Proceedings of the 42nd IEEE Conference on Decision and Control
6084–6089, December 2003
DOI: 10.1109/CDC.2003.1272230

Jacobian linearisation in a geometric setting

Andrew D. Lewis* David R. Tyner†

2003/03/06
Last updated: 2003/12/14

Abstract

Linearisation is a common technique in control applications, putting useful analysis and design methodologies at the disposal of the control engineer. In this paper, linearisation is studied from a differential geometric perspective. First it is pointed out that the "naïve" Jacobian techniques do not make geometric sense along nontrivial reference trajectories, in that they are dependent on a choice of coordinates. A coordinate-invariant setting for linearisation is presented to address this matter. The setting here is somewhat more complicated than that seen in the naïve setting. The controllability of the geometric linearisation is characterised by giving an alternate version of the usual controllability test for time-varying linear systems. The problems of stability, stabilisation, and quadratic optimal control are discussed as topics for future work.

Keywords. Linearisation, differential geometry.

AMS Subject Classifications (2020). 93B18, 93B27.

1. Introduction

Jacobian linearisation[1] is, of course, a standard concept in control theory, useful in studies of controllability, optimal control, and it is still provides the setting for the majority of the control algorithms implemented in practice on nonlinear systems. In this paper we wish to develop a geometric theory of linearisation in a rather general setting. The motivation for this is not so much to broaden the applicability of linearisation techniques, but to better understand the structure of linearisation, and to make explicit some of the choices that are made without mention in the standard practice of linearisation. As we shall see in Section 1.1 below, the standard way of thinking about Jacobian linearisation

*Professor, Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada
Email: andrew.lewis@queensu.ca, URL: http://www.mast.queensu.ca/~andrew/
Research supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.
†Mechanical and Aerospace Engineering, Carleton University, Ottawa, ON, Canada, K1S 5B6
Email: drtyner@gmail.com
Work performed while a graduate student at Queen's University.
Research supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

[1]From now on we will simply say "linearisation" for "Jacobian linearisation," but the reader should understand that we do not mean "feedback linearisation."

has some difficulties as a geometric theory. Our motivation for looking at this matter comes from examples in mechanics where one has a natural geometric understanding—provided by the Jacobi equation of geodesic variation—of what it means to linearise about a reference trajectory. This geometric understanding, as far as the authors have been able to determine, is missing in the general setup for linearisation of control-affine systems. (However, we point out that certain of the ideas are implicit in the paper of Sussmann [1998], although the setup we provide is a little less abstract, and so has more structure.) In order to emphasise the relevant geometry, the setting for control theory we employ is the rather abstract setting of "affine systems" as utilised by Bullo and Lewis [2005]. Here "control" is removed from the formulation and one talks merely of trajectories.

1.1. Motivation. In order to provide a point of reference for our general formulation of control systems and their linearisation, we consider the standard manner in which linearisation is normally carried out. At various points in the paper we refer to this standard strategy as the "naïve" approach to linearisation, since it sweeps under the rug various issues that it is our objective to address.[2] We let $\Omega \subset \mathbb{R}^n$ be an open subset and we let $f_0, f_1, \dots, f_m$ be smooth vector fields, possibly depending measurably on t, on Ω. The control system we consider is then

$$\gamma'(t) = f_0(t, \gamma(t)) + \sum_{a=1}^{m} u^a(t) f_a(t, \gamma(t)), \tag{1.1}$$

where $\gamma\colon I \to \Omega$ is locally absolutely continuous and $u\colon I \to \mathbb{R}^m$ is bounded and measurable, for some interval $I \subset \mathbb{R}$. We fix a reference trajectory γ_{ref} corresponding to a reference control u_{ref}, both defined on $I \subset \mathbb{R}$. To define the linearisation, for each $t \in I$ we define smooth vector fields $f_{a,t}$, $a \in \{0, 1, \dots, m\}$, by $f_{a,t}(x) = f_a(t, x)$. The linearisation of this system is then defined by

$$\xi'(t) = A(t)\xi(t) + B(t)v(t), \tag{1.2}$$

where

$$\begin{aligned} A(t) &= \boldsymbol{D} f_{0,t}(\gamma_{\mathrm{ref}}(t)) + \sum_{a=1}^{m} u^a_{\mathrm{ref}}(t) \boldsymbol{D} f_{a,t}(\gamma_{\mathrm{ref}}(t)) \\ B(t) &= \left[\ f_{1,t}(\gamma(t)) \ \middle|\ \cdots \ \middle|\ f_{m,t}(\gamma(t))\ \right]. \end{aligned}$$

Here $\boldsymbol{D} f_{a,t}$ denotes the Jacobian of the vector field $f_{a,t}$, $a \in \{0, 1, \dots, m\}$. Once one has the linearisation (1.2), one can engage in the standard controllability investigation (see Section 4.2 for restatements of the standard Gramian results). If the linearisation is ascertained to be controllable on $I = [0, \infty[$, then one can stabilise the reference trajectory by stabilising the linearisation using linear feedback [Ikeda, Maeda, and Kodama 1972, Kalman 1960]. Thus one chooses $F\colon I \to \mathrm{L}(\mathbb{R}^m; \mathbb{R}^n)$ ($\mathrm{L}(V; W)$ denotes the linear maps from a vector space V to a vector space W) with the property that the closed-loop system

$$\xi'(t) = \big(A(t) + B(t)F(t)\big)\xi(t)$$

[2] But the reader should not mistake "naïve" with "incorrect." The approach that we refer to as "naïve" is certainly correct, but we are asking "Correct at what?"

is asymptotically stable. If $F^a \in (\mathbb{R}^n)^*$ is the ath row of F^a, then the nonlinear closed-loop system

$$\gamma'(t) = f_0(t, \gamma(t)) + \sum_{a=1}^{m} \big(u^a_{\mathrm{ref}}(t) + (F^a(t)(\gamma(t) - \gamma_{\mathrm{ref}}(t)\big) f_a(t, \gamma(t))$$

has γ_{ref} as a locally asymptotically stable trajectory. In practice one might design F by using optimal control with quadratic cost, the so-called linear quadratic regulator.

The procedure by which one arrives at (1.2), and then performs control theory on the resulting equation, poses some problems when one replaces Ω with a differentiable manifold, at least if one wishes to not submit to working in a specific coordinate chart. For example, the following questions arise.

1. What replaces the Jacobian of the vector fields $f_{a,t}$, $a \in \{0, 1, \dots, m\}$?
2. Since the problem above possesses an open subset of $\mathbb{R}^n$ as its state space, there are certain identifications that one unthinkingly makes. Where does the geometric version of (1.2) live? It cannot live in a vector space, as does (1.2).
3. By virtue of (1.2) living in a vector space, its controllability can be checked using the Gramian. What does it mean for the geometric version of (1.2) to be controllable, and how should one check whether such a system has this property?
4. What does it mean for a reference trajectory to be stable?
5. How does one perform linear state feedback for the geometric version of (1.2). And after one understands this, how can the resulting linear state feedback controller be implemented with the nonlinear system to stabilise a reference trajectory?
6. What does the geometric version of the linear quadratic regulator problem look like?

It is these questions that we begin to address in this paper.

1.2. Organisation. The layout of the paper is as follows. The paper requires certain geometric constructions that are not necessarily part of the usual nonlinear control theoretician's toolbox, and these constructions are reviewed in Section 2. In Section 3 we define the class of control systems we use in the paper, and we define how to linearise these systems. In Section 4 we consider the controllability of linearisations. We begin in Section 4.2 by re-characterising the standard controllability results for (1.2). These re-characterisations have the feature that they may be transfered directly to the setting of our geometric linearisations, and we do this in Section 4.3. Our systematic investigation of linearisation is in its initial stages. In Section 5 we indicate some of the problems that are as yet unresolved in our setting. These problems have to do with the classical approach to stabilising a reference trajectory. The issues involved here include stability, stabilisation, and quadratic optimal control.

2. Background and geometric constructions

Throughout the paper we let M be a n-dimensional Hausdorff manifold with a C^∞ differentiable structure. I will always denote an interval in $\mathbb{R}$. The set of class C^r functions on M is denoted $C^r(M)$. The tangent bundle of M is denoted $\pi_{TM}\colon TM \to M$ and the cotangent bundle by $\pi_{T^*Q}\colon T^*Q \to Q$. If $\phi\colon M \to N$ is a differentiable map between manifolds, its derivative is denoted $T\phi\colon TM \to TN$. For a vector bundle $\pi\colon E \to M$, $\Gamma^r(E)$ denotes the sections of E that are class C^r. If V and W are $\mathbb{R}$-vector spaces, $\mathrm{L}(V;W)$ denotes the set of linear maps from V to W.

2.1. Time-dependent objects on a manifold. We wish to discuss time-dependent vector fields on manifolds in a general way. We do this following Sussmann [1998, §3].

It is convenient to first talk about functions. A ***Carathéodory function*** on M is a map $\phi\colon I \times M \to \mathbb{R}$ with the property that $\phi^t \triangleq \phi(t,\cdot)$ is continuous for each $t \in I$, and $\phi_x \triangleq \phi(\cdot,x)$ is Lebesgue measurable for each $x \in M$. A Carathéodory function ϕ is ***locally integrally bounded*** (**LIB**) if for each compact subset $K \subset M$ there exists a positive locally integrable function $\psi_K\colon I \to \mathbb{R}$ so that $|\phi(t,x)| \leq \psi_K(t)$ for each $x \in K$. A Carathéodory function $\phi\colon I \times M \to \mathbb{R}$ is ***class C^r*** if ϕ^t is class C^r for each $t \in I$ and is ***locally integrally of class C^r*** (**LIC^r**) it is class C^r and if $X_1 \cdots X_r\phi^t$ is LIB for all $t \in I$ and $X_1,\dots,X_r \in \Gamma^\infty(TM)$.

A ***Carathéodory vector field*** on M is a map $X\colon I \times M \to TM$ with the property that $X(t,x) \in T_xM$ and with the property that the function $\alpha \cdot X\colon (t,x) \mapsto \alpha(x)\cdot X(t,x)$ is a Carathéodory function for each $\alpha \in \Gamma^\infty(T^*M)$. For a Carathéodory vector field X on M denote by $X_t\colon M \to TM$ the map $X_t(x) = X(t,x)$. A Carathéodory vector field X on M is ***locally integrally of class C^r*** (**LIC^r**) if $\alpha \cdot X$ is LIC^r for every $\alpha \in \Gamma^\infty(T^*M)$. We denote the set of LIC^r vector fields by $\mathrm{LIC}^r(TM)$. Our interest will primarily be with LIC$^\infty$ vector fields.

The classical theory of time-dependent vector fields with measurable time dependence gives the existence of integral curves for LIC$^\infty$ vector fields [Sontag 1998, Appendix C]. Indeed, an integral curve $\gamma\colon I \to M$ is locally absolutely continuous (LAC) (meaning that for any $\phi \in C^\infty(M)$ the map $t \mapsto \phi\circ\gamma(t)$ is locally absolutely continuous). We denote by $\gamma'(t)$ the tangent vector to γ at $t \in I$, noting that this is defined for almost every $t \in I$. The flow of $X \in \mathrm{LIC}^\infty(TM)$ we denote by Φ^X_{t,t_0}. Thus the curve $\gamma\colon t \mapsto \Phi^X_{t,t_0}(x_0)$ is the integral curve for X with initial condition $\gamma(t_0) = x_0$.

Let $\gamma\colon I \to M$ be a LAC curve. A ***vector field along γ*** is a map $\xi\colon I \to E$ with the property that $\xi(t) \in T_{\gamma(t)}M$. A vector field ξ along γ is ***locally absolutely continuous*** (**LAC**) if it is LAC as a curve in TM. A weaker notion than that of a LAC vector field along γ is that of a ***locally integrable*** (**LI**) vector field along γ, which is a vector field ξ along γ having the property that the function $t \mapsto \alpha(\gamma(t))\cdot\xi(t)$ is locally integrable for every $\alpha \in \Gamma^\infty(T^*M)$.

Let $X \in \mathrm{LIC}^\infty(TM)$ and let $\gamma\colon I \to M$ be an integral curve for X. As described by Sussmann [1998], there is a naturally defined Lie derivative operator along γ that maps LAC sections of TM along γ to LI sections of TM along γ. This operator we denote by $\mathscr{L}^{X,\gamma}$, and define it by

$$\mathscr{L}^{X,\gamma}(V_\gamma)(t) = [X_t, V](\gamma(t)),\ \text{a.e. } t \in I,$$

where $V \in \Gamma^1(TM)$ and V_γ is the LAC section of TM along γ defined by $V_\gamma(t) = V(\gamma(t))$. One easily verifies in coordinates that for a LAC vector field ξ along γ, $\mathscr{L}^{X,\gamma}(\xi)$ is then given in coordinates $(x^1, \dots, x^n)$ by

$$\mathscr{L}^{X,\gamma}(\xi)(t) = \Big(\frac{\mathrm{d}\xi^i}{\mathrm{d}t}(t) - \frac{\partial X_t}{\partial x^j}(\gamma(t))\xi^j(t)\Big)\frac{\partial}{\partial x^i}, \text{ a.e. } t, \tag{2.1}$$

where we use the summation convention. We shall revisit this operation in Sections 2.2 and 4.3.

2.2. Tangent bundle geometry. In this section we describe two ways in which a vector field on M can be lifted to a vector field on TM. It is convenient to make some general definitions.

Let $\pi\colon E \to M$ be a vector bundle. The subbundle $VE \triangleq \ker(T\pi) \subset TE$ is the ***vertical bundle*** of E. We shall be interested in certain vector fields on E. A LIC^∞ vector field X on E is ***linear*** if

1. X is π-projectable (denote the resulting vector field on M by πX) and
2. X is a linear morphism of vector bundles relative to the following diagram:

$$\begin{tikzcd} E \arrow[r, "X"] \arrow[d, "\pi"'] & TE \arrow[d, "T\pi"] \\ M \arrow[r, "\pi X"'] & TM \end{tikzcd}$$

That is to say, the induced mapping from $\pi^{-1}(x)$ to $T\pi^{-1}(\pi X(x))$ is a linear mapping of $\mathbb{R}$-vector spaces.

The flow of a linear vector field has the property that $\Phi^X_{t,t_0}|E_x\colon E_x \to E_{\Phi^{\pi X}_{t,t_0}}$ is a linear transformation.

2.1 Remark: A linear vector field on a vector bundle generalises the notion of a time-varying differential equation in the following manner. Let V be a finite-dimensional $\mathbb{R}$-vector space and consider on V a linear differential equation

$$\xi'(t) = A(t)(\xi(t)),$$

where $A\colon \mathbb{R} \to \mathrm{L}(V;V)$ is locally integrable. We then define a LIC^∞ linear vector field on the trivial bundle $\mathrm{pr}_1\colon \mathbb{R} \times V \to \mathbb{R}$ (pr_1 is projection onto the first factor) by $X_A(\tau, v) = ((\tau, v), (1, A(\tau)(v)))$. Here the projected vector field on the base space is simply $\pi X_A = \frac{\partial}{\partial \tau}$. This special case of a linear vector field has the feature that the vector bundle admits a natural global trivialisation. The lack of this feature in general accounts for some of the additional complexity in our development. Relationships between linear time-varying differential equations and linear flows on vector bundles are considered by, for example, Millionshchikov [1986a, 1986b]. •

Now let us specialise to the tangent bundle. Let $\tilde{X} \in \Gamma^\infty(TM)$. The ***complete lift*** of $\tilde{X}$ is the vector field $\tilde{X}^T \in \Gamma^\infty(TTM)$ defined by

$$\tilde{X}^T(v_x) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} T_x\Phi^{\tilde{X}}_{s,0}(v_x).$$

Now, for $X \in \mathrm{LIC}^\infty(TM)$ we define the ***complete lift*** of X to be the vector field $X^T \in \mathrm{LIC}^\infty(TTM)$ given by $X^T(t, v_x) = X^T_t(v_x)$. In natural coordinates $(x^1, \dots, x^n, v^1, \dots, v^n)$ for TM we have

$$X^T(t,x,v) = X^i(t,x)\frac{\partial}{\partial x^i} + \frac{\partial X}{\partial x^j}(t,x)v^j\frac{\partial}{\partial v^i}. \tag{2.2}$$

Now let us provide an interpretation of the complete lift. Let $\gamma\colon I \to M$ be an integral curve of $X \in \mathrm{LIC}^\infty(TM)$. A ***variation of*** $\boldsymbol{X}$ ***along*** $\boldsymbol{\gamma}$ is a map $\sigma\colon I \times J \to M$ satisfying

1. $J \subset \mathbb{R}$ is an interval for which $0 \in \mathrm{int}(J)$,
2. σ is continuous,
3. the map $I \ni t \mapsto \sigma_s(t) \triangleq \sigma(t,s) \in M$ is an integral curve for X for each $s \in J$,
4. the map $J \ni s \mapsto \sigma^t(s) \triangleq \sigma(t,s) \in M$ is LAC for each $t \in I$,
5. the map $I \ni t \mapsto \frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0} \sigma^t(s) \in TM$ is LAC, and
6. $\sigma_0 = \gamma$.

Corresponding to a variation σ of X along γ we define a LAC vector field V_σ along γ by

$$V_\sigma(t) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \sigma^t(s).$$

With this notation, the following result records some useful properties of the complete lift.

2.2 Proposition: *Let $X\colon I \times M \to TM$ be a LIC^∞ vector field, let $v_{x_0} \in T_xM$, let $t_0 \in I$, and let $\gamma\colon I \to M$ be the integral curve of X satisfying $\gamma(t_0) = x_0$. For a vector field Υ along γ satisfying $\Upsilon(t_0) = v_{x_0}$ the following statements are equivalent:*

(i) *Υ is an integral curve for X^T;*
(ii) *there exists a variation σ of X along γ so that $V_\sigma = \Upsilon$;*
(iii) *$\mathscr{L}^{X,\gamma}(\Upsilon) = 0$.*

Proof: The equivalence of (i) and (ii) will follow from the more general Proposition 3.1 below. Thus we need prove only the equivalence of (ii) and (iii). This, however, follows directly from the coordinate expressions (2.1) and (2.2). ■

Following directly from the computations arising in the proof of the preceding result is the following.

2.3 Corollary: *For $X \in \mathrm{LIC}^\infty(TM)$, X^T is a linear vector field on $\pi_{TM}\colon TM \to M$ and $\pi X^T = X$.*

Corresponding to $X \in \mathrm{LIC}^\infty(TM)$ there is also a natural vertical vector field $\mathrm{vlft}(X)$ on $\pi_{TM}\colon TM \to M$ defined by

$$\mathrm{vlft}(X)(t, v_x) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} (v_x + sX(t,x)).$$

In natural coordinates $(x^1, \dots, x^n, v^1, \dots, v^n)$ for TM we have

$$\mathrm{vlft}(X)(t,x,v) = X^i(t,x)\frac{\partial}{\partial v^i}.$$

If $u_v \in T_xM$ and $v_x \in T_xM$ then we denote $\mathrm{vlft}_{v_x} X = \mathrm{vlft}(U)(v_x)$ where U is any vector field extending u_x.

3. Affine systems and their linearisation

In this section we present a quite general formulation for control-affine systems. As part of the generalisation, we shall see that control is effectively eliminated from the picture. In practice, of course, one will not do this. However, the presence of a choice of drift vector field f_0 and control vector fields $f_1, \dots, f_m$ obfuscates some of the geometry inherent in control-affine systems. It is for this reason that we consider the control-less analogue, affine systems.

3.1. Definitions. We let M be as in Section 2.

A ***time-dependent distribution*** on M is a subset $\mathcal{D} \subset \mathbb{R} \times TM$ with the property that for each $x_0 \in M$ there exists a neighbourhood $\mathcal{N}$ and LIC^∞ vector fields $\mathscr{X} = \{X_1, \dots, X_k\}$ on $\mathcal{N}$ so that

$$\mathcal{D}_{(t,x)} \triangleq \mathcal{D} \cap \big(\{t\} \times T_xM\big) = \Big\{ \sum_{j=1}^{k} u^j X_j(t,x) \;\Big|\; u \in \mathbb{R}^k \Big\}.$$

The vector fields $\mathscr{X}$ are called ***generators*** for $\mathcal{D}$. A ***time-dependent affine subbundle*** on M is a subset $\mathcal{A} \subset \mathbb{R} \times TM$ with the property that for each $x_0 \in M$ there exists a neighbourhood $\mathcal{N}$ and LIC^∞ vector fields $\mathscr{X} = \{X_0, X_1, \dots, X_k\}$ on $\mathcal{N}$ so that

$$\mathcal{A}_{(t,x)} \triangleq \mathcal{A} \cap \big(\{t\} \times T_xM\big) = \Big\{ X_0(t,x) + \sum_{j=1}^{k} u^j X_j(t,x) \;\Big|\; u \in \mathbb{R}^k \Big\}.$$

The vector fields $\mathscr{X}$ are called ***generators*** for $\mathcal{A}$. The ***linear part*** of a time-dependent affine subbundle is the time-dependent distribution $L(\mathcal{A})$ defined by $L(\mathcal{D})_{(t,x)}$ being the subspace of T_xM upon which the affine subspace $\mathcal{A}_{(t,x)}$ is modelled. If $\mathscr{X}$ are generators for $\mathcal{A}$ as above, then the vector fields $\{X_1, \dots, X_k\}$ are ***linear generators*** for $L(\mathcal{A})$. In the setting of Bullo and Lewis [2005], we would now define an "affine system" in $\mathcal{A}$ to be an assignment to each $(t,x) \in \mathbb{R} \times M$ of a subset $\mathscr{A}(t,x)$ of $\mathcal{A}_{(t,x)}$. This is tantamount to specifying the control set for the system. However, in order to focus on the geometry associated with an affine system and its linearisation, we take $\mathscr{A}_{(t,x)} = \mathcal{A}_{(t,x)}$. This essentially means we allow unbounded controls. Therefore, we shall call the time-dependent affine subbundle $\mathcal{A}$ a ***time-dependent affine system***, accepting a slight abuse of notation. A ***trajectory*** for $\mathcal{A}$ is then a LAC curve $\gamma\colon I \to M$ with the property that $\gamma'(t) \in \mathcal{A}_{(t,\gamma(t))}$.

Note that the specification of an affine system does not provide one with the natural notion of a drift vector field and control vector fields. Indeed, a choice of drift vector field *is* a choice, and as can be seen in [Bullo and Lewis 2005], even basic properties like controllability can depend on this choice. For the theory of linearisation, this issue is put in the background, because it is natural to assume the presence of a reference vector field, cf. the discussion of Section 1.1. To be formal about this, a ***reference vector field*** for an affine system $\mathcal{A}$ is a LIC^∞ vector field $X_{\mathrm{ref}} \in \mathrm{LIC}^\infty(TM)$ with the property that $X_{\mathrm{ref}}(t,x) \in \mathcal{A}(t,x)$. Of course, integral curves of X_{ref} are trajectories for $\mathcal{A}$. One may readily show that if $\gamma\colon I \to M$ is a trajectory for $\mathcal{A}$, then there exists a reference vector field X_{ref} for $\mathcal{A}$ for which γ is an integral curve [Sussmann 1998, Proposition 4.1].

3.2. Linearisation about a reference trajectory. We let $\mathcal{A}$ be a time-dependent affine system. We select a reference vector field X_{ref} and a reference LAC integral curve $\gamma_{\mathrm{ref}}\colon I \to M$ for X_{ref}. More generally, one might choose to linearise about a reference trajectory γ_{ref} without making reference to its being an integral curve of a reference vector field. However, the embedding of γ_{ref} as an integral curve of a reference vector field gives additional useful structure, and corresponds more naturally to what one does in practice.

Now we wish to linearise about γ_{ref}. An ***$\mathcal{A}$-variation*** of γ_{ref} is a map $\sigma\colon I \times J \to M$ with the following properties:

1. $J \subset \mathbb{R}$ is an interval for which $0 \in \mathrm{int}(J)$;
2. σ is continuous;
3. the map $I \ni t \mapsto \sigma_s(t) \triangleq \sigma(t, s) \in M$ is a trajectory of $\mathcal{A}$ for each $s \in J$;
4. the map $J \ni s \mapsto \sigma^t(s) \triangleq \sigma(t, s) \in M$ is LAC for each $t \in I$;
5. the map $I \ni t \mapsto \frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0} \sigma^t(s) \in TM$ is LAC;
6. $\sigma_0 = \gamma_{\mathrm{ref}}$.

Given an $\mathcal{A}$-variation σ of γ_{ref} we define a vector field V_σ along γ_{ref} by

$$V_\sigma(t) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \sigma^t(s). \tag{3.1}$$

The vector field V_σ should be thought of as being the result of linearising in the direction of the $\mathcal{A}$-variation σ. Now let us see that these vector fields along γ_{ref} arise as trajectories for a time-dependent affine system on TM. This requires using the geometric constructions of Section 2.2.

We define a time-dependent affine subbundle $\mathcal{A}^T_{\mathrm{ref}}$ on TM as follows. Let $(t, v_x) \in \mathbb{R} \times TM$ and define

$$\mathcal{A}^T_{\mathrm{ref},(t,v_x)} = \{X^T_{\mathrm{ref}}(t, v_x) + \mathrm{vlft}_{v_x}(X) \mid X \in L(\mathcal{A})_{(t,x)}\}.$$

This is obviously a time-dependent affine subbundle since it possesses generators $\{X^T_{\mathrm{ref}}, \mathrm{vlft}(X_1), \dots, \mathrm{vlft}(X_k)\}$ where $\{X_{\mathrm{ref}}, X_1, \dots, X_k\}$ are generators for $\mathcal{A}$.

3.1 Proposition: *Let $\mathcal{A}$ be a time-dependent affine system, and let X_{ref} be a reference vector field with differentiable reference trajectory γ_{ref}, as above. For a vector field Υ along γ_{ref} the following statements are equivalent:*

(i) Υ is a trajectory for $\mathcal{A}^T_{\mathrm{ref}}$;

(ii) there exists an $\mathcal{A}$-variation σ of γ_{ref} so that $V_\sigma = \Upsilon$.

Proof: Let σ be an $\mathcal{A}$-variation giving rise to the vector field V_σ along γ_{ref}. In a set of generators $\{X_{\mathrm{ref}}, X_1, \dots, X_k\}$ for $\mathcal{A}$ we write

$$\sigma_s'(t) = X_{\mathrm{ref}}(\sigma_s(t)) + \sum_{j=1}^k u^j(t) X_j(\sigma_s(t)),$$

since σ_s is a trajectory for $\mathcal{A}$. Differentiating with respect to s at $s = 0$ gives

$$V'_\sigma(t) = X^T_{\mathrm{ref}}(\gamma'_{\mathrm{ref}}(t)) + \sum_{j=1}^k \Big(v_j(t) X_j(\gamma_{\mathrm{ref}}(t)) + u^j(0,t) \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} X_j(\sigma_s(t)) \Big),$$

where $v^j(t) = \frac{\partial u^j}{\partial s}(0,t)$. Since $\sigma_0 = \gamma_{\mathrm{ref}}$ it follows that $u^j(0,t) = 0$, and so it follows that $V'_\sigma(t) \in \mathcal{A}^T_{\mathrm{ref},(t,\gamma_{\mathrm{ref}}(t))}$, as desired.

We take a collection of generators $\{X_{\mathrm{ref}}, X_1, \dots, X_k\}$ for $\mathcal{A}$. We then let σ_j be the $\mathcal{A}$-variation of γ_{ref} satisfying

$$\frac{\mathrm{d}\sigma_s}{\mathrm{d}t}(t) = X_{\mathrm{ref}}(\sigma_s(t)) + s X_j(\sigma_s(t)),$$

noting that the corresponding infinitesimal variation is

$$V_{\sigma_j}(t) = X^T_{\mathrm{ref}}(\gamma'_{\mathrm{ref}}(t)) + \mathrm{vlft}(X_j)(\gamma'_{\mathrm{ref}}(t)).$$

The convexity of the set of variations of a given order (see [Bianchini and Stefani 1993]) now ensures the existence of a variation for any trajectory Υ that covers γ_{ref}. ■

3.3. Linearisation about an equilibrium point. The above developments concerning linearisation about a reference trajectory simplify significantly when dealing with an equilibrium point. Here the development looks a lot more like what one encounters in the non-geometric theory. In this section we explicitly make the necessary connections.

We let $\mathcal{A}$ be a time-dependent affine subbundle on M and we let $X_{\mathrm{ref}}\colon I \times M \to TM$ be a reference vector field for $\mathcal{A}$. A point $x_0 \in M$ is an ***equilibrium point*** for X_{ref} if $X_{\mathrm{ref}}(t, x_0) = 0_{x_0}$ for each $t \in I$. Thus the curve $I \ni t \mapsto x_0 \in M$ is an integral curve for X_{ref}. We consider the developments above in this special case.

First let us see how the complete lift X^T_{ref} reacts to the existence of an equilibrium point for X_{ref}.

3.2 Proposition: *If $x_0 \in M$ is an equilibrium point for a LIC^∞ vector field $X_{\mathrm{ref}}\colon I \times M \to TM$, then $X^T_{\mathrm{ref}}(v_{x_0})$ is vertical for each $v_{x_0} \in T_{x_0}M$. Furthermore, for each $t \in I$ there exists a unique $A(t) \in \mathrm{L}(T_{x_0}M; T_{x_0}M)$ so that $X^T_{\mathrm{ref}}(v_{x_0}) = \mathrm{vlft}(A(t)(v_{x_0}))$, and the map $I \ni t \mapsto A(t) \in \mathrm{L}(T_{x_0}M; T_{x_0}M)$ is Lebesgue-Lebesgue measurable.*

Proof: This follows directly from the coordinate representation (2.2) for the complete lift. ■

Thus the complete lift is vertical-valued on $T_{x_0}M$. Since $V_{x_0}M \simeq T_{x_0}M$ this means that the linearisation is a time-dependent linear affine system on $T_{x_0}M$. Weeding through the definitions we see that

$$\mathcal{A}^T_{\mathrm{ref}}(t, v_{x_0}) = \{A(t)(v_x) + b \mid\ b \in L(\mathcal{A})_{(t,x_0)}\}.$$

Trajectories $\xi\colon I \to T_{x_0}M$ of the linearisation then satisfy

$$\xi'(t) = A(t)(\xi(t)) + b(t) \tag{3.2}$$

for some measurable curve $b\colon I \to T_{x_0}M$ having the property that $b(t) \in L(\mathcal{A})_{(t,x_0)}$. To make this look more like the usual notion of a time-varying linear system, for each $t \in I$

let U be a finite-dimensional $\mathbb{R}$-vector space and let $B(t) \in \mathrm{L}(U; T_{x_0}M)$ have the property that $\mathrm{image}(B(t)) = L(\mathcal{A})_{(t,x)}$. Then we may write the equation governing trajectories as

$$\xi'(t) = A(t)(\xi(t)) + B(t)(u(t))$$

for a measurable curve $u\colon I \to U$. This then recovers the usual notion of a time-dependent linear system.

4. Linear controllability

In Section 3 we constructed a time-dependent affine subbundle $\mathcal{A}^T_{\mathrm{ref}}$ on TM by linearising a time-dependent affine subbundle $\mathcal{A}$ on M along a reference trajectory γ_{ref}. In this section we start by defining the notions of the reachable set of both $\mathcal{A}$ and $\mathcal{A}^T_{\mathrm{ref}}$, and the associated versions of controllability along γ_{ref}. In Section 4.2 the standard results for a time-varying linear system are cast in a geometric manner. Our geometric characterisation of the standard Gramian condition allows us to state in Section 4.3 a completely geometric result that describes the reachable sets for the linearisation.

4.1. Controllability definitions. In this section the reachable sets for both $\mathcal{A}$ and $\mathcal{A}^T_{\mathrm{ref}}$ are defined and for the latter we give two equivalent statements of the reachable set. Then controllability along a reference trajectory is provided for $\mathcal{A}$ and as well the controllability of $\mathcal{A}^T_{\mathrm{ref}}$.

Recall that a trajectory for a time-dependent affine subbundle $\mathcal{A}$ is a LAC curve $\gamma\colon I \to M$ such that $\gamma'(t) \in \mathcal{A}_{(t,\gamma(t))}$. Then the set of trajectories defined on $[t_0, T]$ is denoted by $\mathrm{Traj}(\mathcal{A}, T, t_0)$ and $\mathrm{Traj}(\mathcal{A}, t_0) = \bigcup_{T \geq t_0} \mathrm{Traj}(\mathcal{A}, T, t_0)$. For $x_0 \in M$ and $t \geq t_0$ we define the ***reachable set*** of $\mathcal{A}$ from x_0 as

$$\mathcal{R}_{\mathcal{A}}(x_0, t, t_0) = \{\gamma(t) \mid \ \gamma \in \mathrm{Traj}(\mathcal{A}, t_0),\ \gamma(t_0) = x_0\}.$$

Similarly a ***trajectory*** for the linearised time-dependent affine subbundle $\mathcal{A}^T_{\mathrm{ref}}$ is a LAC curve $\Upsilon\colon I \to TM$ such that $\Upsilon'(t) \in \mathcal{A}^T_{\mathrm{ref}}(t, \Upsilon(t))$. Then the set of trajectories defined on $[t_0, T]$ is denoted by $\mathrm{Traj}(\mathcal{A}^T_{\mathrm{ref}}, T, t_0)$ and $\mathrm{Traj}(\mathcal{A}^T_{\mathrm{ref}}, t_0) = \bigcup_{T \geq t_0} \mathrm{Traj}(\mathcal{A}^T_{\mathrm{ref}}, T, t_0)$. For $v_{x_0} \in TM$ and $t \geq t_0$ we define the ***reachable set*** from v_{x_0} as

$$\mathcal{R}_{\mathcal{A}^T_{\mathrm{ref}}}(v_{x_0}, t, t_0) = \{\Upsilon(t) \mid \ \Upsilon \in \mathrm{Traj}(\mathcal{A}^T_{\mathrm{ref}}, t_0),\ \Upsilon(t_0) = v_{x_0}\}.$$

With these notions of reachable sets, we have the following controllability notions.

4.1 Definition: Let X_{ref} be a reference vector field for $\mathcal{A}$ and let $\gamma_{\mathrm{ref}}\colon I \to M$ be a reference trajectory. Let $x_0 \in M$ and $\gamma_{\mathrm{ref}}(t_0) = x_0$. $\mathcal{A}$ is

(i) ***controllable at t_0*** along γ_{ref} if $\gamma_{\mathrm{ref}}(t) \in \mathrm{int}\,\mathcal{R}_{\mathcal{A}}(x_0, t, t_0)$ for each $t > t_0$ and is

(ii) ***linearly controllable at t_0*** along γ_{ref} if $\mathcal{R}_{\mathcal{A}^T_{\mathrm{ref}}}(0_{x_0}, t, t_0) = T_{\gamma_{\mathrm{ref}}(t)}M$ for each $t > t_0$. •

4.2. Recasting the standard results. Let U and V be $\mathbb{R}$-vector space with $\dim(U) = m$ and $\dim(V) = n$. Let $A\colon \mathbb{R} \to \mathrm{L}(V; V)$ and $B\colon \mathbb{R} \to \mathrm{L}(U; V)$ be continuous and define a time-varying affine subbundle $\mathcal{A}_{(A,B)}$ on V by

$$\mathcal{A}_{(A,B),(t,v)} = \{A(t)v + B(t)u \mid \ u \in U\}.$$

A trajectory ξ of $\mathcal{A}_{(\mathrm{A},\mathrm{B})}$ satisfies

$$\xi'(t) = A(t)\xi(t) + B(t)u(t). \tag{4.1}$$

The solution to (4.1) satisfying $\xi(t_0) = \xi_0$ for $t_0 \in I$ is given by,

$$\xi(t) = \Phi(t, t_0)\xi_0 + \int_{t_0}^{t} \Phi(t, \sigma)B(\sigma)u(\sigma)\mathrm{d}\sigma, \tag{4.2}$$

where $\Phi(t, t_0)$ is the state transition matrix. That is, $t \mapsto \Phi(t, t_0)$ is the solution to the inhomogeneous system $\Phi'(t) = A(t)\Phi(t)$ with initial condition $\Phi(t_0) = \mathrm{id}_V$. We recall that the transition matrix has the following properties:

1. $\Phi(t, \tau)\Phi(\tau, t_0) = \Phi(t, t_0)$;
2. $\Phi(t, \tau)^{-1} = \Phi(\tau, t)$.

We say that $\mathcal{A}_{(A,B)}$ is ***controllable at t_0*** if for each $\xi_0, \xi_1 \in V$ there exists a control $u\colon [t_0, t_1] \to U$ which steers from ξ_0 at time t_0 to ξ_1 at time t_1.

The controllability of a time-varying linear system is determined by the ***controllability Gramian***,

$$W(t_0, t) = \int_{t_0}^{t} \Phi(t_0, \sigma)B(\sigma)B^T(\sigma)\Phi^T(t_0, \sigma)\mathrm{d}\sigma.$$

One shows that $\mathcal{A}_{(A,B)}$ is controllable at t_0 if and only if $W(t_0, t)$ is surjective for $t > t_0$. Note that the notion of a controllability Gramian does not make sense in the geometric setting for linearisation of Section 3. By this we mean that there is no natural way to construct the analogue of $W(t_0, t)$ for the linearisation of a reference vector field X_{ref} along a reference trajectory γ_{ref}. Therefore, we need an alternate characterisation of controllability that *can* be applied in the geometric setting. The following result gives one such characterisation.

4.2 Theorem: *Let V, U, A, and B be as above. Then*

$$\mathrm{image}(W(t_0, t)) = \mathrm{span}_{\mathbb{R}}\Bigg(\bigcup_{\substack{\tau \in [t_0, t] \\ b_\tau \in \mathrm{image}(B(\tau))}} \Phi(t_0, \tau)b_\tau\Bigg).$$

Proof: Let us denote

$$\mathcal{S}_{\mathcal{A}_{(\mathrm{A},\mathrm{B})}}(t_0, t) = \mathrm{span}_{\mathbb{R}}\Bigg(\bigcup_{\substack{\tau \in [t_0, t] \\ b_\tau \in \mathrm{image}(B(\tau))}} \Phi(t_0, \tau)b_\tau\Bigg).$$

Let $v \in \mathrm{image}(W(t_0, t))$. Then there exists a continuous control $u\colon [t_0, t] \to U$ so that

$$v = \int_{t_0}^{t} \Phi(t_0, \sigma)B(\sigma)u(\sigma)\mathrm{d}\sigma$$

(see [Brockett 1970]). Since A, B, and u are continuous, there exists a sequence of partitions $P_i = \{t_0 = t_{1,i}, \dots, t_{k_i,i} = t\}$ of $[t_0, t]$ so that if we define

$$v_i = \sum_{j=2}^{k_i} \Phi(t_0, t_{j,i})B(t_{j,i})u(t_{j,i})(t_{j,i} - t_{j-1,i}),$$

then $\lim_{i\to\infty} v_i = v$. It is clear that for each $i \in \mathbb{N}$, $v_i \in \mathcal{S}_{\mathcal{A}_{(A,B)}}(t_0,t)$. Since $\mathcal{S}_{\mathcal{A}_{(A,B)}}(t_0,t)$ is a subspace of V it follows that $v \in \mathcal{S}_{\mathcal{A}_{(A,B)}}(t_0,t)$.

Now assume that $v \in \mathcal{S}_{\mathcal{A}_{(A,B)}}(t_0,t)$. Choose $t_1,\dots,t_k \in [t_0,t]$ and $b_{t_j} \in \operatorname{image}(B(t_j))$, $j \in \{1,\dots,k\}$, so that

$$\mathcal{S}_{\mathcal{A}_{(A,B)}}(t_0,t) = \operatorname{span}_{\mathbb{R}}(\Phi(t_0,t_1)b_{t_1},\dots,\Phi(t_0,t_k)b_{t_k}).$$

Then we may write

$$v = \sum_{j=1}^{k} c_j \Phi(t_0,t_j) b_{t_j}.$$

We now give a useful characterisation of points in $\operatorname{image}(W(t_0,t))$.

1 Lemma: $\operatorname{image}(W(t_0,t)) = \{\Phi(t_0,t)\tilde{v} \in V \mid \exists\, u\colon [t_0,t] \to U$ *steering zero to* $\tilde{v}\}$.

Proof: By (4.2), the set of points reachable from $0 \in V$ in time t from t_0 is

$$\{\int_{t_0}^{t} \Phi(t,\sigma)B(\sigma)u(\sigma)\mathrm{d}\sigma \mid\ u\colon [t_0,t] \to U \text{ continuous}\}.$$

Thus, if we take a point in this set and apply to it $\Phi(t_0,t)$ we get

$$\Phi(t_0,t)\int_{t_0}^{t} \Phi(t,\sigma)B(\sigma)u(\sigma)\mathrm{d}\sigma = \int_{t_0}^{t} \Phi(t_0,\sigma)B(\sigma)u(\sigma)\mathrm{d}\sigma,$$

by the composition property of the transition matrix. The lemma now follows by comparison with (4.2). ▼

We will now show that one can steer from 0 to $\Phi(t,t_0)v$, and from this, this part of the theorem will follow from Lemma 1. Let $\mu_j \in U$ have the property that $B(t_j)\mu_j = b_{t_j}$ $j \in \{1,\dots,k\}$. Now consider the distributional control $u = \sum_{j=1}^{j} c_j \delta_{t_j}\mu_j$, where δ_{t_j} is the delta-distribution with support $\{t_j\}$. For this control, by (4.2), we have

$$\int_{t_0}^{t} \Phi(t,\sigma)B(\sigma)u(\sigma)\,\mathrm{d}\sigma = \sum_{j=1}^{k} c_j \Phi(t,t_j)b_{t_j} = \Phi(t,t_0)v. \tag{4.3}$$

Thus the distributional control u steers from 0 to $\Phi(t,t_0)v$, as desired. Now we show that we may replace the distributional control u with a sequence of piecewise continuous controls. The following lemma is helpful for this.

2 Lemma: *There exists a sequence of controls* $\{u_i\}_{i\in\mathbb{N}}$ *so that*

$$\lim_{i\to\infty}\int_{t_0}^{t} \Phi(t,\sigma)B(\sigma)u_i(\sigma)\mathrm{d}\sigma = \sum_{j=1}^{k} c_j \Phi(t,t_j)b_{t_j}.$$

Proof: For $j \in \{1,\dots,k\}$ and $i \in \mathbb{N}$ define

$$u_{j,i}(t) = \begin{cases} i\, c_j\, \mu_j, & t \in [t_j, t_j + \frac{1}{i}] \\ 0, & \text{otherwise.} \end{cases}$$

Now note that, using the Peano-Baker series,

$$\begin{aligned}\int_{t_0}^{t}\Phi(t,\sigma)B(\sigma)u_{j,i}(\sigma)\mathrm{d}\sigma &= \Phi(t,t_j+\tfrac{1}{i})\int_{t_0}^{t}\Phi(t_j+\tfrac{1}{i},\sigma)B(\sigma)u_{j,i}(\sigma)\mathrm{d}\sigma\\ &= \Phi(t,t_j+\tfrac{1}{i})\int_{t_j}^{t_j+\frac{1}{i}}\Phi(t_j+\tfrac{1}{i},\sigma)B(\sigma)u_{j,i}(\sigma)\mathrm{d}\sigma\\ &= i\,c_j\,\Phi(t,t_j+\tfrac{1}{i})\int_{t_j}^{t_j+\frac{1}{i}}\Big(\mathrm{id}_V+\int_{\sigma}^{t_j+\frac{1}{i}}A(\sigma_1)\mathrm{d}\sigma_1\\ &\qquad+\int_{\sigma}^{t_j+\frac{1}{i}}A(\sigma_1)\int_{\sigma}^{\sigma_1}A(\sigma_2)\mathrm{d}\sigma_2\mathrm{d}\sigma_1+\dots\Big)B(\sigma)\mu_j\mathrm{d}\sigma.\end{aligned}$$

Because A is continuous all terms in the Peano-Baker series go to zero at least as fast as $(\frac{1}{i})^2$. Thus only the first term remains in the limit, giving

$$\lim_{i\to\infty}\int_{t_0}^{t}\Phi(t,\sigma)B(\sigma)u_{j,i}(\sigma)\mathrm{d}\sigma = c_j\Phi(t,t_j)b_{t_j}.$$

The result now follows by taking $u_i=\sum_{j=1}^{k}u_{j,i}$. ▼

Let $\{u_i\}_{i\in\mathbb{N}}$ be a sequence of controls defined by Lemma 2. For each $i\in\mathbb{N}$ we have

$$\Phi(t_0,t)\int_{t_0}^{t}\Phi(t,\sigma)B(\sigma)u_i(\sigma)\mathrm{d}\sigma\in\mathrm{image}(W(t_0,t))$$

by Lemma 1. Therefore, the limit as $i\to\infty$ is also in $\mathrm{image}(W(t_0,t))$. But by (4.3) we have

$$\lim_{i\to\infty}\Phi(t_0,t)\int_{t_0}^{t}\Phi(t,\sigma)B(\sigma)u_i(\sigma)\mathrm{d}\sigma = v,$$

giving the result. ■

4.3. Controllability of linearisations. Now we use Theorem 4.2 to provide geometric characterisations of the reachable sets for the linearisation of a reference vector field X_{ref} along a reference trajectory γ_{ref}. To state the result, we need a definition. For a LAC curve $\gamma\colon I\to M$ a ***distribution along*** $\boldsymbol{\gamma}$ is a subset $D\subset TM|\mathrm{image}(\gamma)$ with the property that for each $t_0\in I$ there exists a neighbourhood $J\subset I$ of t_0 and LAC vector fields $\xi_1,\dots,\xi_k$ along $\gamma|J$ so that $D_{\gamma(t)}=\mathrm{span}_{\mathbb{R}}(\xi_1(t),\dots,\xi_k(t))$ for each $t\in J$. Let $t_0\in\mathrm{int}(I)$ and denote $T=\sup I$, allowing that $T=\infty$. Let $I_{t_0}=[t_0,T[$. Denote by γ_{t_0} the restriction of γ_{ref} to I_{t_0}. Recalling from Section 2.1 the definition of $\mathscr{L}^{X_{\mathrm{ref}},\gamma_{t_0}}$ we denote by $\langle\mathscr{L}^{X_{\mathrm{ref}},\gamma_{t_0}},L(\mathcal{A})_{t_0}\rangle$ the smallest $\mathscr{L}^{X_{\mathrm{ref}},\gamma_{t_0}}$-invariant distribution along γ_{t_0} that agrees with $L(\mathcal{A})$ at $\gamma_{\mathrm{ref}}(t_0)$.

4.3 Theorem: *Let $\mathcal{A}$ be a time-dependent affine system on M with X_{ref} a reference vector field and $\gamma_{\mathrm{ref}}\colon I\to M$ a differentiable reference trajectory. For $t_0\in I$ and $t>t_0$ the following sets are equal:*

(i) $\mathcal{R}_{\mathcal{A}^T_{\mathrm{ref}}}(0_{x_0},t,t_0)$;

(ii) $\mathrm{span}_{\mathbb{R}}\left(\bigcup_{\substack{\tau\in[t_0,t]\\ v_\tau\in L(\mathcal{A})_{\tau,\gamma_{\mathrm{ref}}(\tau)}}}\Phi^{X^T_{\mathrm{ref}}}_{\tau,t_0}(v_\tau)\right)$;

(iii) $\langle \mathscr{L}^{X_{\mathrm{ref}},\gamma_{t_0}}, L(\mathcal{A})_{t_0}\rangle_{\gamma_{\mathrm{ref}}(t)}$.

Proof: Since γ_{ref} is differentiable we can construct the pull-back bundle $\gamma_{\mathrm{ref}}^*\pi_{TM}\colon \gamma_{\mathrm{ref}}^*TM \to I$ where we recall that

$$\gamma_{\mathrm{ref}}^*TM = \{(t,v) \mid \gamma_{\mathrm{ref}}(t) = \pi_{TM}(v)\}$$

and $\gamma_{\mathrm{ref}}^*\pi_{TM}(t,v) = t$. Thus $\gamma_{\mathrm{ref}}^*TM$ is a vector bundle over I with fibre over $t \in I$ being $T_{\gamma_{\mathrm{ref}}(t)}M$. This bundle may be trivialised since I is contractible, and we denote a particular trivialisation by a vector bundle mapping $\rho\colon \gamma_{\mathrm{ref}}^*\pi_{TM} \to I \times V$, with the diagram

$$\begin{tikzcd} \gamma_{\mathrm{ref}}^*\pi_{TM} \arrow[rr, "\rho"] \arrow[dr, "\gamma_{\mathrm{ref}}^*\pi_{TM}"'] & & I \times V \arrow[dl, "\mathrm{pr}_1"] \\ & I & \end{tikzcd}$$

commuting, with pr_1 the projection onto the first factor.

The following lemma records some useful properties of the representation of trajectories of $\mathcal{A}_{\mathrm{ref}}^T$.

1 Lemma: *(i) There exists a vector bundle endomorphism $A\colon I \times V \to I \times V$ over id_I with the property that $T_{(t,\gamma_{\mathrm{ref}}(t))}\rho(1, X_{\mathrm{ref}}^T(v_{\gamma_{\mathrm{ref}}(t)})) = (1, A(t)\cdot\rho(v_{\gamma_{\mathrm{ref}}(t)}))$.*

(ii) If $X \in \Gamma^\infty(TM)$ then there exists a section ξ_X of $\mathrm{pr}_1\colon I \times V \to I$ so that $T_{(t,\gamma_{\mathrm{ref}}(t))}\rho(0, \mathrm{vlft}(X)(\gamma_{\mathrm{ref}}(t))) = \xi_X(t)$.

Proof: The first assertion follows since X_{ref} is a vector bundle mapping over X. The second part of the lemma is merely a definition of ξ_X. ▼

The lemma tells us that if $v(t) = T\rho(\Upsilon(t))$ for a trajectory Υ for $\mathcal{A}_{\mathrm{ref}}^T$ then we have

$$v'(t) = A(t)v(t) + b(t)$$

where $b(t) \in \mathrm{image}\big(\rho(\mathrm{vlft}(\mathcal{A}_{(t,\gamma_{\mathrm{ref}}(t))}))\big)$. Therefore, the equality of the sets (i) and (ii) follows from Theorem 4.2.

From the definition of the set of (ii), the equivalence of (ii) and (iii) will follow if we can show that the notion of a distribution along γ_{t_0} being invariant under $\mathscr{L}^{X_{\mathrm{ref}},\gamma_{t_0}}$ is equivalent to the notion be being invariant under the flow of X_{ref}^T along γ_{t_0}. Thus we let $\mathcal{D}$ be a distribution along γ_{t_0}, and we claim that $\mathcal{D}$ is invariant under the flow of X_{ref}^T if and only if it is invariant under $\mathscr{L}^{X_{\mathrm{ref}},\gamma_{t_0}}$. This, however, follows from the characterisation of the flow of X_{ref} in Proposition 2.2 in terms of $\mathscr{L}^{X_{\mathrm{ref}},\gamma_{t_0}}$. Indeed, that result states that the flow of X_{ref}^T through $v \in \mathcal{D}_{\gamma_{t_0}(t_0)}$ is obtained by transporting v along γ_{t_0}. ■

4.4 Remark: The set described in part (iii) of the theorem should be thought of as the analogue of “the smallest A-invariant subspace containing image(B)” in the time-invariant linear theory. •

The theorem immediately gives the following corollary, the second part of which follows from the variational cone results of Bianchini and Stefani [1993].

4.5 Corollary: *Let $\mathcal{A}$, X_{ref}, and γ_{ref} be as in Theorem 4.3. Then the following statements hold for $t_0 \in I$:*

(i) *$\mathcal{A}$ is linearly controllable at t_0 along γ_{ref} if the smallest $\mathscr{L}^{X_{\mathrm{ref}},\gamma_{\mathrm{ref}}}$-invariant distribution along γ_{ref} containing $L(\mathcal{A})|\operatorname{image}(\gamma_{\mathrm{ref}})$ is equal to $TM|\operatorname{image}(\gamma_{\mathrm{ref}})$;*

(ii) *if $\mathcal{A}$ is linearly controllable at t_0 along γ_{ref} then it is controllable at t_0 along γ_{ref}.*

5. Some open problems presented by our approach

The geometric characterisation of linear controllability in the preceding section gives an indication of the sort of results one can expect to get in our geometric setting. In this section we outline future work along these lines that will further serve to clarify what gets commonly done in practice.

5.1. Stability. For the study of stability one fixes the reference vector field X_{ref} and a differentiable reference trajectory γ_{ref}. Now one can ask questions concerning the stability of this reference trajectory. Such a discussion will involve a choice of metric, for example arising from a Riemannian metric, on M. One can also talk about the stability of the linearisation X_{ref}. This discussion too will require some sort of metric, namely one on the fibres of $\gamma_{\mathrm{ref}}^*TM$. Again, such a metric will come about naturally if one chooses a Riemannian metric g on M. This then raises the following question.

5.1 Question: Let X_{ref} and γ_{ref} be as above, and choose a Riemannian metric g on M. Consider the stability of the trajectory γ_{ref} relative to the metric d_g defined by g, and consider the stability of the linearisation X_{ref}^T relative to the vector bundle metric γ_{ref}^*g on $\gamma_{\mathrm{ref}}^*TM$. Does uniform asymptotic stability of the linearisation imply exponential stability of the trajectory? •

Note that the existence of Question 5.1 is hidden by the naïve Jacobian linearisation of Section 1.1 because one uses, without thinking about it, the standard Euclidean metric on $\mathbb{R}^n$. However, this may not actually be the proper metric for talking about stability in a particular problem. Perhaps the problem possesses its own natural metric with respect to which stability characterisations ought to be made. Furthermore, if the state manifold is not compact, this choice of metric for measuring stability matters, in that a system may be stable relative to one metric, but not another. Thus Question 5.1 is not vacuous.

5.2. Stabilisation. If a trajectory is not stable with respect to some choice of metric, then one can ask whether it is possible to stabilise it under feedback. This is of course a fairly well understood problem in the setting of Section 1.1 when the linearisation is controllable and one uses linear state feedback. However, in our geometric setting, some problems arise. First let us make sure that we understand what could constitute linear state feedback in our setting. We let X_{ref} be a reference vector field for the affine system $\mathcal{A}$ with $\gamma_{\mathrm{ref}}\colon I \to M$ a differentiable reference trajectory. To avoid complications, suppose that $L(\mathcal{A})$ has constant rank and so is a vector bundle over M. Let $\mathrm{L}(TM; L(\mathcal{A}))$ be the set of vector bundle mappings from TM to $L(\mathcal{A})$ over id_M. A ***linear state feedback*** along γ_{ref} is then a section F of the bundle $\gamma_{\mathrm{ref}}^*\mathrm{L}(TM; L(\mathcal{A}))$. Thus F assigns to each point $t \in I$ a linear map $F(t)\colon T_{\gamma_{\mathrm{ref}}(t)}M \to L(\mathcal{A})_{(t,\gamma_{\mathrm{ref}}(t))}$. For such a linear state feedback, the

closed-loop system is then the linear LIC$^\infty$ vector field on TM defined by $X^F_{\mathrm{ref}}(t, v_x) = X^T_{\mathrm{ref}}(t, v_x) + \mathrm{vlft}(F(t)(v_x))$. Note that the integral curves of X^F_{ref} with initial conditions projecting to γ_{ref} will project fully to γ_{ref}. Therefore, given a linear state feedback F one can then talk about the stability of the linear vector field X^F_{ref} relative to a fibre metric on $\gamma^*_{\mathrm{ref}}TM$, just as in Section 5.1. Therefore it makes sense to pose the following question.

5.2 Question: Let X_{ref} and γ_{ref} be as above and let d_t be a metric on $T_{\gamma_{\mathrm{ref}}(t)}M$. When is it possible to find a linear state feedback F so that the closed-loop linear system X^F_{ref} is stable (asymptotically stable, uniformly asymptotically stable, etc.) relative to the fibre metric d? •

In practice, consistent with Question 5.1, one might choose d as being induced by a Riemannian metric g on M. In this case one might expect that controllability of the linearisation might imply stabilisability of the linearisation, as this is the case in the setting of Section 1.1. However, this conclusion in our general setting does not seem to follow obviously from the same conclusion in the naïve setting, so the work must be done here.

After one stabilises the linearisation, it still remains to see if one can stabilise the actual reference trajectory. In the setting of Section 1.1, this is again done without thought, since the state space is naturally identified with each tangent space, and an implementation of the stabilising control law for the linearisation carries over to a control law in state space in an obvious way. However, this operation is clearly senseless in a geometric setting. Therefore, this raises the following question.

5.3 Question: From a control law that stabilises the linearisation along a reference trajectory, how can one determine a control law for the actual system? And once one has a way of doing this, can one be ensured that stabilisation (in an appropriate manner) of the linearisation will guarantee the stabilisation of the reference trajectory? •

The presence of a Riemannian metric g suggests the following manner in which to implement the stabilising control law for the linearisation on the actual system. Suppose that γ_{ref} is an immersion, and for $t \in I$ consider a neighbourhood $J \subset I$ of t for which $\gamma_{\mathrm{ref}}|J$ is an embedding. Now, let $N \to \gamma_{\mathrm{ref}}(J)$ be the normal bundle of $\gamma_{\mathrm{ref}}(J)$ so that the Riemannian exponential gives a diffeomorphism from a neighbourhood W of the zero section of N to a neighbourhood U in M. The closed-loop system can now be defined in U by $X^{F,g}_{\mathrm{ref}}(t, x) = X_{\mathrm{ref}}(t, x) + \exp(F(t)(\exp^{-1}(x)))$. One can now ask, as in Question 5.3, whether this control stabilises the system if F stabilises the linearisation.

5.3. Quadratic optimal control.

In practice, a common way of designing a stabilising feedback for the linearisation is to use quadratic optimal control, as initially discussed by Kalman [1960], and further by Ikeda, Maeda, and Kodama [1972]. Let us see what the usual theory of optimal control for linear systems looks like when cast in our setting. When a problem is written in the usual form of Section 1.1 the quadratic cost function associated with a controlled trajectory (u, ξ) defined on $[t_0, t_1]$ is typically chosen to be

$$\int_{t_0}^{t_1} \big(Q(t)(\xi(t), \xi(t)) + R(t)(u(t), u(t))\big)\, \mathrm{d}t,$$

where $Q, R\colon [t_0, t_1] \to \mathrm{TS}^2(V)$ are measurable with Q positive-semidefinite and R positive-definite. For our abstract setting, where controls have been eliminated, this changes. Indeed, after some thought one can see that u is embedded in V by B in any case, so the cost function can be thought of as being simply defined on V.

With the preceding comments as motivation, we now consider a reference vector field X_{ref} for an affine system $\mathcal{A}$ along with a differentiable reference trajectory γ_{ref}. We let $\mathrm{TS}^2(TM)$ be the bundle of symmetric $(0,2)$ tensors on M, and we let Q be a LI section of $\gamma_{\mathrm{ref}}^*\mathrm{TS}^2(TM)$ with the property that $Q(t)$ is positive-semidefinite and $Q(t)|L(\mathcal{A})_{(t,\gamma_{\mathrm{ref}}^{\pi}(t))}$ is positive-definite for each $t \in I$. This then leads to the following problem.

5.4 Problem: Let X_{ref}, γ_{ref}, and Q be as above. Find the trajectory Υ for $\mathcal{A}_{\mathrm{ref}}^T$ that projects to γ_{ref} and that minimises the cost

$$J_Q(\Upsilon) = \int_I Q(t)(\Upsilon(t), \Upsilon(t))\,\mathrm{d}t. \qquad \bullet$$

One expects there to be some analogue of the Riccati equation, but one cannot simply use the usual Riccati equation, since there is no natural choice of A and B in our geometric setting. Thus it is an actual problem to understand what form will be given to the solution of the optimal control problem. This then raises a couple of questions.

5.5 Question:

1. Does the solution Υ to Problem 5.4 arise from a linear state feedback as in the standard case?
2. If the answer to the preceding question is, “Yes,” does this linear state feedback stabilise the linearisation?
3. If the answer to the preceding question is, “Yes,” can the stabilising linear feedback be implemented to stabilise the actual reference trajectory? (cf. Question 5.3)

References

Bianchini, R. M. and Stefani, G. [1993] *Controllability along a trajectory: A variational approach*, SIAM Journal on Control and Optimization, **31**(4), pages 900–927, ISSN: 0363-0129, DOI: 10.1137/0331039.

Brockett, R. W. [1970] *Finite Dimensional Linear Systems*, John Wiley and Sons: NewYork, NY, ISBN: 978-0-471-10585-5.

Bullo, F. and Lewis, A. D. [2005] *Low-order controllability and kinematic reductions for affine connection control systems*, SIAM Journal on Control and Optimization, **44**(3), pages 885–908, ISSN: 0363-0129, DOI: 10.1137/S0363012903421182.

Ikeda, M., Maeda, H., and Kodama, S. [1972] *Stabilization of linear systems*, Journal of the Society of Industrial and Applied Mathematics, Series A Control, **10**(4), pages 716–729, DOI: 10.1137/0310052.

Kalman, R. E. [1960] *Contributions to the theory of optimal control*, Boletín de la Sociedad Matemática Mexicana. Segunda Serie, **5**, pages 102–119, URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.26.4070 (visited on 07/10/2014).

Millionshchikov, V. M. [1986a] *Linear systems of differential equations and automorphisms of vector bundles. I*, Izvestiya Natsional′noĭ Akademii Nauk Respubliki Kazakhskoĭ. Seriya Fiziko-Matematicheskaya, **3**, pages 25–30, ISSN: 0002-3191.

Millionshchikov, V. M. [1986b] *Linear systems of differential equations and automorphisms of vector bundles. II*, Izvestiya Natsionalʹnoĭ Akademii Nauk Respubliki Kazakhskoĭ. Seriya Fiziko-Matematicheskaya, **5**, pages 23–27, ISSN: 0002-3191.

Sontag, E. D. [1998] *Mathematical Control Theory, Deterministic Finite Dimensional Systems*, 2nd edition, number 6 in Texts in Applied Mathematics, Springer-Verlag: New York/Heidelberg/Berlin, ISBN: 978-0-387-98489-6.

Sussmann, H. J. [1998] *An introduction to the coordinate-free maximum principle*, in *Geometry of Feedback and Optimal Control*, edited by B. Jakubczyk and W. Respondek, Monographs and Textbooks in Pure and Applied Mathematics, pages 463–557, Dekker Marcel Dekker: New York, NY, ISBN: 978-0-8247-9068-4.