Algorithmic Transverse Feedback Linearization

by

Rollen S. D'Souza

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2022

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Martin Guay, Professor, Dept. of Chemical Engineering Queen's University
Supervisor:	Christopher Nielsen, Professor, Dept. of Electrical and Computer Engineering University of Waterloo
Internal Member:	Daniel Miller, Professor, Dept. of Electrical and Computer Engineering University of Waterloo
Internal-External Member:	Jun Liu, Associate Professor, Dept. of Applied Mathematics University of Waterloo
Other Member:	Andrew Heunis, Professor, Dept. of Electrical and Computer Engineering University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The feedback equivalence problem, that there exists a state and feedback transformation between two control systems, has been used to solve a wide range of problems both in linear and nonlinear control theory. Its significance is in asking whether a particular system can be made equivalent to a possibly simpler system for which the control problem is easier to solve. The equivalence can then be utilized to transform a solution to the simpler control problem into one for the original control system.

Transverse feedback linearization is one such feedback equivalence problem. It is a feedback equivalence problem first introduced by Banaszuk and Hauser for feedback linearizing the dynamics transverse to an orbit in the state-space. In particular, it asks to find an equivalence between the original nonlinear control-affine system and two subsystems: one that is nonlinear but acts tangent to the orbit, and another that is a controllable, linear system and acts transverse to the orbit. If this controllable, linear subsystem is stabilized, the original system converges upon the orbit.

Nielsen and Maggiore generalized this problem to arbitrary smooth manifolds of the state-space, and produced conditions upon which the problem was solvable. Those conditions do not help in finding the specific transformation required to implement the control design, but they did suggest one method to find the required transformation. It relies on the construction of a mathematical object that is difficult to do without system-specific insight.

This thesis proposes an algorithm for transverse feedback linearization that computes the required transformation. Inspired by literature that looked at the feedback linearization and dynamic feedback linearization problems, this work suggests turning to the "dual space" and using a tool known as the derived flag. The algorithm proposed is geometric in nature, and gives a different perspective on, not just transverse feedback linearization, but feedback linearization problems more broadly.

Acknowledgements

This thesis would not have been an option for me had it not been for the incredible support I received from my family (Roque, Daizy and Deepti) during my undergraduate years. It is with their support, especially the support of my mother Daizy, that I was able to focus on what mattered in the moment and convince others that I could succeed in graduate school. Certainly their support throughout my initial master's which transformed into this doctoral program played a key role in my success.

That support alone was not enough, however. Patrick Lam's decision to offer me a position as an undergraduate teaching assistant for an introductory course was a pivotal moment in my life. The responsibility I took on in that moment consumed me, and helped put my life on track in my last, additional year of undergraduate education. It is arguably one of the most significant moments in my life, and I will never forget the memories made with the associated Software Engineering Class of 2021. They deserve a special shout-out for reminding me what I enjoyed most (teaching), and for reimbuing the passion I once held for engineering.

Christopher Nielsen, my supervisor, deserves a special thank you as well, for accepting me even with my spotty undergraduate record. The chance he took gave me an opportunity to remake my career. Not sure either of us expected the outcome we received. On a professional note, thank you to my committee members — Andrew Heunis, Martin Guay, Daniel Miller and Jun Liu — for their useful comments. A special thank you to Andrew Heunis who met with me on a few occassions and whose musings on the academe were always something to look forward to.

In graduate school, I made only a few friends, but they were all very dear to me. My first year of reading group with Phillip, Safoan and Joel was very formative. That was the best reading group I ever had, and I miss it to this very day. Thank you for the great memories we all had together, even though it was short. Thank you to Orlando, whom I shared an office with and spent way too much on coffee with in the process of procrastination. I always looked forward to having Frank and Hassaan pass by the office. A special thank you to Robbert who visited us for a brief time, but worked on a critical side-project of mine. Not only was your contribution significant for my career, but the time we spent together in the office was great.

Unfortunately, the pandemic struck and I was not able to make as many memories as I had hoped with Shaundell and Alex. Nevertheless, there were some fun moments and I am thankful for those rare encounters. Our chats after reading group on Teams were very enjoyable, and kept all of us slightly more sane I hope.

Finally, closer to home, I thank my partner Brooke for giving me company, especially during the worst terms of the pandemic where writing took over. Although the contributions from you are not easily enumerable, one item is worth laughing over: Genshin Impact with you probably did more wonders for my thesis than I would like to admit.

Dedication

To my dear nephews Milan and Luka, who may grow up to wonder what spirit consumed their Uncle Ron during their early years to endure the process of thesis production.

Table of Contents

List of Figures

Chapter 1

Introduction

Nonlinear control theory has a long geometric tradition. Hermann first injected geometric ideas into control theory when studying the accessibility problem [23]. Within decades, a flurry of articles from a number of notable academics used the language of differential geometry to aid in the design of feedback linearizing controllers, ascertain controllability, and find controlled-invariant sets and distributions [8, 21, 22, 24, 25]. These results have since been synthesized into mainstream nonlinear control texts like [27, 42, 43].

The state-space, exact feedback linearization problem is one issue tackled and resolved by taking a geometric approach. The problem asks whether there exists a change of coordinates $\xi = \Phi(x)$ and feedback $u = \alpha(x) + \beta(x)\, v$ that locally transforms the nonlinear control-affine system

$$\dot{x}(t) = f(x(t)) + g(x(t))\, u(t), \tag{1.1}$$

into a linear, controllable system

$$\dot{\xi}(t) = A\, \xi(t) + B\, v(t).$$

This idea is depicted visually in Figure 1.1. Hunt, Su and Meyer presented in [26] necessary and sufficient conditions upon which such a transformation exists. If the conditions are satisfied, then the change of coordinates Φ (and, consequently, the feedback $u = \alpha(x) + \beta(x)\, v$) can be found by solving a large system of partial differential equations (PDEs).

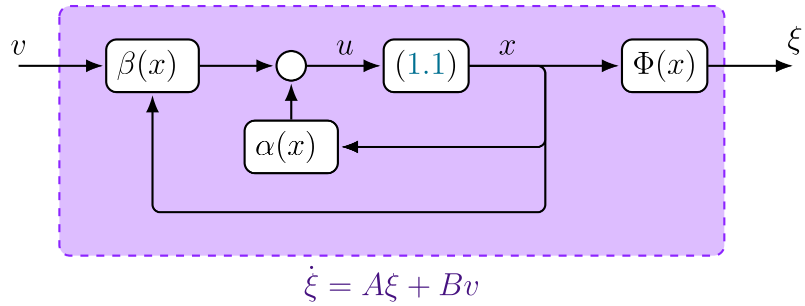


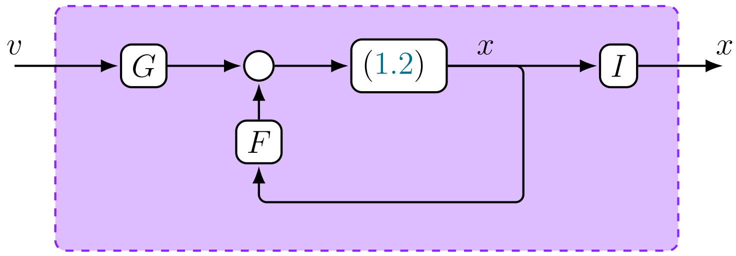
Figure 1.1: *State-space, exact feedback linearization is a feedback equivalence problem wherein a change of coordinates* $\xi := \Phi(x)$ *and feedback* $u = \alpha(x) + \beta(x)\, v$ *transforms the system into a controllable linear system with a new input and new state.*

The difficulty of solving this system of PDEs was clear, and within a few years attempts were made to alleviate the difficulty. The first attempt was made in [15] by Gardner and Shadwick in the case where the controllability indices of (A, B) in the target system are distinct. They completed their work in [16] for the general case. The G.S. (Gardner-Shadwick) algorithm[1], their proposed approach, minimizes the number of integrations required to find the transformation Φ. This remarkable work soon inspired attempts at algorithmic solutions to other nonlinear control design problems such as dynamic feedback linearization [4, 48], differential flatness [44], and even the linearization of nonlinear discrete-time systems [3]. All of these are feedback equivalence problems.

1.1 Feedback Equivalence

The notion of feedback equivalence broadly speaking asks: under what conditions can solutions of one controlled dynamical system be transformed into solutions of another? The transformation that does this is known as a feedback transformation.

[1]The literature is not clear on what "G.S." really stood for. We have taken the liberty of assuming that it stands for the original authors of the algorithm. Had it stood for anything else, the original authors would have stated so.



$\dot{x} = \overline{A}x + Bv$

Figure 1.2: *The pole placement problem of linear control theory is a feedback equivalence problem where the target system is another linear control system with system matrix* $\overline{A} = A + BF$ *that has a desired characteristic polynomial.*

Definition 1.1.1 (Feedback Transformation) . Let $x \in \mathbb{R}^n$, $z \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $v \in \mathbb{R}^m$. A *(regular) feedback transformation* (Φ, α, β) is a diffeomorphism $\Psi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$ that takes the form $(z, v) = \Psi(x, u) = (\Phi(x),\, \beta(x)^{-1}(u - \alpha(x)))$.

Feedback transformations are grounded in practical aspects of control design where $z = \Phi(x)$ changes the (full-state) output of a system block while $u = \alpha(x) + \beta(x)\, v$ transforms a virtual input v into the plant's input u that must be applied to the real system. A number of important modern linear control theory problems may be cast in this framing. For instance, consider the single-input, linear control system,

$$\dot{x}(t) = A\, x(t) + B\, u(t), \tag{1.2}$$

where $A : \mathbb{R}^n \to \mathbb{R}^n$ and $B : \mathbb{R} \to \mathbb{R}^n$. When is it possible to define a state feedback $u = F\,x + G\,v$ so that the eigenvalues of the closed-loop system matrix $A + BF$ are at some desired location? This is the pole placement problem, and it can be cast as a feedback equivalence problem. In particular, we restrict $\alpha(x) = F\,x$, $\beta(x) = G$ and $z = T\,x$ to ensure the new system in terms of state z and virtual input v is a linear time-invariant control system. Figure 1.2 depicts the equivalence problem as a feedback diagram. It is a well-understood result that the necessary and sufficient conditions for the solvability of this equivalence problem is that the system (1.2) is controllable [49]. Controllability ensures that there exists a feedback transformation so that the system (1.2) under the new

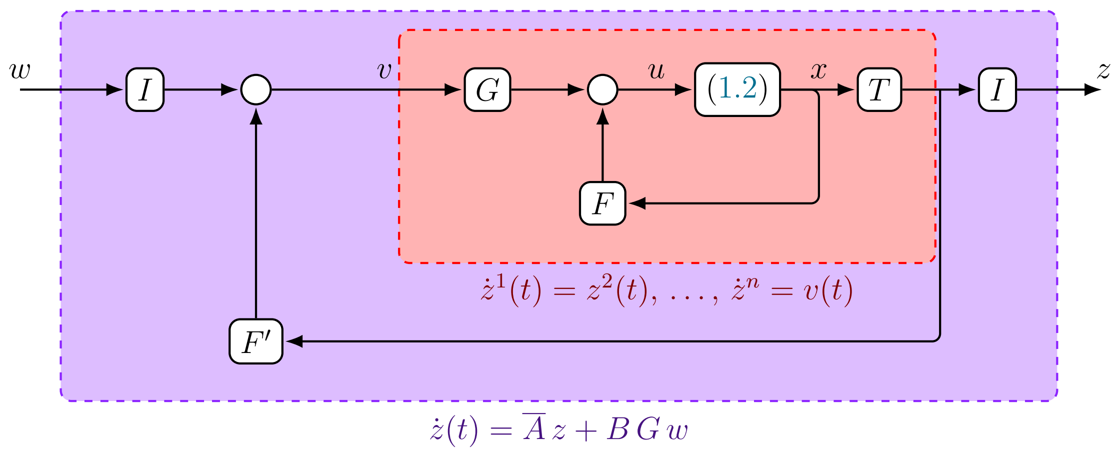


Figure 1.3: *The pole placement problem can be broken down into two equivalence problems that are easier to solve: putting a system into controllable canonical form, and then pole placement of that system so $\overline{A}$ has the desired characteristic polynomial.*

coordinates looks like a chain of integrators,

$$\begin{aligned}\dot{z}^1(t) &= z^2(t),\\ &\vdots\\ \dot{z}^{n-1}(t) &= z^n(t),\\ \dot{z}^n &= v(t).\end{aligned}$$

Such a system is said to be in *Brunovský normal form*. It is easy to place the poles of a system in Brunovský normal form. That is, it is easy to find the required feedback transformation for the system in the new coordinates z with virtual input v that will result in a system matrix with a desired characteristic polynomial. Together, these two feedback transformations solve the pole-placement feedback equivalence problem. This process is depicted in Figure 1.3.

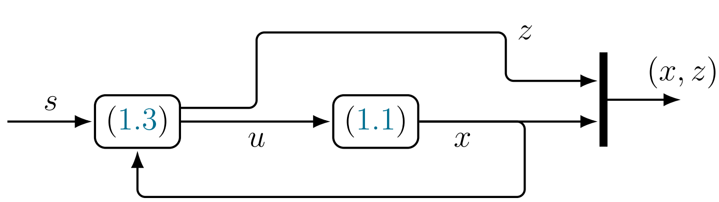
The solvability conditions for feedback equivalence problems can usually be found by observing the properties that must be invariant under such a transformation. For example, for multi-input, linear, time-invariant control systems, the controllability indices of (A, B) are invariant under feedback transformations that preserve the linear, time-invariant property of the system. When the system is *in* the Brunovský normal form, the controllability indices indices are easy to compute, and directly describe the number and length of the integrator chains. This hints that a necessary and sufficient condition for a controllable, multi-input linear system to be equivalent to some chains of integrators, they must have equal controllability indices. The result in [49, Corollary 5.3] says precisely this. It is for this reason, and for its practical power, that feedback equivalence problems are useful. They allow us to determine testable conditions for solvability, and have practical utility in control design as they can be used directly to simplify or change the system dynamics.

This general approach to control problems of linear systems made its way into nonlinear control theory as well, starting with the state-space, exact feedback linearization problem; this feedback equivalence problem was described earlier and depicted in Figure 1.1. We leave the deeper theory of this problem to Section 2.1.

It is uncommon for a nonlinear system to be exactly feedback linearizable. A number of strategies emerged from this observation. Dynamic feedback linearization enlarged the class of systems that could be transformed into controllable, linear systems under feedback transformations by considering the addition of a *dynamic, nonlinear* precompensator. Informally, the dynamic feedback linearization problem asks whether there exists a dynamic, nonlinear precompensator,

$$\begin{aligned} \dot{z}(t) &= q(x(t), z(t)) + r(x(t), z(t))s(t) \\ u(t) &= a(x(t), z(t)) + b(x(t), z(t))s(t) \end{aligned} \tag{1.3}$$

so that the combined nonlinear dynamics from the virtual input s to the combined state (x, z),



could be feedback linearized exactly through a feedback transformation. Figure 1.4 depicts the feedback transformed, dynamically precompensated system that is rendered feedback

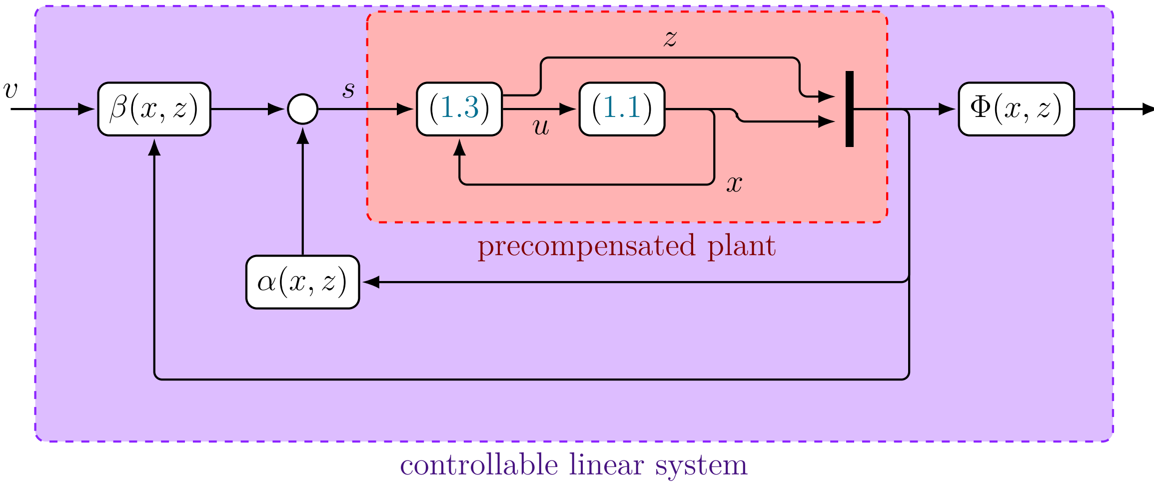


Figure 1.4: *The dynamic feedback linearization problem asks to find a dynamic precompensator that, together with the original nonlinear plant, can be exactly feedback linearized.*

equivalent to a controllable, linear system. Solving the dynamic feedback linearization problem is much more difficult than the exact feedback linearization problem since we must not only find a feedback transformation to feedback linearize a system, but we must make a clever choice of dynamic precompensation (1.3) to make that possible. Necessary conditions for solvability were presented in [47]. Necessary and sufficient conditions were presented a few years later by Guay et al. in [18]. This work was inspired — in a manner similar to this thesis — by the work of Gardner and Shadwick in [16]; Guay et al. bootstrapped off the conditions developed by Gardner and Shadwick to implicitly provide a procedure that computes the required precompensator (1.3) and feedback transformation.

An alternative strategy emerged at the same time that leveraged a property some larger family of systems satisfy: differential flatness [13]. A system (1.1) is differentially flat if there exists integers $k \geq 0$ and $\ell \geq 0$ as well as a change of coordinates of the type $y = \Phi(x, u, \dot{u}, \ldots, u^{(k)})$ and $(x, u) = \Psi(y, \dot{y}, \ldots, y^{(\ell)})$. The output y is called a flat output. The vector y and its first ℓ derivatives along the solutions to the nonlinear control system (1.1) uniquely determine the state and control action. It was long believed that differentially flat systems *are* dynamically feedback linearizable systems, but it was not clear until recently how to take a dynamically feedback linearized system and construct a flat output. Note that the output y in a flat system is only a function of the state and derivatives of the input. It is *not* a function of the internal state of the dynamic precompensator, unlike what is depicted in Figure 1.4 where the output is a function of z as well as x. The equivalence of the dynamic feedback linearization and differential flatness problems was established in [32].

The theory behind flatness takes a more general approach to the equivalence of nonlinear control systems, and can even be applied to systems that are not control-affine. This general approach can be powerful. Knowing Ψ and Φ is more useful than knowing how to dynamically feedback linearize a system when considering planning problems, e.g. trajectory planning. The flatness problem indicates not only what signal to plan a trajectory for, i.e., y, but also to what order of smoothness it must have to ensure there exists an appropriate control action u that renders it invariant. Unfortunately, flatness-based methods do come at the cost of theoretical complexity. More recent articles, such as [17, 38, 44, 45], have made steady progress at applying flatness techniques to derive necessary and sufficient conditions on dynamic feedback linearizability for multi-input control systems that are not necessarily control-affine. However, flatness is difficult to test and use for solving dynamic feedback linearization of a system with many inputs. Moreover, any nonlinear system can be made control-affine by using dynamic extension.

One could simply accept that a nonlinear control system is not feedback linearizable, and instead ask to find the *largest* subsystem that is feedback linearizable. The linearized subsystem can be stabilized with linear state-feedback. The system dynamics restricted to the set where the feedback linearizing output remains identically zero, known as the *zero dynamics manifold*, may not be stable but, assuming it is, the total dynamics can be asymptotically stabilized. This method is known as partial feedback linearization, so as to indicate that only a part of the system is transformed into a controllable, linear system. The problem was first proposed by Krener et al. in [30] and resolved in the single-input case. The conditions upon which the full multi-input problem could be solved were presented in [34] by Marino alongside a procedure to find the required feedback transformation. The feedback equivalence problem for partial feedback linearization is depicted by Figure 1.5.

A closely related problem arose in the late 1990s with application to motion control

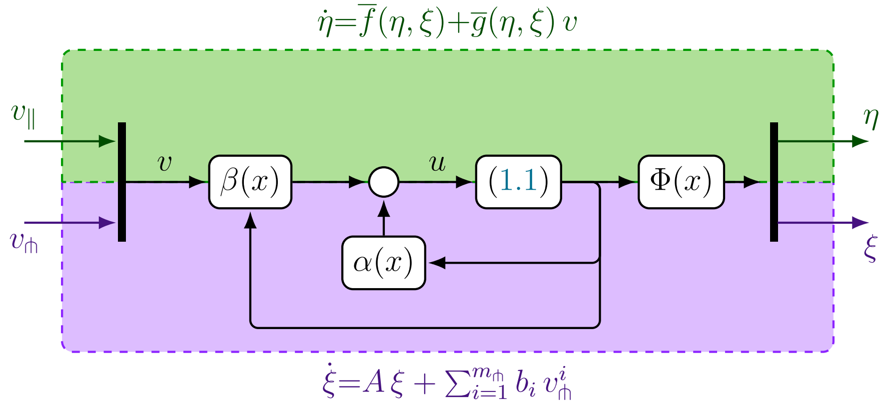


Figure 1.5: *Partial feedback linearization can be viewed as the feedback equivalence problem of finding the largest, linear, controllable subsystem in a nonlinear, control-affine system. The η-dynamics are affected by both $v_{\pitchfork}$ and $v_{\parallel}$ while the linear ξ-dynamics are only affected by $v_{\pitchfork}$.*

problems: given a controlled invariant submanifold[2] N of the state-space, when is it possible to feedback linearize those dynamics that act transverse[3] to N? This problem, known as the transverse feedback linearization (TFL) problem, differs from partial feedback linearization in that it *starts* with a desired zero dynamics manifold N and asks to find a feedback transformation that locally linearizes dynamics transverse to it. Like exact, partial and dynamic feedback linearization, transverse feedback linearization is a feedback equivalence problem: it asks to find a change of coordinates $(\eta, \xi) := \Phi(x)$ and feedback $u = \alpha(x) + \beta(x)\, v$ so that, in the new coordinates, there is a linear subsystem (the subsystem associated with ξ) whose dynamics are precisely the dynamics that are transverse to N. This factors the dynamics of the nonlinear system in a manner suited to the objective.

This feedback equivalence problem is also depicted by Figure 1.5 except for one critical difference. In transverse feedback linearization, the nonlinear control system is expected to be partially feedback linearizable "in the right way." That is, the linearized states ξ are equal to zero if, and only if, the plant's state resides in the target submanifold N. The control engineer *begins* with the set N and finds a partial feedback linearization whose zero

[2] A set $\mathsf{N} \subseteq \mathbb{R}^n$ is controlled invariant if there exists an input u so that if the system starts in N, it evolves in N for all future time.

[3] The dynamics are not tangent to the set.

dynamics manifold — i.e., the submanifold where $\xi = 0$ — is the set N. The algorithm presented in [34] for partial feedback linearization provides no such guarantee that the zero dynamics manifold will coincide with any such desire. The zero dynamics manifold takes on an arbitrary shape dependent on choices taken in the procedure in a non-trivial way. The value of starting with the desired zero dynamics manifold *first*, and asking the transverse dynamics be partially feedback linearized is clear: when the control specification is satisfied by driving the state to the set, for example, asking a robot to follow a path in its output space. As a result, transverse feedback linearization acts as one approach to solving what is known as the set stabilization problem.

1.2 The Set Stabilization Problem

Consider now the multi-input, nonlinear, control-affine system,

$$\dot{x}(t) = f(x(t)) + \sum_{i=1}^{m} g_i(x(t))\, u^i(t), \tag{1.4}$$

where $x(t) \in \mathbb{R}^n$ is the system state at time t and $u^1(t), \ldots, u^m(t) \in \mathbb{R}$ are the control actions at time t. The vector field $f : \mathbb{R}^n \to \mathsf{T}\mathbb{R}^n$ is known as the drift vector field. It determines the behaviour of the system under zero control action. The vector fields g_1, $\ldots$, $g_m : \mathbb{R}^n \to \mathsf{T}\mathbb{R}^n$ are the control vector fields that determine how each input drives the system.

Example 1.1 (Model of a Car) . The motion of a car-like vehicle can be modelled by the kinematic bicycle model with dynamic extension given by

$$\begin{aligned}\dot{x}(t) &= \begin{pmatrix} x^3(t)\cos\left(x^4(t)\right) \\ x^3(t)\sin\left(x^4(t)\right) \\ x^6(t) \\ \frac{x^3(t)}{\ell}\tan\left(x^5(t)\right) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0\\0\\0\\0\\1\\0 \end{pmatrix} u^1(t) + \begin{pmatrix} 0\\0\\0\\0\\0\\1 \end{pmatrix} u^2(t), \\ &=: f(x(t)) + g_1(x(t))\, u^1(t) + g_2(x(t))\, u^2(t),\end{aligned} \tag{1.5}$$

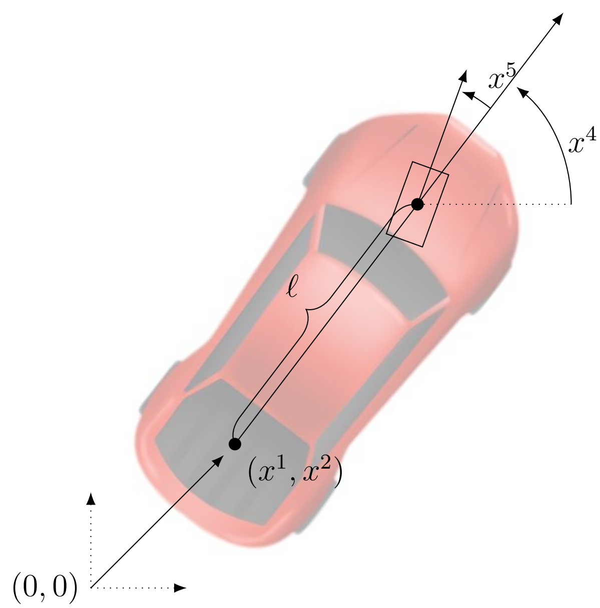


Figure 1.6: *Depiction of the kinematic bicycle model of a vehicle described in Example 1.1.*

where $\ell > 0$ is the distance between the front and rear axles. Part of the state is depicted in Figure 1.6. The position of the car in an inertial frame is $(x^1(t), x^2(t))$ and its heading angle (relative to the x^1-axis) is $x^4(t)$. The forward speed is $x^3(t)$ with a forward acceleration $x^6(t)$ which is controlled by input $u^2(t)$. The state $x^5(t)$ is the steering angle with the angular velocity of the steering angle controlled by input $u^1(t)$. ◀

It is often of great use to have a system's state driven towards a subset of possible states. In some ways this is reflective of the common control design specification that is often found in classical linear control. Indeed, the geometric approach of Wonham [49] is, in his own words, "first characterize solvability as a verifiable property of some constructable state subspace." Then many of the synthesizing approaches therein result in the state being driven towards, or along, these subspaces. The general class of set stabilization techniques involve bringing a system towards a given parameterized set of states without *a priori* imposing a time parametrization. This frees the control designer from accidentally making the problem infeasible. There are a variety of set stabilization techniques, e.g., [6, 19, 29, 40]. One example of a nonlinear control design problem in which set stabilization naturally arises is the so-called "path following" problem illustrated in the coming examples.

Example 1.2 (Path Following for a Car) . Consider the model proposed for vehicular motion in Example 1.1. We would like the *front-axle* of the car to approach a circular path of radius $R > 0$ in the position-plane at a constant forward speed $x^3 = 1$, but do not wish to impose (explicitly) any specific time-parameterization of the motion of the body along the path. This is equivalent to making the state vector $x(t)$ approach the set

$$\mathsf{G} = \left\{x \in \mathbb{R}^5 \colon (x^1 + \ell\cos(x^4))^2 + (x^2 + \ell\sin(x^4))^2 - R^2 = 0, x^3 = 1\right\}.$$

◀

Set stabilization approaches to path following treat the path as a set in its own right, and provide a controller that makes (1) this set invariant and (2) locally (or globally) stable. In doing so the controller brings the system to the path and provides a guarantee, in the absence of disturbances, that the physical system will never leave the objective set. Formally, the set stabilization problem is as follows. Let $\mathsf{N} \subseteq \mathbb{R}^n$ be a subset of the state-space.

Problem 1.2.1 (Set Stabilization)**.** *Find a control law u so that the following holds on some open set U containing N: for any open set $\mathsf{V} \supset \mathsf{N}$ of U, there exists an open set $\mathsf{V}_0 \supseteq \mathsf{N}$ so that solutions of* (1.4) *starting in V_0 stay in V.*

A necessary condition to solve this problem is that the control law render the set N invariant: systems that start in N, stay in N for all future time. When this specification is achievable, i.e., there exists $u^1, \ldots, u^m$ so that N is invariant, we say N is a *controlled-invariant set*.

More often than not, the set stabilization problem is augmented with the demand that the set be made attractive. That is, require solutions stay in the neighbourhood of N and, in some sense, tend towards N. When there exists $u^1, \ldots, u^m$ so that solutions of (1.4) tend towards N, we say that N can be made *attractive*. It is clear that, when N is a one-point set,

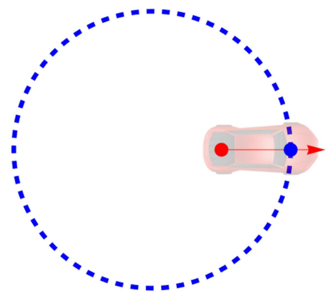


Figure 1.7: *A configuration where the the center of the vehicle's front axle (red disk) resides on the path with the desired forward speed, but its instantaneous velocity (red vector) causes it to leave the path regardless of control action.*

this problem subsumes the classic asymptotic stabilization problem of ensuring dynamics tend to a single-point in the state-space.

As the next example illustrates, one must carefully construct the controlled-invariant set: the controlled-invariant set is normally a smaller set contained in the set of states that meet the control objective.

Example 1.3 (Path Following for a Car) . Recall the objective from Example 1.2 where we wish to bring the vehicle towards a circular path while maintaining regulating a constant forward speed.

Initializing the car on the set G is not sufficient to ensure the car remains on G as time advances, no matter the choice of control input. For example, consider configuration

$$x_0 = (R - \ell, 0, 1, 0, 0, 0).$$

In this configuration, the vehicle is on the east end of the circle, facing east. This configuration resides on the circular path and the vehicle is moving at the desired forward speed, but immediately leaves the path in future time for all possible inputs u as the vehicle has an instantaneous velocity that is not tangent to the path. This is visually depicted by Figure 1.7. We can therefore say that the set of all points on the path G is *not* controlled-invariant.

On the other hand, there does exist a smaller set of configurations, contained in G, where the vehicle resides on the path and can, under appropriate control action, stay on it for all future time; in other words, there exists a controlled-invariant set contained in G that describes motion along the path. It can be shown that the set

$$\mathsf{N} := \left\{x \in \mathbb{R}^5 \colon y_1(x) = y_2(x) = y_3(x) = y_4(x) = y_5(x) = 0\right\} \subset \mathsf{G},$$

where,

$$\begin{aligned}
y_1(x) &:= (x^1 + \ell \cos x^4)^2 + (x^2 + \ell \sin x^4)^2 - R^2, \\
y_2(x) &:= x^3 - 1, \\
y_3(x) &:= x^1 \cos(x^4) + x^2 \sin(x^4), \\
y_4(x) &:= \ell \cos(x^5) - \sin(x^5)(x^1 \sin(x^4) - x^2 \cos(x^4), \\
y_5(x) &:= x^6,
\end{aligned}$$

can be rendered controlled-invariant through the (non-unique) state feedback $u_*(x) = (0, 0)$. The first constraint demands the vehicle's forward axle converge upon the unit circle while the second constraint asks that the vehicle's speed be 1. These are the very same constraints that appeared in our problem specification. The third constraint asks that the vehicle's rear axle itself travel so that its velocity is orthogonal to its position; that is, it achieves circular motion of unknown radius. The fourth constraint can be derived by computing Lie derivatives of the first constraint along solutions of (1.7). One point in this set is

$$x_0 = \begin{pmatrix} -\ell \cos\left(\arcsin\left(\frac{\ell}{R}\right)\right) \\ R - \ell \sin\left(\arcsin\left(\frac{\ell}{R}\right)\right) \\ 1 \\ \arcsin\left(\frac{\ell}{R}\right) \\ -\arctan\left(\frac{\ell}{\sqrt{R^2-\ell^2}}\right) \\ 0 \end{pmatrix},$$

depicted graphically in Figure 1.8. It is less obvious, but this set is the largest, in terms of set inclusion, controlled invariant set that contains x_0 and is contained in G. Configurations

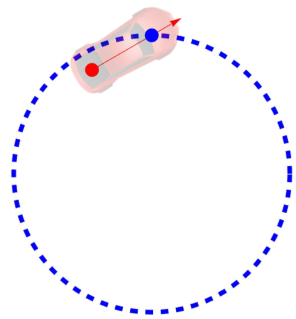


Figure 1.8: *A configuration of the vehicle where the front-axle resides on the path and can be kept on the path for all future time with an appropriate control action. Intuitively, this will correspond to a constant, non-zero steering angle keeping the vehicle in a constant-rate right turn.*

in N describe all possible configurations on the path where the vehicle can be made to remain on the path in for all future time by appropriate control action. ◀

Having a controlled-invariant set, though necessary, is still not sufficient to ensure that the set-stabilization problem is solvable. Making a set attractive itself can be a challenge, and depends on the dynamics in a neighbourhood of the set. Attractiveness is often ensured either by Lyapunov [50], passivity [14, 51], or linearization-based methods [28, 46]. Linearization-based methods either approximate the dynamics transverse to the set with a linear dynamical system (Jacobian-based linearization), or attempt to partially feedback linearize the system in a way so that the resulting linear subsystem's dynamics corresponds to the transverse dynamics to the set.

Transverse feedback linearization is precisely the latter problem. It asks that, given a controlled-invariant set, find a feedback transformation that renders the dynamics transverse to the set linear and controllable. It is at this point that any linear feedback in the new coordinates that renders the linear subsystem asymptotically stable will ensure the set is locally attractive.

1.3 Local Transverse Feedback Linearization

When the target set N is an orbit, necessary and sufficient conditions upon which one could feedback linearize the transverse dynamics were found by Banaszuk and Hauser in [5]. Nielsen and Maggiore then presented conditions for when N is a general submanifold [40]. We now remind the reader of these results, but, before doing so, formally state the local transverse feedback linearization problem (TFL). Take the system dynamics to be (1.4). Let N be a closed, embedded n^*-dimensional submanifold of $\mathbb{R}^n$ and fix a point $x_0 \in \mathsf{N}$. Suppose N is rendered locally, controlled-invariant by a state feedback $u_*(x)$. The local transverse feedback linearization problem is as follows.

Problem 1.3.1 (Local Transverse Feedback Linearization (TFL))**.** *Find a (local) feedback transformation* (Φ, α, β) *defined on an open set* U *of* x_0 *so that, if we write* $(\eta, \xi) = \Phi(x)$ *and* $u = \alpha(x) + \beta(x) \left[v_\parallel^\top \; v_\pitchfork^\top \right]^\top$*, the nonlinear control system* (1.4) *takes the form*

$$\begin{aligned}\dot{\eta}(t) &= \overline{f}(\eta(t), \xi(t)) + \sum_{i=1}^{m_\parallel} \overline{g}_{\parallel,i}(\eta(t), \xi(t))\, v_\parallel^i + \sum_{i=1}^{m_\pitchfork} \overline{g}_{\pitchfork,i}(\eta(t), \xi(t))\, v_\pitchfork^i, \\ \dot{\xi}(t) &= A\,\xi(t) + \sum_{j=1}^{m} b_j\, v_\pitchfork^j(t).\end{aligned} \tag{1.6}$$

and $(A, \left[b_1 \cdots b_m \right])$ *is in Brunovský normal form*[4]*. Moreover, in* (η, ξ)*-coordinates, the manifold* N *is locally given by*

$$\Phi(\mathsf{U} \cap \mathsf{N}) = \left\{ (\eta, \xi) \in \Phi(\mathsf{U}) \colon \xi = 0 \right\}.$$

Linear control design may be used to design a control law $v_\pitchfork$ for the linear subsystem that ensures $\xi(t) \to 0$ as $t \to \infty$ for all initial conditions starting in $\Phi(\mathsf{U})$ as long as the state $x(t)$ resides in U.

It turns out that solving the TFL problem is equivalent to finding an output of suitable vector relative degree that vanishes on N. This was the view championed by Isidori with regards to the problem of feedback linearization; this is discussed further in Chapter 2. In the spirit of this idea, the following theorem was established in [41]. We provide the statement as stated in [40].

[4]The reader should consult Figure 1.5.

Theorem 1.3.1 ([40, Theorem 3.1]). *The local transverse feedback linearization problem is solvable at* x_0 *if, and only if, there exists an open subset* $\mathsf{U} \subseteq \mathbb{R}^n$ *of* x_0*, natural numbers* $\rho_0 \in \mathbb{N}$, $\kappa_1, \ldots, \kappa_{\rho_0} \in \mathbb{N}$*, and a smooth function* $h : \mathsf{U} \to \mathbb{R}^{\rho_0}$ *so that:*

(1) $\mathsf{U} \cap \mathsf{N} \subset h^{-1}(0)$*, and*

(2) the system (1.4) *with output* $y = h(x)$ *yields vector relative degree* $(\kappa_1, \ldots, \kappa_{\rho_0})$ *at* x_0 *with* $\sum_{i=1}^{\rho_0} \kappa_i = n - n^*$.

The theorem shows that the transverse feedback linearization problem is equivalent to the zero dynamics assignment problem with relative degree (see Definition 2.1.3): find an output h for system (1.4) that yields a well-defined relative degree and whose zero dynamics manifold locally coincides with N. The output h is called a *(local) transverse output* with respect to N at x_0, or a transverse output for short. Theorem 1.3.1 does not provide conditions on when the problem can be solved nor does it tell us how to find a transverse output h.

In [40], checkable, necessary and sufficient conditions for the solvability of the TFL problem were presented. Before presenting this theorem, we need to recall the classic control distributions, found as $\mathscr{M}^l$ in [34], given by,

$$\mathscr{G}_x^{(k)} := \operatorname{span}_{\mathbb{R}}\{\operatorname{ad}_f^j g_i|_x : 1 \leq i \leq m, 0 \leq j \leq k\}, \quad k \geq 0, x \in \mathbb{R}^n.$$

When $\mathscr{G}^{(k)}$ is a smooth and regular distribution, the associated $C^\infty(\mathbb{R}^n)$-submodule,

$$\mathcal{G}^{(k)} := \Gamma^\infty(\mathscr{G}^{(k)}) = \operatorname{span}_{C^\infty(\mathbb{R}^n)}\{\operatorname{ad}_f^j g_i \colon 1 \leq i \leq m, 0 \leq j \leq k\} \subseteq \Gamma^\infty(\mathsf{T}\mathbb{R}^n), \quad k \geq 0,$$

is, by construction, finitely, non-degenerately generated. The necessary and sufficient conditions for the solvability of the TFL problem are as follows.

Theorem 1.3.2 ([40, Theorem 3.2]). *Suppose that* $\operatorname{inv}(\mathscr{G}^{(i)})$, $i \in \{1, \ldots, n - n^\star - 1\}$, $0 \leq k \leq n - n^*$ *are smooth and regular distributions in an open set containing* x_0*. The transverse feedback linearization (TFL) problem is solvable at* x_0 *if, and only if,*

(1) $\dim(\mathsf{T}_{x_0}\mathsf{N} + \mathscr{G}_{x_0}^{(n-n^*-1)}) = n$*, and*

(2) there exists an open set $\mathsf{U} \subseteq \mathbb{R}^n$ *containing* x_0 *so that for every* $\kappa \in \{1, \ldots, n - n^* - 1\}$ *and for all* $x \in \mathsf{U} \cap \mathsf{N}$,

$$\dim\left(\mathsf{T}_x\mathsf{N} + \mathrm{inv}(\mathscr{G}^{(\kappa-1)})_x\right) = \dim\left(\mathsf{T}_x\mathsf{N} + \mathscr{G}_x^{(\kappa-1)}\right).$$

Condition (1) is known as the controllability condition, while condition (2) is known as the involutivity condition. The assumptions of Theorem 1.3.2 are checkable, but its proof does not provide a procedure for finding the transverse output even in the single-input case as the next example illustrates.

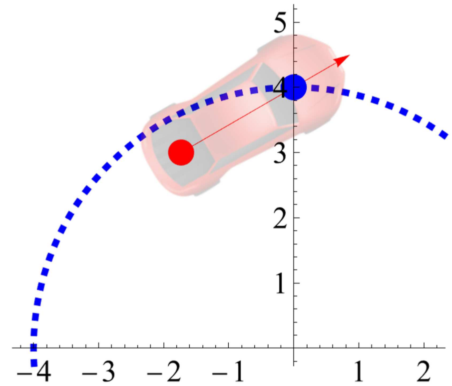
Example 1.4 (Path Following for a Car) . Recall the model from Example 1.1 and the set-stabilization problem from Example 1.3. If we set $u^2(t) = 0$ — i.e. the forward speed is fixed — and take $\ell = 2$, then we can consider a simplified single-input system,

$$\dot{x}(t) = \begin{pmatrix} x^3(t)\cos\left(x^4(t)\right) \\ x^3(t)\sin\left(x^4(t)\right) \\ 0 \\ \frac{x^3(t)}{2}\tan\left(x^5(t)\right) \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} u^1(t). \tag{1.7}$$

Consider the 2-dimensional submanifold ($n^* = 2$),

$$\begin{aligned} \mathsf{N} &= \left\{x \in \mathbb{R}^5 \colon d_1(x) = d_2(x) = d_3(x) = 0\right\}, \\ d_1(x) &= (x^1 + 2\cos x^4)^2 + (x^2 + 2\sin x^4)^2 - 16, \\ d_2(x) &= x^1\ \cos(x^4) + x^2\ \sin(x^4), \\ d_3(x) &= 2\cos(x^5) - \sin(x^5)(x^1\ \sin(x^4) - x^2\ \cos(x^4)), \end{aligned}$$

that can be rendered locally controlled-invariant by $u_*(x) = 0$. We wish to show that, in an open neighbourhood of point $x_0 = (-\sqrt{3}, 3, 1, \pi/6, -\pi/6) \in \mathsf{N}$, the system can be transverse feedback linearized with respect to the set N by using Theorem 1.3.2.



To check the conditions of Theorem 1.3.2 we compute the distributions,

$$\mathscr{G}_x^{(0)} = \operatorname{span}_{\mathbb{R}} \left\{ \left.\frac{\partial}{\partial x^5}\right|_x \right\},$$
$$\mathscr{G}_x^{(1)} = \operatorname{span}_{\mathbb{R}} \left\{ \left.\frac{\partial}{\partial x^5}\right|_x, \sec^2(x^5) \left.\frac{\partial}{\partial x^4}\right|_x \right\},$$
$$\mathscr{G}_x^{(2)} = \operatorname{span}_{\mathbb{R}} \left\{ \left.\frac{\partial}{\partial x^5}\right|_x, \sec^2(x^5) \left.\frac{\partial}{\partial x^4}\right|_x - \sin(x^4) \sec^2(x^5) \left.\frac{\partial}{\partial x^1}\right|_x + \cos(x^4) \sec^2(x^5) \left.\frac{\partial}{\partial x^2}\right|_x \right\}.$$

This is performed using [31, Proposition 8.26] to compute the Lie brackets in local coordinates x^i for $\mathbb{R}^5$. The tangent space can be characterized locally using [31, Proposition 5.38] to be, for $x \in \mathsf{N}$,

$$\begin{aligned} \mathsf{T}_x\mathsf{N} &= \operatorname{span}_{\mathbb{R}} \left\{ \left.\frac{\partial}{\partial x^3}\right|_x, x^2 \left.\frac{\partial}{\partial x^1}\right|_x - x^1 \tan(x^4) \left.\frac{\partial}{\partial x^2}\right|_x - \left.\frac{\partial}{\partial x^4}\right|_x \right\} \\ &= \operatorname{span}_{\mathbb{R}} \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} x^2 \\ -x^1 \tan(x^4) \\ 0 \\ -1 \\ 0 \end{pmatrix} \right\}, \end{aligned} \tag{1.8}$$

in the coordinates for $\mathbb{R}^5$. Condition (1) of Theorem 1.3.2 requires $\dim(\mathsf{T}_{x_0}\mathsf{N} + \mathscr{G}_{x_0}^{(2)}) = 5$.

Evaluating $\mathsf{T}_{x_0}\mathsf{N}$ and $\mathscr{G}^{(2)}_{x_0}$ yields, in coordinates,

$$\mathsf{T}_{x_0}\mathsf{N} = \operatorname{span}_{\mathbb{R}}\left\{\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\\0\\-1\\0\end{pmatrix}\right\}, \qquad \mathscr{G}^{(2)}_{x_0} = \operatorname{span}_{\mathbb{R}}\left\{\begin{pmatrix}1\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}\right\}.$$

It is not hard to see that the sum of the above two vector spaces has dimension 5. So condition (1) holds.

To evaluate condition (2), we compute the involutive closures of $\mathscr{G}^{(0)}$ and $\mathscr{G}^{(1)}$. It turns out both are already involutive. Since both $\mathscr{G}^{(0)}_x$ and $\mathscr{G}^{(1)}_x$ are independent with respect to $\mathsf{T}_x\mathsf{N}$ at every point $x \in \mathsf{N}$ near x_0, the dimension of the sum of the subspaces is equal to the sum of their dimensions. It follows that

$$\dim(\mathsf{T}_x\mathsf{N} + \mathscr{G}^{(0)}_x) = 2, \qquad \dim(\mathsf{T}_x\mathsf{N} + \mathscr{G}^{(1)}_x) = 3,$$

on an open set of N containing x_0. This then establishes condition (2). As a result, (1.7) is transverse feedback linearizable to N at x_0. ◀

The next result connects transverse feedback linearization to partial feedback linearization, but it is not a viable solution to the local transverse feedback linearization problem because it relies on constructing a distribution $\mathscr{D}$ satisfying additional conditions for which a procedure is not known. On the other hand, the theorem provides a guideline for finding the transverse output.

Theorem 1.3.3 ([40, Theorem 3.5])**.** *Suppose that* $\operatorname{inv}(\mathscr{G}^{(i)})$, $i \in \{1, \ldots, n - n^\star - 1\}$, $0 \le k \le n - n^*$ *are smooth and regular distributions in an open set containing* x_0*. The local transverse feedback linearization problem is solvable if, and only if, there exists an involutive smooth and regular distribution* $\mathscr{D}$ *defined on an open set* U *containing* x_0 *so that,*

(i) $\mathscr{D}|_{\mathsf{U}\cap\mathsf{N}} = \mathsf{T}(\mathsf{U} \cap \mathsf{N})$,

(ii) $\mathscr{D}$ *is locally controlled invariant*[5] *under the dynamics* (1.4),

(iii) *For all* $x \in \mathsf{U} \cap \mathsf{N}$, $\dim(\mathsf{T}_x\mathsf{N} + \mathscr{G}_x^{(n-n^\star-1)}) = n$, *and*

(iv) *For every* $i \in \{1, \ldots, n - n^\star - 1\}$, $\mathscr{D} + \mathscr{G}^{(i)}$ *is an involutive distribution on* U.

1.4 Motivation and Statement of Contribution

Indeed, if one begins with a function h whose zero level set *contains* N and happens to yield the correct relative degree, then solving the transverse feedback linearization problem is a straight-forward application of Theorem 1.3.1. On the other hand, if no output is known, then the only result to turn to is Theorems 1.3.2–1.3.3. The procedure, based on Theorems 1.3.2–1.3.3, for computing a transverse output requires the designer to first check the conditions of Theorem 1.3.2. Often, in practice, this is easy. If the conditions hold, then one must find a (non-unique) distribution $\mathscr{D}$ satisfying conditions (i), (ii) and (iv) of Theorem 1.3.3. The distribution $\mathscr{D}$ is guaranteed to exist. Given $\mathscr{D}$ and the distributions $\mathscr{G}^{(i)}$, one computes exact one-forms annihilating the distributions $\mathscr{D}+\mathscr{G}^{(k)}$ at specific indices k where the transverse outputs are expected to appear. The difficulty in this process is in constructing $\mathscr{D}$. It isn't as straightforward as it sounds as the next example demonstrates.

Example 1.5 (Path Following for a Car) . In Example 1.4, we showed that (1.7) is transverse feedback linearizable to N at x_0. It then follows from Theorem 1.3.3 that there exists a distribution $\mathscr{D}$ that satisfies the conditions (i), (ii), and (iv); this distribution will allow us to construct the output function with the correct relative degree.

Recall the characterization of the tangent space for N given in (1.8). Any candidate $\mathscr{D}$ for Theorem 1.3.3 must satisfy condition (i), so a natural choice of candidate is a smooth extension of (1.8),

$$\mathscr{D}_x := \operatorname{span}_{\mathbb{R}} \left\{ x^2 \left.\frac{\partial}{\partial x^1}\right|_x + x^2 \tan(x^4) \left.\frac{\partial}{\partial x^2}\right|_x - \left.\frac{\partial}{\partial x^4}\right|_x , \left.\frac{\partial}{\partial x^3}\right|_x \right\}.$$

[5] A distribution $\mathscr{D}$ is said to be controlled invariant if the corresponding submodule $\Gamma^\infty(\mathscr{D})$ satisfies, $[f, \Gamma^\infty(\mathscr{D})] \subseteq \Gamma^\infty(\mathscr{D}) + \Gamma^\infty(\mathscr{G}^{(0)})$.

It is easy to verify that $\mathscr{D}$ satisfies conditions (i) and (ii). It does not, however, satisfy condition (iv). In particular, we will now show that $\mathrm{inv}(\mathscr{D} + \mathscr{G}^{(1)}) \not\subseteq \mathscr{D} + \mathscr{G}^{(1)}$, contradicting condition (iv). To do this, consider the vector fields $X \in \Gamma^\infty(\mathscr{D})$ and $Y \in \Gamma^\infty(\mathscr{G}^{(1)})$ given by

$$X = x^2 \frac{\partial}{\partial x^1} + x^2 \tan(x^4) \frac{\partial}{\partial x^2} - \frac{\partial}{\partial x^4}, \qquad Y = \sec^2(x^5) \frac{\partial}{\partial x^4}.$$

The Lie bracket of these two vector fields is

$$[X, Y] = x^2 \sec^2(x^4) \sec^2(x^5) \frac{\partial}{\partial x^2}.$$

It is easy to see that if $x^2 \sec^2(x^4) \sec^2(x^5) \neq 0$, then $[X, Y] \notin \mathscr{D} + \mathscr{G}^{(1)}$. The expression $\sec^2(x^4) \sec^2(x^5)$ is necessarily non-zero and x^2 is non-zero on *most* points of N. It is, in particular, non-zero at the point x_0. For the points where it is zero, the distribution fails to be regular. This, of course, does not mean that the problem is not solvable. It just means we have to be more clever. One solution is,

$$\mathscr{D}_x := \mathrm{span}_{\mathbb{R}} \left\{ x^2 \left. \frac{\partial}{\partial x^1} \right|_x - x^1 \left. \frac{\partial}{\partial x^2} \right|_x - \left. \frac{\partial}{\partial x^4} \right|_x, \left. \frac{\partial}{\partial x^3} \right|_x \right\},$$

which does, in fact, satisfy all the conditions of Theorem 1.3.3 near x_0. ◀

As Example 1.5 illustrates, it is not always clear how to choose the distribution $\mathscr{D}$ which satisfies the conditions of Theorem 1.3.3 even when we know that such a distribution exists. This thesis proposes a solution to this problem by considering the dual to those conditions of Theorem 1.3.2 and, in the spirit of [16, 18], leverage them to present an algorithm that constructs the required feedback transformation.

The proposed algorithm possesses a number of features of note. First, it produces its own certificate: the transverse output for which input-output feedback linearization can be performed to locally feedback linearize the dynamics transverse to N. This is in-line with known algorithms for state-space, exact feedback linearization. Secondly, the algorithm provides a geometric intuition for the algebraic adaptation process performed on the derived flag (Section 2.2). We show that the adaptation process amounts to producing a sequence of descending zero dynamics manifolds converging upon the desired zero dynamics manifold N. This not only gives a geometric perspective to our algorithm, but also to the GS algorithm [16] and "Blended Algorithm" [35] since, when $\mathsf{N} = \{x_0\}$ is a one-point set, the TFL problem coincides with the state-space, exact feedback linearization problem [40, Corollary 3.3]. Unlike previously published algorithms, the proposed algorithm begins with a desired zero dynamics chosen *a priori* and finds a virtual output whose zero dynamics manifold coincides with the desired zero dynamics manifold. This is something that *cannot* be done with the GS and Blended algorithms. This is also not possible with the algorithm for partial feedback linearization presented in [34] and used in [33] as the resulting zero dynamics manifold is not fixed before-hand; the resulting zero dynamics is not guaranteed to equal N. This thesis directly addresses this with the proposed algorithm that, by leveraging the conditions under which a desired zero dynamics manifold may be transverse feedback linearized, produces the required output for the desired zero dynamics manifold.

1.5 Organization and Notation

The thesis is organized in the following way. First, a review of the two dominant perspectives on feedback linearization theory is given in Chapter 2: the relative degree view championed by Isidori, and the equivalence of one-forms view championed by Gardner. In that chapter, the differential algebraic objects used throughout the thesis are formally introduced, and we develop some interesting results that connect the ideas of Gardner with those of Isidori. The reader may wish to consult Appendix A to briefly review the differential geometry used throughout the thesis before entering Chapter 2.

After reviewing feedback linearization theory, Chapter 3 proposes the dual conditions for transverse feedback linearization and shows, under additional regularity assumptions, the equivalence between them and the conditions of Theorem 1.3.2 for *single-input* systems. The proof method employed explicitly demonstrates how computations in the space of one-forms relate to computations on the vector fields in the original conditions for transverse feedback linearization.

Chapter 4 is the climax of this thesis, and proves that the dual conditions for transverse feedback linearization — proposed in the preceding chapter — apply just as well to multi-input systems. The proof method used explicitly executes the proposed algorithm, thereby establishing its correctness. Finally, in Chapter 5 we discuss the open problems and interesting questions that arose because of this thesis.

Notation

The set of natural numbers is denoted by $\mathbb{N}$ and the set of real numbers by $\mathbb{R}$. If A is a finite set, then $\mathrm{card}(A)$ denotes its cardinality. If M is a smooth (C^∞) differentiable manifold of dimension m, then we denote by $C^\infty(\mathsf{M})$ the ring of smooth real-valued functions on M. If $p \in \mathsf{M}$, then $\mathsf{T}_p\mathsf{M}$ denotes the tangent space at the point p and its dual, the cotangent space, is denoted $\mathsf{T}^*_p\mathsf{M}$. The tangent and cotangent bundles of M are written TM and $\mathsf{T}^*\mathsf{M}$ respectively. If $(\mathsf{U}; x^1, \ldots, x^m)$ is a chart of M, then for each $p \in \mathsf{U}$ the basis of vectors for $\mathsf{T}_p\mathsf{M}$ induced by the chart is denoted by $\partial/\partial x^1|_p, \ldots, \partial/\partial x^m|_p$. The vector fields $\partial/\partial x^1, \ldots, \partial/\partial x^m$ form a local frame for TM. The unique dual basis for $\mathsf{T}^*_p\mathsf{M}$ induced by the chart is denoted $\mathrm{d}x^1_p, \ldots, \mathrm{d}x^m_p$. If $H : \mathsf{M} \to \mathsf{N}$ is a smooth map between manifolds, then the pushforward at p is $\mathbf{D}H|_p : \mathsf{T}_p\mathsf{M} \to \mathsf{T}_{H(p)}\mathsf{N}$ and the pullback at p is the dual map $\mathbf{D}H|^*_p : \mathsf{T}^*_{H(p)}\mathsf{N} \to \mathsf{T}^*_p\mathsf{M}$.

The set of smooth sections of TM is denoted by $\Gamma^\infty(\mathsf{TM})$, whose elements define vector fields on M, and the set of smooth sections of $\mathsf{T}^*\mathsf{M}$ is denoted by $\Gamma^\infty(\mathsf{T}^*\mathsf{M})$, whose elements define covector fields (smooth one-forms) on M. The set $\Gamma^\infty(\mathsf{TM})$ (resp. $\Gamma^\infty(\mathsf{T}^*\mathsf{M})$) is a real vector space, but can also be endowed with the structure of a module over the ring $C^\infty(\mathsf{M})$. Let $k \in \mathbb{N} \cup \{0\}$. Define the set of smooth k-forms $\Gamma^\infty(\Lambda^k\mathsf{T}^*\mathsf{M})$ as sections of the bundle of k-forms and the space of all smooth forms as $\Gamma^\infty(\Lambda\mathsf{T}^*\mathsf{M}) := \bigoplus_{k=0}^m \Gamma^\infty(\Lambda^k\mathsf{T}^*\mathsf{M})$. Let $X \in \Gamma^\infty(\mathsf{TM})$ and $\omega \in \Gamma^\infty(\Lambda^k\mathsf{T}^*\mathsf{M})$. The notation $X_p \in \mathsf{T}_p\mathsf{M}$ ($\omega_p \in \Lambda^k\mathsf{T}^*_p\mathsf{M}$) denotes the vector X (k-form ω) in the tangent space (forms over the cotangent space) at $p \in \mathsf{M}$.

If $\mathscr{G} \subseteq \mathsf{TM}$ is a distribution, then its restriction to $p \in \mathsf{M}$ is denoted $\mathscr{G}_p \subseteq \mathsf{T}_p\mathsf{M}$. The set of covectors that annihilate vectors in $\mathscr{G}_p$ is $\mathrm{ann}(\mathscr{G}_p) \subseteq \mathsf{T}^*_p\mathsf{M}$. Define $\mathrm{ann}(\mathscr{G}) := \sqcup_{p\in\mathsf{M}} \mathrm{ann}(\mathscr{G}_p) \subseteq \mathsf{T}^*\mathsf{M}$. If $\mathscr{I} \subseteq \mathsf{T}^*\mathsf{M}$ is a codistribution, then its restriction to p is denoted $\mathscr{I}_p \subseteq \mathsf{T}^*_p\mathsf{M}$.

When a distribution $\mathscr{D} \subseteq \mathsf{TM}$ is smooth and regular, it is a subbundle of TM and, as such, can be associated to a $C^\infty(\mathsf{M})$-submodule of vector fields $\mathcal{D} := \Gamma^\infty(\mathscr{D}) \subseteq \Gamma^\infty(\mathsf{TM})$ comprising of smooth sections $X \in \Gamma^\infty(\mathsf{TM})$ that satisfy $X_p \in \mathscr{D}_p$ for all $p \in \mathsf{M}$.

If $\mathcal{D} \subseteq \Gamma^\infty(\mathsf{TM})$ is a $C^\infty(\mathsf{M})$-submodule and $p \in \mathsf{M}$, then to this object we associate a distribution defined pointwise by,

$$\mathscr{D}_p := \{X_p \colon X \in \mathcal{D}\} \subseteq \mathsf{T}_p\mathsf{M}.$$

We say a smooth k-form ω annihilates $\mathcal{D} \subseteq \Gamma^\infty(\mathsf{TM})$ if it evaluates to zero when all of its arguments are vector fields from $\mathcal{D}$. The set of smooth forms that annihilate vector fields in $\mathcal{D}$ is $\operatorname{ann}(\mathcal{D}) \subseteq \Gamma^\infty(\Lambda\mathsf{T}^*\mathsf{M})$. If $\mathcal{I} \subseteq \Gamma^\infty(\Lambda\mathsf{T}^*\mathsf{M})$ and $p \in \mathsf{M}$, then to this object we associate a codistribution defined pointwise by,

$$\mathscr{I}_p := \{\omega_p \colon \omega \in \mathcal{I} \cap \Gamma^\infty(\mathsf{T}^*\mathsf{M})\} \subseteq \mathsf{T}^*_p\mathsf{M}.$$

Given two smooth vector fields $X, Y \in \Gamma^\infty(\mathsf{TM})$, their Lie bracket is $[X, Y] \in \Gamma^\infty(\mathsf{TM})$. If $\mathcal{D}, \mathcal{G} \subseteq \Gamma^\infty(\mathsf{TM})$ are submodules, then

$$[\mathcal{D}, \mathcal{G}] := \{[X, Y] \colon X \in \mathcal{D}, Y \in \mathcal{G}\} \subseteq \Gamma^\infty(\mathsf{TM}).$$

Similarly, if $\mathscr{D}, \mathscr{G} \subseteq \mathsf{TM}$ are distributions, then we can define their Lie bracket pointwise,

$$[\mathscr{D}, \mathscr{G}]_p := \{\, [X, Y]\big|_p \colon X \in \mathcal{D}, Y \in \mathcal{G}\} \subseteq \mathsf{T}_p\mathsf{M}.$$

If $\mathscr{D} \subseteq \mathsf{TM}$ is a distribution, then the involutive closure is $\operatorname{inv}(\mathscr{D})$. As a matter of convenience, repeated Lie brackets are compressed using the following notation. Let X, $Y \in \Gamma^\infty(\mathsf{TM})$ and define $\operatorname{ad}_X^0 Y := Y$ and $\operatorname{ad}_X^1 Y := [X, Y]$. Recursively define $\operatorname{ad}_X^k Y := [X, \operatorname{ad}_X^{k-1} Y]$ for all $k > 1$.

Let $\ell \geq 0$. If $\omega \in \Gamma^\infty(\Lambda^k\mathsf{T}^*\mathsf{M})$ and $\beta \in \Gamma^\infty(\Lambda^\ell\mathsf{T}^*\mathsf{M})$ denote their wedge product by $\omega \wedge \beta \in \Gamma^\infty(\Lambda^{k+\ell}\mathsf{T}^*\mathsf{M})$. The wedge product distributes over the addition of smooth forms, amd endows the space of smooth forms with a graded algebra structure over the ring of smooth functions. If $\omega^1, \ldots, \omega^\ell \in \Gamma^\infty(\Lambda^1\mathsf{T}^*\mathsf{M})$ are smooth one-forms, then $\langle\omega^1, \ldots, \omega^\ell\rangle \subseteq \Gamma^\infty(\Lambda\mathsf{T}^*\mathsf{M})$ denotes the ideal generated by $\omega^1, \ldots, \omega^\ell$ over the aforementioned graded algebra. The exterior derivative of $\omega \in \Gamma^\infty(\Lambda^k \mathsf{T}^*\mathsf{M})$ is $\mathrm{d}\omega \in \Gamma^\infty(\Lambda^{k+1}\mathsf{T}^*\mathsf{M})$. If $\mathcal{I} \subseteq \Gamma^\infty(\Lambda\mathsf{T}^*\mathsf{M})$ is an ideal, then the largest ideal contained in $\mathcal{I}$ that is closed under the exterior derivative is denoted $\mathcal{I}^{(\infty)}$; the ideal $\mathcal{I}^{(\infty)}$ is otherwise known as the differential closure of $\mathcal{I}$. The Lie derivative of ω along a vector field $X \in \Gamma^\infty(\mathsf{TM})$ is the smooth k-form $\mathcal{L}_X\omega \in \Gamma^\infty(\Lambda^k\mathsf{T}^*\mathsf{M})$. Repeated Lie derivatives of order $j > 1$ are defined recursively by $\mathcal{L}_X^j\omega \coloneqq \mathcal{L}_X^{j-1}(\mathcal{L}_X\omega)$.

If $H : \mathsf{M} \to \mathbb{R}^\ell$ is smooth, then it can be written component-wise as $H = (H^1, \ldots, H^\ell)$ where $H^1, \ldots, H^\ell : \mathsf{M} \to \mathbb{R}$ are smooth. The function $H^i : \mathsf{M} \to \mathbb{R}$ may be seen as a smooth zero-form on M and, as such, has an exterior derivative $\mathrm{d}H^i \in \Gamma^\infty(\mathsf{T}^*\mathsf{M})$ and a Lie derivative, along the vector field X, $\mathcal{L}_X H^i : \mathsf{M} \to \mathbb{R}$.

Chapter 2

Feedback Linearization Theory

This chapter provides a consolidated review of the dual theory of feedback linearization which is difficult to find in a single source in the literature. First, in Section 2.1 we review the standard results for state-space, exact feedback linearization, a special case of transverse feedback linearization. We then visit the object known as the derived flag in Section 2.2 and prove a number of useful connections between the flag and the notion of relative degree. These results play a crucial role in the primary proof of this thesis, but, more importantly, provide intuition behind the derived flag construction. Finally, we provide a brief review of the modern results of feedback linearization theory.

2.1 Exact Feedback Linearization

Nonlinear control systems of the form (1.4) are more difficult to design controllers for when compared to controllable, linear control systems,

$$\dot{\xi}(t) = A\,\xi(t) + \sum_{j=1}^{m} b_j\, v^j(t). \tag{2.1}$$

A natural problem to pose is whether, or not, it is possible to transform (1.4) into (2.1) with an appropriate change of variables in an open set containing some objective $x_0 \in \mathbb{R}^n$. This would greatly simplify the control design problem, as one could run through the change of variables and then use linear control design procedures to stabilize the system (2.1). Formally, the problem of finding such a change of variables is the state-space, exact feedback linearization problem.

Problem 2.1.1 (State-Space, Exact Feedback Linearization)**.** *Suppose the set* $\{x_0\} \subseteq \mathbb{R}^n$ *for system* (1.4) *is rendered controlled-invariant by input* $u_0 \in \mathbb{R}^m$*. Find an open set* V *containing* x_0 *and a (local) feedback transformation* (Φ, α, β) *defined on* V *so that*

(1) $\Phi(x_0) = 0$,

(2) $u_0 = \alpha(x_0)$, *and*

(3) letting $\xi = \Phi(x)$ *and* $u = \alpha(x) + \beta(x)\, v$, *solutions to the nonlinear control system* (1.4) *on* V *correspond to solutions of the linear control system* (2.1) *where* $(A, \left[\begin{smallmatrix} b_1 & \cdots & b_m \end{smallmatrix}\right])$ *is in Brunovský normal form.*

In particular, solutions to (1.4) correspond to solutions of a linear control system consisting of m decoupled linear control systems (A_i, b_i') given by,

$$A_i = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 & \cdots & 0 \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & 1 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{\kappa_i \times \kappa_i}, \quad b_i' = \begin{pmatrix} 0 \\ \\ \vdots \\ \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^{\kappa_i \times 1}, \qquad 1 \leq i \leq m, \tag{2.2}$$

where $\kappa_1, \ldots, \kappa_m > 0$ are positive integers satisfying $\sum_{i=1}^m \kappa_i = n$. That is, solutions to (1.4) are rendered equivalent to solutions of a linear control system with m integration chains of lengths $\kappa_1, \ldots, \kappa_m$. This problem was depicted visually in Figure 1.1. The conditions upon which the problem is solvable is written using the language of differential geometry. Specific distributions — an assignment of subspaces to points on a manifold — are constructed through algebraic operations — repeated Lie brackets — on the vector fields f and g. Conditions on these distributions determine the solvability of the feedback linearization problem.

We now review these constructions to formally recall the solution to the feedback linearization problem. Once again, recall the classic control distributions, found as $\mathscr{M}^l$ in [34], given by,

$$\mathscr{G}_x^{(k)} \coloneqq \operatorname{span}_{\mathbb{R}}\{\operatorname{ad}_f^j g_i|_x : 1 \leq i \leq m, 0 \leq j \leq k\}, \quad k \geq 0, x \in \mathbb{R}^n. \tag{2.3}$$

as well as the distributions, which can be seen as a variant of $\mathscr{G}^j$ in [34],

$$\begin{aligned} &\mathscr{S}^{(0)} \coloneqq \mathscr{G}^{(0)}, \\ &\mathscr{S}^{(k)} \coloneqq \mathscr{S}^{(k-1)} + [\mathscr{S}^{(k-1)}, \mathscr{S}^{(k-1)}] + \mathscr{G}^{(k)}, \quad k \geq 1. \end{aligned} \tag{2.4}$$

We have already briefly discussed the control distribution $\mathscr{G}^{(k)}$. The distributions $\mathscr{S}^{(k)}$ were defined in [12] for the single-input case. Although $\mathscr{G}^{(k)}$ and $\mathscr{S}^{(k)}$ may not be equal, their involutive closures are.

Lemma 2.1.1. *For all* $k \geq 0$, $\operatorname{inv}(\mathscr{G}^{(k)}) = \operatorname{inv}(\mathscr{S}^{(k)})$.

Proof. By definition (2.4), we have that $\mathscr{G}^{(k)} \subseteq \mathscr{S}^{(k)}$. It follows directly by the definition of the involutive closure that $\operatorname{inv}(\mathscr{G}^{(k)}) \subseteq \operatorname{inv}(\mathscr{S}^{(k)})$. As a result, it suffices to show the reverse inclusion to show equality.

By definition (2.4), $\mathscr{G}^{(0)} = \mathscr{S}^{(0)}$ so it follows that $\operatorname{inv}(\mathscr{G}^{(0)}) = \operatorname{inv}(\mathscr{S}^{(0)})$. Suppose, by way of induction, that for some $k \geq 0$,

$$\operatorname{inv}(\mathscr{G}^{(k)}) = \operatorname{inv}(\mathscr{S}^{(k)}).$$

Fix $p \in \mathbb{R}^n$ and pick $X_p \in \mathscr{S}_p^{(k+1)}$. By definition (2.4), there exist vectors $A_p \in \mathscr{S}_p^{(k)}$, $B_p \in [\mathscr{S}^k, \mathscr{S}^k]_p$, $C_p \in \mathscr{G}_p^{(k+1)}$ that satisfy

$$X_p = A_p + B_p + C_p.$$

By the definition of the involutive closure, $C_p \in \mathscr{G}_p^{(k+1)} \subseteq \operatorname{inv}(\mathscr{G}^{(k+1)})_p$. By the definition of the involutive closure and the inductive hypothesis,

$$A_p \in \mathscr{S}_p^{(k)} \subseteq \operatorname{inv}(\mathscr{S}^{(k)})_p = \operatorname{inv}(\mathscr{G}^{(k)})_p.$$

Since $\mathrm{inv}(\mathscr{G}^{(k)}) \subseteq \mathrm{inv}(\mathscr{G}^{(k+1)})$, conclude that $A_p \in \mathrm{inv}(\mathscr{G}^{(k+1)})_p$. All that remains is showing that $B_p \in \mathrm{inv}(\mathscr{G}^{(k+1)})_p$. However we know that $[\mathscr{S}^{(k)}, \mathscr{S}^{(k)}] \subseteq \mathrm{inv}(\mathscr{S}^{(k)})$. Conclude that, by the inductive hypothesis,

$$B_p \in [\mathscr{S}^{(k)}, \mathscr{S}^{(k)}]_p \subseteq \mathrm{inv}(\mathscr{S}^{(k)})_p = \mathrm{inv}(\mathscr{G}^{(k)})_p.$$

Again, since $\mathrm{inv}(\mathscr{G}^{(k)}) \subseteq \mathrm{inv}(\mathscr{G}^{(k+1)})$, $B_p \in \mathrm{inv}(\mathscr{G}^{(k+1)})_p$. We therefore conclude that $X_p \in \mathrm{inv}(\mathscr{G}^{(k+1)})_p$. Since $p \in \mathbb{R}^n$ was arbitrary, this proves the result. □

The distribution $\mathscr{S}^{(k)}$ and its associated submodule $\mathcal{S}^{(k)} := \Gamma^{\infty}(\mathscr{S}^{(k)})$ will play a more prominent role when discussing transverse feedback linearization. For now, we focus our attention on the $\mathscr{G}^{(k)}$ distributions.

The following theorem provides the most commonly used conditions to test the solvability of Problem 2.1.1. These conditions are a restating of the conditions Hunt, Su and Meyer presented in [26]. If the conditions are satisfied, then the change of coordinates $\xi := \Phi(x)$ (consequently the feedback $u = \alpha(x) + \beta(x)\, v$) can be found by solving a system of partial differential equations.

Theorem 2.1.2 ([27, Theorem 5.2.3])**.** *The state-space, exact feedback linearization problem is solvable at x_0 if, and only if, there exists an open set V containing x_0 where,*

(1) $\dim(\mathscr{G}^{(n-1)}_{x_0}) = n$,

(2) for every $0 \le i \le n-2$, the distribution $\mathscr{G}^{(i)}$ is involutive, and

(3) for every $0 \le i \le n-1$ and for all $x \in \mathsf{V}$, $\dim(\mathscr{G}^{(i)}_{x_0}) = \dim(\mathscr{G}^{(i)}_{x})$.

The first condition is known as the controllability condition as it is necessary for all states to be reached after n integrations of the control action. The second condition is known as the involutivity condition, and is a technical requirement for sufficiency. The last condition is another technical requirement and is known as the constant dimension condition. Together, these are the necessary and sufficient conditions to solve Problem 2.1.1.

The control distributions $\mathscr{G}^{(k)}$ are not just used to test if Problem 2.1.1 is solvable, but, as Isidori indicates on [27, pg. 233], can be used to construct the required change of coordinates Φ. The procedure is described explicitly in the proof and, in part, inspires the procedure developed in this thesis. Because of its significance, we now build tools required to discuss the algorithm. First, we state a definition that plays a key role in both feedback linearization and in the thesis broadly.

Definition 2.1.3 (Relative Degree) . Let $h \in C^\infty(\mathbb{R}^n)$ be a smooth function. The system (1.4) with output h is said to have *relative degree* $\kappa \in \mathbb{N}$ at x_0 if there exists an open set V of x_0 so that, for all $x \in \mathsf{V}$,

$$\begin{pmatrix} \mathcal{L}_{g_1}\mathcal{L}_f^i h(x) & \cdots & \mathcal{L}_{g_m}\mathcal{L}_f^i h(x) \end{pmatrix} = 0, \qquad 0 \leq i \leq \kappa - 2,$$

and at x_0,

$$\operatorname{rank} \begin{pmatrix} \mathcal{L}_{g_1}\mathcal{L}_f^{\kappa-1} h(x_0) & \cdots & \mathcal{L}_{g_m}\mathcal{L}_f^{\kappa-1} h(x_0) \end{pmatrix} = 1.$$

Relative degree plays an important role in the feedback linearization problem as it indicates how many derivatives with respect to time must be applied to an output before the input appears and, importantly, that the input appears in a manner that permits cancellation of nonlinearities (the coefficient multiplying the control is non-zero). The "top" of the ith integration chain in (2.2) has relative degree κ_i at $\xi(x_0)$.

We can say more about the family of outputs at the top of the integration chains. We expect that not only the input appear at the end of each chain but that the inputs that drive each chain are independent of one another. This motivates the definition of vector relative degree.

Definition 2.1.4 (Vector Relative Degree). Let $h^1, \ldots, h^\ell \in C^\infty(\mathbb{R}^n)$. The system (1.4) with output $(h^1, \ldots, h^\ell)$ is said to have a *vector relative degree* of $(\kappa_1, \ldots, \kappa_\ell) \in \mathbb{N}^m$ if,

$$\begin{pmatrix} \mathcal{L}_{g_1}\mathcal{L}_f^i h^j(x) & \cdots & \mathcal{L}_{g_m}\mathcal{L}_f^i h^j(x) \end{pmatrix} = 0, \qquad 1 \leq j \leq \ell, 0 \leq i \leq \kappa_j - 2,$$

and at x_0,

$$\operatorname{rank} \begin{pmatrix} \mathcal{L}_{g_1}\mathcal{L}_f^{\kappa_1-1}h^1(x_0) & \cdots & \mathcal{L}_{g_m}\mathcal{L}_f^{\kappa_1-1}h^1(x_0) \\ \vdots & & \vdots \\ \mathcal{L}_{g_1}\mathcal{L}_f^{\kappa_\ell-1}h^\ell(x_0) & \cdots & \mathcal{L}_{g_m}\mathcal{L}_f^{\kappa_\ell-1}h^\ell(x_0) \end{pmatrix} = \ell.$$

It is said to have *uniform* vector relative degree if $\kappa_1 = \cdots = \kappa_\ell$.

The matrix

$$\begin{pmatrix} \mathcal{L}_{g_1}\mathcal{L}_f^{\kappa_1-1}h^1(x_0) & \cdots & \mathcal{L}_{g_m}\mathcal{L}_f^{\kappa_1-1}h^1(x_0) \\ \vdots & & \vdots \\ \mathcal{L}_{g_1}\mathcal{L}_f^{\kappa_\ell-1}h^\ell(x_0) & \cdots & \mathcal{L}_{g_m}\mathcal{L}_f^{\kappa_\ell-1}h^\ell(x_0) \end{pmatrix}$$

is known as the decoupling matrix, and appears multiplying the input after an appropriate number of Lie derivatives. The rank condition tells us this matrix has a right inverse.

Isidori points out that solving Problem 2.1.1 is equivalent to finding a family of outputs for (1.4) which would yield a specific vector relative degree. This is asserted in [27, Lemma 5.2.1] which we repeat here for convenience.

Lemma 2.1.5 ([27, Lemma 5.2.1]). *The state-space, exact feedback linearization problem is solvable if, and only if, there exists a family of smooth functions $h^1, \ldots, h^m \in C^\infty(\mathbb{R}^n)$ so that the system (1.4) with output $(h^1, \ldots, h^m)$ has vector relative degree $(\kappa_1, \ldots, \kappa_m)$ at x_0 where $\sum_{i=1}^m \kappa_i = n$.*

The outputs h^i of Lemma 2.1.5 correspond to the top of the integration chains in the linear subsystems (2.2). Constructing the state-transformation, as a result, amounts to differentiating the outputs h^i along the dynamics of (1.4) until the controls appear in the form $-\beta(x)^{-1}\alpha(x) + \beta(x)^{-1}u$; this then defines the feedback transformation.

Isidori's presentation of feedback linearization as a question of finding outputs that yield a vector relative degree hints at a procedure to compute the feedback transformation: one must find the family of m outputs h^i that yield a vector relative degree. At first glance, the connection between vector relative degree and the conditions of Theorem 2.1.2 are not apparent. However, as a consequence of [27, Theorem 4.1.2], the definition of vector relative degree is equivalent to asking that, for all $x \in \mathsf{V}$,

$$\begin{pmatrix} \mathcal{L}_{\mathrm{ad}_f^i g_1} h^j(x) & \cdots & \mathcal{L}_{\mathrm{ad}_f^i g_m} h^j(x) \end{pmatrix} = 0, \qquad 1 \le j \le \ell, 0 \le i \le \kappa_j - 2,$$

and at x_0,

$$\operatorname{rank} \begin{pmatrix} \mathcal{L}_{\mathrm{ad}_f^{\kappa_1 - 1} g_1} h^1(x_0) & \cdots & \mathcal{L}_{\mathrm{ad}_f^{\kappa_1 - 1} g_m} h^1(x_0) \\ \vdots & & \vdots \\ \mathcal{L}_{\mathrm{ad}_f^{\kappa_\ell - 1} g_1} h^\ell(x_0) & \cdots & \mathcal{L}_{\mathrm{ad}_f^{\kappa_\ell - 1} g_m} h^\ell(x_0) \end{pmatrix} = \ell.$$

Observe that the vector fields which we differentiate the h^i along are precisely the vector fields that span, pointwise, the distributions $\mathscr{G}^{(k)}$. This suggests an equivalent characterization for vector relative degree in terms of the differentials of h^i.

Definition 2.1.6 (Vector Relative Degree). Let $h^1, \ldots, h^\ell \in C^\infty(\mathbb{R}^n)$. The system (1.4) with output $(h^1, \ldots, h^\ell)$ is said to have a vector relative degree of $(\kappa_1, \ldots, \kappa_\ell) \in \mathbb{N}^\ell$ at x_0 if, there exists an open set V containing x_0 so that, for all $x \in \mathsf{V}$,

$$\mathrm{d}h^j_x \in \operatorname{ann}(\mathscr{G}^{(i)}_x), \qquad 1 \le j \le \ell, 0 \le i \le \kappa_j - 2,$$

and at x_0,

$$\operatorname{rank} \begin{pmatrix} \mathrm{d}h^1_{x_0}(\mathrm{ad}_f^{\kappa_1 - 1} g_1) & \cdots & \mathrm{d}h^1_{x_0}(\mathrm{ad}_f^{\kappa_1 - 1} g_m) \\ \vdots & & \vdots \\ \mathrm{d}h^m_{x_0}(\mathrm{ad}_f^{\kappa_m - 1} g_1) & \cdots & \mathrm{d}h^\ell_{x_0}(\mathrm{ad}_f^{\kappa_m - 1} g_m) \end{pmatrix} = \ell.$$

Isidori's algorithm for state-space, exact feedback linearization can now be described. Because of Theorem 2.1.2 (1) and the definition of $\mathscr{G}^{(0)}$, there exists a *largest* index $0 \le \kappa_1 \le n$ so that $\dim(\mathscr{G}^{(\kappa_1 - 1)}_{x_0}) < n$. By Theorem 2.1.2 (2), the distribution $\mathscr{G}^{(\kappa_1 - 1)}$ is involutive. By construction, it is finitely generated and, by Theorem 2.1.2 (3), it is constant dimensional pointwise. By Frobenius's Theorem, we can pick the maximal set of *exact* differentials $\mathrm{d}h^1$, $\ldots$, $\mathrm{d}h^\ell$, that annihilate $\mathscr{G}^{(\kappa_1-1)}$ and are pointwise linearly independent.

The integral of these exact differentials correspond to some of the family of output functions that will yield the required vector relative degree. In particular, these are the components of the final output which yield vector relative degree κ_1. The remaining outputs must be computed by walking backwards through the distributions $\mathscr{G}^{(k)}$ and iteratively applying Frobenius's Theorem. Importantly, the known outputs and their Lie derivatives (up to a certain order) will annihilate all of $\mathscr{G}^{(k)}$ for all $k \leq \kappa_1 - 1$. As a result, one must *carefully* find exact differentials that annihilate the submodules but also are linearly independent of the known $\mathrm{d}h^i$ and their Lie derivatives. If they fail to be linearly independent, they will not yield the required vector relative degree. This more, or less, amounts to solving m-systems of partial differential equations. Worse still, at each iteration the size of this system increases and isn't related to the number of outputs that one actually must search for.

The difficulty solving this series of systems of PDEs is clear, and within a few years attempts were made to alleviate the difficulty. The first attempt was made in [15] by Gardner and Shadwick in the case where the controllability indices of the target linear system (A, B) are distinct. They completed their work in [16] for the general case. The matter of how difficult it can be to "decouple" the new outputs from known outputs and how to reduce the PDE size was resolved by this work. The inspiration taken by the authors was to recognize that the decoupling problem in the dual to the space of vector fields — smooth one-forms — could be reduced to an algebraic problem.

It is worth pointing out that working in dual presents another aesthetic advantage. Consider the following question:

> If all we care about is the outputs that yield the right vector relative degree, could we not work instead with the objects that contain the possible candidate outputs instead of the vector fields they must annihilate?

We address this very question in the next section and we will see how the works [16, 48] connect with finding outputs yielding vector relative degree.

2.2 The Derived Flag

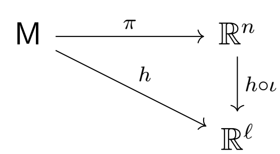
The differentials of outputs h^i that with system (1.4) yield a relative degree are the subject of this section. We will define an object, known as the derived flag, that captures the notion of "eliminating" outputs whose derivative along the solutions of (1.4) contains (1) an input or (2) an already eliminated state affected by an input at a lower order of integration. As a result, this flag aids in the search of functions that yield a desired relative degree.

To do this, the current setting — viewing time t and controls u^i as special relative to the states x^i — is not appropriate even though this is the standard perspective taken in the literature. Unusually, Hermann, Gardner and a few others took the perspective that time and control are independent variables, in their own right, and, instead of searching for the change of coordinates $\xi \coloneqq \Phi(x)$ alone, sought a diffeomorphism that preserved the special status that the time and control must have [15, 22]. This is the perspective we take for the remainder of this thesis as it proves useful.

Define the ambient manifold as the Cartesian product

$$\mathsf{M} \coloneqq \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^n,$$

of time ($\mathbb{R}$), control ($\mathbb{R}^m$) and states ($\mathbb{R}^n$). Let $\pi : \mathsf{M} \to \mathbb{R}^n$ be the projection map $\pi(t, u, x) = x$ and let $\iota : \mathbb{R}^n \to \mathsf{M}$ be the insertion map $\iota(x) = (0, 0, x)$. A smooth function $h : \mathsf{M} \to \mathbb{R}^\ell$ is said to be *a smooth function of the states* if the diagram



commutes. It is convenient to consider vector fields that are ι-related to the vector fields f and g_j in (1.4). Through an abuse of notation, define

$$f \coloneqq \sum_{i=1}^{n} (f^i \circ \pi) \frac{\partial}{\partial x^i}, \qquad g_j \coloneqq \sum_{i=1}^{n} (g_j^i \circ \pi) \frac{\partial}{\partial x^i}, \qquad 1 \leq j \leq m, \tag{2.5}$$

so that $f(\pi(p)) = \mathbf{D}\pi|_p\, f(p)$ and that f on $\mathbb{R}^n$ and f on M are ι-related vector fields. The fact that the vector fields $f \in \Gamma^\infty(\mathsf{T}\mathbb{R}^n)$ as defined in (1.4) and $f \in \Gamma^\infty(\mathsf{TM})$ as defined in (2.5) are related by the insertion ι and projection π justifies this abuse of notation. It will be clear from context which manifold f and g_j reside.

With these constructions, the control system (1.4) is differentially equivalent to the system of differential equations on M,

$$\begin{aligned} \dot{t} &= 1, \\ \dot{x} &= f(x) + \sum_{j=1}^{m} g_j(x)\, u^j. \end{aligned} \tag{2.6}$$

Furthermore, system (1.4) with output $h : \mathbb{R}^n \to \mathbb{R}^m$ yields a vector relative degree at $x_0 \in \mathbb{R}^n$ if, and only if, system (2.6) with output $h \circ \pi : \mathsf{M} \to \mathbb{R}^m$ yields a vector relative degree at $\iota(x_0) \in \mathsf{M}$. Not only this, but the distributions $\mathscr{G}^{(k)}$ and $\mathscr{S}^{(k)}$ defined on M directly correspond to the very same objects $\mathscr{G}^{(k)}$ and $\mathscr{S}^{(k)}$ defined on $\mathbb{R}^n$ in the natural way. That is, $\mathbf{D}\pi|_{\iota(x)} \mathscr{G}^{(k)}_{\iota(x)} = \mathscr{G}^{(k)}_x$.

We can view solutions to the control system (2.6) as integral submanifolds of a distribution. Define the smooth and regular distribution of control directions,

$$\mathscr{U}_p := \operatorname{span}_{\mathbb{R}} \left\{ \left.\frac{\partial}{\partial u^1}\right|_p, \ldots, \left.\frac{\partial}{\partial u^m}\right|_p \right\} \subseteq \mathsf{T}_p\mathsf{M}, \qquad p \in \mathsf{M}, \tag{2.7}$$

which is associated with the $C^\infty(\mathsf{M})$-submodule $\mathcal{U} := \Gamma^\infty(\mathscr{U})$. Additionally, define the smooth and regular distribution

$$\mathscr{D}^{(0)}_p := \operatorname{span}_{\mathbb{R}} \left\{ \left.\frac{\partial}{\partial t}\right|_p + f|_p + \sum_{j=1}^{m} g_j|_p\, u^j \right\} + \mathscr{U}_p, \tag{2.8}$$

which is associated to the $C^\infty(\mathsf{M})$-submodule $\mathcal{D}^{(0)} := \Gamma^\infty(\mathscr{D}^{(0)})$. Observe that the vector field

$$F := \frac{\partial}{\partial t} + f + \sum_{j=1}^{m} g_j u^j \in \mathcal{D}^{(0)} \tag{2.9}$$

is tangent to solutions of (2.6). When u is a sufficiently smooth signal, integral submanifolds of $\mathscr{D}^{(0)}$ determine solutions of (2.6) and, in turn, solutions of (1.4).

Alternatively, the control system (2.6) may be viewed as an exterior differential system on M in the following way. Define the smooth one-forms

$$\omega^i := \mathrm{d}x^i - \left(f^i(x) + \sum_{j=1}^{m} g^i_j(x)\, u^j \right) \mathrm{d}t \in \Gamma^\infty(\mathsf{T}^*\mathsf{M}), \qquad 1 \leq i \leq n. \tag{2.10}$$

The submanifolds of M on which the ideal[1]

$$\mathcal{I}^{(0)} := \langle \omega^1, \ldots, \omega^n \rangle \subseteq \Gamma^\infty(\Lambda\mathsf{T}^*\mathsf{M}), \tag{2.11}$$

vanishes correspond to solutions of the differential equation (2.6) where u is a sufficiently regular signal — corresponding therein to solutions of (1.4). The ideal $\mathcal{I}^{(0)}$ is simply, finitely, non-degenerately generated by construction since it is generated by a finite number of smooth one-forms ω^j that are pointwise linearly independent. It follows that the generators of $\mathcal{I}^{(0)}$ span a smooth and regular codistribution $\mathscr{I}^{(0)} \subseteq \mathsf{T}^*\mathsf{M}$. To this ideal $\mathcal{I}$, we associate the object of importance in this section: the derived flag.

Definition 2.2.1 (The Derived Flag). Let $\mathcal{I}^{(0)} \subseteq \Gamma^\infty(\Lambda\mathsf{T}^*\mathsf{M})$ be an ideal, and define the *derived ideals* by

$$\mathcal{I}^{(k+1)} := \{\omega \in \mathcal{I}^{(k)} \colon \mathrm{d}\omega \in \mathcal{I}^{(k)}\}, \qquad k \geq 0. \tag{2.12}$$

The *derived flag* of $\mathcal{I}^{(0)}$ is the sequence of derived ideals,

$$\{0\} \subseteq \cdots \subseteq \mathcal{I}^{(i+1)} \subseteq \mathcal{I}^{(i)} \subseteq \cdots \subseteq \mathcal{I}^{(1)} \subseteq \mathcal{I}^{(0)}. \tag{2.13}$$

The *length* of the derived flag is the smallest $N \in \mathbb{N}$ such that $\mathcal{I}^{(N)} = \mathcal{I}^{(N+1)}$.

Example 2.1 (Model of a Car). Recall the dynamics (1.5) from Example 1.1. The

[1]See Appendix A.6 for a discussion on exterior differential systems.

system ideal (2.11) for this nonlinear control system is,

$$\begin{aligned}
\mathcal{I}^{(0)} &:= \langle \omega^1, \omega^2, \omega^3, \omega^4, \omega^5, \omega^6 \rangle, \\
\omega^1 &:= \mathrm{d}x^1 - x^3 \cos(x^4)\, \mathrm{d}t, \\
\omega^2 &:= \mathrm{d}x^2 - x^3 \sin(x^4)\, \mathrm{d}t, \\
\omega^3 &:= \mathrm{d}x^3 - x^6\, \mathrm{d}t, \\
\omega^4 &:= \mathrm{d}x^4 - \frac{x^3}{\ell} \tan(x^5)\, \mathrm{d}t, \\
\omega^5 &:= \mathrm{d}x^5 - u^1\, \mathrm{d}t, \\
\omega^6 &:= \mathrm{d}x^6 - u^2\, \mathrm{d}t.
\end{aligned}$$

The derived flag of length 3 is

$$\begin{aligned}
\mathcal{I}^{(1)} &= \langle \mathrm{d}x^1 - x^3 \cos(x^4)\, \mathrm{d}t, \mathrm{d}x^2 - x^3 \sin(x^4)\, \mathrm{d}t, \mathrm{d}x^3 - x^6\, \mathrm{d}t, \mathrm{d}x^4 - \frac{x^3}{\ell} \tan(x^5)\, \mathrm{d}t \rangle, \\
\mathcal{I}^{(2)} &= \langle \mathrm{d}x^1 - x^3 \cos(x^4)\, \mathrm{d}t, \mathrm{d}x^2 - x^3 \sin(x^4)\, \mathrm{d}t \rangle, \\
\mathcal{I}^{(3)} &= \langle 0 \rangle.
\end{aligned}$$

◀

The differential ideal $\mathcal{I}^{(N)}$ is not, in general, the largest differential ideal contained within $\mathcal{I}^{(0)}$. For this reason, we define the largest differential ideal contained within an ideal.

Definition 2.2.2 (The Differential Closure). Let $\mathcal{I} \subseteq \Gamma^\infty(\Lambda \mathsf{T}^*\mathsf{M})$ be an ideal. The largest, differential ideal contained in $\mathcal{I}$ is denoted $\mathcal{I}^{(\infty)}$. The ideal $\mathcal{I}^{(\infty)}$ is said to be the *differential closure* of the ideal $\mathcal{I}$.

The existence of the differential closure is ensured by an argument leveraging Zorn's Lemma (see [10, Lemma 3.9.4]). With an abuse of notation, we denote the differential closure of the ideal $\mathcal{I}^{(0)}$ by $\mathcal{I}^{(\infty)}$. If all the ideals in the derived flag (2.13) are simply, finitely, non-degenerately generated — as was the case for $\mathcal{I}^{(0)}$ — then $\mathcal{I}^{(N)} = \mathcal{I}^{(\infty)}$. Consequently, we make the following convenient assumption about this flag.

Assumption 2.2.3. *The ideals $\mathcal{I}^{(k)}$ and the augmented ideals $\langle \mathcal{I}^{(k)}, \mathrm{d}t\rangle^{(\infty)}$ are locally, simply, finitely, non-degenerately generated, for all $k \geq 0$.*

Assumption 2.2.3 allows us to take the generators of $\mathcal{I}^{(k)}$ and use them as a basis for a smooth and regular codistribution $\mathscr{I}^{(k)} \subseteq \mathsf{T}^*\mathsf{M}$. The next few lemmas are general facts that are useful in the pursuit of finding outputs that yield a desired relative degree.

Definition 2.2.4 (Characteristic Vector Fields) . Let $\mathcal{I}$ be a simply, finitely generated ideal and let $\mathscr{I} \subseteq \mathsf{T}^*\mathsf{M}$ be the associated codistribution. A smooth vector field $Y \in \Gamma^\infty(\mathsf{TM})$ is said to be a *characteristic vector field (of $\mathcal{I}$)* if, for all $p \in \mathsf{M}$, $\mathscr{I}_p \subseteq \operatorname{ann}(\operatorname{span}_\mathbb{R}\{Y_p\}) \subseteq \mathsf{T}_p^*\mathsf{M}$.

Note that the vector field $F \in \mathcal{D}^{(0)}$ defined in (2.9) is a characteristic vector field of $\mathcal{I}^{(0)}$. If a smooth one-form ω annihilates all the characteristic vector fields of an ideal $\mathcal{I}$ that is simply, finitely, non-degenerately generated, then $\omega \in \mathcal{I}$. Using this, we show that the Lie derivative of a smooth one-form in the derived flag along specific vector fields is a smooth one-form that is also in the derived flag.

Lemma 2.2.5. *Let $\kappa \in \mathbb{N}$ and let $Y \in \Gamma^\infty(\mathsf{TM})$ be a characteristic vector field of the $\mathcal{I}^{(0)} \subseteq \Gamma^\infty(\Lambda\mathsf{T}^*\mathsf{M})$. If $\omega \in \mathcal{I}^{(\kappa)}$ is a smooth one-form, then $\mathcal{L}_Y\omega \in \mathcal{I}^{(\kappa-1)}$.*

Proof. By [31, Proposition 14.35],

$$\mathcal{L}_Y\omega(\cdot) = \mathrm{d}(\omega(Y))(\cdot) + \mathrm{d}\omega(Y, \cdot).$$

Since Y is a characteristic vector field of $\mathcal{I}^{(0)}$ and $\omega \in \mathcal{I}^{(\kappa)} \subseteq \mathcal{I}^{(0)}$, we have that $\omega(Y) = 0$. Therefore,

$$\mathcal{L}_Y\omega(\cdot) = \mathrm{d}\omega(Y, \cdot).$$

Since $\omega \in \mathcal{I}^{(\kappa)}$, by definition we have $\mathrm{d}\omega \in \mathcal{I}^{(\kappa-1)}$. Moreover, Y is a characteristic vector field of $\mathcal{I}^{(\kappa-1)} \subseteq \mathcal{I}^{(0)}$ as well. As a result, $\mathrm{d}\omega(Y, X) = 0$ for all characteristic vector fields X of $\mathcal{I}^{(\kappa-1)}$. Conclude then that $\mathcal{L}_Y\omega(\cdot) = \mathrm{d}\omega(Y, \cdot) \in \mathcal{I}^{(\kappa-1)}$. □

Lemma 2.2.5 makes a general statement about smooth one-forms in the derived flag (2.13). The next lemma connects the smooth one-forms in the augmented ideal $\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle$ with those in the derived flag (2.13).

Lemma 2.2.6. *Let $\kappa \geq 0$ and $F \in \mathcal{D}^{(0)}$ be the vector field defined in (2.9). If $\omega \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle$ is a smooth one-form, then $\omega - \omega(F)\,\mathrm{d}t \in \mathcal{I}^{(\kappa)}$.*

Proof. Let $\omega \in \langle \mathcal{I}^{(0)}, \mathrm{d}t\rangle$. Then there exists a $b \in C^\infty(\mathsf{M})$ so that,

$$\omega - b\,\mathrm{d}t \in \mathcal{I}^{(0)}.$$

Since $F \in \mathcal{D}^{(0)}$,

$$0 = \omega(F) - b\,\mathrm{d}t(F) = \omega(F) - b.$$

Therefore $b = \omega(F)$ and $\omega - b\,\mathrm{d}t = \omega - \omega(F)\,\mathrm{d}t \in \mathcal{I}^{(0)}$.

Fix $\kappa \geq 0$. Suppose, by way of induction, that if $\omega \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle$, then $\omega - \omega(F)\,\mathrm{d}t \in \mathcal{I}^{(\kappa)}$. Consider $\omega \in \langle \mathcal{I}^{(\kappa+1)}, \mathrm{d}t\rangle$. By definition, there exists a $b \in C^\infty(\mathsf{M})$ so that,

$$\omega - b\,\mathrm{d}t \in \mathcal{I}^{(\kappa+1)}.$$

By (2.12), $\mathcal{I}^{(\kappa+1)} \subseteq \mathcal{I}^{(\kappa)}$, and so

$$\omega - b\,\mathrm{d}t \in \mathcal{I}^{(\kappa)}.$$

It follows that $\omega \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle$, and, by the inductive hypothesis, we may write

$$\omega - \omega(F)\,\mathrm{d}t \in \mathcal{I}^{(\kappa)}.$$

Taking the difference we have that,

$$(\omega(F) - b)\,\mathrm{d}t \in \mathcal{I}^{(\kappa)}.$$

Since F is a characteristic vector field of $\mathcal{I}^{(0)}$ and $\mathcal{I}^{(\kappa)} \subseteq \mathcal{I}^{(0)}$,

$$0 = (\omega(F) - b)\,\mathrm{d}t(F) = \omega(F) - b.$$

Conclude that

$$\omega - \omega(F)\,\mathrm{d}t \in \mathcal{I}^{(\kappa+1)},$$

completing the proof. □

Using this lemma, we can now show that if the differential of a smooth function $h \in C^\infty(\mathsf{M})$ lives in a particular ideal of the derived flag (2.13), then a Lie derivative along the characteristic vector field F of $\mathcal{I}^{(\kappa)}$ (of a particular order) annihilates the distribution $\mathscr{G}^{(0)}$ defined in (2.3).

Lemma 2.2.7. *Let $\kappa \in \mathbb{N}$, let $h \in C^\infty(\mathsf{M})$ and let $F \in \mathcal{D}^{(0)}$ be the vector field defined in (2.9). If $\mathrm{d}h \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle$, then $\mathcal{L}_F^{\kappa-1}\mathrm{d}h_p \in \mathrm{ann}(\mathscr{G}_p^{(0)})$ for all $p \in \mathsf{M}$.*

Proof. Base Case ($\kappa = 1$): Let $\mathrm{d}h \in \langle \mathcal{I}^{(1)}, \mathrm{d}t \rangle$. Using Lemma 2.2.6, write

$$\mathrm{d}h - \mathrm{d}h(F)\,\mathrm{d}t =: \omega \in \mathcal{I}^{(1)}.$$

Fix $1 \leq i \leq m$. Evaluate this differential form with $g_i \in \Gamma^\infty(\mathsf{TM})$ to find,

$$\mathrm{d}h(g_i) - \mathrm{d}h(F)\,\mathrm{d}t(g_i) = \omega(g_i).$$

Since g_i is only a function of the states, it is time-independent, and so

$$\mathrm{d}h(g_i) = \omega(g_i).$$

Then, using the fact that $g_i = [F, \partial/\partial u^i]$,

$$\mathrm{d}h(g_i) = \omega\left(\left[F, \frac{\partial}{\partial u^i}\right]\right).$$

By Proposition A.5.10,

$$\mathrm{d}h(g_i) = F\left(\omega\left(\frac{\partial}{\partial u^i}\right)\right) - \frac{\partial}{\partial u^i}\Big(\omega(F)\Big) - \mathrm{d}\omega\left(F, \frac{\partial}{\partial u^i}\right).$$

Since $\omega \in \mathcal{I}^{(1)}$, we have that $\omega, \mathrm{d}\omega \in \mathcal{I}^{(0)}$. Both $\partial/\partial u^i \in \mathcal{D}^{(0)}$ and $F \in \mathcal{D}^{(0)}$ are characteristic vector fields of $\mathcal{I}^{(0)}$ and, as a result, it follows that,

$$\mathrm{d}h(g_i) = 0.$$

Since $1 \leq i \leq m$ was arbitrary, we have established that $\mathrm{d}h_p \in \mathrm{ann}(\mathscr{G}_p^{(0)})$ for all $p \in \mathsf{M}$.

Induction ($\kappa \geq 1$): Suppose, by way of induction, that if $\mathrm{d}h \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle$ then $\mathcal{L}_F^{\kappa-1}\mathrm{d}h_p \in \operatorname{ann}(\mathcal{G}_p^{(0)})$ for all $p \in \mathsf{M}$. Let $\mathrm{d}h \in \langle \mathcal{I}^{(\kappa+1)}, \mathrm{d}t \rangle$. Using Lemma 2.2.6, write

$$\mathrm{d}h - \mathcal{L}_F h \, \mathrm{d}t \in \mathcal{I}^{(\kappa+1)}.$$

By definition (2.12), the exterior derivative of $\mathrm{d}h - \mathcal{L}_F h \, \mathrm{d}t$ satisfies,

$$\mathrm{d}t \wedge \mathcal{L}_F \mathrm{d}h \in \mathcal{I}^{(\kappa)}.$$

Since F is a characteristic vector field of $\mathcal{I}^{(\kappa)}$, we can pass F as the first argument of this differential two-form and use Definition A.4.3 of the wedge product to find

$$\mathcal{L}_F \mathrm{d}h - (\mathcal{L}_F \mathrm{d}h(F)) \, \mathrm{d}t \in \mathcal{I}^{(\kappa)}.$$

It immediately follows that $\mathcal{L}_F \mathrm{d}h \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle$. By the inductive hypothesis,

$$\mathcal{L}_F^{\kappa-1} \mathcal{L}_F \mathrm{d}h_p \in \operatorname{ann}(\mathcal{G}_p^{(0)}),$$

for all $p \in \mathsf{M}$. After combining the Lie derivatives, find that $\mathcal{L}_F^{\kappa} \mathrm{d}h_p \in \operatorname{ann}(\mathcal{G}_p^{(0)})$ for all $p \in \mathsf{M}$ as required. □

The next result directly follows from Lemma 2.2.7.

Corollary 2.2.8. *Let $\kappa \in \mathbb{N}$, let $h \in C^\infty(\mathsf{M})$ and let $F \in \mathcal{D}^{(0)}$ be the vector field defined in (2.9). If $\mathrm{d}h \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle$, then*

$$\mathrm{d}h_p, \mathcal{L}_F \mathrm{d}h_p, \ldots, \mathcal{L}_F^{\kappa-1} \mathrm{d}h_p \in \operatorname{ann}(\mathcal{G}_p^{(0)}),$$

for all $p \in \mathsf{M}$.

Proof. Repeatedly apply Lemma 2.2.7 using the observation,

$$\mathrm{d}h \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle \subseteq \langle \mathcal{I}^{(\kappa-1)}, \mathrm{d}t \rangle \subseteq \cdots \subseteq \langle \mathcal{I}^{(1)}, \mathrm{d}t \rangle,$$

to arrive at the conclusion. □

If we now restrict our attention to smooth functions of the state, Corollary 2.2.8 reduces to a condition that looks *almost* like the requirements for a smooth function to yield a relative degree.

Lemma 2.2.9. *Let* $\kappa \in \mathbb{N}$, $h \in C^\infty(\mathsf{M})$ *be a smooth function of the states, and let* $f \in \Gamma^\infty(\mathsf{TM})$ *be the drift vector field of system* (2.6)*. If* $\mathrm{d}h \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle$*, then*

$$\mathrm{d}h_p, \mathcal{L}_f \mathrm{d}h_p, \ldots, \mathcal{L}_f^{\kappa-1} \mathrm{d}h_p \in \mathrm{ann}(\mathscr{G}_p^{(0)}),$$

for all $p \in \mathsf{M}$.

Proof. Let $F \in \mathcal{D}^{(0)}$ be the vector field defined in (2.9). We will show that $\mathcal{L}_F^i \mathrm{d}h = \mathcal{L}_f \mathrm{d}h$ for $1 \leq i \leq \kappa - 1$. Invoke Corollary 2.2.8 to find that

$$\mathrm{d}h_p, \mathcal{L}_F \mathrm{d}h_p, \ldots, \mathcal{L}_F^{\kappa-1} \mathrm{d}h_p \in \mathrm{ann}(\mathscr{G}_p^{(0)}), \tag{2.14}$$

for all $p \in \mathsf{M}$. Observe that, by definition (2.9) of F,

$$\mathcal{L}_F \mathrm{d}h = \mathcal{L}_{\partial/\partial t + f + \sum_{i=1}^m g_j u^j} \mathrm{d}h.$$

Since h is a smooth function of the states,

$$\mathcal{L}_F \mathrm{d}h = \mathcal{L}_f \mathrm{d}h + \sum_{i=1}^m u^j \mathcal{L}_{g_j} \mathrm{d}h.$$

Using (2.14) conclude $\mathcal{L}_F \mathrm{d}h = \mathcal{L}_f \mathrm{d}h$. Suppose, by way of induction, for some $1 \leq i < \kappa - 1$

$$\mathcal{L}_F^i \mathrm{d}h = \mathcal{L}_f^i \mathrm{d}h.$$

Consider $\mathcal{L}_F^{i+1} \mathrm{d}h$. Observe that

$$\mathcal{L}_F^{i+1} \mathrm{d}h = \mathcal{L}_F \mathcal{L}_f^i \mathrm{d}h.$$

Since f and h are smooth functions of the states, $\mathcal{L}_f^i h$ is also a smooth function of the states. As a result,

$$\mathcal{L}_F^{i+1} \mathrm{d}h = \mathcal{L}_f \mathcal{L}_f^i \mathrm{d}h + \sum_{i=1}^m u^j \mathcal{L}_{g_j} \mathcal{L}_f^i \mathrm{d}h.$$

By (2.14), the latter term vanishes resulting in,

$$\mathcal{L}_F^{i+1}\mathrm{d}h = \mathcal{L}_f^{i+1}\mathrm{d}h,$$

completing the proof. □

The connection between the derived flag and relative degree is made clearer by Lemma 2.2.9. It is a necessary condition that if a smooth function of the state h yields a relative degree of κ at p_0 that $\mathrm{d}h \in \langle \mathcal{I}^{(\kappa-1)}, \mathrm{d}t\rangle$ as its Lie derivatives along f up to order $\kappa - 2$ annihilate the control vector fields $g_1, \ldots, g_m$.

It is not clear, however, how one can determine whether a given exact one-form living in $\langle \mathcal{I}^{(\kappa-1)}, \mathrm{d}t\rangle$ is the differential of an output yielding relative degree κ. A fair conjecture is that the differential $\mathrm{d}h$ *not* live in the subsequent augmented ideal of the flag $\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle$. We dedicate the rest of this section to showing that this, in fact, is the right requirement.

Recall from Definition 2.2.2 that the ideal $\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle^{(\infty)}$ is the largest differential ideal contained in $\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle$. We now demonstrate the relationship between the differential ideal $\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle^{(\infty)}$, and the involutive closure of the control distribution $\mathrm{inv}(\mathscr{G}^{(\kappa-1)})$. We separate the claim into two parts that individually rely on different, seemingly unrelated assumptions. First we consider what happens under Assumption 2.2.3.

Lemma 2.2.10. *Let $\kappa \in \mathbb{N}$. If Assumption 2.2.3 holds, then*

$$\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle^{(\infty)} \subseteq \mathrm{ann}\left(\Gamma^\infty(\mathscr{U} \oplus \mathrm{inv}(\mathscr{G}^{(\kappa-1)}))\right).$$

Proof. Let $\kappa \in \mathbb{N}$ and consider the differential ideal, $\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle^{(\infty)}$. By Assumption 2.2.3, this differential ideal is simply, finitely generated. Fix $p \in \mathsf{M}$ and invoke Frobenius's Theorem to find, on some open set U containing p, exact differential forms $\mathrm{d}h^1, \ldots, \mathrm{d}h^\ell$ so that,

$$\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle^{(\infty)} = \langle \mathrm{d}h^1, \ldots, \mathrm{d}h^\ell, \mathrm{d}t\rangle,$$

where we've used the fact that $\mathrm{d}t$ is already known to be exact. Moreover, we can take the $h^1, \ldots, h^\ell$ to be smooth functions of the states. It is already clear that $\mathrm{d}t \in \mathrm{ann}(\Gamma^\infty(\mathscr{U} \oplus \mathrm{inv}(\mathscr{G}^{(\kappa-1)})))$ so we turn to verifying the $\mathrm{d}h^i$ annihilate vector fields of $\mathrm{ann}(\Gamma^\infty(\mathscr{U} \oplus \mathrm{inv}(\mathscr{G}^{(\kappa-1)})))$.

Fix $1 \leq i \leq \ell$. By Lemma 2.2.9, for all $p \in \mathsf{M}$,

$$\mathrm{d}h^i_p, \mathcal{L}_f \mathrm{d}h^i_p, \ldots, \mathcal{L}_f^{\kappa-1}\mathrm{d}h^i_p \in \operatorname{ann}(\mathscr{G}^{(0)}_p), \tag{2.15}$$

Fix $1 \leq j \leq m$ and $0 \leq r \leq \kappa - 1$. Consider $\mathrm{d}h^i(\mathrm{ad}^r_f\, g_j)$. By [27, Lemma 4.1.2][2],

$$\mathrm{d}h^i(\mathrm{ad}^r_f\, g_j) = \sum_{q=0}^{r}(-1)^q \begin{pmatrix} r \\ q \end{pmatrix} \mathcal{L}_f^{r-q}\left(\mathcal{L}_f^q \mathrm{d}h^i\,(g_j)\right).$$

By (2.15), $\mathcal{L}_f^q \mathrm{d}h^i(g_j) = 0$ for all $0 \leq q \leq r$. As a result, $\mathrm{d}h^i(\mathrm{ad}^r_f\, g_j) = 0$. This holds for all $0 \leq r \leq \kappa - 1$ and $1 \leq j \leq m$. As a result, $\mathrm{d}h^i_p \in \operatorname{ann}(\mathscr{G}^{(\kappa-1)}_p)$, for any $p \in \mathsf{M}$. We already have, since the h^i are smooth functions of the state, $\mathrm{d}h^i_p \in \operatorname{ann}(\mathscr{U}_p)$. Since these are exact differential forms, they must also annihilate the involutive closure $\operatorname{inv}(\mathscr{G}^{(\kappa-1)})$. As a result, conclude that

$$\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle^{(\infty)} = \langle \mathrm{d}h^1, \ldots, \mathrm{d}h^\ell, \mathrm{d}t\rangle \subseteq \operatorname{ann}\left(\Gamma^\infty(\mathscr{U} \oplus \operatorname{inv}(\mathscr{G}^{(\kappa-1)}))\right).$$

□

The reverse containment relies on a regularity assumption on the distributions $\mathscr{G}^{(k)}$ and $\mathscr{S}^{(k)}$. Up until now, we have avoided making such an assumption. However, it is standard in the literature to assume certain regularity assumptions about the dimension and smoothness of these distributions.

A $C^\infty(\mathsf{M})$-submodule of vector fields is said to be finitely, non-degenerately generated if there exists a finite set of vector fields which generate the submodule while being linearly independent pointwise. This is precisely the regularity assumption we take on $\mathcal{G}^{(k)}$, $\mathcal{S}^{(k)}$ and their involutivity closures.

Assumption 2.2.11. *The distributions* $\mathscr{G}^{(k)}$, $\operatorname{inv}(\mathscr{G}^{(k)})$ *and* $\mathscr{S}^{(k)}$ *are smooth and regular*[3].

Assumption 2.2.11 ensures that the $C^\infty(\mathsf{M})$-submodules $\mathcal{G}^{(k)} := \Gamma^\infty(\mathscr{G}^{(k)})$ and $\mathcal{S}^{(k)} := \Gamma^\infty(\mathscr{S}^{(k)})$ are locally, finitely, non-degenerately generated. Equipped with Assumption 2.2.11, we can show the reverse containment. The proof is left to Appendix B as it relies on a future result and its proof does not add more insight.

[2] Set $s = 0$, $k = 0$, $r = r$.

[3] It is not necessary to make the assumption on $\operatorname{inv}(\mathscr{S}^{(k)})$ by Lemma 2.1.1.

Lemma 2.2.12. *Let $\kappa \in \mathbb{N}$. If Assumptions 2.2.3 and 2.2.11 holds, then*

$$\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle^{(\infty)} \supseteq \operatorname{ann}\left(\Gamma^\infty(\mathscr{U} \oplus \operatorname{inv}(\mathscr{G}^{(\kappa-1)}))\right).$$

Together, Assumptions 2.2.3 and 2.2.11 along with Lemmas 2.2.10 and 2.2.12 imply equality.

Corollary 2.2.13. *Let $\kappa \in \mathbb{N}$. If Assumptions 2.2.3 and 2.2.11 holds, then*

$$\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle^{(\infty)} = \operatorname{ann}\left(\Gamma^\infty(\mathscr{U} \oplus \operatorname{inv}(\mathscr{G}^{(\kappa-1)}))\right).$$

It turns out that if we relax the discussion to consider only those family of outputs that would yield a *uniform* vector relative degree κ with the system, then the connection between relative degree and the derived flag (2.13) is clear. The next lemma proves the connection in this case.

Lemma 2.2.14. *Suppose Assumptions 2.2.3 and 2.2.11 hold. Take $h^1, \ldots, h^\ell \in C^\infty(\mathsf{M})$, $\ell \leq m$, to be smooth functions of the state. The system (2.6) with output $(h^1, \ldots, h^\ell)$ yields a vector relative degree $(\kappa_1, \ldots, \kappa_1)$ at p_0 if, and only if, there exists an open set $\mathsf{U} \subseteq \mathsf{M}$ containing p_0 where,*

$$\langle \mathrm{d}h^1, \ldots, \mathrm{d}h^\ell \rangle \subseteq \langle \mathcal{I}^{(\kappa_1-1)}, \mathrm{d}t \rangle^{(\infty)} \tag{2.16}$$

and, at p_0,

$$\operatorname{span}_{\mathbb{R}}\{\mathrm{d}h^1_{p_0}, \ldots, \mathrm{d}h^\ell_{p_0}\} \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(\kappa_1)}_{p_0}, \mathrm{d}t_{p_0}\} = \{0\}. \tag{2.17}$$

Proof. Let $h^1, \ldots, h^\ell \in C^\infty(\mathsf{M})$, $\ell \leq m$, be smooth functions of the state with linearly independent differentials at $p_0 \in \mathsf{M}$. First we suppose the conditions (2.16) and (2.17) hold and prove the system (2.6) with output $(h^1, \ldots, h^\ell)$ yields a vector relative degree $(\kappa_1, \ldots, \kappa_1)$ at p_0. For each $i \in \{1, \ldots, \ell\}$, we have that $\mathrm{d}h^i \in \langle \mathcal{I}^{(\kappa_1-1)}, \mathrm{d}t \rangle$, so, by Lemma 2.2.9, for all $p \in \mathsf{M}$,

$$\mathrm{d}h^i_p, \mathcal{L}_f \mathrm{d}h^i_p, \ldots, \mathcal{L}_f^{\kappa_1-2} \mathrm{d}h^i_p \in \operatorname{ann}(\mathscr{G}^{(0)}_p).$$

By [27, Lemma 4.1.2] conclude that,

$$\mathcal{L}_{\mathrm{ad}_f^j g_l} h^i = \mathrm{d}h^i(\mathrm{ad}_f^j g_l) = 0, \quad 0 \leq j \leq \kappa_1 - 2, 1 \leq l \leq m, 1 \leq i \leq \ell.$$

We now show that the associated $\ell \times m$ decoupling matrix has rank ℓ at p_0. In pursuit of contradiction, suppose there exists constants $c_i \in \mathbb{R}$ so that

$$\sum_{i=1}^{\ell} c_i \left(\mathcal{L}_{g_1} \mathcal{L}_f^{\kappa_1 - 1} h^i(p_0) \quad \cdots \quad \mathcal{L}_{g_m} \mathcal{L}_f^{\kappa_1 - 1} h^i(p_0) \right) = 0.$$

Define the smooth function

$$h := \sum_{i=1}^{\ell} c_i h^i \in C^\infty(\mathsf{M}).$$

Since $\mathrm{d}h^i \in \langle \mathcal{I}^{(\kappa_1 - 1)}, \mathrm{d}t \rangle^{(\infty)}$ and $c_i \in \mathbb{R}$, we have $\mathrm{d}h \in \langle \mathcal{I}^{(\kappa_1 - 1)}, \mathrm{d}t \rangle^{(\infty)}$. For any fixed $1 \leq l \leq m$, consider $\mathrm{d}h(\mathrm{ad}_f^{\kappa_1 - 1} g_l)$. By [27, Lemma 4.1.2][4],

$$\mathrm{d}h \left(\mathrm{ad}_f^{\kappa_1 - 1} g_l \right) = \sum_{i=0}^{\kappa_1 - 1} (-1)^i \binom{\kappa_1 - 1}{i} \mathcal{L}_f^{\kappa_1 - 1 - i} \mathcal{L}_{g_l} \mathcal{L}_f^i h.$$

By Lemma 2.2.9, $\mathcal{L}_f^i \mathrm{d}h \in \mathrm{ann}(\mathscr{G}^{(0)})$ for all $0 \leq i \leq \kappa_1 - 2$ and, as a result,

$$\mathrm{d}h \left(\mathrm{ad}_f^{\kappa_1 - 1} g_l \right) = (-1)^{\kappa_1 - 1} \mathcal{L}_{g_l} \mathcal{L}_f^{\kappa_1 - 1} h.$$

Evaluating at p_0

$$\mathrm{d}h_{p_0} \left(\mathrm{ad}_f^{\kappa_1 - 1} g_l(p_0) \right) = (-1)^{\kappa_1 - 1} \mathcal{L}_{g_l} \mathcal{L}_f^{\kappa_1 - 1} h(p_0) = 0.$$

Since $1 \leq l \leq m$ was arbitrary, $\mathrm{d}h_{p_0} \in \mathrm{ann}(\mathscr{G}_{p_0}^{(\kappa_1 - 1)})$. But then, by Lemma 2.2.12 and the fact that $\mathrm{d}h$ is exact,

$$\mathrm{d}h_{p_0} \in \mathrm{span}_{\mathbb{R}} \left\{ \mathscr{I}_{p_0}^{(\kappa_1)}, \mathrm{d}t_{p_0} \right\}$$

which, in tandem with (2.17), implies $\mathrm{d}h_{p_0} = 0$. This contradicts the preliminary assumption that the differentials $\mathrm{d}h^i$ are linearly independent at $p_0 \in \mathsf{U}$. Therefore the decoupling matrix has maximal rank and so it follows that the outputs yield a well-defined vector relative degree.

Now suppose the output $(h^1, \ldots, h^\ell)$ yields a vector relative degree of $(\kappa_1, \ldots \kappa_1)$ at p_0. Fix $1 \leq i \leq m$ and $1 \leq j \leq \ell$. By [27, Lemma 4.1.2] and the definition of relative degree we have

$$\begin{aligned} \mathrm{d}h^j \left(\mathrm{ad}_f^k g_i \right) &= \mathcal{L}_{\mathrm{ad}_f^k g_i} h^j, \\ &= 0, \quad 0 \leq k \leq \kappa_1 - 2. \end{aligned}$$

[4]We consider when $s = 0$, $k = 0$ and $r = \kappa_1 - 1$.

We may conclude $\mathrm{d}h^j \in \mathrm{ann}(\mathscr{G}^{(k)})$ for all $0 \leq k \leq \kappa_1 - 2$. Since $\mathrm{d}h^j$ is exact, it is easy to see that for any two vector fields $X, Y \in \mathcal{G}^{(k)} = \Gamma^\infty(\mathscr{G}^{(k)})$ we have

$$\mathcal{L}_{[X,Y]}h^j = \mathcal{L}_X\left(\mathcal{L}_Y h^j\right) - \mathcal{L}_Y\left(\mathcal{L}_X h^j\right) = 0,$$

So $\mathrm{d}h^j$ annihilates any finite order Lie brackets of vector fields in $\mathcal{G}^{(k)}$. We now show, by induction, that $\mathrm{d}h^j$ annihilates any finite order Lie brackets of vector fields in $\mathcal{S}^{(k)} = \Gamma^\infty(\mathscr{S}^{(k)})$ defined in (2.4). First see that $\mathcal{S}^{(0)} = \mathcal{G}^{(0)}$ so we have $\mathrm{d}h^j \in \mathcal{S}^{(0)}$. Therefore $\mathrm{d}h^j$ also annihilates any finite order Lie brackets of vector fields in $\mathcal{S}^{(0)}$. By way of induction suppose $\mathrm{d}h^j \in \mathcal{S}^{(k)}$ and that it annihilates finite order Lie brackets of vector fields in $\mathcal{S}^k$. Then, by (2.4), the fact that $\mathrm{d}h^j \in \mathrm{ann}(\mathcal{G}^{(k)})$ and $\mathrm{d}h^j \in \mathrm{ann}(\mathcal{S}^{(k)})$ and the fact that it annihilates finite order of Lie brackets for $\mathcal{S}^{(k-1)}$ conclude that $\mathrm{d}h^j \in \mathrm{ann}(\mathcal{S}^{(k)})$. This argument holds until $k = \kappa_1 - 2$. As such we have

$$\mathrm{d}h^j \in \mathrm{ann}(\mathcal{S}^{(k)}), \quad 0 \leq k \leq \kappa_1 - 2,$$

and in particular, since $\mathrm{d}h^j$ is exact,

$$\mathrm{d}h^j \in \mathrm{ann}(\mathrm{inv}(\mathscr{S}^{(\kappa_1-2)})) = \mathrm{ann}(\mathrm{inv}(\mathscr{G}^{(\kappa_1-2)})).$$

Since $\mathrm{d}h^j$ is independent of u, we have that $\mathrm{d}h^j \in \mathrm{ann}(\mathscr{U})$. Together conclude that,

$$\mathrm{d}h^j \in \mathrm{ann}(\Gamma^\infty(\mathscr{U} \oplus \mathrm{inv}(\mathscr{G}^{(\kappa_1-2)}))).$$

By Corollary 2.2.13,

$$\mathrm{d}h^j \in \langle \mathcal{I}^{(\kappa-1)}, \mathrm{d}t\rangle^{(\infty)}$$

Since $j \in \{1, \ldots, m\}$ was arbitrary, this demonstrates that (2.16) holds. We now show that (2.17) holds as well by a contrapositive argument. Suppose that (2.17) does not hold. We will show that the output does not yield a vector relative degree. Given that the condition (2.17) fails, there must exist a linear combination of the $\mathrm{d}h^i$ so that

$$\sum_{i=1}^{\ell} a_i \mathrm{d}h^i_{p_0} \in \mathrm{span}_{\mathbb{R}}\left\{\mathscr{I}^{(\kappa)}_{p_0}, \mathrm{d}t_{p_0}\right\}.$$

By Lemma 2.2.10 and (2.4),

$$\sum_{i=1}^{\ell} a_i \mathrm{d}h^i \in \operatorname{ann}(\mathscr{G}^{(\kappa-1)})$$

and, in particular,

$$\sum_{i=1}^{\ell} a_i \mathrm{d}h^i \left(\mathrm{ad}_f^{\kappa-1} g_j\right) = 0, \quad 1 \leq j \leq m.$$

on U. By [27, Lemma 4.1.2] and because $\sum a_i \mathrm{d}h^i \in \operatorname{ann}(\mathscr{G}^{(k)})$ for $0 \leq k \leq \kappa_2 - 2$,

$$\sum_{i=1}^{\ell} a_i \mathcal{L}_{g_j} \mathcal{L}_f^{\kappa-1} h^i = 0, \quad 1 \leq j \leq m.$$

This contradicts the requirement that the decoupling matrix

$$\begin{pmatrix} \mathcal{L}_{g_1}\mathcal{L}_f^{\kappa_1-1}h^1(p_0) & \cdots & \mathcal{L}_{g_m}\mathcal{L}_f^{\kappa_1-1}h^1(p_0) \\ \vdots & & \vdots \\ \mathcal{L}_{g_1}\mathcal{L}_f^{\kappa_\ell-1}h^\ell(p_0) & \cdots & \mathcal{L}_{g_m}\mathcal{L}_f^{\kappa_\ell-1}h^\ell(p_0) \end{pmatrix}$$

has maximal row rank and so the outputs do not yield a vector relative degree. □

Lemma 2.2.14 only addresses those outputs with uniform relative degree. The non-uniform case involves ensuring the scalar outputs and their Lie derivatives form a basis that generates the ideals in the derived flag adapted to its particular structure. Nevertheless, Lemma 2.2.14 is useful for finding and verifying outputs that yield uniform vector relative degree.

2.3 A Modern Perspective

Having defined and established a number of properties of the derived flag, we now present an alternative perspective on the solution to the state-space, exact feedback linearization problem. To do so, we first define a special indexing of the linearized states in Problem 2.1.1. It can be easily verified that when the system (2.1) is in Brunovský normal form the coordinates ξ can be indexed in the following way. State $\xi^{i,j}$ is the state which, upon j

derivatives along solutions of (2.1), yields the virtual input v^i. The resulting system with these linearized states takes the form,

$$\begin{array}{ccc} \dot{\xi}^{1,\kappa_1}(t) = \xi^{1,\kappa_1-1}(t) & \cdots & \dot{\xi}^{m,\kappa_m}(t) = \xi^{m,\kappa_m-1}(t), \\ \vdots & & \vdots \\ \dot{\xi}^{1,1}(t) = v^1(t) & \cdots & \dot{\xi}^{m,1}(t) = v^m(t), \end{array} \tag{2.18}$$

where $\sum_{i=1}^m \kappa_i = n$. There are m integration chains or towers of lengths $\kappa_1, \ldots, \kappa_m$. With that out of the way, we now remind ourselves of the perspective taken by Hermann in [22].

Hermann saw the feedback linearization problem as a problem of *equivalence*. Suppose, for a moment, we *knew* Problem 2.1.1 was solvable, and we knew the length of the chains that will appear, i.e. the indices $\kappa_1, \ldots, \kappa_m$. The only information missing is the exact expression for the linearized states $\xi^{i,j}$ as a function of the known states x^k. Finding this missing information amounts to searching for a diffeomorphism Ψ between the original coordinates (t, u, x) and the new coordinates (t, v, ξ) with the property that the submodule $\mathcal{D}^{(0)}$ from (2.8) is generated by,

$$\mathcal{D}^{(0)} = \operatorname{span}_{C^\infty}\left\{\frac{\partial}{\partial t} + \sum_{i=1}^{m}\left[v^i\frac{\partial}{\partial \xi^{i,1}} + \sum_{j=2}^{r_i}\xi^{i,j-1}\frac{\partial}{\partial \xi^{i,j}}\right]\right\} \oplus \operatorname{span}_{C^\infty}\left\{\frac{\partial}{\partial v^1}, \ldots, \frac{\partial}{\partial v^m}\right\}. \tag{2.19}$$

The distribution must be preserved under the diffeomorphism and so algebraic properties held by distribution (2.19) must also be held by distribution (2.8). This is significant as there are a number of invariants that *must* hold for the linear system which act naturally through the diffeomorphism and, therefore, must hold of the nonlinear system. For example, the derived flag (2.13) in the coordinates where the system appears linear satisfies the property,

$$\mathcal{I}^{(k)} = \langle \mathrm{d}\xi^{i,j} - \xi^{i,j-1}\,\mathrm{d}t \colon 1 \le i \le m, k < j < \kappa_i \rangle, \qquad k \ge 1,$$

and

$$\langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle = \langle \mathrm{d}t, \{\mathrm{d}\xi^{i,j} \colon 1 \le i \le m, k < j < \kappa_i\} \rangle, \qquad k \ge 1.$$

Notably, the latter ideal is a *differential* ideal generated by exact, smooth one-forms. Since the exterior derivative acts naturally through the diffeomorphism, i.e. Theorem A.5.12, we can inspect these ideals in the original coordinates to deduce what $\mathrm{d}\xi^{i,j}$ are as a function of the $\mathrm{d}x^k$. Integrating these ideals using Frobenius's Theorem produces the required diffeomorphism.

It is unclear (1) whether it is possible to perform this integration, (2) how many integrations must be performed, and (3) which ideals of the derived flag must be integrated. Gardner and Shadwick were the first to recognize this particular property, albeit without presenting this description. Write the generators for the system ideal (2.11) with the same indexing rule used in (2.18) so that, for $j > 2$,

$$\omega^{i,j} = \mathrm{d}\xi^{i,j} - \xi^{i,j-1}\,\mathrm{d}t, \qquad 1 \leq i \leq m,$$

and,

$$\omega^{i,1} = \mathrm{d}\xi^{i,1} - v^i\,\mathrm{d}t, \qquad 1 \leq i \leq m.$$

Observe that, for any fixed $1 \leq i \leq m$ and $j > 1$,

$$\mathrm{d}\omega^{i,j} = \mathrm{d}t \wedge \mathrm{d}\xi^{i,j-1} = \mathrm{d}t \wedge \omega^{i,j-1}.$$

This property is *not* preserved under feedback transformations, but is a desireable property that captures the key property of a controllable, linear system. As a result, we give such a basis a special name.

Definition 2.3.1 (Adapted Basis) . A set of generators $\omega^{i,j}$ for $\mathcal{I}^{(0)}$ is called an *adapted basis* if there exists $\kappa_1, \ldots, \kappa_m \in \mathbb{N}$ so that,

$$\mathrm{d}\omega^{i,1} = \mathrm{d}t \wedge \mathrm{d}v^i, \qquad 1 \leq i \leq m, \tag{2.20}$$

and

$$\mathrm{d}\omega^{i,j} = \mathrm{d}t \wedge \omega^{i,j-1}, \qquad 1 \leq i \leq m, 2 \leq j \leq \kappa_i, \tag{2.21}$$

where $\sum_{i=1}^m \kappa_i = n$.

The generators defined for the original nonlinear system is rarely adapted. The goal is to find a feedback transformation that transforms those generators (2.10) into an adapted basis since, at that point, the linearized states $\xi^{i,j}$ may be recovered. In particular, one needs only to compute the ξ^{i,κ_i} at the top of the integration chain. The rest of the $\xi^{i,j}$ will be Lie derivatives along f of $\xi^{i,j+1}$. This explicitly finds the change of coordinates $\xi = \Phi(x)$.

In search for this transformation, one can instead first seek generators $\omega^{i,j}$ for the original system ideal (2.11) that satisfy (2.20) and (2.21) up to a *congruence* given by,

$$\mathrm{d}\omega^{i,j} - \mathrm{d}t \wedge \omega^{i,j-1} \in \mathcal{I}^{(j)}, \qquad 1 < j \leq \kappa_i. \tag{2.22}$$

Such a basis was what Gardner and Shadwick called an adapted basis. Given such generators, the algorithm proposed by Gardner and Shadwick in [16] turns these congruences into equalities and computes an adapted basis per Definition 2.3.1.

The original algorithm is opaquely presented in [16]. To make matters worse, they do not discuss how to find these generators that satisfy the congruences (2.22) required to use their algorithm which turns out to be non-trivial. Some researchers suggest that this is a major flaw of their algorithm [35]. This is likely a misunderstanding as the existence problem was resolved within a few years. In [48, Theorem 4], the precise method by which one could adapt generators to satisfy these congruences subject to a controllability condition,

$$\mathscr{I}^{(n)} = \{0\}$$

and an involutivity condition,

$$\langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle = \langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}, \qquad 0 \leq k \leq n,$$

was presented. These are the conditions as they were restated in [43, Proposition 12.76]. Together, these conditions allow the rewriting of any generators to satisfy the congruences (2.22) which then permits the use of the GS algorithm presented in [16]. This is the most modern treatment of state-space, exact feedback linearization.

The algorithm both to find generators and to perform the feedback linearization is described in detail in [48] with some simpler examples presented in [36]. Murray demonstrated how the adapted basis can also be used for systems that are not feedback linearizable, e.g. nonholonomic systems, such as to perform path generation [37].

We do not deliberate further except to point out a few observations. All of the works discussed in this section primarily work in the ideals $\mathcal{I}^{(k)}$ of the derived flag (2.13) and *not* of the augmented ideals $\langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle$. However, we know from the end of the preceding section that it is the augmented ideals where the differentials of the outputs explicitly reside (see Lemma 2.2.14). Even in [48], the augmented ideal plays little role but ensuring the congruences can be satisfied by an appropriate rewriting subject to involutivity. This thesis departs from this standard by primarily working with the augmented ideals to better connect the algebraic algorithms on these ideals with the intuition of relative degree.

Chapter 3

Single-Input Systems

In this chapter, we restrict our attention to the problem of finding the feedback transformation for the single-input transverse feedback linearization problem. The single-input case will inform us on the correct dualization of the transverse feedback linearization conditions in the more difficult multi-input case. As such, consider the single-input, nonlinear, control-affine system

$$\dot{x}(t) = f(x(t)) + g(x(t))\, u(t) \tag{3.1}$$

where $f, g \in \Gamma^\infty(\mathsf{T}\mathbb{R}^n)$. This is the single-input variant of the control system (1.4) discussed in Chapters 1 and 2. Let $\mathsf{N} \subset \mathbb{R}^n$ be a closed, embedded n^*-dimensional submanifold passing through the point x_0 that is rendered controlled-invariant by a feedback $u_* : \mathsf{N} \to \mathbb{R}^m$. Fix a point $x_0 \in \mathsf{N}$. In transverse feedback linearization, the controlled invariant submanifold N is treated as a given datum. There exist methods for constructing such sets [7, Section 5]. This set can be viewed as a model of a control specification [20]; the control specification is achieved if the system's state belongs to N (see Section 1.2). This is a common point of view in many control problems [2], e.g., observer design, the synchronization or state agreement problem, path following and regulation. Imposing controlled invariance amounts to requiring that the control specification be feasible.

Section 3.1 states the dual conditions for transverse feedback linearization. We then build a number of useful supporting results in Section 3.2 and then prove the main result in Section 3.3. Finally the chapter concludes with a few simple examples.

3.1 Dual Conditions for Transverse Feedback Linearization

The local transverse feedback linearization problem was already stated, in full, in Section 1.3. As we noted then, this thesis proposes dual conditions with the hope of resolving an algorithm to compute the required feedback transformation. Having established in Chapter 2 how relative degree is related to the derived flag, and recognizing that relative degree plays an important role in transverse feedback linearization as in Theorem 1.3.1, we are inspired to propose new dual conditions for transverse feedback linearization.

Before doing so, let us briefly review the setup. In Section 2.2, we observed that to the single-input, nonlinear control system (3.1) on $\mathbb{R}^n$ we can associate a single-input, nonlinear control system (2.6),

$$\begin{aligned}\dot{t} &= 1,\\ \dot{x} &= f(x) + g(x)\,u.\end{aligned}$$

on the manifold $\mathsf{M} := \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n$ comprised of time ($\mathbb{R}$), control ($\mathbb{R}$) and states ($\mathbb{R}^n$). The control system (2.6) may be viewed as an exterior differential system on M by defining the smooth one-forms (2.10),

$$\omega^j := \mathrm{d}x^i - (f^i(x) + g^i(x)\,u)\,\mathrm{d}t \in \Gamma^\infty(\mathsf{T}^*\mathsf{M}), \qquad 1 \leq i \leq n,$$

and defining the ideal (2.11) over the graded algebra of smooth forms,

$$\mathcal{I}^{(0)} := \langle \omega^1, \dots, \omega^n \rangle \subseteq \Gamma^\infty(\Lambda \mathsf{T}^*\mathsf{M}).$$

Solutions of the control system (2.6) are integral submanifolds of $\mathcal{I}^{(0)}$. To this ideal $\mathcal{I}$, we associate the derived flag

$$\{0\} \subseteq \cdots \subseteq \mathcal{I}^{(i+1)} \subseteq \mathcal{I}^{(i)} \subseteq \cdots \subseteq \mathcal{I}^{(1)} \subseteq \mathcal{I}^{(0)}$$

which is constructed recursively by (2.12) (see Definition 2.2.1).

Solutions to the control system (2.6) may also be viewed as the integral submanifold of the smooth and regular distribution (2.8),

$$\mathscr{D}_p^{(0)} := \mathrm{span}_{\mathbb{R}}\{F_p\} + \mathscr{U}_p,$$

where $F \in \Gamma^\infty(\mathsf{TM})$ was defined in (2.9) to be,

$$F \coloneqq \frac{\partial}{\partial t} + f + gu,$$

and $\mathscr{U}$ is the smooth and regular distribution,

$$\mathscr{U}_p \coloneqq \operatorname{span}_{\mathbb{R}} \left\{ \left.\frac{\partial}{\partial u}\right|_p \right\}.$$

The distribution (2.8) is associated to the $C^\infty(\mathsf{M})$-submodule $\mathcal{D}^{(0)} \coloneqq \Gamma^\infty(\mathscr{D}^{(0)})$. It can be verified that $\mathcal{I}^{(0)} = \operatorname{ann}(\mathcal{D}^{(0)})$. To the smooth and regular distribution $\mathscr{U}$ we associate the $C^\infty(\mathsf{M})$-submodule $\mathcal{U} \coloneqq \Gamma^\infty(\mathscr{U})$.

Furthermore, we established in Lemma 2.2.14 that an output h has relative degree $n - n^*$ at $p_0 \coloneqq (0, u_*(x_0), x_0)$ if, and only if,

$$\mathrm{d}h \in \langle \mathcal{I}^{(n-n^*-1)}, \mathrm{d}t \rangle,$$

and

$$\mathrm{d}h_{p_0} \notin \langle \mathcal{I}^{(n-n^*)}, \mathrm{d}t \rangle^{(\infty)}_{p_0}.$$

Comparing this with the conditions of Theorem 1.3.1, we are motivated to search the ideal $\langle \mathcal{I}^{(n-n^*-1)}, \mathrm{d}t \rangle$ for exact one-forms that are (1) transverse to N — i.e. annihilate tangent vectors — and (2) do not reside in the subsequent ideal $\langle \mathcal{I}^{(n-n^*)}, \mathrm{d}t \rangle^{(\infty)}$.

Before stating the dual conditions, we first translate the data of the problem, N, into an object in M. Define the closed, embedded submanifold

$$\mathsf{L} \coloneqq \{p = (t, u, x) \in \mathsf{M} \colon t = 0, x \in \mathsf{N}\}. \tag{3.2}$$

Note $p_0 = (0,\ u_*(x_0),\ x_0) \in \mathsf{L}$ and, for all $p = (0, u, x) \in \mathsf{L}$,

$$\mathsf{T}_p\mathsf{L} = \operatorname{span}_{\mathbb{R}} \left\{ \left.\frac{\partial}{\partial u}\right|_p \right\} \oplus \mathsf{T}_x\mathsf{N}.$$

Given this, we can now state the conditions upon which we can find the transverse output. The first of these is the controllability condition,

$$\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(n-n^*)}_{p_0}, \mathrm{d}t_{p_0}\} = \operatorname{span}_{\mathbb{R}}\{\mathrm{d}t_{p_0}\}. \tag{Con}$$

The second is an involutivity condition demanding that on an open set $\mathsf{U} \subseteq \mathsf{M}$ containing p_0, for all $p \in \mathsf{U} \cap \mathsf{L}$,

$$\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(k)}, \mathrm{d}t_p\} \subseteq \langle \mathcal{I}^{(k)}, \mathrm{d}t\rangle_p^{(\infty)}, \qquad 0 \leq k \leq n - n^*. \tag{Inv}$$

Lastly, we require that the codistribution in the left-hand side of (Inv) satisfy, for all $p \in \mathsf{U} \cap \mathsf{L}$ and every $0 \leq k \leq n - n^*$,

$$\dim\left(\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_{p_0}^{(k)}, \mathrm{d}t_{p_0}\}\right) = \dim\left(\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(k)}, \mathrm{d}t_p\}\right). \tag{Dim}$$

In this chapter, we prove, subject to the additional regularity Assumption 2.2.3, the following theorem that shows the equivalence of (Con), (Inv) and (Dim) with the conditions of Theorem 1.3.2 in the single-input case only.

Theorem 3.1.1. *Suppose Assumptions 2.2.3 and 2.2.11 hold and $m = 1$. Conditions (Con), (Inv) and (Dim) are equivalent to Theorem 1.3.2 (1) and (2).*

The proof of Theorem 3.1.1 is in Section 3.3. It directly follows from this equivalence that, under the mild regularity Assumptions 2.2.3 and 2.2.11, the controllability condition (Con), involutivity condition (Inv), and constant dimension condition (Dim) together form necessary and sufficient conditions to solve the transverse feedback linearization problem.

Corollary 3.1.2. *Suppose Assumptions 2.2.3 and 2.2.11 hold and $m = 1$. The transverse feedback linearization problem is solvable at x_0 if, and only if, the controllability condition (Con), involutivity condition (Inv) and constant dimension condition (Dim) hold.*

Example 3.1 (Path Following for a Car) . Recall the path following problem posed in Example 1.4. Since Theorem 3.1.1 is true, we expect that the dual conditions (Con), (Inv) and (Dim) hold in an open set of $\mathsf{M} = \mathbb{R} \times \mathbb{R} \times \mathbb{R}^5$ containing $p_0 = (0, u_*(x_0), x_0)$. Let us first verify this.

First we compute the augmented ideals $\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle$ of the derived flag for the system (1.7). In Example 2.1, we computed the derived flag for a similar system (1.5) controlled by two inputs. When input $u^2 = 0$, the derived flag's length is prolonged becoming,

$$\begin{aligned}
\mathcal{I}^{(1)} &= \langle \mathrm{d}x^1 - x^3\cos(x^4)\,\mathrm{d}t, \mathrm{d}x^2 - x^3\sin(x^4)\,\mathrm{d}t, \mathrm{d}x^3, \mathrm{d}x^4 - \frac{x^3}{\ell}\tan(x^5)\,\mathrm{d}t\rangle,\\
\mathcal{I}^{(2)} &= \langle \mathrm{d}x^1 - x^3\cos(x^4)\,\mathrm{d}t, \mathrm{d}x^2 - x^3\sin(x^4)\,\mathrm{d}t, \mathrm{d}x^3\rangle,\\
\mathcal{I}^{(3)} &= \langle \cos(x^4)\,\mathrm{d}x^1 + \sin(x^4)\,\mathrm{d}x^2 - x^3\,\mathrm{d}t, \mathrm{d}x^3\rangle,\\
\mathcal{I}^{(4)} &= \langle \mathrm{d}x^3\rangle.
\end{aligned}$$

Using this, we can compute the augmented ideals $\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle$,

$$\begin{aligned}
\langle \mathcal{I}^{(0)}, \mathrm{d}t\rangle &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}x^4, \mathrm{d}x^5, \mathrm{d}t\rangle\\
\langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}x^4, \mathrm{d}t\rangle,\\
\langle \mathcal{I}^{(2)}, \mathrm{d}t\rangle &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}t\rangle,\\
\langle \mathcal{I}^{(3)}, \mathrm{d}t\rangle &= \langle \mathrm{d}x^3, \cos(x^4)\,\mathrm{d}x^1 - \sin(x^4)\,\mathrm{d}x^2, \mathrm{d}t\rangle,\\
\langle \mathcal{I}^{(4)}, \mathrm{d}t\rangle &= \langle \mathrm{d}x^3, \mathrm{d}t\rangle.
\end{aligned} \tag{3.3}$$

The codistributions $\mathscr{I}^{(\kappa)}$ are defined pointwise by the span of the one-forms that generate $\mathcal{I}^{(\kappa)}$. We also need the differential closures of these augmented ideals for the involutivity condition. They are,

$$\begin{aligned}
\langle \mathcal{I}^{(0)}, \mathrm{d}t\rangle^{(\infty)} &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}x^4, \mathrm{d}x^5, \mathrm{d}t\rangle\\
\langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle^{(\infty)} &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}x^4, \mathrm{d}t\rangle,\\
\langle \mathcal{I}^{(2)}, \mathrm{d}t\rangle^{(\infty)} &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}t\rangle,\\
\langle \mathcal{I}^{(3)}, \mathrm{d}t\rangle^{(\infty)} &= \langle \mathrm{d}x^3, \mathrm{d}t\rangle.
\end{aligned}$$

The only object that remains to compute is the annihilator of the tangent space of L, the lift of N. In this case, since N is generated as a level set of differentially independent functions in an open set containing x_0, the annihilator is spanned by the exterior derivatives of those functions. In particular, on some open set U containing p_0,

$$\mathrm{ann}(\mathsf{TL}) = \mathrm{span}_{C^\infty(\mathsf{U})}\{\mathrm{d}y^1, \mathrm{d}y^2, \mathrm{d}y^3, \mathrm{d}t\}$$

where

$$\begin{aligned}
\mathrm{d}y^1 &:= (x^1 + 2\cos(x^4))\,\mathrm{d}x^1 + (x^2 + 2\sin(x^4))\,\mathrm{d}x^2 - 2(x_1\sin(x^4) - x_2\cos(x^4))\,\mathrm{d}x^4,\\
\mathrm{d}y^2 &:= \cos(x^4)\,\mathrm{d}x^1 + \sin(x^4)\,\mathrm{d}x^2 - (x^1\sin(x^4) - x^2\cos(x^4))\,\mathrm{d}x^4,\\
\mathrm{d}y^3 &:= -\sin(x^4)\sin(x^5)\,\mathrm{d}x^1 + \cos(x^4)\sin(x^5)\,\mathrm{d}x^2\\
&\qquad - ((x^1\sin(x^4) - x^2\cos(x^4))\cos(x^5) + 2\sin(x^5))\,\mathrm{d}x^5.
\end{aligned}$$

At p_0, we have that

$$\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}) = \mathrm{span}_{\mathbb{R}}\{\mathrm{d}x^2 + \sqrt{3}\mathrm{d}x^4, \sqrt{3}\mathrm{d}x^1 + \mathrm{d}x^2 + 4\sqrt{3}\mathrm{d}x^4, \mathrm{d}x^1 - \sqrt{3}\mathrm{d}x^2 + 16\mathrm{d}x^5, \mathrm{d}t\}. \quad (3.4)$$

We are ready to test the dual conditions for transverse feedback linearization. From (3.3), deduce that,

$$\mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(3)}_{p_0}, \mathrm{d}t_{p_0}\} = \mathrm{span}_{\mathbb{R}}\{\mathrm{d}x^3_{p_0}, \mathrm{d}t_{p_0}\}.$$

Combine this with (3.4) to find that

$$\mathrm{ann}\,(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(3)}_{p_0}, \mathrm{d}t_{p_0}\} = \mathrm{span}_{\mathbb{R}}\{\mathrm{d}t_{p_0}\}.$$

In this problem, N is a 2-dimensional submanifold, i.e. $n^* = 2$, and, as such, the controllability (Con) condition holds. The involutivity condition is straightforward to establish in this case. Since,

$$\langle \mathcal{I}^{(0)}, \mathrm{d}t\rangle = \langle \mathcal{I}^{(0)}, \mathrm{d}t\rangle^{(\infty)}, \quad \langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle = \langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle^{(\infty)}, \quad \langle \mathcal{I}^{(2)}, \mathrm{d}t\rangle = \langle \mathcal{I}^{(2)}, \mathrm{d}t\rangle^{(\infty)},$$

and

$$\mathrm{ann}\,(\mathsf{T}_p\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(3)}_p, \mathrm{d}t_p\} = \mathrm{span}_{\mathbb{R}}\{\mathrm{d}t_p\} \subseteq \langle \mathcal{I}^{(3)}, \mathrm{d}t\rangle^{(\infty)},$$

the involutivity condition (Inv) holds automatically. It remains to check the constant dimension condition. For brevity, we leave out this check. It can be verified that the constant dimension condition holds as well. ◀

3.2 Supporting Results

The facts discussed in this section are critical to the proof of the main result, but are also interesting in their own right. They directly connect the distributions used by Isidori, Marino and others in more traditional work with the ideals used in works by Gardner, Shadwick, Schlacher and Schöberl. Under the regularity Assumptions 2.2.3 and 2.2.11, the ideals $\mathcal{I}^{(\kappa)}$ of the derived flag (2.13) are the annihilators of the submodules $\mathcal{S}^{(\kappa-1)} := \Gamma^{\infty}(\mathscr{S}^{(\kappa-1)})$ and a few other vector fields.

Lemma 3.2.1. *Suppose Assumptions 2.2.3 and 2.2.11 hold. Let $\mathcal{I}^{(0)}$ be the system ideal (2.11) corresponding to system (3.1). There exists an open set of M containing p_0 such that in this open set, for any $\kappa \in \{1, \ldots, n - n^{\star}\}$,*

$$\mathcal{I}^{(\kappa)} = \operatorname{ann}\left(\mathcal{D}^{(0)} + \mathcal{S}^{(\kappa-1)}\right).$$

The proof of Lemma 3.2.1 can be found in Appendix B. The vector fields in $\mathcal{D}^{(0)}$ include the vector field F of the system (3.1) and the natural coordinate vector field $\partial/\partial u$ of the control space. The next lemma uses Lemma 3.2.1 to precisely characterize the vector fields annihilated by the augmented ideal $\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle$.

Lemma 3.2.2. *Suppose Assumptions 2.2.3 and 2.2.11 hold. Let $\mathcal{I}^{(0)}$ be the system ideal (2.11) corresponding to system (3.1). There exists an open set on M containing p_0 such that on this open set, for any $\kappa \in \{1, \ldots, n - n^{\star}\}$,*

$$\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle = \operatorname{ann}\left(\mathcal{U} \oplus \mathcal{S}^{(k-1)}\right).$$

The proof of Lemma 3.2.2 can also be found in Appendix B. Using these lemmas, we can connect the known involutivity and controllability conditions for transverse feedback linearization with the dual characterization proposed in this section. First we show what the dual involutivity condition is equivalent to in terms of vector fields.

Lemma 3.2.3. *Suppose Assumptions 2.2.3 and 2.2.11 hold. The involutivity condition (Inv) holds if, and only if, there exists an open set U containing p_0 so that, for every $1 \leq \kappa \leq n - n^{*}$ and for all $p \in \mathsf{U} \cap \mathsf{L}$,*

$$\mathsf{T}_p\mathsf{L} + \operatorname{inv}(\mathscr{G}^{(\kappa-1)})_p = \mathsf{T}_p\mathsf{L} + \mathscr{S}_p^{(\kappa-1)}. \tag{3.5}$$

Proof. Suppose that the involutivity condition (Inv) holds. Then, there exists an open set U containing p_0 so that, for every $0 \leq \kappa \leq n - n^*$ and for all $p \in \mathsf{U} \cap \mathsf{L}$,

$$\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(\kappa)}, \mathrm{d}t_p\} \subseteq \langle \mathfrak{I}^{(\kappa)}, \mathrm{d}t \rangle_p^{(\infty)}.$$

Intersect both sides with $\operatorname{ann}(\mathsf{T}_p\mathsf{L})$ to get,

$$\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(\kappa)}, \mathrm{d}t_p\} \subseteq \operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathfrak{I}^{(\kappa)}, \mathrm{d}t \rangle_p^{(\infty)}.$$

Using the definition of the differential closure, conclude that,

$$\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(\kappa)}, \mathrm{d}t_p\} = \operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathfrak{I}^{(\kappa)}, \mathrm{d}t \rangle_p^{(\infty)}.$$

Compute the annihilator of both sides and invoke Lemma 3.2.2 as well as Corollary 2.2.13 to find this expression is equivalent to,

$$\mathsf{T}_p\mathsf{L} + \mathscr{U}_p + \mathscr{S}_p^{(\kappa-1)} = \mathsf{T}_p\mathsf{L} + \mathscr{U}_p + \operatorname{inv}(\mathscr{G}^{(\kappa-1)})_p.$$

By the construction of the lifted manifold L, we have that $\mathscr{U}_p \subseteq \mathsf{T}_p\mathsf{L}$. As a result,

$$\mathsf{T}_p\mathsf{L} + \mathscr{S}_p^{(\kappa-1)} = \mathsf{T}_p\mathsf{L} + \operatorname{inv}(\mathscr{G}^{(\kappa-1)})_p.$$

The operations performed to arrive here are reversible and, as such, the proof is complete. □

We have not yet established that the dual involutivity condition (Inv) is equivalent to the known involutivity condition,

$$\mathsf{T}_p\mathsf{N} + \operatorname{inv}(\mathscr{G}^{(\kappa-1)})_p = \mathsf{T}_p\mathsf{N} + \mathscr{G}_p^{(\kappa-1)}.$$

It is not obvious, but this condition is equivalent to (Inv) under Assumptions 2.2.3 and 2.2.11. The next lemma, together with Lemma 3.2.3, will establish this.

Lemma 3.2.4. *Suppose Assumption 2.2.11 holds. Let* U *be an open set containing* $p_0 \in \mathsf{L}$. *For all* $p \in \mathsf{U} \cap \mathsf{L}$ *and* $\kappa \in \{1, \ldots, n - n^\star - 1\}$

$$\mathsf{T}_p\mathsf{L} + \operatorname{inv}(\mathscr{G}^{(\kappa-1)})_p = \mathsf{T}_p\mathsf{L} + \mathscr{S}_p^{(\kappa-1)} \tag{3.6}$$

if, and only if, for all $p \in \mathsf{U} \cap \mathsf{L}$ *and every* $\kappa \in \{1, \ldots, n - n^\star - 1\}$

$$\mathsf{T}_p\mathsf{L} + \operatorname{inv}(\mathscr{G}^{(\kappa-1)})_p = \mathsf{T}_p\mathsf{L} + \mathscr{G}_p^{(\kappa-1)}. \tag{3.7}$$

Proof. Sufficiency is the most direct, so we tackle that first. Suppose that there exists an open set U containing p_0 so that, for all $p \in \mathsf{U} \cap \mathsf{L}$ and every $\kappa \in \{1, \ldots, n - n^\star - 1\}$, (3.7) holds. Observe that, by Lemma 2.1.1 and the definition of the involutive closure, we have the inclusion,

$$\mathsf{T}_p\mathsf{L} + \mathrm{inv}(\mathscr{G}^{(\kappa-1)})_p = \mathsf{T}_p\mathsf{L} + \mathrm{inv}(\mathscr{S}^{(\kappa-1)})_p \supseteq \mathsf{T}_p\mathsf{L} + \mathscr{S}_p^{(\kappa-1)}.$$

The reverse inclusion is just as straightforward. Use the definition (2.4) with (3.7) to find,

$$\mathsf{T}_p\mathsf{L} + \mathrm{inv}(\mathscr{G}^{(\kappa-1)})_p = \mathsf{T}_p\mathsf{L} + \mathscr{G}_p^{(\kappa-1)} \subseteq \mathsf{T}_p\mathsf{L} + \mathscr{S}_p^{(\kappa-1)}.$$

Next we show necessity. Suppose that there exists an open set U containing p_0 so that, for all $p \in \mathsf{U} \cap \mathsf{L}$ and every $\kappa \in \{1, \ldots, n - n^\star - 1\}$, (3.6) holds. We make use of induction to show that

$$\mathrm{inv}(\mathscr{G}^{(\kappa-1)})_p \subseteq \mathsf{T}_p\mathsf{L} + \mathscr{G}_p^{(\kappa-1)}.$$

This, along with the fact that $\mathscr{G}^{(\kappa-1)} \subseteq \mathrm{inv}(\mathscr{G}^{(\kappa-1)})$, establishes (3.7).

Base Case ($\kappa = 1$): By (3.6),

$$\mathrm{inv}(\mathscr{G}^{(0)})_p \subseteq \mathsf{T}_p\mathsf{L} + \mathscr{S}_p^{(0)}.$$

Use definition (2.4) to conclude,

$$\mathrm{inv}(\mathscr{G}^{(0)})_p \subseteq \mathsf{T}_p\mathsf{L} + \mathscr{G}_p^{(0)}.$$

Induction ($1 \leq \kappa < n - n^* - 1$): Suppose by way of induction that, for some $\kappa \geq 1$,

$$\mathrm{inv}(\mathscr{G}^{(\kappa-1)})_p \subseteq \mathsf{T}_p\mathsf{L} + \mathscr{G}_p^{(\kappa-1)}. \tag{3.8}$$

Consider $\mathrm{inv}(\mathscr{G}^{(\kappa)})$. By (3.6),

$$\mathrm{inv}(\mathscr{G}^{(\kappa)})_p \subseteq \mathsf{T}_p\mathsf{L} + \mathscr{S}_p^{(\kappa)}.$$

Using definition (2.4) find,

$$\mathrm{inv}(\mathscr{G}^{(\kappa)})_p \subseteq \mathsf{T}_p\mathsf{L} + \mathscr{S}_p^{(\kappa-1)} + [\mathscr{S}^{(\kappa-1)}, \mathscr{S}^{(\kappa-1)}]_p + \mathscr{G}_p^{(\kappa)}.$$

By the definition of the involutive closure,

$$\mathrm{inv}(\mathscr{G}^{(\kappa)})_p \subseteq \mathsf{T}_p\mathsf{L} + \mathrm{inv}(\mathscr{S}^{(\kappa-1)})_p + \mathscr{G}^{(\kappa)}_p.$$

Apply Lemma 2.1.1 so that our expression reads,

$$\mathrm{inv}(\mathscr{G}^{(\kappa)})_p \subseteq \mathsf{T}_p\mathsf{L} + \mathrm{inv}(\mathscr{G}^{(\kappa-1)})_p + \mathscr{G}^{(\kappa)}_p,$$

and invoke the inductive hypothesis (3.8) to arrive at,

$$\mathrm{inv}(\mathscr{G}^{(\kappa)})_p \subseteq \mathsf{T}_p\mathsf{L} + \mathscr{G}^{(\kappa-1)}_p + \mathscr{G}^{(\kappa)}_p.$$

Finally, use the definition (2.3) to arrive at the conclusion,

$$\mathrm{inv}(\mathscr{G}^{(\kappa)})_p \subseteq \mathsf{T}_p\mathsf{L} + \mathscr{G}^{(\kappa)}_p.$$

□

We can now claim that the proposed involutivity condition (Inv) is in-fact equivalent to the known involutivity condition for transverse feedback linearization — up to a projection onto the state-space — as a direct corollary to Lemmas 3.2.3 and 3.2.4.

Corollary 3.2.5. *Suppose Assumptions 2.2.3 and 2.2.11 hold. The involutivity condition (Inv) holds if, and only if, there exists an open set U containing p_0 so that, for all $p \in \mathsf{U} \cap \mathsf{L}$ and every $\kappa \in \{1, \ldots, n - n^\star - 1\}$,*

$$\mathsf{T}_p\mathsf{L} + \mathrm{inv}(\mathscr{G}^{(\kappa-1)})_p = \mathsf{T}_p\mathsf{L} + \mathscr{G}^{(\kappa-1)}_p.$$

Having determined the proposed involutivity condition's equivalence with the known involutivity condition, we now look towards the controllability condition. Unfortunately, the controllability conditions are not equivalent except when the involutivity condition holds. This is expected since the augmented ideals $\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle$ annihilate *more* than just the submodules $\mathcal{G}^{(\kappa-1)}$; they also annihilate vector fields in $\mathcal{S}^{(\kappa-1)}$ which comprises Lie brackets, up to a certain order, of vector fields in $\mathcal{G}^{(\kappa-1)}$. The coming lemma instead concludes this section by translating the controllability condition (Con) into its vector field equivalent.

Lemma 3.2.6. *Suppose Assumptions 2.2.3 and 2.2.11 hold. The controllability condition (Con) holds if, and only if,*

$$\mathsf{T}_{p_0}\mathsf{L} + \mathscr{S}_{p_0}^{(n-n^*-1)} \simeq \mathscr{U}_{p_0} \times \mathbb{R}^n,$$

Proof. Suppose the controllability condition (Con)

$$\text{ann}\,(\mathsf{T}_{p_0}\mathsf{L}) \cap \text{span}_{\mathbb{R}}\{\mathscr{I}_{p_0}^{(n-n^*)}, \mathrm{d}t_{p_0}\} = \text{span}_{\mathbb{R}}\{\mathrm{d}t_{p_0}\},$$

holds. Compute the annihilator of both sides and invoke Lemma 3.2.2 to find,

$$\mathsf{T}_{p_0}\mathsf{L} + \mathscr{U}_{p_0} + \mathscr{S}_{p_0}^{(n-n^*-1)} = \text{span}_{\mathbb{R}}\left\{\frac{\partial}{\partial u^1}, \frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right\}.$$

By construction of L, $\mathscr{U}_{p_0} \subseteq \mathsf{T}_{p_0}\mathsf{L}$ and so

$$\mathsf{T}_{p_0}\mathsf{L} + \mathscr{S}_{p_0}^{(n-n^*-1)} = \text{span}_{\mathbb{R}}\left\{\frac{\partial}{\partial u^1}, \frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right\} \simeq \mathscr{U}_{p_0} \times \mathbb{R}^n.$$

The result directly follows and, since these operations are reversible, the proof is complete. □

3.3 Proof of Theorem 3.1.1

We are now equipped to prove Theorem 3.1.1. The proof is split into two parts, necessity and sufficiency. We will first show that the conditions for Theorem 1.3.2 are necessary for the proposed conditions to hold, then show they are sufficient. Assumptions 2.2.11 and 2.2.3 are taken as standing assumptions for this section and will not be directly called in these proofs although both of them are required to varying degrees throughout the proof. In particular, the use of Lemmas 3.2.5 and 3.2.6 directly rely on these assumptions, the source of which is Corollary 2.2.13 which connects the ideals to the distributions.

Proof. First, we prove the dual conditions are sufficient to solve the single-input transverse feedback linearization problem. Suppose conditions (Con), (Inv) and (Dim) hold. Given the controllability condition (Con), we may use Lemma 3.2.6 to find,

$$\mathsf{T}_{p_0}\mathsf{L} + \mathscr{S}_{p_0}^{(n-n^*-1)} \simeq \mathscr{U}_{p_0} \times \mathbb{R}^n.$$

By definition (2.4),

$$\mathscr{U}_{p_0} \times \mathbb{R}^n \simeq \mathsf{T}_{p_0}\mathsf{L} + \mathscr{S}_{p_0}^{(n-n^*-2)} + [\mathscr{S}^{(n-n^*-2)}, \mathscr{S}^{(n-n^*-2)}]_{p_0} + \mathscr{G}_{p_0}^{(n-n^*-1)}.$$

By the definition of the involutive closure and Lemma 2.1.1,

$$\mathscr{U}_{p_0} \times \mathbb{R}^n \simeq \mathsf{T}_{p_0}\mathsf{L} + \operatorname{inv}(\mathscr{G}^{(n-n^*-2)})_{p_0} + \mathscr{G}_{p_0}^{(n-n^*-1)}.$$

Since we have the involutivity condition (Inv), invoke Corollary 3.2.5 to find,

$$\mathscr{U}_{p_0} \times \mathbb{R}^n \simeq \mathsf{T}_{p_0}\mathsf{L} + \mathscr{G}_{p_0}^{(n-n^*-2)} + \mathscr{G}_{p_0}^{(n-n^*-1)}.$$

Finally use definition (2.3) to conclude,

$$\mathscr{U}_{p_0} \times \mathbb{R}^n \simeq \mathsf{T}_{p_0}\mathsf{L} + \mathscr{G}_{p_0}^{(n-n^*-1)}.$$

Since the subpaces $\mathscr{G}_{p_0}^{(n-n^*-1)}$ and $\mathscr{U}_{p_0}$ are independent and since $\mathsf{T}_{p_0}\mathsf{L} = \mathscr{U}_{p_0} \oplus \mathsf{T}_{x_0}\mathsf{N}$, we may conclude that Theorem 1.3.2 (1),

$$\mathbb{R}^n \simeq \mathsf{T}_{x_0}\mathsf{N} + \mathscr{G}_{x_0}^{(n-n^*-1)},$$

holds. It remains to show Theorem 1.3.2 (2). Before doing so, we first observe that, by the definition (2.3), the fact that system (3.1) has only one input, and Theorem 1.3.2 (1), we must have that, for every $\kappa \in \{1, \ldots, n-n^*\}$,

$$\dim(\mathsf{T}_{x_0}\mathsf{N} + \mathscr{G}_{x_0}^{(\kappa-1)}) = n^* + \kappa.$$

It then follows from the constant dimension condition (Dim) and the very same independence argument given for $\mathscr{U}$ and $\mathscr{G}^{(\kappa-1)}$ that, for every $\kappa \in \{1, \ldots, n-n^*\}$ and $x \in \pi(\mathsf{U}) \cap \mathsf{N}$,

$$\dim(\mathsf{T}_x\mathsf{N} + \mathscr{G}_x^{(\kappa-1)}) = n^* + \kappa, \tag{3.9}$$

shrinking U if necessary. Now we are ready to establish Theorem 1.3.2 (2). Since the involutivity condition (Inv) holds, we can use Corollary 3.2.5 to find,

$$\mathsf{T}_p\mathsf{L} + \operatorname{inv}(\mathscr{G}^{(\kappa-1)})_p = \mathsf{T}_p\mathsf{L} + \mathscr{G}_p^{(\kappa-1)},$$

for all $p \in \mathsf{U} \cap \mathsf{L}$ and every $\kappa \in \{1, \ldots, n - n^\star - 1\}$. It follows that, in the state-space, for all $x \in \pi(\mathsf{U}) \cap \mathsf{N}$,

$$\mathsf{T}_x\mathsf{N} + \mathrm{inv}(\mathscr{G}^{(\kappa-1)})_x = \mathsf{T}_x\mathsf{N} + \mathscr{G}_x^{(\kappa-1)}.$$

Use this alongside (3.9) to arrive at Theorem 1.3.2 (2),

$$\mathsf{T}_x\mathsf{N} + \mathrm{inv}(\mathscr{G}^{(\kappa-1)})_x = \mathsf{T}_x\mathsf{N} + \mathscr{G}_x^{(\kappa-1)} \simeq \mathbb{R}^{n^*+\kappa}.$$

It remains to show that the dual conditions are necessary to solve the single-input transverse feedback linearization problem. Suppose there exists an open subset $\mathsf{V} \subseteq \mathbb{R}^n$ of x_0 where Theorem 1.3.2 (1) and (2) holds. Pick any sufficiently small open subset U of p_0 so that $\pi(\mathsf{U}) = \mathsf{V}$. From Theorem 1.3.2 (1) we have,

$$\mathbb{R}^n \simeq \mathsf{T}_{x_0}\mathsf{N} + \mathscr{G}_{x_0}^{(n-n^*-1)}.$$

By definition (2.4),

$$\mathbb{R}^n \simeq \mathsf{T}_{x_0}\mathsf{N} + \mathscr{G}_{x_0}^{(n-n^*-1)} \subseteq \mathsf{T}_{x_0}\mathsf{N} + \mathscr{S}_{x_0}^{(n-n^*-1)}.$$

Since $\mathsf{T}_{x_0}\mathsf{N} + \mathscr{S}_{x_0}^{(n-n^*-1)} \subseteq \mathsf{T}_{x_0}\mathbb{R}^n \simeq \mathbb{R}^n$, conclude,

$$\mathsf{T}_{x_0}\mathsf{N} + \mathscr{S}_{x_0}^{(n-n^*-1)} \simeq \mathbb{R}^n.$$

It directly follows that, on M,

$$\mathsf{T}_{p_0}\mathsf{L} + \mathscr{S}_{p_0}^{(n-n^*-1)} \simeq \mathscr{U}_{p_0} \times \mathbb{R}^n.$$

We can then use Lemma 3.2.6 to find that the controllability condition (Con) holds.

Next we show that the involutivity condition (Inv) holds. From Theorem 1.3.2 (2) we have, for all $x \in \mathsf{V} \cap \mathsf{N}$ and every $\kappa \in \{1, \ldots, n - n^\star - 1\}$,

$$\mathsf{T}_x\mathsf{N} + \mathrm{inv}(\mathscr{G}^{(\kappa-1)})_x = \mathsf{T}_x\mathsf{N} + \mathscr{G}_x^{(\kappa-1)}.$$

It follows that, for all $p \in \mathsf{U} \cap \mathsf{L}$ and every $\kappa \in \{1, \ldots, n - n^\star - 1\}$,

$$\mathsf{T}_p\mathsf{L} + \mathrm{inv}(\mathscr{G}^{(\kappa-1)})_p = \mathsf{T}_p\mathsf{L} + \mathscr{G}_p^{(\kappa-1)}. \tag{3.10}$$

Conclude that the involutivity condition (Inv) holds by Corollary 3.2.5.

It remains to show the constant dimension condition. Observe that the constant dimension condition, by Lemma 3.2.2, is equivalent to,

$$\dim(\mathsf{T}_{p_0}\mathsf{L} + \mathscr{S}_{p_0}^{(\kappa-1)}) = \dim(\mathsf{T}_p\mathsf{L} + \mathscr{S}_p^{(\kappa-1)}),$$

for all $p \in \mathsf{U} \cap \mathsf{L}$ and every $\kappa \in \{1, \ldots, n - n^\star - 1\}$. Using (3.9) and Lemma 2.1.1, find that,

$$\mathsf{T}_p\mathsf{L} + \mathrm{inv}(\mathscr{S}^{(\kappa-1)})_p = \mathsf{T}_p\mathsf{L} + \mathscr{G}_p^{(\kappa-1)}.$$

Since $\mathscr{G}^{(\kappa-1)} \subseteq \mathscr{S}^{(\kappa-1)} \subseteq \mathrm{inv}(\mathscr{S}^{(\kappa-1)})$, conclude,

$$\mathsf{T}_p\mathsf{L} + \mathscr{S}_p^{(\kappa-1)} = \mathsf{T}_p\mathsf{L} + \mathscr{G}_p^{(\kappa-1)}.$$

The constant dimension condition (Dim) directly follows from this and Theorem 1.3.2 (2).

□

3.4 Illustrative Examples

The stated goal of this thesis is to develop dual conditions for solvability of the transverse feedback linearization problem that ostensibly make it easier to find the transverse output; this can be used to determine the feedback transformation. Returning to this goal, let us consider how one computes the transverse output from the proposed conditions for transverse feedback linearization. Consider the problem of finding an exact smooth one-form $\mathrm{d}h$ so that,

$$\begin{aligned} \mathrm{d}h &\in \langle \mathcal{I}^{(n-n^*-1)}, \mathrm{d}t\rangle^{(\infty)}, \\ \mathrm{d}h\big|_{\mathsf{L}} &\in \mathrm{ann}(\mathsf{TL}). \end{aligned} \tag{3.11}$$

Although it may not appear so, this is a Cauchy problem: it asks to find a smooth function h that is constant (w.l.o.g. zero) on L and solves a system of linear partial differential equations. By Theorem 3.1.1, solvability of this problem is ensured by the controllability (Con), constant dimension (Dim) and involutivity (Inv) conditions. In particular, these conditions are equivalent to the conditions of Theorem 1.3.2 which then ensures the existence of a local transverse output h per Theorem 1.3.1. This local transverse output h vanishes on L by definition so $\mathrm{d}h|_{\mathsf{L}} \in \operatorname{ann}(\mathsf{TL})$. Moreover, h yields a relative degree of $n - n^*$, so, by Lemma 2.2.14, $\mathrm{d}h \in \langle \mathcal{I}^{(n-n^*-1)}, \mathrm{d}t \rangle^{(\infty)}$. As a result, the local transverse output h is a solution to the Cauchy problem (3.11)!

Now we give a sketch of the computation in the single-input case. The locally, simply, finitely, non-degenerately generated differential ideal

$$\langle \mathcal{I}^{(n-n^*-1)}, \mathrm{d}t \rangle^{(\infty)} = \langle \sigma^1, \ldots, \sigma^\ell, \mathrm{d}t \rangle,$$

can be tied to a smooth and regular, involutive distribution,

$$\Delta_p = \operatorname{span}_{\mathbb{R}} \left\{ (X_1)_p, \ldots, (X_{n-\ell})_p, \left.\frac{\partial}{\partial u}\right|_p \right\} \subseteq \mathsf{TM},$$

that satisfies $\operatorname{ann}(\Delta_p) = \langle \mathcal{I}^{(n-n^*-1)}, \mathrm{d}t \rangle_p^{(\infty)}$. Since Δ is involutive, by Frobenius's Theorem, we can assume without loss of generality that the vector fields $X_1, \ldots, X_{n-\ell}, \partial/\partial u$ commute. We then solve the Cauchy problem,

$$\begin{aligned}
\mathcal{L}_{X_1} h &= 0, \\
&\vdots \\
\mathcal{L}_{X_{n-\ell}} h &= 0, \\
\mathcal{L}_{\partial/\partial u} h &= 0, \\
h|_{\mathsf{N}} &= 0.
\end{aligned}$$

for $h \in C^\infty(\mathsf{U})$ that has a non-vanishing differential. This function h can be taken, without loss of generality, to be independent of time t, and is, by construction, independent of u. As a result, we can restrict h onto the state-space $\mathbb{R}^n$ and define an output for the original nonlinear control system (3.1). This output h will yield a relative degree of $n - n^*$ since $0 \neq \mathrm{d}h_{p_0} \in \operatorname{ann}(\mathsf{T}_{p_0}\mathsf{N})$, so, by the controllability condition (Con),

$$\mathrm{d}h_{p_0} \notin \operatorname{span}_{\mathbb{R}} \left\{ \mathscr{I}_{p_0}^{(n-n^*)}, \mathrm{d}t_{p_0} \right\}.$$

The relative degree of h then follows by Lemma 2.2.14. We now apply this method to the path following problem in Example 3.1, as well as a different academic example.

Remark 3.4.1. *Alternatively, Frobenius's Theorem can be used to find the exact generators to the differential ideal* $\langle \mathcal{I}^{(n-n^*-1)}, \mathrm{d}t\rangle^{(\infty)}$. *The transverse output is a function of those smooth functions whose differentials generate this ideal. This is the approach taken in Chapter 4.*

Path Following for a Car

Recall the path following problem posed in Example 3.1. We have already verified the conditions for Theorem 3.1.1, so it now remains to compute the transverse output. From the end of Section 3.1 we recall that we must construct a Cauchy problem associated to the differential ideal $\langle \mathcal{I}^{(n-n^*-1)}, \mathrm{d}t\rangle^{(\infty)}$. In this problem, $n - n^* - 1 = 2$. As a result, we look at the differential ideal,

$$\langle \mathcal{I}^{(2)}, \mathrm{d}t\rangle^{(\infty)} = \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}t\rangle.$$

The distribution annihilated by this differential ideal is,

$$\Delta_p = \mathrm{span}_{\mathbb{R}} \left\{ \left.\frac{\partial}{\partial x^4}\right|_p, \left.\frac{\partial}{\partial x^5}\right|_p, \left.\frac{\partial}{\partial u}\right|_p \right\}.$$

Consider the Cauchy problem

$$\frac{\partial h}{\partial x^4} = 0, \quad \frac{\partial h}{\partial x^5} = 0, \quad \frac{\partial h}{\partial u} = 0, \quad h|_{\mathsf{N}} = 0.$$

A solution to this problem is,

$$h(x) = (x^1)^2 + (x^2)^2 - 12.$$

It is a regular matter to verify that the system (1.7) with output h yields a relative degree of three and the zero dynamics manifold coincides with N in an open set containing x_0.

Academic Example

Consider the nonlinear, control-affine system

$$\dot{x}(t) = \begin{pmatrix} x^3(t) \\ x^3(t) \\ 0 \\ -x^3(t) + x^4(t)x^5(t) \\ -x^1(t) \end{pmatrix} + \begin{pmatrix} -x^2(t) \\ x^1(t) \\ 1 \\ 0 \\ 0 \end{pmatrix} u(t). \tag{3.12}$$

and set

$$\mathsf{N} = \left\{ x \in \mathbb{R}^5 \colon (x^1)^2 + (x^2)^2 - 1 = x^3 = 0 \right\},$$

that is rendered controlled-invariant by input $u_*(x) = 0$. We wish to transverse feedback linearize the system (3.12) with respect to N at $x_0 = (1, 0, 0, 1, 1) \in \mathsf{N}$. Since N has codimension 3, the single transverse output must yield a relative degree of 3. Both functions $(x^1)^2 + (x^2)^2 - 1$ and x^3 that are used to define N fail to yield the required relative degree in an open set containing x_0 and, as a result, cannot act as the transverse output.

However, it turns out that there does exist a transverse output. The dual conditions for transverse feedback linearization hold and, in particular, we have

$$\dim \left(\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(2)}, \mathrm{d}t_p\} \right) = \dim \left(\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(2)}, \mathrm{d}t_{p_0}\} \right) = 2.$$

and

$$\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(3)}, \mathrm{d}t_{p_0}\} = \operatorname{span}_{\mathbb{R}}\{\mathrm{d}t_{p_0}\}.$$

Recall that the outputs of relative degree 3 must reside in differential ideal $\langle \mathcal{I}^{(2)}, \mathrm{d}t \rangle^{(\infty)}$. As a result, we write the system of partial differential equations associated to this ideal and state the Cauchy problem that, when solved, produces the required transverse output. We have that,

$$\langle \mathcal{I}^{(2)}, \mathrm{d}t \rangle^{(\infty)} = \langle \mathrm{d}x^1 - (x^1 - x^2)\mathrm{d}x^3, \mathrm{d}x^2, \mathrm{d}x^4, \mathrm{d}x^5, \mathrm{d}t \rangle,$$

which is associated to the system of partial differential equations,

$$(x^1 - x^2)\frac{\partial h}{\partial x^1} + \frac{\partial h}{\partial x^3} = 0, \quad \frac{\partial h}{\partial u} = 0.$$

The final Cauchy problem is

$$(x^1 - x^2)\frac{\partial h}{\partial x^1} + \frac{\partial h}{\partial x^3} = 0, \quad \frac{\partial h}{\partial u} = 0, \quad h|_{\mathsf{N}} = 0.$$

Using the method of characteristics, a solution to this Cauchy problem is found to be

$$h(x) = (x^2 - x^1)e^{-x^3} - x^2 + \sqrt{1 - (x^2)^2}.$$

The system (3.12) with output h yields a relative degree of three, and h vanishes on N. Therefore, by Theorem 1.3.1, the output h is a transverse output, and may be used with input-output feedback linearization to transverse feedback linearize the dynamics.

Chapter 4

Multi-Input Systems

Having established the conditions for single-input transverse feedback linearization, we now turn to the general multi-input case. Recall the nonlinear, multi-input control-affine system,

$$\dot{x} = f(x) + \sum_{j=1}^{m} g_j(x)\, u^j, \tag{4.1}$$

and its lift onto M in (2.6) given by,

$$\begin{aligned} \dot{t} &= 1, \\ \dot{x} &= f(x) + \sum_{j=1}^{m} g_j(x)\, u^j. \end{aligned} \tag{4.2}$$

Let $\mathsf{N} \subset \mathbb{R}^n$ be a closed, embedded n^*-dimensional submanifold passing through the point x_0 that is rendered controlled-invariant by signal $u_* : \mathsf{N} \to \mathbb{R}^m$. Define the lift of N given by the closed, embedded submanifold

$$\mathsf{L} \coloneqq \{p = (t, u, x) \in \mathsf{M} \colon t = 0, x \in \mathsf{N}\}. \tag{4.3}$$

Fix $p_0 \coloneqq (0,\, u_*(x_0),\, x_0) \in \mathsf{L} \subseteq \mathsf{M}$. In this chapter, we will also lift other submanifolds of the state-space $\mathbb{R}^n$ to M in precisely the same manner.

The conditions for transverse feedback linearization in the multi-input case are precisely the same — (Con), (Inv) and (Dim) — as those presented in Section 3.1 for the single-input case. Although the same proof technique applies in the multi-input case, we seek to provide both a proof *and* an algorithm useful in finding the required feedback transformation.

4.1 The Main Result

The dual conditions for transverse feedback linearization, first stated in Section 3.1, apply just as well to multi-input systems. For convenience, we remind ourselves of the conditions. The first of these is the controllability condition,

$$\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_{p_0}^{(n-n^*)}, \mathrm{d}t_{p_0}\} = \operatorname{span}_{\mathbb{R}}\{\mathrm{d}t_{p_0}\}. \tag{Con}$$

The second is an involutivity condition demanding that on an open set $\mathsf{U} \subseteq \mathsf{M}$ containing p_0, for all $p \in \mathsf{U} \cap \mathsf{L}$,

$$\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(k)}, \mathrm{d}t_p\} \subseteq \langle \mathcal{I}^{(k)}, \mathrm{d}t\rangle_p^{(\infty)}, \qquad 0 \le k \le n - n^*. \tag{Inv}$$

Lastly, we require that the codistribution in the left-hand side of (Inv) satisfies, for all $p \in \mathsf{U} \cap \mathsf{L}$ and every $0 \le k \le n - n^*$,

$$\dim\left(\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_{p_0}^{(k)}, \mathrm{d}t_{p_0}\}\right) = \dim\left(\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(k)}, \mathrm{d}t_p\}\right). \tag{Dim}$$

We now state the main theorem of this thesis: that these conditions may be used to test whether the transverse feedback linearization problem is solvable for multi-input systems.

Theorem 4.1.1 (Main Result)**.** *Suppose Assumptions 2.2.3 and 2.2.11 hold. The system (4.1) is locally transverse feedback linearizable with respect to the closed, embedded and controlled-invariant submanifold $\mathsf{N} \subseteq \mathbb{R}^n$ at x_0 if, and only if, (Con) holds and there exists an open set $\mathsf{U} \subseteq \mathsf{M}$ of p_0 on which conditions (Inv) and (Dim) hold.*

There are a number of constants that appear in the supporting results that follow. First we define the indices, for all $i \ge 0$,

$$\rho_i(p_0) := \dim \frac{\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_{p_0}^{(i)}, \mathrm{d}t_{p_0}\}}{\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_{p_0}^{(i+1)}, \mathrm{d}t_{p_0}\}} \tag{4.4}$$

Using the indices ρ_i define

$$\kappa_i(p_0) \coloneqq \operatorname{card}\{j \colon \rho_j(p_0) \geq i\}, \qquad i \geq 0. \tag{4.5}$$

Unlike in the single-input case, we need not necessarily concern ourselves with ideals in the flag (2.13) up to index $n - n^*$. Observe that

$$\rho_{\kappa_1} = \cdots = \rho_{n-n^*} = 0, \tag{4.6}$$

by definition. If the controllability condition (Con) holds, then we can use the above fact to conclude that

$$\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(\kappa_1)}, \mathrm{d}t_{p_0}\} = \operatorname{span}_{\mathbb{R}}\{\mathrm{d}t_{p_0}\}.$$

As a result, it suffices to look at the ideals in the flag (2.13) with indices up to and including κ_1. When $\kappa_1 = n - n^*$, either there is only one input $m = 1$ or $m - 1$ of the input vector fields g_i act tangent to N.

We call $(\kappa_1, \ldots, \kappa_{n-n^*})$ the *transverse controllability indices* of (4.1) with respect to N at $x_0 = \pi(p_0)$ [40]. Observe that, when the constant dimension condition (Dim) holds, ρ_i and κ_i are constant on an open set of N containing p_0. These indices play a role in the algorithm as they indicate in which ideals components of the transverse output appear. It is fairly straightforward to show that the conditions are necessary for the transverse feedback linearization problem to be solvable. As a result, we now briefly prove the necessity of the dual conditions.

Proof of Theorem 4.1.1 (Necessity). Suppose that, on an open set $\mathsf{V} \subseteq \mathbb{R}^n$ containing x_0 the local transverse feedback linearization problem is solvable. That is, there exists a change of coordinates $(\eta, \xi) \coloneqq \Phi(x)$ and feedback $u \coloneqq \alpha(x) + \beta(x) \begin{bmatrix} v_\parallel & v_\pitchfork \end{bmatrix}^\top$ where, in the new coordinates, the nonlinear control system (4.1) takes the form

$$\begin{aligned} \dot{\eta}(t) &= \overline{f}(\eta(t), \xi(t)) + \sum_{j=1}^{m-\rho_0} \overline{g}_{\parallel,j}(\eta(t), \xi(t))\, v_\parallel^j(t) + \sum_{j=1}^{\rho_0} \overline{g}_{\pitchfork,j}(\eta(t), \xi(t))\, v_\pitchfork^j(t), \\ \dot{\xi}(t) &= A\,\xi(t) + \sum_{j=1}^{\rho_0} b_j\, v_\pitchfork^j(t), \end{aligned} \tag{4.7}$$

and $(A,\, [\, b_1 \cdots b_{\rho_0}\,])$ is in Brunovský normal form, and, furthermore, in (η, ξ)-coordinates the target set N is locally

$$\Phi(\mathsf{V} \cap \mathsf{N}) = \{(\eta, \xi) \in \Phi(\mathsf{V})\colon \xi = 0\}\,.$$

Lifting the dynamics of (4.7) in the same manner as was done to obtain (4.2) gives, on some open set $\mathsf{U} \subseteq \mathsf{M}$ containing p_0,

$$\begin{aligned}
\dot{t} &= 1,\\
\dot{\eta} &= \overline{f}(\eta,\xi) + \sum_{j=1}^{m-\rho_0} \overline{g}_{\parallel,j}(\eta,\xi)\, v_{\parallel}^{j} + \sum_{j=1}^{\rho_0} \overline{g}_{\pitchfork,j}(\eta,\xi)\, v_{\pitchfork}^{j},\\
\dot{\xi} &= A\,\xi + \sum_{j=1}^{\rho_0} b_j\, v_{\pitchfork}^{j},
\end{aligned}$$

and the lifted manifold $\mathsf{L} \subseteq \mathsf{U}$ is locally

$$\mathsf{U} \cap \mathsf{L} = \{(t, v, \eta, \xi) \in \mathsf{U}\colon t = \xi = 0\}\,.$$

By [40, Lemma 4.3], the controllability indices of $(A,\, [\, b_1 \cdots b_{\rho_0}\,])$ equal the transverse controllability indices of (4.4) so $(A,\, [\, b_1 \cdots b_{\rho_0}\,])$ has ρ_0 integration chains of length $\kappa_1 \geq \cdots \geq \kappa_{\rho_0}$. We can therefore index the ξ-coordinates as in (2.18): for each fixed $1 \leq i \leq \rho_0$, write

$$\begin{aligned}
\dot{\xi}^{i,j} &= \xi^{i,j-1}, \quad 2 \leq j \leq \kappa_i,\\
\dot{\xi}^{i,1} &= v_{\pitchfork}^{i}.
\end{aligned}$$

Observe that the generators ω^i of the system ideal (2.11) associated to the $\xi^{i,j}$ take the form

$$\begin{aligned}
\omega^{i,j} &= \mathrm{d}\xi^{i,j} - \xi^{i,j-1}\mathrm{d}t, \quad 2 \leq j \leq \kappa_i,\\
\omega^{i,1} &= \mathrm{d}\xi^{i,1} - v_{\pitchfork}^{i},
\end{aligned}$$

and, as a result, satisfy the adaptation property,

$$\begin{aligned}
\mathrm{d}\omega^{i,j} &= \mathrm{d}t \wedge \omega^{i,j-1}, \quad 2 \leq j \leq \kappa_i,\\
\mathrm{d}\omega^{i,1} &= \mathrm{d}t \wedge v_{\pitchfork}^{i}.
\end{aligned}$$

It is clear then that, for each $p \in \mathsf{U}$, we have the adapted basis structure,

$$\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(k)}, \mathrm{d}t_p\} = \operatorname{span}_{\mathbb{R}}\{\mathrm{d}t_p\} + \operatorname{span}_{\mathbb{R}}\{\mathrm{d}\xi_p^{i,j} \colon 1 \le i \le \rho_0, j > k\}.$$

From this we can deduce the TFL conditions. The constant dimension condition (Dim) follows directly. The controllability condition (Con) follows from considering the index $k = n - n^*$. Observe there are no $\xi^{i,j}$ with $j > n - n^*$ since that would imply the existence of more than $n - n^*$ transverse directions to N. Therefore

$$\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(n-n^*)}, \mathrm{d}t_p\} = \operatorname{span}_{\mathbb{R}}\{\mathrm{d}t_p\}.$$

The involutivity condition (Inv) follows because we have exact generators $\mathrm{d}t$ and $\mathrm{d}\xi^{i,j}$ that generate the codistribution,

$$\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(k)}, \mathrm{d}t_p\}$$

for all $0 \le k \le n - n^*$. □

The rest of this chapter is dedicated to establishing the dual conditions for transverse feedback linearization are sufficient, and to exposing the proposed algorithm for transverse feedback linearization.

4.2 Supporting Results

Lemma 2.2.14 of Section 2.2 established that, under mild regularity assumptions, system (4.2) with output h yielding a uniform vector relative degree at p_0 must have a differential that lives in a specific ideal but not live in the subsequent augmented ideal of the derived flag (2.13). Now we ask a different question: Given an output that yields a uniform vector relative degree, how do we find additional scalar outputs that (1) yield a smaller uniform vector relative degree and (2) combine with previously known scalar outputs to yield a vector relative degree? All the while, we must ensure that (3) the new outputs vanish on the target manifold N.

Points (1) and (2) are classically performed by “adapting” the basis of *exact* generators for the derived flag (2.13) per Definition 2.3.1. This is precisely what we wish to replicate except that the adapted portion of the basis acts transverse to N. This is where point (3) imposes a greater degree of difficulty in the adaptation process precisely because we ask that the differentials of the outputs $\mathrm{d}h^i$ annihilate tangent vectors to N. At a high level, we present three procedures that together will be used to correctly adapt the derived flag. These are:

(a) finding generators that "drop off" when computing the derived flag (2.13),

(b) grouping generators into those that annihilate tangent vectors to L and those that do not, and

(c) rewriting generators so that the induced output has a full rank decoupling matrix.

Note that these steps are performed repeatedly throughout the algorithm. This section presents technical results that demonstrate how to perform steps (a) and (b). Step (c) is presented in the proof of the main result.

Before discussing these subprocedures in any detail, we must know the dimension of the subspace in an ideal of the derived flag (2.13) that annihilates tangent vectors to L. The first lemma shows how the controllability condition (Con) determines this.

Lemma 4.2.1. *If* (Con)*, then for all* $0 \le k \le n - n^*$

$$\dim\left(\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(k)}_{p_0}, \mathrm{d}t_{p_0}\}\right) = 1 + \textstyle\sum_{i=k}^{n-n^*-1} \rho_i,$$

and $n - n^* = \sum_{i=0}^{n-n^*-1} \rho_i$.

Proof. By (Con) we have

$$\dim\left(\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(n-n^*)}_{p_0}, \mathrm{d}t_{p_0}\}\right) = 1$$

so the formula holds for $k = n - n^*$. Suppose, by way of induction, that for some $1 \le k \le n - n^*$

$$\dim\left(\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(k)}_{p_0}, \mathrm{d}t_{p_0}\}\right) = 1 + \textstyle\sum_{i=k}^{n-n^*-1} \rho_i. \tag{4.8}$$

Consider

$$\dim\left(\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(k-1)}_{p_0}, \mathrm{d}t_{p_0}\}\right).$$

By (4.4)

$$\dim(\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(k-1)}_{p_0}, \mathrm{d}t_{p_0}\}) = \dim(\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(k)}_{p_0}, \mathrm{d}t_{p_0}\}) + \rho_{k-1}.$$

Apply the inductive hypothesis (4.8) and conclude

$$\dim\left(\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(k-1)}_{p_0}, \mathrm{d}t_{p_0}\}\right) = 1 + \textstyle\sum_{i=k-1}^{n-n^*-1} \rho_i.$$

We now verify the final fact. First observe

$$\mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(0)}_{p_0}, \mathrm{d}t_{p_0}\} = \mathrm{span}_{\mathbb{R}}\{\mathrm{d}x^1_{p_0}, \ldots, \mathrm{d}x^n_{p_0}, \mathrm{d}t_{p_0}\}.$$

By construction of L we have

$$\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}) \subseteq \mathrm{span}_{\mathbb{R}}\{\mathrm{d}x^1_{p_0}, \ldots, \mathrm{d}x^n_{p_0}, \mathrm{d}t_{p_0}\}.$$

Therefore

$$\dim\left(\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(0)}_{p_0}, \mathrm{d}t_{p_0}\}\right) = \dim\left(\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L})\right) = 1 + n - n^*.$$

Combining this with the formula

$$\dim\left(\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(0)}_{p_0}, \mathrm{d}t_{p_0}\}\right) = 1 + \textstyle\sum_{i=0}^{n-n^*-1} \rho_i,$$

completes the proof. □

The previous result concerned the ideal but not its differential closure. Next we show that, when the involutivity condition holds, we can work with either the ideal or its differential closure as long as we are only concerned with the differentials that annihilate tangent vectors to L.

Proposition 4.2.2. *If* (Inv) *holds on some open set* U *containing* p_0*, then for all* $p \in \mathsf{U} \cap \mathsf{L}$

$$\mathrm{ann}(\mathsf{T}_p\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(k)}_p, \mathrm{d}t_p\} = \mathrm{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathcal{I}^{(k)}, \mathrm{d}t\rangle^{(\infty)}_p, \quad 0 \le k \le n - n^*.$$

Proof. Fix $p \in \mathsf{U} \cap \mathsf{L}$ and intersect both sides of the involutivity condition (Inv) with $\mathrm{ann}(\mathsf{T}_p\mathsf{L})$. Use the fact that $\langle \mathcal{I}^{(k)}, \mathrm{d}t\rangle^{(\infty)} \subseteq \langle \mathcal{I}^{(k)}, \mathrm{d}t\rangle$ to arrive at the equality. □

As mentioned in Section 2.3, finding outputs of vector relative degree amounts to finding an appropriately adapted basis for the derived flag. That is, find generators for $\mathcal{I}^{(0)}$ that act as an adapted basis per Definition 2.3.1. Recall that, when $\mathsf{N} = \{x_0\}$, the transverse feedback linearization problem reduces to the exact state-space feedback linearization problem [40]. As a result, we expect our algorithm to apply just as well to feedback linearization. In the exact feedback linearization algorithm presented in [16], the generators for the differential ideals $\langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}$ are assumed to satisfy,

$$
\begin{aligned}
\langle \mathcal{I}^{(0)}, \mathrm{d}t \rangle^{(\infty)} &= \langle \omega^1, \ldots, \ldots, \omega^n, \mathrm{d}t \rangle, \\
\langle \mathcal{I}^{(1)}, \mathrm{d}t \rangle^{(\infty)} &= \langle \omega^1, \ldots, \omega^{n-\rho_0}, \mathrm{d}t \rangle, \\
&\vdots \\
\langle \mathcal{I}^{(\kappa_1 - 1)}, \mathrm{d}t \rangle^{(\infty)} &= \langle \omega^1, \omega^2, \mathrm{d}t \rangle, \\
\langle \mathcal{I}^{(\kappa_1)}, \mathrm{d}t \rangle^{(\infty)} &= \langle \mathrm{d}t \rangle.
\end{aligned}
\tag{4.9}
$$

The generators “drop off” as the derived flag is computed. This is precisely what is meant by subprocedure (a) and is what is meant by Gardner and Shadwick’s adapted basis. Note that the procedure provided by Tilbury and Sastry in [48, Theorem 4] cannot be used as it relies on the conditions for state-space, *exact* feedback linearization. This is why we explicitly demonstrate that the re-adaptation process, subprocedure (a) in particular, can be applied when required.

Even given an adapted basis, finding exact generators that annihilate tangent vectors to L is itself challenging. Constructing one-forms that annihilate L directly from known exact forms can ruin their exactness. In reality, one must solve a Cauchy problem for every component of the transverse output, as seen in the single-input case in Section 3.4. This is tenable in the single-input case, but is not a satisfying solution in the multi-input case especially when seeking multiple, independent, scalar, transverse outputs. To avoid this, we introduce maps H_k whose differential is the exact generator for $\langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}$. The re-adaptation process mentioned earlier, subprocedure (a), amounts to rewriting the components of these smooth maps so that the image of the pullback $\mathrm{Im}(\mathbf{D}H_k)^* = \langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}$ remains the same.

Remark 4.2.3. *Manipulating the integrals of the exact generators suggests integrating all the differential closures of the augmented ideals in the derived flag* (2.13)*. In Section 4.4, we show that integration is only required for those differential ideals* $\langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}$ *at which* k *is a* distinct *transverse controllability index. This is in keeping with the simplicity of Section 3.4 where integration is only performed at index* $\kappa_1 - 1 = n - n^* - 1$.

We start by showing the existence of the aforementioned maps. The TFL conditions — (Con), (Inv) and (Dim) — allow us to construct the map H_k explicitly with a specific rank deficiency when restricted to L. This deficiency will ultimately be used to construct the transverse outputs.

Lemma 4.2.4. *Suppose Assumption 2.2.3 holds. If* (Con) *holds at* p_0 *and there exists an open set* U *containing* p_0 *where* (Inv) *and* (Dim) *hold, then for every* $0 \leq k \leq n - n^*$ *there exists an integer* $\ell_k \geq 1 + \sum_{i=k}^{n-n^*-1} \rho_i$*, a possibly smaller open set* $\mathsf{V} \subseteq \mathsf{U}$ *containing* p_0*, and a smooth map* $H_k : \mathsf{V} \to \mathbb{R}^{\ell_k}$ *satisfying the characteristic property*

$$\langle \mathrm{d}H_k^1, \ldots, \mathrm{d}H_k^{\ell_k} \rangle = \langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}, \tag{4.10}$$

with constant rank, on $\mathsf{V} \cap \mathsf{L}$, *equal to*

$$\operatorname{rank}\left(H_k|_{\mathsf{V}\cap\mathsf{L}} \right) = \ell_k - \left(1 + \textstyle\sum_{i=k}^{n-n^*-1} \rho_i\right).$$

Proof. By Assumption 2.2.3, the differential ideal $\langle \mathcal{I}^{(k)}, dt \rangle^{(\infty)}$ is simply, finitely, non-degenerately generated so, by Frobenius's Theorem (Theorem A.6.4), there exists ℓ_k exact generators $\mathrm{d}h^1, \ldots, \mathrm{d}h^{\ell_k}$ on some open neighbourhood $\mathsf{V} \subseteq \mathsf{U}$ of p_0. Define a smooth map $H_k : \mathsf{V} \to \mathbb{R}^{\ell_k}$ by

$$H_k(p) \coloneqq \left(h^1(p), \ldots, h^{\ell_k - 1}(p), t \right).$$

Without loss of generality, take h^i to be smooth functions of the state. By construction, H_k satisfies the characteristic property (4.10).

We already know, by Assumption 2.2.3, that H_k has constant rank. It is not directly obvious that $H_k|_{\mathsf{V}\cap\mathsf{L}}$ has constant rank as well. Since (Inv) holds over V, invoke Proposition 4.2.2 to find

$$\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle_p^{(\infty)} = \operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(k)}, \mathrm{d}t_p\}, \tag{4.11}$$

for all $p \in \mathsf{V} \cap \mathsf{L}$. It then follows by (Dim) that, for $p \in \mathsf{V} \cap \mathsf{L}$,

$$\dim \left(\operatorname{ann}(\mathsf{T}_p \mathsf{L}) \cap \langle \mathcal{J}^{(k)}, \mathrm{d}t \rangle_p^{(\infty)} \right) = \text{constant}.$$

Then use the characteristic property (4.10) to determine that $H_k|_{\mathsf{V}\cap\mathsf{L}}$ must have constant rank.

We now proceed by directly computing its rank at a point $p \in \mathsf{V} \cap \mathsf{L}$. Because the rank of $H_k|_{\mathsf{V}\cap\mathsf{L}}$ is constant, it suffices to compute its rank at p_0. Compute the dimension on both sides of (4.11) and invoke Lemma 4.2.1 to find,

$$\dim \left(\operatorname{ann}(\mathsf{T}_{p_0} \mathsf{L}) \cap \langle \mathcal{J}^{(k)}, \mathrm{d}t \rangle_{p_0}^{(\infty)} \right) = 1 + \textstyle\sum_{i=k}^{n-n^*-1} \rho_i.$$

It immediately follows that

$$\operatorname{rank} H_k|_{\mathsf{V}\cap\mathsf{L}} = \ell_k - \left(1 + \textstyle\sum_{i=k}^{n-n^*-1} \rho_i \right)$$

□

The proof of Lemma 4.2.4 did not specifically rely on the way the map H_k was constructed (as the integral of a Frobenius system). The proof holds without modification for *any* map H_k that satisfies the characteristic property (4.10). The next corollary states this fact.

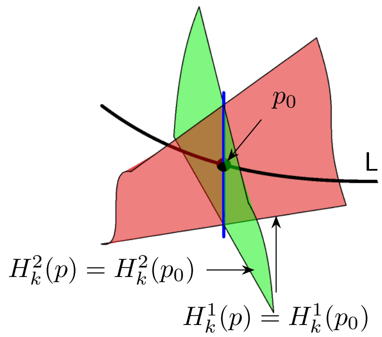
Corollary 4.2.5. *Suppose* (Con) *holds and there exists an open set* U *containing* p_0 *where* (Inv) *and* (Dim) *hold. If a smooth map* $H_k : \mathsf{U} \to \mathbb{R}^{\ell_k}$ *satisfies the characteristic property* (4.10) *then it has constant rank on* $\mathsf{U} \cap \mathsf{L}$ *equal to*

$$\operatorname{rank} \left(H_k|_{\mathsf{U}\cap\mathsf{L}} \right) = \ell_k - \left(1 + \textstyle\sum_{i=k}^{n-n^*-1} \rho_i \right).$$

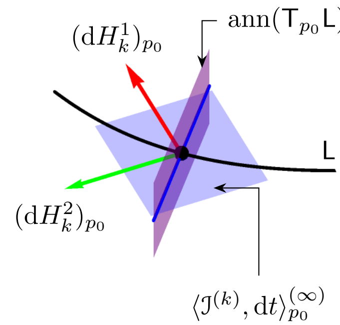
We now ask whether the H_k can be "adapted" to include components that are constant (without loss of generality, zero) on L. In general, level sets of the form

$$\{p \in \mathsf{M} \colon H_k^i(p) = H_k^i(p_0)\}$$

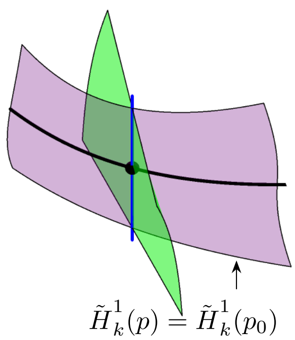
do not contain L, as depicted in Figure 4.1a. Figure 4.1b shows how the corresponding generators for the codistribution $\langle \mathcal{J}^{(k)}, \mathrm{d}t \rangle^{(\infty)}$ do not annihilate vectors tangent to L although



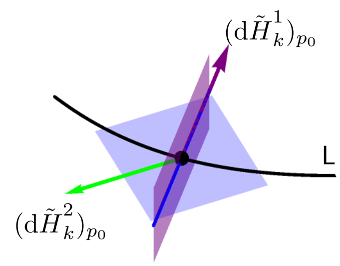
(a) *The intersection of the level sets of H_k^i form an integral submanifold (light blue) of $\langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}$ passing through p_0. Neither level set contains* L.



(b) *The original choice of generators do not annihilate vectors tangent to* L.



(c) *There exists a map $\tilde{H}_k$ so that the zero locus of the leading components contain* L *while preserving the integral submanifold.*



(d) *The smooth one-form $\mathrm{d}\tilde{H}^1$ lives in $\langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}$ and annihilates the tangent space of* L.

Figure 4.1: *An arbitrary set of generators (red, green) for the codistribution $\langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}$ (blue) is adapted to annihilate the tangent space of* L *(black).*

the codistribution has a non-trivial intersection with $\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L})$. We can use the rank deficiency of H_k on L to construct an adaptation of H_k where the leading $1 + \sum_{i=k}^{n-n^*-1} \rho_i$ components have level sets that locally contain L. Figure 4.1c shows the level sets of the newly rewritten $\tilde{H}_k$. At a point $p \in \mathsf{L}$, the leading components' differential lives in $\mathrm{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle_p^{(\infty)}$ as seen in Figure 4.1d.

Lemma 4.2.6. *If $H_k : \mathsf{V} \to \mathbb{R}^{\ell_k}$ is a smooth map satisfying the characteristic prop-*

erty (4.10) and its restriction has constant rank equal to

$$\operatorname{rank}\left(H_k|_{\mathsf{V}\cap\mathsf{L}}\right) = \ell_k - \rho,$$

then there exists a map $\tilde{H}_k : \mathsf{V} \to \mathbb{R}^{\ell_k}$ *which satisfies the characteristic property (4.10) and, for all* $p \in \mathsf{V}\cap\mathsf{L}$, $v_p \in \mathsf{T}_p\mathsf{L}$, $i \in \{1, \dots, \rho\}$, $\mathrm{d}\tilde{H}_k^i(v_p) = 0$.

Proof. Apply [31, Rank Theorem (Proposition 4.12)] and shrink V if necessary to find a coordinate chart φ for L and a coordinate chart ψ for $\mathbb{R}^{\ell_k}$ so the composition $\psi \circ H_k|_{\mathsf{V}\cap\mathsf{L}} \circ \varphi^{-1}$ takes the form

$$\psi \circ H_k|_{\mathsf{V}\cap\mathsf{L}} \circ \varphi^{-1} = (0,\ \dots,\ 0,\ \star,\ \dots,\ \star),$$

with ρ leading zeros. Define $\tilde{H}_k \coloneqq \psi \circ H_k$. The map $\tilde{H}_k$ still satisfies the characteristic property (4.10) but is now "adapted" so that the first ρ components vanish on L and, consequently, their differentials must annihilate tangent vectors to L as required. □

Remark 4.2.7. *Let* $\mathfrak{J}$ *be a differential ideal associated with codistribution* $\mathscr{J}$*, and let* $\mathsf{L} \subseteq \mathsf{M}$ *be a closed, embedded submanifold. Consider the Cauchy problem*

$$\mathrm{d}h \in \mathfrak{J},$$
$$h|_{\mathsf{L}} = 0,$$

locally around a point $p_0 \in \mathsf{L}$*, and ask to find the maximal set of differentially independent functions that solve this problem. Lemma 4.2.6 suggests that a necessary and sufficient condition is that the non-empty condition*

$$\dim\left(\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathscr{J}_{p_0}\right) > 0,$$

and constant dimension condition,

$$\dim\left(\operatorname{ann}(\mathsf{T}_{p}\mathsf{L}) \cap \mathscr{J}_{p}\right) = \dim\left(\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathscr{J}_{p_0}\right),$$

holds for all p *in a sufficiently small neighbourhood of* p_0*. This should be compared with similar results such as [31, Theorem 19.27].*

The established facts ensure that, assuming (Con) holds at p_0 and there exists an open set U so that (Inv) and (Dim) hold, we can write, for all $p \in \mathsf{U} \cap \mathsf{L}$,

$$\begin{aligned}
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathcal{I}^{(0)}, \mathrm{d}t\rangle_p^{(\infty)} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}H_0^1, \ldots, \ldots, \ldots, \mathrm{d}H_0^{n-n^*}, \mathrm{d}t\},\\
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle_p^{(\infty)} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}H_1^1, \ldots, \ldots, \mathrm{d}H_1^{n-n^*-\rho_0}, \mathrm{d}t\},\\
&\;\vdots\\
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathcal{I}^{(\kappa_1-1)}, \mathrm{d}t\rangle_p^{(\infty)} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}H_{\kappa_1-1}^1, \ldots, \mathrm{d}H_{\kappa_1-1}^{\rho_{\kappa_1-1}}, \mathrm{d}t\},\\
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathcal{I}^{(\kappa_1)}, \mathrm{d}t\rangle_p^{(\infty)} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}t\}.
\end{aligned} \tag{4.12}$$

None of the previous results guarantee that components of H_k are also components of H_{k-1} even though we know that

$$\langle \mathrm{d}H_k^1, \ldots, \mathrm{d}H_k^{\ell_k}\rangle = \langle \mathcal{I}^{(k)}, \mathrm{d}t\rangle^{(\infty)} \subseteq \langle \mathcal{I}^{(k-1)}, \mathrm{d}t\rangle^{(\infty)} = \langle \mathrm{d}H_{k-1}^1, \ldots, \mathrm{d}H_{k-1}^{\ell_{k-1}}\rangle$$

The purpose of the coming lemmas is to ensure that we can always rewrite (4.12) as,

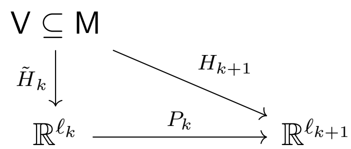
$$\begin{aligned}
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathcal{I}^{(0)}, \mathrm{d}t\rangle_p^{(\infty)} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}H_{\kappa_1-1}^1, \ldots, \ldots, \ldots, \mathrm{d}H_0^{n-n^*}, \mathrm{d}t\},\\
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle_p^{(\infty)} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}H_{\kappa_1-1}^1, \ldots, \ldots, \mathrm{d}H_1^{n-n^*-\rho_0}, \mathrm{d}t\},\\
&\;\vdots\\
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathcal{I}^{(\kappa_1-1)}, \mathrm{d}t\rangle_p^{(\infty)} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}H_{\kappa_1-1}^1, \ldots, \mathrm{d}H_{\kappa_1-1}^{\rho_{\kappa_1-1}}, \mathrm{d}t\},\\
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathcal{I}^{(\kappa_1)}, \mathrm{d}t\rangle_p^{(\infty)} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}t\}.
\end{aligned} \tag{4.13}$$

Pay close attention to the subtle difference between (4.13) and (4.12): if a differential appears as a generator in one ideal, it appears as a generator in all the preceding ideals of the derived flag. We would like the components of our smooth maps to satisfy this same property. To do this, noting that $\ell_k \geq \ell_{k+1}$ because $\langle \mathcal{I}^{(k)}, \mathrm{d}t\rangle^{(\infty)} \supseteq \langle \mathcal{I}^{(k+1)}, \mathrm{d}t\rangle^{(\infty)}$, define the linear projection $P_k : \mathbb{R}^{\ell_{k+1}} \times \mathbb{R}^{\ell_k - \ell_{k+1}} \to \mathbb{R}^{\ell_{k+1}}$ defined by the matrix, $P_k = \left[\, I_{\ell_{k+1}} \; 0 \,\right]$. Then, rewrite the components of H_k so that the following diagram commutes.

$$\begin{array}{ccccccccc}
\mathsf{V} \subseteq \mathsf{M} & \xrightarrow{\mathrm{id}} & \mathsf{V} \subseteq \mathsf{M} & \xrightarrow{\mathrm{id}} & \cdots & \xrightarrow{\mathrm{id}} & \mathsf{V} \subseteq \mathsf{M} & \xrightarrow{\mathrm{id}} & \mathsf{V} \subseteq \mathsf{M}\\
\downarrow{\scriptstyle H_0} & & \downarrow{\scriptstyle H_1} & & & & \downarrow{\scriptstyle H_{n-n^*-1}} & & \downarrow{\scriptstyle H_{n-n^*}}\\
\mathbb{R}^{\ell_0} & \xrightarrow{P_0} & \mathbb{R}^{\ell_1} & \xrightarrow{P_1} & \cdots & \xrightarrow{P_{n-n^*-2}} & \mathbb{R}^{\ell_{n-n^*-1}} & \xrightarrow{P_{n-n^*-1}} & \mathbb{R}^{\ell_{n-n^*}}
\end{array}$$

Maps H_k that make this diagram commute have components that are subsumed in the "larger" maps $H_{k-1}, \ldots, H_0$. This can be done, up to a reordering in the projection, to preserve the fact that the leading components of H_k vanish on L. To prove that such a construction is possible, we need only prove that a smaller adaptation is possible.

Proposition 4.2.8. *Let $k \geq 0$, $\ell_k > \ell_{k+1} > 0$, let $P_k : \mathbb{R}^{\ell_{k+1}} \times \mathbb{R}^{\ell_k - \ell_{k+1}} \to \mathbb{R}^{\ell_{k+1}}$, $P_k(x, y) = x$, and let U be an open set containing p_0. If $H_{k+1} : \mathsf{U} \to \mathbb{R}^{\ell_{k+1}}$ and $H_k : \mathsf{U} \to \mathbb{R}^{\ell_k}$ are smooth maps satisfying the characteristic property (4.10), then on a possibly smaller open set V containing p_0, there exists a smooth map $\tilde{H}_k : \mathsf{V} \to \mathbb{R}^{\ell_k}$ that makes the diagram,*



commute and $\langle \mathrm{d}\tilde{H}_k^1, \ldots, \mathrm{d}\tilde{H}_k^{\ell_k} \rangle = \langle \mathrm{d}H_k^1, \ldots, \mathrm{d}H_k^{\ell_k} \rangle = \langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}$.

Proof. Write $H_{k+1} = (H_{k+1}^1, \ldots, H_{k+1}^{\ell_{k+1}})$. Since $\langle \mathcal{I}^{(k+1)}, \mathrm{d}t \rangle^{(\infty)}$ is contained in $\langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}$ and by the characteristic property (4.10),

$$\langle \mathrm{d}H_{k+1}^1, \ldots, \mathrm{d}H_{k+1}^{\ell_{k+1}} \rangle \subseteq \langle \mathrm{d}H_k^1, \ldots, \mathrm{d}H_k^{\ell_k} \rangle.$$

It follows that we can pick the $\ell_{k+1} - \ell_k$ components of H_k that are differentially independent from the components H_{k+1}^i at p_0. Take these differentially independent components of H_{k+1} to be the last $\ell_{k+1} - \ell_k$ components without loss of generality. Define

$$\tilde{H}_k \coloneqq (H_{k+1}^1, \ldots, H_{k+1}^{\ell_{k+1}}, H_k^{\ell_{k+1} - \ell_k + 1}, \ldots, H_k^{\ell_k}),$$

and observe that, on a sufficiently small open set V containing p_0, this map will satisfy

$$\langle \mathrm{d}\tilde{H}_k^1, \ldots, \mathrm{d}\tilde{H}_k^{\ell_k} \rangle = \langle \mathrm{d}H_k^1, \ldots, \mathrm{d}H_k^{\ell_k} \rangle$$

and $P_k \circ \tilde{H}_k = H_{k+1}$. □

Proposition 4.2.8 implies that, for every $0 \leq k \leq n - n^* - 1$, there exists H_k, H_{k+1} so that $H_{k+1} = P_k \circ H_k$. Geometrically, the level sets of the components of H_{k+1} are subsumed by the level sets of the components of H_k. This is depicted in Figure 4.2. Together with Lemma 4.2.6, we can find a sequence of maps whose leading components vanish on L while making the aforementioned diagram commute (up to a reordering in the projections). The next corollary states this fact.

Corollary 4.2.9. *If* (Con) *holds and there exists an open set* $\mathsf{U} \subseteq \mathsf{M}$ *containing* p_0 *where* (Inv) *and* (Dim) *hold then there exists a possibly smaller open set* $\mathsf{V} \subseteq \mathsf{U}$ *containing* p_0 *and a sequence of smooth maps* $H_0, \ldots, H_{n-n^*}$ *so that:*

(1) each map H_k *satisfies the characteristic property* (4.10),

(2) for all $0 \leq k \leq n - n^* - 1$, $H_{k+1} = P_k \circ H_k$ *where* $P_k : \mathbb{R}^{\ell_{k+1}} \times \mathbb{R}^{\ell_k - \ell_{k+1}} \to \mathbb{R}^{\ell_{k+1}}$ *is a projection onto the leading* ℓ_{k+1} *components of* $\mathbb{R}^{\ell_k}$, *and*

(3) for all $0 \leq k \leq n - n^*$ *the leading* $1 + \sum_{i=k}^{n-n^*-1} \rho_i$ *components of* H_k *vanish on* L.

Corollary 4.2.9 encodes subprocedures (a) and (b) as presented at the start of this section. It assures us that there exists a set of generators which "drop off" on computing the derived flag while explicitly expressing the components with differentials that annihilate tangent vectors to L. Note, however, that we still do not know what the transverse output is.

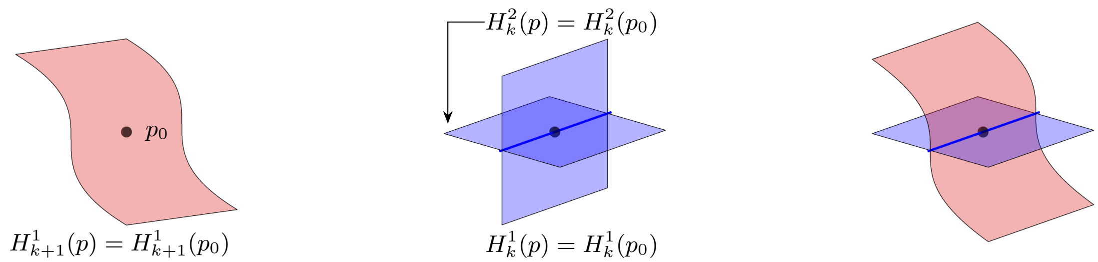


Figure 4.2: *A depiction of Proposition 4.2.8. The level sets of components of* H_k *subsume those of* H_{k+1} *while preserving the image of their differentials.*

4.3 The Proposed Algorithm

It was purported that the proof that the dual conditions are sufficient for transverse feedback linearization employs the proposed algorithm. This is precisely what we aim to show in this section.

The Geometry of the Algorithm

The proposed algorithm produces a flag of (locally) closed, embedded submanifolds,

$$\mathbb{R}^n \supseteq \pi(\mathsf{U}) =: \mathsf{Z}^{(n-n^*+1)} \supseteq \mathsf{Z}^{(n-n^*)} \supseteq \cdots \supseteq \mathsf{Z}^{(2)} \supseteq \mathsf{Z}^{(1)} = \pi(\mathsf{U}) \cap \mathsf{N}.$$

Each manifold $\mathsf{Z}^{(i)}$ is constructed as the local zero dynamics manifold containing x_0 of an incomplete transverse output with respect to N. The scalar outputs used to construct $\mathsf{Z}^{(i)}$ are all the components of the transverse output for N which have relative degree greater than, or equal to, i.

As a result, one can take the following perspective on the algorithm. Informally, the algorithm transverse feedback linearizes the system dynamics with respect to N by finding those scalar outputs that are transverse to N and have the *largest possible* relative degree κ_1. The algorithm proceeds by finding outputs that yield a lower relative degree, are transverse to N but *not* transverse to the zero dynamics manifold $\mathsf{Z}^{(\kappa_1)}$ induced by the already known outputs. The new outputs can then be combined with the known outputs to yield a relative degree with a smaller zero dynamics manifold. The process repeats until the zero dynamics manifold agrees with N locally.

To simplify the discussion involving the outputs and their associated zero dynamics manifolds, we define a special class of controlled-invariant set.

Definition 4.3.1 (Regular Zero Dynamics Manifold) . A closed, embedded submanifold $\mathsf{Z} \subseteq \mathbb{R}^n$ is said to be a **regular zero dynamics manifold of type $(\boldsymbol{\ell}, \boldsymbol{\kappa})$ (at $\boldsymbol{x_0}$)** for (4.1), where $1 \leq \ell \leq n$ and $\kappa \in \mathbb{N}^\ell$, if there exists an open set $\mathsf{U} \subseteq \mathbb{R}^n$ containing x_0 and a smooth function $h : \mathsf{U} \to \mathbb{R}^\ell$ so that

(1) the system (4.1) with output h yields a vector relative degree $\kappa = (\kappa_1, \ldots, \kappa_\ell)$ at x_0, and

(2) the zero dynamics manifold for h coincides with $\mathsf{U} \cap \mathsf{Z}$.

Although not explicit, Definition 4.3.1 implies that a regular zero dynamics manifold Z of type (ℓ, κ) has dimension $n - \sum_{i=1}^{\ell} \kappa_i$. It is also clear from Definition 4.3.1 that regular zero dynamics manifolds are controlled-invariant sets; the converse is clearly not true. Since all of the $\mathsf{Z}^{(i)}$ in the proposed algorithm's flag are constructed as the zero dynamics manifold associated to some output for (4.1), the proposed algorithm produces a flag of regular zero dynamics manifolds containing N. We can restate Theorem 1.3.1 in the language of Definition 4.3.1.

Theorem 4.3.2. *The local transverse feedback linearization problem is solvable at x_0 if, and only if, there exists constants $\rho_0 \in \mathbb{N}$ and $\kappa = (\kappa_1, \ldots, \kappa_{\rho_0}) \in \mathbb{N}^{\rho_0}$ so that N is a regular zero dynamics manifold of type (ρ_0, κ) at x_0.*

Theorem 4.3.2 implies that the proposed algorithm produces a descending flag of manifolds that are transverse feedback linearizable at x_0. This formulation gives geometric intuition behind the algorithm for transverse feedback linearization, but also to algorithms for feedback linearization methods more broadly. The GS algorithm proposed in [16] and Blended algorithm proposed in [35] both involve adapting generators starting with the smallest (last) ideal in the derived flag (2.13) and working backwards to the system ideal $\mathcal{I}^{(0)}$. This can be viewed as constructing a sequence of components for the feedback linearizing output starting with those components that yield the highest relative degree to those that yield the lowest relative degree. Implicitly, this produces a descending flag of regular zero dynamics manifolds that terminates at the one-point set $\{x_0\}$. Partial feedback linearization, which is concerned with finding the largest feedback linearizable subsystem, could be cast in the same light by instead asking that the last regular zero dynamics manifold in the flag be as small (in dimension) as possible.

The Proof

We are ready to prove the dual conditions for transverse feedback linearization are sufficient and, in turn, demonstrate the purported algorithm.

Proof of Theorem 4.1.1 (Sufficiency). Suppose (Con) holds at p_0 and that there exists an open set $\mathsf{U} \subseteq \mathsf{M}$ of p_0 on which conditions (Inv) and (Dim) hold. At the start of the algorithm we set $\mathsf{Z}^{(n-n^*+1)} := \pi(\mathsf{U})$ since there cannot be a scalar transverse output for N with relative degree greater than $n - n^*$.

By Corollary 4.2.9 there exists a sequence of smooth maps $H_0, \ldots, H_{n-n^*}$ defined on an open set $\mathsf{V} \subseteq \mathsf{U}$ containing p_0 that satisfy the characteristic property (4.10), i.e., for any fixed $0 \leq k \leq n - n^*$,

$$\langle \mathrm{d}H_k^1, \ldots, \mathrm{d}H_k^{\ell_k} \rangle = \langle \mathcal{J}^{(k)}, \mathrm{d}t \rangle^{(\infty)}.$$

Furthermore, the leading components of these maps vanish on L so we may write, for any $0 \leq k \leq n - n^*$,

$$H_k = (\underbrace{H_k^1, \ldots, t}_{\text{vanish on } \mathsf{L}}, \underbrace{\ldots, \ldots, \ldots, H_k^{\ell_k}}_{\text{not constant on } \mathsf{L}}).$$

Now recall the discussion surrounding (4.6). We already know that the t component of H_{n-n^*} is the only component that vanishes on L. However, since $\rho_{\kappa_1} = \cdots = \rho_{n-n^*-1} = 0$, we can also say the same thing about the map H_{κ_1}. That is,

$$H_{\kappa_1} = (t, \underbrace{H_{\kappa_1}^2, \ldots, \ldots, H_{\kappa_1}^{\ell_{\kappa_1}}}_{\text{not constant on } \mathsf{L}}), \tag{4.14}$$

As a result, we start our algorithm at index κ_1.

Set $\mathsf{Z}^{(\kappa_1+1)} \coloneqq \mathsf{Z}^{(n-n^*+1)} = \pi(\mathsf{U})$, and consider the map H_{κ_1-1}. By definition, $\rho_{\kappa_1-1} > 0$. Therefore, we can write[1]

$$H_{\kappa_1-1} = (\underbrace{H^1_{\kappa_1-1}, \dots, H^{\rho_{\kappa_1-1}}_{\kappa_1-1}, t}_{\text{vanish on } \mathsf{L}}, \underbrace{\dots, \dots, \dots, H^{\ell_{\kappa_1-1}}_{\kappa_1-1}}_{\text{not constant on } \mathsf{L}}).$$

Take the ρ_{κ_1-1} smooth functions $H^1_{\kappa_1-1}, \dots, H^{\rho_{\kappa_1-1}}_{\kappa_1-1}$, and define the candidate transverse output

$$h \coloneqq (H^1_{\kappa_1-1}, \dots, H^{\rho_{\kappa_1-1}}_{\kappa_1-1}).$$

Now we show that the system (4.2) with output h yields a vector relative degree of $(\kappa_1, \dots, \kappa_1)$ at p_0. Clearly $\mathrm{d}h^i \in \langle \mathcal{I}^{(\kappa_1-1)}, \mathrm{d}t\rangle^{(\infty)}$ by the characteristic property (4.10). Therefore

$$\langle \mathrm{d}h^1, \dots, \mathrm{d}h^{\rho_{\kappa_1-1}}\rangle \subseteq \langle \mathcal{I}^{(\kappa_1-1)}, \mathrm{d}t\rangle^{(\infty)}. \tag{4.15}$$

Observe that $\mathrm{d}h^i \in \mathrm{ann}(\mathsf{TL})$. As a result,

$$\mathrm{span}_{\mathbb{R}}\{\mathrm{d}h^1_{p_0}, \dots, \mathrm{d}h^{\rho_{\kappa_1-1}}_{p_0}\} \subseteq \mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}).$$

Using (4.14), the characteristic property (4.10), and (Con),

$$\mathrm{span}_{\mathbb{R}}\{\mathrm{d}h^1_{p_0}, \dots, \mathrm{d}h^{\rho_{\kappa_1-1}}_{p_0}\} \cap \langle \mathcal{I}^{(\kappa_1)}, \mathrm{d}t\rangle^{(\infty)}_{p_0} \subseteq \mathrm{span}_{\mathbb{R}}\{\mathrm{d}t_{p_0}\}.$$

We already know that the h^i are smooth functions of the state, so we may conclude

$$\mathrm{span}_{\mathbb{R}}\{\mathrm{d}h^1_{p_0}, \dots, \mathrm{d}h^{\rho_{\kappa_1-1}}_{p_0}\} \cap \langle \mathcal{I}^{(\kappa_1)}, \mathrm{d}t\rangle^{(\infty)}_{p_0} = \{0\}.$$

Finally use Proposition 4.2.2 to deduce

$$\mathrm{span}_{\mathbb{R}}\{\mathrm{d}h^1_{p_0}, \dots, \mathrm{d}h^{\rho_{\kappa_1-1}}_{p_0}\} \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(\kappa_1)}_{p_0}, \mathrm{d}t_{p_0}\} = \{0\}. \tag{4.16}$$

The expressions (4.15) and (4.16) are the conditions for Lemma 2.2.14. Since Assumptions 2.2.3 and 2.2.11 hold, conclude that h yields a uniform vector relative degree of $(\kappa_1,$

[1] The components $H^1_{\kappa_1-1}$ up to $H^{\rho_{\kappa_1-1}}_{\kappa_1-1}$ are constant on L and can be treated, without loss of generality, as zero on L. Therefore, we write that these components vanish on L.

$\ldots,\ \kappa_1)$ at p_0. Define the local regular zero dynamics manifold, possibly shrinking U if necessary,

$$\mathsf{Z}^{(\kappa_1)} := \{x \in \pi(\mathsf{U})\colon h(\iota(x)) = \cdots = \mathcal{L}_f^{\kappa_1-1}h(\iota(x)) = 0\},$$

of type $(\rho_{\kappa_1-1}, (\kappa_1, \ldots, \kappa_{\rho_{\kappa_1-1}}))$ at $x_0 = \pi(p_0)$. Fix $0 \le k < \kappa_1 - 1$. Because of (4.15), Lemma 2.2.9 implies that

$$\mathrm{d}h^i, \ldots, \mathcal{L}_f^{\kappa_1-k-1}\mathrm{d}h^i \in \langle \mathcal{I}^{(k)}, \mathrm{d}t\rangle^{(\infty)} = \langle \mathrm{d}H_k^1, \ldots, \mathrm{d}H_k^{\ell_k}\rangle, \qquad 1 \le i \le \rho_{\kappa_1-1}.$$

As a result, we can, without loss of generality, rewrite H_k to take the form

$$H_k = (\underbrace{\overbrace{\ldots,\ldots,\ \ldots,\ldots,}^{\text{other components}}\overbrace{h,\ldots,\mathcal{L}_f^{\kappa_1-k-1}h}^{\text{vanish on } \mathsf{Z}^{(\kappa_1)}},\ t,}_{\text{vanish on } \mathsf{L}}\underbrace{\ldots,\ \ldots,\ \ldots,\ \ldots}_{\text{not constant on } \mathsf{L}}),$$

This process adapts all maps H_k for $k < \kappa_1 - 1$ so that they explicitly include h and its Lie derivatives along f in their components. Perform this operation for each k while ensuring the components of H_k are subsumed by the components of H_{k-1}, i.e., $H_k = P_{k-1} \circ H_{k-1}$, up to a reordering. We say that the sequence $H_0, \ldots, H_{\kappa_1}$ is adapted to L subordinate to the regular zero dynamics manifold $\mathsf{Z}^{(\kappa_1)}$. Observe that $\mathsf{Z}^{(\kappa_1)} \supseteq \mathsf{N}$ since $h|_{\mathsf{N}} = 0$ by construction. It is trivially the case that $\mathsf{Z}^{(\kappa_1)} \subset \mathsf{Z}^{(\kappa_1+1)} = \pi(\mathsf{U})$. The final fact that we simply state is that the only differentials that vanish on L at index $\kappa_1 - 1$ are those that are linearly dependent on the differentials of h. That is,

$$\mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathrm{d}H^1_{\kappa_1-1} \ldots, \mathrm{d}H^{\ell_{\kappa_1-1}}_{\kappa_1-1}\} \subseteq \mathrm{ann}(\mathsf{T}_{p_0}\mathsf{L}^{(\kappa_1)}),$$

where $\mathsf{L}^{(\kappa_1)}$ is the lift of $\mathsf{Z}^{(\kappa_1)}$. This completes the base case. Suppose, by way of induction, that, for some $2 \le k \le \kappa_1$,

(H.1) $\mathsf{Z}^{(k)}$ is a regular zero dynamics manifold of type $(\rho_{k-1}, (\kappa_1, \ldots, \kappa_{\rho_{k-1}}))$ at x_0 satisfying

$$\mathsf{Z}^{(k+1)} \supseteq \mathsf{Z}^{(k)} \supseteq \mathsf{N},$$

(H.2) there exists a smooth function $h : \mathsf{U} \to \mathbb{R}^{\rho_{k-1}}$ so that system (4.1) with output h yields a vector relative degree $(\kappa_1,\ \ldots,\ \kappa_{\rho_{k-1}})$ at x_0 and the zero dynamics coincide locally with $\mathsf{Z}^{(k)}$, (H.3) all maps $H_0, \dots, H_{n-n^*}$ are adapted to L subordinate to the regular zero dynamics manifold $\mathsf{Z}^{(k)}$ using output h and,

(H.4) denoting $\mathsf{L}^{(k)}$ as the lift of $\mathsf{Z}^{(k)}$ we have that[2]

$$\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}} \left\{ \mathrm{d}H^1_{k-1}, \dots, \mathrm{d}H^{\ell_{k-1}}_{k-1} \right\} \subseteq \operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}^{(k)}), \tag{4.17}$$

The goal of this induction is to construct new regular zero dynamics manifold $\mathsf{Z}^{(k-1)}$ that satisfies

(C.1) $\mathsf{Z}^{(k-1)}$ is a regular zero dynamics manifold of type $(\rho_{k-2}, (\kappa_1, \dots, \kappa_{\rho_{k-2}}))$ at x_0 satisfying

$$\mathsf{Z}^{(k)} \supseteq \mathsf{Z}^{(k-1)} \supseteq \mathsf{N},$$

(C.2) there exists a smooth function $h' : \mathsf{U} \to \mathbb{R}^{\rho_{k-2}}$ so that system (4.1) with output h' yields a vector relative degree $(\kappa_1, \dots, \kappa_{\rho_{k-2}})$ at x_0 and the zero dynamics coincide locally with $\mathsf{Z}^{(k-1)}$,

(C.3) all maps $H_0, \dots, H_{n-n^*}$ are adapted to L subordinate to the regular zero dynamics manifold $\mathsf{Z}^{(k-1)}$ using output h' and,

(C.4) denoting $\mathsf{L}^{(k-1)}$ as the lift of $\mathsf{Z}^{(k-1)}$ we have that

$$\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}} \left\{ \mathrm{d}H^1_{k-2}, \dots, \mathrm{d}H^{\ell_{k-2}}_{k-2} \right\} \subseteq \operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}^{(k-1)}), \tag{4.18}$$

Consider the map H_{k-2} whose component differentials generate the differential ideal $\langle \mathcal{I}^{(k-2)}, \mathrm{d}t \rangle^{(\infty)}$. By (H.3) of the inductive hypothesis, H_{k-2} takes the form

$$H_{k-2} = (\underbrace{\overbrace{H^1_{k-2}, \dots, \dots, H^{\rho_{k-2}-\rho_{k-1}}_{k-2}}^{\text{other components}}, \overbrace{h^1, \dots, \mathcal{L}_f^{\kappa_{\rho_{k-1}}-k+1} h^{\rho_{k-1}}}^{\text{vanish on } \mathsf{Z}^{(k)}}, t}_{\text{vanish on } \mathsf{L}}, \underbrace{\dots, \dots, \dots}_{\text{not constant on } \mathsf{L}}).$$

There are now two cases. If $\rho_{k-2} = \rho_{k-1}$, then the number of "other components" that vanish on L is zero. This is because, due to the vector relative degree of h, the ρ_{k-1} new

[2] This condition ensures that $\mathsf{Z}^{(k)} \supseteq \mathsf{N}$ is the *smallest* possible regular zero dynamics manifold of type (ρ, κ) where $\kappa \geq k$.

components $\mathcal{L}_f^{\kappa_1-k+1}h^1,\ldots,\mathcal{L}_f^{\kappa_{\rho_{k-1}}-k+1}h^{\rho_{k-1}}$, appear in H_{k-2}. In this case, set $\mathsf{Z}^{(k-1)}=\mathsf{Z}^{(k)}$ which remains a regular zero dynamics manifold of type $(\rho_{k-2},(\kappa_1,\ldots,\kappa_{\rho_{k-2}}))$ establishing (C.1). The rest of the inductive properties (C.2)–(C.4) follow directly from (H.2)–(H.4) by leaving the output $h' \coloneqq h$ unchanged.

Alternatively, $\rho_{k-2}>\rho_{k-1}$. In this case, there exists precisely $\mu \coloneqq \rho_{k-2}-\rho_{k-1}$ "new" components whose differentials annihilate tangent vectors to L: these are the first components that are differentially indepedent of the Lie derivatives of h yet vanish on L. Take these new component functions, up to a reordering, to be the leading components $H^1_{k-2},\ldots,H^{\mu}_{k-2}$, and define the output

$$q \coloneqq (H^1_{k-2},\ldots,H^{\mu}_{k-2}).$$

We now show that the system (4.2) with candidate output $h' \coloneqq (h,q)$ yields a well-defined vector relative degree of $(\kappa_1,\ldots,\kappa_{\rho_{k-2}})$ at p_0. Since the q^i are component functions of H_{k-2} we have by the characteristic property (4.10) that

$$\langle \mathrm{d}q^1,\ldots,\mathrm{d}q^{\mu}\rangle \subseteq \langle \mathcal{I}^{(k-2)},\mathrm{d}t\rangle^{(\infty)}.$$

We also know from (H.2) of the inductive hypothesis that h^i yields a relative degree of κ_i so the jth Lie derivative of h along f yields a relative degree as well of κ_i-j, for $0\le j\le \kappa_i-1$. Invoke Lemma 2.2.14, subject to Assumptions 2.2.3 and 2.2.11, to find

$$\langle \mathcal{L}_f^{\kappa_1-k+1}\mathrm{d}h^1,\ldots,\mathcal{L}_f^{\kappa_{\rho_{k-1}}-k+1}\mathrm{d}h^{\rho_{k-1}}\rangle \subseteq \langle \mathcal{I}^{(k-2)},\mathrm{d}t\rangle^{(\infty)}.$$

Combine these data to find

$$\langle \mathrm{d}q^1,\ldots,\mathrm{d}q^{\mu},\mathcal{L}_f^{\kappa_1-k+1}\mathrm{d}h^1,\ldots,\mathcal{L}_f^{\kappa_{\rho_{k-1}}-k+1}\mathrm{d}h^{\rho_{k-1}}\rangle \subseteq \langle \mathcal{I}^{(k-2)},\mathrm{d}t\rangle^{(\infty)}. \tag{4.19}$$

Putting that aside, invoke Lemma 2.2.14 once again to find

$$\mathrm{span}_{\mathbb{R}}\left\{\mathcal{L}_f^{\kappa_1-k+1}\mathrm{d}h^1_{p_0},\ldots,\mathcal{L}_f^{\kappa_{\rho_{k-1}}-k+1}\mathrm{d}h^{\rho_{k-1}}_{p_0}\right\}\cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}^{(k-1)}_{p_0},\mathrm{d}t_{p_0}\}=\{0\}.$$

By (H.4) of the inductive hypothesis,

$$\mathrm{span}_{\mathbb{R}}\{\mathrm{d}q^1_{p_0}\ldots,\mathrm{d}q^{\mu}_{p_0}\}\cap\langle \mathrm{d}H^1_{k-1},\ldots,\mathrm{d}H^{\ell_{k-1}}_{k-1}\rangle=\{0\},$$

since the $\mathrm{d}q^i \in \operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L})$ but $\mathrm{d}q^i \notin \operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}^{(k)})$. Combine these data to conjecture that

$$\operatorname{span}_{\mathbb{R}}\left\{\mathrm{d}q^1_{p_0}, \dots, \mathrm{d}q^{\mu}_{p_0}, \mathcal{L}_f^{\kappa_1-k+1}\mathrm{d}h^1_{p_0}, \dots, \mathcal{L}_f^{\kappa_{\rho_{k-1}}-k+1}\mathrm{d}h^{\rho_{k-1}}_{p_0}\right\}$$

has a trivial intersection with $\operatorname{span}\{\mathscr{I}^{(k-1)}_{p_0}, \mathrm{d}t_{p_0}\}$. Suppose, in search of a contradiction, there is a linear combination

$$\sum_{i=1}^{\mu} a_i\,\mathrm{d}q^i_{p_0} + \sum_{i=1}^{\rho_{k-1}} b_i\,\mathcal{L}_f^{\kappa_i-k+1}\mathrm{d}h^i_{p_0} \in \operatorname{span}_{\mathbb{R}}\left\{\mathscr{I}^{(k-1)}_{p_0}, \mathrm{d}t_{p_0}\right\}.$$

If there is an $a_i \neq 0$, then this form does not live in $\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}^{(k)})$ which contradicts (H.4) of the inductive hypothesis. Therefore $a_i = 0$ for all $1 \le i \le \mu$. We now show that $b_i = 0$ for all $1 \le i \le \rho_{k-1}$. Suppose

$$\sum_{i=1}^{\rho_{k-1}} b_i\,\mathcal{L}_f^{\kappa_i-k+1}\mathrm{d}h^i_{p_0} \in \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(k-1)}_{p_0}, \mathrm{d}t_{p_0}\},$$

Then, by Lemma 2.2.14, the system (4.2) with output

$$(\mathcal{L}_f^{\kappa_1-k+1}h^1, \dots, \mathcal{L}_f^{\kappa_{\rho_{k-1}}-k+1}h^{\rho_{k-1}})$$

does *not* yield a uniform vector relative degree at p_0. This immediately contradicts the well-defined vector relative degree for h. Therefore $b_i = 0$ for all $1 \le i \le \rho_{k-1}$. As a result, conclude that system (4.2) with output $h' = (h, q)$ yields a vector relative degree at p_0. The vector relative degree must be $(\kappa_1, \dots, \kappa_{\rho_{k-2}})$. This demonstrates (C.2). Define the local regular zero dynamics manifold

$$\mathsf{Z}^{(k-1)} := \{x \in \mathsf{Z}^{(k)} \colon q(\iota(x)) = \cdots = \mathcal{L}_f^{\kappa_{\rho_{k-2}}-1}q(\iota(x)) = 0\},$$

of type $(\rho_{k-2}, (\kappa_1, \dots, \kappa_{\rho_{k-2}}))$. By construction $\mathsf{Z}^{(k-1)} \subset \mathsf{Z}^{(k)}$ and, since $q|_{\mathsf{N}} = 0$, $\mathsf{Z}^{(k-1)} \supseteq \mathsf{N}$. This establishes (C.1). To establish (C.3), we adapt, exactly as in the base case, the maps $H_0, \dots, H_{n-n^*}$ to L subordinate to $\mathsf{Z}^{(k-1)}$ so that all the Lie derivatives of $h' = (h, q)$ appear explicitly. It remains to show (C.4). First observe that the components of

$$H_{k-2} = (\underbrace{\overbrace{q^1, \dots, \dots, q^{\mu}}^{\text{other components}}, \overbrace{h^1, \dots, \mathcal{L}_f^{\kappa_{\rho_{k-1}}-k+1}h^{\rho_{k-1}}}^{\text{vanish on } \mathsf{Z}^{(k)}}, t}_{\text{vanish on } \mathsf{L}, \mathsf{Z}^{(k-1)}}, \underbrace{\dots, \dots, \dots}_{\text{not constant on } \mathsf{L}}),$$

that vanish on L constitute a component of h, a Lie derivative of h, or q. It follows that, using the characteristic property 4.10,

$$\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \langle \mathcal{I}^{(k-2)}, \mathrm{d}t \rangle^{(\infty)} \subseteq \operatorname{ann}(\mathsf{T}_p\mathsf{L}^{(k-1)}).$$

Use (Inv) with Proposition 4.2.2 to conclude that (C.4) holds. This completes the induction.

The inductive algorithm proceeds until the regular zero dynamics manifold $\mathsf{Z}^{(1)}$ of type $(\rho_0, (\kappa_1, \ldots, \kappa_{\rho_0}))$ is produced at step $k = 2$. By Lemma 4.2.1, $\mathsf{Z}^{(1)}$ is an n^*-dimensional submanifold with codimension $n - n^*$. It contains the n^*-dimensional submanifold N and so $\mathsf{Z}^{(1)} = \mathsf{N}$ is a regular zero dynamics manifold of type $(\rho_0, (\kappa_1, \ldots, \kappa_{\rho_0}))$. Theorem 4.3.2 implies that N is transverse feedback linearizable at x_0. A by-product of this algorithm is that the final output h is the transverse output. □

Simplifying the Algorithm

The algorithm used in the proof inspires a shortened, but equivalent, algorithm that produces a transverse output under the conditions for TFL — (Con), (Inv), (Dim). The procedure is presented in Algorithm 1. One difference from the proof is that integration only happens at iterations where ρ_{k-1} differs from ρ_k. These are indices corresponding to the *distinct* transverse controllability indices. Another difference is the lack of re-adaptation of all the maps $H_0, \ldots, H_{n-n^*}$ throughout the algorithm. In fact, the vast majority of the maps H_k are not constructed. This is *not* an oversight. The adaptation process is embedded in Line 13 where H_{k-1} is adapted to have the known output h and its Lie derivatives appear explicitly. This alongside the fact that the algorithm runs from larger to smaller indices ensure an appropriately adapted basis is constructed.

4.4 Illustrative Examples

In this section we explore a number of examples that illustrate the evaluation of the conditions for transverse feedback linearization, and the algorithm proposed in Section (4.3).

Algorithm 1 The Transverse Feedback Linearization algorithm.

procedure TFL PROCEDURE($f, g, \mathsf{N}, x_0, u_*$)
    Compute $\rho_0(p_0), \ldots, \rho_{n-n^*}(p_0)$ ▷ as in (4.4)
    Compute $\kappa_1(p_0), \ldots, \kappa_m(p_0)$ ▷ as in (4.5)
    $\mathsf{Z}^{(\kappa_1+1)} \leftarrow \mathbb{R}^n$
    Initialize $h \leftarrow (\quad)$
    for $k \leftarrow \kappa_1, \ldots, 1$ **do**
        if $\rho_{k-1} = \rho_k$ **then**
            $\mathsf{Z}^{(k)} \leftarrow \mathsf{Z}^{(k+1)}$
            continue
        end if
        $\mu \leftarrow \rho_{k-1} - \rho_k$
        Construct H_{k-1} to satisfy (4.10) ▷ integrate $\langle \mathcal{J}^{(k-1)}, \mathrm{d}t \rangle^{(\infty)}$
        Rewrite H_{k-1}, while preserving (4.10), so that

$$H_{k-1} = (\underbrace{H^1_{k-1}, \ldots, H^\mu_{k-1}, h^1, \ldots, \ldots, \mathcal{L}_f^{\kappa_{\rho_k}-k} h^{\rho_k}, t}_{\text{vanish on } \mathsf{L}}, \ldots).$$

        $h \leftarrow (h, H^1_{k-1}, \ldots, H^\mu_{k-1})$
        $\mathsf{Z}^{(k)} \leftarrow$ zero dynamics of h.
    end for
end procedure

Academic Example

Consider the nonlinear, control-affine system

$$\dot{x}(t) = \begin{pmatrix} -x^2(t) \\ x^1(t) \\ x^3(t)\, x^4(t) \\ 0 \\ x^6(t) \\ -x^3(t)\, x^5(t) + x^6(t) + x^7(t) \\ x^5(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ x^3(t) \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} u^1(t) + \begin{pmatrix} -x^2(t) \\ 0 \\ 0 \\ 0 \\ -x^1(t) \\ x^1(t) \\ x^1(t) \end{pmatrix} u^2(t) \tag{4.20}$$

and the closed, embedded 2-dimensional submanifold

$$\mathsf{N} := \left\{x \in \mathbb{R}^7 \colon (x^1)^2 + (x^2)^2 - x^3 = x^4 = x^5 = x^6 = x^7 = 0\right\}, \tag{4.21}$$

rendered controlled-invariant by $u_*(x) = 0$. Fix a point $x_0 = (2,\ 0,\ 4,\ 0,\ 0,\ 0,\ 0) \in \mathsf{N}$. In light of Theorem 1.3.1, if (4.20) is transverse feedback linearizable with respect to N at x_0, then there exists either a single scalar function yielding a relative degree 5 at x_0 that vanishes on N or two scalar functions yielding a vector relative degree $(\kappa_1, 5-\kappa_1)$ at x_0 that simultaneously vanish on N. Natural candidates can be picked out of the functions that define N since they satisfy (1) of Theorem 1.3.1. Unfortunately, all of the scalar functions used to define N in (4.21) either yield a relative degree of 1 at x_0 or do not yield a relative degree at all. As a result, we cannot directly use them to form an output that satisfies (2) of Theorem 1.3.1.

However, as we now show using our dual TFL conditions, the transverse feedback linearization problem is solvable at $x_0 \in \mathsf{N}$. The ideals $\mathcal{I}^{(i)}$ in (2.13) for system (4.20) are

$$\begin{aligned}
\mathcal{I}^{(0)} &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}x^4, \mathrm{d}x^5, \mathrm{d}x^6, \mathrm{d}x^7 \rangle \\
\mathcal{I}^{(1)} &= \langle x^1\,\mathrm{d}x^1 + x^2\,\mathrm{d}x^7, \mathrm{d}x^2 - \mathrm{d}x^7, \mathrm{d}x^3 - x^3\,\mathrm{d}x^4, \mathrm{d}x^5 + \mathrm{d}x^7, \mathrm{d}x^6 - \mathrm{d}x^7 \rangle \\
\mathcal{I}^{(2)} &= \langle \beta^1, \beta^2 \rangle, \\
\mathcal{I}^{(3)} &= \langle 0 \rangle,
\end{aligned}$$

where β^1, $\beta^2 \in \Gamma^\infty(\mathsf{T}^*\mathsf{M})$ are smooth one-forms whose expressions we omit for clarity. The derived flag (2.13) terminates at $\mathcal{I}^{(3)}$. Immediately we see that the controllability condition (Con) holds since

$$\langle \mathcal{I}^{(n-n^*)}, \mathrm{d}t \rangle = \langle \mathcal{I}^{(5)}, \mathrm{d}t \rangle = \langle \mathcal{I}^{(3)}, \mathrm{d}t \rangle = \langle \mathrm{d}t \rangle.$$

Next we check the constant dimensionality condition (Dim). Observe that, for any $p \in \mathsf{L}$

in a sufficiently small open set containing p_0,

$$\begin{aligned}
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(0)}_p, \mathrm{d}t_p\} &= \operatorname{ann}(\mathsf{T}_p\mathsf{L}),\\
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(1)}_p, \mathrm{d}t_p\} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}x^6 - \mathrm{d}x^7, \mathrm{d}x^5 + \mathrm{d}x^7,\\
&\qquad -2x^1\,\mathrm{d}x^1 - 2x^2\,\mathrm{d}x^2 + \mathrm{d}x^3 - x^3\,\mathrm{d}x^4 - 2x^2\,\mathrm{d}x^7\mathrm{d}t\}. \qquad (4.22)\\
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(2)}_p, \mathrm{d}t_p\} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}x^5 + \mathrm{d}x^7, \mathrm{d}t\}.\\
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(3)}_p, \mathrm{d}t_p\} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}t\}.
\end{aligned}$$

Thus the constant dimensionality condition (Dim) holds. It remains to check the involutivity condition (Inv) holds. Note that the ideals $\langle \mathcal{I}^{(0)}, \mathrm{d}t\rangle$, $\langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle$ and $\langle \mathcal{I}^{(3)}, \mathrm{d}t\rangle$ are all differential ideals. Therefore, it suffices to check that

$$\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(2)}_p, \mathrm{d}t_p\} \subseteq \langle \mathcal{I}^{(2)}, \mathrm{d}t\rangle^{(\infty)}_p.$$

Using `Maple`, we directly compute the derived flag for $\langle \mathcal{I}^{(2)}, \mathrm{d}t\rangle$ and verify it converges to

$$\langle \mathcal{I}^{(2)}, \mathrm{d}t\rangle^{(\infty)} = \langle \mathrm{d}x^5 + \mathrm{d}x^7, \mathrm{d}t\rangle.$$

Comparing this with the expression for $\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}\{\mathscr{I}^{(2)}_p, \mathrm{d}t_p\}$ above, we see that the involutivity condition (Inv) holds. As a result, the transverse feedback linearization problem is solvable for (4.20) with respect to N at x_0, and we can execute Algorithm 1.

We start by computing the indices ρ and κ as required by Line 2. Using (4.22), deduce

$$\rho_0(p_0) = 2, \quad \rho_1(p_0) = 2, \quad \rho_2(p_0) = 1, \quad \rho_3(p_0) = 0,$$

and,

$$\kappa_1(p_0) = 3, \quad \kappa_2(p_0) = 2.$$

Let

$$\mathsf{U} := \left\{(t, u, x) \in \mathsf{M} \colon x^1 > 0, x^3 > -3\right\}$$

and set $\mathsf{Z}^{(\kappa_1+1)} = \mathsf{Z}^{(4)} := \pi(\mathsf{U})$. The ideals are simply, finitely, non-degenerately generated on U and it can be verified that the conditions (Dim) and (Inv) hold over U.

The iteration begins at $k = \kappa_1 = 3$. Observe that $\rho_2 > \rho_3$. It follows that we expect to find $\mu = \rho_2 - \rho_3 = 1$ new scalar output that will yield a relative degree of 3 at p_0 and is constant on L. We proceed by integrating the differential ideal $\langle \mathcal{I}^{(2)}, \mathrm{d}t\rangle^{(\infty)}$ to find the map $H_2 = (x^5 + x^7, t)$. Right away we see that the new component that is constant (i.e. vanishes) on L is the first component and so we define our candidate partial transverse output $h \coloneqq x^5 + x^7$. We then define the local regular zero dynamics manifold $\mathsf{Z}^{(3)}$ to be the zero dynamics of system (4.20) with output h. Explicitly,

$$\mathsf{Z}^{(3)} \coloneqq \left\{ x \in \mathbb{R}^n \colon x^5 + x^7 = x^5 + x^6 = -x^3\, x^5 + 2\, x^6 + x^7 = 0 \right\}.$$

The next iteration of the algorithm looks at index $k = 2$. Here, we again note that $\rho_1 > \rho_2$ and $\mu = \rho_1 - \rho_2 = 1$. We expect to see one new scalar output with relative degree 2 at p_0 that is constant on L. Integrate $\langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle^{(\infty)}$ to find

$$H_1 = (H_1^1, \dots, H_1^5) = (\overbrace{x^5 + x^7, x^5 + x^6}^{h, \mathcal{L}_f h}, \frac{1}{2}(x^1)^2 + x^2\, x^7 - 2, x^2, x^3\, e^{-x^4} - 4, t),$$

where we highlight the fact that, as expected, the known output h and its Lie derivative $\mathcal{L}_f h$ appear explicitly in the first two components. Unfortunately, it is not obvious (at first glance) what the new scalar output is since none of the other components are constant on L besides the trivial t component. However, Lemma 4.2.6 states that a rewriting for H_1 where four components vanish on L is possible. The simplest strategy is to restrict H_1 to L and eliminate the effect of the coordinates algebraically. We perform this to find that the algebraic combination of components of H_1,

$$2\, H_1^3 + (H_1^4)^2 - H_1^5,$$

vanish on L. Using this, rewrite H_1 as

$$H_1 = (\underbrace{(x^1)^2 + (x^2)^2 + 2\, x^2\, x^7 - x^3\, e^{-x^4}, \overbrace{x^5 + x^7, x^5 + x^6}^{h, \mathcal{L}_f h}, t}_{\text{vanish on } \mathsf{L}}, \underbrace{x^2, x^3\, e^{-x^4} - 4}_{\text{not constant on } \mathsf{L}}).$$

Define the new candidate output $h \coloneqq (x^5 + x^7, H_1^1)$ and the induced local, regular zero dynamics manifold, shrinking U as necessary,

$$\mathsf{Z}^{(2)} \coloneqq \left\{ x \in \mathsf{Z}^{(1)} \colon H_1^1(x) = \mathcal{L}_f H_1^1(x) = 0 \right\}.$$

The final iteration of the algorithm at index $k = 1$ is skipped since $\rho_0 = \rho_1$. Set $\mathsf{Z}^{(1)} := \mathsf{Z}^{(2)}$. The algorithm asserts that $\mathsf{Z}^{(1)} = \mathsf{N}$ locally. The transverse output is

$$h(x) = \begin{bmatrix} x^5 + x^7 \\ (x^1)^2 + (x^2)^2 + 2\,x^2\,x^7 - x^3\,e^{-x^4} \end{bmatrix},$$

and it is a regular matter to verify that the system (4.20) with output h yields a vector relative degree of $(3, 2)$ at x_0 while locally vanishing on N.

Path Following for a Car

We now recall Example 1.3 where the objective was to design a control law that brings the vehicle's front axle towards a circle of radius $R > 0$ with a forward speed of 1. The nonlinear control-affine system is,

$$\dot{x}(t) = \begin{pmatrix} x^3(t)\,\cos(x^4(t)) \\ x^3(t)\,\sin(x^4(t)) \\ x^6(t) \\ \frac{x^3(t)}{\ell}\,\tan(x^5(t)) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} u^1(t) + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} u^2(t), \tag{4.23}$$

and we wish to make the controlled-invariant set,

$$\mathsf{N} := \left\{ x \in \mathbb{R}^6 \colon y_1(x) = y_2(x) = y_3(x) = y_4(x) = y_5(x) = 0 \right\},$$

locally attractive where,

$$\begin{aligned} y_1(x) &:= (x^1 + \ell\cos(x^4))^2 + (x^2 + \ell\sin(x^4))^2 - R^2, \\ y_2(x) &:= x^3 - 1, \\ y_3(x) &:= x^1\,\cos(x^4) + x^2\,\sin(x^4), \\ y_4(x) &:= \ell\cos(x^5) - \sin(x^5)\left(x^1\sin(x^4) - x^2\cos(x^4)\right), \\ y_5(x) &:= x^6. \end{aligned}$$

In Example 1.3, we noted that the point

$$x_0 = \left(-\ell \cos \arcsin \frac{\ell}{R}, R - \ell \sin \arcsin \frac{\ell}{R}, 1, \arcsin \frac{\ell}{R}, -\arctan \frac{\ell}{\sqrt{R^2 - \ell^2}}, 0\right),$$

is in N and N is rendered controlled-invariant by state feedback $u_*(x) = (0, 0)$. We aim to transverse feedback linearize the dynamics (4.23) at x_0 with N. First, we must check the conditions for transverse feedback linearizability.

In a sufficiently small open set of $p_0 = (t, u_*(x_0), x_0)$, the annihilator of the tangent space of L is,

$$\operatorname{ann}(\mathsf{T}_p\mathsf{L}) = \operatorname{span}_{\mathbb{R}}\{\mathrm{d}x^1 + x^2\,\mathrm{d}x^4, \mathrm{d}x^2 - x^1\,\mathrm{d}x^4, \mathrm{d}x^3, \mathrm{d}x^5, \mathrm{d}x^6, \mathrm{d}t\}. \tag{4.24}$$

We computed the derived flag (2.13) associated to the lift of the dynamics (4.23) in Example 2.1, so it remains to compute the augmented ideals $\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle$ and their differential closures. They are

$$\begin{aligned} \langle \mathcal{I}^{(0)}, \mathrm{d}t\rangle &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}x^4, \mathrm{d}x^5, \mathrm{d}t\rangle, \\ \langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}x^4, \mathrm{d}t\rangle, \\ \langle \mathcal{I}^{(2)}, \mathrm{d}t\rangle &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}t\rangle, \\ \langle \mathcal{I}^{(3)}, \mathrm{d}t\rangle &= \langle \mathrm{d}t\rangle, \end{aligned} \tag{4.25}$$

and

$$\begin{aligned} \langle \mathcal{I}^{(0)}, \mathrm{d}t\rangle^{(\infty)} &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}x^4, \mathrm{d}x^5, \mathrm{d}t\rangle, \\ \langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle^{(\infty)} &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}x^4, \mathrm{d}t\rangle, \\ \langle \mathcal{I}^{(2)}, \mathrm{d}t\rangle^{(\infty)} &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}t\rangle, \\ \langle \mathcal{I}^{(3)}, \mathrm{d}t\rangle^{(\infty)} &= \langle \mathrm{d}t\rangle. \end{aligned} \tag{4.26}$$

Equipped with these objects, we are now ready to verify the conditions for transverse feedback linearizability at p_0. The controllability condition (Con) asks that, at p_0,

$$\operatorname{ann}(\mathsf{T}_{p_0}\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(4)}_{p_0}, \mathrm{d}t_{p_0}\} = \operatorname{span}_{\mathbb{R}}\{\mathrm{d}t_{p_0}\}.$$

Because $\operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(4)}_{p_0}, \mathrm{d}t_{p_0}\} \subseteq \operatorname{span}_{\mathbb{R}}\{\mathscr{I}^{(3)}_{p_0}, \mathrm{d}t_{p_0}\} = \operatorname{span}_{\mathbb{R}}\{\mathrm{d}t_{p_0}\}$, the controllability condition holds.

Next we compute the intersection of (4.24) with the augmented ideals and use this to compute the transverse controllability indices (4.5). Observe that, combining (4.24) and (4.25), for all $p \in \mathsf{L}$ in a sufficiently small open set containing p_0,

$$\begin{aligned}
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(0)}, \mathrm{d}t_p\} &= \operatorname{ann}(\mathsf{T}_p\mathsf{L}), \\
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(1)}, \mathrm{d}t_p\} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}x_p^1 + x^2\,\mathrm{d}x_p^4, \mathrm{d}x_p^2 - x^1\,\mathrm{d}x_p^4, \mathrm{d}x_p^3, \mathrm{d}t_p\}, \\
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(2)}, \mathrm{d}t_p\} &= \operatorname{span}_{\mathbb{R}}\{x^1\,\mathrm{d}x_p^1 + x^2\,\mathrm{d}x_p^2, \mathrm{d}t_p\}, \\
\operatorname{ann}(\mathsf{T}_p\mathsf{L}) \cap \operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(3)}, \mathrm{d}t_p\} &= \operatorname{span}_{\mathbb{R}}\{\mathrm{d}t_p\},
\end{aligned} \tag{4.27}$$

As a result, the constant dimension condition (Dim) holds, and we have that

$$\rho_0(p_0) = 2, \rho_1(p_0) = 2, \rho_2(p_0) = 1, \rho_3(p_0) = 0,$$

as well as,

$$\kappa_1(p_0) = 2, \kappa_2(p_0) = 1.$$

Comparing (4.26) with (4.27), we also see that the involutivity condition (Inv) holds. We therefore may proceed in finding the transverse output. The indices indicate that we must find two scalar outputs that with system (4.23) yield a relative degree of three and two at p_0.

We have already presented the algorithm in detail as it applies to an academic example. For this example, we now show a number of shortcuts implied by the intuition of the algorithm that are useful in solving real problems by hand.

The iteration of the algorithm begins at $k = \kappa_1 = 3$ where we have $\rho_2 > \rho_3$. It follows that we expect to find $\mu = \rho_2 - \rho_3 = 1$ new scalar output that yields a relative degree of two at p_0 and is constant on L. Importantly, the differential of this output must live in the differential ideal

$$\langle \mathcal{I}^{(2)}, \mathrm{d}t \rangle^{(\infty)} = \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}t \rangle,$$

which is generated by exact one-forms. Since the output cannot be a function of time, we expect that the scalar output we are searching for must take the form $h_1(x) = \alpha(x^1, x^2)$.

Looking at the constraints y_i that define N, we observe that,

$$\begin{aligned} y_1(x) &= (x^1 + \ell\,\cos(x^4))^2 + (x^2 + \ell\,\sin(x^4))^2 - R^2, \\ &= (x^1)^2 + (x^2)^2 + 2\,\ell\,(x^1\,\cos(x^4) + x^2\,\sin(x^4)) - R^2 + \ell^2, \\ &= (x^1)^2 + (x^2)^2 + 2\,\ell\,y_2(x) - R^2 + \ell^2. \end{aligned}$$

On L, $y_2(x) = 0$, so,

$$y_1\big|_{\mathsf{L}} = (x^1)^2 + (x^2)^2 - R^2 + \ell^2.$$

Given that $y_1\big|_{\mathsf{L}} = 0$ by definition, we can consider the smooth extension $\alpha : \mathsf{M} \to \mathbb{R}$ of $y_1\big|_{\mathsf{L}}$,

$$\alpha(x^1, x^2) = (x^1)^2 + (x^2)^2 - R^2 + \ell^2.$$

Additionally, by construction, it satisfies the requirement,

$$\mathrm{d}\alpha \in \langle \mathrm{d}x^1, \mathrm{d}x^2 \rangle \subseteq \langle \mathcal{I}^{(2)}, \mathrm{d}t \rangle^{(\infty)}.$$

Using α we can rewrite the generators for $\langle \mathcal{I}^{(2)}, \mathrm{d}t \rangle^{(\infty)}$ as,

$$\langle \mathcal{I}^{(2)}, \mathrm{d}t \rangle^{(\infty)} = \langle \mathrm{d}\alpha, x^2\,\mathrm{d}x^1 - x^1\,\mathrm{d}x^2 \rangle.$$

Define the local regular zero dynamics manifold of system (4.23) with output $h \coloneqq \alpha$ by,

$$\mathsf{Z}^{(3)} \coloneqq \left\{ x \in \mathbb{R}^n \colon (x^1)^2 + (x^2)^2 - R^2 + \ell^2 = x^3\,\left(x^1\,\cos(x^4) + x^2\,\sin(x^4)\right) = \mathcal{L}_f^2\alpha = 0 \right\},$$

where

$$\mathcal{L}_f^2\alpha = x^3\,x^6\,\left(x^1\,\cos(x^4) + x^2\,\sin(x^4)\right) + (x^3)^2\left(1 + \frac{x^2\,\cos(x^4) - x^1\,\sin(x^4)}{\ell}\,\tan(x^5)\right).$$

The iteration of the algorithm proceeds at step $k = \kappa_1 = 2$ where we have $\rho_1 > \rho_2$. It follows that we expect to find $\mu = \rho_1 - \rho_2 = 1$ new scalar output that yields a relative degree of two at p_0 and is constant on L. The differential of this output must live in the differential ideal

$$\langle \mathcal{I}^{(1)}, \mathrm{d}t \rangle^{(\infty)} = \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3, \mathrm{d}x^4, \mathrm{d}t \rangle,$$

which is generated by exact one-forms.

By inspection, one integral of $\mathrm{d}x^3$ is $x^3 - 1$ which is a function that is zero on L. However, it is important that we ensure $\mathrm{d}x^3$ is differentially independent from the known output $h = \alpha$ and its Lie derivatives. This is a key component of the proposed algorithm. An equivalent condition is to ask that $\mathrm{d}x^3_p \notin \mathrm{ann}(\mathsf{T}_p\mathsf{Z}^{(3)})$ for all $p \in \mathsf{Z}^{(3)}$. This works because we know, by construction, that differentials of the known output h live in $\mathrm{ann}(\mathsf{T}_p\mathsf{Z}^{(3)})$.

The condition holds. As a result, the candidate output

$$h(x) = \begin{bmatrix} h^1(x) \\ h^2(x) \end{bmatrix} := \begin{bmatrix} \alpha(x) \\ x^3 - 1 \end{bmatrix} = \begin{bmatrix} (x^1)^2 + (x^2)^2 - R^2 + \ell^2 \\ x^3 - 1 \end{bmatrix}$$

for system (4.23) yields a vector relative degree of $(3, 2)$ at p_0 and vanishes on L. Therefore h is a local transverse output and solves the local transverse feedback linearization problem for system (4.23) with respect to N at x_0. In particular, define,

$$\begin{aligned} \xi^{1,3} &:= h^1, & \xi^{2,2} &:= h^2, \\ \xi^{1,2} &:= \mathcal{L}_f h^1, & \xi^{2,1} &:= \mathcal{L}_f h^2, \\ \xi^{1,1} &:= \mathcal{L}_f^2 h^1. \end{aligned} \tag{4.28}$$

Make any choice of differentially independent coordinate η^1 to complete the state transformation. For the feedback transformation, write

$$\begin{pmatrix} u^1 \\ u^2 \end{pmatrix} := \begin{pmatrix} \mathcal{L}_{g_1}\mathcal{L}_f^2 h^1 & \mathcal{L}_{g_2}\mathcal{L}_f^2 h^1 \\ \mathcal{L}_{g_1}\mathcal{L}_f h^2 & \mathcal{L}_{g_2}\mathcal{L}_f h^2 \end{pmatrix}^{-1} \left[-\begin{pmatrix} \mathcal{L}_f^3 h^1 \\ \mathcal{L}_f^2 h^2 \end{pmatrix} + \begin{pmatrix} v^1_\pitchfork \\ v^2_\pitchfork \end{pmatrix} \right], \tag{4.29}$$

where $v^1_\pitchfork$, $v^2_\pitchfork$ are the new virtual inputs, both of which act transversal to the set N. It can be verified that both (4.28) and (4.29) can be used on an open set of any point of N. The size of this open set is entirely determined by the choice of η^1.

Chapter 5

Future Work and Open Problems

Chapter 4 established the purported goal of this thesis: an algorithmic procedure to find the local transverse output (see Theorem 1.3.1) that solves the local transverse feedback linearization problem (Problem 1.3.1). The algorithm — Algorithm 1 — amounts to iteratively constructing a sequence of adapted Cauchy problems whose solution produces the required transverse output. Since state-space, exact feedback linearization is equivalent to the local transverse feedback linearization problem with $\mathsf{N} = \{x_0\}$, the algorithm applies just as well to the ubiquitous state-space, exact feedback linearization problem.

Equipped with an algorithm to compute a local transverse output, one might be tempted to stitch together a "global" transverse output through the use of a partition of unity; this in the hopes of solving a "global" transverse feedback linearization problem. We will see in Section 5.1 that this is not possible in general, but a careful understanding of the issues may yield progress to solving that problem.

Another interesting problem is whether the local transverse feedback linearization problem can be solved when full state-feedback cannot be employed, i.e. only an output is available. In [39], the dual space was used to address the problem in the single-input case. We discuss in Section 5.2 this problem and briefly discuss how Algorithm 1 may inform a solution.

5.1 Global Transverse Feedback Linearization

Consider the multi-input, nonlinear control system,

$$\dot{x}(t) = f(x(t)) + \sum_{i=1}^{m} g_i(x(t))\, u^i(t), \tag{5.1}$$

where $x(t) \in \mathbb{R}^n$, and let $\mathsf{N} \subseteq \mathbb{R}^n$ be a closed, embedded n^*-dimensional submanifold of the state-space. Suppose N is rendered controlled-invariant by input signal $u_*(x)$. The global transverse feedback linearization problem is as follows.

Problem 5.1.1 (Global Transverse Feedback Linearization (GTFL))**.** *Find an open set* U *containing* N, *and a feedback transformation* (Φ, α, β) *defined on* U,

(i) The restriction of Φ *to* N *is* $\Phi|_{\mathsf{N}} : \eta \mapsto (\eta, 0)$, *and*

(ii) if we let $(\eta, \xi) := \Phi(x)$, *then the nonlinear control system* (1.4) *is differentially equivalent to,*

$$\begin{aligned} \dot{\eta}(t) &= \overline{f}(\eta(t), \xi(t)) + \sum_{i=1}^{m_{\parallel}} \overline{g}_{\parallel,i}(\eta(t), \xi(t))\, v_{\parallel}^i + \sum_{i=1}^{m_{\pitchfork}} \overline{g}_{\pitchfork,i}(\eta(t), \xi(t))\, v_{\pitchfork}^i, \\ \dot{\xi}(t) &= A\,\xi(t) + \sum_{j=1}^{m_{\pitchfork}} b_j\, v_{\pitchfork}^j(t), \end{aligned} \tag{5.2}$$

where $(A,\, b_1,\, \ldots,\, b_{m_{\pitchfork}})$ *is in Brunovský normal form.*

A necessary condition to solve Problem 5.1.1 is that there exists a transverse output that yields a relative degree of $n - n^*$ at all points of N and vanishes on N. This is not sufficient.

Conditions upon which Problem 5.1.1 is solvable for single-input systems were suggested by Nielsen & Maggiore in [41, Theorem 4.4]. Their theorem shows that the vector field conditions for transverse feedback linearization, i.e. conditions of Theorem 1.3.2, are sufficient when N is parallelizable. These are only sufficient conditions, but, in principle, they can be used to explicitly construct the transverse output that solves the global transverse feedback linearization problem; albeit, the construction of the transverse output involves a very involved computation of flows.

It is not difficult to show that orientability of N is a necessary condition for the system (5.1) to be globally transverse feedback linearizable with respect to N.

Proposition 5.1.1. *If system* (5.1) *is globally transverse feedback linearizable with respect to* N, *then* N *is orientable.*

Proof. Suppose system (5.1) is globally transverse feedback linearizable with respect to N. Let U and $\Phi : \mathsf{U} \to \mathsf{N} \times \mathbb{R}^{n-n^*}$ be the open set and diffeomorphism that solves the problem. The orientation of $\mathbb{R}^n$ induces an orientation on U. Define $\xi : \mathsf{U} \subseteq \mathbb{R}^n \to \mathbb{R}^{n-n^*}$ to be smooth function that satisfies $\Phi(p) = (\eta(p), \xi(p)) \in \mathsf{N} \times \mathbb{R}^{n-n^*}$. By Problem 5.1.1 (i), $\xi|_{\mathsf{N}} = 0$. Additionally, since Φ is a diffeomorphism, ξ has constant rank $n-n^*$. It follows that $\mathsf{N} = \xi^{-1}(0)$ is a regular level set of ξ, and, by [31, Proposition 15.23], N is orientable. □

However, it turns out that we actually must impose a stronger condition on N: parallelizability. This is to ensure that the feedback transformation is regular in the sense that the controls v^i act in linearly independent directions everywhere on N.

As of yet, it is not known under what conditions Problem 5.1.1 is solvable when $m > 1$ (multi-input), let alone how to construct the transverse output in that case. Recognizing that it is certainly necessary to have a transverse output, one may be tempted to use the methods of Chapter 4 to construct a candidate solution to Problem 5.1.1. Unfortunately, the method used to construct the transverse output locally in Chapter 4 does not generalize well here.

For one, it is taken that a differentially closed form is exact in order to construct the smooth map $H_k : \mathsf{M} \to \mathbb{R}^{\ell_k}$ that satisfies the characteristic property,

$$\langle \mathrm{d}H_k^1, \ldots, \mathrm{d}H_k^{\ell_k} \rangle = \langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}.$$

The components of H_k are formed by integrating the differentially closed generators of $\langle \mathcal{I}^{(k)}, \mathrm{d}t \rangle^{(\infty)}$. It is not always true that a differentially closed form is exact, i.e. can be integrated to form a smooth function. In particular, one would require that the open set

U, where the form is differentially closed, have a trivial cohomology. In local transverse feedback linearization, U can be taken to be a sufficiently small open ball which does have a trivial cohomology. As the next example demonstrates, working locally may result in solutions that cannot be extended to solve the global problem.

Example 5.1 (Failure to Extend Local Transverse Outputs) . Let $x(t) \in \mathbb{R}^3$. Consider the nonlinear control system,

$$\dot{x}(t) = \begin{pmatrix} x^2(t) \\ -x^1(t) \\ -\frac{1}{2} \end{pmatrix} + \begin{pmatrix} \frac{x^1(t)\, x^3(t)}{1-(x^3(t))^2} \\ \frac{x^2(t)\, x^3(t)}{1-(x^3(t))^2} \\ 1 \end{pmatrix} u^1(t) \tag{5.3}$$

and suppose we want to globally transverse feedback linearize this system with respect to the unit circle laying in the $z = 0$ plane; that is, the 1-dimensional submanifold $\mathsf{N} :=$ $\mathsf{S}^1 \times \{0\} \subseteq \mathbb{R}^3$. The augmented ideals $\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t\rangle$ are given by,

$$\begin{aligned}
\langle \mathcal{I}^{(0)}, \mathrm{d}t\rangle &= \langle \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3\rangle, \\
\langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle &= \langle \mathrm{d}x^1 - \frac{x^1\, x^3}{1-(x^3)^2}\,\mathrm{d}x^3, \mathrm{d}x^2 - \frac{x^2\, x^3}{1-(x^3)^2}\,\mathrm{d}x^3\rangle, \\
\langle \mathcal{I}^{(2)}, \mathrm{d}t\rangle &= \langle \mathrm{d}t\rangle.
\end{aligned}$$

For any $p \in \mathsf{L} := \{0\} \times \mathbb{R} \times \mathsf{N}$,

$$\begin{aligned}
\mathrm{ann}(\mathsf{T}_p\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}_p^{(0)}, \mathrm{d}t_p\} &= \mathrm{ann}(\mathsf{T}_p\mathsf{L}), \\
\mathrm{ann}(\mathsf{T}_p\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}_p^{(1)}, \mathrm{d}t_p\} &= \mathrm{span}_{\mathbb{R}}\{x^1\,\mathrm{d}x^1_p + x^2\,\mathrm{d}x^2_p, \mathrm{d}t_p\}, \\
\mathrm{ann}(\mathsf{T}_p\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\{\mathscr{I}_p^{(2)}, \mathrm{d}t_p\} &= \mathrm{span}_{\mathbb{R}}\{\mathrm{d}t_p\}
\end{aligned}$$

The controllability condition (Con), involutivity condition (Inv) and constant dimension condition (Dim) do hold for all $p \in \mathsf{L}$. Suppose we wanted to find the transverse output as described in Algorithm 1. In this case, there is only one input and, therefore, only one scalar output to find, and its relative degree must be $n - n^* = 2$ at every point of N. This output's differential lives in the differential ideal, $\langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle^{(\infty)}$. In local transverse feedback linearization, we would use Lemma 4.2.4 which uses Frobenius's Theorem to find exact generators for the differential ideal $\langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle^{(\infty)}$. Unfortunately, Frobenius's Theorem only finds exact generators for the differential ideal *locally*. Worse still, the adaptation process of Lemma 4.2.6 can further restrict the open set and the resulting generators may not always extend naturally to an open set containing L. Let us see how.

Consider the point $x_0 = (0, 1, 0) \in \mathsf{N}$ and the lifted point $p_0 = (0, 1/2, 0, 1, 0) \in \mathsf{L}$. We will use Frobenius's Theorem to find exact generators for the differential ideal $\langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle^{(\infty)}$, and then use these exact generators to define a smooth function H_1 whose component function's differentials generate the ideal. The differential ideal is exactly generated by

$$\langle \mathcal{I}^{(1)}, \mathrm{d}t\rangle^{(\infty)} = \left\langle \sqrt{1-(x^3)^2}\,\mathrm{d}x^2 + \frac{x^2\,x^3}{\sqrt{1-(x^3)^2}}\,\mathrm{d}x^3, \frac{1}{x^2}\,\mathrm{d}x^1 - \frac{x^1}{(x^2)^2}\,\mathrm{d}x^2 \right\rangle.$$

Clearly, this holds only on the convex, open set

$$\mathsf{U} := \left\{(t, u, x) \in \mathsf{M} \colon x^2 > 0, -1 < x^3 < 1\right\}.$$

We can integrate these locally exact one-forms on U to find a smooth function

$$H_1(x) := \left(x^2\,\sqrt{1-(x^3)^2}, \frac{x^1}{x^2}\right),$$

whose component function's differentials generate the differential ideal.

The restriction $H_1|_{\mathsf{U}\cap\mathsf{L}}$ has constant rank 1 and, as a result, we may use Lemma 4.2.6 to rewrite H_1 to have one transverse component with respect to L. For instance, one solution is,

$$\tilde{H}_1(x) = \left(x^2\,\sqrt{1-(x^3)^2}, \frac{\sqrt{1-(x^2)^2(1-(x^3)^2)}}{x^2\,\sqrt{1-(x^3)^2}} - \frac{x^1}{x^2}\right),$$

where the second component is constant on $\mathsf{U}\cap\mathsf{L}$. As a result, the local transverse feedback linearization problem is solvable at x_0 using the output

$$h(x) = \frac{\sqrt{1-(x^2)^2(1-(x^3)^2)}}{x^2\,\sqrt{1-(x^3)^2}} - \frac{x^1}{x^2}.$$

Unfortunately, this output cannot be extended onto a larger open set containing N. Instead consider the smooth function

$$h(x) = (x^1)^2 + (x^2)^2 - 1 - (x^3)^2.$$

This function vanishes on N, and, with system (5.3), yields a relative degree of 2 at every point of N. ◀

Example 5.1 shows that the methods applied in Chapter 4 force us to work locally, and the resulting output function may not extend naturally to solve the global problem. A standard technique in smooth differential geometry is to take local solutions to a problem and use the partition of unity to build a global solution. Suppose we found a finite open cover $\{\mathsf{U}_i\}$ of N, and solved the local transverse feedback linearization problem on these open sets. That is, we found a family of outputs $h_i : \mathsf{U}_i \to \mathbb{R}^{n-n^*}$ that with system (5.1) yield the required relative degree on all points of $\mathsf{U}_i \cap \mathsf{N}$. Is it possible to stitch these outputs together using a partition of unity to form a transverse output that is defined on $\bigcup_i \mathsf{U}_i$? This is not always possible, as the next example demonstrates, even when the global transverse feedback linearization problem is solvable.

Example 5.2 (Stitching Local Transverse Outputs) . Consider the linear control system

$$\dot{x}(t) = \begin{pmatrix} x^1(t) \\ 0 \\ -x^2(t) \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} u^1(t), \tag{5.4}$$

and suppose we want to globally transverse feedback linearize this system with respect to the x^1-axis,

$$\mathsf{N} := \{(x^1, 0, 0) \in \mathbb{R}^3 \colon x^1 \in \mathbb{R}\}.$$

Consider the open cover $\{\mathsf{U}_1, \mathsf{U}_2\}$ of N given by

$$\mathsf{U}_1 := \{x \in \mathbb{R}^3 \colon x^1 < 1\}, \quad \mathsf{U}_2 := \{x \in \mathbb{R}^3 \colon x^1 > -1\},$$

and define the smooth functions

$$h^1 : \mathsf{U}_1 \to \mathbb{R}, \qquad\qquad h^2 : \mathsf{U}_2 \to \mathbb{R}$$
$$x \mapsto x^2 - x^3 \qquad\qquad x \mapsto e^{x^3 - x^2} - 1$$

Both functions solve the local transverse feedback linearization problem in the open sets they are defined on since they vanish on N, and yield the required relative degree of 2 at every point of $\mathsf{U}_i \cap \mathsf{N}$. In fact, both functions can be extended globally in the natural way to solve the global transverse feedback linearization problem. However, it is not always the case that we will be lucky to find functions that do extend naturally. We now show that a partition of unity cannot be used to construct a global transverse output out of these local transverse outputs.

Take any partition of unity $\lambda_1, \lambda_2 : \mathsf{U}_1 \cup \mathsf{U}_2 \to \mathbb{R}$ subordinate to the open cover U_1 and U_2. Define

$$h(x) \coloneqq \lambda_1(x)\, h^1(x) + \lambda_2(x)\, h^2(x).$$

Substituting our expression for h_1 and h_2 and using the partition of unity property find,

$$h(x) = e^{x^3 - x^2} - 1 + \lambda_1(x) \left(x^2 - x^3 - e^{x^3 - x^2} + 1 \right).$$

Compute the exterior derivative and observe that, on N,

$$\mathrm{d}h|_{\mathsf{U}_1 \cap \mathsf{U}_2 \cap \mathsf{N}} = (1 - 2\lambda_1)\, \mathrm{d}x^3 - (1 - 2\lambda_1)\, \mathrm{d}x^2.$$

Since λ_1 is a bump function that varies from 0 to 1 over $\mathsf{U}_1 \cap \mathsf{U}_2$, there exists a point on $\mathsf{U}_1 \cap \mathsf{U}_2 \cap \mathsf{N}$ where $\mathrm{d}h$ vanishes. This contradicts the requirement that h yields a relative degree at all points of N. ◀

The problem observed in Example 5.2 is a problem of orientation. Output h^2 induces a different orientation onto N than output h^1. Flipping the sign of h^2, or considering the function $e^{x^2 - x^3} - 1$, solves this issue, but a second, more insidious problem, arises.

Example 5.3 (Stitching Local Transverse Outputs ctd.) . Again consider the linear control system (5.4) and the open cover U_1, U_2. Define the smooth functions,

$$h^1 : \mathsf{U}_1 \to \mathbb{R}, \qquad h^2 : \mathsf{U}_2 \to \mathbb{R}$$
$$x \mapsto x^2 - x^3 \qquad x \mapsto \exp(h^1(x)) - 1$$

Unlike in the last example, let us consider a class of partitions. Suppose $\lambda_1 : \mathsf{U}_1 \cup \mathsf{U}_2 \to \mathbb{R}$ is a smooth transition function that satisfies,

$$\lambda_1|_{\mathsf{U}_1 \setminus \mathsf{U}_2} = 1, \quad \lambda_1|_{\mathsf{U}_2 \setminus \mathsf{U}_1} = 0,$$

and is *only* a function of x^1 on $\mathsf{U}_1 \cap \mathsf{U}_2$. That is, $\mathrm{d}\lambda_1|_{\mathsf{U}_1 \cap \mathsf{U}_2} = \beta(x^1)\,\mathrm{d}x^1$. Set $\lambda_2 \coloneqq 1 - \lambda_1$. Together, λ_1 and λ_2 form a partition of unity subordinate to the open cover $\mathsf{U}_1, \mathsf{U}_2$. Define the smooth function

$$h(x) \coloneqq \lambda_1(x)\, h^1(x) + \lambda_2(x)\, h^2(x).$$

Substituting our expressions for λ_i and $h^2(x)$ find that, on $\mathsf{U}_1 \cap \mathsf{U}_2$,

$$h(x) = \exp(h^1(x)) + 1 + (h^1(x) - \exp(h^1(x)) + 1)\,\lambda_1(x).$$

Let us compute the Lie derivative of h along the control vector field $g(x) = \partial/\partial x^1 + \partial/\partial x^2 + \partial/\partial x^3$ for system (5.4). Observe that,

$$\begin{aligned} \mathcal{L}_g h &= \mathrm{d}h(g), \\ &= \exp(h^1)\,\mathrm{d}h^1(g) + \lambda_1\,(1 - \exp(h^1))\mathrm{d}h^1(g) + (h^1 - \exp(h^1) + 1)\,\mathrm{d}\lambda_1(g). \end{aligned}$$

Since h^1 yields a relative degree of 2, we know that $\mathcal{L}_g h^1 = \mathrm{d}h^1(g) = 0$. As a result,

$$\mathcal{L}_g h = (h^1 - \exp(h^1) + 1)\,\mathrm{d}\lambda_1(g).$$

Using the fact that $\mathrm{d}\lambda_1|_{\mathsf{U}_1 \cap \mathsf{U}_2} = \beta(x^1)\,\mathrm{d}x^1$, we have

$$\mathcal{L}_g h = (h^1 - \exp(h^1) + 1)\,\beta(x^1)\,\mathrm{d}x^1(g) = (h^1 - \exp(h^1) + 1)\,\beta(x^1).$$

This is non-zero in an open set containing N. However, we require that the differential $\mathrm{d}h$ annihilate the control vector field in order to yield a well-defined relative degree of 2. Therefore, system (5.4) with output h does not yield the required relative degree of 2 and fails to solve the global transverse feedback linearization problem. ◀

Example 5.3 illustrates that the partitions of unity themselves must yield a relative degree that is greater than or equal to the relative degree of the desired output. Otherwise, the partition of unity will destroy the relative degree property satisfied by the local transverse outputs.

To summarize, we have discussed a range of problems that appear in the attempt to solve the global transverse feedback linearization problem. In general, the problem is only solvable for parallelizable manifolds N. Solvability conditions are unknown in the multi-input case, let alone the algorithm to compute the required feedback transformation.

Even if we relaxed the problem to ask us to find a global transverse output, the techniques used in Chapter 4 force us to work locally. We have seen that these local solutions do not always extend to be a global transverse output. Worse still, even when the local transverse feedback linearization problem is solved on every open set of a finite open cover of N, a partition of unity can fail to stitch these solutions together; this is either because the stitched solution fails to induce a consistent orientation, or because the relative degree property is destroyed. Further research in this area will require a careful, methodical construction that keeps track of the decided orientation of N, and ensures the relative degree is preserved.

5.2 Transverse Feedback Linearization with Partial Information

Consider the multi-input system with output,

$$\begin{aligned} \dot{x}(t) &= f(x(t)) + \sum_{i=1}^{m} g_i(x(t))\, u^i(t) \\ y(t) &= h(x(t)). \end{aligned} \tag{5.5}$$

Assuming that only the output y is measurable, if h is not a diffeomorphism, then we cannot use the entire state x to perform feedback linearization. In particular, the transversal states $\xi = \phi(x)$ and virtual control $v = \alpha(x) + \beta(x)u$ may depend on the entire state x and, of

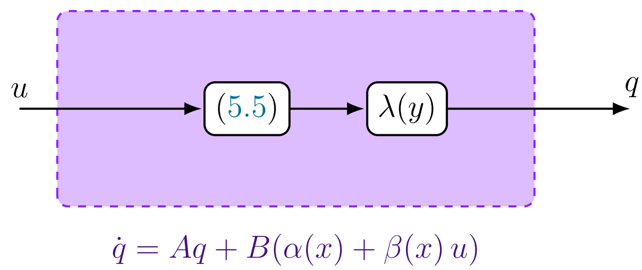


Figure 5.1: *When only the measurable outputs are available for feedback, the system can still be put in a form that is almost transverse feedback linearized with respect to a controlled-invariant set* N *by finding a clever state transformation.*

course, most (if not all) transversal states must be known in order to correctly place poles of the resulting linearized system.

This problem is important for economical and technological reasons. It is often the case that the states themselves are not observed but outputs, i.e. functions of states, that are. These are a function of the available sensors and may be restricted due to constraints in the design of the system. Like in Chapter 4, this problem can be cast as a problem of finding a transverse output.

Problem 5.2.1 (Transverse Feedback Linearization with Partial Information)**.** *Given system* (5.5) *find a transverse output* $q : \mathbb{R}^n \to \mathbb{R}^{\rho_0}$ *that can be written as* $q = \lambda \circ h$ *where* $q : \mathbb{R}^p \to \mathbb{R}^{\rho_0}$ *is a smooth function.*

The equivalence perspective on this problem is depicted in Figure 5.1. Unlike Figure 1.5, no feedback transformation is performed, but the transversal part of the state transformation must be a function of the output y. Currently the result is only established in its single-output formulation [39]. Given the observable outputs $h = (h^1, \ldots, h^\ell)$ define $\mathscr{W}$ as the distribution that satisfies,

$$\operatorname{ann}(\Gamma^\infty(\mathscr{W})) = \langle \mathrm{d}h^1, \ldots, \mathrm{d}h^\ell \rangle.$$

Then we have the following theorem, proven in [39], that solves Problem 5.2.1.

Theorem 5.2.1 ([39, Theorem 5.3])**.** *Suppose that* $\operatorname{inv}(\mathcal{G}^{(n-n^\star-2)}) + \mathscr{W}$ *is regular at* $x_0 \in \mathsf{N}$ *and suppose* $m = 1$. *Then Problem 5.2.1 is solvable at* x_0 *if and only if* *(a)* $\mathsf{T}_{x_0}\mathsf{N} \oplus \mathscr{G}_{x_0}^{(n-n^*-1)} = \mathsf{T}_{x_0}\mathbb{R}^n$ *and,*

(b) there exists an open neighbourhood U *of* $x_0 \in \mathbb{R}^n$ *so that, for all* $x \in \mathsf{U} \cap \mathsf{N}$,

$$\dim\left(\mathsf{T}_x\mathsf{N} \oplus \mathscr{G}_x^{(n-n^\star-2)}\right) = \dim\left(\mathsf{T}_x\mathsf{N} \oplus \mathrm{inv}(\mathscr{G}^{(n-n^\star-2)} + \mathscr{W})_x\right).$$

The multi-input result is not known. However, the results of this thesis provide useful information for those who endeavour to resolve this. Turning to the algorithm for transverse feedback linearization in Algorithm 1, observe that a necessary condition is that the "new" components of the transverse output computed in Line 13 must have differentials that live in the differential ideal $\langle \mathrm{d}h^1, \ldots, \mathrm{d}h^\ell \rangle$. Unfortunately, it is possible that choices made in previous steps of the algorithm can obstruct the ability to find such components in later steps. There is no obvious resolution to this problem. A likely path to success may hinge on specially constructing the regular zero dynamics manifolds in such a way so that the future steps of the algorithm proceed successfully while pulling transverse outputs from smooth functions of the h^i.

A simpler problem asks that the transverse output be observable: informally, the transverse output must be a smooth function combination of the measurable outputs and their Lie derivatives with respect to f. The space of observable states is characterized by the minimal codistribution $\mathscr{O}$ that is invariant under f, g_1, ..., g_m and contains the measurable outputs $\mathrm{d}h^i$ [27]. Suppose $\mathscr{O}$ is associated with a simply, finitely generated, differential ideal,

$$\mathcal{O} = \langle \mathrm{d}b^1, \ldots, \mathrm{d}b^r \rangle.$$

It is a necessary condition that

$$\mathrm{ann}(\mathsf{T}_p\mathsf{L}) \cap \mathrm{span}_{\mathbb{R}}\left\{\mathscr{I}_p^{(\kappa)}, \mathrm{d}t_p\right\} \subseteq \mathscr{O}_p,$$

since Line 13 of Algorithm 1 can be modified to ensure the transverse components come from the differential ideal $\mathcal{O}$. Unfortunately, this doesn't solve the original problem since it is possible for the new components to be Lie derivatives of the measureable output and not the measureable output itself.

Bibliography

[1] Ralph Abraham, Jerrold E. Marsden, and Tudor Ratiu. *Manifolds, Tensor Analysis, and Applications*. Springer New York, 1988. DOI: 10.1007/978-1-4612-1029-0.

[2] Vincent Andrieu, Bayu Jayawardhana, and Laurent Praly. "Transverse Exponential Stability and Applications". In: *IEEE Transactions on Automatic Control* 61.11 (Nov. 2016), pp. 3396–3411. DOI: 10.1109/tac.2016.2528050.

[3] Eduardo Aranda-Bricaire, Ülle Kotta, and Claude H. Moog. "Linearization of Discrete-Time Systems". In: *SIAM Journal on Control and Optimization (SICON)* 34.6 (Nov. 1996), pp. 1999–2023. DOI: 10.1137/s0363012994267315.

[4] Eduardo Aranda-Bricaire, Claude H. Moog, and Jean-Baptiste Pomet. "A linear algebraic framework for dynamic feedback linearization". In: *IEEE Transactions on Automatic Control* 40.1 (1995), pp. 127–132. DOI: 10.1109/9.362886.

[5] Andrzej Banaszuk and John Hauser. "Feedback linearization of transverse dynamics for periodic orbits". In: *Systems & Control Letters* 26.2 (Sept. 1995), pp. 95–105. DOI: 10.1016/0167-6911(94)00110-H.

[6] Bernhard Bischof et al. "A Path/Surface Following Control Approach to Generate Virtual Fixtures". In: *IEEE Transactions on Robotics* 34.6 (Dec. 2018), pp. 1577–1592. DOI: 10.1109/tro.2018.2861913.

[7] Franco Blanchini. "Set invariance in control". In: *Automatica* 35.11 (Nov. 1999), pp. 1747–1767. DOI: 10.1016/s0005-1098(99)00113-2.

[8] Roger W. Brockett, Richard S. Millman, and Hector J. Sussmann. *Differential geometric control theory : proceedings of the conference held at Michigan Technological University, June 28–July 2, 1982.* Progress in mathematics ; v. 27. Boston: Birkhaeuser, 1983.

[9] Robert L. Bryant et al. *Exterior differential systems.* Ed. by S. S. Chern et al. Vol. 18. Mathematical Sciences Research Institute Publications. Springer New York, 2013. ISBN: 9781461397168.

[10] Francesco Bullo and Andrew D. Lewis. *Geometric Control of Mechanical Systems.* Ed. by J. E. Marsden, L. Sirovich, and M. Golubitsky. Vol. 49. Texts in Applied Mathematics. Springer New York, 2005. ISBN: 9781441919687.

[11] Lawrence Conlon. *Differentiable Manifolds.* Birkhäuser Boston, 2001. DOI: 10.1007/978-0-8176-4767-4.

[12] Rollen S. D'Souza and Christopher Nielsen. "An exterior differential characterization of single-input local transverse feedback linearization". In: *Automatica* 127 (May 2021), p. 109493. DOI: 10.1016/j.automatica.2021.109493.

[13] Michel Fliess et al. "Flatness and defect of non-linear systems: introductory theory and examples". In: *International Journal of Control* 61.6 (June 1995), pp. 1327–1361. DOI: 10.1080/00207179508921959.

[14] Kenji Fujimoto and Mitsuru Taniguchi. "Passive path following control for port-Hamiltonian systems". In: *2008 47th IEEE Conference on Decision and Control.* IEEE, 2008. DOI: 10.1109/cdc.2008.4739242.

[15] Robert B. Gardner and William F. Shadwick. "An algorithm for feedback linearization". In: *Differential Geometry and its Applications* 1.2 (Sept. 1991), pp. 153–158. DOI: 10.1016/0926-2245(91)90028-8.

[16] Robert B. Gardner and William F. Shadwick. "The GS algorithm for exact linearization to Brunovsky normal form". In: *IEEE Transactions on Automatic Control* 37.2 (1992), pp. 224–230. DOI: 10.1109/9.121623.

[17] Conrad Gstöttner, Bernd Kolar, and Markus Schöberl. "Necessary and Sufficient Conditions for the Linearizability of Two-Input Systems by a Two-Dimensional Endogenous Dynamic Feedback". In: *International Journal of Control* (Dec. 2021), pp. 1–. DOI: 10.1080/00207179.2021.2015542.

[18] Martin Guay, P. James Mclellan, and David W. Bacon. "A condition for dynamic feedback linearization of control-affine nonlinear systems". In: *International Journal of Control* 68.1 (Jan. 1997), pp. 87–106. DOI: 10.1080/002071797223749.

[19] Mohamed I. El-Hawwary and Manfredi Maggiore. "Global path following for the unicycle and other results". In: *2008 American Control Conference*. IEEE, June 2008. DOI: 10.1109/acc.2008.4587035.

[20] Mohamed I. El-Hawwary and Manfredi Maggiore. "Reduction theorems for stability of closed sets with application to backstepping control design". In: *Automatica* 49.1 (Jan. 2013), pp. 214–222. DOI: 10.1016/j.automatica.2012.09.013.

[21] Robert Hermann. "Invariants for feedback equivalence and Cauchy characteristic multifoliations of nonlinear control systems". In: *Acta Applicandae Mathematicae* 11.2 (Feb. 1988), pp. 123–153. DOI: 10.1007/bf00047284.

[22] Robert Hermann. "Nonlinear feedback control and systems of partial differential equations". In: *Acta Applicandae Mathematicae* 17.1 (Nov. 1989), pp. 41–94. DOI: 10.1007/bf00052493.

[23] Robert Hermann. "On the Accessibility Problem in Control Theory". In: *International Symposium on Nonlinear Differential Equations and Nonlinear Mechanics*. Elsevier, 1963, pp. 325–332. DOI: 10.1016/b978-0-12-395651-4.50035-0.

[24] Robert Hermann. "The theory of equivalence of Pfaffian systems and input systems under feedback". In: *Mathematical Systems Theory* 15.1 (Dec. 1981), pp. 343–356. DOI: 10.1007/bf01786990.

[25] Ronald M. Hirschorn. "$(A, \mathcal{B})$-Invariant Distributions and Disturbance Decoupling of Nonlinear Systems". In: *SIAM Journal on Control and Optimization (SICON)* 19.1 (Jan. 1981), pp. 1–19. DOI: 10.1137/0319001.

[26] Louis R. Hunt, Renjeng Su, and George Meyer. “Design for Multi-Input Nonlinear Systems”. In: *Differential Geometric Control Theory*. Ed. by Roger W. Brockett, Richard S. Millman, and Hector J. Sussmann. Birkhäuser, 1983, pp. 268–298.

[27] Alberto Isidori. *Nonlinear Control Systems*. Third. Springer-Verlag London, 1995. ISBN: 978-1-4471-3909-6.

[28] Nilay Kant and Ranjan Mukherjee. “Energy-Based Orbital Stabilization of Underactuated Systems Using Impulse Controlled Poincaré Maps”. In: *2021 American Control Conference (ACC)*. IEEE, May 2021. DOI: 10.23919/acc50511.2021.9483236.

[29] Yuri A. Kapitanyuk, Anton V. Proskurnikov, and Ming Cao. “A Guiding Vector-Field Algorithm for Path-Following Control of Nonholonomic Mobile Robots”. In: *IEEE Transactions on Control Systems Technology* 26.4 (July 2018), pp. 1372–1385. DOI: 10.1109/tcst.2017.2705059.

[30] Arthur Krener, Alberto Isidori, and Witold Respondek. “Partial and robust linearization by feedback”. In: *The 22nd IEEE Conference on Decision and Control*. IEEE, 1983. DOI: 10.1109/cdc.1983.269814.

[31] John M. Lee. *Introduction to Smooth Manifolds*. Second. Springer New York, Aug. 26, 2012. 724 pp. ISBN: 1441999817.

[32] Jean Lévine. “On the Equivalence Between Differential Flatness and Dynamic Feedback Linearizability”. In: *IFAC Proceedings Volumes* 40.20 (2007), pp. 338–343. DOI: 10.3182/20071017-3-br-2923.00056.

[33] Shunjie Li, Claude H. Moog, and Witold Respondek. “Maximal feedback linearization and its internal dynamics with applications to mechanical systems on”. In: *International Journal of Robust and Nonlinear Control* 29.9 (Mar. 2019), pp. 2639–2659. DOI: 10.1002/rnc.4507.

[34] Riccardo Marino. “On the largest feedback linearizable subsystem”. In: *Systems & Control Letters* 6.5 (Jan. 1986), pp. 345–351. DOI: 10.1016/0167-6911(86)90130-1.

[35] Rachit Mehra et al. “Feedback linearization of Single-Input and Multi-Input Control System”. In: *IFAC Proceedings Volumes* 47.3 (2014), pp. 5479–5484. DOI: 10.3182/20140824-6-za-1003.02270.

[36] Richard M. Murray. “Applications and extensions of Goursat normal form to control of nonlinear systems”. In: *Proceedings of 32nd IEEE Conference on Decision and Control*. IEEE. DOI: 10.1109/cdc.1993.325849.

[37] Richard M. Murray. “Nilpotent Bases for a Class of Non-Integrable Distributions with Applications to Trajectory Generation for Nonholonomic Systems”. In: *Math. Control Signals Systems* 7 (1994), pp. 58–75.

[38] F. Nicolau and W. Respondek. “Normal Forms for Flat Two-input Control Systems Linearizable via a Two-fold Prolongation”. In: *IFAC-PapersOnLine* 53.2 (2020), pp. 5441–5446. DOI: 10.1016/j.ifacol.2020.12.1545.

[39] Christopher Nielsen. “Transverse feedback linearization with partial information for single-input systems”. In: *SIAM Journal on Control and Optimization (SICON)* 52.5 (Sept. 2014), pp. 3002–3021. DOI: 10.1137/120900149.

[40] Christopher Nielsen and Manfredi Maggiore. “On Local Transverse Feedback Linearization”. In: *SIAM Journal on Control and Optimization (SICON)* 47.5 (Jan. 2008), pp. 2227–2250. DOI: 10.1137/070682125.

[41] Christopher Nielsen and Manfredi Maggiore. “Output stabilization and maneuver regulation: A geometric approach”. In: *Systems & Control Letters* 55.5 (May 2006), pp. 418–427. DOI: 10.1016/j.sysconle.2005.09.006.

[42] Henk Nijmeijer and Arjan Van Der Schaft. *Nonlinear Dynamical Control Systems*. Springer New York, Feb. 2016. 440 pp. ISBN: 038797234X.

[43] Shankar Sastry. *Nonlinear Systems*. Springer-Verlag New York, 1999. 700 pp. ISBN: 1441931325.

[44] Markus Schöberl and Kurt Schlacher. “On an implicit triangular decomposition of nonlinear control systems that are 1-flat—A constructive approach”. In: *Automatica* 50.6 (June 2014), pp. 1649–1655. DOI: 10.1016/j.automatica.2014.04.007.

[45] Markus Schöberl and Kurt Schlacher. "On calculating flat outputs for Pfaffian systems by a reduction procedure - Demonstrated by means of the VTOL example". In: *2011 9th IEEE International Conference on Control and Automation (ICCA)*. IEEE, Dec. 2011. DOI: 10.1109/icca.2011.6137922.

[46] Anton S. Shiriaev, Leonid B. Freidovich, and Sergei V. Gusev. "Transverse Linearization for Controlled Mechanical Systems With Several Passive Degrees of Freedom". In: *IEEE Transactions on Automatic Control* 55.4 (Apr. 2010), pp. 893–906. DOI: 10.1109/tac.2010.2042000.

[47] Willem M. Sluis. "A necessary condition for dynamic feedback linearization". In: *Systems & Control Letters* 21.4 (Oct. 1993), pp. 277–283. DOI: 10.1016/0167-6911(93)90069-i.

[48] Dawn Tilbury and Shankar S. Sastry. "On Goursat normal forms, prolongations, and control systems". In: *Proceedings of 33rd IEEE Conference on Decision and Control*. IEEE, 1994. DOI: 10.1109/cdc.1994.411123.

[49] Walter M. Wonham. *Linear Multivariable Control: A Geometric Approach*. Springer Nature, Dec. 1985. 334 pp. ISBN: 0387960716.

[50] Bowen Yi et al. "Path following of a class of underactuated mechanical systems via immersion and invariance-based orbital stabilization". In: *International Journal of Robust and Nonlinear Control* 30.18 (Oct. 2020), pp. 8521–8544. DOI: 10.1002/rnc.5258.

[51] Songhao Yin and Masaki Yamakita. "Passive velocity field control approach to bicycle robot path following". In: *2016 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*. IEEE, Sept. 2016. DOI: 10.1109/sice.2016.7749208.

Appendix A

Elements of Differential Geometry

Critical to the discussion of transverse feedback linearization is the notion of a surface. The particular object we concern ourselves with is that of a manifold, which is a sort of generalization of a surface and is the fundamental concern of the topic of differential geometry.

A.1 Submanifolds of Euclidean Space

Before diving into the more abstract topic of manifolds, we first build on the familiar geometry found in $\mathbb{R}^n$. The tools we build generalize in a natural way to the more abstract smooth manifold structure.

Definition A.1.1 (Smooth Submanifold of $\mathbb{R}^n$). Let $0 \leq m < n$ be integers. A subset $\mathsf{M} \subset \mathbb{R}^n$ is called an *(smooth) m-dimensional embedded submanifold* of $\mathbb{R}^n$ if for every $p \in \mathsf{M}$ there exists an open set $\mathsf{U} \subseteq \mathbb{R}^n$ containing p in $\mathbb{R}^n$, an open subset V of $\mathbb{R}^m$ and a diffeomorphism $\varphi : \mathsf{U} \cap M \to \mathsf{V}$. The pair $\mathsf{U} \cap \mathsf{M}, \varphi$ is called a *chart*.

This definition says, in effect, that M is an m-dimensional embedded submanifold if locally, near every point $p \in \mathsf{M}$, M looks like an open subset of $\mathbb{R}^m$. A submanifold in this thesis always refers to an embedded submanifold of some Euclidean space. Another way to construct submanifolds is through the pre-image of smooth functions. Let U be an open set of $\mathbb{R}^n$ and let $H : \mathsf{U} \to \mathbb{R}^k$ be a smooth map. The function H is said to be a *submersion at* $p \in \mathsf{U}$ if the linear map $\mathbf{D}H|_p : \mathbb{R}^n \to \mathbb{R}^k$ is onto (has rank k). If H is a submersion at every point $p \in \mathsf{U}$ then H is called a *submersion*.

Definition A.1.2 (Regular Values). A point $y \in \mathbb{R}^k$ is called a *regular value* of $H : \mathsf{U} \to \mathbb{R}^k$ if H is a submersion at every point $p \in H^{-1}(y)$.

The next theorem ties the notion of regular values to Definition A.1.1.

Theorem A.1.3 ([31, Corollary 5.14])**.** *Let $H : \mathsf{U} \to \mathbb{R}^k$ and $m \coloneqq n - k$. If y is a regular value of H then the set $H^{-1}(y)$ is an m-dimensional embedded submanifold.*

Many interesting examples of manifolds arise as the pre-images of submersions. This is not an accident since, as the next result states, locally, every submanifold arises this way.

Theorem A.1.4 ([31, Proposition 5.16])**.** *Let M be an m-dimensional submanifold of $\mathbb{R}^n$ and let $k \coloneqq n - m$. For every $p \in \mathsf{M}$ there exists a neighbourhood U of p in $\mathbb{R}^n$ and a submersion $H : \mathsf{U} \to \mathbb{R}^k$ so that $\mathsf{U} \cap \mathsf{M}$ equals $H^{-1}(0)$.*

Remark A.1.5. *Take $\mathsf{M} = \mathbb{R}^n$ and consider the nonlinear control system,*

$$\dot{x}(t) = f(x(t)) + \sum_{i=1}^{m} g(x(t))\, u^i(t), \tag{A.1}$$

where $x(t) \in \mathsf{M}$. Let $\mathsf{N} \subset \mathsf{M}$ be a closed, embedded n^*-dimensional submanifold containing a point $x_0 \in \mathsf{N}$. In this thesis, the problem of finding a transverse output for N amounts to finding a submersion $H : \mathsf{U} \subseteq \mathsf{M} \to \mathbb{R}^{n-n^*}$, where the open set U contains x_0, so that, upon an appropriate choice of feedback, the dynamics of $\xi(t) \coloneqq H(x(t))$ are linear and controllable.

A.2 The Tangent Space

To every point on a surface in $\mathbb{R}^3$ we can naturally attach a tangent plane; this notion allows us to characterize the directions one may travel along the surface. This notion can be generalized to submanifolds. We start by fixing some notation.

Definition A.2.1 (Geometric Tangent Space) . Let $p \in \mathbb{R}^n$. The *(geometric) tangent space to* $\mathbb{R}^n$ *at* p is the set of pairs

$$\mathsf{T}_p\mathbb{R}^n \coloneqq \{(p, v) : v \in \mathbb{R}^n\} .$$

The identification $\mathsf{T}_p\mathbb{R}^n \simeq \mathbb{R}^n$, $(p, v) \mapsto v$ makes $\mathsf{T}_p\mathbb{R}^n$ a n-dimensional real vector space. Specifically, for v_1, $v_2 \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$, the vector space operations on $\mathsf{T}_p\mathbb{R}^n$ are defined as

$$(p, v_1) + (p, v_2) \coloneqq (p, v_1 + v_2), \qquad \lambda(p, v_1) \coloneqq (p, \lambda v_1).$$

Next we give an intuitive definition of the tangent space to a submanifold. It is not the most general definition because it leverages the fact that M sits inside $\mathbb{R}^n$.

Definition A.2.2 . Let M be an m-dimensional submanifold of $\mathbb{R}^n$, let $p \in \mathsf{M}$ and let (U, φ) be a coordinate chart with $p \in \mathsf{U}$. The *tangent space to* M *at* p, denoted $\mathsf{T}_p\mathsf{M}$, is the image of the linear map

$$\mathbf{D}\varphi^{-1}\Big|_{\varphi(p)} : \mathsf{T}_{\varphi(p)}\mathbb{R}^m \to \mathsf{T}_p\mathbb{R}^n.$$

In other words, $w \in \mathsf{T}_p\mathbb{R}^n$ is in $\mathsf{T}_p\mathsf{M}$ if, and only if $w = \mathbf{D}\varphi^{-1}|_q v$ for some $v \in \mathsf{T}_{\varphi(p)}\mathbb{R}^m$. Since φ is a diffeomorphism onto its image, $\mathbf{D}\varphi^{-1}|_q$ is one-to-one, and so $\mathsf{T}_p\mathsf{M}$ is an m-dimensional subspace of $\mathsf{T}_p\mathbb{R}^n$.

One problem with Definition A.2.2 is that it appears to depend on the coordinate chart (U, φ). This is not the case and it is easy to show that if (V, ϕ) is another coordinate chart with $p \in \mathsf{V}$, then $\operatorname{Im} \mathbf{D}\phi^{-1}|_{\phi(p)} = \operatorname{Im} \mathbf{D}\varphi^{-1}|_{\varphi(p)}$.

An equivalent definiton of $\mathsf{T}_p\mathsf{M}$ can be given using the submersion from Theorem A.1.4. Let $H : \mathsf{U} \to \mathbb{R}^k$ be the submersion whose existence is asserted by Theorem A.1.4 and where U is an open subset of $\mathbb{R}^n$ containing p and $H(p) = 0$. Since H is a submersion at all points of $\mathsf{U} \cap \mathsf{M}$, the map $\mathbf{D}H|_p : \mathsf{T}_p\mathbb{R}^n \to \mathsf{T}_0\mathbb{R}^k$ is onto and therefore its kernel has dimension m. It turns out that we can define $\mathsf{T}_p\mathsf{M}$ by

$$\mathsf{T}_p\mathsf{M} = \operatorname{Ker} \mathbf{D}H|_p .$$

It is possible to show that this definition is also independent of the particular submersion H. With these ideas in place, we can define the differential of a map between manifolds. Let M be a submanifold of $\mathbb{R}^n$, let N be a submanifold of $\mathbb{R}^r$, and let $H : \mathsf{M} \to \mathsf{N}$ be a smooth map. By definition there exists an open set U of M in $\mathbb{R}^n$ and a smooth map $\tilde{H} : \mathsf{U} \to \mathbb{R}^r$ extending H. We define, for each $p \in M$,

$$\mathbf{D}H|_p : \mathsf{T}_p\mathsf{M} \to \mathsf{T}_{H(p)}\mathsf{N}$$

to be the restriction of $\mathbf{D}\tilde{H}|_p : \mathsf{T}_p\mathbb{R}^n \to \mathsf{T}_{\tilde{H}(p)}\mathbb{R}^r$ to $\mathsf{T}_p\mathsf{M}$. It can be shown that this map is independent of the chosen extension.

Next we characterize tangent spaces in a way that generalizes to the case where there is no ambient Euclidean space. Let $C^\infty(p)$ denote the space of maps that are smooth in an open set containing $p \in \mathsf{M}$. Suppose that $(p, v) \in \mathsf{T}_p\mathbb{R}^n$, and let $H \in C^\infty(p)$. Then we can associate to v a linear map $L_v : C^\infty(p) \to \mathbb{R}$ defined as

$$L_v H(p) = \sum_{i=1}^{n} v^i \left. \frac{\partial H}{\partial x^i} \right|_p$$

where v^i is the i-th component of v. This map is just the directional derivative of H at p in the direction of v. The essential properties of this directional derivative are captured in the following definition.

Definition A.2.3 (Derivations). A *derivation* at a point $p \in \mathbb{R}^n$ is an operator $X_p : C^\infty(p) \to \mathbb{R}$ that satisfies, for all $\alpha, \beta \in \mathbb{R}$ and $G, H \in C^\infty(p)$,

(i) Linearity over $\mathbb{R}$: $X_p(\alpha G + \beta H) = \alpha X_p(G) + \beta X_p(H)$ and,

(ii) Leibniz Rule: $X_p(GH) = X_p(G)H(p) + G(p)X_p(H)$.

The set of all derivations at a point $p \in \mathbb{R}^n$ can be given a (real) vector space structure. Specifically, for derivations X_1, X_2 at p and $\lambda \in \mathbb{R}$, the vector space operations are defined, for $H \in C^\infty(p)$, as

$$(X_1 + X_2)(H) \coloneqq X_1(H) + X_2(H), \qquad (\lambda X_1)(H) \coloneqq \lambda X_1(H).$$

It can be shown [31, Proposition 3.2] that this vector space is isomorphic to the vector space $\mathsf{T}_p\mathbb{R}^n$ from Definition A.2.1. Indeed, the isomorphism is the map which takes v to its directional derivative L_v. Thus, with an abuse of notation, we also let the symbol $\mathsf{T}_p\mathbb{R}^n$ denote the vector space of derivations at p and write the canonical basis for this vector space as $\{\partial/\partial x^1, \ldots, \partial/\partial x^n\}$. With this in place we give our final, most general, definition of the tangent space to a point on a manifold.

Definition A.2.4 (Abstract Tangent Space) . Let M be an m-dimensional submanifold of $\mathbb{R}^n$. A linear map $X_p : C^\infty(\mathsf{M}) \to \mathbb{R}$ is called a *derivation at p* if it satisfies, for all $G, H \in C^\infty(\mathsf{M})$,

$$X_p(GH) = X_p(G)\, H(p) + G(p)\, X_p(H).$$

The set of all derivations at p is the *tangent space to* M *at p*, denoted $\mathsf{T}_p\mathsf{M}$. An element of $\mathsf{T}_p\mathsf{M}$ is called a *tangent vector at p*.

A.3 Vector Bundles and Distributions

We have now seen how a vector space — the tangent space, which captures the notion of "tangency" — can be attached to any point p of a smooth manifold M. This generalized the idea of a tangent plane to any fixed point on a surface in $\mathbb{R}^3$. It turns out that we can bundle together these vector spaces with the manifold M and endow this space with a smooth manifold structure. The structure views this space locally as a product between M and $\mathbb{R}^m$. In this section, we review this notion.

Recall that a *topological space* $\mathscr{E}$ is a set equipped with a topology $\mathscr{B}$ which consists of all sets which are classified as "open" on $\mathscr{E}$. The topology is closed under possibly uncountably infinite union and under finite intersections. A map is between topological spaces is said to be *continuous* if the pre-image of any open set in the codomain is open in the domain. A homeomorphism is a continuous bijection with a continuous inverse. Finally, any open subset $\mathsf{V} \subseteq \mathscr{E}$ of a topological space can be endowed a natural topology, known as the subspace topology, which consists of open sets of the form $\mathsf{V} \cap \mathsf{R}$ where R is open in $\mathscr{E}$. Equipped with this knowledge of topology, we can define a smooth vector bundle.

Definition A.3.1 ((Smooth) Vector Bundle) . Let M be an m-dimensional submanifold of $\mathbb{R}^n$. A *(smooth) vector bundle over* M *(of rank k)* is a smooth manifold $\mathscr{E}$ together with a surjective smooth map $\pi : \mathscr{E} \to \mathsf{M}$ so that,

(i) for each $p \in \mathsf{M}$, the fiber $\mathscr{E}_p := \pi^{-1}(p)$ is endowed with a k-dimensional (real) vector space structure, and

(ii) for each $p \in \mathsf{M}$, there exists an open set $\mathsf{U} \subseteq \mathsf{M}$ containing p and a diffeomorphism $\phi : \pi^{-1}(\mathsf{U}) \to \mathsf{U} \times \mathbb{R}^k$ that satisfies,

- if $p : \mathsf{U} \times \mathbb{R}^k \to \mathsf{U}$ is the standard projection then $p \circ \phi = \pi|_{\mathsf{U}}$, and
- for each $q \in \mathsf{U}$, the restriction $\phi|_{\mathscr{E}_q}$ is a vector space isomorphism.

Sometimes we will write a vector bundle as $\pi : \mathscr{E} \to \mathsf{M}$. A common theme throughout this thesis is the "smooth" assignment of vector subspaces of the vector spaces in the vector bundle to points on the manifold. To define this notion of smoothness, we first establish how to pick a smooth assignment of vectors to points on the manifold.

Definition A.3.2 (Smooth Sections of a Vector Bundle) . Let M be an m-dimensional submanifold of $\mathbb{R}^n$ and take $\pi : \mathscr{E} \to \mathsf{M}$ to be a vector bundle. A smooth section of $\mathscr{E}$ is a smooth map $X : \mathsf{M} \to \mathscr{E}$ between manifolds that satisfies $\pi \circ X = \mathrm{id}_{\mathsf{M}}$.

The set of all smooth sections of a vector bundle $\mathscr{E}$ is denoted $\Gamma^\infty(\mathscr{E})$. Given $X_1, X_2 \in \Gamma^\infty(\mathscr{E})$ we can define addition by the smooth map,

$$(X_1 + X_2)(p) := X_1(p) + X_2(p), \qquad p \in \mathsf{M},$$

and scalar multiplication by,

$$(\lambda X_1)(p) := \lambda\, X_1(p), \qquad p \in \mathsf{M}, \lambda \in \mathbb{R},$$

thereby endowing $\Gamma^\infty(\mathscr{E})$ with a real vector space structure. Alternatively, $\Gamma^\infty(\mathscr{E})$ can be given a $C^\infty(\mathsf{M})$-module structure by defining scalar multiplication as,

$$(f X_1)(p) := f(p)\, X_1(p), \qquad p \in \mathsf{M}, f \in C^\infty(\mathsf{M}).$$

It is, in general, not the case that a subspace of the smooth sections of a vector bundle can be viewed as a vector bundle in its own right.

Definition A.3.3 (Generalized Subbundle of a Vector Bundle). Let M be an m-dimensional submanifold of $\mathbb{R}^n$ and let $\pi : \mathscr{E} \to \mathsf{M}$ be a vector bundle. A subset $\mathscr{D} \subset \mathscr{E}$ is a *generalized subbundle of* $\mathscr{E}$ if, for any fixed $p \in \mathsf{M}$,

(i) the set $\mathscr{E}_p \cap \mathscr{D}$ is a vector subspace of $\mathscr{E}_p$, and

(ii) there exists an open set $\mathsf{U} \subseteq \mathsf{M}$ containing p and a family of smooth sections $\{X_\alpha\}$ of $\mathscr{E}$ so that, for all $q \in \mathsf{U}$,

$$\mathscr{E}_q \cap \mathscr{D} = \operatorname{span}_{\mathbb{R}} \{X_\alpha(q)\} .$$

The generalized subbundle is said to be *locally generated* and the $\{X_\alpha\}$ are called *local generators*.

A generalized subbundle of $\mathscr{E}$ assigns to each point $p \in \mathsf{M}$ a vector subspace of $\mathscr{E}_p$. The converse is not true. Also observe that Definition A.3.3 (i) does not restrict the dimension of the vector subspaces to be consistent (constant) over M.

Important to this thesis will be this consistency, which we describe as the notion of regularity. It will often be required that the dimension of all the vector spaces assigned to points on M are the same. The next definition incorporates this requirement.

Definition A.3.4 (Subbundle of a Vector Bundle). Let M be an m-dimensional submanifold of $\mathbb{R}^n$ and let $\pi : \mathscr{E} \to \mathsf{M}$ be a vector bundle. A generalized subbundle $\mathscr{D} \subset \mathscr{E}$ is a *subbundle of* $\mathscr{E}$ *(of rank* ℓ*)* if, for all $p \in \mathsf{M}$,

$$\dim(\mathscr{E}_p \cap \mathscr{D}) = \ell.$$

A related object found throughout the nonlinear control is the notion of a distribution. In the literature, a distribution begins *first* with assigning vector subspaces to points on the manifold. Given a vector bundle, a distribution $\mathscr{D} \subset \mathscr{E}$ assigns, to every point $p \in \mathsf{M}$, a vector subspace of $\mathscr{E}_p$. Clearly a distribution is, by itself, not a generalized subbundle.

A *smooth distribution* is simply a distribution that is a generalized subbundle of the tangent bundle TM, i.e. near every point there exists smooth sections that generate the vector subspaces in a neighbourhood of the point. Like a generalized subbundle, we say smooth distributions are locally generated. It is not difficult to see that a finite family of smooth sections $X_1, \ldots, X_\ell \in \Gamma^\infty(\mathsf{TM})$ that span a submodule of $\Gamma^\infty(\mathsf{TM})$ can be used to generate a smooth distribution (generalized subbundle).

A smooth and regular distribution is a subbundle of the tangent bundle. Such distributions are not only locally generated, but are locally, *finitely* generated. Unlike what was discussed earlier, it is not the case that a finite family of smooth sections $X_1, \ldots, X_\ell \in \Gamma^\infty(\mathsf{TM})$ can be used to generate a smooth and regular distribution (subbundle). On the other hand, we need only impose that the X_i are linearly independent pointwise — $X_1(p), \ldots, X_\ell(p) \in \mathscr{E}_p$ are linearly independent for all $p \in \mathsf{M}$ — to ensure that they generate a smooth and regular distribution. In this case, we say that the submodule generated by $X_1, \ldots, X_\ell \in \Gamma^\infty(\mathsf{TM})$ is *locally, finitely, non-degenerately generated.* Since a smooth and regular distribution $\mathscr{D}$ is a subbundle, the set $\Gamma^\infty(\mathscr{D}) \subseteq \Gamma^\infty(\mathsf{TM})$ is a subspace that is locally, finitely, non-degenerately generated.

A.4 The Space of Alternating Tensors

The dual space of $\mathsf{T}_p\mathsf{M}$, called the cotangent space, takes a prominent role in this thesis. Elements of the cotangent space consume vectors of the tangent space and produce a real number.

Definition A.4.1 (Cotangent Space). Let M be an m-dimensional submanifold of $\mathbb{R}^n$ and let $p \in \mathsf{M}$. The *cotangent space to* M *at* p is the dual vector space

$$\mathsf{T}_p^*\mathsf{M} := (\mathsf{T}_p\mathsf{M})^*\,.$$

An element of $\mathsf{T}_p^*\mathsf{M}$ is called a *cotangent vector at* p.

This structure can be further generalized to consider multi-input maps. A map $\omega_p : (\mathsf{T}_p\mathsf{M})^k \to \mathbb{R}$ is called $(0,k)$-*tensor at* p if it is a linear function in each of its parameters assuming all others are fixed, i.e.,

$$\begin{aligned}\omega_p(v_1, \dots, v_{\ell-1}, \alpha v_\ell + \beta v'_\ell, v_{\ell+1}, \dots, v_k) &= \alpha\, \omega_p(v_1, \dots, v_{\ell-1}, v_\ell, v_{\ell+1}, \dots, v_k) \\ &\quad + \beta\, \omega_p(v_1, \dots, v_{\ell-1}, v'_\ell, v_{\ell+1}, \dots, v_k).\end{aligned}$$

Some texts also call it a *covariant k-tensor at p*. Consider the set of all such covariant k-tensors at p, denoted $T^{(0,k)}(\mathsf{T}_p\mathsf{M})$. This is a vector space over $\mathbb{R}$ if we define addition and scalar multiplication, for $\omega_p, \beta_p \in T^{(0,k)}(\mathsf{T}_p\mathsf{M})$, by,

$$\begin{aligned}(\omega_p + \beta_p)(v_1, \dots, v_k) &\coloneqq \omega_p(v_1, \dots, v_k) + \beta_p(v_1, \dots, v_k) \\ (\alpha\omega_p)(v_1, \dots, v_k) &\coloneqq \alpha\, \omega_p(v_1, \dots, v_k).\end{aligned}$$

The space of $(0,0)$-tensors at p is defined to be the field $\mathbb{R}$ and the space of $(0,1)$-tensors at p is the cotangent space $\mathsf{T}^*_p\mathsf{M}$. A $(0,k)$-tensor at p, $\omega_p \in T^{(0,k)}(\mathsf{T}_p\mathsf{M})$, is called *alternating*, or an *alternating $(0,k)$-tensor at p* if its value changes sign whenever two arguments are swapped:

$$\omega_p(v_1, \dots, v_i, \dots, v_j, \dots, v_k) = -\omega_p(v_1, \dots, v_j, \dots, v_i, \dots, v_k).$$

Alternating $(0,k)$-tensors are also called *k-forms*. The space of k-forms at p, denoted $\Lambda^k(\mathsf{T}_p\mathsf{M})$, is a vector subspace of $T^{(0,k)}(\mathsf{T}_p\mathsf{M})$.

Proposition A.4.2. *The dimension of $\Lambda^k(\mathsf{T}_p\mathsf{M})$, for $k > m$, where m is the dimension of manifold M, is 0.*

Proposition A.4.2 implies that the "largest" non-trivial form at a point p on an m-dimensional manifold is an m-form. As a result, the space of all forms at p, denoted $\Lambda(\mathsf{T}_p\mathsf{M})$ is characterized by a *finite*, homogeneous, external direct sum of the individual spaces $\Lambda^k(\mathsf{T}_p\mathsf{M})$:

$$\Lambda(\mathsf{T}_p\mathsf{M}) \coloneqq \bigoplus_{k=0}^{n} \Lambda^k(\mathsf{T}_p\mathsf{M}).$$

where again $\Lambda^0(\mathsf{T}_p\mathsf{M}) = \mathbb{R}$. This space can be given a graded algebra structure. First we define a product between forms called a wedge product. In order to define it, recall that if $n \in \mathbb{N}$, then S_n denotes the symmetric group. Elements of S_n are bijections $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$. The parity of the bijection is denoted $\operatorname{sgn}(\sigma)$.

Definition A.4.3 (The Wedge Product) . Let $k, \ell \in \{1, \dots, m\}$ and say $\omega_p \in \Lambda^k(\mathsf{T}_p\mathsf{M})$ and $\beta_p \in \Lambda^\ell(\mathsf{T}_p\mathsf{M})$. Then the *wedge product* $\omega_p \wedge \beta_p \in \Lambda^{k+\ell}(\mathsf{T}_p\mathsf{M})$ is defined, for all $v_1, \dots, v_{k+\ell} \in \mathsf{T}_p\mathsf{M}$, as

$$(\omega_p \wedge \beta_p)(v_1, \dots, v_{k+\ell}) = \frac{1}{k!\,\ell!} \sum_{\sigma \in S_{k+\ell}} \operatorname{sgn}(\sigma)\, \omega_p(v_{\sigma(1)}, \dots, v_{\sigma(k)})\, \beta_p(v_{\sigma(k+1)}, \dots, v_{\sigma(k+\ell)}).$$

The following proposition establishes the properties that allow the wedge product to give a graded algebra structure to $\Lambda(\mathsf{T}_p\mathsf{M})$.

Proposition A.4.4 ([31, Proposition 14.11])**.** *Suppose* $\omega_p,\ \omega_p' \in \Lambda^k(\mathsf{T}_p\mathsf{M})$, $\beta_p,\ \beta_p' \in \Lambda^\ell(\mathsf{T}_p\mathsf{M})$ *and* $\gamma \in \Lambda^q(\mathsf{T}_p\mathsf{M})$. *Then,*

(a) Bilinearity*: For* $\alpha, \alpha' \in \mathbb{R}$, $(\alpha\, \omega_p + \alpha'\, \omega_p') \wedge \beta_p = \alpha\, (\omega_p \wedge \beta_p) + \alpha'\, (\omega_p' \wedge \beta_p)$,

(b) Associativity*:* $\omega_p \wedge (\beta_p \wedge \gamma) = (\omega_p \wedge \beta_p) \wedge \gamma$,

(c) Anticommutativity*:* $\omega_p \wedge \beta_p = (-1)^{k\ell} \beta_p \wedge \omega_p$.

The graded algebra $\Lambda(\mathsf{T}_p\mathsf{M})$, is sometimes called the *exterior algebra*. A consequence of Proposition A.4.4 is that a basis for each individual vector space $\Lambda^k(\mathsf{T}_p\mathsf{M})$ can be obtained directly from a basis for $\Lambda^1(\mathsf{T}_p\mathsf{M})$. In particular,

Proposition A.4.5. *Let* $\Lambda^1(\mathsf{T}_p\mathsf{M}) = \operatorname{span}_{\mathbb{R}}\{\omega_p^1, \dots, \omega_p^n\}$. *Then for* $k \in \{2, \dots, m\}$

$$\Lambda^k(\mathsf{T}_p\mathsf{M}) = \operatorname{span}_{\mathbb{R}} \left\{ \omega_p^{\sigma(1)} \wedge \cdots \wedge \omega_p^{\sigma(k)} \colon \sigma \in S_k, \sigma(1) < \sigma(2) < \cdots < \sigma(k) \right\}.$$

A.5 Vector Fields and Differential Forms

A vector field on an open subset $\mathsf{U} \subseteq \mathbb{R}^n$ is a function X which assigns, to each $p \in \mathsf{U}$, an element X_p of $\mathsf{T}_p\mathsf{U}$. The vector field is smooth if, for every $H \in C^\infty(U)$, the function $\mathcal{L}_X H$, defined pointwise by $(\mathcal{L}_X H)(p) \coloneqq X_p(H)$, is smooth. This has an obvious generalization to embedded submanifolds of $\mathbb{R}^n$.

Definition A.5.1 (Vector Fields on Manifolds) . Let M be a manifold. A *vector field* on M is a function X which assigns to each $p \in M$ an element X_p of $\mathsf{T}_p\mathsf{M}$. The vector field is smooth if, for all $H \in C^\infty(\mathsf{M})$, the function $\mathcal{L}_X H : \mathsf{M} \to \mathbb{R}$ defined pointwise by $(\mathcal{L}_X H)(p) \coloneqq X_p(H)$ is smooth.

The notation $\mathcal{L}_X H$ itself plays an important role in the thesis, as it computes how the function H changes under the flow of X.

Definition A.5.2 (Lie Derivative) . Let M be a manifold, X be a smooth vector field, and $H \in C^\infty(\mathsf{M})$. The Lie derivative of H along the vector field X is the smooth function $\mathcal{L}_X H$ defined pointwise by $(\mathcal{L}_X H)(p) \coloneqq X_p(H)$ for all $p \in \mathsf{M}$.

An equivalent characterization of smooth vector fields can be given using the notion of smooth sections. To do this, we first construct a vector bundle using the tangent space.

Definition A.5.3 (Tangent Bundle) . Let M be an m-dimensional submanifold of $\mathbb{R}^n$ and let $p \in \mathsf{M}$. The *tangent bundle of* M is the disjoint union,

$$\mathsf{TM} \coloneqq \bigsqcup_{p \in \mathsf{M}} \mathsf{T}_p\mathsf{M},$$

and can be made into a vector bundle over M of rank m.

The tangent bundle *bundles* together the tangent spaces at each and every point of the manifold. The tangent bundle is an important example of a smooth vector bundle. A smooth vector field $X : \mathsf{M} \to \mathsf{TM}$ is simply a smooth section of the tangent bundle. It is for this reason that we let $\Gamma^\infty(\mathsf{TM})$ denote the set of all smooth vector fields on a manifold M.

This set can be given different algebraic structures the most important of which, from the perspective of this thesis, is that of a (real) Lie algebra.

Definition A.5.4 (Lie Bracket of Vector Fields). For X, $Y \in \Gamma^\infty(\mathsf{TM})$, the vector field $[X, Y] \in \Gamma^\infty(\mathsf{TM})$ is the vector field that satisfies, for all $H \in C^\infty(M)$,

$$\mathcal{L}_{[X,Y]}H = \mathcal{L}_X(\mathcal{L}_Y H) - \mathcal{L}_Y(\mathcal{L}_X H).$$

It is called the *Lie bracket* of X and Y.

The Lie bracket of two smooth vector fields X, $Y \in \Gamma^\infty(\mathsf{TM})$ can be computed easily in local coordinates using the following proposition.

Proposition A.5.5 ([31, Proposition 8.26])**.** *Let* M *be an* m*-dimensional manifold. Suppose* $(\mathsf{U};\, x^1, \ldots, x^m)$ *is a local coordinate chart. Let* $X, Y \in \Gamma^\infty(\mathsf{TM})$ *given locally by,*

$$X = \sum_{i=1}^{m} X^i \frac{\partial}{\partial x^i},$$
$$Y = \sum_{i=1}^{m} Y^i \frac{\partial}{\partial x^i}.$$

The smooth vector field $[X, Y]$ *is locally — on* U *— given by,*

$$[X, Y] = \sum_{i=1}^{m} \left[\sum_{j=1}^{m} X^i \left(\frac{\partial Y^j}{\partial x^i} \right) - Y^i \left(\frac{\partial X^j}{\partial x^i} \right) \right] \frac{\partial}{\partial x^i}$$

The set $\Gamma^\infty(\mathsf{TM})$, viewed as a real vector space, together with the Lie bracket constitutes a (real) Lie algebra. Given smooth distributions $\mathscr{D}_1, \mathscr{D}_2 \subseteq \mathsf{TM}$, their Lie bracket $[\mathscr{D}_1, \mathscr{D}_2] \subseteq \mathsf{TM}$ is defined pointwise by,

$$[\mathscr{D}_1, \mathscr{D}_2]_p := \{[X, Y]_p \colon X \in \Gamma^\infty(\mathscr{D}_1), Y \in \Gamma^\infty(\mathscr{D}_2)\}.$$

Generally, $[\mathscr{D}_1, \mathscr{D}_1] \nsubseteq \mathscr{D}_1$. A smooth distribution is *involutive* if $[\mathscr{D}_1, \mathscr{D}_1] \subseteq \mathscr{D}_1$.

When a distribution $\mathscr{D} \subseteq \mathsf{TM}$ is not involutive, there exists a smallest such involutive distribution $\mathrm{inv}(\mathscr{D}) \supseteq \mathscr{D}$ called the *involutive closure* of $\mathscr{D}$. The existence of the involutive closure is ensured by an argument utilizing Zorn's Lemma (see [10, Lemma 3.9.4]).

Next we define an object which is algebraically dual to a vector field. A *differential one-form* on an open subset $\mathsf{U} \subseteq \mathbb{R}^n$ is a function ω which assigns, to each $p \in \mathsf{U}$, an element ω_p of $\mathsf{T}_p^*\mathsf{U}$. Observe that at any given fixed point $p \in \mathsf{M}$, ω_p is a one-form over the vector space of tangent vectors at p. The differential one-form is smooth if, for every $X \in \Gamma^\infty(\mathsf{T}\mathbb{R}^n)$, the function $\omega(X) : \mathbb{R}^n \to \mathbb{R}$ defined pointwise by $\omega(X)(p) := \omega_p(X_p)$, is smooth. This again has an obvious generalization to embedded submanifolds of $\mathbb{R}^n$.

Definition A.5.6 (Smooth One-Forms) . Let M be a manifold. A *differential one-form* on M is a function ω which assigns to each $p \in \mathsf{M}$ an element ω_p of $\mathsf{T}_p^*\mathsf{M}$. The one-form is smooth if, for all $X \in \Gamma^\infty(\mathsf{TM})$, the function $\omega(X) : \mathsf{M} \to \mathbb{R}$ defined by $\omega(X)(p) := \omega_p(X_p)$ is smooth. Such a form ω is simply said to be a *smooth one-form*.

Again, an equivalent characterization of smooth one-forms can be given using the notion of smooth sections. Like before, we first construct a vector bundle of the related vector space.

Definition A.5.7 (Cotangent Bundle) . Let M be an m-dimensional submanifold of $\mathbb{R}^n$ and let $p \in \mathsf{M}$. The *cotangent bundle of* M is the disjoint union,

$$\mathsf{T}^*\mathsf{M} := \bigsqcup_{p \in \mathsf{M}} \mathsf{T}_p^*\mathsf{M},$$

and can be made into a vector bundle over M of rank m.

A smooth one-form $\omega : \mathsf{M} \to \mathsf{T}^*\mathsf{M}$ is simply a smooth section of the cotangent bundle. It is for this reason that we let $\Gamma^\infty(\mathsf{T}^*\mathsf{M})$ denote the set of all smooth one-forms on a manifold M. The set $\Gamma^\infty(\mathsf{T}^*\mathsf{M})$ can also be viewed as a module over the ring of smooth functions $C^\infty(\mathsf{M})$ like the space of vector fields $\Gamma^\infty(\mathsf{TM})$. We leave the discussion of this to a later section.

In Section A.4, we constructed pointwise a graded algebra of forms over the vector space of tangent vectors. It turns out we can construct a graded algebra over the space of differential forms as well, defining higher order differential forms, by leveraging the algebraic structure available at any fixed point $p \in \mathsf{M}$.

Definition A.5.8 (Smooth k-Forms) . A *differential k-form* is a function ω which assigns to $p \in \mathsf{M}$ an element of $\Lambda^k(\mathsf{T}_p\mathsf{M})$. The differential k-form is *smooth* if, for all $X^1, \ldots, X^k \in \Gamma^\infty(\mathsf{TM})$, the function $\omega(X^1, \ldots, X^k) : \mathsf{M} \to \mathbb{R}$ defined pointwise by

$$\omega(X^1, \ldots, X^k)(p) \coloneqq \omega_p(X^1_p, \ldots, X^k_p)$$

is smooth. The set of differential k-forms on M is denoted $\Gamma^\infty(\Lambda^k(\mathsf{TM}))$ and the set of all differential forms is denoted $\Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{M}))$.

The algebra of differential forms can be equipped with an operation that takes a smooth k-form and produces a smooth $(k+1)$-form, called the exterior derivative. To start, we define the differential of a smooth function. If $(\mathsf{U};\, x^1, \ldots, x^m)$ is a chart of M, then for each $p \in \mathsf{U}$ the basis of vectors for $\mathsf{T}^*_p\mathsf{M}$ induced by the chart is denoted by $\mathrm{d}x^1_p, \ldots, \mathrm{d}x^m_p$. Let $\varphi : \mathsf{U} \to \mathbb{R}^m$, $q \mapsto (x^1(q), \ldots, x^m(q))$. The exterior derivative of $H \in C^\infty(\mathsf{M})$ at p is the smooth one-form $\mathrm{d}H \in \Gamma^\infty(\mathsf{T}^*\mathsf{M})$ given pointwise by the expression

$$\mathrm{d}H_p = \sum_{k=1}^{m} \left.\frac{\partial\left(H \circ \varphi^{-1}\right)}{\partial x^k}\right|_{\varphi(p)} \mathrm{d}x^k_p.$$

This definition gives a construction for the smooth one-form $\mathrm{d}H$ on the open set U. We can then define, recursively, the exterior derivative operator on the set of differential forms.

Definition A.5.9 (Exterior Derivative) . The *exterior derivative* of a smooth k-form written in components as

$$F = \sum_{j \in J} F^j \,\mathrm{d}x^{j_1} \wedge \ldots \wedge \mathrm{d}x^{j_k},$$

where $J \subset \{1, \ldots, m\}^k$ is a set of multi-indices and $F^j : \mathsf{M} \to \mathbb{R}$ are smooth functions, is given by the smooth $(k+1)$-form

$$\mathrm{d}F = \sum_{j \in J} \mathrm{d}F^j \wedge \mathrm{d}x^{j_1} \wedge \ldots \wedge \mathrm{d}x^{j_k}.$$

Exterior differentiation observes a number of important properties. For one, let F be a smooth k-form and G a smooth ℓ-form. Then,

$$\mathrm{d}(F \wedge G) = \mathrm{d}F \wedge G + (-1)^k F \wedge \mathrm{d}G.$$

Moreover, by the symmetry of mixed partial derivatives, we have that $\mathrm{d}(\mathrm{d}F) = 0$ for *any* smooth form F. Another useful property of the exterior derivative is described in the following proposition.

Proposition A.5.10 ([31, Proposition 14.32])**.** *Let* $k \geq 0$, $X^1, \ldots, X^{k+1} \in \Gamma^\infty(\mathsf{TM})$ *and* $F \in \Gamma^\infty(\Lambda^k(\mathsf{TM}))$. *Then we have*

$$\begin{aligned}\mathrm{d}F\left(X^1, \ldots X^{k+1}\right) = \sum_{\ell=1}^{k} X^\ell \left(F(X^1, \ldots, \widehat{X}^\ell, \ldots, X^{k+1})\right) \\ + \sum_{0 \leq i < j \leq k} (-1)^{i+j} F([X^i, X^j], X^1, \ldots, \widehat{X}^i, \ldots, \widehat{X}^j, \ldots, X^{k+1}),\end{aligned}$$

where $\widehat{X}^i$ *denotes that the argument is deleted.*

Finally, we close this section with a useful fact relating the smooth forms between manifolds. Let $H : \mathsf{M} \to \mathsf{N}$ be a smooth map.

Definition A.5.11 . A vector field X on M and a vector field Y on N are *H-related* if, for all $p \in \mathsf{M}$,

$$\mathbf{D}H|_p \, X_p = Y_{H(p)}.$$

If $H : \mathsf{M} \to \mathsf{N}$ is a *diffeomorphism*, then to each $X \in \Gamma^\infty(\mathsf{TM})$ we can uniquely define an H-related vector field $Y \in \Gamma^\infty(\mathsf{TN})$ by letting, for each $q \in \mathsf{N}$,

$$Y_q \coloneqq \mathbf{D}H|_{H^{-1}(q)} \, X_{H^{-1}(q)}. \tag{A.2}$$

The unique H-related vector field defined by (A.2) is called the *pushforward* of X by H and is denoted by $H_* X$. A similar, and arguably nicer, statement can be made for smooth one-forms and, even more generally, smooth forms. In particular, we can relax the requirement that H be a diffeomorphism.

Let $H : \mathsf{M} \to \mathsf{N}$ be an *arbitrary* smooth map between manifolds. For each $p \in \mathsf{M}$ let $(\mathbf{D}H|_p)^* : \mathsf{T}^*_{H(p)}\mathsf{N} \to \mathsf{T}^*_p\mathsf{M}$ denote the dual map of $\mathbf{D}H|_p : \mathsf{T}_p\mathsf{M} \to \mathsf{T}_{H(p)}\mathsf{N}$. Then for any smooth one-form ω on N, and any $p \in \mathsf{M}$, define a smooth one-form $H^*\omega$ on M pointwise by

$$(H^*\omega)_p \coloneqq \left(\mathbf{D}H|_p\right)^* \omega_{H(p)}. \tag{A.3}$$

In particular, using the definition of the dual map, we have that for each $X_p \in \mathsf{T}_p\mathsf{M}$,

$$(H^*\omega)_p(X_p) = \omega_{H(p)} \left(\mathbf{D}H|_p \, X_p\right).$$

The unique, smooth one-form on M defined by (A.3) is called the *pullback* of ω by H.

This pullback extends naturally to arbitrary smooth forms. Given a smooth k-form ω we define the *pullback of* ω as the smooth k-form $H^*\omega$ on M that is given, for all $X^1, \dots, X^k \in \Gamma^\infty(\mathsf{TM})$ and $p \in \mathsf{M}$, by the expression

$$(H^*\omega)_p(X^1, \dots, X^k) \coloneqq \left(\omega_{H(p)}\right) \left(\mathbf{D}H|_p \, X^1_p, \dots, \mathbf{D}H|_p \, X^k_p\right). \tag{A.4}$$

The map $H^* : \Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{N})) \to \Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{M}))$, defined using (A.4), is also called the *pullback map* and can be seen as simply an extension of the pullback map (A.3). Remarkably, the next theorem shows that this map acts naturally with respect to the exterior derivative and, as such, is a homomorphism of algebras.

Theorem A.5.12 ([1, Theorem 6.4.4])**.** *Let $H : \mathsf{M} \to \mathsf{N}$ be a smooth map between manifolds. Then the pullback map $H^* : \Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{N})) \to \Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{M}))$ satisfies for all ω, $\beta \in \Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{N}))$,*

(i) $H^*(\omega \wedge \beta) = H^*\omega \wedge H^*\beta$ *and,*

(ii) $H^*(\mathrm{d}\omega) = \mathrm{d}\left(H^*\omega\right).$

A.6 Exterior Differential Systems

Recall that vector spaces at points can be bundled together to form what is known as a vector bundle. This can be done first by individually bundling the vector spaces of k-forms

$\Lambda^k(\mathsf{T}^*\mathsf{M}) \coloneqq \sqcup_{p\in\mathsf{M}}\Lambda^k(\mathsf{T}^*\mathsf{M}p)$. Then, take the direct sum of these *manifolds* to produce $\Lambda(\mathsf{T}^*\mathsf{M}) \coloneqq \oplus_{k=1}^{m}\Lambda^k(\mathsf{T}^*\mathsf{M})$. Smooth sections of this vector bundle are smooth forms over M. The set of smooth forms $\Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{M}))$ can be viewed as a graded algebra over the ring of smooth functions $C^\infty(\mathsf{M})$. A graded subalgebra $\mathcal{I} \subseteq \Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{M}))$ is called an *ideal* over the graded subalgebra if, for all $\omega \in \mathcal{I}$ and $\beta \in \Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{M}))$, $\beta \wedge \omega \in \mathcal{I}$.

Let $\mathcal{I} \subseteq \Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{M}))$ be an ideal over the graded subalgebra of smooth forms. It is said to be *locally, simply, finitely generated* if for any point $p \in \mathsf{M}$ there exists an open set U containing p and a finite set of smooth one-forms $\omega^1, \ldots, \omega^\ell \in \Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{U}))$ such that any smooth form $\beta \in \mathcal{I}$ can be written locally as,

$$\beta|_{\mathsf{U}} = \sum_{i=1}^{\ell} c \wedge \omega^i.$$

The smooth one-forms $\{\omega^1, \ldots, \omega^\ell\}$ are called local generators. The ideal $\mathcal{I}$ is *non-degenerately generated* if the local generators are linearly independent pointwise. Notice that a locally, simply, finitely, non-degenerately generated ideal $\mathcal{I}$ can be associated with a smooth and regular distribution $\mathscr{I} \subseteq \mathsf{T}^*\mathsf{M}$. Equipped with this notion, we can define what an exterior differential system is. For our purposes, the following definition suffices.

Definition A.6.1 (Exterior Differential System). An *exterior differential system* is a locally, simply, finitely generated ideal $\mathcal{I} \subseteq \Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{M}))$ over the graded algebra of smooth forms $\Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{M}))$.

The term "system" evokes a desire for a solution. The next definition describes what precisely is considered to be a solution to an exterior differential system.

Definition A.6.2. An *integral submanifold* to an exterior differential system $\mathcal{I}$ is an immersed submanifold $\mathsf{N} \subseteq \mathsf{M}$ whose immersion $H : \mathsf{N} \to \mathsf{M}$ is such that $H^*\omega = 0$ for all $\omega \in \mathcal{I}$.

Often we seek an integral submanifold to an exterior differential system that passes through a particular point $p_0 \in \mathsf{M}$. These are stated to be the integral submanifolds passing through p_0. The one-point manifold $\mathsf{N} = \{p_0\} \subseteq \mathsf{M}$ is an integral manifold to *every* exterior differential system since the pullback maps all inputs to 0. In fact, given any non-trivial integral submanifold N to $\mathcal{I}$, any submanifold $\overline{\mathsf{N}} \subseteq \mathsf{N}$ is also an integral submanifold to $\mathcal{I}$. This fact motivates the search for *maximal* integral manifold passing through the point p_0. When it exists, the maximal integral submanifold to an exterior differential system is considered the solution. At times it is convenient to specifically ask for an integral manifold of a given dimension where it is clear what the upper bound on dimension is. To build some intuition, let us consider a few examples.

Example A.1 (ODE as EDS). Let $\mathsf{M} = \mathbb{R}^3$ with the standard smooth manifold structure and coordinates (t, x^1, x^2). Consider the system of ordinary differential equations

$$\frac{dx^1}{dt} = x^2,$$
$$\frac{dx^2}{dt} = -x^1.$$

Tied to this is the exterior differential system $\mathcal{I} \coloneqq \langle \mathrm{d}x^1 - x^2\,\mathrm{d}t, \mathrm{d}x^2 + x^1\,\mathrm{d}t \rangle$. Fix $p_0 = (t_0, x_0^1, x_0^2) \in \mathsf{M}$. Define the submanifold

$$\mathsf{N}_{p_0} \coloneqq \left\{ (t - t_0, x_0^1 \cos(t - t_0) + x_0^2 \sin(t - t_0), -x_0^1 \sin(t - t_0) + x_0^2 \cos(t - t_0)) \colon t \in \mathbb{R} \right\}.$$

The submanifold N_{p_0} has an immersion given by $H : \mathsf{N}_{p_0} \to \mathbb{R}^3$, $t \mapsto (x^1, x^2, x^3)$ with pullback, for all a, b, $c \in \mathbb{R}$,

$$H^*(a\,\mathrm{d}x^1 + b\,\mathrm{d}x^2 + c\,\mathrm{d}t) \coloneqq \left[c + (a\,x_0^2 - b\,x_0^1) \cos(t - t_0) - (a\,x_0^1 + b\,x_0^2) \sin(t - t_0) \right] \mathrm{d}t.$$

It can also be shown that,

$$H^*(\mathrm{d}x^1 - x^2\,\mathrm{d}t) = 0, \quad H^*(\mathrm{d}x^2 + x^1\,\mathrm{d}t) = 0.$$

It follows that N_{p_0} is an integral submanifold of $\mathcal{I}$ passing through p_0. Moreover, it is the maximal integral submanifold passing through p_0 since a 2-dimensional submanifold of $\mathbb{R}^3$ would limit the kernel of H^* to a maximum dimension of 1. Since $\mathrm{d}x^1 - x^2\,\mathrm{d}t$ and $\mathrm{d}x^2 + x^1\,\mathrm{d}t$ are linearly independent, we have that the maximal integral submanifold must have dimension 1. ◀

Example A.2 . Let $\mathsf{M} = \mathbb{R}^3$ with the standard smooth manifold structure and coordinates (x^1, x^2, x^3). Define the exterior differential system $\mathcal{I} \coloneqq \langle x^3\,\mathrm{d}x^1 - x^2\,\mathrm{d}x^2\rangle \subseteq \Gamma^\infty(\Lambda\mathsf{T}^*\mathsf{M})$. Based on the intuition of Example A.1, one would expect to have a maximal integral submanifold of dimension 2. Suppose there exists an immersed integral submanifold N of dimension 2 passing through $p_0 \in \mathsf{M}$ with immersion $H : \mathsf{N} \to \mathsf{M}$. Then the one-form $\omega \coloneqq x^3\,\mathrm{d}x^1 - x^2\,\mathrm{d}x^2 \in \mathcal{I}$ satisfies $H^*\omega = 0$. Additionally, we have that the exterior derivative $\mathrm{d}(H^*\omega) = 0$. By Theorem A.5.12, $H^*(\mathrm{d}\omega) = 0$. As a result

$$0 = H^*\mathrm{d}\omega = H^*\left(\mathrm{d}x^3 \wedge \mathrm{d}x^2\right).$$

Again, use Theorem A.5.12 to find

$$0 = H^*\mathrm{d}x^2 \wedge H^*\mathrm{d}x^3.$$

This suggests that there exist non-zero constants $a, b \in \mathbb{R}$ so that at the point $p_0 \in \mathsf{M}$,

$$0 = \left(\mathbf{D}H|_{p_0}\right)^*\left(a\,\mathrm{d}x^2_{p_0} + b\,\mathrm{d}x^3_{p_0}\right).$$

Suppose $p_0^3 \neq 0$. Then the covectors $p_0^3\,\mathrm{d}x^1 - p_0^2\,\mathrm{d}x^2$ and $a\,\mathrm{d}x^2_{p_0} + b\,\mathrm{d}x^3_{p_0}$ are linearly independent and in the kernel of $\mathbf{D}H|_{p_0}^*$. This is a contradiction since we have shown that $\mathbf{D}H|_{p_0}^*$ has a kernel of dimension 2 at p_0 but the rank of $\mathbf{D}H|_{p_0}^*$ has to be at least two (H is an injective map and the dimension of N is 2). As a result, there cannot exist an integral submanifold of dimension 2 passing through points where $p_0^3 \neq 0$.

When $p_0^3 = b = 0$, the above contradiction does not appear. However, an integral submanifold of dimension 2 passing through such a point also fails to exist since such an integral submanifold must pass through points of the form $q \coloneqq (q^1, q^2, q^3)$ where $q^3 \neq 0$. The above contradiction then appears at this point q. It follows that there are no 2-dimensional integral submanifolds to $\mathcal{I}$ through any point of $\mathbb{R}^3$.

On the other hand, if N is 1-dimensional there is no contradiction, but also there is no unique integral submanifold of dimension 1 passing through any fixed point $p_0 \in \mathsf{M}$. ◀

Example A.1 and A.2 demonstrate that there is an additional property about the ideal

$\mathcal{I}$ and the forms that generate it that is required to ensure a maximal integral submanifold exists through some point p_0. This property, we define now, is that the ideal in Example A.1 is said to be a differential ideal. This property is the key requirement in the existence and uniqueness of the maximal integral manifold.

Definition A.6.3 (Differential Ideal) . An ideal $\mathcal{I}$ is said to be *differentially closed* if it is closed under the exterior derivative operator $\mathrm{d} : \Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{M})) \to \Gamma^\infty(\Lambda(\mathsf{T}^*\mathsf{M}))$. A differentially closed ideal $\mathcal{I}$ is called a *differential ideal*.

Differential ideals play a critical role in guaranteeing the uniqueness of the maximal integral submanifolds when $\mathcal{I}$ is generated by one-forms. Intuitively, this is because the exterior derivative cannot introduce new, unaccounted for constraints as it did in Example A.2. The standard result that relates differential ideals to the uniqueness of the maximal integral manifold is Frobenius's Theorem, which we state now.

Theorem A.6.4 (Frobenius's Theorem (Differential Forms))**.** *Let $\mathcal{I}$ be a differential ideal on an m-dimensional manifold M that is simply, finitely and non-degenerately generated by r smooth 1-forms. Then, in a sufficiently small neighbourhood of a point $p_0 \in \mathsf{M}$, there exists coordinates $y^1, \ldots, y^m$ for M so that*

$$\mathcal{I} = \langle \mathrm{d}y^1, \ldots, \mathrm{d}y^{m-r} \rangle.$$

It is important to take a moment to observe the significance of this result. Frobenius's theorem shows that, given a differentially closed ideal generated by smooth 1-forms, there exists a local coordinate chart $\varphi : \mathsf{U} \to \mathbb{R}^m$, $p \mapsto (y^1(p), \ldots, y^m(p))$, where the differentials of the first $m - r$ coordinates generate the ideal. Immediately, one can construct the family of integral submanifolds in U by letting $c \in \mathbb{R}^{m-r}$ and defining

$$\mathsf{N}_c \coloneqq \varphi^{-1}\left(\left\{(c^1, \ldots, c^{m-r}, y^{m-r+1}(p), \ldots, y^m(p)) \colon p \in \mathsf{U}\right\}\right).$$

The integral submanifold passing through p_0 is given by,

$$\mathsf{N}_{(y^1(p_0), \ldots, y^{m-r}(p_0))} \coloneqq \varphi^{-1}\left(\left\{(y^1(p_0), \ldots, y^{m-r}(p_0), y^{m-r+1}(p), \ldots, y^m(p)) \colon p \in \mathsf{U}\right\}\right).$$

The proof of Theorem A.6.4 can be found in [9, Theorem 1.1]. The following example illustrates how the proof of Frobenius' theorem may be used. The method is laborious in comparison to traditional methods, but its generality makes it a powerful tool.

Example A.3 . Let $\mathsf{M} = \mathbb{R}^3$ with the standard smooth manifold structure with coordinates x^1, x^2 and x^3. Consider the simply, finitely and non-degenerately generated ideal

$$\mathcal{I} \coloneqq \langle e^{x^2} \mathrm{d}x^1 + e^{-x^1} \mathrm{d}x^2 \rangle \eqqcolon \langle \omega^1 \rangle.$$

We would like to find the maximal integral manifolds, of dimension 2, to this ideal. First we check that $\mathcal{I}$ is a differential ideal. The exterior derivative of the generator ω^1 is

$$-\left(e^{-x^1} + e^{x^2}\right) \mathrm{d}x^1 \wedge \mathrm{d}x^2 = -\frac{e^{-x^1} + e^{x^2}}{e^{-x^1}} \mathrm{d}x^1 \wedge \omega^1 \in \mathcal{I}.$$

Indeed $\mathcal{I}$ is a differential ideal, so we know there exists coordinates where the integral manifolds to $\mathcal{I}$ are level sets of the last coordinate. We can compute these coordinates by running through the proof of Frobenius's theorem. First fix a base point $q \in \mathsf{M}$. Now, since $r = 2$, we have to first construct a new *differential* ideal, containing our ideal, that has a codimension of 1. Define $\mathcal{J} \coloneqq \langle \omega^1, \mathrm{d}x^3 \rangle$. This is a differential ideal containing $\mathcal{I}$ with codimension $r = 1$. We can use the base case of Frobenius's theorem to construct coordinates where $\mathcal{J} = \langle \mathrm{d}a^2, \mathrm{d}a^3 \rangle$. The distribution annihilated by $\mathcal{J}$ is given pointwise by

$$\mathscr{D}_x = \operatorname{span}_{\mathbb{R}} \left\{ e^{-x^1} \left. \frac{\partial}{\partial x^1} \right|_x - e^{x^2} \left. \frac{\partial}{\partial x^2} \right|_x \right\}.$$

This is a smooth distribution with a local flow $\Theta : D \subseteq \mathbb{R} \times \mathsf{M} \to \mathsf{M}$, given by

$$\Theta(t, q) = \left(\ln\left(t + e^{q^1}\right), -\ln\left(t + e^{q^2}\right), q^3 \right).$$

From Flow Box Theorem (see [11, Theorem 4.4.1]), we can define a diffeomorphism Φ by

$$\Phi(a^1, a^2, a^3) \coloneqq \Theta(a^1, (q^1, a^2, a^3))$$

where the open set this is defined on is determined by the base point q^1. The smooth inverse is given by

$$\Phi^{-1}(x^1, x^2, x^3) = \left(e^{x^1} - 1, \ln\left(e^{-x^2} - e^{x^1} + 1\right), x^3\right).$$

In the new coordinates we have that $\mathcal{J} = \langle \mathrm{d}a^2, \mathrm{d}a^3 \rangle$. By chance, $\Phi^*\omega^1 = h(a)\mathrm{d}a^2$ where $h(a) \neq 0$ on an open set containing $h(q)$. We can immediately write $\mathcal{I} = \langle \mathrm{d}a^2 \rangle$ and this then gives an expression for the integral manifolds to $\mathcal{I}$. The integral manifold passing through $q \in \mathsf{M}$ is given locally by

$$\mathsf{N}_q \coloneqq \left\{ x \in M \;\middle|\; e^{-x^2} - e^{x^1} = e^{-q^2} - e^{q^1} \right\}.$$

◀

Not all ideals $\mathcal{J}$ are differentially closed. Some of these times, it is useful to find the largest differential ideal contained in $\mathcal{J}$. This construction can be viewed as a dual to the involutive closure $\operatorname{inv}(\mathcal{D})$ of a submodule $\mathcal{D}$ of vector fields. We denote the largest differentially closed ideal contained in $\mathcal{J}$ by $\mathcal{J}^{(\infty)} \subseteq \mathcal{J}$. Like $\operatorname{inv}(\mathcal{D})$, the existence of $\mathcal{J}^{(\infty)}$ object is ensured by an argument utilizing Zorn's Lemma. The proof follows closely that of [10, Lemma 3.9.4].

Appendix B

Technical Proofs

This appendix presents the proofs of a number of technical supporting results used throughout the thesis that were involved and not critical to the discussion at hand. The first few of these results come from Section 3.2.

Lemma (Lemma 3.2.1)**.** *Suppose Assumptions 2.2.3 and 2.2.11 hold. There exists an open set on* M *containing* p_0 *such that on this open set, for any* $k \in \{1, \ldots, n - n^\star\}$,

$$\mathcal{I}^{(k)} = \operatorname{ann}\left(\mathcal{D}^{(0)} + \mathcal{S}^{(k-1)}\right).$$

Proof. By Assumptions 2.2.3 and 2.2.11, the ideals $\mathcal{I}^{(k)}$ are simply, finitely, non-degenerately generated and the submodules $\mathcal{D}^{(0)}$ and $\mathcal{S}^{(k-1)}$ are finitely, non-degenerately generated. The latter fact ensures that the annihilators $\operatorname{ann}(\mathcal{D}^{(0)})$ and $\operatorname{ann}(\mathcal{S}^{(k-1)})$ are (locally) simply, finitely generated ideals. It therefore suffices to verify that the two ideals are locally equal in their space of one-forms.

Base Case ($k = 1$): Pick $\omega \in \mathcal{I}^{(1)}$. Since $\mathcal{I}^{(1)} \subset \mathcal{I}^{(0)}$, $\omega \in \mathcal{D}^{(0)}$. It suffices to show that ω annihilates vector fields in $\mathcal{S}^{(0)} = \mathcal{G}^{(0)}$. Let $1 \leq i \leq m$ and consider $\omega(g_i)$. Observe that,

$$\omega(g_i) = \omega\left(\left[F, \frac{\partial}{\partial u^i}\right]\right).$$

Use Proposition A.5.10 to find,

$$\omega(g_i) = F\left(\omega\left(\frac{\partial}{\partial u^i}\right)\right) + \frac{\partial}{\partial u^i}\left(\omega\left(F\right)\right) - \mathrm{d}\omega\left(F, \frac{\partial}{\partial u^i}\right).$$

Since $\omega \in \mathcal{I}^{(0)} = \operatorname{ann}(\mathcal{D}^{(0)})$,

$$\omega(g_i) = -\mathrm{d}\omega\left(F, \frac{\partial}{\partial u^i}\right).$$

Use (2.12) and $\omega \in \mathcal{I}^{(1)}$ to conclude that $\mathrm{d}\omega \in \mathcal{I}^{(0)}$. Therefore,

$$\omega(g_i) = 0.$$

This was shown for arbitrary $1 \leq i \leq m$ and so $\omega \in \mathcal{G}^{(0)} = \mathcal{S}^{(0)}$. Therefore $\mathcal{I}^{(1)} \subseteq \operatorname{ann}(\mathcal{D}^{(0)} + \mathcal{S}^{(0)})$. Now pick $\omega \in \operatorname{ann}(\mathcal{D}^{(0)} + \mathcal{S}^{(0)})$. Since $\omega \in \mathcal{D}^{(0)}$, $\omega \in \mathcal{I}^{(0)}$. We now show that ω is differentially closed in $\mathcal{I}^{(0)}$. Pick arbitrary vector fields $X, Y \in \mathcal{D}^{(0)}$ and consider $\mathrm{d}\omega(X, Y)$. By Cartan's formula,

$$\mathrm{d}\omega(X, Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y]).$$

Since $\omega \in \operatorname{ann}(\mathcal{D}^{(0)})$,

$$\mathrm{d}\omega(X, Y) = -\omega([X, Y]).$$

Now observe that $[X, Y] \subseteq \mathcal{D}^{(0)} + \mathcal{G}^{(0)}$ since either both $X, Y \in \mathcal{U}$ — which is involutive — or we have at least one of X or $Y = F$. In the latter case, the Lie bracket produces vector fields g_i as demonstrated earlier in the base case. As a result, $\omega([X, Y]) = 0$. Therefore $\mathrm{d}\omega(X, Y) = 0$ and we can conclude that $\omega \in \mathcal{I}^{(1)}$. This completes the base case.

Induction ($k \geq 1$): Suppose, by way of induction, that for some $k \geq 1$

$$\mathcal{I}^{(k)} = \operatorname{ann}\left(\mathcal{D}^{(0)} + \mathcal{S}^{(k-1)}\right). \tag{B.1}$$

First we show that $\mathcal{I}^{(k+1)} \subseteq \operatorname{ann}(\mathcal{D}^{(0)} + \mathcal{S}^{(k)})$. Pick a smooth one-form $\omega \in \mathcal{I}^{(k+1)}$. If the one-form ω is in

$$\operatorname{ann}(\mathcal{D}^{(0)}) \cap \operatorname{ann}(\mathcal{S}^{(k-1)}) \cap \operatorname{ann}\left(\left[\mathcal{S}^{(k-1)}, \mathcal{S}^{(k-1)}\right]\right) \cap \operatorname{ann}(\mathcal{G}^{(k)}),$$

then ω lives in

$$\operatorname{ann}\left(\mathcal{D}^{(0)} + \mathcal{S}^{(k-1)} + \left[\mathcal{S}^{(k-1)}, \mathcal{S}^{(k-1)}\right] + \mathcal{G}^{(k)}\right).$$

It follows by (2.4) that $\omega \in \operatorname{ann}(\mathcal{D}^{(0)} + \mathcal{S}^{(k)})$. To this end, by definition $\omega \in \mathcal{I}^{(k)}$ and therefore by the induction hypothesis (B.1)

$$\omega \in \operatorname{ann}\left(\mathcal{D}^{(0)} + \mathcal{S}^{(k-1)}\right). \tag{B.2}$$

Now let X_1, $X_2 \in \mathcal{S}^{(k-1)}$ be arbitrary smooth vector fields and let $X = [X_1, X_2]$. Evaluating the differential form ω on the vector field X and using Proposition A.5.10 gives

$$\begin{aligned}\omega(X) &= \omega([X_1, X_2]),\\ &= X_1(\omega(X_2)) - X_2(\omega(X_1)) - \mathrm{d}\omega(X_1, X_2).\end{aligned}$$

By (B.2), $\omega \in \operatorname{ann}(\mathcal{S}^{(k-1)})$ so

$$\omega(X) = -\mathrm{d}\omega(X_1, X_2).$$

As $\omega \in \mathcal{I}^{(k+1)}$ it follows that ω is closed in $\mathcal{I}^{(k)}$ under exterior derivative: $\mathrm{d}\omega \in \mathcal{I}^{(k)}$. Then, by the inductive hypothesis (B.1), $\mathrm{d}\omega \in \operatorname{ann}(\mathcal{D}^{(0)} + \mathcal{S}^{(k-1)})$. So it follows that $\mathrm{d}\omega \in \operatorname{ann}(\mathcal{S}^{(k-1)})$ and we arrive at $\omega(X) = 0$. As X was a general vector field of $[\mathcal{S}^{(k-1)}, \mathcal{S}^{(k-1)}]$ we have established that $\omega \in \operatorname{ann}([\mathcal{S}^{(k-1)}, \mathcal{S}^{(k-1)}])$. Combining with (B.2) gives

$$\omega \in \operatorname{ann}\left(\mathcal{D}^{(0)} + \mathcal{S}^{(k-1)} + [\mathcal{S}^{(k-1)}, \mathcal{S}^{(k-1)}]\right).$$

It remains to show that $\omega \in \operatorname{ann}(\mathcal{G}^{(k)})$. It follows from (2.3) that $\mathcal{G}^{(k)} = \mathcal{G}^{(k-1)} + [\mathcal{D}^{(0)}, \mathcal{G}^{(k-1)}]$. However, by (B.2), $\omega \in \operatorname{ann}(\mathcal{G}^{(k-1)})$. As a result it suffices to check that $\omega \in \operatorname{ann}([\mathcal{D}^{(0)}, \mathcal{G}^{(k-1)}])$. Let $X \in \mathcal{G}^{(k-1)}$ and $Y \in \mathcal{D}^{(0)}$ be smooth vector fields. Evaluating $\omega([Y, X])$ and using Proposition A.5.10 gives

$$\omega([Y, X]) = -\mathrm{d}\omega(Y, X) - X(\omega(Y)) + Y(\omega(X)).$$

By the inductive hypothesis (B.1) and (B.2) $\omega \in \operatorname{ann}(\mathcal{D}^{(0)} + \mathcal{G}^{(k-1)})$, so

$$\omega([Y, X]) = -\mathrm{d}\omega(Y, X).$$

But we also know that $\mathrm{d}\omega \in \operatorname{ann}(\mathcal{D}^{(0)} + \mathcal{S}^{(k-1)})$ and $\mathcal{G}^{(k-1)} \subseteq \mathcal{S}^{(k-1)}$ therefore $\mathrm{d}\omega(Y, X) = 0$ and

$$\omega([Y, X]) = 0.$$

Since Y, X were arbitrary $\omega \in \operatorname{ann}(\mathcal{G}^{(k)})$. As a result

$$\omega \in \operatorname{ann}\left(\mathcal{D}^{(0)} + \mathcal{S}^{(k-1)} + [\mathcal{S}^{(k-1)}, \mathcal{S}^{(k-1)}] + \mathcal{G}^{(k)}\right).$$

Then using (2.4) we get

$$\omega \in \operatorname{ann}\left(\mathcal{D}^{(0)} + \mathcal{S}^{(k)}\right)$$

and we conclude that

$$\mathcal{I}^{(k+1)} \subseteq \operatorname{ann}\left(\mathcal{D}^{(0)} + \mathcal{S}^{(k)}\right).$$

To complete the proof we must establish the reverse inclusion $\mathcal{I}^{(k+1)} \supseteq \operatorname{ann}(\mathcal{D}^{(0)}+\mathcal{S}^{(k)})$. Pick any smooth one-form $\omega \in \operatorname{ann}(\mathcal{D}^{(0)} + \mathcal{S}^{(k)})$. Since $\mathcal{S}^{(k-1)} \subseteq \mathcal{S}^{(k)}$, $\omega \in \operatorname{ann}(\mathcal{D}^{(0)} + \mathcal{S}^{(k-1)})$. It then follows by the inductive hypothesis that $\omega \in \mathcal{I}^{(k)}$. Consider the two-form $\mathrm{d}\omega$; observe that, since $\omega \in \mathcal{I}^{(k)}$, $\mathrm{d}\omega \in \mathcal{I}^{(k-1)}$. As $\mathcal{I}^{(k-1)} \subseteq \operatorname{ann}(\mathcal{D}^{(0)})$, $\mathrm{d}\omega \in \operatorname{ann}(\mathcal{D}^{(0)})$. Arbitrarily pick smooth vector fields X_1, $X_2 \in \mathcal{S}^{(k-1)}$ and observe that, by Proposition A.5.10, $\mathrm{d}\omega(X_1, X_2)$ equals

$$-\omega([X_1, X_2]) + X_1(\omega(X_2)) - X_2(\omega(X_1)).$$

Since $\omega \in \operatorname{ann}(\mathcal{S}^{(k-1)})$,

$$\mathrm{d}\omega(X_1, X_2) = -\omega([X_1, X_2]).$$

Observe that, by (2.4), $[X_1, X_2] \in [\mathcal{S}^{(k-1)}, \mathcal{S}^{(k-1)}] \subseteq \mathcal{S}^{(k)}$. Using this in combination with $\omega \in \operatorname{ann}(\mathcal{D}^{(0)} + \mathcal{S}^{(k)})$ it follows that

$$\mathrm{d}\omega(X_1, X_2) = 0,$$

But then $\mathrm{d}\omega \in \mathcal{I}^{(k)}$ by the inductive hypothesis. Since ω, $\mathrm{d}\omega \in \mathcal{I}^{(k)}$, we have that ω is closed under exterior derivative in $\mathcal{I}^{(k)}$ so we conclude that $\omega \in \mathcal{I}^{(k+1)}$. This establishes that $\mathcal{I}^{(k+1)}$ and $\operatorname{ann}(\mathcal{D}^{(0)} + \mathcal{S}^{(k)})$ are equal in their homogeneous component of one-forms and the proof is complete. □

Lemma (Lemma 3.2.2)**.** *Suppose Assumptions 2.2.3 and 2.2.11 hold. There exists an open set on* M *containing* p_0 *such that on this open set, for any* $\kappa \in \{1, \ldots, n - n^\star\}$,

$$\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle = \operatorname{ann}\left(\mathcal{U} \oplus \mathcal{S}^{(\kappa-1)}\right).$$

Proof. For any fixed $\kappa \in \mathbb{N}$, the inclusion

$$\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle \subseteq \operatorname{ann}\left(\mathcal{U} \oplus \mathcal{S}^{(\kappa-1)}\right),$$

follows directly from Lemma 3.2.1 and properties of annihilators. By Assumption 2.2.11, there exists an open set $\mathsf{U} \subseteq \mathsf{M}$ containing p_0 and a finite number of generators $X_1, \ldots, X_\ell \in \Gamma^\infty(\mathsf{TU})$ that non-degenerately generate,

$$\mathcal{S}^{(\kappa-1)} = \operatorname{span}_{C^\infty(\mathsf{U})} \{X_1, \ldots, X_\ell\}.$$

Using this, we know that at $p \in \mathsf{U}$, the associated codistribution satisfy,

$$\mathscr{U}_p + \mathscr{S}_p^{(\kappa-1)} = \operatorname{span}_{\mathbb{R}} \left\{ \left.\frac{\partial}{\partial u^1}\right|_p, \dots, \left.\frac{\partial}{\partial u^m}\right|_p, X_1|_p, \dots, X_\ell|_p \right\}.$$

Contrast that with,

$$\begin{aligned} \mathscr{D}_p^{(0)} + \mathscr{S}_p^{(\kappa-1)} &= \operatorname{span}_{\mathbb{R}} \{F|_p\} \oplus \mathcal{U}_p \oplus \mathcal{S}_p^{(\kappa-1)} \\ &= \operatorname{span}_{\mathbb{R}} \left\{ F|_p, \left.\frac{\partial}{\partial u^1}\right|_p, \dots, \left.\frac{\partial}{\partial u^m}\right|_p, X_1|_p, \dots, X_\ell|_p \right\}, \end{aligned}$$

and observe that

$$\dim\left(\mathscr{U}_p + \mathscr{S}_p^{(\kappa-1)}\right) = \dim\left(\mathscr{D}_p^{(0)} + \mathscr{S}_p^{(\kappa-1)}\right) - 1. \tag{B.3}$$

This implies that

$$\dim\left(\operatorname{ann}\left(\mathscr{U}_p + \mathscr{S}_p^{(\kappa-1)}\right)\right) = \dim\left(\operatorname{ann}\left(\mathscr{D}_p^{(0)} + \mathscr{S}_p^{(\kappa-1)}\right)\right) + 1$$

Using Lemma 3.2.1 find,

$$\dim\left(\operatorname{ann}\left(\mathscr{U}_p + \mathscr{S}_p^{(\kappa-1)}\right)\right) = \dim\left(\mathscr{I}^{(\kappa)}\right) + 1.$$

But,

$$\dim(\operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(\kappa)}, \mathrm{d}t_p\}) = \dim(\mathscr{I}_p^{(\kappa)}) + 1.$$

It follows that,

$$\dim\left(\operatorname{ann}\left(\mathscr{U}_p + \mathscr{S}_p^{(\kappa-1)}\right)\right) = \dim\left(\operatorname{span}_{\mathbb{R}}\{\mathscr{I}_p^{(\kappa)}, \mathrm{d}t_p\}\right).$$

□

Lemma (Lemma 2.2.12)**.** *Let $\kappa \in \mathbb{N}$. If Assumptions 2.2.3 and 2.2.11 hold, then*

$$\langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle^{(\infty)} \supseteq \operatorname{ann}\left(\Gamma^\infty(\mathscr{U} \oplus \operatorname{inv}(\mathscr{G}^{(\kappa-1)}))\right)$$

Proof. Fix $p \in \mathsf{M}$. Since $\mathrm{inv}(\mathscr{G}^{(\kappa-1)})$ is a smooth and regular distribution, so is the direct sum $\mathscr{U} \oplus \mathrm{inv}(\mathscr{G}^{(\kappa-1)})$. Invoke Frobenius's Theorem on an open neighbourhood $\mathsf{U} \subseteq \mathsf{M}$ of p to find a new local coordinate system (t, u, z) where, without loss of generality[1],

$$\mathscr{U}_q \oplus \mathrm{inv}(\mathscr{G}^{(\kappa-1)})_q = \mathrm{span}_{\mathbb{R}} \left\{ \left.\frac{\partial}{\partial u^1}\right|_q, \ldots, \left.\frac{\partial}{\partial u^m}\right|_q, \left.\frac{\partial}{\partial z^{\ell+1}}\right|_q, \ldots, \left.\frac{\partial}{\partial z^n}\right|_q \right\}, \qquad q \in \mathsf{U}.$$

It directly follows that

$$\mathrm{ann}\left(\Gamma^\infty(\mathscr{U} \oplus \mathrm{inv}(\mathscr{G}^{(\kappa-1)}))\right) = \langle \mathrm{d}t, \mathrm{d}z^1, \ldots, \mathrm{d}z^\ell \rangle.$$

Now we must show the containment

$$\mathrm{ann}\left(\Gamma^\infty(\mathscr{U} \oplus \mathrm{inv}(\mathscr{G}^{(\kappa-1)}))\right) \subseteq \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle^{(\infty)}.$$

Clearly $\mathrm{d}t \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle^{(\infty)}$. It suffices to show that $\mathrm{d}z^i \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle^{(\infty)}$ for $1 \leq i \leq \ell$. Fix $1 \leq i \leq \ell$ and consider $\mathrm{d}z^i \in \mathrm{ann}(\mathscr{U} \oplus \mathrm{inv}(\mathscr{G}^{(\kappa-1)}))$. By Lemma 2.1.1,

$$\mathrm{ann}\left(\Gamma^\infty(\mathscr{U} \oplus \mathrm{inv}(\mathscr{G}^{(\kappa-1)}))\right) = \mathrm{ann}\left(\Gamma^\infty(\mathscr{U} \oplus \mathrm{inv}(\mathscr{S}^{(\kappa-1)}))\right),$$

therefore

$$\mathrm{d}z^i \in \mathrm{ann}\left(\Gamma^\infty(\mathscr{U} \oplus \mathrm{inv}(\mathscr{S}^{(\kappa-1)}))\right).$$

Since $\mathscr{U} \oplus \mathrm{inv}(\mathscr{S}^{(\kappa-1)}) \supseteq \mathscr{U} \oplus \mathscr{S}^{(\kappa-1)}$,

$$\mathrm{d}z^i \in \mathrm{ann}\left(\Gamma^\infty(\mathscr{U} \oplus \mathscr{S}^{(\kappa-1)})\right).$$

By Lemma 3.2.2, $\mathrm{d}z^i \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle$. Since $\mathrm{d}z^i$ is closed,

$$\mathrm{d}z^i \in \langle \mathcal{I}^{(\kappa)}, \mathrm{d}t \rangle^{(\infty)}.$$

□

[1]The time variable t and control variables u need not change in this coordinate change since $\mathscr{U}$ is already generated by coordinate vector fields and because neither $\mathscr{U}$ nor $\mathscr{G}^{(\kappa-1)}$ are time-variant.