OBSTRUCTIONS TO MOTION PLANNING BY THE CONTINUATION METHOD

by

David Scott Cameron Amiss

A thesis submitted to the Graduate Program in Chemical Engineering

in conformity with the requirements for the

Degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

December 2012

Abstract

The subject of this thesis is the motion planning algorithm known as the continuation method. To solve motion planning problems, the continuation method proceeds by lifting curves in state space to curves in control space; the lifted curves are the solutions of special initial value problems called path-lifting equations. To validate this procedure, three distinct obstructions must be overcome.

The first obstruction is that the endpoint maps of the control system under study must be twice continuously differentiable. By extending a result of A. Margheri, we show that this differentiability property is satisfied by an inclusive class of time-varying fully nonlinear control systems.

The second obstruction is the existence of singular controls, which are simply the singular points of a fixed endpoint map. Rather than attempting to completely characterize such controls, we demonstrate how to isolate control systems for which no controls are singular. To this end, we build on the work of S. A. Vakhrameev to obtain a necessary and sufficient condition. In particular, this result accommodates time-varying fully nonlinear control systems.

The final obstruction is that the solutions of path-lifting equations may not exist globally. To study this problem, we work under the standing assumption that the control system under study is control-affine. By extending a result of Y. Chitour, we show that the question of global existence can be resolved by examining Lie bracket configurations and momentum functions.

Finally, we show that if the control system under study is completely unobstructed with respect to a fixed motion planning problem, then its corresponding endpoint map is a fiber bundle. In this sense, we obtain a necessary condition for unobstructed motion planning by the continuation method.

For Azadeh

Acknowledgements

First and foremost, I sincerely thank Dr. Martin Guay. Although the coffee runs played a role, this thesis owes its existence primarily to his guidance, his enthusiasm, and his steadfast belief in my abilities.

For generously offering their time and expertise, I thank committee members Dr. Manfredi Maggiore, Dr. Abdol-Reza Mansouri, Dr. P. James McLellan, Dr. Xiang Li, and Dr. Serdar Yüksel. In particular, I must acknowledge the advice of Dr. Mansouri and Dr. McLellan, which guided this thesis at an early stage. I also thank Dr. Mansouri for many illuminating discussions, mathematical and otherwise, over the last few years.

For sharing the highs and lows of graduate life, I thank my officemates Samandeep Dhaliwal, Sean Dougherty, Kai Höffner, Nicolas Hudon, Michael Liba, Ehsan Moshksar, Heung-Youl Derek Park, and Walter Weiss de la Vega. Kai and Nick deserve very special recognition, as they were integral parts of my life as a Ph.D. student.

For their friendship and support, I thank Nazanin Alavi Tabari, Déborah Geist, Bahman Gharesifard, Sina Khosravi Maram, Ali Naji Almassi, Mohsen Omrani, Amir Tahmasebi, Shahram Yousefi, and Reza Zamani.

Finally, but beyond all measure, I thank Azadeh Moghtaderi for her love and understanding. With the greatest joy, it is to her that I dedicate this thesis.

Table of contents

Abstract ii

Dedication iii

Acknowledgements iv

Table of contents v

List of figures xi

Chapter 1 Introduction 1

1.1 Motion planning problems 1

1.1.1 The continuation method 4

1.1.2 The first obstruction: Ill-posed path-lifting equations . . 6

1.1.3 The second obstruction: Singular controls 6

1.1.4 The third obstruction: State explosions 7

1.2 Problem statement 8

1.3 Organization and contributions 9

Chapter 2 Preliminaries 13

2.1 Notation and preliminary results 13

2.1.1 Sets and maps 13

2.1.2 Vector spaces, linear maps, and multilinear maps 14

2.1.3 Measure and integration 17
2.1.4 Derivatives 18
2.1.4.1 Total derivatives 19
2.1.4.2 Partial derivatives 19
2.1.4.3 Mixed partial derivatives 21
2.1.5 The chain rule and Leibniz rule 22
2.1.6 L^p and locally L^p spaces 24
2.1.6.1 L^p spaces 24
2.1.6.2 Locally L^p spaces 25
2.1.6.3 Hölder's inequality for k-fold products 26
2.1.7 Manifolds 26
2.2 Initial value problems evolving on open subsets of Euclidean spaces 28
2.3 Initial value problems evolving on finite-dimensional manifolds . 38

Chapter 3 Control systems **54**
3.1 Control systems 55
3.2 C^q_p and C^q_q-polynomial control systems evolving on open subsets of Euclidean spaces 63
3.2.1 Basic definitions and properties 63
3.2.2 A special property of total derivatives 69
3.2.3 Differentiability 71
3.3 C^q_p and C^q_q-polynomial control systems evolving on finite-dimensional manifolds 72
3.3.1 Basic definitions and properties 73
3.3.2 A special property of total derivatives 74
3.3.3 Differentiability 75

Chapter 4 The continuation method **81**

4.1 Right inverses 82

4.2 Initial value problems evolving on Banach manifolds 86

4.3 The continuation method 91

4.3.1 Path-lifting equations 91

4.3.2 The continuation method 97

4.3.2.1 Basic definitions 97

4.3.2.2 The general case 98

4.3.2.3 The case where no controls are singular 100

Chapter 5 Operations on time-varying vector fields 101

5.1 Vertical lifts 102

5.1.1 Basic definitions and properties 102

5.1.2 Composition with second differentials 105

5.2 Tangent lifts 106

5.2.1 A local result concerning total derivatives 106

5.2.2 Basic definitions and properties 107

5.3 Cotangent lifts 109

5.3.1 A local result concerning adjoint total derivatives 109

5.3.2 Basic definitions and properties 110

5.4 Pullbacks 112

5.5 Pullbacks involving lifts 118

5.6 A particular global flow 120

5.7 Time derivatives of pullbacks 121

5.8 Pullbacks involving lifts: An integral formula 125

5.9 Scalar parameter derivatives of pullbacks 126

5.9.1 Time-varying vector fields with scalar parameters 126

5.9.2 Scalar parameter derivatives of pullbacks 129

Chapter 6 Differentials of endpoint maps 133

6.1 First variations . . . 134

6.1.1 Basic definitions and properties . . . 135

6.1.2 First variations compute differentials . . . 139

6.2 Examples . . . 141

6.3 Discussion . . . 143

6.3.1 L^∞ controls . . . 143

6.3.2 Related results . . . 144

Chapter 7 Intrinsic quadratic differentials of anchored endpoint maps 146

7.1 Intrinsic quadratic differentials . . . 147

7.1.1 Basic definitions and properties . . . 148

7.1.2 Locally constant-rank maps . . . 152

7.2 Second variations . . . 153

7.2.1 Basic definitions and properties . . . 154

7.2.2 Second variations compute intrinsic quadratic differentials 158

7.3 Examples . . . 164

7.4 Discussion . . . 165

7.4.1 L^∞ controls . . . 166

7.4.2 Related results . . . 166

7.4.3 Intrinsic differentials of higher order . . . 166

Chapter 8 Constant-rank conditions 167

8.1 First-order Pontryagin cones . . . 169

8.2 A sufficient constant-rank condition . . . 170

8.3 Symmetric Lebesgue points . . . 173

8.4 A characterization of first-order Pontryagin cones . . . 175

8.5 Containment lemmas 179

8.6 A necessary and sufficient constant-rank condition 181

8.7 Subimmersivity 190

Chapter 9 Sublinear growth **192**

9.1 Cotangent lifts of control systems 194

9.2 Lie derivatives, momentum functions, and switching functions . 195

9.3 Sublinear growth 196

9.4 An alternative characterization of sublinear growth 199

9.5 The main result 203

9.5.1 The Υ-condition 203

9.5.2 The locally Lipschitz-AC chain rule 205

9.5.3 The main result 206

Chapter 10 A necessary condition for unobstructed motion planning by the continuation method **219**

10.1 Fiber bundles 220

10.2 PLE-completeness 221

10.3 Initial value problems evolving on Banach manifolds 222

10.4 The main results 225

10.5 Unobstructed control systems 235

Chapter 11 Examples **237**

11.1 Hirschorn's system 237

11.2 An augmented version of Hirschorn's system 241

11.3 A special class of control-affine systems 245

Chapter 12 Conclusions **248**

12.1 Summary 248

12.2 Future work 249
12.2.1 Numerical implementation 250
12.2.2 Relax the sublinear growth conditions 250
12.2.3 Beyond submersivity 251

Bibliography **252**

A Proofs **265**

A.1 Lemma 5.2.6 265
A.2 Lemma 5.9.9 267
A.3 Lemma 5.9.10 268
A.4 Lemma 5.9.11 269
A.5 Lemma 8.2.3 270

List of figures

1.1 An illustration of lifting π to $\mathbf{\Pi}$ through $\mathrm{End}^{\Sigma}_{x_0}$ 3

2.1 An illustration of the proof of Theorem 2.3.25 53

3.1 An illustration of the proof of Theorem 3.3.11 80

4.1 An illustration of $\mathrm{Lift}(q) \in \mathrm{Hom}(T_{F(q)}R, T_qQ)$ 92

4.2 An illustration of lifting $\dot{\pi}(t)$ for each $t \in [0,1]$ 93

5.1 An illustration of the pointwise vertical lift operation 104

8.1 An illustration of the proof of Theorem 8.4.3 176

9.1 An illustration of the compact subset K^* 204

9.2 An illustration of the proof of Lemma 9.5.10 215

9.3 An illustration of the proof of Theorem 9.5.12 218

10.1 An illustration of a fiber bundle $p : E \to B$ 221

Chapter 1

Introduction

1.1 Motion planning problems

In motion planning problems, one is faced with a system whose behaviour can be manipulated, and a high-level description of some desired system behaviour. The problem is to translate the high-level description into low-level instructions. The low-level instructions, when implemented, should produce the desired system behaviour. In this generality, the label "motion planning problem" applies in a very wide variety of situations, as described by LaValle [2006, Chapter 1].

In this thesis, we are exclusively interested in the case where the system is a deterministic nonlinear control system evolving in continuous time. For example, consider a vehicle moving in an obstacle-rich environment. One might be faced with the problem of steering the vehicle from an initial point to a target point, in a way that avoids the obstacles. In this case, the low-level instructions are represented by a control action which effects the point-to-point maneuver. Another example is provided by a robotic manipulator, such as a multi-jointed arm with a grasping tool. One might be faced with the problem of taking the manipulator from an initial pose to a desired final pose through a sequence of joint positions. In this case, the low-level instructions are represented by a control action which effects the desired sequence of joint positions.

Consider a control system Σ with dynamical description

$$\dot{x}(t) = f(t, x(t), \boldsymbol{u}(t)).$$

We assume that $t \in J = [a, b]$, the state $x(t)$ evolves in a connected Riemannian manifold M, and the control $\boldsymbol{u}$ belongs to a collection $\mathscr{U}$ of maps of the form

$$\boldsymbol{u} : J \to \mathbb{R}^r.$$

The $\boldsymbol{u}$-controlled trajectory of Σ with initial condition (t_0, x_0) is denoted by

$$\mu^{\Sigma}(\cdot, t_0, x_0, \boldsymbol{u}) : J \to M.$$

This curve represents the state evolution of Σ, starting from state x_0 at time t_0, as it is actuated by the control $\boldsymbol{u}$. We assume that an initial state $x_0 \in M$ is fixed, and that Σ is completely controllable from x_0 on J. The latter assumption means, precisely, that for each state $x \in M$ there exists a control $\boldsymbol{u} \in \mathscr{U}$ such that

$$\mu^{\Sigma}(b, a, x_0, \boldsymbol{u}) = x.$$

In other words, $\boldsymbol{u}$ takes Σ from state x_0 at time a to state x at time b.

The x_0-*anchored motion planning problem* (MPP) for Σ is posed as follows:

Problem: For each $x \in M$, find $\boldsymbol{u} \in \mathscr{U}$ such that $\mu^{\Sigma}(b, a, x_0, \boldsymbol{u}) = x$.

A number of exact and approximate solutions of this problem can be found in the literature. These solutions take a wide variety of analytical approaches, based on the theory of differential flatness [Martin 1992, Lamiraux and Laumond 2000], the notion of dynamic extensions [Sussmann 1991], the technical apparatus of exterior differential systems [Tilbury et al. 1995], the imposition of hierarchical relations between two control systems [Tabuada and Pappas 2005], the application of highly

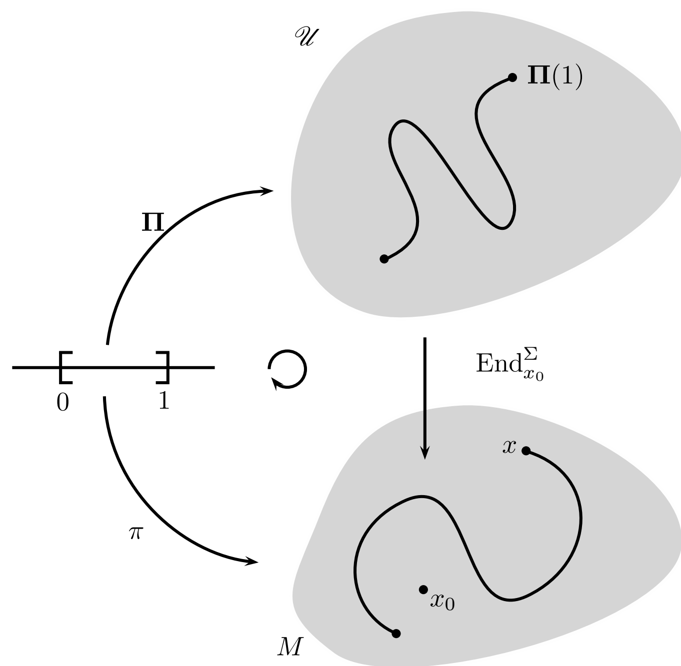

Figure 1.1: An illustration of lifting π to $\mathbf{\Pi}$ through $\mathrm{End}^{\Sigma}_{x_0}$

oscillatory inputs [Liu 1997, Leonard and Krishnaprasad 1995, Murray and Sastry 1993], the use of nilpotent approximating systems [Lafferriere 1991, Lafferriere and Sussmann 1991, 1993], the notion of non-singular loops [Sontag 1995], and the use of the continuation method [Sussmann 1992, 1993, Chitour 1996, Chitour and Sussmann 1998, Chitour 2002, 2006].

Broadly speaking, the continuation method is the subject of this thesis. Prior work on the continuation method was carried out under the assumption that Σ is a driftless C^∞ control-affine system; that is, the dynamical description of Σ is

$$\dot{x}(t) = \sum_{i=1}^{r} u^i(t) f_i(x(t)), \qquad \boldsymbol{u} = (u^1, \ldots, u^r),$$

relative to C^∞ vector fields $f_1, \ldots, f_r$. However, as we will see in the next section, the continuation method does not rely fundamentally on this assumption.

1.1.1 The continuation method

We begin by introducing the x_0-anchored endpoint map

$$\mathrm{End}_{x_0}^{\Sigma} : \mathscr{U} \to M$$

that sends $\boldsymbol{u}$ to

$$\mathrm{End}_{x_0}^{\Sigma}(\boldsymbol{u}) = \mu^{\Sigma}(b, a, x_0, \boldsymbol{u}).$$

In terms of this map, the x_0-anchored MPP can be recast as follows:

PROBLEM: For each $x \in M$, find $\boldsymbol{u} \in \mathscr{U}$ such that $\mathrm{End}_{x_0}^{\Sigma}(\boldsymbol{u}) = x$.

In Figure 1.1, we illustrate the following general procedure: Given $x \in M$,

1. Choose a curve $\pi : [0,1] \to M$ with $\pi(1) = x$,
2. Lift π to a curve $\mathbf{\Pi} : [0,1] \to \mathscr{U}$ through $\mathrm{End}_{x_0}^{\Sigma}$, in the sense that

$$\mathrm{End}_{x_0}^{\Sigma} \circ \mathbf{\Pi} = \pi, \qquad \text{and}$$

3. Choose $\boldsymbol{u} = \mathbf{\Pi}(1)$.

By construction, the control $\boldsymbol{u}$ satisfies

$$\mathrm{End}_{x_0}^{\Sigma}(\boldsymbol{u}) = \mathrm{End}_{x_0}^{\Sigma} \circ \mathbf{\Pi}(1) = \pi(1) = x.$$

The continuation method follows this general procedure. In rough terms, the lifted curve $\mathbf{\Pi}$ is constructed by leveraging differentiability of $\mathrm{End}_{x_0}^{\Sigma}$ while respecting its singular points. To make this more precise, we assume that

- $\mathscr{U}$ is the Hilbert space $L^2(J, \mathbb{R}^r)$,
- $\mathrm{End}_{x_0}^{\Sigma}$ is C^1 (that is, continuously differentiable) so that its differential $T\mathrm{End}_{x_0}^{\Sigma}$ is well-defined, and

- $\mathscr{U}_{x_0}^{\text{sing}}$ is the set of singular points of $\text{End}_{x_0}^{\Sigma}$.

Provided that π and $\boldsymbol{u}_0 \in \mathscr{U}$ are such that

- The constraint $\text{image}(\pi) \subseteq \text{End}_{x_0}^{\Sigma}(\mathscr{U} \setminus \mathscr{U}_{x_0}^{\text{sing}})$ is satisfied and
- $\boldsymbol{u}_0$ is contained in $(\text{End}_{x_0}^{\Sigma})^{-1}(\pi(0)) \cap (\mathscr{U} \setminus \mathscr{U}_{x_0}^{\text{sing}})$,

the lifted curve $\mathbf{\Pi}$ is taken to be the C^1 solution of the *path-lifting equation*

$$\begin{cases} \dot{\mathbf{\Pi}}(t) = T\text{End}_{x_0}^{\Sigma}(\mathbf{\Pi}(t))^{\#} \cdot H_\pi(t, \text{End}_{x_0}^{\Sigma} \circ \mathbf{\Pi}(t)), \quad \mathbf{\Pi}(t) \in \mathscr{U} \setminus \mathscr{U}_{x_0}^{\text{sing}}, \quad t \in [0,1] \\ \mathbf{\Pi}(0) = \boldsymbol{u}_0. \end{cases} \tag{1.1}$$

Here, H_π is a time-varying vector field on M such that

$$H_\pi(t, \pi(t)) = \dot{\pi}(t)$$

and the superscript $\#$ denotes the Moore–Penrose pseudoinverse.

Despite the rather complicated appearance of (1.1), it arises naturally from the problem at hand. To see this, suppose that $\mathbf{\Pi}$ is the solution of (1.1). Using the chain rule and the fact that $T\text{End}_{x_0}^{\Sigma}(\mathbf{\Pi}(t))^{\#}$ is a right inverse of $T\text{End}_{x_0}^{\Sigma}(\mathbf{\Pi}(t))$,

$$\dot{\overbrace{\text{End}_{x_0}^{\Sigma} \circ \mathbf{\Pi}}}(t) = T\text{End}_{x_0}^{\Sigma}(\mathbf{\Pi}(t)) \cdot \dot{\mathbf{\Pi}}(t) = H_\pi(t, \text{End}_{x_0}^{\Sigma} \circ \mathbf{\Pi}(t)).$$

Since $H_\pi(t, \pi(t)) = \dot{\pi}(t)$, we see that $\text{End}_{x_0}^{\Sigma} \circ \mathbf{\Pi}$ and π are solutions of H_π with

$$\text{End}_{x_0}^{\Sigma} \circ \mathbf{\Pi}(0) = \text{End}_{x_0}^{\Sigma}(\boldsymbol{u}_0) = \pi(0).$$

The uniqueness of solutions yields the desired result, namely that

$$\text{End}_{x_0}^{\Sigma} \circ \mathbf{\Pi} = \pi.$$

The preceding discussion demonstrates the plausibility of constructing $\mathbf{\Pi}$ as the solution of (1.1), although we have sidestepped three issues. These issues are so fundamental that they constitute obstructions to the continuation method.

1.1.2 The first obstruction: Ill-posed path-lifting equations

The first obstruction is that path-lifting equations may be ill-posed. That is, the solution of (1.1) may not exist, and, even if it does exist, it may not be unique. One can show, however, that this difficulty is alleviated whenever $\mathrm{End}_{x_0}^{\Sigma}$ is C^2 (that is, twice continuously differentiable). In the literature on the continuation method, a result known as Bismut's theorem is used to ensure high-order differentiability of $\mathrm{End}_{x_0}^{\Sigma}$; see [Chitour 2006]. Bismut's theorem implies that $\mathrm{End}_{x_0}^{\Sigma}$ is C^∞ whenever Σ is a driftless C^∞ control-affine system

$$\dot{x}(t) = \sum_{i=1}^{r} u^i(t) f_i(x(t))$$

and $f_1(x), \ldots, f_r(x)$ are linearly independent for each $x \in M$. From a control-theoretic point of view, Bismut's theorem is quite restrictive. Other than the fact that it only applies to driftless control-affine systems, the linear independence condition puts topological restrictions on M. For example, it is not hard to see that M cannot be a product manifold where one of the factors is an even-dimensional sphere. These limitations motivate the search for a less restrictive high-order differentiability result.

1.1.3 The second obstruction: Singular controls

The second obstruction to the continuation method is the existence of singular controls. Recall that the curve π must satisfy the constraint

$$\mathrm{image}(\pi) \subseteq \mathrm{End}_{x_0}^{\Sigma}(\mathscr{U} \setminus \mathscr{U}_{x_0}^{\mathrm{sing}}). \tag{1.2}$$

One can show [Bellaïche 1996] that if Σ is a driftless C^∞ control-affine system and the set $\mathscr{U} \setminus \mathscr{U}_{x_0}^{\mathrm{sing}}$ is nonempty, then this constraint is rendered trivial. That is,

$$\mathrm{End}_{x_0}^{\Sigma}(\mathscr{U} \setminus \mathscr{U}_{x_0}^{\mathrm{sing}}) = M.$$

For other control systems, however, verifying (1.2) requires a complete characterization of the set $\mathscr{U}_{x_0}^{\text{sing}}$. Although specialized results exist [Chitour and Sussmann 1998, Popa and Wen 2000, Chitour 2002, Chelouah and Chitour 2003], the problem of completely characterizing $\mathscr{U}_{x_0}^{\text{sing}}$ does not appear to be tractable, in general. Because of this difficulty, it seems that the only way forward is to restrict attention to control systems for which $\mathscr{U}_{x_0}^{\text{sing}}$ is minimized. Since the issue of minimizing $\mathscr{U}_{x_0}^{\text{sing}}$ is also connected with the third obstruction to the continuation method, our discussion of this issue is continued in the next section.

1.1.4 The third obstruction: State explosions

The third obstruction to the continuation method is that the solution of a given path-lifting equation (1.1) may not be defined on $[0, 1]$. In general, the solution $\mathbf{\Pi}$ may be defined on the interval $[0, \delta)$ for some $\delta \in (0, 1]$. Since the continuation method hinges on the ability to choose

$$\boldsymbol{u} = \mathbf{\Pi}(1),$$

this possibility must be ruled out. It turns out that, if

$$\operatorname{image}(\pi) \subseteq M \setminus \overline{\operatorname{End}_{x_0}^{\Sigma}(\mathscr{U}_{x_0}^{\text{sing}})}, \tag{1.3}$$

then $\mathbf{\Pi}$ is defined on $[0, \delta)$ if and only if a "state explosion" occurs. That is,

$$\lim_{t \nearrow \delta} \|\dot{\mathbf{\Pi}}(t)\| = \infty.$$

Consequently, one can show that $\mathbf{\Pi}$ is defined on $[0, 1]$ whenever the Moore–Penrose pseudoinverses $T\operatorname{End}_{x_0}^{\Sigma}(\boldsymbol{u})^{\#}$ satisfy a suitable sublinear growth condition.

This approach, which we will use in this thesis, engenders its own difficulties. As in the preceding section, verifying that (1.3) is satisfied requires a complete characterization of the set $\mathscr{U}_{x_0}^{\text{sing}}$, and one is again forced to restrict attention to control systems for which $\mathscr{U}_{x_0}^{\text{sing}}$ is minimized. This explains why the literature concerning the continuation method has been focused on strongly bracket-generating (SBG) driftless control-affine systems [Sussmann 1993, Chitour 1996, Chitour and Sussmann 1998, Chitour 2006]. Indeed, if Σ is an SBG driftless control-affine system with $r < n$, then $\mathscr{U}_{x_0}^{\text{sing}} = \{\mathbf{0}\}$. This renders (1.2) trivial and drastically simplifies the task of verifying (1.3). A result of Sussmann [1993] tells us that the SBG condition impinges on sublinear growth conditions as well. To be precise, this result states that if Σ is an SBG driftless control-affine system, then the Moore–Penrose pseudoinverses $T\mathrm{End}_{x_0}^{\Sigma}(\boldsymbol{u})^{\#}$ satisfy a sublinear growth condition uniformly for

$$\boldsymbol{u} \in (\mathrm{End}_{x_0}^{\Sigma})^{-1}(\mathrm{image}(\pi)).$$

An alternative proof of this result can be found in [Chitour 2006].

1.2 Problem statement

To validate the continuation method as a solution of the x_0-anchored MPP, we must show that the three obstructions are overcome. Doing so reduces to solving three subproblems, each of which corresponds to an obstruction. To reiterate, solving these subproblems is equivalent to demonstrating that

1. The x_0-anchored endpoint map $\mathrm{End}_{x_0}^{\Sigma}$ is C^2,
2. The set $\mathscr{U}_{x_0}^{\text{sing}}$ of singular controls is minimized, and
3. The solution of each path-lifting equation (1.1) is defined on $[0, 1]$.

As described above, each subproblem has a satisfying answer provided that Σ is an SBG driftless control-affine system. In the sense that this result applies to a class of control systems evolving on general manifolds, it is the only general result which validates the continuation method. More specialized results exist as well. Indeed, it is known that the three obstructions are at least partially overcome for a simple model of a front wheel-driven car [Chitour and Sussmann 1998], and for certain left-invariant control systems evolving on compact Lie groups [Chitour 2002]. In each of the cited works, the control systems under study are driftless control-affine systems.

The overall goal of this thesis is to study the validity of the continuation method outside of the realm of driftless control-affine systems. This investigation is motivated by two simple facts: First, the continuation method can be applied—at least in principle—to fully nonlinear systems. Second, many control systems encountered in practice are not driftless control-affine systems. Obviously, control-affine systems with drift are automatically excluded, as are control systems which are time-varying. Furthermore, even a very simple physical control system may not be a control-affine system, as evidenced by the dielectrophoretic control systems studied by Melnyk and Chang [2010]. With this in mind, our objective is to examine each of the above-mentioned subproblems individually, and under minimal prior assumptions about the nature of Σ.

1.3 Organization and contributions

The following is a brief chapter-by-chapter summary of this thesis, with an emphasis on the original contributions contained in each chapter.

Chapter 2: Preliminaries

In this chapter, we establish notation and review preliminary material. This material deals primarily with the basic theory of initial value problems.

Chapter 3: Control systems

In this chapter, we establish essential material concerning control systems. We begin by recalling the theory of C_p^q and C_q^q-polynomial control systems evolving on open subsets of Euclidean spaces. Then, we extend this theory to accommodate control systems evolving on finite-dimensional manifolds. The extended theory includes a high-order differentiability result which subsumes Bismut's theorem. This result can be used to overcome the first obstruction to the continuation method.

Chapter 4: The continuation method

In this chapter, we present the continuation method in full detail. By generalizing key features of the "classical" continuation method, we obtain a continuation method which does not rely fundamentally on Moore–Penrose pseudoinverses.

Chapter 5: Operations on time-varying vector fields

In this chapter, we review four operations on time-varying vector fields. These operations are the vertical lift, tangent lift, cotangent lift, and the pullback of one time-varying vector field by the global flow of another. In anticipation of Chapter 6, we derive a number of new identities. These identities provide reductive formulas for pullbacks involving lifts, an explicit formula for the global flow of $X + Y$, where X is a tangent lift and Y is a vertical lift, and explicit formulas for time derivatives and scalar parameter derivatives of pullbacks.

Chapter 6: Differentials of endpoint maps

In this chapter, we explicitly compute the differentials of the anchored endpoint maps of Σ. In fact, we compute the differentials of the so-called "endpoint maps" of Σ, which are functionally dependent on the initial state as well as on the control. In contrast with similar results derived by Vakhrameev [1991b], our computations do not rely on the chronological calculus formalism, and accommodate weakly regular, time-varying, fully nonlinear control systems. The analytical approach involves linearized lifts of Σ to the tangent bundle TM, together with the identities derived in Chapter 5.

Chapter 7: Intrinsic quadratic differentials of anchored endpoint maps

In this chapter, we explicitly compute the intrinsic quadratic differentials of the anchored endpoint maps of Σ. Roughly speaking, the results in this chapter carry the analysis of Chapter 6 to the second order. In particular, the analytical approach involves bilinearized lifts of Σ to the second tangent bundle TTM, together with the identities derived in Chapter 5.

Chapter 8: Constant-rank conditions

In this chapter, we study constant-rank conditions; that is, conditions which ensure that the anchored endpoint maps of Σ are constant-rank. After recalling a sufficient constant-rank condition derived by Vakhrameev [1991b], we describe how this condition can be used to overcome the second obstruction to the continuation method. Building on the work of Vakhrameev [1991b], we then derive a necessary and sufficient constant-rank condition. Mirroring the analysis in Chapters 6 and 7, the computations in this chapter do not rely on the chronological calculus formalism, and accommodate weakly regular, time-varying, fully nonlinear control systems.

Chapter 9: Sublinear growth

In this chapter, we study the third obstruction to the continuation method, under the standing assumption that Σ is a control-affine system. We present a sublinear growth condition, and explain how it can be used to overcome the third obstruction. Then, by extending a result of Chitour [1996], we show how the sublinear growth condition can be verified by examining the Lie bracket configuration and momentum functions of Σ.

Chapter 10: A necessary condition for unobstructed motion planning by the continuation method

In this chapter, we demonstrate the following fact: If Σ is a control system which is completely unobstructed with respect to the x_0-anchored MPP, then its x_0-anchored endpoint map is a fiber bundle over M. This result indicates that such unobstructed control systems are quite exceptional within the class of all control systems.

Chapter 11: Examples

In this chapter, we illustrate the theory developed in the preceding chapters. This is accomplished by applying the theory to several academic control-affine systems.

Chapter 12: Conclusions

In this chapter, we summarize the major contributions of this thesis and indicate several avenues for future work.

Chapter 2

Preliminaries

This chapter serves to recall preliminary material, including our notational conventions, essential preliminary results, and the basic theory of initial value problems.

2.1 Notation and preliminary results

In this section, we fix our notation, and quickly review definitions and results which will be used throughout this thesis.

2.1.1 Sets and maps

The symbols $\varnothing$, $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{R}$ have their standard meanings, being the empty set, the natural numbers, the integers, and the real numbers, respectively. We denote by $\mathbb{Z}_{\geq 0}$ the set of nonnegative integers. Similarly, we denote by $\mathbb{R}_{<0}$, $\mathbb{R}_{>0}$, and $\mathbb{R}_{\geq x}$ the sets of negative real numbers, positive real numbers, and real numbers greater than or equal to $x \in \mathbb{R}$, respectively. For convenience, we define

$$\mathbb{N}^* = \mathbb{N} \cup \{\infty\} \qquad \text{and} \qquad \mathbb{R}^*_{\geq 1} = \mathbb{R}_{\geq 1} \cup \{\infty\}.$$

By convention, $\infty - 1 = \infty + 1 = \infty$, $\frac{1}{\infty} = 0$, and $x \leq \infty$ for each $x \in \mathbb{R}$. Without further qualification, an ***interval*** is a nonempty connected subset of $\mathbb{R}$. Typically, we will interpret the elements of an interval as being time instants.

Given a set S, the identity map on S is denoted by $\mathrm{id}_S : S \to S$. The image of a map $f : S \to T$ is denoted by $\mathrm{image}(f)$, and the restriction of f to a subset $S' \subseteq S$ is denoted by $f|S'$. Given sets $S_1, \ldots, S_n$, projection onto the ith factor of the product $S_1 \times \cdots \times S_n$ is denoted by pr_i. If $S_1, \ldots, S_n$ are topological spaces, then each pr_i is continuous and open; the property of being open means that the image of any open subset of S_i under pr_i is open.

The term ***function*** is exclusively reserved for $\mathbb{R}$-valued maps. A ***curve*** in a topological space T is a continuous map $\gamma : I \to T$ whose domain is an interval. Finally, the terms "C^0" and "continuous" are synonymous.

2.1.2 Vector spaces, linear maps, and multilinear maps

All vector spaces considered in this thesis are real vector spaces. Given a vector space E, its origin is denoted by 0_E unless specified otherwise. For example, the origin of $\mathbb{R}$ is denoted by 0 instead of $0_{\mathbb{R}}$. The dimension of E is denoted by $\dim(E)$. Elements of $\mathbb{R}^n$ are written in boldface, as are $\mathbb{R}^n$-valued maps, whenever $n \geq 2$ or n is indeterminate. Depending on context, a vector $\boldsymbol{x} \in \mathbb{R}^n$ will be written as

$$\boldsymbol{x} = (x^1, \ldots, x^n) \qquad \text{or} \qquad \boldsymbol{x} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}.$$

Given a Banach space E, its norm is denoted by $\|\cdot\|$. When we wish to emphasize E, we will write $\|\cdot\|_E$ instead of $\|\cdot\|$. In particular, $\mathbb{R}^n$ will always have the Euclidean norm $\|\cdot\|_{\mathbb{R}^n}$. Given Banach spaces E and F, $\mathrm{Hom}(E, F)$ denotes the Banach space of all continuous linear maps $\lambda : E \to F$ with the ***operator norm***

$$\|\lambda\| = \sup\{\|\lambda \cdot e\|_F \ : \ \|e\|_E = 1\}.$$

Equivalently,

$$\|\lambda\| = \inf\{C \in \mathbb{R}_{>0} \ : \ \|\lambda \cdot e\|_F \leq C\|e\|_E \text{ for each } e \in E\}.$$

In particular, when $E = \mathbb{R}^n$ and $F = \mathbb{R}^m$, each $\lambda \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ can be identified with its matrix representation with respect to the standard bases. When $F = \mathbb{R}$, we obtain the continuous dual space of E, denoted by[1]

$$E^* = \mathrm{Hom}(E, \mathbb{R}).$$

The image of $e \in E$ under $\lambda \in \mathrm{Hom}(E, F)$ is denoted by $\lambda \cdot e$, and the kernel and image of λ are denoted by $\ker(\lambda)$ and $\mathrm{image}(\lambda)$, respectively. The cokernel of λ is denoted by $\mathrm{coker}(\lambda)$. By definition, $\mathrm{coker}(\lambda)$ is the quotient vector space

$$\mathrm{coker}(\lambda) = F \,/\, \overline{\mathrm{image}(\lambda)},$$

where $\overline{\mathrm{image}(\lambda)}$ denote the closure of $\mathrm{image}(\lambda)$ in F. Note that if F is finite-dimensional, then $\mathrm{image}(\lambda)$ is closed and $\mathrm{coker}(\lambda) = F \,/\, \mathrm{image}(\lambda)$. The rank of λ is denoted by $\mathrm{rank}(\lambda)$. By definition, $\mathrm{rank}(\lambda) = \dim(\mathrm{image}(\lambda))$.

Given Banach spaces $E_1, \ldots, E_k$, their direct sum is denoted by $E_1 \oplus \cdots \oplus E_k$. The latter space is a Banach space with the norm

$$\|(e_1, \ldots, e_k)\| = \sum_{i=1}^{k} \|e_i\|_{E_i}.$$

Given a Hilbert space E, its inner product is denoted by $\langle \cdot, \cdot \rangle$. When we wish to emphasize E, we will write $\langle \cdot, \cdot \rangle_E$ instead of $\langle \cdot, \cdot \rangle$. By the Riesz representation theorem, there is a canonical vector space isomorphism $E \cong E^*$. Given Hilbert spaces E and F, the adjoint of $\lambda \in \mathrm{Hom}(E, F)$ is denoted by λ^*. By definition,

$$\lambda^* \in \mathrm{Hom}(F^*, E^*) \cong \mathrm{Hom}(F, E)$$

[1]This notation does not conflict with $\mathbb{N}^*$ and $\mathbb{R}^*_{\geq 1}$, as $\mathbb{N}$ and $\mathbb{R}_{\geq 1}$ are not vector spaces.

satisfies

$$\langle \lambda \cdot e, f \rangle_F = \langle e, \lambda^* \cdot f \rangle_E$$

for each $e \in E$ and each $f \in F$. In particular, when $E = \mathbb{R}^n$ and $F = \mathbb{R}^m$, each $\lambda^* \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ can be identified with the transpose of the matrix representation of λ with respect to the standard bases.

To accommodate high-order derivatives, we require notation concerning continuous multilinear maps. Given Banach spaces $E_1, \ldots, E_k, F$,

$$\mathrm{Hom}(E_1, \ldots, E_k, F)$$

denotes the Banach space of all continuous k-multilinear maps

$$\lambda : E_1 \times \cdots \times E_k \to F$$

with the ***operator norm***

$$\|\lambda\| = \sup\{\|\lambda \cdot (e_1, \ldots, e_k)\|_F \ : \ \|e_1\|_{E_1} = \cdots = \|e_k\|_{E_k} = 1\}.$$

Equivalently,

$$\|\lambda\| = \inf\{C \in \mathbb{R}_{>0} \ : \ \|\lambda \cdot (e_1, \ldots, e_k)\|_F \leq C\|e_1\|_{E_1} \cdots \|e_k\|_{E_k}\}.$$

The image of $(e_1, \ldots, e_k)$ under $\lambda \in \mathrm{Hom}(E_1, \ldots, E_k, F)$ is denoted by

$$\lambda \cdot (e_1, \ldots, e_k).$$

The symmetric subspace of $\mathrm{Hom}(E_1, \ldots, E_k, F)$ is denoted by $\mathrm{Sym}(E_1, \ldots, E_k, F)$. Recall that $\lambda \in \mathrm{Sym}(E_1, \ldots, E_k, F)$ if and only if λ is invariant under all permutations of its arguments. In particular, when $E_1 = \cdots = E_k = E$, we write

$$\mathrm{Hom}^k(E, F) = \mathrm{Hom}(E, \ldots, E, F) \qquad \text{and}$$
$$\mathrm{Sym}^k(E, F) = \mathrm{Sym}(E, \ldots, E, F).$$

Finally, we use the following standard conventions:

$$\mathrm{Hom}^0(E,F) = \mathrm{Sym}^0(E,F) = F \qquad \text{and}$$
$$\mathrm{Hom}^1(E,F) = \mathrm{Sym}^1(E,F) = \mathrm{Hom}(E,F).$$

In particular, when $E = \mathbb{R}^n$ and $F = \mathbb{R}^m$, each $\lambda \in \mathrm{Hom}^k(\mathbb{R}^n, \mathbb{R}^m)$ can be identified with an element of $\mathbb{R}^{n^k m}$ via its tensor representation with respect to the standard bases. For further details, we refer to [Abraham et al. 1988, Chapter 2].

2.1.3 Measure and integration

We now recall basic facts related to measure and integration. This material will first come into play in Section 2.2, when we consider initial value problems whose right-hand sides are measurable in the time variable.

Recall that a measurable space $(S, \mathfrak{S})$ is comprised of a set S and a σ-algebra $\mathfrak{S}$ of subsets of S. When considered as a measure space, an interval I will always have the Lebesgue σ-algebra $\mathfrak{L}_I$. Similarly, a topological space T will always have the Borel σ-algebra $\mathfrak{B}_T$. Given two measurable spaces $(S, \mathfrak{S})$ and $(S', \mathfrak{S}')$, a map $f : S \to S'$ is measurable if and only if $A \in \mathfrak{S}'$ implies that $f^{-1}(A) \in \mathfrak{S}$. When we wish to emphasize the roles of the σ-algebras $\mathfrak{S}$ and $\mathfrak{S}'$, we will say that f is ***$(\mathfrak{S}, \mathfrak{S}')$-measurable***.

Suppose that I is an interval and T is a topological space. We denote by

$$\mathrm{Meas}(I, T)$$

the quotient space obtained from the set of all measurable maps of I into T, modulo the following equivalence relation: Two measurable maps $u, v : I \to T$ are declared to be equivalent if and only if $u(t) = v(t)$ for a.a. (almost all) $t \in I$. By a standard abuse of notation, we will identify an equivalence class $u \in \mathrm{Meas}(I, T)$ with any representative of u whenever it is convenient. Thus we refer to elements of $\mathrm{Meas}(I,T)$ as ***maps*** (or ***functions*** when $T \subseteq \mathbb{R}$). Operations on maps are extended to equivalence classes by working with representatives.

All integrals appearing in this thesis are Lebesgue integrals. Suppose that

$$f : I \to E,$$

where E is an n-dimensional vector space. One can show that f is measurable if and only if $\lambda \cdot f$ is measurable for each $\lambda \in E^*$. Equivalently, f is measurable if and only if its component functions $f^1, \dots, f^n : I \to \mathbb{R}$ with respect to some basis for E are measurable. By [Hunter 2010, Theorem 6.24], f is integrable if and only if

$$\int_I \|f(\sigma)\| \, \mathrm{d}\sigma < \infty,$$

where $\|\cdot\|$ is any choice of norm on E. In this case, we have

$$\left\| \int_I f(\sigma) \, \mathrm{d}\sigma \right\| \leq \int_I \|f(\sigma)\| \, \mathrm{d}\sigma.$$

If $t_0, t \in I$ and $t_0 \leq t$, we use the standard notation

$$\int_{t_0}^{t} f(\sigma) \, \mathrm{d}\sigma = \int_{[t_0,t]} f(\sigma) \, \mathrm{d}\sigma.$$

By convention, if $t_0 > t$, then

$$\int_{t_0}^{t} f(\sigma) \, \mathrm{d}\sigma = -\int_{t}^{t_0} f(\sigma) \, \mathrm{d}\sigma.$$

For further details, we refer to [Cohn 1980] and [Hunter 2010, Chapter 6.A].

2.1.4 Derivatives

We now recall basic facts concerning high-order total derivatives, partial derivatives, and mixed partial derivatives. This material will first come into play in Section 2.2, when we consider initial value problems whose right-hand sides are k times differentiable in the state variable. Also, this material will be used in Chapter 3 to isolate certain special control systems.

All derivatives appearing in this thesis are Fréchet derivatives.

2.1.4.1 Total derivatives

In this section, E, F are Banach spaces, U is a nonempty open subset of E, and

$$f : U \to F.$$

If f is C^k, where $k \in \mathbb{N}^*$, then the kth-order total derivative of f is denoted by $\boldsymbol{D}^k f$. By definition, $\boldsymbol{D}^k f : U \to \mathrm{Hom}^k(E, F)$ and one can show that $\boldsymbol{D}^k f$ actually takes values in $\mathrm{Sym}^k(E, F)$. We use the following standard conventions:

$$\boldsymbol{D}^0 f = f \qquad \text{and} \qquad \boldsymbol{D}^1 f = \boldsymbol{D} f.$$

In particular, if f is C^1 and $E = \mathbb{R}$, then each $\boldsymbol{D} f(e)$ can be identified with an element of F via the canonical vector space isomorphism $\iota : \mathrm{Hom}(\mathbb{R}, F) \to F$ that sends λ to $\iota(\lambda) = \lambda \cdot 1$. In this case, $\boldsymbol{D} f(e)$ is denoted equivalently by

$$\boldsymbol{D} f(e), \qquad \dot{f}(e), \qquad \text{and} \qquad \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_e f.$$

The notion of the total derivative can also be extended to the case where U is an interval of the form $[a, b]$, $[a, b)$, or $(a, b]$. For example, if $U = [a, b]$ and $e \in (a, b)$, then $\boldsymbol{D} f(e)$ is defined as above. If $e \in \{a, b\}$, then $\boldsymbol{D} f(e)$ is understood to be a one-sided derivative. The other cases $U = [a, b)$ and $U = (a, b]$ are analogous.

For products of maps, we have the following result: Suppose that $F_1, \ldots, F_n$ are Banach spaces and $f = f^1 \times \cdots \times f^n : U \to F_1 \oplus \cdots \oplus F_n$. Then f is C^k if and only if each of its component maps $f^i : U \to F_i$ is C^k. In this case,

$$\boldsymbol{D}^k f = \boldsymbol{D}^k f^1 \times \cdots \times \boldsymbol{D}^k f^n.$$

2.1.4.2 Partial derivatives

In this section, $T_1, \ldots, T_{i-1}, T_{i+1}, \ldots, T_n$ are nonempty sets, E_i and F are Banach spaces, U_i is a nonempty open subset of E_i, and

$$f : T_1 \times \cdots \times T_{i-1} \times U_i \times T_{i+1} \times \cdots \times T_n \to F.$$

Suppose that

$$\boldsymbol{t} = (t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n) \in T_1 \times \cdots \times T_{i-1} \times T_{i+1} \times \cdots \times T_n$$

is fixed and the partial map

$$e_i \mapsto P_{\boldsymbol{t}} = f(t_1, \ldots, t_{i-1}, e_i, t_{i+1}, \ldots, t_n)$$

of U_i into F is C^k, where $k \in \mathbb{N}^*$. We denote the kth-order total derivative of $P_{\boldsymbol{t}}$ by

$$\boldsymbol{D}_i^k f(t_1, \ldots, t_{i-1}, \cdot, t_{i+1}, \ldots, t_n) : U_i \to \mathrm{Hom}^k(E_i, F).$$

We use the following standard conventions:

$$\boldsymbol{D}_i^0 f(t_1, \ldots, t_{i-1}, \cdot, t_{i+1}, \ldots, t_n) = f(t_1, \ldots, t_{i-1}, \cdot, t_{i+1}, \ldots, t_n)$$

and

$$\boldsymbol{D}_i^1 f(t_1, \ldots, t_{i-1}, \cdot, t_{i+1}, \ldots, t_n) = \boldsymbol{D}_i f(t_1, \ldots, t_{i-1}, \cdot, t_{i+1}, \ldots, t_n).$$

If the partial map $P_{\boldsymbol{t}}$ is C^k for each $\boldsymbol{t}$, where $k \in \mathbb{N}^*$, then the kth-order partial derivative of f in its ith independent variable is the map

$$\boldsymbol{D}_i^k f : T_1 \times \cdots \times T_{i-1} \times U_i \times T_{i+1} \times \cdots \times T_n \to \mathrm{Hom}^k(E_i, F)$$

that sends $(t_1, \ldots, t_{i-1}, e_i, t_{i+1}, \ldots, t_n)$ to

$$\boldsymbol{D}_i^k f(t_1, \ldots, t_{i-1}, e_i, t_{i+1}, \ldots, t_n).$$

The next proposition relates the total and partial derivatives of f.

Proposition 2.1.1. *Suppose that $E_1, \ldots, E_n$ are Banach spaces, $U = U_1 \times \cdots \times U_n$ is a nonempty product open subset of $E_1 \oplus \cdots \oplus E_n$, and $f : U \to F$. Then f is C^k, where $k \in \mathbb{N}^*$, if and only if $\boldsymbol{D}_i f$ exists and is C^{k-1} for each $1 \le i \le n$. Furthermore, if f is C^1, then*

$$\boldsymbol{D} f(e) \cdot (e_1, \ldots, e_n) = \sum_{i=1}^{n} \boldsymbol{D}_i f(e) \cdot e_i$$

for each $e \in U$ and each $(e_1, \ldots, e_n) \in E_1 \oplus \cdots \oplus E_n$.

For further details, we refer to [Abraham et al. 1988, Chapter 2].

2.1.4.3 Mixed partial derivatives

Proposition 2.1.1 implies that partial derivatives can be iterated. Here, we describe a particular situation which will arise in this thesis, particularly in Chapter 3. In this section, I is an interval, E_1, E_2, F are Banach spaces, $U_1 \times U_2$ is a nonempty product open subset of $E_1 \oplus E_2$, and

$$f : I \times U_1 \times U_2 \to F.$$

Suppose that $t \in I$ is fixed and the partial map

$$(e_1, e_2) \mapsto P_t = f(t, e_1, e_2)$$

of $U_1 \times U_2$ into F is C^k, where $k \in \mathbb{N}^*$. Furthermore, suppose that $1 \leq i, j \leq k$ are such that $i + j \leq k$. By Proposition 2.1.1, $\boldsymbol{D}_2 P_t$ exists and is C^{k-1}. Iterating,

$$\boldsymbol{D}_2^j P_t : U_1 \times U_2 \to \mathrm{Hom}^j(E_2, F)$$

exists and is C^{k-j}. Similarly, $\boldsymbol{D}_1 \boldsymbol{D}_2^j P_t$ exists and is $C^{k-(1+j)}$. Iterating,

$$\boldsymbol{D}_1^i \boldsymbol{D}_2^j P_t : U_1 \times U_2 \to \mathrm{Hom}^i(E_1, \mathrm{Hom}^j(E_2, F))$$

exists and is $C^{k-(i+j)}$. By construction,

$$\boldsymbol{D}_1^i \boldsymbol{D}_2^j P_t = \boldsymbol{D}_2^i \boldsymbol{D}_3^j f(t, \cdot, \cdot).$$

In the remainder of this thesis, we will use the notation appearing on the right-hand side of the above equation. To maintain notational consistency with the literature, particularly with the notation used by Margheri [1996], we write

$$\begin{aligned}
&\boldsymbol{D}_2^i \boldsymbol{D}_3^j f(t, e_1, e_2) \cdot (e_{1,1}, \ldots, e_{1,i}, e_{2,1}, \ldots, e_{2,j}) \\
&= (\boldsymbol{D}_2^i \boldsymbol{D}_3^j f(t, e_1, e_2) \cdot (e_{1,1}, \ldots, e_{1,i})) \cdot (e_{2,1}, \ldots, e_{2,j})
\end{aligned}$$

for each $(e_1, e_2) \in U_1 \times U_2$, each $(e_{1,1}, \ldots, e_{1,i}) \in E_1^i$, and each $(e_{2,1}, \ldots, e_{2,j}) \in E_2^j$.

For further details, we refer to [Abraham et al. 1988, Chapter 2].

2.1.5 The chain rule and Leibniz rule

The chain rule and Leibniz rule require some additional terminology, which we now establish. Our presentation of this material follows [Pötzsche 2010]. If S is a set, then its power set (that is, the set of all subsets of S) is denoted by 2^S and its cardinality is denoted by $\mathrm{card}(S)$.

Definition 2.1.2. Suppose that $j, k \in \mathbb{N}$. A j***-partition*** of $\{1, \ldots, k\}$ is a j-tuple

$$(N_1, \ldots, N_j), \qquad N_i \subseteq \{1, \ldots, k\}, \qquad \mathrm{card}(N_i) \geq 0,$$

whose components form a partition of $\{1, \ldots, k\}$. We denote the set of all j-partitions of $\{1, \ldots, k\}$ by $P_j(k)$. For each $(N_1, \ldots, N_j) \in P_j(k)$, we write

$$N_i = \{n_i^1, \ldots, n_i^{\mathrm{card}(N_i)}\}.$$

An ***ordered*** j***-partition*** of $\{1, \ldots, k\}$ is a j-tuple $(N_1, \ldots, N_j) \in P_j(k)$ such that

- $\mathrm{card}(N_i) \geq 1$ for each $1 \leq i \leq j$ and
- $\max(N_i) < \max(N_{i+1})$ for each $1 \leq i \leq j - 1$.

We denote the set of all ordered j-partitions of $\{1, \ldots, k\}$ by $P_j^{\mathrm{ord}}(k)$.

Remark 2.1.3. By definition, the components N_i of a j-partition of $\{1, \ldots, k\}$ may be empty. This facilitates the expression of the Leibniz rule, as we will see below.

The next example illustrates these definitions.

Example 2.1.4. The set of all 2-partitions of $\{1\}$ is

$$P_2(1) = \{(\varnothing, \{1\}), (\{1\}, \varnothing)\}.$$

The set of all 2-partitions of $\{1, 2\}$ is

$$P_2(2) = \{(\varnothing, \{1, 2\}), (\{1\}, \{2\}), (\{2\}, \{1\}), (\{1, 2\}, \varnothing)\}.$$

The set of all ordered 1-partitions of $\{1,2\}$ is

$$P_1^{\mathrm{ord}}(2) = \{(\{1,2\})\}.$$

Finally, the set of all ordered 2-partitions of $\{1,2\}$ is

$$P_2^{\mathrm{ord}}(2) = \{(\{1\},\{2\})\}.$$

Theorem 2.1.5. *(**Chain rule**) Suppose that E, F, G are Banach spaces, U is a nonempty open subset of E, V is a nonempty open subset of F, $f : U \to F$ and $g : V \to G$ are C^k, where $k \in \mathbb{N}^*$, and* $\mathrm{image}(f) \subseteq V$. *Then $g \circ f : U \to G$ is C^k and*

$$\begin{aligned}
&\boldsymbol{D}^k(g\circ f)(e)\cdot(e_1,\ldots,e_k)\\
&= \sum_{j=1}^{k} \sum_{(N_1,\ldots,N_j)\in P_j^{\mathrm{ord}}(k)} \boldsymbol{D}^j g(f(e))\cdot\left(\boldsymbol{D}^{\mathrm{card}(N_1)} f(e)\cdot e_{N_1},\ldots,\boldsymbol{D}^{\mathrm{card}(N_j)} f(e)\cdot e_{N_j}\right),
\end{aligned}$$

where for each N_i, e_{N_i} denotes the $\mathrm{card}(N_i)$*-tuple*

$$e_{N_i} = \left(e_{n_i^1},\ldots,e_{n_i^{\mathrm{card}(N_i)}}\right).$$

Proof. See [Rybakowski 1991]. ∎

Example 2.1.6. If f, g are C^2, then

$$\begin{aligned}
&\boldsymbol{D}^2(g\circ f)(e)\cdot(e_1,e_2)\\
&= \boldsymbol{D}g(f(e))\cdot\left(\boldsymbol{D}^2 f(e)\cdot(e_1,e_2)\right) + \boldsymbol{D}^2 g(f(e))\cdot(\boldsymbol{D}f(e))\cdot e_1, \boldsymbol{D}f(e)\cdot e_2).
\end{aligned}$$

This is easily seen, using the expressions from Example 2.1.4.

Theorem 2.1.7. *(**Leibniz rule**) Suppose that $E, E_1, \ldots, E_n, F$ are Banach spaces, U is a nonempty open subset of E, $f_i : U \to E_i$ is C^k for each $1 \le i \le n$, where $k \in \mathbb{N}^*$, and $\lambda \in$* $\mathrm{Hom}(E_1,\ldots,E_n,F)$. *Then $\Lambda = \lambda\circ(f_1\times\cdots\times f_n) : U \to F$ is C^k and*

$$\boldsymbol{D}^k\Lambda(e)\cdot(e_1,\ldots,e_k)$$

$$= \sum_{(N_1,\dots,N_n)\in P_n(k)} \lambda \cdot (\boldsymbol{D}^{\mathrm{card}(N_1)} f_1(e) \cdot e_{N_1}, \dots, \boldsymbol{D}^{\mathrm{card}(N_n)} f_n(e) \cdot e_{N_n}),$$

where for each N_i, e_{N_i} denotes the $\mathrm{card}(N_i)$*-tuple*

$$e_{N_i} = \left(e_{n_i^1}, \dots, e_{n_i^{\mathrm{card}(N_i)}}\right).$$

Proof. See [Abraham et al. 1988, pages 95–96 and Exercise 2.4C] ∎

Example 2.1.8. Suppose that $n = 2$ and f_1, f_2 are C^1. Then

$$\boldsymbol{D}\Lambda(e) \cdot e_1 = \lambda \cdot (\boldsymbol{D}f_1(e) \cdot e_1, f_2(e)) + \lambda \cdot (f_1(e), \boldsymbol{D}f_2(e) \cdot e_1).$$

On the other hand, suppose that f_1, f_2 are C^2. Then

$$\begin{aligned}
&\boldsymbol{D}^2\Lambda(e) \cdot (e_1, e_2) \\
&= \lambda \cdot (f_1(e), \boldsymbol{D}^2 f_2(e) \cdot (e_1, e_2)) + \lambda \cdot (\boldsymbol{D}f_1(e) \cdot e_1, \boldsymbol{D}f_2(e) \cdot e_2) + \\
&\quad \lambda \cdot (\boldsymbol{D}f_1(e) \cdot e_2, \boldsymbol{D}f_2(e) \cdot e_1) + \lambda \cdot (\boldsymbol{D}^2 f_1(e) \cdot (e_1, e_2), f_2(e)).
\end{aligned}$$

Again, this is easily seen, using the expressions from Example 2.1.4.

For further details, we refer to [Pötzsche 2010].

2.1.6 L^p and locally L^p spaces

In this section, we recall basic facts about L^p spaces. Throughout this thesis, these spaces will serve as control spaces—an instance of this appeared in Chapter 1, where we chose an L^2 space of controls. Throughout this section, I is an interval, E is a finite-dimensional vector space, and $\|\cdot\|_E$ is any choice of norm on E.

2.1.6.1 L^p spaces

By definition, $u \in L^p(I, E)$, where $p \in \mathbb{R}_{\geq 1}$, if and only if $u \in \mathrm{Meas}(I, E)$ and

$$\int_I \|u(\sigma)\|_E^p \, \mathrm{d}\sigma < \infty.$$

The vector space $L^p(I, E)$ is a Banach space with the p-norm

$$\|u\|_p = \left(\int_I \|u(\sigma)\|_E^p \, \mathrm{d}\sigma\right)^{1/p}.$$

By definition, $u \in L^\infty(I, E)$ if and only if $u \in \mathrm{Meas}(I, E)$ and there exists a compact subset K of E such that $u(t) \in K$ for a.a. $t \in I$. The vector space $L^\infty(I, E)$ is a Banach space with the ∞-norm

$$\|u\|_\infty = \inf\{C \in \mathbb{R}_{>0} \;:\; \|u(t)\|_E < C \text{ for a.a. } t \in I\}.$$

Provided that I is compact, there is a chain of inclusions

$$L^\infty(I, E) \subseteq L^q(I, E) \subseteq L^p(I, E) \subseteq L^1(I, E), \qquad q \geq p. \tag{2.1}$$

More generally, (2.1) holds when the measure of I is finite.

For further details, we refer to [Cohn 1980, Chapter 3].

2.1.6.2 Locally L^p spaces

By definition, $u \in L^p_{\mathrm{loc}}(I, E)$, where $p \in \mathbb{R}^*_{\geq 1}$, if and only if $u \in \mathrm{Meas}(I, E)$ and

$$u|K \in L^p(K, E)$$

for each compact subinterval $K \subseteq I$. We say that elements of $L^1_{\mathrm{loc}}(I, E)$ are ***locally integrable***. Clearly, if I is compact, then there is no distinction between $L^p_{\mathrm{loc}}(I, E)$ and $L^p(I, E)$. From (2.1), we obtain the chain of inclusions

$$L^\infty_{\mathrm{loc}}(I, E) \subseteq L^q_{\mathrm{loc}}(I, E) \subseteq L^p_{\mathrm{loc}}(I, E) \subseteq L^1_{\mathrm{loc}}(I, E), \qquad q \geq p. \tag{2.2}$$

To handle nonnegative functions, we define

$$L^p_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0}) = \{u \in L^p_{\mathrm{loc}}(I, \mathbb{R}) \;:\; u(t) \geq 0 \text{ for a.a. } t \in I\}.$$

Given a compact subinterval $K \subseteq I$, the projection map

$$\pi^p_K : L^p_{\mathrm{loc}}(I, E) \to L^p(K, E)$$

is defined by

$$\pi_K^p(u) = u|K.$$

One can show that π_K^p is a surjective continuous linear map, where $L_{\text{loc}}^p(I, E)$ has its natural topology as a Fréchet space, and thus π_K^p is an open map.

For further details, we refer to [Trèves 1966].

2.1.6.3 Hölder's inequality for k-fold products

The next theorem is Hölder's inequality for k-fold products of maps. Recall from Section 2.1.1 that $\frac{1}{\infty} = 0$ as a matter of convention.

Theorem 2.1.9. ***(Hölder's inequality)*** *Suppose that* $p_1, \ldots, p_k, q \in \mathbb{R}_{\geq 1}^*$,

$$\sum_{j=1}^{k} \frac{1}{p_j} = \frac{1}{q},$$

and $u_j \in L_{\text{loc}}^{p_j}(I, E)$ *for each* $1 \leq j \leq k$. *Then* $u_1 \cdots u_k \in L_{\text{loc}}^q(I, E)$ *and*

$$\|(u_1 \cdots u_k)|K\|_q \leq \|u_1|K\|_{p_1} \cdots \|u_k|K\|_{p_k}.$$

for each compact subinterval $K \subseteq I$.

Proof. See [Bogachev 2007, Chapter 2]. ∎

Note that Hölder's inequality reduces to the well-known Cauchy–Schwarz inequality when $k = 2$, $p_1 = 2$, $p_2 = 2$, and $q = 1$. In this case, we have

$$\int_K \|u_1(\sigma)u_2(\sigma)\|_E \, \mathrm{d}\sigma \leq \left(\int_K \|u_1(\sigma)\|_E^2 \, \mathrm{d}\sigma\right)^{1/2} \left(\int_K \|u_2(\sigma)\|_E^2 \, \mathrm{d}\sigma\right)^{1/2}.$$

2.1.7 Manifolds

As described in Chapter 1, the objects of primary interest in this thesis are control systems evolving on finite-dimensional manifolds. We also have a primary interest in maps between manifolds modelled on Banach spaces (which may be infinte-dimensional). For example, the x_0-anchored endpoint map of a control system, as described in Chapter 1, is such a map. To treat the finite- and infinite-dimensional cases simultaneously, we use the language of Banach manifolds. All Banach manifolds considered in this thesis are nonempty, real, positive-dimensional, Hausdorff, and C^∞, unless specified otherwise. For example, at several junctures we will strengthen these hypotheses by assuming that a given Banach manifold is second-countable or C^ω (that is, real-analytic). Note that these conventions apply to vector bundles as well, since vector bundles are special instances of Banach manifolds.

Suppose that Q is a Banach manifold modelled on a Banach space E_Q. The dimension of Q is denoted by $\dim(Q)$; by definition, $\dim(Q) = \dim(E_Q)$. The tangent and cotangent bundles of Q, which are themselves Banach manifolds modelled on the direct sum $E_Q \oplus E_Q$, are denoted by

$$\pi_{TQ} : TQ \to Q \qquad \text{and} \qquad \pi_{T^*Q} : T^*Q \to Q,$$

respectively. The tangent space to Q at $q \in Q$ (that is, the fiber of TQ over q) is denoted by T_qQ. A generic element of T_qQ is written v_q. For notational economy, the origin of T_qQ is denoted by 0_q instead of 0_{T_qQ}. If Q is an open submanifold of E_Q, then each tangent space T_qQ is canonically isomorphic to E_Q; we will use this fact implicitly throughout this thesis. The cotangent space to Q at $q \in Q$ (that is, the fiber of T^*Q over q) is denoted by T_q^*Q. By construction, $T_q^*Q = (T_qQ)^*$. A generic element of T_q^*Q is written p_q. The second tangent bundle of Q, which is itself a Banach manifold modelled on $E_Q \oplus E_Q \oplus E_Q \oplus E_Q$, is denoted by

$$\pi_{TTQ} : TTQ \to TQ.$$

The canonical involution of TTQ is denoted by $s_Q : TTQ \to TTQ$.

Any chart (U, φ) on Q induces natural charts on the bundles TQ, T^*Q, and TTQ, denoted by $(TU, T\varphi)$, $(T^*U, T^*\varphi)$, and $(TTU, TT\varphi)$, respectively. In this way, any compatible atlas $\mathscr{A}_Q$ on Q induces natural atlases on the bundles TQ, T^*Q, and TTQ. These are denoted by $T\mathscr{A}_Q$, $T^*\mathscr{A}_Q$, and $TT\mathscr{A}_Q$, respectively.

Suppose now that $F : Q \to R$, where Q and R are Banach manifolds. If F is C^k, where $k \in \mathbb{N}^*$, then its differential $TF : TQ \to TR$ is C^{k-1}. Given $q \in Q$, the restriction of TF to the tangent space T_qQ is denoted by $TF(q)$. By definition,

$$TF(q) \in \mathrm{Hom}(T_qQ, T_{F(q)}R).$$

When it is notationally unwieldy to write $TF(q)$, we simply write TF. If (U, φ) and (V, ψ) are charts on Q and R, respectively, such that $F(U) \subseteq V$, then we say that (U, φ) and (V, ψ) are F***-compatible***. For such charts,

$$F_{\psi,\varphi} = \psi \circ F \circ \varphi^{-1}$$

denotes the ***local representative*** of F in (U, φ) and (V, ψ). For example, the local representative of the canonical involution s_Q in $(TTU, TT\varphi)$ and $(TTU, TT\varphi)$ is

$$(s_Q)_{TT\varphi,TT\varphi}(\boldsymbol{q}, \boldsymbol{v}, \boldsymbol{Q}, \boldsymbol{V}) = (\boldsymbol{q}, \boldsymbol{Q}, \boldsymbol{v}, \boldsymbol{V}).$$

Finally, given a Riemannian manifold $Q = (Q, g)$, each tangent space T_qQ is a Hilbert space with the inner product induced by g. Since T_qQ is canonically isomorphic to T_q^*Q, each T_q^*Q is a Hilbert space. The natural pairing of $v_q \in T_qQ$ and $p_q \in T_q^*Q$ is denoted by $\langle p_q, v_q \rangle = p_q \cdot v_q$.

For further details, we refer to [Abraham et al. 1988, Chapters 3 and 5].

2.2 Initial value problems evolving on open subsets of Euclidean spaces

As we will see in Chapter 3, each of the controlled trajectories of a control system is the solution of an initial value problem. In this section, we review fundamental results from the theory of initial value problems evolving on open subsets of Euclidean spaces. These results concern the nature of solutions, their existence and uniqueness, and the functional dependence of solutions on initial conditions. Our presentation follows Sontag [1998, Appendix C]. In Section 2.3, this material is extended to accommodate initial value problems evolving on finite-dimensional manifolds.

Our standing assumptions throughout this section are that

- I is an interval,
- V is a nonempty open subset of $\mathbb{R}^n$, and
- $\boldsymbol{f} : I \times V \to \mathbb{R}^n$.

Throughout this section, $\boldsymbol{\xi} : \mathrm{dom}(\boldsymbol{\xi}) \to V$ is a curve.

Here we consider initial value problems of the form

$$\begin{cases} \dot{\boldsymbol{\xi}}(t) = \boldsymbol{f}(t, \boldsymbol{\xi}(t)), \quad \boldsymbol{\xi}(t) \in V, \quad t \in I \\ \boldsymbol{\xi}(t_0) = \boldsymbol{x}_0, \end{cases} \tag{2.3}$$

where $(t_0, \boldsymbol{x}_0) \in I \times V$. At the outset, it is reasonable to say that $\boldsymbol{\xi}$ is a solution of (2.3) if and only if $\mathrm{dom}(\boldsymbol{\xi})$ is a subinterval of I containing t_0, the map

$$t \mapsto \boldsymbol{f}(t, \boldsymbol{\xi}(t))$$

is an element of $L^1_{\mathrm{loc}}(\mathrm{dom}(\boldsymbol{\xi}), \mathbb{R}^n)$ and

$$\boldsymbol{\xi}(t) = \boldsymbol{x}_0 + \int_{t_0}^{t} \boldsymbol{f}(\sigma, \boldsymbol{\xi}(\sigma)) \, \mathrm{d}\sigma \tag{2.4}$$

for each $t \in \mathrm{dom}(\boldsymbol{\xi})$. The fundamental theorem of calculus, which we now review, equates this property to local absolute continuity and to differentiability.

Definition 2.2.1. Suppose that $\mathrm{dom}(\boldsymbol{\xi})$ is compact. We say that $\boldsymbol{\xi}$ is ***absolutely continuous*** (AC) if for each $\varepsilon \in \mathbb{R}_{>0}$ there exists $\delta \in \mathbb{R}_{>0}$ such that

$$\sum_{i=1}^{N} \|\boldsymbol{\xi}(d_i) - \boldsymbol{\xi}(c_i)\|_{\mathbb{R}^n} < \varepsilon$$

whenever $N \in \mathbb{N}$ and $[c_i, d_i]$, $1 \leq i \leq n$, are disjoint subintervals of $\mathrm{dom}(\boldsymbol{\xi})$ with

$$\sum_{i=1}^{N}(d_i - c_i) < \delta.$$

On the other hand, suppose that $\mathrm{dom}(\boldsymbol{\xi})$ is not compact. We say that $\boldsymbol{\xi}$ is ***locally absolutely continuous*** (LAC) if $\boldsymbol{\xi}|K$ is AC for each compact subinterval K of $\mathrm{dom}(\boldsymbol{\xi})$. Clearly, if $\mathrm{dom}(\boldsymbol{\xi})$ is compact, then there is no distinction between $\boldsymbol{\xi}$ being AC and $\boldsymbol{\xi}$ being LAC.

The LAC property is stable under composition with locally Lipschitz maps.

Lemma 2.2.2. *Suppose that $\boldsymbol{F} : V \to \mathbb{R}^k$ is locally Lipschitz and $\boldsymbol{\xi}$ is LAC. Then $\boldsymbol{F} \circ \boldsymbol{\xi} : \mathrm{dom}(\boldsymbol{\xi}) \to \mathbb{R}^k$ is LAC.*

Proof. Choose a compact subinterval $K \subseteq \mathrm{dom}(\boldsymbol{\xi})$. We must show that $\boldsymbol{F} \circ \boldsymbol{\xi}|K$ is AC. To this end, observe that $\boldsymbol{\xi}(K)$ is compact since $\boldsymbol{\xi}$ is continuous. Since $\boldsymbol{F}$ is locally Lipschitz, $\boldsymbol{F}|\boldsymbol{\xi}(K)$ is Lipschitz. That is, there exists $C \in \mathbb{R}_{\geq 0}$ such that

$$\|\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{y})\|_{\mathbb{R}^k} \leq C\|\boldsymbol{x} - \boldsymbol{y}\|_{\mathbb{R}^n}$$

for each $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\xi}(K)$. Without loss of generality, we may assume that $C \neq 0$. Since $\boldsymbol{\xi}|K$ is AC, for each $\varepsilon \in \mathbb{R}_{>0}$ there exists $\delta \in \mathbb{R}_{>0}$ such that

$$\sum_{i=1}^{N} \|\boldsymbol{\xi}(d_i) - \boldsymbol{\xi}(c_i)\|_{\mathbb{R}^n} < \frac{\varepsilon}{C}$$

whenever $N \in \mathbb{N}$ and $[c_i, d_i]$, $1 \leq i \leq N$ are disjoint subintervals of K satisfying

$$\sum_{i=1}^{N}(d_i - c_i) < \delta.$$

Thus $\boldsymbol{F} \circ \boldsymbol{\xi}|K$ is AC, since for each $\varepsilon \in \mathbb{R}_{>0}$ there exists $\delta \in \mathbb{R}_{>0}$ such that

$$\sum_{i=1}^{N} \|\boldsymbol{F} \circ \boldsymbol{\xi}(d_i) - \boldsymbol{F} \circ \boldsymbol{\xi}(c_i)\|_{\mathbb{R}^k} \leq C \sum_{i=1}^{N} \|\boldsymbol{\xi}(d_i) - \boldsymbol{\xi}(c_i)\|_{\mathbb{R}^n} < C\,\frac{\varepsilon}{C} = \varepsilon$$

whenever $N \in \mathbb{N}$ and $[c_i, d_i]$, $1 \leq i \leq N$ are disjoint subintervals of K satisfying

$$\sum_{i=1}^{N}(d_i - c_i) < \delta.$$

We conclude that $\boldsymbol{F} \circ \boldsymbol{\xi}$ is LAC. ∎

The next theorem is the fundamental theorem of calculus.

Theorem 2.2.3. *The following properties are equivalent:*

1. $\boldsymbol{\xi}$ *is LAC;*

2. $\dot{\boldsymbol{\xi}}(t)$ *exists for a.a.* $t \in \mathrm{dom}(\boldsymbol{\xi})$, $t \mapsto \dot{\boldsymbol{\xi}}(t) \in L^1_{\mathrm{loc}}(\mathrm{dom}(\boldsymbol{\xi}), \mathbb{R}^n)$, *and*

$$\boldsymbol{\xi}(t) = \boldsymbol{\xi}(t_0) + \int_{t_0}^{t} \dot{\boldsymbol{\xi}}(\sigma)\,\mathrm{d}\sigma$$

   for each $t, t_0 \in \mathrm{dom}(\boldsymbol{\xi})$*;*

3. *There exists* $\boldsymbol{P} \in L^1_{\mathrm{loc}}(\mathrm{dom}(\boldsymbol{\xi}), \mathbb{R}^n)$ *such that*

$$\boldsymbol{\xi}(t) = \boldsymbol{\xi}(t_0) + \int_{t_0}^{t} \boldsymbol{P}(\sigma)\,\mathrm{d}\sigma \tag{2.5}$$

   for each $t, t_0 \in \mathrm{dom}(\boldsymbol{\xi})$*.*

Furthermore, if $\boldsymbol{\xi}$ *is LAC, then* $\boldsymbol{P}$ *is uniquely determined by the fact that* $\dot{\boldsymbol{\xi}}(t) = \boldsymbol{P}(t)$ *for a.a.* $t \in \mathrm{dom}(\boldsymbol{\xi})$*.*

Proof. See [Leoni 2009, Chapter 3]. ∎

We now formally define initial value problems and their solutions.

Definition 2.2.4. Suppose that $(t_0, \boldsymbol{x}_0) \in I \times V$. The triple $(\boldsymbol{f}, t_0, \boldsymbol{x}_0)$ is said to be an ***initial value problem*** evolving on V, with ***right-hand side*** $\boldsymbol{f}$ and ***initial condition*** $(t_0, \boldsymbol{x}_0)$. We say that $\boldsymbol{\xi}$ is a ***solution*** of $(\boldsymbol{f}, t_0, \boldsymbol{x}_0)$ if

- $\mathrm{dom}(\boldsymbol{\xi})$ is a relatively open subinterval[2] of I containing t_0,
- The map $t \mapsto \boldsymbol{f}(t, \boldsymbol{\xi}(t))$ is an element of $L^1_{\mathrm{loc}}(\mathrm{dom}(\boldsymbol{\xi}), \mathbb{R}^n)$, and
- $\boldsymbol{\xi}(t) = \boldsymbol{x}_0 + \int_{t_0}^{t} \boldsymbol{f}(\sigma, \boldsymbol{\xi}(\sigma))\,\mathrm{d}\sigma$ for each $t \in \mathrm{dom}(\boldsymbol{\xi})$.

The next lemma characterizes solutions in an alternative way.

[2]That is, $\mathrm{dom}(\boldsymbol{\xi})$ can be written as the intersection of I with an open interval.

Lemma 2.2.5. *The curve* $\boldsymbol{\xi}$ *is a solution of* $(\boldsymbol{f}, t_0, \boldsymbol{x}_0)$ *if and only if*

1. $\mathrm{dom}(\boldsymbol{\xi})$ *is a relatively open subinterval of* I *containing* t_0*,*
2. $\boldsymbol{\xi}$ *is LAC,*
3. $\boldsymbol{\xi}(t_0) = \boldsymbol{x}_0$*, and* $\dot{\boldsymbol{\xi}}(t) = \boldsymbol{f}(t, \boldsymbol{\xi}(t))$ *for a.a.* $t \in \mathrm{dom}(\boldsymbol{\xi})$*.*

Proof. This follows immediately from Theorem 2.2.3. ∎

Definition 2.2.6. Suppose that $(t_0, \boldsymbol{x}_0) \in I \times V$ and $\boldsymbol{\xi}$ is a solution of $(\boldsymbol{f}, t_0, \boldsymbol{x}_0)$. We say that $\boldsymbol{\xi}$ is ***maximally-defined*** if it has the following property: If

$$\tilde{\boldsymbol{\xi}} : \mathrm{dom}(\tilde{\boldsymbol{\xi}}) \to V$$

is another solution of $(\boldsymbol{f}, t_0, \boldsymbol{x}_0)$, then

$$\mathrm{dom}(\tilde{\boldsymbol{\xi}}) \subseteq \mathrm{dom}(\boldsymbol{\xi}) \qquad \text{and} \qquad \tilde{\boldsymbol{\xi}}(t) = \boldsymbol{\xi}(t)$$

for each $t \in \mathrm{dom}(\tilde{\boldsymbol{\xi}})$. Clearly, such a solution is unique.

Definition 2.2.7. We say that $\boldsymbol{f}$ is ***solvable*** if there exists a maximally-defined solution of $(\boldsymbol{f}, t_0, \boldsymbol{x}_0)$ for each $(t_0, \boldsymbol{x}_0) \in I \times V$. If $\boldsymbol{f}$ is solvable, then the maximally-defined solution of $(\boldsymbol{f}, t_0, \boldsymbol{x}_0)$ is denoted by

$$\mu^{\boldsymbol{f}}(\cdot, t_0, \boldsymbol{x}_0) : I^{\boldsymbol{f}}(t_0, \boldsymbol{x}_0) \to V.$$

Two basic properties of maximally-defined solutions are given next.

Proposition 2.2.8. *Suppose that* $\boldsymbol{f}$ *is solvable and* $(t_0, \boldsymbol{x}_0) \in I \times V$*. Then*

1. *For each* $t \in I^{\boldsymbol{f}}(t_0, \boldsymbol{x}_0)$*, we have* $I^{\boldsymbol{f}}(t_0, \boldsymbol{x}_0) = I^{\boldsymbol{f}}(t, \mu^{\boldsymbol{f}}(t, t_0, \boldsymbol{x}_0))$*, and*
2. *For each* $t, s \in I^{\boldsymbol{f}}(t_0, \boldsymbol{x}_0)$*, we have* $\mu^{\boldsymbol{f}}(t, t_0, \boldsymbol{x}_0) = \mu^{\boldsymbol{f}}(t, s, \mu^{\boldsymbol{f}}(s, t_0, \boldsymbol{x}_0))$*.*

To study the functional dependence of maximally-defined solutions on initial conditions, we aggregate these solutions into a single map.

Definition 2.2.9. Suppose that $\boldsymbol{f}$ is solvable. Define

$$\operatorname{dom}(\Phi^{\boldsymbol{f}}) = \{(t, t_0, \boldsymbol{x}_0) \in I \times I \times V \ : \ t \in I^{\boldsymbol{f}}(t_0, \boldsymbol{x}_0)\}.$$

The ***global flow*** of $\boldsymbol{f}$ is the map $\Phi^{\boldsymbol{f}} : \operatorname{dom}(\Phi^{\boldsymbol{f}}) \to V$ that sends $(t, t_0, \boldsymbol{x}_0)$ to

$$\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{x}_0) = \mu^{\boldsymbol{f}}(t, t_0, \boldsymbol{x}_0).$$

We must also consider the global flow of $\boldsymbol{f}$ with its first two independent variables fixed. For each $(t, t_0) \in I \times I$, define

$$\operatorname{dom}(\Phi^{\boldsymbol{f}}_{t,t_0}) = \{\boldsymbol{x}_0 \in V \ : \ (t, t_0, \boldsymbol{x}_0) \in \operatorname{dom}(\Phi^{\boldsymbol{f}})\}.$$

By a mild abuse of notation, we denote by

$$\Phi^{\boldsymbol{f}}_{t,t_0} : \operatorname{dom}(\Phi^{\boldsymbol{f}}_{t,t_0}) \to V$$

the map that sends $\boldsymbol{x}_0$ to $\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{x}_0)$.

Remark 2.2.10. Suppose that $\boldsymbol{f}$ is solvable and $(t_0, \boldsymbol{x}_0) \in I \times V$. Then for each $t, s \in I^{\boldsymbol{f}}(t_0, \boldsymbol{x}_0)$, we have

$$\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{x}_0) = \Phi^{\boldsymbol{f}}_{t,s} \circ \Phi^{\boldsymbol{f}}_{s,t_0}(\boldsymbol{x}_0).$$

This follows immediately from Proposition 2.2.8.

We now turn to verifying that $\boldsymbol{f}$ is solvable.

Definition 2.2.11. Suppose that E is a normed vector space and

$$g : I \times V \to E.$$

We say that g is ***locally integrably bounded*** if for each compact subset $K \subseteq V$, there exists $\alpha \in L^1_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0})$ such that

$$\|g(t, \boldsymbol{x})\|_E \leq \alpha(t)$$

for a.a. $t \in I$ and each $\boldsymbol{x} \in K$. We say that g is ***locally integrably Lipschitz*** if for each compact subset $K \subseteq V$, there exists $\beta \in L^1_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0})$ such that

$$\|g(t, \boldsymbol{x}) - g(t, \boldsymbol{y})\|_E \leq \beta(t)\, \|\boldsymbol{x} - \boldsymbol{y}\|_{\mathbb{R}^n}$$

for a.a. $t \in I$ and each $\boldsymbol{x}, \boldsymbol{y} \in K$.

Remark 2.2.12. In later chapters, it will be useful to have slightly different formulations of the conditions described in Definition 2.2.11. First, g is locally integrably bounded if and only if it has the following property: For each compact subinterval $J \subseteq I$ and each compact subset $K \subseteq V$, there exists $\alpha \in L^1(J, \mathbb{R}_{\geq 0})$ such that

$$\|g(t, \boldsymbol{x})\|_E \leq \alpha(t)$$

for a.a. $t \in J$ and each $\boldsymbol{x} \in K$. Similarly, g is locally integrably Lipschitz if and only if it has the following property: For each compact subinterval $J \subseteq I$ and each compact subset $K \subseteq V$, there exists $\beta \in L^1(J, \mathbb{R}_{\geq 0})$ such that

$$\|g(t, \boldsymbol{x}) - g(t, \boldsymbol{y})\|_E \leq \beta(t)\, \|\boldsymbol{x} - \boldsymbol{y}\|_{\mathbb{R}^n}$$

for a.a. $t \in J$ and each $\boldsymbol{x}, \boldsymbol{y} \in K$.

Definition 2.2.13. We say that $\boldsymbol{f}$ satisfies ***Carathéodory conditions*** if

- For each $t \in I$, the map $\boldsymbol{x} \mapsto \boldsymbol{f}(t, \boldsymbol{x})$ of V into $\mathbb{R}^n$ is continuous,
- For each $\boldsymbol{x} \in V$, the map $t \mapsto \boldsymbol{f}(t, \boldsymbol{x})$ of I into $\mathbb{R}^n$ is measurable,
- $\boldsymbol{f}$ is locally integrably bounded, and
- $\boldsymbol{f}$ is locally integrably Lipschitz.

Theorem 2.2.14. *If $\boldsymbol{f}$ satisfies Carathéodory conditions, then it is solvable.*

Proof. See [Sontag 1998, Appendix C]. ∎

Remark 2.2.15. It is natural to ask if it is strictly necessary to introduce initial value problems whose right-hand sides are measurable in the time variable. The answer is affirmative: For the purposes of this thesis, we cannot strengthen the assumption that each map $t \mapsto \boldsymbol{f}(t, \boldsymbol{x})$ is measurable. Roughly speaking, this is a consequence of the fact that the continuation method requires measurable controls.

Remark 2.2.16. The first two criteria of Definition 2.2.13 ensure that for each measurable map $\boldsymbol{\nu} : \mathrm{dom}(\boldsymbol{\nu}) \to V$, where $\mathrm{dom}(\boldsymbol{\nu})$ is a subinterval of I, the map

$$t \mapsto \boldsymbol{f}(t, \boldsymbol{\nu}(t))$$

is measurable. Indeed, this follows from [Aliprantis and Border 2006, Lemma 4.51], which tells us that $\boldsymbol{f}$ is $(\mathfrak{L}_I \otimes \mathfrak{B}_V, \mathfrak{B}_{\mathbb{R}^n})$-measurable. Here, $\mathfrak{L}_I \otimes \mathfrak{B}_V$ denotes the product of the σ-algebras $\mathfrak{L}_I$ and $\mathfrak{B}_V$; see [Cohn 1980] for definitions.

The next result gives a condition under which maximally-defined solutions are defined on the entire interval I. Note this result only applies when $V = \mathbb{R}^n$.

Lemma 2.2.17. *Suppose that $V = \mathbb{R}^n$, $\boldsymbol{f}$ satisfies Carathéodory conditions, and there exists $\beta \in L^1_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0})$ such that*

$$\left\| \boldsymbol{f}(t, \boldsymbol{x}) - \boldsymbol{f}(t, \boldsymbol{y}) \right\|_{\mathbb{R}^n} \leq \beta(t) \left\| \boldsymbol{x} - \boldsymbol{y} \right\|_{\mathbb{R}^n}$$

for a.a. $t \in I$ and each $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Then the following property holds: For each $(t_0, \boldsymbol{x}_0) \in I \times V$, the maximally-defined solution of $(\boldsymbol{f}, t_0, \boldsymbol{x}_0)$ satisfies

$$\mu^{\boldsymbol{f}}(t_0, \boldsymbol{x}_0) = I.$$

Proof. See [Sontag 1998, Proposition C.3.8]. ∎

Remark 2.2.18. In later chapters, it will be useful to have a slightly different formulation of the Lipschitz condition described in Lemma 2.2.17. The map $\boldsymbol{f}$

satisfies this condition if and only it has the following property: For each compact subinterval $J \subseteq I$, there exists $\beta \in L^1(J, \mathbb{R}_{\geq 0})$ such that

$$\|\boldsymbol{f}(t, \boldsymbol{x}) - \boldsymbol{f}(t, \boldsymbol{y})\|_{\mathbb{R}^n} \leq \beta(t)\,\|\boldsymbol{x} - \boldsymbol{y}\|_{\mathbb{R}^n}$$

for a.a. $t \in J$ and each $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$.

We now consider differentiability. The next definition follows Sussmann [1998].

Definition 2.2.19. We say that $\boldsymbol{f}$ is ***locally integrably*** C^k, where $k \in \mathbb{Z}^*_{\geq 0}$, if

- For each $t \in I$, the map $\boldsymbol{x} \mapsto \boldsymbol{f}(t, \boldsymbol{x})$ of V into $\mathbb{R}^n$ is C^k,
- For each $\boldsymbol{x} \in V$, the map $t \mapsto \boldsymbol{f}(t, \boldsymbol{x})$ of I into $\mathbb{R}^n$ is measurable,
- For each $0 \leq j \leq k$, the map $\boldsymbol{D}_2^j \boldsymbol{f}$ is locally integrably bounded.[3]

Of course, if $\boldsymbol{f}$ is locally integrably C^k, then it is locally integrably C^j for $0 \leq j \leq k$.

Remark 2.2.20. In addition to the criteria described in Definition 2.2.19, some authors require that for each $\boldsymbol{x} \in V$ and each $1 \leq j \leq k$, the map

$$t \mapsto \boldsymbol{D}_2^j \boldsymbol{f}(t, \boldsymbol{x})$$

of I into $\mathrm{Hom}^j(\mathbb{R}^n, \mathbb{R}^n)$ is measurable; see [Grasse 2008, Definition 2.6]. Although this extra requirement leads to a hierarchical definition, where $\boldsymbol{f}$ and its partial derivatives $\boldsymbol{D}_2^j \boldsymbol{f}$ satisfy identical conditions, it is redundant. Indeed, each map

$$t \mapsto \boldsymbol{f}(t, \boldsymbol{x}) = (f^1(t, \boldsymbol{x}), \ldots, f^n(t, \boldsymbol{x}))$$

is assumed to be measurable, hence each of its component functions f^i is measurable. Now consider the n^{j+1} component functions

$$t \mapsto \boldsymbol{D}_2^j f^i(t, \boldsymbol{x}) \cdot (e_{i_1}, \ldots, e_{i_n}) = \frac{\partial^j f^i}{\partial x^{i_1} \cdots \partial x^{i_n}}(t, \boldsymbol{x}), \qquad 1 \leq i, i_1, \ldots, i_n \leq n$$

[3] In the literature, some authors refer to these three conditions as C^k Carathéodory conditions; see, for example, [Grasse 2008].

of the tensor representation of $t \mapsto \boldsymbol{D}_2^j \boldsymbol{f}(t, \boldsymbol{x})$ with respect to the standard bases. Since each component function is obtained as the pointwise limit of measurable functions, it is measurable [Cohn 1980]. Thus $t \mapsto \boldsymbol{D}_2^j \boldsymbol{f}(t, \boldsymbol{x})$ is measurable.

In particular, this implies that if $\boldsymbol{f}$ is locally integrably C^k, where $k \in \mathbb{N}^*$, then for each $0 \leq j \leq k$ the jth-order partial derivative $\boldsymbol{D}_2^j \boldsymbol{f}$ is locally integrably C^0.

Remark 2.2.21. The map $\boldsymbol{f}$ satisfies Carathéodory conditions whenever $\boldsymbol{f}$ is locally integrably C^1. Indeed, the first three criteria of Definition 2.2.13 are clearly satisfied. By [Sontag 1998, Proposition C.3.4], the fourth criterion is also satisfied. In particular, $\boldsymbol{f}$ has a well-defined global flow whenever $\boldsymbol{f}$ is locally integrably C^1. We will use this fact implicitly in the remainder of this section.

We now address how global flows give rise to diffeomorphisms.

Theorem 2.2.22. *Suppose that $\boldsymbol{f}$ is locally integrably C^k, where $k \in \mathbb{N}^*$. Then*

1. $\mathrm{dom}(\Phi^{\boldsymbol{f}})$ *is an open subset of* $I \times I \times V$*,*
2. $\Phi^{\boldsymbol{f}}$ *is continuous, and*
3. $\Phi^{\boldsymbol{f}}_{t,t_0}$ *is* C^k *for each* $(t, t_0) \in I \times I$*.*

The first two conclusions hold if $\boldsymbol{f}$ merely satisfies Carathéodory conditions.

Proof. See [McShane 1944, Chapter IX, Section 69.4]. ∎

Corollary 2.2.23. *Suppose that $\boldsymbol{f}$ is locally integrably C^k, where $k \in \mathbb{N}^*$, and*

$$(t, t_0) \in I \times I.$$

Then the following conclusions hold:

1. $\Phi^{\boldsymbol{f}}_{t,t_0}$ *is a* C^k *diffeomorphism of* $\mathrm{dom}(\Phi^{\boldsymbol{f}}_{t,t_0})$ *onto* $\mathrm{dom}(\Phi^{\boldsymbol{f}}_{t_0,t})$*;*
2. *The inverse of* $\Phi^{\boldsymbol{f}}_{t,t_0}$ *is* $(\Phi^{\boldsymbol{f}}_{t,t_0})^{-1} = \Phi^{\boldsymbol{f}}_{t_0,t}$*.*

Proof. By Theorem 2.2.22, it follows that

$$V_0 = \mathrm{dom}(\Phi^{\boldsymbol{f}}_{t,t_0})$$

is open in V and $\Phi^{\boldsymbol{f}}_{t,t_0}$ is C^k. Choose $\boldsymbol{x} \in V_0$. By construction, $t \in I^{\boldsymbol{f}}(t_0, \boldsymbol{x})$, hence

$$I^{\boldsymbol{f}}(t_0, \boldsymbol{x}) = I^{\boldsymbol{f}}(t, \mu^{\boldsymbol{f}}(t, t_0, \boldsymbol{x})) = I^{\boldsymbol{f}}(t, \Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{x}))$$

by Proposition 2.2.8. In particular, $t_0 \in I^{\boldsymbol{f}}(t, \Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{x}))$. By definition,

$$\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{x}) \in \tilde{V}_0 = \mathrm{dom}(\Phi^{\boldsymbol{f}}_{t_0,t}).$$

This shows that $\Phi^{\boldsymbol{f}}_{t,t_0}$ maps V_0 into $\tilde{V}_0$. It is injective, since

$$\Phi^{\boldsymbol{f}}_{t_0,t} \circ \Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{x}) = \Phi^{\boldsymbol{f}}_{t_0,t_0}(\boldsymbol{x}) = \boldsymbol{x}.$$

Exchanging the roles of t and t_0 completes the proof. ∎

Remark 2.2.24. Analogues of Theorem 2.2.22 and Corollary 2.2.23 can be found throughout the control theory literature. For example, these results appear in the work of Sussmann [1999], Agrachev and Sachkov [2004, Chapter 2], Barbu and Lefter [2005, Section 2.3], and Bullo and Lewis [2005a, Theorem A.1].

2.3 Initial value problems evolving on finite-dimensional manifolds

In this section, we extend the results of the previous section to accommodate initial value problems evolving on finite-dimensional manifolds. Although the material in this section is well-known, we are not aware of any source which provides detailed proofs. For this reason, as well as the importance of this section to later chapters, we believe that the detailed treatment in this section is warranted.

Our standing assumptions in this section are that

- I is an interval and
- M is an n-dimensional manifold.

Throughout this section, $\xi : \mathrm{dom}(\xi) \to M$ is a curve.

In Section 2.2, we saw that the solution of an initial value problem is a locally absolutely continuous (LAC) curve. To extend these notions, we first define what it means for the curve ξ to be LAC. Our approach follows Kawan [2009].

Definition 2.3.1. We say that ξ is ***locally absolutely continuous*** (LAC) if for each $t \in \mathrm{dom}(\xi)$, there exist

- A compact subinterval $I_t \subseteq \mathrm{dom}(\xi)$ such that $t \in \mathrm{int}(I_t)$, where $\mathrm{int}(I_t)$ denotes the interior of I_t relative to $\mathrm{dom}(\xi)$, and
- A chart $(V_t, \boldsymbol{\psi}_t)$ on M such that $\xi(I_t) \subseteq V_t$ and $\boldsymbol{\psi}_t \circ \xi | I_t$ is AC.

One can show that Definition 2.3.1 is equivalent to Definition 2.2.1 when M is an open submanifold of $\mathbb{R}^n$. Furthermore, one can show that Definition 2.3.1 is well-defined, in the sense that it does not depend on the particular choice of chart $(V_t, \boldsymbol{\psi}_t)$. Both of these assertions are proven in [Kawan 2009, Proposition 1.1.6].

In the literature, one can find several equivalent characterizations of what it means for a map into M to be AC and LAC. (There is sometimes no distinction drawn between these two terms, since the notion of locality is already built in by means of charts or bump functions.) For this material, we refer to [Klingenberg 1995, Sussmann 1998, Ballard 2002, Bullo and Lewis 2005a, Lewis 2009].

One can show that the LAC property is stable under composition with locally Lipschitz maps. However, we only require this result for composition with C^1 maps.

Lemma 2.3.2. *Suppose that $F : M \to N$ is C^1, where N is a finite-dimensional manifold, and ξ is LAC. Then $F \circ \xi : \mathrm{dom}(\xi) \to N$ is LAC.*

Proof. Suppose that $t \in \mathrm{dom}(\xi)$ and I_t, $(V_t, \boldsymbol{\psi}_t)$ are prescribed as in Definition 2.3.1. Choose a chart $(W, \boldsymbol{\chi})$ on N such that $F \circ \xi(t) \subseteq W$. Shrinking I_t and V_t, we can assume without loss of generality that

- $F \circ \xi | I_t \subseteq W$ and
- $(V_t, \boldsymbol{\psi}_t)$ and $(W, \boldsymbol{\chi})$ are F-compatible.

Since $F \circ \xi | I_t = (\boldsymbol{\chi} \circ F \circ \boldsymbol{\psi}_t^{-1}) \circ (\boldsymbol{\psi}_t \circ \xi | I_t)$ and C^1 maps are locally Lipschitz, it follows from Lemma 2.2.2 that $F \circ \xi | I_t$ is AC. This completes the proof. ∎

By Theorem 2.2.3, LAC curves in Euclidean spaces are differentiable almost everywhere. We now demonstrate that LAC curves in M share this property.

Lemma 2.3.3. *Suppose that ξ is LAC. Then $\dot{\xi}(t)$ exists for a.a. $t \in \mathrm{dom}(\xi)$.*

Proof. By [Barreira and Valls 2012, Definition 1.80], $\dot{\xi}(t)$ exists if and only if

$$\dot{\widehat{\boldsymbol{\psi} \circ \xi}}(t)$$

exists for some chart $(V, \boldsymbol{\psi})$ on M such that $\xi(t) \in V$. Now, for each $t \in \mathrm{dom}(\xi)$, let I_t and $(V_t, \boldsymbol{\psi}_t)$ be prescribed as in Definition 2.3.1. Then

$$\{\zeta_t = \mathrm{int}(I_t)\}_{t \in \mathrm{dom}(\xi)}$$

is an open cover of $\mathrm{dom}(\xi)$. Since $\mathrm{dom}(\xi)$ is second-countable, this cover can be reduced to a countable subcover $\{\zeta_{t_i}\}_{i \in \mathbb{N}}$. For each index $i \in \mathbb{N}$,

$$\boldsymbol{\xi}_i = \boldsymbol{\psi}_{t_i} \circ \xi | I_{t_i}$$

is AC, and we have the chain of implications

$$\begin{aligned}
\boldsymbol{\xi}_i \text{ is AC} &\implies \dot{\boldsymbol{\xi}}_i(t) \text{ exists for a.a. } t \in I_{t_i} \\
&\implies \dot{\boldsymbol{\xi}}_i(t) \text{ exists for a.a. } t \in \zeta_{t_i} \subseteq I_{t_i} \\
&\implies \dot{\boldsymbol{\xi}}_i(t) \text{ exists for each } t \in \zeta_{t_i} \setminus Z_i, \text{ where } Z_i \subseteq \zeta_{t_i} \text{ has measure zero}
\end{aligned}$$

$$\implies \dot{\xi}(t) \text{ exists for each } t \in \zeta_{t_i} \setminus Z_i.$$

Thus $\dot{\xi}(t)$ exists for each

$$t \in \bigcup_{i\in\mathbb{N}} (\zeta_{t_i} \setminus Z_i) = \bigcup_{i\in\mathbb{N}} \zeta_{t_i} \setminus \bigcup_{i\in\mathbb{N}} Z_i = \mathrm{dom}(\xi) \setminus \bigcup_{i\in\mathbb{N}} Z_i.$$

Since a countable union of sets of measure zero has measure zero [Cohn 1980], we conclude that $\dot{\xi}(t)$ exists for a.a. $t \in \mathrm{dom}(\xi)$. This completes the proof. ∎

We now consider time-varying vector fields on M.

Definition 2.3.4. Consider a map $X : I \times M \to TM$. We say that X is a ***time-varying vector field*** on M if $\pi_{TM} \circ X(t,x) = x$ for each $(t,x) \in I \times M$. The set of all such maps is denoted by $\mathscr{V}(I,M)$. Given a chart $(V,\boldsymbol{\psi})$ on M, the ***local representative*** of X in $(V,\boldsymbol{\psi})$ is the map $X_{\boldsymbol{\psi}} : I \times \boldsymbol{\psi}(V) \to \mathbb{R}^n$ defined by

$$X_{\boldsymbol{\psi}}(t,\boldsymbol{x}) = T\boldsymbol{\psi}(\boldsymbol{\psi}^{-1}(\boldsymbol{x})) \cdot X(t,\boldsymbol{\psi}^{-1}(\boldsymbol{x})).$$

Of course, one can also form the local representative of any vector field X on M by identifying X with the time-varying vector field $X(t,x) = X(x)$.

In what follows, $X \in \mathscr{V}(I,M)$.

Definition 2.3.5. Suppose that $(t_0,x_0) \in I \times M$. The triple (X,t_0,x_0) is said to be an ***initial value problem*** evolving on M, with ***right-hand side*** X and ***initial condition*** (t_0,x_0). We say that ξ is a ***solution*** of (X,t_0,x_0) if

- $\mathrm{dom}(\xi)$ is a relatively open subinterval of I containing t_0,
- ξ is LAC,
- $\xi(t_0) = x_0$, and $\dot{\xi}(t) = X(t,\xi(t))$ for a.a. $t \in \mathrm{dom}(\xi)$.

The next lemma relates solutions of X to solutions of $X_{\boldsymbol{\psi}}$.

Lemma 2.3.6. *Suppose that*

- $(t_0, x_0) \in I \times M$,
- $(V, \boldsymbol{\psi})$ *is a chart on* M *such that* $x_0 \in V$,
- ξ *is LAC, and*
- $\mathrm{image}(\xi) \subseteq V$.

Then ξ *is a solution of* (X, t_0, x_0) *if and only if*

$$\boldsymbol{\psi} \circ \xi$$

is a solution of $(X_{\boldsymbol{\psi}}, t_0, \boldsymbol{\psi}(x_0))$.

Proof. Suppose that ξ is a solution of (X, t_0, x_0). Since $\boldsymbol{\psi} \circ \xi(t_0) = \boldsymbol{\psi}(x_0)$ and

$$\begin{aligned}
\dot{\widehat{\boldsymbol{\psi} \circ \xi}}(t) &= T\boldsymbol{\psi}(\xi(t)) \circ \dot{\xi}(t) \\
&= T\boldsymbol{\psi}(\xi(t)) \circ X(t, \xi(t)) \\
&= T\boldsymbol{\psi}(\boldsymbol{\psi}^{-1} \circ \boldsymbol{\psi} \circ \xi(t)) \circ X(t, \boldsymbol{\psi}^{-1} \circ \boldsymbol{\psi} \circ \xi(t)) \\
&= X_{\boldsymbol{\psi}}(t, \boldsymbol{\psi} \circ \xi(t))
\end{aligned}$$

for a.a. $t \in \mathrm{dom}(\xi)$, it follows that $\boldsymbol{\psi} \circ \xi$ is a solution of $(X_{\boldsymbol{\psi}}, t_0, \boldsymbol{\psi}(x_0))$. Conversely, suppose that $\boldsymbol{\psi} \circ \xi$ is a solution of $(X_{\boldsymbol{\psi}}, t_0, \boldsymbol{\psi}(x_0))$. Since $\xi(t_0) = x_0$ and

$$\begin{aligned}
\dot{\xi}(t) &= \dot{\widehat{\boldsymbol{\psi}^{-1} \circ \boldsymbol{\psi} \circ \xi}}(t) \\
&= T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi} \circ \xi(t)) \cdot \dot{\widehat{\boldsymbol{\psi} \circ \xi}}(t) \\
&= T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi} \circ \xi(t)) \cdot T\boldsymbol{\psi}(\xi(t)) \cdot X(t, \xi(t)) \\
&= T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi} \circ \xi(t)) \cdot X_{\boldsymbol{\psi}}(t, \boldsymbol{\psi} \circ \xi(t)) \\
&= X(t, \xi(t))
\end{aligned}$$

for a.a. $t \in \mathrm{dom}(\xi)$, it follows that ξ is a solution of (X, t_0, x_0). ∎

Definition 2.3.7. Suppose that $(t_0, x_0) \in I \times M$ and ξ is a solution of (X, t_0, x_0). We say that ξ is ***maximally-defined*** if it has the following property: If

$$\tilde{\xi} : \mathrm{dom}(\tilde{\xi}) \to M$$

is another solution of (X, t_0, x_0), then

$$\mathrm{dom}(\tilde{\xi}) \subseteq \mathrm{dom}(\xi) \qquad \text{and} \qquad \tilde{\xi}(t) = \xi(t)$$

for each $t \in \mathrm{dom}(\tilde{\xi})$. Clearly, such a solution is unique.

Definition 2.3.8. We say that X is ***solvable*** if there exists a maximally-defined solution of (X, t_0, x_0) for each $(t_0, x_0) \in I \times M$. If X is solvable, then the maximally-defined solution of (X, t_0, x_0) is denoted by

$$\mu^X(\cdot, t_0, x_0) : I^X(t_0, x_0) \to M.$$

We have the following analogue of Proposition 2.2.8.

Proposition 2.3.9. *Suppose that X is solvable and $(t_0, x_0) \in I \times M$. Then*

1. *For each $t \in I^X(t_0, x_0)$, we have $I^X(t_0, x_0) = I^X(t, \mu^X(t, t_0, x_0))$, and*
2. *For each $t, s \in I^X(t_0, x_0)$, we have $\mu^X(t, t_0, x_0) = \mu^X(t, s, \mu^X(s, t_0, x_0))$.*

Proof. Identical to the proof of [Grasse 1979, Proposition 3.4.10]. ∎

Definition 2.3.10. Suppose that X is solvable. Define

$$\mathrm{dom}(\Phi^X) = \{(t, t_0, x_0) \in I \times I \times M \ : \ t \in I^X(t_0, x_0)\}.$$

The ***global flow*** of X is the map $\Phi^X : \mathrm{dom}(\Phi^X) \to M$ that sends (t, t_0, x_0) to

$$\Phi^X_{t,t_0}(x_0) = \mu^X(t, t_0, x_0).$$

We must also consider the global flow of X with its first two independent variables fixed. For each $(t, t_0) \times I \times I$, define

$$\mathrm{dom}(\Phi^X_{t,t_0}) = \{x_0 \in M \ : \ (t, t_0, x_0) \in \mathrm{dom}(\Phi^X)\}.$$

By a mild abuse of notation, we denote by

$$\Phi^X_{t,t_0} : \mathrm{dom}(\Phi^X_{t,t_0}) \to M$$

the map that sends x_0 to $\Phi^X_{t,t_0}(x_0)$.

Remark 2.3.11. Suppose that X is solvable and $(t_0, x_0) \in I \times M$. Then for each $t, s \in I^X(t_0, x_0)$, we have

$$\Phi^X_{t,t_0}(x_0) = \Phi^X_{t,s} \circ \Phi^X_{s,t_0}(x_0).$$

This follows immediately from Proposition 2.3.9.

We now turn to verifying that X is solvable.

Definition 2.3.12. We say that X is ***locally solvable*** if $X_{\boldsymbol{\psi}}$ is solvable for each chart $(V, \boldsymbol{\psi})$ on M.

One can show that X is locally solvable if and only if $X_{\boldsymbol{\psi}}$ is solvable for each chart $(V, \boldsymbol{\psi}) \in \mathscr{A}_M$, where $\mathscr{A}_M$ is a compatible atlas on M.

We now show that if X is locally solvable, then X is solvable as well. To this end, it will be useful to refer to the following supporting lemmas.

Lemma 2.3.13. *Suppose that X is locally solvable,*

- $(t_0, x_0) \in I \times M$,
- $(V, \boldsymbol{\psi})$ *is a chart on M such that* $x_0 \in V$,
- ξ *is LAC, and*
- $\operatorname{image}(\xi) \subseteq V$.

Then ξ is a solution of (X, t_0, x_0) if and only if

$$\boldsymbol{\psi} \circ \xi(t) = \mu^{X_{\boldsymbol{\psi}}}(t, t_0, \boldsymbol{\psi}(x_0))$$

for each $t \in \operatorname{dom}(\xi)$.

Proof. This follows immediately from Lemma 2.3.6. ∎

The next result is obvious in light of Lemma 2.3.13.

Lemma 2.3.14. *Suppose that X is locally solvable, $(V, \boldsymbol{\psi})$ is a chart on M,*

$$I_0 \times I_0 \times V_0 \subseteq \mathrm{dom}(\Phi^{X_{\boldsymbol{\psi}}}),$$

and $M_0 = \boldsymbol{\psi}^{-1}(V_0)$. Then

1. $I_0 \times I_0 \times M_0 \subseteq \mathrm{dom}(\Phi^X)$ *and*

2. $\Phi^X | I_0 \times I_0 \times M_0 = \boldsymbol{\psi}^{-1} \circ \Phi^{X_{\boldsymbol{\psi}}} \circ (\mathrm{id}_{I_0} \times \mathrm{id}_{I_0} \times \boldsymbol{\psi}|M_0)$.

Lemma 2.3.15. *Suppose that S and T are topological spaces, the maps*

$$f, g : S \to T$$

are continuous, and T is Hausdorff. Define

$$\mathrm{eq}(f, g) = \{s \in S \ : \ f(s) = g(s)\}.$$

Then $\mathrm{eq}(f, g)$ *is closed in S.*

Proof. We will show that $S \smallsetminus \mathrm{eq}(f, g)$ is open in S. If $\mathrm{eq}(f, g) = S$, then there is nothing to prove. Suppose that $\mathrm{eq}(f, g) \neq S$, and let $s \in S \smallsetminus \mathrm{eq}(f, g)$. By construction, $f(s) \neq g(s)$. Since T is Hausdorff, there exist disjoint neighbourhoods W_f and W_g of $f(s)$ and $g(s)$, respectively. Since f and g are continuous,

$$f^{-1}(W_f) \cap g^{-1}(W_g)$$

is a neighbourhood of s. We claim that

$$f^{-1}(W_f) \cap g^{-1}(W_g) \subseteq S \smallsetminus \mathrm{eq}(f, g).$$

Indeed, since W_f and W_g are disjoint, $f(\tilde{s}) \neq g(\tilde{s})$ for each

$$\tilde{s} \in f^{-1}(W_f) \cap g^{-1}(W_g).$$

We conclude that $S \smallsetminus \mathrm{eq}(f, g)$ is open in S. ∎

We say that the set $\mathrm{eq}(f, g)$ is the ***equalizer*** of f and g.

Theorem 2.3.16. *Suppose that X is locally solvable. Then X is solvable.*

Proof. Choose $(t_0, x_0) \in I \times M$. The proof is divided into three parts.

PART 1: In this part, we show that any two solutions of (X, t_0, x_0) coincide on a relatively open subinterval of I containing t_0. Suppose that

$$\xi_i : \mathrm{dom}(\xi_i) \to M, \qquad i \in \{1, 2\},$$

are two solutions of (X, t_0, x_0). Choose a chart $(V, \boldsymbol{\psi})$ on M such that $x_0 \in V$ and let $i \in \{1, 2\}$. Since ξ_i is continuous, there exists a relatively open subinterval S_i of $\mathrm{dom}(\xi_i)$ containing t_0 such that $\xi_i(S_i) \subseteq V$. Clearly, $\xi_i|S_i$ is a solution of (X, t_0, x_0). By Lemma 2.3.13,

$$(\boldsymbol{\psi} \circ \xi_i|S_i)(t) = \mu^{X_{\boldsymbol{\psi}}}(t, t_0, \boldsymbol{\psi}(x_0))$$

for each $t \in S_i$. Thus $\boldsymbol{\psi} \circ \xi_1(t) = \boldsymbol{\psi} \circ \xi_2(t)$ for each $t \in S_1 \cap S_2$.

PART 2: In this part, we show that any two solutions of (X, t_0, x_0) coincide on the intersection of their domains. Suppose that

$$\xi_i : \mathrm{dom}(\xi_i) \to M, \qquad i \in \{1, 2\},$$

are two solutions of (X, t_0, x_0). Define

$$\xi_i^r = \xi_i|(\mathrm{dom}(\xi_1) \cap \mathrm{dom}(\xi_2))$$

for each i, and consider the equalizer of ξ_1^r and ξ_2^r:

$$\mathrm{eq}(\xi_1^r, \xi_2^r) = \{t \in \mathrm{dom}(\xi_1) \cap \mathrm{dom}(\xi_2) \;:\; \xi_1^r(t) = \xi_2^r(t)\}.$$

We must show that $\mathrm{eq}(\xi_1^r, \xi_2^r)$ and $\mathrm{dom}(\xi_1) \cap \mathrm{dom}(\xi_2)$ coincide. Clearly, $\mathrm{eq}(\xi_1^r, \xi_2^r)$ is nonempty since it contains t_0. Since M is Hausdorff, $\mathrm{eq}(\xi_1^r, \xi_2^r)$ is closed in

$$\mathrm{dom}(\xi_1) \cap \mathrm{dom}(\xi_2)$$

by Lemma 2.3.15. Since $\mathrm{dom}(\xi_1) \cap \mathrm{dom}(\xi_2)$ is connected, it follows that

$$\mathrm{eq}(\xi_1^r, \xi_2^r) = \mathrm{dom}(\xi_1) \cap \mathrm{dom}(\xi_2)$$

if and only if $\mathrm{eq}(\xi_1^r, \xi_2^r)$ is open in $\mathrm{dom}(\xi_1) \cap \mathrm{dom}(\xi_2)$. To see that $\mathrm{eq}(\xi_1^r, \xi_2^r)$ is indeed open in $\mathrm{dom}(\xi_1) \cap \mathrm{dom}(\xi_2)$, let $s \in \mathrm{eq}(\xi_1^r, \xi_2^r)$. Since ξ_1^r and ξ_2^r are solutions of $(X, s, \xi_1^r(s)) = (X, s, \xi_2^r(s))$, Part 1 tells us that ξ_1^r and ξ_2^r coincide on a relatively open subinterval of I containing s. Hence ξ_1^r and ξ_2^r coincide on a relatively open subinterval of $\mathrm{dom}(\xi_1) \cap \mathrm{dom}(\xi_2)$ containing s. We conclude that $\mathrm{eq}(\xi_1^r, \xi_2^r)$ is open in $\mathrm{dom}(\xi_1) \cap \mathrm{dom}(\xi_2)$.

PART 3: In this part, we construct the maximally-defined solution of (X, t_0, x_0). To begin, we denote by $\{\xi_\alpha : \mathrm{dom}(\xi_\alpha) \to M\}_{\alpha \in A}$ the set of all solutions of (X, t_0, x_0). This set is nonempty, since if $(V, \boldsymbol{\psi})$ is a chart on M such that $x_0 \in V$, then $\boldsymbol{\psi}^{-1} \circ \mu^{X_{\boldsymbol{\psi}}}(t, t_0, \boldsymbol{\psi}(x_0))$ is a solution of (X, t_0, x_0) by Lemma 2.3.13. Now set

$$I^X(t_0, x_0) = \bigcup_{\alpha \in A} \mathrm{dom}(\xi_\alpha),$$

and define $\mu^X(\cdot, t_0, x_0) : I^X(t_0, x_0) \to M$ by

$$\mu^X(t, t_0, x_0) = \xi_\alpha(t)$$

whenever $t \in \mathrm{dom}(\xi_\alpha)$. By Part 2, $\mu^X(\cdot, t_0, x_0)$ is well-defined. To complete the proof, it is enough to show that $\mu^X(\cdot, t_0, x_0)$ is a solution of (X, t_0, x_0). First, $I^X(t_0, x_0)$ is an interval since each $\mathrm{dom}(\xi_\alpha)$ contains t_0.[4] Being a union of relatively open subintervals of I, $I^X(t_0, x_0)$ is itself a relatively open subinterval of I. Second, let us show that $\mu^X(\cdot, t_0, x_0)$ is a LAC curve. Suppose that $t \in I^X(t_0, x_0)$ and $(V_t, \boldsymbol{\psi}_t)$ is a chart on M such that

$$x_t = \mu^X(t, t_0, x_0) \in V_t.$$

By Lemma 2.3.13 and the definition of $\mu^X(\cdot, t_0, x_0)$, we have

$$\mu^X(\tau, t_0, x_0) = \boldsymbol{\psi}_t^{-1} \circ \mu^{X_{\boldsymbol{\psi}_t}}(\tau, t, \boldsymbol{\psi}_t(x_t))$$

[4]For an explicit proof of this fact, see [Munkres 2000, Theorem 23.3].

for each $\tau \in I^{X_{\psi_t}}(t, \boldsymbol{\psi}_t(x_t))$. From this expression, we conclude that $\mu^X(\cdot, t_0, x_0)$ is a LAC curve. Finally, it is evident $\mu^X(t_0, t_0, x_0) = x_0$, so it remains to show that

$$\dot{\mu}^X(t, t_0, x_0) = X(t, \mu^X(t, t_0, x_0))$$

for a.a. $t \in I^X(t_0, x_0)$. This follows from a covering argument, proceeding along the same lines as the argument used in the proof of Lemma 2.3.3. ∎

In later chapters, the next result plays a subtle but important role. Roughly speaking, this result tells us that if X is solvable, then its maximally-defined solutions are insensitive to the time behaviour of X on a set of measure zero. This will come into play in Chapter 3, when we consider control systems.

Lemma 2.3.17. *Suppose that*

- $X_1, X_2 \in \mathscr{V}(I, M)$ *are solvable,*
- $X_1(t, x) = X_2(t, x)$ *for a.a.* $t \in I$ *and each* $x \in M$, *and*
- $(t_0, x_0) \in I \times M$.

Then $I^{X_1}(t_0, x_0) = I^{X_2}(t_0, x_0)$ *and*

$$\mu^{X_1}(t, t_0, x_0) = \mu^{X_2}(t, t_0, x_0)$$

for each $t \in I^{X_1}(t_0, x_0)$.

Proof. For brevity, we write $\mu_i = \mu^{X_i}(\cdot, t_0, x_0)$ for each $i \in \{1, 2\}$. By definition,

$$\mu_i(t_0) = x_0 \qquad \text{and} \qquad \dot{\mu}_i(t) = X_i(t, \mu_i(t))$$

for a.a. $t \in I^{X_i}(t_0, x_0)$. Since $X_1(t, x) = X_2(t, x)$ for a.a. $t \in I$ and each $x \in M$,

$$\dot{\mu}_1(t) = X_2(t, \mu_1(t))$$

for a.a. $t \in I^{X_1}(t_0, x_0)$. In other words, μ_1 is a solution of (X_2, t_0, x_0). Since μ_2 is maximally-defined, we have $I^{X_1}(t_0, x_0) \subseteq I^{X_2}(t_0, x_0)$ and

$$\mu^{X_1}(t, t_0, x_0) = \mu^{X_2}(t, t_0, x_0)$$

for each $t \in I^{X_1}(t_0, x_0)$. Exchanging the roles of the indices completes the proof. ∎

Definition 2.3.18. We say that X satisfies ***Carathéodory conditions*** if $X_{\boldsymbol{\psi}}$ satisfies Carathéodory conditions for each chart $(V, \boldsymbol{\psi})$ on M.

One can show that X satisfies Carathéodory conditions if and only if $X_{\boldsymbol{\psi}}$ satisfies Carathéodory conditions for each chart $(V, \boldsymbol{\psi}) \in \mathscr{A}_M$, where $\mathscr{A}_M$ is a compatible atlas on M.

The next result extends Theorem 2.2.14.

Lemma 2.3.19. *If X satisfies Carathéodory conditions, then it is solvable.*

Proof. This follows immediately from Theorems 2.2.14 and 2.3.16. ∎

Definition 2.3.20. We say that X is ***locally integrably*** C^k, where $k \in \mathbb{Z}^*_{\geq 0}$, if $X_{\boldsymbol{\psi}}$ is locally integrably C^k for each chart $(V, \boldsymbol{\psi})$ on M. Of course, if X is locally integrably C^k, then it is locally integrably C^j for $0 \leq j \leq k$.

Again, one can show that X is locally integrably C^k if and only if $X_{\boldsymbol{\psi}}$ is locally integrably C^k for each chart $(V, \boldsymbol{\psi}) \in \mathscr{A}_M$, where $\mathscr{A}_M$ is a compatible atlas on M.

The next three remarks will be used implicitly throughout this thesis.

Remark 2.3.21. Suppose that X is locally integrably C^0 and $x \in M$. Define the map $\gamma : I \to T_xM$ by $\gamma(t) = X(t, x)$. Then $\gamma \in L^1_{\mathrm{loc}}(I, T_xM)$.

Remark 2.3.22. Suppose that X is locally integrably C^k, where $k \in \mathbb{Z}^*_{\geq 0}$. Then each frozen-time vector field $X_t : M \to TM$ defined by $X_t(x) = X(t, x)$ is C^k.

Remark 2.3.23. Suppose that X is locally integrably C^1. Then X satisfies Carathéodory conditions and, consequently, is solvable.

We now extend Theorem 2.2.22. To do so, we need the following lemma.

Lemma 2.3.24. *Suppose that X is locally integrably C^k, where $k \in \mathbb{N}^*$, and*

$$(t_*, x_*) \in I \times M.$$

Then there exists a product neighbourhood

$$I_* \times I_* \times M_*$$

of (t_*, t_*, x_*) in $I \times I \times M$ such that

1. $I_* \times I_* \times M_* \subseteq \mathrm{dom}(\Phi^X)$,

2. $\Phi^X | I_* \times I_* \times M_*$ *is continuous, and*

3. $\Phi^X_{t,t_0} | M_*$ *is C^k for each $(t, t_0) \in I_* \times I_*$.*

Proof. Suppose that $(V, \boldsymbol{\psi})$ is a chart on M such that $x_* \in V$. By Theorem 2.2.22,

1. $\mathrm{dom}(\Phi^{X_{\boldsymbol{\psi}}})$ is an open subset of $I \times I \times \boldsymbol{\psi}(V)$,

2. $\Phi^{X_{\boldsymbol{\psi}}}$ is continuous, and

3. $\Phi^{X_{\boldsymbol{\psi}}}_{t,t_0}$ is C^k for each $(t, t_0) \in I \times I$.

In particular, since $(t_*, t_*, \boldsymbol{\psi}(x_*)) \in \mathrm{dom}(\Phi^{X_{\boldsymbol{\psi}}})$, there exists a product neighbourhood $I_* \times I_* \times V_*$ of $(t_*, t_*, \boldsymbol{\psi}(x_*))$ in $I \times I \times \boldsymbol{\psi}(V)$ such that

$$I_* \times I_* \times V_* \subseteq \mathrm{dom}(\Phi^{X_{\boldsymbol{\psi}}}).$$

Setting $M_* = \boldsymbol{\psi}^{-1}(V_*)$, we see that $I_* \times I_* \times M_*$ is the desired product neighbourhood of (t_*, t_*, x_*). Indeed, Lemma 2.3.14 tells us that

1. $I_* \times I_* \times M_* \subseteq \mathrm{dom}(\Phi^X)$ and

2. $\Phi^X | I_* \times I_* \times M_* = \boldsymbol{\psi}^{-1} \circ \Phi^{X_{\boldsymbol{\psi}}} \circ (\mathrm{id}_{I_*} \times \mathrm{id}_{I_*} \times \boldsymbol{\psi} | M_*)$.

The conclusions of the lemma follow by composition. ∎

The next result extends Theorem 2.2.22.

Theorem 2.3.25. *Suppose that X is locally integrably C^k, where $k \in \mathbb{N}^*$. Then*

1. $\mathrm{dom}(\Phi^X)$ *is an open subset of* $I \times I \times M$*,*
2. Φ^X *is continuous, and*
3. Φ^X_{t,t_0} *is* C^k *for each* $(t, t_0) \in I \times I$*.*

Proof. We follow the proof of Grasse [1979, Theorem 3.4.12]. Given

$$(t_*, x_*) \in I \times M,$$

we say that $s \in I^X(t_*, x_*)$ has the C^k ***neighbourhood property*** (relative to the particular choice of (t_*, x_*)) if there exists a product neighbourhood

$$I^1_s \times I^2_s \times M_s$$

of (s, t_*, x_*) in $I \times I \times M$ such that

1. $I^1_s \times I^2_s \times M_s \subseteq \mathrm{dom}(\Phi^X)$,
2. $\Phi^X | I^1_s \times I^2_s \times M_s$ is continuous, and
3. $\Phi^X_{t,t_0} | M_s$ is C^k for each $(t, t_0) \in I^1_s \times I^2_s$.

We say that such a neighbourhood is a C^k ***neighbourhood*** for s. The set of all $s \in I^X(t_*, x_*)$ that have the C^k neighbourhood property is denoted by NP^k_*.

To complete the proof, it is enough to show that

$$\mathrm{NP}^k_* = I^X(t_*, x_*) \tag{2.6}$$

for each $(t_*, x_*) \in I \times M$. To this end, choose $(t_*, x_*) \in I \times M$. By Lemma 2.3.24, NP_*^k is nonempty since it contains t_*, and NP_*^k is open in $I^X(t_*, x_*)$ by construction. Since $I^X(t_*, x_*)$ is connected, it follows that

$$\begin{aligned} \mathrm{NP}_*^k = I^X(t_*, x_*) &\iff \mathrm{NP}_*^k \text{ is closed in } I^X(t_*, x_*) \\ &\iff \overline{\mathrm{NP}_*^k} = \mathrm{NP}_*^k \\ &\iff \overline{\mathrm{NP}_*^k} \subseteq \mathrm{NP}_*^k, \end{aligned} \tag{2.7}$$

where $\overline{\mathrm{NP}_*^k}$ denotes the closure of NP_*^k in $I^X(t_*, x_*)$.

To prove (2.7), choose $s \in \overline{\mathrm{NP}_*^k}$. By Lemma 2.3.24, there exists a product neighbourhood $I_s \times I_s \times M_s$ of $(s, s, \Phi^X_{s,t_*}(x_*))$ in $I \times I \times M$ such that

1. $I_s \times I_s \times M_s \subseteq \mathrm{dom}(\Phi^X)$,
2. $\Phi^X | I_s \times I_s \times M_s$ is continuous, and
3. Φ^X_{t,t_0} is C^k for each $(t, t_0) \in I_s \times I_s$.

Since s is a limit point of NP_*^k in $I^X(t_*, x_*)$, it can be approached arbitrarily closely by elements of $I_s \cap \mathrm{NP}_*^k$. Thus there exists

$$\rho \in I_s \cap \mathrm{NP}_*^k$$

such that $\Phi^X_{\rho,t_*}(x_*) \in M_s$. This is illustrated in Figure 2.1.

Since $\rho \in \mathrm{NP}_*^k$, there exists a C^k neighbourhood $I^1_\rho \times I^2_\rho \times M_\rho$ for ρ. Write

$$F^X_s = \Phi^X | I_s \times I_s \times M_s \qquad \text{and} \qquad G^X_\rho = \Phi^X | I^1_\rho \times I^2_\rho \times M_\rho.$$

Shrinking $I^1_\rho \times I^2_\rho \times M_\rho$, we can assume without loss of generality that

$$\mathrm{image}(G^X_\rho) \subseteq M_s.$$

We claim that $I_s \times I^2_\rho \times M_\rho$ is a C^k neighbourhood for s. Consider the map

$$H^X_s = \Phi^X | I_s \times I^2_\rho \times M_\rho.$$

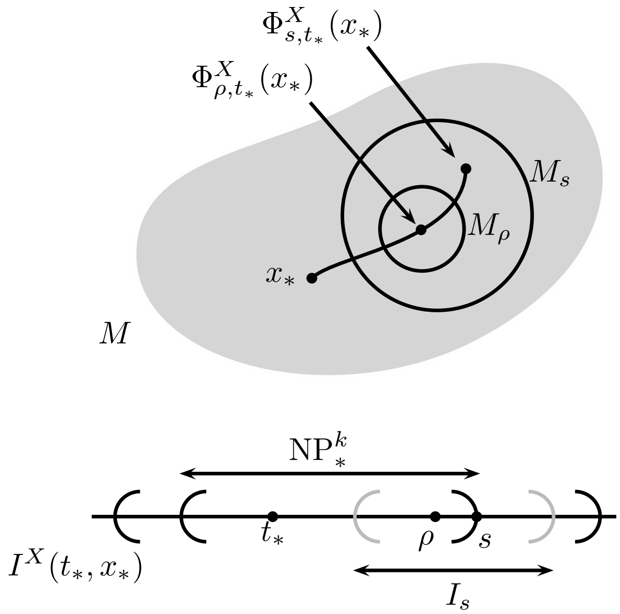

Figure 2.1: An illustration of the proof of Theorem 2.3.25

Since

$$G^X_s(\beta, \beta_0, x) = F^X_s(\beta, \rho, G^X_\rho(\rho, \beta_0, x)),$$

the criteria of the C^k neighbourhood property are satisfied by composition. This proves the claim, and the proof is complete. ∎

Corollary 2.3.26. *Suppose that X is locally integrably C^k, where $k \in \mathbb{N}^*$, and*

$$(t, t_0) \in I \times I.$$

Then the following conclusions hold:

1. *Φ^X_{t,t_0} is a C^k diffeomorphism of $\mathrm{dom}(\Phi^X_{t,t_0})$ onto $\mathrm{dom}(\Phi^X_{t_0,t})$;*
2. *The inverse of Φ^X_{t,t_0} is $(\Phi^X_{t,t_0})^{-1} = \Phi^X_{t_0,t}$.*

Proof. Identical to the proof of Corollary 2.2.23. ∎

Chapter 3

Control systems

In this chapter, we establish background material concerning control systems. As indicated in Chapter 1, the control systems of interest in this thesis are deterministic nonlinear control systems evolving in continuous time.

This chapter is organized in the following way. In Section 3.1, we establish a number of general definitions. In Section 3.2, we recall the theory of C^q_p and C^q_q-polynomial control systems evolving on open subsets of Euclidean spaces, following Margheri [1996]. Under certain conditions, the anchored endpoint maps of such control systems are C^q. This fact is instrumental in overcoming the first obstruction to the continuation method. Indeed, as explained in Chapter 1, the first obstruction is simply that the anchored endpoint maps of a given control system may fail to be C^2. Finally, in Section 3.3, we extend the theory of C^q_p and C^q_q-polynomial control systems to accommodate control systems evolving on finite-dimensional manifolds.

Our standing assumptions throughout this chapter are that

- I is an interval and
- M is an n-dimensional manifold.

Furthermore, for each $k \in \mathbb{Z}_{\geq 0}$, the map

$$\mathrm{ev}_k : \mathrm{Hom}^k(\mathbb{R}^r, \mathbb{R}^n) \times \underbrace{\mathbb{R}^r \times \cdots \times \mathbb{R}^r}_{k \text{ factors}} \to \mathbb{R}^n$$

is defined by

$$\mathrm{ev}_k \cdot (\boldsymbol{\lambda}, \boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_k) = \boldsymbol{\lambda} \cdot (\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_k).$$

Note that $\mathrm{ev}_0 = \mathrm{id}_{\mathbb{R}^n}$ and $\mathrm{ev}_k \in \mathrm{Hom}(\mathrm{Hom}^k(\mathbb{R}^r, \mathbb{R}^n), \mathbb{R}^r, \ldots, \mathbb{R}^r, \mathbb{R}^n)$.

3.1 Control systems

Roughly speaking, a control system evolving on M is a pair

$$\Sigma = (f, \mathscr{U})$$

comprised of a time-varying and parameter-dependent vector field

$$f : I \times M \times \mathbb{R}^r \to TM$$

and a set $\mathscr{U}$ of controls. Each control $\boldsymbol{u} \in \mathscr{U}$ describes a way in which the parameter in $\mathbb{R}^r$ can be manipulated over time to produce a controlled trajectory of Σ. Since the controlled trajectories of Σ are maximally-defined solutions of initial value problems evolving on M, the substitution of each $\boldsymbol{u} \in \mathscr{U}$ into f must yield a time-varying vector field which is solvable. In what follows, this intuitive description is made precise. The next definition follows Grasse [1979].

Definition 3.1.1. Suppose that $f : I \times M \times \mathbb{R}^r \to TM$. We say that f is a ***controllable time-varying vector field***[1] on M if $\pi_{TM} \circ f(t, x, \boldsymbol{\omega}) = x$ for each

[1] More generally, one can consider the case $f : I \times M \times \Omega \to TM$, where Ω is a separable metrizable space. This is discussed in detail by Grasse and Sussmann [1990], Sontag [1998] and Grasse [2003]. We will not need this degree of generality, however, since our interest lies in controllable time-varying vector fields on M such that $(x, \boldsymbol{\omega}) \mapsto f(t, x, \boldsymbol{\omega})$ is differentiable.

$(t, x, \boldsymbol{\omega}) \in I \times M \times \mathbb{R}^r$. The set of all such maps is denoted by $\mathscr{V}(I, M, \mathbb{R}^r)$. Given a chart $(V, \boldsymbol{\psi})$ on M, the ***local representative*** of f in $(V, \boldsymbol{\psi})$ is the map

$$f_{\boldsymbol{\psi}} : I \times \boldsymbol{\psi}(V) \times \mathbb{R}^r \to \mathbb{R}^n$$

defined by

$$f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{\omega}) = T\boldsymbol{\psi}(\boldsymbol{\psi}^{-1}(\boldsymbol{x})) \cdot f(t, \boldsymbol{\psi}^{-1}(\boldsymbol{x}), \boldsymbol{\omega}).$$

In what follows, $f \in \mathscr{V}(I, M, \mathbb{R}^r)$ and $\mathscr{U} \subseteq \mathrm{Meas}(I, \mathbb{R}^r)$ is nonempty. Recall from Section 2.1.3 that $\mathrm{Meas}(I, \mathbb{R}^r)$ consists of equivalence classes of maps.

Definition 3.1.2. Suppose that $\boldsymbol{u} \in \mathscr{U}$. For each representative $\tilde{\boldsymbol{u}}$ of the equivalence class $\boldsymbol{u}$, define the time-varying vector field $f^{\tilde{\boldsymbol{u}}} \in \mathscr{V}(I, M)$ by

$$f^{\tilde{\boldsymbol{u}}}(t, x) = f(t, x, \tilde{\boldsymbol{u}}(t)).$$

We say that $\boldsymbol{u}$ is ***f-admissible*** if $f^{\tilde{\boldsymbol{u}}}$ is solvable for each representative $\tilde{\boldsymbol{u}}$ of $\boldsymbol{u}$. Similarly, we say that $\boldsymbol{u}$ is ***C^k f-admissible***, where $k \in \mathbb{N}^*$, if $f^{\tilde{\boldsymbol{u}}}$ is locally integrably C^k for each representative $\tilde{\boldsymbol{u}}$ of $\boldsymbol{u}$.

By Remark 2.3.23, $\boldsymbol{u}$ is f-admissible whenever it is C^1 f-admissible.

Lemma 3.1.3. *Suppose that $\boldsymbol{u} \in \mathscr{U}$ is f-admissible,*

$$(t_0, x_0) \in I \times M,$$

and $\tilde{\boldsymbol{u}}_1$, $\tilde{\boldsymbol{u}}_2$ are representatives of $\boldsymbol{u}$. Then $I^{f^{\tilde{\boldsymbol{u}}_1}}(t_0, x_0) = I^{f^{\tilde{\boldsymbol{u}}_2}}(t_0, x_0)$ and

$$\mu^{f^{\tilde{\boldsymbol{u}}_1}}(t, t_0, x_0) = \mu^{f^{\tilde{\boldsymbol{u}}_2}}(t, t_0, x_0)$$

for each $t \in I^{f^{\tilde{\boldsymbol{u}}_1}}(t_0, x_0)$.

Proof. By construction, we have

$$f^{\tilde{\boldsymbol{u}}_1}(t, x) = f^{\tilde{\boldsymbol{u}}_2}(t, x)$$

for a.a. $t \in I$ and each $x \in M$. Invoking Lemma 2.3.17 completes the proof. ∎

In light of Lemma 3.1.3, in the remainder of this thesis we will simply write $f^{\boldsymbol{u}}$ instead of $f^{\tilde{\boldsymbol{u}}}$, with the understanding that $\boldsymbol{u}$ is simply a placeholder for any representative $\tilde{\boldsymbol{u}}$ of $\boldsymbol{u}$. We now formally define control systems.

Definition 3.1.4. Suppose that each $\boldsymbol{u} \in \mathscr{U}$ is f-admissible. The pair

$$\Sigma = (f, \mathscr{U})$$

is said to be a ***control system*** evolving on M, with ***time domain*** I, ***state space*** M, and ***control space*** $\mathscr{U}$. The constituent elements of I, M, and $\mathscr{U}$ are said to be ***times***, ***states***, and ***controls***, respectively. Finally, given a chart $(V, \boldsymbol{\psi})$ on M, the ***local representative*** of Σ in $(V, \boldsymbol{\psi})$ is the control system

$$\Sigma_{\boldsymbol{\psi}} = (f_{\boldsymbol{\psi}}, \mathscr{U}).$$

In what follows, $\Sigma = (f, \mathscr{U})$ is a control system evolving on M.

Definition 3.1.5. We say that Σ is ***time-invariant*** if f is not functionally dependent on its first independent variable. When we wish to emphasize the fact that Σ is not time-invariant, we will say that Σ is ***time-varying***.

Definition 3.1.6. We say that

- Σ is a C^k control system, where $k \in \mathbb{N}^*$, if each $\boldsymbol{u} \in \mathscr{U}$ is C^k f-admissible, and
- Σ ***uses L^p controls***, where $p \in \mathbb{R}^*_{\geq 1}$, if $\mathscr{U} = L^p_{\mathrm{loc}}(I, \mathbb{R}^r)$.

Definition 3.1.7. Suppose that $(t_0, x_0, \boldsymbol{u}) \in I \times M \times \mathscr{U}$. The maximally-defined solution of $(f^{\boldsymbol{u}}, t_0, x_0)$ is denoted by

$$\mu^\Sigma(\cdot, t_0, x_0, \boldsymbol{u}) : I^\Sigma(t_0, x_0, \boldsymbol{u}) \to M.$$

We say that $\mu^\Sigma(\cdot, t_0, x_0, \boldsymbol{u})$ is the ***$\boldsymbol{u}$-controlled trajectory*** of Σ with ***initial condition*** (t_0, x_0).

We now consider global flows and endpoint maps. In Section 2.3, we defined the global flow of a time-varying vector field on M. This object was introduced to study the functional dependence of maximally-defined solutions on the initial condition. In a similar way, we can study the functional dependence of the controlled trajectories of Σ on the initial condition and control, by aggregating these controlled trajectories into a single map.

Definition 3.1.8. Define

$$\operatorname{dom}(\Phi^\Sigma) = \{(t, t_0, x_0, \boldsymbol{u}) \in I \times I \times M \times \mathscr{U} \ : \ t \in I^\Sigma(t_0, x_0, \boldsymbol{u})\}.$$

The ***global flow*** of Σ is the map $\Phi^\Sigma : \operatorname{dom}(\Phi^\Sigma) \to M$ that sends $(t, t_0, x_0, \boldsymbol{u})$ to

$$\Phi^\Sigma_{t,t_0}(x_0, \boldsymbol{u}) = \mu^\Sigma(t, t_0, x_0, \boldsymbol{u}).$$

We say that Σ is ***complete*** if $\operatorname{dom}(\Phi^\Sigma) = I \times I \times M \times \mathscr{U}$.

Next, we define the endpoint map and anchored endpoint maps of Σ. To do so, it is convenient to introduce time-restricted versions of Σ.[2] In what follows,

$$(t, t_0) \in I \times I$$

is such that $t_0 \leq t$.

Definition 3.1.9. Suppose that Σ uses L^p controls. Define

$$\Sigma|[t_0, t] = (f|[t_0, t] \times M \times L^p([t_0, t], \mathbb{R}^r), L^p([t_0, t], \mathbb{R}^r))$$

Clearly, $\Sigma|[t_0, t]$ is itself a control system that uses L^p controls.

Note that if $I = [t_0, t]$, then there is no distinction between $\Sigma|[t_0, t]$ and Σ.

[2]The concept of time restriction is used, albeit implicitly, by Margheri [1996]; see the comment near the bottom of [Margheri 1996, page 193].

Definition 3.1.10. Suppose that Σ uses L^p controls. Define

$$\mathrm{dom}(\mathrm{End}^{\Sigma|[t_0,t]}) = \{(x_0, \boldsymbol{u}) \in M \times L^p([t_0,t], \mathbb{R}^r) \ : \ (t, t_0, x_0, \boldsymbol{u}) \in \mathrm{dom}(\Phi^{\Sigma|[t_0,t]})\}.$$

The ***endpoint map***[3] of $\Sigma|[t_0,t]$ is the map

$$\mathrm{End}^{\Sigma|[t_0,t]} : \mathrm{dom}(\mathrm{End}^{\Sigma|[t_0,t]}) \to M$$

that sends $(x_0, \boldsymbol{u})$ to

$$\mathrm{End}^{\Sigma|[t_0,t]}(x_0, \boldsymbol{u}) = \Phi^{\Sigma|[t_0,t]}_{t,t_0}(x_0, \boldsymbol{u}).$$

The namesake of this map is clear, since the value of $\mathrm{End}^{\Sigma|[t_0,t]}(x_0, \boldsymbol{u})$ is simply the right endpoint $\mu^{\Sigma|[t_0,t]}(t, t_0, x_0, \boldsymbol{u})$ of the $\boldsymbol{u}$-controlled trajectory $\mu^{\Sigma|[t_0,t]}(\cdot, t_0, x_0, \boldsymbol{u})$.

Definition 3.1.11. Suppose that Σ uses L^p controls and $x_0 \in M$. Define

$$\mathrm{dom}(\mathrm{End}^{\Sigma|[t_0,t]}_{x_0}) = \{\boldsymbol{u} \in L^p([t_0,t], \mathbb{R}^r) \ : \ (x_0, \boldsymbol{u}) \in \mathrm{dom}(\mathrm{End}^{\Sigma|[t_0,t]})\}.$$

The x_0***-anchored endpoint map*** of $\Sigma|[t_0,t]$ is the map

$$\mathrm{End}^{\Sigma|[t_0,t]}_{x_0} : \mathrm{dom}(\mathrm{End}^{\Sigma|[t_0,t]}_{x_0}) \to M$$

that sends $\boldsymbol{u}$ to

$$\mathrm{End}^{\Sigma|[t_0,t]}_{x_0}(\boldsymbol{u}) = \mathrm{End}^{\Sigma|[t_0,t]}(x_0, \boldsymbol{u})$$

The value of $\mathrm{End}^{\Sigma|[t_0,t]}_{x_0}(\boldsymbol{u})$ is simply the right endpoint $\mu^{\Sigma|[t_0,t]}(t, t_0, x_0, \boldsymbol{u})$ of the $\boldsymbol{u}$-controlled trajectory $\mu^{\Sigma|[t_0,t]}(\cdot, t_0, x_0, \boldsymbol{u})$. Since the left endpoint of

$$\mu^{\Sigma|[t_0,t]}(\cdot, t_0, x_0, \boldsymbol{u})$$

is x_0, we regard each such controlled trajectory as being anchored at x_0.

Next, we state a useful relationship between $\mathrm{dom}(\Phi^{\Sigma})$ and $\mathrm{dom}(\Phi^{\Sigma|[t_0,t]})$.

[3]This map is sometimes called the input-to-state map of Σ, as in [Sontag 1998].

Remark 3.1.12. Suppose that Σ uses L^p controls and

$$I_0 \times I_0 \times M_0 \times \mathscr{U}_0 \subseteq \operatorname{dom}(\Phi^\Sigma)$$

is a product open subset of $I \times I \times M \times \mathscr{U}$. Then

$$M_0 \times \tilde{\mathscr{U}_0} \subseteq \operatorname{dom}(\Phi^{\Sigma|[t_0,t]})$$

is a (possibly empty) product open subset of $M \times L^p([t_0,t],\mathbb{R}^r)$, where

$$\tilde{\mathscr{U}_0} = \pi^p_{[t_0,t]}(\mathscr{U}_0)$$

and $\pi^p_{[t_0,t]}$ is the projection map defined as in Section 2.1.6.

The next lemma is an analogue of Lemma 2.3.14 for control systems.

Lemma 3.1.13. *Suppose that $(V,\boldsymbol{\psi})$ is a chart on M,*

$$I_0 \times I_0 \times V_0 \times \mathscr{U}_0 \subseteq \operatorname{dom}(\Phi^{\Sigma_{\boldsymbol{\psi}}}),$$

and $M_0 = \boldsymbol{\psi}^{-1}(V_0)$. Then

1. $I_0 \times I_0 \times M_0 \times \mathscr{U}_0 \subseteq \operatorname{dom}(\Phi^\Sigma)$,
2. $\Phi^\Sigma | I_0 \times I_0 \times M_0 \times \mathscr{U}_0 = \boldsymbol{\psi}^{-1} \circ \Phi^{\Sigma_{\boldsymbol{\psi}}} \circ (\mathrm{id}_{I_0} \times \mathrm{id}_{I_0} \times \boldsymbol{\psi}|M_0 \times \mathrm{id}_{\mathscr{U}_0})$, *and*
3. *We have*

$$\operatorname{End}^{\Sigma|[t_0,t]}|M_0 \times \tilde{\mathscr{U}_0} = \boldsymbol{\psi}^{-1} \circ \operatorname{End}^{\Sigma_{\boldsymbol{\psi}}|[t_0,t]} \circ (\boldsymbol{\psi}|M_0 \times \mathrm{id}_{\tilde{\mathscr{U}_0}})$$

for each $(t,t_0) \in I_0 \times I_0$ such that $t_0 \leq t$, where

$$\tilde{\mathscr{U}_0} = \pi^p_{[t_0,t]}(\mathscr{U}_0).$$

Next, we define a notion of complete controllability.

Definition 3.1.14. Suppose that $x_0 \in M$. We say that Σ is ***completely controllable from*** x_0 ***on*** $[t_0,t]$ if the map $\operatorname{End}^{\Sigma|[t_0,t]}_{x_0}$ is surjective. Of course, if $I = [t_0,t]$, then Σ is completely controllable from x_0 on I if and only if $\operatorname{End}^\Sigma_{x_0}$ is surjective.

Conditions which imply complete controllability are derived by Nikitin [1994], Jurdjevic [1997], Ayala et al. [2009], and Jouan [2011a,b]. As one might expect, these conditions impose certain restrictions on Σ. For example, the conditions derived by Jouan [2011a] assume that M is at least compact, and possibly a Lie group, whereas the conditions derived by Jurdjevic [1997] assume that Σ is a C^ω control-affine system. We do not dwell on these conditions here, since for our purposes it is enough to assume that Σ is completely controllable without further qualification. This is because we are interested in studying obstructions to the continuation method under the standing assumption of complete controllability from a fixed initial state.

Due to their special structure, control-affine systems play an important role in later chapters. Before a formal definition, let us consider the following example.

Example 3.1.15. Suppose that

- $f_0, f_1, \ldots, f_r$ are C^k vector fields on M, where $k \in \mathbb{N}^*$,
- $f(t, x, \boldsymbol{\omega}) = f_0(x) + \sum_{i=1}^{r} \omega^i f_i(x)$,
- $\boldsymbol{u} \in L^1_{\text{loc}}(I, \mathbb{R}^r)$,
- $\tilde{\boldsymbol{u}}$ is a representative of $\boldsymbol{u}$, and
- $(V, \boldsymbol{\psi})$ is a chart on M.

By linearity,

$$f_{\boldsymbol{\psi}}^{\tilde{\boldsymbol{u}}}(t, \boldsymbol{x}) = (f_0)_{\boldsymbol{\psi}}(\boldsymbol{x}) + \sum_{i=1}^{r} \tilde{u}^i(t)(f_i)_{\boldsymbol{\psi}}(\boldsymbol{x}).$$

Clearly, $f_{\boldsymbol{\psi}}^{\tilde{\boldsymbol{u}}}$ satisfies the first two criteria of Definition 2.2.19. For the third criterion, choose a compact subset $K \subseteq \boldsymbol{\psi}(V)$ and define $C \in \mathbb{R}_{\geq 0}$ by

$$C = \sup_{\substack{0 \leq i \leq r,\, 0 \leq j \leq k \\ \boldsymbol{x} \in K}} \{\|\boldsymbol{D}^j (f_i)_{\boldsymbol{\psi}}(\boldsymbol{x})\|\}.$$

Choose $0 \leq j \leq k$. We have

$$\begin{aligned}
\|\boldsymbol{D}_2^j f_{\boldsymbol{\psi}}^{\tilde{\boldsymbol{u}}}(t, \boldsymbol{x})\| &\leq \|\boldsymbol{D}^j (f_0)_{\boldsymbol{\psi}}(\boldsymbol{x})\| + \sum_{i=1}^{r} |\tilde{u}^i(t)| \, \|\boldsymbol{D}^j (f_i)_{\boldsymbol{\psi}}(\boldsymbol{x})\| \\
&\leq C \left(1 + \sum_{i=1}^{r} |\tilde{u}^i(t)|\right) \\
&\leq C \left(1 + \sqrt{r} \|\tilde{\boldsymbol{u}}(t)\|_{\mathbb{R}^r}\right)
\end{aligned}$$

for a.a. $t \in I$ and each $\boldsymbol{x} \in K$. Consequently, $\boldsymbol{u}$ is C^k f-admissible.

Definition 3.1.16. We say that Σ is a C^k ***control-affine system***, where $k \in \mathbb{N}^*$, if there exist C^k vector fields $f_0, f_1, \ldots, f_r$ on M such that

$$f(t, x, \boldsymbol{\omega}) = f_0(x) + \sum_{i=1}^{r} \omega^i f_i(x). \tag{3.1}$$

If M is C^ω (that is, real-analytic), then the definition of a C^ω control-affine system is totally analogous. The vector field f_0 is called the ***drift vector field*** of Σ, while the vector fields $f_1, \ldots, f_r$ are called the ***control vector fields*** of Σ. If $f_0(x) = 0_x$ for each $x \in M$, then we say that Σ is ***driftless***. In this case, f can be written as

$$f(t, x, \boldsymbol{\omega}) = \sum_{i=1}^{r} \omega^i f_i(x).$$

When we wish to emphasize the fact that Σ is not driftless, we will say that Σ is a C^k control-affine system ***with drift***.

As a point of terminology, note that a C^k control-affine system Σ is not necessarily a C^k control system. However, as indicated by Example 3.1.15, this will be the case whenever each element of $\mathscr{U}$ is locally integrable.

In the next section, we consider special classes of control systems.

3.2 C^q_p and C^q_q-polynomial control systems evolving on open subsets of Euclidean spaces

In this section, we recall the theory of C^q_p and C^q_q-polynomial control systems evolving on open subsets of Euclidean spaces, following Margheri [1996]. Under certain conditions, the anchored endpoint maps of such control systems are C^q. As explained at the beginning of this chapter, this fact is instrumental in overcoming the first obstruction to the continuation method. This point will be made clearer in Chapter 4, when we present the continuation method in full detail. Throughout this section,

- V is a nonempty open subset of $\mathbb{R}^n$,
- $\Sigma = (\boldsymbol{f}, \mathscr{U})$ is a control system evolving on V, and
- $p \in \mathbb{R}_{\geq 1}$ and $q \in \mathbb{N}$ are fixed, subject to the requirement that $p \geq q$.

3.2.1 Basic definitions and properties

In this section, we will encounter the expression

$$\frac{p}{p-p} = \frac{p}{0}.$$

Wherever this expression appears, it is understood to be equal to ∞. In particular, in the following definition, the function α_p is an element of $L^\infty_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0})$.

Definition 3.2.1. We say that $\boldsymbol{f}$ is C^q_p if

- For each $t \in I$, the map $(\boldsymbol{x}, \boldsymbol{\omega}) \mapsto \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega})$ of $V \times \mathbb{R}^r$ into $\mathbb{R}^n$ is C^q,
- For each $(\boldsymbol{x}, \boldsymbol{\omega}) \in V \times \mathbb{R}^r$, the map $t \mapsto \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega})$ of I into $\mathbb{R}^n$ is measurable, and

- For each compact subset $K \subseteq V$, there exist

$$\alpha_k \in L^{\frac{p}{p-k}}_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0}), \qquad k \in \{0, 1, \ldots, \lfloor p \rfloor, p\},$$

  such that for each $0 \leq i, j \leq q$ satisfying $i + j \leq q$, we have

$$\|\boldsymbol{D}_2^i \boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega})\| \leq \sum_{k \in \{0,1,\ldots,\lfloor p \rfloor - j, p-j\}} \alpha_{j+k}(t) \|\boldsymbol{\omega}\|^k_{\mathbb{R}^r} \tag{3.2}$$

  for a.a. $t \in I$ and each $(\boldsymbol{x}, \boldsymbol{\omega}) \in K \times \mathbb{R}^r$.

Of course, if $\boldsymbol{f}$ is C^q_p, then $\boldsymbol{f}$ is C^j_p for each $1 \leq j \leq q$.

Remark 3.2.2. Suppose that $\boldsymbol{f}$ is C^q_p and $0 \leq i, j \leq q$ are such that $i + j \leq q$. For each $t \in I$, the map $(\boldsymbol{x}, \boldsymbol{\omega}) \mapsto \boldsymbol{D}_2^i \boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega})$ is continuous by construction. Arguing as in Remark 2.2.20, one can show that for each $(\boldsymbol{x}, \boldsymbol{\omega}) \in V \times \mathbb{R}^r$, the map $t \mapsto \boldsymbol{D}_2^i \boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega})$ is measurable. By Remark 2.2.16, if

$$\boldsymbol{\nu} : \mathrm{dom}(\boldsymbol{\nu}) \to V \times \mathbb{R}^r$$

is measurable, where $\mathrm{dom}(\boldsymbol{\nu})$ is a subinterval of I, then the map

$$t \mapsto \boldsymbol{D}_2^i \boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{\nu}(t))$$

is measurable. We will use this fact in Chapters 6 and 7.

Definition 3.2.1 is complicated by the fact that p may not be an integer. When $p \in \mathbb{N}$, we have $\lfloor p \rfloor - j = p - j$ and (3.2) becomes

$$\|\boldsymbol{D}_2^i \boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega})\| \leq \sum_{k=0}^{p-j} \alpha_{j+k}(t) \|\boldsymbol{\omega}\|^k_{\mathbb{R}^r}.$$

This occurs, for example, when $\boldsymbol{f}$ is C^q_q-polynomial in the following sense.

Definition 3.2.3. We say that $\boldsymbol{f}$ is ***C^q_q-polynomial*** if there exist maps

$$\boldsymbol{P}_k : I \times V \to \mathrm{Sym}^k(\mathbb{R}^r, \mathbb{R}^n), \qquad 0 \leq k \leq q,$$

such that

- For each $t \in I$, the map $\boldsymbol{x} \mapsto \boldsymbol{P}_k(t, \boldsymbol{x})$ of V into $\mathrm{Sym}^k(\mathbb{R}^r, \mathbb{R}^n)$ is C^q,
- For each $\boldsymbol{x} \in V$, the map $t \mapsto \boldsymbol{P}_k(t, \boldsymbol{x})$ of I into $\mathrm{Sym}^k(\mathbb{R}^r, \mathbb{R}^n)$ is measurable,
- For each compact subset $K \subseteq V$, there exist

$$\alpha_k \in L^{\frac{q}{q-k}}_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0}), \qquad 0 \leq k \leq q,$$

  such that for each $0 \leq i \leq q$, we have

$$\|\boldsymbol{D}_2^i \boldsymbol{P}_k(t, \boldsymbol{x})\| \leq \alpha_k(t) \tag{3.3}$$

  for a.a. $t \in I$ and each $\boldsymbol{x} \in K$, and
- $\boldsymbol{f}$ can be written as

$$\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega}) = \sum_{k=0}^{q} \boldsymbol{P}_k(t, \boldsymbol{x}) \cdot \underbrace{(\boldsymbol{\omega}, \ldots, \boldsymbol{\omega})}_{k \text{ copies of } \boldsymbol{\omega}}. \tag{3.4}$$

Next, we show that $\boldsymbol{f}$ is C^q_q-polynomial if it has a C^q control-affine form.

Example 3.2.4. Suppose that $\boldsymbol{f}_0, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_r : V \to \mathbb{R}^n$ are C^q and

$$\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega}) = \boldsymbol{f}_0(\boldsymbol{x}) + \sum_{i=1}^{r} \omega^i \boldsymbol{f}_i(\boldsymbol{x}).$$

Define maps $\boldsymbol{P}_k : I \times V \to \mathrm{Sym}^k(\mathbb{R}^r, \mathbb{R}^n)$, where $0 \leq k \leq q$, by setting

- $\boldsymbol{P}_0(t, \boldsymbol{x}) = \boldsymbol{f}_0(\boldsymbol{x})$,
- $\boldsymbol{P}_1(t, \boldsymbol{x}) \cdot \boldsymbol{\omega} = \sum_{i=1}^{r} \omega^i \boldsymbol{f}_i(\boldsymbol{x})$, and
- $\boldsymbol{P}_k$ to be identically equal to zero for $2 \leq k \leq q$.

We have

$$\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega}) = \boldsymbol{P}_0(t, \boldsymbol{x}) + \boldsymbol{P}_1(t, \boldsymbol{x}) \cdot \boldsymbol{\omega}$$

and a straightforward verification shows that $\boldsymbol{f}$ is C^q_q-polynomial.

The next lemma appears, without proof, in [Margheri 1996]. For the sake of completeness, we provide a full proof here.

Lemma 3.2.5. *Suppose that $\boldsymbol{f}$ is C^q_q-polynomial. Then $\boldsymbol{f}$ is C^q_q.*

Proof. Observe that

$$\boldsymbol{f}(t,\boldsymbol{x},\boldsymbol{\omega}) = \sum_{k=0}^{q} \mathrm{ev}_k \cdot (\boldsymbol{P}_k(t,\boldsymbol{x}), \boldsymbol{\omega}, \dots, \boldsymbol{\omega}),$$

where ev_k is the evaluation map defined at the beginning of this chapter. For each $t \in I$, the maps $\boldsymbol{x} \mapsto \boldsymbol{P}_k(t,\boldsymbol{x})$ are C^q by definition. It follows from the Leibniz rule that for each $t \in I$, the map

$$(\boldsymbol{x},\boldsymbol{\omega}) \mapsto \boldsymbol{f}(t,\boldsymbol{x},\boldsymbol{\omega})$$

is C^q. Similarly, for each $\boldsymbol{x} \in V$, the maps $t \mapsto \boldsymbol{P}_k(t,\boldsymbol{x})$ are measurable by definition. Thus for each $(\boldsymbol{x},\boldsymbol{\omega}) \in V \in \mathbb{R}^r$, the map

$$t \mapsto \boldsymbol{f}(t,\boldsymbol{x},\boldsymbol{\omega})$$

is measurable by composition. We now show that the third criterion of Definition 3.2.1 is satisfied. To this end, let $0 \leq j,k \leq q$. By the Leibniz rule, we have

$$\begin{aligned}
&\boldsymbol{D}^j \left[\boldsymbol{\omega} \mapsto \mathrm{ev}_k \cdot (\boldsymbol{P}_k(t,\boldsymbol{x}), \boldsymbol{\omega}, \dots, \boldsymbol{\omega})\right](\boldsymbol{\omega}) \cdot (\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_j) \\
&= \sum_{(N_1,\dots,N_{k+1}) \in P_{k+1}(j)} \mathrm{ev}_k \cdot (\boldsymbol{D}^{\mathrm{card}(N_1)}[\boldsymbol{\omega} \mapsto \boldsymbol{P}_k(t,\boldsymbol{x})](\boldsymbol{\omega}) \cdot \boldsymbol{\omega}_{N_1}, \\
&\qquad\qquad \boldsymbol{D}^{\mathrm{card}(N_2)} \mathrm{id}_{\mathbb{R}^r}(\boldsymbol{\omega}) \cdot \boldsymbol{\omega}_{N_2}, \dots, \boldsymbol{D}^{\mathrm{card}(N_{k+1})} \mathrm{id}_{\mathbb{R}^r}(\boldsymbol{\omega}) \cdot \boldsymbol{\omega}_{N_{k+1}}).
\end{aligned}$$

If the $(k+1)$-tuple $(N_1, \dots, N_{k+1})$ is such that

- $\mathrm{card}(N_1) \geq 1$ or
- $\mathrm{card}(N_i) \geq 2$ for $2 \leq i \leq k+1$,

then the corresponding term in the above sum is $\boldsymbol{0}_{\mathbb{R}^n}$. Thus we can restrict our attention to $(k+1)$-tuples such that

- $\operatorname{card}(N_1) = 0$ and
- $0 \leq \operatorname{card}(N_i) \leq 1$ for $2 \leq i \leq k+1$.

If $k < j$, then there are no such $(k+1)$-tuples. On the other hand, if $k \geq j$, then there are $\binom{k}{j}$ such $(k+1)$-tuples, each of which is equal to

$$\mathrm{ev}_k \cdot (\boldsymbol{P}_k(t, \boldsymbol{x}), \boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_j, \boldsymbol{\omega}, \ldots, \boldsymbol{\omega})$$

by the symmetry of $\boldsymbol{P}_k(t, \boldsymbol{x})$. Thus

$$\boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega}) \cdot (\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_j) = \sum_{k=j}^{q} \binom{k}{j} \mathrm{ev}_k \cdot (\boldsymbol{P}_k(t, \boldsymbol{x}), \boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_j, \boldsymbol{\omega}, \ldots, \boldsymbol{\omega}).$$

Now let $0 \leq i, j \leq q$ satisfy $i + j \leq q$. By the Leibniz rule,

$$\begin{aligned}
&\boldsymbol{D}_2^i \boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega}) \cdot (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i, \boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_j) \\
&= \sum_{k=j}^{q} \binom{k}{j} \mathrm{ev}_k \cdot (\boldsymbol{D}_2^i \boldsymbol{P}_k(t, \boldsymbol{x}) \cdot (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i), \boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_j, \boldsymbol{\omega}, \ldots, \boldsymbol{\omega}) \\
&= \sum_{k=j}^{q} \binom{k}{j} (\boldsymbol{D}_2^i \boldsymbol{P}_k(t, \boldsymbol{x}) \cdot (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i)) \cdot (\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_j, \boldsymbol{\omega}, \ldots, \boldsymbol{\omega})
\end{aligned}$$

and consequently

$$\|\boldsymbol{D}_2^i \boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega})\| \leq \sum_{k=j}^{q} \binom{k}{j} \|\boldsymbol{D}_2^i \boldsymbol{P}_k(t, \boldsymbol{x})\| \, \|\boldsymbol{\omega}\|_{\mathbb{R}^r}^{k-j}. \tag{3.5}$$

To complete the proof, choose a compact subset $K \subseteq V$, and let the functions α_k be prescribed as in Definition 3.2.3. Using (3.5), we have

$$\|\boldsymbol{D}_2^i \boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega})\| \leq \sum_{k=j}^{q} \binom{k}{j} \alpha_k(t) \, \|\boldsymbol{\omega}\|_{\mathbb{R}^r}^{k-j} = \sum_{k=0}^{q-j} \binom{j+k}{j} \alpha_{j+k}(t) \, \|\boldsymbol{\omega}\|_{\mathbb{R}^r}^{k}$$

for a.a. $t \in I$ and each $(\boldsymbol{x}, \boldsymbol{\omega}) \in K \times \mathbb{R}^r$. This completes the proof. ∎

The next lemma connects $\boldsymbol{f}$ being C_p^q to C^q $\boldsymbol{f}$-admissibility.

Lemma 3.2.6. *Suppose that $\boldsymbol{f}$ is C_p^q and*

$$\boldsymbol{u} \in L_{\mathrm{loc}}^p(I, \mathbb{R}^r).$$

Then $\boldsymbol{u}$ is C^q $\boldsymbol{f}$-admissible.

Proof. Suppose that $\tilde{\boldsymbol{u}}$ is a representative of $\boldsymbol{u}$. We must show that $\boldsymbol{f}^{\tilde{\boldsymbol{u}}}$ is locally integrably C^q in the sense of Definition 2.2.19. For each $t \in I$ and each $\boldsymbol{\omega} \in \mathbb{R}^r$, the map $\boldsymbol{x} \mapsto \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega})$ is C^q by definition. It follows that for each $t \in I$, the map

$$\boldsymbol{x} \mapsto \boldsymbol{f}^{\tilde{\boldsymbol{u}}}(t, \boldsymbol{x})$$

is C^q. Similarly, for each $\boldsymbol{x} \in V$, the map

$$t \mapsto \boldsymbol{f}(t, \boldsymbol{x}, \tilde{\boldsymbol{u}}(t))$$

is measurable by composition. Indeed, it can be written as $\boldsymbol{f} \circ (\mathrm{id}_I \times \kappa_{\boldsymbol{x}} \times \tilde{\boldsymbol{u}})$, where $\kappa_{\boldsymbol{x}} : I \to V$ denotes the constant map with value $\boldsymbol{x}$. We now show that the third criterion of Definition 2.2.19 is satisfied. Choose a compact subset $K \subseteq V$, and let the functions α_k be prescribed as in Definition 3.2.1. Recall that

$$\alpha_k \in L_{\mathrm{loc}}^{\frac{p}{p-k}}(I, \mathbb{R}_{\geq 0}), \qquad k \in \{0, 1, \ldots, \lfloor p \rfloor, p\}.$$

Choose $0 \leq i \leq q$. We have

$$\|\boldsymbol{D}_2^i \boldsymbol{f}^{\tilde{\boldsymbol{u}}}(t, \boldsymbol{x})\| = \|\boldsymbol{D}_2^i \boldsymbol{f}(t, \boldsymbol{x}, \tilde{\boldsymbol{u}}(t))\| \leq \sum_{k \in \{0,1,\ldots,\lfloor p \rfloor, p\}} \alpha_k(t) \|\tilde{\boldsymbol{u}}(t)\|_{\mathbb{R}^r}^k$$

for a.a. $t \in I$ and each $\boldsymbol{x} \in K$. We now examine the terms in the above sum. Choose $k \in \{0, 1, \ldots, \lfloor p \rfloor, p\}$. We claim that the function

$$t \mapsto \alpha_k(t) \|\tilde{\boldsymbol{u}}(t)\|_{\mathbb{R}^r}^k$$

is an element of $L_{\mathrm{loc}}^1(I, \mathbb{R}_{\geq 0})$. To see this, observe that

$$\alpha_k \in L_{\mathrm{loc}}^{\frac{p}{p-k}}(I, \mathbb{R}_{\geq 0}) \qquad \text{and} \qquad \|\tilde{\boldsymbol{u}}\|_{\mathbb{R}^r}^k \in L_{\mathrm{loc}}^{\frac{p}{k}}(I, \mathbb{R}_{\geq 0}).$$

Since

$$\frac{p-k}{p} + \frac{k}{p} = \frac{p}{p} = 1,$$

the claim follows from Hölder's inequality. Thus $\alpha : I \to \mathbb{R}$ defined by

$$\alpha(t) = \sum_{k \in \{0,1,\dots,\lfloor p \rfloor, p\}} \alpha_k(t) \|\tilde{\boldsymbol{u}}(t)\|_{\mathbb{R}^r}^k$$

is an element of $L^1_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0})$. We have shown that

$$\|\boldsymbol{D}_2^i \boldsymbol{f}^{\tilde{\boldsymbol{u}}}(t, \boldsymbol{x})\| \leq \alpha(t)$$

for a.a. $t \in I$ and each $\boldsymbol{x} \in K$. This completes the proof. ∎

Definition 3.2.7. We say that Σ is

- C^q_p, if $\boldsymbol{f}$ is C^q_p and Σ uses L^p controls, and
- ***C^q_q-polynomial***, if $\boldsymbol{f}$ is C^q_q-polynomial and Σ uses L^q controls.

Lemma 3.2.8. *Suppose that Σ is C^q_p. Then Σ is a C^q control system. In particular, if Σ is C^q_q-polynomial, then it is a C^q control system.*

Proof. This follows immediately from Lemmas 3.2.5 and 3.2.6. ∎

The importance of the next definition will be made clear below.

Definition 3.2.9. Suppose that Σ is C^q_p. We say that Σ is ***nice*** if

- $p > q$ or
- Σ is C^q_q-polynomial.

Example 3.2.10. Suppose that Σ is a C^q control-affine system. By Example 3.2.4, f is C^q_q-polynomial. We conclude that Σ is nice whenever Σ uses L^q controls.

3.2.2 A special property of total derivatives

In this section, we prove a technical lemma. This lemma will come into play in Chapters 6, 7, and 8, in connection with the first and second variations of Σ.

Lemma 3.2.11. *Suppose that*

- Σ *is* C^q_p,
- $0 \leq j \leq q$,
- $\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_j \in \mathscr{U} = L^p_{\mathrm{loc}}(I, \mathbb{R}^r)$, *and*
- $\tilde{\boldsymbol{u}}_0, \tilde{\boldsymbol{u}}_1, \ldots, \tilde{\boldsymbol{u}}_j$ *are representatives of* $\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_j$, *respectively.*

Define $\boldsymbol{g} : I \times V \to \mathbb{R}^n$ *by*

$$\boldsymbol{g}(t, \boldsymbol{x}) = \boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \tilde{\boldsymbol{u}}_0(t)) \cdot (\tilde{\boldsymbol{u}}_1(t), \ldots, \tilde{\boldsymbol{u}}_j(t)).$$

Then $\boldsymbol{g}$ *is locally integrably* C^{q-j}.

Proof. Observe that

$$\boldsymbol{g}(t, \boldsymbol{x}) = \mathrm{ev}_j(\boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \tilde{\boldsymbol{u}}_0(t)), \tilde{\boldsymbol{u}}_1(t), \ldots, \tilde{\boldsymbol{u}}_j(t)),$$

where ev_j is the evaluation map defined at the beginning of this chapter. For each $t \in I$ and each $\boldsymbol{\omega} \in \mathbb{R}^r$, the map $\boldsymbol{x} \mapsto \boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega})$ is C^{q-j} by Proposition 2.1.1. It follows from the Leibniz rule that for each $t \in I$, the map

$$\boldsymbol{x} \mapsto \boldsymbol{g}(t, \boldsymbol{x})$$

is C^{q-j}. Similarly, for each $\boldsymbol{x} \in V$, the map

$$t \mapsto \boldsymbol{g}(t, \boldsymbol{x})$$

is measurable by composition. Indeed, it can be written as

$$\mathrm{ev}_j(\boldsymbol{D}_3^j \boldsymbol{f} \circ (\mathrm{id}_I \times \kappa_{\boldsymbol{x}} \times \tilde{\boldsymbol{u}}_0), \tilde{\boldsymbol{u}}_1, \ldots, \tilde{\boldsymbol{u}}_j),$$

where $\kappa_{\boldsymbol{x}} : I \to V$ denotes the constant map with value $\boldsymbol{x}$. We now show that the third criterion of Definition 2.2.19 is satisfied. Choose a compact subset $K \subseteq V$, and let the functions α_k be prescribed as in Definition 3.2.1. Recall that

$$\alpha_k \in L^{\frac{p}{p-k}}_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0}), \qquad k \in \{0, 1, \ldots, \lfloor p \rfloor, p\}.$$

Choose $0 \le i \le q - j$. We have

$$\begin{aligned}
\|\boldsymbol{D}_2^i \boldsymbol{g}(t, \boldsymbol{x})\| &= \|\boldsymbol{D}_2^i \boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \tilde{\boldsymbol{u}}_0(t)) \cdot (\tilde{\boldsymbol{u}}_1(t), \dots, \tilde{\boldsymbol{u}}_j(t))\| \\
&\le \|\boldsymbol{D}_2^i \boldsymbol{D}_3^j \boldsymbol{f}(t, \boldsymbol{x}, \tilde{\boldsymbol{u}}_0(t))\| \, \|\tilde{\boldsymbol{u}}_1(t)\|_{\mathbb{R}^r} \cdots \|\tilde{\boldsymbol{u}}_j(t)\|_{\mathbb{R}^r} \\
&\le \sum_{k \in \{0,1,\dots,\lfloor p \rfloor - j, p-j\}} \alpha_{j+k}(t) \|\tilde{\boldsymbol{u}}_0(t)\|_{\mathbb{R}^r}^k \|\tilde{\boldsymbol{u}}_1(t)\|_{\mathbb{R}^r} \cdots \|\tilde{\boldsymbol{u}}_j(t)\|_{\mathbb{R}^r}
\end{aligned}$$

for a.a. $t \in I$ and each $\boldsymbol{x} \in K$. We now examine the terms in the above sum. Choose $k \in \{0, 1, \dots, \lfloor p \rfloor - j, p - j\}$. We claim that the function

$$t \mapsto \alpha_{j+k}(t) \|\tilde{\boldsymbol{u}}_0(t)\|_{\mathbb{R}^r}^k \, \|\tilde{\boldsymbol{u}}_1(t)\|_{\mathbb{R}^r} \cdots \|\tilde{\boldsymbol{u}}_j(t)\|_{\mathbb{R}^r}$$

is an element of $L^1_{\mathrm{loc}}(I, \mathbb{R}_{\ge 0})$. To see this, observe that

$$\alpha_{j+k} \in L_{\mathrm{loc}}^{\frac{p}{p-(j+k)}}(I, \mathbb{R}_{\ge 0}), \quad \|\tilde{\boldsymbol{u}}_0\|_{\mathbb{R}^r}^k \in L_{\mathrm{loc}}^{\frac{p}{k}}(I, \mathbb{R}_{\ge 0}), \quad \text{and} \quad \|\tilde{\boldsymbol{u}}_i\|_{\mathbb{R}^r} \in L_{\mathrm{loc}}^{p}(I, \mathbb{R}_{\ge 0}).$$

Since

$$\frac{p-(j+k)}{p} + \frac{k}{p} + \sum_{\ell=1}^{j} \frac{1}{p} = \frac{p-(j+k)}{p} + \frac{k}{p} + \frac{j}{p} = 1,$$

the claim follows from Hölder's inequality. Thus $\alpha : I \to \mathbb{R}$ defined by

$$\alpha(t) = \sum_{k \in \{0,1,\dots,\lfloor p \rfloor - j, p-j\}} \alpha_{j+k}(t) \|\tilde{\boldsymbol{u}}_0(t)\|_{\mathbb{R}^r}^k \, \|\tilde{\boldsymbol{u}}_1(t)\|_{\mathbb{R}^r} \cdots \|\tilde{\boldsymbol{u}}_j(t)\|_{\mathbb{R}^r}$$

is an element of $L^1_{\mathrm{loc}}(I, \mathbb{R}_{\ge 0})$. We have shown that

$$\|\boldsymbol{D}_2^i \boldsymbol{g}(t, \boldsymbol{x})\| \le \alpha(t)$$

for a.a. $t \in I$ and each $\boldsymbol{x} \in K$. This completes the proof. ∎

3.2.3 Differentiability

In this section, we recall two differentiability results.

Theorem 3.2.12. *Suppose that Σ is a nice C^q_p control system. Then*

1. $\mathrm{dom}(\Phi^{\Sigma})$ *is an open subset of* $I \times I \times V \times L^p_{\mathrm{loc}}(I, \mathbb{R}^r)$,

2. Φ^{Σ} *is continuous, and*

3. $\mathrm{End}^{\Sigma|[t_0,t]}$ *is* C^q *for each* $(t, t_0) \in I \times I$ *such that* $t_0 \leq t$.

Proof. This result is a composite of [Bianchini and Margheri 1996, Theorem 3.2] and [Margheri 1996, Theorem 3.2]. ∎

Corollary 3.2.13. *Suppose that* Σ *is a nice* C^q_p *control system and* $\boldsymbol{x}_0 \in V$*. Then the map* $\mathrm{End}^{\Sigma|[t_0,t]}_{\boldsymbol{x}_0}$ *is* C^q *for each* $(t, t_0) \in I \times I$ *such that* $t_0 \leq t$.

Remark 3.2.14. To prove the third conclusion of Theorem 3.2.12, Margheri [1996] used a converse of Taylor's theorem. The argument requires that the mixed partial derivatives of $\boldsymbol{f}$ satisfy local growth conditions, and this explains the prominent role of such conditions in Definition 3.2.1. The extent to which these conditions are necessary is also explored by Margheri [1996].

Remark 3.2.15. By definition, nice C^q_p control systems use L^p controls, where

$$p \neq \infty.$$

For control systems that use L^∞ controls, an approximate analogue of Theorem 3.2.12 can be found in [Grasse 1979]. Here, one replaces the assumption that Σ is C^q_p by the assumption that Σ is quasi-C^q. Interestingly, the results of Grasse [1979] concerning quasi-C^q control systems subsume differentiability results derived independently by a number of authors, including Lee and Markus [1986], Bonnard and Kupka [1993], and Sontag [1998]. This fact is not widely recognized.

3.3 C^q_p and C^q_q-polynomial control systems evolving on finite-dimensional manifolds

In this section, we extend the theory of C^q_p and C^q_q-polynomial control systems to accommodate control systems evolving on finite-dimensional manifolds.

Throughout this section,

- $\Sigma = (f, \mathscr{U})$ is a control system evolving on M,
- $p \in \mathbb{R}_{\geq 1}$ and $q \in \mathbb{N}$ are fixed, subject to the requirement that $p \geq q$.

3.3.1 Basic definitions and properties

Definition 3.3.1. We say that f is C^q_p if $f_{\boldsymbol{\psi}}$ is C^q_p for each chart $(V, \boldsymbol{\psi})$ on M.

One can show that f is C^q_p if and only if $f_{\boldsymbol{\psi}}$ is C^q_p for each chart $(V, \boldsymbol{\psi}) \in \mathscr{A}_M$, where $\mathscr{A}_M$ is a compatible atlas on M.

Remark 3.3.2. Suppose that f is C^q_p. By construction, for each $t \in I$ the map

$$(x, \boldsymbol{\omega}) \mapsto f(t, x, \boldsymbol{\omega})$$

of $M \times \mathbb{R}^r$ into TM is C^q.

Definition 3.3.3. We say that f is ***C^q_q-polynomial*** if $f_{\boldsymbol{\psi}}$ is C^q_q-polynomial for each chart $(V, \boldsymbol{\psi})$ on M.

Again, one can show that f is C^q_q-polynomial if and only if $f_{\boldsymbol{\psi}}$ is C^q_q-polynomial for each chart $(V, \boldsymbol{\psi}) \in \mathscr{A}_M$, where $\mathscr{A}_M$ is a compatible atlas on M.

Example 3.3.4. Suppose that Σ is a C^q control-affine system,

$$f(t, x, \boldsymbol{\omega}) = f_0(x) + \sum_{i=1}^{r} \omega^i f_i(x),$$

and $(V, \boldsymbol{\psi})$ is a chart on M. Observe that

$$f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{\omega}) = (f_0)_{\boldsymbol{\psi}}(\boldsymbol{x}) + \sum_{i=1}^{r} \omega^i (f_i)_{\boldsymbol{\psi}}(\boldsymbol{x}).$$

By Example 3.2.4, $f_{\boldsymbol{\psi}}$ is C^q_q-polynomial, and consequently f is C^q_q-polynomial.

Definition 3.3.5. We say that Σ is

- C^q_p, if f is C^q_p and Σ uses L^p controls, and

- C^q_q***-polynomial***, if f is C^q_q-polynomial and Σ uses L^q controls.

Lemma 3.3.6. *Suppose that Σ is C^q_p. Then Σ is a C^q control system. In particular, if Σ is C^q_q-polynomial, then it is a C^q control system.*

Proof. This follows immediately from Lemma 3.2.8. ∎

Definition 3.3.7. Suppose that Σ is C^q_p. We say that Σ is ***nice*** if

- $p > q$ or

- Σ is C^q_q-polynomial.

Example 3.3.8. Suppose that Σ is a C^q control-affine system. By Example 3.3.4, f is C^q_q-polynomial. We conclude that Σ is nice whenever Σ uses L^q controls.

3.3.2 A special property of total derivatives

In this section, we extend Lemma 3.2.11.

Lemma 3.3.9. *Suppose that*

- *Σ is C^q_p,*

- *$0 \leq j \leq q$,*

- *$\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_j \in \mathscr{U} = L^p_{\mathrm{loc}}(I, \mathbb{R}^r)$, and*

- *$\tilde{\boldsymbol{u}}_0, \tilde{\boldsymbol{u}}_1, \ldots, \tilde{\boldsymbol{u}}_j$ are representatives of $\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_j$, respectively.*

Define $g : I \times M \to TM$ by

$$g(t, x) = \boldsymbol{D}^j_3 f(t, x, \tilde{\boldsymbol{u}}_0(t)) \cdot (\tilde{\boldsymbol{u}}_1(t), \ldots, \tilde{\boldsymbol{u}}_j(t)).$$

Then g is locally integrably C^{q-j}. Note that g is a well-defined time-varying vector field on M, since $\boldsymbol{D}^j_3 f(t, x, \tilde{\boldsymbol{u}}_0(t)) \in \mathrm{Hom}^j(\mathbb{R}^r, T_xM)$ by definition.

Proof. Suppose that $(V, \boldsymbol{\psi})$ is a chart on M. We must show that $g_{\boldsymbol{\psi}}$ is locally integrably C^{q-j} in the sense of Definition 2.2.19. By the Leibniz rule, we have

$$\begin{aligned} g_{\boldsymbol{\psi}}(t, \boldsymbol{x}) &= T\boldsymbol{\psi}(\boldsymbol{\psi}^{-1}(\boldsymbol{x})) \circ \boldsymbol{D}_3^j f(t, \boldsymbol{\psi}^{-1}(\boldsymbol{x}), \tilde{\boldsymbol{u}}_0(t)) \cdot (\tilde{\boldsymbol{u}}_1(t), \dots, \tilde{\boldsymbol{u}}_j(t)) \\ &= \boldsymbol{D}^j[\boldsymbol{\omega} \mapsto T\boldsymbol{\psi}(\boldsymbol{\psi}^{-1}(\boldsymbol{x})) \circ f(t, \boldsymbol{\psi}^{-1}(\boldsymbol{x}), \boldsymbol{\omega})](\tilde{\boldsymbol{u}}_0(t)) \cdot (\tilde{\boldsymbol{u}}_1(t), \dots, \tilde{\boldsymbol{u}}_j(t)) \\ &= \boldsymbol{D}_3^j f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \tilde{\boldsymbol{u}}_0(t)) \cdot (\tilde{\boldsymbol{u}}_1(t), \dots, \tilde{\boldsymbol{u}}_j(t)) \end{aligned}$$

for each $(t, \boldsymbol{x}) \in I \times \boldsymbol{\psi}(V)$. By Lemma 3.2.11, $g_{\boldsymbol{\psi}}$ is locally integrably C^{q-j}. ∎

3.3.3 Differentiability

In this section, we extend Theorem 3.2.12. To do so, we need the following lemma.

Lemma 3.3.10. *Suppose that Σ is a nice C_p^q control system and*

$$(t_*, x_*, \boldsymbol{u}_*) \in I \times M \times L_{\mathrm{loc}}^p(I, \mathbb{R}^r).$$

Then there exists a product neighbourhood

$$I_* \times I_* \times M_* \times \mathscr{U}_*$$

of $(t_*, t_*, x_*, \boldsymbol{u}_*)$ in $I \times I \times M \times L_{\mathrm{loc}}^p(I, \mathbb{R}^r)$ such that

1. $I_* \times I_* \times M_* \times \mathscr{U}_* \subseteq \mathrm{dom}(\Phi^\Sigma)$,
2. $\Phi^\Sigma | I_* \times I_* \times M_* \times \mathscr{U}_*$ *is continuous, and*
3. $\mathrm{End}^{\Sigma|[t_0,t]} | M_* \times \tilde{\mathscr{U}}_*$ *is C^q for each $(t, t_0) \in I_* \times I_*$ such that $t_0 \leq t$, where*

$$\tilde{\mathscr{U}}_* = \pi_{[t_0,t]}^p(\mathscr{U}_*).$$

Note that the third conclusion is well-defined by Remark 3.1.12.

Proof. Suppose that $(V, \boldsymbol{\psi})$ is a chart on M such that $x_* \in V$. By Theorem 3.2.12,

1. $\mathrm{dom}(\Phi^{\Sigma_{\boldsymbol{\psi}}})$ is an open subset of $I \times I \times \boldsymbol{\psi}(V) \times L_{\mathrm{loc}}^p(I, \mathbb{R}^r)$,

2. $\Phi^{\Sigma_{\psi}}$ is continuous, and

3. $\mathrm{End}^{\Sigma_{\psi}|[t_0,t]}$ is C^q for each $(t,t_0) \in I \times I$ such that $t_0 \leq t$.

In particular, since $(t_*, t_*, \boldsymbol{\psi}(x_*), \boldsymbol{u}_*) \in \mathrm{dom}(\Phi^{\Sigma_{\psi}})$, there exists a product neighbourhood $I_* \times I_* \times V_* \times \mathscr{U}_*$ of $(t_*, t_*, \boldsymbol{\psi}(x_*), \boldsymbol{u}_*)$ in

$$I \times I \times \boldsymbol{\psi}(V) \times L^p_{\mathrm{loc}}(I, \mathbb{R}^r)$$

such that

$$I_* \times I_* \times V_* \times \mathscr{U}_* \subseteq \mathrm{dom}(\Phi^{\Sigma_{\psi}}).$$

Setting $M_* = \boldsymbol{\psi}^{-1}(V_*)$, we see that $I_* \times I_* \times M_* \times \mathscr{U}_*$ is the desired product neighbourhood of $(t_*, t_*, x_*, \boldsymbol{u}_*)$. Indeed, Lemma 3.1.13 tells us that

1. $I_* \times I_* \times M_* \times \mathscr{U}_* \subseteq \mathrm{dom}(\Phi^{\Sigma})$,

2. $\Phi^{\Sigma}|I_* \times I_* \times M_* \times \mathscr{U}_* = \boldsymbol{\psi}^{-1} \circ \Phi^{\Sigma_{\psi}} \circ (\mathrm{id}_{I_*} \times \mathrm{id}_{I_*} \times \boldsymbol{\psi}|M_* \times \mathrm{id}_{\mathscr{U}_*})$, and

3. We have

$$\mathrm{End}^{\Sigma|[t_0,t]}|M_* \times \tilde{\mathscr{U}}_* = \boldsymbol{\psi}^{-1} \circ \mathrm{End}^{\Sigma_{\psi}|[t_0,t]} \circ (\boldsymbol{\psi}|M_* \times \mathrm{id}_{\tilde{\mathscr{U}}_*})$$

   for each $(t,t_0) \in I_* \times I_*$ such that $t_0 \leq t$, where

$$\tilde{\mathscr{U}}_* = \pi^p_{[t_0,t]}(\mathscr{U}_*).$$

The conclusions of the lemma follow by composition. ∎

Theorem 3.3.11. *Suppose that Σ is a nice C^q_p control system. Then*

1. $\mathrm{dom}(\Phi^{\Sigma})$ *is an open subset of* $I \times I \times M \times L^p_{\mathrm{loc}}(I, \mathbb{R}^r)$*,*

2. Φ^{Σ} *is continuous, and*

3. $\mathrm{End}^{\Sigma|[t_0,t]}$ *is* C^q *for each* $(t,t_0) \in I \times I$ *such that* $t_0 \leq t$*.*

Proof. We follow the proof of Theorem 2.3.25. Given

$$(t_*, x_*, \boldsymbol{u}_*) \in I \times M \times L^p_{\mathrm{loc}}(I, \mathbb{R}^r),$$

we say that $s \in I^\Sigma(t_*, x_*, \boldsymbol{u}_*)$ has the C^q ***neighbourhood property*** (relative to the particular choice of $(t_*, x_*, \boldsymbol{u}_*)$) if there exists a product neighbourhood

$$I^1_s \times I^2_s \times M_s \times \mathscr{U}_s$$

of $(s, t_*, x_*, \boldsymbol{u}_*)$ in $I \times I \times M \times L^p_{\mathrm{loc}}(I, \mathbb{R}^r)$ such that

1. $I^1_s \times I^2_s \times M_s \times \mathscr{U}_s \subseteq \mathrm{dom}(\Phi^\Sigma)$,
2. $\Phi^\Sigma | I^1_s \times I^2_s \times M_s \times \mathscr{U}_s$ is continuous, and
3. $\mathrm{End}^{\Sigma|[t_0,t]} | M_s \times \tilde{\mathscr{U}}_s$ is C^q for each $(t, t_0) \in I^1_s \times I^2_s$ such that $t_0 \leq t$, where

$$\tilde{\mathscr{U}}_s = \pi^p_{[t_0,t]}(\mathscr{U}_s).$$

We say that such a neighbourhood is a C^q ***neighbourhood*** for s. The set of all $s \in I^\Sigma(t_*, x_*, \boldsymbol{u}_*)$ that have the C^q neighbourhood property is denoted by NP^q_*.

To complete the proof, it is enough to show that

$$\mathrm{NP}^q_* = I^\Sigma(t_*, x_*, \boldsymbol{u}_*)$$

for each $(t_*, x_*, \boldsymbol{u}_*) \in I \times M \times L^p_{\mathrm{loc}}(I, \mathbb{R}^r)$. To this end, choose

$$(t_*, x_*, \boldsymbol{u}_*) \in I \times M \times L^p_{\mathrm{loc}}(I, \mathbb{R}^r).$$

By Lemma 3.3.10, NP^q_* is nonempty since it contains t_*, and NP^q_* is open in $I^\Sigma(t_*, x_*, \boldsymbol{u}_*)$ by construction. Since $I^\Sigma(t_*, x_*, \boldsymbol{u}_*)$ is connected, it follows that

$$\begin{aligned} \mathrm{NP}^q_* = I^\Sigma(t_*, x_*, \boldsymbol{u}_*) &\iff \mathrm{NP}^q_* \text{ is closed in } I^\Sigma(t_*, x_*, \boldsymbol{u}_*) \\ &\iff \overline{\mathrm{NP}^q_*} = \mathrm{NP}^q_* \\ &\iff \overline{\mathrm{NP}^q_*} \subseteq \mathrm{NP}^q_*, \end{aligned} \tag{3.6}$$

where $\overline{\mathrm{NP}^q_*}$ denotes the closure of NP^q_* in $I^\Sigma(t_*, x_*, \boldsymbol{u}_*)$.

To prove (3.6), choose $s \in \overline{\mathrm{NP}^q_*}$. If $s = t_*$, then there is nothing to prove. Suppose that $s \neq t_*$. By Lemma 3.3.10, there exists a product neighbourhood

$$I_s \times I_s \times M_s \times \mathscr{U}_s$$

of $(s, s, \Phi^\Sigma_{s,t_*}(x_*, \boldsymbol{u}_*), \boldsymbol{u}_*)$ in $I \times I \times M \times L^p_{\mathrm{loc}}(I, \mathbb{R}^r)$ such that

1. $I_s \times I_s \times M_s \times \mathscr{U}_s \subseteq \mathrm{dom}(\Phi^\Sigma)$,
2. $\Phi^\Sigma | I_s \times I_s \times M_s \times \mathscr{U}_s$ is continuous, and
3. $\mathrm{End}^{\Sigma|[t_0,t]} | M_s \times \mathscr{U}_s$ is C^q for each $(t, t_0) \in I_s \times I_s$ such that $t_0 \leq t$, where

$$\tilde{\mathscr{U}_s} = \pi^p_{[t_0,t]}(\mathscr{U}_s).$$

Since s is a limit point of NP^q_* in $I^\Sigma(t_*, x_*, \boldsymbol{u}_*)$, it can be approached arbitrarily closely by elements of $I_s \cap \mathrm{NP}^q_*$. Thus there exists

$$\rho \in I_s \cap \mathrm{NP}^q_*$$

such that $\Phi^\Sigma_{\rho,t_*}(x_*, \boldsymbol{u}_*) \in M_s$. This is illustrated in Figure 3.1.

Since $\rho \in \mathrm{NP}^q_*$, there exists a C^q neighbourhood $I^1_\rho \times I^2_\rho \times M_\rho \times \mathscr{U}_\rho$ for ρ. Write

$$F^\Sigma_s = \Phi^\Sigma | I_s \times I_s \times M_s \times \mathscr{U}_s \qquad \text{and} \qquad G^\Sigma_\rho = \Phi^\Sigma | I^1_\rho \times I^2_\rho \times M_\rho \times \mathscr{U}_\rho.$$

Shrinking $I^1_\rho \times I^2_\rho \times M_\rho \times \mathscr{U}_\rho$, we can assume without loss of generality that

- $I^2_\rho \cap I_s = \varnothing$ (note that this is possible since we have assumed that $s \neq t_*$),
- $\mathrm{image}(G^\Sigma_\rho) \subseteq M_s$, and
- $\mathscr{U}_\rho \subseteq \mathscr{U}_s$.

We claim that $I_s \times I^2_\rho \times M_\rho \times \mathscr{U}_\rho$ is a C^q neighbourhood for s. Consider the map

$$H^\Sigma_s = \Phi^X | I_s \times I^2_\rho \times M_\rho \times \mathscr{U}_\rho.$$

Since

$$H_s^\Sigma(\beta, \beta_0, x, \boldsymbol{u}) = F_s^\Sigma(\beta, \rho, G_\rho^\Sigma(\rho, \beta_0, x, \boldsymbol{u}), \boldsymbol{u}),$$

the first two criteria of the C^q neighbourhood property are satisfied by composition. For the third criterion, choose $(\beta, \beta_0) \in I_s \times I_\rho^2$ such that $\beta_0 \leq \beta$. There are two cases to consider, depending on the precise relationship between β and ρ.

- Case 1: Suppose that $\beta_0 \leq \rho \leq \beta$. Since

  $$\mathrm{End}^{\Sigma|[\beta_0,\beta]}(x, \boldsymbol{u}) = \mathrm{End}^{\Sigma|[\rho,\beta]}(\mathrm{End}^{\Sigma|[\beta_0,\rho]}(x, \pi^p_{[\beta_0,\rho]}(\boldsymbol{u})), \pi^p_{[\rho,\beta]}(\boldsymbol{u}))$$

  for each $(x, \boldsymbol{u}) \in M_\rho \times \tilde{\mathscr{U}}_\rho$, where

  $$\tilde{\mathscr{U}}_\rho = \pi^p_{[\beta_0,\beta]}(\mathscr{U}_\rho),$$

  it follows that $\mathrm{End}^{\Sigma|[\beta_0,\beta]}|M_\rho \times \tilde{\mathscr{U}}_\rho$ is C^q by composition.

- Case 2: Suppose that $\beta_0 \leq \beta \leq \rho$. Shrinking I_s, we may assume without loss of generality that $\beta \in I_\rho^1$. Thus $\mathrm{End}^{\Sigma|[\beta_0,\beta]}|M_\rho \times \tilde{\mathscr{U}}_\rho$ is automatically C^q, where

  $$\tilde{\mathscr{U}}_\rho = \pi^p_{[\beta_0,\beta]}(\mathscr{U}_\rho).$$

This proves the claim, and the proof is complete. ∎

Corollary 3.3.12. *Suppose that Σ is a nice C^q_p control system and $x_0 \in M$. Then the map End^Σ is C^q for each $(t, t_0) \in I \times I$ such that $t_0 \leq t$.*

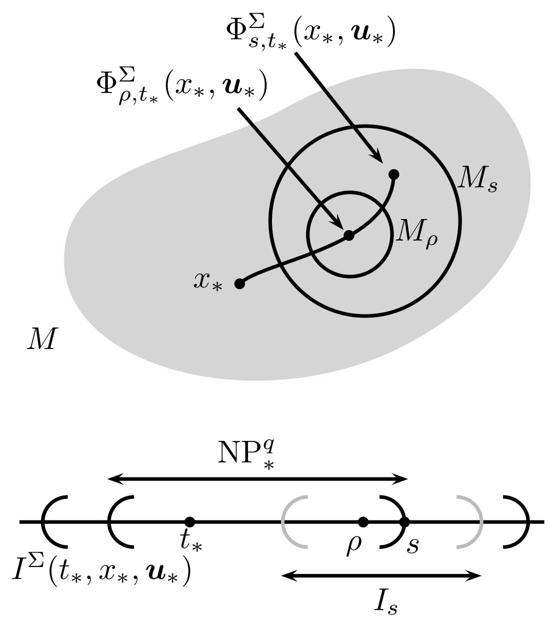

Figure 3.1: An illustration of the proof of Theorem 3.3.11

Chapter 4

The continuation method

Consider a control system

$$\Sigma = (f, \mathscr{U})$$

evolving on an n-dimensional manifold M, and let $x_0 \in M$. In Chapter 1, we indicated that the continuation method solves the x_0-anchored motion planning problem for Σ by lifting curves in M to curves in $\mathscr{U}$. In this chapter, we present the continuation method in full detail. In contrast to the "classical" continuation method of Sussmann [1993], the continuation method presented in this chapter does not rely fundamentally on Moore–Penroses pseudoinverses to lift curves.

This chapter is organized in the following way. In Section 4.1, we establish the theory of right inverses, which play the role of generalized Moore–Penrose pseudoinverses. In Section 4.2, we briefly review the required theory of initial value problems evolving on Banach manifolds. Finally, in Section 4.3, we present the continuation method. The presentation encompasses the general case and the simplified case where no controls are singular.

Our standing assumptions throughout this chapter are that

- Q is a Banach manifold modelled on a Banach space E_Q,

- R is a second-countable ℓ-dimensional manifold, and
- $F : Q \to R$ is a C^k submersion, where $k \in \mathbb{N}^*$.

The developments in this chapter are quite general, being phrased in terms of F. In Section 4.3, we specialize to control systems. This is accomplished by replacing the map F with the anchored endpoint map of a control system.

4.1 Right inverses

In this section, we establish the theory of right inverses, approximately following Earle and Eells [1967]. A right inverse is a special type of vector bundle map whose domain is the pullback of a vector bundle. Accordingly, we begin this section by recalling the definition of the pullback of a vector bundle by a C^k map. For all details concerning vector bundles, we refer to [Abraham et al. 1988].

Definition 4.1.1. Suppose that $\pi_E : E \to R$ is a vector bundle. Define

$$\pi_{F^*(E)} : F^*(E) \to Q$$

by $\pi_{F^*(E)}(q, e) = q$, where

$$F^*(E) = \{(q, e) \in Q \times E \ : \ F(q) = \pi_E(e)\}.$$

Then $\pi_{F^*(E)}$ is a C^k vector bundle over Q, called the ***pullback of*** E ***by*** F.

For each $q \in Q$, we have a canonical vector space isomorphism

$$\pi^{-1}_{F^*(E)}(q) \cong \pi_E^{-1}(F(q)).$$

These isomorphisms will be used implicitly in the remainder of this chapter.

As a special case, the tangent bundle $\pi_{TR} : TR \to R$ is a vector bundle over R, and thus $F^*(TR)$ is a C^k vector bundle over Q. Since $\pi^{-1}_{F^*(TR)}(q) \cong T_{F(q)}R$ for each

$q \in Q$, we can regard TF as a C^{k-1} vector bundle map

$$TF : TQ \to F^*(TR)$$

over id_Q. Recall from [Abraham et al. 1988, Theorem 3.5.18] that

$$\ker(TF) = \bigcup_{q \in Q} \ker(TF(q))$$

is a C^{k-1} vector bundle over Q. Since F is a submersion,

$$0 \to \ker(TF) \to TQ \xrightarrow{TF} F^*(TR) \to 0 \tag{4.1}$$

is a short exact sequence of C^{k-1} vector bundle maps over id_Q.

Definition 4.1.2. A ***right inverse*** of TF is a vector bundle map

$$TF^\dagger : F^*(TR) \to TQ$$

over id_Q such that $TF(q) \circ TF^\dagger(q) = \mathrm{id}_{T_{F(q)}R}$ for each $q \in Q$. Using more algebraic language, a right inverse of TF is a splitting of (4.1) at the third arrow.

In the remainder of this section, $TF^\dagger$ denotes a right inverse of TF.

Definition 4.1.3. We say that $TF^\dagger$ is

- ***Locally Lipschitz***, if it is locally Lipschitz as a vector bundle map, and
- C^j, where $j \in \mathbb{Z}^*_{\geq 0}$, if it is C^j as a vector bundle map.

Of course, if $TF^\dagger$ is C^1, then it is locally Lipschitz.

Proposition 4.1.4. *The right inverse $TF^\dagger$ is locally Lipschitz whenever the following criterion is satisfied: If (U, φ) and $(V, \boldsymbol{\psi})$ are F-compatible charts on Q and R, respectively, then the map $TF^\dagger_{\varphi,\boldsymbol{\psi}} : \varphi(U) \to \mathrm{Hom}(\mathbb{R}^\ell, E_Q)$ defined by*

$$TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) = T\varphi(\varphi^{-1}(q)) \circ TF^\dagger(\varphi^{-1}(q)) \circ T\boldsymbol{\psi}^{-1}(F_{\boldsymbol{\psi},\varphi}(q))$$

is locally Lipschitz.[1]

Proof. By an extension of [Abraham et al. 1988, Definition 3.4.2], $TF^\dagger$ is locally Lipschitz as a vector bundle map if and only if the following criterion is satisfied: For each point $\rho_0 \in F^*(TR)$, there exist $TF^\dagger$-compatible vector bundle charts (A, α) and (B, β) on $F^*(TR)$ and TQ, respectively, such that

- $\rho_0 \in A$,
- $\alpha : A \to A' \times \mathbb{R}^\ell$,
- $\beta : B \to B' \times E_Q$, and
- $(TF^\dagger)_{\beta,\alpha}(q, \boldsymbol{v}) = (G(q), H(q) \cdot \boldsymbol{v})$,

where $G : A' \to B'$ and $H : A' \to \mathrm{Hom}(\mathbb{R}^\ell, E_Q)$ are locally Lipschitz.

Choose $(q_0, v_{F(q_0)}) \in F^*(TR)$ and suppose that (U, φ) and $(V, \boldsymbol{\psi})$ are F-compatible charts on Q and R, respectively, such that $q_0 \in U$. Define the map

$$T\zeta^* : \pi^{-1}_{F^*(TR)}(U) \to \varphi(U) \times \mathbb{R}^\ell$$

by

$$T\zeta^*(q, v_{F(q)}) = (\varphi(q), T\boldsymbol{\psi}(F(q)) \cdot v_{F(q)}).$$

Thus $(\pi^{-1}_{F^*(TR)}(U), T\zeta^*)$ and $(TU, T\varphi)$ are $TF^\dagger$-compatible vector bundle charts on $F^*(TR)$ and TQ, respectively, such that $(q_0, v_{F(q_0)}) \in \pi^{-1}_{F^*(TR)}(U)$. Since

$$(T\zeta^*)^{-1}(q, \boldsymbol{v}) = (\varphi^{-1}(q), T\boldsymbol{\psi}^{-1}(F_{\boldsymbol{\psi},\varphi}(q)) \cdot \boldsymbol{v}),$$

the local representative

$$(TF^\dagger)_{T\varphi, T\zeta^*} : \varphi(U) \times \mathbb{R}^\ell \to \varphi(U) \times E_Q$$

[1] That is, for each $q_0 \in \varphi(U)$, there exist a neighbourhood U_0 of q_0 and $C \in \mathbb{R}_{\geq 0}$ such that

$$\|TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) - TF^\dagger_{\varphi,\boldsymbol{\psi}}(\tilde{q})\| \leq C\, \|q - \tilde{q}\|_{E_Q}$$

for each $q, \tilde{q} \in U_0$.

is given by

$$
\begin{aligned}
(TF^\dagger)_{T\varphi,T\zeta^*}(q,\boldsymbol{v}) &= T\varphi(\varphi^{-1}(q)) \circ TF^\dagger \circ (T\zeta^*)^{-1}(q,\boldsymbol{v}) \\
&= T\varphi(\varphi^{-1}(q)) \circ TF^\dagger(\varphi^{-1}(q)) \circ T\boldsymbol{\psi}^{-1}(F_{\boldsymbol{\psi},\varphi}(q)) \cdot \boldsymbol{v}) \\
&= (q, TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) \cdot \boldsymbol{v}).
\end{aligned}
$$

This completes the proof. ∎

Proposition 4.1.5. *The right inverse $TF^\dagger$ is C^j, where $j \in \mathbb{Z}^*_{\geq 0}$, whenever the following criterion is satisfied: If (U,φ) and $(V,\boldsymbol{\psi})$ are F-compatible charts on Q and R, respectively, then the map $TF^\dagger_{\varphi,\boldsymbol{\psi}} : \varphi(U) \to \operatorname{Hom}(\mathbb{R}^\ell, E_Q)$ defined by*

$$TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) = T\varphi(\varphi^{-1}(q)) \circ TF^\dagger(\varphi^{-1}(q)) \circ T\boldsymbol{\psi}^{-1}(F_{\boldsymbol{\psi},\varphi}(q))$$

is C^j.

Proof. Analogous to the proof of Proposition 4.1.4. ∎

By assuming that Q and R are Riemannian, we obtain the next result.

Proposition 4.1.6. *Suppose that Q and R are Riemannian. Define the vector bundle map $TF^\# : F^*(TR) \to TQ$ by setting $TF^\#(q)$ to be the Moore–Penrose pseudoinverse of $TF(q)$. That is,*

$$TF^\#(q) = TF(q)^* \circ (TF(q) \circ TF(q)^*)^{-1}$$

for each $q \in Q$.[2] *Then $TF^\#$ is a C^{k-1} right inverse of TF.*

Proof. Choose $q \in Q$. We begin by showing that

$$(TF(q) \circ TF(q)^*)^{-1}$$

is well-defined. Since Q is Riemannian, we can form the orthogonal complement

$$\ker(TF(q))^\perp = \{v \in T_qQ \ : \ \langle v, \tilde{v} \rangle = 0 \text{ for each } \tilde{v} \in \ker(TF(q))\}.$$

[2]Here, the assumption that Q and R are Riemannian is used to form the adjoints $TF(q)^*$.

Clearly, the restriction of $TF(q)$ to $\ker(TF(q))^{\perp}$ is injective. Using the canonical vector space isomorphisms from [Bachman and Narici 2000, Section 20.3], we have

- $\ker(TF(q)^*) \cong \{0_{F(q)}\}$ and
- $\operatorname{image}(TF(q)^*) \cong \ker(TF(q))^{\perp}$.

Thus $TF(q)^*$ is injective and maps $T_{F(q)}R$ onto $\ker(TF(q))^{\perp}$. Consequently,

$$TF(q) \circ TF(q)^*$$

is injective by composition and $TF(q) \circ TF(q)^* \in \operatorname{Hom}(T_{F(q)}R, T_{F(q)}R)$ is bijective.

To show that $TF^{\#}$ is C^{k-1}, we use Proposition 4.1.5. Suppose that (U, φ) and $(V, \boldsymbol{\psi})$ are F-compatible charts on Q and R, respectively. Then

$$\begin{aligned} TF^{\#}_{\varphi,\boldsymbol{\psi}}(q) &= T\varphi(\varphi^{-1}(q)) \circ TF^{\#}(\varphi^{-1}(q)) \circ T\boldsymbol{\psi}^{-1}(F_{\boldsymbol{\psi},\varphi}(q)) \\ &= TF_{\boldsymbol{\psi},\varphi}(q)^* \circ (TF_{\boldsymbol{\psi},\varphi}(q) \circ TF_{\boldsymbol{\psi},\varphi}(q)^*)^{-1} \end{aligned}$$

for each $q \in \varphi(U)$. Recall from [Abraham et al. 1988, Lemma 2.5.5] that the map[3]

$$\mathcal{I} : \mathrm{GL}(\mathbb{R}^{\ell}) \to \mathrm{GL}(\mathbb{R}^{\ell})$$

that sends λ to $\mathcal{I}(\lambda) = \lambda^{-1}$ is C^{∞}. By composition, $TF^{\#}_{\varphi,\boldsymbol{\psi}}$ is C^{k-1}. ∎

Definition 4.1.7. The right inverse $TF^{\#} : F^*(TR) \to TQ$, defined as in Proposition 4.1.6 above, is called the ***Moore–Penrose pseudoinverse*** of TF.

In Section 4.3, right inverses are incorporated into initial value problems.

4.2 Initial value problems evolving on Banach manifolds

In this section, we briefly review the required theory of initial value problems evolving on Banach manifolds. We emphasize that the material in this section

[3] Here, $\mathrm{GL}(\mathbb{R}^{\ell}) \subseteq \operatorname{Hom}(\mathbb{R}^{\ell})$ denotes the set of linear automorphisms of $\mathbb{R}^{\ell}$.

does not follow from the material in Section 2.3 concerning initial value problems evolving on finite-dimensional manifolds. Indeed, the latter material does not generalize in a straightforward way to the cases considered in this section, as this generalization requires the introduction of strong measurability and the Bochner integral; see [Kuttler 1998, Chapter 23]. In any case, it is not necessary to generalize the material in Section 2.3. For the purposes of the continuation method, it is enough to consider initial value problems whose right-hand sides are continuous. Throughout this section, I is an interval and $\xi : \mathrm{dom}(\xi) \to Q$ is a curve.

Definition 4.2.1. Consider a map $X : I \times Q \to TQ$. We say that X is a ***time-varying vector field*** on Q if $\pi_{TQ} \circ X(t, q) = q$ for each $(t, q) \in I \times Q$. The set of all such maps is denoted by $\mathscr{V}(I, Q)$. Given a chart (U, φ) on Q, the ***local representative*** of X in (U, φ) is the map $X_\varphi : I \times \varphi(U) \to E_Q$ defined by

$$X_\varphi(t, q) = T\varphi(\varphi^{-1}(q)) \cdot X(t, \varphi^{-1}(q)).$$

In what follows, $X \in \mathscr{V}(I, Q)$.

Definition 4.2.2. Suppose that $(t_0, q_0) \in I \times Q$. The triple (X, t_0, q_0) is said to be an ***initial value problem*** evolving on Q, with ***right-hand side*** X and ***initial condition*** (t_0, q_0). We say that ξ is a ***solution*** of (X, t_0, q_0) if

- $\mathrm{dom}(\xi)$ is a relatively open subinterval of I containing t_0,
- ξ is C^1,
- $\xi(t_0) = q_0$, and $\dot{\xi}(t) = X(t, \xi(t))$ for each $t \in \mathrm{dom}(\xi)$.

Definition 4.2.3. Suppose that $(t_0, q_0) \in I \times Q$ and ξ is a solution of (X, t_0, q_0). We say that ξ is ***maximally-defined*** if it has the following property: If

$$\tilde{\xi} : \mathrm{dom}(\tilde{\xi}) \to Q$$

is another solution of (X, t_0, q_0), then $\mathrm{dom}(\tilde{\xi}) \subseteq \mathrm{dom}(\xi)$ and

$$\tilde{\xi}(t) = \xi(t)$$

for each $t \in \mathrm{dom}(\tilde{\xi})$. Clearly, such a solution is unique.

Next, we establish a suitable Lipschitz condition on X.

Definition 4.2.4. Suppose that U is a nonempty open subset of E_Q and

$$f : I \times U \to E_Q.$$

We say that f is ***locally Lipschitz*** if it is continuous and for each $(t_0, u_0) \in I \times U$, there exist a product neighbourhood $I_0 \times U_0$ of (t_0, u_0) and $C \in \mathbb{R}_{\geq 0}$ such that

$$\|f(t, u) - f(t, \tilde{u})\|_{E_Q} \leq C \, \|u - \tilde{u}\|_{E_Q}$$

for each $t \in I_0$ and each $u, \tilde{u} \in U_0$.

Definition 4.2.5. We say that X is ***locally Lipschitz*** if X_φ is locally Lipschitz for each chart (U, φ) on Q.

One can show that X is locally Lipschitz if and only if X_φ is locally Lipschitz for each chart $(U, \varphi) \in \mathscr{A}_Q$, where $\mathscr{A}_Q$ is a compatible atlas on Q.

Theorem 4.2.6. *Suppose that X is locally Lipschitz. Then there exists a maximally-defined solution of (X, t_0, q_0) for each $(t_0, q_0) \in I \times Q$.*

Proof. This follows from [Amann 1990, Theorem 7.6], together with a globalization procedure analogous to the one employed in Section 2.3. ∎

Provided that X is locally Lipschitz, the maximally-defined solution of the initial value problem (X, t_0, q_0) is denoted by

$$\mu^X(\cdot, t_0, q_0) : I^X(t_0, q_0) \to Q.$$

One can also define the global flow of X, although we do not require this notion.

We conclude this section with the next proposition, which describes how maximally-defined solutions behave near the boundaries of their domains of definition.

Proposition 4.2.7. *Suppose that U is a nonempty open subset of E_Q,*

$$f : I \times U \to E_Q$$

is locally Lipschitz, and $(t_0, u_0) \in I \times U$. Define

$$I^f_-(t_0, u_0) = \inf(I^f(t_0, u_0)) \qquad \textit{and} \qquad I^f_+(t_0, u_0) = \sup(I^f(t_0, u_0)).$$

If $I^f_+(t_0, u_0) < \sup(I)$, then there are two mutually exclusive possibilities:

1. $\lim_{t \nearrow I^f_+(t_0,u_0)} \|\dot{\mu}^f(t, t_0, u_0)\|_{E_Q} = \lim_{t \nearrow I^f_+(t_0,u_0)} \|f(t, \mu^f(t, t_0, u_0))\|_{E_Q} = \infty$;
2. $u_+ = \lim_{t \nearrow I^f_+(t_0,u_0)} \mu^f(t, t_0, u_0)$ *exists in E_Q and $u_+ \notin U$.*

Similarly, if $\inf(I) < I^f_-(t_0, u_0)$, then there are two mutually exclusive possibilities:

1. $\lim_{t \searrow I^f_-(t_0,u_0)} \|\dot{\mu}^f(t, t_0, u_0)\|_{E_Q} = \lim_{t \searrow I^f_-(t_0,u_0)} \|f(t, \mu^f(t, t_0, u_0))\|_{E_Q} = \infty$;
2. $u_- = \lim_{t \searrow I^f_-(t_0,u_0)} \mu^f(t, t_0, u_0)$ *exists in E_Q and $u_- \notin U$.*

Proof. Suppose that $I^f_+(t_0, u_0) < \sup(I)$. If

$$\lim_{t \nearrow I^f_+(t_0,u_0)} \|f(t, \mu^f(t, t_0, u_0))\|_{E_Q} = \infty,$$

then there is nothing to prove. Suppose that

$$C = \lim_{t \nearrow I^f_+(t_0,u_0)} \|f(t, \mu^f(t, t_0, u_0))\|_{E_Q} < \infty.$$

By continuity and the triangle inequality, there exists $t_* \in [t_0, I^f_+(t_0, u_0))$ such that

$$\|f(t, \mu^f(t, t_0, u_0))\|_{E_Q} \leq 2C$$

for each $t \in [t_*, I_+^f(t_0, u_0))$. It follows that

$$\begin{aligned}
\|\mu^f(t, t_0, u_0) - \mu^f(\tilde{t}, t_0, u_0)\|_{E_Q} &= \left\| \int_{\tilde{t}}^{t} f(\sigma, \mu^f(\sigma, t_0, u_0))\, \mathrm{d}\sigma \right\|_{E_Q} \\
&\leq \operatorname{sign}(t - \tilde{t}) \int_{\tilde{t}}^{t} \|f(\sigma, \mu^f(\sigma, t_0, u_0))\|_{E_Q}\, \mathrm{d}\sigma \\
&\leq 2C \operatorname{sign}(t - \tilde{t}) \int_{\tilde{t}}^{t} \mathrm{d}\sigma \\
&\leq 2C \operatorname{sign}(t - \tilde{t})(t - \tilde{t}) \\
&= 2C\, |t - \tilde{t}| \qquad (4.2)
\end{aligned}$$

for each $t, \tilde{t} \in [t_*, I_+^f(t_0, u_0))$. Now consider a sequence $\{t_n\}_{n\in\mathbb{N}}$ in $[t_*, I_+^f(t_0, u_0))$ converging to $I_+^f(t_0, u_0)$. Since $\{t_n\}_{n\in\mathbb{N}}$ is convergent, it is Cauchy. By (4.2), the sequence $\{\mu^f(t_n, t_0, u_0)\}_{n\in\mathbb{N}}$ in E_Q is also Cauchy. Thus

$$u_+ = \lim_{t \nearrow I_+^f(t_0, u_0)} \mu^f(t, t_0, u_0) = \lim_{n\to\infty} \mu^f(t_n, t_0, u_0)$$

exists in E_Q. To complete the proof, suppose that $u_+ \in U$. Since $I^f(I_+^f(t_0, u_0), u_+)$ is a relatively open subinterval of I containing $I_+^f(t_0, u_0)$, we have

$$I \cap (I_+^f(t_0, u_0) - \delta, I_+^f(t_0, u_0) + \delta) \subseteq I^f(I_+^f(t_0, u_0), u_+)$$

for some $\delta \in \mathbb{R}_{>0}$. In particular, we have

$$I_+^{\text{ext}} = I \cap [I_+^f(t_0, u_0), I_+^f(t_0, u_0) + \delta) \subseteq I^f(I_+^f(t_0, u_0), u_+).$$

Now define $\mu : I^f(t_0, u_0) \cup I_+^{\text{ext}} \to U$ by

$$\mu(t) = \begin{cases} \mu^f(t, t_0, u_0), & t \in I^f(t_0, u_0), \\ \mu^f(t, I_+^f(t_0, u_0), u_+), & t \in I_+^{\text{ext}}. \end{cases}$$

Clearly, μ is a solution of (f, t_0, u_0), and $I^f(t_0, u_0)$ is properly contained in

$$I^f(t_0, u_0) \cup I_+^{\text{ext}}.$$

This contradicts the fact that $\mu^f(\cdot, t_0, u_0)$ is the maximally-defined solution of (f, t_0, u_0). Hence $u_+ \notin U$. The second assertion of the proof is analogous. ∎

Example 4.2.8. Suppose that $f : [0,1] \times U \to E_Q$ is locally Lipschitz, $t_0 = 0$, $u_0 \in U$, and $I^f(0, u_0) = [0, \delta)$ for some $\delta \in (0, 1]$. Then

$$I_+^f(0, u_0) = \delta < 1 = \sup([0, 1]).$$

Proposition 4.2.7 completely characterizes this phenomenon, by telling us that exactly one of the following possibilities has occurred:

1. $\lim_{t \nearrow \delta} \|\dot{\mu}^f(t, 0, u_0)\|_{E_Q} = \infty$;
2. $u_+ = \lim_{t \nearrow \delta} \mu^f(t, 0, u_0)$ exists in E_Q and $u_+ \notin U$.

Note that if $U = E_Q$, then it must be the case that

$$\lim_{t \nearrow \delta} \|\dot{\mu}^f(t, t_0, u_0)\|_{E_Q} = \infty.$$

In Chapter 9, Proposition 4.2.7 is used in the context of sublinear growth.

4.3 The continuation method

In this section, we present the continuation method, beginning in Section 4.3.1 with general definitions and results concerning path-lifting equations. In Section 4.3.2, we specialize the material of Section 4.3.1 to control systems.

4.3.1 Path-lifting equations

Throughout this section, $TF^\dagger$ is a locally Lipschitz right inverse of TF and

$$\pi : [0, 1] \to R$$

is a C^1 curve such that $\text{image}(\pi) \subseteq \text{image}(F)$.

Definition 4.3.1. Suppose that

$$\Pi : [0, 1] \to Q$$

is a C^1 curve. We say that Π is a C^1 ***lift of*** π ***with respect to*** F if

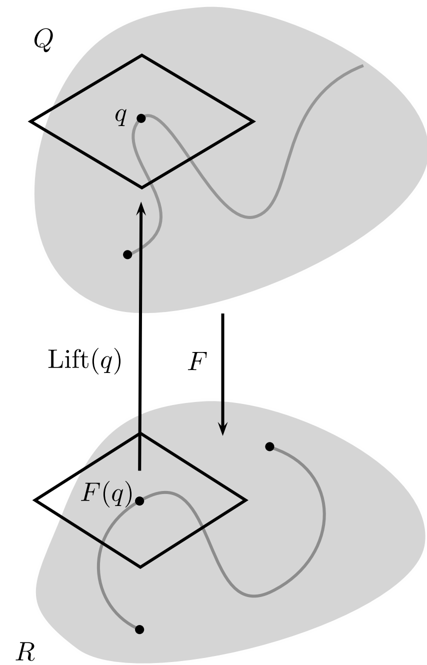

Figure 4.1: An illustration of $\mathrm{Lift}(q) \in \mathrm{Hom}(T_{F(q)}R, T_qQ)$

- J is a relatively open[4] subinterval of $[0, 1]$ containing 0 and
- $F \circ \Pi = \pi|J$.

If $J = [0, 1]$, then Π is ***total***. On the other hand, if $J \neq [0, 1]$, then Π is ***partial***.

We now turn to path-lifting equations. Since F is a submersion, for each $t \in [0, 1]$ there exists $v_q \in TQ$ such that $TF(q) \cdot v_q = \dot{\pi}(t)$. We would like to have a systematic way to select the tangent vector v_q. To this end, we take the following approach: For each $q \in Q$, prescribe $\mathrm{Lift}(q) \in \mathrm{Hom}(T_{F(q)}R, T_qQ)$ such that

$$TF(q) \circ \mathrm{Lift}(q) \cdot \dot{\pi}(t) = \dot{\pi}(t).$$

This is illustrated in Figure 4.1. Collectively, the maps $\mathrm{Lift}(q)$ comprise a right inverse of TF. Intuitively speaking, if this right inverse is regular enough, then we

[4]That is, J can be written as the intersection of $[0, 1]$ with an open interval.

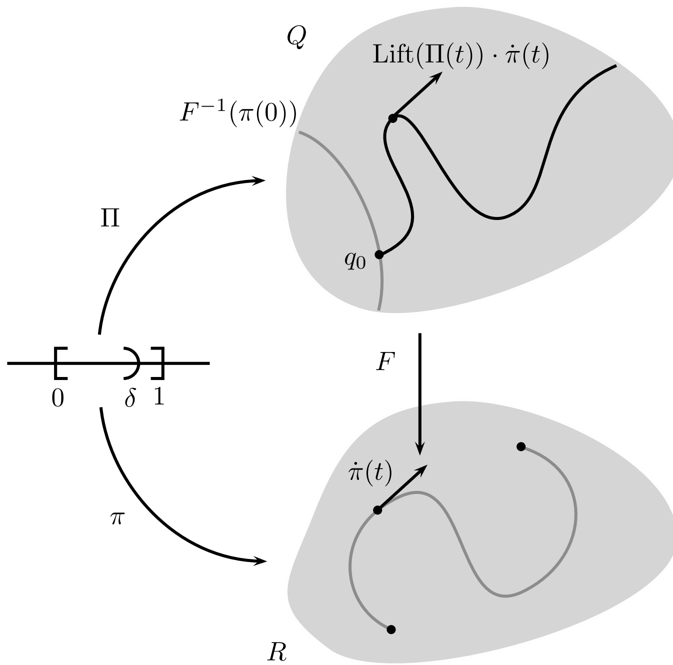

Figure 4.2: An illustration of lifting $\dot{\pi}(t)$ for each $t \in [0, 1]$

can construct Π as the maximally-defined solution of the initial value problem

$$\begin{cases} \dot{\Pi}(t) = \mathrm{Lift}(\Pi(t)) \cdot \dot{\pi}(t), \quad \Pi(t) \in Q, \quad t \in [0, 1] \\ \Pi(0) = q_0, \end{cases} \tag{4.3}$$

provided that q_0 is chosen in an appropriate way. In this sense, one can think of Π as being constructed from "infinitesimal samples" of π which are lifted to TQ. This is illustrated in Figure 4.2.

In the next proposition, we begin to make this intuitive description precise. The time-varying vector field H_π is incorporated to extend the "tangent vector field" along π to a vector field on R, and the assumption that H_π is time-varying is included to handle the possibility that π may not be injective.

Proposition 4.3.2. *There exists $H_\pi \in \mathscr{V}([0, 1], R)$ such that*

1. *H_π is locally Lipschitz,*

2. *$\dot{\pi}(t) = H_\pi(t, \pi(t))$ for each $t \in [0, 1]$, and*

3. *The time-varying vector field* $H_\pi^\dagger \in \mathscr{V}([0,1], Q)$ *defined by*

$$H_\pi^\dagger(t,q) = TF^\dagger(q) \cdot H_\pi(t, F(q))$$

is locally Lipschitz.

Proof. The first two assertions are proven in [Chitour 2006, Section 5.3]. In the cited work, the construction of H_π requires that R admits C^∞ partitions of unity; this fact underlies our standing assumption that R is second-countable.

To show that $H_\pi^\dagger$ is locally Lipschitz, we use Proposition 4.1.4. Suppose that (U, φ) and $(V, \boldsymbol{\psi})$ are F-compatible charts on Q and R, respectively. Then

$$\begin{aligned}
(H_\pi^\dagger)_\varphi(t,q) &= T\varphi(\varphi^{-1}(q)) \cdot H_\pi^\dagger(t, \varphi^{-1}(q)) \\
&= T\varphi(\varphi^{-1}(q)) \circ TF^\dagger(\varphi^{-1}(q)) \cdot H_\pi(t, F \circ \varphi^{-1}(q)) \\
&= T\varphi(\varphi^{-1}(q)) \circ TF^\dagger(\varphi^{-1}(q)) \circ \\
&\qquad T\boldsymbol{\psi}^{-1}(F_{\boldsymbol{\psi},\varphi}(q)) \circ T\boldsymbol{\psi}(F \circ \varphi^{-1}(q)) \cdot H_\pi(t, F \circ \varphi^{-1}(q)) \\
&= TF_{\varphi,\boldsymbol{\psi}}^\dagger(q) \circ T\boldsymbol{\psi}(\boldsymbol{\psi}^{-1} \circ F_{\boldsymbol{\psi},\varphi}(q)) \cdot H_\pi(t, \boldsymbol{\psi}^{-1} \circ F_{\boldsymbol{\psi},\varphi}(q)) \\
&= TF_{\varphi,\boldsymbol{\psi}}^\dagger(q) \cdot (H_\pi)_{\boldsymbol{\psi}}(t, F_{\boldsymbol{\psi},\varphi}(q))
\end{aligned}$$

for each $(t,q) \in [0,1] \times \varphi(U)$. Now choose $(t_0, q_0) \in [0,1] \times \varphi(U)$. Then

- There exists a neighbourhood U_0 of q_0 and $C_1 \in \mathbb{R}_{\geq 0}$ such that

  $$\|TF_{\varphi,\boldsymbol{\psi}}^\dagger(q) - TF_{\varphi,\boldsymbol{\psi}}^\dagger(\tilde{q})\| \leq C_1 \|q - \tilde{q}\|_{E_Q}$$

  for each $q, \tilde{q} \in U_0$, and

- There exists a product neighbourhood $I_0 \times V_0$ of

  $$(t_0, \boldsymbol{\psi} \circ F(q_0))$$

  and $C_2 \in \mathbb{R}_{\geq 0}$ such that

  $$\|(H_\pi)_{\boldsymbol{\psi}}(t, \boldsymbol{r}) - (H_\pi)_{\boldsymbol{\psi}}(t, \tilde{\boldsymbol{r}})\|_{\mathbb{R}^\ell} \leq C_2 \|\boldsymbol{r} - \tilde{\boldsymbol{r}}\|_{\mathbb{R}^\ell}$$

  for each $t \in I_0$ and each $\boldsymbol{r}, \tilde{\boldsymbol{r}} \in V_0$.

Shrinking U_0 and $I_0 \times V_0$, we may assume without loss of generality that

- $F_{\boldsymbol{\psi},\varphi}(U_0) \subseteq V_0$,
- $\|(H_\pi)_{\boldsymbol{\psi}}(t, \boldsymbol{r})\|_{\mathbb{R}^\ell} \leq C_0$ for each $(t, \boldsymbol{r}) \in I_0 \times V_0$, where $C_0 \in \mathbb{R}_{\geq 0}$, and
- $\|TF^\dagger_{\varphi,\boldsymbol{\psi}}(q)\| \leq C_0'$ for each $q \in U_0$, where $C_0' \in \mathbb{R}_{\geq 0}$.

Using the above properties, we have

$$\begin{aligned}
&\|(H^\dagger_\pi)_\varphi(t, q) - (H^\dagger_\pi)_\varphi(t, \tilde{q})\|_{E_Q} \\
&= \|TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) \cdot (H_\pi)_{\boldsymbol{\psi}}(t, F_{\boldsymbol{\psi},\varphi}(q)) - TF^\dagger_{\varphi,\boldsymbol{\psi}}(\tilde{q}) \cdot (H_\pi)_{\boldsymbol{\psi}}(t, F_{\boldsymbol{\psi},\varphi}(\tilde{q}))\|_{E_Q} \\
&= \|[TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) - TF^\dagger_{\varphi,\boldsymbol{\psi}}(\tilde{q})] \cdot (H_\pi)_{\boldsymbol{\psi}}(t, F_{\boldsymbol{\psi},\varphi}(\tilde{q})) \\
&\qquad + TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) \cdot [(H_\pi)_{\boldsymbol{\psi}}(t, F_{\boldsymbol{\psi},\varphi}(q)) - (H_\pi)_{\boldsymbol{\psi}}(t, F_{\boldsymbol{\psi},\varphi}(\tilde{q}))]\|_{E_Q} \\
&\leq C_1 \|q - \tilde{q}\|_{E_Q} \|(H_\pi)_{\boldsymbol{\psi}}(t, F_{\boldsymbol{\psi},\varphi}(\tilde{q}))\|_{\mathbb{R}^\ell} + C_2 \|TF^\dagger_{\varphi,\boldsymbol{\psi}}(q)\| \|F_{\boldsymbol{\psi},\varphi}(q) - F_{\boldsymbol{\psi},\varphi}(\tilde{q})\|_{\mathbb{R}^\ell} \\
&\leq C_0 C_1 \|q - \tilde{q}\|_{E_Q} + C_0' C_2 \|F_{\boldsymbol{\psi},\varphi}(q) - F_{\boldsymbol{\psi},\varphi}(\tilde{q})\|_{\mathbb{R}^\ell}
\end{aligned}$$

for each $t \in I_0$ and each $q, \tilde{q} \in U_0$. Since $F_{\boldsymbol{\psi},\varphi}$ is C^1, it is locally Lipschitz. In particular, by further shrinking U_0, we may assume without loss of generality that

$$\|F_{\boldsymbol{\psi},\varphi}(q) - F_{\boldsymbol{\psi},\varphi}(\tilde{q})\|_{\mathbb{R}^\ell} \leq C_0'' \|q - \tilde{q}\|_{E_Q}$$

for each $q, \tilde{q} \in U_0$, where $C_0'' \in \mathbb{R}_{\geq 0}$. Setting $C = \max\{C_0 C_1, C_0' C_0'' C_2\}$, we have

$$\|(H^\dagger_\pi)_\varphi(t, q) - (H^\dagger_\pi)_\varphi(t, \tilde{q})\|_{E_Q} \leq C \|q - \tilde{q}\|_{E_Q}$$

for each $t \in I_0$ and each $q, \tilde{q} \in U_0$. This completes the proof. ∎

For our purposes, the most important consequence of Proposition 4.3.2 is that there exist unique maximally-defined solutions of the initial value problems

$$\begin{cases} \dot{\Pi}(t) = TF^\dagger(\Pi(t)) \cdot H_\pi(t, F \circ \Pi(t)), \quad \Pi(t) \in Q, \quad t \in [0, 1] \\ \Pi(0) = q_0, \end{cases} \tag{4.4}$$

and

$$\begin{cases} \dot{r}(t) = H_\pi(t, r(t)), \quad r(t) \in R, \quad t \in [0,1] \\ r(0) = r_0 \end{cases} \tag{4.5}$$

for each $q_0 \in Q$ and each $r_0 \in R$. For reasons which will be made clear below, we say that (4.4) is a ***path-lifting equation*** (PLE) for F. Here, the right inverse $TF^\dagger$ is understood. When we wish to emphasize the role of $TF^\dagger$, we will say that (4.4) is a PLE for F ***relative to*** $TF^\dagger$. On the other hand, when we wish to emphasize the data π and q_0, we will refer to (4.4) as the (π, q_0)-PLE for F.

The next lemma tells us that certain solutions of path-lifting equations are C^1 lifts of π with respect to F. Note that these lifts may be partial lifts, in general.

Lemma 4.3.3. *Suppose that $q_0 \in F^{-1}(\pi(0))$. Then*

$$\mu^{H_\pi^\dagger}(\cdot, 0, q_0) : I^{H_\pi^\dagger}(0, q_0) \to Q$$

is a C^1 lift of π with respect to F, where $H_\pi^\dagger$ is defined as in Proposition 4.3.2.

Proof. Set $\Pi = \mu^{H_\pi^\dagger}(\cdot, 0, q_0)$. By definition, $I^{H_\pi}(0, q_0)$ is a relatively open subinterval of $[0,1]$ containing 0. Since $F \circ \Pi(0) = F(q_0) = \pi(0)$ and

$$\begin{aligned} \dot{\widehat{F \circ \Pi}}(t) &= TF(\Pi(t)) \cdot \dot{\Pi}(t) \\ &= TF(\Pi(t)) \circ H_\pi^\dagger(t, \Pi(t)) \\ &= TF(\Pi(t)) \circ TF^\dagger(\Pi(t)) \cdot H_\pi(t, F \circ \Pi(t)) \\ &= \mathrm{id}_{T_{F \circ \Pi(t)} R} \cdot H_\pi(t, F \cdot \Pi(t)) \\ &= H_\pi(t, F \circ \Pi(t)) \end{aligned}$$

for each $t \in I^{H_\pi}(t_0, q_0)$, it follows that $F \circ \Pi$ is a solution of the initial-value problem

$$(H_\pi, 0, \pi(0)).$$

However, π is the maximally-defined solution of $(H_\pi, 0, \pi(0))$ by construction, since

$$\dot{\pi}(t) = H_\pi(t, \pi(t))$$

for each $t \in [0,1]$. This implies that

$$F \circ \Pi = \pi | I^{H_\pi^\dagger}(0, q_0).$$

We conclude that Π is a C^1 lift of π with respect to F. ∎

With this material established, we now describe the continuation method.

4.3.2 The continuation method

Throughout this section,

- M is a second-countable n-dimensional manifold,
- $\Sigma = (f, \mathscr{U})$ is a nice C^1_p control system evolving on M (see Definition 3.3.7),
- The time domain of Σ is $J = [a, b]$, so that $\mathscr{U} = L^p(J, \mathbb{R}^r)$, and
- Σ is completely controllable from a fixed initial state $x_0 \in M$ on J.

By Corollary 3.3.12, the map

$$\mathrm{End}^{\Sigma}_{x_0} : \mathrm{dom}(\mathrm{End}^{\Sigma}_{x_0}) \subseteq \mathscr{U} \to M$$

is C^1. For notational economy, we write

$$\mathscr{U}_{x_0} = \mathrm{dom}(\mathrm{End}^{\Sigma}_{x_0}).$$

4.3.2.1 Basic definitions

We begin by recalling that in Chapter 1, the ***x_0-anchored motion planning problem*** (x_0-anchored MPP) for Σ was posed in the following way:

Problem: For each $x \in M$, find $\boldsymbol{u} \in \mathscr{U}_{x_0}$ such that $\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}) = x$.

To attack this problem, we would like to use the material of Section 4.3, replacing the submersion $F : Q \to R$ by $\mathrm{End}^{\Sigma}_{x_0}$. However, the latter map may not be a submersion. To remedy this, we will work with a "submersive version" of $\mathrm{End}^{\Sigma}_{x_0}$ obtained by restriction. Next, we establish the requisite terminology and notation.

Definition 4.3.4. We say that a control $\boldsymbol{u} \in \mathscr{U}_{x_0}$ is ***singular*** (with respect to x_0) if it is a singular point of $\mathrm{End}^{\Sigma}_{x_0}$. That is, $\boldsymbol{u}$ is singular if and only if

$$\operatorname{rank}(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u})) < n.$$

The set of all singular controls is denoted by $\mathscr{U}^{\mathrm{sing}}_{x_0}$. If $\boldsymbol{u}$ is not singular, then it is ***regular***. We denote the set of all regular controls by $\mathscr{U}^{\mathrm{reg}}_{x_0}$.

In the remainder of this section, we assume that $\mathscr{U}^{\mathrm{reg}}_{x_0}$ is nonempty. By [Abraham et al. 1988, Proposition 3.6.10], $\mathscr{U}^{\mathrm{reg}}_{x_0}$ is an open submanifold of $L^p(J, \mathbb{R}^r)$.

Definition 4.3.5. Define $\underline{\mathrm{End}}^{\Sigma}_{x_0} : \mathscr{U}^{\mathrm{reg}}_{x_0} \to M$ by

$$\underline{\mathrm{End}}^{\Sigma}_{x_0}(\boldsymbol{u}) = \mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}),$$

so that $\underline{\mathrm{End}}^{\Sigma}_{x_0}$ is simply the restriction of $\mathrm{End}^{\Sigma}_{x_0}$ to $\mathscr{U}^{\mathrm{reg}}_{x_0}$. We say that $\underline{\mathrm{End}}^{\Sigma}_{x_0}$ is the ***desingularized*** x_0-anchored endpoint map of Σ.

By construction, $\underline{\mathrm{End}}^{\Sigma}_{x_0}$ is a C^1 submersion.

4.3.2.2 The general case

We now give an algorithmic description of the continuation method. Throughout this section, $T\underline{\mathrm{End}}^{\Sigma,\dagger}_{x_0}$ is a locally Lipschitz right inverse of $T\underline{\mathrm{End}}^{\Sigma}_{x_0}$.

Example 4.3.6. Suppose that M is Riemannian and Σ is C^2_2-polynomial. By Corollary 3.3.12, the map $\underline{\mathrm{End}}^{\Sigma}_{x_0}$ is C^2. Since $\mathscr{U}^{\mathrm{reg}}_{x_0}$ is an open submanifold of

$$\mathscr{U} = L^2(J, \mathbb{R}^r),$$

it is Riemannian when endowed with the induced metric. By Proposition 4.1.6, the Moore–Penrose pseudoinverse $T\underline{\mathrm{End}}^{\Sigma,\#}_{x_0}$ is C^1 and thus locally Lipschitz. This scenario includes the "classical" continuation method of Sussmann [1993]. Indeed, if Σ is a driftless C^∞ control-affine system that uses L^2 controls, then Σ is C^2_2-polynomial by Example 3.2.10.

The continuation method solves the x_0-anchored MPP in the following way.

Algorithm 4.3.7. (CONTINUATION METHOD) Given $x \in M$,

1. Choose a C^1 curve $\pi : [0,1] \to M$ such that $\pi(1) = x$.
2. Given that $\mathrm{image}(\pi) \subseteq \mathrm{image}(\underline{\mathrm{End}}^\Sigma_{x_0})$,

   a. Choose $\boldsymbol{u}_0 \in (\underline{\mathrm{End}}^\Sigma_{x_0})^{-1}(\pi(0))$.

   b. If the maximally-defined solution $\mathbf{\Pi}$ of the $(\pi, \boldsymbol{u}_0)$-PLE for $\underline{\mathrm{End}}^\Sigma_{x_0}$ is defined on $[0,1]$, then choose $\boldsymbol{u} = \mathbf{\Pi}(1)$. By Lemma 4.3.3, the control $\boldsymbol{u}$ satisfies

$$\underline{\mathrm{End}}^\Sigma_{x_0}(\boldsymbol{u}) = \mathrm{End}^\Sigma_{x_0}(\boldsymbol{u}) = \mathrm{End}^\Sigma_{x_0} \circ \mathbf{\Pi}(1) = \pi(1) = x.$$

If we cannot choose π such that

$$\mathrm{image}(\pi) \subseteq \mathrm{image}(\underline{\mathrm{End}}^\Sigma_{x_0})$$

is satisfied, then the algorithm fails. For example, this occurs if $x \notin \mathrm{image}(\underline{\mathrm{End}}^\Sigma_{x_0})$.

Remark 4.3.8. Before moving on, let us explain the absence of $\mathrm{image}(\underline{\mathrm{End}}^\Sigma_{x_0})$ in the "classical" continuation method of Sussmann [1993], as this is somewhat difficult to discern from the literature. Suppose that Σ is a driftless C^∞ control-affine system that uses L^2 controls. The key fact is that

$$\mathrm{image}(\underline{\mathrm{End}}^\Sigma_{x_0}) = M,$$

which follows from the assumption that $\mathscr{U}^{\mathrm{reg}}_{x_0}$ is nonempty; see [Bellaïche 1996, Lemmas 2.1 and 2.2]. For instance, $\mathscr{U}^{\mathrm{reg}}_{x_0}$ is nonempty whenever

- Σ is C^ω and satisfies the Lie algebra rank condition [Sontag 1992] or
- Σ is strongly bracket-generating [Sussmann 1993].

4.3.2.3 The case where no controls are singular

Evidently, the existence of singular controls complicates the continuation method to a large degree. In particular, verifying that the constraint

$$\operatorname{image}(\pi) \subseteq \operatorname{image}(\underline{\mathrm{End}}^{\Sigma}_{x_0})$$

is satisfied necessitates a complete characterization of $\mathscr{U}^{\mathrm{sing}}_{x_0}$. In Chapter 1, we called this difficulty the second obstruction to the continuation method. In this section, we indicate how the continuation method simplifies in the case where $\mathscr{U}^{\mathrm{sing}}_{x_0}$ is empty. Throughout this section, $T\underline{\mathrm{End}}^{\Sigma,\dagger}_{x_0}$ is a locally Lipschitz right inverse of $T\underline{\mathrm{End}}^{\Sigma}_{x_0}$.

Note that if $\mathscr{U}^{\mathrm{sing}}_{x_0} = \varnothing$, then $\underline{\mathrm{End}}^{\Sigma}_{x_0} = \mathrm{End}^{\Sigma}_{x_0}$ and

$$\operatorname{image}(\underline{\mathrm{End}}^{\Sigma}_{x_0}) = M.$$

Furthermore, the continuation method simplifies in the following way.

Algorithm 4.3.9. (Simplified continuation method) Given $x \in M$,

1. Choose a C^1 curve $\pi : [0,1] \to M$ with $\pi(1) = x$.

2. Choose $\boldsymbol{u}_0 \in (\mathrm{End}^{\Sigma}_{x_0})^{-1}(\pi(0))$.

3. If the maximally-defined solution $\boldsymbol{\Pi}$ of the $(\pi, \boldsymbol{u}_0)$-PLE for $\mathrm{End}^{\Sigma}_{x_0}$ is defined on $[0,1]$, then choose $\boldsymbol{u} = \boldsymbol{\Pi}(1)$. By Lemma 4.3.3, the control $\boldsymbol{u}$ satisfies

$$\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}) = \mathrm{End}^{\Sigma}_{x_0} \circ \boldsymbol{\Pi}(1) = \pi(1) = x.$$

In Chapter 11, we will study a control system for which $\mathscr{U}^{\mathrm{sing}}_{x_0}$ is empty.

Chapter 5

Operations on time-varying vector fields

Consider a control system

$$\Sigma = (f, \mathscr{U})$$

evolving on an n-dimensional manifold M, and let $x_0 \in M$. In subsequent chapters, we will derive expressions for the differential of End^{Σ} and the intrinsic quadratic differentials of $\mathrm{End}^{\Sigma}_{x_0}$ (a formal definition of intrinsic quadratic differentials is delayed until Chapter 7). As we will see, these differentials are related to the controlled trajectories of certain control systems derived from Σ. These control systems, called the first and second variations of Σ, are constructed by lifting time-varying vector fields on M to TM and TTM. The requisite lifting operations are the subject of this chapter, including their basic properties and their interplay with one another.

This chapter is organized in the following way. In Section 5.1, we define vertical lifts and record their relevant properties. In Sections 5.2 and 5.3, we repeat this process for tangent lifts and cotangent lifts, respectively. In Section 5.4, we define the pullback of one time-varying vector field by the global flow of another, and recall the nonlinear variation of constants formula. Finally, in Sections 5.5, 5.6, 5.7, 5.8, and 5.9, we derive a number of new and useful identities. These identities provide reductive formulas for pullbacks involving lifts, an explicit formula for the global flow of $X + Y$, where X is a tangent lift and Y is a vertical lift, and explicit formulas for time derivatives and scalar parameter derivatives of pullbacks. These identities will play an important role in later chapters, although they are also interesting in their own right.

Our standing assumptions throughout this chapter are that

- I is an interval,
- Q is an ℓ-dimensional manifold,
- $\Xi, \Upsilon \in \mathscr{V}(I, Q)$ are time-varying vector fields on Q, and
- $\chi \in \mathscr{V}(I, Q, \mathbb{R}^r)$ is a controllable time-varying vector field on Q.

At many junctures, it will be necessary to assume that Ξ, Υ, and χ have stronger regularity properties. These assumptions will be clearly stated.

5.1 Vertical lifts

5.1.1 Basic definitions and properties

We begin by defining vertical subspaces. Given $v_q \in T_qQ$, we say that

$$V_{v_q}TQ = \ker(T\pi_{TQ}(v_q))$$

is the ***vertical subspace*** at v_q, and $\mathrm{vlft}_{v_q} \in \mathrm{Hom}(T_qQ, V_{v_q}TQ)$, defined by

$$\mathrm{vlft}_{v_q} \cdot \tilde{v}_q = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_0 (v_q + s\tilde{v}_q),$$

is the ***pointwise vertical lift*** at v_q. For details, see [Crampin and Pirani 1986, Chapter 13.2]. Applying each vlft_{v_q}, one obtains the vertical lifts of Ξ and χ.

Definition 5.1.1. We say that $\mathrm{vlft}(\Xi) \in \mathscr{V}(I, TQ)$, defined by

$$\mathrm{vlft}(\Xi)(t, v_q) = \mathrm{vlft}_{v_q} \cdot \Xi(t, q),$$

is the ***vertical lift*** of Ξ. Of course, one can also form the vertical lift of any vector field Ξ on Q by identifying Ξ with the time-varying vector field

$$\Xi(t, q) = \Xi(q).$$

Under this identification, we have $\mathrm{vlft}(\Xi)_t = \mathrm{vlft}(\Xi_t)$.

Definition 5.1.2. We say that $\mathrm{vlft}(\chi) \in \mathscr{V}(I, TQ, \mathbb{R}^r)$, defined by

$$\mathrm{vlft}(\chi)(t, v_q, \boldsymbol{\omega}) = \mathrm{vlft}_{v_q} \cdot \chi(t, q, \boldsymbol{\omega}),$$

is the ***vertical lift*** of χ.

If $(TTV, TT\boldsymbol{\psi})$ is a natural chart on TTQ, then $\alpha_{v_q} \in V_{v_q}TQ$ if and only if

$$TT\boldsymbol{\psi}(v_q) \cdot \alpha_{v_q} = (\boldsymbol{q}, \boldsymbol{v}, \mathbf{0}_{\mathbb{R}^\ell}, \tilde{\boldsymbol{v}})$$

for some $\tilde{\boldsymbol{v}} \in \mathbb{R}^\ell$. In particular, if $T\boldsymbol{\psi}(q) \cdot \tilde{v}_q = (\boldsymbol{q}, \tilde{\boldsymbol{v}})$, then

$$\begin{aligned} TT\boldsymbol{\psi}(v_q) \circ \mathrm{vlft}_{v_q} \cdot \tilde{v}_q &= TT\boldsymbol{\psi}(v_q) \circ \mathrm{vlft}_{v_q} \circ T\boldsymbol{\psi}^{-1}(\boldsymbol{q}) \cdot \tilde{\boldsymbol{v}} \\ &= (\boldsymbol{q}, \boldsymbol{v}, \mathbf{0}_{\mathbb{R}^\ell}, \tilde{\boldsymbol{v}}). \end{aligned}$$

From this observation, we see that vlft_{v_q} is a canonical vector space isomorphism, and that the local representatives of $\mathrm{vlft}(\Xi)$ and $\mathrm{vlft}(\chi)$ are given as follows.

Lemma 5.1.3. *We have*

$$\mathrm{vlft}(\Xi)_{T\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{v}) = \Big(\mathbf{0}_{\mathbb{R}^\ell}, \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\Big)$$

for each natural chart $(TV, T\boldsymbol{\psi})$ on TQ.

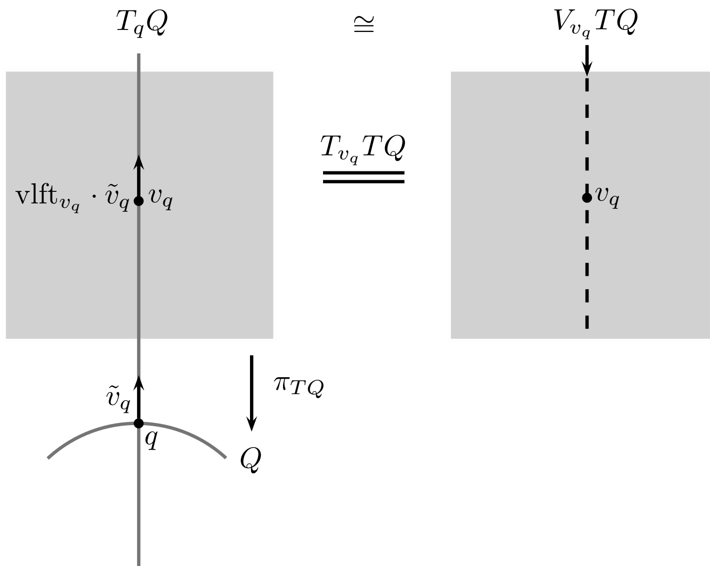

Figure 5.1: An illustration of the pointwise vertical lift operation

Lemma 5.1.4. *We have*

$$\mathrm{vlft}(\chi)_{T\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{v}, \boldsymbol{\omega}) = \Big(\mathbf{0}_{\mathbb{R}^\ell}, \chi_{\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{\omega})\Big)$$

for each natural chart $(TV, T\boldsymbol{\psi})$ on TQ.

The next two results follow immediately.

Lemma 5.1.5. *Suppose that Ξ is locally integrably C^k, where $k \in \mathbb{Z}^*_{\geq 0}$. Then its vertical lift* $\mathrm{vlft}(\Xi)$ *is locally integrably C^k.*

Theorem 5.1.6. *Suppose that Ξ is locally integrably C^0. Then its vertical lift* $\mathrm{vlft}(\Xi)$ *is solvable. Furthermore,*

$$\mathrm{dom}(\Phi^{\mathrm{vlft}(\Xi)}) = I \times I \times TQ$$

and $\Phi^{\mathrm{vlft}(\Xi)}$ sends (t, t_0, v_{q_0}) to

$$\Phi^{\mathrm{vlft}(\Xi)}_{t,t_0}(v_{q_0}) = v_{q_0} + \int_{t_0}^{t} \Xi(\sigma, q_0)\, \mathrm{d}\sigma.$$

In particular, each tangent space $T_{q_0}Q$ is invariant under $\Phi^{\mathrm{vlft}(\Xi)}_{t,t_0}$.

Remark 5.1.7. Theorem 5.1.6 holds under more relaxed conditions. It is enough that for each $q_0 \in Q$, there exists a chart $(V, \boldsymbol{\psi})$ on Q and $\alpha \in L^1_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0})$ such that $q_0 \in V$ and $\|\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{\psi}(q_0))\|_{\mathbb{R}^\ell} \leq \alpha(t)$ for a.a. $t \in I$. We will not need this degree of generality, however.

5.1.2 Composition with second differentials

The next lemma gives a relationship between first differentials, second differentials, and pointwise vertical lifts (more precisely, inverse pointwise vertical lifts).

Lemma 5.1.8. *Suppose that*

- *Q_1 and Q_2 are manifolds of dimension ℓ_1 and ℓ_2, respectively,*
- *$F : Q_1 \to Q_2$ is C^2, and*
- *$\alpha_{v_q} \in V_{v_q}TQ_1$.*

Then $TTF(v_q) \cdot \alpha_{v_q} \in V_{TF(q)\cdot v_q}TQ_2$ and

$$\mathrm{vlft}^{-1}_{TF(q)\cdot v_q} \circ TTF(v_q) \cdot \alpha_{v_q} = TF(q) \circ \mathrm{vlft}^{-1}_{v_q} \cdot \alpha_{v_q}.$$

Proof. Suppose that $(TTV_1, TT\boldsymbol{\psi}_1)$ and $(TTV_2, TT\boldsymbol{\psi}_2)$ are TTF-compatible natural charts on TTQ_1 and TTQ_2, respectively, such that $\alpha_{v_q} \in TTV_1$. By assumption,

$$TT\boldsymbol{\psi}_1(v_q) \cdot \alpha_{v_q} = (\boldsymbol{q}, \boldsymbol{v}, \boldsymbol{0}_{\mathbb{R}^{\ell_1}}, \tilde{\boldsymbol{v}})$$

for some $\tilde{\boldsymbol{v}} \in \mathbb{R}^{\ell_1}$. Using [Abraham et al. 1988, Exercise 2.4H], we have

$$TTF_{TT\boldsymbol{\psi}_2, TT\boldsymbol{\psi}_1}(\boldsymbol{q}, \boldsymbol{v}, \boldsymbol{0}_{\mathbb{R}^{\ell_1}}, \tilde{\boldsymbol{v}}) = (F_{\boldsymbol{\psi}_2,\boldsymbol{\psi}_1}(\boldsymbol{q}), \boldsymbol{D}F_{\boldsymbol{\psi}_2,\boldsymbol{\psi}_1}(\boldsymbol{q}) \cdot \boldsymbol{v}, \boldsymbol{0}_{\mathbb{R}^{\ell_2}}, \boldsymbol{D}F_{\boldsymbol{\psi}_2,\boldsymbol{\psi}_1}(\boldsymbol{q}) \cdot \tilde{\boldsymbol{v}}).$$

Thus $TTF(v_q) \cdot \alpha_{v_q} \in V_{TF(q)\cdot v_q}TQ_2$. Using this result, we compute

$$T\boldsymbol{\psi}_2(F(q)) \circ \mathrm{vlft}^{-1}_{TF(q)\cdot v_q} \circ TTF(v_q) \cdot \alpha_{v_q} = (F_{\boldsymbol{\psi}_2,\boldsymbol{\psi}_1}(\boldsymbol{q}), \boldsymbol{D}F_{\boldsymbol{\psi}_2,\boldsymbol{\psi}_1}(\boldsymbol{q}) \cdot \tilde{\boldsymbol{v}}).$$

Another straightforward computation yields

$$T\boldsymbol{\psi}_2(F(q)) \circ TF(q) \circ \mathrm{vlft}_{v_q}^{-1} \cdot \alpha_{v_q} = (F_{\boldsymbol{\psi}_2,\boldsymbol{\psi}_1}(\boldsymbol{q}), \boldsymbol{D}F_{\boldsymbol{\psi}_2,\boldsymbol{\psi}_1}(\boldsymbol{q}) \cdot \tilde{\boldsymbol{v}}).$$

This completes the proof. ∎

5.2 Tangent lifts

5.2.1 A local result concerning total derivatives

Before defining tangent lifts, we recall a local result concerning the total derivatives of global flows. In what follows, V is a nonempty open subset of $\mathbb{R}^\ell$ and

$$\boldsymbol{f} : I \times V \to \mathbb{R}^\ell$$

is locally integrably C^1.

Theorem 5.2.1. *The map $\boldsymbol{g} : I \times V \times \mathbb{R}^\ell \to \mathbb{R}^\ell \oplus \mathbb{R}^\ell$, defined by*

$$\boldsymbol{g}(t, \boldsymbol{q}, \boldsymbol{v}) = \Big(\boldsymbol{f}(t, \boldsymbol{q}), \boldsymbol{D}_2\boldsymbol{f}(t, \boldsymbol{q}) \cdot \boldsymbol{v}\Big),$$

is solvable. Furthermore,

$$\mathrm{dom}(\Phi^{\boldsymbol{g}}) = \{(t, t_0, \boldsymbol{q}_0, \boldsymbol{v}_0) \ : \ (t_0, \boldsymbol{q}_0, \boldsymbol{v}_0) \in I \times V \times \mathbb{R}^\ell \textit{ and } t \in I^{\boldsymbol{f}}(t_0, \boldsymbol{q}_0)\}$$

and $\Phi^{\boldsymbol{g}}$ sends $(t, t_0, \boldsymbol{q}_0, \boldsymbol{v}_0)$ to

$$\Phi^{\boldsymbol{g}}_{t,t_0}(\boldsymbol{q}_0, \boldsymbol{v}_0) = \Big(\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0), \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0) \cdot \boldsymbol{v}_0\Big).$$

In particular, for each $(t_0, \boldsymbol{q}_0) \in I \times V$, the map

$$t \mapsto \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)$$

of $I^{\boldsymbol{f}}(t_0, \boldsymbol{q}_0)$ into $\mathrm{Hom}(\mathbb{R}^\ell, \mathbb{R}^\ell)$ is LAC and satisfies

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_t \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0) = \boldsymbol{D}_2\boldsymbol{f}(t, \Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)) \circ \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)$$

for a.a. $t \in I^{\boldsymbol{f}}(t_0, \boldsymbol{q}_0)$.

Proof. See [McShane 1983, Chapter III, Theorem 6-1]. ∎

5.2.2 Basic definitions and properties

We now define tangent lifts. Throughout this section, we assume that Ξ is locally integrably C^1. Thus the frozen-time vector field Ξ_t is C^1 for each $t \in I$. Recall from Section 2.1.7 that s_Q denotes the canonical involution of TTQ.

Definition 5.2.2. We say that $\mathrm{tlft}(\Xi) \in \mathscr{V}(I, TQ)$, defined by

$$\mathrm{tlft}(\Xi)(t, v_q) = s_Q \circ T\Xi_t(q) \cdot v_q,$$

is the ***tangent lift*** of Ξ.[1] Of course, one can also form the tangent lift of any C^1 vector field Ξ on Q by identifying Ξ with the time-varying vector field

$$\Xi(t, q) = \Xi(q).$$

Under this identification, we have $\mathrm{tlft}(\Xi)_t = \mathrm{tlft}(\Xi_t)$.

The next lemma describes the local representatives of $\mathrm{tlft}(\Xi)$.

Lemma 5.2.3. *We have*

$$\mathrm{tlft}(\Xi)_{T\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{v}) = \Big(\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}), \boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}) \cdot \boldsymbol{v}\Big)$$

for each natural chart $(TV, T\boldsymbol{\psi})$ on TQ.

Proof. Choose a natural chart $(TV, T\boldsymbol{\psi})$ on TQ. For each $t \in I$, the local representative of $F_t = s_Q \circ T\Xi_t$ in $(TV, T\boldsymbol{\psi})$ and $(TTV, TT\boldsymbol{\psi})$ is given by

$$(F_t)_{TT\boldsymbol{\psi}, T\boldsymbol{\psi}}(\boldsymbol{q}, \boldsymbol{v}) = \Big(\boldsymbol{q}, \boldsymbol{v}, \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}), \boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}) \cdot \boldsymbol{v}\Big).$$

Using this result, we have

$$\mathrm{tlft}(\Xi)_{T\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{v}) = \Big(\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}), \boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}) \cdot \boldsymbol{v}\Big).$$

This completes the proof. ∎

[1] The tangent lift of Ξ is known by various other names throughout the literature, such as the canonical lift, complete lift, natural lift, variational vector field, prolongation, and flow prolongation of Ξ. These names are used in standard references on differential geometry, such as [Kolář et al. 1993], as well as in the control theory literature [Crouch and van der Schaft 1987].

The next result follows from Theorem 5.2.1 and Lemma 5.2.3.

Theorem 5.2.4. *The tangent lift* $\mathrm{tlft}(\Xi)$ *is solvable. Furthermore,*

$$\mathrm{dom}(\Phi^{\mathrm{tlft}(\Xi)}) = \{(t, t_0, v_{q_0}) \;:\; (t_0, v_{q_0}) \in I \times T_{q_0}Q \text{ and } t \in I^{\Xi}(t_0, q_0)\}$$

and $\Phi^{\mathrm{tlft}(\Xi)}$ *sends* (t, t_0, v_{q_0}) *to*

$$\Phi^{\mathrm{tlft}(\Xi)}_{t,t_0}(v_{q_0}) = T\Phi^{\Xi}_{t,t_0}(q_0) \cdot v_{q_0}.$$

In particular, $t \mapsto \Phi^{\mathrm{tlft}(\Xi)}_{t,t_0}(v_{q_0})$ *is a vector field along* $t \mapsto \Phi^{\Xi}_{t,t_0}(q_0)$.

Remark 5.2.5. Given $(t, t_0) \in I \times I$, the preceding theorem implies that

$$\mathrm{dom}(\Phi^{\mathrm{tlft}(\Xi)}_{t,t_0}) = T\mathrm{dom}(\Phi^{\Xi}_{t,t_0}).$$

If Ξ is locally integrably C^k, where $k \in \mathbb{N}^*$, and Q_0 is open in $\mathrm{dom}(\Phi^{\Xi}_{t,t_0})$, then

$$\Phi^{\mathrm{tlft}(\Xi)}_{t,t_0}|TQ_0 = T\Phi^{\Xi}_{t,t_0}|TQ_0$$

is a C^{k-1} diffeomorphism of TQ_0 onto its image with inverse

$$\Phi^{\mathrm{tlft}(\Xi)}_{t_0,t}|\Phi^{\mathrm{tlft}(\Xi)}_{t,t_0}(TQ_0) = T\Phi^{\Xi}_{t_0,t}|\Phi^{\mathrm{tlft}(\Xi)}_{t,t_0}(TQ_0).$$

This fact, which follows from Corollary 2.3.26, will be used implicitly in the remainder of this chapter.

Tangent lifts lose one degree of differentiability in the following sense.

Lemma 5.2.6. *Suppose that* Ξ *is locally integrably* C^k*, where* $k \in \mathbb{N}^*$*. Then its tangent lift* $\mathrm{tlft}(\Xi)$ *is locally integrably* C^{k-1}*.*

Proof. See Section A.1. ∎

To conclude this section, let us note that the discussion in this section is not entirely comprehensive, being limited to those properties of tangent lifts which are germane to our purposes. For additional properties of tangent lifts, we refer to [Kolář et al. 1993] and [Bullo and Lewis 2005b, Sections S1.2.1 and S1.3.4].

5.3 Cotangent lifts

5.3.1 A local result concerning adjoint total derivatives

Before defining cotangent lifts, we derive a local result concerning adjoint total derivatives of global flows. This result is dual to Theorem 5.2.1, in a certain sense. In what follows, V is a nonempty open subset of $\mathbb{R}^\ell$ and

$$\boldsymbol{f} : I \times V \to \mathbb{R}^\ell$$

is locally integrably C^1.

Theorem 5.3.1. *The map $\boldsymbol{h} : I \times V \times \mathbb{R}^\ell \to \mathbb{R}^\ell \oplus \mathbb{R}^\ell$, defined by*

$$\boldsymbol{h}(t, \boldsymbol{q}, \boldsymbol{p}) = \Big(\boldsymbol{f}(t, \boldsymbol{q}), -\boldsymbol{D}_2\boldsymbol{f}(t, \boldsymbol{q})^* \cdot \boldsymbol{p}\Big),$$

is solvable. Furthermore,

$$\operatorname{dom}(\Phi^{\boldsymbol{h}}) = \{(t, t_0, \boldsymbol{q}_0, \boldsymbol{p}_0) \ : \ (t_0, \boldsymbol{q}_0, \boldsymbol{p}_0) \in I \times V \times \mathbb{R}^\ell \textit{ and } t \in I^{\boldsymbol{f}}(t_0, \boldsymbol{q}_0)\}$$

and $\Phi^{\boldsymbol{h}}$ sends $(t, t_0, \boldsymbol{q}_0, \boldsymbol{p}_0)$ to

$$\Phi^{\boldsymbol{h}}_{t,t_0}(\boldsymbol{q}_0, \boldsymbol{p}_0) = \Big(\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0), \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t_0,t}(\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0))^* \cdot \boldsymbol{p}_0\Big).$$

In particular, for each $(t_0, \boldsymbol{q}_0) \in I \times V$, the map

$$t \mapsto \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t_0,t}(\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0))^*$$

of $I^{\boldsymbol{f}}(t_0, \boldsymbol{q}_0)$ into $\operatorname{Hom}(\mathbb{R}^\ell, \mathbb{R}^\ell)$ is LAC and satisfies

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_t \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t_0,t}(\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0))^* = -\boldsymbol{D}_2\boldsymbol{f}(t, \Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0))^* \circ \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t_0,t}(\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0))^*$$

for a.a. $t \in I^{\boldsymbol{f}}(t_0, \boldsymbol{q}_0)$.

Proof. Choose $(t_0, \boldsymbol{q}_0) \in I \times V$. By Corollary 2.2.23, we have

$$\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t_0,t}(\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0))^* = \left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^{-1}\right)^* = \left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^*\right)^{-1}.$$

Recall from [Abraham et al. 1988, Lemma 2.5.5] that the map[2]

$$\mathcal{I} : \mathrm{GL}(\mathbb{R}^\ell) \to \mathrm{GL}(\mathbb{R}^\ell)$$

that sends λ to $\mathcal{I}(\lambda) = \lambda^{-1}$ is C^∞ and $\boldsymbol{D}\mathcal{I}(\lambda) \cdot \mu = -\lambda^{-1} \circ \mu \circ \lambda^{-1}$. These facts, together with Lemma 2.2.2 and Theorem 5.2.1, imply that the map

$$t \mapsto \left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^*\right)^{-1}$$

of $I^{\boldsymbol{f}}(t_0, \boldsymbol{q}_0)$ into $\mathrm{Hom}(\mathbb{R}^\ell, \mathbb{R}^\ell)$ is LAC and satisfies

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_t \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t_0,t}(\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0))^* \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_t \left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^*\right)^{-1} \\
&= -\left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^*\right)^{-1} \circ \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_t \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^* \circ \left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^*\right)^{-1} \\
&= -\left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^*\right)^{-1} \circ \left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_t \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)\right)^* \circ \left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^*\right)^{-1} \\
&= -\left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^*\right)^{-1} \circ (\boldsymbol{D}_2\boldsymbol{f}(t, \Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)) \circ \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0))^* \circ \left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^*\right)^{-1} \\
&= -\left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^*\right)^{-1} \circ \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^* \circ \boldsymbol{D}_2\boldsymbol{f}(t, \Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0))^* \circ \left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^*\right)^{-1} \\
&= -\boldsymbol{D}_2\boldsymbol{f}(t, \Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0))^* \circ \left(\boldsymbol{D}\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0)^*\right)^{-1} \\
&= -\boldsymbol{D}_2\boldsymbol{f}(t, \Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0))^* \circ \boldsymbol{D}\Phi^{\boldsymbol{f}}_{t_0,t}(\Phi^{\boldsymbol{f}}_{t,t_0}(\boldsymbol{q}_0))^*
\end{aligned}$$

for a.a. $t \in I^{\boldsymbol{f}}(t_0, \boldsymbol{q}_0)$. This completes the proof. ∎

5.3.2 Basic definitions and properties

We now define cotangent lifts. Throughout this section, we assume that Ξ is locally integrably C^1 and T^*Q is a symplectic manifold with its canonical symplectic structure. We use the symplectic structure on T^*Q to form Hamiltonian vector fields in the following way. Choose $t \in I$, and recall from Section 2.1.7 that the

[2] Here, $\mathrm{GL}(\mathbb{R}^\ell) \subseteq \mathrm{Hom}(\mathbb{R}^\ell)$ denotes the set of all linear automorphisms of $\mathbb{R}^\ell$.

natural pairing of $v_q \in T_qQ$ and $p_q \in T_q^*Q$ is denoted by $\langle p_q, v_q\rangle$. Then

$$\mathcal{H}_{\Xi_t} : T^*Q \to \mathbb{R}$$

is the C^1 function defined by

$$\mathcal{H}_{\Xi_t}(p_q) = \langle p_q, \Xi_t(q)\rangle$$

and $\overrightarrow{\mathcal{H}}_{\Xi_t} : T^*Q \to TT^*Q$ is its associated Hamiltonian vector field. If $(T^*V, T^*\boldsymbol{\psi})$ is a natural (hence, symplectic) chart on T^*Q, then

$$(\overrightarrow{\mathcal{H}}_{\Xi_t})_{T^*\boldsymbol{\psi}}(\boldsymbol{p}, \boldsymbol{q}) = ((\Xi_t)_{\boldsymbol{\psi}}(\boldsymbol{q}), -\boldsymbol{D}\Xi_t(\boldsymbol{q})^* \cdot \boldsymbol{p}) = (\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}), -\boldsymbol{D}_2\Xi(t, \boldsymbol{q})^* \cdot \boldsymbol{p}) \quad (5.1)$$

For all details concerning symplectic manifolds, including the canonical symplectic structure on T^*Q, we refer to [Abraham et al. 1988, Chapter 8].

Definition 5.3.2. We say that $\operatorname{ctlft}(\Xi) \in \mathscr{V}(I, T^*Q)$, defined by

$$\operatorname{ctlft}(\Xi)(t, p_q) = \overrightarrow{\mathcal{H}}_{\Xi_t}(p_q),$$

is the ***cotangent lift*** of Ξ.[3] Of course, one can also form the cotangent lift of any C^1 vector field Ξ on Q by identifying Ξ with the time-varying vector field

$$\Xi(t, q) = \Xi(q).$$

Under this identification, we have $\operatorname{ctlft}(\Xi)_t = \operatorname{ctlft}(\Xi_t)$.

The next lemma describes the local representatives of $\operatorname{ctlft}(\Xi)$.[4]

[3] In the literature, the cotangent lift of Ξ is also known as the Hamiltonian lift and variational covector field of Ξ. In particular, this is true in some of the literature concerning the continuation method; see [Chitour and Sussmann 1998, Section 3] and [Chitour 2006, Section 2.1].

[4] In the literature, the local representative of $\operatorname{ctlft}(\Xi)$ in a natural chart $(T^*V, T^*\boldsymbol{\psi})$ on T^*Q is sometimes expressed as

$$\operatorname{ctlft}(\Xi)_{T^*\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{p}) = \left(\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}), -\boldsymbol{p} \cdot \boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\right). \quad (5.2)$$

The idea is that $T^*\boldsymbol{\psi}$ is a bijection of T^*V onto $\boldsymbol{\psi}(V) \times (\mathbb{R}^\ell)^*$, where $(\mathbb{R}^\ell)^*$ is identified with the space of ℓ-dimensional row vectors. Thus $\operatorname{ctlft}(\Xi)_{T^*\boldsymbol{\psi}}$ is a map of $I \times V \times (\mathbb{R}^\ell)^*$ into $\mathbb{R}^\ell \oplus (\mathbb{R}^\ell)^*$. In particular, the expression (5.2) appears in some of the literature concerning the continuation method; see [Chitour and Sussmann 1998, Section 3] and [Chitour 2006, Section 2.1].

Lemma 5.3.3. *We have*

$$\mathrm{tlft}(\Xi)_{T^*\boldsymbol{\psi}}(t,\boldsymbol{q},\boldsymbol{p}) = \Big(\Xi_{\boldsymbol{\psi}}(t,\boldsymbol{q}), -\boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(t,\boldsymbol{q})^* \cdot \boldsymbol{p}\Big)$$

for each natural (hence, symplectic) chart $(T^*V, T^*\boldsymbol{\psi})$ *on* T^*Q.

Proof. This is simply a restatement of (5.1). ∎

The next result follows from Theorem 5.3.1 and Lemma 5.3.3.

Theorem 5.3.4. *The cotangent lift* $\mathrm{ctlft}(\Xi)$ *is solvable. Furthermore,*

$$\mathrm{dom}(\Phi^{\mathrm{ctlft}(\Xi)}) = \{(t, t_0, p_{q_0}) \ : \ (t_0, p_{q_0}) \in I \times T^*_{q_0}Q \ \textit{and} \ t \in I^{\Xi}(t_0, q_0)\}$$

and $\Phi^{\mathrm{ctlft}(\Xi)}$ *sends* (t, t_0, p_{q_0}) *to*

$$\Phi^{\mathrm{ctlft}(\Xi)}_{t,t_0}(p_{q_0}) = T\Phi^{\Xi}_{t_0,t}(\Phi^{\Xi}_{t,t_0}(q_0))^* \cdot p_{q_0}.$$

In particular, $t \mapsto \Phi^{\mathrm{ctlft}(\Xi)}_{t,t_0}(p_{q_0})$ *is a covector field along* $t \mapsto \Phi^{\Xi}_{t,t_0}(q_0)$.

Cotangent lifts lose one degree of differentiability in the following sense.

Lemma 5.3.5. *Suppose that* Ξ *is locally integrably* C^k*, where* $k \in \mathbb{N}^*$*. Then its cotangent lift* $\mathrm{ctlft}(\Xi)$ *is locally integrably* C^{k-1}.

Proof. Identical to the proof of Lemma 5.2.6. ∎

To conclude this section, let us mention that additional properties of cotangent lifts can be found in [Bullo and Lewis 2005b, Sections S1.2.2 and S1.3.4].

5.4 Pullbacks

In this section, we define the pullback of one time-varying vector field by the global flow of another. For our purposes, the importance of the pullback stems from its appearance in the nonlinear variation of constants formula, which describes the global flow of a time-varying vector field of the form $\Xi + \Upsilon$. The nonlinear variation of constants formula, presented below in Proposition 5.4.2, will play an important role in later chapters via the identities derived in Sections 5.5, 5.6, 5.7, 5.8, and 5.9. Throughout this section,

$$I = [a, b].$$

We begin with the following basic observation: Suppose that Ξ is locally integrably C^k, where $k \in \mathbb{N}^*$, and $q_0 \in \mathrm{dom}(\Phi^{\Xi}_{b,a})$. Then there exists a neighbourhood Q_0 of q_0 such that $\Phi^{\Xi}_{t,a}|Q_0$ is a C^k diffeomorphism of Q_0 onto its image for each $t \in I$. In fact, one can choose Q_0 to be any open subset of $\mathrm{dom}(\Phi^{\Xi}_{b,a})$ that contains q_0, since $\mathrm{dom}(\Phi^{\Xi}_{b,a}) \subseteq \mathrm{dom}(\Phi^{\Xi}_{t,a})$ for each $t \in I$.[5] This fact will be used implicitly in the remainder of this chapter. Indeed, it appears almost ubiquitously in the definitions and results that follow.

Definition 5.4.1. Suppose that Ξ is locally integrably C^1, $q_0 \in \mathrm{dom}(\Phi^{\Xi}_{b,a})$, and Q_0 is a neighbourhood of q_0 such that $\Phi^{\Xi}_{t,a}|Q_0$ is a C^1 diffeomorphism of Q_0 onto its image for each $t \in I$. The time-varying vector field

$$\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon) \in \mathscr{V}(I, Q_0)$$

defined by

$$\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon)(t, q) = T\Phi^{\Xi}_{a,t}(\Phi^{\Xi}_{t,a}(q_0)) \cdot \Upsilon(t, \Phi^{\Xi}_{t,a}(q_0))$$

is called the ***pullback*** of Υ by Ξ.[6] It is clear that

$$\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon)(t, q_0) \in T_{q_0}Q \cong T_{q_0}Q_0,$$

and thus $\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon)$ is well-defined.

[5] This conclusion does not rely fundamentally on the assumption that $I = [a, b]$. In fact, the same conclusion holds if $[a, b] \subseteq I$. We will not need this degree of generality, however.

[6] The "Ad" notation stems from similarities with Lie theory; see [Fulton and Harris 1991].

The next result is the nonlinear variation of constants formula.

Proposition 5.4.2. *Suppose that Ξ is locally integrably C^{k+1}, where $k \in \mathbb{Z}^*_{\geq 0}$, $q_0 \in \operatorname{dom}(\Phi^{\Xi}_{b,a})$, and Q_0 is a neighbourhood of q_0 such that $\Phi^{\Xi}_{t,a}|Q_0$ is a C^{k+1} diffeomorphism of Q_0 onto its image for each $t \in I$. Furthermore, suppose that Υ is locally integrably C^k. Then*

1. *$\Xi + \Upsilon$ is locally integrably C^k,*
2. *$\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon)$ is locally integrably C^k, and*
3. *The global flow of $\Xi + \Upsilon$ satisfies*

$$\Phi^{\Xi+\Upsilon}_{t,a}(q_0) = \Phi^{\Xi}_{t,a} \circ \Phi^{\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon)}_{t,a}(q_0)$$

whenever the left- and right-hand sides of the above equation are well-defined.

Proof. See [Bullo and Lewis 2005a, Proposition 9.6]. ∎

Finally, we prove three simple but useful technical lemmas.

Lemma 5.4.3. *Suppose that Ξ is locally integrably C^1, Υ is locally integrably C^1, and $s \in I$. Define the time-varying vector field $Z \in \mathscr{V}(I, Q)$ by*

$$Z(t, q) = [\Xi_t, \Upsilon_s](q).$$

Then Z is locally integrably C^0.

Proof. Choose a chart $(V, \boldsymbol{\psi})$ on Q. We must show that $Z_{\boldsymbol{\psi}}$ is locally integrably C^0 in the sense of Definition 2.2.19. To this end, observe that

$$\begin{aligned} Z_{\boldsymbol{\psi}}(t, \boldsymbol{q}) &= \boldsymbol{D}(\Upsilon_s)_{\boldsymbol{\psi}}(\boldsymbol{q}) \cdot (\Xi_t)_{\boldsymbol{\psi}}(\boldsymbol{q}) - \boldsymbol{D}(\Xi_t)_{\boldsymbol{\psi}}(\boldsymbol{q}) \cdot (\Upsilon_s)_{\boldsymbol{\psi}}(\boldsymbol{q}) \\ &= \boldsymbol{D}(\Upsilon_s)_{\boldsymbol{\psi}}(\boldsymbol{q}) \cdot \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}) - \boldsymbol{D}_2 \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}) \cdot (\Upsilon_s)_{\boldsymbol{\psi}}(\boldsymbol{q}). \end{aligned}$$

Since Ξ and Υ are locally integrably C^1, the first two criteria of Definition 2.2.19 are satisfied by composition. We now show that the third criterion is satisfied.

Choose a compact subset $K \subseteq \boldsymbol{\psi}(V)$. Since Υ is locally integrably C^1, Υ_s is a C^1 vector field on Q. Hence there exists $C \in \mathbb{R}_{\geq 0}$ such that

$$\|(\Upsilon_s)_{\boldsymbol{\psi}}(\boldsymbol{q})\|_{\mathbb{R}^\ell} \leq C \qquad \text{and} \qquad \|\boldsymbol{D}(\Upsilon_s)_{\boldsymbol{\psi}}(\boldsymbol{q})\| \leq C$$

for each $\boldsymbol{q} \in K$. Similarly, since Ξ is locally integrably C^1, there exists a function $\alpha \in L^1_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0})$ such that

$$\|\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\|_{\mathbb{R}^\ell} \leq \alpha(t) \qquad \text{and} \qquad \|\boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| \leq \alpha(t)$$

for a.a. $t \in I$ and each $\boldsymbol{q} \in K$. Thus

$$\begin{aligned}
\|Z_{\boldsymbol{\psi}}(t, \boldsymbol{q})\|_{\mathbb{R}^\ell} &\leq \|\boldsymbol{D}(\Upsilon_s)_{\boldsymbol{\psi}}(\boldsymbol{q})\| \, \|\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\|_{\mathbb{R}^\ell} + \|\boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| \, \|(\Upsilon_s)_{\boldsymbol{\psi}}(\boldsymbol{q})\|_{\mathbb{R}^\ell} \\
&\leq C\alpha(t) + C\alpha(t) \\
&= 2C\alpha(t)
\end{aligned}$$

for a.a. $t \in I$ and each $\boldsymbol{q} \in K$. This completes the proof. ∎

The next proposition states that linear maps commute with integration.

Proposition 5.4.4. *Suppose that E, F are finite-dimensional vector spaces,*

$$\lambda \in \mathrm{Hom}(E, F),$$

and $\gamma \in L^1(I, E)$. Then $\lambda \circ \gamma \in L^1(I, F)$ and

$$\lambda \cdot \int_I \gamma(\sigma) \, \mathrm{d}\sigma = \int_I \lambda \cdot \gamma(\sigma) \, \mathrm{d}\sigma.$$

Proof. See [Hunter 2010, Chapter 6.A]. ∎

Lemma 5.4.5. *Suppose that Ξ is locally integrably C^1 and Υ is locally integrably C^1. Define the time-varying vector field $Z \in \mathscr{V}(I, Q)$ by*

$$Z(t, q) = \int_a^t [\Xi_\sigma, \Upsilon_t](q) \, \mathrm{d}\sigma.$$

Then Z is well-defined and is locally integrably C^0.

Proof. By Lemma 5.4.3, Z is well-defined. Choose a chart $(V, \boldsymbol{\psi})$ on Q. We must show that $Z_{\boldsymbol{\psi}}$ is locally integrably C^0 in the sense of Definition 2.2.19. To this end, observe that

$$\begin{aligned}
Z_{\boldsymbol{\psi}}(t, \boldsymbol{q}) &= T\boldsymbol{\psi}(\boldsymbol{\psi}^{-1}(\boldsymbol{q})) \cdot Z(t, \boldsymbol{\psi}^{-1}(\boldsymbol{q})) \\
&= T\boldsymbol{\psi}(\boldsymbol{\psi}^{-1}(\boldsymbol{q})) \cdot \int_a^t [\Xi_\sigma, \Upsilon_t](\boldsymbol{\psi}^{-1}(\boldsymbol{q}))\, \mathrm{d}\sigma \\
&= \int_a^t T\boldsymbol{\psi}(\boldsymbol{\psi}^{-1}(\boldsymbol{q})) \cdot [\Xi_\sigma, \Upsilon_t](\boldsymbol{\psi}^{-1}(\boldsymbol{q}))\, \mathrm{d}\sigma \\
&= \int_a^t \boldsymbol{D}(\Upsilon_t)_{\boldsymbol{\psi}}(\boldsymbol{q}) \cdot (\Xi_\sigma)_{\boldsymbol{\psi}}(\boldsymbol{q})\, \mathrm{d}\sigma - \int_a^t \boldsymbol{D}(\Xi_\sigma)_{\boldsymbol{\psi}}(\boldsymbol{q}) \cdot (\Upsilon_t)_{\boldsymbol{\psi}}(\boldsymbol{q})\, \mathrm{d}\sigma \\
&= \int_a^t \boldsymbol{D}_2\Upsilon_{\boldsymbol{\psi}}(t, \boldsymbol{q}) \cdot \Xi_{\boldsymbol{\psi}}(\sigma, \boldsymbol{q})\, \mathrm{d}\sigma - \int_a^t \boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(\sigma, \boldsymbol{q}) \cdot \Upsilon_{\boldsymbol{\psi}}(t, \boldsymbol{q})\, \mathrm{d}\sigma \\
&= \boldsymbol{D}_2\Upsilon_{\boldsymbol{\psi}}(t, \boldsymbol{q}) \cdot \int_a^t \Xi_{\boldsymbol{\psi}}(\sigma, \boldsymbol{q})\, \mathrm{d}\sigma - \int_a^t \boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(\sigma, \boldsymbol{q})\, \mathrm{d}\sigma \cdot \Upsilon_{\boldsymbol{\psi}}(t, \boldsymbol{q})
\end{aligned}$$

by Proposition 5.4.4. Since Ξ and Υ are locally integrably C^1, the first two criteria of Definition 2.2.19 are satisfied by composition.[7] We now show that the third criterion is satisfied. Choose a compact subset $K \subseteq \boldsymbol{\psi}(V)$. Since Υ is locally integrably C^1, there exists $\alpha \in L^1_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0})$ such that

$$\|\Upsilon_{\boldsymbol{\psi}}(t, \boldsymbol{q})\|_{\mathbb{R}^\ell} \leq \alpha(t) \qquad \text{and} \qquad \|\boldsymbol{D}_2\Upsilon_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| \leq \alpha(t)$$

for a.a. $t \in I$ and each $\boldsymbol{q} \in K$. Similarly, since Ξ is locally integrably C^1, there exists $\beta \in L^1_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0})$ such that

$$\|\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\|_{\mathbb{R}^\ell} \leq \beta(t) \qquad \text{and} \qquad \|\boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| \leq \beta(t)$$

for a.a. $t \in I$ and each $\boldsymbol{q} \in K$. Thus

$$\begin{aligned}
&\|Z_{\boldsymbol{\psi}}(t, \boldsymbol{q})\|_{\mathbb{R}^\ell} \\
&\leq \|\boldsymbol{D}_2\Upsilon_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| \int_a^t \|\Xi_{\boldsymbol{\psi}}(\sigma, \boldsymbol{q})\|_{\mathbb{R}^\ell}\, \mathrm{d}\sigma + \int_a^t \|\boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(\sigma, \boldsymbol{q})\|\, \mathrm{d}\sigma\, \|\Upsilon_{\boldsymbol{\psi}}(t, \boldsymbol{q})\|_{\mathbb{R}^\ell}
\end{aligned}$$

[7] Here we are using the fact that the maps

$$t \mapsto \int_a^t \Xi_{\boldsymbol{\psi}}(\sigma, \boldsymbol{q})\, \mathrm{d}\sigma \qquad \text{and} \qquad t \mapsto \int_a^t \boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(\sigma, \boldsymbol{q})\, \mathrm{d}\sigma$$

of I into $\mathbb{R}^\ell$ are AC, hence measurable.

$$\begin{aligned}
&\leq \alpha(t) \int_a^t \beta(\sigma)\,\mathrm{d}\sigma + \alpha(t) \int_a^t \beta(\sigma)\,\mathrm{d}\sigma \\
&\leq \alpha(t) \int_a^b \beta(\sigma)\,\mathrm{d}\sigma + \alpha(t) \int_a^b \beta(\sigma)\,\mathrm{d}\sigma \\
&= \alpha(t)\|\beta\|_1 + \alpha(t)\|\beta\|_1 \\
&= 2\alpha(t)\|\beta\|_1
\end{aligned}$$

for a.a. $t \in I$ and each $\boldsymbol{q} \in K$. This completes the proof. ∎

Lemma 5.4.6. *Suppose that Ξ is locally integrably C^1, $q_0 \in \mathrm{dom}(\Phi^{\Xi}_{b,a})$, and Q_0 is a neighbourhood of q_0 such that $\Phi^{\Xi}_{t,a}|Q_0$ is a C^1 diffeomorphism of Q_0 onto its image for each $t \in I$. Furthermore, suppose that Υ is locally integrably C^1 and $s \in I$. Define the map $\gamma : I \to T_{q_0}Q$ by*

$$\gamma(t) = \mathrm{Ad}^{\Xi}_{Q_0}([\Xi_t, \Upsilon_s])(t, q_0).$$

Then $\gamma \in L^1(I, T_{q_0}Q)$.

Proof. Define the time-varying vector field $Z \in \mathscr{V}(I, Q)$ by

$$Z(t, q) = [\Xi_t, \Upsilon_s](q).$$

By Proposition 5.4.2 and Lemma 5.4.3, $\mathrm{Ad}^{\Xi}_{Q_0}(Z)$ is locally integrably C^0. To complete the proof, just observe that

$$\begin{aligned}
\gamma(t) &= \mathrm{Ad}^{\Xi}_{Q_0}([\Xi_t, \Upsilon_s])(t, q_0) \\
&= T\Phi^{\Xi}_{a,t}(\Phi^{\Xi}_{t,a}(q_0)) \cdot [\Xi_t, \Upsilon_s](\Phi^{\Xi}_{t,a}(q_0)) \\
&= T\Phi^{\Xi}_{a,t}(\Phi^{\Xi}_{t,a}(q_0)) \cdot Z(t, \Phi^{\Xi}_{t,a}(q_0)) \\
&= \mathrm{Ad}^{\Xi}_{Q_0}(Z)(t, q_0)
\end{aligned}$$

for each $t \in I$. ∎

5.5 Pullbacks involving lifts

In this section, we derive reductive formulas for pullbacks involving vertical lifts and tangent lifts. Throughout this section,

$$I = [a, b].$$

Lemma 5.5.1. *Suppose that Ξ is locally integrably C^2, $q_0 \in \mathrm{dom}(\Phi^{\Xi}_{b,a})$, and Q_0 is a neighbourhood of q_0 such that $\Phi^{\Xi}_{t,a}|Q_0$ is a C^2 diffeomorphism of Q_0 onto its image for each $t \in I$. Furthermore, suppose that Υ is locally integrably C^1. Define $X = \mathrm{tlft}(\Xi)$ and $Y = \mathrm{tlft}(\Upsilon)$. Then*

$$\mathrm{Ad}^{X}_{TQ_0}(Y)(t, v_{q_0}) = \mathrm{tlft}(\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon))(t, v_{q_0})$$

for each $(t, v_{q_0}) \in I \times T_{q_0}Q$.

Proof. Observe that

- $\mathrm{Ad}^{X}_{TQ_0}(Y)$ is well-defined by Lemma 5.2.6 and
- $\mathrm{tlft}(\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon))$ is well-defined by Proposition 5.4.2.

By [Abraham et al. 1988, Exercise 3.3B] and Theorem 5.2.4, we have

$$\begin{aligned}
\mathrm{Ad}^{X}_{TQ_0}(Y)(t, v_{q_0}) &= T\Phi^{X}_{a,t}(\Phi^{X}_{t,a}(v_{q_0})) \circ Y(t, \Phi^{X}_{t,a}(v_{q_0})) \\
&= T\Phi^{X}_{a,t}(\Phi^{X}_{t,a}(v_{q_0})) \circ \mathrm{tlft}(\Upsilon)(t, \Phi^{X}_{t,a}(v_{q_0})) \\
&= T\Phi^{X}_{a,t}(\Phi^{X}_{t,a}(v_{q_0})) \circ s_{Q_0} \circ T\Upsilon_t(\Phi^{\Xi}_{t,a}(v_{q_0})) \circ \Phi^{X}_{t,a}(v_{q_0}) \\
&= TT\Phi^{\Xi}_{a,t} \circ s_{Q_0} \circ T\Upsilon_t(\Phi^{\Xi}_{t,a}(v_{q_0})) \circ T\Phi^{\Xi}_{t,a}(q_0) \cdot v_{q_0} \\
&= s_{Q_0} \circ TT\Phi^{\Xi}_{a,t} \circ T\Upsilon_t(\Phi^{\Xi}_{t,a}(v_{q_0})) \circ T\Phi^{\Xi}_{t,a}(q_0) \cdot v_{q_0} \\
&= s_{Q_0} \circ T(T\Phi^{\Xi}_{a,t} \circ \Upsilon_t \circ \Phi^{\Xi}_{t,a})(q_0) \cdot v_{q_0} \\
&= s_{Q_0} \circ T(\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon)_t)(q_0) \cdot v_{q_0} \\
&= \mathrm{tlft}(\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon))(t, v_{q_0})
\end{aligned}$$

for each $(t, v_{q_0}) \in I \times T_{q_0}Q$. This completes the proof. ∎

Lemma 5.5.2. *Suppose that Ξ is locally integrably C^2, $q_0 \in \operatorname{dom}(\Phi^{\Xi}_{b,a})$, and Q_0 is a neighbourhood of q_0 such that $\Phi^{\Xi}_{t,a}|Q_0$ is a C^2 diffeomorphism of Q_0 onto its image for each $t \in I$. Define $X = \operatorname{tlft}(\Xi)$ and $Y = \operatorname{vlft}(\Upsilon)$. Then*

$$\operatorname{Ad}^X_{TQ_0}(Y)(t, v_{q_0}) = \operatorname{vlft}(\operatorname{Ad}^{\Xi}_{Q_0}(\Upsilon))(t, v_{q_0})$$

for each $(t, v_{q_0}) \in I \times T_{q_0}Q$.

Proof. Observe that $\operatorname{Ad}^X_{TQ_0}(Y)$ is well-defined by Lemma 5.2.6. By [Lee 2003, Proposition 3.12], Theorem 5.2.4, and linearity of each $\Phi^X_{a,t}|T_{\Phi^{\Xi}_{t,a}(q_0)}Q$, we have

$$\begin{aligned}
\operatorname{Ad}^X_{TQ_0}(Y)(t, v_{q_0}) &= T\Phi^X_{a,t}(\Phi^X_{t,a}(v_{q_0})) \cdot Y(t, \Phi^X_{t,a}(v_{q_0})) \\
&= T\Phi^X_{a,t}(\Phi^X_{t,a}(v_{q_0})) \cdot \operatorname{vlft}(\Upsilon)(t, \Phi^X_{t,a}(v_{q_0})) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_0 \Phi^X_{a,t}\left(\Phi^X_{t,a}(v_{q_0}) + s\Upsilon(t, \Phi^{\Xi}_{t,a}(q_0))\right) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_0 v_{q_0} + \Phi^X_{a,t}(s\Upsilon(t, \Phi^{\Xi}_{t,a}(q_0))) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_0 v_{q_0} + s\Phi^X_{a,t}(\Upsilon(t, \Phi^{\Xi}_{t,a}(q_0))) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_0 v_{q_0} + sT\Phi^{\Xi}_{a,t}(\Phi^{\Xi}_{t,a}(q_0)) \cdot \Upsilon(t, \Phi^{\Xi}_{t,a}(q_0)) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_0 v_{q_0} + s\operatorname{Ad}^{\Xi}_{Q_0}(\Upsilon)(t, q_0) \\
&= \operatorname{vlft}(\operatorname{Ad}^{\Xi}_{Q_0}(\Upsilon))(t, v_{q_0})
\end{aligned}$$

for each $(t, v_{q_0}) \in I \times T_{q_0}Q$. This completes the proof. ∎

Lemma 5.5.3. *Suppose that Ξ is locally integrably C^1. Define $X = \operatorname{vlft}(\Xi)$ and $Y = \operatorname{vlft}(\Upsilon)$. Then*

$$\operatorname{Ad}^X_{TQ}(Y)(t, v_q) = Y(t, v_q)$$

for each $(t, v_q) \in I \times TQ$.

Proof. Observe that $\mathrm{Ad}^X_{TQ}(Y)$ is well-defined by Lemma 5.1.5 and Theorem 5.1.6. By [Lee 2003, Proposition 3.12], we have

$$\begin{aligned}
\mathrm{Ad}^X_{TQ}(Y)(t, v_q) &= T\Phi^X_{a,t}(\Phi^X_{t,a}(v_q)) \cdot Y(t, \Phi^X_{t,a}(v_q)) \\
&= T\Phi^{\mathrm{vlft}(\Xi)}_{a,t}(\Phi^X_{t,a}(v_q)) \cdot \mathrm{vlft}(\Upsilon)(t, \Phi^{\mathrm{vlft}(\Xi)}_{t,a}(v_q)) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_0 \Phi^{\mathrm{vlft}(\Xi)}_{a,t}(\Phi^{\mathrm{vlft}(\Xi)}_{t,a}(v_q) + s\Upsilon(t, q)) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_0 \Phi^{\mathrm{vlft}(\Xi)}_{a,t}\left(v_q + \int_a^t \Xi(\sigma, q)\,\mathrm{d}\sigma + s\Upsilon(t, q)\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_0 \left(v_q + \int_a^t \Xi(\sigma, q)\,\mathrm{d}\sigma + s\Upsilon(t, q) + \int_t^a \Xi(\sigma, q)\,\mathrm{d}\sigma\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_0 \left(v_q + \int_a^t \Xi(\sigma, q)\,\mathrm{d}\sigma + s\Upsilon(t, q) - \int_a^t \Xi(\sigma, q)\,\mathrm{d}\sigma\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_0 (v_q + s\Upsilon(t, q)) \\
&= \mathrm{vlft}(\Upsilon)(t, v_q) \\
&= Y(t, v_q)
\end{aligned}$$

for each $(t, v_q) \in I \times TQ$. This completes the proof. ∎

The case where $X = \mathrm{vlft}(\Xi)$ and $Y = \mathrm{tlft}(\Upsilon)$ is treated below in Section 5.8.

5.6 A particular global flow

In this section, we compute the global flow of $X + Y$, where X is a tangent lift and Y is a vertical lift. Throughout this section,

$$I = [a, b].$$

Lemma 5.6.1. *Suppose that Ξ is locally integrably C^2, $q_0 \in \mathrm{dom}(\Phi^{\Xi}_{b,a})$, and Q_0 is a neighbourhood of q_0 such that $\Phi^{\Xi}_{t,a}|Q_0$ is a C^2 diffeomorphism of Q_0 onto its image for each $t \in I$. Furthermore, suppose that Υ is locally integrably C^0. Define $X = \mathrm{tlft}(\Xi)$ and $Y = \mathrm{vlft}(\Upsilon)$. If*

$$\Phi^{X+Y}_{t,a}(v_{q_0})$$

is well-defined for each $(t, v_{q_0}) \in I \times T_{q_0}Q$*, then*

$$\Phi_{t,a}^{X+Y}(v_{q_0}) = T\Phi_{t,a}^{\Xi}(q_0) \cdot v_{q_0} + \int_a^t T\Phi_{t,a}^{\Xi}(q_0) \cdot \mathrm{Ad}_{Q_0}^{\Xi}(\Upsilon)(\sigma, q_0)\, \mathrm{d}\sigma$$

for each $(t, v_{q_0}) \in I \times T_{q_0}Q$*.*

Proof. By Proposition 5.4.2, $\mathrm{Ad}_{Q_0}^{\Xi}(\Upsilon)$ is locally integrably C^0. By Theorem 5.1.6,

$$\Phi_{t,a}^{\mathrm{vlft}(\mathrm{Ad}_{Q_0}^{\Xi}(\Upsilon))}(v_{q_0}) = v_{q_0} + \int_a^t \mathrm{Ad}_{Q_0}^{\Xi}(\Upsilon)(\sigma, q_0)\, \mathrm{d}\sigma$$

for each $(t, v_{q_0}) \in I \times T_{q_0}Q$. Using this fact, together with Theorem 5.2.4, Proposition 5.4.2, Proposition 5.4.4, Lemma 5.5.2, and linearity of each $\Phi_{t,a}^{X}|T_{q_0}Q$, we have

$$\begin{aligned}
\Phi_{t,a}^{X+Y}(v_{q_0}) &= \Phi_{t,a}^{X} \circ \Phi_{t,a}^{\mathrm{Ad}_{TQ_0}^{X}(Y)}(v_{q_0}) \\
&= \Phi_{t,a}^{\mathrm{tlft}(\Xi)} \circ \Phi_{t,a}^{\mathrm{Ad}_{TQ_0}^{\mathrm{tlft}(\Xi)}(\mathrm{vlft}(\Upsilon))}(v_{q_0}) \\
&= \Phi_{t,a}^{\mathrm{tlft}(\Xi)} \circ \Phi_{t,a}^{\mathrm{vlft}(\mathrm{Ad}_{Q_0}^{\Xi}(\Upsilon))}(v_{q_0}) \\
&= \Phi_{t,a}^{\mathrm{tlft}(\Xi)} \cdot \left(v_{q_0} + \int_a^t \mathrm{Ad}_{Q_0}^{\Xi}(\Upsilon)(\sigma, q_0)\, \mathrm{d}\sigma \right) \\
&= T\Phi_{t,a}^{\Xi}(q_0) \cdot v_{q_0} + T\Phi_{t,a}^{\Xi}(q_0) \cdot \int_a^t \mathrm{Ad}_{Q_0}^{\Xi}(\Upsilon)(\sigma, q_0)\, \mathrm{d}\sigma \\
&= T\Phi_{t,a}^{\Xi}(q_0) \cdot v_{q_0} + \int_a^t T\Phi_{t,a}^{\Xi}(q_0) \cdot \mathrm{Ad}_{Q_0}^{\Xi}(\Upsilon)(\sigma, q_0)\, \mathrm{d}\sigma
\end{aligned}$$

for each $(t, v_{q_0}) \in I \times T_{q_0}Q$. This completes the proof. ∎

5.7 Time derivatives of pullbacks

In this section, we compute time derivatives of pullbacks. This leads to useful expressions for pullbacks, which are phrased in terms of integrated Lie brackets. While the developments in this section are similar to those in [Abraham et al. 1988, Section 4.2], the cited work only deals with C^1 time-varying vector fields. Our computational approach is also completely different from that of Abraham et al. [1988], as it involves tangent lifts. Throughout this section,

$$I = [a, b].$$

Lemma 5.7.1. *Suppose that Ξ is locally integrably C^2, $q_0 \in \operatorname{dom}(\Phi^{\Xi}_{b,a})$, and Q_0 is a neighbourhood of q_0 such that $\Phi^{\Xi}_{t,a}|Q_0$ is a C^2 diffeomorphism of Q_0 onto its image for each $t \in I$. Furthermore, suppose that $v_{q_0} \in T_{q_0}Q$. Then*

$$\begin{aligned}&\left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t T\Phi^{\Xi}_{a,\tau}(\Phi^{\Xi}_{\tau,a}(q_0)) \cdot T\Phi^{\Xi}_{t,a}(q_0) \cdot v_{q_0}\\ &= -TT\Phi^{\Xi}_{a,t}(w) \circ s_{Q_0} \circ T\Xi_t(\Phi^{\Xi}_{t,a}(q_0)) \circ T\Phi^{\Xi}_{t,a}(q_0) \cdot v_{q_0}\end{aligned}$$

for a.a. $t \in I$, where $w = T\Phi^{\Xi}_{t,a}(q_0) \cdot v_{q_0}$.

Proof. Define $\gamma : I \to TQ$ by

$$\gamma(t) = T\Phi^{\Xi}_{a,t}(\Phi^{\Xi}_{t,a}(q_0)) \circ T\Phi^{\Xi}_{t,a}(q_0) \cdot v_{q_0}.$$

Clearly, $\dot{\gamma}(t) = 0$ for each $t \in I$. By the chain rule, we have

$$\dot{\gamma}(t) = \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t T\Phi^{\Xi}_{a,\tau}(\Phi^{\Xi}_{\tau,a}(q_0)) \circ T\Phi^{\Xi}_{t,a}(q_0) \cdot v_{q_0} + TT\Phi^{\Xi}_{a,t}(w) \circ \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t T\Phi^{\Xi}_{\tau,a}(q_0) \cdot v_{q_0}$$

for a.a. $t \in I$, where $w = T\Phi^{\Xi}_{t,a}(q_0) \cdot v_{q_0}$. By Theorem 5.2.4, the second term is equal to

$$\begin{aligned}&TT\Phi^{\Xi}_{a,t}(w) \circ \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t T\Phi^{\Xi}_{\tau,a}(q_0) \cdot v_{q_0}\\ &= TT\Phi^{\Xi}_{a,t}(w) \circ \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \Phi^{\mathrm{tlft}(\Xi)}_{\tau,a}(v_{q_0})\\ &= TT\Phi^{\Xi}_{a,t}(w) \circ \mathrm{tlft}(\Xi)(t, T\Phi^{\Xi}_{t,a}(q_0) \cdot v_{q_0})\\ &= TT\Phi^{\Xi}_{a,t}(w) \circ s_{Q_0} \circ T\Xi_t(\Phi^{\Xi}_{t,a}(q_0)) \circ T\Phi^{\Xi}_{t,a}(q_0) \cdot v_{q_0}.\end{aligned}$$

This completes the proof. ∎

Lemma 5.7.2. *Suppose that Ξ is locally integrably C^2, $q_0 \in \operatorname{dom}(\Phi^{\Xi}_{b,a})$, and Q_0 is a neighbourhood of q_0 such that $\Phi^{\Xi}_{t,a}|Q_0$ is a C^2 diffeomorphism of Q_0 onto its image for each $t \in I$. Furthermore, suppose that Υ is locally integrably C^1 and $s \in I$. Then*

$$\left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon_s)(\tau, q_0) = \mathrm{Ad}^{\Xi}_{Q_0}([\Xi_t, \Upsilon_s])(t, q_0)$$

for a.a. $t \in I$.

Proof. By definition,

$$\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon_s)(t,q_0) = T\Phi^{\Xi}_{a,t}(\Phi^{\Xi}_{t,a}(q_0)) \circ \Upsilon_s \circ \Phi^{\Xi}_{t,a}(q_0)$$

for each $t \in I$. By the chain rule,

$$\begin{aligned}
&\left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon_s)(\tau,q_0) \\
&= TT\Phi^{\Xi}_{a,t}(w) \circ \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \Upsilon_s \circ \Phi^{\Xi}_{\tau,a}(q_0) + \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t T\Phi^{\Xi}_{a,\tau}(\Phi^{\Xi}_{\tau,a}(q_0)) \circ \Upsilon_s \circ \Phi^{\Xi}_{t,a}(q_0)
\end{aligned}$$

for a.a. $t \in I$, where $w = \Upsilon_s \circ \Phi^{\Xi}_{t,a}(q_0)$. The first term is equal to

$$TT\Phi^{\Xi}_{a,t}(w) \circ \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \Upsilon_s \circ \Phi^{\Xi}_{\tau,a}(q_0) = TT\Phi^{\Xi}_{a,t}(w) \circ T\Upsilon_s(\Phi^{\Xi}_{t,a}(q_0)) \circ \Xi_t \circ \Phi^{\Xi}_{t,a}(q_0).$$

Invoking Lemma 5.7.1 with $v_{q_0} = \mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon_s)(t,q_0)$, the second term is equal to

$$\begin{aligned}
&\left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t T\Phi^{\Xi}_{a,\tau}(\Phi^{\Xi}_{\tau,a}(q_0)) \circ \Upsilon_s \circ \Phi^{\Xi}_{t,a}(q_0) \\
&= -TT\Phi^{\Xi}_{a,t}(w) \circ s_{Q_0} \circ T\Xi_t(\Phi^{\Xi}_{t,a}(q_0)) \circ \Upsilon_s \circ \Phi^{\Xi}_{t,a}(q_0).
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon_s)(\tau,q_0) \\
&= TT\Phi^{\Xi}_{a,t}(w) \circ [T\Upsilon_s(\Phi^{\Xi}_{t,a}(q_0)) \circ \Xi_t \circ \Phi^{\Xi}_{t,a}(q_0) \\
&\qquad - s_{Q_0} \circ T\Xi_t(\Phi^{\Xi}_{t,a}(q_0)) \circ \Upsilon_s \circ \Phi^{\Xi}_{t,a}(q_0)] \\
&= TT\Phi^{\Xi}_{a,t}(w) \cdot K(q_0).
\end{aligned}$$

By [Abraham et al. 1988, Exercise 4.2K], we have $K(q_0) \in V_wTQ$ and

$$K(q_0) = \mathrm{vlft}_w \circ [\Xi_t, \Upsilon_s] \circ \Phi^{\Xi}_{t,a}(q_0).$$

Finally, we invoke Lemma 5.1.8 to obtain

$$\begin{aligned}
&\mathrm{vlft}^{-1}_{\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon_s)(t,q_0)} \circ \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon_s)(\tau,q_0) \\
&= \mathrm{vlft}^{-1}_{T\Phi^{\Xi}_{a,t}(\Phi^{\Xi}_{t,a}(q_0))\circ\Upsilon_s\circ\Phi^{\Xi}_{t,a}(q_0)} \circ \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon_s)(\tau,q_0)
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{vlft}^{-1}_{T\Phi^{\Xi}_{a,t}(\Phi^{\Xi}_{t,a}(q_0))\cdot w} \circ TT\Phi^{\Xi}_{a,t}(w) \cdot K(q_0) \\
&= T\Phi^{\Xi}_{a,t}(\Phi^{\Xi}_{t,a}(q_0)) \circ \mathrm{vlft}^{-1}_{w} \circ K(q_0) \\
&= T\Phi^{\Xi}_{a,t}(\Phi^{\Xi}_{t,a}(q_0)) \circ \mathrm{vlft}^{-1}_{w} \circ \mathrm{vlft}_{w} \circ [\Xi_t, \Upsilon_s] \circ \Phi^{\Xi}_{t,a}(q_0) \\
&= T\Phi^{\Xi}_{a,t}(\Phi^{\Xi}_{t,a}(q_0)) \circ [\Xi_t, \Upsilon_s] \circ \Phi^{\Xi}_{t,a}(q_0) \\
&= \mathrm{Ad}^{\Xi}_{Q_0}([\Xi_t, \Upsilon_s])(t, q_0).
\end{aligned}
$$

This completes the proof.[8] ∎

The next corollary is the integral form of Lemma 5.7.2.

Corollary 5.7.3. *Suppose that Ξ is locally integrably C^2, $q_0 \in \mathrm{dom}(\Phi^{\Xi}_{b,a})$, and Q_0 is a neighbourhood of q_0 such that $\Phi^{\Xi}_{t,a}|Q_0$ is a C^2 diffeomorphism of Q_0 onto its image for each $t \in I$. Furthermore, suppose that Υ is locally integrably C^1 and $s \in I$. Then*

$$
\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon_s)(t, q_0) = \Upsilon_s(q_0) + \int_a^t \mathrm{Ad}^{\Xi}_{Q_0}([\Xi_\sigma, \Upsilon_s])(\sigma, q_0)\, \mathrm{d}\sigma
$$

for each $t \in I$. In particular, with $s = t$, we have

$$
\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon)(t, q_0) = \Upsilon_t(q_0) + \int_a^t \mathrm{Ad}^{\Xi}_{Q_0}([\Xi_\sigma, \Upsilon_t])(\sigma, q_0)\, \mathrm{d}\sigma
$$

for each $t \in I$.

Proof. By Lemma 5.4.6, the map

$$
t \mapsto \mathrm{Ad}^{\Xi}_{Q_0}([\Xi_t, \Upsilon_s])(t, q_0)
$$

of I into $T_{q_0}Q$ is an element of $L^1(I, T_{q_0}Q)$. Integrating, we obtain

$$
\begin{aligned}
\mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon_s)(t, q_0) &= \mathrm{Ad}^{\Xi}_{Q_0}(\Upsilon_s)(a, q_0) + \int_a^t \mathrm{Ad}^{\Xi}_{Q_0}([\Xi_\sigma, \Upsilon_s])(\sigma, q_0)\, \mathrm{d}\sigma \\
&= \Upsilon_s(q_0) + \int_a^t \mathrm{Ad}^{\Xi}_{Q_0}([\Xi_\sigma, \Upsilon_s])(\sigma, q_0)\, \mathrm{d}\sigma
\end{aligned}
$$

for each $t \in I$. This completes the proof. ∎

[8] Here, we are using the fact that each pointwise vertical lift is a canonical vector space isomorphism.

To conclude this section, we note that the results of Lemma 5.7.2 and Corollary 5.7.3 match certain formulas appearing in the proof of [Bullo and Lewis 2005a, Proposition 9.9].

5.8 Pullbacks involving lifts: An integral formula

In this section, we derive an integral formula for pullbacks of the general form $\mathrm{Ad}^X(Y)$, where X is a vertical lift and Y is a tangent lift. Throughout this section,

$$I = [a, b].$$

Lemma 5.8.1. *Suppose that Ξ is locally integrably C^2, $q_0 \in \mathrm{dom}(\Phi^{\Xi}_{b,a})$, and Q_0 is a neighbourhood of q_0 such that $\Phi^{\Xi}_{t,a}|Q_0$ is a C^2 diffeomorphism of Q_0 onto its image for each $t \in I$. Furthermore, suppose that Υ is locally integrably C^2. Define $X = \mathrm{vlft}(\Xi)$ and $Y = \mathrm{tlft}(\Upsilon)$. Then*

$$\mathrm{Ad}^X_{TQ_0}(Y)(t, v_{q_0}) = \mathrm{tlft}(\Upsilon)(t, v_{q_0}) + \int_a^t \mathrm{vlft}([\Xi_\sigma, \Upsilon_t])(v_{q_0})\,\mathrm{d}\sigma$$

for each $(t, v_{q_0}) \in I \times T_{q_0}Q$.

Proof. Observe that

- X is locally integrably C^2 by Lemma 5.1.5 and
- Y is locally integrably C^1 by Lemma 5.2.6.

Invoking Corollary 5.7.3, we have

$$\begin{aligned}
\mathrm{Ad}^X_{TQ_0}(Y)(t, v_{q_0}) &= Y_t(v_{q_0}) + \int_a^t \mathrm{Ad}^X_{TQ_0}([X_\sigma, Y_t])(\sigma, v_{q_0})\,\mathrm{d}\sigma \\
&= \mathrm{tlft}(\Upsilon)(t, v_{q_0}) + \int_a^t \mathrm{Ad}^{\mathrm{vlft}(\Xi)}_{TQ_0}([\mathrm{vlft}(\Xi)_\sigma, \mathrm{tlft}(\Upsilon)_t])(\sigma, v_{q_0})\,\mathrm{d}\sigma \\
&= \mathrm{tlft}(\Upsilon)(t, v_{q_0}) + \int_a^t \mathrm{Ad}^{\mathrm{vlft}(\Xi)}_{TQ_0}([\mathrm{vlft}(\Xi_\sigma), \mathrm{tlft}(\Upsilon_t)])(\sigma, v_{q_0})\,\mathrm{d}\sigma
\end{aligned}$$

for each $(t, v_{q_0}) \in I \times T_{q_0}Q$. By [Grabowski and Urbański 1995, Theorem 2.4],

$$[\mathrm{vlft}(A), \mathrm{tlft}(B)] = \mathrm{vlft}([A, B]).$$

whenever A is a C^1 vector field on Q and B is a C^2 vector field on Q. Thus

$$\mathrm{Ad}^X_{TQ_0}(Y)(t, v_{q_0}) = \mathrm{tlft}(\Upsilon)(t, v_{q_0}) + \int_a^t \mathrm{Ad}^{\mathrm{vlft}(\Xi)}_{TQ_0}(\mathrm{vlft}([\Xi_\sigma, \Upsilon_t]))(\sigma, v_{q_0})\, \mathrm{d}\sigma.$$

By Lemma 5.5.3, we have

$$\begin{aligned}\mathrm{Ad}^X_{TQ_0}(Y)(t, v_{q_0}) &= \mathrm{tlft}(\Upsilon)(t, v_{q_0}) + \int_a^t \mathrm{vlft}([\Xi_\sigma, \Upsilon_t])(\sigma, v_{q_0})\, \mathrm{d}\sigma \\ &= \mathrm{tlft}(\Upsilon)(t, v_{q_0}) + \int_a^t \mathrm{vlft}([\Xi_\sigma, \Upsilon_t])(v_{q_0})\, \mathrm{d}\sigma.\end{aligned}$$

This completes the proof. ∎

5.9 Scalar parameter derivatives of pullbacks

In this section, we compute parameter derivatives of pullbacks, where the parameter lies in an interval R. This leads to useful expressions for parameter derivatives of pullbacks, which are phrased in terms of integrated Lie brackets. While the developments in this section are similar to those in [Tretiyak 1997], the cited work only deals with C^∞ dependence on scalar parameters. Our computational approach is also completely different, as it is not based on the chronological calculus formalism [Agrachev and Sachkov 2004]. Throughout this section,

$$I = [a, b].$$

5.9.1 Time-varying vector fields with scalar parameters

We begin by establishing the necessary terminology and notation.

Definition 5.9.1. Suppose that V is a nonempty open subset of $\mathbb{R}^\ell$ and

$$\boldsymbol{f} : I \times V \times R \to \mathbb{R}^\ell.$$

We say that $\boldsymbol{f}$ is ***locally integrably*** $C^{j,k}$, where $j, k \in \mathbb{N}^*$, if

- For each $r \in R$, the map $(t, \boldsymbol{q}) \mapsto \boldsymbol{f}(t, \boldsymbol{q}, r)$ of $I \times V$ into $\mathbb{R}^\ell$ is locally integrably C^j,
- For each $t \in I$, the map $(\boldsymbol{q}, r) \mapsto \boldsymbol{f}(t, \boldsymbol{q}, r)$ of $V \times R$ into $\mathbb{R}^\ell$ is C^k, and
- For each $r \in R$, the map

  $$(t, \boldsymbol{q}) \mapsto \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \boldsymbol{f}(t, \boldsymbol{q}, \rho)$$

  of $I \times V$ into $\mathbb{R}^\ell$ is locally integrably C^{k-1}.

Definition 5.9.2. Suppose that $X : I \times Q \times R \to TQ$. We say that X is a ***time-varying vector field on*** Q ***with scalar parameters*** if $\pi_{TQ} \circ X(t, q, r) = q$ for each $(t, q, r) \in I \times Q \times R$. The set of all such maps is denoted by $\mathscr{V}(I, Q, R)$. We obtain $X^r \in \mathscr{V}(I, Q)$ by freezing the scalar parameter at $r \in R$. That is,

$$X^r(t, q) = X(t, q, r).$$

Given a chart $(V, \boldsymbol{\psi})$ on Q, the ***local representative*** of X in $(V, \boldsymbol{\psi})$ is the map

$$X_{\boldsymbol{\psi}} : I \times \boldsymbol{\psi}(V) \times R \to \mathbb{R}^\ell$$

defined by

$$X_{\boldsymbol{\psi}}(t, \boldsymbol{q}, r) = T\boldsymbol{\psi}(\boldsymbol{\psi}^{-1}(\boldsymbol{q})) \cdot X(t, \boldsymbol{\psi}^{-1}(\boldsymbol{q}), r).$$

In what follows, $X \in \mathscr{V}(I, Q, R)$.

Definition 5.9.3. We say that X is ***locally integrably*** $C^{j,k}$, where $j, k \in \mathbb{N}^*$, if $X_{\boldsymbol{\psi}}$ is locally integrably $C^{j,k}$ for each chart $(V, \boldsymbol{\psi})$ on Q.

One can show that X is locally integrably $C^{j,k}$ if and only if $X_{\boldsymbol{\psi}}$ is locally integrably $C^{j,k}$ for each chart $(V, \boldsymbol{\psi}) \in \mathscr{A}_Q$, where $\mathscr{A}_Q$ is a compatible atlas on Q.

Furthermore, if X is locally integrably $C^{j,k}$, there exists a maximally-defined solution of (X^r, t_0, q_0) for each $(t_0, q_0, r) \in I \times Q \times R$. Indeed, this follows from the fact that each time-varying vector field X^r is locally integrably C^j by definition.

Definition 5.9.4. Suppose that X is locally integrably $C^{j,k}$. Define

$$\mathrm{dom}(\Phi^X) = \{(t, t_0, q_0, r) \in I \times I \times Q \times R \ : \ t \in I^{X^r}(t_0, q_0)\}.$$

The ***global flow*** of X is the map $\Phi^X : \mathrm{dom}(\Phi^X) \to Q$ that sends (t, t_0, q_0, r) to

$$\Phi^{X^r}_{t,t_0}(q_0) = \mu^{X^r}(t, t_0, q_0).$$

Theorem 5.9.5. *Suppose that X is locally integrably $C^{j,k}$. Then*

1. $\mathrm{dom}(\Phi^X)$ *is an open subset of* $I \times I \times Q \times R$*,*
2. *The map* Φ^X *is continuous,*
3. $\Phi^{X^r}_{t,t_0}$ *is* C^j *for each* $(t, t_0, r) \in I \times I \times R$*, and*
4. *The map* $\rho \mapsto \Phi^{X^\rho}_{t,t_0}(q_0)$ *of*

$$\{r \in R \ : \ (t, t_0, q_0, r) \in \mathrm{dom}(\Phi^X)\}$$

into Q *is* C^k *for each* $(t, t_0, q_0) \in I \times I \times Q$*.*

Proof. This follows from [McShane 1944, Chapter IX, Section 69.4], together with a globalization procedure analogous to the one employed in Section 2.3. ∎

Using the above theorem, one can show that if $q_0 \in \mathrm{dom}(\Phi^{X^{r_0}}_{b,a})$ for some $r_0 \in R$, then there exists a product neighbourhood $Q_0 \times R_0$ of (q_0, r_0) such that $\Phi^{X^r}_{t,a}|Q_0$ is a C^j diffeomorphism of Q_0 onto its image for each $(t, r) \in I \times R_0$, and the map

$$\rho \mapsto \Phi^{X^\rho}_{t,a}(q_0)$$

of R_0 into Q is C^k for each $t \in I$. This fact, which follows from an argument similar to the one used in Section 5.4 to obtain the neighbourhood Q_0, will be used implicitly in the remainder of this chapter.

Definition 5.9.6. Suppose that X is locally integrably $C^{j,k}$. We say that

$$\mathrm{tlft}(X) \in \mathscr{V}(I, TQ, R),$$

defined by

$$\mathrm{tlft}(X)(t, q, r) = \mathrm{tlft}(X^r)(t, q),$$

is the ***tangent lift*** of X.

Definition 5.9.7. Suppose that X is locally integrably $C^{j,k}$. Define

$$Z_X \in \mathscr{V}(I, Q, R)$$

by

$$Z_X(t, q, r) = \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r X(t, q, \rho).$$

By construction, Z_X^r is locally integrably C^{k-1} for each $r \in R$.

One can easily verify that $\mathrm{tlft}(X)$ and Z_X have the following properties:

- If X is locally integrably $C^{j+1,k+1}$, then $\mathrm{tlft}(X)$ is locally integrably $C^{j,k}$;
- For each $r \in R$, we have $\mathrm{tlft}(X)^r = \mathrm{tlft}(X^r)$;
- For each $r \in R$, we have $Z^r_{\mathrm{tlft}(X)} = \mathrm{tlft}(Z^r_X)$ provided that $k \geq 2$.

5.9.2 Scalar parameter derivatives of pullbacks

We now compute derivatives. As above, $X \in \mathscr{V}(I, Q, R)$.

Lemma 5.9.8. *Suppose that X is locally integrably $C^{2,1}$, $q_0 \in \mathrm{dom}(\Phi^{X^{r_0}}_{b,a})$ for some $r_0 \in R$, $Q_0 \times R_0$ is a product neighbourhood of (q_0, r_0) such that $\Phi^{X^r}_{t,a}|Q_0$ is a C^2 diffeomorphism of Q_0 onto its image for each $(t, r) \in I \times R_0$, and the map*

$$\rho \mapsto \Phi^{X^\rho}_{t,a}(q_0)$$

of R_0 into Q is C^1 for each $t \in I$. Then

$$\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \Phi^{X^\rho}_{t,a}(q_0) = \int_a^t T\Phi^{X^r}_{t,a}(q_0) \cdot \mathrm{Ad}^{X^r}_{Q_0}(Z^r_X)(\sigma, q_0)\,\mathrm{d}\sigma$$

for each $(t, r) \in I \times R_0$.

Proof. For each $r \in R$, define $A^r_X \in \mathscr{V}(I, TQ)$ by

$$A^r_X(t, v_q) = \mathrm{tlft}(X^r)(t, v_q) + \mathrm{vlft}(Z^r_X)(t, v_q).$$

The local representative of A^r_X in a natural chart $(TV, T\boldsymbol{\psi})$ on TQ is

$$(A^r_X)_{T\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{v}) = \begin{pmatrix} X^r_{\boldsymbol{\psi}}(t, \boldsymbol{q}) \\ \boldsymbol{D}_2 X^r_{\boldsymbol{\psi}}(t, \boldsymbol{q}) \cdot \boldsymbol{v} + (Z^r_X)_{\boldsymbol{\psi}}(t, \boldsymbol{q}) \end{pmatrix}. \tag{5.3}$$

Using (5.3), together with [McShane 1983, Theorem III.6-2], it follows that

$$\Phi^{A^r_X}_{t,a}(v_{q_0}) = \Phi^{\mathrm{tlft}(X^r)+\mathrm{vlft}(Z^r_X)}_{t,a}(v_{q_0})$$

is well-defined for each $(t, v_{q_0}, r) \in I \times T_{q_0}Q \times R_0$ and

$$\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \Phi^{X^\rho}_{t,a}(q_0) = \Phi^{A^r_X}_{t,a}(0_{q_0})$$

for each $(t, r) \in I \times R_0$. By construction, X^r is locally integrably C^2 and Z^r_X is locally integrably C^0. By Lemma 5.6.1, we have

$$\begin{aligned} \Phi^{A^r_X}_{t,a}(0_{q_0}) &= \Phi^{\mathrm{tlft}(X^r)+\mathrm{vlft}(Z^r_X)}_{t,a}(0_{q_0}) \\ &= T\Phi^{X^r}_{t,a}(q_0) \cdot 0_{q_0} + \int_a^t T\Phi^{X^r}_{t,a}(q_0) \cdot \mathrm{Ad}^{X^r}_{Q_0}(Z^r_X)(\sigma, q_0)\,\mathrm{d}\sigma \\ &= \int_a^t T\Phi^{X^r}_{t,a}(q_0) \cdot \mathrm{Ad}^{X^r}_{Q_0}(Z^r_X)(\sigma, q_0)\,\mathrm{d}\sigma. \end{aligned}$$

This completes the proof. ∎

Lemma 5.9.9. *Suppose that X is locally integrably $C^{3,2}$, $q_0 \in \mathrm{dom}(\Phi^{X^{r_0}}_{b,a})$ for some $r_0 \in R$, $Q_0 \times R_0$ is a product neighbourhood of (q_0, r_0) such that $\Phi^{X^r}_{t,a}|Q_0$ is a C^3 diffeomorphism of Q_0 onto its image for each $(t, r) \in I \times R_0$, and the map*

$$\rho \mapsto \Phi^{X^\rho}_{t,a}(q_0)$$

of R_0 into Q is C^2 for each $t \in I$. Furthermore, suppose that $v_{q_0} \in T_{q_0}Q$. Then

$$\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r T\Phi^{X^\rho}_{a,t}(\Phi^{X^\rho}_{t,a}(q_0)) \cdot T\Phi^{X^r}_{t,a}(q_0) \cdot v_{q_0} = -\int_a^t s_{Q_0} \circ T(\mathrm{Ad}^{X^r}_{Q_0}(Z^r_X)_\sigma)(q_0) \cdot v_{q_0}\,\mathrm{d}\sigma$$

for each $(t, r) \in I \times R_0$.

Proof. See Section A.2. ∎

In what follows, $Y \in \mathscr{V}(I, Q, R)$.

Lemma 5.9.10. *Suppose that X is locally integrably $C^{3,2}$, $q_0 \in \mathrm{dom}(\Phi^{X^{r_0}}_{b,a})$ for some $r_0 \in R$, $Q_0 \times R_0$ is a product neighbourhood of (q_0, r_0) such that $\Phi^{X^r}_{t,a}|Q_0$ is a C^3 diffeomorphism of Q_0 onto its image for each $(t, r) \in I \times R_0$, and the map*

$$\rho \mapsto \Phi^{X^\rho}_{t,a}(q_0)$$

of R_0 into Q is C^2 for each $t \in I$. Furthermore, suppose that Y is locally integrably $C^{1,1}$, $s \in I$, and $\mathfrak{r} \in R_0$. Then

$$\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}^{X^\rho}_{Q_0}(Y^{\mathfrak{r}}_s)(t, q_0) = \int_a^t \left[\mathrm{Ad}^{X^r}_{Q_0}(Z^r_X)_\sigma, \mathrm{Ad}^{X^r}_{Q_0}(Y^{\mathfrak{r}}_s)_t\right](q_0)\,\mathrm{d}\sigma$$

for each $(t, r) \in I \times R_0$.

Proof. See Section A.3. ∎

Lemma 5.9.11. *Suppose that X is locally integrably $C^{3,2}$, $q_0 \in \mathrm{dom}(\Phi^{X^{r_0}}_{b,a})$ for some $r_0 \in R$, $Q_0 \times R_0$ is a product neighbourhood of (q_0, r_0) such that $\Phi^{X^r}_{t,a}|Q_0$ is a C^3 diffeomorphism of Q_0 onto its image for each $(t, r) \in I \times R_0$, and the map*

$$\rho \mapsto \Phi^{X^\rho}_{t,a}(q_0)$$

of R_0 into Q is C^2 for each $t \in I$. Furthermore, suppose that Y is locally integrably $C^{1,1}$ and $s \in I$. Then

$$\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}^{X^\rho}_{Q_0}(Y^\rho_s)(t, q_0) = \mathrm{Ad}^{X^r}_{Q_0}((Z^r_Y)_s)(t, q_0) + \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}^{X^\rho}_{Q_0}(Y^r_s)(t, q_0)$$

or equivalently

$$\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}_{Q_0}^{X^\rho}(Y_s^\rho)(t,q_0) = \mathrm{Ad}_{Q_0}^{X^r}((Z_Y^r)_s)(t,q_0) + \int_a^t \left[\mathrm{Ad}_{Q_0}^{X^r}(Z_X^r)_\sigma, \mathrm{Ad}_{Q_0}^{X^r}(Y_s^r)_t\right](q_0)\,\mathrm{d}\sigma$$

for each $(t,r) \in I \times R_0$*. In particular, with* $s = t$*, we have*

$$\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}_{Q_0}^{X^\rho}(Y^\rho)(t,q_0) = \mathrm{Ad}_{Q_0}^{X^r}(Z_Y^r)(t,q_0) + \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}_{Q_0}^{X^\rho}(Y^r)(t,q_0)$$

or equivalently

$$\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}_{Q_0}^{X^\rho}(Y^\rho)(t,q_0) = \mathrm{Ad}_{Q_0}^{X^r}(Z_Y^r)(t,q_0) + \int_a^t \left[\mathrm{Ad}_{Q_0}^{X^r}(Z_X^r)_\sigma, \mathrm{Ad}_{Q_0}^{X^r}(Y^r)_t\right](q_0)\,\mathrm{d}\sigma$$

for each $(t,r) \in I \times R_0$*.*

Proof. See Section A.4. ∎

Remark 5.9.12. When we invoke Lemma 5.9.11 in Chapter 8, we will write

$$\begin{aligned}&\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}_{Q_0}^{X^\rho}(Y^\rho)(t,q_0)\\ &= \mathrm{Ad}_{Q_0}^{X^r}\left(\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r Y^\rho\right)(t,q_0) + \int_a^t \left[\mathrm{Ad}_{Q_0}^{X^r}\left(\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r X^\rho\right)_\sigma, \mathrm{Ad}_{Q_0}^{X^r}(Y^r)_t\right](q_0)\,\mathrm{d}\sigma.\end{aligned}$$

This formulation is less precise, but eliminates the need to write Z_X and Z_Y.

To conclude this chapter, let us briefly indicate where these identities come into play in subsequent chapters. In Chapters 6 and 7, we make heavy use of the identities dealing with pullbacks involving lifts. In Chapter 8, the most important identities are the ones dealing with derivatives of pullbacks.

Chapter 6

Differentials of endpoint maps

Consider a control system

$$\Sigma = (f, \mathscr{U})$$

evolving on an n-dimensional manifold M, and let $x_0 \in M$. In Chapter 4, we saw that the continuation method solves the x_0-anchored motion planning problem for Σ by lifting curves in M to curves in $\mathscr{U}$. The candidate curves in M must satisfy

$$\operatorname{image}(\pi) \subseteq \operatorname{image}(\underline{\mathrm{End}}^{\Sigma}_{x_0}). \tag{6.1}$$

Verifying that this constraint is satisfied is difficult, in the sense that it necessitates a complete characterization of $\mathscr{U}^{\mathrm{sing}}_{x_0}$.[1] In Chapter 1, we called this difficulty the second obstruction to the continuation method. It is clear that, in order to understand the nature of $\mathscr{U}^{\mathrm{sing}}_{x_0}$, one must begin by computing the differential

$$T\mathrm{End}^{\Sigma}_{x_0},$$

or, more generally, the differential $T\mathrm{End}^{\Sigma}$. In this chapter, we compute $T\mathrm{End}^{\Sigma}$, using a novel approach which is not based on the chronological calculus formalism [Agrachev and Sachkov 2004]. In Chapter 8, we use the resulting expression for $T\mathrm{End}^{\Sigma}$ to derive a necessary and sufficient constant-rank condition.

[1] Recall from Chapter 4 that the domain of $\underline{\mathrm{End}}^{\Sigma}_{x_0}$ is $\mathscr{U} \setminus \mathscr{U}^{\mathrm{sing}}_{x_0}$.

This chapter is organized in the following way. In Section 6.1, we begin by constructing the first variations of Σ. Roughly speaking, each first variation can be viewed as a linearized version of Σ evolving on TM. We then show how the first variations of Σ can be used to compute $T\mathrm{End}^{\Sigma}$. In Section 6.2, we present three examples. Finally, in Section 6.3, we discuss how the contents of this chapter relate to the established literature.

Our standing assumptions in this chapter are that

- M is an n-dimensional manifold,
- Σ is a nice C^1_p control system evolving on M (see Definition 3.3.7), and
- The time domain of Σ is $J = [a, b]$, so that $\mathscr{U} = L^p(J, \mathbb{R}^r)$.

By Corollary 3.3.12, the map

$$\mathrm{End}^{\Sigma} : \mathrm{dom}(\mathrm{End}^{\Sigma}) \subseteq M \times \mathscr{U} \to M$$

is C^1.

6.1 First variations

To explicitly compute $T\mathrm{End}^{\Sigma}$, it will be necessary to evaluate the global flows of time-varying vector fields on open submanifolds of TM. To make this evaluation, we will invoke Lemma 5.6.1. To ensure that its hypotheses are satisfied, Σ must satisfy the additional criterion contained in the next definition.

Definition 6.1.1. We say that Σ has ***first variations*** if, in addition to being a nice C^1_p control system, Σ is also a C^2 control system.

In the remainder of this section, we assume that Σ has first variations and

$$(x_0, \boldsymbol{u}_0) \in \mathrm{dom}(\mathrm{End}^{\Sigma}).$$

We say that $\boldsymbol{u}_0$ is the ***zeroth-order reference control***. There exists a neighbourhood M_0 of x_0 such that $\Phi_{t,a}^{f_{\boldsymbol{u}_0}}$ is a C^2 diffeomorphism of M_0 onto its image for each $t \in J$. This follows from the assumption that Σ is a C^2 control system, together with an argument identical to the one used in Section 5.4 to obtain the neighbourhood Q_0.

6.1.1 Basic definitions and properties

In this section, we construct the first variation of Σ along the zeroth-order reference control $\boldsymbol{u}_0$. We do so using tangent and vertical lifts. To the best of our knowledge, this idea originates in the work of Crouch and van der Schaft [1987]. For more recent work involving the same type of construction, we refer to [Bullo and Lewis 2007]. The major difference between the analysis in this chapter and the cited work is that the latter only concerns C^∞ control-affine systems.

We require the following additional notation: Define the controllable time-varying vector field $\boldsymbol{D}_3 f_{\boldsymbol{u}_0} \in \mathscr{V}(J, M, \mathbb{R}^r)$ by

$$\boldsymbol{D}_3 f_{\boldsymbol{u}_0}(t, x, \boldsymbol{\omega}) = \boldsymbol{D}_3 f(t, x, \boldsymbol{u}_0(t)) \cdot \boldsymbol{\omega}.$$

On the right-hand side, $\boldsymbol{u}_0$ denotes any representative of $\boldsymbol{u}_0$. Using Lemma 3.1.3, it is not hard to see that the considerations in this chapter involving $\boldsymbol{D}_3 f_{\boldsymbol{u}_0}$ are independent of the particular choice of representative. Note that $\boldsymbol{D}_3 f_{\boldsymbol{u}_0}$ is well-defined, since

$$\boldsymbol{D}_3 f(t, x, \boldsymbol{u}_0(t)) \in \mathrm{Hom}(\mathbb{R}^r, T_x M)$$

by definition.

Definition 6.1.2. The ***first variation of*** Σ ***along*** $\boldsymbol{u}_0$ is the pair

$$T\Sigma_{\boldsymbol{u}_0} = (T f_{\boldsymbol{u}_0}, \mathscr{U}),$$

where the controllable time-varying vector field $Tf_{\boldsymbol{u}_0} \in \mathscr{V}(J, TM, \mathbb{R}^r)$ is defined by

$$Tf_{\boldsymbol{u}_0}(t, v_x, \boldsymbol{\omega}) = \mathrm{tlft}(f^{\boldsymbol{u}_0})(t, v_x) + \mathrm{vlft}(\boldsymbol{D}_3 f_{\boldsymbol{u}_0})(t, v_x, \boldsymbol{\omega}).$$

Here, tlft and vlft are the tangent and vertical lift operations from Chapter 5.

In a natural chart $(TV, T\boldsymbol{\psi})$ on TM, we have

$$\mathrm{tlft}(f^{\boldsymbol{u}_0})_{T\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{v}) = \begin{pmatrix} f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \\ \boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{v} \end{pmatrix}$$

and

$$\mathrm{vlft}(\boldsymbol{D}_3 f_{\boldsymbol{u}_0})_{T\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{v}, \boldsymbol{\omega}) = \begin{pmatrix} \boldsymbol{0}_{\mathbb{R}^n} \\ \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{\omega} \end{pmatrix}.$$

By linearity,

$$(Tf^{\boldsymbol{u}}_{\boldsymbol{u}_0})_{T\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{v}) = \begin{pmatrix} f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \\ \boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{v} + \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{u}(t) \end{pmatrix}. \tag{6.2}$$

We now show that $T\Sigma_{\boldsymbol{u}_0}$ is a control system in the sense of Definition 3.1.4.

Lemma 6.1.3. *The pair $T\Sigma_{\boldsymbol{u}_0}$ is a control system.*

Proof. We must show that each $\boldsymbol{u} \in \mathscr{U}$ is $Tf_{\boldsymbol{u}_0}$-admissible. To this end, suppose that $\boldsymbol{u} \in \mathscr{U}$, $\mathscr{A}_M$ is a compatible atlas on M, $(TV, T\boldsymbol{\psi}) \in T\mathscr{A}_M$, and

$$(t_0, \boldsymbol{x}_0, \boldsymbol{v}_0) \in J \times \boldsymbol{\psi}(V) \times \mathbb{R}^n.$$

Provided that it exists, it is clear that the maximally-defined solution

$$\boldsymbol{\zeta} = (\boldsymbol{\xi}, \boldsymbol{\nu}) : \mathrm{dom}(\boldsymbol{\zeta}) \to \boldsymbol{\psi}(V) \times \mathbb{R}^n$$

of $((Tf^{\boldsymbol{u}}_{\boldsymbol{u}_0})_{T\boldsymbol{\psi}}, t_0, \boldsymbol{x}_0, \boldsymbol{v}_0)$ must satisfy the following three properties:

- $\mathrm{dom}(\boldsymbol{\zeta}) \subseteq J^{\Sigma_{\boldsymbol{\psi}}}(t_0, \boldsymbol{x}_0, \boldsymbol{u}_0)$,
- $\boldsymbol{\xi}(t) = \Phi^{\Sigma_{\boldsymbol{\psi}}}_{t,t_0}(\boldsymbol{x}_0, \boldsymbol{u}_0)$ for each $t \in \mathrm{dom}(\boldsymbol{\zeta})$, and

- $\boldsymbol{\nu}$ is the maximally-defined solution of $(\boldsymbol{g}, t_0, \boldsymbol{v}_0)$, where

$$\boldsymbol{g} : J^{\Sigma_{\boldsymbol{\psi}}}(t_0, \boldsymbol{x}_0, \boldsymbol{u}_0) \times \mathbb{R}^n \to \mathbb{R}^n$$

is defined by

$$\boldsymbol{g}(t, \boldsymbol{v}) = \boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot \boldsymbol{v} + \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot \boldsymbol{u}(t). \tag{6.3}$$

If we can show that there exists a maximally-defined solution of $(\boldsymbol{g}, t_0, \boldsymbol{v}_0)$, then the proof will be complete. By Theorem 2.2.14, it is enough to show that $\boldsymbol{g}$ satisfies Carathéodory conditions. Recalling Remark 3.2.2, it is clear that the first two criteria of Definition 2.2.13 are satisfied. We now show that $\boldsymbol{g}$ is locally integrably bounded and locally integrably Lipschitz. Suppose that J_0 is a compact subinterval of $J^{\Sigma_{\boldsymbol{\psi}}}(t_0, \boldsymbol{x}_0, \boldsymbol{u}_0)$ and K is a compact subset of $\mathbb{R}^n$. Invoking Lemma 3.2.11, there exists $\alpha \in L^1(J_0, \mathbb{R}_{\geq 0})$ such that

$$\|\boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t))\| \leq \alpha(t) \quad \text{and} \quad \|\boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot \boldsymbol{u}(t)\|_{\mathbb{R}^n} \leq \alpha(t)$$

for a.a. $t \in J_0$. Setting $C = \sup_{\boldsymbol{v} \in K} \|\boldsymbol{v}\|_{\mathbb{R}^n}$, it follows that

$$\|\boldsymbol{g}(t, \boldsymbol{v})\|_{\mathbb{R}^n} \leq (C + 1)\alpha(t)$$

for a.a. $t \in J_0$ and each $\boldsymbol{v} \in K$. Furthermore, by linearity, we have

$$\begin{aligned}
\|\boldsymbol{g}(t, \boldsymbol{v}) - \boldsymbol{g}(t, \boldsymbol{w})\|_{\mathbb{R}^n} &= \|\boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot (\boldsymbol{v} - \boldsymbol{w})\|_{\mathbb{R}^n} \\
&\leq \|\boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t))\| \, \|\boldsymbol{v} - \boldsymbol{w}\|_{\mathbb{R}^n} \\
&\leq \alpha(t) \, \|\boldsymbol{v} - \boldsymbol{w}\|_{\mathbb{R}^n}
\end{aligned}$$

for a.a. $t \in J_0$ and each $\boldsymbol{v}, \boldsymbol{w} \in K$. In fact, it is clear that this upper bound holds regardless of whether or not $\boldsymbol{v}, \boldsymbol{w}$ are constrained to lie in K. That is, the upper bound holds for a.a. $t \in J_0$ and each $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$. Since J_0 and K were chosen arbitrarily, we conclude that $\boldsymbol{g}$ satisfies Carathéodory conditions. ∎

The next lemma states that, locally, the controlled trajectories of $T\Sigma_{\boldsymbol{u}_0}$ are defined on the same interval as the corresponding $\boldsymbol{u}_0$-controlled trajectories of Σ.

Lemma 6.1.4. *Suppose that $\boldsymbol{u} \in \mathscr{U}$, $(TV, T\boldsymbol{\psi})$ is a natural chart on TM, and*

$$(t_0, \boldsymbol{x}_0, \boldsymbol{v}_0) \in J \times \boldsymbol{\psi}(V) \times \mathbb{R}^n.$$

Then the maximally-defined solution

$$\boldsymbol{\zeta} = (\boldsymbol{\xi}, \boldsymbol{\nu}) : \operatorname{dom}(\boldsymbol{\zeta}) \to \boldsymbol{\psi}(V) \times \mathbb{R}^n$$

of $((Tf^{\boldsymbol{u}}_{\boldsymbol{u}_0})_{T\boldsymbol{\psi}}, t_0, \boldsymbol{x}_0, \boldsymbol{v}_0)$ satisfies

$$\operatorname{dom}(\boldsymbol{\zeta}) = J^{\Sigma_{\boldsymbol{\psi}}}(t_0, \boldsymbol{x}_0, \boldsymbol{u}_0).$$

Proof. In the proof of Lemma 6.1.3, we showed that $\boldsymbol{\nu}$ is the maximally-defined solution of $(\boldsymbol{g}, t_0, \boldsymbol{v}_0)$, where $\boldsymbol{g}$ was defined by (6.3). We also showed that for each compact subinterval J_0 of $J^{\Sigma_{\boldsymbol{\psi}}}(t_0, \boldsymbol{x}_0, \boldsymbol{u}_0)$, there exists $\alpha \in L^1(J_0, \mathbb{R}_{\geq 0})$ such that

$$\|\boldsymbol{g}(t, \boldsymbol{v}) - \boldsymbol{g}(t, \boldsymbol{w})\|_{\mathbb{R}^n} \leq \alpha(t)$$

for a.a. $t \in J_0$ and each $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$. By Lemma 2.2.17,

$$\operatorname{dom}(\boldsymbol{\zeta}) = J^{\Sigma_{\boldsymbol{\psi}}}(t_0, \boldsymbol{x}_0, \boldsymbol{u}_0).$$

This completes the proof. ∎

Recalling that $(x_0, \boldsymbol{u}_0) \in \operatorname{dom}(\operatorname{End}^{\Sigma})$, the next result follows immediately.

Proposition 6.1.5. *We have*

$$\pi_{TM}^{-1}(x_0) \times \mathscr{U} = T_{x_0}M \times \mathscr{U} \subseteq \operatorname{dom}(\operatorname{End}^{T\Sigma_{\boldsymbol{u}_0}}).$$

Furthermore,

$$\pi_{TM} \circ \operatorname{End}^{T\Sigma_{\boldsymbol{u}_0}}(v_{x_0}, \boldsymbol{u}) = \operatorname{End}^{\Sigma}(x_0, \boldsymbol{u}_0)$$

for each $(v_{x_0}, \boldsymbol{u}) \in T_{x_0}M \times \mathscr{U}$.

6.1.2 First variations compute differentials

It was indicated at the beginning of this chapter that the first variations of Σ can be used to compute $T\mathrm{End}^{\Sigma}$. In this section, we make this statement precise.

Theorem 6.1.6. *We have*

$$\mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(v_{x_0},\boldsymbol{u}) = T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0)\cdot v_{x_0} + \int_a^b T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0)\cdot \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})(\sigma,x_0)\,\mathrm{d}\sigma$$

for each $(v_{x_0},\boldsymbol{u})\in T_{x_0}M\times\mathscr{U}$.

Proof. Choose $(v_{x_0},\boldsymbol{u})\in T_{x_0}M\times\mathscr{U}$. By construction,

$$\begin{aligned}\mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(v_{x_0},\boldsymbol{u}) &= \Phi^{T\Sigma_{\boldsymbol{u}_0}}_{b,a}(v_{x_0},\boldsymbol{u})\\ &= \Phi^{Tf^{\boldsymbol{u}}_{\boldsymbol{u}_0}}_{b,a}(v_{x_0})\\ &= \Phi^{\mathrm{tlft}(f^{\boldsymbol{u}_0})+\mathrm{vlft}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})}_{b,a}(v_{x_0})\end{aligned}$$

Since Σ has first variations, $f^{\boldsymbol{u}_0}$ is locally integrably C^2. Furthermore, $\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0}$ is locally integrably C^0 by Lemma 3.3.9. Invoking Lemma 5.6.1, we conclude that

$$\mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(v_{x_0},\boldsymbol{u}) = T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0)\cdot v_{x_0} + \int_a^b T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0)\cdot \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})(\sigma,x_0)\,\mathrm{d}\sigma.$$

This completes the proof. ∎

The control system $T\Sigma_{\boldsymbol{u}_0}$ has the following linearity property.

Corollary 6.1.7. *We have*

$$\mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(v_{x_0}+\tilde{v}_{x_0},\boldsymbol{u}+\tilde{\boldsymbol{u}}) = \mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(v_{x_0},\boldsymbol{u}) + \mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(\tilde{v}_{x_0},\tilde{\boldsymbol{u}})$$

for each $(v_{x_0},\boldsymbol{u}),(\tilde{v}_{x_0},\tilde{\boldsymbol{u}})\in T_{x_0}M\times\mathscr{U}$.

Proof. This follows immediately from Theorem 6.1.6 and the fact that

$$\begin{aligned}\mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}+\tilde{\boldsymbol{u}}}_{\boldsymbol{u}_0}) &= \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0} + \boldsymbol{D}_3 f^{\tilde{\boldsymbol{u}}}_{\boldsymbol{u}_0})\\ &= \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0}) + \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\tilde{\boldsymbol{u}}}_{\boldsymbol{u}_0})\end{aligned}$$

for each $\boldsymbol{u},\tilde{\boldsymbol{u}}\in\mathscr{U}$. ∎

Remark 6.1.8. In [Amiss and Guay 2011b], Theorem 6.1.6 is derived under unnecessarily strong conditions on Σ. Namely, it is assumed that Σ is a nice C_p^3 control system. This hypothesis can be relaxed, as in this chapter.

The next lemma relates $T\mathrm{End}^\Sigma(x_0, \boldsymbol{u}_0)$ to the controlled trajectories of $T\Sigma_{\boldsymbol{u}_0}$.

Lemma 6.1.9. *We have*

$$T\mathrm{End}^\Sigma(x_0, \boldsymbol{u}_0) \cdot (v_{x_0}, \boldsymbol{u}) = \mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(v_{x_0}, \boldsymbol{u}) \tag{6.4}$$

for each $(v_{x_0}, \boldsymbol{u}) \in T_{(x_0, \boldsymbol{u}_0)}(M \times \mathscr{U}) \cong T_{x_0}M \oplus \mathscr{U}$.

Proof. Choose $(v_{x_0}, \boldsymbol{u}) \in T_{x_0}M \oplus \mathscr{U}$ and charts $(V_0, \boldsymbol{\psi}_0)$, $(V_1, \boldsymbol{\psi}_1)$ on M such that $x_0 \in V_0$ and $x_1 = \mathrm{End}^\Sigma_{x_0}(\boldsymbol{u}_0) \in V_1$. By (6.2) and [Margheri 1996, Equation 3.12],

$$\begin{aligned}
&T\boldsymbol{\psi}_1(x_1) \circ \mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(v_{x_0}, \boldsymbol{u}) \\
&= \begin{pmatrix} \boldsymbol{\psi}_1(x_1) \\ \boldsymbol{D}(\boldsymbol{\psi}_1 \circ \mathrm{End}^\Sigma \circ (\boldsymbol{\psi}_0^{-1} \times \mathrm{id}_{\mathscr{U}}))(\boldsymbol{\psi}_0(x_0), \boldsymbol{u}_0) \cdot (T\boldsymbol{\psi}_0(x_0) \cdot v_{x_0}, \boldsymbol{u}) \end{pmatrix}.
\end{aligned}$$

This completes the proof. ∎

The next theorem, which is the main result in this chapter, combines the results of Theorem 6.1.6 and Lemma 6.1.9 to compute the linear map $T\mathrm{End}^\Sigma(x_0, \boldsymbol{u}_0)$.

Theorem 6.1.10. *We have*

$$\begin{aligned}
&T\mathrm{End}^\Sigma(x_0, \boldsymbol{u}_0) \cdot (v_{x_0}, \boldsymbol{u}) \\
&= T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot v_{x_0} + \int_a^b T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})(\sigma, x_0)\,\mathrm{d}\sigma
\end{aligned}$$

for each $(v_{x_0}, \boldsymbol{u}) \in T_{(x_0, \boldsymbol{u}_0)}(M \times \mathscr{U}) \cong T_{x_0}M \oplus \mathscr{U}$.

In particular, the preceding theorem tells us that

$$\begin{aligned}
T_1\mathrm{End}^\Sigma(x_0, \boldsymbol{u}_0) \cdot v_{x_0} &= T\mathrm{End}^\Sigma(x_0, \boldsymbol{u}_0) \cdot (v_{x_0}, \boldsymbol{0}_{\mathscr{U}}) \\
&= T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot v_{x_0}
\end{aligned}$$

and

$$\begin{aligned} T_2\mathrm{End}^{\Sigma}(x_0, \boldsymbol{u}_0) \cdot \boldsymbol{u} &= T\mathrm{End}^{\Sigma}(x_0, \boldsymbol{u}_0) \cdot (0_{x_0}, \boldsymbol{u}) \\ &= T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0) \cdot \boldsymbol{u} \\ &= \int_a^b T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})(\sigma, x_0)\,\mathrm{d}\sigma \end{aligned}$$

for each $(v_{x_0}, \boldsymbol{u}) \in T_{x_0}M \oplus \mathscr{U}$. Here, $T_1\mathrm{End}^{\Sigma}$ and $T_2\mathrm{End}^{\Sigma}$ denote the partial differentials of End^{Σ}; see [Lee 2009, Section 2.4] for definitions.

The first variations of Σ can be regarded as linearized versions of Σ. Thus Theorem 6.1.10 can be regarded as a coordinate-free *linearizations compute differentials* principle, in the sense of [Sontag 1998, Section 2.8].

6.2 Examples

In this section, we illustrate Theorem 6.1.10 by way of three examples.

Example 6.2.1. Suppose that Σ is a linear system with $M = \mathbb{R}^n$ and

$$\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega}) = \boldsymbol{A} \cdot \boldsymbol{x} + \boldsymbol{B} \cdot \boldsymbol{\omega}$$

for matrices $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n\times r}$. In this case, the well-known variation of constants formula yields

$$\Phi^{\Sigma}_{t,a}(\boldsymbol{x}, \boldsymbol{u}) = \exp((t-a)\boldsymbol{A}) \cdot \boldsymbol{x} + \int_a^t \exp((t-\sigma)\boldsymbol{A}) \circ \boldsymbol{B} \cdot \boldsymbol{u}(\sigma)\,\mathrm{d}\sigma,$$

where exp denotes the matrix exponential [Sontag 1998]. In particular,

$$\mathrm{End}^{\Sigma}(\boldsymbol{x}, \boldsymbol{u}) = \exp((b-a)\boldsymbol{A}) \cdot \boldsymbol{x} + \int_a^b \exp((b-\sigma)\boldsymbol{A}) \circ \boldsymbol{B} \cdot \boldsymbol{u}(\sigma)\,\mathrm{d}\sigma$$

and one computes directly that

$$T\mathrm{End}^{\Sigma}(\boldsymbol{x}_0, \boldsymbol{u}_0) \cdot (\boldsymbol{v}, \boldsymbol{u}) = \exp((b-a)\boldsymbol{A}) \cdot \boldsymbol{v} + \int_a^b \exp((b-\sigma)\boldsymbol{A}) \circ \boldsymbol{B} \cdot \boldsymbol{u}(\sigma)\,\mathrm{d}\sigma.$$

To see that this coincides with the result of Theorem 6.1.10, observe that

$$T\Phi^{\boldsymbol{f}^{\boldsymbol{u}_0}}_{t,a}(\boldsymbol{x}) \cdot \boldsymbol{v} = \exp((t-a)\boldsymbol{A}) \cdot \boldsymbol{v}$$

and, furthermore, that $\boldsymbol{D}_3 \boldsymbol{f}^{\boldsymbol{u}}_{\boldsymbol{u}_0}(t, \boldsymbol{x}) = \boldsymbol{B} \cdot \boldsymbol{u}(t)$ for each $\boldsymbol{u} \in L^p(J, \mathbb{R}^r)$. Thus

$$\mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})(t, \boldsymbol{x}_0) = \exp((a-t)\boldsymbol{A}) \circ \boldsymbol{B} \cdot \boldsymbol{u}(t).$$

By Theorem 6.1.10,

$$\begin{aligned}
&T\mathrm{End}^{\Sigma}(\boldsymbol{x}_0, \boldsymbol{u}_0) \cdot (\boldsymbol{v}, \boldsymbol{u}) \\
&= T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(\boldsymbol{x}_0) \cdot \boldsymbol{v} + \int_a^b T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(\boldsymbol{x}_0) \cdot \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})(\sigma, \boldsymbol{x}_0)\, \mathrm{d}\sigma \\
&= \exp((b-a)\boldsymbol{A}) \cdot \boldsymbol{v} + \int_a^b \exp((b-a)\boldsymbol{A}) \circ \exp((a-\sigma)\boldsymbol{A}) \circ \boldsymbol{B} \cdot \boldsymbol{u}(\sigma)\, \mathrm{d}\sigma \\
&= \exp((b-a)\boldsymbol{A}) \cdot \boldsymbol{v} + \int_a^b \exp((b-\sigma)\boldsymbol{A}) \circ \boldsymbol{B} \cdot \boldsymbol{u}(\sigma)\, \mathrm{d}\sigma,
\end{aligned}$$

as expected.

In the next two examples, Σ is a C^2 control-affine system with $\mathscr{U} = L^2(J, \mathbb{R}^r)$. Such a control system has first variations; see Examples 3.1.15 and 3.2.4.

Example 6.2.2. Suppose that Σ is driftless and

$$f(t, x, \boldsymbol{\omega}) = \sum_{i=1}^{r} \omega^i f_i(x)$$

for C^2 vector fields $f_1, \ldots, f_r$. We have

$$\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0}(t, x) = \sum_{i=1}^{r} u^i(t) f_i(x).$$

By Theorem 6.1.10,

$$\begin{aligned}
&T\mathrm{End}^{\Sigma}(x_0, \boldsymbol{u}_0) \cdot (v_{x_0}, \boldsymbol{u}) \\
&= T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot v_{x_0} + \int_a^b T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})(\sigma, x_0)\, \mathrm{d}\sigma \\
&= T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot v_{x_0} + \int_a^b \sum_{i=1}^{r} u^i(\sigma) T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot T\Phi^{f^{\boldsymbol{u}_0}}_{a,\sigma}(\Phi^{f^{\boldsymbol{u}_0}}_{\sigma,a}(x_0)) \cdot f_i(\Phi^{f^{\boldsymbol{u}_0}}_{\sigma,a}(x_0))\, \mathrm{d}\sigma
\end{aligned}$$

$$= T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot v_{x_0} + \int_a^b \sum_{i=1}^r u^i(\sigma) T\Phi^{f^{\boldsymbol{u}_0}}_{b,\sigma}(\Phi^{f^{\boldsymbol{u}_0}}_{\sigma,a}(x_0)) \cdot f_i(\Phi^{f^{\boldsymbol{u}_0}}_{\sigma,a}(x_0))\, \mathrm{d}\sigma.$$

In particular,

$$\begin{aligned} T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0) \cdot \boldsymbol{u} &= T_2\mathrm{End}^{\Sigma}(x_0, \boldsymbol{u}_0) \cdot \boldsymbol{u} \\ &= \int_a^b \sum_{i=1}^r u^i(\sigma) T\Phi^{f^{\boldsymbol{u}_0}}_{b,\sigma}(\Phi^{f^{\boldsymbol{u}_0}}_{\sigma,a}(x_0)) \cdot f_i(\Phi^{f^{\boldsymbol{u}_0}}_{\sigma,a}(x_0))\, \mathrm{d}\sigma. \end{aligned}$$

This result matches the expressions for $T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)$ presented, without proof, in [Sussmann 1993], [Chitour and Sussmann 1998], and [Chitour 2006].

Example 6.2.2 is easily extended to C^2 control-affine systems with drift.

Example 6.2.3. Suppose that

$$f(t, x, \boldsymbol{\omega}) = f_0(x) + \sum_{i=1}^r \omega^i f_i(x)$$

for C^2 vector fields $f_0, f_1, \ldots, f_r$. As in Example 6.2.2, we have

$$\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0}(t, x) = \sum_{i=1}^r u^i(t) f_i(x).$$

By Theorem 6.1.10,

$$\begin{aligned} &T\mathrm{End}^{\Sigma}(x_0, \boldsymbol{u}_0) \cdot (v_{x_0}, \boldsymbol{u}) \\ &= T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot v_{x_0} + \int_a^b \sum_{i=1}^r u^i(\sigma) T\Phi^{f^{\boldsymbol{u}_0}}_{b,\sigma}(\Phi^{f^{\boldsymbol{u}_0}}_{\sigma,a}(x_0)) \cdot f_i(\Phi^{f^{\boldsymbol{u}_0}}_{\sigma,a}(x_0))\, \mathrm{d}\sigma. \end{aligned}$$

Although they are formally similar, the difference between this example and Example 6.2.2 is that here, $\Phi^{f^{\boldsymbol{u}_0}}$ depends on the drift vector field f_0.

6.3 Discussion

6.3.1 L^∞ controls

The scope of this chapter was limited to nice C^1_p control systems, which, by definition, use L^p controls for $p \neq \infty$. To adapt the results in this chapter to accommodate control systems that use L^∞ controls, it is enough to replace the assumption that Σ is a nice C^1_p control system by the assumption that Σ is a quasi-C^1 control system in the sense of Grasse [1979]. Indeed, the assumption that Σ is quasi-C^1 ensures that End is C^1. For a proof of this fact, we refer to [Grasse 1979, Chapter 3].

6.3.2 Related results

Results similar to Theorem 6.1.10 can be found in the literature. For example, we refer to [Agrachev and Sachkov 2004, Section 20.3]. For this reason, we are obliged to point out how the analysis in this chapter differs from the cited work. The major difference is that the approach of Agrachev and Sachkov [2004] employs the chronological calculus formalism, which requires that Σ meets more stringent smoothness and completeness conditions—in particular, Σ must be a complete, time-invariant, C^∞ control system. We do not use this formalism in our approach. As a result, we require only that Σ meets minimal regularity conditions. The notion of minimality is encapsulated by the requirement that Σ has first variations.

To conclude this section, let us explain how Theorem 6.1.10 makes contact with the field of sub-Riemannian geometry. With respect to sub-Riemannian geometry, suppose that $\mathscr{F} \subseteq TM$ is a C^∞ distribution of rank r. We denote by

$$\mathrm{Hor}^2_{x_0}(\mathscr{F})$$

the set of all LAC curves $\gamma : J \to M$ which are horizontal to $\mathscr{F}$, are L^2 with respect to any sub-Riemannian metric on M, and satisfy $\gamma(0) = x_0$; for definitions, we refer to [Montgomery 2002]. Bismut's theorem, which is one of the basic results of sub-Riemannian geometry, states that $\mathrm{Hor}^2_{x_0}(\mathscr{F})$ is a Hilbert manifold and the endpoint map $E : \mathrm{Hor}^2_{x_0}(\mathscr{F}) \to M$, defined by $E(\gamma) = \gamma(b)$, is C^∞. Moreover, its differential TE can be explicitly computed, as in [Montgomery 2002, Appendix D].

With respect to control theory, suppose that Σ is a complete, driftless, C^∞

control-affine system with $\mathscr{U} = L^2(J, \mathbb{R}^r)$. Further, let us suppose that

$$f(t, x, \boldsymbol{\omega}) = \sum_{i=1}^{r} \omega^i f_i(x)$$

for C^∞ vector fields $f_1, \ldots, f_r$ and that $f_1(x), \ldots, f_r(x)$ are linearly independent for each $x \in M$. If $\mathscr{F}$ is the C^∞ distribution of rank r spanned by the vector fields $f_1, \ldots, f_r$, then the linear independence condition allows $\mathscr{U}$ to be identified with $\mathrm{Hor}^2_{x_0}(\mathscr{F})$. Under this identification, E coincides with the map $\mathrm{End}^\Sigma_{x_0}$. As described in Chapter 1, this fact is used in the literature to ensure that $\mathrm{End}^\Sigma_{x_0}$ is C^2, or, in other words, that the first obstruction to the continuation method is overcome. Furthermore, it is not hard to see that TE, as computed in [Montgomery 2002, Appendix D], coincides with the result of Example 6.2.2. However, we emphasize that our approach does not rely on the assumptions that Σ is complete and C^∞, nor does it rely on any type of linear independence condition.

Chapter 7

Intrinsic quadratic differentials of anchored endpoint maps

Consider a control system

$$\Sigma = (f, \mathscr{U})$$

evolving on an n-dimensional manifold M, and let $x_0 \in M$. In Chapter 6, we computed the differential $T\mathrm{End}^{\Sigma}$. In this chapter, we carry the analysis of Chapter 6 to the second order. In particular, we compute the intrinsic quadratic differentials

$$\mathcal{Q}\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)$$

of $\mathrm{End}^{\Sigma}_{x_0}$, using a novel approach which is not based on the chronological calculus formalism [Agrachev and Sachkov 2004]. In Chapter 8, we use the resulting expressions for $\mathcal{Q}\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)$ to derive a necessary and sufficient constant-rank condition. This leads to conditions which ensure that $\mathscr{U}^{\mathrm{sing}}_{x_0}$ is empty, or, in other words, that the second obstruction to the continuation method is overcome.

This chapter is organized in the following way. We begin in Section 7.1 by establishing the basic theory of intrinsic quadratic differentials. In Section 7.2, we construct the second variations of Σ, mirroring the developments in Chapter 6.

Roughly speaking, each second variation can be viewed as a bilinearized version of Σ evolving on TTM. We then show how the second variations of Σ can be used to compute the intrinsic quadratic differentials of $\mathrm{End}_{x_0}^{\Sigma}$. In Section 7.3, we present two examples. Finally, in Section 7.4, we briefly discuss how the contents of this chapter relate to the established literature.

Our standing assumptions in this chapter are that

- M is a second-countable n-dimensional manifold,
- Σ is a nice C_p^2 control system evolving on M (see Definition 3.3.7), and
- The time domain of Σ is $J = [a, b]$, so that $\mathscr{U} = L^p(J, \mathbb{R}^r)$.

By Corollary 3.3.12, the map

$$\mathrm{End}_{x_0}^{\Sigma} : \mathrm{dom}(\mathrm{End}_{x_0}^{\Sigma}) \subseteq \mathscr{U} \to M$$

is C^2.

7.1 Intrinsic quadratic differentials

In this section, we establish the basic theory of intrinsic quadratic differentials[1] (IQDs), independently of any considerations involving Σ. Throughout this section,

$$F : Q \to R$$

is C^2, where Q is an open submanifold of a Banach space E and R is a second-countable ℓ-dimensional manifold. In rough terms, the IQDs of F represent the coordinate-invariant part of the second total derivatives of F, taken in charts.

[1] As a historical remark, it seems that IQDs first appeared in the work of Arnol'd [1968]. IQDs are also developed by Porteous [1971], Agrachev and Sachkov [2004], Fehér and Kőműves [2006], and Reis and Weiss [2010]. In particular, the material in this section is inspired by the treatments of Arnol'd [1968] and Agrachev and Sachkov [2004]. An important caveat is that there is no standardized terminology concerning IQDs. Throughout the literature, IQDs are also referred to as Porteous quadratic differentials, Porteous quadratic derivatives, generalized Hessians, and Hessians. The latter term is used, in particular, in the control theory literature; see, for example, [Agrachev and Gamkrelidze 1985, Agrachev 1990, Agrachev and Gamkrelidze 1991, Vakhrameev 1991b, 1996, Agrachev and Sachkov 2004].

7.1.1 Basic definitions and properties

We begin by recalling some basic properties of continuous dual spaces.

Lemma 7.1.1. *Suppose that $e \in E$ and $\lambda \cdot e = 0$ for each $\lambda \in E^*$. Then $e = 0_E$.*

Proof. By the Hahn–Banach theorem [Abraham et al. 1988] there exists $\lambda_0 \in E^*$ such that $\lambda_0 \cdot e = \|e\|_E$. In particular, $\lambda_0 \cdot e = \|e\|_E = 0$ so $e = 0_E$. ∎

Corollary 7.1.2. *Suppose that $e, \tilde{e} \in E$ and*

$$\lambda \cdot e = \lambda \cdot \tilde{e}$$

for each $\lambda \in E^*$. Then $e = \tilde{e}$.

Proof. Clearly, $\lambda \cdot e = \lambda \cdot \tilde{e}$ for each $\lambda \in E^*$ if and only if $\lambda \cdot (e - \tilde{e}) = 0$ for each $\lambda \in E^*$. Invoking Lemma 7.1.1, we conclude that $e - \tilde{e} = 0_E$. ∎

For each $q \in Q$, the ***annihilator*** of $\operatorname{image}(TF(q))$ is

$$\operatorname{image}(TF(q))^0 = \{\lambda \in T^*_{F(q)}R \ : \ \lambda \cdot v_{F(q)} = 0 \text{ for each } v_{F(q)} \in \operatorname{image}(TF(q))\}.$$

In the remainder of this section, we use the canonical vector space isomorphisms

$$\operatorname{image}(TF(q))^0 \cong \ker(TF(q)^*) \cong \operatorname{coker}(TF(q))^*$$

from [Bachman and Narici 2000, Section 17.4] and [Conway 1990, Theorem 10.2].

We now define the IQD of F at a point $q \in Q$, which will be denoted by $\mathcal{Q}F(q)$. Following Agrachev and Sachkov [2004], our definition does not directly specify the value of $\mathcal{Q}F(q) \in \operatorname{coker}(TF(q))$. Rather, we specify how $\mathcal{Q}F(q)$ is acted on by elements of the continuous dual space $\operatorname{coker}(TF(q))^*$. By Corollary 7.1.2, this specification will uniquely determine $QF(q)$ as an element of $\operatorname{coker}(TF(q))$.

Definition 7.1.3. Suppose that $q \in Q$. The ***intrinsic quadratic differential*** (IQD) of F at q is the map

$$\mathcal{Q}F(q) : \ker(TF(q)) \times \ker(TF(q)) \to \operatorname{coker}(TF(q))$$

defined as follows. For tangent vectors

$$(v_q, \tilde{v}_q) \in \ker(TF(q)) \times \ker(TF(q)),$$

choose C^∞ vector fields X and $\tilde{X}$ on Q such that $X(q) = v_q$ and $\tilde{X}(q) = \tilde{v}_q$. For each $\lambda \in \operatorname{coker}(TF(q))^*$, choose a C^∞ function $g_\lambda : R \to \mathbb{R}$ such that

$$Tg_\lambda(F(q)) = \lambda.$$

Then $\mathscr{Q}F(q) \cdot (v_q, \tilde{v}_q) + \operatorname{image}(TF(q))$ is specified by the requirement that

$$\lambda \cdot (\mathscr{Q}F(q) \cdot (v_q, \tilde{v}_q) + \operatorname{image}(TF(q))) = \mathscr{L}_X \mathscr{L}_{\tilde{X}}(g_\lambda \circ F)(q)$$

for each $\lambda \in \operatorname{coker}(TF(q))^*$, where $\mathscr{L}$ denotes the Lie derivative.

This definition relies on several choices. The next lemma tells us that $\mathscr{Q}F(q)$ is well-defined, in the sense that it is independent of these choices.

Lemma 7.1.4. *Suppose that $q \in Q$, $(v_q, \tilde{v}_q) \in \ker(TF(q)) \times \ker(TF(q))$, and the value of $\mathscr{Q}F(q) \cdot (v_q, \tilde{v}_q)$ is prescribed as in Definition 7.1.3, relative to particular choices of X, $\tilde{X}$, and g_λ. Then $\mathscr{Q}F(q) \cdot (v_q, \tilde{v}_q)$ is well-defined.*

Proof. We must show that for each $\lambda \in \operatorname{coker}(TF(q))^*$, the value of

$$\mathscr{L}_X \mathscr{L}_{\tilde{X}}(g_\lambda \circ F)(q)$$

depends only on $X(q) = v_q$, $\tilde{X}(q) = \tilde{v}_q$, and $Tg_\lambda(F(q)) = \lambda$.

Suppose that (U, φ) and $(V, \boldsymbol{\psi})$ are F-compatible charts on Q and R, respectively, such that $q \in U$. For each $\lambda \in \operatorname{coker}(TF(q))^*$, it holds that

$$\mathscr{L}_{\tilde{X}}(g_\lambda \circ F)(\varphi^{-1}(x)) = \boldsymbol{D}(g_\lambda \circ \boldsymbol{\psi}^{-1} \circ F_{\boldsymbol{\psi},\varphi})(x) \cdot \tilde{X}_\varphi(x)$$

for each $x \in \varphi(U)$; see [Abraham et al. 1988]. Writing $\boldsymbol{r} = \boldsymbol{\psi}(F(q))$, we have

$$\mathscr{L}_X \mathscr{L}_{\tilde{X}}(g_\lambda \circ F)(q)$$

$$
\begin{aligned}
&= \boldsymbol{D}^2(g_\lambda \circ \boldsymbol{\psi}^{-1} \circ F_{\boldsymbol{\psi},\varphi})(\varphi(q)) \cdot (X_\varphi(\varphi(q)), \tilde{X}_\varphi(\varphi(q))) \\
&\qquad + \boldsymbol{D}(g_\lambda \circ \boldsymbol{\psi}^{-1} \circ F_{\boldsymbol{\psi},\varphi})(\varphi(q)) \circ \boldsymbol{D}\tilde{X}_\varphi(\varphi(q)) \cdot X_\varphi(\varphi(q)) \\
&= \boldsymbol{D}(g_\lambda \circ \boldsymbol{\psi}^{-1})(\boldsymbol{r}) \cdot (\boldsymbol{D}^2 F_{\boldsymbol{\psi},\varphi}(\varphi(q)) \cdot (X_\varphi(\varphi(q)), \tilde{X}_\varphi(\varphi(q)))) \\
&\qquad + \boldsymbol{D}^2(g_\lambda \circ \boldsymbol{\psi}^{-1})(\boldsymbol{r}) \cdot (\boldsymbol{D}F_{\boldsymbol{\psi},\varphi}(\varphi(q)) \cdot X_\varphi(\varphi(q)), \boldsymbol{D}F_{\boldsymbol{\psi},\varphi}(\varphi(q)) \cdot \tilde{X}_\varphi(\varphi(q))) \\
&\qquad + \boldsymbol{D}(g_\lambda \circ \boldsymbol{\psi}^{-1} \circ F_{\boldsymbol{\psi},\varphi})(\varphi(q)) \circ \boldsymbol{D}\tilde{X}_\varphi(\varphi(q)) \cdot X_\varphi(\varphi(q)) \\
&= \boldsymbol{D}(g_\lambda \circ \boldsymbol{\psi}^{-1})(\boldsymbol{r}) \cdot (\boldsymbol{D}^2 F_{\boldsymbol{\psi},\varphi}(\varphi(q)) \cdot (X_\varphi(\varphi(q)), \tilde{X}_\varphi(\varphi(q)))) \\
&\qquad + \boldsymbol{D}(g_\lambda \circ \boldsymbol{\psi}^{-1} \circ F_{\boldsymbol{\psi},\varphi})(\varphi(q)) \circ \boldsymbol{D}\tilde{X}_\varphi(\varphi(q)) \cdot X_\varphi(\varphi(q))
\end{aligned}
$$

by the Leibniz rule, chain rule, and the fact that $v_q, \tilde{v}_q \in \ker(TF(q))$. Hence

$$
\begin{aligned}
&\mathscr{L}_X \mathscr{L}_{\tilde{X}}(g_\lambda \circ F)(q) \\
&= Tg_\lambda(F(q)) \circ T\boldsymbol{\psi}^{-1}(\boldsymbol{r}) \cdot (\boldsymbol{D}^2 F_{\boldsymbol{\psi},\varphi}(\varphi(q)) \cdot (X_\varphi(\varphi(q)), \tilde{X}_\varphi(\varphi(q)))) \\
&\qquad + Tg_\lambda(F(q)) \circ T\boldsymbol{\psi}^{-1}(\boldsymbol{r}) \circ \boldsymbol{D}F_{\boldsymbol{\psi},\boldsymbol{\varphi}}(\varphi(q)) \circ \boldsymbol{D}\tilde{X}_\varphi(\varphi(q)) \cdot X_\varphi(\varphi(q)) \\
&= \lambda \circ T\boldsymbol{\psi}^{-1}(\boldsymbol{r}) \cdot (\boldsymbol{D}^2 F_{\boldsymbol{\psi},\varphi}(\varphi(q)) \cdot (X_\varphi(\varphi(q)), \tilde{X}_\varphi(\varphi(q)))) \\
&\qquad + \lambda \circ T\boldsymbol{\psi}^{-1}(\boldsymbol{r}) \circ \boldsymbol{D}F_{\boldsymbol{\psi},\boldsymbol{\varphi}}(\varphi(q)) \circ \boldsymbol{D}\tilde{X}_\varphi(\varphi(q)) \cdot X_\varphi(\varphi(q)).
\end{aligned}
$$

This computation shows that the value of $\mathscr{L}_X \mathscr{L}_{\tilde{X}}(g_\lambda \circ F)(q)$ depends on $X(q) = v_q$, $\tilde{X}(q) = \tilde{v}_q$, and $Tg_\lambda(F(q)) = \lambda$, as well as on the map germ of $\tilde{X}$ at q. To complete the proof, it is enough to show that

$$
\mathscr{L}_X \mathscr{L}_{\tilde{X}}(g_\lambda \circ F)(q) = \mathscr{L}_{\tilde{X}} \mathscr{L}_X(g_\lambda \circ F)(q). \tag{7.1}
$$

Indeed, suppose that (7.1) holds. By exchanging the roles of X and $\tilde{X}$, we conclude that the value of $\mathscr{L}_X \mathscr{L}_{\tilde{X}}(g_\lambda \circ F)(q)$ does not depend on the map germ of $\tilde{X}$ at q. To see that (7.1) holds, observe that

$$
\begin{aligned}
&\mathscr{L}_X \mathscr{L}_{\tilde{X}}(g_\lambda \circ F)(q) - \mathscr{L}_{\tilde{X}} \mathscr{L}_X(g_\lambda \circ F)(q) \\
&= (\mathscr{L}_X \mathscr{L}_{\tilde{X}}(g_\lambda \circ F) - \mathscr{L}_{\tilde{X}} \mathscr{L}_X(g_\lambda \circ F))(q) \\
&= \mathscr{L}_{[X,\tilde{X}]}(g_\lambda \circ F)(q)
\end{aligned}
$$

$$
\begin{aligned}
&= T(g_\lambda \circ F)(q) \cdot [X, \tilde{X}](q) \\
&= Tg_\lambda(F(q)) \circ TF(q) \cdot [X, \tilde{X}](q) \\
&= 0, \qquad (7.2)
\end{aligned}
$$

since $\lambda \in \operatorname{coker}(TF(q))^* \cong \operatorname{image}(TF(q))^0$. ∎

From the proof of Lemma 7.1.4, one can deduce several useful properties of the map $\mathcal{Q}F(q)$. These properties are stated in the following corollaries.

Corollary 7.1.5. *Suppose that $q \in Q$ and (U, φ), $(V, \boldsymbol{\psi})$ are F-compatible charts on Q and R, respectively, such that $q \in U$. Then $\mathcal{Q}F(q) \cdot (v_q, \tilde{v}_q)$ is equal to*

$$
T\boldsymbol{\psi}^{-1}(\boldsymbol{r}) \cdot (\boldsymbol{D}^2 F_{\boldsymbol{\psi},\varphi}(\varphi(q)) \cdot (X_\varphi(\varphi(q)), \tilde{X}_\varphi(\varphi(q)))) + \operatorname{image}(TF(q))
$$

for each $(v_q, \tilde{v}_q) \in \ker(TF(q)) \times \ker(TF(q))$, where $\boldsymbol{r} = \boldsymbol{\psi}(F(q))$ and the C^∞ vector fields $X, \tilde{X}$ are prescribed as in Definition 7.1.3.

Corollary 7.1.6. *We have $\mathcal{Q}F(q) \in \operatorname{Hom}^2(\ker(TF(q)), \operatorname{coker}(TF(q)))$.*

The next result was also pointed out by Agrachev and Gamkrelidze [1991].

Corollary 7.1.7. *Suppose that $q \in Q$ and $R = \mathbb{R}^\ell$. Then the map*

$$
\mathcal{Q}F(q) : T_qE \times T_qE \cong E \times E \to \operatorname{coker}(TF(q)),
$$

prescribed as in Definition 7.1.3 up to the extension of its domain, is well-defined in the sense that its values are invariant under linear automorphisms of R. More precisely, suppose that (U, φ) is a chart on Q such that $q \in U$, and $\boldsymbol{\psi}$, $\tilde{\boldsymbol{\psi}}$ are linear automorphisms of R. Then $\mathcal{Q}F(q) \cdot (v_q, v_q)$ is equal to

$$
\begin{aligned}
&T\boldsymbol{\psi}^{-1}(\boldsymbol{r}) \cdot (\boldsymbol{D}^2 F_{\boldsymbol{\psi},\varphi}(\varphi(q)) \cdot (X_\varphi(\varphi(q)), \tilde{X}_\varphi(\varphi(q)))) + \operatorname{image}(TF(q)) \\
&\quad = T\tilde{\boldsymbol{\psi}}^{-1}(\tilde{\boldsymbol{r}}) \cdot (\boldsymbol{D}^2 F_{\tilde{\boldsymbol{\psi}},\varphi}(\varphi(q)) \cdot (X_\varphi(\varphi(q)), \tilde{X}_\varphi(\varphi(q)))) + \operatorname{image}(TF(q))
\end{aligned}
$$

for each $(v_q, \tilde{v}_q) \in E \times E$, where $\boldsymbol{r} = \boldsymbol{\psi}(F(q))$, $\tilde{\boldsymbol{r}} = \tilde{\boldsymbol{\psi}}(F(q))$, and the C^∞ vector fields $X, \tilde{X}$ are prescribed as in Definition 7.1.3.

Next, we examine the link between IQDs and locally constant-rank maps.

7.1.2 Locally constant-rank maps

For our purposes, the most important property of intrinsic quadratic differentials is that they give us a way of detecting if F is locally constant-rank. To explain this, we begin by recalling a theorem from [Margalef-Roig and Outerelo Domínguez 1992, Chapter 5.1] concerning the local character of locally constant-rank maps.

Definition 7.1.8. We say that F is ***constant-rank*** if

$$\operatorname{rank}(TF(q)) = \operatorname{rank}(TF(\tilde{q})) \tag{7.3}$$

for each $q, \tilde{q} \in Q$. In this case, the ***rank*** of F is $\operatorname{rank}(F) = \operatorname{rank}(TF(q))$. If (7.3) is satisfied whenever $q, \tilde{q}$ are contained in the same connected component of Q, then we say that F is ***locally constant-rank***. Clearly, if Q is connected, then F is locally constant-rank if and only if F is constant-rank.

Theorem 7.1.9. *Suppose that F is constant-rank and $q \in Q$. Then there exist*

1. *F-compatible charts (U, φ) and $(V, \boldsymbol{\psi})$ on Q and R, respectively, such that $q \in U$ and*
2. *A continuous linear map $\boldsymbol{\Lambda} \in \operatorname{Hom}(E, \mathbb{R}^{\ell})$*

such that $\operatorname{rank}(\boldsymbol{\Lambda}) = \operatorname{rank}(F)$ *and* $F_{\boldsymbol{\psi},\varphi}(x) = \boldsymbol{\Lambda}(x)$ *for each* $x \in \varphi(U)$.

Thus F can be locally linearized by suitable choices of charts.

Lemma 7.1.10. *Suppose that F is locally constant-rank and $q \in Q$. Then*

$$\mathcal{Q}F(q) \cdot (v_q, \tilde{v}_q) = 0_{\operatorname{coker}(TF(q))} = 0_{F(q)} + \operatorname{image}(TF(q))$$

for each $(v_q, \tilde{v}_q) \in \ker(TF(q)) \times \ker(TF(q))$.

Proof. Suppose that (U, φ), $(V, \boldsymbol{\psi})$, and $\boldsymbol{\Lambda}$ are prescribed as in Theorem 7.1.9. By Corollary 7.1.5, Theorem 7.1.9, and the fact that the second total derivative of a continuous linear map is identically equal to zero, we have

$$
\begin{aligned}
&\mathscr{Q}F(q) \cdot (v_q, \tilde{v}_q) \\
&= T\boldsymbol{\psi}^{-1}(\boldsymbol{r}) \cdot (\boldsymbol{D}^2 F_{\boldsymbol{\psi},\varphi}(\varphi(q)) \cdot (X_\varphi(\varphi(q)), \tilde{X}_\varphi(\varphi(q)))) + \operatorname{image}(TF(q)) \\
&= T\boldsymbol{\psi}^{-1}(\boldsymbol{r}) \cdot (\boldsymbol{D}^2 \boldsymbol{\Lambda}(\varphi(q)) \cdot (X_\varphi(\varphi(q)), \tilde{X}_\varphi(\varphi(q)))) + \operatorname{image}(TF(q)) \\
&= T\boldsymbol{\psi}^{-1}(\boldsymbol{r}) \cdot \boldsymbol{0}_{\mathbb{R}^\ell} + \operatorname{image}(TF(q)) \\
&= 0_{F(q)} + \operatorname{image}(TF(q)) \\
&= 0_{\operatorname{coker}(TF(q))}
\end{aligned}
$$

for each $(v_q, \tilde{v}_q) \in \ker(TF(q)) \times \ker(TF(q))$. ∎

At this point, we have established the general theory of intrinsic quadratic differentials. In the remainder of this chapter, we specialize to the case where F is one of the anchored endpoint maps of Σ. That is, $F = \operatorname{End}_{x_0}^{\Sigma}$ for some $x_0 \in M$.

7.2 Second variations

To explicitly compute $\mathscr{Q}\operatorname{End}_{x_0}^{\Sigma}$, it will be necessary to evaluate the global flows of time-varying vector fields on open submanifolds of TTM. To make this evaluation, we will invoke several lemmas from Chapter 5. To ensure that their hypotheses are satisfied, Σ must satisfy the additional criteria contained in the next definition.

Definition 7.2.1. We say that Σ has ***second variations*** if, in addition to being a nice C_p^2 control system,

- Σ is a C^3 control system and
- $\boldsymbol{D}_3 f_{\boldsymbol{u}_0}^{\boldsymbol{u}}$ is locally integrably C^2 for each $\boldsymbol{u}_0, \boldsymbol{u} \in \mathscr{U}$, where the controllable time-varying vector field $\boldsymbol{D}_3 f_{\boldsymbol{u}_0} \in \mathscr{V}(J, M, \mathbb{R}^r)$ is defined as in Section 6.1.

In the remainder of this section, we assume that Σ has second variations and

$$\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End}^{\Sigma}_{x_0}).$$

We say that $\boldsymbol{u}_0$ is the ***zeroth-order reference control***. There exists a neighbourhood M_0 of x_0 such that $\Phi^{f^{\boldsymbol{u}_0}}_{t,a}$ is a C^3 diffeomorphism of M_0 onto its image for each $t \in J$. This follows from the assumption that Σ is a C^3 control system, together with an argument identical to the one used in Section 5.4 to obtain the neighbourhood Q_0. For reasons which will be made clear below, we also fix

$$\boldsymbol{u}_1 \in \mathscr{U}.$$

We say that $\boldsymbol{u}_1$ is the ***first-order reference control***.

7.2.1 Basic definitions and properties

In this section, we construct the second variation of Σ along the zeroth- and first-order reference controls $\boldsymbol{u}_0$ and $\boldsymbol{u}_1$. As in Section 6.1, we do so using tangent and vertical lifts. In addition to the notation established in Section 6.1, we define the controllable time-varying vector field $\boldsymbol{D}^2_3 f_{\boldsymbol{u}_0,\boldsymbol{u}_1} \in \mathscr{V}(J, M, \mathbb{R}^r)$ by

$$\boldsymbol{D}^2_3 f_{\boldsymbol{u}_0,\boldsymbol{u}_1}(t, x, \boldsymbol{\omega}) = \boldsymbol{D}^2_3 f(t, x, \boldsymbol{u}_0(t)) \cdot (\boldsymbol{u}_1(t), \boldsymbol{\omega}).$$

On the right-hand side, $\boldsymbol{u}_0$ and $\boldsymbol{u}_1$ denote any representatives of $\boldsymbol{u}_0$ and $\boldsymbol{u}_1$, respectively. Using Lemma 3.1.3, it is not hard to see that the considerations in this chapter involving $\boldsymbol{D}^2_3 f_{\boldsymbol{u}_0,\boldsymbol{u}_1}$ are independent of the particular choices of representatives. Note that $\boldsymbol{D}^2_3 f_{\boldsymbol{u}_0,\boldsymbol{u}_1}$ is well-defined, since

$$\boldsymbol{D}^2_3 f(t, x, \boldsymbol{u}_0(t)) \in \mathrm{Hom}^2(\mathbb{R}^r, T_x M).$$

Definition 7.2.2. The ***second variation of*** Σ ***along*** $\boldsymbol{u}_0$ ***and*** $\boldsymbol{u}_1$ is the pair

$$TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1} = (TTf_{\boldsymbol{u}_0,\boldsymbol{u}_1}, \mathscr{U}),$$

where $TTf_{\boldsymbol{u}_0,\boldsymbol{u}_1} \in \mathscr{V}(J, TTM, \mathbb{R}^r)$ is defined by

$$TTf_{\boldsymbol{u}_0,\boldsymbol{u}_1}(t, \alpha_{v_x}, \boldsymbol{\omega}) = \mathrm{tlft}(\mathrm{tlft}(f^{\boldsymbol{u}_0}) + \mathrm{vlft}(\boldsymbol{D}_3 f^{\boldsymbol{u}_1}_{\boldsymbol{u}_0}))(t, \alpha_{v_x}) \\ + \mathrm{vlft}(\mathrm{tlft}(\boldsymbol{D}_3 f_{\boldsymbol{u}_0}) + \mathrm{vlft}(\boldsymbol{D}_3^2 f_{\boldsymbol{u}_0,\boldsymbol{u}_1}))(t, \alpha_{v_x}, \boldsymbol{\omega}).$$

Here, tlft and vlft are the tangent and vertical lift operations from Chapter 5.

In a natural chart $(TTV, TT\boldsymbol{\psi})$ on TTM, we have

$$\mathrm{tlft}(\mathrm{tlft}(f^{\boldsymbol{u}_0}) + \mathrm{vlft}(\boldsymbol{D}_3 f^{\boldsymbol{u}_1}_{\boldsymbol{u}_0}))_{TT\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{v}, \boldsymbol{X}, \boldsymbol{V}) \\ = \begin{pmatrix} f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \\ \boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{v} + \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{u}_1(t) \\ \boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{X} \\ \hdashline \boldsymbol{D}_2^2 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot (\boldsymbol{v}, \boldsymbol{X}) + \boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{V} + \\ \boldsymbol{D}_2 \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot (\boldsymbol{u}_1(t), \boldsymbol{X}) \end{pmatrix}$$

and

$$\mathrm{vlft}(\mathrm{tlft}(\boldsymbol{D}_3 f_{\boldsymbol{u}_0}) + \mathrm{vlft}(\boldsymbol{D}_3^2 f_{\boldsymbol{u}_0,\boldsymbol{u}_1}))_{TT\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{v}, \boldsymbol{X}, \boldsymbol{V}, \boldsymbol{\omega}) \\ = \begin{pmatrix} \boldsymbol{0}_{\mathbb{R}^n} \\ \boldsymbol{0}_{\mathbb{R}^n} \\ \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{\omega} \\ \boldsymbol{D}_2 \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot (\boldsymbol{\omega}, \boldsymbol{v}) + \boldsymbol{D}_3^2 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot (\boldsymbol{u}_1(t), \boldsymbol{\omega}) \end{pmatrix}.$$

By linearity,

$$(TTf^{\boldsymbol{u}}_{\boldsymbol{u}_0,\boldsymbol{u}_1})_{TT\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{v}, \boldsymbol{X}, \boldsymbol{V})$$

$$= \begin{pmatrix} f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \\ \boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{v} + \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{u}_1(t) \\ \boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{X} + \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{u}(t) \\ - - - - - - - - - - - - - - - \\ \boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot \boldsymbol{V} + \boldsymbol{D}_2^2 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot (\boldsymbol{v}, \boldsymbol{X}) + \\ \boldsymbol{D}_2 \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot (\boldsymbol{u}(t), \boldsymbol{v}) + \boldsymbol{D}_2 \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot (\boldsymbol{u}_1(t), \boldsymbol{X}) + \\ \boldsymbol{D}_3^2 f_{\boldsymbol{\psi}}(t, \boldsymbol{x}, \boldsymbol{u}_0(t)) \cdot (\boldsymbol{u}_1(t), \boldsymbol{u}(t)) \end{pmatrix}. \tag{7.4}$$

We now show that $TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}$ is a control system in the sense of Definition 3.1.4.

Lemma 7.2.3. *The pair $TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}$ is a control system.*

Proof. We must show that each $\boldsymbol{u} \in \mathscr{U}$ is $TTf_{\boldsymbol{u}_0,\boldsymbol{u}_1}$-admissible. To this end, suppose that $\boldsymbol{u} \in \mathscr{U}$, $\mathscr{A}_M$ is a compatible atlas on M, $(TTV, TT\boldsymbol{\psi}) \in TT\mathscr{A}_M$, and

$$(t_0, \boldsymbol{x}_0, \boldsymbol{v}_0, \boldsymbol{X}_0, \boldsymbol{V}_0) \in J \times \boldsymbol{\psi}(V) \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n.$$

Provided that it exists, it is clear that the maximally-defined solution

$$\boldsymbol{\zeta} = (\boldsymbol{\xi}, \boldsymbol{\nu}, \boldsymbol{\Xi}, \boldsymbol{\Upsilon}) : \mathrm{dom}(\boldsymbol{\zeta}) \to \boldsymbol{\psi}(V) \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n,$$

of $((TTf^{\boldsymbol{u}}_{\boldsymbol{u}_0,\boldsymbol{u}_1})_{T\boldsymbol{\psi}}, t_0, \boldsymbol{x}_0, \boldsymbol{v}_0, \boldsymbol{X}_0, \boldsymbol{V}_0)$ must satisfy the following five properties:

- $\mathrm{dom}(\boldsymbol{\zeta}) \subseteq J^{\Sigma_{\boldsymbol{\psi}}}(t_0, \boldsymbol{x}_0, \boldsymbol{u}_0)$,
- $\boldsymbol{\xi}(t) = \Phi^{\Sigma_{\boldsymbol{\psi}}}_{t,t_0}(\boldsymbol{x}_0, \boldsymbol{u}_0)$ for each $t \in \mathrm{dom}(\boldsymbol{\zeta})$,
- $\boldsymbol{\nu}$ is the maximally-defined solution of $(\boldsymbol{g}, t_0, \boldsymbol{v}_0)$, where

  $$\boldsymbol{g} : J^{\Sigma_{\boldsymbol{\psi}}}(t_0, \boldsymbol{x}_0, \boldsymbol{u}_0) \times \mathbb{R}^n \to \mathbb{R}^n$$

  is defined by

  $$\boldsymbol{g}(t, \boldsymbol{v}) = \boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot \boldsymbol{v} + \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot \boldsymbol{u}_1(t),$$

- $\boldsymbol{\Xi}$ is the maximally-defined solution of $(\tilde{\boldsymbol{g}}, t_0, \boldsymbol{X}_0)$, where

  $$\tilde{\boldsymbol{g}} : J^{\Sigma_{\boldsymbol{\psi}}}(t_0, \boldsymbol{x}_0, \boldsymbol{u}_0) \times \mathbb{R}^n \to \mathbb{R}^n$$

  is defined by

  $$\boldsymbol{g}(t, \boldsymbol{X}) = \boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot \boldsymbol{X} + \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot \boldsymbol{u}(t),$$

  and, finally,

- $\boldsymbol{\Upsilon}$ is the maximally-defined solution of $(\boldsymbol{h}, t_0, \boldsymbol{V}_0)$, where

  $$\boldsymbol{h} : J^{\Sigma_{\boldsymbol{\psi}}}(t_0, \boldsymbol{x}_0, \boldsymbol{u}_0) \times \mathbb{R}^n \to \mathbb{R}^n$$

  is defined by

  $$\begin{aligned} &\boldsymbol{h}(t, \boldsymbol{V}) \\ &= \boldsymbol{D}_2 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot \boldsymbol{V} + \boldsymbol{D}_2^2 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot (\boldsymbol{\nu}(t), \boldsymbol{\Xi}(t)) \\ &+ \boldsymbol{D}_2 \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot (\boldsymbol{u}(t), \boldsymbol{\nu}(t)) + \boldsymbol{D}_2 \boldsymbol{D}_3 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot (\boldsymbol{u}_1(t), \boldsymbol{\Xi}(t)) \\ &+ \boldsymbol{D}_3^2 f_{\boldsymbol{\psi}}(t, \boldsymbol{\xi}(t), \boldsymbol{u}_0(t)) \cdot (\boldsymbol{u}_1(t), \boldsymbol{u}(t)). \end{aligned}$$

If we can show that there exist maximally-defined solutions of $(\boldsymbol{g}, t_0, \boldsymbol{v}_0)$, $(\tilde{\boldsymbol{g}}, t_0, \boldsymbol{X}_0)$, and $(\boldsymbol{h}, t_0, \boldsymbol{V}_0)$, then the proof will be complete. In each case, this follows from an argument analogous to the one used in the proof of Lemma 6.1.3. ∎

The next lemma states that, locally, the controlled trajectories of $TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}$ are defined on the same interval as the corresponding $\boldsymbol{u}_0$-controlled trajectories of Σ.

Lemma 7.2.4. *Suppose that* $\boldsymbol{u} \in \mathscr{U}$, $(TTV, TT\boldsymbol{\psi})$ *is a natural chart on* TTM, *and*

$$(t_0, \boldsymbol{x}_0, \boldsymbol{v}_0, \boldsymbol{X}_0, \boldsymbol{V}_0) \in J \times \boldsymbol{\psi}(V) \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n.$$

Then the maximally-defined solution

$$\boldsymbol{\zeta} = (\boldsymbol{\xi}, \boldsymbol{\nu}, \boldsymbol{\Xi}, \boldsymbol{\Upsilon}) : \mathrm{dom}(\boldsymbol{\zeta}) \to \boldsymbol{\psi}(V) \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$$

of $((TTf^{\boldsymbol{u}}_{\boldsymbol{u}_0,\boldsymbol{u}_1})_{T\boldsymbol{\psi}}, t_0, \boldsymbol{x}_0, \boldsymbol{v}_0, \boldsymbol{X}_0, \boldsymbol{V}_0)$ *satisfies*

$$\mathrm{dom}(\boldsymbol{\zeta}) = J^{\Sigma_{\boldsymbol{\psi}}}(t_0, \boldsymbol{x}_0, \boldsymbol{u}_0).$$

Proof. Analogous to the proof of Lemma 6.1.4. ∎

Recalling that $\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End}^{\Sigma}_{x_0})$, the next result follows immediately.

Proposition 7.2.5. *We have*

$$\pi^{-1}_{TTM}(T_{x_0}M) \times \mathscr{U} \subseteq \mathrm{dom}(\mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}).$$

Furthermore,

1. $\pi_{TTM} \circ \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(\alpha_{v_{x_0}}, \boldsymbol{u}) = \mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(v_{x_0}, \boldsymbol{u}_1)$ *and*
2. $\pi_{TTM} \circ s_M \circ \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(\alpha_{v_{x_0}}, \boldsymbol{u}) = \mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(\pi_{TTM} \circ s_M(\alpha_{v_{x_0}}), \boldsymbol{u})$

for each $(\alpha_{v_{x_0}}, \boldsymbol{u}) \in \pi^{-1}_{TTM}(T_{x_0}M) \times \mathscr{U}$.

7.2.2 Second variations compute intrinsic quadratic differentials

It was indicated at the beginning of this chapter that the second variations of Σ can be used to compute IQDs. In this section, we make this statement precise.

Lemma 7.2.6. *Suppose that* $\boldsymbol{u}_1 \in \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))$*. Then*

$$\mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}}, \boldsymbol{u}) \in V_{0_{x_1}}TM$$

for each $\boldsymbol{u} \in \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))$*, where* $x_1 = \mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)$*.*

Proof. Choose $\boldsymbol{u} \in \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))$. By Theorem 6.1.10 and Proposition 7.2.5,

$$\begin{aligned}\pi_{TTM} \circ s_M \circ \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}}, \boldsymbol{u}) &= \mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(\pi_{TTM} \circ s_M(0_{0_{x_0}}), \boldsymbol{u}) \\ &= \mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(0_{x_0}, \boldsymbol{u})\end{aligned}$$

$$
\begin{aligned}
&= T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)\cdot \boldsymbol{u} \\
&= 0_{x_1}.
\end{aligned}
$$

By [Abraham et al. 1988, Exercise 3.3B], the diagram

$$
\begin{tikzcd}
TTM \arrow[rr, shift left, "s_M"] \arrow[dr, "\pi_{TTM}"'] & & TTM \arrow[ll, shift left, "s_M"] \arrow[dl, "T\pi_{TM}"] \\
& TM &
\end{tikzcd}
\tag{7.5}
$$

commutes. It follows that

$$
T\pi_{TM}(\mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(0_{x_0},\boldsymbol{u}_1))\cdot \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}},\boldsymbol{u}) = 0_{x_1}.
$$

Again using Theorem 6.1.10, we see that

$$
\begin{aligned}
\mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}},\boldsymbol{u}) &\in \ker(T\pi_{TM}(\mathrm{End}^{T\Sigma_{\boldsymbol{u}_0}}(0_{x_0},\boldsymbol{u}_1))) \\
&= \ker(T\pi_{TM}(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)\cdot \boldsymbol{u}_1) \\
&= \ker(T\pi_{TM}(0_{x_1})),
\end{aligned}
$$

the latter space being equal to $V_{0_{x_1}}TM$ by definition. $\blacksquare$

Recall from Section 5.1 that each pointwise vertical lift

$$
\mathrm{vlft}_{v_x} : T_xM \to V_{v_x}TM
$$

is a canonical vector space isomorphism, thus invertible.

Theorem 7.2.7. *Suppose that* $\boldsymbol{u}_1 \in \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))$*. Then*

$$
\begin{aligned}
&\mathrm{vlft}^{-1}_{0_{x_1}} \circ \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}},\boldsymbol{u}) \\
&\quad = T\Phi^{f_{\boldsymbol{u}_0}}_{b,a}(x_0)\cdot \int_a^b \mathrm{Ad}^{f_{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3^2 f^{\boldsymbol{u}}_{\boldsymbol{u}_0,\boldsymbol{u}_1})(s,x_0)\,\mathrm{d}s \\
&\qquad + T\Phi^{f_{\boldsymbol{u}_0}}_{b,a}(x_0)\cdot \int_a^b\int_a^s [\mathrm{Ad}^{f_{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})_\sigma, \mathrm{Ad}^{f_{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}_1}_{\boldsymbol{u}_0})_s](x_0)\,\mathrm{d}\sigma\,\mathrm{d}s
\end{aligned}
$$

for each $\boldsymbol{u} \in \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))$*, where* $x_1 = \mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)$*.*

Proof. Choose $\boldsymbol{u} \in \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))$. For notational economy, we write

$$0 = 0_{0_{x_0}}, \quad f = f^{\boldsymbol{u}_0}, \qquad g = \boldsymbol{D}_3 f^{\boldsymbol{u}_1}_{\boldsymbol{u}_0}, \quad \tilde{g} = \boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0}, \quad h = \boldsymbol{D}_3^2 f^{\boldsymbol{u}}_{\boldsymbol{u}_0,\boldsymbol{u}_1},$$

and suppress subscripts on pullbacks throughout the proof. Since Σ has second variations, f is locally integrably C^3 and $g, \tilde{g}$ are locally integrably C^2. Furthermore, h is locally integrably C^0 by Lemma 3.3.9. By construction,

$$\begin{aligned}
\mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0, \boldsymbol{u}) &= \Phi^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}_{b,a}(0, \boldsymbol{u}) \\
&= \Phi^{TTf^{\boldsymbol{u}}_{\boldsymbol{u}_0,\boldsymbol{u}_1}}_{b,a}(0) \\
&= \Phi^{\mathrm{tlft}(\mathrm{tlft}(f)+\mathrm{vlft}(g))+\mathrm{vlft}(\mathrm{tlft}(\tilde{g})+\mathrm{vlft}(h))}_{b,a}(0).
\end{aligned}$$

Invoking Proposition 5.4.2, together with Lemmas 5.5.1, 5.5.2, and 5.5.3, we obtain

$$\begin{aligned}
&\Phi^{\mathrm{tlft}(\mathrm{tlft}(f)+\mathrm{vlft}(g))+\mathrm{vlft}(\mathrm{tlft}(\tilde{g})+\mathrm{vlft}(h))}_{b,a}(0) \\
&= \Phi^{\mathrm{tlft}(\mathrm{tlft}(f))+\mathrm{tlft}(\mathrm{vlft}(g))+\mathrm{vlft}(\mathrm{tlft}(\tilde{g}))+\mathrm{vlft}(\mathrm{vlft}(h))}_{b,a}(0) \\
&= \Phi^{\mathrm{tlft}(\mathrm{tlft}(f))}_{b,a} \circ \Phi^{\mathrm{Ad}^{\mathrm{tlft}(\mathrm{tlft}(f))}(\mathrm{tlft}(\mathrm{vlft}(g))+\mathrm{vlft}(\mathrm{tlft}(\tilde{g}))+\mathrm{vlft}(\mathrm{vlft}(h)))}_{b,a}(0) \\
&= \Phi^{\mathrm{tlft}(\mathrm{tlft}(f))}_{b,a} \circ \Phi^{\mathrm{Ad}^{\mathrm{tlft}(\mathrm{tlft}(f))}(\mathrm{tlft}(\mathrm{vlft}(g)))+\mathrm{Ad}^{\mathrm{tlft}(\mathrm{tlft}(f))}(\mathrm{vlft}(\mathrm{tlft}(\tilde{g})))+\mathrm{Ad}^{\mathrm{tlft}(\mathrm{tlft}(f))}(\mathrm{vlft}(\mathrm{vlft}(h)))}_{b,a}(0) \\
&= \Phi^{\mathrm{tlft}(\mathrm{tlft}(f))}_{b,a} \circ \Phi^{\mathrm{vlft}(\mathrm{tlft}(\mathrm{Ad}^f(g)))+\mathrm{tlft}(\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g})))+\mathrm{vlft}(\mathrm{vlft}(\mathrm{Ad}^f(h)))}_{b,a}(0) \\
&= \Phi^{\mathrm{tlft}(\mathrm{tlft}(f))}_{b,a} \circ \Phi^{\mathrm{tlft}(\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g})))+\mathrm{vlft}(\mathrm{tlft}(\mathrm{Ad}^f(g)))+\mathrm{vlft}(\mathrm{vlft}(\mathrm{Ad}^f(h)))}_{b,a}(0).
\end{aligned}$$

Iterating this procedure, we see that

$$\Phi^{\mathrm{tlft}(\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g})))+\mathrm{vlft}(\mathrm{tlft}(\mathrm{Ad}^f(g)))+\mathrm{vlft}(\mathrm{vlft}(\mathrm{Ad}^f(h)))}_{b,a}(0) = \Phi^{\mathrm{tlft}(\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g})))}_{b,a}(Z),$$

where

$$\begin{aligned}
Z &= \Phi^{\mathrm{Ad}^{\mathrm{tlft}(\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g})))}(\mathrm{vlft}(\mathrm{tlft}(\mathrm{Ad}^f(g))))+\mathrm{Ad}^{\mathrm{tlft}(\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g})))}(\mathrm{vlft}(\mathrm{vlft}(\mathrm{Ad}^f(h))))}_{b,a}(0) \\
&= \Phi^{\mathrm{vlft}(\mathrm{Ad}^{\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g}))}(\mathrm{tlft}(\mathrm{Ad}^f(g))))+\mathrm{vlft}(\mathrm{Ad}^{\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g}))}(\mathrm{vlft}(\mathrm{Ad}^f(h))))}_{b,a}(0) \\
&= \Phi^{\mathrm{vlft}(\mathrm{Ad}^{\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g}))}(\mathrm{tlft}(\mathrm{Ad}^f(g))))+\mathrm{vlft}(\mathrm{vlft}(\mathrm{Ad}^f(h)))}_{b,a}(0) \\
&= \int_a^b \mathrm{Ad}^{\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g}))}(\mathrm{tlft}(\mathrm{Ad}^f(g)))(s, 0_{x_0})\, \mathrm{d}s + \int_a^b \mathrm{vlft}(\mathrm{Ad}^f(h))(s, 0_{x_0})\, \mathrm{d}s
\end{aligned}$$

$$= Z_1 + Z_2.$$

To compute Z_1, we invoke Lemma 5.8.1 to obtain

$$Z_1 = \int_a^b \mathrm{tlft}(\mathrm{Ad}^f(g))(s, 0_{x_0})\,\mathrm{d}s + \int_a^b \int_a^s \mathrm{vlft}([\mathrm{Ad}^f(\tilde{g})_\sigma, \mathrm{Ad}^f(g)_s])(0_{x_0})\,\mathrm{d}\sigma\,\mathrm{d}s.$$

Since $\boldsymbol{u}_1 \in \ker(T\mathrm{End}^\Sigma_{x_0}(\boldsymbol{u}_0))$, Theorem 6.1.10 yields

$$\int_a^b \mathrm{Ad}^f(g)(s, x_0)\,\mathrm{d}s = T\Phi^f_{a,b}(x_0) \cdot T\mathrm{End}^\Sigma_{x_0}(\boldsymbol{u}_0) \cdot \boldsymbol{u}_1 = T\Phi^f_{a,b}(x_0) \cdot 0_{x_1} = 0_{x_0}.$$

Consequently,

$$\int_a^b \mathrm{tlft}(\mathrm{Ad}^f(g))(s, 0_{x_0})\,\mathrm{d}s = 0.$$

This shows that

$$Z_1 + Z_2 = \int_a^b \mathrm{vlft}(\mathrm{Ad}^f(h))(s, 0_{x_0})\,\mathrm{d}s + \int_a^b \int_a^s \mathrm{vlft}([\mathrm{Ad}^f(\tilde{g})_\sigma, \mathrm{Ad}^f(g)_s])(0_{x_0})\,\mathrm{d}\sigma\,\mathrm{d}s$$

is contained in $V_{0_{x_0}}TM$. Thus

$$\begin{aligned}
&\Phi^{\mathrm{tlft}(\mathrm{tlft}(f))}_{b,a} \circ \Phi^{\mathrm{tlft}(\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g})))}_{b,a}(Z_1 + Z_2) \\
&= TT\Phi^f_{b,a} \circ \Phi^{\mathrm{tlft}(\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g})))}_{b,a}(Z_1 + Z_2) \\
&= TT\Phi^f_{b,a} \circ s_M \circ \Phi^{\mathrm{vlft}(\mathrm{tlft}(\mathrm{Ad}^f(\tilde{g})))}_{b,a} \circ s_M(Z_1 + Z_2) \\
&= TT\Phi^f_{b,a} \circ s_M \circ \Phi^{\mathrm{vlft}(\mathrm{tlft}(\mathrm{Ad}^f(\tilde{g})))}_{b,a}(Z_1 + Z_2) \\
&= TT\Phi^f_{b,a} \circ s_M \left(Z_1 + Z_2 + \int_a^b \mathrm{tlft}(\mathrm{Ad}^f(\tilde{g}))(s, 0_{x_0})\,\mathrm{d}s\right).
\end{aligned}$$

Since $\boldsymbol{u} \in \ker(T\mathrm{End}^\Sigma_{x_0}(\boldsymbol{u}_0))$, Theorem 6.1.10 yields

$$\int_a^b \mathrm{Ad}^f(\tilde{g})(s, x_0)\,\mathrm{d}s = 0_{x_0}.$$

Consequently,

$$\int_a^b \mathrm{tlft}(\mathrm{Ad}^f(\tilde{g}))(s, 0_{x_0})\,\mathrm{d}s = 0.$$

This shows that

$$\begin{aligned}\Phi^{\mathrm{tlft}(\mathrm{tlft}(f))}_{b,a} \circ \Phi^{\mathrm{tlft}(\mathrm{vlft}(\mathrm{Ad}^f(\tilde{g})))}_{b,a}(Z_1+Z_2) &= TT\Phi^f_{b,a} \circ s_M(Z_1+Z_2)\\ &= TT\Phi^f_{b,a} \cdot (Z_1+Z_2).\end{aligned}$$

Finally, using Lemma 5.1.8 and Proposition 5.4.4, we obtain

$$\begin{aligned}&\mathrm{vlft}^{-1}_{0_{x_1}} \circ \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0,\boldsymbol{u})\\ &= \mathrm{vlft}^{-1}_{0_{x_1}} \circ TT\Phi^f_{b,a} \cdot (Z_1+Z_2)\\ &= T\Phi^f_{b,a}(x_0) \circ \mathrm{vlft}^{-1}_{0_{x_0}} \cdot (Z_1+Z_2)\\ &= T\Phi^f_{b,a}(x_0) \cdot \int_a^b \mathrm{Ad}^f(h)(s,x_0)\,\mathrm{d}s\\ &\qquad + T\Phi^f_{b,a}(x_0) \cdot \int_a^b \int_a^s [\mathrm{Ad}^f(\tilde{g})_\sigma, \mathrm{Ad}^f(g)_s](x_0)\,\mathrm{d}\sigma\,\mathrm{d}s.\end{aligned}$$

This completes the proof. ∎

The control system $TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}$ has the following bilinearity property.

Corollary 7.2.8. *We have*

$$\mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}}, \boldsymbol{u}+\tilde{\boldsymbol{u}}) = \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}}, \boldsymbol{u}) + \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}}, \tilde{\boldsymbol{u}})$$

and

$$\mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1+\tilde{\boldsymbol{u}}}}(0_{0_{x_0}}, \boldsymbol{u}) = \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}}, \boldsymbol{u}) + \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1+\tilde{\boldsymbol{u}}}}(0_{0_{x_0}}, \boldsymbol{u})$$

for each $\boldsymbol{u}, \tilde{\boldsymbol{u}} \in \mathscr{U}$.

Proof. Analogous to the proof of Corollary 6.1.7. ∎

The next lemma relates $\mathscr{Q}\mathrm{End}^\Sigma_{x_0}(\boldsymbol{u}_0)$ to the controlled trajectories of $TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}$.

Lemma 7.2.9. *We have*

$$\begin{aligned}&\mathscr{Q}\mathrm{End}^\Sigma_{x_0}(\boldsymbol{u}_0) \cdot (\boldsymbol{u}_1, \boldsymbol{u}_2)\\ &\qquad = \mathrm{vlft}^{-1}_{0_{x_1}} \circ \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}}, \boldsymbol{u}_2) + \mathrm{image}(T\mathrm{End}^\Sigma_{x_0}(\boldsymbol{u}_0))\end{aligned}$$

for each $(\boldsymbol{u}_1, \boldsymbol{u}_2) \in \ker(T\mathrm{End}^\Sigma_{x_0}(\boldsymbol{u}_0)) \times \ker(T\mathrm{End}^\Sigma_{x_0}(\boldsymbol{u}_0))$, *where* $x_1 = \mathrm{End}^\Sigma_{x_0}(\boldsymbol{u}_0)$.

Proof. Choose $(\boldsymbol{u}_1, \boldsymbol{u}_2) \in \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)) \times \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))$ and a chart $(V, \boldsymbol{\psi})$ on M such that $x_1 \in V$. By (7.4) and [Margheri 1996, Equations 3.12 and 3.13],

$$\begin{aligned} TT\boldsymbol{\psi} \circ \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}}, \boldsymbol{u}_2) &= \begin{pmatrix} \boldsymbol{\psi}(x_1) \\ \boldsymbol{D}(\boldsymbol{\psi} \circ \mathrm{End}^{\Sigma}_{x_0})(\boldsymbol{u}_0) \cdot \boldsymbol{u}_1 \\ \boldsymbol{D}(\boldsymbol{\psi} \circ \mathrm{End}^{\Sigma}_{x_0})(\boldsymbol{u}_0) \cdot \boldsymbol{u}_2 \\ \boldsymbol{D}^2(\boldsymbol{\psi} \circ \mathrm{End}^{\Sigma}_{x_0})(\boldsymbol{u}_0) \cdot (\boldsymbol{u}_1, \boldsymbol{u}_2) \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{\psi}(x_1) \\ \boldsymbol{0}_{\mathbb{R}^n} \\ \boldsymbol{0}_{\mathbb{R}^n} \\ \boldsymbol{D}^2(\boldsymbol{\psi} \circ \mathrm{End}^{\Sigma}_{x_0})(\boldsymbol{u}_0) \cdot (\boldsymbol{u}_1, \boldsymbol{u}_2) \end{pmatrix}. \end{aligned}$$

Using Lemma 5.1.8 and Corollary 7.1.5,

$$\begin{aligned} &\mathcal{Q}\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0) \cdot (\boldsymbol{u}_1, \boldsymbol{u}_2) \\ &= T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi}(x_1)) \circ \boldsymbol{D}^2(\boldsymbol{\psi} \circ \mathrm{End}^{\Sigma}_{x_0})(\boldsymbol{u}_0) \cdot (\boldsymbol{u}_1, \boldsymbol{u}_2) + \mathrm{image}(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)) \\ &= T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi}(x_1)) \circ \mathrm{vlft}^{-1}_{(\boldsymbol{\psi}(x_1), \boldsymbol{0}_{\mathbb{R}^n})} \circ TT\boldsymbol{\psi} \circ \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}}, \boldsymbol{u}_2) \\ &\qquad + \mathrm{image}(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)) \\ &= T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi}(x_1)) \circ T\boldsymbol{\psi}(x_1) \circ \mathrm{vlft}^{-1}_{0_{x_1}} \circ \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}}, \boldsymbol{u}_2) \\ &\qquad + \mathrm{image}(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)) \\ &= \mathrm{vlft}^{-1}_{0_{x_1}} \circ \mathrm{End}^{TT\Sigma_{\boldsymbol{u}_0,\boldsymbol{u}_1}}(0_{0_{x_0}}, \boldsymbol{u}_2) + \mathrm{image}(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)). \end{aligned}$$

This completes the proof. ∎

The next theorem, which is the main result in this chapter, combines the results of Theorem 7.2.7 and Lemma 7.2.9 to compute the bilinear map $\mathcal{Q}\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)$.

Theorem 7.2.10. *We have*

$$\begin{aligned} &\mathcal{Q}\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0) \cdot (\boldsymbol{u}_1, \boldsymbol{u}_2) \\ &= T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot \int_a^b \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3^2 f^{\boldsymbol{u}_2}_{\boldsymbol{u}_0,\boldsymbol{u}_1})(s, x_0)\, \mathrm{d}s \end{aligned}$$

$$+ \, T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot \int_a^b \int_a^s [\mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}_2}_{\boldsymbol{u}_0})_\sigma, \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}_1}_{\boldsymbol{u}_0})_s](x_0)\,\mathrm{d}\sigma\,\mathrm{d}s$$
$$+ \, \mathrm{image}(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))$$

for each $(\boldsymbol{u}_1, \boldsymbol{u}_2) \in \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)) \times \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))$.

This result matches the expression for $\mathcal{Q}\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)$ derived by Agrachev and Sachkov [2004]. In particular, see [Agrachev and Sachkov 2004, pages 307–308].

The second variations of Σ can be regarded as bilinearized versions of Σ. Thus Theorem 7.2.10 can be regarded as a coordinate-free *bilinearizations compute intrinsic quadratic differentials* principle, in the spirit of [Sontag 1998, Section 2.8]. To the best of our knowledge, the results in this section constitute the first explicit demonstration of such a principle.

7.3 Examples

In this section, we illustrate Theorem 7.2.10 by way of two examples.

Example 7.3.1. Suppose that Σ is a linear system with $M = \mathbb{R}^n$ and

$$\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega}) = \boldsymbol{A} \cdot \boldsymbol{x} + \boldsymbol{B} \cdot \boldsymbol{\omega}$$

for matrices $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n \times r}$. In this case,

$$\mathrm{End}^{\Sigma}_{\boldsymbol{x}}(\boldsymbol{u}) = \exp((b-a)\boldsymbol{A}) \cdot \boldsymbol{x} + \int_a^b \exp((b-\sigma)\boldsymbol{A}) \circ \boldsymbol{B} \cdot \boldsymbol{u}(\sigma)\,\mathrm{d}\sigma$$

and one computes directly that

$$\boldsymbol{D}^2\mathrm{End}^{\Sigma}_{\boldsymbol{x}_0}(\boldsymbol{u}_0) \cdot (\boldsymbol{u}_1, \boldsymbol{u}_2) = \boldsymbol{0}_{\mathbb{R}^n}.$$

To see that this coincides with the result of Theorem 7.2.10, observe that

$$\boldsymbol{D}_3^2 \boldsymbol{f}^{\boldsymbol{u}}_{\boldsymbol{u}_0, \boldsymbol{u}_1}(t, \boldsymbol{x}) = \boldsymbol{0}_{\mathbb{R}^n}$$

Furthermore, we have

$$\mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})(t, \boldsymbol{x}_0) = \exp((a-t)\boldsymbol{A}) \circ \boldsymbol{B} \cdot \boldsymbol{u}(t).$$

Thus $[\mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}_2}_{\boldsymbol{u}_0})_\sigma, \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}_1}_{\boldsymbol{u}_0})_s](\sigma, \boldsymbol{x}_0) = \boldsymbol{0}_{\mathbb{R}^n}$. By Theorem 7.2.10,

$$\mathscr{Q}\mathrm{End}^{\Sigma}_{\boldsymbol{x}_0}(\boldsymbol{u}_0) \cdot (\boldsymbol{u}_1, \boldsymbol{u}_2) = \boldsymbol{0}_{\mathbb{R}^n} + \mathrm{image}(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))$$

for each $(\boldsymbol{u}_1, \boldsymbol{u}_2) \in \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)) \times \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))$, as expected.

In the next example, Σ is a C^3 control-affine system with $\mathscr{U} = L^2(J, \mathbb{R}^r)$. Such a control system has second variations; see Examples 3.1.15 and 3.2.4.

Example 7.3.2. Suppose that

$$f(t, x, \boldsymbol{\omega}) = f_0(x) + \sum_{i=1}^{r} \omega^i f_i(x)$$

for C^3 vector fields $f_0, f_1, \ldots, f_r$. We have $\boldsymbol{D}_3^2 f^{\boldsymbol{u}}_{\boldsymbol{u}_0, \boldsymbol{u}_1}(t, x) = 0_x$. By Theorem 7.2.10,

$$\begin{aligned}
&\mathscr{Q}\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0) \cdot (\boldsymbol{u}_1, \boldsymbol{u}_2) \\
&= T\Phi^{f^{\boldsymbol{u}_0}}_{b,a}(x_0) \cdot \int_a^b \int_a^s [\mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}_2}_{\boldsymbol{u}_0})_\sigma, \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}_1}_{\boldsymbol{u}_0})_s](x_0) \,\mathrm{d}\sigma \,\mathrm{d}s \\
&\quad + \mathrm{image}(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))
\end{aligned}$$

for each $(\boldsymbol{u}_1, \boldsymbol{u}_2) \in \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)) \times \ker(T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0))$. In particular, this result matches the expression for $\mathscr{Q}\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}_0)$ derived in [Agrachev and Sachkov 2004, Section 20.4]. The latter expression is derived under a certain technical assumption which is satisfied, in particular, when Σ is control-affine.

7.4 Discussion

Here we briefly discuss the results obtained in this chapter.

7.4.1 L^∞ controls

To adapt the results in this chapter to accommodate control systems that use L^∞ controls, it is enough to replace the assumption that Σ is a nice C_p^2 control system by the assumption that Σ is a quasi-C^2 control system in the sense of Grasse [1979]. For more information, we refer to [Grasse 1979, Chapter 3].

7.4.2 Related results

Results similar to Theorem 7.2.10 can be found in the literature. For example, we refer to [Agrachev and Sachkov 2004, Section 20.3]. In contrast with the cited work, we do not use the chronological calculus formalism in our approach. Instead, our approach employs the geometry of TTM and, as a result, requires only that Σ meets minimal regularity conditions. The notion of minimality is encapsulated by the requirement that Σ has second variations.

7.4.3 Intrinsic differentials of higher order

Our approach to computing the IQDs of $\mathrm{End}_{x_0}^{\Sigma}$ suggests a general, albeit computationally unwieldy, approach to computing intrinsic differentials of higher order; see, for example, [Porteous 1971]. However, it is not clear what role these intrinsic differentials play in control theory, if any.

Chapter 8

Constant-rank conditions

Consider a control system

$$\Sigma = (f, \mathscr{U})$$

evolving on an n-dimensional manifold M, and let $x_0 \in M$. In Chapter 4, we saw that the continuation method solves the x_0-anchored motion planning problem for Σ by lifting curves in M to curves in $\mathscr{U}$. As described at the beginning of Chapter 6, the candidate curves in M must satisfy

$$\mathrm{image}(\pi) \subseteq \mathrm{image}(\underline{\mathrm{End}}^{\Sigma}_{x_0}). \tag{8.1}$$

Verifying that this constraint is satisfied is difficult, in the sense that it requires a complete characterization of $\mathscr{U}^{\mathrm{sing}}_{x_0}$. In this chapter, we isolate control systems for which $\mathscr{U}^{\mathrm{sing}}_{x_0}$ is empty. For such control systems, (8.1) is trivially satisfied, and thus the second obstruction to the continuation method is overcome. Here, the point of departure is the simple observation that $\mathscr{U}^{\mathrm{sing}}_{x_0}$ is empty if and only if $\mathrm{End}^{\Sigma}_{x_0}$ is a submersion. Since a submersion is a special type of constant-rank map, this observation motivates the search for *constant-rank conditions*; that is, conditions which ensure that $\mathrm{End}^{\Sigma}_{x_0}$ is constant-rank.

This chapter is organized in the following way. We begin in Section 8.1 by establishing preliminary material concerning first-order Pontryagin cones. In Section

8.2, we recall a sufficient constant-rank condition derived by Vakhrameev [1991b]. We then describe how this condition can be used to check if $\mathrm{End}^{\Sigma}_{x_0}$ is a submersion. In Section 8.3, we recall the basic theory of symmetric Lebesgue points. In Section 8.4, we characterize first-order Pontryagin cones in a way that is amenable to computation. In Section 8.5, we prove two simple containment lemmas. In Section 8.6, we derive a necessary and sufficient constant-rank condition which extends [Vakhrameev 1991b, Theorem 1.1]. Our approach uses the results derived in Chapters 5, 6, and 7, and, consequently, is not based on the chronological calculus formalism [Agrachev and Sachkov 2004]. Finally, in Section 8.7, we briefly connect the results of Section 8.6 to the property of subimmersivity.

Our standing assumptions in this chapter are that

- M is a second-countable n-dimensional manifold,
- $\Sigma = (f, \mathscr{U})$ is a nice C^2_p control system evolving on M (see Definition 3.3.7),
- The time domain of Σ is $J = [a, b]$, so that $\mathscr{U} = L^p(J, \mathbb{R}^r)$,
- Σ has second variations in the sense of Definition 7.2.1, and
- $x_0 \in M$.

By Corollary 3.3.12, the map

$$\mathrm{End}^{\Sigma}_{x_0} : \mathrm{dom}(\mathrm{End}^{\Sigma}_{x_0}) \subseteq \mathscr{U} \to M$$

is C^2. Since Σ and x_0 are fixed, we write

$$\mathrm{End} = \mathrm{End}^{\Sigma}_{x_0}$$

and

$$\mu^{\boldsymbol{u}_0}(t) = \mu^{\Sigma}(t, a, x_0, \boldsymbol{u}_0) = \Phi^{f^{\boldsymbol{u}_0}}_{t,a}(x_0)$$

for each $\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End})$. Using this notation, we have $\mathrm{End}(\boldsymbol{u}_0) = \mu^{\boldsymbol{u}_0}(b)$.

For each $\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End})$, there exists a neighbourhood M_0 of x_0 such that $\Phi^{f^{\boldsymbol{u}_0}}_{t,a}|M_0$ is a C^3 diffeomorphism onto its image for each $t \in J$. This follows from the fact that Σ is a C^3 control system, together with an argument identical to the one used in Section 5.4 to obtain the neighbourhood Q_0. We emphasize that M_0 depends on $\boldsymbol{u}_0$, although this is not reflected in the notation. Since $\boldsymbol{u}_0$ will always be understood from context, this should not be the source of any confusion.

8.1 First-order Pontryagin cones

In this section, we define a collection of subspaces which encode the rank of $T\mathrm{End}(\boldsymbol{u}_0)$ for each $\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End})$. Note that, although we have assumed that Σ is a nice C^2_p control system, the contents of this section also apply in the case where Σ is merely a nice C^1_p control system. Indeed, in the latter case, End is C^1. The next definition follows Bonnard and Caillau [2006, Definition 1.1].

Definition 8.1.1. Suppose that $\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End})$. We say that

$$\mathsf{PC}^{\Sigma}_{x_0}(\boldsymbol{u}_0) = T\Phi^{f^{\boldsymbol{u}_0}}_{a,b}(\mathrm{End}(\boldsymbol{u}_0)) \cdot \mathrm{image}(T\mathrm{End}(\boldsymbol{u}_0))$$

is the ***first-order Pontryagin cone along*** $\mu^{\boldsymbol{u}_0}$.[1]

Since Σ and x_0 are fixed, we write

$$\mathsf{PC}(\boldsymbol{u}_0) = \mathsf{PC}^{\Sigma}_{x_0}(\boldsymbol{u}_0)$$

in this chapter. Next, we show that $\mathsf{PC}(\boldsymbol{u}_0)$ encodes the rank of End at $\boldsymbol{u}_0$.

Lemma 8.1.2. *Suppose that* $\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End})$*. Then*

$$\dim(\mathsf{PC}(\boldsymbol{u}_0)) = \mathrm{rank}(T\mathrm{End}(\boldsymbol{u}_0)).$$

[1] Of course, $\mathsf{PC}^{\Sigma}_{x_0}(\boldsymbol{u}_0)$ is trivially a cone since it is a vector subspace of $T_{x_0}M$. When controls are U-valued, where U is a proper subset of $\mathbb{R}^r$, one can define Pontryagin cones in an analogous way [Bonnard and Caillau 2006]. In this situation, however, Pontryagin cones are not subspaces.

Proof. Since $\Phi_{a,b}^{f^{\boldsymbol{u}_0}} | \Phi_{b,a}^{f^{\boldsymbol{u}_0}}(M_0)$ is a C^3 diffeomorphism onto M_0, its differential

$$T\Phi_{a,b}^{f^{\boldsymbol{u}_0}}(\mathrm{End}(\boldsymbol{u}_0))$$

is a vector space isomorphism. Since such isomorphisms preserve dimension,

$$\begin{aligned}
\dim(\mathsf{PC}(\boldsymbol{u}_0)) &= \dim(T\Phi_{a,b}^{f^{\boldsymbol{u}_0}}(\mathrm{End}(\boldsymbol{u}_0)) \cdot \mathrm{image}(T\mathrm{End}(\boldsymbol{u}_0))) \\
&= \dim(\mathrm{image}(T\mathrm{End}(\boldsymbol{u}_0))) \\
&= \mathrm{rank}(T\mathrm{End}(\boldsymbol{u}_0)).
\end{aligned}$$

This completes the proof. ∎

8.2 A sufficient constant-rank condition

In this section, Σ is a C^∞ control-affine system with

$$f(t, x, \boldsymbol{\omega}) = f_0(x) + \sum_{i=1}^{r} \omega^i f_i(x).$$

We assume that Σ is complete and uses L^2 controls, so that

$$\mathrm{dom}(\mathrm{End}) = \mathscr{U} = L^2(J, \mathbb{R}^r).$$

Since the vector fields f_i are C^∞, we can take their iterated Lie brackets of arbitrarily high order. Such Lie brackets will be written compactly using the following standard notation [Sontag 1998]: For C^∞ vector fields X, Y on M, we define

$$\mathrm{ad}_X^0(Y) = Y$$

and inductively $\mathrm{ad}_X^k(Y) = [X, \mathrm{ad}_X^{k-1}(Y)]$ for $k \geq 1$.

The next definition, taken from [Vakhrameev 1991b], can be viewed as a nonlinear analogue of the Cayley–Hamilton theorem [Sontag 1998, Chapter 3.2].

Definition 8.2.1. We say that Σ satisfies the ***local finite definiteness condition*** if for each $x_* \in M$ there exist a neighbourhood V of x_* in M, $\Delta \in \mathbb{Z}_{\geq 0}$, and C^∞ functions $P_{j,k,\ell} : V \to \mathbb{R}$, such that

$$\mathrm{ad}_{f_0}^{\Delta+1}(f_j)(x) = \sum_{k=0}^{\Delta} \sum_{\ell=1}^{r} P_{j,k,\ell}(x) \mathrm{ad}_{f_0}^{k}(f_\ell)(x) \tag{8.2}$$

for each $x \in V$ and each $1 \leq j \leq r$. If (8.2) holds with $V = M$, then we say that Σ satisfies the ***global finite definiteness condition*** with ***degree*** Δ. Of course, if Σ satisfies the global finite definiteness condition with degree Δ, then it satisfies the local finite definiteness condition.

To expand on the remark made prior to Definition 8.2.1, we note that if $M = \mathbb{R}^n$ and Σ is a time-invariant linear system, then Σ satisfies the global finite definiteness condition with degree $\Delta = n - 1$ by virtue of the Cayley–Hamilton theorem.

Example 8.2.2. Observe that Σ satisfies the global finite definiteness condition with degree 0 if and only if there exist C^∞ functions $P_{j,k} : M \to \mathbb{R}$ such that

$$[f_0, f_j] = \sum_{k=1}^{r} P_{j,k} f_k \tag{8.3}$$

for each $1 \leq j \leq r$. For instance, (8.3) is satisfied whenever $[f_0, f_j] \equiv 0$ for each $1 \leq j \leq r$, since in this case we can choose each function $P_{j,k}$ to be identically equal to 0. One may also consider the case where $r = n$ and the vector fields $f_1, \ldots, f_n$ constitute a global frame for TM; that is,

$$f_1(x), \ldots, f_n(x)$$

are linearly independent for each $x \in M$. In this case, (8.3) is satisfied whenever $P_{j,k}$ are the components of $[f_0, f_j]$ in this global frame; see [Lee 2003, Chapter 5].

Lemma 8.2.3. *Suppose that Σ is a C^ω control-affine system. Then Σ satisfies the local finite definiteness condition.*

Proof. See Section A.5. ∎

The next definition is also due to Vakhrameev [1991b], who studied constant-rank conditions in their own right and also in relation to bang-bang theorems.

Definition 8.2.4. We say that Σ satisfies the ***local bang-bang condition*** if for each $x_* \in M$, there exist a neighbourhood V of x_* in M and C^∞ functions

$$Q^{\Delta}_{i,j,k,\ell} : V \to \mathbb{R},$$

such that

$$[f_i, \mathrm{ad}^{\Delta}_{f_0} f_j](x) = \sum_{k=0}^{\Delta} \sum_{\ell=1}^{r} Q^{\Delta}_{i,j,k,\ell}(x) \mathrm{ad}^{k}_{f_0} f_\ell(x) \tag{8.4}$$

for each $x \in V$, each $\Delta \in \mathbb{Z}_{\geq 0}$, and each $1 \leq i, j \leq r$. If (8.4) holds with $V = M$, then we say that Σ satisfies the ***global bang-bang condition***. Of course, if Σ satisfies the global bang-bang condition, then it satisfies the local bang-bang condition.

Example 8.2.5. If $[f_0, f_j] \equiv 0$ for each $1 \leq j \leq r$, then Σ clearly satisfies the global bang-bang condition. In fact, under the same condition, Σ satisfies the global finite definiteness condition with degree 0, as explained in Example 8.2.2.

Theorem 8.2.6. *Suppose that*

- Σ *satisfies the local finite definiteness condition and*
- Σ *satisfies the local bang-bang condition.*

Then End *is constant-rank.*

Proof. See [Vakhrameev 1991b, Theorem 1.3] and [Vakhrameev 1995, p. 2608]. ∎

Corollary 8.2.7. *Suppose that*

- Σ *satisfies the local finite definiteness condition,*

- Σ *satisfies the local bang-bang condition, and*
- *There exists* $\boldsymbol{u}_0 \in \mathscr{U}$ *such that* $\dim(\mathsf{PC}(\boldsymbol{u}_0)) = n$.

Then End *is a* C^2 *submersion.*

Proof. This follows immediately from Lemma 8.1.2 and Theorem 8.2.6. ∎

Using Corollary 8.2.7, we can verify that End is a submersion by first examining the Lie bracket configuration of Σ, and then examining the dependence of $\mathsf{PC}(\boldsymbol{u}_0)$ on $\boldsymbol{u}_0$. Clearly, this is a considerable improvement over a direct verification. As we will see in Section 8.4, it is possible to go somewhat further and characterize $\mathsf{PC}(\boldsymbol{u}_0)$ in a more computable fashion.

8.3 Symmetric Lebesgue points

In this section, we recall the basic theory of symmetric Lebesgue points and their associated Lebesgue values, following Mikkola [2002, Appendix B.5]. Throughout this section, E is a finite-dimensional vector space.

Suppose that $g : J \to E$ is continuous at the point $t \in (a, b)$. Clearly,

$$\lim_{\delta \searrow 0} \frac{1}{\delta} \int_{t-\delta}^{t} \|g(s) - g(t)\| \, \mathrm{d}s = 0 \tag{8.5}$$

and

$$\lim_{\delta \searrow 0} \frac{1}{\delta} \int_{t}^{t+\delta} \|g(s) - g(t)\| \, \mathrm{d}s = 0, \tag{8.6}$$

where $\| \cdot \|$ is any choice of norm on E. Now suppose that $g \in L^1(J, E)$. By definition, g is an equivalence class of maps. Thus the value of g at $t \in J$ is not well-defined. Nevertheless, one can consider the points $t \in (a, b)$ at which we can associate an element $g(t) \in E$ such that (8.5) and (8.6) hold. These points are called the symmetric Lebesgue points of g. The end result is that one can work with the symmetric Lebesgue points of g as if they were points of continuity of g, at least with respect to the validity of (8.5) and (8.6).

Definition 8.3.1. Suppose that $t \in (a,b)$. We say that t is a ***symmetric Lebesgue point*** of g if there exists $e \in E$ such that

$$\lim_{\delta \searrow 0} \frac{1}{\delta} \int_{t-\delta}^{t} \|g(s) - e\| \,\mathrm{d}s = 0 \tag{8.7}$$

and

$$\lim_{\delta \searrow 0} \frac{1}{\delta} \int_{t}^{t+\delta} \|g(s) - e\| \,\mathrm{d}s = 0, \tag{8.8}$$

where $\|\cdot\|$ is any choice of norm on E. The set of all symmetric Lebesgue points of g is denoted by $\mathrm{Leb}(g)$.

Note that if $t \in \mathrm{Leb}(g)$ and $e \in E$ satisfies (8.7) and (8.8), then

$$\lim_{\delta \searrow 0} \frac{1}{\delta} \int_{t-\delta}^{t} g(s) \,\mathrm{d}s = e \tag{8.9}$$

and

$$\lim_{\delta \searrow 0} \frac{1}{\delta} \int_{t}^{t+\delta} g(s) \,\mathrm{d}s = e. \tag{8.10}$$

For example, to prove the first assertion, simply observe that

$$\lim_{\delta \searrow 0} \left\| \frac{1}{\delta} \int_{t-\delta}^{t} g(s) \,\mathrm{d}s - e \right\| \leq \lim_{\delta \searrow 0} \frac{1}{\delta} \int_{t-\delta}^{t} \|g(s) - e\| \,\mathrm{d}s = 0.$$

This shows, in particular, that e is the unique element of E which satisfies (8.7) and (8.8). We say that e is the ***Lebesgue value*** of g at t, and we write

$$g(t) = e.$$

We stress that this notation does not create any ambiguity. Indeed, the value of g at t is not well-defined, as mentioned above.

The following facts about symmetric Lebesgue points will be useful.

- Suppose that $t \in \mathrm{Leb}(g)$. Then for each $\varepsilon \in \mathbb{R}_{>0}$, there exists a neighbourhood $U(t, \varepsilon)$ of t in (a, b) such that the inequalities
$$\left\| \int_{t-\delta}^{t} g(s)\, \mathrm{d}s - \delta g(t) \right\| < \delta\varepsilon \quad \text{and} \quad \left\| \int_{t}^{t+\tilde{\delta}} g(s)\, \mathrm{d}s - \tilde{\delta} g(t) \right\| < \tilde{\delta}\varepsilon$$
hold whenever $\delta, \tilde{\delta} \in \mathbb{R}_{>0}$ are such that $t-\delta, t+\tilde{\delta} \in U(t, \varepsilon)$. This is a consequence of (8.9) and (8.10). Combining the above inequalities, it follows that
$$\left\| \int_{t-\delta}^{t+\tilde{\delta}} g(s)\, \mathrm{d}s - (\delta + \tilde{\delta}) g(t) \right\| < (\delta + \tilde{\delta})\varepsilon$$
whenever $\delta, \tilde{\delta} \in \mathbb{R}_{>0}$ are such that $t - \delta, t + \tilde{\delta} \in U(t, \varepsilon)$.

- The set $(a, b) \setminus \mathrm{Leb}(g)$ has measure zero; see, for example, [Hewitt and Stromberg 1965, Lemma 18.4]. For our purposes, the important consequence of this fact is that $\mathrm{Leb}(g)$ is dense in (a, b), and, in turn, $\mathrm{Leb}(g)$ is dense in $J = [a, b]$.[2]

- Finally, if g is continuous, in the sense that it has a continuous representative, then $\mathrm{Leb}(g) = (a, b)$.

8.4 A characterization of first-order Pontryagin cones

In this section, we characterize the vector subspaces $\mathsf{PC}(\boldsymbol{u}_0)$ in a way that is amenable to computation. We require the following specialized notation. For each $\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End})$ and each $\boldsymbol{u} \in \mathscr{U}$, define $\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u} \in L^1(J, T_{x_0}M)$ by
$$(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(t) = \mathrm{Ad}_{M_0}^{f^{\boldsymbol{u}_0}} (\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})(t, x_0),$$
where $\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0}$ is prescribed as in Chapter 6. The fact that $\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u}$ is integrable follows from Lemma 3.3.9 and Proposition 5.4.2.

[2] It is enough to show that each nonempty open subset of (a, b) has a nonempty intersection with $\mathrm{Leb}(g)$. Suppose that $U \subseteq (a, b)$ is such a subset. Since U is nonempty and open in (a, b), it must contain an open subinterval of (a, b). This implies that the measure of U is positive, so U cannot be contained in $(a, b) \setminus \mathrm{Leb}(g)$. Hence $U \cap \mathrm{Leb}(g) \neq \varnothing$, and $\mathrm{Leb}(g)$ is dense in (a, b).

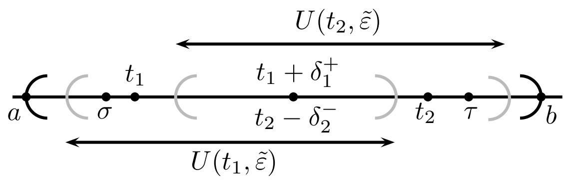

Figure 8.1: An illustration of the proof of Theorem 8.4.3

The notion of a simple chain will be useful.

Definition 8.4.1. Suppose that T is a topological space and $t, \tilde{t} \in T$. A ***simple chain*** connecting t and $\tilde{t}$ is a sequence $U_1, \dots, U_k$ of open subsets of T such that $t \in U_1$, $\tilde{t} \in U_k$, and $U_i \cap U_j \neq \varnothing$ if and only if $|i - j| \leq 1$.

Theorem 8.4.2. *Suppose that T is a topological space, $t, \tilde{t}$ lie in the same connected component of T, and $\mathscr{C}$ is an open cover of T. Then there exists a simple chain $U_1, \dots, U_k$ connecting t and $\tilde{t}$ with the property that $U_i \in \mathscr{C}$ for each $1 \leq i \leq k$.*

Proof. See [Willard 1970, Theorem 26.15]. ∎

The next theorem bears close similarity to results appearing, without proof, in the work of Agrachev and Gamkrelidze [1985, 1991], Agrachev and Vakhrameev [1984, 1986], and Vakhrameev [1991a,b, 1995, 1996, 1998].

Theorem 8.4.3. *Suppose that $\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End})$. Then $\mathsf{PC}(\boldsymbol{u}_0)$ coincides with*

$$\mathsf{Q}(\boldsymbol{u}_0) = \mathrm{span}\left\{(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(t) \ : \ \boldsymbol{u} \in \mathscr{U},\, t \in \mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})\right\}.$$

Proof. First observe that by Theorem 6.1.10,

$$\begin{aligned} T\Phi^{f^{\boldsymbol{u}_0}}_{a,b}(\mathrm{End}(\boldsymbol{u}_0)) \circ T\mathrm{End}(\boldsymbol{u}_0) \cdot \boldsymbol{u} &= \int_a^b \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})(s, x_0)\,\mathrm{d}s \\ &= \int_a^b (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s \end{aligned} \tag{8.11}$$

for each $\boldsymbol{u} \in \mathscr{U}$. Choose

$$q = \sum_{j=1}^{k} C_j (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u}_j)(t_j) \in \mathsf{Q}(\boldsymbol{u}_0)$$

and let $\{\delta_n\}_{n\in\mathbb{N}}$ be a sequence of positive real numbers such that $\delta_n \to 0$ as $n \to \infty$. Without loss of generality, we can assume that for each $n \in \mathbb{N}$, we have

- $[t_j, t_j + \delta_n] \subseteq (a, b)$ for each $1 \leq j \leq k$ and
- $[t_i, t_i + \delta_n] \cap [t_j, t_j + \delta_n] = \varnothing$ for each $1 \leq i, j \leq k$ with $i \neq j$.

Consider the sequence $\{\boldsymbol{w}_n\}_{n\in\mathbb{N}}$ in $\mathscr{U}$ defined by

$$\boldsymbol{w}_n(t) = \begin{cases} \delta_n^{-1} C_j \boldsymbol{u}_j(t), & t \in [t_j, t_j + \delta_n], \\ \mathbf{0}_{\mathbb{R}^r}, & \text{otherwise.} \end{cases}$$

We say that each $\boldsymbol{w}_n$ is a *control variation*. Using (8.11),

$$\begin{aligned} \lim_{n\to\infty} T\Phi^{f^{\boldsymbol{u}_0}}_{a,b}(\mathrm{End}(\boldsymbol{u}_0)) \circ T\mathrm{End}(\boldsymbol{u}_0) \cdot \boldsymbol{w}_n &= \lim_{n\to\infty} \int_a^b (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{w}_n)(s)\,\mathrm{d}s \\ &= \sum_{j=1}^k C_j \lim_{n\to\infty} \frac{1}{\delta_n} \int_{t_j}^{t_j+\delta_n} (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u}_j)(s)\,\mathrm{d}s \\ &= \sum_{j=1}^k C_j (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u}_j)(t_j). \end{aligned}$$

Consequently, $q \in \overline{\mathsf{PC}(\boldsymbol{u}_0)} = \mathsf{PC}(\boldsymbol{u}_0)$ and $\mathsf{Q}(\boldsymbol{u}_0) \subseteq \mathsf{PC}(\boldsymbol{u}_0)$.

Now choose $\boldsymbol{u} \in \mathscr{U}$. Recall that for each $t \in \mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})$ and each $\varepsilon \in \mathbb{R}_{>0}$, there exists a neighbourhood $U(t, \varepsilon)$ of t in (a, b) such that

$$\left\| \int_{t-\delta}^{t+\tilde{\delta}} (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s - (\delta + \tilde{\delta})(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(t) \right\| \leq (\delta + \tilde{\delta})\varepsilon \tag{8.12}$$

whenever $\delta, \tilde{\delta} \in \mathbb{R}_{>0}$ are such that $t - \delta, t + \tilde{\delta} \in U(t, \varepsilon)$. Now let $\varepsilon \in \mathbb{R}_{>0}$. Since $\mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})$ is dense in J and the maps

$$t \mapsto \int_a^t (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s \quad \text{and} \quad t \mapsto \int_t^b (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s$$

are continuous, there exist $\sigma, \tau \in \mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})$ such that

$$\left\| \int_a^\sigma (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s \right\|, \left\| \int_\tau^b (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s \right\| < \frac{\varepsilon}{3}.$$

Set $\tilde{\varepsilon} = \varepsilon/(3(b - a))$ and consider the set

$$\mathscr{C} = \{U(t, \tilde{\varepsilon}) \ : \ t \in \mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})\}.$$

Since $\mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})$ is dense in J, $\mathscr{C}$ is an open cover of (a,b). Relative to $\mathscr{C}$, choose a simple chain $U(t_1,\tilde{\varepsilon}),\dots,U(t_k,\tilde{\varepsilon})$ connecting σ to τ as in Theorem 8.4.2. By definition, $\sigma \in U(t_1,\tilde{\varepsilon})$ and $\tau \in U(t_k,\tilde{\varepsilon})$. Without loss of generality, we can assume that $\sigma < t_1$ and $t_k < \tau$. This allows us to write

$$\int_\sigma^\tau (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s = \sum_{j=1}^k \int_{t_j-\delta_j^-}^{t_j+\delta_j^+} (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s,$$

where $\delta_j^-, \delta_j^+ \in \mathbb{R}_{>0}$, $1 \le j \le k$ are such that

- $t_1 - \delta_1^- = \sigma$, $t_k + \delta_k^+ = \tau$,
- $t_j + \delta_j^+ = t_{j+1} - \delta_{j+1}^-$ for each $1 \le j \le k-1$, and
- $t_j + \delta_j^+ = t_{j+1} - \delta_{j+1}^- \in U(t_j,\tilde{\varepsilon}) \cap U(t_{j+1},\tilde{\varepsilon})$ for each $1 \le j \le k-1$.

By construction,

$$\sum_{j=1}^k (\delta_j^- + \delta_j^+) < \tau - \sigma < b - a.$$

Using (8.12), we can now complete the proof. We have

$$\begin{aligned}
&\left\| \int_a^b (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s - \sum_{j=1}^k (\delta_j^- + \delta_j^+)(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(t_j) \right\| \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \left\| \int_\sigma^\tau (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s - \sum_{j=1}^k (\delta_j^- + \delta_j^+)(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(t_j) \right\| \\
&= \frac{2\varepsilon}{3} + \left\| \sum_{j=1}^k \int_{t_j-\delta_j^-}^{t_j+\delta_j^+} (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s - (\delta_j^- + \delta_j^+)(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(t_j) \right\| \\
&\le \frac{2\varepsilon}{3} + \sum_{j=1}^k \left\| \int_{t_j-\delta_j^-}^{t_j+\delta_j^+} (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s - (\delta_j^- + \delta_j^+)(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(t_j) \right\| \\
&< \frac{2\varepsilon}{3} + \tilde{\varepsilon} \sum_{j=1}^k (\delta_j^- + \delta_j^+) \\
&< \frac{2\varepsilon}{3} + \tilde{\varepsilon}(b-a) \\
&= \varepsilon.
\end{aligned}$$

Consequently,

$$T\Phi^{f\boldsymbol{u}_0}_{a,b}(\mathrm{End}(\boldsymbol{u}_0)) \circ T\mathrm{End}(\boldsymbol{u}_0) \cdot \boldsymbol{u} = \int_a^b (\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(s)\,\mathrm{d}s \in \overline{\mathsf{Q}(\boldsymbol{u}_0)} = \mathsf{Q}(\boldsymbol{u}_0)$$

and $\mathsf{PC}(\boldsymbol{u}_0) \subseteq \mathsf{Q}(\boldsymbol{u}_0)$. This completes the proof. ∎

8.5 Containment lemmas

In this section, we prove two technical lemmas dealing with containment.

Lemma 8.5.1. *Suppose that* $\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End})$ *and* $\boldsymbol{u} \in \mathscr{U}$*. Then*

$$(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(t) \in \mathsf{PC}(\boldsymbol{u}_0)$$

for a.a. $t \in J$*.*

Proof. This follows immediately from Theorem 8.4.3 and the fact that

$$(a, b) \setminus \mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})$$

has measure zero. ∎

For the next lemma, we require the following specialized notation. For each $\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End})$ and each $\boldsymbol{u}, \tilde{\boldsymbol{u}} \in \mathscr{U}$, define $\lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}}) \in L^1(J, T_{x_0}M)$ by

$$(\lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}}))(t) = \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}^2_3 f^{\tilde{\boldsymbol{u}}}_{\boldsymbol{u}_0, \boldsymbol{u}})(t, x_0),$$

where $\boldsymbol{D}^2_3 f^{\tilde{\boldsymbol{u}}}_{\boldsymbol{u}_0, \boldsymbol{u}}$ is prescribed as in Chapter 7. We also define

$$\Lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}}) \in L^1(J, T_{x_0}M)$$

by

$$(\Lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}}))(t) = \int_a^t [\mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\boldsymbol{u}}_{\boldsymbol{u}_0})_\sigma, \mathrm{Ad}^{f^{\boldsymbol{u}_0}}_{M_0}(\boldsymbol{D}_3 f^{\tilde{\boldsymbol{u}}}_{\boldsymbol{u}_0})_t](x_0)\, \mathrm{d}\sigma.$$

The fact that $\lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}})$ is integrable follows from Lemma 3.3.9 and Proposition 5.4.2, while the fact that $\Lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}})$ is integrable follows from Lemma 3.3.9, Proposition 5.4.2, and Lemma 5.4.5.

Recall from Section 7.1 that if $M = \mathbb{R}^n$, then each intrinsic quadratic differential

$$\mathscr{Q}\mathrm{End}(\boldsymbol{u}_0)$$

is well-defined on $\mathscr{U} \times \mathscr{U}$, in the sense that it its values are invariant under linear automorphisms of M. In the proof of the next lemma, we use this fact to construct control variations as in the proof of Theorem 8.4.3.

Lemma 8.5.2. *Suppose that $M = \mathbb{R}^n$, the map* End *is locally constant-rank,*

$$\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End}),$$

and $\boldsymbol{u}, \tilde{\boldsymbol{u}} \in \mathscr{U}$. Then

$$(\lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}}))(t) + (\Lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}}))(t) \in \mathsf{PC}(\boldsymbol{u}_0)$$

for a.a. $t \in J$.

Proof. Since End is locally constant-rank, we have

$$\mathscr{Q}\mathrm{End}(\boldsymbol{u}_0) \cdot (\boldsymbol{w}, \tilde{\boldsymbol{w}}) = 0_{\mathrm{coker}(T\mathrm{End}(\boldsymbol{u}_0))} \cong \mathrm{image}(T\mathrm{End}(\boldsymbol{u}_0))$$

for each $(\boldsymbol{w}, \tilde{\boldsymbol{w}}) \in \mathscr{U} \times \mathscr{U}$. It follows from Theorem 7.2.10 that

$$\int_a^b (\lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{w}, \tilde{\boldsymbol{w}}))(s)\, \mathrm{d}s + \int_a^b (\Lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{w}, \tilde{\boldsymbol{w}}))(s)\, \mathrm{d}s$$

is contained in

$$T\Phi^{f^{\boldsymbol{u}_0}}_{a,b}(\mathrm{End}(\boldsymbol{u}_0)) \cdot \mathrm{image}(T\mathrm{End}(\boldsymbol{u}_0)) = \mathsf{PC}(\boldsymbol{u}_0)$$

for each $(\boldsymbol{w}, \tilde{\boldsymbol{w}}) \in \mathscr{U} \times \mathscr{U}$. Choose

$$t \in \mathrm{Leb}(\lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}})) \cap \mathrm{Leb}(\Lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}}))$$

and let $\{\delta_n\}_{n\in\mathbb{N}}$ be a sequence of positive real numbers such that $\delta_n \to 0$ as $n \to \infty$. Without loss of generality, we can assume that for each $n \in \mathbb{N}$,

$$[t, t + \delta_n] \subseteq J.$$

Consider the sequences $\{\boldsymbol{w}_n\}_{n\in\mathbb{N}}, \{\tilde{\boldsymbol{w}}_n\}_{n\in\mathbb{N}}$ in $\mathscr{U}$ defined by

$$\boldsymbol{w}_n(s) = \begin{cases} \frac{\boldsymbol{u}(s)}{\sqrt{\delta_n}}, & s \in [t, t+\delta_n], \\ \mathbf{0}_{\mathbb{R}^r}, & \text{otherwise,} \end{cases}$$

and

$$\tilde{\boldsymbol{w}}_n(s) = \begin{cases} \frac{\tilde{\boldsymbol{u}}(s)}{\sqrt{\delta_n}}, & s \in [t, t+\delta_n], \\ \mathbf{0}_{\mathbb{R}^r}, & \text{otherwise.} \end{cases}$$

Using these control variations, we have

$$\begin{aligned} \lim_{n\to\infty} \int_a^b (\lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{w}_n, \tilde{\boldsymbol{w}}_n))(s)\,\mathrm{d}s &= \lim_{n\to\infty} \frac{1}{\delta_n} \int_t^{t+\delta_n} (\lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}}))(s)\,\mathrm{d}s \\ &= (\lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}}))(t) \end{aligned}$$

and similarly

$$\lim_{n\to\infty} \int_a^b (\Lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{w}_n, \tilde{\boldsymbol{w}}_n))(s)\,\mathrm{d}s = (\Lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}}))(t).$$

Consequently,

$$\lambda^2_{\boldsymbol{u}_0}(t) \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}}) + \Lambda^2_{\boldsymbol{u}_0}(t) \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}}) \in \overline{\mathsf{PC}(\boldsymbol{u}_0)} = \mathsf{PC}(\boldsymbol{u}_0).$$

To complete the proof, it is enough to recall the fact that

$$(a, b) \setminus (\mathrm{Leb}(\lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}})) \cap \mathrm{Leb}(\Lambda^2_{\boldsymbol{u}_0} \cdot (\boldsymbol{u}, \tilde{\boldsymbol{u}})))$$

has measure zero. ∎

8.6 A necessary and sufficient constant-rank condition

In this section, we derive an extension of [Vakhrameev 1991b, Theorem 1.1]. The latter theorem is a necessary and sufficient constant-rank condition whose applicability is limited to complete, time-invariant, C^∞ control systems. The extended version accommodates weakly regular, time-varying, fully nonlinear control systems.

We will require the notion of a time-varying subspace. In what follows, E is a finite-dimensional vector space. If $\tilde{E}$ is a vector subspace of E, then

$$\mathrm{proj}_{\tilde{E}} \in \mathrm{Hom}(E)$$

denotes the orthogonal projection of E onto $\tilde{E}$.

Definition 8.6.1. An assignment $[0,1] \ni t \mapsto S(t)$, where $S(t)$ is a ρ-dimensional vector subspace of E, is called a ρ-dimensional ***time-varying subspace*** of E. We say that S is C^k, where $k \in \mathbb{N}^*$, if for each $t_0 \in [0,1]$, there exist

- A relatively open subinterval $I \subseteq [0,1]$ such that $t_0 \in I$ and
- C^k maps $b_1, \ldots, b_\rho : I \to E$ such that

$$S(t) = \mathrm{span}\{b_1(t), \ldots, b_\rho(t)\}$$

  for each $t \in I$.

Lemma 8.6.2. *Suppose that $t \mapsto S(t)$ is a ρ-dimensional time-varying subspace of E and define $\mathrm{proj}_S : [0,1] \to \mathrm{Hom}(E)$ by*

$$\mathrm{proj}_S(t) = \mathrm{proj}_{S(t)}.$$

Then S is C^k, where $k \in \mathbb{N}^*$, if and only if proj_S is C^k.

Proof. For definiteness, let us suppose that E is ℓ-dimensional. By choosing a basis for E, we can identify E and $\mathrm{Hom}(E)$ with $\mathbb{R}^\ell$ and $\mathbb{R}^{\ell \times \ell}$, respectively.

Suppose that S is C^k, where $k \in \mathbb{N}^*$. We must show that proj_S is C^k or, equivalently, that for each $t_0 \in [0,1]$, there exists a relatively open subinterval $I \subseteq [0,1]$ such that $t_0 \in I$ and $\mathrm{proj}_S|I$ is C^k. To this end, let $t_0 \in [0,1]$, let I and $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\rho$ be prescribed as in Definition 8.6.1, and define $\boldsymbol{B} : I \to \mathbb{R}^{\ell \times \ell}$ by

$$\boldsymbol{B}(t) = [\boldsymbol{b}_1(t) \cdots \boldsymbol{b}_\rho(t)].$$

Clearly, $\boldsymbol{B}$ is C^k. Since

$$(\text{proj}_S|I)(t) = \boldsymbol{B}(t)(\boldsymbol{B}(t)^*\boldsymbol{B}(t))^{-1}\boldsymbol{B}(t)^*$$

for each $t \in I$, we conclude that $\text{proj}_S|I$ is C^k.

Conversely, suppose that proj_S is C^k. We must show that S is C^k. To this end, let $t_0 \in [0,1]$ and let $\boldsymbol{b}_1^0, \dots, \boldsymbol{b}_\rho^0 \in S(t_0)$ be linearly independent. Given $1 \leq i \leq \rho$, let $\boldsymbol{b}_i : [0,1] \to \mathbb{R}^\ell$ be the maximally-defined solution of the initial value problem

$$\begin{cases} \dot{\boldsymbol{\xi}}(t) = \dot{\widehat{\text{proj}_S}}(t) \cdot \boldsymbol{\xi}(t), \quad \boldsymbol{\xi}(t) \in \mathbb{R}^\ell, \quad t \in [0,1] \\ \boldsymbol{\xi}(t_0) = \boldsymbol{b}_i^0. \end{cases}$$

The fact that $\boldsymbol{b}_i$ is defined on $[0,1]$ follows from Lemma 2.2.17. Furthermore, $\boldsymbol{b}_i$ is C^k, since the right-hand side

$$(t, \boldsymbol{x}) \mapsto \dot{\widehat{\text{proj}_S}}(t) \cdot \boldsymbol{x}$$

is clearly C^{k-1}. By continuity, there exists a relatively open subinterval $I \subseteq [0,1]$ such that $t_0 \in I$ and $\boldsymbol{b}_1(t), \dots, \boldsymbol{b}_\rho(t)$ are linearly independent for each $t \in I$. To finish the proof, we now show that

$$\boldsymbol{b}_1(t), \dots, \boldsymbol{b}_\rho(t) \in S(t)$$

for each $t \in I$. Given $1 \leq i \leq \rho$, consider the map $\boldsymbol{c}_i : I \to \mathbb{R}^\ell$ defined by

$$\boldsymbol{c}_i(t) = \boldsymbol{A}(t) \cdot \boldsymbol{b}_i(t) = (\text{id}_{\mathbb{R}^\ell} - \text{proj}_S(t)) \cdot \boldsymbol{b}_i(t).$$

By the Leibniz rule, $\boldsymbol{c}_i$ is C^k. Differentiating the identity

$$\boldsymbol{A}(t) \circ \text{proj}_S(t) = \boldsymbol{0}_{\mathbb{R}^{\ell \times \ell}},$$

we have

$$\boldsymbol{A}(t) \circ \dot{\widehat{\text{proj}_S}}(t) = -\dot{\boldsymbol{A}}(t) \circ \text{proj}_S(t)$$

and hence

$$\begin{aligned}
\dot{\boldsymbol{c}}_i(t) &= \dot{\boldsymbol{A}}(t)\cdot\boldsymbol{b}_i(t) + \boldsymbol{A}(t)\cdot\dot{\boldsymbol{b}}_i(t)\\
&= \dot{\boldsymbol{A}}(t)\cdot\boldsymbol{b}_i(t) + \boldsymbol{A}(t)\circ\widehat{\dot{\mathrm{proj}}_S}(t)\cdot\boldsymbol{b}_i(t)\\
&= \dot{\boldsymbol{A}}(t)\cdot\boldsymbol{b}_i(t) - \dot{\boldsymbol{A}}(t)\circ\mathrm{proj}_S(t)\cdot\boldsymbol{b}_i(t)\\
&= \dot{\boldsymbol{A}}(t)\circ(\mathrm{id}_{\mathbb{R}^\ell} - \mathrm{proj}_S(t))\cdot\boldsymbol{b}_i(t)\\
&= \dot{\boldsymbol{A}}(t)\cdot\boldsymbol{c}_i(t)
\end{aligned}$$

for each $t\in I$. Thus $\boldsymbol{c}_i$ is the maximally-defined solution of

$$\begin{cases}\dot{\boldsymbol{\xi}}(t) = \dot{\boldsymbol{A}}(t)\cdot\boldsymbol{\xi}(t), \quad \boldsymbol{\xi}(t)\in\mathbb{R}^\ell, \quad t\in I\\ \boldsymbol{\xi}(t_0) = (\mathrm{id}_{\mathbb{R}^\ell} - \mathrm{proj}_S(t_0))\cdot\boldsymbol{b}_i^0.\end{cases}$$

By construction, $\boldsymbol{b}_i^0\in S(t_0)$, so that

$$(\mathrm{id}_{\mathbb{R}^\ell} - \mathrm{proj}_S(t_0))\cdot\boldsymbol{b}_i^0 = \mathbf{0}_{\mathbb{R}^\ell}.$$

We conclude that $\boldsymbol{c}_i$ is identically equal to $\mathbf{0}_{\mathbb{R}^\ell}$. Equivalently,

$$\boldsymbol{b}_i(t)\in S(t)$$

for each $t\in I$. This completes the proof. ∎

The next theorem extends [Vakhrameev 1991b, Theorem 1.1]. To avail ourselves of Lemma 8.5.2, we include the hypothesis that $M=\mathbb{R}^n$.

Theorem 8.6.3. *Suppose that $M=\mathbb{R}^n$. Then* End *is locally constant-rank if and only if* $\mathsf{PC}(\boldsymbol{u}_0) = \mathsf{PC}(\tilde{\boldsymbol{u}}_0)$ *whenever $\boldsymbol{u}_0, \tilde{\boldsymbol{u}}_0$ are contained in the same connected component of* dom(End).

Proof. Suppose that $\mathsf{PC}(\boldsymbol{u}_0) = \mathsf{PC}(\tilde{\boldsymbol{u}}_0)$ whenever $\boldsymbol{u}_0, \tilde{\boldsymbol{u}}_0$ are contained in the same connected component of dom(End). By Lemma 8.1.2,

$$\mathrm{rank}(T\mathrm{End}(\boldsymbol{u}_0)) = \mathrm{rank}(T\mathrm{End}(\tilde{\boldsymbol{u}}_0))$$

for each such $\boldsymbol{u}_0, \tilde{\boldsymbol{u}}_0$. Consequently, End is locally constant-rank.

Conversely, suppose that End is locally constant-rank and $\boldsymbol{u}_0, \tilde{\boldsymbol{u}}_0$ are contained in the same connected component $\mathscr{K}$ of dom(End). Set

$$\rho = \dim(\mathsf{PC}(\boldsymbol{u})) = \operatorname{rank}(T\mathrm{End}(\boldsymbol{u})), \qquad \boldsymbol{u} \in \mathscr{K}.$$

We must show that $\mathsf{PC}(\boldsymbol{u}_0) = \mathsf{PC}(\tilde{\boldsymbol{u}}_0)$. To this end, let

$$\boldsymbol{\gamma} : [0,1] \to \mathscr{K}$$

be a C^∞ curve in $\mathscr{K}$ such that $\boldsymbol{\gamma}(0) = \boldsymbol{u}_0$ and $\boldsymbol{\gamma}(1) = \tilde{\boldsymbol{u}}_0$. Consider the map

$$[0,1] \ni t \mapsto S(t) = \mathsf{PC}(\boldsymbol{\gamma}(t))$$

We claim that S is a C^1 ρ-dimensional time-varying subspace of $T_{x_0}M$. To see this, let $t_0 \in [0,1]$ and let $b_1^0, \dots, b_\rho^0 \in S(t_0)$ be linearly independent. Then

$$b_i^0 = T\Phi_{a,b}^{f^{\boldsymbol{\gamma}(t_0)}}(\mathrm{End}(\boldsymbol{\gamma}(t_0))) \circ T\mathrm{End}(\boldsymbol{\gamma}(t_0)) \cdot \boldsymbol{u}_i$$

for some $\boldsymbol{u}_i \in \mathscr{U}$. Consider the maps $b_1, \dots, b_\rho : [0,1] \to T_{x_0}M$ defined by

$$\begin{aligned} b_i(t) &= T\Phi_{a,b}^{f^{\boldsymbol{\gamma}(t)}}(\mathrm{End}(\boldsymbol{\gamma}(t))) \circ T\mathrm{End}(\boldsymbol{\gamma}(t)) \cdot \boldsymbol{u}_i \\ &= (T_1\mathrm{End}^\Sigma(x_0, \boldsymbol{\gamma}(t)))^{-1} \circ T\mathrm{End}(\boldsymbol{\gamma}(t)) \cdot \boldsymbol{u}_i. \end{aligned}$$

It follows from Theorem 3.3.11 that each b_i is C^1 by composition. By continuity, there exists a relatively open subinterval $I \subseteq [0,1]$ such that $t_0 \in I$ and $b_1(t), \dots, b_\rho(t)$ are linearly independent for each $t \in I$. By construction, $b_1(t), \dots, b_\rho(t) \in S(t)$ for each $t \in I$. Thus S is a C^1 ρ-dimensional time-varying subspace of $T_{x_0}M$.

Suppose now that the tangent vectors $v_1, \dots, v_n \in T_{x_0}M$ are linearly independent. For each $1 \leq i \leq n$, define the map $V_i : [0,1] \to T_{x_0}M$ by

$$V_i(t) = \mathrm{proj}_S(t) \cdot v_i = \mathrm{proj}_{\mathsf{PC}(\boldsymbol{\gamma}(t))} \cdot v_i.$$

In what follows, the following easily-verified observations will be crucial:

- $\mathsf{PC}(\boldsymbol{\gamma}(t)) = \operatorname{span}\{V_1(t), \dots, V_n(t)\}$ for each $t \in [0,1]$;

- By the Leibniz rule and Lemma 8.6.2, each V_i is C^1;

- Each V_i can be written in the form

  $$V_i(t) = T\Phi^{f^{\boldsymbol{\gamma}(t)}}_{a,b}(\mathrm{End}(\boldsymbol{\gamma}(t))) \circ T\mathrm{End}(\boldsymbol{\gamma}(t)) \cdot \boldsymbol{U}_i(t),$$

  where $\boldsymbol{U}_i : [0,1] \to \mathscr{U}$ is C^1. Alternatively, by Theorem 6.1.10, each V_i can be written in the form[3]

  $$V_i(t) = \int_a^b \mathrm{Ad}^{f^{\boldsymbol{\gamma}(t)}} \left(\boldsymbol{D}_3 f^{\boldsymbol{U}_i(t)}_{\boldsymbol{\gamma}(t)}\right)(s, x_0)\, \mathrm{d}s,$$

  where $\boldsymbol{U}_i : [0,1] \to \mathscr{U}$ is C^1;

- By Lemma 3.3.9 and Proposition 5.4.2, the map

  $$(t,s) \mapsto \mathrm{Ad}^{f^{\boldsymbol{\gamma}(t)}} \left(\boldsymbol{D}_3 f^{\boldsymbol{U}_i(t)}_{\boldsymbol{\gamma}(t)}\right)(s, x_0)$$

  of $[0,1] \times J$ into $T_{x_0}M$ satisfies the hypotheses of [Dudley 2002, Section 4.3, Exercise 10], at least locally in its first independent variable, and thus

  $$\begin{aligned} \dot{V}_i(t) &= \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \int_a^b \mathrm{Ad}^{f^{\boldsymbol{\gamma}(\tau)}} \left(\boldsymbol{D}_3 f^{\boldsymbol{U}_i(\tau)}_{\boldsymbol{\gamma}(\tau)}\right)(s, x_0)\, \mathrm{d}s \\ &= \int_a^b \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \mathrm{Ad}^{f^{\boldsymbol{\gamma}(\tau)}} \left(\boldsymbol{D}_3 f^{\boldsymbol{U}_i(\tau)}_{\boldsymbol{\gamma}(\tau)}\right)(s, x_0)\, \mathrm{d}s \end{aligned}$$

  for each $t \in [0,1]$;

- By Lemma 3.3.9, the map $F : J \times M \times [0,1] \to TM$ defined by

  $$F(t,x,\tau) = f(t,x,\boldsymbol{\gamma}(\tau)(t))$$

  is a locally integrably $C^{3,2}$ time-varying vector field on M with scalar parameters.

- By Lemma 3.3.9, the map $G : J \times M \times [0,1] \to TM$ defined by

  $$G(t,x,\tau) = \boldsymbol{D}_3 f(t,x,\boldsymbol{\gamma}(\tau)(t)) \cdot \boldsymbol{U}_i(\tau)(t)$$

  is a locally integrably $C^{1,1}$ time-varying vector field on M with scalar parameters.

[3] Here, and in the rest of the proof, the subscript M_0 is suppressed on pullbacks.

By construction, we have

$$F^{\tau} = f^{\boldsymbol{\gamma}(\tau)} \qquad \text{and} \qquad G^{\tau} = \boldsymbol{D}_3 f_{\boldsymbol{\gamma}(\tau)}^{\boldsymbol{U}_i(\tau)}$$

for each $\tau \in [0,1]$. For concreteness, we write $f^{\boldsymbol{\gamma}(\tau)}$ and $\boldsymbol{D}_3 f_{\boldsymbol{\gamma}(\tau)}^{\boldsymbol{U}_i(\tau)}$ instead of F^{τ} and G^{τ}, respectively. Invoking Lemma 5.9.11, we obtain

$$\begin{aligned}\dot{V}_i(t) &= \int_a^b \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \mathrm{Ad}^{f^{\boldsymbol{\gamma}(\tau)}}\left(\boldsymbol{D}_3 f_{\boldsymbol{\gamma}(\tau)}^{\boldsymbol{U}_i(\tau)}\right)(s,x_0)\,\mathrm{d}s \\ &= \underbrace{\int_a^b \mathrm{Ad}^{f^{\boldsymbol{\gamma}(t)}}\left(\left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \boldsymbol{D}_3 f_{\boldsymbol{\gamma}(\tau)}^{\boldsymbol{U}_i(\tau)}\right)(s,x_0)\,\mathrm{d}s}_{(\dagger)} + \underbrace{\int_a^b \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \mathrm{Ad}^{f^{\boldsymbol{\gamma}(\tau)}}\left(\boldsymbol{D}_3 f_{\boldsymbol{\gamma}(t)}^{\boldsymbol{U}_i(t)}\right)(s,x_0)\,\mathrm{d}s}_{(\ddagger)}\end{aligned}$$

for each $t \in [0,1]$. To evaluate $(\dagger)$, observe that

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \boldsymbol{D}_3 f_{\boldsymbol{\gamma}(\tau)}^{\boldsymbol{U}_i(\tau)}(s,x) &= \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t \boldsymbol{D}_3 f(s,x,\boldsymbol{\gamma}(\tau)(s)) \cdot \boldsymbol{U}_i(\tau)(s) \\ &= \boldsymbol{D}_3^2 f(s,x,\boldsymbol{\gamma}(t)(s)) \cdot (\dot{\boldsymbol{\gamma}}(t)(s), \boldsymbol{U}_i(t)(s)) \\ &\qquad + \boldsymbol{D}_3 f(s,x,\boldsymbol{\gamma}(t)(s)) \cdot \dot{\boldsymbol{U}}_i(t)(s) \\ &= \boldsymbol{D}_3^2 f_{\boldsymbol{\gamma}(t),\dot{\boldsymbol{\gamma}}(t)}^{\boldsymbol{U}_i(t)}(s,x) + \boldsymbol{D}_3 f_{\boldsymbol{\gamma}(t)}^{\dot{\boldsymbol{U}}_i(t)}(s,x)\end{aligned}$$

for each $(s,x) \in J \times M$. Thus

$$(\dagger) = \int_a^b (\lambda^2_{\boldsymbol{\gamma}(t)} \cdot (\dot{\boldsymbol{\gamma}}(t), \boldsymbol{U}_i(t)))(s)\,\mathrm{d}s + \int_a^b (\lambda^1_{\boldsymbol{\gamma}(t)} \cdot \dot{\boldsymbol{U}}_i(t))(s)\,\mathrm{d}s.$$

To evaluate $(\ddagger)$, we invoke Lemma 5.9.11 again, concluding that

$$(\ddagger) = \int_a^b \int_a^s \left[\mathrm{Ad}^{f^{\boldsymbol{\gamma}(t)}}\left(\left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t f^{\boldsymbol{\gamma}(\tau)}\right)_\sigma, \mathrm{Ad}^{f^{\boldsymbol{\gamma}(t)}}\left(\boldsymbol{D}_3 f_{\boldsymbol{\gamma}(t)}^{\boldsymbol{U}_i(t)}\right)_s\right](x_0)\,\mathrm{d}\sigma\,\mathrm{d}s.$$

Observe that

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t f^{\boldsymbol{\gamma}(\tau)}(s,x) &= \left.\frac{\mathrm{d}}{\mathrm{d}\tau}\right|_t f(s,x,\boldsymbol{\gamma}(\tau)(s)) \\ &= \boldsymbol{D}_3 f(s,x,\boldsymbol{\gamma}(t)(s)) \cdot \dot{\boldsymbol{\gamma}}(t)(s) \\ &= \boldsymbol{D}_3 f_{\boldsymbol{\gamma}(t)}^{\dot{\boldsymbol{\gamma}}(t)}(s,x)\end{aligned}$$

for each $(s,x) \in J \times M$. Thus

$$(\ddagger) = \int_a^b (\Lambda^2_{\boldsymbol{\gamma}(t)} \cdot (\dot{\boldsymbol{\gamma}}(t), \boldsymbol{U}_i(t)))(s)\,\mathrm{d}s.$$

Altogether, we have shown that

$$\begin{aligned}\dot{V}_i(t) = &\int_a^b (\lambda^1_{\boldsymbol{\gamma}(t)} \cdot \dot{\boldsymbol{U}}_i(t))(s)\,\mathrm{d}s \\ &+ \int_a^b (\lambda^2_{\boldsymbol{\gamma}(t)} \cdot (\dot{\boldsymbol{\gamma}}(t), \boldsymbol{U}_i(t)))(s) + (\Lambda^2_{\boldsymbol{\gamma}(t)} \cdot (\dot{\boldsymbol{\gamma}}(t), \boldsymbol{U}_i(t)))(s)\,\mathrm{d}s\end{aligned}$$

for each $t \in [0,1]$. It follows from Lemmas 8.5.1 and 8.5.2 that

$$\dot{V}_i(t) \in \mathsf{PC}(\boldsymbol{\gamma}(t))$$

for each $t \in [0,1]$.[4] Since $\mathsf{PC}(\boldsymbol{\gamma}(t)) = \operatorname{span}\{V_1(t), \ldots, V_n(t)\}$ for each $t \in [0,1]$,

$$\dot{V}_i(t) = \sum_{i=1}^n A_i^j(t) V_j(t) \tag{8.13}$$

for continuous functions $A_i^j : [0,1] \to \mathbb{R}$.

To complete the proof, suppose that z_{x_0} lies in the annihilator

$$\mathsf{PC}(\boldsymbol{\gamma}(0))^0 = \mathsf{PC}(\boldsymbol{u}_0)^0.$$

For each $1 \le i \le n$, define $Z_i : [0,1] \to \mathbb{R}$ by

$$Z_i(t) = \langle z_{x_0}, V_i(t) \rangle.$$

Then each Z_i satisfies $Z_i(0) = \langle z_{x_0}, V_i(0) \rangle = 0$ and, using (8.13), we have

$$\begin{aligned}\dot{Z}_i(t) &= \langle z_{x_0}, \dot{V}_i(t) \rangle \\ &= \left\langle z_{x_0}, \sum_{j=1}^n A_i^j(t) V_j(t) \right\rangle\end{aligned}$$

[4] This relies on the following fact: Suppose that E is a finite-dimensional vector space and $\tilde{E}$ is a vector subspace of E. If $\xi : J \to E$ is integrable and satisfies $\xi(t) \in \tilde{E}$ for a.a. $t \in J$, then

$$\int_a^b \xi(s)\,\mathrm{d}s \in \tilde{E}.$$

$$= \sum_{j=1}^{n} A_i^j(t)\langle z_{x_0}, V_j(t)\rangle$$
$$= \sum_{j=1}^{n} A_i^j(t) Z_j(t).$$

Thus $\boldsymbol{Z} = (Z_1, \dots, Z_n)$ is the maximally-defined solution of the initial value problem

$$\begin{cases} \dot{\boldsymbol{\xi}}(t) = \boldsymbol{A}(t) \cdot \boldsymbol{\xi}(t), \quad \boldsymbol{\xi}(t) \in \mathbb{R}^n, \quad t \in [0,1] \\ \boldsymbol{\xi}(0) = \boldsymbol{0}_{\mathbb{R}^n}, \end{cases}$$

where $\boldsymbol{A}(t)$ is the matrix with $A_i^j(t)$ in the (i,j)th position. It follows that $\boldsymbol{Z}$ is identically equal to $\boldsymbol{0}_{\mathbb{R}^n}$ on $[0,1]$, or, equivalently, that $z_{x_0} \in \mathsf{PC}(\boldsymbol{\gamma}(t))^0$ for each $t \in [0,1]$. Since z_{x_0} was chosen arbitrarily, we have shown that

$$\mathsf{PC}(\boldsymbol{u}_0)^{\perp} \subseteq \mathsf{PC}(\boldsymbol{\gamma}(t))^{\perp}$$

for each $t \in [0,1]$. Equivalently [Leung 1973, Theorem 7.7], we have shown that

$$\mathsf{PC}(\boldsymbol{\gamma}(t)) \subseteq \mathsf{PC}(\boldsymbol{u}_0)$$

for each $t \in [0,1]$. In particular, setting $t = 1$, we have shown that

$$\mathsf{PC}(\tilde{\boldsymbol{u}}_0) = \mathsf{PC}(\boldsymbol{\gamma}(1)) \subseteq \mathsf{PC}(\boldsymbol{u}_0).$$

Since End is locally constant-rank, $\dim(\mathsf{PC}(\boldsymbol{u}_0)) = \dim(\mathsf{PC}(\tilde{\boldsymbol{u}}_0))$ and consequently

$$\mathsf{PC}(\boldsymbol{u}_0) = \mathsf{PC}(\tilde{\boldsymbol{u}}_0).$$

This completes the proof. ∎

Remark 8.6.4. Interestingly, the proof of [Vakhrameev 1991b, Theorem 1.1] does not rely on the assumption that $M = \mathbb{R}^n$. Since the details of the argument are not forthcoming, it is not clear to us how Vakhrameev's proof proceeds without this assumption. More precisely, it is not clear to us how to verify that

$$\ker(T\mathrm{End}(\boldsymbol{u}_0)) \times \ker(T\mathrm{End}(\boldsymbol{u}_0))$$

contains a sufficiently rich family of control variations. In the case where Σ is C_2^2-polynomial, and thus $\mathscr{U} = L^2(J, \mathbb{R}^r)$, the easily-verified fact that

$$\ker(T\mathrm{End}(\boldsymbol{u}_0)) \cong L^2(J, \mathbb{R}^r)$$

may be pertinent. However, this remains speculative.

Corollary 8.6.5. *Suppose that* $M = \mathbb{R}^n$ *and* $\mathrm{dom}(\mathrm{End})$ *is connected. Then* End *is constant-rank if and only if* $\mathsf{PC}(\boldsymbol{u}_0) = \mathsf{PC}(\tilde{\boldsymbol{u}}_0)$ *for each* $\boldsymbol{u}_0, \tilde{\boldsymbol{u}}_0 \in \mathrm{dom}(\mathrm{End})$. *In particular,* End *is a submersion if and only if*

- $\mathsf{PC}(\boldsymbol{u}_0) = \mathsf{PC}(\tilde{\boldsymbol{u}}_0)$ *for each* $\boldsymbol{u}_0, \tilde{\boldsymbol{u}}_0 \in \mathrm{dom}(\mathrm{End})$ *and*
- *There exists* $\boldsymbol{u}_0 \in \mathrm{dom}(\mathrm{End})$ *such that* $\dim(\mathsf{PC}(\boldsymbol{u}_0)) = n$.

For example, dom(End) is connected whenever Σ is complete, since in this case

$$\mathrm{dom}(\mathrm{End}) = \mathscr{U}$$

by definition.

8.7 Subimmersivity

Finally, we briefly connect the results of Section 8.6 to subimmersivity.

Definition 8.7.1. Suppose that Q and R are Banach manifolds and

$$F : Q \to R$$

is C^k for $k \in \mathbb{N}^*$. We say that F is a ***subimmersion at*** $q \in Q$ if there exist a neighbourhood U of q in Q, a Banach manifold P, a C^k submersion $s : U \to P$, and a C^k immersion $i : P \to R$ such that $F|U = i \circ s$. We say that F is a ***subimmersion*** if it is a subimmersion at each $q \in Q$.

Corollary 8.7.2. *Suppose that* $M = \mathbb{R}^n$*. Then* End *is a subimmersion if and only if* $\mathsf{PC}(\boldsymbol{u}_0) = \mathsf{PC}(\tilde{\boldsymbol{u}}_0)$ *whenever* $\boldsymbol{u}_0, \tilde{\boldsymbol{u}}_0$ *are contained in the same connected component of* dom(End)*.*

Proof. Since M is finite-dimensional, End is a subimmersion if and only if it is locally constant-rank; see [Abraham et al. 1988, Proposition 3.5.16]. ∎

Chapter 9

Sublinear growth

Consider a control system

$$\Sigma = (f, \mathscr{U})$$

evolving on an n-dimensional manifold M, and let $x_0 \in M$. In Chapter 4, we saw that the continuation method attempts to solve the x_0-anchored motion planning problem for Σ by lifting curves in M to curves in $\mathscr{U}$; the lifted curves are maximally-defined solutions of path-lifting equations (PLEs). For this procedure to be successful, it is essential that each such solution is defined on $[0, 1]$. In general, this is not guaranteed. Indeed, each PLE is an initial value problem, which opens up the possibility that its maximally-defined solution is defined on a proper subinterval of $[0, 1]$. In Chapter 1, we called this phenomenon the third obstruction to the continuation method. In this chapter, we derive conditions which ensure that this obstruction is overcome. For tractability, the analysis is carried out under the assumption that Σ is a control-affine system that uses L^2 controls.

This chapter is organized in the following way. In Section 9.1, we introduce the cotangent lift of Σ and record several of its basic properties. In Section 9.2, we establish terminology surrounding Lie derivatives, momentum functions, and switching functions. In Section 9.3, we establish a certain sublinear growth condition, then explain its relevance to the study of PLEs. An alternative formulation of this condition, phrased in terms of switching functions, is the subject of Section 9.4. In Section 9.5, we prove a theorem on sublinear growth. This theorem, which extends a result of Chitour [1996], constitutes the main result in this chapter.

Our standing assumptions in this chapter are that

- M is a second-countable n-dimensional Riemannian manifold,
- $\Sigma = (f, \mathscr{U})$ is a C^2 control-affine system evolving on M, where

$$f(t, x, \boldsymbol{\omega}) = f_0(x) + \sum_{i=1}^{r} \omega^i f_i(x),$$

- The time domain of Σ is $J = [a, b]$,
- Σ uses L^2 controls, so that $\mathscr{U} = L^2(J, \mathbb{R}^r)$,
- Σ is complete, and
- Σ is completely controllable from a fixed initial state x_0 on J.

By Corollary 3.3.12, the map

$$\mathrm{End}_{x_0}^{\Sigma} : \mathscr{U} \to M$$

is C^2. We define

$$M_\Delta = M \smallsetminus \overline{\mathrm{End}_{x_0}^{\Sigma}(\mathscr{U}_{x_0}^{\mathrm{sing}})},$$

where the overline denotes closure in M. The reasons for including the closure operation in the definition of M_Δ will be made clear below. Finally, we define

$$x_{\boldsymbol{u}} = \mathrm{End}_{x_0}^{\Sigma}(\boldsymbol{u})$$

for each $\boldsymbol{u} \in \mathscr{U}$.

9.1 Cotangent lifts of control systems

In this brief section, we introduce the cotangent lift of Σ and record several of its basic properties. For further information, we refer to [Sussmann 1998].

Definition 9.1.1. The ***cotangent lift***[1] of Σ is the C^1 control-affine system

$$\Sigma^* = (f^*, \mathscr{U})$$

evolving on T^*M, where $f^* \in \mathscr{V}(J, T^*M, \mathbb{R}^r)$ is defined by

$$f^*(t, p_x, \boldsymbol{\omega}) = f_0^*(p_x) + \sum_{i=1}^{r} \omega^i f_i^*(p_x).$$

Recall from Section 5.3 that f_i^* is the cotangent lift of f_i.

Note that Σ^* is a C^1 control-affine system by Lemma 5.3.5. From Theorem 5.3.4 we obtain the next lemma, which describes the controlled trajectories of Σ^*.

Lemma 9.1.2. *For each $\boldsymbol{u} \in \mathscr{U}$ and each $p_{x_{\boldsymbol{u}}} \in T^*_{x_{\boldsymbol{u}}}M$, we have*

1. $J^{\Sigma^*}(b, p_{x_{\boldsymbol{u}}}, \boldsymbol{u}) = J$,
2. $\mu^{\Sigma^*}(t, b, p_{x_{\boldsymbol{u}}}, \boldsymbol{u}) = T\Phi^{f\boldsymbol{u}}_{a,t}(\Phi^{f\boldsymbol{u}}_{t,a}(x_0))^* \circ T\Phi^{f\boldsymbol{u}}_{b,a}(x_0)^* \cdot p_{x_{\boldsymbol{u}}}$ *for each* $t \in J$, *and*
3. $\pi_{T^*M} \circ \mu^{\Sigma^*}(t, b, p_{x_{\boldsymbol{u}}}, \boldsymbol{u}) = \mu^{\Sigma}(t, a, x_0, \boldsymbol{u})$ *for each* $t \in J$.

This lemma gives an explicit form for the $\boldsymbol{u}$-controlled trajectory of Σ^* with initial condition $(b, p_{x_{\boldsymbol{u}}})$. Since b is the right endpoint of the interval J, it is sometimes referred to as a terminal condition; see, for example, [Sussmann 1993].

[1] In the literature, the cotangent lift of Σ is also known as the Hamiltonian lift of Σ. In particular, this is true in some of the literature concerning the continuation method; see [Chitour and Sussmann 1998, Section 3] and [Chitour 2006, Section 2.1].

9.2 Lie derivatives, momentum functions, and switching functions

In this section, we establish terminology surrounding Lie derivatives, momentum functions, and switching functions, following Sussmann [1993].

Definition 9.2.1. Suppose that $G : T^*M \to \mathbb{R}$ is C^1 and $0 \leq i \leq r$. The f_i^****-Lie derivative*** of G is the continuous function $\mathscr{L}_{f_i^*}G : T^*M \to \mathbb{R}$ defined by

$$\mathscr{L}_{f_i^*}G(p_x) = TG(p_x) \cdot f_i^*(p_x).$$

Now let $\boldsymbol{f}^*$ denote the r-tuple $(f_1^*, \ldots, f_r^*)$ of C^1 vector fields on T^*M. The $\boldsymbol{f}^*$***-Lie derivative*** of G is the continuous map $\mathscr{L}_{\boldsymbol{f}^*}G : T^*M \to \mathbb{R}^r$ defined by

$$\mathscr{L}_{\boldsymbol{f}^*}G(p_x) = (\mathscr{L}_{f_1^*}G(p_x), \ldots, \mathscr{L}_{f_r^*}G(p_x)).$$

The next lemma describes how these Lie derivatives arise in this chapter.

Lemma 9.2.2. *Suppose that $G : T^*M \to \mathbb{R}$ is C^1 and $\mu^* : \mathrm{dom}(\mu^*) \to T^*M$ is a $\boldsymbol{u}$-controlled trajectory of Σ^*. Then the function $G \circ \mu^*$ is LAC and*

$$\dot{\widehat{G \circ \mu^*}}(t) = \mathscr{L}_{f_0^*}G(\mu^*(t)) + \langle \boldsymbol{u}(t), \mathscr{L}_{\boldsymbol{f}^*}G(\mu^*(t))\rangle_{\mathbb{R}^r} \tag{9.1}$$

for a.a. $t \in \mathrm{dom}(\mu^*)$.

Proof. By Lemma 2.3.2, $G \circ \mu^*$ is LAC. By the chain rule, we have

$$\begin{aligned}
\dot{\widehat{G \circ \mu^*}}(t) &= TG(\mu^*(t)) \cdot \dot{\mu}^*(t) \\
&= TG(\mu^*(t)) \cdot \left(f_0^*(\mu^*(t)) + \sum_{i=1}^{r} u^i(t) f_i^*(\mu^*(t)) \right) \\
&= TG(\mu^*(t)) \cdot f_0^*(\mu^*(t)) + \sum_{i=1}^{r} u^i(t) TG(\mu^*(t)) \cdot f_i^*(\mu^*(t)) \\
&= \mathscr{L}_{f_0^*}G(\mu^*(t)) + \sum_{i=1}^{r} u^i(t) \mathscr{L}_{f_i^*}G(\mu^*(t)) \\
&= \mathscr{L}_{f_0^*}G(\mu^*(t)) + \langle \boldsymbol{u}(t), \mathscr{L}_{\boldsymbol{f}^*}G(\mu^*(t))\rangle_{\mathbb{R}^r}
\end{aligned}$$

for a.a. $t \in \mathrm{dom}(\mu^*)$. This completes the proof. ∎

We now define the momentum functions and switching functions of Σ. For insight into the origin of these terms, we refer to [Loomis and Sternberg 1990, Chapter 13] and [Vakhrameev and Topunov 2002], respectively.

Definition 9.2.3. Suppose that $0 \le i \le r$. The i***th momentum function of*** Σ is the C^2 function $\varphi_i^\Sigma : T^*M \to \mathbb{R}$ defined by

$$\varphi_i^\Sigma(p_x) = \langle p_x, f_i(x) \rangle.$$

The ***momentum vector*** of Σ is the C^2 map $\boldsymbol{\varphi}^\Sigma : T^*M \to \mathbb{R}^r$ defined by

$$\boldsymbol{\varphi}^\Sigma(p_x) = \left(\varphi_1^\Sigma(p_x), \ldots, \varphi_r^\Sigma(p_x)\right).$$

Now suppose that $0 \le i, j \le r$. The (i,j)***th first derived momentum function of*** Σ is the C^1 function $\psi_{i,j}^\Sigma : T^*M \to \mathbb{R}$ defined by

$$\psi_{i,j}^\Sigma(p_x) = \langle p_x, [f_j, f_i](x) \rangle.$$

Note that the order of the indices i and j is reversed on the right-hand side.

Lemma 9.2.4. *Suppose that* $0 \le i, j \le r$. *Then* $\mathscr{L}_{f_i^*}\varphi_j^\Sigma = \psi_{i,j}^\Sigma$.

Proof. See [Chitour 1996, Proposition 3]. ∎

Definition 9.2.5. Suppose that $0 \le i \le r$ and $\mu^* : \mathrm{dom}(\mu^*) \to T^*M$ is a controlled trajectory of Σ^*. We say that the function $\varphi_i^\Sigma \circ \mu^*$ is a ***switching function*** of Σ.

9.3 Sublinear growth

In this section, we establish a certain sublinear growth condition, then explain its relevance to the study of PLEs. Recall from Chapter 4 that

$$\underline{\mathrm{End}}_{x_0}^\Sigma : \mathscr{U}_{x_0}^{\mathrm{reg}} \to M$$

is the desingularized x_0-anchored endpoint map of Σ, and that

$$T\underline{\mathrm{End}}^{\Sigma}_{x_0}(\cdot)^{\#}$$

is the Moore–Penrose pseudoinverse of $T\underline{\mathrm{End}}^{\Sigma}_{x_0}$. Since $\underline{\mathrm{End}}^{\Sigma}_{x_0}$ is a C^2 submersion, it follows from Proposition 4.1.6 that $T\underline{\mathrm{End}}^{\Sigma}_{x_0}(\cdot)^{\#}$ is C^1.

In the remainder of this section, K is a nonempty compact subset of M_{Δ}.

Definition 9.3.1. If there exists $C \in \mathbb{R}_{>0}$ such that

$$\|T\underline{\mathrm{End}}^{\Sigma}_{x_0}(\boldsymbol{u})^{\#}\| \leq C(1 + \|\boldsymbol{u}\|)$$

for each $\boldsymbol{u} \in (\underline{\mathrm{End}}^{\Sigma}_{x_0})^{-1}(K)$, then we say that the Moore–Penrose pseudoinverse of $T\underline{\mathrm{End}}^{\Sigma}_{x_0}$ has ***sublinear growth over*** K.

Note that in the above definition, $\|T\underline{\mathrm{End}}^{\Sigma}_{x_0}(\boldsymbol{u})^{\#}\|$ denotes the operator norm of

$$T\underline{\mathrm{End}}^{\Sigma}_{x_0}(\boldsymbol{u})^{\#} \in \mathrm{Hom}(T_{x_{\boldsymbol{u}}}M, \mathscr{U}),$$

while $\|\boldsymbol{u}\|$ denotes the norm of $\boldsymbol{u}$ as an element of $\mathscr{U} = L^2(J, \mathbb{R}^r)$.

We now explain the relevance of the sublinear growth condition.

Lemma 9.3.2. *Suppose that the Moore–Penrose pseudoinverse of $T\underline{\mathrm{End}}^{\Sigma}_{x_0}$ has sublinear growth over K and $\pi : [0,1] \to M$ is a C^1 curve such that*

$$\mathrm{image}(\pi) \subseteq K \cap \mathrm{image}(\underline{\mathrm{End}}^{\Sigma}_{x_0}).$$

Then for each choice of control $\boldsymbol{u}_0 \in (\underline{\mathrm{End}}^{\Sigma}_{x_0})^{-1}(\pi(0))$, the maximally-defined solution of the $(\pi, \boldsymbol{u}_0)$-PLE for $\underline{\mathrm{End}}^{\Sigma}_{x_0}$ is defined on $[0,1]$.

Proof. Choose a control $\boldsymbol{u}_0 \in (\underline{\mathrm{End}}^{\Sigma}_{x_0})^{-1}(\pi(0))$, and let $\boldsymbol{\Pi}$ be the maximally-defined solution of the $(\pi, \boldsymbol{u}_0)$-PLE for $\underline{\mathrm{End}}^{\Sigma}_{x_0}$. To reach a contradiction, let us assume that $\boldsymbol{\Pi}$ is defined on $[0, \delta)$ for $\delta \in (0, 1]$. By Proposition 4.2.7, either

1. $\lim_{t \nearrow \delta} \|\dot{\boldsymbol{\Pi}}(t)\| = \infty$, or

2. $\boldsymbol{u}_+ = \lim_{t \nearrow \delta} \boldsymbol{\Pi}(t)$ exists in $\mathscr{U}$ and $\boldsymbol{u}_+ \in \mathscr{U}_{x_0}^{\mathrm{sing}}$.

Since H_π is continuous, there exists $B \in \mathbb{R}_{\geq 0}$ such that

$$\|H_\pi(t, \pi(t))\| \leq B$$

for each $t \in [0, 1]$. Since $\mathrm{image}(\pi) \subseteq K$ and $\underline{\mathrm{End}}_{x_0}^{\Sigma} \circ \boldsymbol{\Pi} = \pi|[0, \delta)$, it follows that

$$\mathrm{image}(\boldsymbol{\Pi}) \subseteq (\underline{\mathrm{End}}_{x_0}^{\Sigma})^{-1}(K).$$

With $C \in \mathbb{R}_{>0}$ prescribed as in Definition 9.3.1, we have

$$\begin{aligned}
\|\boldsymbol{\Pi}(t)\| &= \left\| \boldsymbol{u}_0 + \int_0^t \dot{\boldsymbol{\Pi}}(\sigma)\, \mathrm{d}\sigma \right\| \\
&\leq \|\boldsymbol{u}_0\| + \int_0^t \|\dot{\boldsymbol{\Pi}}(\sigma)\|\, \mathrm{d}\sigma \\
&= \|\boldsymbol{u}_0\| + \int_0^t \|T\underline{\mathrm{End}}_{x_0}^{\Sigma}(\boldsymbol{\Pi}(\sigma))^{\#} \cdot H_\pi(\sigma, \pi(\sigma))\|\, \mathrm{d}\sigma \\
&\leq \|\boldsymbol{u}_0\| + \int_0^t \|T\underline{\mathrm{End}}_{x_0}^{\Sigma}(\boldsymbol{\Pi}(\sigma))^{\#}\|\, \|H_\pi(\sigma, \pi(\sigma))\|\, \mathrm{d}\sigma \\
&\leq \|\boldsymbol{u}_0\| + \int_0^t BC(1 + \|\boldsymbol{\Pi}(\sigma)\|)\, \mathrm{d}\sigma \\
&\leq \|\boldsymbol{u}_0\| + BCt + \int_0^t BC\|\boldsymbol{\Pi}(\sigma)\|\, \mathrm{d}\sigma \\
&\leq \|\boldsymbol{u}_0\| + BC + \int_0^t BC\|\boldsymbol{\Pi}(\sigma)\|\, \mathrm{d}\sigma
\end{aligned}$$

for each $t \in [0, \delta)$. By the Bellman–Gronwall inequality [Sontag 1998], we have

$$\|\boldsymbol{\Pi}(t)\| \leq C_0 = (\|\boldsymbol{u}_0\| + BC)\mathrm{e}^{BC}$$

for each $t \in [0, \delta)$. But this implies that

$$\begin{aligned}
\|\dot{\boldsymbol{\Pi}}(t)\| &= \|T\underline{\mathrm{End}}_{x_0}^{\Sigma}(\boldsymbol{\Pi}(t))^{\#} \cdot H_\pi(t, \pi(t))\| \\
&\leq \|T\underline{\mathrm{End}}_{x_0}^{\Sigma}(\boldsymbol{\Pi}(t))^{\#}\|\, \|H_\pi(t, \pi(t))\| \\
&\leq BC(1 + \|\boldsymbol{\Pi}(t)\|) \\
&\leq BC(1 + C_0)
\end{aligned}$$

for each $t \in [0, \delta)$. Hence

$$\lim_{t \nearrow \delta} \|\dot{\mathbf{\Pi}}(t)\| < \infty.$$

Thus $\boldsymbol{u}_+$ exists in $\mathscr{U}$ and $\boldsymbol{u}_+ \in \mathscr{U}_{x_0}^{\text{sing}}$. By continuity,

$$\lim_{t \nearrow \delta} \underline{\mathrm{End}}_{x_0}^{\Sigma} \circ \mathbf{\Pi}(t) = \lim_{t \nearrow \delta} \mathrm{End}_{x_0}^{\Sigma} \circ \mathbf{\Pi}(t) = \mathrm{End}_{x_0}^{\Sigma}(\boldsymbol{u}_+) \in \mathrm{End}_{x_0}^{\Sigma}(\mathscr{U}_{x_0}^{\text{sing}}).$$

However,

$$\lim_{t \nearrow \delta} \underline{\mathrm{End}}_{x_0}^{\Sigma} \circ \mathbf{\Pi}(t) = \lim_{t \nearrow \delta} \pi(t) = \pi(\delta) \in K \subseteq M_\Delta.$$

We have shown that $\lim_{t \nearrow \delta} \underline{\mathrm{End}}_{x_0}^{\Sigma} \circ \mathbf{\Pi}(t)$ is contained in

$$\mathrm{End}_{x_0}^{\Sigma}(\mathscr{U}_{x_0}^{\text{sing}}) \qquad \text{and} \qquad M_\Delta = M \setminus \overline{\mathrm{End}_{x_0}^{\Sigma}(\mathscr{U}_{x_0}^{\text{sing}})},$$

which is a contradiction. This completes the proof. ∎

Although the preceding lemma gives a way of checking if the maximally-defined solution of the $(\pi, \boldsymbol{u}_0)$-PLE is defined on $[0, 1]$, its hypotheses are difficult to verify. Next, we begin to reduce the hypotheses of the lemma to a computable form.

9.4 An alternative characterization of sublinear growth

In this section, we give an alternative characterization of the sublinear growth condition, phrased in terms of switching functions. The alternative characterization is more tangible from a control-theoretic point of view, since it is phrased in terms of the controlled trajectories of the cotangent lift of Σ.

We begin by establishing notation for unit covectors. We define

$$S_x^* M = \{p_x \in T_x^* M \;:\; \|p_x\| = 1\}$$

for each $x \in M$. Recalling that $x_{\boldsymbol{u}} = \mathrm{End}_{x_0}^{\Sigma}(\boldsymbol{u})$, we have the following lemma.

Lemma 9.4.1. *Suppose that $\boldsymbol{u} \in \mathscr{U}^{\mathrm{reg}}_{x_0}$. Then*

$$\|T\underline{\mathrm{End}}^{\Sigma}_{x_0}(\boldsymbol{u})^{\#}\| = \left(\inf_{p_{x_{\boldsymbol{u}}} \in S^*_{x_{\boldsymbol{u}}} M} \|T\underline{\mathrm{End}}^{\Sigma}_{x_0}(\boldsymbol{u})^* \cdot p_{x_{\boldsymbol{u}}}\|^2 \right)^{-\frac{1}{2}}. \tag{9.2}$$

Proof. See the discussion preceding [Sussmann 1993, Equation 10]. ∎

Remark 9.4.2. Since M is finite-dimensional, $S^*_{x_{\boldsymbol{u}}} M$ is compact. Consequently, there exists an element $p^{\min}_{x_{\boldsymbol{u}}} \in S^*_{x_{\boldsymbol{u}}} M$ which achieves the infimum in (9.2).

Equation (9.2) suggests that we should examine the adjoints $T\underline{\mathrm{End}}^{\Sigma}_{x_0}(\boldsymbol{u})^*$ to determine if the Moore–Penrose pseudoinverse of $T\underline{\mathrm{End}}^{\Sigma}_{x_0}$ has sublinear growth over a nonempty compact subset of M_Δ. To this end, we will compute the adjoints

$$T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u})^*$$

for each $\boldsymbol{u} \in \mathscr{U}$. Doing so computes the adjoints $T\underline{\mathrm{End}}^{\Sigma}_{x_0}(\boldsymbol{u})^*$ as well. Indeed, since $\underline{\mathrm{End}}^{\Sigma}_{x_0}$ is obtained from $\mathrm{End}^{\Sigma}_{x_0}$ by restriction to $\mathscr{U}^{\mathrm{reg}}_{x_0}$, it is clear that

$$T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u})^* = T\underline{\mathrm{End}}^{\Sigma}_{x_0}(\boldsymbol{u})^*$$

whenever $\boldsymbol{u} \in \mathscr{U}^{\mathrm{reg}}_{x_0}$.

For each $1 \leq i \leq r$ and each $\boldsymbol{u} \in \mathscr{U}$, define the map

$$\Omega^i_{x_0}(\boldsymbol{u}) : T^*_{x_{\boldsymbol{u}}} M \to \mathscr{U}$$

by setting

$$\begin{aligned}(\Omega^i_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}})(t) &= \varphi^{\Sigma}_i(\mu^{\Sigma^*}(t, b, p_{x_{\boldsymbol{u}}}, \boldsymbol{u})) \\ &= \left\langle \mu^{\Sigma^*}(t, b, p_{x_{\boldsymbol{u}}}, \boldsymbol{u}), f_i(\mu^{\Sigma}(t, a, x_0, \boldsymbol{u})) \right\rangle \end{aligned} \tag{9.3}$$

for each $t \in J$. By Lemma 9.1.2, (9.3) is well-defined. Furthermore, each

$$\Omega^i_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}}$$

is LAC by Lemma 2.3.2, and thus is an element of $\mathscr{U}$. The next lemma tells us that each $\Omega^i_{x_0}(\boldsymbol{u})$ is continuous and linear, as well.

Lemma 9.4.3. *Suppose that $1 \leq i \leq r$ and $\boldsymbol{u} \in \mathscr{U}$. Then*

$$\Omega^i_{x_0}(\boldsymbol{u}) \in \operatorname{Hom}(T^*_{x_{\boldsymbol{u}}} M, \mathscr{U}).$$

Proof. By Lemma 9.1.2, we have

$$\begin{aligned}
&(\Omega^i_{x_0}(\boldsymbol{u}) \cdot (C p_{x_{\boldsymbol{u}}} + \tilde{C} \tilde{p}_{x_{\boldsymbol{u}}}))(t) \\
&= \left\langle \mu^{\Sigma^*}(t, b, C p_{x_{\boldsymbol{u}}} + \tilde{C} \tilde{p}_{x_{\boldsymbol{u}}}, \boldsymbol{u}), f_i(\mu^{\Sigma}(t, a, x_0, \boldsymbol{u})) \right\rangle \\
&= \left\langle T\Phi^{f^{\boldsymbol{u}}}_{a,t}(\Phi^{f^{\boldsymbol{u}}}_{t,a}(x_0))^* \circ T\Phi^{f^{\boldsymbol{u}}}_{b,a}(x_0)^* \cdot (C p_{x_{\boldsymbol{u}}} + \tilde{C} \tilde{p}_{x_{\boldsymbol{u}}}), f_i(\mu^{\Sigma}(t, a, x_0, \boldsymbol{u})) \right\rangle \\
&= C \left\langle T\Phi^{f^{\boldsymbol{u}}}_{a,t}(\Phi^{f^{\boldsymbol{u}}}_{t,a}(x_0))^* \circ T\Phi^{f^{\boldsymbol{u}}}_{b,a}(x_0)^* \cdot p_{x_{\boldsymbol{u}}}, f_i(\mu^{\Sigma}(t, a, x_0, \boldsymbol{u})) \right\rangle \\
&\qquad + \tilde{C} \left\langle T\Phi^{f^{\boldsymbol{u}}}_{a,t}(\Phi^{f^{\boldsymbol{u}}}_{t,a}(x_0))^* \circ T\Phi^{f^{\boldsymbol{u}}}_{b,a}(x_0)^* \cdot \tilde{p}_{x_{\boldsymbol{u}}}, f_i(\mu^{\Sigma}(t, a, x_0, \boldsymbol{u})) \right\rangle \\
&= C(\Omega^i_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}})(t) + \tilde{C}(\Omega^i_{x_0}(\boldsymbol{u}) \cdot \tilde{p}_{x_{\boldsymbol{u}}})(t)
\end{aligned}$$

for each $t \in J$. Since $T^*_{x_{\boldsymbol{u}}} M$ is finite-dimensional, $\Omega^i_{x_0}(\boldsymbol{u})$ is automatically continuous; see [Bachman and Narici 2000, Theorem 14.7]. This completes the proof. ∎

For each $\boldsymbol{u} \in \mathscr{U}$, define the map $\boldsymbol{\Omega}_{x_0}(\boldsymbol{u}) : T^*_{x_{\boldsymbol{u}}} M \to \mathscr{U}$ by setting

$$\begin{aligned}
(\boldsymbol{\Omega}_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}})(t) &= \boldsymbol{\varphi}^{\Sigma}(\mu^{\Sigma^*}(t, b, p_{x_{\boldsymbol{u}}}, \boldsymbol{u})) \\
&= \big((\boldsymbol{\Omega}^1_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}})(t), \ldots, (\boldsymbol{\Omega}^r_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}})(t)\big)
\end{aligned}$$

for each $t \in J$. The next lemma is an obvious consequence of Lemma 9.4.3.

Lemma 9.4.4. *Suppose that $\boldsymbol{u} \in \mathscr{U}$. Then $\boldsymbol{\Omega}_{x_0}(\boldsymbol{u}) \in \operatorname{Hom}(T^*_{x_{\boldsymbol{u}}} M, \mathscr{U})$.*

Somewhat less obvious is that $\boldsymbol{\Omega}_{x_0}(\boldsymbol{u})$ is the adjoint of $T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u})$.

Lemma 9.4.5. *Suppose that $\boldsymbol{u} \in \mathscr{U}$. Then $\boldsymbol{\Omega}_{x_0}(\boldsymbol{u}) = T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u})^*$.*

Proof. Choose $\boldsymbol{u} \in \mathscr{U}$. We must show that

$$\langle \boldsymbol{\Omega}_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}}, \tilde{\boldsymbol{u}} \rangle = \left\langle p_{x_{\boldsymbol{u}}}, T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}) \cdot \tilde{\boldsymbol{u}} \right\rangle$$

for each $p_{x_{\boldsymbol{u}}} \in T^*_{x_{\boldsymbol{u}}}M$ and each $\tilde{\boldsymbol{u}} \in \mathscr{U}$. Recall from Example 6.2.3 that

$$T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}) \cdot \tilde{\boldsymbol{u}} = \int_a^b \sum_{i=1}^r \tilde{u}^i(\sigma) T\Phi^{f\boldsymbol{u}}_{b,a}(x_0) \circ T\Phi^{f\boldsymbol{u}}_{a,\sigma}(\Phi^{f\boldsymbol{u}}_{\sigma,a}(x_0)) \cdot f_i(\Phi^{f\boldsymbol{u}}_{\sigma,a}(x_0)) \, \mathrm{d}\sigma.$$

In what follows, we write

$$\mu_\sigma = \Phi^{f\boldsymbol{u}}_{\sigma,a}(x_0) = \mu^{\Sigma}(\sigma, a, x_0, \boldsymbol{u}).$$

Invoking Proposition 5.4.4 and Lemma 9.1.2, we have

$$\begin{aligned}
\langle p_{x_{\boldsymbol{u}}}, T\mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u}) \cdot \tilde{\boldsymbol{u}} \rangle &= \left\langle p_{x_{\boldsymbol{u}}}, \int_a^b \sum_{i=1}^r \tilde{u}^i(\sigma) T\Phi^{f\boldsymbol{u}}_{b,a}(x_0) \circ T\Phi^{f\boldsymbol{u}}_{a,\sigma}(\mu_\sigma) \cdot f_i(\mu_\sigma) \, \mathrm{d}\sigma \right\rangle \\
&= \int_a^b \sum_{i=1}^r \left\langle p_{x_{\boldsymbol{u}}}, \tilde{u}^i(\sigma) T\Phi^{f\boldsymbol{u}}_{b,a}(x_0) \circ T\Phi^{f\boldsymbol{u}}_{a,\sigma}(\mu_\sigma) \cdot f_i(\mu_\sigma) \right\rangle \mathrm{d}\sigma \\
&= \int_a^b \sum_{i=1}^r \tilde{u}^i(\sigma) \left\langle p_{x_{\boldsymbol{u}}}, T\Phi^{f\boldsymbol{u}}_{b,a}(x_0) \circ T\Phi^{f\boldsymbol{u}}_{a,\sigma}(\mu_\sigma) \cdot f_i(\mu_\sigma) \right\rangle \mathrm{d}\sigma \\
&= \int_a^b \sum_{i=1}^r \tilde{u}^i(\sigma) \left\langle T\Phi^{f\boldsymbol{u}}_{a,\sigma}(\mu_\sigma)^* \circ T\Phi^{f\boldsymbol{u}}_{b,a}(x_0)^* \cdot p_{x_{\boldsymbol{u}}}, f_i(\mu_\sigma) \right\rangle \mathrm{d}\sigma \\
&= \int_a^b \sum_{i=1}^r \tilde{u}^i(\sigma) \left\langle \mu^{\Sigma^*}(\sigma, b, p_{x_{\boldsymbol{u}}}, \boldsymbol{u}), f_i(\mu_\sigma) \right\rangle \mathrm{d}\sigma \\
&= \int_a^b \sum_{i=1}^r \tilde{u}^i(\sigma) (\Omega^i_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}})(\sigma) \, \mathrm{d}\sigma \\
&= \langle \boldsymbol{\Omega}_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}}, \tilde{\boldsymbol{u}} \rangle
\end{aligned}$$

for each $p_{x_{\boldsymbol{u}}} \in T^*_{x_{\boldsymbol{u}}}M$ and each $\tilde{\boldsymbol{u}} \in \mathscr{U}$. In the last line, we have used the fact that

$$\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle = \int_a^b \sum_{i=1}^r u^i_1(\sigma) u^i_2(\sigma) \, \mathrm{d}\sigma$$

by definition. This completes the proof. ∎

In the next lemma, we derive the alternative characterization.

Lemma 9.4.6. *Suppose that K is a nonempty compact subset of M_Δ. Then the Moore–Penrose pseudoinverse of $T\underline{\mathrm{End}}^{\Sigma}_{x_0}$ has sublinear growth over K if and only if there exists $C \in \mathbb{R}_{>0}$ such that*

$$C \le \|\boldsymbol{\Omega}_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}}\|(1 + \|\boldsymbol{u}\|)$$

for each choice of control $\boldsymbol{u} \in (\underline{\mathrm{End}}^{\Sigma}_{x_0})^{-1}(K)$ and each unit covector $p_{x_{\boldsymbol{u}}} \in S^*_{x_{\boldsymbol{u}}}M$.

Proof. By Lemmas 9.4.1 and 9.4.5, the Moore–Penrose pseudoinverse of $T\underline{\mathrm{End}}^{\Sigma}_{x_0}$ has sublinear growth over K if and only if there exists $C_0 \in \mathbb{R}_{>0}$ such that

$$\begin{aligned}
&\left(\inf_{p_{x_{\boldsymbol{u}}} \in S^*_{x_{\boldsymbol{u}}} M} \|T\underline{\mathrm{End}}^{\Sigma}_{x_0}(\boldsymbol{u})^* \cdot p_{x_{\boldsymbol{u}}}\|^2\right)^{-\frac{1}{2}} \leq C_0(1 + \|\boldsymbol{u}\|) \\
\iff\ & \frac{1}{C_0^2} \leq \inf_{p_{x_{\boldsymbol{u}}} \in S^*_{x_{\boldsymbol{u}}} M} \|T\underline{\mathrm{End}}^{\Sigma}_{x_0}(\boldsymbol{u})^* \cdot p_{x_{\boldsymbol{u}}}\|^2 (1 + \|\boldsymbol{u}\|)^2 \\
\iff\ & \frac{1}{C_0^2} \leq \inf_{p_{x_{\boldsymbol{u}}} \in S^*_{x_{\boldsymbol{u}}} M} \|\boldsymbol{\Omega}_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}}\|^2 (1 + \|\boldsymbol{u}\|)^2 \\
\iff\ & \frac{1}{C_0^2} \leq \|\boldsymbol{\Omega}_{x_0}(\boldsymbol{u}) \cdot p^{\min}_{x_{\boldsymbol{u}}}\|^2 (1 + \|\boldsymbol{u}\|)^2 \\
\iff\ & \frac{1}{C_0} \leq \|\boldsymbol{\Omega}_{x_0}(\boldsymbol{u}) \cdot p^{\min}_{x_{\boldsymbol{u}}}\| (1 + \|\boldsymbol{u}\|)
\end{aligned}$$

for each $\boldsymbol{u} \in (\underline{\mathrm{End}}^{\Sigma}_{x_0})^{-1}(K)$, where each $p^{\min}_{x_{\boldsymbol{u}}}$ is prescribed as in Remark 9.4.2. ∎

This lemma plays a crucial role in the next section.

9.5 The main result

Although the conditions of Lemma 9.4.6 are phrased in terms of the controlled trajectories of the cotangent lift of Σ, it is not obvious that these conditions can be verified in practice. In this section, we show that this is indeed possible, by imposing extra conditions on Σ. These extra conditions are partly Lie-algebraic.

9.5.1 The Υ-condition

In this section, we introduce a technical condition on Lie derivatives.

Definition 9.5.1. Suppose that $K \subseteq M$ is compact. We define K^* by

$$K^* = \left\{p_x \in T^*M \ :\ x \in K \text{ and } \|p_x\| \in \left[\tfrac{1}{2}, \tfrac{3}{2}\right]\right\}. \tag{9.4}$$

Clearly, K^* is itself compact. See Figure 9.1 for an illustration.

In the remainder of this section, K is a nonempty compact subset of M.

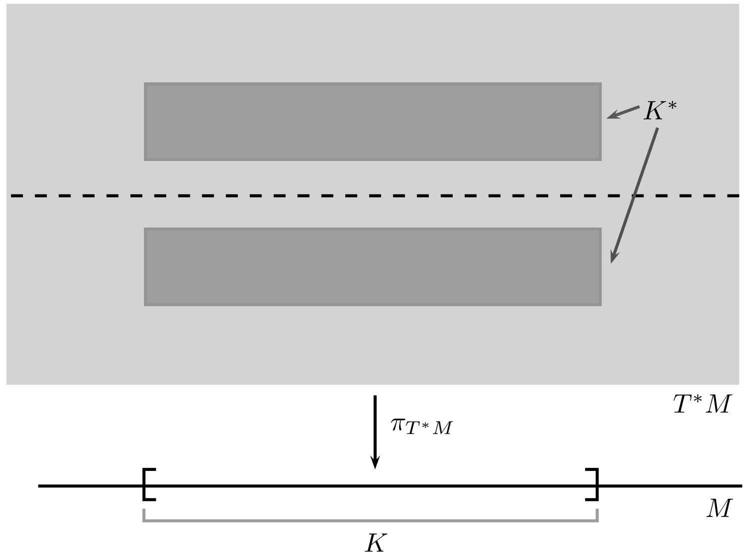

Figure 9.1: An illustration of the compact subset K^*

Definition 9.5.2. Suppose that $G : T^*M \to \mathbb{R}$ is C^1. If there exist

- A (possibly empty) compact subset $\mathscr{N}$ of K^* such that

$$\overline{\rho} = \inf \left\{ 1, \inf_{p_x \in K^* \smallsetminus \mathscr{N}} \|\boldsymbol{\varphi}^{\Sigma}(p_x)\|_{\mathbb{R}^r} \right\} > 0,$$

- Continuous functions $\alpha_i : \mathscr{N} \to \mathbb{R}$ such that

$$\mathscr{L}_{f_0^*} G(p_x) = \sum_{i=1}^{r} \alpha_i(p_x) \varphi_i^{\Sigma}(p_x) \tag{9.5}$$

  for each $p_x \in \mathscr{N}$, and

- C^1 functions $\beta_i : \mathscr{N} \to \mathbb{R}$ and continuous maps $\boldsymbol{\gamma}_i : \mathscr{N} \to \mathbb{R}^r$ such that

$$\mathscr{L}_{\boldsymbol{f}^*} G(p_x) = \sum_{i=1}^{r} \beta_i(p_x) \mathscr{L}_{\boldsymbol{f}^*} \varphi_i^{\Sigma}(p_x) + \varphi_i^{\Sigma}(p_x) \boldsymbol{\gamma}_i(p_x) \tag{9.6}$$

  for each $p_x \in \mathscr{N}$,

then we say that G satisfies the Υ***-condition over*** K. Here, each β_i is C^1 in the sense that it has a C^1 extension to an open submanifold of T^*M containing $\mathscr{N}$.

In the above definition, an infimum taken over $\varnothing$ is equal to ∞.

Remark 9.5.3. Our version of the Υ-condition extends the Υ-condition for driftless control-affine systems. In the literature, the driftless version of the Υ-condition can be found in [Chitour 1996, Proposition 5] and [Chitour and Sussmann 1998].

Remark 9.5.4. From the above remarks, it is clear that if

$$\inf_{p_x \in K^*} \|\boldsymbol{\varphi}^{\Sigma}(p_x)\|_{\mathbb{R}^r} > 0,$$

then each C^1 function $G : T^*M \to \mathbb{R}$ satisfies the Υ-condition over K with $\mathscr{N} = \varnothing$.

9.5.2 The locally Lipschitz-AC chain rule

In the next section, we will need the following version of the chain rule, valid for the composition of a locally Lipschitz function and an AC curve in $\mathbb{R}^k$. This result, which is a composite of [Leoni 2009, Theorem 3.44] and [Leoni 2009, Theorem 4.54], requires the notion of a purely $\mathscr{H}^1$-unrectifiable subset of $\mathbb{R}^k$: If $\mathscr{H}^1$ denotes the 1-dimensional Hausdorff measure on $\mathbb{R}^k$, then the set $E \subseteq \mathbb{R}^k$ is said to be ***purely $\mathscr{H}^1$-unrectifiable*** if

$$\mathscr{H}^1(E \cap \operatorname{image}(\boldsymbol{\ell})) = 0$$

for each Lipschitz curve $\boldsymbol{\ell} : \mathbb{R} \to \mathbb{R}^k$.

Theorem 9.5.5. *Suppose that $F : \mathbb{R}^k \to \mathbb{R}$ is locally Lipschitz and $\boldsymbol{\xi} : J \to \mathbb{R}^k$ is AC. Then $F \circ \boldsymbol{\xi} : J \to \mathbb{R}$ is AC and the following chain rules hold, depending on the value of k:*

1. *If $k = 1$, then*

$$\dot{\widehat{F \circ \boldsymbol{\xi}}}(t) = \dot{F}(\boldsymbol{\xi}(t))\dot{\boldsymbol{\xi}}(t)$$

for a.a. $t \in \operatorname{dom}(\boldsymbol{\xi})$, where we set $\dot{F}(\boldsymbol{\xi}(t))\dot{\boldsymbol{\xi}}(t) = 0$ whenever $\dot{\boldsymbol{\xi}}(t) = 0$;

2. If $k \geq 2$ and the set

$$\left\{\boldsymbol{x} \in \mathbb{R}^k \;:\; F \textit{ is not differentiable at } \boldsymbol{x}\right\}$$

is purely $\mathscr{H}^1$-unrectifiable, then

$$\dot{\widehat{F \circ \boldsymbol{\xi}}}(t) = \boldsymbol{D}F(\boldsymbol{\xi}(t)) \cdot \dot{\boldsymbol{\xi}}(t)$$

for a.a. $t \in \operatorname{dom}(\boldsymbol{\xi})$, where we set $\boldsymbol{D}F(\boldsymbol{\xi}(t)) \cdot \dot{\boldsymbol{\xi}}(t) = 0$ whenever $\dot{\boldsymbol{\xi}}(t) = 0$.

9.5.3 The main result

We now derive the main results of this chapter. Throughout this section, K is a nonempty compact subset of M. We begin by specializing Definition 8.2.1.

Definition 9.5.6. If there exist C^1 functions $P_{i,j} : K \to \mathbb{R}$ such that

$$[f_i, f_0](x) = \sum_{j=1}^{r} P_{i,j}(x) f_j(x) \tag{9.7}$$

for each $1 \leq j \leq r$ and each $x \in K$, then we say that Σ satisfies the ***finite definiteness condition with degree*** $\Delta = 0$ ***on*** K. Here, each $P_{i,j}$ is C^1 in the sense that it has a C^1 extension to an open submanifold of M containing K.

The next proposition extends [Chitour 1996, Proposition 5].[2]

Proposition 9.5.7. *Suppose that*

- *$G : T^*M \to \mathbb{R}$ is C^1,*
- *G satisfies the Υ-condition over K, and*

[2]We note that the cited result is stated in a slightly inaccurate way, since it asserts the existence of a constant C which is universal, in the sense that it is independent of the data μ^* and $[c, d]$. (Here, we are using the notation of Proposition 9.5.7.) On the contrary, the constant C depends strongly on μ^* and $[c, d]$, and this is clear from an inspection of the proofs of [Chitour 1996, Proposition 5] (see also [Chitour 2006, Proposition 4]). The important fact is that the constant C depends only on Σ, G, and the absolute difference Δ of $G(\mu^*(d))$ and $G(\mu^*(c))$. As we will see in Theorem 9.5.12 below, one is ultimately interested in restricting μ^* and $[c, d]$ such that Δ is constant. In this way, one effectively obtains the universal constant C, although we stress that it is only universal after restricting μ^* and $[c, d]$.

- *Σ satisfies the finite definiteness condition with degree $\Delta = 0$ on K.*

Then the following conclusion holds: If $\mu^* : J \to T^*M$ is a $\boldsymbol{u}$-controlled trajectory of Σ^* and $[c,d]$ is a subinterval of J such that $\mu^*(t) \in K^*$ for each $t \in [c,d]$, then there exists $C \in \mathbb{R}_{>0}$ such that

$$\begin{aligned}\Delta &= |G(\mu^*(d)) - G(\mu^*(c))| \\ &\le C\sqrt{\int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}^2 \,\mathrm{d}\sigma}\left(1 + \sqrt{\int_c^d \|\boldsymbol{u}(\sigma)\|_{\mathbb{R}^r}^2 \,\mathrm{d}\sigma}\right).\end{aligned}$$

Furthermore, C depends only on Σ (precisely, on the vector fields $f_0, f_1, \ldots, f_r$ and their cotangent lifts), G, and Δ.

Proof. If $\Delta = 0$, then there is nothing to prove, since we can simply choose $C = 1$. Suppose therefore that $\Delta \neq 0$. Since G satisfies the Υ-condition over K, we let $\mathscr{N}$, $\overline{\rho}$, α_i, β_i, and $\boldsymbol{\gamma}_i$ be prescribed as in Definition 9.5.2. Define $C_1 \in \mathbb{R}_{\ge 0}$ by

$$C_1 = \sup_{p_x \in K^*} \left\{ |\mathscr{L}_{f_0^*} G(p_x)|, \|\mathscr{L}_{\boldsymbol{f}^*} G(p_x)\|_{\mathbb{R}^r} \right\},$$

and let $C_2 \in \mathbb{R}_{\ge 0}$ be the maximum of

$$\sup_{\substack{p_x \in \mathscr{N} \\ 1 \le i \le r}} \left\{ |\alpha_i(p_x)|, |\beta_i(p_x)|, |\mathscr{L}_{f_0^*}\beta_i(p_x)|, \|\mathscr{L}_{\boldsymbol{f}^*}\beta_i(p_x)\|_{\mathbb{R}^r} \right\}$$

and

$$\sup_{\substack{p_x \in \mathscr{N} \\ 1 \le i \le r}} \left\{ |\mathscr{L}_{f_0^*}\varphi_i^\Sigma(p_x)|, \|\mathscr{L}_{\boldsymbol{f}^*}\varphi_i^\Sigma(p_x)\|_{\mathbb{R}^r}, \|\boldsymbol{\gamma}_i(p_x)\|_{\mathbb{R}^r} \right\}.$$

Now let μ^* and $[c,d]$ be given as in the statement of the proposition. We define auxiliary functions as follows: For each $\rho \in (0, \overline{\rho})$, define $\nu_\rho^0 : \mathbb{R} \to \mathbb{R}$ by

$$\nu_\rho^0(t) = \begin{cases} -1, & t < -\rho \\ t\rho^{-1}, & -\rho \le t \le \rho \\ 1, & t > \rho \end{cases}$$

and define $\nu_\rho : [c, d] \to \mathbb{R}$ by

$$\nu_\rho(t) = \nu_\rho^0(\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r}).$$

These functions have the following obvious properties:

- $0 \leq \nu_\rho^0(t) \leq 1$ for each $t \in \mathbb{R}_{\geq 0}$ and
- $0 \leq \nu_\rho^0(t) \leq t\rho^{-1}$ for each $t \in \mathbb{R}_{\geq 0}$, so that

$$0 \leq \nu_\rho(t) \leq \frac{\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r}}{\rho} \tag{9.8}$$

  for each $t \in [c, d]$.

By Lemma 2.3.2, $G \circ \mu^* : [c, d] \to \mathbb{R}$ is AC. Consequently

$$G(\mu^*(d)) - G(\mu^*(c)) = \int_c^d \dot{\widehat{G \circ \mu^*}}(\sigma)\, \mathrm{d}\sigma.$$

Using (9.1), we have

$$\dot{\widehat{G \circ \mu^*}}(t) = \mathscr{L}_{f_0^*} G(\mu^*(t)) + \langle \boldsymbol{u}(t), \mathscr{L}_{\boldsymbol{f}^*} G(\mu^*(t)) \rangle_{\mathbb{R}^r}$$

for a.a. $t \in [c, d]$. Thus

$$\begin{aligned} G(\mu^*(d)) - G(\mu^*(c)) &= \int_c^d \mathscr{L}_{f_0^*} G(\mu^*(\sigma))\, \mathrm{d}\sigma + \int_c^d \langle \boldsymbol{u}(\sigma), \mathscr{L}_{\boldsymbol{f}^*} G(\mu^*(\sigma)) \rangle_{\mathbb{R}^r}\, \mathrm{d}\sigma \\ &= I_1 + I_2, \end{aligned}$$

where

$$I_1 = \int_c^d \nu_\rho(\sigma) \mathscr{L}_{f_0^*} G(\mu^*(\sigma))\, \mathrm{d}\sigma + \int_c^d \nu_\rho(\sigma) \langle \boldsymbol{u}(\sigma), \mathscr{L}_{\boldsymbol{f}^*} G(\mu^*(\sigma)) \rangle_{\mathbb{R}^r}\, \mathrm{d}\sigma$$

and

$$I_2 = \int_c^d (1 - \nu_\rho(\sigma)) \mathscr{L}_{f_0^*} G(\mu^*(\sigma))\, \mathrm{d}\sigma + \int_c^d (1 - \nu_\rho(\sigma)) \langle \boldsymbol{u}(\sigma), \mathscr{L}_{\boldsymbol{f}^*} G(\mu^*(\sigma)) \rangle_{\mathbb{R}^r}\, \mathrm{d}\sigma.$$

It follows that $\Delta \leq |I_1| + |I_2|$. In the remainder of the proof we will establish upper bounds on $|I_1|$ and $|I_2|$ by manipulating ρ, which is indeterminate until Step 3.

STEP 1 (I_1): Using (9.8) and the Cauchy–Schwarz inequality,

$$\begin{aligned} |I_1| &\leq \int_c^d \nu_\rho(\sigma)\,|\mathscr{L}_{f_0^*} G(\mu^*(\sigma))|\,\mathrm{d}\sigma + \int_c^d \nu_\rho(\sigma)\,|\langle \boldsymbol{u}(\sigma), \mathscr{L}_{\boldsymbol{f}^*} G(\mu^*(\sigma))\rangle_{\mathbb{R}^r}|\,\mathrm{d}\sigma \\ &\leq \frac{C_1}{\rho}\int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma + \frac{C_1}{\rho}\int_c^d \|\boldsymbol{u}(\sigma)\|_{\mathbb{R}^r}\,\|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma. \end{aligned}$$

The next two steps of the proof are concerned with I_2. Note that for each $t \in [c,d]$, we have $1-\nu_\rho(t) \neq 0$ if and only if $\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r} < \rho$. Furthermore,

$$\begin{aligned} \|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r} < \rho &\implies \|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r} < \overline{\rho} \\ &\implies \mu^*(t) \notin K^* \setminus \mathscr{N} \\ &\implies \mu^*(t) \in \mathscr{N}. \end{aligned}$$

Thus whenever the term $1-\nu_\rho(t)$ appears as a multiplicative factor in an integrand (or finite sum), we can assume that t is such that $|\varphi_i^\Sigma(\mu^*(t))| \leq \|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r} < \rho$ and $\mu^*(t)$ is contained in $\mathscr{N}$, since otherwise the contribution to the integral (or finite sum) is zero. We will use this fact implicitly in the remainder of the proof. In particular, it follows that

$$I_2 = J_1 + J_2 + J_3,$$

where

$$\begin{aligned} J_1 &= \sum_{i=1}^r \int_c^d (1-\nu_\rho(\sigma))\alpha_i(\mu^*(\sigma))\varphi_i^\Sigma(\mu^*(\sigma))\,\mathrm{d}\sigma, \\ J_2 &= \sum_{i=1}^r \int_c^d (1-\nu_\rho(\sigma))\beta_i(\mu^*(\sigma))\left\langle \boldsymbol{u}(\sigma), \mathscr{L}_{\boldsymbol{f}^*}\varphi_i^\Sigma(\mu^*(\sigma))\right\rangle_{\mathbb{R}^r}\,\mathrm{d}\sigma, \qquad \text{and} \\ J_3 &= \sum_{i=1}^r \int_c^d (1-\nu_\rho(\sigma))\varphi_i^\Sigma(\mu^*(\sigma))\,\langle \boldsymbol{u}(\sigma), \boldsymbol{\gamma}_i(\mu^*(\sigma))\rangle_{\mathbb{R}^r}\,\mathrm{d}\sigma. \end{aligned}$$

STEP 2.1 (J_1): Directly, we have

$$|J_1| \leq rC_2 \int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma.$$

STEP 2.2 (J_2): Using (9.1), we have

$$J_2 = \sum_{i=1}^r \int_c^d (1-\nu_\rho(\sigma))\beta_i(\mu^*(\sigma))\left(\dot{\widehat{\varphi_i^\Sigma \circ \mu^*}}(\sigma) - \mathscr{L}_{f_0^*}\varphi_i^\Sigma(\mu^*(\sigma))\right)\mathrm{d}\sigma$$

$$= J_{2,1} - J_{2,2},$$

where

$$J_{2,1} = \sum_{i=1}^{r} \int_c^d (1 - \nu_\rho(\sigma))\beta_i(\mu^*(\sigma))\dot{\widehat{\varphi_i^\Sigma \circ \mu^*}}(\sigma)\,\mathrm{d}\sigma \qquad \text{and}$$
$$J_{2,2} = \sum_{i=1}^{r} \int_c^d (1 - \nu_\rho(\sigma))\beta_i(\mu^*(\sigma))\mathscr{L}_{f_0^*}\varphi_i^\Sigma(\mu^*(\sigma))\,\mathrm{d}\sigma.$$

By the AC integration by parts formula [Leoni 2009, Corollary 3.37], we have

$$J_{2,1} = J_{2,1,1} - J_{2,1,2} + J_{2,1,3},$$

where

$$J_{2,1,1} = \sum_{i=1}^{r}(1 - \nu_\rho(d))\beta_i(\mu^*(d))\varphi_i^\Sigma(\mu^*(d)) - \sum_{i=1}^{r}(1 - \nu_\rho(c))\beta_i(\mu^*(c))\varphi_i^\Sigma(\mu^*(c)),$$
$$J_{2,1,2} = \sum_{i=1}^{r} \int_c^d \dot{\nu}_\rho(\sigma)\beta_i(\mu^*(\sigma))\varphi_i^\Sigma(\mu^*(\sigma))\,\mathrm{d}\sigma, \qquad \text{and}$$
$$J_{2,1,3} = \sum_{i=1}^{r} \int_c^d (1 - \nu_\rho(\sigma))\dot{\widehat{\beta_i \circ \mu^*}}(\sigma)\varphi_i^\Sigma(\mu^*(\sigma))\,\mathrm{d}\sigma.$$

For $|J_{2,1,1}|$, it is clear that

$$|J_{2,1,1}| \leq 2rC_2\rho.$$

For $|J_{2,1,2}|$, we begin by noting that the Euclidean norm $\|\cdot\|_{\mathbb{R}^r}$ is locally Lipschitz and $\boldsymbol{\varphi}^\Sigma \circ \mu^*$ is AC. It is well-known that

$$\{\boldsymbol{x} \in \mathbb{R}^r \ : \ \|\cdot\|_{\mathbb{R}^r} \text{ is not differentiable at } \boldsymbol{x}\} = \{\mathbf{0}_{\mathbb{R}^r}\}.$$

This set is purely $\mathscr{H}^1$-unrectifiable, since the 1-dimensional Hausdorff measure of a set with at most one element is 0. Invoking Theorem 9.5.5, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r} = \frac{\left\langle \boldsymbol{\varphi}^\Sigma(\mu^*(t)), \dot{\widehat{\boldsymbol{\varphi}^\Sigma \circ \mu^*}}(t)\right\rangle_{\mathbb{R}^r}}{\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r}}$$
$$= \frac{\sum_{i=1}^{r} \varphi_i^\Sigma(\mu^*(t))\,\dot{\widehat{\varphi_i^\Sigma \circ \mu^*}}(t)}{\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r}}$$

for a.a. $t \in [c,d]$. Another invocation of Theorem 9.5.5 yields

$$\begin{aligned}\dot{\nu}_\rho(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\nu^0_\rho(\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r})\\ &= \frac{\sum_{i=1}^r \varphi_i^\Sigma(\mu^*(t))\,\dot{\widehat{\varphi_i^\Sigma \circ \mu^*}}(t)}{\rho\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r}}\end{aligned}$$

for a.a. $t \in [c,d]$. Here $\dot{\nu}^0_\rho(t) = \rho^{-1}$, since

$$\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r} < \rho$$

by construction. Using (9.1), we have

$$\dot{\nu}_\rho(t) = \frac{\sum_{i=1}^r \varphi_i^\Sigma(\mu^*(t))\mathscr{L}_{f_0^*}\varphi_i^\Sigma(\mu^*(t))}{\rho\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r}} + \frac{\sum_{i=1}^r \varphi_i^\Sigma(\mu^*(t))\langle \boldsymbol{u}(t), \mathscr{L}_{\boldsymbol{f}^*}\varphi_i^\Sigma(\mu^*(t))\rangle_{\mathbb{R}^r}}{\rho\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r}}$$

for a.a. $t \in [c,d]$. Using the Cauchy–Schwarz inequality,

$$\begin{aligned}|\dot{\nu}_\rho(t)| &\le \frac{rC_2\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r}}{\rho\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r}} + \frac{rC_2\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r}\|\boldsymbol{u}(t)\|_{\mathbb{R}^r}}{\rho\|\boldsymbol{\varphi}^\Sigma(\mu^*(t))\|_{\mathbb{R}^r}}\\ &= \frac{rC_2}{\rho}(1 + \|\boldsymbol{u}(t)\|_{\mathbb{R}^r})\end{aligned}$$

for a.a. $t \in [c,d]$. Thus

$$|J_{2,1,2}| \le \frac{r^2C_2^2}{\rho}\int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma + \frac{r^2C_2^2}{\rho}\int_c^d \|\boldsymbol{u}(\sigma)\|_{\mathbb{R}^r}\,\|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma.$$

Now let us consider $|J_{2,1,3}|$. Using (9.1), we have

$$\begin{aligned}J_{2,1,3} = &\sum_{i=1}^r \int_c^d (1-\nu_\rho(\sigma))\mathscr{L}_{f_0^*}\beta_i(\mu^*(\sigma))\varphi_i^\Sigma(\mu^*(\sigma))\,\mathrm{d}\sigma\\ &+ \sum_{i=1}^r \int_c^d (1-\nu_\rho(\sigma))\langle \boldsymbol{u}(\sigma), \mathscr{L}_{\boldsymbol{f}^*}\beta_i(\mu^*(\sigma))\rangle_{\mathbb{R}^r}\varphi_i^\Sigma(\mu^*(\sigma))\,\mathrm{d}\sigma.\end{aligned}$$

Using the Cauchy–Schwarz inequality,

$$\begin{aligned}|J_{2,1,3}| \le rC_2 &\int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma\\ &+ rC_2\int_c^d \|\boldsymbol{u}(\sigma)\|_{\mathbb{R}^r}\,\|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma.\end{aligned}$$

Finally, we consider $J_{2,2}$. Recall from Lemma 9.2.4 that $\mathscr{L}_{f_0^*}\varphi_i^\Sigma = \psi_{0,i}^\Sigma$. Using this fact, together with the assumption that Σ satisfies the finite definiteness condition with degree $\Delta = 0$ on K, there exist C^1 functions $P_{i,j} : K \to \mathbb{R}$ such that

$$\begin{aligned}
\mathscr{L}_{f_0^*}\varphi_i^\Sigma(p_x) &= \psi_{0,i}^\Sigma(p_x) \\
&= \langle p_x, [f_i, f_0](x)\rangle \\
&= \left\langle p_x, \sum_{j=1}^{r} P_{i,j}(x) f_j(x) \right\rangle \\
&= \sum_{j=1}^{r} P_{i,j}(x) \langle p_x, f_j(x)\rangle \\
&= \sum_{j=1}^{r} P_{i,j}(x) \varphi_j^\Sigma(p_x)
\end{aligned}$$

for each $p_x \in \mathscr{N}$. Writing $\mu = \pi_{T^*M} \circ \mu^*$, we have

$$J_{2,2} = \sum_{i=1}^{r}\sum_{j=1}^{r}\int_c^d (1 - \nu_\rho(\sigma))\beta_i(\mu^*(\sigma))P_{i,j}(\mu(\sigma))\varphi_j^\Sigma(\mu^*(\sigma))\,\mathrm{d}\sigma.$$

Consequently,

$$|J_{2,2}| \leq r^2 C_2 C_3 \int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma,$$

where

$$C_3 = \sup_{\substack{x \in K \\ 1 \leq i,j \leq r}} \{|P_{i,j}(x)|\}.$$

Step 2.3 (J_3): Using the Cauchy–Schwarz inequality,

$$|J_3| \leq r C_2 \int_c^d \|\boldsymbol{u}(\sigma)\|_{\mathbb{R}^r}\,\|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma.$$

Step 3 (Summary of Steps 1 and 2): We have shown above that there exist constants $C_1, C_2, C_3 \in \mathbb{R}_{\geq 0}$ such that

$$\begin{aligned}
\Delta &\leq |I_1| + |J_1| + \sum_{i=1}^{3} |J_{2,1,i}| + |J_{2,2}| + |J_3| \\
&\leq 2rC_2\rho + r^2 C_2 C_3 \int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma
\end{aligned}$$

$$+\left(\frac{C_1}{\rho}+2rC_2+\frac{r^2C_2^2}{\rho}\right)\int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma$$
$$+\left(\frac{C_1}{\rho}+2rC_2+\frac{r^2C_2^2}{\rho}\right)\int_c^d \|\boldsymbol{u}(\sigma)\|_{\mathbb{R}^r}\|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma. \tag{9.9}$$

By construction, these constants depend only on Σ and G, and do not depend on μ^* or $[c,d]$. To complete the proof, define $C_* \in \mathbb{R}_{>0}$ by

$$C_* = \sup\left\{C_1, 2rC_2, r^2C_2^2, r^2C_2C_3, \frac{\Delta}{\overline{\rho}}\right\}.$$

Note that C_* *does* depend on μ^* and $[c,d]$, by way of its dependence on Δ, and the fact that $C_* \in \mathbb{R}_{>0}$ is due to our earlier assumption that Δ is nonzero. Then (9.9), together with the fact that $\rho < \overline{\rho} \leq 1$, implies that

$$\Delta \leq C_*\rho + \frac{4C_*}{\rho}\int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma + \frac{4C_*}{\rho}\int_c^d \|\boldsymbol{u}(\sigma)\|_{\mathbb{R}^r}\|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma. \tag{9.10}$$

By definition, $\overline{\rho} \geq \frac{\Delta}{C_*}$. Making the specific choice

$$\rho = \frac{\Delta}{2C_*} \leq \frac{\overline{\rho}}{2} \in (0, \overline{\rho}),$$

we have $C_*\rho = \frac{\Delta}{2}$. Using (9.10) and the Cauchy–Schwarz inequality,

$$\begin{aligned}
\Delta &\leq \frac{16C_*^2}{\Delta}\int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma \\
&\quad + \frac{16C_*^2}{\Delta}\int_c^d \|\boldsymbol{u}(\sigma)\|_{\mathbb{R}^r}\|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}\,\mathrm{d}\sigma \\
&\leq \frac{16C_*^2\sqrt{d-c}}{\Delta}\sqrt{\int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}^2\,\mathrm{d}\sigma} \\
&\quad + \frac{16C_*^2}{\Delta}\sqrt{\int_c^d \|\boldsymbol{u}(\sigma)\|_{\mathbb{R}^r}^2\,\mathrm{d}\sigma}\sqrt{\int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}^2\,\mathrm{d}\sigma} \\
&\leq \frac{16C_*^2\sqrt{b-a}}{\Delta}\sqrt{\int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}^2\,\mathrm{d}\sigma} \\
&\quad + \frac{16C_*^2}{\Delta}\sqrt{\int_c^d \|\boldsymbol{u}(\sigma)\|_{\mathbb{R}^r}^2\,\mathrm{d}\sigma}\sqrt{\int_c^d \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}^2\,\mathrm{d}\sigma}.
\end{aligned}$$

Setting

$$C = \sup\left\{\frac{16C_*^2}{\Delta}, \frac{16C_*^2\sqrt{b-a}}{\Delta}\right\}$$

and gathering terms completes the proof. ∎

We now turn to proving the main result of this chapter. To begin, we recall some general notions concerning subsets of metric spaces and their δ-enlargements. Suppose that $X = (X, d_X)$ is a metric space and S is a nonempty subset of X. For each point $x \in X$, the ***point-set distance from** x **to** S* is defined by

$$d_X(x, S) = \inf_{s \in S} d_X(x, s).$$

For each $\delta \in \mathbb{R}_{>0}$, the ***open** δ**-enlargement of*** S is the set

$$B_{<\delta}(S) = \{x \in X \ : \ d_X(x, S) < \delta\},$$

and the ***closed** δ**-enlargement of*** S is the set

$$B_{\leq\delta}(S) = \{x \in X \ : \ d_X(x, S) \leq \delta\}.$$

By [Prasolov 2006, Theorem 2.1], the map $f_S : X \to \mathbb{R}_{\geq 0}$ defined by

$$f_S(x) = d_X(x, S)$$

is continuous. Thus, $B_{<\delta}(S)$ is open and $B_{\leq\delta}(S)$ is closed.

Lemma 9.5.8. *Suppose that X is locally compact and K is a nonempty compact subset of X. Then there exists $\delta_0 \in \mathbb{R}_{>0}$ such that $\overline{B_{<\delta_0}(K)}$ is compact.*

Proof. See [Bridges 1998, Section 3.3]. ∎

Corollary 9.5.9. *Suppose that X is locally compact and S is a nonempty compact subset of X. Then there exists $\delta \in \mathbb{R}_{>0}$ such that $B_{\leq\delta}(S)$ is compact.*

Proof. Choose $\delta = \delta_0/2$, where δ_0 is prescribed as in Lemma 9.5.8. Then $B_{\leq\delta}(K)$ is compact, since it is closed and contained in the compact set $\overline{B_{<\delta_0}(S)}$. ∎

In particular, Lemma 9.5.8 and Corollary 9.5.9 apply whenever X is an open submanifold of M, where X is viewed as a locally compact metric space with the Riemannian distance. In the remainder of this section, we assume that $K \subseteq M_\Delta$.

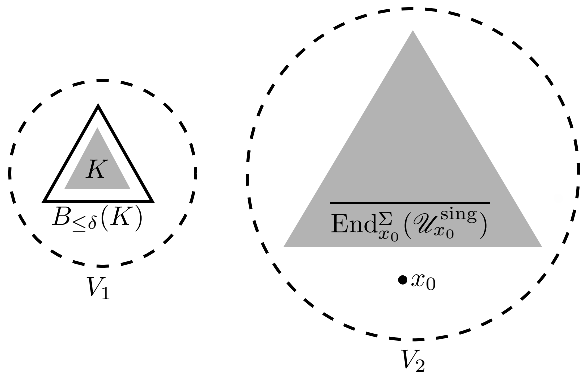

Figure 9.2: An illustration of the proof of Lemma 9.5.10

Lemma 9.5.10. *Suppose that $x_0 \notin K$. Then there exists $\delta \in \mathbb{R}_{>0}$ such that*

1. *$B_{\leq\delta}(K)$ is compact,*
2. *$x_0 \notin B_{\leq\delta}(K)$, and*
3. *$B_{\leq\delta}(K) \subseteq M_\Delta$.*

Proof. By assumption, K and

$$L = \{x_0\} \cup \overline{\mathrm{End}^{\Sigma}_{x_0}(\mathscr{U}^{\mathrm{sing}}_{x_0})}$$

are disjoint closed subsets of M. Since M is a metric space endowed with the Riemannian distance, there exist disjoint open submanifolds V_1 and V_2 of M containing K and L, respectively; see, for example, [Munkres 2000, Theorem 32.2]. Invoking Corollary 9.5.9 with $X = V_1$, there exists $\delta \in \mathbb{R}_{>0}$ such that $B_{\leq\delta}(K)$ is compact and contained in V_1. This is illustrated in Figure 9.2. ∎

Remark 9.5.11. Note that the proof of Lemma 9.5.10 relies fundamentally on the closure operation, which is built into the definition of M_Δ.

The next theorem constitutes the main result in this chapter. We use the following notation: For each $\delta \in \mathbb{R}_{>0}$, we write $K_\delta = B_{\leq\delta}(K)$. Define

$$K_1^* = \{p_x \in T^*M \ : \ x \in K \text{ and } \|p_x\| = 1\}.$$

Clearly, K_1^* is compact and contained in each $K_\delta^* = B_{\leq\delta}(K)^*$. By construction,

$$\overline{T^*M \setminus K_\delta^*} \subseteq T^*M \setminus K_1^*.$$

By [Lee 2003, Proposition 2.26], there exists a C^∞ bump function

$$G : T^*M \to \mathbb{R}$$

for $\overline{T^*M \setminus K_\delta^*}$ supported in $T^*M \setminus K_1^*$. In particular,

- $G(p_x) = 1$ for each $p_x \in \overline{T^*M \setminus K_\delta^*}$ and
- $G(p_x) = 0$ for each $p_x \in K_1^*$.

Although the existence of such a bump function G is guaranteed, the next theorem requires, in addition, that G satisfies the Υ-condition over K_δ.

Theorem 9.5.12. *Suppose that*

- $x_0 \notin K$*,*
- $\delta \in \mathbb{R}_{>0}$ *satisfies the conclusions of Lemma 9.5.10,*
- *There exists a* C^∞ *bump function* $G : T^*M \to \mathbb{R}$ *for* $\overline{T^*M \setminus K_\delta^*}$ *supported in* $T^*M \setminus K_1^*$ *that satisfies the* Υ*-condition over* K_δ*, and*
- Σ *satisfies the finite definiteness condition with degree* $\Delta = 0$ *on* K_δ*.*

Then the Moore–Penrose pseudoinverse of $T\underline{\mathrm{End}}_{x_0}^{\Sigma}$ *has sublinear growth over* K*.*

Proof. Choose $\boldsymbol{u} \in (\underline{\mathrm{End}}_{x_0}^{\Sigma})^{-1}(K)$ and $p_{x_{\boldsymbol{u}}} \in S_{x_{\boldsymbol{u}}}^* M$. For brevity, we write

$$\mu^*(t) = \mu^{\Sigma^*}(t, p_{x_{\boldsymbol{u}}}, \boldsymbol{u})$$

in what follows. Note that

- $G(\mu^*(a)) = 1$, since $\mu^*(a) \in \overline{T^*M \setminus K_\delta^*}$ and
- $G(\mu^*(b)) = G(p_{x_{\boldsymbol{u}}}) = 0$, since $S_{x_{\boldsymbol{u}}}^* M \subseteq K_1^*$.

Thus there exists $c \in J$ such that $G(\mu^*(c)) = 1/2$ and $\mu^*(t) \in K_\delta^*$ for each $t \in [c, b]$. See Figure 9.3. By Proposition 9.5.7, there exists $C \in \mathbb{R}_{>0}$ such that

$$\begin{aligned}
\frac{1}{2} &= \left| G(p_{x_{\boldsymbol{u}}}) - \frac{1}{2} \right| \\
&= |G(\mu^*(b)) - G(\mu^*(c))| \\
&\leq C \sqrt{\int_c^b \|\boldsymbol{\varphi}^\Sigma(\mu^*(\sigma))\|_{\mathbb{R}^r}^2 \, \mathrm{d}\sigma} \left(1 + \sqrt{\int_c^b \|\boldsymbol{u}(\sigma)\|_{\mathbb{R}^r}^2 \, \mathrm{d}\sigma} \right) \\
&\leq C \|\boldsymbol{\varphi}^\Sigma \circ \mu^*\| (1 + \|\boldsymbol{u}\|) \\
&= C \|\boldsymbol{\Omega}_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}}\| (1 + \|\boldsymbol{u}\|).
\end{aligned}$$

Furthermore, C depends only on Σ, G, and the fact that

$$\frac{1}{2} = |G(\mu^*(b)) - G(\mu^*(c))| \, .$$

In particular, C does not depend on $\boldsymbol{u}$ and $p_{x_{\boldsymbol{u}}}$. Hence

$$0 < \frac{1}{2C} \leq \|\boldsymbol{\Omega}_{x_0}(\boldsymbol{u}) \cdot p_{x_{\boldsymbol{u}}}\| (1 + \|\boldsymbol{u}\|)$$

for each $\boldsymbol{u} \in (\underline{\mathrm{End}}_{x_0}^\Sigma)^{-1}(K)$ and each $p_{x_{\boldsymbol{u}}} \in S_{x_{\boldsymbol{u}}}^* M$. Now invoke Lemma 9.4.6. ∎

We will see a concrete application of Theorem 9.5.12 in Chapter 11.

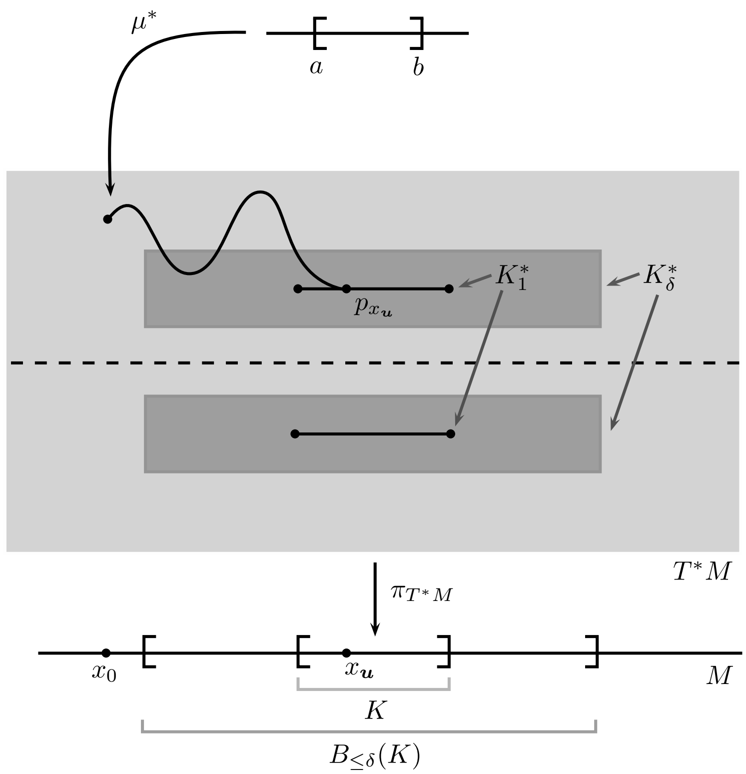

Figure 9.3: An illustration of the proof of Theorem 9.5.12

Chapter 10

A necessary condition for unobstructed motion planning by the continuation method

In the preceding chapters, we derived conditions which ensure that the three obstructions to the continuation method are overcome. Here, we assume that the obstructions are overcome, and then study a particular consequence of this assumption. Thus the subject of this chapter is a necessary condition for unobstructed motion planning by the continuation method.

This chapter is organized in the following way. In Section 10.1, we recall basic facts about fiber bundles. In Section 10.2, we establish two notions of PLE-completeness. In Section 10.3, we briefly review the required theory of initial value problems evolving on Banach manifolds. In Section 10.4, we derive general topological results concerning PLE-completeness. Finally, in Section 10.5, we apply the results of Section 10.4 to control systems, yielding the necessary condition.

Our standing assumptions in this chapter are that

- Q is a Banach manifold modelled on a Banach space E_Q,

- R is a second-countable ℓ-dimensional manifold,
- $F : Q \to R$ is a surjective C^1 submersion, and
- $TF^\dagger$ is a locally Lipschitz right inverse of TF.

10.1 Fiber bundles

In this section, we recall basic facts about fiber bundles. This material is well-known, and more information can be found, for example, in [Husemoller 1994].

Definition 10.1.1. Suppose that E and B are topological spaces. A ***fiber bundle*** is a surjective continuous map $p : E \to B$ that is ***locally trivializable*** in the following sense: For each $b \in B$, there exist

- A neighbourhood V_b of b in B,
- A topological space F_b called the ***fiber over*** b, and
- A homeomorphism $\Psi_b : p^{-1}(V_b) \to F_b \times V_b$, such that

$$\mathrm{pr}_2 \circ \Psi_b = p|p^{-1}(V_b).$$

This definition is illustrated in Figure 10.1.

Definition 10.1.2. Suppose that E and B are Banach manifolds, and $k \in \mathbb{N}^*$. A C^k ***fiber bundle*** is a surjective C^k submersion $p : E \to B$ that is C^k ***locally trivializable*** in the following sense: For each $b \in B$, there exist

- A neighbourhood V_b of b in B,
- A C^k Banach manifold F_b called the ***fiber above*** b, and
- A C^k diffeomorphism $\Psi_b : p^{-1}(V_b) \to F_b \times V_b$, such that

$$\mathrm{pr}_2 \circ \Psi_b = p|p^{-1}(V_b).$$

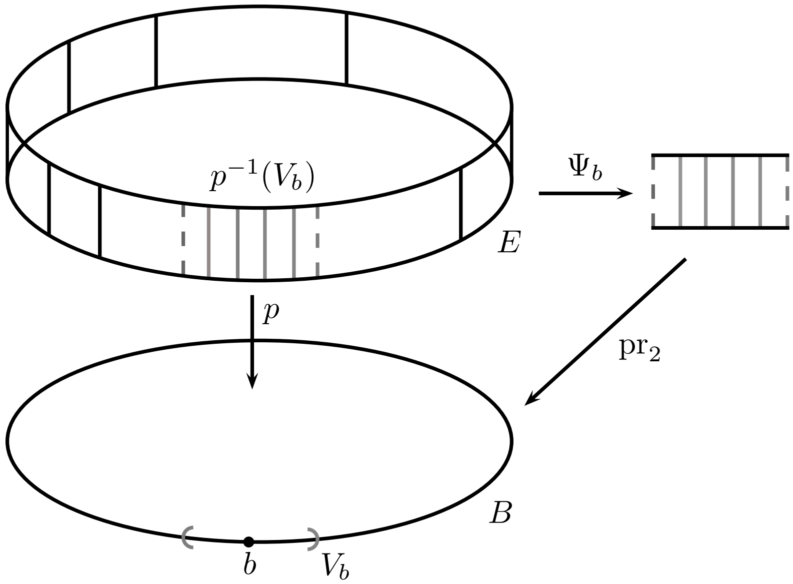

Figure 10.1: An illustration of a fiber bundle $p : E \to B$

Of course, if p is a C^k fiber bundle, then it is a fiber bundle.

10.2 PLE-completeness

In this section, we establish two notions of PLE-completeness. We begin by recalling from Chapter 4 that the (π, q_0)-path-lifting equation, or simply (π, q_0)-PLE for F, is the initial value problem

$$\begin{cases} \dot{\Pi}(t) = TF^\dagger(\Pi(t)) \cdot H_\pi(t, F \circ \Pi(t)), \quad \Pi(t) \in Q, \quad t \in [0,1] \\ \Pi(0) = q_0, \end{cases} \tag{10.1}$$

where $\pi : [0,1] \to R$ is a C^1 curve, $q_0 \in F^{-1}(\pi(0))$, H_π is prescribed as in Proposition 4.3.2, and $TF^\dagger$ is a fixed locally Lipschitz right inverse of TF.

Definition 10.2.1. We say that F is ***PLE-complete*** if for each C^1 curve

$$\pi : [0,1] \to R$$

and each $q_0 \in F^{-1}(\pi(0))$, the maximally-defined solution of the (π, q_0)-PLE for F is defined on $[0, 1]$. Here, the right inverse $TF^\dagger$ is understood. When we wish to emphasize the role of $TF^\dagger$, we will say that F is PLE-complete ***relative to*** $TF^\dagger$.

Next, we establish a weaker notion of PLE-completeness.

Definition 10.2.2. Suppose that $r \in R$. A ***coordinate ball centered at*** r is a chart $(V, \boldsymbol{\psi})$ on R such that

- $r \in V$,
- $\boldsymbol{\psi}(V)$ is an open ball in $\mathbb{R}^\ell$ centered at $\mathbf{0}_{\mathbb{R}^\ell}$, and
- $\boldsymbol{\psi}(r) = \mathbf{0}_{\mathbb{R}^\ell}$.

Definition 10.2.3. Suppose that $(V, \boldsymbol{\psi})$ is a chart on R. Relative to this chart, a ***line segment*** is a curve $L : [0, 1] \to \boldsymbol{\psi}(V)$ such that

$$L(t) = t\boldsymbol{x} + (1 - t)\boldsymbol{y}$$

for some $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\psi}(V)$. We say that F is ***PLE-complete for line segments*** if for each $r \in R$, there exists a coordinate ball $(V, \boldsymbol{\psi})$ centered at r with the following property: For each line segment $L : [0, 1] \to \boldsymbol{\psi}(V)$ and each

$$q_0 \in F^{-1}(\boldsymbol{\psi}^{-1} \circ L(0)),$$

the maximally-defined solution of the $(\boldsymbol{\psi}^{-1} \circ L, q_0)$-PLE for F is defined on $[0, 1]$.

Note that if F is PLE-complete, then it is PLE-complete for line segments.

10.3 Initial value problems evolving on Banach manifolds

In this section, we briefly review the required theory of initial value problems evolving on Banach manifolds. This material is distinct from the material of Section

4.2, since we are interested in initial value problems whose right-hand sides are functionally dependent on an extra parameter which lies in a locally compact metric space Λ. Although there is a considerable overlap with Section 4.2, we have elected to present this material in a systematic way for the sake of completeness. Throughout this section, $\xi : \mathrm{dom}(\xi) \to Q$ is a curve.

Definition 10.3.1. Suppose that $X : Q \times \Lambda \to TQ$. We say that X is a ***parameter-dependent vector field*** on Q if $\pi_{TQ} \circ X(q, \lambda) = q$ for each $(q, \lambda) \in Q \times \Lambda$. The set of all such maps is denoted by $\mathscr{V}(Q, \Lambda)$. Given a chart (U, φ) on Q, the ***local representative*** of X in (U, φ) is the map $X_\varphi : \varphi(U) \times \Lambda \to E_Q$ defined by

$$X_\varphi(q, \lambda) = T\varphi(\varphi^{-1}(q)) \cdot X(\varphi^{-1}(q), \lambda).$$

In what follows, $X \in \mathscr{V}(Q, \Lambda)$.

Definition 10.3.2. Suppose that $(q_0, \lambda) \in Q \times \Lambda$. The triple (X, q_0, λ) is said to be an ***initial value problem*** evolving on Q, with ***right-hand side*** X, ***initial condition*** q_0, and ***parameter*** λ. We say that ξ is a ***solution*** of (X, q_0, λ) if

- $\mathrm{dom}(\xi)$ is a relatively open subinterval of $\mathbb{R}$ containing 0,
- ξ is C^1,
- $\xi(0) = q_0$, and $\dot{\xi}(t) = X(\xi(t), \lambda)$ for each $t \in \mathrm{dom}(\xi)$.

Definition 10.3.3. Suppose that $(q_0, \lambda) \in Q \times \Lambda$ and ξ is a solution of (X, q_0, λ). We say that ξ is ***maximally-defined*** if it has the following property: If

$$\tilde{\xi} : \mathrm{dom}(\tilde{\xi}) \to Q$$

is another solution of (X, q_0, λ), then $\mathrm{dom}(\tilde{\xi}) \subseteq \mathrm{dom}(\xi)$ and

$$\tilde{\xi}(t) = \xi(t)$$

for each $t \in \mathrm{dom}(\tilde{\xi})$. Clearly, such a solution is unique.

Next, we establish a suitable Lipschitz condition on X.

Definition 10.3.4. Suppose that U is a nonempty open subset of E_Q and

$$f : U \times \Lambda \to E_Q.$$

We say that f is ***locally Lipschitz*** if it is continuous and for each $(u_*, \lambda_*) \in U \times \Lambda$, there exist a product neighbourhood $U_* \times \Lambda_*$ of (u_*, λ_*) and $C_* \in \mathbb{R}_{\geq 0}$ such that

$$\|f(u, \lambda) - f(\tilde{u}, \lambda)\|_{E_Q} \leq C_* \|u - \tilde{u}\|_{E_Q}$$

for each $u, \tilde{u} \in U_*$ and each $\lambda \in \Lambda_*$.

Note that if $\Lambda \subseteq \mathbb{R}^p$ is open and f is C^1, then f is locally Lipschitz.

Definition 10.3.5. We say that X is ***locally Lipschitz*** if X_φ is locally Lipschitz for each chart (U, φ) on Q.

One can show that X is locally Lipschitz if and only if X_φ is locally Lipschitz for each chart $(U, \varphi) \in \mathscr{A}_Q$, where $\mathscr{A}_Q$ is a compatible atlas on Q.

Theorem 10.3.6. *Suppose that X is locally Lipschitz. Then there exists a maximally-defined solution of (X, q_0, λ) for each $(q_0, \lambda) \in Q \times \Lambda$.*

Proof. This follows from [Amann 1990, Theorem 7.6], together with a globalization procedure analogous to the one employed in Section 2.3. ∎

Provided that X is locally Lipschitz, the maximally-defined solution of the initial value problem (X, q_0, λ) is denoted by

$$\mu^X(\cdot, q_0, \lambda) : I^X(q_0, \lambda) \to Q.$$

Since $I^X(q_0, \lambda)$ is an open subinterval of $\mathbb{R}$ containing 0, it can be written as

$$I^X(q_0, \lambda) = (I^X_-(q_0, \lambda), I^X_+(q_0, \lambda))$$

for $I^X_-(q_0, \lambda) \in \mathbb{R}_{<0}$ and $I^X_+(q_0, \lambda) \in \mathbb{R}_{>0}$.

Definition 10.3.7. Suppose that X is locally Lipschitz. Define

$$\mathrm{dom}(\Phi^X) = \{(t, q_0, \lambda) \in \mathbb{R} \times Q \times \Lambda \ : \ t \in I^X(q_0, \lambda)\}.$$

The ***global flow*** of X is the map $\Phi^X : \mathrm{dom}(\Phi^X) \to Q$ that sends (t, q_0, λ) to

$$\Phi^X_t(q_0, \lambda) = \mu^X(t, q_0, \lambda).$$

Theorem 10.3.8. *Suppose that X is locally Lipschitz. Then*

1. $\mathrm{dom}(\Phi^X)$ *is an open subset of* $\mathbb{R} \times Q \times \Lambda$ *and*

2. Φ^X *is continuous.*

Furthermore, if $\Lambda \subseteq \mathbb{R}^p$ is open and X is C^k, where $k \in \mathbb{N}^*$, then Φ^X is C^k.

Proof. This follows from [Amann 1990, Theorem 8.3], together with a globalization procedure analogous to the one employed in Section 2.3. ∎

10.4 The main results

In this section, we derive the main results of this chapter. Throughout this section, $\mathscr{A}_Q$ is a compatible atlas on Q. The following technical lemma will be useful.

Lemma 10.4.1. *Suppose that $r \in R$ and $(V, \boldsymbol{\psi})$ is a coordinate ball centered at r. Then $W = F^{-1}(V)$ is an open submanifold of Q and the parameter-dependent vector field $X \in \mathscr{V}(W, \mathbb{R}^\ell)$ defined by*

$$X(q, \boldsymbol{\xi}) = TF^\dagger(q) \circ T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi} \circ F(q)) \cdot \boldsymbol{\xi}$$

is locally Lipschitz.

Proof. Consider the compatible atlas $\mathscr{A}_W$ on W defined by

$$\mathscr{A}_W = \{(U \cap W, \varphi | U \cap W) \ : \ (U, \varphi) \in \mathscr{A}_Q \text{ and } U \cap W \neq \varnothing\}.$$

Choose $(U, \varphi) \in \mathscr{A}_W$. We must show that X_φ is locally Lipschitz. To this end, let $(q_*, \boldsymbol{\xi}_*) \in \varphi(U) \times \mathbb{R}^\ell$. By construction, $F(U) \subseteq V$. Furthermore, since W is an open submanifold of Q, (U, φ) is also a chart on Q. Thus the map

$$TF^\dagger_{\varphi,\boldsymbol{\psi}} : \varphi(U) \to \mathrm{Hom}(\mathbb{R}^\ell, E_Q),$$

defined as in Proposition 4.1.4, is locally Lipschitz. Thus there exists a neighbourhood U_* of q_* and $C_* \in \mathbb{R}_{\geq 0}$ such that

$$\|TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) - TF^\dagger_{\varphi,\boldsymbol{\psi}}(\tilde{q})\| \leq C_* \|q - \tilde{q}\|_{E_Q}$$

for each $q, \tilde{q} \in U_*$. Returning to X_φ, we have

$$\begin{aligned}
X_\varphi(q, \boldsymbol{\xi}) &= T\varphi(\varphi^{-1}(q)) \cdot X(\varphi^{-1}(q), \boldsymbol{\xi}) \\
&= T\varphi(\varphi^{-1}(q)) \circ TF^\dagger(\varphi^{-1}(q)) \circ T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi} \circ F \circ \varphi^{-1}(q)) \cdot \boldsymbol{\xi} \\
&= T\varphi(\varphi^{-1}(q)) \circ TF^\dagger(\varphi^{-1}(q)) \circ T\boldsymbol{\psi}^{-1}(F_{\boldsymbol{\psi},\varphi}(q)) \cdot \boldsymbol{\xi} \\
&= TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) \cdot \boldsymbol{\xi}
\end{aligned}$$

for each $(q, \boldsymbol{\xi}) \in \varphi(U) \times \mathbb{R}^\ell$. This implies that X_φ is continuous at $(q_*, \boldsymbol{\xi}_*)$, since

$$\begin{aligned}
&\|X_\varphi(q, \boldsymbol{\xi}) - X_\varphi(\tilde{q}, \tilde{\boldsymbol{\xi}})\|_{E_Q} \\
&= \|TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) \cdot \boldsymbol{\xi} - TF^\dagger_{\varphi,\boldsymbol{\psi}}(\tilde{q}) \cdot \tilde{\boldsymbol{\xi}}\|_{E_Q} \\
&= \|TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) \cdot \boldsymbol{\xi} - TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) \cdot \tilde{\boldsymbol{\xi}} + TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) \cdot \tilde{\boldsymbol{\xi}} - TF^\dagger_{\varphi,\boldsymbol{\psi}}(\tilde{q}) \cdot \tilde{\boldsymbol{\xi}}\|_{E_Q} \\
&= \|TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) \cdot (\boldsymbol{\xi} - \tilde{\boldsymbol{\xi}}) + (TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) - TF^\dagger_{\varphi,\boldsymbol{\psi}}(\tilde{q})) \cdot \tilde{\boldsymbol{\xi}}\|_{E_Q} \\
&\leq \|TF^\dagger_{\varphi,\boldsymbol{\psi}}(q)\| \, \|\boldsymbol{\xi} - \tilde{\boldsymbol{\xi}}\|_{\mathbb{R}^\ell} + \|TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) - TF^\dagger_{\varphi,\boldsymbol{\psi}}(\tilde{q})\| \, \|\tilde{\boldsymbol{\xi}}\|_{\mathbb{R}^\ell} \\
&\leq \|TF^\dagger_{\varphi,\boldsymbol{\psi}}(q)\| \, \|\boldsymbol{\xi} - \tilde{\boldsymbol{\xi}}\|_{\mathbb{R}^\ell} + C_* \|q - \tilde{q}\| \, \|\tilde{\boldsymbol{\xi}}\|_{\mathbb{R}^\ell}
\end{aligned}$$

for each $(q, \boldsymbol{\xi}), (\tilde{q}, \tilde{\boldsymbol{\xi}}) \in U_* \times \mathbb{R}^\ell$. Furthermore,

$$\begin{aligned}
\|X_\varphi(q, \boldsymbol{\xi}) - X_\varphi(\tilde{q}, \boldsymbol{\xi})\|_{E_Q} &= \|TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) \cdot \boldsymbol{\xi} - TF^\dagger_{\varphi,\boldsymbol{\psi}}(\tilde{q}) \cdot \boldsymbol{\xi}\|_{E_Q} \\
&= \|(TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) - TF^\dagger_{\varphi,\boldsymbol{\psi}}(\tilde{q})) \cdot \boldsymbol{\xi}\|_{E_Q}
\end{aligned}$$

$$\leq \|TF^{\dagger}_{\varphi,\boldsymbol{\psi}}(q) - TF^{\dagger}_{\varphi,\boldsymbol{\psi}}(\tilde{q})\| \, \|\boldsymbol{\xi}\|_{\mathbb{R}^\ell}$$
$$\leq C_* \|q - \tilde{q}\|_{E_Q} \, \|\boldsymbol{\xi}\|_{\mathbb{R}^\ell}$$
$$\leq C_*(1 + \|\boldsymbol{\xi}_*\|_{\mathbb{R}^\ell})\|q - \tilde{q}\|_{E_Q}$$

for each $q, \tilde{q} \in U_*$ and each $\boldsymbol{\xi} \in B_{<1}(\boldsymbol{\xi}_*)$. Since $(q_*, \boldsymbol{\xi}_*) \in \varphi(U) \times \mathbb{R}^\ell$ was chosen arbitrarily, we conclude that X_φ is locally Lipschitz. This completes the proof. ∎

The next result is based on [Gutú and Jaramillo 2004, Theorem 2.3], but differs from the cited theorem in two major respects. First of all, the result of Gutú and Jaramillo [2004] is phrased in terms of a continuation property with respect to arbitrary C^1 lifts of C^1 curves $\pi : [0,1] \to R$. This continuation property implies that F is, in our terminology, PLE-complete with respect to any locally Lipschitz right inverse of TF. It is therefore enough to show that a locally Lipschitz right inverse of TF exists. In contrast, our C^1 lifts are not arbitrary. Indeed, they are solutions of PLEs, which involve a particular choice of locally Lipschitz right inverse. Second, PLEs incorporate the time-varying vector fields H_π. Handling these differences is fairly straightforward, as we will see below.

Theorem 10.4.2. *Suppose that F is PLE-complete. Then F is a fiber bundle.*

Proof. Since F is surjective and continuous, it remains to show that F is locally trivializable. To this end, let $r \in R$, let $(V, \boldsymbol{\psi})$ be a coordinate ball centered at r, and define $W = F^{-1}(V)$. Then $X \in \mathscr{V}(W, \mathbb{R}^\ell)$, defined by

$$X(q, \boldsymbol{\xi}) = TF^{\dagger}(q) \circ T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi} \circ F(q)) \cdot \boldsymbol{\xi}$$

is locally Lipschitz by Lemma 10.4.1. Now define $\hat{F} : W \to \boldsymbol{\psi}(V)$ by

$$\hat{F} = \boldsymbol{\psi} \circ F,$$

and let $(q_0, \boldsymbol{\xi}) \in W \times \mathbb{R}^\ell$. Clearly, $\hat{F}(\mu^X(0, q_0, \boldsymbol{\xi})) = \hat{F}(q_0)$, and for each

$$t \in (I^X_-(q_0, \boldsymbol{\xi}), I^X_+(q_0, \boldsymbol{\xi}))$$

the tangent vector to the composite curve $\hat{F} \circ \mu^X(\cdot, q_0, \boldsymbol{\xi})$ at t is equal to

$$\begin{aligned}
&T\hat{F}(\mu^X(t, q_0, \boldsymbol{\xi})) \cdot \dot{\mu}^X(t, q_0, \boldsymbol{\xi}) \\
&= T\hat{F}(\mu^X(t, q_0, \boldsymbol{\xi})) \cdot X(\mu^X(t, q_0, \boldsymbol{\xi}), \boldsymbol{\xi}) \\
&= T\hat{F}(\mu^X(t, q_0, \boldsymbol{\xi})) \circ TF^{\dagger}(\mu^X(t, q_0, \boldsymbol{\xi})) \circ T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi} \circ F(\mu^X(t, q_0, \boldsymbol{\xi}))) \cdot \boldsymbol{\xi} \\
&= T\boldsymbol{\psi}(F(\mu^X(t, q_0, \xi))) \circ TF(\mu^X(t, q_0, \boldsymbol{\xi})) \circ TF^{\dagger}(\mu^X(t, q_0, \boldsymbol{\xi})) \circ \\
&\qquad T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi} \circ F(\mu^X(t, q_0, \boldsymbol{\xi}))) \cdot \boldsymbol{\xi} \\
&= T\boldsymbol{\psi}(F(\mu^X(t, q_0, \boldsymbol{\xi}))) \circ T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi} \circ F(\mu^X(t, q_0, \boldsymbol{\xi}))) \cdot \boldsymbol{\xi} \\
&= \boldsymbol{\xi}.
\end{aligned}$$

It follows that for each $(q_0, \boldsymbol{\xi}) \in W \times \mathbb{R}^\ell$ and each $t \in (I_-^X(q_0, \boldsymbol{\xi}), I_+^X(q_0, \boldsymbol{\xi}))$,

$$\hat{F}(\mu^X(t, q_0, \boldsymbol{\xi})) = \hat{F}(q_0) + t\boldsymbol{\xi}. \tag{10.2}$$

We now prove two technical sublemmas.

Lemma 10.4.3. *For each $q_0 \in W$, we have*

$$[-1, 0] \subseteq (I_-^X(q_0, \hat{F}(q_0)), I_+^X(q_0, \hat{F}(q_0))).$$

Proof. Choose $q_0 \in W$, and suppose that

$$I_-^X(q_0, \hat{F}(q_0)) \in (-1, 0).$$

Define the line segment $L_1 : [0, 1] \to \boldsymbol{\psi}(V)$ by

$$L_1(t) = (1 - t)\hat{F}(q_0).$$

Note that

- L_1 is well-defined, since $\boldsymbol{\psi}(V)$ is convex, and
- $\boldsymbol{\psi} \circ F(q_0) = \hat{F}(q_0) = L_1(0)$, which implies that $q_0 \in F^{-1}(\boldsymbol{\psi}^{-1} \circ L_1(0))$.

Since F is PLE-complete, the maximally-defined solution of

$$\begin{cases} \dot{\omega}(t) = TF^{\dagger}(\omega(t)) \circ H_{\boldsymbol{\psi}^{-1} \circ L_1}(t, F \circ \omega(t)), \quad \omega(t) \in Q, \quad t \in [0,1] \\ \omega(0) = q_0 \end{cases} \tag{10.3}$$

is defined on $[0,1]$. We denote this solution by $\omega : [0,1] \to Q$. By Lemma 4.3.3, we have $F \circ \omega = \boldsymbol{\psi}^{-1} \circ L_1$. Writing $\boldsymbol{\mathcal{L}}_1 = \boldsymbol{\psi}^{-1} \circ L_1$ for brevity, we compute

$$\begin{aligned} \dot{\omega}(t) &= TF^{\dagger}(\omega(t)) \cdot H_{\mathcal{L}_1}(t, F \circ \omega(t)) \\ &= TF^{\dagger}(\omega(t)) \cdot H_{\mathcal{L}_1}(t, \boldsymbol{\mathcal{L}}_1(t)) \\ &= TF^{\dagger}(\omega(t)) \cdot \dot{\boldsymbol{\mathcal{L}}}_1(t) \\ &= TF^{\dagger}(\omega(t)) \circ T\boldsymbol{\psi}^{-1}(L_1(t)) \cdot \dot{L}_1(t) \\ &= -TF^{\dagger}(\omega(t)) \circ T\boldsymbol{\psi}^{-1}(L_1(t)) \cdot \hat{F}(q_0) \\ &= -TF^{\dagger}(\omega(t)) \circ T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi} \circ \boldsymbol{\psi}^{-1} \circ L_1(t)) \cdot \hat{F}(q_0) \\ &= -TF^{\dagger}(\omega(t)) \circ T\boldsymbol{\psi}^{-1}(\boldsymbol{\psi} \circ F(\omega(t))) \cdot \hat{F}(q_0) \\ &= -X(\omega(t), \hat{F}(q_0)) \end{aligned}$$

for each $t \in [0,1]$. Consider now the curve

$$\Omega : [0, -I_{-}^{X}(q_0, \hat{F}(q_0))) \to Q$$

defined by

$$\Omega(t) = \mu^X(-t, q_0, \hat{F}(q_0)).$$

Since $\Omega(0) = q_0$ and

$$\begin{aligned} \dot{\Omega}(t) &= -\dot{\mu}^X(-t, q_0, \hat{F}(q_0)) \\ &= -X(\mu^X(-t, q_0, \hat{F}(q_0)), \hat{F}(q_0)) \\ &= -X(\Omega(t), \hat{F}(q_0)) \end{aligned}$$

for each $t \in [0, -I_{-}^{X}(q_0, \hat{F}(q_0)))$, we see that Ω is a solution of (10.3). Thus

$$\Omega(t) = \omega(t)$$

for each $t \in [0, -I_-^X(q_0, \hat{F}(q_0)))$. By continuity, we have

$$\begin{aligned}
\lim_{t \searrow I_-^X(q_0, \hat{F}(q_0))} \mu^X(t, q_0, \hat{F}(q_0)) &= \lim_{t \nearrow -I_-^X(q_0, \hat{F}(q_0))} \mu^X(-t, q_0, \hat{F}(q_0)) \\
&= \lim_{t \nearrow -I_-^X(q_0, \hat{F}(q_0))} \Omega(t) \\
&= \lim_{t \nearrow -I_-^X(q_0, \hat{F}(q_0))} \omega(t) \\
&= \omega(-I_-^X(q_0, \hat{F}(q_0))).
\end{aligned}$$

Clearly, this contradicts the fact that $\mu^X(\cdot, q_0, \hat{F}(q_0))$ is the maximally-defined solution of $(X, q_0, \hat{F}(q_0))$. Hence $I_-^X(q_0, \hat{F}(q_0)) \notin (-1, 0)$. This completes the proof. ∎

Lemma 10.4.4. *For each* $(q_0, \boldsymbol{\xi}) \in \hat{F}^{-1}(\mathbf{0}_{\mathbb{R}^\ell}) \times \boldsymbol{\psi}(V)$*, we have*

$$[0, 1] \subseteq (I_-^X(q_0, \boldsymbol{\xi}), I_+^X(q_0, \boldsymbol{\xi})).$$

Proof. Choose $(q_0, \boldsymbol{\xi}) \in \hat{F}^{-1}(\mathbf{0}_{\mathbb{R}^\ell}) \times \boldsymbol{\psi}(V)$, and suppose that

$$I_+^X(q_0, \boldsymbol{\xi}) \in (0, 1).$$

Define the line segment $L_2 : [0, 1] \to \boldsymbol{\psi}(V)$ by

$$L_2(t) = t\boldsymbol{\xi}.$$

Note that

- L_2 is well-defined, since $\boldsymbol{\psi}(V)$ is an open ball centered at $\mathbf{0}_{\mathbb{R}^\ell}$, and
- $\boldsymbol{\psi} \circ F(q_0) = \hat{F}(q_0) = \mathbf{0}_{\mathbb{R}^\ell} = L_2(0)$, which implies that $q_0 \in F^{-1}(\boldsymbol{\psi}^{-1} \circ L_2(0))$.

Since F is PLE-complete, the maximally-defined solution of

$$\begin{cases} \dot{\omega}(t) = TF^\dagger(\omega(t)) \cdot H_{\boldsymbol{\psi}^{-1} \circ L_2}(t, F \circ \omega(t)), \quad \omega(t) \in Q, \quad t \in [0, 1] \\ \omega(0) = q_0 \end{cases} \tag{10.4}$$

is defined on $[0,1]$. We denote this solution by $\omega : [0,1] \to Q$. By Lemma 4.3.3, we have $F \circ \omega = \boldsymbol{\psi}^{-1} \circ L_2$. A straightforward computation yields

$$\dot{\omega}(t) = X(\omega(t), \boldsymbol{\xi})$$

for each $t \in [0,1]$. Consider now the curve $\Omega : [0, I_+^X(q_0, \boldsymbol{\xi})) \to Q$ defined by

$$\Omega(t) = \mu^X(t, q_0, \boldsymbol{\xi}).$$

Since $\Omega(0) = q_0$ and

$$\dot{\Omega}(t) = X(\Omega(t), \boldsymbol{\xi})$$

for each $t \in [0, I_+^X(q_0, \boldsymbol{\xi}))$, we see that Ω is a solution of (10.4). Thus

$$\Omega(t) = \omega(t)$$

for each $t \in [0, I_+^X(q_0, \boldsymbol{\xi}))$. By continuity,

$$\begin{aligned}
\lim_{t \nearrow I_+^X(q_0, \boldsymbol{\xi})} \mu^X(t, q_0, \boldsymbol{\xi}) &= \lim_{t \nearrow I_+^X(q_0, \boldsymbol{\xi})} \Omega(t) \\
&= \lim_{t \nearrow I_+^X(q_0, \boldsymbol{\xi})} \omega(t) \\
&= \omega(I_+^X(q_0, \boldsymbol{\xi})).
\end{aligned}$$

Clearly, this contradicts the fact that $\mu^X(\cdot, q_0, \boldsymbol{\xi})$ is the maximally-defined solution of $(X, q_0, \boldsymbol{\xi})$. Hence $I_+^X(q_0, \boldsymbol{\xi}) \notin (0,1)$. This completes the proof. ∎

In summary, Lemma 10.4.3 guarantees for each $q_0 \in W$, the expression

$$\mu^X(-1, q_0, \hat{F}(q_0))$$

is well-defined. In particular, for each $q_0 \in W$, we have

$$\hat{F}(\mu^X(-1, q_0, \hat{F}(q_0))) = \hat{F}(q_0) - \hat{F}(q_0) = \mathbf{0}_{\mathbb{R}^\ell} \tag{10.5}$$

by (10.2). Similarly, Lemma 10.4.4 guarantees that for each

$$(q_0, \boldsymbol{\xi}) \in \hat{F}^{-1}(\mathbf{0}_{\mathbb{R}^\ell}) \times \boldsymbol{\psi}(V),$$

the expression $\mu^X(1, q_0, \boldsymbol{\xi})$ is well-defined. Using these results, we now complete the proof that F is locally trivializable.

Consider the map $A_r^X : \hat{F}^{-1}(\mathbf{0}_{\mathbb{R}^\ell}) \times \boldsymbol{\psi}(V) \to W$, defined by

$$A_r^X(q_0, \boldsymbol{\xi}) = \Phi_1^X(q_0, \boldsymbol{\xi}) = \mu^X(1, q_0, \boldsymbol{\xi}).$$

To see that A_r^X is a homeomorphism, first observe that it is continuous, since it is the composition of $(q_0, \boldsymbol{\xi}) \mapsto (1, q_0, \boldsymbol{\xi})$ and $\Phi^X|\{1\} \times \hat{F}^{-1}(\mathbf{0}_{\mathbb{R}^\ell}) \times \boldsymbol{\psi}(V)$, the latter map being continuous by Theorem 10.3.8. To see that A_r^X is injective, suppose that $A_r^X(q_0, \boldsymbol{\xi}) = A_r^X(\tilde{q}_0, \tilde{\boldsymbol{\xi}})$, so that $\mu^X(1, q_0, \boldsymbol{\xi}) = \mu^X(1, \tilde{q}_0, \tilde{\boldsymbol{\xi}})$. By (10.2),

$$\begin{aligned}
\hat{F}(\mu^X(1, q_0, \boldsymbol{\xi})) = \hat{F}(\mu^X(1, \tilde{q}_0, \tilde{\boldsymbol{\xi}}) &\iff \hat{F}(q_0) + \boldsymbol{\xi} = \hat{F}(\tilde{q}_0) + \tilde{\boldsymbol{\xi}} \\
&\iff \mathbf{0}_{\mathbb{R}^\ell} + \boldsymbol{\xi} = \mathbf{0}_{\mathbb{R}^\ell} + \tilde{\boldsymbol{\xi}} \\
&\iff \boldsymbol{\xi} = \tilde{\boldsymbol{\xi}}.
\end{aligned}$$

Hence $q_0 = \tilde{q}_0$ and A_r^X is injective. To see that A_r^X is surjective, let $q_0 \in W$. By Lemma 10.4.3 and (10.5), $\mu^X(-1, q_0, \hat{F}(q_0)) \in \hat{F}^{-1}(\mathbf{0}_{\mathbb{R}^\ell})$ and

$$A_r^X(\mu^X(-1, q_0, \hat{F}(q_0)), \hat{F}(q_0)) = q_0.$$

Hence A_r^X is surjective. We claim that the inverse of A_r^X is the map

$$B_r^X : W \to \hat{F}^{-1}(\mathbf{0}_{\mathbb{R}^\ell}) \times \boldsymbol{\psi}(V)$$

defined by

$$B_r^X(q_0) = (\Phi_{-1}^X(q_0, \hat{F}(q_0)), \hat{F}(q_0)) = (\mu^X(-1, q_0, \hat{F}(q_0)), \hat{F}(q_0)).$$

Continuity of B_r^X follows from continuity of $\hat{F}$ and $\Phi^X|\{-1\} \times \mathrm{graph}(\hat{F})$, where

$$\mathrm{graph}(\hat{F}) = \{(q_0, \hat{F}(q_0)) \ : \ q_0 \in W\} \subseteq W \times \boldsymbol{\psi}(V)$$

denotes the graph of $\hat{F}$. To see that $B_r^X = (A_r^X)^{-1}$, choose

$$(q_0, \boldsymbol{\xi}) \in \hat{F}^{-1}(\mathbf{0}_{\mathbb{R}^\ell}) \times \boldsymbol{\psi}(V).$$

By (10.2), we have

$$\begin{aligned}
B_r^X \circ A_r^X(q_0, \boldsymbol{\xi}) &= B_r^X(\Phi_1^X(q_0, \boldsymbol{\xi})) \\
&= (\Phi_{-1}^X(\Phi_1^X(q_0, \boldsymbol{\xi})), \hat{F}(\Phi_1^X(q_0, \boldsymbol{\xi}))) \\
&= (q_0, \hat{F}(\Phi_1^X(q_0, \boldsymbol{\xi}))) \\
&= (q_0, \hat{F}(\mu^X(1, q_0, \boldsymbol{\xi}))) \\
&= (q_0, \hat{F}(q_0) + \boldsymbol{\xi}) \\
&= (q_0, \mathbf{0}_{\mathbb{R}^\ell} + \boldsymbol{\xi}) \\
&= (q_0, \boldsymbol{\xi}).
\end{aligned}$$

Hence W is homeomorphic to

$$\hat{F}^{-1}(\mathbf{0}_{\mathbb{R}^\ell}) \times \boldsymbol{\psi}(V) = F^{-1}(r) \times \boldsymbol{\psi}(V).$$

To complete the proof, consider the map $\Psi_r : W \to F^{-1}(r) \times V$ defined by

$$\Psi_r(q_0) = (\mathrm{id} \times \boldsymbol{\psi}^{-1}) \circ B_r^X,$$

where $\mathrm{id} = \mathrm{id}_{F^{-1}(r)}$. Clearly, Ψ_r is a homeomorphism, and

$$\begin{aligned}
\mathrm{pr}_2 \circ \Psi_r(q_0) &= \mathrm{pr}_2 \circ (\mathrm{id} \times \boldsymbol{\psi}^{-1}) \circ B_r^X(q_0) \\
&= \mathrm{pr}_2 \circ (\mathrm{id} \times \boldsymbol{\psi}^{-1})(\Phi_{-1}^X(q_0, \hat{F}(q_0)), \hat{F}(q_0)) \\
&= \mathrm{pr}_2 \circ (\Phi_{-1}^X(q_0, \hat{F}(q_0)), \boldsymbol{\psi}^{-1} \circ \hat{F}(q_0)) \\
&= \mathrm{pr}_2 \circ (\Phi_{-1}^X(q_0, \hat{F}(q_0)), F(q_0)) \\
&= F(q_0)
\end{aligned}$$

for each $q_0 \in W$. In other words, $\mathrm{pr}_2 \circ \Psi_r = F|W = F|F^{-1}(V)$. Since r was chosen arbitrarily, we conclude that F is locally trivializable. This completes the proof. ∎

Corollary 10.4.5. *Suppose that F is PLE-complete for line segments. Then F is a fiber bundle.*

Proof. This is clear from the proof of Theorem 10.4.2, in which it was only necessary to consider (π, q_0)-PLEs, where π is a line segment. ∎

The next results are C^k versions of Theorem 10.4.2 and Corollary 10.4.5.

Theorem 10.4.6. *Suppose that F is C^k, where $k \in \mathbb{N}^*$, $TF^\dagger$ is a C^k right inverse of TF, and F is PLE-complete. Then F is a C^k fiber bundle.*

Proof. We freely use the notation established in the proof of Theorem 10.4.2. First of all, since F is a C^k submersion, $F^{-1}(r)$ is a C^k submanifold of Q; see, for example, [Margalef-Roig and Outerelo Domínguez 1992, Theorem 4.2.1]. Note also that $\operatorname{graph}(\hat{F})$ is a C^k submanifold of $W \times \boldsymbol{\psi}(V)$ by [Margalef-Roig and Outerelo Domínguez 1992, Proposition 3.3.10]. If we can show that X is C^k, then the proof will be complete. Indeed, if X is C^k, then its global flow Φ^X is C^k by Theorem 10.3.8. Thus the restrictions of Φ^X to the C^k submanifolds

$$\{1\} \times F^{-1}(r) \times \boldsymbol{\psi}(V) \qquad \text{and} \qquad \{-1\} \times \operatorname{graph}(\hat{F})$$

of $\mathbb{R} \times W \times \mathbb{R}^\ell$ are also C^k. Inspecting the proof of Theorem 10.4.2, we see that this implies that Ψ_r is a C^k diffeomorphism, and hence F is a C^k fiber bundle.

Now choose $(U, \varphi) \in \mathscr{A}_W$, where $\mathscr{A}_W$ is the compatible atlas on W defined as in Lemma 10.4.1. By construction, $F(U) \subseteq V$. Furthermore, since W is an open submanifold of Q, (U, φ) is also a chart on Q. Thus the map

$$TF^\dagger_{\varphi,\boldsymbol{\psi}} : \varphi(U) \to \operatorname{Hom}(\mathbb{R}^\ell, E_Q),$$

defined as in Proposition 4.1.4, is C^k. As in Lemma 10.4.1, we have

$$X_\varphi(q, \boldsymbol{\xi}) = TF^\dagger_{\varphi,\boldsymbol{\psi}}(q) \cdot \boldsymbol{\xi}$$

for each $(q, \boldsymbol{\xi}) \in \varphi(U) \times \mathbb{R}^\ell$. By the Leibniz rule, X_φ is C^k. Since (U, φ) was chosen arbitrarily, we conclude that X is C^k. This completes the proof. ∎

Corollary 10.4.7. *Suppose that F is C^k, where $k \in \mathbb{N}^*$, $TF^\dagger$ is a C^k right inverse of TF, and F is PLE-complete for line segments. Then F is a C^k fiber bundle.*

We will see a concrete application of Corollary 10.4.7 in Chapter 11.

10.5 Unobstructed control systems

We now apply the results of Section 10.4 to control systems. In this section,

- M is a second-countable n-dimensional manifold,
- $\Sigma = (f, \mathscr{U})$ is a control system evolving on M,
- The time domain of Σ is $J = [a, b]$,
- Σ uses L^p controls, so that $\mathscr{U} = L^p(J, \mathbb{R}^r)$,
- Σ is complete, and
- Σ is completely controllable from a fixed initial state $x_0 \in M$ on J.

The next definition codifies what it means for a control system to be unobstructed with respect to the x_0-anchored motion planning problem.

Definition 10.5.1. We say that Σ is C^k ***unobstructed***, where $k \in \mathbb{N}^*$, if

- $\mathrm{End}^\Sigma_{x_0} : \mathscr{U} \to M$ is C^k,
- $\mathscr{U}^{\mathrm{sing}}_{x_0} = \varnothing$, and
- $\mathrm{End}^\Sigma_{x_0}$ is PLE-complete relative to $T\mathrm{End}^{\Sigma,\dagger}_{x_0}$,

where $T\mathrm{End}^{\Sigma,\dagger}_{x_0}$ is a C^k right inverse of $T\mathrm{End}^\Sigma_{x_0}$.

The next two results seem to indicate that unobstructed control systems are quite exceptional within the class of all control systems, in the sense that they are distinguished by their topological structure. In prior work on the continuation method, this point was not addressed.

Theorem 10.5.2. *Suppose that* Σ *is* C^k *unobstructed, where* $k \in \mathbb{N}^*$*. Then* $\mathrm{End}_{x_0}^{\Sigma}$ *is a* C^k *fiber bundle.*

Proof. This is simply an application of Theorem 10.4.6. ∎

Corollary 10.5.3. *Suppose that* Σ *uses* L^2 *controls,*

- $\mathrm{End}_{x_0}^{\Sigma} : \mathscr{U} \to M$ *is* C^{k+1}*, where* $k \in \mathbb{N}^*$*,*
- $\mathscr{U}_{x_0}^{\mathrm{sing}} = \varnothing$*, and*
- *For each compact subset* K *of* M*, the Moore–Penrose pseudoinverse of* $T\mathrm{End}_{x_0}^{\Sigma}$ *has sublinear growth over* K*.*

Then $\mathrm{End}_{x_0}^{\Sigma}$ *is a* C^k *fiber bundle.*

Proof. By Lemma 9.3.2, $\mathrm{End}_{x_0}^{\Sigma}$ is PLE-complete relative to the Moore–Penrose pseudoinverse of $T\mathrm{End}_{x_0}^{\Sigma}$. Since $\mathrm{End}_{x_0}^{\Sigma}$ is C^{k+1}, the Moore–Penrose pseudoinverse of $T\mathrm{End}_{x_0}^{\Sigma}$ is C^k by Proposition 4.1.6. Thus Σ is C^k unobstructed. ∎

To conclude this chapter, let us remark that the results in this section are modest contributions to an emerging literature [Vakhrameev 1991a, Zhong 1993, Vakhrameev 1996, Kızıl 2008, Dominy and Rabitz 2011] on the topological structure of endpoint maps. Within this literature, our results are distinguished by the topological structure obtained, in that we obtain fiber bundles and not fibrations. This is interesting, since fiber bundles constitute a special class of fibrations.

Chapter 11

Examples

In previous chapters, we presented a number of theoretical results concerning the three obstructions to the continuation method. In this chapter, we illustrate several of these results by applying them to academic control-affine systems.

11.1 Hirschorn's system

Here we consider the control-affine system from [Hirschorn 1990, Example 2.5]. Although it meets the criteria of Corollary 8.2.7 for an appropriate choice of initial state, this control system does not satisfy local or global finite definiteness conditions with degree $\Delta = 0$. Consequently, the results of Chapter 9 do not apply. Despite this limitation, the analysis in this section is interesting in its own right, as it demonstrates that an underactuated control-affine system can give rise to a submersive anchored endpoint map. This phenomenon has not been previously reported in the literature, to the best of our knowledge.

We begin by setting $M = \mathbb{R}_{>0} \times \mathbb{R}$. In what follows, we consider M as an open Riemannian submanifold of $\mathbb{R}^2$. A generic element of M is written

$$\boldsymbol{x} = \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}.$$

Consider the C^ω control-affine system $\Sigma = (\boldsymbol{f}, \mathscr{U})$ evolving on M, where

- The controllable time-varying vector field $\boldsymbol{f} \in \mathscr{V}([0,1], M, \mathbb{R})$ is defined by

$$\boldsymbol{f}(t, \boldsymbol{x}, \omega) = \boldsymbol{f}_0(\boldsymbol{x}) + \omega \boldsymbol{f}_1(\boldsymbol{x}) = \begin{pmatrix} 0 \\ \ln(x^1) \end{pmatrix} + \omega \begin{pmatrix} x^1 \\ 0 \end{pmatrix}$$

  and

- $\mathscr{U} = L^2([0,1], \mathbb{R})$.

We claim that Σ is complete. To see this, observe that for each $\boldsymbol{x} \in M$ and each $u \in \mathscr{U}$, the u-controlled trajectory of Σ with initial condition $(0, \boldsymbol{x})$ is

$$\boldsymbol{\mu}^\Sigma(t, 0, \boldsymbol{x}, u) = \begin{pmatrix} x^1 \mathrm{e}^{\int_0^t u(\sigma)\,\mathrm{d}\sigma} \\ x^2 + \int_0^t \ln\left(x^1 \mathrm{e}^{\int_0^\tau u(\sigma)\,\mathrm{d}\sigma}\right)\,\mathrm{d}\tau \end{pmatrix}.$$

This is clearly well-defined for each $t \in [0,1]$. Using the results of Hirschorn [1990], we see that Σ is completely controllable from each $\boldsymbol{x} \in M$ on $[0,1]$.[1]

Our next objective is to show that $\mathrm{End}^\Sigma_{\boldsymbol{x}_0}$ is a submersion, where

$$\boldsymbol{x}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

To this end, we use the results of Chapter 8. A straightforward computation shows that $[\boldsymbol{f}_0, [\boldsymbol{f}_0, \boldsymbol{f}_1]]$ is identically equal to $\boldsymbol{0}_{\mathbb{R}^2}$, so that Σ satisfies the global finite definiteness condition with degree $\Delta = 1$. Furthermore, $[\boldsymbol{f}_1, \boldsymbol{f}_1]$ and $[\boldsymbol{f}_1, [\boldsymbol{f}_0, \boldsymbol{f}_1]]$ are identically equal to $\boldsymbol{0}_{\mathbb{R}^2}$. This implies that Σ satisfies the global bang-bang condition. Now let $u_0 \in \mathscr{U}$ be the function

$$u_0(t) = t.$$

Recall from Section 8.4 that for each $u \in \mathscr{U}$, the map

$$\lambda^1_{u_0} \cdot u \in L^1([0,1], T_{\boldsymbol{x}_0} M)$$

[1] In fact, a much stronger conclusion holds, namely that Σ is strongly controllable from $\boldsymbol{x}$. Roughly speaking, this means that Σ is completely controllable from $\boldsymbol{x}$ on $[0, \varepsilon]$ for each $\varepsilon \in \mathbb{R}_{>0}$.

is defined by

$$\begin{aligned}(\lambda^1_{u_0}\cdot u)(t) &= \mathrm{Ad}_M^{f^{u_0}}(\boldsymbol{D}_3 f^u_{u_0})(t,\boldsymbol{x}_0)\\ &= T\Phi^{\boldsymbol{f}^{u_0}}_{0,t}(\boldsymbol{\mu}^{u_0}(t))\circ \boldsymbol{D}_3\boldsymbol{f}(t,\boldsymbol{\mu}^{u_0}(t),u_0(t))\cdot u(t),\end{aligned}$$

where

$$\boldsymbol{\mu}^{u_0}(t)=\boldsymbol{\mu}^{\Sigma}(t,0,\boldsymbol{x}_0,u_0)=\begin{pmatrix}\mathrm{e}^{\frac{t^2}{2}}\\ \frac{t^3}{6}\end{pmatrix}.$$

To evaluate this map, first note that

$$\Phi^{\boldsymbol{f}^{u_0}}_{t,0}(\boldsymbol{x})=\begin{pmatrix}x^1\mathrm{e}^{\frac{t^2}{2}}\\ x^2+t\ln(x^1)+\frac{t^3}{6}\end{pmatrix}$$

and thus

$$T\Phi^{\boldsymbol{f}^{u_0}}_{0,t}(\boldsymbol{x})=\begin{pmatrix}\mathrm{e}^{-\frac{t^2}{2}} & 0\\ -\frac{t\mathrm{e}^{-\frac{t^2}{2}}}{x^1} & 1\end{pmatrix}$$

for each $t\in[0,1]$ and each $\boldsymbol{x}\in M$. In particular, this means that

$$T\Phi^{\boldsymbol{f}^{u_0}}_{0,t}(\boldsymbol{\mu}^{u_0}(t))=\begin{pmatrix}\mathrm{e}^{-\frac{t^2}{2}} & 0\\ -t\mathrm{e}^{-t^2} & 1\end{pmatrix}.$$

Since Σ is control-affine, it is clear that

$$\boldsymbol{D}_3\boldsymbol{f}(t,\boldsymbol{\mu}^{u_0}(t),u_0(t))\cdot u(t)=\begin{pmatrix}u(t)\mathrm{e}^{\frac{t^2}{2}}\\ 0\end{pmatrix}$$

and consequently

$$(\lambda^1_{u_0}\cdot u)(t)=\begin{pmatrix}\mathrm{e}^{-\frac{t^2}{2}} & 0\\ -t\mathrm{e}^{-\frac{t^2}{2}} & 1\end{pmatrix}\cdot\begin{pmatrix}u(t)\mathrm{e}^{\frac{t^2}{2}}\\ 0\end{pmatrix}=\begin{pmatrix}u(t)\\ -u(t)t\mathrm{e}^{-\frac{t^2}{2}}\end{pmatrix}.$$

Note that if u is continuous, then $\lambda^1_{u_0}\cdot u$ is continuous, and thus each $t\in(0,1)$ is a symmetric Lebesgue point of $\lambda^1_{u_0}$. In other words, $\mathrm{Leb}(\lambda^{u_0})=(0,1)$. By Theorem 8.4.3, the first-order Pontryagin cone along μ^{u_0} is

$$\mathsf{PC}^{\Sigma}_{\boldsymbol{x}_0}(u_0)=\mathrm{span}\{(\lambda^1_{u_0}\cdot u)(t)\;:\;u\in\mathscr{U},\,t\in\mathrm{Leb}(\lambda^1_{u_0}\cdot u)\}$$

$$= \operatorname{span}\left\{ \begin{pmatrix} u(t) \\ -u(t)t\mathrm{e}^{-\frac{t^2}{2}} \end{pmatrix} \;:\; u \in \mathscr{U},\, t \in \operatorname{Leb}(\lambda^1_{u_0} \cdot u) \right\}.$$

Choosing $u \equiv 1$ and $t = 1/2$, we see that

$$\begin{pmatrix} 1 \\ -\frac{\mathrm{e}^{-\frac{1}{8}}}{2} \end{pmatrix} \in \mathsf{PC}^{\Sigma}_{\boldsymbol{x}_0}(u_0).$$

Now let $\{t_n\}_{n\in\mathbb{N}}$ be a sequence of real numbers in $(0, 1)$ such that $t_n \to 0$ as $n \to \infty$. Choosing $u \equiv 1$, we see that the sequence of vectors

$$\left\{ \begin{pmatrix} 1 \\ -t_n\mathrm{e}^{-\frac{t_n^2}{2}} \end{pmatrix} \right\}_{n\in\mathbb{N}}$$

in $\mathsf{PC}^{\Sigma}_{\boldsymbol{x}_0}(u_0)$ satisfies

$$\begin{pmatrix} 1 \\ -t_n\mathrm{e}^{-\frac{t_n^2}{2}} \end{pmatrix} \to \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

as $n \to \infty$. Since $\mathsf{PC}^{\Sigma}_{\boldsymbol{x}_0}(u_0)$ is closed,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathsf{PC}^{\Sigma}_{\boldsymbol{x}_0}(u_0).$$

The fact that

$$\begin{pmatrix} 1 \\ -\frac{\mathrm{e}^{-\frac{1}{8}}}{2} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

are linearly independent yields

$$\mathsf{PC}^{\Sigma}_{\boldsymbol{x}_0}(u_0) = \mathbb{R}^2.$$

Invoking Corollary 8.2.7, we conclude that $\operatorname{End}^{\Sigma}_{\boldsymbol{x}_0}$ is a C^2 submersion.

Finally, observe that $[\boldsymbol{f}_0, \boldsymbol{f}_1]$ is identically equal to

$$\begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

It follows that Σ cannot satisfy finite definiteness conditions with degree $\Delta = 0$, as $[\boldsymbol{f}_0, \boldsymbol{f}_1](\boldsymbol{x})$ cannot be written as a scalar multiple of $\boldsymbol{f}_1(\boldsymbol{x})$ for any $\boldsymbol{x} \in M$.

11.2 An augmented version of Hirschorn's system

Here we consider a control-affine system Σ obtained by augmenting Hirschorn's system with an additional control vector field. One effect of the additional control vector field is that Σ satisfies local finite definiteness conditions with degree $\Delta = 0$.

Consider the C^ω control-affine system $\Sigma = (\boldsymbol{f}, \mathscr{U})$ evolving on M, where

- M is the Riemannian manifold $\mathbb{R}_{>0} \times \mathbb{R}$, as in Section 11.1,
- The controllable time-varying vector field $\boldsymbol{f} \in \mathscr{V}([0,1], M, \mathbb{R}^2)$ is defined by

$$\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{\omega}) = \boldsymbol{f}_0(\boldsymbol{x}) + \omega^1 \boldsymbol{f}_1(\boldsymbol{x}) + \omega^2 \boldsymbol{f}_2(\boldsymbol{x}) = \begin{pmatrix} 0 \\ \ln(x^1) \end{pmatrix} + \omega^1 \begin{pmatrix} x^1 \\ 0 \end{pmatrix} + \omega^2 \begin{pmatrix} 0 \\ x^2 \end{pmatrix},$$

  and

- $\mathscr{U} = L^2([0,1], \mathbb{R}^2)$.

Note that $\operatorname{span}\{\boldsymbol{f}_1(\boldsymbol{x}), \boldsymbol{f}_2(\boldsymbol{x})\} = \mathbb{R}^2$ if and only if $x^2 \neq 0$.

We claim that Σ is complete. To see this, observe that for each $\boldsymbol{x} \in M$ and each $\boldsymbol{u} \in \mathscr{U}$, the $\boldsymbol{u}$-controlled trajectory of Σ with initial condition $(0, \boldsymbol{x})$ is

$$\boldsymbol{\mu}^\Sigma(t, 0, \boldsymbol{x}, \boldsymbol{u}) = \begin{pmatrix} x^1 \mathrm{e}^{\int_0^t u^1(\sigma)\,\mathrm{d}\sigma} \\ \mathrm{e}^{\int_0^t u^2(\sigma)\,\mathrm{d}\sigma} \left(x^2 + \int_0^t \ln\left(x^1 \mathrm{e}^{\int_0^\tau u^1(\sigma)\,\mathrm{d}\sigma} \right) \mathrm{e}^{-\int_0^\tau u^2(\tau)\,\mathrm{d}\tau}\,\mathrm{d}\tau \right) \end{pmatrix}.$$

This is clearly well-defined for each $t \in [0,1]$. Furthermore, it follows immediately from Section 11.1 that Σ is completely controllable from each $\boldsymbol{x} \in M$ on $[0,1]$.

Our next objective is to show that $\mathrm{End}^\Sigma_{\boldsymbol{x}_0}$ is a submersion, where

$$\boldsymbol{x}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

We proceed as in the previous section. Straightforward computations show that $[\boldsymbol{f}_0, [\boldsymbol{f}_0, \boldsymbol{f}_1]]$ and $[\boldsymbol{f}_0, [\boldsymbol{f}_0, \boldsymbol{f}_2]]$ are identically equal to $\boldsymbol{0}_{\mathbb{R}^2}$, so that Σ satisfies the global finite definiteness condition with degree $\Delta = 1$. Furthermore,

- $[\boldsymbol{f}_1, \boldsymbol{f}_1]$, $[\boldsymbol{f}_2, \boldsymbol{f}_2]$, and $[\boldsymbol{f}_1, \boldsymbol{f}_2]$ are identically equal to $\mathbf{0}_{\mathbb{R}^2}$,
- $[\boldsymbol{f}_1, [\boldsymbol{f}_0, \boldsymbol{f}_1]]$ is identically equal to $\mathbf{0}_{\mathbb{R}^2}$,
- $[\boldsymbol{f}_1, [\boldsymbol{f}_0, \boldsymbol{f}_2]] = -[\boldsymbol{f}_0, \boldsymbol{f}_1]$,
- $[\boldsymbol{f}_2, [\boldsymbol{f}_0, \boldsymbol{f}_1]] = -[\boldsymbol{f}_0, \boldsymbol{f}_1]$, and
- $[\boldsymbol{f}_2, [\boldsymbol{f}_0, \boldsymbol{f}_2]] = -[\boldsymbol{f}_0, \boldsymbol{f}_2]$.

This implies that Σ satisfies the global bang-bang condition. Now let $\boldsymbol{u}_0 \in \mathscr{U}$ be the map

$$\boldsymbol{u}_0(t) = \begin{pmatrix} t \\ 0 \end{pmatrix}.$$

For each $\boldsymbol{u} \in \mathscr{U}$, consider the map $\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u} \in L^1([0,1], T_{\boldsymbol{x}_0}M)$ that sends t to

$$(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(t) = T\Phi^{\boldsymbol{f}^{\boldsymbol{u}_0}}_{0,t}(\boldsymbol{\mu}^{\boldsymbol{u}_0}(t)) \circ \boldsymbol{D}_3\boldsymbol{f}(t, \boldsymbol{\mu}^{\boldsymbol{u}_0}(t), \boldsymbol{u}_0(t)) \cdot \boldsymbol{u}(t),$$

where

$$\boldsymbol{\mu}^{\boldsymbol{u}_0}(t) = \boldsymbol{\mu}^{\Sigma}(t, 0, \boldsymbol{x}_0, \boldsymbol{u}_0) = \begin{pmatrix} \mathrm{e}^{\frac{t^2}{2}} \\ \frac{t^3}{6} \end{pmatrix}.$$

To evaluate this map, first note that

$$\Phi^{\boldsymbol{f}^{\boldsymbol{u}_0}}_{t,0}(\boldsymbol{x}) = \begin{pmatrix} x^1\mathrm{e}^{\frac{t^2}{2}} \\ x^2 + t\ln(x^1) + \frac{t^3}{6} \end{pmatrix}$$

and thus

$$T\Phi^{\boldsymbol{f}^{\boldsymbol{u}_0}}_{0,t}(\boldsymbol{x}) = \begin{pmatrix} \mathrm{e}^{-\frac{t^2}{2}} & 0 \\ -\frac{t\mathrm{e}^{-\frac{t^2}{2}}}{x^1} & 1 \end{pmatrix}$$

for each $t \in [0,1]$ and each $\boldsymbol{x} \in M$. In particular, this means that

$$T\Phi^{\boldsymbol{f}^{\boldsymbol{u}_0}}_{0,t}(\boldsymbol{\mu}^{\boldsymbol{u}_0}(t)) = \begin{pmatrix} \mathrm{e}^{-\frac{t^2}{2}} & 0 \\ -t\mathrm{e}^{-t^2} & 1 \end{pmatrix}.$$

Since Σ is control-affine, it is clear that

$$\boldsymbol{D}_3\boldsymbol{f}(t,\boldsymbol{\mu}^{\boldsymbol{u}_0}(t),\boldsymbol{u}_0(t))\cdot\boldsymbol{u}(t)=\begin{pmatrix}\mathrm{e}^{\frac{t^2}{2}} & 0\\ 0 & \frac{t^3}{6}\end{pmatrix}\cdot\begin{pmatrix}u^1(t)\\ u^2(t)\end{pmatrix}=\begin{pmatrix}u^1(t)\mathrm{e}^{\frac{t^2}{2}}\\ u^2(t)\frac{t^3}{6}\end{pmatrix}$$

and consequently

$$(\lambda^1_{\boldsymbol{u}_0}\cdot\boldsymbol{u})(t)=\begin{pmatrix}\mathrm{e}^{-\frac{t^2}{2}} & 0\\ -t\mathrm{e}^{-t^2} & 1\end{pmatrix}\cdot\begin{pmatrix}u^1(t)\mathrm{e}^{\frac{t^2}{2}}\\ u^2(t)\frac{t^3}{6}\end{pmatrix}=\begin{pmatrix}u^1(t)\\ -u^1(t)t\mathrm{e}^{-\frac{t^2}{2}}\end{pmatrix}+\begin{pmatrix}0\\ u^2(t)\frac{t^3}{6}\end{pmatrix}.$$

Note that if $\boldsymbol{u}$ is continuous, then $\lambda^1_{\boldsymbol{u}_0}\cdot\boldsymbol{u}$ is continuous and thus $\mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0}\cdot\boldsymbol{u})=(0,1)$. By Theorem 8.4.3, the first-order Pontryagin cone along $\mu^{\boldsymbol{u}_0}$ is

$$\begin{aligned}\mathsf{PC}^{\Sigma}_{\boldsymbol{x}_0}(\boldsymbol{u}_0)&=\mathrm{span}\{(\lambda^1_{\boldsymbol{u}_0}\cdot\boldsymbol{u})(t)\ :\ \boldsymbol{u}\in\mathscr{U},\,t\in\mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0}\cdot\boldsymbol{u})\}\\ &=\mathrm{span}\left\{\begin{pmatrix}u^1(t)\\ -u^1(t)t\mathrm{e}^{-\frac{t^2}{2}}\end{pmatrix}+\begin{pmatrix}0\\ u^2(t)\frac{t^3}{6}\end{pmatrix}\ :\ \boldsymbol{u}\in\mathscr{U},\,t\in\mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0}\cdot\boldsymbol{u})\right\}\end{aligned}$$

Choosing $t=1/2$,

$$\boldsymbol{u}\equiv\begin{pmatrix}1\\0\end{pmatrix},\qquad\text{and}\qquad\boldsymbol{u}\equiv\begin{pmatrix}0\\1\end{pmatrix},$$

we see that

$$\begin{pmatrix}1\\ -\frac{\mathrm{e}^{-\frac{1}{8}}}{2}\end{pmatrix},\begin{pmatrix}0\\ \frac{1}{48}\end{pmatrix}\in\mathsf{PC}^{\Sigma}_{\boldsymbol{x}_0}(\boldsymbol{u}_0).$$

The fact that

$$\begin{pmatrix}1\\ -\frac{\mathrm{e}^{-\frac{1}{8}}}{2}\end{pmatrix}\qquad\text{and}\qquad\begin{pmatrix}0\\ \frac{1}{48}\end{pmatrix}$$

are linearly independent yields

$$\mathsf{PC}^{\Sigma}_{\boldsymbol{x}_0}(\boldsymbol{u}_0)=\mathbb{R}^2.$$

Invoking Corollary 8.2.7, we conclude that $\mathrm{End}^{\Sigma}_{\boldsymbol{x}_0}$ is a C^2 submersion.

With respect to finite definiteness conditions with degree $\Delta=0$, observe that

- $[\boldsymbol{f}_0, \boldsymbol{f}_1](\boldsymbol{x}) = \binom{0}{-1}$ and
- $[\boldsymbol{f}_0, \boldsymbol{f}_2](\boldsymbol{x}) = \binom{0}{\ln(x^1)}$

for each $\boldsymbol{x} \in M$. Define $\mathscr{Z} = \{\boldsymbol{x} \in M \ : \ x^2 = 0\}$. Clearly, $\mathscr{Z}$ is closed in M and

$$[\boldsymbol{f}_0, \boldsymbol{f}_1](\boldsymbol{x}) = -\frac{1}{x^2}\boldsymbol{f}_2(\boldsymbol{x}) \qquad \text{and} \qquad [\boldsymbol{f}_0, \boldsymbol{f}_2](\boldsymbol{x}) = \frac{\ln(x^1)}{x^2}\boldsymbol{f}_2(\boldsymbol{x})$$

for each $\boldsymbol{x} \in M \setminus \mathscr{Z}$. If K is a nonempty compact subset of $M \setminus \mathscr{Z}$, then $\boldsymbol{x}_0 \notin K$ by construction. Furthermore, Σ satisfies the local finite definiteness condition with degree $\Delta = 0$ on K. Using this observation, we have the following lemma.

Lemma 11.2.1. *Suppose that K is a nonempty compact subset of $M \setminus \mathscr{Z}$. Then the Moore–Penrose pseudoinverse of $T\mathrm{End}^{\Sigma}_{\boldsymbol{x}_0}$ has sublinear growth over K.*

Proof. Choose $\delta \in \mathbb{R}_{>0}$ such that $B_{\leq\delta}(K) \subseteq M \setminus \mathscr{Z}$ and the conclusions of Lemma 9.5.10 are satisfied. By construction, each $\boldsymbol{x} \in B_{\leq\delta}(K)$ is such that $x^2 \neq 0$. Thus $\boldsymbol{f}_1(\boldsymbol{x})$ and $\boldsymbol{f}_2(\boldsymbol{x})$ are linearly independent for each $\boldsymbol{x} \in B_{\leq\delta}(K)$. It follows[2] that

$$\inf_{(\boldsymbol{x},\boldsymbol{p}) \in B_{\leq\delta}(K)^*} \|\boldsymbol{\varphi}^{\Sigma}(\boldsymbol{x}, \boldsymbol{p})\|_{\mathbb{R}^2} > 0. \tag{11.1}$$

By Remark 9.5.4, each C^1 function $G : T^*M \to \mathbb{R}$ satisfies the Υ-condition over $B_{\leq\delta}(K)$. Thus the hypotheses of Theorem 9.5.12 are satisfied. Invoking Theorem 9.5.12, we conclude that the Moore–Penrose pseudoinverse of $T\mathrm{End}^{\Sigma}_{\boldsymbol{x}_0}$ has sublinear growth over K. ∎

Remark 11.2.2. This preceding proof illustrates how the hypotheses of Theorem 9.5.12 can be verified in practice. However, with respect to asserting the existence of a bump function that satisfies the conditions of Theorem 9.5.12, *the situation encountered in the proof constitutes the simplest possible case*—that is, the case where $\boldsymbol{f}_1(\boldsymbol{x})$ and $\boldsymbol{f}_2(\boldsymbol{x})$ are linearly independent for each $\boldsymbol{x} \in B_{\leq\delta}(K)$. In other cases, one must use the full strength of the Υ-condition, which does not require linearly independent control vector fields.

[2] Here we are identifying T^*M with the product space $M \times \mathbb{R}^2$.

In the next section, we move away from specific examples.

11.3 A special class of control-affine systems

Consider the C^ω control-affine system $\Sigma = (f, \mathscr{U})$ evolving on M, where

- M is a connected, compact, n-dimensional, C^ω Riemannian manifold,
- The controllable time-dependent vector field $f \in \mathscr{V}([0,1], M, \mathbb{R}^n)$ is given by

$$f(t, x, \boldsymbol{\omega}) = f_0(x) + \sum_{i=1}^{n} \omega^i f_i(x)$$

  where $f_1, \ldots, f_n$ constitute a global frame[3] for M, and
- $\mathscr{U} = L^2([0,1], \mathbb{R}^n)$.

We further assume that the fundamental group of M does not contain any elements of infinite order; for definitions, we refer to [Lee 2000, Chapter 7]. For example, this topological restriction is satisfied whenever M is the k-dimensional sphere $\mathbb{S}^k$, where $k \in \{1, 3, 7\}$.

Since M is compact, Σ is complete; this follows from an argument along the lines of [Lee 2003, Lemma 17.10]. Using the well-known fact that an arbitrary family of C^ω vector fields is Lie-determined, it follows from [Jurdjevic 1997, Chapter 4, Theorem 2] and [Jurdjevic 1997, Chapter 3, Theorem 13] that Σ is completely controllable from each $x \in M$ on $[0,1]$. To clarify the role played by the restriction on the fundamental group of M, we note that this restriction is an essential hypothesis of [Jurdjevic 1997, Chapter 3, Theorem 13].

We now show that $\mathrm{End}^{\Sigma}_{x_0}$ is a submersion, where the initial state $x_0 \in M$ is chosen arbitrarily. Since $f_1, \ldots, f_n$ constitute a global frame for M, it is clear that Σ satisfies the global finite definiteness condition with degree $\Delta = 0$. Furthermore, Σ

[3]This means that $\operatorname{span}\{f_1(x), \ldots, f_n(x)\} = T_xM$ for each $x \in M$. By definition, M is parallelizable if and only if it admits a global frame; see [Lee 2003, Chapter 5] for more information.

satisfies the global bang-bang condition. Now choose an arbitrary control $\boldsymbol{u}_0 \in \mathscr{U}$. For each $\boldsymbol{u} \in \mathscr{U}$, consider the map $\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u} \in L^1([0,1], T_{x_0}M)$ that sends t to

$$(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(t) = T\Phi^{f^{\boldsymbol{u}_0}}_{0,t}(\mu^{\boldsymbol{u}_0}(t)) \circ \boldsymbol{D}_3\boldsymbol{f}(t, \mu^{\boldsymbol{u}_0}(t), \boldsymbol{u}_0(t)) \cdot \boldsymbol{u}(t),$$

where

$$\mu^{\boldsymbol{u}_0}(t) = \mu^{\Sigma}(t, 0, x_0, \boldsymbol{u}_0).$$

To evaluate this map, first observe that

$$\boldsymbol{D}_3 f(t, \mu^{\boldsymbol{u}_0}(t), \boldsymbol{u}_0(t)) \cdot \boldsymbol{u}(t) = \sum_{i=1}^{n} u^i(t) f_i(\mu^{\boldsymbol{u}_0}(t))$$

and

$$\begin{aligned}(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(t) &= T\Phi^{f^{\boldsymbol{u}_0}}_{0,t}(\mu^{\boldsymbol{u}_0}(t)) \circ \boldsymbol{D}_3\boldsymbol{f}(t, \mu^{\boldsymbol{u}_0}(t), \boldsymbol{u}_0(t)) \cdot \boldsymbol{u}(t) \\ &= T\Phi^{f^{\boldsymbol{u}_0}}_{0,t}(\mu^{\boldsymbol{u}_0}(t)) \cdot \sum_{i=1}^{n} u^i(t) f_i(\mu^{\boldsymbol{u}_0}(t)) \\ &= \Phi^{\mathrm{tlft}(f^{\boldsymbol{u}_0})}_{0,t}\left(\sum_{i=1}^{n} u^i(t) f_i(\mu^{\boldsymbol{u}_0}(t))\right) \\ &= \sum_{i=1}^{n} u^i(t) \Phi^{\mathrm{tlft}(f^{\boldsymbol{u}_0})}_{0,t}(f_i(\mu^{\boldsymbol{u}_0}(t))).\end{aligned}$$

Note that if $\boldsymbol{u}$ is continuous, then $\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u}$ is continuous and thus $\mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u}) = (0,1)$. By Theorem 8.4.3, the first-order Pontryagin cone along $\mu^{\boldsymbol{u}_0}$ is

$$\begin{aligned}&\mathsf{PC}^{\Sigma}_{x_0}(\boldsymbol{u}_0) \\ &= \mathrm{span}\{(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})(t) \ : \ \boldsymbol{u} \in \mathscr{U}, \, t \in \mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})\} \\ &= \mathrm{span}\left\{T\Phi^{f^{\boldsymbol{u}_0}}_{0,t}(\mu^{\boldsymbol{u}_0}(t)) \cdot \sum_{i=1}^{n} u^i(t) f_i(\mu^{\boldsymbol{u}_0}(t)) \ : \ \boldsymbol{u} \in \mathscr{U}, \, t \in \mathrm{Leb}(\lambda^1_{\boldsymbol{u}_0} \cdot \boldsymbol{u})\right\}.\end{aligned}$$

Choosing $\boldsymbol{u} \equiv \boldsymbol{e}_1, \dots, \boldsymbol{e}_n$, where $\boldsymbol{e}_i$ is the ith standard basis vector, we see that

$$T\Phi^{f^{\boldsymbol{u}_0}}_{0,t}(\mu^{\boldsymbol{u}_0}(t)) \cdot f_i(\mu^{\boldsymbol{u}_0}(t)) \in \mathsf{PC}^{\Sigma}_{x_0}(\boldsymbol{u}_0)$$

for each $t \in (0,1)$ and each $1 \leq i \leq n$. Since each $\Phi^{f^{\boldsymbol{u}_0}}_{0,t}$ is a C^∞ diffeomorphism,

$$\mathsf{PC}^{\Sigma}_{x_0}(\boldsymbol{u}_0) = T_{x_0}M.$$

Invoking Corollary 8.2.7, we conclude that $\mathrm{End}^{\Sigma}_{x_0}$ is a C^2 submersion.

We have the following analogue of Lemma 11.2.1.

Lemma 11.3.1. *Suppose that K is a nonempty compact subset of M such that $x_0 \notin K$. Then the Moore–Penrose pseudoinverse of $T\mathrm{End}^{\Sigma}_{x_0}$ has sublinear growth over K.*

Proof. Identical to the proof of Lemma 11.2.1. ∎

We now briefly consider topological ramifications, using the results derived in Chapter 10. Consider the map

$$\mathfrak{E} : \mathrm{dom}(\mathfrak{E}) = \mathscr{U} \setminus (\mathrm{End}^{\Sigma}_{x_0})^{-1}(x_0) \to M \setminus \{x_0\}$$

obtained from $\mathrm{End}^{\Sigma}_{x_0}$ by restriction. That is,

$$\mathfrak{E}(\boldsymbol{u}) = \mathrm{End}^{\Sigma}_{x_0}(\boldsymbol{u})$$

for each $\boldsymbol{u} \in \mathrm{dom}(\mathfrak{E})$. By construction, $\mathfrak{E}$ is a surjective C^2 submersion. It follows from Lemma 9.3.2 and Lemma 11.3.1 that $\mathfrak{E}$ is PLE-complete relative to $T\mathfrak{E}^{\#}$, which is C^1. Invoking Theorem 10.4.6, we conclude that $\mathfrak{E}$ is a C^1 fiber bundle. Although Σ is simple from a control-theoretic point of view, an interesting feature of this result is that the drift vector field f_0 does not play any role whatsoever.

Chapter 12

Conclusions

12.1 Summary

This thesis presented a number of results pertaining to the continuation method for motion planning. More precisely, the analysis in this thesis dealt with technical obstructions to the continuation method. After describing the three distinct obstructions, we demonstrated that they can be overcome. In each case, we accomplished this task by constraining the control system under study. The constraints were imposed by restricting the control system's dynamical description, its Lie bracket configuration, and its momentum functions.

The major contributions of this thesis are as follows:

- An extended theory of C_p^q and C_q^q-polynomial control systems that accommodates control systems evolving on finite-dimensional manifolds (Chapter 3);
- An extended continuation method that incorporates arbitrary locally Lipschitz right inverses in lieu of Moore–Penrose pseudoinverses (Chapter 4);
- ◦ A number of identities involving time-varying vector fields. These identities provide reductive formulas for pullbacks involving lifts, an explicit formula for

the global flow of $X+Y$, where X is a tangent lift and Y is a vertical lift, and explicit formulas for time and parameter derivatives of pullbacks (Chapter 5);

- ◦ An explicit formula for the differentials of endpoint maps (Chapter 6);
- ◦ An explicit formula for the intrinsic quadratic differentials of anchored endpoint maps (Chapter 7);
- ◦ A necessary and sufficient constant-rank condition, which is applicable to control systems evolving on Euclidean spaces (Chapter 8);
- • A general theorem on sublinear growth, which is applicable to control-affine systems with drift (Chapter 9);
- • A topological necessary condition for unobstructed motion planning by the continuation method (Chapter 10).

In contrast with related work, the results marked with a "◦" have two distinguishing features—they are not derived using the chronological calculus formalism, and accommodate weakly regular, time-varying, fully nonlinear control systems.

In the next section, we indicate some possible avenues for future work.

12.2 Future work

Although the results presented in this thesis have shed additional light on the continuation method, the method is not yet a viable solution of practical motion planning problems. In our opinion, there are three fundamental challenges:

- The method is quite elaborate, as it relies on a number of potentially non-trivial mathematical constructions;
- The results pertaining to sublinear growth have extremely restrictive hypotheses;

- Out of necessity, one must restrict attention to control systems with few singular controls. Unfortunately, such control systems are not very abundant.

In what follows, we describe some possible ways to address these challenges.

12.2.1 Numerical implementation

Despite the elaborate nature of the continuation method, it does not defy numerical implementation. In fact, the method has been implemented by Alouges et al. [2010] towards a solution of the rolling-body problem. Their specific approach uses Galerkin methods, in which the solutions of path-lifting equations are computed by means of finite-dimensional reductions. (It must be noted that the application of Galerkin methods is only justified in light of [Chitour 2006, Theorem 1]. Roughly speaking, the latter result states that the finite-dimensional reductions are well-posed and converge, in a suitable sense, to the solution of the original path-lifting equation.) The work of Alouges et al. [2010] suggests the possibility of a more general implementation of the continuation method. This would permit exploratory numerical studies, particularly those involving the application of the continuation method to practical motion planning problems.

12.2.2 Relax the sublinear growth conditions

Although the results pertaining to sublinear growth have extremely restrictive hypotheses, we believe that this phenomenon does not reflect an inherent limitation of the continuation method. On the contrary, it seems to be an artifact produced by the imposition of sublinear growth conditions. One alternative is to use the results derived by Lee and O'Regan [1993], which imply that sublinear growth conditions can be replaced, with no loss of generality, by Wintner-type growth conditions. By relaxing the sublinear growth conditions to the full extent allowed by the Wintner theory, it may be possible to relax the results of Chapter 9 in a corresponding way.

12.2.3 Beyond submersivity

In this thesis, we focused on "submersive" control systems. That is, we focused on control systems which give rise to submersive anchored endpoint maps. The appeal of submersive control systems lies in the fact that they possess no singular controls. Broadening the domain of inquiry to non-submersive control systems presents the difficult problem of characterizing singular controls; see, for example, [Chitour et al. 2008]. This problem was taken up by Popa and Wen [2000], who described an algorithmic technique to characterize the singular controls of control systems in multi-chain form and control systems which are finitely generated in a certain sense. This technique was successfully applied to a number of interesting examples, including Dubins' car and the control of a knife edge in point contact with a plane. Thus the control systems considered by Popa and Wen [2000] are a natural point of departure for any investigations aiming to go beyond submersivity.

Bibliography

R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*, volume 75 of *Applied Mathematical Sciences*. Springer-Verlag, New York, second edition, 1988.

A. Agrachev and R. Gamkrelidze. Quasi-extremality for control systems. *Journal of Mathematical Sciences*, 55(4):1849–1864, 1991. Translated from Itogi Nauki i Tekhniki, Seriya Sovremennye Problemy Matematiki, Noveishie Dostizheniya, volume 35, pages 109–134, 1989.

A. Agrachev and Y. Sachkov. *Control Theory from the Geometric Viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer, Berlin, 2004.

A. A. Agrachev. Quadratic mappings in geometric control theory. *Journal of Mathematical Sciences*, 51(6):2667–2734, 1990.

A. A. Agrachev and R. V. Gamkrelidze. The index of extremality and quasiextremal controls. *Doklady Akademii Nauk SSSR*, 284(4):777–781, 1985.

A. A. Agrachev and S. A. Vakhrameev. Nonlinear control systems of constant rank and bang-bang conditions for extremal controls. *Doklady Akademii Nauk SSSR*, 279(2):265–269, 1984.

A. A. Agrachev and S. A. Vakhrameev. Systems, linear with respect to the control, of constant rank and conditions for the extremal controls to be switchable.

Akademiya Nauk SSSR i Moskovskoe Matematicheskoe Obshchestvo. Uspekhi Matematicheskikh Nauk, 41(6(252)):163–164, 1986.

C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, Berlin, third edition, 2006.

F. Alouges, Y. Chitour, and R. Long. A motion planning algorithm for the rolling-body problem. *IEEE Transactions on Robotics*, 26(5):827–836, 2010.

H. Amann. *Ordinary Differential Equations: An Introduction to Nonlinear Analysis*, volume 13 of *de Gruyter Studies in Mathematics*. Walter de Gruyter, Berlin, 1990.

S. C. Amiss and M. Guay. A necessary condition for path-finding by the homotopy continuation method. In *Proceedings of the 48th IEEE Conference on Decision and Control*, pages 2117–2124, Shanghai, 2009.

S. C. Amiss and M. Guay. Intrinsic quadratic differentials of the endpoint maps of nonlinear systems with L^p controls. In *Proceedings of the 18th IFAC World Congress*, pages 13551–13556, Milan, 2011a.

S. C. Amiss and M. Guay. On the endpoint maps of fully nonlinear control systems that use locally L^p controls. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 2716–2722, Atlanta, 2011b.

S. C. Amiss and M. Guay. A constant-rank theorem for time-varying control systems. 2012a. Submitted to the 2013 American Control Conference.

S. C. Amiss and M. Guay. Formulas for derivatives of global flows and pullbacks. 2012b. Submitted to the 2013 American Control Conference.

S. C. Amiss and M. Guay. Motion planning by the continuation method for control-affine systems. In *Proceedings of the 2012 American Control Conference*, pages 1767–1772, Montréal, 2012c.

S. C. Amiss and M. Guay. Motion planning by the homotopy continuation method for control-affine systems: Sublinear growth conditions. In *Proceedings of the 51st IEEE Conference on Decision and Control*, Maui, 2012d. To appear.

V. I. Arnol'd. Singularities of smooth mappings. *Russian Mathematical Surveys*, 23 (1):1–43, 1968.

V. Ayala, J. Ayala-Hoffmann, and I. de Azevedo Tribuzy. Controllability of invariant control systems at uniform time. *Kybernetika (Prague)*, 45(3):405–416, 2009.

G. Bachman and L. Narici. *Functional Analysis*. Dover Publications, Inc., Mineola, Dover edition, 2000.

P. Ballard. Dynamics of rigid bodies systems with unilateral or frictional constraints: Formulation and well-posedness. In D. Y. Gao and R. W. Ogden, editors, *Advances in Mechanics and Mathematics, 2002*, volume 1 of *Advances in Mechanics and Mathematics*, pages 3–87. Kluwer Academic Publishers, Dordrecht, 2002.

V. Barbu and C. Lefter. Optimal control of ordinary differential equations. In A. Cañada, P. Drábek, and A. Fonda, editors, *Handbook of Differential Equations: Ordinary Differential Equations*, volume 2, pages 1–75. Elsevier B. V., Amsterdam, 2005.

L. Barreira and C. Valls. *Ordinary Differential Equations: Qualitative Theory*, volume 137 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2012. Translated from the 2010 Portuguese original.

A. Bellaïche. The tangent space in sub-Riemannian geometry. In A. Bellaïche and J.-J. Risler, editors, *Sub-Riemannian Geometry*, volume 144 of *Progress in Mathematics*, pages 4–78. Birkhäuser, Basel, 1996.

R. M. Bianchini and A. Margheri. First-order differentiability of the flow of a system with L^p controls. *Journal of Optimization Theory and Applications*, 89 (2):293–310, 1996.

V. I. Bogachev. *Measure Theory. Vol. I, II.* Springer-Verlag, Berlin, 2007.

B. Bonnard and J.-B. Caillau. Introduction to nonlinear optimal control. In *Advanced Topics in Control Systems Theory*, volume 328 of *Lecture Notes in Control and Information Science*, pages 1–60. Springer, London, 2006.

B. Bonnard and I. Kupka. Théorie des singularités de l'application entrée/sortie et optimalité des trajectoires singulières dans le problème du temps minimal. *Forum Mathematicum*, 5:111–159, 1993.

D. S. Bridges. *Foundations of Real and Abstract Analysis*, volume 174 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1998.

F. Bullo and A. D. Lewis. *Geometric Control of Mechanical Systems: Modeling, Analysis, and Design for Simple Mechanical Control Systems*, volume 49 of *Texts in Applied Mathematics.* Springer-Verlag, New York, 2005a.

F. Bullo and A. D. Lewis. Supplementary chapters for *Geometric Control of Mechanical Systems* [Bullo and Lewis 2005a], 2005b. URL `http://motion.mee.ucsb.edu/book-gcms/`.

F. Bullo and A. D. Lewis. Reduction, linearization, and stability of relative equilibria for mechanical systems on Riemannian manifolds. *Acta Applicandae Mathematicae*, 99(1):53–95, 2007.

A. Chelouah and Y. Chitour. On the motion planning of rolling surfaces. *Forum Mathematicum*, 15:727–758, 2003.

Y. Chitour. *Applied and Theoretical Aspects of the Controllability of Nonholonomic Systems*. PhD thesis, Rutgers University, 1996.

Y. Chitour. Path planning on compact Lie groups using a homotopy method. *Systems & Control Letters*, 47:383–391, 2002.

Y. Chitour. A continuation method for motion-planning problems. *ESAIM: Control, Optimisation and Calculus of Variations*, 12:139–168, 2006.

Y. Chitour and H. J. Sussmann. Line-integral estimates and motion planning using a continuation method. In J. Baillieul, S. Sastry, and H. J. Sussmann, editors, *Essays on Mathematical Robotics*, volume 104 of *IMA Volumes in Mathematics and its Applications*, pages 91–125. Institute for Mathematics and its Applications, Minneapolis, 1998.

Y. Chitour, F. Jean, and E. Trélat. Singular trajectories of control-affine systems. *SIAM Journal on Control and Optimization*, 47(2):1078–1095, 2008.

D. L. Cohn. *Measure Theory*. Birkhäuser, Boston, 1980.

J. B. Conway. *A Course in Functional Analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.

M. Crampin and F. A. E. Pirani. *Applicable Differential Geometry*, volume 59 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1986.

P. E. Crouch and A. J. van der Schaft. *Variational and Hamiltonian Control Systems*, volume 101 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin Heidelberg, 1987.

J. Dominy and H. Rabitz. Dynamic homotopy and landscape dynamical set topology in quantum control. Preprint, `arXiv:1102.4360v2 [quant-ph]`, 2011.

R. M. Dudley. *Real Analysis and Probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002.

C. J. Earle and J. Eells. Foliations and fibrations. *Journal of Differential Geometry*, 1:61–69, 1967.

L. M. Fehér and B. Kőműves. On second order Thom–Boardman singularities. *Fundamenta Mathematicae*, 191:249–264, 2006.

W. Fulton and J. Harris. *Representation Theory: A First Course*, volume 129 of *Graduate Texts in Mathematics (Readings in Mathematics)*. Springer-Verlag, New York, 1991.

J. Grabowski and P. Urbański. Tangent lifts of Poisson and related structures. *Journal of Physics. A. Mathematical and General*, 28(23):6743–6777, 1995.

K. A. Grasse. *Controllability and Accessibility in Nonlinear Systems*. PhD thesis, University of Illinois at Urbana-Champaign, 1979.

K. A. Grasse. Admissibility of trajectories for control systems related by smooth mappings. *Mathematics of Control, Signals, and Systems*, 16:120–140, 2003.

K. A. Grasse. Admissible simulation relations, set-valued feedback, and controlled invariance. *Mathematics of Control, Signals, and Systems*, 20(3):199–226, 2008.

K. A. Grasse and H. J. Sussmann. Global controllability by nice controls. In H. J. Sussmann, editor, *Nonlinear Controllability and Optimal Control*, volume 133 of *Monographs and Textbooks in Applied Mathematics*, pages 33–79. Marcel Dekker, Inc., New York, 1990.

O. Gutú and J. A. Jaramillo. Fibrations on Banach manifolds. *Pacific Journal of Mathematics*, 215(2):313–329, 2004.

E. Hewitt and K. Stromberg. *Real and Abstract Analysis. A Modern Treatment of the Theory of Functions of a Real Variable.* Springer-Verlag, New York, 1965.

R. M. Hirschorn. Global controllability of locally linearizable systems. *SIAM Journal on Control and Optimization*, 28(3):540–551, 1990.

J. K. Hunter. Notes on partial differential equations. Unpublished lecture notes, 2010. URL www.math.ucdavis.edu/~hunter/pdes/.

D. Husemoller. *Fibre Bundles*, volume 20 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, third edition, 1994.

P. Jouan. Finite-time and exact-time controllability on compact manifolds. *Journal of Mathematical Sciences*, 177:402–410, 2011a.

P. Jouan. Invariant measures and controllability of finite systems on compact manifolds. *ESAIM: Control, Optimisation and Calculus of Variations, E-First*, 2011b.

V. Jurdjevic. *Geometric Control Theory*, volume 52 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1997.

C. Kawan. *Invariance Entropy for Control Systems.* PhD thesis, Universität Augsburg, 2009.

E. Kızıl. Universal covering of driftless control systems. *Journal of Dynamical and Control Systems*, 14(4):453–464, 2008.

W. P. A. Klingenberg. *Riemannian Geometry*, volume 1 of *de Gruyter Studies in Mathematics.* Walter de Gruyter & Co., Berlin, second edition, 1995.

I. Kolář, P. W. Michor, and J. Slovák. *Natural Operations in Differential Geometry*. Springer-Verlag, Berlin, 1993.

K. L. Kuttler. *Modern Analysis*. Studies in Advanced Mathematics. CRC Press, LLC, Boca Raton, 1998.

G. Lafferriere. A general strategy for computing steering controls of systems without drift. In *Proceedings of the 30th IEEE Conference on Decision and Control*, pages 1115–1120, Brighton, 1991.

G. Lafferriere and H. Sussmann. Motion planning for controllable systems without drift. In *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, pages 1148–1153, Sacramento, 1991.

G. Lafferriere and H. J. Sussmann. Motion planning for controllable systems without drift. In Z. Li and J. F. Canny, editors, *Nonholonomic Motion Planning*, pages 235–270. Kluwer Academic Publishers, Dordrecht, 1993.

F. Lamiraux and J.-P. Laumond. Flatness and small-time controllability of multibody mobile robots: Application to motion planning. *IEEE Transactions on Automatic Control*, 45(10):1878–1881, 2000.

S. M. LaValle. *Planning Algorithms*. Cambridge University Press, Cambridge, 2006.

E. B. Lee and L. Markus. *Foundations of Optimal Control Theory*. Robert E. Krieger Publishing Co., Inc., Melbourne, second edition, 1986.

J. M. Lee. *Introduction to Topological Manifolds*, volume 202 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000.

J. M. Lee. *Introduction to Smooth Manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2003.

J. M. Lee. *Manifolds and differential geometry*, volume 107 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2009.

J. W. Lee and D. O'Regan. Existence results for differential equations in Banach spaces. *Commentationes Mathematicae Universitatis Carolinae*, 34(2):239–251, 1993.

N. E. Leonard and P. S. Krishnaprasad. Motion control of drift-free, left-invariant systems on Lie groups. *IEEE Transactions on Automatic Control*, 40(9):1539–1554, 1995.

G. Leoni. *A First Course in Sobolev Spaces*, volume 105 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2009.

K.-T. Leung. *Linear Algebra and Geometry*. Hong Kong University Press, Hong Kong, 1973.

A. D. Lewis. Incomplete notes on geometric control theory. Third International Summer School on Geometry, Mechanics and Control (L'Ametlla del Mar, Tarragona, Spain), 2009. URL http://webpages.ull.es/users/gmcnet/Summer-School09/Documentos/andrews-notes.pdf.

W. Liu. An approximation algorithm for nonholonomic systems. *SIAM Journal on Control and Optimization*, 35(4):1328–1365, 1997.

L. H. Loomis and S. Sternberg. *Advanced Calculus*. Jones and Bartlett Publishers, Inc., Boston, revised edition, 1990.

J. Margalef-Roig and E. Outerelo Domínguez. *Differential Topology*, volume 173 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1992. With a preface by Peter W. Michor.

A. Margheri. Flow regularity and optimality conditions with controls in L^p. *Mathematics of Control, Signals, and Systems*, 9(3):189–206, 1996.

P. Martin. *Contribution à l'Étude des Systèmes Différentiellement Plats*. PhD thesis, École des Mines, Paris, 1992.

E. J. McShane. *Integration*. Princeton University Press, Princeton, 1944.

E. J. McShane. *Unified Integration*, volume 107 of *Pure and Applied Mathematics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1983.

M. P. Melnyk and D. E. Chang. Time optimal control of a dielectrophoretic system. *Asian Journal of Control*, 13(4):480–491, 2010.

K. M. Mikkola. *Infinite-Dimensional Linear Systems, Optimal Control and Algebraic Riccati Equations*. PhD thesis, Helsinki University of Technology, 2002.

R. Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, 2002.

J. R. Munkres. *Topology*. Prentice Hall, Upper Saddle River, second edition, 2000.

R. M. Murray and S. S. Sastry. Nonholonomic motion planning: Steering using sinusoids. *IEEE Transactions on Automatic Control*, 38(5):700–716, 1993.

S. Nikitin. *Global Controllability and Stabilization of Nonlinear Systems*, volume 20 of *Series on Advances in Mathematics for Applied Sciences*. World Scientific Publishing Co., Inc., River Edge, 1994.

D. Popa and J. T. Wen. Singularity computation for iterative control of nonlinear affine systems. *Asian Journal of Control*, 2(2):57–75, 2000.

I. R. Porteous. Simple singularities of maps. In *Proceedings of Liverpool Singularities Symposium, I (1969/70)*, volume 192 of *Lecture Notes in Mathematics*, pages 286–307. Springer, Berlin, 1971.

C. Pötzsche. *Geometric Theory of Discrete Nonautonomous Dynamical Systems*, volume 2002 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin Heidelberg, 2010.

V. V. Prasolov. *Elements of Combinatorial and Differential Topology*, volume 74 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2006. Translated from the 2004 Russian original by Olga Sipacheva.

R. Reis and M. Weiss. Smooth maps to the plane and Pontryagin classes. Part I: Local aspects. Preprint, `arXiv:1008.5104v1 [math.AT]`, 2010.

K. Rybakowski. Formulas for higher-order Fréchet derivatives of composite maps, implicitly defined maps and solutions of differential equations. *Nonlinear Analysis: Theory, Methods & Applications*, 16(6):517–532, 1991.

E. D. Sontag. Universal nonsingular controls. *Systems & Control Letters*, 19: 221–224, 1992.

E. D. Sontag. Control of systems without drift via generic loops. *IEEE Transactions on Automatic Control*, 40(7):1210–1219, 1995.

E. D. Sontag. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, volume 6 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 1998.

H. J. Sussmann. Local controllability and motion planning for some classes of systems with drift. In *Proceedings of the 30th IEEE Conference on Decision and Control*, pages 1110–1114, Brighton, 1991.

H. J. Sussmann. New differential geometric methods in nonholonomic path finding. In A. Isidori and T.-J. Tarn, editors, *Systems, Models and Feedback: Theory and Applications (Capri, 1992)*, volume 12 of *Progress in Systems and Control Theory*, pages 365–384. Birkhäuser, Boston, 1992.

H. J. Sussmann. A continuation method for nonholonomic path-finding problems. In *Proceedings of the 32nd IEEE Conference on Decision and Control*, pages 2718–2723, San Antonio, 1993.

H. J. Sussmann. An introduction to the coordinate-free Maximum Principle. In *Geometry of Feedback and Optimal Control*, volume 207 of *Pure and Applied Mathematics: A series of monographs and textbooks*, pages 463–557. Marcel Dekker, Inc., New York, 1998.

H. J. Sussmann. Geometry and optimal control. In *Mathematical Control Theory*, pages 140–198. Springer, New York, 1999.

P. Tabuada and G. J. Pappas. Hierarchical trajectory generation for a class of nonlinear systems. *Automatica*, 41:701–708, 2005.

D. Tilbury, R. M. Murray, and S. S. Sastry. Trajectory generation for the n-trailer problem using Goursat normal form. *IEEE Transactions on Automatic Control*, 40(5):802–819, 1995.

A. I. Tretiyak. Sufficient conditions for local controllability and high-order necessary conditions for optimality. A differential-geometric approach. *Journal of Mathematical Sciences*, 85(3):1899–2001, 1997. Translated from Itogi Nauki i Tekhniki, Seriya Sovremennaya Matematika i Ee Prilozheniya. Tematicheskie Obzory, Vol. 23, Dinamicheskie Sistemy-4, 1995.

F. Trèves. *Linear Partial Differential Equations with Constant Coefficients: Existence, Approximation and Regularity of Solutions*, volume 6 of *Mathematics and its Applications*. Gordon and Breach Science Publishers, New York, 1966.

S. A. Vakhrameev. Hilbert manifolds with angles of finite codimension, and optimal control theory. *Journal of Soviet Mathematics*, 53(2):176–223, 1991a.

S. A. Vakhrameev. Smooth control systems of constant rank and linearizable systems. *Journal of Mathematical Sciences*, 55(4):1864–1891, 1991b. Translated from Itogi Nauki i Tekhniki, Seriya Sovremennye Problemy Matematiki, Noveishie Dostizheniya, volume 35, pages 135–178, 1989.

S. A. Vakhrameev. Geometrical and topological methods in optimal control theory. *Journal of Mathematical Sciences*, 76(5):2555–2719, 1995.

S. A. Vakhrameev. On some properties of the path spaces for smooth control systems of constant rank. *Journal of Mathematical Sciences*, 78(5):612–625, 1996.

S. A. Vakhrameev. Bang-bang theorems and related questions. *Trudy Matematicheskogo Instituta Imeni V. A. Steklova. Rossiĭskaya Akademiya Nauk*, 220:49–112, 1998.

S. A. Vakhrameev and M. V. Topunov. On the variational characterization of switching points. *Journal of Mathematical Sciences*, 112(5):4558–4574, 2002. Translated from Itogi Nauki i Tekhniki, Seriya Sovremennaya Matematika i Ee Prilozheniya. Tematicheskie Obzory, volume 83, Dinamicheskie Sistemy-11, 2001.

S. Willard. *General Topology*. Dover, Mineola, 1970.

G. Zhong. Horizontal path spaces and Carnot–Carathéodory metrics. *Pacific Journal of Mathematics*, 161:255–286, 1993.

Appendix A

Proofs

A.1 Lemma 5.2.6

Proof. Suppose that $\mathscr{A}_Q$ is a compatible atlas on Q and $(TV, T\boldsymbol{\psi}) \in T\mathscr{A}_Q$. We must show that $\mathrm{tlft}(\Xi)_{T\boldsymbol{\psi}}$ is locally integrably C^{k-1} in the sense of Definition 2.2.19. To this end, observe that

$$\mathrm{tlft}(\Xi)_{T\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{v}) = \Big(\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}), \mathrm{ev}(\boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}), \boldsymbol{v})\Big),$$

where $\mathrm{ev} : \mathrm{Hom}(\mathbb{R}^\ell, \mathbb{R}^\ell) \times \mathbb{R}^\ell \to \mathbb{R}^\ell$ is defined by

$$\mathrm{ev}(\boldsymbol{\lambda}, \boldsymbol{v}) = \boldsymbol{\lambda} \cdot \boldsymbol{v}.$$

It is not hard to see that

$$\mathrm{ev} \in \mathrm{Hom}(\mathrm{Hom}(\mathbb{R}^\ell, \mathbb{R}^\ell), \mathbb{R}^\ell, \mathbb{R}^\ell).$$

For each $t \in I$, the map $\boldsymbol{q} \mapsto \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})$ is C^k by definition. It follows from the Leibniz rule that for each $t \in I$, the map

$$(\boldsymbol{q}, \boldsymbol{v}) \mapsto \mathrm{tlft}(\Xi)_{T\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{v})$$

is C^{k-1}. Similarly, for each $\boldsymbol{q} \in \boldsymbol{\psi}(V)$, the map $t \mapsto \boldsymbol{D}_2\Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})$ is measurable by Remark 2.2.20. Thus for each $(\boldsymbol{q}, \boldsymbol{v}) \in \boldsymbol{\psi}(V) \times \mathbb{R}^\ell$, the map

$$t \mapsto \operatorname{tlft}(\Xi)_{T\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{v})$$

is measurable by composition. We now show that the third criterion of Definition 2.2.19 is satisfied. Choose $0 \leq j \leq k-1$. Regarding $\operatorname{tlft}(\Xi)_{T\boldsymbol{\psi}}$ as a map of

$$I \times (\boldsymbol{\psi}(V) \times \mathbb{R}^\ell) \subseteq I \times (\mathbb{R}^\ell \oplus \mathbb{R}^\ell)$$

into $\mathbb{R}^\ell \oplus \mathbb{R}^\ell$, we must show that

$$\boldsymbol{D}_2^j \operatorname{tlft}(\Xi)_{T\boldsymbol{\psi}} : I \times (\boldsymbol{\psi}(V) \times \mathbb{R}^\ell) \to \operatorname{Hom}^j(\mathbb{R}^\ell \oplus \mathbb{R}^\ell, \mathbb{R}^\ell \oplus \mathbb{R}^\ell)$$

is locally integrably bounded. By the Leibniz rule,

$$\begin{aligned}
&\boldsymbol{D}_2^j \operatorname{tlft}(\Xi)_{T\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{v}) \cdot (\boldsymbol{q}_1, \boldsymbol{v}_1, \ldots, \boldsymbol{q}_j, \boldsymbol{v}_j) \\
&= \left(\begin{array}{c} \boldsymbol{D}_2^j \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}) \cdot (\boldsymbol{q}_1, \ldots, \boldsymbol{q}_j) \\ \hdashline \boldsymbol{D}_2^{j+1} \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}) \cdot (\boldsymbol{v}, \boldsymbol{q}_1, \ldots, \boldsymbol{q}_j) + \\ \sum_{i=1}^j \boldsymbol{D}_2^j \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q}) \cdot (\boldsymbol{q}_1, \ldots, \boldsymbol{q}_{i-1}, \boldsymbol{v}_i, \boldsymbol{q}_{i+1}, \ldots, \boldsymbol{q}_j) \end{array} \right).
\end{aligned} \tag{A.1}$$

This implies that

$$\begin{aligned}
&\|\boldsymbol{D}_2^j \operatorname{tlft}(\Xi)_{T\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{v})\| \\
&\leq \|\boldsymbol{D}_2^j \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| + \|\boldsymbol{v}\|_{\mathbb{R}^\ell} \|\boldsymbol{D}_2^{j+1} \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| + \sum_{i=1}^j \|\boldsymbol{D}_2^j \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\|
\end{aligned}$$

To complete the proof, let $K \subseteq \boldsymbol{\psi}(V) \times \mathbb{R}^\ell$ be compact. Since $\Xi_{\boldsymbol{\psi}}$ is locally integrably C^k and $\operatorname{pr}_1(K)$ is compact, there exist $\alpha, \beta \in L^1_{\mathrm{loc}}(I, \mathbb{R}_{\geq 0})$ such that

$$\|\boldsymbol{D}_2^j \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| \leq \alpha(t) \qquad \text{and} \qquad \|\boldsymbol{D}_2^{j+1} \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| \leq \beta(t)$$

for a.a. $t \in I$ and each $\boldsymbol{q} \in \operatorname{pr}_1(K)$. On the other hand, since $\operatorname{pr}_2(K) \subseteq \mathbb{R}^\ell$ is compact, and hence bounded, there exists $C \in \mathbb{R}_{\geq 0}$ such that $\|\boldsymbol{v}\|_{\mathbb{R}^\ell} \leq C$ for each

$\boldsymbol{v} \in \mathrm{pr}_2(K)$. Together with (A.1), these observations imply that

$$
\begin{aligned}
&\|\boldsymbol{D}_2^j \mathrm{tlft}(\Xi)_{T\boldsymbol{\psi}}(t, \boldsymbol{q}, \boldsymbol{v})\| \\
&\leq \|\boldsymbol{D}_2^j \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| + C\|\boldsymbol{D}_2^{j+1} \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| + \sum_{i=1}^{j} \|\boldsymbol{D}_2^j \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| \\
&= (j+1)\|\boldsymbol{D}_2^j \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| + C\|\boldsymbol{D}_2^{j+1} \Xi_{\boldsymbol{\psi}}(t, \boldsymbol{q})\| \\
&\leq (j+1)\alpha(t) + C\beta(t)
\end{aligned}
$$

for a.a. $t \in I$ and each $(\boldsymbol{q}, \boldsymbol{v}) \in K$. This completes the proof. ∎

A.2 Lemma 5.9.9

Proof. For each $t \in I$, define $\gamma_t : R_0 \to TQ$ by

$$\gamma_t(\rho) = T\Phi_{a,t}^{X^\rho}(\Phi_{t,a}^{X^\rho}(q_0)) \circ T\Phi_{t,a}^{X^\rho}(q_0) \cdot v_{q_0}.$$

Clearly, $\dot{\gamma}_t(r) = 0$ for each $(t, r) \in I \times R_0$. By the chain rule, we have

$$\dot{\gamma}_t(r) = \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r T\Phi_{a,t}^{X^\rho}(\Phi_{t,a}^{X^\rho}(q_0)) \circ T\Phi_{t,a}^{X^r}(q_0) \cdot v_{q_0} + TT\Phi_{a,t}^{X^r}(w) \circ \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r T\Phi_{t,a}^{X^\rho}(q_0) \cdot v_{q_0}$$

for each $(t, r) \in I \times R_0$, where $w = T\Phi_{t,a}^{X^r}(q_0) \cdot v_{q_0}$. By Theorem 5.2.4, the second term is

$$\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r T\Phi_{t,a}^{X^\rho}(q_0) \cdot v_{q_0} = \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \Phi_{t,a}^{\mathrm{tlft}(X^\rho)}(q_0) \cdot v_{q_0} = \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \Phi_{t,a}^{\mathrm{tlft}(X)^\rho}(v_{q_0}).$$

By Theorem 5.2.4, Proposition 5.4.4, Lemma 5.5.1, and Lemma 5.9.8, we have

$$
\begin{aligned}
\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \Phi_{t,a}^{\mathrm{tlft}(X)^\rho}(v_{q_0}) &= \int_a^t T\Phi_{t,a}^{\mathrm{tlft}(X)^r}(v_{q_0}) \cdot \mathrm{Ad}_{TQ_0}^{\mathrm{tlft}(X)^r}(Z^r_{\mathrm{tlft}(X)})(\sigma, v_{q_0})\,\mathrm{d}\sigma \\
&= \int_a^t T\Phi_{t,a}^{\mathrm{tlft}(X^r)}(v_{q_0}) \cdot \mathrm{Ad}_{TQ_0}^{\mathrm{tlft}(X^r)}(Z^r_{\mathrm{tlft}(X)})(\sigma, v_{q_0})\,\mathrm{d}\sigma \\
&= \int_a^t T\Phi_{t,a}^{\mathrm{tlft}(X^r)}(v_{q_0}) \cdot \mathrm{Ad}_{TQ_0}^{\mathrm{tlft}(X^r)}(\mathrm{tlft}(Z^r_X))(\sigma, v_{q_0})\,\mathrm{d}\sigma \\
&= \int_a^t T\Phi_{t,a}^{\mathrm{tlft}(X^r)}(v_{q_0}) \cdot \mathrm{tlft}(\mathrm{Ad}_{Q_0}^{X^r}(Z^r_X))(\sigma, v_{q_0})\,\mathrm{d}\sigma \\
&= \int_a^t T\Phi_{t,a}^{\mathrm{tlft}(X^r)}(v_{q_0}) \circ s_{Q_0} \circ T(\mathrm{Ad}_{Q_0}^{X^r}(Z^r_X)_\sigma)(q_0) \cdot v_{q_0}\,\mathrm{d}\sigma
\end{aligned}
$$

$$
\begin{aligned}
&= \int_a^t TT\Phi^{X^r}_{t,a}(v_{q_0}) \circ s_{Q_0} \circ T(\mathrm{Ad}^{X^r}_{Q_0}(Z^r_X)_\sigma)(q_0) \cdot v_{q_0}\,\mathrm{d}\sigma \\
&= TT\Phi^{X^r}_{t,a}(v_{q_0}) \cdot \int_a^t s_{Q_0} \circ T(\mathrm{Ad}^{X^r}_{Q_0}(Z^r_X)_\sigma)(q_0) \cdot v_{q_0}\,\mathrm{d}\sigma.
\end{aligned}
$$

This completes the proof. ∎

A.3 Lemma 5.9.10

Proof. By definition,

$$
\mathrm{Ad}^{X^r}_{Q_0}(Y^{\mathfrak{r}}_s)(t, q_0) = T\Phi^{X^r}_{a,t}(\Phi^{X^r}_{t,a}(q_0)) \circ Y^{\mathfrak{r}}_s \circ \Phi^{X^r}_{t,a}(q_0)
$$

for each $(t, r) \in I \times R_0$. By the chain rule,

$$
\begin{aligned}
&\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}^{X^\rho}_{Q_0}(Y^{\mathfrak{r}}_s)(t, q_0) \\
&= TT\Phi^{X^r}_{a,t}(w) \circ \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r Y^{\mathfrak{r}}_s \circ \Phi^{X^\rho}_{t,a}(q_0) + \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r T\Phi^{X^\rho}_{a,t}(\Phi^{X^\rho}_{t,a}(q_0)) \circ Y^{\mathfrak{r}}_s \circ \Phi^{X^r}_{t,a}(q_0)
\end{aligned}
$$

for each $(t, r) \in I \times R_0$, where $w = Y^{\mathfrak{r}}_s \circ \Phi^{X^r}_{t,a}(q_0)$. By Proposition 5.4.4 and Lemma 5.9.8, the first term is equal to

$$
\begin{aligned}
&TT\Phi^{X^r}_{a,t}(w) \circ \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r Y^{\mathfrak{r}}_s \circ \Phi^{X^\rho}_{t,a}(q_0) \\
&= TT\Phi^{X^r}_{a,t}(w) \circ TY^{\mathfrak{r}}_s(\Phi^{X^r}_{t,a}(q_0)) \circ \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \Phi^{X^\rho}_{t,a}(q_0) \\
&= TT\Phi^{X^r}_{a,t}(w) \circ TY^{\mathfrak{r}}_s(\Phi^{X^r}_{t,a}(q_0)) \circ \int_a^t T\Phi^{X^r}_{t,a}(q_0) \cdot \mathrm{Ad}^{X^r}_{Q_0}(Z^r_X)(\sigma, q_0)\,\mathrm{d}\sigma \\
&= \int_a^t TT\Phi^{X^r}_{a,t}(w) \circ TY^{\mathfrak{r}}_s(\Phi^{X^r}_{t,a}(q_0)) \circ T\Phi^{X^r}_{t,a}(q_0) \cdot \mathrm{Ad}^{X^r}_{Q_0}(Z^r_X)(\sigma, q_0)\,\mathrm{d}\sigma \\
&= \int_a^t T(T\Phi^{X^r}_{a,t} \circ Y^{\mathfrak{r}}_s \circ \Phi^{X^r}_{t,a})(q_0) \cdot \mathrm{Ad}^{X^r}_{Q_0}(Z^r_X)(\sigma, q_0)\,\mathrm{d}\sigma \\
&= \int_a^t T(\mathrm{Ad}^{X^r}(Y^{\mathfrak{r}}_s)_t)(q_0) \cdot \mathrm{Ad}^{X^r}_{Q_0}(Z^r_X)(\sigma, q_0)\,\mathrm{d}\sigma \\
&= \int_a^t T(\mathrm{Ad}^{X^r}(Y^{\mathfrak{r}}_s)_t)(q_0) \cdot \mathrm{Ad}^{X^r}_{Q_0}(Z^r_X)_\sigma(q_0)\,\mathrm{d}\sigma.
\end{aligned}
$$

Invoking Lemma 5.9.9 with $v_{q_0} = \mathrm{Ad}^{X^r}_{Q_0}(Y^{\mathfrak{r}}_s)(t, q_0)$, the second term is equal to

$$
\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r T\Phi^{X^\rho}_{a,t}(\Phi^{X^\rho}_{t,a}(q_0)) \circ Y^{\mathfrak{r}}_s \circ \Phi^{X^r}_{t,a}(q_0)
$$

$$= -\int_a^t s_{Q_0} \circ T(\mathrm{Ad}_{Q_0}^{X^r}(Z_X^r)_\sigma)(q_0) \cdot \mathrm{Ad}_{Q_0}^{X^r}(Y_s^{\mathfrak{r}})(t, q_0)\, \mathrm{d}\sigma$$
$$= -\int_a^t s_{Q_0} \circ T(\mathrm{Ad}_{Q_0}^{X^r}(Z_X^r)_\sigma)(q_0) \cdot \mathrm{Ad}_{Q_0}^{X^r}(Y_s^{\mathfrak{r}})_t(q_0)\, \mathrm{d}\sigma.$$

Consequently,

$$\begin{aligned}
&\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}_{Q_0}^{X^\rho}(Y_s^{\mathfrak{r}})(t, q_0) \\
&= \int_a^t T(\mathrm{Ad}_{Q_0}^{X^r}(Y_s^{\mathfrak{r}})_t)(q_0) \cdot \mathrm{Ad}_{Q_0}^{X^r}(Z_X^r)_\sigma(q_0) \\
&\qquad - \ s_{Q_0} \circ T(\mathrm{Ad}_{Q_0}^{X^r}(Z_X^r)_\sigma)(q_0) \cdot \mathrm{Ad}_{Q_0}^{X^r}(Y_s^{\mathfrak{r}})_t(q_0)\, \mathrm{d}\sigma \\
&= \int_a^t K_\sigma(q_0)\, \mathrm{d}\sigma.
\end{aligned}$$

By [Abraham et al. 1988, Exercise 4.2K], we have

$$K_\sigma(q_0) \in V_\Omega TQ,$$

where $\Omega = \mathrm{Ad}_{Q_0}^{X^r}(Y_s^{\mathfrak{r}})_t(q_0)$, and

$$K_\sigma(q_0) = \mathrm{vlft}_\Omega \circ \left[\mathrm{Ad}_{Q_0}^{X^r}(Z_X^r)_\sigma, \mathrm{Ad}_{Q_0}^{X^r}(Y_s^{\mathfrak{r}})_t\right](q_0).$$

Finally, we invoke Proposition 5.4.4 to obtain

$$\begin{aligned}
\mathrm{vlft}_\Omega^{-1} \cdot \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}_{Q_0}^{X^\rho}(Y_s^{\mathfrak{r}})(t, q_0) &= \mathrm{vlft}_\Omega^{-1} \cdot \int_a^t K_\sigma(q_0)\, \mathrm{d}\sigma \\
&= \int_a^t \mathrm{vlft}_\Omega^{-1} \cdot K_\sigma(q_0)\, \mathrm{d}\sigma \\
&= \int_a^t \left[\mathrm{Ad}_{Q_0}^{X^r}(Z_X^r)_\sigma, \mathrm{Ad}_{Q_0}^{X^r}(Y_s^{\mathfrak{r}})_t\right](q_0)\, \mathrm{d}\sigma.
\end{aligned}$$

This completes the proof.[1] ∎

A.4 Lemma 5.9.11

Proof. Using the chain rule, it is not hard to see that

$$\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}_{Q_0}^{X^\rho}(Y_s^\rho)(t, q_0) = \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}_{Q_0}^{X^r}(Y_s^\rho)(t, q_0) + \left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}_{Q_0}^{X^\rho}(Y_s^r)(t, q_0)$$

[1]Here, we are using the fact that each pointwise vertical lift is a canonical vector space isomorphism.

for each $(t,r) \in I \times R_0$. By linearity, the first term is equal to

$$\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}_{Q_0}^{X^r}(Y_s^\rho)(t,q_0) = \mathrm{Ad}_{Q_0}^{X^r}((Z_Y^r)_s).$$

By Lemma 5.9.10, the second term is equal to

$$\left.\frac{\mathrm{d}}{\mathrm{d}\rho}\right|_r \mathrm{Ad}_{Q_0}^{X^\rho}(Y_s^r)(t,q_0) = \int_a^t \left[\mathrm{Ad}_{Q_0}^{X^r}(Z_X^r)_\sigma, \mathrm{Ad}_{Q_0}^{X^r}(Y_s^r)_t\right](q_0)\,\mathrm{d}\sigma.$$

This completes the proof. ∎

A.5 Lemma 8.2.3

Proof. In this proof, we use the following notation:

- $\mathscr{V}^\omega(M)$ is the set of C^ω vector fields on M;
- $C^\omega(V,\mathbb{R})$ is the ring of C^ω functions on an open submanifold V of M.

Recall that $\mathscr{V}^\omega(M)$ is a $C^\omega(M,\mathbb{R})$-module [Bullo and Lewis 2005a, Section 3.9.4], and consider the submodule $\mathscr{M}$ of $\mathscr{V}^\omega(M)$ generated by the set

$$\{\mathrm{ad}_{f_0}^k(f_\ell) \;:\; k \in \mathbb{Z}_{\geq 0},\, 1 \leq \ell \leq r\}.$$

More explicitly, each element $Y \in \mathscr{M}$ is of the form

$$Y = \sum_{k=0}^{\deg(Y)} \sum_{\ell=1}^{r} P_{k,\ell}^Y \mathrm{ad}_{f_0}^k(f_\ell),$$

where $\deg(Y) \in \mathbb{Z}_{\geq 0}$ and $P_{k,\ell}^Y \in C^\omega(M,\mathbb{R})$. For each open submanifold V of M, we define $\mathscr{M}|V = \{Y|V \;:\; Y \in \mathscr{M}\}$. By [Lewis 2009, Theorem 2.4.28], $\mathscr{M}$ is locally finitely generated. That is, for each $x_* \in M$, there exist a neighbourhood V of x_* in M, $\rho \in \mathbb{Z}_{\geq 0}$, and $Y_1, \ldots, Y_\rho \in \mathscr{M}$, such that each $Y \in \mathscr{M}|V$ can be written as

$$Y = \sum_{i=1}^{\rho} C_i Y_i|V$$

for functions $C_i \in C^\omega(V, \mathbb{R})$. Given x_*, V, ρ, and $Y_1, \ldots, Y_\rho$ in this way, set

$$\Delta = \max\{\deg(Y_1), \ldots, \deg(Y_\rho)\}.$$

By construction, $\mathrm{ad}_{f_0}^{\Delta+1}(f_j)|V \in \mathscr{M}|V$, hence

$$\begin{aligned}
\mathrm{ad}_{f_0}^{\Delta+1}(f_j)|V &= \sum_{i=1}^{\rho} C_j^i Y_i|V \\
&= \sum_{i=1}^{\rho} C_j^i \sum_{k=0}^{\deg(Y_i)} \sum_{\ell=1}^{r} P_{k,\ell}^{Y_i}|V \,\mathrm{ad}_{f_0}^k(f_\ell)|V.
\end{aligned}$$

For $\deg(Y_i) < k \leq \Delta$, let $P_{k,\ell}^{Y_i} \in C^\omega(M, \mathbb{R})$ be identically equal to 0. Then

$$\begin{aligned}
\mathrm{ad}_{f_0}^{\Delta+1}(f_j)|V &= \sum_{i=1}^{\rho} C_j^i \sum_{k=0}^{\Delta} \sum_{\ell=1}^{r} P_{k,\ell}^{Y_i}|V \mathrm{ad}_{f_0}^k(f_\ell)|V \\
&= \sum_{k=0}^{\Delta} \sum_{\ell=1}^{r} \left(\sum_{i=1}^{\rho} C_j^i P_{k,\ell}^{Y_i}|V \right) \mathrm{ad}_{f_0}^k(f_\ell)|V \\
&= \sum_{k=0}^{\Delta} \sum_{\ell=1}^{r} P_{j,k,\ell} \mathrm{ad}_{f_0}^k(f_\ell)|V,
\end{aligned}$$

where

$$P_{j,k,\ell} = \sum_{i=1}^{\rho} C_j^i P_{k,\ell}^{Y_i}|V.$$

Since x_* was chosen arbitrarily, Σ satisfies the local finite definiteness condition. ∎