
$$\times\big[-\langle\mathcal{I}_{g(t_s)}^{-1}\big(\mathbf{G}(g(t_s))(N,.)\big),f_1(e,u^o(t_s))\rangle^{-1}$$
$$\times\langle T^*L_{g(t^1)}\circ T^*R_{exp((t_s-t^1)f(e,u))}\big(\mathbf{G}(g(t_s))(N,.)\big),$$
$$f_1(e,u_1)-f_1(e,u^o(t^1))\rangle$$
$$\times TL_{g(t_s)}\big(f_2(e,u^o(t_s))-f_1(e,u^o(t_s))\big)$$
$$+TR_{exp(t_s-t^1)f_1(e,u^o)}\circ TL_{g(t^1)}\big(f_1(e,u_1)-f_1(e,u^o(t^1))\big)\big]\rangle\ , \tag{4.81}$$

equivalently

$$
\begin{aligned}
&0 \leq -\langle \mathcal{I}^{-1}_{g(t_s)}\big(\mathbf{G}(g(t_s))(N,.)\big), f_1(e,u^o(t_s))\rangle^{-1}\\
&\times\langle T^*L_{g(t^1)} \circ T^*R_{exp((t_s-t^1)f(e,u))}\big(\mathbf{G}(g(t_s))(N,.)\big),\\
&f_1(e,u_1) - f_1(e,u^o(t^1))\rangle\\
&\times\langle dh(g^o(t_f)), TR_{exp(t_f-t_s)f_2(e,u^o)} \circ TL_{g(t_s)}\\
&\big(f_2(e,u^o(t_s)) - f_1(e,u^o(t_s))\big)\rangle\\
&+\langle dh(g^o(t_f)), TR_{exp(t_f-t_s)f_2(e,u^o)}\\
&\circ TR_{exp(t_s-t^1)f_1(e,u^o)} \circ TL_{g(t^1)}\big(f_1(e,u_1) - f_1(e,u^o(t^1))\big)\rangle,
\end{aligned}
\tag{4.82}
$$

Let us denote μ by

$$
\begin{aligned}
\mu = &-\langle \mathcal{I}^{-1}_{g(t_s)} big(\mathbf{G}(g(t_s))(N,.)\big), f_1(e,u^o(t_s))\rangle^{-1}\\
&\langle dh(g^o(t_f)), TR_{exp(t_f-t_s)f_2(e,u^o)} TL_{g(t_s)}\big(f_2(e,u^o(t_s)) - f_1(e,u^o(t_s))\big)\rangle,
\end{aligned}
\tag{4.83}
$$

therefore

$$
\begin{aligned}
0 \leq &\langle dh(g^o(t_f)), TR_{exp(t_f-t_s)f_2(e,u^o)}\\
&\circ TR_{exp(t_s-t^1)f_1(e,u^o)} \circ TL_{g(t^1)}\big(f_1(e,u_1) - f_1(e,u^o(t^1))\big)\rangle + \mu\langle T^*L_{g(t^1)}\\
&\circ T^*R_{exp((t_s-t^1)f(e,u))}\big(\mathbf{G}(g(t_s))(N,.)\big), f_1(e,u_1) - f_1(e,u^o(t^1))\rangle.
\end{aligned}
\tag{4.84}
$$

Similar to step 1 we have

$$
\begin{aligned}
&\langle dh(g^o(t_f)), TR_{exp(t_f-t_s)f_2(e,u^o)} \circ TR_{exp(t_s-t^1)f_1(e,u^o)} \circ TL_{g(t^1)}\\
&\big(f_1(e,u_1) - f_1(e,u^o(t^1))\big)\rangle = \langle T^*L_{g(t^1)} \circ T^*R_{exp(t_s-t^1)f_1(e,u^o)}\\
&\circ T^*R_{exp(t_f-t_s)f_2(e,u^o)}\big(dh(g^o(t_f))\big), \big(f_1(e,u_1) - f_1(e,u^o(t^1))\big)\rangle,
\end{aligned}
\tag{4.85}
$$

Combining (4.84) and (4.85) we have

$$0 \leq \langle T^*L_{g(t^1)} \circ T^*R_{exp(t_s-t^1)f_1(e,u^o)} \circ T^*R_{exp(t_f-t_s)f_2(e,u^o)}\big(dh(g^o(t_f))\big),$$
$$\big(f_1(e,u_1) - f_1(e,u^o(t^1))\big)\rangle + \mu\langle T^*L_{g(t^1)}$$
$$\circ T^*R_{exp((t_s-t^1)f(e,u))}\big(\mathbf{G}(g(t_s))(N,.)\big), f_1(e,u_1) - f_1(e,u^o(t^1))\rangle. \tag{4.86}$$

The adjoint process λ is defined as follows:

$$\lambda(t) = T^*L_{g(t)} \circ T^*R_{exp(t_s-t)f_1(e,u^o)}$$
$$\circ T^*R_{exp(t_f-t_s)f_2(e,u^o)}\big(dh(g^o(t_f))\big) + \mu T^*L_{g(t)} \circ T^*R_{exp((t_s-t)f(e,u))}\big(\mathbf{G}(g(t_s))(N,.)\big). \tag{4.87}$$

At time $t = t_s$ we have

$$\lambda(t_s^{-o}) = \lambda(t_s^{+o}) + \mu T^*L_{g(t_s)}(\mathbf{G}(g(t_s))(N,.)) \in \mathcal{L}^*. \tag{4.88}$$

It only remains to show

$$\frac{dg}{dt} = TL_{g(t)}\big(\frac{\partial H_{q_i}}{\partial \lambda}\big), \quad \frac{d\lambda}{dt} = -(ad)^*_{\frac{\partial H_{q_i}}{\partial \lambda}}(\lambda(t)), \;\; t \in [t_i, t_{i+1}), q_i \in Q. \tag{4.89}$$

The first part of (4.89) is obvious by the definition of $H_{q_i} := \langle \lambda, TL_{g^{-1}(t)}f_{q_i}(g(t),u)\rangle$, since f_{q_i} is left invariant and $\frac{dg}{dt} = TL_{g(t)} \circ TL_{g^{-1}(t)}f_{q_i}(g(t),u) = f_{q_i}(g(t),u)$.

Step 3: In order to invoke results from [**38**], it is sufficient to show that for the constructed adjoint variable $\lambda(.)$, we have $\lambda(t) = Ad^*_{g(t)}(\lambda(0))$ where Ad^* is defined below. For a given $g \in G$ we define the *conjugate map* $I_g : G \to G$ as follows (see [**2, 19**]):

$$I_g(h) = g \star h \star g^{-1}. \tag{4.90}$$

The *adjoint map* $Ad_g : \mathcal{L} \to \mathcal{L}$ is defined by

$$Ad_g = TI_g = TL_g \circ TR_{g^{-1}}, \tag{4.91}$$

where the dual of the adjoint map Ad^*_g is calculated as $Ad^*_g = T^*L_g \circ T^*R_{g^{-1}}$. As is obtained in the step 1, $\lambda(t) = -T^*L_{g(t)} \circ T^*R_{exp((t_f-t)f(e,u))}(dh) \in \mathcal{L}^*$, then in order to show the second claim of (4.89), it is enough to show that $\lambda(t) = Ad^*_{g(t)}(\lambda(0))$ where without loss of generality we set $t_s = 0$ and $\lambda(0) = \lambda(t_s)$. Therefore we should show

$$T^*L_{g(t)} \circ T^*R_{exp((t_f-t)f(e,u))}(dh) = T^*L_{g(t)} \circ T^*R_{g^{-1}(t)} \circ T^*L_{g(0)} \circ T^*R_{exp((t_f)f(e,u))}(dh). \tag{4.92}$$

Employing the group operation we have

$$g(t_f) = g(0) \star exp(t_f f_2(e, u^o)) = g(0) \star g(t) \star g^{-1}(t) \star exp(t_f f_2(e, u^o)), \tag{4.93}$$

and also

$$g(t_f) = g(t) \star exp((t_f - t)f_2(e, u^o)), \tag{4.94}$$

then

$$R_{exp((t_f-t)f_2(e,u^o))}(g(t)) = R_{exp(t_f f_2(e,u^o))} \circ L_{g(0)} \circ R_{g^{-1}(t)}(g(t)), \quad \forall g(t) \in G, \tag{4.95}$$

which implies

$$T^*R_{exp((t_f-t)f_2(e,u^o))} = T^*R_{g^{-1}(t)} \circ T^*L_{g(0)} \circ T^*_{g(0)}R_{exp(t_f f_2(e,u^o))}, \tag{4.96}$$

which shows (4.92). As is shown in [**38**], $\lambda(t) = Ad^*_{g(t)}(\lambda(0))$ implies

$$\frac{d\lambda}{dt} = -ad^*_{\frac{dg}{dt}}(\lambda(t)) = -ad^*_{\frac{\partial H_i}{\partial \lambda}}(\lambda(t)), \tag{4.97}$$

and completes the proof. Same argument holds for $\lambda(t), \quad t_0 \leq t < t_s$. □

4.4.2. Interior Optimal Switching State. Here we specify a hypothesis for MHOCP which expresses the HMP statement based on a differential form of the hybrid value function.

A5: For an MHOCP, the value function $v(g,t),\ g \in G, t \in (t_0, t_f)$, is assumed to be differentiable at the optimal switching state $g^o(t_s^-)$ in the switching manifold $\mathcal{N}$, where the optimal switching state is an interior point of the attainable switching states on the switching manifold.

We note that ***A5*** rules out MHOCPs derived from BHOCPs. The following theorem gives the HMP statement for an accessible MHOCP satisfying ***A5***.

THEOREM 4.4. *Consider a hybrid system satisfying the hypotheses presented in* ***A1, A2, A3, A5*** *on a Lie group G and an embedded switching submanifold $\mathcal{N} \subset G$. Then corresponding to the optimal control and optimal state trajectory $u^o(t), g^o(t)$, there exists a nontrivial $\lambda^o \in \mathcal{L}^*$ along the optimal state trajectory such that:*

$$H_{q_i}(g^o(t), \lambda^o(t), u^o(t)) \le H_{q_i}(g^o(t), \lambda^o(t), u_1),$$
$$\forall u_1 \in U, t \in [t_0, t_f], i = 1, 2, \tag{4.98}$$

and at the optimal switching state and switching time $g^o(t_s), t_s$ we have

$$\lambda(t_s^{-o}) = \lambda(t_s^{+o}) + \mu T^* L_{g(t_s)}(dv(g_s^o, t_s)) \in \mathcal{L}^*. \tag{4.99}$$

and the continuity of the Hamiltonian is given as follows

$$H_{q_i}(g^o(t_s^-), \lambda^o(t_s^-), u^o(t_s^-)) = H_{q_{i+1}}(g^o(t_s), \lambda^o(t_s), u^o(t_s)). \tag{4.100}$$

The adjoint variable λ satisfies

$$\frac{dg}{dt} = TL_{g(t)}\left(\frac{\partial H_{q_i}}{\partial \lambda}\right), \quad \frac{d\lambda}{dt} = -(ad)^*_{\frac{\partial H_{q_i}}{\partial \lambda}}(\lambda(t)), t \in [t_i, t_{i+1}), q_i \in Q, \tag{4.101}$$

where

$$H_{q_i}(g, \lambda, u) := \langle \lambda, TL_{g^{-1}} f_{q_i}(g, u) \rangle, \tag{4.102}$$

and

$$dv(g^o(t_s^-), t_s) = \frac{\partial v(g_s^o, t_s^-)}{\partial g} \in T^*_{g^o(t_s^-)} G. \tag{4.103}$$

□

Proof. The proof is a repetition of the proof of Theorem 4.5 where $\mathbf{G}(g)(N, .)$ is replaced by $dv(g,t)$ where $\mathbf{G}$ is the Riemannian metric associated with the inner product $\mathbf{I}$, see Lemma 4.2. As shown in Theorem 4.3, the adjoint process discontinuity is given by

$$\lambda(t_s^-) = \lambda(t_s) + \mu T^* L_{g(t_s)} dv(g(t_s), t_s). \tag{4.104}$$

□

4.5. Exp-Gradient HMP Algorithm

In this section we introduce an algorithm which is based upon the HMP algorithm first introduced in [**66**] and then extended on Riemannian manifolds in [**77**]. The algorithm presented in [**77**] is an extension of the Steepest decent algorithm along the geodesics on Riemannian manifolds. As known (see [**37**]), geodesics are defined as length minimizing curves on Riemannian manifolds. The solution of the Euler-Lagrange variational problem associated with the length minimizing problem shows that all the geodesics on $\mathcal{M}$ connecting $\gamma(a), \gamma(b) \in \mathcal{M}$ must satisfy the following system of ordinary differential equations:

$$\ddot{x}_i(s) + \sum_{j,k=1}^{n} \Gamma^i_{j,k} \dot{x}_j(s) \dot{x}_k(s) = 0, \quad i = 1, ..., n, \tag{4.105}$$

where

$$\Gamma^i_{j,k} = \frac{1}{2} \sum_{l=1}^{n} g^{il} (g^{\mathcal{M}}_{jl,k} + g^{\mathcal{M}}_{kl,j} - g^{\mathcal{M}}_{jk,l}), \quad g^{\mathcal{M}}_{jl,k} = \frac{\partial}{\partial x_k} g^{\mathcal{M}}_{jl}, \tag{4.106}$$

where $g^{\mathcal{M}}$ is the Riemannain metric corresponding to $\mathcal{M}$ and all the indices i, j, k here run from 1 up to $n = dim(\mathcal{M})$ and $[g^{ij}] = [g^{\mathcal{M}}_{ij}]^{-1}$.

In order to introduce the gradient of the value function on a Lie group G we employ the notion of inner product on a finite dimensional Lie algebra $\mathcal{L}$ defined in Section 4.4. For a given value function $v : G \to R$ on a Lie group G we have

$$dv|_g := \frac{\partial v}{\partial g} \in T^*_g G. \tag{4.107}$$

The gradient of v, i.e. ∇v, is defined as

$$\langle dv, X_g \rangle = \mathbf{G}(g)(\nabla v, X_g), \ \ \forall X_g \in T_g G, \tag{4.108}$$

which can be written as

$$\begin{aligned} \langle dv, X_g \rangle &= \langle dv, TL_g X \rangle = \langle T^* L_g dv, X \rangle \\ &= \mathbf{I}(TL_{g^{-1}} \nabla v, TL_{g^{-1}} X_g), \ \ \forall X_g \in T_g G. \end{aligned} \tag{4.109}$$

We call $TL_{g^{-1}} \nabla v$ the *projected gradient* of v on $\mathcal{L}$. Similar to the geodesic gradient flow defined on Riemannian manifold $\mathcal{M}$ in [**77**], we introduce *Exp-Gradient Flow* on Lie groups as follows:

Definition 4.4. *(Exp-Gradient Flow) Let $\theta^0 = 0$, and $g(\theta^0) = g^0 \in G$, then for all $0 \leq k$ and all g^k such that $TL_{{g^k}^{-1}} \nabla v(g^k) \neq 0$, define*

$$\begin{aligned} \gamma_{g^k}(\theta) = g(\theta) \ = g^k \star exp(-\theta TL_{{g^k}^{-1}} \nabla v(g^k)), \\ \theta \in [\theta^k, \theta^{k+1}), \quad g(\theta) \in G, \end{aligned} \tag{4.110}$$

where

$$\theta^{k+1} = \sup_s \{s; \frac{dv(g(t))}{dt} \leq 0, \ \ t \in [\theta^k, s), s \in [\theta^k, \theta^k + 1)\}. \tag{4.111}$$

□

Over the interval of existence $[0, \omega)$ we denote the total flow induced by (4.110)) as

$$\varphi(\theta, g^0) = \Pi_{i=1}^{n} \psi_i(\theta^{i-1}, \theta^i, g^{i-1}) \circ \psi_n(\theta^n, \theta, g^n), \tag{4.112}$$

where

$$\psi_i(\theta^{i-1}, \theta^i, g^{i-1}) = \gamma_{g^{i-1}}(\theta^i - \theta^{i-1}), \;\; \gamma_{g^0}(\theta^1 - \theta^0) = \gamma_{g^0}(\theta^1), \theta^0 = 0, \tag{4.113}$$

$\theta^i - \theta^{i-1}$ is the elapsed time between the switching times θ^i, θ^{i-1} to the next iteration and n is the index number of the last switching before the instant θ. By the continuity of geodesic flows $\{\psi_i, 1 \leq i < \infty\}$, φ is a continuous map on $[0, \omega)$. In the notation of topological dynamics, and in particular Lasalle Theory (see e.g. [**20**,**60**]), the limit set of the initial state x^0 is denoted as $\Omega(g^0)$, where

$$y \in \Omega(g^0) \Rightarrow \exists \theta_n, n \geq 1, \quad s.t. \quad lim_{n \to \infty} g(\theta_n) = y, \tag{4.114}$$

when $lim_{n\to\infty}(\theta_n) = \omega$. Note the sequence $\{\theta_n\}$ is in general distinct from $\{\theta^n\}$.

H1: There exists $0 < b < \infty$ such that the associated sublevel set $\mathcal{N}_b = \{g \in M; \quad v(g) < b\}$ is (i) open (ii) connected, (iii) contains a strict local minimum g_* which is the only local minimum in $\mathcal{N}_b$, (iv) $\mathcal{N}_b$ has compact closure and (v) $\mathcal{N}_{g_*} \subset \mathcal{N}_b$.

Without loss of generality, we assume $\mathcal{N}_{g_*} \subset \mathcal{N}_{b-\epsilon}$ for some $\epsilon > 0$, then by selecting $g^0 \in \mathcal{N}_{g_*} \subset \mathcal{N}_{b-\epsilon} \subset \mathcal{N}_b$ we prove $\omega = \infty$ by the following lemma:

Lemma 4.8. *For an initial state $g^0 \in \mathcal{N}_{g_*}$, the existence interval of the flow defined in (4.112) goes to ∞.*

Proof. By ***H1*** we have $\mathcal{N}_{g_*} \subset \mathcal{N}_{b-\epsilon}$. Choose $0 < \theta < \theta'$ then if θ is not a switching time by the construction of ϕ, i.e. (4.111)

$$v(\varphi(\theta', g^0)) \leq v(\varphi(\theta, g^0)) \leq v(g^0) < b - \epsilon < b. \tag{4.115}$$

We need to prove the statement above when θ is a switching time. The derivative from the right of the flow φ which is the combination of the flows defined in (4.110) at the switching state g^k is given by

$$\begin{aligned} \frac{dv(g^k(\theta))}{d\theta}|_{\theta=0} &= \langle dv, -TL_{g^k}TL_{g^{k^{-1}}}\nabla v\rangle \\ &= -\langle T^*L_{g^k}dv, TL_{g^{k^{-1}}}\nabla v\rangle \\ &= -\mathbf{I}(TL_{g^{k^{-1}}}\nabla v, TL_{g^{k^{-1}}}\nabla v) < 0. \end{aligned} \tag{4.116}$$

It follows by the construction of φ in 4.112, for all $0 < \theta < \theta^{'}$, that

$$v(\varphi(\theta^{'}, g^0)) \leq v(\varphi(\theta, g^0)) \leq v(g^0) < b - \epsilon < b, \tag{4.117}$$

and hence for $\Phi^+ := \{\varphi(\theta, g^0); 0 \leq \theta < \omega\}$

$$\overline{\Phi^+} \subset \overline{\mathcal{N}}_{b-\epsilon} \subset \mathcal{N}_b. \tag{4.118}$$

So the flow φ is defined everywhere in $\overline{\mathcal{N}_{b-\epsilon}}$, where $\mathcal{N}_b$ has compact closure. Hence for all $g \in \mathcal{N}_{b-\epsilon}$ we have an extension of φ in $\mathcal{N}_b$, therefore the maximum interval of existence of $\varphi(., g^0)$ in $\mathcal{N}_b$ is infinite. □

THEOREM 4.5. *Subject to the hypothesis* ***H1*** *on $\mathcal{N}_b$ and with an initial state g^0 such that $g^0 \in \mathcal{N}_{b-\epsilon} \subset M$, $0 < \epsilon < b$, either the Geodesic-Gradient flow, φ, reaches an equilibrium after a finite number of switchings, or it satisfies*

$$\varphi(\theta, g^0) \to \Omega(g^0) \subset v^{-1}(c), \quad c \in R, \tag{4.119}$$

as $\theta \to \infty$, for some $c \in R$, where

$$\forall y \in \Omega(g^0), \quad \frac{dv(y)}{d\theta}|_{\theta=0} = 0, \tag{4.120}$$

and, furthermore, the switching sequence $\{g\}_0^\infty = \{g^0, g^1, \cdots, \}$ converges to the limit point $g_* \in \Omega(g^0) \subset \mathcal{N}_b$, where g_* is the unique element of $\mathcal{N}_b$ such that $\nabla_M^\gamma v(g_*) = 0$.

PROOF. The first statement of the theorem is immediate by Definition 4.4. To prove the second statement, similar to the proof of the Lasalle Theorem, we proceed by showing that $v(.)$ is constant on the set $\Omega(x^0)$. The precompactness of Φ^+ ((i): $\overline{\Phi^+} \subset \overline{\mathcal{N}}_b$ (ii): there does not exist $\theta_i \to \omega, i \to \infty$, such that $\varphi(\theta_i, g^0) \to \partial\mathcal{N}_b$, i.e. $\overline{\Phi^+} \bigcap \partial\mathcal{N}_b = \oslash$), imply $\Omega(g^0) \neq \oslash$, see [**20**]. By the definition of $\Omega(g^0)$ we have

$$\forall y \in \Omega(g^0) \Rightarrow \exists \theta_n, n \geq 1, \quad s.t, \quad \varphi(\theta_n, g^0) \to y, \; \theta_n \to \infty, \tag{4.121}$$

and since $v(.) \in C^1$,

$$lim_{n\to\infty} v(g(\theta_n)) = lim_{n\to\infty} v(\varphi(\theta_n, g^0)) = v(y) =: c. \tag{4.122}$$

Now choose $y^{'} \in \Omega(g^0), y^{'} \neq y$, then by the existence of a convergent sequence $g(\theta^{'}_n)$ to $y^{'}$ we have

$$\begin{aligned} &\forall \epsilon > 0 \Rightarrow \exists n, n_i, k \quad s.t \quad \theta_n < \theta^{'}_{n_i} < \theta_{n+k} \\ &c - \epsilon < v(g(\theta_{n+k})) \leq v(g(\theta^{'}_{n_i})) \leq v(g(\theta_n)) < c + \epsilon, \end{aligned} \tag{4.123}$$

i.e. $v(y^{'}) = c$, hence $\Omega(x^0) \subset v^{-1}(c)$. To prove stationarity, i.e. (4.120), we observe that $\Omega(x^0)$ is positive invariant under the flow φ, i.e.

$$\varphi(\theta, \Omega(g^0)) \subset \Omega(g^0), \quad \theta > 0. \tag{4.124}$$

This follows from the continuity of $\varphi(.,.)$, see [**20**]. Differentiability from the right for all $g \in \varphi(\theta, g^0), 0 < \theta$, implies

$$\begin{aligned} \frac{dv}{d\theta}|_{\theta=0} &= lim_{\theta\to 0^+} \frac{v(\varphi(\theta, y)) - v(\varphi(0, y))}{\theta} \\ &= lim_{\theta\to 0^+} \frac{c - c}{\theta} = 0, \quad y \in \Omega(g^0), \end{aligned} \tag{4.125}$$

since $\varphi(\theta, y) \in \Omega(g^0)$ by (4.124) and $v(\Omega(g^0)) = c$ by (4.123).

It remains to prove the statement for the sequence of the switching states $\{g\}_0^\infty = \{g^0, g^1, \cdots\}$. The switching sequence $\{g\}_0^\infty$ consists of the switching points on $\varphi(\theta, g^0)$ which by (4.110) is an infinite sequence.

The precompactness of Φ^+ with respect to $\mathcal{N}_b$ implies the existence of a convergent subsequence of $\{g\}_0^\infty$ such that

$$lim_{i\to\infty}\varphi(\theta_i^n, g^0) = g^* \in \Omega(g^0), \Omega(g^0) \subset \overline{\Phi^+} \subset \overline{\mathcal{N}_{b-\epsilon}}. \tag{4.126}$$

Since $v \in C^\infty(\mathcal{N}_b)$

$$lim_{i\to\infty}\nabla v(\varphi(\theta_i^n, g^0)) = \nabla v(g^*), \tag{4.127}$$

and

$$lim_{i\to\infty}\frac{dv(\varphi(\theta_i^n, g^0))}{d\theta}|_{\theta=0} = \frac{dv(g^*)}{d\theta}|_{\theta=0}. \tag{4.128}$$

But since the state $\varphi(\theta_i^n, g^0))$ is a switching state chosen from the switching sequence $\{g\}_0^\infty$,

$$\frac{dv(\varphi(\theta_i^n, g^0))}{d\theta}|_{\theta=0} = -\mathbf{I}(TL_{\varphi(\theta_i^n,g^0)^{-1}}\nabla v, TL_{\varphi(\theta_i^n,g^0)^{-1}}\nabla v), \tag{4.129}$$

As is stated in (4.126), the limit point g^* is an element of the limit set $\Omega(g^0)$, therefore by (4.125) we have

$$\frac{dv(g^*)}{d\theta}|_{\theta=0} = 0. \tag{4.130}$$

From (4.127)-(4.129) we have

$$\begin{aligned} 0 &= \frac{dv(x^*)}{d\theta}|_{\theta=0} = lim_{i\to\infty}\frac{dv(\varphi(\theta_i^n, x^0))}{d\theta}|_{\theta=0} \\ &= lim_{i\to\infty}\big(-\mathbf{I}(TL_{\varphi(\theta_i^n,g^0)^{-1}}\nabla v, TL_{\varphi(\theta_i^n,g^0)^{-1}}\nabla v)\big). \end{aligned} \tag{4.131}$$

Hence

$$\nabla v(g^*) = 0, \tag{4.132}$$

or equivalently

$$dv|_{g^*} = 0. \tag{4.133}$$

But by ***H1***, g_* is the unique point in $\mathcal{N}_{b-\epsilon} \subset \mathcal{N}_b$ for which this holds, hence all subsequences of $\{g\}_0^\infty$ converge to $g_* = g^*$ an hence so does the sequence. □

Definition 4.5. *(Conceptual EG-HMP Algorithm)*
Consider the hybrid system (4.29) with two phases and the performance function $v(.)$.

Generate the Exp-Gradient flow, (4.110)-(4.112), on G with $\nabla v(g), \quad g \in M$, evaluated by (4.109).

Stopping rule: for a given $0 < \beta$, if $\boldsymbol{I}(TL_{g^{k-1}}\nabla v, TL_{g^{k-1}}\nabla v) < \beta$ stop. □

Theorem 1. *Assume **H1** holds for $\mathcal{N}_b \subset G$, for the HOCP with the performance function $v(.)$, then the EG-HMP with data (G, v, β) halts at $g^{k(\beta)}(\beta)$, where either $g^{k(\beta)}(\beta)$ is a finite equilibrium point of the Geodesic-Gradient flow, and hence $\nabla v(g^{k(\beta)}(\beta)) = 0$ and $g^{k(\beta)}(\beta) = g_*$, where g_* is the unique point of $\mathcal{N}_b \subset G$ such that $||\nabla v(g_*)|| = 0$, or $g^{k(\beta)}(\beta)$ is such that*

$$g^{k(\beta)}(\beta) \to g_*, \quad k(\beta) \to \infty, \quad as \quad \beta \to 0. \tag{4.134}$$

Proof. The first statement is immediate by Definition 4.5. The second holds since $v(.)$ has a unique local minimum at g_*, and $v(.) \in C^1(\mathcal{N}_b)$ with $\nabla v(g_*) = 0$,

$$\rho_\beta(g_*) := \sup \left\{ d_G(g, g_*); \mathbf{I}(TL_{g^{-1}}\nabla v, TL_{g^{-1}}\nabla v) < \beta, x \in G \right\}, \tag{4.135}$$

where $d(.,.)$ is the geodesic distance on G, is such that $\rho_\beta(g_*) \to 0$ as $\beta \to 0$, hence $g^{k(\beta)}(\beta) \to g_*, \quad as \quad \beta \to 0.$ □

4.6. Satellite Example

In this section we give a conceptual example on $SO(3)$ to clarify the notion of left invariant hybrid systems optimal control.

As known $SO(3)$ is the rotation group in R^3 which is given by

$$SO(3) = \{g \in GL(3) | \quad g.g^T = I, \; det(g) = 1\}, \tag{4.136}$$

where $GL(n)$ is the set of nonsingular $n \times n$ matrices. The Lie algebra of $SO(3)$ which is denoted by $so(3)$ is given by (see [**82**])

$$so(3) = \{X \in M(3) | \quad X + X^T = 0\}, \tag{4.137}$$

where $M(n)$ is the space of all $n \times n$ matrices. The Lie group operation $\star$ is given by the matrix multiplication and consequently TL_{g_2} is also given by the matrix multiplication $g_2 X, \;\; X \in T_{g_1} G$.

A left invariant dynamical system on $SO(3)$ is given by

$$\dot{g}(t) = gX, \quad g(0) = g_0, X \in so(3). \tag{4.138}$$

The Lie algebra bilinear operator is defined as the commuter of matrices, i.e.

$$[X, Y] = XY - YX, \quad X, Y \in so(3). \tag{4.139}$$

The kinematic equations expressing the state trajectory $g(.)$ for a satellite is given by

$$\begin{aligned} &\dot{g}(t) = g(t)X(t), \quad \dot{\hat{X}}(t) + \mathbb{I}^{-1}(\hat{X}(t) \times \mathbb{I}\hat{X}(t)) = \mathbb{I}^{-1}\tau(t), \\ &g(t) \in SO(3), X(t) \in so(3), \end{aligned} \tag{4.140}$$

where $\hat{.} : so(3) \to \mathbb{R}^3$ is an isomorphism, $\mathbb{I}$ is the inertia tensor and $\tau(t) \in \mathbb{R}^3$ is the input torque. For more details of the modelling above see [**19**], Page 281. The second part of (4.140) is the *controlled Euler-Poincare equation* and (4.140) is the geodesic equation on G in the presence of external forces, see [**5**, **19**].

A controlled left invariant system on $SO(3)$ is defined as

$$\dot{g}(t) = g(t)\begin{pmatrix} 0 & u_1(t) & u_3(t) \\ -u_1(t) & 0 & u_2(t) \\ -u_3(t) & -u_2(t) & 0 \end{pmatrix},$$
$$g(t) \in SO(3), (u_1, u_2, u_3) \in R^3. \tag{4.141}$$

The Lie algebra $so(3)$ is spanned by $e_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$. One can check that

$$[e_1, e_2] = e_3,\ [e_1, e_3] = -e_2,\ [e_2, e_3] = e_1. \tag{4.142}$$

By the controllability results presented in [**38**] since all the Lie algebras generated by $(e_1, e_2), (e_2, e_3), (e_1, e_3)$ span the tangent space of the Lie group then all the systems derived by each pair of controls are controllable. Here we define a hybrid system on $SO(3)$ as follows: The continuous dynamics are given by

$$\dot{g}_1(t) = g_1(t)\begin{pmatrix} 0 & u_1(t) & 0 \\ -u_1(t) & 0 & u_2(t) \\ 0 & -u_2(t) & 0 \end{pmatrix}, \quad t \in [t_0, t_s)$$
$$\dot{g}_2(t) = g_2(t)\begin{pmatrix} 0 & u_1(t) & u_3(t) \\ -u_1(t) & 0 & 0 \\ -u_3(t) & 0 & 0 \end{pmatrix}, \quad t \in [t_s, t_f]$$
$$g_1(t), g_2(t) \in SO(3), (u_1, u_2, u_3) \in R^3, \tag{4.143}$$

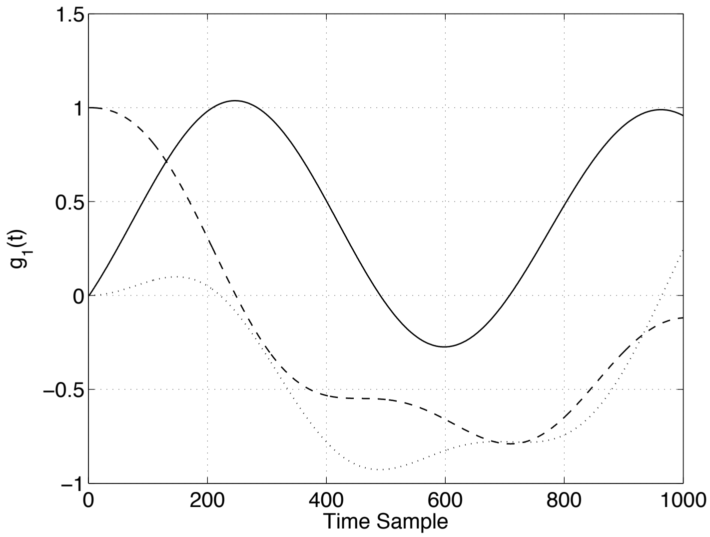

FIGURE 4.1. Hybrid State Trajectory Phase 1

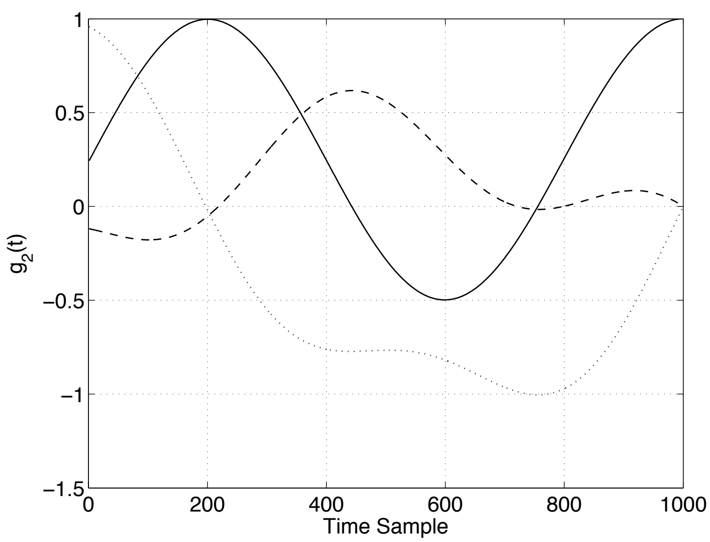

FIGURE 4.2. Hybrid State Trajectory Phase 2

where

$$J_1 = \frac{1}{2}\int_{t_0}^{t_s} u_1^2(t) + u_2^2(t)dt, J_2 = \frac{1}{2}\int_{t_0}^{t_s} u_1^2(t) + u_3^2(t)dt. \tag{4.144}$$

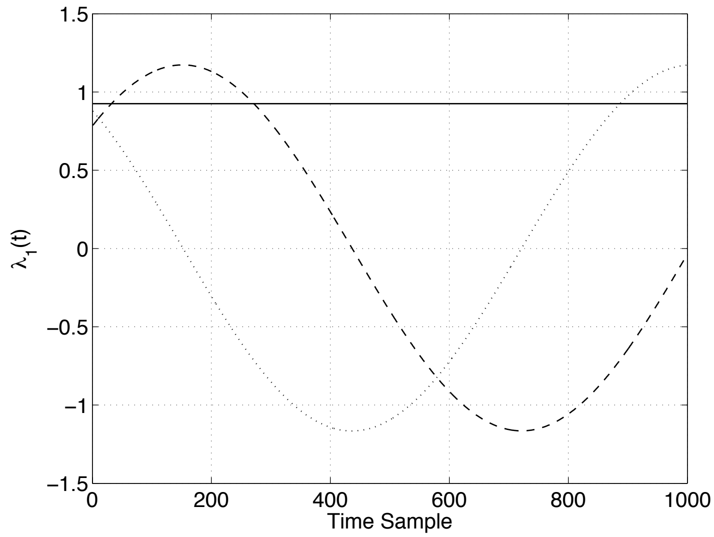

FIGURE 4.3. Hybrid Adjoint Trajectory Phase 1

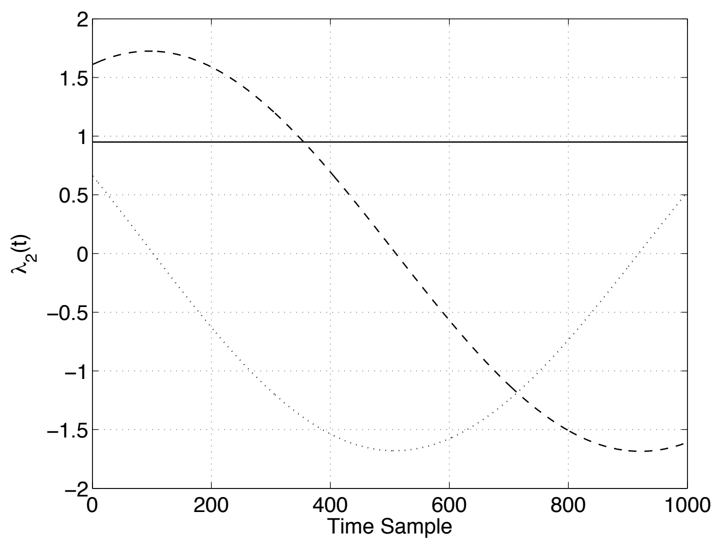

FIGURE 4.4. Hybrid Adjoint Trajectory Phase 2

The Hamiltonians corresponding to the left invariant dynamics are

$$H_1(\lambda, u_1, u_2) = \langle \lambda, u_1 e_1 + u_2 e_2 \rangle + \frac{1}{2}(u_1^2 + u_2^2), \tag{4.145}$$

$$H_2(\lambda, u_1, u_3) = \langle \lambda, u_1 e_1 + u_3 e_3 \rangle + \frac{1}{2}(u_1^2 + u_3^2), \tag{4.146}$$

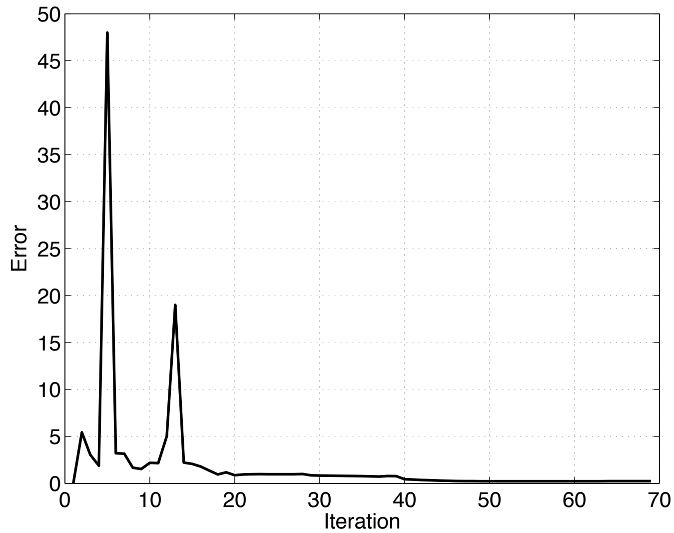

FIGURE 4.5. EX-HMP Convergence

where $\lambda = \lambda_1 e_1^* + \lambda_2 e_2^* + \lambda_3 e_3^*$ and $\langle e_i^*, e_j \rangle = \delta_{i,j},\ \ i,j = 1,2,3$. By the Minimum Principle, the optimal controls are obtained as

$$u_1^*(t) = -\lambda_1(t), u_2^*(t) = -\lambda_2(t), \quad t \in [t_0, t_s), \tag{4.147}$$

$$u_1^*(t) = -\lambda_1(t), u_3^*(t) = -\lambda_3(t), \quad t \in [t_s, t_f]. \tag{4.148}$$

We can put the elements $X \in \mathcal{L}$ into one to one correspondence with the vectors in R^3 via the unique coefficients of the linear expansion of any X in terms of $e_i,\ \ i = 1,2,3$, i.e.

$$X = \sum_{i=1}^{3} \alpha_i e_i, \qquad X \to (\alpha_1, \alpha_2, \alpha_3), \tag{4.149}$$

which yields in particular $e_i \to \mathbf{e}_i \in R^3, \quad i = 1,2,3$.

Using the identification above and by (4.142) we can identify the ad operators as 3×3 matrices as follows:

$$ad_{e_1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, ad_{e_2} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \tag{4.150}$$

$$ad_{e_3} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{4.151}$$

where for example

$$ad_{e_1}(e_2) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \mathbf{e}_3. \tag{4.152}$$

We note that (4.152) and related expressions display the correspondence between the operators ad_{e_i} and rotations around the ith axis in $R^3, i = 1, 2, 3$.

Equations (4.145) and (4.146) together imply

$$\frac{\partial H_1}{\partial \lambda} = u_1 e_1 + u_2 e_2, \frac{\partial H_2}{\partial \lambda} = u_1 e_1 + u_3 e_3, \tag{4.153}$$

therefore

$$ad_{\frac{\partial H_1}{\partial \lambda}} = u_1 ad_{e_1} + u_2 ad_{e_2} = \begin{pmatrix} 0 & 0 & u_2 \\ 0 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}, ad_{\frac{\partial H_2}{\partial \lambda}} = \begin{pmatrix} 0 & -u_3 & 0 \\ u_3 & 0 & -u_1 \\ 0 & u_1 & 0 \end{pmatrix}. \tag{4.154}$$

Hence the differential equations corresponding to the adjoint variable λ are given by

$$\begin{aligned}\dot{\lambda}_1(t) &= \lambda_3(t)u_2^*(t),\\ \dot{\lambda}_2(t) &= -\lambda_3(t)u_1^*(t),\\ \dot{\lambda}_3(t) &= -\lambda_1(t)u_2^*(t) + \lambda_2(t)u_1^*(t), \quad t \in [t_0, t_s),\end{aligned} \tag{4.155}$$

$$\begin{aligned}\dot{\lambda}_1(t) &= -\lambda_2(t)u_3^*(t),\\ \dot{\lambda}_2(t) &= \lambda_1(t)u_3^*(t) - \lambda_3(t)u_1^*(t),\\ \dot{\lambda}_3(t) &= \lambda_2(t)u_1^*(t), \quad t \in [t_s, t_f].\end{aligned} \tag{4.156}$$

Definition 4.6. *For a finite dimensional Lie algebra $so(3)$, we define the Killing Form B as*

$$B(X,Y) = tr(ad_X ad_Y), \;\; X, Y \in so(3). \tag{4.157}$$

□

The Killing Form is invariant in the sense that

$$B([X,Y],Z) = B(X,[Y,Z]). \tag{4.158}$$

Now corresponding to B we introduce an inner product I_B on $so(3)$ as

$$I_B(X,Y) = -tr(ad_X ad_Y). \tag{4.159}$$

Lemma 4.2 implies that B induces a left invariant metric on G. By (4.150)-(4.151) we have

$$I_B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{4.160}$$

By the realization above, $T^*L_g dv = \lambda_1 e_1^* + \lambda_2 e_2^* + \lambda_3 e_3^* \in so^*(3)$ implies that

$$TL_{g^{-1}}\nabla v = \frac{\lambda_1}{2}e_1 + \frac{\lambda_2}{2}e_2 + \frac{\lambda_3}{2}e_3 \in so(3). \tag{4.161}$$

The algorithm initiates from $t_0 = 0$, $t_f = 10$, $g_0 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, $t_s = 5.8s$ and $g_s = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $g_f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. The algorithm converges to $g_s = \begin{pmatrix} 0.3039 & 0.9574 & -0.1194 \\ -0.3688 & 0.1508 & -0.9165 \\ -0.8604 & 0.3156 & 0.3988 \end{pmatrix}$ and $t_s = 5.9733$. The state trajectory and adjoint variables are shown in Figures 4.1-4.4 and Figure 4.5 shows the convergence of the Exp-HMP algorithm.

CHAPTER 5

The Geometry and Deformation of Switching Manifolds

5.1. Problem Formulation

The hypotheses of Theorem 2.1 restrict attention to hybrid systems without controlled discrete state switchings, that is to say, switching occurs if and only if the state trajectory meets a switching manifold.

Now let us consider switching manifolds as structures defined by time and state variables, as in the definition of $\mathcal{N} = M$, and in addition, depending upon a parameter $\alpha \in R^m$. Such a definition gives us the ability to parametrize families of manifold configurations by changing α. It is assumed that any local equation description of M, generically denoted by $m(x,t,\alpha)$ is such that $m(x,t,\alpha), x \in R^n, t \in R, \alpha \in R^m$, depends C^2-smoothly on its arguments. Similar to the method in [**66**], the optimal cost variation is obtained here by perturbing the manifold parameter around the nominal parameter. First consider the nominal manifold described locally by

$$m_{p,q}(x,t,\alpha) = 0, \quad x \in R^n, \alpha \in R^m, p, q \in Q \tag{5.1}$$

and the perturbed family by

$$m_{p,q}(x,t,\beta) = 0, \quad x \in R^n, \quad \beta \in \mathcal{N}(\alpha) \subset R^m, \tag{5.2}$$

where $\mathcal{N}(\alpha)$ is an open neighbourhood of α. Changes in the nominal switching manifold will generally result in deviations of the nominal hitting time and state of the optimal trajectory on the switching manifold from those of the nominal manifold. In this setting, the motivating problem of this work is to find values of α which infimize the total α dependent value function

$$inf_{\alpha\in\mathcal{R}^m} V(t_0,t_f,h_0,\alpha), \quad x \in R^n, \alpha \in R^m, \tag{5.3}$$

where

$$V(t_0,t_f,h_0,\alpha) = \inf_{u\in\mathcal{U}} J(t_0,t_f,h_0,u,\alpha), \tag{5.4}$$

which $J(t_0,t_f,h_0,u,\alpha)$ denotes the cost function corresponding to the switching manifold with parameter α.

A0: It is assumed that the discrete states (σ_0,σ_1) along the associated optimal trajectories do not depend upon α as α varies in R^m. It is further assumed that the minimizing values $(x^\alpha(.), u^\alpha(.), t^\alpha)$, where t^α, for (3.4) is unique. □

5.2. Definition of the Sensitivity Function

In this section we formulate the differentials of the state and adjoint variables with respect to the switching time. Here for both the nominal and perturbed manifolds, the controls are taken to be optimal. Let us write $x^\alpha(.), x^\beta(.)$ for the optimal state trajectories corresponding to the nominal and perturbed parameters respectively, and let $u^\alpha(.), u^\beta(.)$ be the associated optimal controls, i.e.

$$u^\alpha(.) = Argmin_{u\in\mathcal{U}} J(t_0,t_f,h_0,u,\alpha). \tag{5.5}$$

and similarly for $u^\beta(.)$. Define the state and adjoint variables variations for the nominal and perturbed manifold parameters as:

$$y(t) = \lim_{\delta t^\alpha \to 0} \frac{\delta x(t)}{\delta t^\alpha}, \qquad t \in [0, t_f] \tag{5.6}$$

$$z(t) = \lim_{\delta t^\alpha \to 0} \frac{\delta \lambda(t)}{\delta t^\alpha}, \qquad t \in [0, t_f] \tag{5.7}$$

where

$$\delta x(t) := x^\beta(t) - x^\alpha(t), \quad \delta\lambda(t) := \lambda^\beta(t) - \lambda^\beta(t), \quad \delta t^\alpha := t^\beta - t^\alpha. \tag{5.8}$$

□

We note that in the definition above $\delta t^\alpha \to 0$ does not necessarily imply $\beta \to \alpha$. The following result deals with the restricted single switching case.

THEOREM 5.1. *Under the standing assumption, **A0**, and the hypotheses of Theorem 2.1, consider a hybrid system possessing two modes:*

$$\dot{x}_1 = f_1(x_1(t), u_1(t)), \qquad t \in [0, t^\alpha]. \tag{5.9}$$

$$\dot{x}_2 = f_2(x_2(t), u_2(t)), \qquad t \in (t^\alpha, t_f], \tag{5.10}$$

and for which the cost function is defined as

$$J(0, t_f, h_0, u, \alpha) = \int_0^{t^\alpha} l_1(x_1(t), u_1(t))dt + \int_{t^\alpha}^{t_f} l_2(x_2(t), u_2(t))dt + h(x(t_f)) \tag{5.11}$$

where t^α is the switching time which is assumed to be unique.
Let the switching manifold be locally defined by $m(x, t, \alpha) = 0, x \in R^n, \alpha \in R^m$; then for the nominal manifold parameter α, the optimal state and adjoint variable variations with respect to the switching time are given by

$$y(t) = \int_0^t F_{1(x,\lambda)}(y(\tau), z(\tau))d\tau, \qquad t \in [0, t^\alpha), \tag{5.12}$$

$$y(t) = \int_0^{t^\alpha} F_{1(x,\lambda)}(y(\tau), z(\tau))d\tau + R_1 + \int_{t^\alpha}^{t} F_{2(x,\lambda)}(y(\tau), z(\tau))d\tau, \qquad t \in [t^\alpha, t_f], \tag{5.13}$$

and by

$$\begin{aligned} z(t) \quad &= \frac{\partial^2 h}{\partial x^2} y(t_f) + \int_{t^\alpha}^{t_f} H^2_{2(x,\lambda)}(y(\tau), z(\tau))d\tau + \int_t^{t^\alpha} H^2_{1(x,\lambda)}(y(\tau), z(\tau))d\tau \\ &+ \bar{H}_{(1,2)}(x, \lambda) + \frac{\partial p^\alpha \nabla_x m(x^\alpha, t^\alpha)}{\partial x^\alpha}(y(t^\alpha) + f_1(x(t^\alpha), \lambda_1^\alpha(t^\alpha))) \\ &+ \frac{\partial p^\alpha \nabla_x m(x^\alpha, t^\alpha)}{\partial \lambda_1}(z^-(t^\alpha) - \frac{\partial H_1}{\partial x}(x^\alpha(t^\alpha), \lambda_1^\alpha(t^\alpha))) \\ &+ \frac{\partial p^\alpha \nabla_x m(x^\alpha, t^\alpha)}{\partial \lambda_2}(z^+(t^\alpha) - \frac{\partial H_2}{\partial x}(x^\alpha(t^\alpha), \lambda_2^\alpha(t^\alpha))) \\ &+ \frac{\partial p^\alpha \nabla_x m(x^\alpha, t^\alpha)}{\partial t^\alpha}, \qquad t \in [0, t^\alpha), \end{aligned} \tag{5.14}$$

$$z(t) \quad = \quad \int_t^{t_f} H^2_{2(x,\lambda)}(y(\tau), z(\tau))d\tau + \frac{\partial^2 h}{\partial x^2}(x^\alpha(t_f))y(t_f), \qquad t \in [t^\alpha, t_f], \tag{5.15}$$

where p^α is the adjoint variable discontinuity parameter and

$$R_1 = f_1(x^\alpha(t^\alpha), \lambda_1^\alpha(t^\alpha)) - f_2(x^\alpha(t^\alpha), \lambda_2^\alpha(t^\alpha)), \tag{5.16}$$

$$F_{1(x,\lambda)}(y(t), z(t)) = \nabla_{(x,\lambda)} f_1(x^\alpha(t^\alpha), \lambda_1^\alpha(t^\alpha)).[y(t), z(t)]^T, \tag{5.17}$$

$$\bar{H}_{(1,2)}(x, \lambda) = \frac{\partial H_1}{\partial x}(x^\alpha(t^\alpha), \lambda_1^\alpha(t^\alpha)) - \frac{\partial H_2}{\partial x}(x^\alpha(t^\alpha), \lambda_2^\alpha(t^\alpha)). \tag{5.18}$$

$$H^2_{i(x,\lambda)}(y(t), z(t)) = \frac{\partial \nabla_{(x,\lambda)} H_i}{\partial x}(x^\alpha(t), \lambda_i^\alpha(t)).[y(t), z(t)]^T, \quad i = 1, 2. \tag{5.19}$$

Proof. See Appendix B □

5.3. Sensitivity of the Optimal Cost Functions

This section connects the previous formulation of $z(.), y(.)$ to the variation of the optimal hybrid cost (i.e. value function) with respect to manifold parameter variation. Using (5.4) we define the new function $\tilde{V}$ in order to study the variation of value function; we set

$$\tilde{V}(t, t_0, t_f, h_0) := inf_{u \in \mathcal{U}}(\int_0^t l_1(x(s), u(s))ds + \int_t^{t_f} l_2(x(s), u(s))ds + h(x(t_f)), \tag{5.20}$$

where t denotes the switching time from discrete state $q = 1$ to $q = 2$, but $x(t)$ is not assumed to lie on $m(., ., \alpha) = 0$. By the uniqueness assumption on the minimizing t^α in **A0**, the switching time t^α is

$$t^\alpha = t(\alpha) = Argmin_{0 \le t \le t_f}(\tilde{V}(t, t_0, t_f, h_0); \quad x(t) \in m_{p,q}(x, t, \alpha)), \tag{5.21}$$

where in this chapter $t(.)$ is assumed to be C^1 with respect to α. Then the value function corresponding to the manifold parameter α is

$$V(t_0, t_f, h_0, \alpha) = \tilde{V}(t, t_0, t_f, h_0)|_{t=t^\alpha}. \tag{5.22}$$

So

$$V(t_0, t_f, h_0, \alpha) = \int_0^{t^\alpha} l_1(x^\alpha(t), u^\alpha(t))dt + \int_{t^\alpha}^{t_f} l_2(x^\alpha(t), u^\alpha(t))dt + h(x^\alpha(t_f)), \tag{5.23}$$

and similarly for the manifold parameter β. Then we have the expression (5.24) below based on $y(.), z(.)$ for the optimal cost variation.

THEOREM 5.2 ([**71**]). *Subject to **A0** and the hypotheses of Theorem 2.1, the optimal cost variation with respect to the switching time, t^α of the hybrid system*

(2.3)

$$\begin{aligned}\frac{\partial \tilde{V}}{\partial t}|_{t=t^\alpha} \quad = \quad & \int_0^{t^\alpha} \frac{\partial l_1}{\partial x}((x^\alpha(t),\lambda^\alpha(t))y(t)+\frac{\partial l_1}{\partial \lambda}((x^\alpha(t),\lambda^\alpha(t))z(t)dt \\ & +(l_1(x^\alpha(t^\alpha),\lambda^\alpha(t^\alpha))-l_2(x^\alpha(t^\alpha),\lambda^\alpha(t^\alpha))) \\ & +\int_{t^\alpha}^{t_f} \frac{\partial l_2}{\partial x}((x^\alpha(t),\lambda^\alpha(t))y(t)+\frac{\partial l_2}{\partial \lambda}((x^\alpha(t),\lambda^\alpha(t))z(t)dt \\ & +\frac{\partial h}{\partial x}y(t_f), \end{aligned} \tag{5.24}$$

where $l_i(x^\alpha(.),\lambda^\alpha(.)), i = 1,2,$ is expressed as a function of optimal adjoint process $\lambda^\alpha(.)$.

Proof. The proof parallels the proof of Theorem 5.1 by extending the hybrid cost function along the nominal trajectory x^α. □

We may now compute the optimal cost variation as a function of the switching manifold parameter. The first step is to apply the chain rule to the value function variation to obtain (see [**71**])

$$\frac{\partial V}{\partial \alpha} = \frac{\partial V}{\partial t}|_{t=t^\alpha}\frac{\partial t^\alpha}{\partial \alpha} \tag{5.25}$$

where $\frac{\partial V}{\partial t}|_{t=t^\alpha}$ is the optimal cost derivative with respect to the switching time as presented in (5.24) and the second term is the switching time derivative with respect to the manifold parameter α. Taking derivatives of the nominal manifold equation with respect to the manifold parameter α, we have

$$\frac{\partial m(x,\alpha,t^\alpha)}{\partial x}\frac{\partial x}{\partial t^\alpha}\frac{\partial t^\alpha}{\partial \alpha}+\frac{\partial m(x,\alpha,t^\alpha)}{\partial t^\alpha}\frac{\partial t^\alpha}{\partial \alpha}+\frac{\partial m(x,\alpha,t^\alpha)}{\partial \alpha}=0 \tag{5.26}$$

So

$$\frac{\partial t^\alpha}{\partial \alpha} = -(\frac{\partial m(x,\alpha,t^\alpha)}{\partial x}\frac{\partial x}{\partial t^\alpha}+\frac{\partial m(x,\alpha,t^\alpha)}{\partial t^\alpha})^{-1}\frac{\partial m(x,\alpha,t^\alpha)}{\partial \alpha}, \tag{5.27}$$

subject to the assumption that $(\frac{\partial m(x,\alpha,t^\alpha)}{\partial x}\frac{\partial x}{\partial t^\alpha}+\frac{\partial m(x,\alpha,t^\alpha)}{\partial t^\alpha})$ is non zero and it is shown that

$$\frac{\partial x}{\partial t^\alpha}=y(t^\alpha)+f_1(x(t^\alpha),u(t^\alpha)),\quad 0\le t^\alpha\le t_f. \tag{5.28}$$

We conclude that the value function variation with respect to the manifold parameter perturbation is computable via (5.24),(5.25),(5.26),(5.27) and (5.28).

5.4. Example

Here we present an example in order to illustrate the results above. In the case below, since analytic solutions are not available, the optimal switching time and state for the nominal α are obtained numerically via the HMPC algorithm [**66**]. Consider the hybrid system with two modes given by:

$$\dot{x}(t)=x(t)+u(t),\qquad t\in[0,t^\alpha), \tag{5.29}$$

$$\dot{x}(t)=-x(t)+u(t),\qquad t\in[t^\alpha,2], \tag{5.30}$$

where the switching manifold is the following time varying structure:

$$m(x(t),\alpha,t)=x-t-\alpha=0,\quad t\in[0,2] \tag{5.31}$$

The cost function for the hybrid system is chosen to be

$$J(0,2,h_0,u)=\frac{1}{2}\int_0^{t^\alpha}u^2(t)dt+\frac{1}{2}\int_{t^\alpha}^2 u^2(t)dt, \tag{5.32}$$

where $L=1$ and $h_0=(0,1)$. Let us define the corresponding Hamiltonian function in the first discrete state to be

$$H_1(x(t),u(t),\lambda(t))=\frac{1}{2}u_1^2(t)+\lambda_1(t)(x_1(t)+u_1(t)),\qquad t\in[0,t^\alpha). \tag{5.33}$$

The optimal control is then given as the minimizer of the Hamiltonian function:

$$u^*(t)=Argmin_{u\in U}H(x^*(t),\lambda^*(t),u),\quad t\in[0,t^\alpha), \tag{5.34}$$

and we get the following equation relating the optimal control and optimal adjoint process

$$u_1^*(t) + \lambda_1^*(t) = 0, \quad t \in [0, t^\alpha). \tag{5.35}$$

The optimal state and optimal adjoint differential equations are

$$\dot{x^*}_1(t) = x_1^*(t) - \lambda^*(t), \qquad t \in [0, t^\alpha), \tag{5.36}$$

$$\dot{\lambda^*}_1(t) = -\lambda_1^*(t), \qquad t \in [0, t^\alpha). \tag{5.37}$$

In the second discrete state the Hamiltonian function is

$$H_2(x(t), u(t), \lambda(t)) = \frac{1}{2}u_2^2(t) + \lambda_2(t)(-x_2(t) + u_2(t)), \qquad t \in [t^\alpha, 2]. \tag{5.38}$$

As in the first discrete state the optimal control is derived by minimizing the Hamiltonian function, giving

$$u_2^*(t) + \lambda_2^*(t) = 0, \quad t \in [t^\alpha, 2], \tag{5.39}$$

and the optimal state and adjoint differential equations are then

$$\dot{x^*}_2(t) = -x_2^*(t) - \lambda_2^*(t), \qquad t \in [t^\alpha, 2], \tag{5.40}$$

$$\dot{\lambda^*}_2(t) = \lambda_2^*(t), \qquad t \in [t^\alpha, 2]. \tag{5.41}$$

By taking derivative of $y(.)$ and $z(.)$, given by Theorem 5.1, with resect to time we obtain

$$\dot{y}(t) = y(t) - z(t), \qquad t \in [0, t^\alpha), \tag{5.42}$$

$$\dot{y}(t) = -y(t) - z(t), \qquad t \in [t^\alpha, 2], \tag{5.43}$$

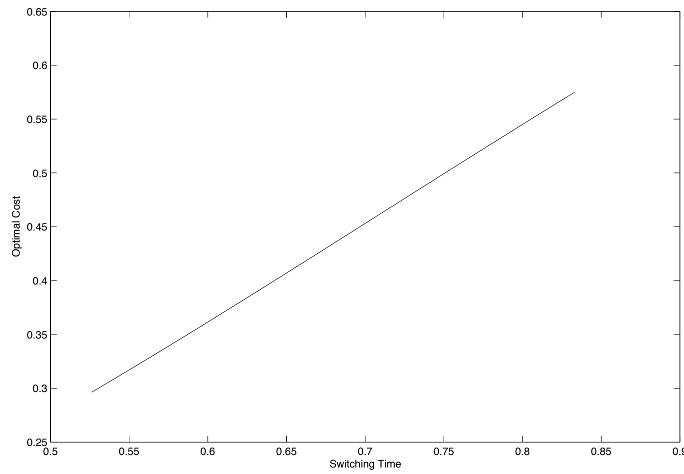

Figure 5.1. Optimal Cost as a Function of Switching Time

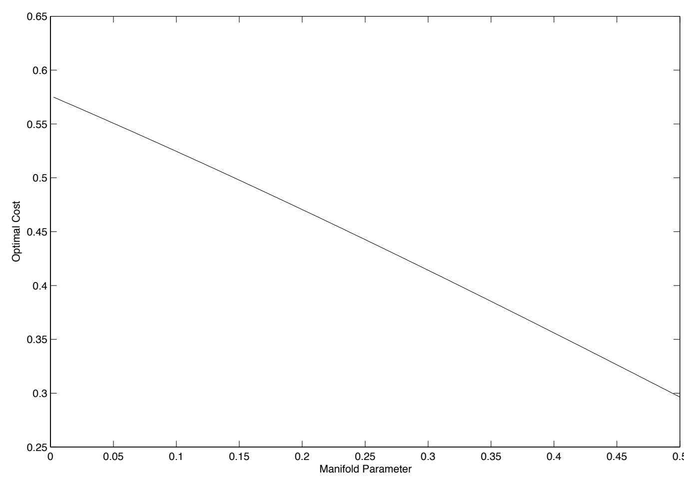

Figure 5.2. Optimal Cost as a Function of Manifold Parameter

$$\dot{z}(t) = -z(t), \qquad t \in [0, t^{\alpha}), \tag{5.44}$$

$$\dot{z}(t) = z(t), \qquad t \in [t^{\alpha}, 2], \tag{5.45}$$

$$y(0) = y(2) = 0. \tag{5.46}$$

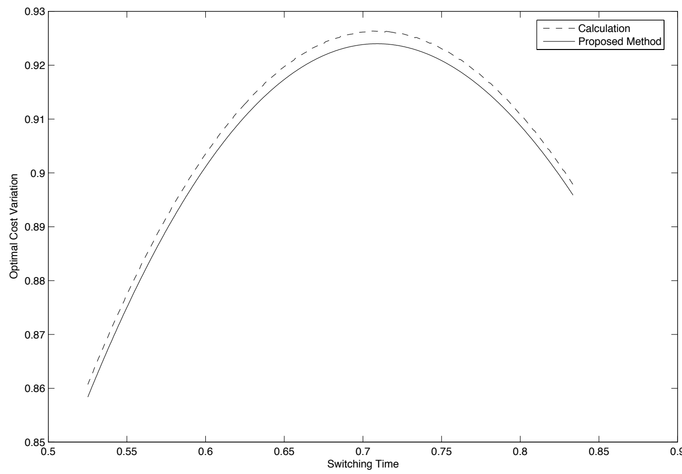

FIGURE 5.3. Optimal Cost Derivative versus Switching Time

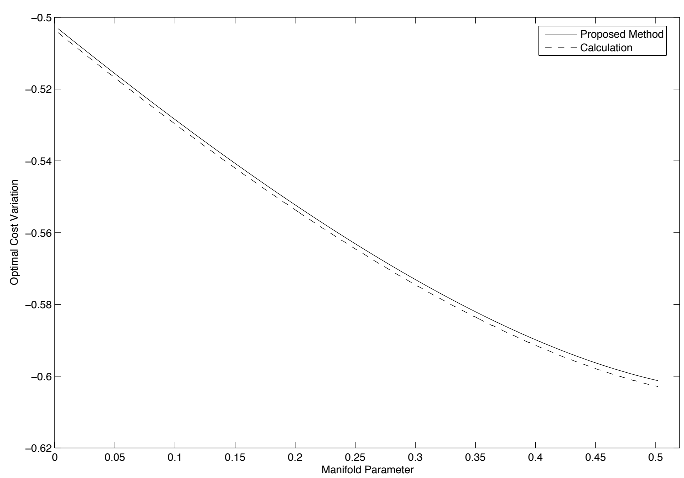

FIGURE 5.4. Optimal Cost Derivative versus Manifold Parameter

For this numerical example we consider $x(0) = 1, x(2) = 0$; then for $\alpha = 0$, for instance, the HMP algorithm [**66**] yields the optimal switching time and state values

$$t^{\alpha} = 0.834, \quad x^{\alpha} = 0.834, \tag{5.47}$$

and the value function at $\alpha = 0$ is

$$V(0, 2, h_0, 0) = 0.5624. \tag{5.48}$$

The adjoint variables for the first and second system phases at the switching time are as follows:

$$\lambda_1(t^\alpha) = 0.6825, \quad \lambda_2(t^\alpha) = 0.1796 \tag{5.49}$$

And furthermore

$$\frac{\partial p_1^\alpha \nabla_x m(x^\alpha, t^\alpha)}{\partial \lambda_1}(z^-(t^\alpha) - \frac{\partial H_1}{\partial x}(x^\alpha(t^\alpha)) = 6.4830.10^{-4},$$
$$\frac{\partial p_1^\alpha \nabla_x m(x^\alpha, t^\alpha)}{\partial \lambda_2}(z^+(t^\alpha) - \frac{\partial H_2}{\partial x}(x^\alpha(t^\alpha)) = 1.3744, \tag{5.50}$$
$$\frac{\partial p_1^\alpha \nabla_x m(x^\alpha, t^\alpha)}{\partial x^\alpha}(y(t^\alpha) + f_1(x(t^\alpha), \lambda_1^\alpha(t^\alpha))) = 1.0157. \tag{5.51}$$

The boundary conditions are respectively:

$$y(0) = y(2) = 0, \tag{5.52}$$

$$y^-(t^\alpha) = y^+(t^\alpha) - 1.1657, \tag{5.53}$$

and

$$0.9994z^-(t^\alpha) = 2.3744z^+(t^\alpha) + 1.0157y^-(t^\alpha) + 1.2624. \tag{5.54}$$

The final solution for the state and adjoint variation function are

$$y(t) = 0.9989\frac{e^{-t} - e^t}{2}, \qquad t \in [0, .834), \tag{5.55}$$

and

$$y(t) = 0.2327e^{-t+0.834} - 0.025(e^{t-0.834} - e^{-t+0.834}), \qquad t \in [0.834, 2]. \tag{5.56}$$

with

$$z(t) = 0.9989e^{-t}, \qquad t \in [0, .834), \tag{5.57}$$

$$z(t) = 0.05e^{t-0.834}, \qquad t \in [.834, 2]. \tag{5.58}$$

The state and adjoint variables variations have been computed so far with respect to the switching time variation. This is motivated by the simplicity of the use of the chain rule in (5.25) for the computation of $\frac{\partial V}{\partial \alpha}$. The optimal cost variation based on the switching time variation is then given by

$$\begin{aligned}
\frac{\partial \tilde{V}}{\partial t}|_{t=t^\alpha} &= \int_0^{0.834} \lambda(\tau)z(\tau)d\tau + \int_{0.834}^2 \lambda(\tau)z(\tau)d\tau + \frac{1}{2}(\lambda_-^2(0.834) - \lambda_+^2(0.834)) \\
&= \int_0^{0.834} 1.5715 \times 0.9989e^{-2\tau}d\tau \\
&\quad + \int_{0.834}^2 0.1796 \times 0.05e^{2(\tau-.834)}d\tau + 0.2168,
\end{aligned} \tag{5.59}$$

where $t^\alpha = 0.834$ and $l_1(x(t), u(t)) = \frac{1}{2}u^2(t) = \frac{1}{2}\lambda_1^2(t)$ and $l_2(x(t), u(t)) = \frac{1}{2}u^2(t) = \frac{1}{2}\lambda_2^2(t)$. Now

$$\frac{\partial t^\alpha}{\partial \alpha} = -(1 - \frac{\partial x}{\partial t^\alpha})^{-1}, \tag{5.60}$$

where

$$\frac{\partial x}{\partial t^\alpha} = y(t^\alpha) + f_1(x^\alpha, \lambda^\alpha) = -.9331 + .834 - .6825 = -0.7816. \tag{5.61}$$

So we see that at $\alpha = 0$, $\frac{\partial V}{\partial \alpha}$ can be easily computed and so it is plausible that starting from an initial nominal value for α one could find the optimal α by a gradient method. In Figure 5.1 the optimal cost function is displayed as a function of the switching time t^α. In this example we vary α between 0 to 0.5 and the optimal cost is then obtained as a function of the manifold parameter.

Figure 5.2 shows the optimal cost plotted against the manifold parameter α, while Figure 5.3 displays the optimal cost variation (i.e. derivative) displayed as a function of the switching time t^α. In Figure 5.4 the optimal cost variation (i.e. derivative) as a function of the manifold parameter α is shown. In both Figures 5.3 and 5.4 numerical results obtained from both the method above and by direct calculation are indicated by the solid and dashed lines respectively. The close approximation of each curve by the other corroborates the theory presented in this chapter. As Figure 5.2 indicates the optimal hybrid value function is attained on the boundary of the manifold parameter set, i.e. $\alpha = 0.5$. In general a local optimal manifold parameter α^o satisfies at least one of the following conditions:

$$\frac{\partial v}{\partial x}(x(t(\alpha^o)), t(\alpha^o)) = 0, \text{or} \quad \frac{\partial m}{\partial \alpha}(x(t(\alpha^o)), t(\alpha^o)) = 0. \tag{5.62}$$

5.5. Geometrical Representation of Switching Manifolds and Optimality

5.5.1. Geometrical Preliminaries. As stated in the literature, for a hybrid system with state space R^{n+1}, time invariant switching manifolds are defined as n dimensional surfaces. Considering hybrid trajectories in state$\times$time space, time variant and time invariant switching manifolds are both defined as $n+1$ dimensional surfaces in R^{n+2}, see e.g. [**65**]. Based on the discussion above, we define an n dimensional smooth differentiable manifold M as a surface for which the local coordinate chart and the associated neighbourhood will be generically defined by (φ, V) , where φ is a local homeomorphism, see [**14**, **40**].

$$\varphi : V_M \subset M \to R^n, \quad x = \varphi^{-1} : U_{R^n} \subset R^n \to M \subset R^{n+1}, \tag{5.63}$$

where V_M and U_{R^n} are open sets in M and R^n respectively. The definition above is helpful as we mostly work with x instead of φ. The tangent space $T_x M$, at $x \in M$, is then represented as a linear space spanned by the vectors $[\frac{\partial x}{\partial x^i}, \quad i = 1, ..., n]$, see [**40**]. Using the same notation as in [**40**], by the assumption of differentiability of x, the fundamental forms of the switching manifold in the local coordinates are defined as:

$$g_{ij} = \langle \frac{\partial x}{\partial x^i}, \frac{\partial x}{\partial x^j} \rangle, \quad i, j = 1, ..., n, \tag{5.64}$$

$$H_{ij} = \langle N, \frac{\partial^2 x}{\partial x^i \partial x^j} \rangle = -\langle \frac{\partial N}{\partial x^i}, \frac{\partial x}{\partial x^j} \rangle, \quad i, j = 1, ..., n. \tag{5.65}$$

where $\langle .,. \rangle$ is the inner product in R^{n+1}, g and h are the first and second fundamental forms and N is the unit length normal vector, [**40**]. By reversing the direction of the definition above, equation (5.64) can be interpreted as the Riemannian metric defined for general Riemannian manifolds, that is to say, manifolds which are not necessarily embedded in Euclidean space, [**40**]). Let Y be a differentiable vector field, defined on an open set of R^{n+1}, and let X be a fixed directional vector at some fixed point p of this open set. Then the directional derivative of the vector field Y in the direction X is defined as

$$D_X Y|_p := \lim_{t \to 0} \frac{1}{t}(Y(p + tX) - Y(p)). \tag{5.66}$$

Here the vector fields X, Y are not necessarily tangential to the manifold and are defined as $X, Y \in T_x(R^{n+1}) \cong R^{n+1}$. Moreover, in the special case where $X, Y \in T_x M$, the directional derivative is itself not necessarily tangential to the manifold; to be specific the tangential derivative on the manifold, i.e. the covariant derivative of Y in the X direction, is defined to be the tangential component of the directional derivative defined above. Then the covariant derivative has the following definition

$$\nabla_X Y := D_X Y - \langle D_X Y, N \rangle N. \tag{5.67}$$

The last term above is the normal component of the directional derivative which is subtracted from the total vector, $D_X Y$. Now we briefly introduce the Christoffel symbols (see [**14, 40**]). Consider a tangential vector field Y as

$$Y = \sum_{j=1}^{n} \eta^j \frac{\partial x}{\partial x^j}. \tag{5.68}$$

The covariant derivative of Y in direction X, where $X = \sum_{i=1}^{n} \xi^i \frac{\partial x}{\partial x^i}$ is

$$\nabla_X Y = \sum_{i=1}^{n} \xi^i (\sum_{j=1}^{n} \nabla_{\frac{\partial x}{\partial x^i}} (\eta^j \frac{\partial x}{\partial x^j})). \tag{5.69}$$

Based on the tangential property of the covariant derivative which is given by (5.67), we define

$$\nabla_{\frac{\partial x}{\partial x^i}} \frac{\partial x}{\partial x^j} := \sum_k \Gamma_{ij}^k \frac{\partial x}{\partial x^k}, \tag{5.70}$$

and

$$\Gamma_{ij,k} := \langle \nabla_{\frac{\partial x}{\partial x^i}} \frac{\partial x}{\partial x^j}, \frac{\partial x}{\partial x^k} \rangle, \tag{5.71}$$

where $\Gamma_{ij,k}$ and Γ_{ij}^k are called first and second kind of the Christoffel symbols, see [**40**]. Using the notation above we get:

$$\Gamma_{ij,k} = \sum_{m=1}^{n} \Gamma_{ij}^m g_{mk}. \tag{5.72}$$

Hence (see [**40**])

$$\nabla_X Y = \sum_{i,k=1}^{n} \xi^i (\frac{\partial \eta^k}{\partial x^i} + \sum_{j=1}^{n} \eta^j \Gamma_{ij}^k) \frac{\partial x}{\partial x^k}. \tag{5.73}$$

5.5.1.1. *Necessary Conditions for Optimality.* Here in this subsection we represent the results presented in [**80**] and formulate the necessary conditions of optimality of autonomous hybrid systems consisting of two modes associated with a time invariant switching manifold. The *hybrid value function* is defined as follows:

$$v(x,t) = inf_{u \in \mathcal{U}} J(t_0, t_f, h_0; x_{t_s}, t_s, u)|_{t_s = t, x_{t_s} = x}, x_{t_s} \in M.$$

This formulation shows that the hybrid value function depends on the switching state and switching time on the switching manifold M. The next lemma gives necessary conditions for optimality of HSOC problems.

LEMMA 5.1. *Let* $(x(t_s), t_s) = (x(t_s^o), t_s^o)$ *be the optimal switching state and time subject to the hypotheses for the HSOC problem and the hypotheses of Theorem 5.1, then* $(x(t_s), t_s)$ *satisfies the local stationarity conditions.*

$$\frac{\partial v(x,t)}{\partial t}|_{(x(t_s),t_s)} = 0, \tag{5.74}$$

$$\nabla_x v(x,t)|_{(x(t_s),t_s)} \perp T_{x(t_s)} M, \tag{5.75}$$

where $T_{x(t_s)}M$ is the tangent space at the switching state on the manifold M.

Proof. By the Principle of Optimality we know that the variation of the value function around the optimal time state pair must be positive therefore

$$v(x,t) - v(x(t_s),t_s) \geq 0, \quad x \in \mathcal{B}_{x(t_s)}(\eta), t \in (t_s - \epsilon, t_s + \epsilon), \tag{5.76}$$

where $\mathcal{B}_{x(t_s)}(\eta)$ is defined by the induced topology on the manifold $m(x) = 0$. The value function variation is written as the summation of variations around the optimal switching time and the optimal state. The point here is that the state variation must be defined with respect to the switching manifold, then:

$$\begin{aligned} v(x,t) - v(x(t_s),t_s) \quad &= \nabla_x v(x(t_s),t_s).(x - x(t_s)) \\ &+ \frac{\partial v(x,t)}{\partial t}|_{(x(t_s),t_s)} \delta t + o(\delta x) + o(\delta t) \geq 0, \end{aligned} \tag{5.77}$$

Fixing $x = x(t_s)$ and varying t we have

$$\frac{\partial v(x,t)}{\partial t}|_{(x(t_s),t_s)} = 0. \tag{5.78}$$

Setting $t = t_s$ i.e. $\delta t = 0$, (5.77) yields $\nabla_x v(x(t_s),t_s).\psi = \sum_{j=1}^{n} \psi_j \frac{\partial v(x(t_s),t_s)}{\partial x_j} \geq 0$, $\forall \psi \in T_{x(t_s)} m(x)$ which completes the proof since $\psi, -\psi \in T_{x(t_s)} m(x)$. □

It should be noted that switching manifolds in this chapter are considered time invariant and we can see very close similarities between (5.75) and (3.1). The next lemma presents the relation between these equations.

Lemma 5.2 ([**1**,**72**]). *For the HSOC problem defined in Theorem 5.1 the following relations hold:*

$$\frac{\partial v(x,t)}{\partial t}|_{(x(t_s),t_s)} = H_1(t_s^-) - H_2(t_s), \tag{5.79}$$

$$\nabla_x v(x(t_s), t_s) = \lambda_2(t_s) - \lambda_1(t_s^-). \tag{5.80}$$

□

The results of Lemma 5.1 and the chain rule imply

$$\frac{\partial v(x^o, t^o)}{\partial x^i} = \langle \nabla_x v(x^o, t^o), \frac{\partial x}{\partial x^i} \rangle = 0. \tag{5.81}$$

The following theorem is a generalization of a theorem which appeared in [**80**].

THEOREM 5.3. *At the optimal switching state x^o and optimal switching time t^o we have*

$$\begin{aligned} -H_{ik} &= \mu \frac{\partial x^T}{\partial x^i} \frac{\partial^2 v(x^o, t^o)}{\partial x^2} \frac{\partial x}{\partial x^k} + T_i \frac{\partial x}{\partial x^k} \\ &= \mu \frac{\partial x^T}{\partial x^k} \frac{\partial^2 v(x^o, t^o)}{\partial x^2} \frac{\partial x}{\partial x^i} + T_k \frac{\partial x}{\partial x^i}, \end{aligned} \tag{5.82}$$

*where $T_i, T_k \in T_x M$ and μ is the discontinuity parameter appearing in the adjoint process boundary condition at the switching time, i.e. $\lambda_2 - \lambda_1^- = \mu^{-1} N(x^o)$, see [**58, 66**].*

PROOF. Applying Lemma 5.1 we have

$$\nabla_x v(x^o, t^o) = \mu^{-1} N(x^o), \quad \mu \in R. \tag{5.83}$$

Since the normal vector $N(x)$ is unit length we have

$$\frac{d}{dx^i}(N^T(x).N(x)) = N^T(x).N_{x^i}(x) + N^T_{x^i}(x).N(x) = 0, \quad i = 1, ...n, \tag{5.84}$$

where

$$N_{x^i}(x) = \lim_{\delta x^i \to 0} \frac{N(x + \delta x^i) - N(x)}{\delta x^i}, \quad i = 1, ...n. \tag{5.85}$$

Then $\langle N(x), N_{x^i}(x) \rangle = N^T(x).N_{x^i}(x) = N^T_{x^i}(x).N(x) = 0, \quad i = 1, ...n$, and,

$$N^T(x).\frac{\partial x}{\partial x^i} = 0, \quad i = 1, ...n. \tag{5.86}$$

The derivative of the last equation with respect to x^i implies

$$\langle N_{x^i}(x), \frac{\partial x}{\partial x^j}\rangle = N_{x^i}^T(x).\frac{\partial x}{\partial x^j} = -N^T(x).\frac{\partial^2 x}{\partial x^i x^j} = -H_{ij}, \quad i,j = 1,...n. \tag{5.87}$$

perturbing x^i to $x^i + \delta x^i$ gives $\nabla_x v(x^o + \delta x, t^o)$ in the new coordinate $x^o + \delta x$. Since by the assumed uniqueness of the optimal trajectory the new switching state is not optimal, the perturbed vector is not necessarily normal to the switching manifold . The normal vector at $x^o + \delta x$ is estimated as follows:

$$N(x^o + \delta x^o) = (\mu + \delta\mu)\nabla_x v(x^o + \delta x, t^o) + \delta\psi_{x^i}, \ \ \delta\psi_{x^i} \in T_{x^o+\delta x^o}M \tag{5.88}$$

hence

$$N_{x^i} = \lim_{\delta x^i \to 0} \frac{(\mu + \delta\mu)\nabla_x v(x^o + \delta x, t^o) + \delta\psi_{x^i} - \mu\nabla_x v(x^o, t^o)}{\delta x^i}, \tag{5.89}$$

where $\delta\psi_{x^i} \in T_{x^o+\delta x^i}M, \quad i = 1, ..., n.$

Finally, (5.87) together with (5.89) yield

$$-H_{ik} = \mu_{x^i}\nabla_x v(x^o, t^o).\frac{\partial x}{\partial x^k} + \mu\frac{\partial^T x}{\partial x^i}.\nabla^2_{(x,x)}v(x^o, t^o).\frac{\partial x}{\partial x^k} + T_i.\frac{\partial x}{\partial x^k}, \tag{5.90}$$

where

$$T_i = \lim_{x^i \to 0} \frac{\delta\psi_{x^i}}{\delta x^i}, \quad i = 1, ..., n. \tag{5.91}$$

The optimality of (x^o, t^o), i.e (5.75), yields

$$\nabla_x v(x^o, t^o).\frac{\partial x}{\partial x^k} = 0, \quad k = 1, ..., n, \tag{5.92}$$

and hence

$$-H_{ik} = \mu\frac{\partial^T x}{\partial x^i}.\nabla^2_{(x,x)}v(x^o, t^o).\frac{\partial x}{\partial x^k} + T_i.\frac{\partial x}{\partial x^k}, \tag{5.93}$$

which proves the theorem statement. By changing the role of i and k and symmetrical properties of the Hessian matrix and second fundamental forms, the second statement is proved. □

PROPOSITION 1. *In the local coordinates of the optimal switching state x^o we have*

$$\frac{\partial x^T}{\partial x^i}\frac{\partial^2 v(x^o,t^o)}{\partial x^2}\frac{\partial x}{\partial x^k} = \frac{\partial^2 v(x^o,t^o)}{\partial x^i \partial x^k} - \mu^{-1}H_{ik}, \;\; i,k = 1,...,n. \tag{5.94}$$

PROOF. In the local coordinates of x^o we have

$$\frac{\partial v(x,t^o)}{\partial x^i} = \frac{\partial v(x,t^o)^T}{\partial x}\frac{\partial x}{\partial x^i}, \tag{5.95}$$

taking derivative with respect to x^k we obtain

$$\frac{\partial^2 v(x,t^o)}{\partial x^i \partial x^k} = \frac{\partial}{\partial x^k}(\frac{\partial v(x,t^o)}{\partial x})^T\frac{\partial x}{\partial x^i} + (\frac{\partial v(x,t^o)^T}{\partial x})\frac{\partial^2 x}{\partial x^i x^k}. \tag{5.96}$$

where we have

$$\frac{\partial}{\partial x^k}(\frac{\partial v(x^o,t^o)}{\partial x})^T = (\frac{\partial^T x}{\partial x^k})\frac{\partial^2 v(x^o,t^o)}{\partial x^2}. \tag{5.97}$$

(5.96),(5.99),(5.83) and (5.87) imply (5.94) and completes the proof. □

The vectors T_i, T_k which appear in the statement of Theorem 5.3 are still unknown and needed to be computed. The following lemma and theorem give us a mathematical description of those vectors in a specific coordinate (coordinates of the sensitivity function of the value function) system around the optimal switching state.

LEMMA 5.3. *The local components of the vector $T_i = \sum_{j=1}^n t_{ij}\frac{\partial x}{\partial x^j}, \quad i = 1,..,n$ in Theorem 5.3 satisfy*

$$t_{ij} = -\mu\frac{\partial \eta^i}{\partial x^j}, \quad i,j = 1,...,n, \tag{5.98}$$

where $\bar{\eta} = [\eta^1,...,\eta^n]$ is the tangential local representation of the sensitivity function of the value function with respect to the switching state at the optimal switching state.

PROOF. As shown in the proof of Theorem 5.3, $T_i \in T_{x^o}M$, $i = 1, ..., n$, then by the symmetryl properties of the Hessian matrix, the definition of the second fundamental forms H_{ij} and (5.93) we have

$$\langle T_i, \frac{\partial x}{\partial x^k} \rangle = T_i . \frac{\partial x}{\partial x^k} = T_k . \frac{\partial x}{\partial x^i} = \langle T_k, \frac{\partial x}{\partial x^i} \rangle \quad i, k = 1, ..., n. \tag{5.99}$$

The sensitivity of the value function with respect to state x is defined in a neighbourhood of the optimal state x^o with the following local representation:

$$\frac{\partial v(x, t^o)}{\partial x} = \sum_{j=1}^{n} \eta^j \frac{\partial x}{\partial x^j} + \eta^{n+1} N, \tag{5.100}$$

The derivative of (5.100) with respect to x^k is

$$\frac{\partial}{\partial x^k} \left(\frac{\partial v(x, t^o)}{\partial x} \right) = \frac{\partial^2 v(x, t^o)}{\partial x^2} \frac{\partial x}{\partial x^k}$$

$$= \sum_{j=1}^{n} \left(\frac{\partial \eta^j}{\partial x^k} \frac{\partial x}{\partial x^j} + \eta^j \frac{\partial^2 x}{\partial x^j \partial x^k} \right) + \frac{\partial \eta^{n+1}}{\partial x^k} N + \eta^{n+1} \frac{\partial N}{\partial x^k}. \tag{5.101}$$

Based on the *Gauss formula*, see [**40**], we have

$$\frac{\partial^2 x}{\partial x^j \partial x^k} = \sum_{l=1}^{n} \Gamma^l_{jk} . \frac{\partial x}{\partial x^l} + H_{jk} . N \quad j, k = 1, ..., n, \tag{5.102}$$

Applying the inner product with $\frac{\partial x}{\partial x^i}$ on (5.101) and using (5.102) it follows

$$\frac{\partial x^T}{\partial x^i} \frac{\partial^2 v(x, t^o)}{\partial x^2} \frac{\partial x}{\partial x^k} = \sum_{j=1}^{n} \left(\frac{\partial \eta^j}{\partial x^k} g_{ij} + \eta^j \sum_{l=1}^{n} \Gamma^l_{jk} g_{li} \right) - \eta^{n+1} H_{ki}. \tag{5.103}$$

Hence from (5.93) and (5.99) we have

$$-\mu^{-1} H_{ik} - \mu^{-1} T_k . \frac{\partial x}{\partial x^i} = \sum_{j=1}^{n} \left(\frac{\partial \eta^j}{\partial x^k} g_{ij} + \eta^j \sum_{l=1}^{n} \Gamma^l_{jk} g_{li} \right) - \eta^{n+1} H_{ki}, \tag{5.104}$$

where

$$T_k . \frac{\partial x}{\partial x^i} = \sum_{j=1}^{n} t_{kj} g_{ij}. \tag{5.105}$$

Using (5.105),(5.104) and the fact that at the optimal switching state $\bar{\eta} = [\eta^1, ..., \eta^n] = 0$ (tangential components of the sensitivity function), $\mu^{-1} = \eta^{n+1}$ (Lemma 5.1) we get

$$t_{kj} = -\mu \frac{\partial \eta^j}{\partial x^k}. \tag{5.106}$$

The same argument evidently holds with i replacing k in the statement. □

It should be mentioned here that all the results above are valid only at the optimal switching state which optimize (minimize) the hybrid value function on the switching manifold. The following theorem gives the behaviour of the tangential component of the value function sensitivity on a neighbourhood of the optimal state at the switching manifold.

THEOREM 5.4. *Consider a vector field* $Y = \frac{\partial v(x,t^o)}{\partial x}$ *on a neighbourhood of an optimal switching state* x^o *of the switching manifold, then there exist local coordinates,* $x^1, ..., x^n$*, in which the tangential components of the defined vector field satisfy the following differential equations:*

$$\frac{\partial \bar{\eta}}{\partial x^k} = G^{-T} \begin{pmatrix} \Omega_{1,1,k} & \dots & \Omega_{1,n,k} \\ \vdots & \vdots & \vdots \\ \Omega_{n,1,k} & \dots & \Omega_{n,n,k} \end{pmatrix}^T \bar{\eta} + G^{-T}[0, ..., 2_k, 0, ..., 0]^T, k = 1, ..., n, \tag{5.107}$$

where

$$\bar{\eta} = [\eta^1, ..., \eta^n], \quad G = [g_{ij}], \tag{5.108}$$

and

$$\Omega_{i,j,k} = \frac{1}{2} \sum_{l=1}^{n} (\Gamma^l_{ik} g_{lj} + \Gamma^l_{jk} g_{li}), \quad i, j, k = 1, ..., n, \tag{5.109}$$

and $[0, ..., 2_k, 0, ..., 0]$ *has 2 in the kth position.*

PROOF. Using the statement of Proposition 1 and (5.103) we have

$$\frac{\partial^2 v(x,t^o)}{\partial x^i \partial x^k} - \mu^{-1} H_{ik} = \sum_{j=1}^{n} (\frac{\partial \eta^j}{\partial x^k} g_{ij} + \eta^j \sum_{l=1}^{n} \Gamma^l_{jk} g_{li}) - \eta^{n+1} H_{ki}.$$

By Lemma 5.1 we have $\mu^{-1} = \eta^{n+1}$ and the symmetry properties of the second fundamental form give $H_{ik} = H_{ki}$, then

$$\frac{\partial^2 v(x,t^o)}{\partial x^i \partial x^k} = \sum_{j=1}^{n} (\frac{\partial \eta^j}{\partial x^k} g_{ij} + \eta^j \sum_{l=1}^{n} \Gamma^l_{jk} g_{li}). \tag{5.110}$$

By an application of the Morse Lemma, [**50**,**52**], and the positivity assumption of the value function Hessian matrix, there exist a system of local coordinates around the optimal state x^o such that

$$\frac{\partial^2 v(x,t^o)}{\partial x^i \partial x^k} = 2\delta_{ik}, \quad i,k = 1,...,n. \tag{5.111}$$

where δ_{ik} is the Kronecker delta. (5.110) and (5.111) together prove the theorem statement. □

The coordinates in the Theorem 5.4 are the Morse coordinates which appear in the Morse theory for critical points of a scalar function defined on a manifold [**50**,**52**]. One direct result of Theorem 5.4 is that if two value functions are minimized at the same point x on a manifold and have the same Morse coordinates then they both have same tangential components on the switching manifold. Combining the results of Lemma 5.1, Theorem 5.4 and (5.110) we get:

$$[t_{li}]_{l=1,...,n} = -\mu G^{-1}[0,0,...,0,2_i,0,,...,0]^T, \quad i = 1,...,n. \tag{5.112}$$

In order to compute the tangential vector $\bar{\eta}$ we apply a method similar to the state transition matrix in the linear systems. The main point here is that we have n different coordinate variations, consequently we propose the following method which starts from the initial condition 0 and will finally attains the coordinates $x^1,...,x^n$ in the neighbourhood of the optimal state x. Let us denote $\varphi(x^1,...,x^n,x^1(x),...,x^n(x))$

as the state transition matrix corresponding to the linear system described in Theorem 5.4. Starting from the optimal point we have [**72**]

$$\begin{aligned}\bar{\eta}(x^1, 0, ..., 0) &= \varphi(x^1, 0, ..., 0)\bar{\eta}(0, ..., 0) \\ &+ \int_0^{x^1} \varphi(x^1, 0, ..., 0, \sigma, 0, ..., 0) G^{-T}[2, 0, ..., 0]^T d\sigma, \\ &= \int_0^{x^1} \varphi(x^1, 0, ..., 0, \sigma, 0, ..., 0) G^{-T}[2, 0, ..., 0]^T d\sigma, \end{aligned} \tag{5.113}$$

now applying the same method to the new initial condition $\bar{\eta}(x^1, 0, ..., 0)$ we have

$$\begin{aligned}\bar{\eta}(x^1, x^2, 0, ..., 0) &= \varphi(x^1, x^2, 0, ..., x^1, 0, ..., 0)\bar{\eta}(x^1, 0, ..., 0) \\ &+ \int_0^{x^2} \varphi(x^1, x^2, ..., x^1, \sigma, 0, ..., 0) G^{-T}[0, 2, 0, ..., 0]^T d\sigma,\end{aligned}$$

and finally we have

$$\begin{aligned}&\bar{\eta}(x^1, x^2, ..., x^n) = \\ &\varphi(x^1, x^2, ..., x^n, x^1, , ..., u^{n-1}, 0)\bar{\eta}(x^1, x^2, ..., u^{n-1}, 0) \\ &+ \int_0^{x^n} \varphi(x^1, x^2, ..., x^n, x^1, x^2, ..., u^{n-1}, \sigma) G^{-T}[0, ..., 2]^T d\sigma.\end{aligned}$$

Remark: In this chapter the value function arguments are x, t which are assumed to be independent of the local representation of the switching manifold which is also assumed to be time invariant and independent of the switching time t. Perturbing the switching state x to $x + \delta x$ generally results in a change to the optimal switching time t. Employing the hypotheses introduced in the first section, the switching state x and switching time t are considered independently therefore the sensitivity formulas derived before are valid as long as x and t are treated independently. If the value function is defined as the infimized cost with respect to the switching time for a given state on the switching manifold by, i.e.

$$\tilde{v}(x) = inf_{t \in [0, t_f]} v(x, t), \tag{5.114}$$

then the sensitivity of $\tilde{v}(.)$ with respect to the switching state x is

$$\frac{d\tilde{v}(x)}{dx} = \frac{\partial v(x,t)}{\partial x} + \frac{\partial v(x,t)}{\partial t}\frac{\partial t}{\partial x}, \tag{5.115}$$

where $\tilde{v}(x) = v(x, t(x))$. At the optimal (x^o, t^o), we have $\frac{\partial v(x^o,t^o)}{\partial t} = 0$ then

$$\frac{d\tilde{v}(x^o)}{dx} = \frac{\partial v(x^o,t^o)}{\partial x} = \mu^{-1}N. \tag{5.116}$$

Henceforth the argument of Theorem 5.3 holds for $\tilde{v}(.)$. Finally, Theorem 5.4 statment holds for the value function $\tilde{v}(.)$.

5.6. Local Variation of the Value Function

In this section we present a numerical method in order to compute the second order derivative of the value function $v(x,t)$ with respect to the coordinates $x^1, ..., x^n$. As proved in Lemma 5.1, the sensitivity of the value function with respect to the state variable is the discontinuity of the adjoint process at the switching time. Based on a method in [**66**], the optimal state and adjoint process are solutions of a two point boundary value problem which satisfy the boundary conditions at the final time and on the switching manifold. Here in this chapter the hybrid problem is considered in the context of the fixed end point optimal control therefore the total hybrid trajectory is a sequence of solutions of two point boundary value problems between the switching state pairs on switching manifolds, see [**66**].

By the controllability assumption of the hybrid system, we only perturb the ith coordinate of the optimal switching state x^o where the switching time is fixed. The state and adjoint variable variations with respect to the corresponding perturbation are defined similar to [**71**] as follows:

$$y_i^x(t) = \lim_{\delta x^i \to 0} \frac{\delta x(t)}{\delta x^i}, \qquad i = 1, ..., n, t \in [0, t_f], \tag{5.117}$$

$$z_i^x(t) = \lim_{\delta x^i \to 0} \frac{\delta \lambda(t)}{\delta x^i}, \qquad i = 1, ..., n, t \in [0, t_f], \tag{5.118}$$

where $y_i^x(t), z_i^x(t)$ satisfy the following differential equations:

$$\dot{y}_i^x(t) = \frac{\partial f_q}{\partial x} y_i^x(t) + \frac{\partial f_q}{\partial \lambda} z_i^x(t), \qquad t \in [t_{s-1}, t_s), \tag{5.119}$$

$$\dot{z}_i^x(t) = -\frac{\partial^2 H_q}{\partial x^2} y_i^x(t) - \frac{\partial^2 H_q}{\partial \lambda \partial x} z_i^x(t), \qquad t \in [t_{s-1}, t_s), \tag{5.120}$$

where H_q is the Hamiltonian function in the time interval $[t_{s-1}, t_s)$. Since the final state is given, the boundary conditions for (5.117) and (5.118) are given as follows:

$$y_i^x(0) = y_i^x(t_f) = 0, \quad y_i^x(t_s) = \frac{\partial x}{\partial x^i}. \tag{5.121}$$

The following Lemma gives the second derivative of the value function evaluated at optimal switching states on switching manifolds.

Lemma 5.4. *At an optimal switching state, the local Hessian matrix of the value function of the hybrid system satisfies the following equations ($i, k = 1, ..., n$):*

$$\begin{aligned} \frac{\partial^2 v(x,t)}{\partial x^i \partial x^k} &= y_i^x(t_s)(z_k^x(t_s) - z_k^x(t_s^-)) + \mu^{-1} H_{ik} \\ &= y_k^x(t_s)(z_i^x(t_s) - z_i^x(t_s^-)) + \mu^{-1} H_{ki} \end{aligned} \tag{5.122}$$

where $(z_k^x(t_s) - z_k^x(t_s^-))$ is the discontinuity of the z solution of (5.119),(5.120).

Proof. The chain rule in ordinary differentiation gives

$$\frac{\partial(\lambda_2(t_s) - \lambda_1(t_s^-))}{\partial x^i} = \frac{\partial(\lambda_2(t_s) - \lambda_1(t_s^-))}{\partial x} . \frac{\partial x}{\partial x^i}. \tag{5.123}$$

As given by Lemma 5.1, $\nabla_x v(x^o, t_s) = \lambda_2(t_s) - \lambda_1(t_s^-)$ hence

$$\begin{aligned} \frac{\partial x^T}{\partial x^i} . \nabla^2_{(x,x)} v(x^o, t_s) . \frac{\partial x}{\partial x^k} &= (\frac{\partial x^T}{\partial x^i})(\frac{\partial(\lambda_2(t_s) - \lambda_1(t_s^-))}{\partial x^k}) \\ &= (\frac{\partial x^T}{\partial x^k})(\frac{\partial(\lambda_2(t_s) - \lambda_1(t_s^-))}{\partial x^i}), \end{aligned}$$

and by the definition, $y_i(t) = \frac{\partial x(t)}{\partial x^i}$ and $z_i(t) = \frac{\partial \lambda(t)}{\partial x^i}$, therefore

$$y_i^x(t_s)(z_k^x(t_s) - z_k^x(t_s^-)) = y_k^x(t_s)(z_i^x(t_s) - z_i^x(t_s^-)), \tag{5.124}$$

where together with Proposition 1, the lemma statement is proved. □

Corollary 1. *(i): If in (5.122) the positive Hessian Condition* $\left[\frac{\partial^2 v(x,t)}{\partial x^i \partial x^j}\right] \geq 0, i, j = 1, ..., n$, $(\left[\frac{\partial^2 v(x,t)}{\partial x^i \partial x^j}\right] > 0, i, j = 1, ..., n)$ *holds, the local stationary state* x^o *is locally (strictly locally) optimal on* M.
(ii): The local optimality of $x^o \in M$ *implies (5.122) holds.*

We observe that equation (5.122) gives the Hessian matrix in any given parametrization of the switching manifold at the optimal switching state.

5.7. Example

Consider a hybrid system consisting of two modes with the following dynamics:

$$S_1 \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} u \tag{5.125}$$

$$S_2 \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} u, \tag{5.126}$$

for which the cost function and boundary conditions are defined as:

$$J = \frac{1}{2}\int_0^{10} u^2(t)dt, \quad x_0 = (0, 0, 0), \quad x_f = (4, 1, 3).$$

The simulation is performed with respect to the following switching manifold:

$$M = \{(x_1, x_2, x_3); \quad x_1^2 + 2x_2^2 - x_3 = 4\}, \tag{5.127}$$

By direct calculation we have

$$H = \frac{1}{\sqrt{(1 + 4x_1^2 + 16x_2^2)}} \begin{pmatrix} -2 & 0 \\ 0 & -4 \end{pmatrix}, \tag{5.128}$$

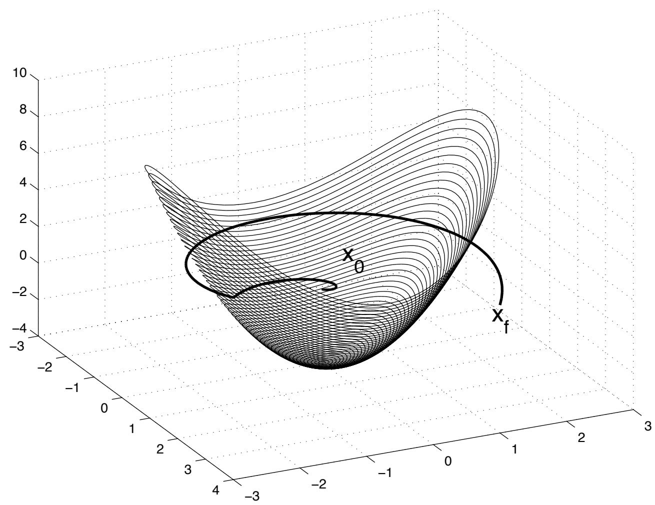

FIGURE 5.5. State Trajectory of Example 1

By applying the *GG-HMP* algorithm introduced in [**77**] to the example above, the optimal switching time and state are computed as follows:

$$t^o = 6.1900s, \quad x^o = [-1.1771, -0.8307, -1.23420], \quad \mu = -27.8, \tag{5.129}$$

where at x^o we have

$$H^o = \begin{bmatrix} -0.4770 & 0 \\ 0 & -0.9539 \end{bmatrix}, \tag{5.130}$$

(GG-HMP algorithm denotes the *Geodesic-Gradient Hybrid Maximum Principle* algorithm whose convergence properties are proven in [**77**]). The evolution of the state trajectory in the example is shown in Figure 5.5 and the Hessian of the value function at x^o is obtained by direct calculation as

$$\frac{\partial^2 v(x^o, t^o)}{\partial x^2} = \begin{bmatrix} 2.8235 & -1.5010 & 0.7973 \\ -1.4993 & 4.5418 & -1.4824 \\ 0.7987 & -1.4827 & 2.8875 \end{bmatrix}. \tag{5.131}$$

Employing Theorem 5.3, Proposition 1 and the solution of (5.117,5.118) we get

$$T_1 = [16.0188, 40.2344, -171.4023], \tag{5.132}$$
$$T_2 = [-155.4266, 207.9389, -325.0340].$$

Inserting T_1, T_2 and $\frac{\partial^2 v(x^o,t^o)}{\partial x^2}$ in (5.82) we obtain

$$[H_{ij}] = \begin{bmatrix} -0.4458 & -0.0521 \\ 0.0500 & -1.0430 \end{bmatrix} \cong \begin{bmatrix} -0.4770 & 0 \\ 0 & -0.9539 \end{bmatrix}, \tag{5.133}$$

which (to two decimal places) is consistent with (5.130). The lack of symmetry in (5.133) results from the lack of symmetry in (5.131) which is due to the level of precision chosen in the numerical calculation.

5.8. Local Deformation of Switching Manifolds

In this section we analyze the deformation of switching manifolds with the objective of reducing the HOCP value function via changes in the design of the switching manifold configuration. This problem can be studied within the two different frameworks of global or local deformations and both will be considered in this subsection. The question is how we can deform the switching manifold to reduce the value function. This problem can be studied in two different classes of global or local deformations.

In principle, the study of the global deformation of a switching manifold necessitates a global parametrization for manifolds. As discussed in Section 5.1 and [**71**], if we parametrize the manifold perturbation by a variable α we can proceed by a gradient algorithm to find the optimal parameter α^o which minimizes the value function. Here we assume that the nominal manifold is parametrized around the nominal optimal switching state by x, and we search for a perturbation in the parametrization space to reduce the value function. This perturbation is defined by the following formula:

$$x^{\epsilon}(x^1, ..., x^n) = x(x^1, ..., x^n) + \epsilon\beta(x^1, ..., x^n), \tag{5.134}$$

where

$$\beta(x^1, ..., x^n) \in R^{n+1}, \quad \beta(x^1, ..., x^n) = 0 \quad \forall (x^1, ..., x^n) \notin U_{R^n}. \tag{5.135}$$

One possible perturbation candidate is changing the switching manifold along the normal direction of the switching manifold, therefore, the perturbed parametrization is given as follows:

$$x^{\epsilon}(x^1, ..., x^n) = x(x^1, ..., x^n) + \epsilon\theta(x^1, ..., x^n).N(x^1, ..., x^n). \tag{5.136}$$

The following lemma shows that by choosing small enough ϵ we can always find a normal deformation to reduce the hybrid value function.

Lemma 5.5. *For a nominal parametrization x around the nominal optimal switching state x of a nominal switching manifold for a given HOCP such that $\lambda - \lambda^- = pN \neq 0$, there exists a parameter ϵ and a function θ by which the perturbation of the nominal parametrization with*
$x^{\epsilon}(x^1, ..., x^n) = x(x^1, ..., x^n) + \epsilon\theta(x^1, ..., x^n).N(x^1, ..., x^n)$, results in a deformation of the switching manifold which yields a decrease in the HOCP value function for all $\delta, 0 \leq \delta < \epsilon$.

Proof. Here we give a proof in three dimensional Euclidean spaces. A proof for general n dimensional Euclidean spaces can be obtained by the generalization of the presented method. The switching manifold is defined in R^3 then the new parametrization is given as

$$x^{\epsilon}(x^1, x^2) = x(x^1, x^2) + \epsilon\theta(x^1, x^2).N(x^1, x^2). \tag{5.137}$$

Consider (x, t) as the nominal optimal state and time (for the simplicity in the notations we drop the superscript o). Consequently the perturbed optimal switching state and switching time will be $(x+\delta x, t+\delta t)$. Expanding the value function around the nominal switching state, time implies

$$v(x+\delta x(\epsilon), t+\delta t(\epsilon)) = v(x,t) + \frac{\partial v(x,t)}{\partial x}\delta x(\epsilon) + \frac{\partial v(x,t)}{\partial t}\delta t(\epsilon) + O(\delta t^2) + O(\delta x^2). \tag{5.138}$$

Lemma 5.1 gives $v(x+\delta x, t+\delta t) = v(x,t) + \frac{\partial v(x,t)}{\partial x}\delta x + o(\delta t) + o(\delta x)$ since $\frac{\partial v(x,t)}{\partial t} = 0$. The new coordinates of the perturbed optimal state is given as $(x^1+\delta x^1, x^2+\delta x^2)$ where $\delta x = x^\epsilon - x$ is as follows:

$$\begin{aligned}
\delta x &= x^\epsilon(x^1+\delta x^1, x^2+\delta x^2) - x(x^1, x^2) \\
&= x(x^1+\delta x^1, x^2+\delta x^2) + \epsilon\theta(x^1+\delta x^1, x^2+\delta x^2)N(x^1+\delta x^1, x^2+\delta x^2) \\
&\quad -x(x^1, x^2) \\
&= x_{x^1}(x^1, x^2)\delta x^1 + x_{x^2}(x^1, x^2)\delta x^2 + \epsilon\theta_{x^1}(x^1, x^2)N(x^1, x^2)\delta x^1 \\
&\quad +\epsilon\theta_{x^2}(x^1, x^2)N(x^1, x^2)\delta x^2 \\
&\quad +\epsilon\theta(x^1, x^2)N_{x^1}(x^1, x^2)\delta x^1 + \epsilon\theta(x^1, x^2)N_{x^2}(x^1, x^2)\delta x^2 + \epsilon\theta(x^1, x^2)N(x^1, x^2) \\
&\quad +o(\delta x^1) + o(\delta x^2).
\end{aligned} \tag{5.139}$$

By the first order optimality properties of the hybrid value function given in Lemma 5.1 and geometrical properties of the switching manifold we have

$$\frac{\partial v(x,t)}{\partial x} \perp T_xM,\ x_{x^1},\ x_{x^2},\ N_{x^1},\ N_{x^2} \in T_xM, \tag{5.140}$$

where $x_{x^i} = \frac{\partial x}{\partial x^i}, i = 1, 2$.

Theorem 3.1 and Lemma 5.2 together imply

$$\begin{aligned}
v(x+\delta x, t+\delta t) &= v(x,t) + (\lambda - \lambda^-)(\epsilon\theta_{x^1}(x^1, x^2)N(x^1, x^2)\delta x^1 \\
&\quad +\epsilon\theta_{x^2}(x^1, x^2)N(x^1, x^2)\delta x^2 + \epsilon\theta(x^1, x^2)N(x^1, x^2)).
\end{aligned}$$

We shall choose the function θ such that

$$\theta_{x^1}(x^1, x^2) = \theta_{x^2}(x^1, x^2) = 0, \quad x^o = x(x^1, x^2). \tag{5.141}$$

Since $\lambda - \lambda^- = pN(x^1, x^2), p \in R$, therefore

$$\begin{aligned} v(x + \delta x, t + \delta t) &= v(x, t) + \epsilon p \theta(x^1, x^2)||N(x^1, x^2)||^2 \\ &= v(x, t) + \epsilon p \theta(x^1, x^2). \end{aligned} \tag{5.142}$$

The statement of lemma is proved if we choose $||\epsilon||$ as small as (5.138) is valid and sign of ϵ as $\epsilon p \theta(x^1, x^2) < 0$. □

Remark: One important consequence of the previous lemma is that for all autonomous hybrid control systems there exists a normal perturbation around the nominal optimal point such as to reduce the hybrid value function unless $p = 0$, which is an optimal controlled switching point of hybrid systems, see [**80**].

Subject to the smoothness constraints, the arbitrary class of functions generically denoted θ are chosen to be stationary with respect to epsilon (at the nominal optimal state x). Such functions are usually referred to in differential geometry as bump functions (at a given nominal point).

5.8.1. Simulation Results. Here we consider a simple example in two dimensional Euclidean space to confirm the existence of the normal deformation of switching manifolds. Consider the hybrid system consisting of two phases with the following dynamics:

$$S_1 \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} u, \tag{5.143}$$

$$S_2 \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u, \tag{5.144}$$

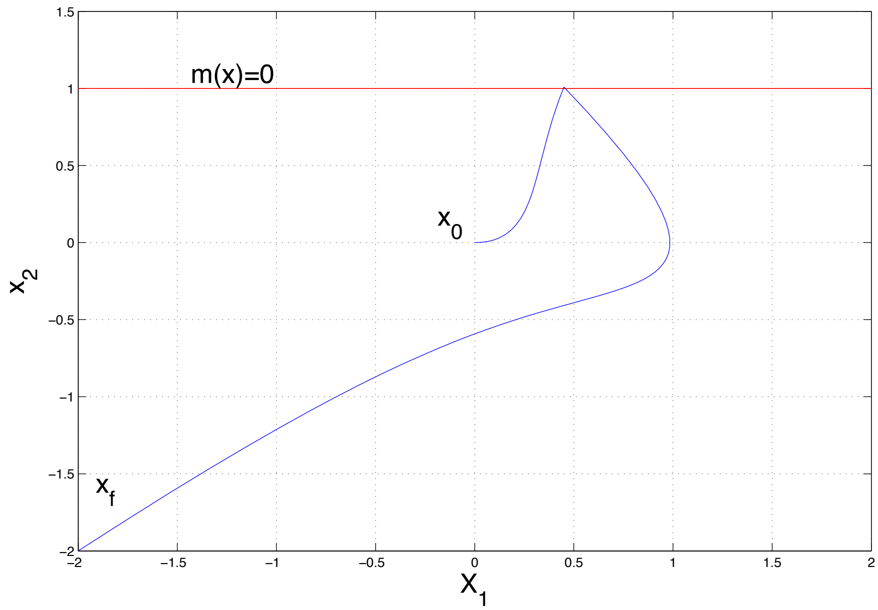

FIGURE 5.6. Hybrid Optimal Trajectory

The dynamics are both controllable and the boundary conditions and cost function are given as follows:

$$x(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ \ x(10) = \begin{pmatrix} -2 \\ -2 \end{pmatrix}, \ \ J = \frac{1}{2}\int_0^{10} x^2 dt. \tag{5.145}$$

The switching manifold is considered as a simple one dimensional line in two dimensional space:

$$(x, y) = (x, 1) \in R^2, \quad x \in R. \tag{5.146}$$

The hybrid state trajectory and switching manifold are shown in Figure 5.6. The HMP algorithm, see [**66**], gives the following results for the nominal switching manifold:

$$t^o = 3.69, \ \ x^o = (0.4498, 1.0088), \ \ J^o = 0.3082, \ \ p = 2.6875. \tag{5.147}$$

Based on (5.142) we can choose $\theta(x) = 1$ which satisfies (5.141) at the nominal optimal switching state $x_s = (0.4498, 1.0088)$ so just by choosing small negative ϵ we

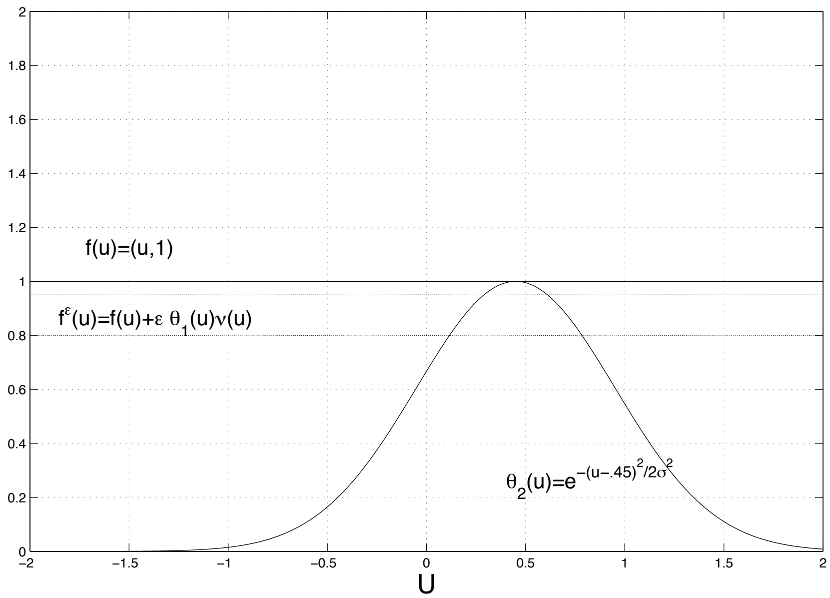

FIGURE 5.7. Deformation Functions

will end up a hybrid system with a lower value function. By choosing $\epsilon = -.038$ we have the following results:

$$t^o = 3.7092, \quad x^o = (0.4041, 0.9084), \quad J^o = 0.2517, \tag{5.148}$$

which obviously gives a less cost function than the nominal one. Another candidate for the deformation function is a bell function which also satisfies (5.141) and is given by
$\theta(x) = exp(\frac{-(x-x_s(1))^2}{2\sigma^2})$. Figures 5.7 and 5.8 show the perturbed and nominal hybrid trajectories for the perturbed and nominal switching manifold applying $\theta(x) = exp(\frac{-(x-x_s(1))^2}{2\sigma^2})$ as a deformation function. By choosing $\epsilon = -.03, x = .6$ results in the following parameters:

$$t^o = 3.5957, \quad , x^o = (0.4707, 0.9749), \quad J^o = 0.2783, \tag{5.149}$$

where again the new optimal cost function is lower than the nominal optimal cost, $J^o = 0.3082$.

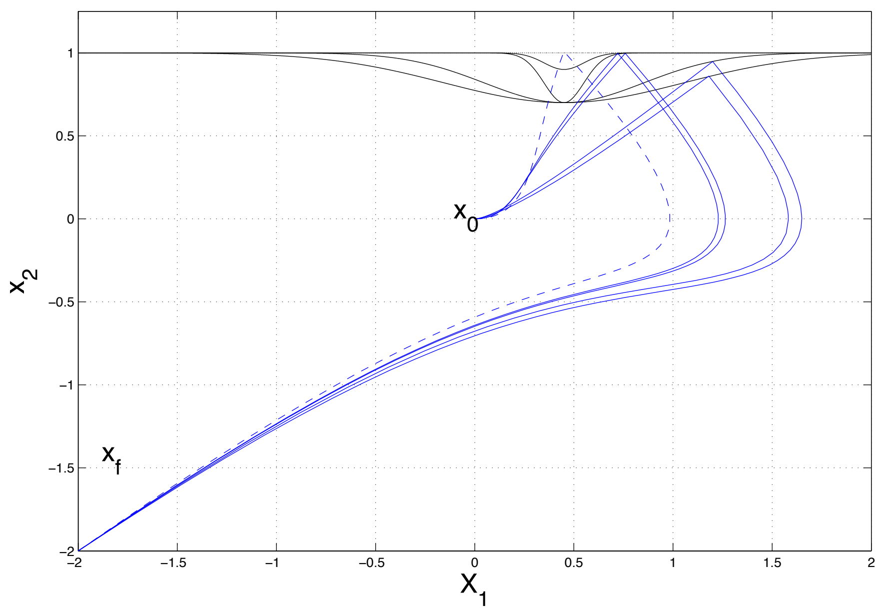

FIGURE 5.8. Hybrid Optimal Trajectory for Perturbed Systems

5.9. Global Deformation of Switching Manifolds

In this section we introduce hybrid systems for which the total cost is the summation of the cost defined by (3.4) and the energy of the switching manifold deformation mapping (as originally introduced in [**73**]). As shown in Lemma 5.5, for a given switching manifold it is in general possible to decrease the value function by a local perturbation around the optimal switching state x^o. In practice, changing the switching manifold may impose an extra cost. Consider an automatic gear changing system which changes at a certain speed. Changing the speed level at which the switching happens requires a change in the mechanical structure of the gear box which may not be feasible or may make its manufacture more expensive.

This motivates us to include an extra term in hybrid cost which corresponds to the cost of deformations of switching manifolds. This extra cost depends on the nature of the hybrid system and may differ from one system to another. We shall assign a positive cost for changing the nominal switching manifold M to a new switching manifold $F(M)$. Figure 5.9 provides a picture of such deformations in a three dimensional space.

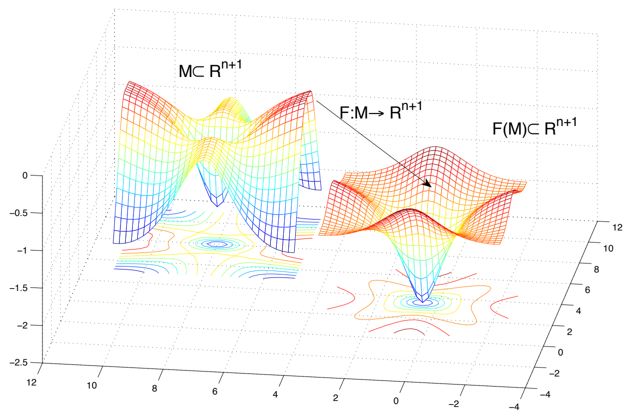

FIGURE 5.9. Switching Manifold Deformation

We now define an energy function for the mapping $F : M \to N, \quad F \in C^\infty(M, N)$. We note that in this section N is the image Riemannian manifold and is not the normal vector to the switching manifold M. In general, we consider (M, g) to be an m dimensional domain Riemannian manifold with the corresponding Riemannian metric g and (N, h) to be an n dimensional image Riemannian manifold with the corresponding Riemannian metric h. The push forward of the mapping F at point $x \in M$ is defined to be the linear map from T_xM to $T_{F(x)}N$ given by:

$$TF := dF_x : T_xM \to T_{F(x)}N. \tag{5.150}$$

The local coordinates on T_xM at $x \in M$ and at T_yN at y are given respectively by $\{(\frac{\partial}{\partial x^i})\}_{i=1,...,m}$ and $\{(\frac{\partial}{\partial y^\alpha})\}_{\alpha=1,...,n}$, the push forward of F applied on a base tangent vector is given locally as follows:

$$dF_x((\frac{\partial}{\partial x^i})) = \sum_{\alpha=1}^{n}(\frac{\partial F^\alpha}{\partial x^i})(x)(\frac{\partial}{\partial y^\alpha}), \quad i = 1, ..., m. \tag{5.151}$$

The space $Hom(T_xM, T_{F(x)}N)$, i.e. the space of all linear maps from T_xM to $T_{F(x)}N$, is linearly isomorphic to $T^*_xM \otimes T_{F(x)}N$, (see [**37**]), therefore dF_x can be locally described as an element of $T^*_xM \otimes T_{F(x)}N$.

$$dF_x \in Hom(T_xM, T_{F(x)}N) \cong T^*_xM \otimes T_{F(x)}N, \tag{5.152}$$

where

$$dF_x = \sum_{i=1}^{m}\sum_{\alpha=1}^{n}(\frac{\partial F^\alpha}{\partial x^i})(x)(dx^i) \otimes (\frac{\partial}{\partial y^\alpha}). \tag{5.153}$$

The norm g^* on T^*_xM and h on N induce an inner product on the tensor product space $T^*_xM \otimes T_{F(x)}N$ given by

$$\langle dx^i \otimes (\frac{\partial}{\partial y^\alpha}), dx^j \otimes (\frac{\partial}{\partial y^\beta})\rangle = g^{ij}h_{\alpha\beta}(F(x)), \tag{5.154}$$

where

$$g^* = [g_{ij}]^{-1}, \quad i, j = 1, ..., n. \tag{5.155}$$

Employing (5.153) and (5.154), the norm squared of dF_x is defined as follows:

$$\begin{aligned} |dF_x|^2 &= \langle dF_x, dF_x\rangle \\ &= \sum_{i,j=1}^{m}\sum_{\alpha,\beta=1}^{n} g^{ij}(x)h_{\alpha\beta}(F(x))(\frac{\partial F^\alpha}{\partial x^i})(x)(\frac{\partial F^\beta}{\partial x^j})(x), \end{aligned}$$

and the mapping dF is defined as the following section:

$$dF \in \Gamma(T^*M \otimes F^{-1}TN), \tag{5.156}$$

where $F^{-1}TN$ is the induced tangent bundle by F and Γ is a cross section of the vector bundle $T^*M \otimes F^{-1}TN$. The norm squared of dF over the manifold M is written as

$$|dF|^2 = \sum_{i,j=1}^{m}\sum_{\alpha,\beta=1}^{n} g^{ij}h_{\alpha\beta}(F)(\frac{\partial F^\alpha}{\partial x^i})(\frac{\partial F^\beta}{\partial x^j}), \tag{5.157}$$

where $h_{\alpha\beta}$ is the metric on N. The energy density of F is finally defined as

$$e(F)(x) = \frac{1}{2}|dF|^2(x), \quad x \in M, \tag{5.158}$$

with the corresponding energy functional

$$E(F) = \int_M e(F) d\mu_g, \tag{5.159}$$

where μ_g is the Lebesgue measure defined by g_{ij} on M, see [**37**].

The minimization of the energy functional E with respect to a parametrized mapping may be analyzed using the variational methods, [**37**, **53**].

Consider a *variation* $\hat{F}$ of F as follows:

$$\hat{F} : I \times M \to N, \quad I = (-\epsilon, \epsilon) \in R,$$
$$\hat{F}(x, 0) = F(x) \in N, \quad \forall x \in M, \hat{F} \in C^\infty(I \times M, N). \tag{5.160}$$

We shall write $F_s = \hat{F}(., s), s \in (-\epsilon, \epsilon)$. *Harmonic maps* are solutions to the differential equation$\frac{d}{ds}E(F_s)|_{s=0} = 0$.

The corresponding *tension field* $\tau(F)$ in the induced bundle $F^{-1}TN$ is written as (see [**37**, **53**]):

$$\tau(F)^\alpha := \sum_{i,j=1}^{m} g^{ij} (\frac{\partial^2 F^\alpha}{\partial x^i \partial x^j} - \sum_{k=1}^{m} \Gamma_{ij}^k \frac{\partial F^\alpha}{\partial x^k} + \sum_{\beta,\gamma=1}^{n} \Gamma_{\beta\gamma}^\alpha(F) \frac{\partial F^\beta}{\partial x^i} \frac{\partial F^\gamma}{\partial x^j}),$$
$$\alpha = 1, ..., n, \tag{5.161}$$

where $\Gamma_{ij}^k, \Gamma_{\beta\gamma}^\alpha$ are the Christoffel symbols defined on (M, g) and (N, h) respectively. The following theorem gives an expression for the first variation of energy of maps from M to N.

THEOREM 5.5 ([**37**] , Eq 8.1.13, [**53**]). *Let* $F_s = \hat{F}(., s)$ *be a* C^∞ *variation of* $F = F_0 : M \to N$, *then*

$$\frac{d}{ds}E(F_s)|_{s=0} = -\int_M \langle Y, \tau(F) \rangle d\mu_g, \tag{5.162}$$

where $Y = \frac{d}{ds} F_s|_{s=0}$ *is a variation vector field of* F *and* $\tau(F)$ *is the tension field of* F *defined in (5.161).*

□

5.10. Extended Hybrid Optimal Control Problems

Employing the notion of the energy of mappings among manifolds, we modify the cost function defined in (3.4) to the following formula which includes the energy of deformations:

$$J^F(t_0, t_f, u) := \int_{t_0}^{t_s} l_{q_1}(x_{q_1}(s), u(s))ds + \int_{t_s}^{t_f} l_{q_2}(x_{q_2}(s), u(s))ds$$
$$+h(x_{q_L}(t_f)) + ||E(F_{M\to N}) - E(I_{M\to N})||^2, \tag{5.163}$$

where $I_{M\to N}$ is the identity map from M to N. For a given F_s, the corresponding value function is defined as:

$$v(x, t, F_s) := inf_{u\in U} J^{F_s}(t_0, t_f, u)|_{x_s=x, t_s=t},$$
$$x \in M, t, s \in R, F_s \in C^\infty(M, N). \tag{5.164}$$

Since the first two terms on the right hand side of (5.163) are computable separately, the value function defined in (5.164) can be split into two independent terms as follows:

$$v(x, t, F) = v_F(x, t) + ||E(F) - E(I_{M\to N})||^2, \tag{5.165}$$

where $v_F(x, t)$ is the value function of the hybrid system without considering the deformation cost. In order to be able to compute the sensitivity of $v(x, t, F_s)$ at (x, t, F_s) we use the variation of F, introduced in (5.160). The sensitivity of $v(x, t, F_s)$ with respect to (x, t, F_s) is given as follows:

$$\frac{\partial v(x, t, F_s)}{\partial x} = \frac{\partial v_{F_s}(x, t)}{\partial x} + \frac{\partial ||E(F_s) - E(I_{M\to N})||^2}{\partial x}, \tag{5.166}$$

and similarly for t and s we have

$$\frac{\partial v(x,t,F_s)}{\partial t} = \frac{\partial v_{F_s}(x,t)}{\partial t} + \frac{\partial ||E(F_s) - E(I_{M\to N})||^2}{\partial t}, \tag{5.167}$$

$$\frac{\partial v(x,t,F_s)}{\partial s} = \frac{\partial v_{F_s}(x,t)}{\partial s} + \frac{\partial ||E(F_s) - E(I_{M\to N})||^2}{\partial s}. \tag{5.168}$$

By definition, $E(F_s)$ does not depend on x and t therefore the second terms in (5.166) and (5.167) vanish; hence

$$\frac{\partial v(x,t,F_s)}{\partial x} = \frac{\partial v_{F_s}(x,t)}{\partial x} = \frac{\partial v(F_s(x),t)}{\partial F_s(x)} W(x), \tag{5.169}$$

where $v(F_s(x),t)$ is the hybrid value function given at the switching time t and the switching state $F_s(x) \in F(M)$, therefore

$$\frac{\partial v(F_s(x),t)}{\partial F_s(x)} = \lambda - \lambda^-, \quad W(x) := \frac{\partial F_s(x)}{\partial x}, \tag{5.170}$$

where $W(x)$ is the Jacobian matrix of the coordinate change and $\lambda - \lambda^-$ is the adjoint process discontinuity at $F_s(x)$.

By the results of Lemma 5.2, the first term on the right hand side of Equation (5.168) is given by the following lemma:

Lemma 5.6. *For the extended hybrid cost function defined in (5.163),* $\frac{\partial v_F(x,t)}{\partial s}$ *is given by*

$$\frac{\partial v_{F_s}(x,t)}{\partial s}|_{s=0} = \langle (\lambda - \lambda^-), Y(x) \rangle, \tag{5.171}$$

where $(\lambda - \lambda^-)$ *is the adjoint process discontinuity at* $F_0(x) = F(x) \in F(M)$ *and* $Y(x) = \frac{dF_s(x)}{ds}|_{s=0}$.

Proof. In the case of general Riemannian manifolds M and N, the sensitivity function defined in (5.166) is a linear map from the tangent space of M to R. By definition 2.1, $x \in R^{n+1}$ and so in our analysis $N = R^{n+1}$. Since by the definition

(5.164), t and x are chosen independently of s, we have

$$\frac{\partial v_{F_s}(x,t)}{\partial s}|_{s=0} = \lim_{s\to 0} \frac{v(x,t,F_s(x)) - v(x,t,F_0(x))}{s}, \tag{5.172}$$

where $F_s(x)$ is the switching state on $F_s(M)$. Applying the chain rule we have

$$\frac{\partial v_{F_s}(x,t)}{\partial s} = \frac{\partial v(x,t,F_s(x))}{\partial F_s(x)}|_{s=0}.\frac{dF_s(x)}{ds}. \tag{5.173}$$

Lemma 5.2 gives

$$\frac{\partial v(x,t,F_s(x))}{\partial F_s(x)}|_{s=0} = \lambda - \lambda^-, \tag{5.174}$$

hence (5.174) and $Y(x) = \frac{dF_s(x)}{ds}|_{s=0}$ yield (5.171). □

The following theorem gives the optimal mapping characteristics for the extended hybrid cost function defined in (5.163).

THEOREM 5.6. *A mapping F_0 which passes through the optimal controlled switching state, at which necessarily, $\lambda - \lambda^- = 0$, and which is a harmonic map, i.e. a map satisfying $\frac{d}{ds}E(F_s)|_{s=0} = 0$, satisfies the necessary conditions for a solution to the extended HOCP.*

PROOF. We observe that Theorem 5.5 implies that the second term of the right hand side of (5.168) is given by:

$$\frac{\partial ||E(F_s) - E(I_{M\to N})||^2}{\partial s} = -2(E(F_s) - E(I_{M\to N})).\int_M \langle Y, \tau(F)\rangle d\mu_g. \tag{5.175}$$

Now if (x,t,F_0) corresponds to the minimum of $v(x,t,F)$ then

$$\frac{\partial v(x,t,F_0)}{\partial x} \perp T_xM, \quad \frac{\partial v(x,t,F_0)}{\partial t} = 0, \tag{5.176}$$

and

$$\frac{\partial v(x,t,F_s)}{\partial s}|_{s=0} = 0, \tag{5.177}$$

where (5.176) holds by Lemma 5.1.

(5.177), Lemma 5.6 and Theorem 5.5 yield

$$0 = \frac{\partial v(x,t,F_s)}{\partial s}|_{s=0} = (\lambda - \lambda^-).Y(x) - 2(E(F_0) - E(I_{M\to N})).\int_M \langle Y, \tau(F_0)\rangle d\mu_g, \tag{5.178}$$

and consequently

$$(\lambda - \lambda^-).Y(x) = 2(E(F_0) - E(I_{M\to N})) \times \int_M \langle Y, \tau(F_0)\rangle d\mu_g. \tag{5.179}$$

We observe that equation (5.179) must be satisfied for all C^∞ *variational vector fields* $Y(.)$, the left hand side of (5.179) only depends on $Y(.)$ at the optimal x and the right hand side of (5.179) is independent of x.

Clearly $\lambda - \lambda^- = 0, \tau(F_0) = 0$, satisfy both (5.176) and (5.177). Hence the necessary condition (5.179) for optimally with respect to joint hybrid control and switching manifold perturbations is satisfied for the mapping F_0 for which $F_0(M)$ passes through the optimal *controlled* switching state, i.e $\lambda - \lambda^- = 0$ (see [**80**]), and for which F_0 is a harmonic map, i.e. $\tau(F_0) = 0$, (see [**37**]). □

Remark: It should be noted that such a mapping always exists since, for example in R^3, a simple shift is a harmonic map and it is always possible to obtain a shift for which the initial manifold M passes through one of the states for which $\lambda - \lambda^- = 0$.

CHAPTER 6

Future Research

Proposed Research on Hybrid Minimum Principle

In the Hybrid Minimum Principle discussed in Chapters 2, it is assumed that (i) all dynamical equations are defined on a same ambient Riemannian manifold $\mathcal{M}$; and (ii) all switching manifolds $\mathcal{N}$ are smoothly embedded orientable submanifolds of $\mathcal{M}$ (see [**66**, **75**]).

These assumption can be relaxed so as to generalize the HMP framework and in particular to extend the HMP results to hybrid systems with non-smooth switching manifolds.

Proposed Research on GG-HMP and NG-HMP Algorithms and HMP on Lie Groups

So far, the GG-HMP and NG-HMP algorithms mainly focus on the minimization of the hybrid value function along geodesics. It is of interest and of potential value to extend these algorithms to hybrid systems where switching manifolds do not have pre-specified geometric structures. In that case geodesic curves are necessarily a priori defined.

The GG-HMP algorithm was modified for hybrid systems in such a way that flows along exponential curves were utilized in the specification of the algorithm [**76**]. In this connection, the general extension of the NG-HMP algorithm to hybrid systems on Lie groups is a most important candidate for the further application and development of optimization techniques on Riemannian manifolds.

Proposed Research on Hybrid Dynamic Programming on Manifolds

It is known that the sufficient conditions for optimality can be derived by applying dynamic programming to optimal control problems. A generalization of the standard dynamic programming approach to hybrid systems was introduced by Shaikh and Caines in [**64**]. In this case the hybrid value function is defined by the optimization of hybrid cost function with respect to continuous and discrete controls; in principle, this hybrid dynamic programming approach can be extended to hybrid systems defined on Riemannian manifolds by employing the language of differential geometry and in this connection we note a generalization of classic dynamic programming to optimal control problems defined on manifolds is given in [**23**].

Proposed Research on Hybrid Systems with Uncertainty

All the analyses presented in this thesis are based upon full information about the continuous and discrete dynamics. The analysis of hybrid systems where uncertainty appears in both the continuous and the discrete dynamics constitutes a generalization of the hybrid systems framework considered in this thesis. These problems can be divided into two classes: (i) robust hybrid systems, and (ii) stochastic hybrid systems.

In the first case, within the framework of this thesis, one would analyze hybrid systems on manifolds where uncertain parameters are constrained to lie within prespecified sets; in this setting the use of robust control methods such as H_∞ and μ-synthesis is to be investigated to obtain optimal performance under dynamical uncertainties.

In the second case, uncertainty is modelled within the framework of stochastic dynamical systems, see [**35**, **36**, **48**]. Consequently the fundamental methods of Stochastic Dynamic Programming and the Stochastic Minimum Principle would be used for the construction of a theory of the optimization of stochastic hybrid systems on manifolds.

A paradigm problem illustrating both cases above is that of a controlled rotating satellite (see Section 4.6) whose mechanical parameters are not exactly known and which is subject to random disturbances in the functioning of its thrusters.

Second Order Variation of the Energy of Deformation maps

In Chapter 5 we obtained the necessary conditions for the optimal deformation of switching manifolds for autonomous hybrid systems in order to minimize the extended hybrid cost function. However results presented by Theorem 5.6 do not guarantee the minimization of the extended hybrid cost function introduced in (5.163). Similar to standard optimization problems, sufficient conditions of optimality could be given by the second order variation of the extended hybrid cost function with respect to the optimization variables. In this connection, an investigation of the second order variation of the extended hybrid value function is an important candidate for future research, in particular this could be based upon the analysis of the second order variation of the energy of maps presented in [**53**], Section 3.5.

REFERENCES

[1] Y. H. A. Bryson. *Applied Optimal Control: Optimization, Estimation and Control*. Taylor and Francis, 1975.

[2] R. Abraham and J. Marsden. *Foundations of Mechanics*. AddisonWesley, 1978.

[3] A. Agrachev and Y. Sachkov. *Control Theory from the Geometric Viewpoint*. Springer, 2004.

[4] F. Alvarez, J. Bolte, and J. Munier. *A Unifying Local Convergence Result for Newton's Method in Riemannian Manifold*. Research Report, INRIA, France, 2004.

[5] V. Arnold. *Mathematical Methods of Classical Mechanics*. Springer, 1989.

[6] M. Barbero-Linan and C. Muoz-Lecanda. Geometric approach to pontryagin's maximum principle. *Acta Applicandae Mathematicae*, 108(2):429–485, 2009.

[7] A. Bemporad, A. Giua, and C. Seatzu. An iterative algorithm for the optimal control of continuous time switched linear systems. *Proc. 6th International Workshop on Discrete Event Systems*, pages 335–340, 2002.

[8] A. Bemporad, A. Giua, and C. Seatzu. A master-slave algorithm for the optimal control of continuous time switched affine systems. *Proc. 41th IEEE Conf. Decision and Control*, pages 1976–1981, Las Vegas, december, 2002.

[9] A. Bemporad, A. Giua, and C. Seatzu. Synthesis of state feedback optimal controllers for continuous time switched linear systems. *Proc. 41th IEEE Conf. Decision and Control*, pages 3182–3187, Las Vegas, december, 2002.

[10] A. Bloch. *Nonholonomic Mechanics and Control*. Springer-Verlag, 2003.

[11] M. Boccadoro, P. Valigi, and Y. Wardi. A method for the design of optimal switching surfaces for autonomous hybrid systems. In *10th International Conference on Hybrid Systems HSCC LNCS*, pages 650–655, 2007.

[12] M. Boccadoro, Y. Wardi, M. Egersted, and E. Verriest. Optimal control of switching surfaces in hybrid dynamical systems. *Journal of Discrete Event Dynamic Systems*, 15(3):433–448, 2005.

[13] V. G. Boltyanskii. Sufficient conditions for optimality and the justification of the dynamic programming method. *SIAM J. Control and Optimization*, 4(2):326–361, 1966.

[14] W. Boothby. *An introduction to differential manifolds and Riemannian geometry*. Academic Press, 1986.

[15] M. Branicky, V. Borkar, and S. Mitter. A unified framework for hybrid control: Model and optimal control theory. *IEEE Trans Automatic Control*, 43(3):31–45, 1998.

[16] M. Branicky and S. Mitter. Algorithms for optimal hybrid control. *Proc. 34th IEEE Conf. Decision and Control*, pages 2661–2666, New Orleans, December, 1995.

[17] R. W. Brockett. System theory on group manifolds and coset spaces. *SIAM J. Control and Optimization,*, 10(2):265–284, 1972.

[18] R. W. Brockett. Lie theory and control systems defined on spheres. *SIAM J. Appl. Math.*, 23(2):213–225, 1973.

[19] F. Bullo and A. Lewis. *Geometric Control of Mechanical Systems: Modelling, Analysis, and Design for Mechanical Control Systems*. Springer, 2005.

[20] P. E. Caines. *Lecture Notes on Nonlinear Systems*. Department of Electrical and Computer Engineering, McGill University, 2000.

[21] F. Cardetti. On properties of linear control systems on lie groups. *PhD thesis, Department of Mathematics, Louisiana State University*, 2002.

[22] C. Cassandras, D. Pepyne, and Y. Wardi. Optimal control of a class of hybrid systems. *IEEE Trans. Automatic Control*, 46(3):398–415, 2001.

[23] I. Chryssochoos and R. Vinter. Optimal control problems on manifolds: a dynamic programming approach. *Journal of Mathematical Analysis and Applications*, 287(1):118–140, 2003.

[24] F. Clarke and R. Vinter. The relationship between the maximum principle and dynamic programming. *SIAM J. Control and Optimization*, 25(5):1291–1311, 1987.

[25] R. P. de la Barriere. *Optimal Control Theory*. Dover Publications, 1967.

[26] A. Dmitruk and M. Kaganovich. The hybrid maximum principle is a consequence of pontryagin maximum principle. *System and Control Letters*, 57:964–970, 2008.

[27] M. Egerstedt, Y. Wardi, and F. Delmotte. Optimal control of switching times in switched dynamical systems. *Proc. 42th IEEE Conf. Decision and Control*, pages 2138–2143, Maui, December, 2003.

[28] H. Flanders. *Differential Forms with Applications to the Physical Sciences*. Academic Press, 1963.

[29] D. Gabay. Minimizing a differentiable function over a differentiable manifold. *Journal of Optimization Theory and Applications*, 37:177–219, 1982.

[30] M. Garavello and B. Piccoli. Hybrid necessary principles. *SIAM J. Math. Anal*, 43:1867–1887, 2005.

[31] A. Giua, C. Seatzu, and C. V. D. Mee. Optimal control of autonomous linear systems switched with a preassigned finite sequence. *Proc. 2001 IEEE International Symposium on intelligent Control*, pages 144–149, 2001.

[32] A. Giua, C. Seatzu, and C. V. D. Mee. Optimal control of switched autonomous linear systems. *Proc. 40th IEEE Conf. Decision and Control*, pages 2472–2477, Orlando, december, 2001.

[33] S. Hedlund and A. Rantzer. Convex dynamic programming for hybrid systems. *IEEE Trans. Automatic Control*, 47(9):1536–1540, 2002.

[34] S. Hedlund and A. Rantzer. Optimal control of hybrid systems. *Proc. 38th IEEE Conf. Decision and Control*, pages 3972–3977, Phoenix, December, 1999.

[35] A. P. Henk and J. Lygeros. Stochastic hybrid systems. *Lecture Notes in Control and Information Sciences*, 337, 2006.

[36] J. P. Hespanha. Stochastic hybrid systems: Application to communication networks. *Hybrid Systems: Computation and Control Lecture Notes in Computer Science*, pages 47–56, 2004.

[37] J. Jost. *Reimannian Geometry and Geometrical Analysis*. Springer, 2004.

[38] V. Jurdjevic. *Geometric Control Theory*. Cambridge Univ. Press, 1997.

[39] V. Jurdjevic. Integrable hamiltonian systems on complex lie groups. *American Mathematical Society*, 178(838), 2005.

[40] W. Kuhnel. *Differential Geometry Curves - Surfaces-Manifolds*. American Mathematical Society, 2005.

[41] E. Lee and L. Markus. *Foundation of Optimal Control*. Dover Books on Advanced Mathematics, 1972.

[42] J. Lee. *Riemannian Manifolds, An Introduction to Curvature*. Springer, 1997.

[43] J. Lee. *Introduction to Smooth Manifolds*. Springer, 2002.

[44] B. Lincoln and B. Bernhardsson. Lqr optimization of linear system switching. *IEEE Trans. Automatic Control*, 47(10):1701–1705, 2002.

[45] J. Lu, L. Liao, A. Nerode, and J. H. Taylor. Optimal control of systems with continuous and discrete states. *Proc. 32th IEEE Conf. Decision and Control*, pages 2292–2297, San Antonio, December, 1993.

[46] D. Luenberger. The gradient projection method along geodesics. *Management Science*, 18(11):620–631, 1972.

[47] R. Luus. Optimal control of batch reactors by iterative dynamic programming. *Journal of Process Control*, 4(4):218–226, 1994.

[48] J. Lygeros. Stochastic hybrid systems: Theory and applications. *Proc. 47th IEEE Conf. Decision and Control*, pages 40–42, 2008.

[49] R. Mahony and J. Manton. Geometry of the newton method on non-compact lie-groups. *Journal of Global Optimization*, 23(3):309–327, 9 2002.

[50] Y. Matsumoto. *An Introduction to Morse Theory*. American Mathematical Society, 2002.

[51] B. Miller and E. Rubinovich. *Impulsive Control in Continuous and Discrete-Continuous Systems*. Kluwer Academic/Plenum Publishers, 2003.

[52] J. Milnor. *Morse Theory*. Ann. Math Studies Princeton, 1963.

[53] S. Nishikawa. *Variational Problems in Geometry*. American Mathematical Society, 2002.

[54] B. Passenberg, P. Kock, and O. Stursberg. Combined time and fuel optimal driving of trucks based on a hybrid model. *Proc of the European Control Conference*, pages 4955–4960, 2009.

[55] B. Piccoli. Necessary conditions for hybrid optimization. *Proc. 38th IEEE Conf. Decision and Control*, pages 410–415, Phoenix, December, 1999.

[56] B. Piccoli. Hybrid systems and optimal control. *Proc. 37th IEEE Conf. Decision and Control*, pages 13–18, Tampa, December, 1998.

[57] P. Riedinger, J. Daafouz, and C. Iung. Suboptimal switched control in context of singular arcs. In *42th IEEE Int. Conf. Decision and Control*, pages 6254–6259, 2003.

[58] P. Riedinger, C. Iung, and F. Krutz. Linear quadratic optimization of hybrid systems. In *38th IEEE Int. Conf. Decision and Control*, pages 3059–3064, 1999.

[59] W. Rudin. *Real and Complex Analysis*. New York: McGraw-Hill, 1974.

[60] J. L. Salle. The stability of dynamical systems. In *CBMS-NSF Regional Conference Series in Applied Mathematics*, 1976.

[61] A. Schild, X. Ding, M. Egerstedt, and J. Lunze. Desing of optimal switching surfaces for switched autonomous systems. In *42th IEEE Int Conf Decision and Control*, 2009.

[62] A. Schild and J. Lunze. Stabilization of limit cycles of discreetly controlled continuous systems by controlling switching surfaces. In *10th International Conference on Hybrid Systems HSCC LNCS 4416*, pages 515–528, 2007.

[63] I. Segal and R. Kunze. *Integrals and Operators*. Springer, 1978.

[64] M. Shaikh. Optimal control of hybrid systems: Theory and algorithms. *Department of Electrical and Computer Engineering, McGill University, Ph.D Thesis*, 2004.

[65] M. Shaikh and P. E. Caines. On the optimal control of hybrid systems: Analysis and algorithms for trajectory and schedule optimization. In *42nd IEEE Int. Conf. Decision Control*, volume 43, pages 2144–2149, 1, 2003.

[66] M. Shaikh and P. E. Caines. On the hybrid optimal control problem: Theory and algorithms. *IEEE Trans. Automatic Control*, 52(9):1587–1603, Corrigendum: 54 (6) (2009) 1428., 2007.

[67] S. Smith. Optimization techniques on riemannian manifolds. *Fields Institute Communications*, 3(3):113–135, 1994.

[68] O. Stursberg and S. Engell. Optimal control of switched continuous systems using mixed integer programming. *Proc. 15th IFAC World Congress on Automatic Control*, Barcelona, July, 2002.

[69] H. Sussmann. A maximum principle for hybrid optimal control problems. In *Proc. 38th IEEE Int. Conf. Decision and Control*, pages 425–430, 1999.

[70] F. Taringoo and P. Caines. Linked sections to "on the optimal control of impulsive hybrid systems on riemannian manifolds". http://arxiv.org/abs/1209.4067, 2012.

[71] F. Taringoo and P. E. Caines. The sensitivity of hybrid systems optimal cost functions with respect to switching manifold parameters. *Hybrid Systems: Computation and Control, Springer Verlag, San Fransisco, April, 2009. Eds: R. Majumdar and P. Tabuada*, pages 475–479, 2009.

[72] F. Taringoo and P. E. Caines. The extension of the hybrid maximum principle on reimannian manifolds: Theory and algorithms. *Technical Report, McGill University*, 2010.

[73] F. Taringoo and P. E. Caines. *Geometry and Deformation of Switching Manifolds for Autonomous Hybrid Systems*. 19th International Symposium on Mathematical Theory of Networks and Systems, Budapest, Hungary, 2010.

[74] F. Taringoo and P. E. Caines. Gradient geodesic and newton geodesic hmp algorithms for the optimization of hybrid systems. In *IFAC Annual Reviews in Control*, volume 35, pages 187–198, 2011.

[75] F. Taringoo and P. E. Caines. *On the Optimal Control of Impulsive Hybrid Systems on Riemannian Manifolds*. Submitted to SIAM Journal on Control and Optimization, 2012.

[76] F. Taringoo and P. E. Caines. The exponential gradient hmp algorithm for the optimization of hybrid systems on lie groups. In *4th IFAC Conference on Analysis and Design of Hybrid Systems*, Eindhoven, June 6-8, 2012.

[77] F. Taringoo and P. E. Caines. Gradient-geodesic hmp algorithms for the optimization of hybrid systems based on the geometry of switching manifolds. In *49th IEEE Conference on Decision and Control*, pages 1534–1539, Georgia, USA, December, 2010.

[78] F. Taringoo and P. E. Caines. On the extension of the hybrid minimum principle to hybrid systems on lie groups. In *American Control Conference*, Montreal, Canada, June 2012.

[79] F. Taringoo and P. E. Caines. On the extension of the hybrid maximum principle to riemannian manifolds. In *50th IEEE Conference on Decision and Control*, pages 3301–3306, Orlando, USA, December, 2011.

[80] F. Taringoo and P. E. Caines. Geometrical properties of optimal hybrid system trajectories and the optimization of switching manifolds. In *3rd IFAC Conference on Analysis and Design of Hybrid Systems*, Zaragoza, September, 2009.

[81] D. Tyner. Geometric jacobian linearization. *PhD thesis, Department of Mathematics and Statistics, Queens University*, 2007.

[82] V. Varadarajan. *Lie groups, Lie algebras, and their representations*. Springer, 1984.

[83] H. Witsenhausen. A class of hybrid state continuous time dynamic systems. *IEEE Trans. Automatic Control*, 11(9):161–167, 1966.

[84] X. Xu and J. Antsaklis. Optimal control of switched systems based on parametrization of the switching instants. *IEEE Trans. Automatic Control*, 49(1):2–16, 2004.

[85] X. Xu and J. Antsaklis. An approach for solving general switched systems with internally forced switchings. *Proc. American Control Conference*, pages 148–153, Anchorage, 2002.

[86] X. Xu and J. Antsaklis. An approach for solving general switched linear quadratic optimal control problems. *Proc. 40th IEEE Conf. Decision and Control*, pages 2478–2483, Orlando, December, 2001.

[87] J. Zabczyk. *Mathematical Control Theory: An Introduction*. Birkhauser, 1992.

[88] P. Zhang and C. Cassandras. An improved forward algorithm for optimal control of a class of hybrid systems. *IEEE Trans. Automatic Control*, 47(10):1735–1739, 2002.

APPENDIX A

Proofs and Extended Results for Chapter 2

A.1. Proof of Lemma 2.5

PROOF. Since $\mathcal{S}$ is a smooth embedded submanifold of $\mathcal{M}$ the inclusion $i : \mathcal{S} \to \mathcal{M}$ is a topological embedding and hence its rank is constant (see [**43**]). By the Rank Theorem for Manifolds (see [**42**]), i may be locally given as

$$i(x_1, ..., x_{n-1}) = (x_1, ..., x_{n-1}, 0).$$

Hence, $\mathcal{S}$ is locally homeomorphic to $\mathbb{R}^{n-1}$. As stated in Subsection 2.1.2, $\Phi^{(t_s,t_0)}_{\pi,f_{q_0}}(x_0)$ converges to $x^o(t_s) \in \mathcal{S}$ as $\epsilon \downarrow 0$ (see [**30**]), therefore $\Phi^{(t_s,t_0)}_{\pi,f_{q_0}}(x_0)$ converges into any neighbourhood of $x^o(t_s) \in \mathcal{S}$ as $\epsilon \downarrow 0$. Let us denote the coordinate domain neighbourhood given by the Rank Theorem as $U_{x^o(t_s^-)}$, where $x^o(t_s^-) = \Phi^{(t_s^-,t^1)}_{f_{q_0}}(x(t^1)) \in \mathcal{S}$.

Consider $0 < \delta t$ such that $\Phi^{(t_s+\delta t,t_s)}_{f_{q_0}}(x^o(t_s^-)) \in U_{x^o(t_s^-)}$. In the local coordinate system around $x^o(t_s^-)$ defined above, the switching manifold $\mathcal{S}$ separates $U_{x^o(t_s^-)}$ into two subsets $U^1_{x^o(t_s^-)}, U^2_{x^o(t_s^-)}$, where $U^1_{x^o(t_s^-)} = \{x \in U_{x^o(t_s^-)}, x_n < 0\}$ and $U^2_{x^o(t_s^-)} = \{x \in U_{x^o(t_s^-)}, x_n > 0\}$. For definiteness, we assume that first, the state trajectory enters $U^1_{x^o(t_s^-)}$ and second, it enters $U^2_{x^o(t_s^-)}$ after meeting the switching manifold;

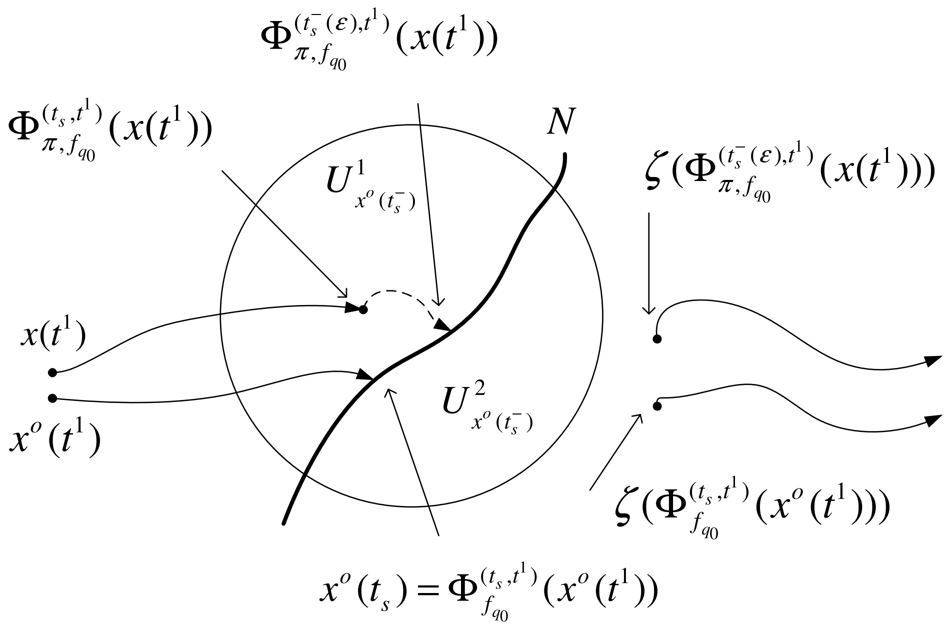

Figure A.1. Nominal and Perturbed State Trajectories

therefore $\Phi_{f_{q_0}}^{(t_s+\delta t,t_s)}(x^o(t_s^-)) \in U^2_{x^o(t_s^-)}$ for all sufficiently small $\delta t > 0$. The convergence of $\Phi_{\pi,f_{q_0}}^{(t_s+\delta t,t_s)}(x(t_s))$ to $\Phi_{f_{q_0}}^{(t_s+\delta t,t_s)}(x^o(t_s^-))$ implies that for sufficiently small ϵ, $\Phi_{\pi,f_{q_0}}^{(t_s+\delta t,t_s)}(x(t_s)) \in U^2_{x^o(t_s^-)}$, hence by the continuity of the trajectory there exists a switching time, $t_s(\epsilon)$, such that $\Phi_{\pi,f_{q_0}}^{(t_s(\epsilon),t_s)}(x(t_s)) \in \mathcal{S}$, see Figure A.1.

Furthermore by the continuity of the state trajectory, we may choose $0 \leq \epsilon$ sufficiently small that $\Phi_{\pi,f_{q_0}}^{(t_s(\epsilon),t_s)}(x(t_s)) \in U_{x^o(t_s^-)}$. Let us define $\Psi : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ by $\Psi(\epsilon, t) = x_n \circ \Phi_{\pi,f_{q_0}}^{(t,t^1)}(x(t^1))$, where x_n is the last coordinate function. Hence, the differentiability of Ψ with respect to t is immediate by the construction of Ψ since $\frac{d\Psi(\epsilon,t)}{dt}|_{t_s(\epsilon)} = f_{q_0}^n(x(t_s(\epsilon)))$, where $f_{q_0}^n$ is the corresponding coefficient of the last coordinate of f_{q_0}.

In order to show the differentiability of Ψ with respect to ϵ the following needle variation is applied

$$u_\pi(t,\epsilon) = \begin{cases} u^o(t) & t \le t^1 - \epsilon \\ u_1 & t^1 - \epsilon \le t \le t^1 \\ u^o(t) & t^1 < t \le t_s \\ u^o(t_s) & t_s \le t < t_s(\epsilon) \end{cases}, \tag{A.1}$$

We recall that $\Phi^{(t,s),x}_{\pi,f_q}(\epsilon) := \Phi^{(t,s)}_{f_q^{u_\pi(t,\epsilon)}}(x(s))$ then one can verify that, by the results of Proposition 2.2, the needle variation control $u_\pi(t,\epsilon)$, given in (2.14), results in the following tangent perturbation vector at t^1, where $\epsilon \in [0,\epsilon_0)$ for some $\epsilon_0 > 0$.

$$\begin{aligned} \frac{d}{d\epsilon}\Phi^{(t^1,s),x}_{\pi,f_{q_0}}|_\epsilon &= \lim_{\delta\to 0}\frac{\Phi^{(t^1,s),x}_{\pi,f_{q_0}}(\epsilon+\delta) - \Phi^{(t^1,s),x}_{\pi,f_{q_0}}(\epsilon)}{\delta} \\ &= T\Phi^{(t^1,t^1-\epsilon)}_{\pi,f_{q_0}}\Big(f_{q_0}(x(t^1-\epsilon),u_1) - f_{q_0}(x(t^1-\epsilon),u(t^1-\epsilon))\Big) \\ &\in T_{x(t^1)}\mathcal{M} = T_{\Phi^{(t^1,s),x}_{\pi,f_{q_0}}(\epsilon)}\mathcal{M}. \end{aligned} \tag{A.2}$$

That implies the differentiability of Ψ on $[0,\epsilon_0)$, where (see Figure A.2)

$$\frac{d}{d\epsilon}\Phi^{(t,t^1),x}_{\pi,f_{q_0}}|_\epsilon = T\Phi^{(t,t^1)}_{f_{q_0}}(\frac{d}{d\epsilon}\Phi^{(t^1,s),x}_{\pi,f_{q_0}}|_\epsilon), \quad t \in [t^1, t_s^-(\epsilon)).$$

The transversality hypothesis at the intersection of the state trajectory and the switching manifold implies that $f^n_{q_0}(x(t_s(\epsilon))) \neq 0$; then by employing the Implicit Function Theorem (see [**43**], Theorem 7.9) we have

$$\Psi(\epsilon, t_s(\epsilon)) = 0 \Rightarrow \exists\kappa : \mathbb{R} \to \mathbb{R}, \quad s.t. \quad \kappa(\epsilon) = t_s(\epsilon),$$

and κ and Ψ both are C^1; then the derivative of $\kappa(.)$ with respect to ϵ is given as

$$\frac{d\kappa(\epsilon)}{d\epsilon} = -(\frac{\partial\Psi}{\partial t})^{-1}|_{t=t_s(\epsilon)}.\frac{\partial\Psi}{\partial\epsilon} = -f^{n^{-1}}_{q_0}(x(t_s(\epsilon))).T^n\Phi^{(t_s(\epsilon),t^1)}_{f_{q_0}}(\frac{d}{d\epsilon}\Phi^{(t^1,s),x}_{\pi,f_{q_0}}|_\epsilon), \tag{A.3}$$

where $T^n\Phi^{(t_s(\epsilon),t^1)}_{f_{q_0}}(\frac{d}{d\epsilon}\Phi^{(t^1,s),x}_{\pi,f_{q_0,u_1}}|_\epsilon)$ is the nth coordinate of $T\Phi^{(t_s(\epsilon),t^1)}_{f_{q_0}}(\frac{d}{d\epsilon}\Phi^{(t^1,s),x}_{\pi,f_{q_0}}|_\epsilon)$.

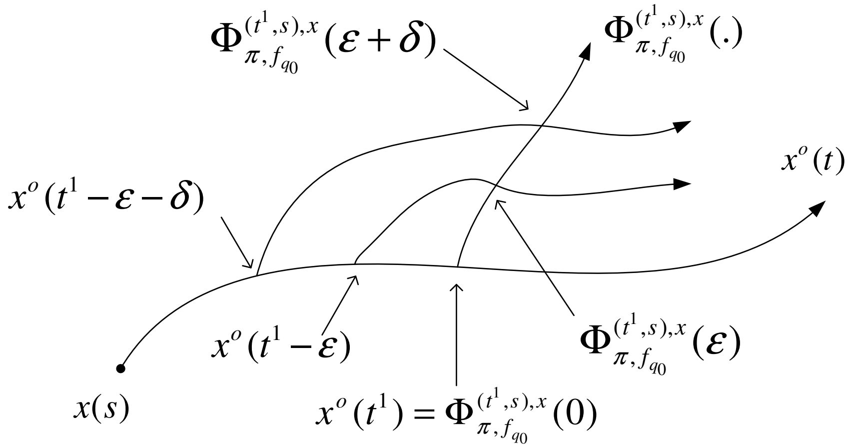

FIGURE A.2. Nominal and Perturbed State Trajectories

This completes the proof of differentiability of $t_s(\epsilon)$ with respect to ϵ. The proof for the differentiability of $t_s(\epsilon)$ in the case where $t_s(\epsilon) < t_s$ parallels the proof given above. □

A.2. Proof of Lemma 2.6

PROOF. Without loss of generality assume $t_s \le t_s(\epsilon)$, then

$$\Phi^{(t_s(\epsilon),t^1),x}_{\pi,f_{q_1}}(\epsilon) = \zeta \circ \Phi^{(t_s^-(\epsilon),t^1),x}_{\pi,f_{q0}}(\epsilon), \qquad t^1 \in [t_0,t_s), \tag{A.4}$$

where $\Phi^{(t_s^-(\epsilon),t^1),x}_{\pi,f_{q0}}(\epsilon) = \Phi^{(t_s^-(\epsilon),t_s)}_{\pi,f_{q0}} \circ \Phi^{(t_s^-,t^1),x}_{\pi,f_{q0}}(\epsilon)$ and $x(t^1)=x$, then in a local coordinate system of $x(t_s)$ we have

$$\zeta \circ \Phi^{(t_s^-(\epsilon),t_s)}_{\pi,f_{q0}} \circ \Phi^{(t_s^-,t^1),x}_{\pi,f_{q0}}(\epsilon) - \Phi^{(t_s(\epsilon),t_s)}_{f_{q_1}} \circ \zeta \circ \Phi^{(t_s^-,t^1),x}_{f_{q0}}(\epsilon) =$$
$$\Big\{\zeta\big(\int_{t_s}^{t_s(\epsilon)} f_{q0}(x_\epsilon(t),u^o(t_s))dt + x_\epsilon(t_s)\big)\Big\} - \Big\{\int_{t_s}^{t_s(\epsilon)} f_{q_1}(x^o(t),u^o(t))dt + \zeta(x^o(t_s^-))\Big\}. \tag{A.5}$$

Since $u(t)=u^o(t_s), t\in[t_s,t_s^-(\epsilon))$, $f_{q0}(x_\epsilon(t),u^o(t_s))$ is differentiable with respect to t. Hence by the Taylor expansion of ζ around $x_\epsilon(t_s)$ and the Mean Value Theorem we have

$$\left\{\zeta\big(\int_{t_s}^{t_s^-(\epsilon)} f_{q_0}(x_\epsilon(t), u^o(t_s))dt + x_\epsilon(t_s)\big)\right\} = \\ \zeta(x_\epsilon(t_s)) + (t_s(\epsilon) - t_s) \times T\zeta.f_{q_0}(x_\epsilon(\hat{t}), u^o(t_s)) + o(\delta x), \tag{A.6}$$

where $\hat{t} \in (t_s, t_s^-(\epsilon))$. Applying the Taylor expansion of ζ around $x^o(t_s)$ imples

$$\zeta(x_\epsilon(t_s)) - \zeta(x^o(t_s^-)) = T\zeta(x_\epsilon(t_s) - x^o(t_s^-)) + o(\delta x), \tag{A.7}$$

where by the definition of the derivatives we have

$$\begin{aligned} \frac{d\Phi_{\pi,f_{q_1}}^{(t_s(\epsilon),t^1),x}}{d\epsilon}\Big|_{\epsilon=0} &= \lim_{\epsilon\downarrow 0} \frac{\Phi_{\pi,f_{q_1}}^{(t_s(\epsilon),t^1),x} - \Phi_{f_{q_1}}^{(t_s(\epsilon),t^1),x}}{\epsilon} \\ &= \lim_{\epsilon\downarrow 0} \frac{\zeta \circ \Phi_{\pi,f_{q_0}}^{(t_s^-(\epsilon),t_s)} \circ \Phi_{\pi,f_{q_0}}^{(t_s^-,t^1),x} - \Phi_{f_{q_1}}^{(t_s(\epsilon),t_s)} \circ \zeta \circ \Phi_{f_{q_0}}^{(t_s^-,t^1),x}}{\epsilon}, \end{aligned} \tag{A.8}$$

therefore as $\epsilon \downarrow 0$, Lemma 2.5 and (A.8) together yield

$$\begin{aligned} \frac{d\Phi_{\pi,f_{q_1}}^{(t_s(\epsilon),t^1),x}}{d\epsilon}\Big|_{\epsilon=0} &= T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-,t^1)} \circ \frac{d\Phi_{\pi,f_{q_0}}^{(t^1,t_0),x_0}}{d\epsilon}\Big|_{\epsilon=0} \\ &\quad + \frac{dt_s(\epsilon)}{d\epsilon}\Big|_{\epsilon=0}.\big(T\zeta(f_{q_0}((x^o(t_s^-), u^o(t_s^-))) - f_{q_1}(x^o(t_s), u^o(t_s))\big). \end{aligned} \tag{A.9}$$

Lemma 2.2 and Proposition 2.2 complete the proof for the case $t_s \le t_s(\epsilon)$. The same argument holds for $t_s(\epsilon) < t_s$ with a sign change for $\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}$. It should be noted that the derivative in (A.8) gives the state variation at $t_s(\epsilon)$, therefore the nominal flow is subtracted from the perturbed one up to $t_s(\epsilon)$. □

A.3. Proof of Theorem 2.2

PROOF. We split the proof into the following steps: First, the needle variation is applied at t^1, where $t_s < t^1 \le t_f$ and t_s is the optimal switching time on the switching manifold from q_0 to q_1, hence there is no switching phenomena after t^1. At this stage the proof is same as the proof presented in [**3**] and [**6**]. Second, the needle variation is applied at t^1, $t_0 \leq t^1 < t_s$. Third, we show that the constructed adjoint variable, λ, satisfies the Hamiltonian equations and, fourth, the continuity of the Hamiltonian at the optimal switching state and time $(x^o(t_s), t_s)$ is obtained.

Step 1: Choose the following control needle variation:

$$u_\pi(t,\epsilon) = \begin{cases} u_1 & t^1 - \epsilon \leq t \leq t^1 \\ u^o(t) & \text{elsewhere} \end{cases}, \tag{A.10}$$

where $t_s < t^1 \leq t_f, u_1 \in U$. By Lemma 2.2 the state variation at t^1 is $[f_{q_1}(x^o(t^1), u_1) - f_{q_1}(x^o(t^1), u^o(t^1))] \in T_{x^o(t^1)}\mathcal{M}$. By the definition of K_{t_f} we have

$$T\Phi_{f_{q_1}}^{(t_f,t^1)}([f_{q_1}(x^o(t^1), u_1) - f_{q_1}(x^o(t^1), u^o(t^1))]) \in K_{t_f}^1 \subset T_{x^o(t_f)}\mathcal{M}. \tag{A.11}$$

Lemma 2.7 implies that

$$0 \leq \langle dh(x^o(t_f)), T\Phi_{f_{q_1}}^{(t_f,t^1)}([f_{q_1}(x^o(t^1), u_1) - f_{q_1}(x^o(t^1), u^o(t^1))])\rangle\ , \tag{A.12}$$

and by Proposition 2.1

$$\begin{aligned} 0 &\leq \langle dh(x^o(t_f)), T\Phi_{f_{q_1}}^{(t_f,t^1)} f_{q_1}(x^o(t^1), u_1) - f_{q_1}(x^o(t^1), u^o(t^1))\rangle \\ &= \langle T^*\Phi_{f_{q_1}}^{(t_f,t^1)} dh(x^o(t_f)), f_{q_1}(x^o(t^1), u_1) - f_{q_1}(x^o(t^1), u^o(t^1))\rangle, \\ &\qquad t_s < t^1 < t_f. \end{aligned} \tag{A.13}$$

Therefore

$$\begin{aligned} &\langle T^*\Phi_{f_{q_1}}^{(t_f,t^1)} dh(x^o(t_f)), f_{q_1}(x^o(t^1), u^o(t^1))\rangle \\ &\leq \langle T^*\Phi_{f_{q_1}}^{(t_f,t^1)} dh(x^o(t_f)), f_{q_1}(x^o(t^1), u_1))\rangle, \quad t_s < t^1 < t_f, \end{aligned} \tag{A.14}$$

for all $u_1 \in U$ and setting $p^o(t) := T^*\Phi_{f_{q_1}}^{(t_f,t)} dh(x^o(t_f))$ yields a trajectory $p^o(.)$ satisfying the minimization statement of the theorem.

Step 2: Here we use the needle variation before the optimal switching time t_s i.e:

$$u_\pi(t,\epsilon) = \begin{cases} u_1 & t^1-\epsilon \le t \le t^1 \\ u^o(t) & \text{elsewhere} \end{cases}, \tag{A.15}$$

where $t^1 < t_s, u_1 \in U$. Similar to the first step, the derivative of the state trajectory with respect to ϵ at t^1 is obtained as $[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))] \in T_x\mathcal{M}$ and $\frac{d}{d\epsilon}\Phi^{(t,s),x}_{\pi,f_{q_0}}|_{\epsilon=0} = \Psi(t), t \in [t^1,t_s)$. In order to use the method introduced in the first step, we describe the evolution of the perturbed state, $\Phi^{(t,s),x}_{\pi,f_q}$, after the switching time. Note that each elementary control variation, $u_\pi(t^1,\epsilon)$, results in a different switching time t_s which depends upon both of ϵ and u_1. Now let us consider a state mapping from $x(t^1)$ to the switching state $x(t_s^-(\epsilon))$ induced by the needle control variation; then the state variation at the optimal switching state $x^o(t_s^-)$ is obtained as the push forward of

$$\Phi^{(t_s^-(.),t^1),x}_{\pi,f_{q_0}} : [0,\tau] \to \mathcal{S}, \quad x \in \mathcal{M}, \quad x(t_s(\epsilon)) \in \mathcal{S}, \tag{A.16}$$

where $\Phi^{(t_s^-(\epsilon),t^1),x}_{\pi,f_{q_0}} := \Phi^{(t_s^-(\epsilon),t^1)}_{\pi,f_{q_0}}(x(t^1)) \in \mathcal{S}$ and $t_s(\epsilon)$ is the switching time corresponding to the selected ϵ. Here we have two possibilities, (i): $t_s \le t_s(\epsilon)$ and (ii): $t_s(\epsilon) < t_s$. The corresponding control needle variations for these two possibilities are given as follows:

$$(i): t_s \le t_s(\epsilon), \quad u_\pi(t,\epsilon) = \begin{cases} u^o(t) & t \le t^1-\epsilon \\ u_1 & t^1-\epsilon \le t \le t^1 \\ u^o(t) & t^1 < t \le t_s \\ u^o(t_s) & t_s \le t < t_s(\epsilon) \end{cases}, \tag{A.17}$$

and

$$(ii): t_s(\epsilon) < t_s, \quad u_\pi(t,\epsilon) = \begin{cases} u^o(t) & t \le t^1-\epsilon \\ u_1 & t^1-\epsilon \le t \le t^1 \\ u^o(t) & t^1 < t < t_s(\epsilon) \\ u^o(t_s) & t_s(\epsilon) \le t \le t_s \end{cases}. \tag{A.18}$$

Notice that $u^o(t_s)$ in (i) corresponds to f_{q_0} under the optimal control and in (ii) corresponds to f_{q_1} under the optimal control. The right differentiability of $t_s(\epsilon)$ with respect to ϵ at 0 by Lemma 2.5 (since the needle variation is defined for $0 \leq \epsilon$) and Lemma 2.6, in case (i), together imply

$$\begin{aligned}\frac{d\Phi^{(t_s^-(\epsilon),t^1),x}_{\pi,f_{q_0}}}{d\epsilon}|_{\epsilon=0} &= \big(\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}\big).f_{q_0}((x^o(t_s^-),u^o(t_s^-))) \\ &\quad + T\Phi^{(t_s^-,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))] \in T_{x^o(t_s^-)}\mathcal{S} \subset T_{x^o(t_s^-)}\mathcal{M}.\end{aligned} \tag{A.19}$$

And in case (ii) we have

$$\begin{aligned}\frac{d\Phi^{(t_s^-(\epsilon),t^1),x}_{\pi,f_{q_0}}}{d\epsilon}|_{\epsilon=0} &= -\big(\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}\big).f_{q_0}((x^o(t_s^-),u^o(t_s^-))) \\ &\quad + T\Phi^{(t_s^-,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))] \in T_{x^o(t_s^-)}\mathcal{S} \subset T_{x^o(t_s^-)}\mathcal{M}.\end{aligned} \tag{A.20}$$

In the first case, (2.53) and (A.19) together yield

$$\begin{aligned}\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0} &= -\langle dN_{x^o(t_s^-)}, f_{q_0}((x^o(t_s^-),u^o(t_s^-)))\rangle^{-1} \\ &\quad \times\langle dN_{x^o(t_s^-)}, T\Phi^{(t_s,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))]\rangle,\end{aligned} \tag{A.21}$$

and in the second case, (2.53) and (A.20) together yield

$$\begin{aligned}\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0} &= \langle dN_{x^o(t_s^-)}, f_{q_0}((x^o(t_s^-),u^o(t_s^-)))\rangle^{-1} \\ &\quad \times\langle dN_{x^o(t_s^-)}, T\Phi^{(t_s,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))]\rangle,\end{aligned} \tag{A.22}$$

where due to the transversality assumption in Definition 2.2, $\langle dN_{x^o(t_s^-)}, f_{q_0}((x^o(t_s^-), u^o(t_s^-)))\rangle \neq 0$.

(We notice that (A.21) coincides with (A.3) since in the coordinate system given in the proof of Lemma 2.5, $dN_{x^o(t_s^-)} = (0, ..., 0, 1)$ therefore

$$\langle dN_{x^o(t_s^-)}, f_{q_0}((x^o(t_s^-), u^o(t_s^-)))\rangle^{-1}\langle dN_{x^o(t_s^-)}, T\Phi^{(t_s,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1), u_1) \\ -f_{q_0}(x^o(t^1), u^o(t^1))]\rangle = f^{n^{-1}}_{q_0}(x(t_s)).T^n\Phi^{t_s,t^1}_{f_{q_0}}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))].)$$

Based on (A.19) and (A.20), we have $\frac{d\Phi^{(t_s^-(\epsilon),t^1),x}_{\pi,f_{q_0}}}{d\epsilon}|_{\epsilon=0} \in T_{x^o(t_s^-)}\mathcal{S}$, where for the simplicity of notation Ti is omitted. The variation of the state trajectory at t_s is obtained by evaluating $T\zeta$ on $(\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}).f_{q_0}((x^o(t_s^-), u^o(t_s^-)))+T\Phi^{(t_s,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1), u_1)-f_{q_0}(x^o(t^1), u^o(t^1))]$, where by definition, $T\zeta : T\mathcal{M} \to T\mathcal{M}$ is the push forward of ζ. Therefore

$$T\zeta\Bigg((\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}).f_{q_0}((x^o(t_s^-), u^o(t_s^-))) \;+\; T\Phi^{(t_s^-,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1), u_1) \\ - \; f_{q_0}(x^o(t^1), u^o(t^1))]\Bigg) \in T_{x^o(t_s)}\mathcal{M}. \quad \text{(A.23)}$$

Parallel to the results in [**66**], and following Lemma 2.6, in case (i), the state variation at t_s is

$$\frac{d\Phi^{(t_s,t_s(\epsilon))}_{\pi,f_{q_1}} \circ \Phi^{(t_s^-(\epsilon),t^1)}_{\pi,f_{q_0}}(x(t^1))}{d\epsilon}|_{\epsilon=0} \;=\; T\zeta \circ T\Phi^{(t_s^-,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))] \\ +\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}[T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-))) - f_{q_1}(x^o(t_s^-), u^o(t_s^-))] \in T_{x^o(t_s)}\mathcal{M}, \quad \text{(A.24)}$$

and in case (ii)

$$\frac{d\Phi^{(t_s(\epsilon),t_s)}_{\pi,f_{q_1}} \circ \Phi^{(t_s^-,t^1)}_{\pi,f_{q_0}}(x(t^1))}{d\epsilon}|_{\epsilon=0} = T\zeta \circ T\Phi^{(t_s^-,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))]$$
$$+\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}[f_{q_1}(x^o(t_s),u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-),u^o(t_s^-)))] \in T_{x^o(t_s)}\mathcal{M}. \tag{A.25}$$

Due to the sign change in (A.21) and (A.22), both of the cases (i) and (ii) give the same results as in (A.25) and (A.24) respectively. Henceforth, we only consider the second case. (A.22) and (A.25) together imply

$$\frac{d\Phi^{(t_s(\epsilon),t_s)}_{\pi,f_{q_1}} \circ \Phi^{(t_s^-,t^1)}_{\pi,f_{q_0}}(x(t^1))}{d\epsilon}|_{\epsilon=0} = T\zeta \circ T\Phi^{(t_s^-,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))]$$
$$+\langle dN_{x^o(t_s^-)}, f_{q_0}((x^o(t_s^-),u^o(t_s^-)))\rangle^{-1}\langle dN_{x^o(t_s^-)}, T\Phi^{(t_s^-,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1),u_1)$$
$$-f_{q_0}(x^o(t^1),u^o(t^1))]\rangle[f_{q_1}(x^o(t_s),u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-),u^o(t_s^-)))] \in T_{x^o(t_s)}\mathcal{M}, \tag{A.26}$$

where $T\Phi^{(t_f,t_s)}_{f_{q_1}}\Big(\frac{d\Phi^{(t_s(\epsilon),t_s)}_{\pi,f_{q_1}} \circ \Phi^{(t_s^-,t^1)}_{\pi,f_{q_0}}(x(t^1))}{d\epsilon}|_{\epsilon=0}\Big) \in K^2_{t_f}$ and by Lemma 2.7, we have

$$0 \le \langle dh(x^o(t_f)), T\Phi^{(t_f,t_s)}_{f_{q_1}}\Big(\frac{d\Phi^{(t_s(\epsilon),t_s)}_{\pi,f_{q_1}} \circ \Phi^{(t_s^-,t^1)}_{\pi,f_{q_0}}(x(t^1))}{d\epsilon}|_{\epsilon=0}\Big)\rangle, \tag{A.27}$$

therefore

$$0 \le \Big\langle dh(x^o(t_f)), T\Phi^{(t_f,t_s)}_{f_{q_1}}\Big\{T\zeta \circ T\Phi^{(t_s^-,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))]$$
$$+\langle dN_{x^o(t_s^-)}, f_{q_0}((x^o(t_s^-),u^o(t_s^-)))\rangle^{-1}$$
$$\times\langle dN_{x^o(t_s^-)}, T\Phi^{(t_s^-,t^1)}_{f_{q_0}}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))]\rangle$$
$$[f_{q_1}(x^o(t_s),u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-),u^o(t_s^-)))]\Big\}\Big\rangle, \tag{A.28}$$

which implies

$$
\begin{aligned}
0 \leq \langle T^{*}\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-,t^1)}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))] \\
+\langle dN_{x^o(t_s^-)}, f_{q_0}((x^o(t_s^-),u^o(t_s^-)))\rangle^{-1} \times \big\{\langle dN_{x^o(t_s^-)}, T\Phi_{f_{q_0}}^{(t_s^-,t^1)} \\
[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))]\rangle\big\} \\
\times[f_{q_1}(x^o(t_s),u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-),u^o(t_s^-)))]\rangle.
\end{aligned}
\tag{A.29}
$$

By the linearity of push-forwards (see [**43**]), (A.29) becomes

$$
\begin{aligned}
0 \leq \langle T^{*}\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), T\zeta \circ T\Phi_{f_{q_0}}^{(t_s^-,t^1)}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))]\rangle \\
+\langle dN_{x^o(t_s^-)}, f_{q_0}((x^o(t_s^-),u^o(t_s^-)))\rangle^{-1} \\
\times\langle dN_{x^o(t_s^-)}, T\Phi_{f_{q_0}}^{(t_s^-,t^1)}[f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))]\rangle \\
\times\langle T^{*}\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), f_{q_1}(x^o(t_s),u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-),u^o(t_s^-)))\rangle,
\end{aligned}
\tag{A.30}
$$

where one may write this as

$$
\begin{aligned}
0 \leq \langle T^{*}\Phi_{f_{q_0}}^{(t_s^-,t^1)} \circ T^{*}\zeta \circ T^{*}\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))\rangle \\
+\mu\langle T^{*}\Phi_{f_{q_0}}^{(t_s^-,t^1)} dN_{x^o(t_s^-)}, f_{q_0}(x^o(t^1),u_1) - f_{q_0}(x^o(t^1),u^o(t^1))\rangle,
\end{aligned}
\tag{A.31}
$$

where

$$
\begin{aligned}
\mu &= \langle dh(x^o(t_f)), T\Phi_{f_{q_1}}^{(t_f,t_s)}[f_{q_1}((x^o(t_s),u^o(t_s)) - T\zeta(f_{q_0}((x^o(t_s^-),u^o(t_s^-)))]\rangle \\
&\quad \times\langle dN_{x^o(t_s^-)}, f_{q_0}((x^o(t_s^-),u^o(t_s^-)))\rangle^{-1} \in R.
\end{aligned}
\tag{A.32}
$$

Applying Proposition 2.1 to (A.31) on $[t^1, t_s^-]$, we have

$$\begin{aligned}&\langle T^*\Phi_{f_{q_0}}^{(t_s^-,t^1)} \circ T^*\zeta \circ T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f))\\&+\mu T^*\Phi_{f_{q_0}}^{(t_s^-,t^1)} dN_{x^o(t_s^-)}, f_{q_0}(x^o(t^1), u^o(t^1))\rangle\\&\leq \langle T^*\Phi_{f_{q_0}}^{(t_s^-,t^1)} \circ T^*\zeta \circ T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f))\\&+\mu T^*\Phi_{f_{q_0}}^{(t_s^-,t^1)} dN_{x^o(t_s^-)}, f_{q_0}(x^o(t^1), u_1)\rangle;\end{aligned} \tag{A.33}$$

then, as in the first step, define

$$\begin{aligned}p^o(t) := T^*\Phi_{f_{q_0}}^{(t_s^-,t)} \circ T^*\zeta \circ T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f))\\+\mu T^*\Phi_{f_{q_0}}^{(t_s^-,t)} dN_{x^o(t_s^-)}, \quad t \in [t_0, t_s).\end{aligned} \tag{A.34}$$

Since $T^*\Phi_{f_{q_0}}^{(t_s^-,t_s)} = I$, choosing $t^1 = t_s$ gives

$$p^o(t_s^-) = T^*\zeta(p^o(t_s)) + \mu dN_{x^o(t_s^-)}. \tag{A.35}$$

Following (2.28) in the non-hybrid case, the Hamiltonian function is defined as

$$\begin{aligned}H_{q_0}(x^o(t), p^o(t), u^o(t)) &= \langle\big\{T^*\Phi_{f_{q_0}}^{(t_s^-,t)} \circ T^*\zeta \circ T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f))\\&+ \mu T^*\Phi_{f_{q_0}}^{(t_s^-,t)} dN_{x^o(t_s^-)}\big\}, f_{q_0}(x^o(t), u^o(t)\rangle, \quad t \in [t_0, t_s).\end{aligned} \tag{A.36}$$

Step 3: We need to show $\lambda^o(t) = (x^o(t), p^o(t)) = (x^o(t), T^*\Phi_{f_{q_1}}^{(t_f,t)} dh(x^o(t_f)))$, $t \in [t_0, t_s)$ and
$\lambda^o(t) = (x^o(t), T^*\Phi_{f_{q_0}}^{(t_s^-,t)} \circ T^*\zeta \circ T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)) + \mu T^*\Phi_{f_{q_0}}^{(t_s^-,t)} dN_{x^o(t_s^-)}$, $t \in [t_s, t_f]$ satisfy (2.31). By the definition of Hamiltonian functions given by (2.27) and (2.28), it is obvious that $\dot{x}(t) = \frac{\partial H_i}{\partial p}$, $i = 0, 1$. To prove $\dot{p^o}(t) = -\frac{\partial H_i}{\partial x}$, $i = 0, 1$, first we use the adjoint curve expression $\lambda(t), t \in [t_s, t_f]$ given by (A.14). Therefore we have

$$\dot{p^o}(t) = \frac{d}{dt} T^*\Phi_{f_{q_1}}^{(t_f,t)} dh(x^o(t_f))), \tag{A.37}$$

where together with Lemma 2.3 and 2.22 implies

$$\dot{p^o}(t) = \left[-(\frac{\partial f^i_{q_1}}{\partial x^j} p^j) \right]^n_{i,j=1} = -\frac{\partial H_{q_1}(x^o(t), p^o(t))}{\partial x}. \qquad \text{(A.38)}$$

Same argument holds for $p^o(t) = T^*\Phi^{(t_s^-,t)}_{f_{q_0}} \circ T^*\zeta \circ T^*\Phi^{(t_f,t_s)}_{f_{q_1}} dh(x^o(t_f))$
$+ \mu T^*\Phi^{(t_s^-,t)}_{f_{q_0}} dN_{x^o(t_s^-)}, \quad t \in [t_0, t_s).$

Step 4: Here we complete the proof by obtaining the continuity of the Hamiltonian at the optimal switching time t_s. In [**66**], the Hamiltonian continuity based on the control needle variation approach is derived only for controlled switching hybrid systems. We give a continuity proof in the case of autonomous switching hybrid systems via the following algebraic steps.
Notice that

$$\begin{aligned}
&\langle T^*\Phi^{(t_f,t_s)}_{f_{q_1}} dh(x^o(t_f)), [f_{q_1}((x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}((x^o(t_s^-), u^o(t_s^-)))] \rangle = \\
&\langle T^*\Phi^{(t_f,t_s)}_{f_{q_1}} dh(x^o(t_f)), \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s)) \rangle^{-1} \times \\
&\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s)) \rangle \times \\
&[f_{q_1}((x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}((x^o(t_s^-), u^o(t_s^-)))] \rangle.
\end{aligned} \qquad \text{(A.39)}$$

Therefore by (A.39) we have

$$
\begin{aligned}
H_{q_1}(x^o(t_s), p^o(t_s), u^o(t_s)) &= \langle p(t_s), f_{q_1}((x^o(t_s), u^o(t_s))\rangle \\
&= \langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), f_{q_1}((x^o(t_s), u^o(t_s))\rangle \quad \text{by A.14} \\
&= \langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), T\zeta(f_{q_0}((x^o(t_s^-), u^o(t_s^-)))\rangle \\
&+ \langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), \\
&\quad \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s))\rangle^{-1} \\
&\times [f_{q_1}((x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}((x^o(t_s^-), u^o(t_s^-)))]\rangle \\
&\times \langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s))\rangle \quad \text{by A.39} \\
&= \langle T^*\zeta \circ T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), f_{q_0}((x^o(t_s^-), u^o(t_s^-))\rangle \\
&+ \mu\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s))\rangle \quad \text{by A.32,} \qquad (A.40)
\end{aligned}
$$

hence by the definition of p in (A.14) and (A.34) we have

$$
\langle p(t_s), f_{q_1}((x^o(t_s), u^o(t_s))\rangle = \langle p(t_s^-), f_{q_0}((x^o(t_s^-), u^o(t_s^-))\rangle, \qquad (A.41)
$$

which gives the continuity of the Hamiltonian at the switching time t_s. □

It should be noted that setting $\zeta = I$ above subsumes the results obtained in [**66**] for non-impulsive autonomous hybrid systems.

A.4. Proof of Theorem 2.3

The following results for the variation of the hybrid value function is presented and then a complete proof of Theorem 2.3 is provided.

Since $\mathcal{S}$ is an embedded submanifold of $\mathcal{M}$, necessarily there exists an embedding i from $\mathcal{S}$ to $i(\mathcal{S}) \subset \mathcal{M}$. The push-forward of i is given as

$$
Ti : T_x\mathcal{S} \to T_x\mathcal{M}. \qquad (A.42)
$$

For any tangent vector $X \in T_x\mathcal{S}$, the image vector $Ti(X) \in T_x\mathcal{M}$ is a tangent vector on $\mathcal{M}$. There exists a local coordinate representation of X, i.e. $X = \sum_{j=1}^{n} X^j \frac{\partial}{\partial x^j}$, such that $X \in T_x\mathcal{S}$ if and only if $X^j = 0, \quad j > k$, where k is the dimension of $\mathcal{S}$, see [**43**]. The following lemma gives a relation between $dv(x^o(t_s^-), t_s) = \sum_{j=1}^{n} \frac{\partial v(x^o(t_s), t_s)}{\partial x^j} dx^j \in T^*_{x^o(t_s^-)}\mathcal{M}$ where $v(x^o(t_s), t_s)$ is smooth by ***A4*** and a tangent vector $X \in T_{x^o(t_s^-)}\mathcal{M}$ which is also a tangent vector in $T_{x^o(t_s^-)}\mathcal{S}$ in the local coordinate system given above. The statement of the following lemma is given for a general embedded submanifold $\mathcal{S}$ which is not necessarily $n-1$ dimensional.

Lemma A.1. *Consider an MHOCP with a single switching from the discrete state q_0 to the discrete state q_1 at the unique switching time t_s on the optimal trajectory $x^o(.)$ and a k dimensional embedded switching manifold $\mathcal{S} \subset \mathcal{M}$ satisfying **A1**-**A4**; then at the optimal switching state $x^o(t_s) \in \mathcal{S}$ and switching time t_s, we have*

$$\langle dv(x^o(t_s^-), t_s), X \rangle = 0, \quad \forall X \in Ti(T_{x^o(t_s^-)}\mathcal{S}). \tag{A.43}$$

Proof. Since $X \in Ti(T_{x^o(t_s^-)}\mathcal{S})$ there exists $X_\mathcal{S} \in T_{x^o(t_s^-)}\mathcal{S}$ such that $X = Ti(X_\mathcal{S})$. By applying the same extension method employed in Lemma 2.7, we extend $X_\mathcal{S}$ to a vector field $X'_\mathcal{S} \in \mathcal{X}(\mathcal{S})$ such that $X'_\mathcal{S}(x^o(t_s^-)) = X_\mathcal{S}$.

Let us denote the induced Riemannian metric from $\mathcal{M}$ to $\mathcal{S}$ as $g_\mathcal{S}$. By the fundamental theorem of existence of geodesics and the Taylor expansion on Riemannian manifolds we have

$$\begin{aligned} v((exp_{x^o(t_s^-)}\theta X_\mathcal{S}), t_s) &= v(x^o(t_s^-), t_s) + \theta(\nabla'_{X'_\mathcal{S}} v)(x^o(t_s^-), t_s) + o(\theta), \\ &\quad 0 < \theta < \theta^*, \end{aligned} \tag{A.44}$$

where ∇' is the Levi-Civita connection of $\mathcal{S}$ with respect to the induced metric $g_\mathcal{S}$. Since $\mathcal{S}$ is an embedded submanifold of $\mathcal{M}$, the inclusion map is a full rank homeomorphism from $\mathcal{S}$ to $i(\mathcal{S})$, therefore, for each $X \in Ti(T_{x^o(t_s^-)}\mathcal{S})$ the corresponding $X_\mathcal{S}$ is unique. The vector space property of $T_{x^o(t_s^-)}\mathcal{S}$ implies $-X_\mathcal{S} \in T_{x^o(t_s)}\mathcal{S}$, hence, by the optimality of $x^o(t_s^-)$ on $\mathcal{S}$ and the accessibility of $\dot{x}(t) = f_{q_0}(x, u)$, an application of (A.44) to v along $-X_\mathcal{S}$ gives

$$\nabla'_{X'_\mathcal{S}} v = 0, \quad \forall X'_\mathcal{S} \in T_{x^o(t_s^-)}\mathcal{S}. \tag{A.45}$$

(A.45) and (2.41) (ii) together imply

$$\frac{\partial v}{\partial x^j}(x^o(t_s^-), t_s) = 0, \quad j = 1, ..., k, \tag{A.46}$$

where k is the dimension of $\mathcal{S}$. In the local coordinates of $x^o(t_s) \in \mathcal{M}$, (A.46) yields

$$\langle dv(x^o(t_s^-), t_s), X \rangle = \langle \sum_{j=1}^{n} \frac{\partial v(x^o(t_s^-), t_s)}{\partial x^j} dx^j, \sum_{j=1}^{n} X^j \frac{\partial}{\partial x^j} \rangle, \tag{A.47}$$

where (A.46) together with $X^j = 0, j > k$ completes the proof. □

The proof of Theorem 2.3 is then given as follows:

PROOF. The proof is parallel to the proof of Theorem 2.2 where by Lemma A.1, $dN_{x^o(t_s^-)}$ is replaced by $dv(x^o(t_s^-), t_s)$. □

A.5. Proof of Theorem 2.4

PROOF. The first step of the proof of Theorem 2.2 is unchanged. For the control needle variation before the optimal switching time t_s, i.e. step 2, in case (i): $t_s \le t_s(\epsilon)$, we have

$$\begin{aligned}\frac{d\Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t^1), x}}{d\epsilon}|_{\epsilon=0} \oplus \frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0} \frac{\partial}{\partial t_s} &= (\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}) \\ &\times f_{q_0}((x^o(t_s^-), u^o(t_s^-))) + T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) \\ &- f_{q_0}(x^o(t^1), u^o(t^1))] \oplus \frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0} \frac{\partial}{\partial t_s} \in T_{(x^o(t_s), t_s)}\mathcal{S}.\end{aligned} \tag{A.48}$$

And in case (ii), i.e. $t_s(\epsilon) \le t_s$, we have

$$\begin{aligned}\frac{d\Phi_{\pi, f_{q_0}}^{(t_s^-(\epsilon), t^1), x}}{d\epsilon}|_{\epsilon=0} \oplus \frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0} \frac{\partial}{\partial t_s} &= -\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0} \\ &\times f_{q_0}((x^o(t_s^-), u^o(t_s^-))) + T\Phi_{f_{q_0}}^{(t_s^-, t^1)}[f_{q_0}(x^o(t^1), u_1) \\ &- f_{q_0}(x^o(t^1), u^o(t^1))] \oplus \frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0} \frac{\partial}{\partial t_s} \in T_{(x^o(t_s), t_s)}\mathcal{S}.\end{aligned} \tag{A.49}$$

Therefore by (2.67) we have

$$\langle dN_{(x^o(t_s^-),t_s)}, \frac{d\Phi_{\pi,f_{q_0}}^{(t_s^-(\epsilon),t^1),x}}{d\epsilon}|_{\epsilon=0} \oplus \frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}\frac{\partial}{\partial t_s}\rangle =$$
$$\langle dN_x^o(t_s^-), \frac{d\Phi_{\pi,f_{q_0}}^{(t_s^-(\epsilon),t^1),x}}{d\epsilon}|_{\epsilon=0}\rangle + \frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0}\langle dN_{t_s}, \frac{\partial}{\partial t}\rangle = 0, \tag{A.50}$$

and finally in case (i) we have

$$\begin{aligned}\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0} &= -\Big(\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s))\rangle + \langle dN_{t_s}, \frac{\partial}{\partial t}\rangle\Big)^{-1} \\ &\times\Big\langle dN_{x^o(t_s^-)}, T\Phi_{f_{q_0}}^{(t_s^-,t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))]\Big\rangle,\end{aligned} \tag{A.51}$$

and in case (ii)

$$\begin{aligned}\frac{dt_s(\epsilon)}{d\epsilon}|_{\epsilon=0} &= \Big(\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s))\rangle + \langle dN_{t_s}, \frac{\partial}{\partial t}\rangle\Big)^{-1} \\ &\times\Big\langle dN_{x^o(t_s^-)}, T\Phi_{f_{q_0}}^{(t_s^-,t^1)}[f_{q_0}(x^o(t^1), u_1) - f_{q_0}(x^o(t^1), u^o(t^1))]\Big\rangle,\end{aligned} \tag{A.52}$$

where μ in (A.34) is given as

$$\begin{aligned}\mu = &\Big\langle dh(x^o(t_f)), T\Phi_{f_{q_1}}^{(t_f,t_s)}[f_{q_1}((x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}((x^o(t_s^-), u^o(t_s^-)))]\Big\rangle \\ &\times\Big(\langle dN_{x^o(t_s^-)}, f_{q_0}((x^o(t_s^-), u^o(t_s^-)))\rangle + \langle dN_{t_s}, \frac{\partial}{\partial t}\rangle\Big)^{-1}.\end{aligned} \tag{A.53}$$

Following the steps of the proof of Theorem 2.2 we have

$$\begin{aligned}&p^o(t_s^-) = T^*\zeta(p^o(t_s)) + \mu dN_{x^o(t_s^-)}, \\ &p^o(t_s^-) \in T^*_{x^o(t_s^-)}\mathcal{M}, \quad p^o(t_s) \in T^*_{x^o(t_s)}\mathcal{M}, \\ &x^o(t_s) = \zeta(x^o(t_s^-)),\end{aligned} \tag{A.54}$$

where

$$\begin{aligned} p^o(t) &:= T^*\Phi_{f_{q_0}}^{(t_s^-,t)} \circ T^*\zeta \circ T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)) \\ &\quad +\mu T^*\Phi_{f_{q_0}}^{(t_s^-,t)} dv(x^o(t_s^-),t_s), \quad t \in [t_0, t_s), \end{aligned} \tag{A.55}$$

and

$$p^o(t) := T^*\Phi_{f_{q_1}}^{(t_f,t)} dh(x^o(t_f)), \quad t \in [t_s, t_f]. \tag{A.56}$$

Step 3 in the proof of Theorem 2.2 also holds for time varying switching cases. To analyze the possible discontinuity of the Hamiltonian we employ the same method as that used in step 4 of the proof of Theorem 2.2. Therefore

$$\begin{aligned} &\langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), [f_{q_1}((x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-)))] \rangle = \\ &\Big\langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), \big[\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-))\rangle \\ &+\langle dN_{t_s}, \frac{\partial}{\partial t}\rangle\big]^{-1} \times \Big(\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-))\rangle + \langle dN_{t_s}, \frac{\partial}{\partial t}\rangle\Big) \\ &\times [f_{q_1}((x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-)))] \Big\rangle, \end{aligned} \tag{A.57}$$

which implies

$$
\begin{aligned}
&H_{q_1}(x^o(t_s), p^o(t_s), u^o(t_s)) = \langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), f_{q_1}((x^o(t_s), u^o(t_s))\rangle \quad \text{by A.14} \\
&= \langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), T\zeta(f_{q_0}((x^o(t_s^-), u^o(t_s^-)))\rangle \\
&+ \Big\langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), \\
&\big\{\big(\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s))\rangle + \langle dN_{t_s}, \frac{\partial}{\partial t}\rangle\big)^{-1} \\
&\times [f_{q_1}((x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}((x^o(t_s^-), u^o(t_s^-)))]\big\}\Big\rangle \\
&\times \big(\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s))\rangle + \langle dN_{t_s}, \frac{\partial}{\partial t}\rangle\big) \quad \text{by A.57} \\
&= \langle T^*\zeta \circ T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), f_{q_0}((x^o(t_s^-), u^o(t_s^-))\rangle \\
&+ \mu\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s))\rangle + \mu\langle dN_{t_s}, \frac{\partial}{\partial t}\rangle \quad \text{by A.53,}
\end{aligned}
\tag{A.58}
$$

and finally we have

$$
\langle p^o(t_s)), f_{q_1}((x^o(t_s), u^o(t_s))\rangle = \langle p^o(t_s^-)), f_{q_0}((x^o(t_s^-), u^o(t_s^-))\rangle + \mu\langle dN_{t_s}, \frac{\partial}{\partial t}\rangle,
\tag{A.59}
$$

or equivalently

$$
H_{q_0}(x^o(t_s^-), p^o(t_s^-), u^o(t_s^-)) = H_{q_1}(x^o(t_s), p^o(t_s), u^o(t_s)) - \mu\langle dN_{t_s}, \frac{\partial}{\partial t}\rangle,
\tag{A.60}
$$

which completes the proof. □

A.5.1. Interior Optimal Switching States, Time Varying Switching Manifolds and Impulsive Jumps. In this section we extend Theorem 2.4 to MHOCPs satisfying ***A4*** where the switching manifold $\mathcal{S}$ and the impulsive jump $\hat{\zeta}$ are both time varying. The results here are consistent with the results presented in [**57**].

In the case where the switching manifold is a time variant submanifold $\mathcal{S} \subset \mathcal{M} \times \mathbb{R}$, we have

$$\hat{d}v(x,t) \in T^*_{(x,t)}(\mathcal{M} \times \mathbb{R}) = T^*_x\mathcal{M} \oplus T^*_t\mathbb{R}, \tag{A.61}$$

where locally

$$\hat{d}v(x^o(t_s^-),t_s) = \sum_{j=1}^{n} \frac{\partial v(x^o(t_s^-),t_s)}{\partial x^j} dx^j + D_t^* v(x^o(t_s^-),t_s)dt \in T^*_{x^o(t_s^-)}\mathcal{M} \oplus T^*_{t_s}\mathbb{R}. \tag{A.62}$$

The following lemma is an extension of Lemma A.1 on time varying switching manifolds.

Lemma A.2. *For an MHOCP with a single switching from the discrete state q_0 to the discrete state q_1 at the unique switching time t_s on the optimal trajectory $x^o(.)$ and an embedded time varying switching manifold $\mathcal{S} \subset \mathcal{M} \times \mathbb{R}$ of dimension $k \leq dim(\mathcal{M})$; then at the optimal switching state and time $(x^o(t_s^-),t_s) \in \mathcal{S}$,*

$$\langle \hat{d}v(x^o(t_s^-),t_s), X\rangle = 0, \quad \forall X \in T_{(x^o(t_s^-),t_s)}\mathcal{S}. \tag{A.63}$$

Proof. The proof is parallel to the proof of Lemma A.1 concerning the extra variable t_s. □

Theorem A.1. *Consider an impulsive MHOCP satisfying hypotheses **A1-A4**; then corresponding to the optimal control and optimal trajectory u^o, x^o, there exists a nontrivial adjoint trajectory $\lambda^o(.) = (x^o(.), p^o(.)) \in T^*\mathcal{M}$ defined along the optimal state trajectory such that:*

$$H_{q_i}(x(t),p^o(t),u^o(t)) \leq H_{q_i}(x(t),p^o(t),u_1), \quad \forall u_1 \in U, t \in [t_0,t_f], i = 0,1, \tag{A.64}$$

and the corresponding optimal adjoint variable $\lambda^o(.) \in T^*\mathcal{M}$*, locally given as* $\lambda^o(.) = (x^o(.), p^o(.))$*, satisfies*

$$\dot{\lambda}^o(t) = \overrightarrow{H}_{q_i}(\lambda^o(t)), \quad t \in [t_0, t_f], i = 0, 1. \tag{A.65}$$

At the optimal switching state and switching time, $(x^o(t_s), t_s)$*, we have*

$$\begin{aligned} &p^o(t_s^-) = T^*\zeta(p^o(t_s)) + \mu dv(x^o(t_s^-), t_s), \\ &p^o(t_s^-) \in T^*_{x^o(t_s^-)}\mathcal{M}, \quad p^o(t_s) \in T^*_{x^o(t_s)}\mathcal{M}, \\ &x^o(t_s) = \zeta(x^o(t_s^-)), \end{aligned} \tag{A.66}$$

$$x^o(t_0) = x_0, \;\; p^o(t_f) = dh(x^o(t_f)) \in T^*_{x^o(t_f)}\mathcal{M}, \;\; dh = \sum_{i=1}^{n} \frac{\partial h}{\partial x^i} dx^i \in T^*_x\mathcal{M}, \tag{A.67}$$

where $\mu \in R$*,*

$$T^*\hat{\zeta} = T^*\zeta \oplus D_t^*\zeta : T^*\mathcal{M} \to T^*\mathcal{M} \oplus T^*\mathbb{R}, \tag{A.68}$$

and

$$T^*\zeta : T^*\mathcal{M} \to T^*\mathcal{M}, \quad D_t^*\zeta : T^*\mathcal{M} \to T^*\mathbb{R}. \tag{A.69}$$

The discontinuity of the Hamiltonian at $(x^o(t_s^-), t_s)$*, is given as follows:*

$$\begin{aligned} &H_{q_0}(x^o(t_s^-), p^o(t_s^-), u^o(t_s^-)) = \\ &H_{q_1}(x^o(t_s), p^o(t_s), u^o(t_s)) - D_t^*\zeta(p^o(t_s)) - \mu D_t^* v(x^o(t_s^-), t_s). \end{aligned} \tag{A.70}$$

Proof. The proof is parallel to that of Theorem 2.4 and employs the results of Lemma A.2. □

A.6. Proof of Theorem 2.5

PROOF. The proof closely parallels the proof of Theorem 2.4 where

$$\begin{aligned} p^o(t) := T^*\Phi_{f_{q_0}}^{(t_s^-,t)} \circ T^*\zeta \circ T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)) \\ +\mu T^*\Phi_{f_{q_0}}^{(t_s^-,t)} dN_{x^o(t_s^-)}, \quad t \in [t_0, t_s), \end{aligned} \tag{A.71}$$

and

$$p^o(t) := T^*\Phi_{f_{q_1}}^{(t_f,t)} dh(x^o(t_f)), \quad t \in [t_s, t_f], \tag{A.72}$$

where

$$\begin{aligned} \mu = \langle dh(x^o(t_f)), T\Phi_{f_{q_1}}^{(t_f,t_s)}[f_{q_1}(x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-))) \\ -D_t\zeta(x^o(t_s), t_s)]\rangle \times \big(\langle dN_{x^o(t_s^-)}, f_{q_0}((x^o(t_s^-), u^o(t_s^-)))\rangle + \langle dN_{t_s}, \frac{\partial}{\partial t}\rangle\big)^{-1}. \end{aligned} \tag{A.73}$$

It should be noted that $D_t\zeta(x^o(t_s), t_s)(\frac{\partial}{\partial t}) \in T\mathcal{M}$ and for simplicity we drop $\frac{\partial}{\partial t}$. To prove the Hamiltonian discontinuity we have

$$\begin{aligned} &\Big\langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), \Big\{ f_{q_1}(x^o(t_s), u^o(t_s)) \\ &-T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-))) - D_t\zeta(x^o(t_s), t_s) \Big\} \Big\rangle \\ &= \Big\langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), \big(\langle dv(x^o(t_s^-), t_s), f_{q_0}(x^o(t_s^-), u^o(t_s^-))\rangle \\ &+\langle dN_{t_s}, \frac{\partial}{\partial t}\rangle\big)^{-1} \times \big(\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-))\rangle + \langle dN_{t_s}, \frac{\partial}{\partial t}\rangle\big) \\ &\times[f_{q_1}(x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-))) - D_t\zeta(x^o(t_s), t_s)] \Big\rangle, \end{aligned} \tag{A.74}$$

which implies

$$
\begin{aligned}
&H_{q_1}(x^o(t_s), p^o(t_s), u^o(t_s)) = \left\langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), f_{q_1}(x^o(t_s), u^o(t_s))\right\rangle \quad \text{by A.14}\\
&= \left\langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), T\zeta(f_{q_0}((x^o(t_s^-), u^o(t_s^-)))\right\rangle\\
&+\langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), D_t\zeta(x^o(t_s), t_s)\rangle\\
&+\left\langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), \left(\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s))\rangle + \langle dN_{t_s}, \frac{\partial}{\partial t}\rangle\right)^{-1}\right.\\
&\left.\times[f_{q_1}(x^o(t_s), u^o(t_s)) - T\zeta(f_{q_0}(x^o(t_s^-), u^o(t_s^-)))]\right\rangle\\
&\times\left(\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s), u^o(t_s))\rangle + \langle dN_{t_s}, \frac{\partial}{\partial t}\rangle\right) \quad \text{by A.74}\\
&= \langle T^*\zeta \circ T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), f_{q_0}(x^o(t_s^-), u^o(t_s^-))\rangle\\
&+\mu\langle dN_{x^o(t_s^-)}, f_{q_0}(x^o(t_s^-), u^o(t_s^-))\rangle\\
&+D_t^*\zeta(T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f))) + \mu\langle dN_{t_s}, \frac{\partial}{\partial t}\rangle \quad \text{by A.73}\\
&= H_{q_0}(x^o(t_s^-), p^o(t_s^-), u^o(t_s^-)) + D_t^*\zeta(T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f))) + \mu\langle dN_{t_s}, \frac{\partial}{\partial t}\rangle,
\end{aligned}
\tag{A.75}
$$

where by the definition of pullbacks (see [**43**])

$$\langle T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f)), D_t\zeta(x^o(t_s), t_s)\rangle = D_t^*\zeta(T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f))) \in R, \tag{A.76}$$

and $p^o(t_s) = T^*\Phi_{f_{q_1}}^{(t_f,t_s)} dh(x^o(t_f))$. The remaining of the proof is similar to that of (A.59) and (A.60). □

APPENDIX B

Proofs and Extended Results for Chapter 5

B.1. Proof of Theorem 5.1

Proof. The state variation is derived by subtracting the perturbed and nominal states as follows

$$\begin{aligned} x^{\beta}(t_f) - x^{\alpha}(t_f) &= \int_0^{t^{\alpha}} \big(f_1(x^{\beta}(\tau), u^{\beta}(\tau)) - f_1(x^{\alpha}(\tau), u^{\alpha}(\tau))\big) d\tau \\ &+ \int_{t^{\alpha}}^{t^{\beta}} \big(f_1(x^{\beta}(\tau), u^{\beta}(\tau)) - f_2(x^{\alpha}(\tau), u^{\alpha}(\tau))\big) d\tau \\ &+ \int_{t^{\beta}}^{t_f} \big(f_2(x^{\beta}(\tau), u^{\beta}(\tau)) - f_2(x^{\alpha}(\tau), u^{\alpha}(\tau))\big) d\tau. \end{aligned} \tag{B.1}$$

The problem here is the dependency on the optimal controls in the integral equations. To be able to expand the terms above we replace the optimal controls with the corresponding optimal adjoint variables. The formal optimal control definition is given by

$$u^*(t) = Argmin_{u \in \mathcal{U}} H(x, u, \lambda), \tag{B.2}$$

where $\mathcal{U}$ is the set of admissible controls. (B.2) suggests to write the optimal control as a function of optimal state and adjoint process, so $u^*(.)$ at time t is computed by

$x^*(.), \lambda^*(.)$ at time t. Here we write the optimal control by

$$u^*(t) = u^*(x^*(t), \lambda^*(t)). \tag{B.3}$$

Rewriting (B.1) we have (for simplicity in our notation we drop $*$)

$$\begin{aligned} x^\beta(t_f) - x^\alpha(t_f) &= \int_0^{t^\alpha} \big(f_1(x^\beta(\tau), \lambda^\beta(\tau)) - f_1(x^\alpha(\tau), \lambda^\alpha(\tau))\big) d\tau \\ &+ \int_{t^\alpha}^{t^\beta} \big(f_1(x^\beta(\tau), \lambda^\beta(\tau)) - f_2(x^\alpha(\tau), \lambda^\alpha(\tau))\big) d\tau \\ &+ \int_{t^\beta}^{t_f} \big(f_2(x^\beta(\tau), \lambda^\beta(\tau)) - f_2(x^\alpha(\tau), \lambda^\alpha(\tau))\big) d\tau. \end{aligned} \tag{B.4}$$

Now expanding (B.4) around the nominal trajectories x^α, λ^α we have

$$\begin{aligned} x^\beta(t_f) - x^\alpha(t_f) &= \int_0^{t^\alpha} \frac{\partial f_1}{\partial x}(x^\alpha(\tau), \lambda^\alpha(\tau))\delta x + \frac{\partial f_1}{\partial \lambda}(x^\alpha(\tau), \lambda^\alpha(\tau))\delta\lambda d\tau \\ &+ \big(f_1(x^\beta(\tau), \lambda^\beta(\tau)) - f_2(x^\alpha(\tau), \lambda^\alpha(\tau))\big)\delta t \\ &+ \int_{t^\beta}^{t_f} \frac{\partial f_2}{\partial x}(x^\alpha(\tau), \lambda^\alpha(\tau))\delta x + \frac{\partial f_2}{\partial \lambda}(x^\alpha(\tau), \lambda^\alpha(\tau))\delta\lambda d\tau + o(\delta x) + o(\delta\lambda). \end{aligned} \tag{B.5}$$

So in the limit we have

$$\begin{aligned} y(t_f) &= \int_0^{t^\alpha} \frac{\partial f_1}{\partial x}(x^\alpha(\tau), \lambda^\alpha(\tau))y(\tau) + \frac{\partial f_1}{\partial \lambda}(x^\alpha(\tau), \lambda^\alpha(\tau))z(\tau) d\tau \\ &+ \big(f_1(x^\beta(\tau), \lambda^\beta(\tau)) - f_2(x^\alpha(\tau), \lambda^\alpha(\tau))\big) \\ &+ \int_{t^\alpha}^{t_f} \frac{\partial f_2}{\partial x}(x^\alpha(\tau), \lambda^\alpha(\tau))y(\tau) + \frac{\partial f_2}{\partial \lambda}(x^\alpha(\tau), \lambda^\alpha(\tau))z(\tau) d\tau. \end{aligned} \tag{B.6}$$

Therefore we have

$$y(t) = \int_0^{t} \frac{\partial f_1}{\partial x}(x^\alpha(\tau), \lambda^\alpha(\tau))y(\tau) + \frac{\partial f_1}{\partial \lambda}(x^\alpha(\tau), \lambda^\alpha(\tau))z(\tau) d\tau \quad t \in [0, t^\alpha), \tag{B.7}$$

and

$$
\begin{aligned}
y(t) &= \int_0^{t^\alpha} \frac{\partial f_1}{\partial x}(x^\alpha(\tau),\lambda^\alpha(\tau))y(\tau) + \frac{\partial f_1}{\partial \lambda}(x^\alpha(\tau),\lambda^\alpha(\tau))z(\tau)d\tau \\
&+ \big(f_1(x^\beta(\tau),\lambda^\beta(\tau)) - f_2(x^\alpha(\tau),\lambda^\alpha(\tau))\big) \\
&+ \int_{t^\alpha}^{t} \frac{\partial f_2}{\partial x}(x^\alpha(\tau),\lambda^\alpha(\tau))y(\tau) + \frac{\partial f_2}{\partial \lambda}(x^\alpha(\tau),\lambda^\alpha(\tau))z(\tau)d\tau, \quad t \in [t^\alpha, t_f].
\end{aligned}
\tag{B.8}
$$

For the adjoint variables the evolution equations are given as follows:

$$
\lambda^{\alpha,\beta}(t) = \lambda^{\alpha,\beta}(t_f) + \int_t^{t_f} \frac{\partial H_2}{\partial x}(x^{\alpha,\beta}(\tau),\lambda^{\alpha,\beta}(\tau))d\tau, \quad t \in (t^{\alpha,\beta}, t_f]. \tag{B.9}
$$

Therefore $z(.)$ satisfies the following equation

$$
\begin{aligned}
z(t) &= \int_t^{t_f} \frac{\partial^2 H_2}{\partial x^2}(x^\alpha(\tau),\lambda^\alpha(\tau))y(\tau)d\tau \\
&+ \int_t^{t_f} \frac{\partial^2 H_2}{\partial x \partial \lambda}(x^\alpha(\tau),\lambda^\alpha(\tau))z(\tau) + \frac{\partial^2 h}{\partial x^2} y(t_f) \quad t \in (t^\alpha, t_f].
\end{aligned}
\tag{B.10}
$$

For $t < t^\alpha$ equations need little modification on the switching time in order to be computable. Writing the backward equation for the adjoint variables we have

$$
\lambda^{\alpha,\beta}(t) = \lambda^{\alpha,\beta}(t^{\alpha,\beta}) + \int_t^{t^{\alpha,\beta}} \frac{\partial H_1}{\partial x}(x^{\alpha,\beta}(\tau),\lambda^{\alpha,\beta}(\tau))d\tau, \quad t \in [t_0, t^{\alpha,\beta}). \tag{B.11}
$$

Applying the results of Theorem 3.1 on the adjoint process discontinuity at switching times we have

$$
\begin{aligned}
\lambda^\beta(t) - \lambda^\alpha(t) &= \lambda^\beta(t^\beta) - \lambda^\alpha(t^\alpha) - p^\beta \nabla_x m(x^\beta, t^\beta) + p^\alpha \nabla_x m(x^\alpha, t^\alpha) \\
&+ \int_t^{t^\alpha} \frac{\partial H_1}{\partial x}(x^\beta(\tau),\lambda^\beta(\tau)) - \frac{\partial H_1}{\partial x}(x^\alpha(\tau),\lambda^\alpha(\tau))d\tau \\
&+ \int_{t^\alpha}^{t^\beta} \frac{\partial H_1}{\partial x}(x^\beta(\tau),\lambda^\beta(\tau))d\tau,
\end{aligned}
\tag{B.12}
$$

and

$$\begin{aligned} \lambda^\beta(t^\beta) - \lambda^\alpha(t^\alpha) &= \int_{t^\beta}^{t_f} \frac{\partial H_2}{\partial x}(x^\beta(\tau), \lambda^\beta(\tau)) - \frac{\partial H_2}{\partial x}(x^\alpha(\tau), \lambda^\alpha(\tau)) d\tau \\ &- \int_{t^\alpha}^{t^\beta} \frac{\partial H_2}{\partial x}(x^\alpha(\tau), \lambda^\alpha(\tau)) d\tau + \frac{\partial h}{\partial x}(x^\beta(t_f)) - \frac{\partial h}{\partial x}(x^\alpha(t_f)). \end{aligned} \tag{B.13}$$

The discontinuity parameters p^α and p^β are parameters of $(x^{\alpha,\beta}, \lambda^{\alpha,\beta}, t^{\alpha,\beta})$. Therefore

$$\begin{aligned} p^\beta \nabla_x m(x^\beta, t^\beta) - p^\alpha \nabla_x m(x^\alpha, t^\alpha) &= \frac{\partial(p\nabla_x m(x,t))}{\partial x}(x^\alpha, t^\alpha)\delta x + \frac{\partial(p\nabla_x m(x,t))}{\partial \lambda_1}(x^\alpha, t^\alpha)\delta\lambda_1 \\ &+ \frac{\partial(p\nabla_x m(x,t))}{\partial \lambda_2}(x^\alpha, t^\alpha)\delta\lambda_2 + \frac{\partial(p\nabla_x m(x,t))}{\partial t}(x^\alpha, t^\alpha)\delta t \\ &+ o(\delta x) + o(\delta\lambda_1) + o(\delta\lambda_2) + o(\delta t), \end{aligned} \tag{B.14}$$

where

$$\delta\lambda_1 = \lambda^\beta(t^{-\beta}) - \lambda^\alpha(t^{-\alpha}), \tag{B.15}$$

$$\delta\lambda_2 = \lambda^\beta(t^{+\beta}) - \lambda^\alpha(t^{+\alpha}), \tag{B.16}$$

and

$$\delta x = x^\beta(t^\beta) - x^\alpha(t^\alpha). \tag{B.17}$$

It is shown that

$$\begin{aligned} \frac{\partial x}{\partial t^\alpha} &= y(t^\alpha) + f_1(x^\alpha(t^\alpha), \lambda^\alpha(t^\alpha)), \\ \frac{\partial \lambda_1}{\partial t^\alpha} &= z(t^{-\alpha}) - \frac{\partial H_1}{\partial x}(x^\alpha(t^\alpha), \lambda^\alpha(t^\alpha)), \\ \frac{\partial \lambda_2}{\partial t^\alpha} &= z(t^\alpha) - \frac{\partial H_2}{\partial x}(x^\alpha(t^\alpha), \lambda^\alpha(t^\alpha)). \end{aligned} \tag{B.18}$$

Employing (B.14,B.18) in (B.12) the proof is obtained. □

Document Log:

Manuscript Version 1 — 9 October 2012

Typeset by $\mathcal{AMS}$-LaTeX — 9 October 2012

Farzin Taringoo

Center for Intelligent Machines, McGill University, 3480 University Street, Montréal (Québec) H3A 0E9, Canada

E-mail address: `taringoo@cim.mcgill.ca`

Typeset by $\mathcal{AMS}$-LaTeX