EXTERIOR DIFFERENTIAL SYSTEMS IN CONTROL AND ROBOTICS

GEORGE J. PAPPAS, JOHN LYGEROS, DAWN TILBURY, AND SHANKAR SASTRY

Contents

1. Introduction

The vast majority of the mathematically oriented literature in the areas of robotics and control has been heavily influenced by a differential geometric point of view. For nonlinear systems in particular, most of the research has concentrated on the analysis of the Lie algebras associated with controllability, reachability and observability. In recent years, however, a small but influential trend has begun in the literature on the use of other methods, such as differential algebra [9, 8, 10] and exterior differential systems [13, 11] for the analysis of nonlinear control systems and nonlinear implicit systems. In this paper we survey some key results from the theory of exterior differential systems and their application to current and challenging problems in robotics and control.

The area of exterior differential systems has a long history. The early theory in this area sprung from the work of Darboux, Lie, Engel, Pfaff, Carnot and Caratheodory on the structure of systems with non-integrable linear constraints on the velocities of their configuration variables, the so-called nonholonomic control systems (for a good development of this see [3]). This was followed by the work of Goursat and Cartan, which is considered to contain some of the finest achievements of the mid-part of this century on exterior differential systems. In parallel has been an effort to develop connections between exterior differential systems and the calculus of variations (see [13]).

Our attention was first attracted to exterior differential systems through their applications in path planning for nonholonomic control systems. Our initial results were for the problem of steering a car with trailers [34], [22], the so-called "parallel parking a car with N trailers problem." This involved the transformation of the system of nonholonomic rolling without slipping constraints on each pair of wheels into a canonical form, the so-called Goursat normal form. This program continued with another example, the parallel parking of a fire truck [5], which in turn was generalized to a multi-steering N trailer system. In [36] we showed how the multi-steering N trailer system could be converted into a generalized Goursat normal form, which was easy to steer. The full analysis of the system from the exterior differential systems point of view was made in [35].

In parallel with this activity in nonholonomic motion planning there has been considerable activity in the nonlinear control community on the problem of exactly linearizing a nonlinear control system using (possibly dynamic) state feedback and change of coordinates. The first results in this direction were necessary and sufficient conditions for exact linearization of a nonlinear control system using static state feedback. The conditions were obtained using techniques from differential geometry (for a full discussion of this see [17, 23]). It was shown that a system that satisfies these conditions can be transformed into a special canonical form, the so called Brunovsky normal form. As pointed out by Gardner and Shadwick in [11], this normal form is very close to the Goursat normal form for exterior differential systems. The problem of dynamic state feedback linearization, on the other hand, remained largely open, despite some early results by [6]. In his dissertation work Sluis [26] attempted to extend the exterior differential approach in this direction.

This tutorial paper is divided into three parts. Section 2 contains the necessary mathematical background on algebra and geometry for defining exterior differential systems. Section 3 describes some of the important normal forms for exterior differential systems: the Engel, Pfaff, Caratheodory, Goursat and extended Goursat normal forms. It is shown how certain important robotic systems can be converted to these normal forms. Section 4 discusses some of the connections between the exterior differential systems formalism, specialized to the case of control systems, and the vector field approach currently popular in nonlinear control. Finally, in Section 5 we highlight directions for future research and open problems.

2. Introduction to Exterior Differential Systems

In this section we will introduce the concept of an exterior differential system. To this end, we first introduce multilinear algebra, including the tensor and wedge products, and exterior algebra. Then we review some results from differential geometry including tangent spaces and vector fields. Once we have defined the exterior derivative, we will study many of its important properties. We then review the Frobenius theorem, for both vector fields and forms, and finally define an exterior differential system. Some tools which will be used to analyze these systems in the following sections will also be presented.

2.1. **Multilinear Algebra.**

2.1.1. *The Dual Space of a Vector Space.* Many of the ideas underlying the theory of multilinear algebra involve duality and the notion of the dual space to a vector space.

Definition 1. *Let $(V, \mathbb{R})$ denote a finite dimensional vector space over $\mathbb{R}$. The* dual space *associated with $(V, \mathbb{R})$ is defined as the space of all linear mappings $f : V \to \mathbb{R}$. The dual space of V is denoted as V^* and the elements of V^* are called* covectors. *V^* is a vector space over $\mathbb{R}$ with $dim(V^*) = dim(V)$ for the operations of addition and scalar multiplication defined by:*

$$
\begin{aligned}
(\alpha + \beta)(v) &= \alpha(v) + \beta(v) \\
(c\alpha)(v) &= c \cdot \alpha(v)
\end{aligned}
$$

Furthermore, if $\{v_1, \ldots, v_n\}$ is a set of basis vectors for V, then the set of linear functions $\phi^i : V \to \mathbb{R},\ 1 \le i \le n$, defined by:

$$\phi^i(v_j) = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

form a basis of V^* called the dual basis.

Example. Let $V = \mathbb{R}^n$ with the standard basis $e_1, \ldots, e_n$ and let $\phi^1, \ldots, \phi^n$ be the dual basis. If

$$x \in \mathbb{R}^n = \sum_{j=1}^{n} x_j e_j$$

then evaluating each function in the dual basis at x gives

$$\phi^i(x) = \phi^i(\sum_{j=1}^{n} x_j e_j) = \sum_{j=1}^{n} x_j \phi^i(e_j) = x_i$$

Since the functions $\phi^1, \ldots, \phi^n$ form a basis for V^*, a general covector in $(\mathbb{R}^n)^*$ is of the form $f = \alpha_1 \phi^1 + \ldots + \alpha_n \phi^n$. Evaluating this covector at the point x gives

$f(x) = \alpha_1 x_1 + \ldots + \alpha_n x_n$. If we think of a vector as a column matrix and a covector as a row matrix, then

$$f(x) = [\alpha_1 \ldots \alpha_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$\square$

Definition 2. *Given a subspace $W \subset V$ its* annihilator *is the subspace $W^\perp \subset V^*$ defined by*

$$W^\perp := \{\alpha \in V^* \mid \alpha(v) = 0 \; \forall \; v \in W\}$$

Given a subspace $X \subset V^*$, its annihilator is the subspace $X^\perp \subset V$ defined by

$$X^\perp := \{v \in V \mid \alpha(v) = 0 \; \forall \; \alpha \in X\}$$

A linear mapping between any two vector spaces $F : V_1 \to V_2$ induces a linear mapping between their dual spaces.

Definition 3. *Given a linear mapping $F : V_1 \to V_2$, its* dual map *is the linear mapping $F^* : V_2^* \to V_1^*$ defined by*

$$(F^*(\alpha))(v) = \alpha(F(v)), \; \forall \; \alpha \in V_2^*, \; v \in V_1$$

2.1.2. *Tensors.* Let $V_1, \ldots, V_k$ be a collection of real vector spaces. A function $f : V_1 \times \ldots \times V_k \to \mathbb{R}$ is said to be linear in the ith variable if the function $T : V_i \to \mathbb{R}$ defined for fixed v_j, $j \neq i$ as $T(v_i) = f(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_k)$ is linear. The function f is called multilinear if it is linear in each variable. A multilinear function $T : V^k \to \mathbb{R}$ is called a covariant tensor of order k or simply a k-tensor. The set of all k-tensors on V is denoted $\mathcal{L}^k(V)$. Note that $\mathcal{L}^1(V) = V^*$, the dual space of V. Therefore, we can think of covariant tensors as generalized covectors.

Example. The inner product of two vectors is an example of a 2-tensor. Another important example of a multilinear function is the determinant. If $x_1, x_2, \ldots, x_n$ are n column vectors in $\mathbb{R}^n$ then

$$f(x_1, x_2, \ldots, x_n) = det[x_1 \; x_2 \; \ldots \; x_n]$$

is multilinear by the properties of the determinant. $\square$

As in the case of V^*, each $\mathcal{L}^k(V)$ can be made into a vector space.

Theorem 1. *If for $S, T \in \mathcal{L}^k(V)$ and $c \in \mathbb{R}$ we define addition and scalar multiplication by:*

$$(S+T)(v_1, \ldots, v_k) = S(v_1, \ldots, v_k) + T(v_1, \ldots, v_k)$$
$$(cT)(v_1, \ldots, v_k) = c \cdot T(v_1, \ldots, v_k)$$

then the set of all k-tensors on V, $\mathcal{L}^k(V)$, is a real vector space.

Proof. See Munkres [20, page 220]. $\square$

Because of their multilinear structure, two tensors are equal if they agree on any set of basis elements.

Theorem 2. *Let $a_1, \ldots, a_n$ be a basis for V. Let $f, g : V^k \to \mathbb{R}$ be k-tensors on V. If $f(a_{i_1}, \ldots, a_{i_k}) = g(a_{i_1}, \ldots, a_{i_k})$ for every k-tuple $I = (i_1, \ldots, i_k) \in \{1, 2, \ldots, n\}^k$, then $f = g$.*

Proof. See Munkres [20, page 221]. ☐

Theorem 2 allows us to construct a basis for the space $\mathcal{L}^k(V)$.

Theorem 3. *Let $a_1, \ldots, a_n$ be a basis for V. Let $I = (i_1, \ldots, i_k) \in \{1, 2, \ldots, n\}^k$. Then there is a unique tensor ϕ^I on V such that for every k-tuple $J = (j_1, \ldots, j_k) \in \{1, 2, \ldots, n\}^k$*

$$\phi^I(a_{j_1}, \ldots, a_{j_k}) = \begin{cases} 0 & \text{if } I \neq J \\ 1 & \text{if } I = J \end{cases}$$

The collection of all such ϕ^I forms a basis for $\mathcal{L}^k(V)$.

Proof. Uniqueness follows from Theorem 2. To construct the functions ϕ^I, we start with a basis for V^*, $\phi^i : V \to \mathbb{R}$, defined by $\phi^i(a_j) = \delta_{ij}$ We then define each ϕ^I by:

$$\phi^I = \phi^{i_1}(v_1) \cdot \phi^{i_2}(v_2) \cdot \cdots \cdot \phi^{i_k}(v_k) \tag{1}$$

and claim that these ϕ^I form a basis for $\mathcal{L}^k(V)$. To show this, we select an arbitrary k-tensor $f \in \mathcal{L}^k(V)$, and define the scalars $\alpha_I := f(a_{i_1}, \ldots, a_{i_k})$. Next, we define a k-tensor

$$g = \sum_J \alpha_J \phi^J \tag{2}$$

where $J \in \{1, \ldots, n\}^k$. Then by Theorem 2, $f \equiv g$. ☐

Since there are n^k distinct k-tuples from the set $\{1, \ldots, n\}$ the space $\mathcal{L}^k(V)$ has dimension n^k.

Example. Let $V = \mathbb{R}^n$ with the standard basis $e_1, \ldots, e_n$, and let $\phi^1, \ldots, \phi^n$ be the dual basis. For every

$$x = \sum_{j=1}^n x_j e_j \in \mathbb{R}^n$$

evaluating each function in the dual basis at x gives

$$\phi^i(x) = \phi^i(\sum_{j=1}^n x_j e_j) = \sum_{j=1}^n x_j \phi^i(e_j) = x_i$$

Similarly, if $I = (i_1, \ldots, i_k)$ then evaluating the basis vectors for $\mathcal{L}^k(V)$ at $(x^1, \ldots, x^k)$ gives

$$\phi^I(x^1, \ldots, x^k) = \phi^{i_1}(x^1) \cdot \phi^{i^2}(x^2) \cdot \ldots \cdot \phi^{i_k}(x^k) = x^1_{i_1} \cdot \ldots \cdot x^k_{i_k}$$

Since the tensors $\phi^1, \ldots, \phi^n$ form a basis for V^*, evaluating a general 1-tensor $f \in V^*$ at $x \in V$ gives $f(x) = \alpha_1 x_1 + \ldots + \alpha_n x_n$. Likewise, evaluating a general 2-tensor at $(x^1, x^2) \in V^2$ gives

$$g(x^1, x^2) = \sum_{i,j=1}^n \alpha_{ij} x^1_i x^2_j = (x^1)^T D x^2$$

and evaluating a general k-tensor at $(x^1, \ldots, x^k) \in V^n$ gives

$$g(x^1, x^2, \ldots, x^k) = \sum_{i_1, \ldots, i_k = 1}^{n} \alpha_{i_1, \ldots, i_k} x^1_{i_1} \ldots x^k_{i_k}$$

where $\alpha_{i_1, \ldots, i_k} = g(e_{i_1}, \ldots, e_{i_k})$.

□

2.1.3. *Tensor Products.*

Definition 4. *Let $f \in \mathcal{L}^k(V)$ and $g \in \mathcal{L}^\ell(V)$. The* tensor product *$f \otimes g$ of f and g is a tensor in $\mathcal{L}^{k+\ell}(V)$ defined by*

$$(f \otimes g)(v_1, \ldots, v_{k+\ell}) := f(v_1, \ldots, v_k) \cdot g(v_{k+1}, \ldots, v_{k+\ell})$$

Theorem 4. *Let f, g, h be tensors on V and $c \in \mathbb{R}$. The following properties hold:*

1. *Associativity* $f \otimes (g \otimes h) = (f \otimes g) \otimes h$
2. *Homogeneity* $cf \otimes g = c(f \otimes g) = f \otimes cg$
3. *Distributivity* $(f + g) \otimes h = f \otimes h + g \otimes h$
4. *Given a basis $a_1, \ldots, a_n$ for V, the corresponding basis tensors satisfy $\phi^I = \phi^{i_1} \otimes \phi^{i_2} \otimes, \ldots, \otimes \phi^{i_k}$*

Proof. See Munkres [20, page 224]. □

We can also define the tensor product of two subspaces $U, W \subset V^*$ by:

$$U \otimes W := span\{x \in \mathcal{L}^2(V) \mid x = u \otimes w,\ u \in U,\ w \in W\}$$

From Theorem 3 we can conclude that $V^* \otimes V^* = \mathcal{L}^2(V)$. More generally:

$$\underbrace{V^* \otimes \ldots \otimes V^*}_{k-times} = \bigotimes^k V^* = \mathcal{L}^k(V)$$

2.1.4. *Alternating Tensors.* Before introducing alternating tensors, we present some facts about permutations.

Definition 5. *A* permutation *of the set of integers $\{1, 2, \ldots, k\}$ is an injective function σ mapping this set into itself. The set of all permutations σ is a group under function composition called the* symmetric group *on $\{1, \ldots, k\}$ and is denoted by S_k. Given $1 \le i < k$, a permutation e_i is called* elementary *if given some $i \in \{1, 2, \ldots, k\}$ we have*

$$\begin{aligned} e_i(j) &= j \qquad for \qquad j \neq i, i+1 \\ e_i(i) &= i + 1 \\ e_i(i+1) &= i \end{aligned}$$

An elementary permutation leaves the set intact except for consecutive elements i and $i + 1$ which are switched. The space S_k is of cardinality $k!$; its elements can be written as the composition of elementary permutations.

Definition 6. *Let $\sigma \in S_k$. Consider the set of all pairs of integers i, j from the set $\{1, \ldots, k\}$ for which $i < j$ and $\sigma(i) > \sigma(j)$. Each such pair is called an* inversion *in σ. The* sign *of σ is defined to be the number -1 if the number of inversions is odd and $+1$ if it is even. We call σ an* odd *or* even *permutation respectively. The sign of σ is denoted by $sgn(\sigma)$.*

The following theorem helps us calculate the sign of permutations.

Theorem 5. *Let $\sigma, \tau \in S_k$. Then*
1. *If σ is the composition of m elementary permutations then $sgn(\sigma) = (-1)^m$*
2. $sgn(\sigma \circ \tau) = sgn(\sigma) \cdot sgn(\tau)$
3. $sgn(\sigma^{-1}) = sgn(\sigma)$
4. *If $p \neq q$, and if τ is the permutation that exchanges p and q and leaves all other integers fixed, then $sgn(\tau) = -1$*

Proof. See Munkres [20, page 228]. □

We are now ready to define alternating tensors.

Definition 7. *Let f be an arbitrary k-tensor on V. If σ is a permutation of $\{1, \ldots, k\}$, we define f^σ by the equation*

$$f^\sigma(v_1, \ldots, v_k) = f(v_{\sigma(1)}, \ldots, v_{\sigma(k)})$$

Since f is linear in each of its variables, so is f^σ. The tensor f is said to be symmetric *if $f = f^e$ for each elementary permutation e, and it is said to be* alternating *if $f = -f^e$ for every elementary permutation e.*

We will denote the set of all alternating k-tensors on V by $\Lambda^k(V^*)$. The reason for this notation will be apparent when we introduce the wedge product in the next section. One can verify that the sum of two alternating tensors is alternating and a scalar multiple of an alternating tensor is alternating. Therefore, $\Lambda^k(V^*)$ is a linear subspace of the space $\mathcal{L}^k(V)$ of all k-tensors on V In the special case of $\mathcal{L}^1(V)$, elementary permutations cannot be performed and therefore every 1-tensor is vacuously alternating, i.e. $\Lambda^1(V^*) = \mathcal{L}^1(V) = V^*$. For completeness, we define $\Lambda^0(V^*) = \mathbb{R}$.

Example. Elementary tensors are not alternating but the linear combination:

$$f = \phi^i \otimes \phi^j - \phi^j \otimes \phi^i$$

is alternating. To see this, let $V = \mathbb{R}^n$ and let ϕ^i be the usual dual basis. Then

$$f(x, y) = x_i y_j - x_j y_i = det \begin{bmatrix} x_i & y_i \\ x_j & y_j \end{bmatrix}$$

and it is easily seen that $f(x, y) = -f(y, x)$. Similarly, the function

$$g(x, y, z) = det \begin{bmatrix} x_i & y_i & z_i \\ x_j & y_j & z_j \\ x_k & y_k & z_k \end{bmatrix}$$

is an alternating 3-tensor. □

We are interested in obtaining a basis for the linear space $\Lambda^k(V^*)$. We start with the following lemma.

Lemma 6. *Let f be a k-tensor on V and $\sigma, \tau \in S_k$ be permutations. Then*
1. *The transformation $f \longrightarrow f^\sigma$ is a linear transformation from $\mathcal{L}^k(V^*)$ to $\mathcal{L}^k(V^*)$. It has the property that for all $\sigma, \tau \in S_k$*

$$(f^\sigma)^\tau = f^{\tau \circ \sigma}$$

2. *The tensor f is alternating if and only if $f^\sigma = sgn(\sigma) \cdot f$ for all $\sigma \in S_k$.*
3. *If f is alternating and if $v_p = v_q$ with $p \neq q$ then $f(v_1, \ldots, v_k) = 0$.*

Proof. The linearity property is obvious since $(af+bg)^\sigma = af^\sigma + bg^\sigma$. Now

$$\begin{aligned}(f^\sigma)^\tau(v_1,\dots,v_k) &= f^\sigma(v_{\tau(1)},\dots,v_{\tau(k)})\\ &= f(v_{\tau(\sigma(1))},\dots,v_{\tau(\sigma(k))})\\ &= f^{\tau\circ\sigma}(v_1,\dots,v_k)\end{aligned}$$

Let σ be an arbitrary permutation, $\sigma = \sigma_1\circ\sigma_2\circ\cdots\circ\sigma_m$, where each σ_i is an elementary permutation. Then:

$$\begin{aligned}f^\sigma &= f^{\sigma_1\circ\sigma_2\circ\cdots\circ\sigma_m}\\ &= ((\dots(f^{\sigma_m})\dots)^{\sigma_2})^{\sigma_1}\\ &= (-1)^m\cdot f\\ &= sgn(\sigma)\cdot f\end{aligned}$$

Now suppose $v_p = v_q$ and $p\neq q$. Let τ be a permutation that exchanges p and q. Since $v_p = v_q$, $f^\tau(v_1,\dots,v_k) = f(v_1,\dots,v_k)$. Since f is an alternating tensor and $sgn(\tau) = -1$, $f^\tau(v_1,\dots,v_k) = -f(v_1,\dots,v_k)$. Therefore $f(v_1,\dots,v_k) = 0$. □

Lemma 6 implies that if $k > n$, the space $\Lambda^k(V^*)$ is trivial since one of the basis elements must appear in the k-tuple more than once. Hence for $k>n$, $\Lambda^k(V^*) = 0$. We have also seen that for $k=1$ we have $\Lambda^1(V^*) = \mathcal{L}^1(V) = V^*$ and therefore one can use the dual basis as a basis for $\Lambda^1(V^*)$. In order to specify an alternating tensor for $1 < k \leq n$ we simply need to define it on an ascending k-tuple of basis elements since, from Lemma 6, every other combination can be obtained by permuting the k-tuple.

Theorem 7. *Let $a_1, a_2, \dots, a_n$ be a basis for V. If f, g are alternating k-tensors on V and if*

$$f(a_{i_1}, a_{i_2}, \dots, a_{i_k}) = g(a_{i_1}, a_{i_2}, \dots, a_{i_k})$$

for every ascending k-tuple of integers $(i_1,\dots,i_k)\in\{1,2,\dots,n\}^k$ then $f=g$.

Proof. See Munkres [20, page 231]. □

Theorem 8. *Let $a_1,\dots,a_n$ be a basis for V. Let $I = (i_1,\dots,i_k)\in\{1,2,\dots,n\}^k$ be an ascending k-tuple. There is a unique alternating k-tensor ψ^I on V such that for every ascending k-tuple $J = (j_1,\dots,j_k)\in\{1,2,\dots,n\}^k$*

$$\psi^I(a_{j_1},\dots,a_{j_k}) = \begin{cases} 0 & \textit{if } J\neq I\\ 1 & \textit{if } J = I\end{cases}$$

The tensors ψ^I form a basis for $\Lambda^k(V^*)$ and satisfy the formula:

$$\psi^I = \sum_{\sigma\in S_k} sgn(\sigma)(\phi^I)^\sigma$$

Proof. In Munkres [20, pages 232–233]. □

The tensors ψ^I are called elementary alternating k-tensors on V corresponding to the basis $a_1,\dots,a_n$ of V. Every alternating k-tensor f may be uniquely expressed as $f = \sum_J d_J\psi^J$ where J indicates that summation extends over all ascending k-tuples. The dimension of $\Lambda^1(V^*)$ is simply n; its basis is the standard basis for V^*. If $k>1$, then we need to find the number of possible ascending k-tuples from the set $\{1,2,\dots,n\}$. Since if we choose k elements from a set of n elements there is only one way to put them in ascending order, the number of ascending k-tuples, and therefore the dimension of $\Lambda^k(V^*)$, is:

$$dim(\Lambda^k(V^*)) = \left(\begin{array}{c} n \\ k \end{array} \right) = \frac{n!}{k!(n-k)!}$$

2.1.5. *The Wedge Product.* Just as we defined the tensor product operation in the set of all tensors on a vector space V, we can define an analogous product operation, the wedge product, in the space of alternating tensors. The tensor product alone will not suffice, since even if $f \in \Lambda^k(V^*)$ and $g \in \Lambda^l(V^*)$ are alternating, their tensor product $f \otimes g \in \mathcal{L}^{k+l}(V)$ need not be alternating. We therefore construct an alternating operator taking k-tensors to alternating k-tensors.

Theorem 9. *For any tensor $f \in \mathcal{L}^k(V)$, define $Alt : \mathcal{L}^k(V) \rightarrow \Lambda^k(V^*)$ by:*

$$Alt(f) = \frac{1}{k!} \sum_{\sigma \in S_k} sgn(\sigma) f^{\sigma} \tag{3}$$

Then $Alt(f) \in \Lambda^k(V^*)$ and if $f \in \Lambda^k(V^*)$ then $Alt(f) = f$.

Proof. The fact that $Alt(f) \in \Lambda^k(V^*)$ is a consequence of Lemma 6, parts (1) and (2). Simply expanding the summation for $f \in \Lambda^k(V^*)$ yields that $Alt(f) = f$. □

Example. Let $f(x,y)$ be any 2-tensor. By using the alternating operator we obtain,

$$Alt(f) = \frac{1}{2}(f(x,y) - f(y,x))$$

which is clearly alternating. Similarly for any 3-tensor $g(x,y,z)$ we have:

$$Alt(g) = \frac{1}{6}(g(x,y,z) + g(y,z,x) + g(z,x,y) - g(y,x,z) - g(z,y,x) - g(x,z,y))$$

which can be easily checked to be alternating. □

Definition 8. *Given $f \in \Lambda^k(V^*)$ and $g \in \Lambda^l(V^*)$, we define the* wedge product *or* exterior product*, $f \wedge g \in \Lambda^{k+l}(V^*)$ by*

$$f \wedge g = \frac{(k+l)!}{k!l!} Alt(f \otimes g).$$

The somewhat complicated normalization constant is required as we would like the wedge product to be associative and $Alt(f) = f$ if f is already alternating. Since alternating tensors of order zero are elements of $\mathbb{R}$, we define the wedge product of an alternating 0-tensor and any alternating k-tensor by the usual scalar multiplication. The following theorem lists some important properties of the wedge product.

Theorem 10. *Let $f \in \Lambda^k(V^*)$, $g \in \Lambda^l(V^*)$ and $h \in \Lambda^m(V^*)$. Then:*

1. *Associativity* $f \wedge (g \wedge h) = (f \wedge g) \wedge h$
2. *Homogeneity* $cf \wedge g = c(f \wedge g) = f \wedge cg$
3. *Distributivity* $(f+g) \wedge h = f \wedge h + g \wedge h$
   $h \wedge (f+g) = h \wedge f + h \wedge g$
4. *Skew-commutativity*[1], $g \wedge f = (-1)^{kl} f \wedge g$

[1] also called anti-commutativity

Proof. Properties (2), (3) and (4) follow directly from the definitions of the alternating operator and the tensor product. Associativity, property (1), requires a few more manipulations (see Spivak [30, pages 80–81]). □

Example. Let $f(x) \in \Lambda^1(V^*)$ and $g(y,z) \in \Lambda^2(V^*)$. Then:

$$\begin{aligned} f \wedge g = &\frac{(2+1)!}{2!1!}\frac{1}{3!}(f(x)\otimes g(y,z) + f(y)\otimes g(z,x) + \\ &+ f(z)\otimes g(x,y) - f(y)\otimes g(x,z) - f(z)\otimes g(y,x) - f(x)\otimes g(z,y)) \end{aligned}$$

We can also check that:

$$f \wedge f = \frac{(1+1)!}{1!1!}\frac{1}{2!}(f(x)\otimes f(x) - f(x)\otimes f(x)) = 0$$

which can also been seen from the skew commutativity of exterior multiplication. □

An elegant basis for $\Lambda^k(V^*)$ can be formed using of the dual basis for V.

Theorem 11. *Given a basis $a_1, \dots, a_n$ for vector space V, let $\phi^1, \dots, \phi^n$ denote its dual basis and ψ^I the corresponding elementary alternating tensors. Then if $I = (i_1, \dots, i_k)$ is any ascending k-tuple of integers,*

$$\psi^I = \phi^{i_1} \wedge \phi^{i_2} \wedge \dots \wedge \phi^{i_k}.$$

Proof. May be deduced from the construction of the elementary alternating tensors in Theorem 8. □

By Theorem 11, any alternating k-tensor $f \in \Lambda^k(V^*)$ may be expressed in terms of the dual basis $\phi^1, \dots, \phi^n$ as:

$$f = \sum_J d_{j_1,\dots,j_k}\phi^{j_1} \wedge \phi^{j_2} \wedge \dots \wedge \phi^{j_k}$$

for all ascending k-tuples $J = (j_1, \dots, j_k)$ and some scalars, $d_{j_1,\dots,j_k}$. If we require the coefficients to be skew-symmetric, $d_{i_1,\dots,i_l,i_{l+1},\dots,i_k} = -d_{i_1,\dots,i_{l+1},i_l,\dots,i_k}$ for all $l \in \{1, \dots, k-1\}$ we can extend this summation over all k-tuples.

$$f = \frac{1}{k!}\sum_{i_1,\dots,i_k=1}^{n} d_{i_1,\dots,i_k}\phi^{i_1} \wedge \phi^{i_2} \wedge \dots \wedge \phi^{i_k} \tag{4}$$

The wedge product provides a convenient way to check whether a set of 1-tensors is linearly independent.

Theorem 12. *If $\omega^1, \dots, \omega^k$ are 1-tensors over V then*

$$\omega^1 \wedge \omega^2 \wedge \dots \wedge \omega^k = 0$$

if and only if $\omega^1, \dots, \omega^k$ are linearly dependent.

Proof. Suppose that $\omega^1, \dots, \omega^k$ are linearly independent, and pick $\alpha^{k+1}, \dots, \alpha^n$ to complete a basis for V^*. From Theorem 11 we know that $\omega^1 \wedge \omega^2 \wedge \dots \wedge \omega^k$ is a basis element for $\Lambda^k(V^*)$. Therefore, it must be nonzero. If $\omega^1, \dots, \omega^k$ are linearly dependent, then at least one of the them can be written as a linear combination of the others. Without loss of generality, assume that:

$$\omega^k = \sum_{i=1}^{k-1} c_i \omega^i$$

From this we get that:

$$\omega^1 \wedge \omega^2 \wedge \ldots \wedge \omega^k = \omega^1 \wedge \omega^2 \wedge \ldots \wedge \omega^{k-1} \wedge (\sum_{i=1}^{k-1} c_i \omega^i) = 0$$

by the skew-commutativity of the wedge product. □

Theorem 12 allows us to give a geometric interpretation to a nonzero k-tensor

$$\omega^1 \wedge \omega^2 \wedge \ldots \wedge \omega^k \neq 0$$

by associating it with the subspace

$$W := span\{\omega^1, \ldots, \omega^k\} \subset V^*.$$

An obvious question that arises is what happens if we select a different basis for W.

Theorem 13. *Given a subspace $W \subset V^*$ and two sets of linearly-independent 1-tensors which span W, there exists a nonzero scalar $c \in \mathbb{R}$ such that*

$$c \cdot \omega^1 \wedge \omega^2 \ldots \wedge \omega^k = \alpha^1 \wedge \alpha^2 \wedge \ldots \wedge \alpha^k \neq 0$$

Proof. Each α^i can be written as a linear combination of the ω^i

$$\alpha^i = \sum_{j=1}^{k} a_{ij} \omega^j.$$

Therefore, the product

$$\alpha^1 \wedge \alpha^2 \wedge \ldots \wedge \alpha^k = (\sum_{j=1}^{k} a_{1j} \omega^j) \wedge \ldots \wedge (\sum_{j=1}^{k} a_{kj} \omega^j)$$

Multiplying this out gives

$$\alpha^1 \wedge \alpha^2 \wedge \ldots \wedge \alpha^k = \sum_{i_1, \ldots, i_k = 1}^{n} b_{i_1, \ldots, i_k} \omega^{i_1} \wedge \omega^{i_2} \wedge \ldots \wedge \omega^{i_k}.$$

The claim follows by Theorem 12, and the skew commutativity of the wedge product. □

Definition 9. *A k-tensor $\xi \in \Lambda^k(V^*)$ is* decomposable *if there exist $x^1, x^2, \ldots, x^k \in \Lambda^1(V^*)$ such that $\xi = x^1 \wedge x^2 \wedge \ldots \wedge x^k$.*

Note that, if ξ is decomposable, then we must have $\xi \wedge \xi = 0$. The reason is that we should be able to express ξ as $\xi = \alpha^1 \wedge \alpha^2 \wedge \ldots \wedge \alpha^k$ for some basis vectors $\{\alpha^1, \alpha^2, \ldots, \alpha^n\}$ and therefore $\xi \wedge \xi = \alpha^1 \wedge \alpha^2 \wedge \ldots \wedge \alpha^k \wedge \alpha^1 \wedge \alpha^2 \wedge \ldots \wedge \alpha^k = 0$. Not all $\xi \in \Lambda^k(V^*)$ are decomposable, as demonstrated in the following example.

Example. Let $\xi = \phi^1 \wedge \phi^2 + \phi^3 \wedge \phi^4 \in \Lambda^2((\mathbb{R}^4)^*)$. Then, $\xi \wedge \xi = 2\phi^1 \wedge \phi^2 \wedge \phi^3 \wedge \phi^4 \neq 0$ Therefore ξ is not decomposable. Note that $\xi \wedge \xi = 0$ is a necessary but not a sufficient condition for ξ to be decomposable. For example if ξ is an odd alternating tensors (say of dimension $2k+1$):

$$\xi \wedge \xi = (-1)^{(2k+1)^2} \xi \wedge \xi = 0$$

□

If an alternating k-tensor ξ is not decomposable, it may still be possible to factor out a 1-tensor from every term in the summation which defines it.

Example. Let $\xi = \phi^1 \wedge \phi^2 \wedge \phi^5 + \phi^3 \wedge \phi^4 \wedge \phi^5 \in \Lambda^3((\mathbb{R}^5)^*)$. From the previous example, we know that this tensor is not decomposable, but the 1-tensor ϕ^5 can clearly be factored from every term

$$\xi = (\phi^1 \wedge \phi^2 + \phi^3 \wedge \phi^4) \wedge \phi^5 = \hat{\xi} \wedge \phi^5$$

□

Definition 10. *Let $\xi \in \Lambda^k(V^*)$. The subspace $L_\xi \subset V^*$ defined by:*

$$L_\xi := \{\omega \in V^* \mid \xi = \hat{\xi} \wedge \omega \text{ for some } \hat{\xi} \in \Lambda^{k-1}(V^*)\}$$

is called the divisor space *of ξ. Any $\omega \in L_\xi$ is called a* divisor *of ξ.*

Theorem 14. *A 1-tensor $\omega \in V^*$ is a divisor of $\xi \in \Lambda^k(V^*)$ if and only if $\omega \wedge \xi \equiv 0$.*

Proof. Pick a basis $\phi^1, \phi^2, \ldots, \phi^n$ for V^* such that $\omega = \phi^1$. With respect to this basis, ξ can be written as

$$\xi = \sum_J^n d_{j_1,\ldots,j_k} \phi^{j_1} \wedge \phi^{j_2} \wedge \ldots \wedge \phi^{j_k} \tag{5}$$

for all ascending k-tuples $J = (j_1, \ldots, j_k)$ and some scalars, $d_{j_1,\ldots,j_k}$. If ω is a divisor of ξ, then it must be contained in each nonzero term of this summation. Therefore $\omega \wedge \xi \equiv 0$. On the other hand, if $\omega \wedge \xi \equiv 0$, then every nonzero term of ξ must contain ω. Otherwise, we would have $\omega \wedge \phi^{j_1} \wedge \ldots \wedge \phi^{j_k} = \phi^1 \wedge \phi^{j_1} \wedge \ldots \wedge \phi^{j_k}$ for $j_1, \ldots, j_k \neq 1$ which is a basis element of $\Lambda^{k+1}(V^*)$ and therefore nonzero. □

If we select a basis $\phi^1, \phi^2, \ldots, \phi^n$ for V^* such that $span\{\phi^1, \phi^2, \ldots, \phi^l\} = L_\xi$, then, ξ can be written as $\xi = \hat{\xi} \wedge \phi^1 \wedge \ldots \wedge \phi^l$, where $\hat{\xi} \in \Lambda^{k-l}(V^*)$ is not decomposable and involves only $\phi^{l+1}, \ldots, \phi^n$.

2.1.6. *The Interior Product.*

Definition 11. *The* interior product *is a linear mapping $\lrcorner : V \times \mathcal{L}^k(V) \to \mathcal{L}^{k-1}(V)$ which operates on a a vector $v \in V$ and a tensor $T \in \mathcal{L}^k(V)$ and produces a tensor $(v \lrcorner T) \in \mathcal{L}^{k-1}(V)$ defined by*

$$(v \lrcorner T)(v_1, \ldots, v_{k-1}) := T(v, v_1, \ldots, v_{k-1})$$

Theorem 15. *Let $a, b, c, d \in \mathbb{R}$ be real numbers; $v, w \in V$ be vectors; $g, h \in \mathcal{L}^k(V)$ be k-tensors; and $r \in \Lambda^s(V^*)$, $f \in \Lambda^m(V^*)$ be alternating tensors. Then we have the following identities.*

1. *Bilinearity* $\begin{array}{rcl} (av + bw) \lrcorner g & = & a(v \lrcorner g) + b(w \lrcorner g) \\ v \lrcorner (cg + dh) & = & c(v \lrcorner g) + d(v \lrcorner h) \end{array}$
2. $v \lrcorner (f \wedge r) = (v \lrcorner f) + (-1)^m f \wedge (v \lrcorner g)$

Proof. See Abraham et al. [1, page 429]. □

Theorem 16. *Let $a_1, \ldots, a_n$ be a basis for V. Then the value of an alternating k-tensor $\omega \in \Lambda^k(V^*)$ is independent of a basis element a_i if and only if $a_i \lrcorner \omega \equiv 0$.*

Proof. Let $\phi^1, \ldots, \phi^n$ be the dual basis to $a_1, \ldots, a_n$. Then ω can be written with respect to the dual basis as

$$\omega = \sum_J d_J \phi^{j_1} \wedge \phi^{j_2} \wedge \ldots \wedge \phi^{j_k} = \sum_J d_J \psi^J$$

where the sum is taken oven all ascending k-tuples J. If a basis element ψ^J does not contain ϕ_i, then clearly $a_i \lrcorner \psi^J \equiv 0$. If a basis element contains ϕ_i, then $a_i \lrcorner \phi^{j_1} \wedge \phi^{j_2} \wedge \ldots \wedge \phi^{j_k} \not\equiv 0$ because a_i can always be matched with ϕ_i through a permutation which only affects the sign. Consequently, $(a_I \lrcorner \omega) \equiv 0$ if and only if the coefficients d_J of all the terms containing ϕ^i are zero. □

Definition 12. *Let $\omega \in \Lambda^k(V^*)$ be an alternating k-tensor. The space consisting of all vectors of which the value of ω is independent is called the* associated space *of ω:*

$$A_\omega := \{v \in V | v \lrcorner \omega \equiv 0\}$$

The dual associated space *of ω is defined as $A_\omega^\perp \subset V^*$.*

Recall that the divisor space L_ω of an alternating k-tensor ω contains all the 1-tensors which can be factored from every term of ω. The dual associated space $A_\omega^\perp$ contains all the 1-tensors which are contained in at least one term of ω. Therefore, $L_\omega \subset A_\omega^\perp$.

Theorem 17. *The following statements are equivalent:*

1. *An alternating k-tensor $\omega \in \Lambda^k(V^*)$ is decomposable.*
2. *The divisor space L_ω has dimension k.*
3. *The dual associated space $A_\omega^\perp$ has dimension k.*
4. *$L_\omega = A_\omega^\perp$.*

Proof. (1) $\Leftrightarrow$ (2) If ω is decomposable, then there exists a set of basis vectors $\phi^1, \phi^2, \ldots, \phi^n$ for V^* such that $\omega = \phi^1 \wedge \ldots \wedge \phi^k$. Therefore $L_\omega = span\{\phi^1, \phi^2, \ldots, \phi^k\}$ which has dimension k. Conversely, if L_ω has dimension k, then k terms can be factored from ω. Since ω is a k-tensor, it must be decomposable.

(1) $\Leftrightarrow$ (3) Let $a_1, \ldots, a_n$ be the basis of V which is dual to $\phi^1, \phi^2, \ldots, \phi^n$. Since $\omega = \phi^1 \wedge \ldots \wedge \phi^k$, ω is not a function of $a_{k+1}, \ldots, a_n$. Therefore,

$$A_\omega = span\{a_{k+1}, \ldots, a_n\}.$$

This implies that $A_\omega^\perp$ has dimension k. Conversely, if $A_\omega^\perp$ has dimension k, then A_ω has dimension $n-k$ which means that ω is an alternating k-tensor which is a function of k variables. Therefore, it must have the form $\omega = \phi^1 \wedge \ldots \wedge \phi^k$, for some linearly independent $\phi^1, \phi^2, \ldots, \phi^k$ in V^*.

(2)&(3) $\Leftrightarrow$ (4) It is always true that $L_\omega \subset A_\omega^\perp$. Therefore if $dim(L_\omega) = dim(A_\omega^\perp)$ then $L_\omega = A_\omega^\perp$. It is also always true that $0 \leq dim(L_\omega) \leq k$ and $k \leq dim(A_\omega^\perp) \leq n$. Therefore, $L_\omega = A_\omega^\perp$ implies that $dim(L_\omega) = dim(A_\omega^\perp) = k$. □

2.1.7. *The Pull Back of a Linear Transformation.* Let T be a linear map from a vector space V to a vector space W. Assume that there exists a multilinear function f on W. Using the above, we can define a multilinear function on V as follows:

Definition 13. *Let $T : V \to W$ be a linear transformation. The* dual *or* pull back transformation

$$T^* : \mathcal{L}^k(W) \to \mathcal{L}^k(V)$$

is defined for all $f \in \mathcal{L}^k(W)$ by

$$(T^* f)(v_1, \ldots, v_k) := f(T(v_1), \ldots, T(v_k))$$

Note that $T^* f$ is multilinear since T is a linear transformation.

Theorem 18. *Let $T : V \to W$ be a linear transformation, and let*

$$T^* : \mathcal{L}^k(W) \to \mathcal{L}^k(V)$$

be the dual transformation. Then

1. *T^* is linear.*
2. *$T^*(f \otimes g) = T^*f \otimes T^*g$.*
3. *If $S : W \to X$ is linear, then $(S \circ T)^* f = T^*(S^* f)$.*

Proof. See Munkres [20, page 225]. □

Theorem 19. *Let $T : V \longrightarrow W$ be a linear transformation. If f is an alternating tensor on W then $T^* f$ is an alternating tensor on V, and*

$$T^*(f \wedge g) = T^* f \wedge T^* g$$

Proof. See Abraham et. al. [1, page 420]. □

2.1.8. *Algebras and Ideals.* In Sections 2.1.5 and 2.1.6, we introduced the wedge product and interior product and demonstrated some of their properties. We now look more closely at the algebraic structure these operations impart to the space of alternating tensors. We begin by introducing some algebraic structures which will be used in the development of the exterior algebra.

Definition 14. *An* algebra, *$(V, \odot)$, is a vector space V together with a multiplication operation $\odot : V \times V \to V$ which for every scalar $\alpha \in \mathbb{R}$ and $a, b \in V$ satisfies $\alpha(a \odot b) = (\alpha a) \odot b = a \odot (\alpha b)$. If there exists an element $e \in V$ such that for all $x \in V, x \odot e = e \odot x = x$ then e is unique and is called the* identity element.

Definition 15. *Given an algebra $(V, \odot)$, a subspace $W \subset V$ is called an* algebraic ideal[2] *if $x \in W, y \in V$ implies that $x \odot y, y \odot x \in W$*

Recall that if W is an ideal and $x, y \in W$ then $x + y \in W$ since W is a subspace.

Example. The set of all polynomials with real-valued coefficients, $\mathbb{R}[s]$, is a vector space over $\mathbb{R}$ with vector addition and scalar multiplication defined by

$$(P_1 + P_2)(s) = P_1(s) + P_2(s)$$
$$(\alpha \cdot P)(s) = \alpha \cdot P(s)$$

If we define multiplication by

$$(P_1 \cdot P_2)(s) = P_1(s) \cdot P_2(s)$$

then $\mathbb{R}[s]$ is also an algebra. In $\mathbb{R}[s]$, the set of all polynomials with a zero at $s = -2$ is an algebraic ideal. This is true because for all $P_1(s), P_2(s) \in \mathbb{R}[s]$ which satisfy $P_1(-2) = P_2(-2) = 0$ we have that:

$$P_1(-2) + P_2(-2) = 0, \quad \alpha \cdot P_1(-2) = 0, \quad P_1(-2) \cdot P_2(-2) = 0$$

Furthermore for all $P(s), R(s) \in \mathbb{R}[s]$ with $P(-2) = 0$ we have that

$$P(-2) \cdot R(-2) = 0$$

□

[2]For readers who are familiar with algebra, the algebraic ideal is the ideal of the algebra considered as a ring. Furthermore, since this ring has an identity, any ideal must be a subspace of the algebra considered as a vector space.

It can be easily verified that the intersection of ideals is also an ideal. Using this fact we have the following definition.

Definition 16. *Let $(V, \odot)$ be an algebra. Let the set $A := \{a_i \in V, 1 \leq i \leq K\}$ be any finite collection of linearly independent elements in V. Let S be the set of all ideals containing A*

$$S := \{I \subset V | I \text{ is an ideal and } A \subset I\}.$$

The ideal I_A generated by A *is defined as*

$$I_A = \bigcap_{I \in S} I$$

and is the minimal ideal in S containing A.

If $(V, \odot)$ has an identity element, then the ideal generated by a finite set of elements can be represented in a simple form.

Theorem 20. *Let $(V, \odot)$ be an algebra with an identity element, $A := \{a_i \in V, 1 \leq i \leq K\}$ a finite collection of elements in V, and I_A the ideal generated by A. Then for each $x \in I_A$, there exist vectors $v_1, \ldots, v_K$ such that*

$$x = v_1 \odot a_1 + v_2 \odot a_2 + \ldots + v_K \odot a_K$$

Proof. See Hungerford [16, pages 123–124]. □

Definition 17. *Let $(V, \odot)$ be an algebra, and $I \subset V$ an ideal. Two vectors $x, y \in V$ are said to be* equivalent mod I *if and only if $x - y \in I$. This equivalence is denoted*

$$x \equiv y \text{ mod } I$$

If $(V, \odot)$ has an identity element the above definition implies that $x \equiv y \text{ mod } I$ if and only if:

$$x - y = \sum_{i=1}^{K} \theta_i \odot \alpha_i$$

for some $\theta_K \in V$. It is customary to denote this as $x \equiv y \text{ mod } \alpha_1, \ldots, \alpha_K$. where the mod operation is implicitly performed over the ideal generated by $\alpha_1, \ldots, \alpha_K$.

2.1.9. *The Exterior Algebra of a Vector Space.* Although the space $\Lambda^k(V^*)$ is a vector space with a multiplication operation, the wedge product of two alternating k-tensors is not a k-tensor. Therefore $\Lambda^k(V^*)$ is not an algebra under the wedge product. If we consider, however, the direct sum of the space of all alternating tensors we obtain,

$$\Lambda(V^*) = \Lambda^0(V^*) \oplus \Lambda^1(V^*) \oplus \Lambda^2(V^*) \oplus \ldots \oplus \Lambda^n(V^*)$$

Any $\xi \in \Lambda(V^*)$ may be written as $\xi = \xi_0 + \xi_1 + \ldots + \xi_n$ where each $\xi_p \in \Lambda^p(V^*)$. $\Lambda(V^*)$ is clearly a vector space, and is also closed under exterior multiplication. It is therefore an algebra.

Definition 18. $(\Lambda(V^*), \wedge)$ *is an algebra, called the* exterior algebra *over V^*.*

Since $(\Lambda(V^*), \wedge)$ has the identity element $1 \in \Lambda^0(V^*)$, Theorem 20 implies that the ideal generated by a finite set $\Sigma := \{\alpha^i \in \Lambda(V^*), 1 \leq i \leq K\}$ can be written as:

$$I_\Sigma = \{\pi \in \Lambda(V^*) | \ \pi = \sum_{i=1}^{K} \theta^i \wedge \alpha^i, \theta^i \in \Lambda(V^*)\}$$

Given an arbitrary set Σ of linearly independent generators, it may also be possible to generate I_Σ with a smaller set of generators Σ'.

2.1.10. *Systems of Exterior Equations.* In the preceding sections we have developed an algebra of alternating multilinear functions over a vector space. We will now apply these ideas to solve a system of equations in the form

$$\alpha^1 = 0, \dots, \alpha^K = 0$$

where each $\alpha^i \in \Lambda(V^*)$ is an alternating tensor. First we need to clarify what constitutes a "solution" to these equations.

Definition 19. *A* system of exterior equations on V *is a finite set of linearly independent equations*

$$\alpha^1 = 0, \dots, \alpha^K = 0$$

where each $\alpha^i \in \Lambda^k(V^*)$ *for some* $1 \leq k \leq n$. *A* solution to a system of exterior equations *is any subspace* $W \subset V$ *such that*

$$\alpha^1|_W \equiv 0, \dots, \alpha^K|_W \equiv 0$$

where $\alpha|_W$ *stands for* $\alpha(v_1, \dots, v_k)$ *for all* $v_1, \dots, v_k \in W$

A system of exterior equations generally does not have a unique solution since any subspace $W_1 \subset W$ will satisfy $\alpha|_{W_1} \equiv 0$ if $\alpha|_W \equiv 0$. A central fact concerning systems of exterior equations is given by the following theorem:

Theorem 21. *Given a system of exterior equations* $\alpha^1 = 0, \dots, \alpha^K = 0$ *and the corresponding ideal* I_Σ *generated by the collection of alternating tensors* $\Sigma := \{\alpha^1, \dots, \alpha^K\}$, *a subspace* W *solves the system of exterior equations if and only if it also satisfies* $\pi|_W \equiv 0$ *for every* $\pi \in I_\Sigma$.

Proof. Clearly, $\pi|_W \equiv 0$ for every $\pi \in I_\Sigma$ implies $\alpha^i|_W = 0$ as $\alpha^i \in I_\Sigma$. Conversely, if $\pi \in I_\Sigma$, then $\pi = \sum_{i=1}^K \theta^i \wedge \alpha^i$ for some $\theta^i \in \Lambda(V^*)$. Therefore, $\alpha^i|_W = 0$ implies that $\pi|_W \equiv 0$. □

This result allows us to treat the system of exterior equations, the set of generators for the ideal, and the algebraic ideal as essentially equivalent objects. We may sometimes abuse notation and confuse a system of equations with its corresponding generator set and a generator set with its corresponding ideal. When it is important to distinguish them, we will explicitly write out the system of exterior equations, denote the set of generators by Σ and the ideal which they generate by I_Σ.

Recall that an algebraic ideal was defined in a coordinate-free way as a subspace of the algebra satisfying certain closure properties. Thus the ideal has an intrinsic geometric meaning, and we can think of two sets of generators as representing the same system of exterior equations if they generate the same algebraic ideal.

Definition 20. *Two sets of generators,* Σ_1 *and* Σ_2 *which generate the same ideal, i.e.* $I_{\Sigma_1} = I_{\Sigma_2}$, *are said to be* algebraically equivalent.

We will exploit this notion of equivalence to represent a system of exterior equations in a simplified form. In order to do this, we need a few preliminary definitions.

Definition 21. *Let* Σ *be a system of exterior equations and* I_Σ *the ideal which it generates. The* associated space of the ideal I_Σ *is defined by:*

$$A(I_\Sigma) := \{v \in V | v \,\lrcorner\, \alpha \in I_\Sigma \;\forall\, \alpha \in I_\Sigma\}$$

that is, for all v in the associated space and α in the ideal, $v \lrcorner \alpha \equiv 0 \mod I_\Sigma$. The dual associated space *or* retracting space *of the ideal is defined by: $A(I_\Sigma)^\perp$ and denoted by $C(I_\Sigma) \subset V^*$.*

Once we have determined the retracting space $C(I_\Sigma)$, we can find an algebraically equivalent system Σ' which is a subset of $\Lambda(C(I_\Sigma))$, the exterior algebra over the retracting space.

Theorem 22. *Let Σ be a system of exterior equations and I_Σ its corresponding algebraic ideal. Then there exists an algebraically equivalent system Σ' such that $\Sigma' \subset \Lambda(C(I_\Sigma))$.*

Proof. Let $v_1, \ldots, v_n$ be a basis for V, and $\phi^1, \ldots, \phi^n$ be the dual basis, selected such that $v_{r+1}, \ldots, v_n$ span $A(I_\Sigma)$. Consequently $\phi^1, \ldots, \phi^r$ must span $C(I_\Sigma)$.

The proof is by induction. First, let α be any one-tensor in I_Σ. With respect to the chosen basis, α can be written as

$$\alpha = \sum_{i=1}^{n} a_i \phi^i$$

Since $v \lrcorner \alpha \equiv 0 \mod I_\Sigma$ for all $v \in A(I_\Sigma)$ by the definition of the associated space, we must have $a_i = 0$ for $i = r+1, \ldots, n$. Therefore,

$$\alpha = \sum_{i=1}^{r} a_i \phi^i.$$

So all the 1-tensors in Σ are contained in $\Lambda^1(C(I_\Sigma))$.

Now suppose that all tensors of degree $\leq k$ in I_Σ are contained in $\Lambda(C(I_\Sigma))$. Let α be any $k+1$ tensor in I_Σ. Consider the tensor

$$\alpha' = \alpha - \phi^{r+1} \wedge (v_{r+1} \lrcorner \alpha)$$

The term $v_{r+1} \lrcorner \alpha$ is a k-tensor in I_Σ by the definition of associated space, and thus, by the induction hypothesis, it must be in $\Lambda(C(I_\Sigma))$. The wedge product of this term with ϕ^{r+1} is also clearly in $\Lambda(C(I_\Sigma))$. Furthermore,

$$v_{r+1} \lrcorner \alpha' = v_{r+1} \lrcorner \alpha - (v_{r+1} \lrcorner \phi^{r+1}) \wedge (v_{r+1} \lrcorner \alpha) + \phi^{r+1} \wedge (v_{r+1} \lrcorner (v_{r+1} \lrcorner \alpha)) \equiv 0$$

By Theorem 16, α' has no terms involving ϕ^{r+1}.

If we now replace α with α' the ideal generated will be unchanged since

$$\theta \wedge \alpha = \theta \wedge \alpha' + \theta \wedge \phi^{r+1} \wedge (v_{r+1} \lrcorner \alpha)$$

and $v_{r+1} \lrcorner \alpha \in I_\Sigma$.

We can continue this process for $v_{r+2}, \ldots, v_n$ to produce an $\hat{\alpha}$ which is a generator of I_Σ and is an element of $\Lambda(C(I_\Sigma))$. □

Example. Let $v_1, \ldots, v_6$ be a basis for $\mathbb{R}^6$, and let $\theta^1, \ldots, \theta^6$ be the dual basis. Consider the system of exterior equations

$$\alpha^1 = \theta^1 \wedge \theta^3 = 0,$$

$$\alpha^2 = \theta^1 \wedge \theta^4 = 0,$$

$$\alpha^3 = \theta^1 \wedge \theta^2 - \theta^3 \wedge \theta^4 = 0,$$

$$\alpha^4 = \theta^1 \wedge \theta^2 \wedge \theta^5 - \theta^3 \wedge \theta^4 \wedge \theta^6 = 0$$

Let I_Σ be the ideal generated by $\Sigma = \{\alpha^1, \alpha^2, \alpha^3, \alpha^4\}$ and $A(I_\Sigma)$ the associated space of I_Σ. Because I_Σ contains no 1-tensors, we can infer that for all $v \in A(I_\Sigma)$ $v \lrcorner \alpha^1 = 0$, $v \lrcorner \alpha^2 = 0$, and $v \lrcorner \alpha^3 = 0$. Expanding the first equation, we get

$$v \lrcorner \alpha^1 = v \lrcorner (\theta^1 \wedge \theta^3) = (v \lrcorner \theta^1) \wedge \theta^3 + (-1)^1 \theta^1 \wedge (v \lrcorner \theta^3) = \theta^1(v)\theta^3 - \theta^3(v)\theta^1 = 0$$

which implies that $\theta^1(v) = 0$ and $\theta^3(v) = 0$. Similarly,

$$\begin{aligned} v \lrcorner \alpha^2 &= \theta^1(v)\theta^4 - \theta^4(v)\theta^1 = 0 \\ v \lrcorner \alpha^3 &= \theta^1(v)\theta^2 - \theta^2(v)\theta^1 - \theta^3(v)\theta^4 + \theta^4(v)\theta^3 = 0 \end{aligned}$$

implying that $\theta^2(v) = 0$ and $\theta^4(v) = 0$. Therefore, we can conclude that

$$A(I_\Sigma) \subset span\{v_5, v_6\}.$$

Evaluating the expression $v \lrcorner \alpha^4 \subset I_\Sigma$ gives

$$\begin{aligned} v \lrcorner \alpha^4 &= (v \lrcorner (\theta^1 \wedge \theta^2)) \wedge \theta^5 + (-1)^2 (\theta^1 \wedge \theta^2) \wedge (v \lrcorner \theta^5) \\ &\quad - (v \lrcorner (\theta^3 \wedge \theta^4)) \wedge \theta^6 - (-1)^2 (\theta^3 \wedge \theta^4) \wedge (v \lrcorner \theta^6) \\ &= \theta^5(v)\theta^1 \wedge \theta^2 - \theta^6(v)\theta^3 \wedge \theta^4 \\ &= a(\theta^1 \wedge \theta^3) + b(\theta^1 \wedge \theta^4) + c(\theta^1 \wedge \theta^2 - \theta^3 \wedge \theta^4) \end{aligned}$$

Equating coefficients, we find that

$$\theta^5(v) = \theta^6(v) = c, \; \forall\, v \in A(I_\Sigma).$$

Now v must be of the form $v = xv_5 + yv_6$, so we get

$$\theta^5(xv_5 + yv_6) = x = c$$

$$\theta^6(xv_5 + yv_6) = y = c.$$

Therefore, $A(I_\Sigma) = span\{(v_5 + v_6)\}$. If we select as a new basis for $\mathbb{R}^6$ the vectors

$$w_i = v_i, \; i = 1, \ldots, 4, \; w_5 = v_5 - v_6, \; w_6 = v_5 + v_6$$

then the new dual basis becomes

$$\gamma^i = \theta^i, \; i = 1, \ldots, 4, \; \gamma^5 = \frac{\theta^5 - \theta^6}{2}, \; \gamma^6 = \frac{\theta^5 + \theta^6}{2}.$$

With respect to this new basis, the retracting space $C(I_\Sigma)$ is given by

$$C(I_\Sigma) = span\{\gamma^1, \ldots, \gamma^5\}$$

In these coordinates, the generator set becomes

$$\Sigma' = \{\gamma^1 \wedge \gamma^3, \; \gamma^1 \wedge \gamma^4, \; \gamma^1 \wedge \gamma^2 - \gamma^3 \wedge \gamma^4, \; \gamma^1 \wedge \gamma^2 \wedge \gamma^5\} \subset \Lambda(C(I_\Sigma))$$

□

We conclude this section on exterior algebra with a theorem which will allows us to find the dimension of the retracting space in the special case where the generators of the ideal are a collection of 1-tensors together a single alternating 2-tensor.

Theorem 23. *Let I_Σ be an ideal generated by the set*

$$\Sigma = \{\omega^1, \ldots, \omega^s, \Omega\}$$

where $\omega^i \in V^*$ and $\Omega \in \Lambda^2(V^*)$. Let r be the smallest integer such that

$$(\Omega)^{r+1} \wedge \omega^1 \wedge \cdots \wedge \omega^s = 0$$

Then the retracting space $C(I_\Sigma)$ is of dimension $2r + s$.

Proof. See Bryant et al. [3, pages 11–12]. □

2.2. **Differential Geometry and Forms.** Since the tangent space to a differentiable manifold at each point is a vector space, we can apply to it the multilinear algebra presented in the previous section. Before doing this, we need to review some basic facts from differential geometry. The reader may wish to consult numerous books on the subject such as [1, 20, 31].

2.2.1. *Differentiable Manifolds.*

Definition 22. *A* manifold *M of dimension n is a metric space*[3] *which is locally homeomorphic to $\mathbb{R}^n$.*

A simple example of a manifold is $\mathbb{R}^n$ itself. Other examples are the circle S^1 and the sphere S^2. The circle is a one dimensional manifold while the sphere is a two dimensional manifold. Other examples of manifolds are the n-torus, $T^n = S^1 \times S^1 \times \cdots \times S^1$ and $SO(n)$, the space of unitary $n \times n$ matrices of determinant 1. A subset N of manifold M which is itself a manifold is called a submanifold of M. Any open subset N of a manifold M is clearly a submanifold since if M is locally homeomorphic to $\mathbb{R}^n$ then so is N.

In order to perform calculus on manifolds, a differentiable structure is needed. A coordinate chart on a manifold M is a pair (U, x) where U is an open set of M and x is a homeomorphism of U onto an open set of $\mathbb{R}^n$. The function x is also called a coordinate function and can be written as $(x^1, \ldots, x^n)$ where $x^i : M \longrightarrow \mathbb{R}$. If $p \in U$ then $x(p) = (x^1(p), \ldots, x^n(p))$ is called the set of local coordinates in the chart (U, x). When doing operations on a manifold, we must ensure that our results are consistent regardless of the particular chart we use. We must therefore impose some compatibility conditions. Two charts (U, x) and (V, y) with $U \cap V \neq \emptyset$, are called C^∞ compatible if the map

$$y \circ x^{-1} : x(U \cap V) \subset \mathbb{R}^n \longrightarrow y(U \cap V) \subset \mathbb{R}^n \tag{6}$$

is a C^∞ function. A C^∞ atlas on a manifold M is a collection of C^∞ compatible charts (U_α, x_α) indexed by $\alpha \in A$ such that the open sets U_α cover the manifold M. An atlas is called maximal if it is not contained in any other atlas.

Definition 23. *A* differentiable *or* smooth manifold *is a manifold with a maximal, C^∞ atlas.*

With this differential structure we can perform calculus on the manifold M. In particular let $f : M \longrightarrow \mathbb{R}$ be a function. If (U, x) is a chart on M then the function

$$\hat{f} = f \circ x^{-1} : x(U) \subset \mathbb{R}^n \longrightarrow \mathbb{R} \tag{7}$$

is called the local representative of f in the chart (U, x). We define the map f to be C^∞ or smooth if its local representative $\hat{f}$ is C^∞. Note that if f is C^∞ in one chart, then it must be C^∞ in every chart since the charts are C^∞ compatible and the atlas is maximal. Therefore these results are intrinsic to the manifold and do not depend on the particular homeomorphism chosen. Similarly, if we have a map $f : M \longrightarrow N$, where M,N are differentiable manifolds, the local representation of f given charts (U, x) of M and (V, y) of N is

$$\hat{f} = y \circ f \circ x^{-1}, \tag{8}$$

[3] Readers familiar with topology may replace metric space with Hausdorff, second countable topological space.

defined only if $f(U) \cap V \neq \emptyset$. Again f is a C^∞ map if $\hat{f}$ is a C^∞ map. Let $f : M \longrightarrow N$ be a map between two manifolds. The map f is called a diffeomorphism if both f and f^{-1} are smooth. In this case, manifolds M and N are called diffeomorphic.

Example. We have seen that $\mathbb{R}^n$ is an example of a trivial but important manifold. The differentiable structure on $\mathbb{R}^n$ consists of the chart $(\mathbb{R}^n, i)$ where i is the identity function on $\mathbb{R}^n$ as well as all other charts that are C^∞ compatible with it. We denote the standard coordinates on $\mathbb{R}^n$ as $\{r^1, \dots, r^n\}$.

The sphere, S^2 can be given a differentiable structure as follows. Consider the charts (U_N, p_N) and (U_S, p_S) where U_N is the sphere minus the North pole, U_S is the sphere minus the South pole and p_N, p_S are the stereographic projections of the sphere to the plane from the North and South poles respectively. One can show that these charts are compatible. We can then extend our atlas to a maximal one by considering all other charts that are compatible with $(U_N, p_N), (U_S, p_S)$. □

2.2.2. *Tangent Spaces.* Let p be a point on a manifold M. Let $C^\infty(p)$ denote the set of all smooth functions defined on a neighborhood of p. The set $C^\infty(p)$ is a vector space over $\mathbb{R}$ since the sum of two smooth functions and the scalar multiple of a smooth function are themselves smooth functions.

Definition 24. *A* tangent vector X_p *at* $p \in M$ *is an operator from* $C^\infty(p)$ *to* $\mathbb{R}$ *which satisfies the following properties, for* $f, g \in C^\infty(p)$ *and* $a, b \in \mathbb{R}$*:*

1. *Linearity* $X_p(a \cdot f + b \cdot g) = a \cdot X_p(f) + b \cdot X_p(g)$
2. *Derivation* $X_p(f \cdot g) = f(p) \cdot X_p(g) + X_p(f) \cdot g(p)$

The set of all tangent vectors at $p \in M$ *is called the* tangent space *of* M *at* p *and is denoted by* T_pM*.*

The tangent space T_pM becomes a vector space over $\mathbb{R}$ if for tangent vectors X_p, Y_p and real numbers c_1, c_2 we define addition and scalar multiplication as

$$(c_1 \cdot X_p + c_2 \cdot Y_p)(f) = c_1 \cdot X_p(f) + c_2 \cdot Y_p(f) \tag{9}$$

for any smooth function f in the neighborhood of p. The collection of all tangent spaces of the manifold,

$$TM = \bigcup_{p \in M} T_pM \tag{10}$$

is called the tangent bundle.

Example. Given the standard differentiable structure on $\mathbb{R}^n$, the standard tangent vectors to $\mathbb{R}^n$ at any point p are

$$\frac{\partial}{\partial r^1} \cdots \frac{\partial}{\partial r^n} \tag{11}$$

Thus given any smooth function $f(r^1, \dots, r^n) : U \longrightarrow \mathbb{R}$ where U is a neighborhood of p, we have

$$\frac{\partial}{\partial r^i}(f) = \frac{\partial f}{\partial r^i} \tag{12}$$

for $i = 1, \dots, n$. □

Now let M be a manifold and let (U,x) be a chart containing the point p. In this chart we can associate the following tangent vectors

$$\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n} \tag{13}$$

defined by

$$\frac{\partial}{\partial x^i}(f) = \frac{\partial (f \circ x^{-1})}{\partial r^i} \tag{14}$$

for any smooth function $f \in C^\infty(p)$.

Theorem 24. *Let M be an n-dimensional manifold and let T_pM be the tangent space at $p \in M$. Then T_pM is an n-dimensional vector space and if (U,x) is a local chart around p then the tangent vectors*

$$\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n} \tag{15}$$

form a basis for T_pM.

Proof. See Spivak [31, page 107]. $\square$

From this Theorem we can see that if X_p is a tangent vector at p then

$$X_p = \sum_{i=1}^{n} a_i \frac{\partial}{\partial x^i} \tag{16}$$

where $a_1, \ldots, a_n$ are real numbers. The above formula indicates that a tangent vector is an operator which simply takes the directional derivative of function in the direction of $[a_1, \ldots, a_n]$.

Now let M and N be smooth manifolds and $f : M \longrightarrow N$ be a smooth map. Let $p \in M$ and let $q = f(p) \in N$. We wish to transport tangent vectors from T_pM to T_qN using the map f. The natural way to do this is by defining a map $f_* : T_pM \longrightarrow T_qN$ by

$$(f_*(X_p))(g) = X_p(g \circ f) \tag{17}$$

for smooth functions g in the neighborhood of q. One can easily check [1] that $f_*(X_p)$ is a linear operator and a derivation and thus a tangent vector. The map $f_* : T_pM \longrightarrow T_{f(p)}N$ is called the push forward map of f.

Proposition 25. *Let $f : M \longrightarrow N$ and $g : N \longrightarrow K$. Then*

$$(g \circ f)_* = g_* \circ f_* \tag{18}$$

Proof. See Spivak [31, page 101]. $\square$

We now arrive at the important concept of a vector field on a manifold.

Definition 25. *Let M be a manifold. A* vector field *on M is a continuous function F which associates a tangent vector from T_pM to each point p of M. Such functions are called* sections *of the tangent bundle TM. If F is of class C^∞, it is called a smooth section of TM or a smooth vector field. An* integral curve *of a vector field F is a curve $c : (-\varepsilon, \varepsilon) \longrightarrow M$ such that*

$$\dot{c}(t) = F(c(t)) \in T_{c(t)}M$$

for all $t \in (-\varepsilon, \varepsilon)$.

A local expression for a vector field F in the chart (U, x) is

$$F(p) = \sum_{i=1}^{n} a_i(p) \frac{\partial}{\partial x^i} \tag{19}$$

The vector field F is C^∞ if and only if the scalar functions $a_i : M \longrightarrow \mathbb{R}$ are C^∞.

2.2.3. *Tensor Fields.* Since the tangent space to a manifold at a point is a vector space, we can apply all the multilinear algebra that we presented in the previous section to it. The dual space of T_pM at each $p \in M$ is called the cotangent space to the manifold M at p and is denoted by T_p^*M. The collection of all cotangent spaces,

$$T^*M := \bigcup_{p \in M} T_p^*M \tag{20}$$

is called the cotangent bundle. Similarly, we can form the bundles

$$\mathcal{L}^k(M) := \bigcup_{p \in M} \mathcal{L}^k(T_pM) \tag{21}$$

$$\Lambda^k(M) := \bigcup_{p \in M} \Lambda^k(T_p^*M) \tag{22}$$

Tensor fields are constructed on a manifold M by assigning to each point p of the manifold a tensor. A k-tensor field on M is a section of $\mathcal{L}^k(M)$, i.e. a function ω assigning to every $p \in M$ a k-tensor $\omega(p) \in \mathcal{L}^k(T_pM)$. At some point $p \in M$, $\omega(p)$ is a function mapping k-tuples of tangent vectors of T_pM to $\mathbb{R}$, that is $\omega(p)(X_1, X_2, \ldots, X_k) \in \mathbb{R}$ is a multi-linear function of tangent vectors $X_1, \ldots, X_k \in T_pM$. In particular, if ω is a section of $\Lambda^k(M)$ then ω is called a differential form of order k or a k-form on M. In this case, $\omega(p)$ is an alternating k-tensor at each point $p \in M$. The space of all k-forms on a manifold M will be denoted by $\Omega^k(M)$ and the space of all forms on M is simply

$$\Omega(M) := \Omega^0(M) \oplus \cdots \oplus \Omega^n(M) \tag{23}$$

At each point $p \in M$, let

$$\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n} \tag{24}$$

be the basis for T_pM. Let the 1-forms ϕ^i be the dual basis to these basis tangent vectors, i.e.

$$\phi^i(p)(\frac{\partial}{\partial x^j}) = \delta_{ij} \tag{25}$$

Recall that the forms $\phi^I = \phi^{i_1} \otimes \phi^{i_2} \otimes \ldots \otimes \phi^{i_k}$ for multi-index $I = (i_1, \ldots, i_k)$ form a basis for $\mathcal{L}^k(T_pM)$. Similarly, given an ascending multi-index $I = (i_1, \ldots, i_k)$, the k-forms $\psi^I = \phi^{i_1} \wedge \phi^{i_2} \wedge \ldots \wedge \phi^{i_k}$ form a basis for $\Lambda^k(T_pM)$. If ω is a k-tensor on M, it can be uniquely written as

$$\omega(p) = \sum_I b_I(p) \phi^I(p) \tag{26}$$

for multi-indices I and scalar functions $b_I(p)$. The k-form α can be written uniquely as

$$\alpha(p) = \sum_I c_I(p)\psi^I(p) \tag{27}$$

for ascending multi-indices I and scalar functions c_I. The k-tensor ω and k-form α are of class C^∞ if and only if the functions b_I and c_I are of class C^∞ respectively. Given two forms $\omega \in \Omega^k(M), \theta \in \Omega^l(M)$, we have,

$$\omega = \sum_I b_I\psi^I \tag{28}$$

$$\theta = \sum_J c_J\psi^J \tag{29}$$

$$\omega \wedge \theta = \sum_I \sum_J b_I c_I \psi^I \wedge \psi^J \tag{30}$$

Recall that we have defined $\Lambda^0(T_pM) = \mathbb{R}$. As a result, the space of differential forms of order 0 on M is simply the space of all functions $f : M \longrightarrow \mathbb{R}$ and the wedge product of $f \in \Omega^0(M)$ and $\omega \in \Omega^k(M)$, is defined as

$$(w \wedge f)(p) = (f \wedge w)(p) = f(p) \cdot w(p) \tag{31}$$

2.2.4. *The Exterior Derivative.* Recall that a 0-form on a manifold M is a function $f : M \to \mathbb{R}$. The differential df of a 0-form f is defined pointwise as the 1-form,

$$df(p)(X_p) = X_p(f) \tag{32}$$

It acts on a vector field X_p to give the directional derivative of f in the direction of X_p at p. As X_p is a linear operator, the operator d is also linear, that is if a, b are real numbers,

$$d(af + bg) = a \cdot df + b \cdot dg \tag{33}$$

The operator d provides a new way of expressing the elementary 1-forms $\phi^i(p)$ on T_pM. Let $x : M \longrightarrow \mathbb{R}^n$ be the coordinate function in a neighborhood of p. Consider the differentials of the coordinate functions

$$dx^i(p)(X_p) = X_p(x^i) \tag{34}$$

If we evaluate the differentials dx^i at the basis tangent vectors of T_pM we obtain,

$$dx^i(p)(\frac{\partial}{\partial x^j}) = \delta_{ij} \tag{35}$$

and therefore the $dx^i(p)$ are the dual basis of T_pM. Since the $\phi^i(p)$ are also the dual basis, $dx^i(p) = \phi^i(p)$. Thus the differentials $dx^i(p)$ span $\mathcal{L}^1(T_pM)$ and from our previous results, any k-tensor ω can be uniquely written as

$$\omega(p) = \sum_I b_I(p)dx^I(p) = \sum_I b_I(p)dx^{i_1} \otimes \cdots \otimes dx^{i_k} \tag{36}$$

for multi-indices $I = \{i_1, i_2, \dots, i_k\}$. Similarly, any k-form can be uniquely written as

$$\omega(p) = \sum_I c_I(p)dx^I(p) = \sum_I c_I(p)dx^{i_1} \wedge \cdots \wedge dx^{i_k} \tag{37}$$

for ascending multi-indices $I = \{i_1, i_2, \dots, i_k\}$. Using this basis now we have that for a 0-form,

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} dx^i \tag{38}$$

More generally, we can define an operator $d : \Omega^k(M) \longrightarrow \Omega^{k+1}(M)$ which takes k-forms to $(k+1)$-forms

Definition 26. *Let ω be a k-form on a manifold M whose representation in a chart (U, x) is given by*

$$\omega = \sum_I \omega_I dx^I \tag{39}$$

for ascending multi-indices I. The exterior derivative *or* differential operator*, d, is a linear map taking the k-form ω to the (k+1)-form $d\omega$ by*

$$d\omega = \sum_I d\omega_I \wedge dx^I \tag{40}$$

Notice that the ω_I are smooth functions (0-forms) whose differential $d\omega_I$ has already been defined as

$$d\omega_I = \sum_{j=1}^{n} \frac{\partial \omega_I}{\partial x^j} dx^j \tag{41}$$

Therefore, for any k-form ω,

$$d\omega = \sum_I \sum_{j=1}^{n} \frac{\partial \omega_I}{\partial x^j} dx^j \wedge dx^I \tag{42}$$

From the definition, this operator is certainly linear. We now prove that this differential operator is a true generalization of the operator taking 0-forms to 1-forms, satisfies some important properties, and is the unique operator with those properties.

Theorem 26. *Let M be a manifold and let $p \in M$. Then the exterior derivative is the unique linear operator*

$$d : \Omega^k(M) \to \Omega^{k+1}(M) \tag{43}$$

for $k \geq 0$, that satisfies,

1. *If f is a 0-form, then df is the 1-form*

$$df(p)(X_p) = X_p(f) \tag{44}$$

2. *If $\omega^1 \in \Omega^k(M), \omega^2 \in \Omega^\ell(M)$ then*

$$d(\omega^1 \wedge \omega^2) = d\omega^1 \wedge \omega^2 + (-1)^k \omega^1 \wedge d\omega^2 \tag{45}$$

3. *For every form ω, $d(d\omega) = 0$*

Proof. Property (1) can be easily checked from the definition of the exterior derivative. For property (2), it suffices to consider the case $\omega^1 = f dx^I$ and $\omega^2 = g dx^J$ in some chart (U, x), because of linearity of the exterior derivative.

$$\begin{aligned} d(\omega^1 \wedge \omega^2) &= d(fg) \wedge dx^I \wedge dx^J \\ &= g df \wedge dx^I \wedge dx^J + f dg \wedge dx^I \wedge dx^J \\ &= d\omega^1 \wedge \omega^2 + (-1)^k f dx^I \wedge dg \wedge dx^J \\ &= d\omega^1 \wedge \omega^2 + (-1)^k \omega^1 \wedge d\omega^2 \end{aligned}$$

For property (3), it again suffices to consider the case $\omega = f dx^I$ because of linearity. Since f is a 0-form,

$$d(df) = d(\sum_{j=1}^{n} \frac{\partial f}{\partial x^j}) dx^j = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial}{\partial x^i} \frac{\partial f}{\partial x^j} dx^i \wedge dx^j \tag{46}$$

$$= \sum_{i<j} (\frac{\partial}{\partial x^i} \frac{\partial f}{\partial x^j} - \frac{\partial}{\partial x^j} \frac{\partial f}{\partial x^i}) dx^i \wedge dx^j = 0 \tag{47}$$

We therefore have $d(df) = 0$ by the equality of mixed partial derivatives and the fact that $dx^i \wedge dx^i = 0$. If $\omega = f dx^I$ is a k-form, then $d\omega = df \wedge dx^I + f \wedge d(dx^I)$ by property (2), and since

$$d(dx^I) = d(1 \wedge dx^I) = d(1) \wedge dx^I = 0 \tag{48}$$

we get

$$d(d\omega) = d(df) \wedge dx^I - df \wedge d(dx^I) = 0 \tag{49}$$

To show that d is the unique such operator, we assume that d' is another linear operator with the same properties and then show that $d = d'$. Consider again a k-form $\omega = f dx^I$. Since d' satisfies property (2) we have

$$d'(f dx^I) = d' f \wedge dx^I + f \wedge d'(dx^I) \tag{50}$$

From the above formula we see that if we can show that $d'(dx^I) = 0$ then we will get

$$d'(f dx^I) = d' f \wedge dx^I = df \wedge dx^I = d(f dx^I) \tag{51}$$

because $d' f = df$ by property (1), and that will complete the proof. We therefore want to show that

$$d'(dx^{i_1} \wedge \ldots \wedge dx^{i_k}) = 0 \tag{52}$$

But since both d and d' satisfy property (1) we have

$$dx^I = dx^{i_1} \wedge \cdots \wedge dx^{i_k} = d'x^{i_1} \wedge \cdots \wedge d'x^{i_k} = d'x^I \tag{53}$$

since the coordinate functions x^i are 0-forms. Then

$$d'(dx^{i_1} \wedge \cdots \wedge dx^{i_k}) = d'(d'x^{i_1} \wedge \cdots \wedge d'x^{i_k}) = 0 \tag{54}$$

since d' satisfies property (3). $\square$

Now let $f : M \longrightarrow N$ be a smooth map between two manifolds. We have seen that the push forward map, f_*, is a linear transformation from T_pM to $T_{f(p)}N$. Therefore given tensors or forms on $T_{f(p)}N$ we can use the pull back transformation[4], f^*, in order to define tensors or forms on T_pM. The next theorem shows that the exterior derivative and the pull back transformation commute.

Theorem 27. *Let $f : M \longrightarrow N$ be a smooth map between manifolds. If ω is a k-form on N then*

$$f^*(d\omega) = d(f^*\omega) \tag{55}$$

Proof. See Spivak [31, pages 295–296]. ☐

The d-operator may be used to define two classes of forms of particular interest.

Definition 27. *A k-form $\omega \in \Omega^k(M)$ is said to be* closed *if $d\omega \equiv 0$.*

Definition 28. *A k-form $\omega \in \Omega^k(M)$ with $k > 0$ is* exact *if there exists a $(k-1)$-form θ such that $\omega = d\theta$. A 0-form is exact on any open set if it is constant on that set.*

Clearly, since every exact form is closed. However, not all closed forms are exact.

2.2.5. *The Interior Product.* We can define the interior product of a tensor field and a vector field pointwise as the interior product of a tensor and a tangent vector.

Definition 29. *Given a k-form $\omega \in \Omega^k(M)$ and a vector field X the* interior product *or* anti-derivation *of ω with X is a $(k-1)$ form defined pointwise by*

$$(X(p) \lrcorner \omega(p))(v_1, \ldots, v_{k-1}) = \omega(p)(X(p), v_1, \ldots, v_{k-1}) \tag{56}$$

Definition 30. *Given a function $h : M \longrightarrow \mathbb{R}$, the* Lie derivative *of h along the vector field X is denoted as $L_X h$ and is defined by*

$$L_X h = X(h) = X \lrcorner dh \tag{57}$$

The Lie derivative is simply the directional derivative of the function h along the vector field X.

Definition 31. *Given two vector fields X and Y, their* Lie bracket *is defined to be the vector field such that for each $h \in C^\infty(p)$ we have*

$$[X, Y](h) = X(Y(h)) - Y(X(h)) = X \lrcorner d(Y \lrcorner dh) - Y \lrcorner d(X \lrcorner dh) \tag{58}$$

In particular, if we choose the coordinate functions x^i, we get

$$[X, Y](x^i) = [X, Y]_i = \sum_j \frac{\partial Y_i}{\partial x^j} X_j - \sum_j \frac{\partial X_i}{\partial x^j} Y_j \tag{59}$$

and we therefore obtain

$$[X, Y](x) = \frac{\partial Y}{\partial x} X(x) - \frac{\partial X}{\partial x} Y(x) \tag{60}$$

The Lie bracket is skew symmetric

$$[X, Y] = -[Y, X] \tag{61}$$

[4] to be consistent with our previous notation, we should write $(f_*)^*$ to denote the pull back of f_*. Notation is abused, however, and we simply denote it by f^*.

and satisfies the Jacobi identity

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0 \tag{62}$$

The following lemma establishes a relation between the exterior derivative and Lie brackets.

Lemma 28. *(Cartan's Magic Formula). Let $\omega \in \Omega^1(M)$ and X, Y smooth vector fields. Then*

$$d\omega(X,Y) = X(\omega(Y)) - Y(\omega(x)) - \omega([X,Y])$$

Proof. Because of linearity, it is adequate to consider $\omega = fdg$ where f, g are functions. The left hand side of the above formula is

$$\begin{aligned} d\omega(X,Y) &= df \wedge dg(X,Y) \\ &= df(X)\cdot dg(Y) - df(Y)\cdot dg(X) \\ &= X(f)\cdot Y(g) - Y(f)\cdot X(g) \end{aligned}$$

while the right hand side is

$$\begin{aligned} X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]) &= X(fY(g)) - Y(fX(g)) - f(XY(g) - YX(g)) \\ &= X(f)\cdot Y(g) - Y(f)\cdot X(g) \end{aligned}$$

which completes the proof. □

2.2.6. *Distributions and Codistributions.* Recall that a vector field is a map which assigns a tangent vector to each point on the manifold. In the case of multiple vector fields, one may assign a number of tangent vectors at a point and look at the subspace of the tangent space spanned by these vectors. This assignment, which places at each point of the manifold a subspace of the tangent space at that point, is called a distribution and is denoted by $\Delta(p) = span\{f_1(p), \ldots, f_d(p)\}$ or, if we drop the dependence on the point p,

$$\Delta = span\{f_1, \ldots, f_d\} \tag{63}$$

Since distributions are subspaces one can define the sum or intersection of two distributions as the sum or intersection of the respective subspaces. If the vector fields are smooth, we call $\Delta(p)$ a smooth distribution. The dimension of the distribution at a point is defined to be the dimension of the subspace $\Delta(p)$. A distribution is said to regular if its dimension does not vary with p. A vector field f belongs to a distribution Δ if $f(p) \in \Delta(p)$ for all p.

A distibution is involutive if given any two vector fields f_1 and f_2 belonging to the distribution, their Lie bracket also belongs to the distribution, i.e.

$$f_1, f_2 \in \Delta \Longrightarrow [f_1, f_2] \in \Delta \tag{64}$$

or more compactly, $[\Delta, \Delta] \subset \Delta$. A distribution Δ is called integrable if there exists a submanifold N of M such that the tangent space of N at x equals $\Delta(x)$. The submanifold N is called the integral manifold of the distribution Δ. The following theorem provides us with a condition under which a distribution is integrable.

Theorem 29. *(Frobenius Theorem for distributions) A regular distribution $\Delta(x)$ is integrable if and only if it is involutive.*

Proof. See Spivak [31, page 261]. □

Similarly, one may also assign to each point of the manifold a set of 1-forms. The span of these 1-forms at each point will be a subspace of the cotangent space T_p^*M. This assignment is called a codistribution and is denoted by $\Theta(p) = span\{\omega_1(p), \dots, \omega_d(p)\}$ or, dropping the dependence on the point p,

$$\Theta = span\{\omega_1, \dots, \omega_d\} \tag{65}$$

where $\omega_1, \dots, \omega_d$ are the 1-forms which generate this codistribution.

There is a notion of duality between distributions and codistributions which allows us to construct codistributions from distributions and vice versa. Given a distribution Δ, for each p in a neighborhood U, consider all the 1-forms which pointwise annihilate all vectors in $\Delta(p)$,

$$\Delta^\perp(p) = span\{\omega(p) \in T_p^*M : \omega(p)(f) = 0 \; \forall f \in \Delta(p)\} \tag{66}$$

Clearly, $\Delta^\perp(p)$ is a subspace of T_p^*M and is therefore a codistribution. We call $\Delta^\perp$ the annihilator or dual of Δ. Conversely, given a codistribution Θ, we construct the annihilating or dual distribution pointwise as

$$\Theta^\perp(p) = span\{v \in T_pM : \omega(p)(v) = 0 \; \forall \omega(p) \in \Omega(p)\} \tag{67}$$

If N is an integral manifold of a distribution Δ and v is a vector in the distribution Δ at a point p (and consequently in T_pN), then for any $\alpha \in \Delta^\perp$, $\alpha(p)(v) = 0$. Notice that this must also be true for any integral curve of the distribution. Therefore given a codistribution $\Theta = span\{\omega_1, \dots, \omega_s\}$, an integral curve of the codistribution is a curve $c(t)$ whose tangent $c'(t)$ at each point satisfies, for $i = 1, \dots, s$,

$$\omega_i(c(t))(c'(t)) = 0 \tag{68}$$

Example. Consider the following kinematic model of a unicycle

$$\begin{aligned} \dot{x} &= u_1 \cos\theta \\ \dot{y} &= u_1 \sin\theta \\ \dot{\theta} &= u_2 \end{aligned}$$

which can be written as

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix} u_1 + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u_2 = f_1 u_1 + f_2 u_2 \tag{69}$$

The corresponding control distribution is

$$\Delta(x) = span\{ \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \} \tag{70}$$

while the dual codistribution is

$$\Delta^\perp = span\{\omega\} \tag{71}$$

where $\omega = \sin\theta dx - \cos\theta dy + 0 d\theta$, the nonholonomic constraint of rolling without slipping. $\square$

2.3. Exterior Differential Systems.

2.3.1. *The Exterior Algebra On a Manifold.* The space of all forms on a manifold M,

$$\Omega(M) = \Omega^0(M) \oplus \cdots \oplus \Omega^n(M)$$

together with the wedge product is called the exterior algebra on M. An algebraic ideal of this algebra is defined as in Section 2.1.8 as a subspace I such that if $\alpha \in I$ then $\alpha \wedge \beta \in I$ for any $\beta \in \Omega(M)$.

Definition 32. *An ideal $I \subset \Omega(M)$ is said to be* closed *with respect to exterior differentiation if and only if*

$$\alpha \in I \Longrightarrow d\alpha \in I$$

or more compactly $dI \subset I$. A algebraic ideal which is closed with respect to exterior differentiation is called a differential ideal.

A finite collection of forms, $\Sigma := \{\alpha^1, \ldots, \alpha^K\}$ generates an algebraic ideal

$$I_\Sigma := \{\omega \in \Omega(M) \mid \omega = \sum_{i=1}^{K} \theta^i \wedge \alpha^i \ \text{ for some } \ \theta^i \in \Omega(M)\}.$$

We can also talk about the differential ideal generated by Σ.

Definition 33. *Let S_d denote the collection of all differential ideals containing Σ. The* differential ideal generated by Σ *is defined as the smallest differential ideal containing Σ*

$$\mathcal{I}_\Sigma := \bigcap_{I \in S_d} I$$

Theorem 30. *Let Σ be a finite collection of forms, and let $\mathcal{I}_\Sigma$ denote the differential ideal generated by Σ. Define the collection*

$$\Sigma' = \Sigma \cup d\Sigma$$

and denote the algebraic ideal which it generates by $I_{\Sigma'}$. Then

$$\mathcal{I}_\Sigma = I_{\Sigma'}$$

Proof. By definition, $\mathcal{I}_\Sigma$ is closed with respect to exterior differentiation, so $\Sigma' \subset \mathcal{I}_\Sigma$. Consequently, $I_{\Sigma'} \subset \mathcal{I}_\Sigma$. The ideal $I_{\Sigma'}$ is a closed with respect to exterior differentiation and contains Σ by construction. Therefore, from the definition of $\mathcal{I}_\Sigma$ we have that $\mathcal{I}_\Sigma \subset I_{\Sigma'}$. □

The associated space and retracting space of an ideal in $\Omega(M)$ are defined pointwise as in section 2.1.10. The associated space of $\mathcal{I}_\Sigma$ is called the Cauchy characteristic distribution and is denoted $A(\mathcal{I}_\Sigma)$.

2.3.2. *Exterior Differential Systems.* In Section 2.1.10 we introduced systems of exterior equations on a vector space V and characterized their solutions as subspaces of V. We are now ready to define a similar notion for a collection of differential forms defined on a manifold M. The basic problem will be to study the integral submanifolds of M which satisfy the constraints represented by the exterior differential system.

Definition 34. *An* exterior differential system *is a finite collection of equations*

$$\alpha^1 = 0, \dots, \alpha^r = 0$$

where each $\alpha^i \in \Omega^k(M)$ is a smooth k-form. A solution to an exterior differential system is any submanifold N of M which satisfies $\alpha^i(x)|_{T_xN} \equiv 0$. for all $x \in N$ and all $i \in \{1, \dots, r\}$.

An exterior differential system can be viewed pointwise as a system of exterior equations on T_pM. In view of this, one might expect that a solution would be defined as a distribution on the manifold. The trouble with this approach is that most distributions are not integrable, and we want our solution set to be a collection of integral submanifolds. Therefore, we will restrict our solution set to integrable distributions.

Theorem 31. *Given an exterior differential system*

$$\alpha^1 = 0, \dots, \alpha^K = 0 \tag{72}$$

and the corresponding differential ideal $\mathcal{I}_\Sigma$ generated by the collection of forms

$$\Sigma := \{\alpha^1, \dots, \alpha^K\} \tag{73}$$

An integral submanifold N of M solves the system of exterior equations if and only if it also solves the equation $\pi = 0$ for every $\pi \in \mathcal{I}_\Sigma$.

Proof. If an integral submanifold N of M is a solution to Σ, then for all $x \in N$ and all $i \in 1, \dots, K$

$$\alpha^i(x)|_{T_xN} \equiv 0.$$

Taking the exterior derivative gives

$$d\alpha^i(x)|_{T_xN} \equiv 0.$$

Therefore, the submanifold also satisfies the exterior differential system

$$\alpha^1 = 0, \dots, \alpha^K = 0, d\alpha^1 = 0, \dots, d\alpha^K = 0$$

From Theorem 30 we know that the differential ideal generated by Σ is equal to the algebraic ideal generated by the above system. Therefore, Theorem 21 tells us that every solution N to Σ is also a solution for every element of $\mathcal{I}_\Sigma$. Conversely, if N solves the equation $\pi = 0$ for every $\pi \in \mathcal{I}_\Sigma$ then in particular it must solve Σ. $\square$

This theorem allows us to either work with the generators of an ideal or with the ideal itself. In fact some authors define exterior differential systems as differential ideals of $\Omega(M)$. Because a set of generators Σ generates both a differential ideal $\mathcal{I}_\Sigma$ and a algebraic ideal I_Σ, we can define two different notions of equivalence for exterior differential systems.

Definition 35. *Two exterior differential systems, Σ_1 and Σ_2, are said to be* algebraically equivalent *if they generate the same algebraic ideal. i.e. $I_{\Sigma_1} = I_{\Sigma_2}$.*

Definition 36. *Two exterior differential systems, Σ_1 and Σ_2, are said to be* equivalent *if they generate the same differential ideal. i.e. $\mathcal{I}_{\Sigma_1} = \mathcal{I}_{\Sigma_2}$.*

Intuitively, we want to think of two exterior differential systems as equivalent if they have the same solution set. Therefore, we will usually discuss equivalence in terms of this second definition.

2.3.3. *Pfaffian Systems.* Pfaffian systems are of particular interest because they can be used to represent a set of first-order ordinary differential equations.

Definition 37. *An exterior differential system of the form*

$$\alpha^1 = \alpha^2 = \cdots = \alpha^s = 0$$

where the α^i are independent 1-forms on an n-dimensional manifold is called a Pfaffian system *of codimension $n - s$. If $\{\alpha^1, \ldots, \alpha^n\}$ is a basis for $\Omega^1(M)$, then the set $\{\alpha^{s+1}, \ldots, \alpha^n\}$ is called a* complement *to the Pfaffian system.*

An independence condition is a one-form τ which is required to be nonzero along integral curves of the Pfaffian system. That is, if $\alpha^i(c(t))(c'(t)) = 0$, then $\tau(c(t))(c'(t)) \neq 0$ The 1-forms $\alpha^1, \ldots, \alpha^s$, generate the algebraic ideal

$$I = \{\sigma \in \Omega(M) : \sigma \wedge \alpha^1 \wedge \cdots \wedge \alpha^s = 0\}$$

The algebraic ideal generated by the 1-forms α^i is also a differential ideal if the following conditions are satisfied.

Definition 38. *A set of linearly independent 1-forms $\alpha^1, \ldots, \alpha^s$ in the neighborhood of a point is said to satisfy the* Frobenius condition *if one of the following equivalent conditions hold:*

1. *$d\alpha^i$ is a linear combination of $\alpha^1, \ldots, \alpha^s$.*
2. *$d\alpha^i \wedge \alpha^1 \wedge \cdots \wedge \alpha^s = 0$ for $1 \leq i \leq s$.*
3. *$d\alpha^i = \sum_{j=1}^{s} \theta^j \wedge \alpha^j$*

When $d\alpha^i$ is a linear combination of $\alpha^1, \ldots, \alpha^s$ the following expression is frequently used

$$d\alpha^i \equiv 0 \; mod \; \alpha^1, \ldots, \alpha^s \; 1 \leq i \leq s$$

where the mod operation is implicitly performed over the algebraic ideal generated by the α^i.

Example. We will illustrate the above concepts for the unicycle. Recall that the unciycle can be described by the following codistribution

$$I = \{\omega\}$$

where

$$\omega = \sin\theta dx - \cos\theta dy + 0 d\theta$$

The exterior derivative of ω is

$$d\omega = \cos\theta \; d\theta \wedge dx + \sin\theta \; d\theta \wedge dy$$

and therefore

$$d\omega \wedge \omega = -\cos^2 d\theta \wedge dx \wedge dy + \sin^2 d\theta \wedge dy \wedge dx = -dx \wedge dy \wedge d\theta \neq 0$$

Since the second condition of Definition 38 is not satisfied, I is not a differential ideal. □

Theorem 32. *(Frobenius Theorem for codistributions) Let I be an algebraic ideal generated by the independent 1-forms $\alpha^1, \ldots, \alpha^s$ which satisfies the Frobenius condition. Then in a neighborhood of x there exist functions h^i with $1 \leq i \leq s$ such that*

$$I = \{\alpha^1, \ldots, \alpha^s\} = \{dh^1, \ldots, dh^s\}$$

Proof. See Bryant et al. [3, pages 27–29]. □

For more general exterior differential systems we have the following integrability results.

Theorem 33. *If the Cauchy characteristic distribution $A(\mathcal{I}_\Sigma)$ of $\mathcal{I}_\Sigma$ has constant dimension r in a neighborhood, then the distribution $A(\mathcal{I}_\Sigma)$ is integrable.*

Proof. See Bryant et al. [3, page 31]. □

Theorem 34. *Let $\mathcal{I}$ be a differential ideal whose retracting space $C(\mathcal{I})$ has a constant dimension $n-r$. There is a neighborhood in which there are coordinates $y^1,\ldots,y^n$ such that $\mathcal{I}$ has a set of generators which are forms in $y^1,\ldots,y^{n-r}$.*

Proof. See Bryant et al. [3, pages 31–33]. □

2.3.4. *Derived flags.* If the algebraic ideal generated by a Pfaffian system does not satisfy the Frobenius condition, then it is not a differential ideal. However, there may exist a differential ideal which is a subset of the algebraic ideal. This subideal will can be found by taking the derived flag of the Pfaffian system. Let $I^{(0)} = \{\omega^1,\ldots,\omega^s\}$ be the algebraic ideal generated by independent 1-forms $\omega^1,\ldots,\omega^s$. We define $I^{(1)}$ as

$$I^{(1)} = \{\lambda \in I^{(0)} : d\lambda \equiv 0 \; mod \; I^{(0)}\} \subset I^{(0)}$$

The ideal $I^{(1)}$ is called the first derived system. The analog of the first derived system from the distribution point of view is given by the following theorem.

Theorem 35. *If $I^{(0)} = \Delta^\perp$ then $I^{(1)} = (\Delta + [\Delta,\Delta])^\perp$.*

Proof. Let $I^{(0)}$ be spanned by 1-forms $\omega^1,\ldots,\omega^s$ and let Δ be its annihilating distribution. By definition we have that

$$I^{(1)} = \{\omega \in I^{(0)} : d\omega \equiv 0 \; mod \; I^{(0)}\}$$

Let $\eta \in I^{(1)}$. Therefore $d\eta \equiv 0 \; mod \; I^{(0)}$ which means that

$$d\eta = \sum_{j=1}^{s} \theta^j \wedge \omega^j$$

for some forms θ^j. Now let X, Y be vector fields in Δ. Since Δ is the annihilating distribution of $I^{(0)}$, $\omega^j(X) = \omega^j(Y) = 0$. Also, $\eta \in I^{(1)} \subset I^{(0)}$, and therefore $\eta(X) = \eta(Y) = 0$. Now, using the expression for $d\eta$,

$$\begin{aligned} d\eta(X,Y) &= \sum_{j=1}^{s} \theta^j \wedge \omega^j(X,Y) \\ &= \sum_{j=1}^{s} \theta^j(X)\omega^j(Y) - \theta^j(Y)\omega^j(X) \\ &= 0 \end{aligned}$$

Cartan's magic formula gives

$$d\eta(X,Y) = X\eta(Y) - Y\eta(X) - \eta([X,Y]) = 0$$

and therefore

$$\eta([X,Y]) = 0$$

which means that η annihilates any vector fields belonging in $[\Delta, \Delta]$ in addition to any vector fields in Δ. Therefore $\eta \in (\Delta + [\Delta, \Delta])^{\perp}$ and thus

$$I^{(1)} \subset (\Delta + [\Delta, \Delta])^{\perp}$$

To show the other inclusion, let $\eta \in (\Delta + [\Delta, \Delta])^{\perp}$ and let X,Y be vector fields in Δ. Cartan's magic formula gives

$$d\eta(X, Y) = X\eta(Y) - Y\eta(X) - \eta([X, Y]) = 0$$

and therefore $d\eta = 0 \ mod \ I^{(0)}$ which means that $\eta \in I^{(1)}$. Thus $(\Delta + [\Delta, \Delta])^{\perp} \subset I^{(1)}$ and therefore $(\Delta + [\Delta, \Delta])^{\perp} = I^{(1)}$. □

One may inductively continue this procedure of obtaining derived systems and define

$$I^{(2)} = \{\lambda \in I^{(1)} : d\lambda \equiv 0 \ mod \ I^{(1)}\} \subset I^{(1)}$$

or in general

$$I^{(k+1)} = \{\lambda \in I^{(k)} : d\lambda \equiv 0 \ mod \ I^{(k)}\} \subset I^{(k)}$$

This procedure results in a nested sequence of codistibutions

$$I^{(k)} \subset I^{(k-1)} \subset \cdots \subset I^{(1)} \subset I^{(0)} \tag{74}$$

We can also generalize Theorem 35. If we define $\Delta_0 = (I^{(0)})^{\perp}$, $\Delta_1 = (I^{(1)})^{\perp}$, and in general $\Delta_k = (I^{(k)})^{\perp}$, then it is not hard to show that if $I^{(k)} = \Delta_k^{\perp}$ then $I^{(k+1)} = (\Delta_k + [\Delta_k, \Delta_k])^{\perp}$. The proof of this fact is similar to the proof of Theorem 35 but uses a more general form of Cartan's magic formula. The sequence of decreasing codistributions (74), called the derived flag of $I^{(0)}$, is associated with a sequence of increasing distributions, called the filtration of Δ_0,

$$\Delta_k \supset \Delta_{k-1} \supset \cdots \supset \Delta_1 \supset \Delta_0$$

If the dimension of each codistribution is constant then there will be an integer N such that $I^{(N)} = I^{(N+1)}$. This integer N is called the derived length of I. A basis for a codistribution I is simply a set of generators for I. A basis of 1-forms α^j for I is said to be adapted to the derived flag if a basis for each derived system $I^{(j)}$ can be chosen to be some subset of the α^j's. The codistribution $I^{(N)}$ is always integrable by definition since

$$dI^{(N)} \equiv 0 \ mod \ I^{(N)}$$

Codistribution $I^{(N)}$ is the largest integrable subsystem in I. Therefore, if $I^{(N)} \neq \{0\}$ then there exist functions $h^1, \ldots, h^r$ such that $\{dh^1, \ldots, dh^r\} \subset I$. As a result, if a Pfaffian system contains an integrable subsystem $I^{(N)} \neq 0$ which is spanned by the 1-forms $dh^1, \ldots, dh^r$, then the integral curves of the system are constrained to satisfy the following equations for some constants k_i.

$$dh^i = 0 \Longrightarrow h^i = k_i \text{ for } 1 \leq i \leq r$$

or equivalently, trajectories of the system must lie on the manifold,

$$M = \{x : h^i(x) = k_i \text{ for } 1 \leq i \leq r\}$$

In particular, this implies that if $I^{(N)} \neq 0$, it is not possible to find an integral curve of the Pfaffian system which connects a configuration $x(0) = x_0$ to another configuration $x(1) = x_1$ unless the initial and final configurations satisfy

$$h^i(x_0) = h^i(x_1) \text{ for } 1 \leq i \leq r$$

Example. Consider the rolling penny system. In addition to the three configuration variables of the unicycle, we also have an angle ϕ describing the orientation of Lincoln's head. The model in this case, assuming for simplicity that the penny has unit radius, is given by

$$\begin{aligned}\dot{x} &= u_1 \cos\theta\\ \dot{y} &= u_1 \sin\theta\\ \dot{\theta} &= u_2\\ \dot{\phi} &= -u_1\end{aligned}$$

which can be written in vector field notation as

$$\begin{bmatrix}\dot{x}\\ \dot{y}\\ \dot{\theta}\\ \dot{\phi}\end{bmatrix} = \begin{bmatrix}\cos\theta\\ \sin\theta\\ 0\\ -1\end{bmatrix} u_1 + \begin{bmatrix}0\\ 0\\ 1\\ 0\end{bmatrix} u_2 = f_1 u_1 + f_2 u_2$$

The annihilating codistribution to the distribution $\Delta_0 = \{f_1, f_2\}$ can be easily determined to be

$$I = \Delta^{\perp} = \{\alpha^1, \alpha^2\}$$

where

$$\alpha^1 = \cos\theta dx + \sin\theta dy + 0d\theta + 1d\phi$$

$$\alpha^2 = \sin\theta dx - \cos\theta dy + 0d\theta + 0d\phi$$

To compute the derived systems, we must first take the exterior derivatives of the constraints.

$$\begin{aligned}d\alpha^2 &= \cos\theta d\theta \wedge dx + \sin\theta d\theta \wedge dy\\ d\alpha^1 &= -\sin\theta d\theta \wedge dx + \cos\theta d\theta \wedge dy\\ d\alpha^2 \wedge \alpha^1 \wedge \alpha^2 &= d\theta \wedge dx \wedge dy \wedge d\phi\\ d\alpha^1 \wedge \alpha^2 &= \sin\theta\cos\theta(d\theta \wedge dx \wedge dy + d\theta \wedge dy \wedge dx) = 0\\ d\alpha^1 \wedge \alpha^1 \wedge \alpha^2 &= 0\end{aligned}$$

From these wedge products, we can see that

$$\begin{aligned}d\alpha^1 &= 0 \; mod \; \alpha^1, \alpha^2\\ d\alpha^2 &\neq 0 \; mod \; \alpha^1, \alpha^2\end{aligned}$$

and thus the first derived system is spanned by α^1,

$$I^{(1)} = \{\alpha^1\}$$

It can be easily checked that

$$d\alpha^1 \wedge \alpha^1 \neq 0$$

and thus

$$I^{(2)} = \{0\}$$

The derived flag of the system is given by the decreasing sequence of codistributions

$$I^{(0)} = \{\alpha^1, \alpha^2\}$$
$$I^{(1)} = \{\alpha^1\}$$
$$I^{(2)} = \{0\}$$

Note that the basis is adapted to the derived flag. Because $I^{(2)} = \{0\}$, an integrable subsystem does not exist. The system is not constrained to move on some submanifold of $\mathbb{R}^4$. □

3. Normal Forms

Now that we have defined an exterior differential system, and introduced some analysis tools, we are ready to study some important normal forms for exterior differential systems. We will restrict ourselves to Pfaffian systems. The first normal form which we introduce, the Pfaffian form, is restricted to systems of only one equation. The Engel form applies to two equations on a four-dimensional space, and the Goursat form is for $n-2$ equations on an n-dimensional space. The extended Goursat normal form is defined for systems with codimension greater than two. The Goursat normal forms can be thought of as the generalization of linear systems. Their study will lead us to the study of linearization of control systems in Section 4.

3.1. **The Goursat Normal Form.**

3.1.1. *Systems of one equation.* We will first study Pfaffian systems of codimension $n-1$, or systems consisting of a single equation

$$\alpha = 0$$

where α is a 1-form on a manifold M. In some chart (U, x) of a point $p \in M$ the equation can be expressed as

$$a_1(x)dx^1 + a_2(x)dx^2 + \cdots + a_n(x)dx^n = 0$$

In order to understand the integral manifolds of this equation we will attempt to express α in a normal form by performing a coordinate transformation.

Definition 39. *Let $\alpha \in \Omega^1(M)$. The integer r defined by*

$$\begin{array}{rcl} (d\alpha)^r \wedge \alpha & \neq & 0 \\ (d\alpha)^{r+1} \wedge \alpha & = & 0 \end{array}$$

is called the rank *of α.*

The following theorem allows us, under a rank condition, to write α in a normal form.

Theorem 36. *(Pfaff) Let $\alpha \in \Omega^1(M)$ have constant rank r in a neighborhood of p. Then there exists a coordinate chart (U, z) such that in these coordinates, $\alpha = dz^1 + z^2dz^3 + \ldots + z^{2r}dz^{2r+1}$.*

Proof. Let $\mathcal{I}$ be the differential ideal generated by α. From Theorem 23 the retracting space of $\mathcal{I}$ is of dimension $2r + 1$. By Theorem 34 there exist local coordinates

$y^1, \ldots, y^n$ such that $\mathcal{I}$ has a set of generators in $y^1, \ldots, y^{2r+1}$. Then, by dimension count, any function f_1 of those $2r+1$ coordinates results in

$$(d\alpha)^r \wedge \alpha \wedge df_1 = 0$$

Now let $\mathcal{I}_1$ be the ideal generated by $\{df_1, \alpha, d\alpha\}$. If $r = 0$ then the result follows from the Frobenius Theorem. If $r > 0$, the forms df_1 and α must be linearly independent since α is not integrable. Applying Theorem 23 to $\mathcal{I}_1$, let r_1 be the smallest integer such that

$$(d\alpha)^{r_1+1} \wedge \alpha \wedge df_1 = 0$$

Clearly, $r_1 + 1 \leq r$. Furthermore, the equality sign must hold because $(d\alpha)^r \wedge \alpha \neq 0$. Applying Theorem 34 to $\mathcal{I}_1$ there exists a function f_2 such that

$$(d\alpha)^{r-1} \wedge \alpha \wedge df_1 \wedge df_2 = 0$$

Repeating this process, we find r functions $f_1, f_2, \ldots, f_r$ satisfying

$$d\alpha \wedge \alpha \wedge df_1 \wedge df_2 \wedge \ldots \wedge df_r = 0$$

$$\alpha \wedge df_1 \wedge df_2 \wedge \ldots \wedge df_r \neq 0$$

Finally, let $\mathcal{I}_r$ be the ideal $\{df_1, \ldots, df_r, \alpha, d\alpha\}$. Its retraction space $\mathcal{C}(\mathcal{I}_r)$ is of dimension $r+1$. There is a function f_{r+1} such that

$$\alpha \wedge df_1 \wedge df_2 \wedge \ldots \wedge df_{r+1} = 0$$

$$df_1 \wedge df_2 \wedge \ldots \wedge df_{r+1} \neq 0$$

By modifying α by a factor, we can write

$$\alpha = df_{r+1} + g_1 df_1 + \ldots + g_r df_r.$$

Because $(d\alpha)^r \wedge \alpha \neq 0$, the functions $f_1, \ldots, f_{r+1}, g_1, \ldots, g_r$ are independent. The result then follows by setting

$$z^1 \;=\; f_{r+1} \qquad z^{2i} = g_i \qquad z^{2i+1} = f_i$$

for $1 \leq i \leq r$. $\square$

Example. Consider the unicycle example described by the codistribution $I = \{\alpha\}$ where $\alpha = \sin\theta dx - \cos\theta dy$. We can immediately see that

$$d\alpha = \cos\theta d\theta \wedge dx + \sin\theta d\theta \wedge dy$$

and that

$$\begin{aligned} d\alpha \wedge \alpha &= d\theta \wedge dy \wedge dx \neq 0 \\ (d\alpha)^2 \wedge \alpha &= 0 \end{aligned}$$

Therefore α has rank 1 and by Pfaff's Theorem there exist coordinates z^1, z^2, z^3 such that

$$\alpha = dz^1 + z^2 dz^3$$

In this example we trivially obtain,

$$\alpha = dy + (-\tan\theta)dx$$

$\square$

The following theorem is similar to Pfaff's theorem and simply expresses the result in a more symmetric form.

Theorem 37. *Given any $\alpha \in \Omega^1(M)$ with constant rank r in a neighborhood U of p, there exist coordinates $z, y^1, \ldots, y^r, x^1, \ldots, x^r$ such that*

$$\alpha = dz + \frac{1}{2}\sum_{i=1}^{r}(y^i dx^i - x^i dy^i)$$

Proof. The following coordinate transformation

$$\begin{aligned} z^1 &= z - \frac{1}{2}\sum_{i=1}^{r} x^i y^i \\ z^{2i} &= y^i \qquad 1 \le i \le r \\ z^{2i+1} &= x^i \qquad 1 \le i \le r \end{aligned}$$

reduces the above Theorem to Pfaff's Theorem. □

The Pfaffian system $\alpha = 0$ on a manifold M is said to have the local accessibility property if every point $x \in M$ has a neighborhood U such that every point in U can be joined to x by an integral curve. The following theorem answers the question of when does this Pfaffian system have the local accessibility property.

Theorem 38. *(Caratheodory) The Pfaffian system,*

$$\alpha = 0$$

where α has constant rank, has the local accessibility property if and only if

$$\alpha \wedge d\alpha \neq 0$$

Proof. The above condition simply says that the rank of α must be greater than or equal to 1. If α has zero rank then $d\alpha \wedge \alpha = 0$ and therefore by Frobenius Theorem we can write

$$\alpha = dh = 0$$

for some function h. The integral curves are of the form $h = c$ for any arbitrary constant c. Since we can only join points $p, q \in M$ for which $h(p) = h(q)$, we do not have the local accessibility property.

Conversely, let α have rank $r \geq 1$. From Theorem 37, we can find coordinates $z, x^1, \ldots, x^r, y^1, \ldots, y^r, u^1, \ldots, u^s$ in some neighborhood U with $2r + s + 1 = \dim M$ such that

$$\alpha = dz + \frac{1}{2}\sum_{i=1}^{r}(y^i dx^i - x^i dy^i) = 0$$

and therefore

$$dz = \frac{1}{2}\sum_{i=1}^{r}(x^i dy^i - y^i dx^i)$$

Given any two points $p, q \in U$ we must find integral curves $c : [0,1] \longrightarrow U$ with $c(0) = p$ and $c(1) = q$. Since we are working locally, we can assume that the initial point p is the origin: $z(p) = x^i(p) = y^i(p) = u^i(p) = 0$. Let the final point q be defined by $z(q) = z_1, x^i(q) = x_1^i, y^i(q) = y_1^i, u^i(q) = u_1^i$. Because the expression of the one-form α does not depend on the u^i coordinates, we can choose the curve tu_1^i to connect the u^i coordinates of p and q.

In the (x^i, y^i) plane there are many curves $(x^i(t), y^i(t))$ which join the origin with the desired point (x_1^i, y_1^i). We need to find one which steers the z coordinate to z_1. In order to satisfy the equation $\alpha = 0$, we must have that

$$dz = \frac{1}{2}\sum_{i=1}^{r}(x^i dy^i - y^i dx^i)$$

Integrating this equation gives

$$z(t) = \frac{1}{2}\int_0^t \sum_{i=1}^{r}(x^i \frac{dy^i}{dt} - y^i \frac{dx^i}{dt})dt = \frac{1}{2}\sum_{i=1}^{r} A_i$$

where A_i is the area enclosed by the curve $(x^i(t), y^i(t))$ and the chord joining the origin with (x_1^i, y_1^i). To reach the point q, the curve $(x^i(t), y^i(t))$ must satisfy $z(1) = z_1$. Geometrically, it is clear that a curve $(x^i(t), y^i(t))$ linking the points p and q while enclosing the area prescribed by z_1 will always exist. Thus, the integral curve $c(t)$ given by

$$(z(t), x^1(t), \ldots, x^r(t), y^1(t), \ldots, y^r(t), tu^1(t), \ldots, tu^s(t))$$

has $c(0) = p$ and $c(1) = q$ and satisfies the equation $\alpha = 0$, and the system therefore has the local accessibility property. $\square$

3.1.2. *Codimension two systems.* We now consider Pfaffian systems of codimension two. We are again interested in performing coordinate changes so that the generators of these Pfaffian systems are in some normal form.

Theorem 39. *(Engel) Let I be a two dimensional codistribution*

$$I = \{\alpha^1, \alpha^2\}$$

of four variables. If the derived flag satisfies

$$dim I^{(1)} = 1$$

$$dim I^{(2)} = 0$$

then there exist coordinates z^1, z^2, z^3, z^4 such that

$$I = \{dz^4 - z^3 dz^1, dz^3 - z^2 dz^1\}$$

Proof. Choose a basis for I which is adapted to the derived flag; that is $I^{(0)} = I = \{\alpha^1, \alpha^2\}$, $I^{(1)} = \{\alpha^1\}$, and $I^{(2)} = \{0\}$. Choose α^3 and α^4 to complete the basis. Since $I^{(2)} = \{0\}$ we have

$$d\alpha^1 \wedge \alpha^1 \neq 0$$

while

$$(d\alpha^1)^2 \wedge \alpha^1 = 0$$

since it is a 5-form on a 4-dimensional space. Therefore α^1 has rank 1. By Pfaff's Theorem, we know that there exists a coordinate change so that

$$\alpha^1 = dz^4 - z^3 dz^1$$

Taking the exterior derivative, we have that

$$d\alpha^1 = -dz^3 \wedge dz^1 = dz^1 \wedge dz^3$$

Now, since $\alpha^1 \in I^{(1)}$, the definition of the first derived system will imply that

$$d\alpha^1 \wedge \alpha^1 \wedge \alpha^2 = 0$$

and thus

$$dz^1 \wedge dz^3 \wedge \alpha^1 \wedge \alpha^2 = 0$$

Therefore α^2 must be a linear combination of dz^1, dz^3 and α^1:

$$\alpha^2 = a(x)dz^3 + b(x)dz^1 \; mod \; \alpha^1$$

By definition, this means that

$$\alpha^2 + \lambda(x)\alpha^1 = a(x)dz^3 + b(x)dz^1$$

Now if either $a(x) = 0$ or $b(x) = 0$, then $d\alpha^2 \wedge \alpha^1 \wedge \alpha^2 = 0$ and thus the flag assumptions are violated. Thus $a(x) \neq 0$ and therefore

$$\frac{1}{a(x)}\alpha^2 + \frac{\lambda(x)}{a(x)}\alpha^1 = dz^3 + \frac{b(x)}{a(x)}dz^1$$

and if we set $z^2 = -\frac{b(x)}{a(x)}$ then

$$\frac{1}{a(x)}\alpha^2 + \frac{\lambda(x)}{a(x)}\alpha^1 = dz^3 - z^2 dz^1$$

and thus

$$I = \{\alpha^1, \alpha^2\} = \{\alpha^1, \frac{1}{a(x)}\alpha^2 + \frac{\lambda(x)}{a(x)}\alpha^1\} = \{dz^4 - z^3 dz^1, dz^3 - z^2 dz^1\}$$

□

It should be noted that the only place the dimension assumption is used in the proof is to guarantee that $(d\alpha^1)^2 \wedge \alpha^1 = 0$. If α^1 has rank 1 this equality holds by definition.

Corollary 40. *Let $I = \{\alpha^1, \alpha^2\}$ be a two dimensional codistribution. If the derived flag satisfies $dim I^{(1)} = 1$ and $dim I^{(2)} = 0$ and $\alpha^1 \in I^{(1)}$ has rank 1, then there exist coordinates z^1, z^2, z^3, z^4 such that*

$$I = \{dz^4 - z^3 dz^1, dz^3 - z^2 dz^1\}$$

Proof. The corollary follows by the proof of Engel's Theorem. □

Example. Consider again the penny rolling on a plane. The system is described by the codistribution $I = \{\alpha^1, \alpha^2\}$ where

$$\begin{aligned} \alpha^1 &= \cos\theta dx + \sin\theta dy + d\phi \\ \alpha^2 &= \sin\theta dx - \cos\theta dy \end{aligned}$$

In Example 2.3.4, we saw that the derived flag for this system is given by

$$\begin{aligned} I^{(0)} &= \{\alpha^1, \alpha^2\} \\ I^{(1)} &= \{\alpha^1\} \\ I^{(2)} &= \{0\} \end{aligned}$$

and thus satisfies the conditions of Engel's Theorem. After some calculations we obtain

$$d\alpha^1 \wedge \alpha^1 = -dx \wedge dy \wedge d\theta + \sin\theta d\theta \wedge dx \wedge d\phi + \cos\theta d\theta \wedge dy \wedge d\phi$$

Since $(d\alpha^1)^2 \wedge \alpha^1 = 0$ the rank of α^1 is 1. Following the proof of Pfaff's Theorem we know that there exists a function f_1 such that

$$d\alpha^1 \wedge \alpha^1 \wedge df_1 = 0$$

We can easily see that the function $f_1 = \theta$ is a solution to this equation. Since the rank of α^1 is 1, we must now search for a function f_2 such that

$$\alpha^1 \wedge df_1 \wedge df_2 = 0$$

Let $f_2 = f_2(x, y, \theta, \phi)$. Then it may be verified that a solution to this system of equations is

$$f_2(x, y, \theta, \phi) = x \cos\theta + y \sin\theta + \phi$$

Therefore, following again the proof of Pfaff's Theorem we may now choose $z^1 = f_1$ and $z^4 = f_2$ so that

$$\alpha^1 = dz^4 - z^3 dz^1$$

where z^3 can be found from the above equation to be

$$z^3 = -x \sin\theta + y \cos\theta$$

We will now try to transform α^2 into the normal form. Following the proof of Engel's Theorem, we have that

$$\alpha^2 = [a(x, y, \theta, \phi) dz^3 + b(x, y, \theta, \phi) dz^1] \; mod \; \alpha^1$$

We must now determine the functions a and b. Simple calculations show that the following choices

$$\begin{aligned} a(x, y, \theta, \phi) &= -1 \\ b(x, y, \theta, \phi) &= -x \cos\theta - y \sin\theta \end{aligned}$$

will satisfy the equation. Therefore by Engel's Theorem, if we set

$$z^2 = -x \cos\theta - y \sin\theta = -\frac{b(x, y, \theta, \phi)}{a(x, y, \theta, \phi)}$$

we may express α^1,α^2 in the following normal form

$$\begin{aligned} \alpha^1 &= dz^4 - z^3 dz^1 \\ \alpha^2 &= dz^3 - z^2 dz^1 \end{aligned}$$

If we look at the differential equation expressed in these new coordinates we obtain

$$\begin{aligned} \dot{z}^4 &= z^3 \dot{z}^1 \\ \dot{z}^3 &= z^2 \dot{z}^1 \end{aligned}$$

and z^1, z^2 are free. The annihilating distribution is given by

$$\Delta = \left\{ \begin{bmatrix} 1 \\ 0 \\ z^2 \\ z^3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\} = \{g_1, g_2\}$$

If we set $\dot{z}^1 = u_1, \dot{z}^2 = u_2$ the distribution has the form

$$\begin{aligned} \dot{z}^4 &= z^3 u_1 \\ \dot{z}^3 &= z^2 u_1 \\ \dot{z}^2 &= u_2 \\ \dot{z}^1 &= u_1 \end{aligned}$$

or $\dot{z} = g_1 u_1 + g_2 u_2$. The advantage of performing this coordinate transformation is that our system can be expressed in this simple form. In particular if we set

$u_1 = 1$ then the system has been transformed to a linear system in Brunovsky canonical form. This allows us to use the powerful analysis tools that exist for linear systems. Engel's Theorem therefore gives the conditions under which a system of four configuration variables with two constraints can be "linearized". □

Engel's Theorem can be generalized to a system with n configuration variables and $n-2$ constraints. This powerful theorem was proved by Goursat.

Theorem 41. *(Goursat Normal Form) Let I be a Pfaffian system spanned by s 1-forms,*

$$I = \{\alpha^1, \dots, \alpha^s\}$$

on a space of dimension $n = s+2$. Suppose that there exists an integrable form π with $\pi \neq 0$ mod I satisfying the Goursat congruences,

$$\begin{aligned} d\alpha^i &\equiv -\alpha^{i+1} \wedge \pi \; mod \; \alpha^1, \dots, \alpha^i \quad 1 \leq i \leq s-1 \\ d\alpha^s &\neq 0 \; mod \; I \end{aligned} \tag{75}$$

Then there exists a coordinate system $z^1, z^2, \dots, z^n$ in which the Pfaffian system is in Goursat normal form:

$$I = \{dz^3 - z^2 dz^1, dz^4 - z^3 dz^1, \dots, dz^n - z^{n-1} dz^1\}$$

Proof. The Goursat congruences can be expressed as

$$\begin{aligned} d\alpha^1 &\equiv -\alpha^2 \wedge \pi \; mod \; \alpha^1 \\ d\alpha^2 &\equiv -\alpha^3 \wedge \pi \; mod \; \alpha^1, \alpha^2 \\ &\vdots \\ d\alpha^{s-1} &\equiv -\alpha^s \wedge \pi \; mod \; \alpha^1, \alpha^2, \dots, \alpha^s \\ d\alpha^s &\equiv -\alpha^{s+1} \wedge \pi \; mod \; \alpha^1, \alpha^2, \dots, \alpha^s \end{aligned}$$

where $\alpha^{s+1} \notin I$. It can be shown that $\{\alpha^{s+1}, \pi\}$ must form a complement to I. This basis satisfies the Goursat congruences and is adapted to the derived flag of I:

$$\begin{aligned} I^{(0)} &= \{\alpha^1, \alpha^2, \dots, \alpha^s\} \\ I^{(1)} &= \{\alpha^1, \dots, \alpha^{s-1}\} \\ &\vdots \\ I^{(s-1)} &= \{\alpha^1\} \\ I^{(s)} &= \{0\} \end{aligned}$$

From the Goursat congruences,

$$d\alpha^1 = -\alpha^2 \wedge \pi \; mod \; \alpha^1$$

which means that

$$d\alpha^1 = -\alpha^2 \wedge \pi + \alpha^1 \wedge \eta$$

for some one-form η. But then we have that

$$\begin{aligned} d\alpha^1 \wedge \alpha^1 &= -\alpha^2 \wedge \pi \wedge \alpha^1 \neq 0 \\ (d\alpha^1)^2 \wedge \alpha^1 &= 0 \end{aligned}$$

which means that α^1 has rank 1. We can therefore apply Pfaff's Theorem and express α^1 as

$$\alpha^1 = dz^n - z^{n-1}dz^1$$

for some choice of z^1, z^{n-1}, z^n. Furthermore, by Corollary 40 we can express α^2 as

$$\alpha^2 = dz^{n-1} - z^{n-2}dz^1 \tag{76}$$

In these new coordinates we have

$$d\alpha^1 \wedge \alpha^1 = -dz^{n-1} \wedge dz^1 \wedge dz^n$$

Now we have that

$$d\alpha^1 \wedge \alpha^1 \wedge \pi = \pi \wedge (-dz^{n-1} \wedge dz^1 \wedge dz^n) = \pi \wedge (-\alpha^2 \wedge \pi \wedge \alpha^1) = 0$$

and therefore π is a linear combination of dz^1, dz^{n-1}, dz^n. Noting that $dz^{n-1} \equiv z^{n-2}dz^1 \mod \alpha^1, \alpha^2$,

$$\begin{aligned}\pi =& adz^1 + bdz^{n-1} + cdz^n \\ =& adz^1 + bz^{n-2}dz^1 + cz^{n-1}dz^1 \; mod \; \alpha^1, \alpha^2 \\ =& \psi dz^1 \; mod \; \alpha^1, \alpha^2\end{aligned}$$

where $\psi = a + bz^{n-2} + cz^{n-1}$ is nonzero since we have assumed that $\pi \neq 0 \; mod \; I$. From the Goursat congruences we have that

$$d\alpha^2 = -\alpha^3 \wedge \pi \; mod \; \alpha^1, \alpha^2$$

while from equation (76) we have

$$d\alpha^2 = -dz^{n-2} \wedge dz^1$$

and thus

$$-dz^{n-2} \wedge dz^1 = -\alpha^3 \wedge \pi \; mod \; \alpha^1, \alpha^2$$

which means that

$$\alpha^3 = \lambda(x) dz^{n-2} \; mod \; dz^1, \alpha^1, \alpha^2$$

for nonzero function $\lambda(x)$. Therefore we can rewrite this as

$$\alpha^3 = dz^{n-2} - \frac{1}{\lambda(x)} dz^1 \; mod \; \alpha^1, \alpha^2$$

and if we set $z^{n-3} = \frac{1}{\lambda(x)}$ we have

$$\alpha^3 = dz^{n-2} - z^{n-3}dz^1 \; mod \; \alpha^1, \alpha^2$$

and we can therefore let

$$\alpha^3 = dz^{n-2} - z^{n-3}dz^1$$

If we inductively continue this procedure using the Goursat congruences we obtain

$$\begin{aligned}\alpha^4 &= dz^{n-3} - z^{n-4}dz^1 \\ &\vdots \\ \alpha^s &= dz^3 - z^2dz^1\end{aligned}$$

Now from the Goursat congruences we have that

$$d\alpha^s \neq 0 \; mod \; I$$

and therefore

$$\alpha^1 \wedge \alpha^2 \wedge \cdots \wedge \alpha^s \wedge \ d\alpha^s \neq 0$$

If we substitute the α^i in the above expression we obtain

$$dz^1 \wedge dz^2 \wedge \cdots \wedge dz^n \neq 0$$

and therefore the functions $z^1, \ldots, z^n$ can serve as a local coordinate system. □

The following example illustrates the power of the Goursat's Theorem by applying it in order to linearize a nonlinear system. A more systematic approach to the feedback linearization problem can be found in the paper by Gardner and Shadwick [11]. Note that the integral curves of a system in Goursat normal form are completely determined by two arbitrary functions in one variable and their derivatives. For example, once $z^1(\tau)$ and $z^s(\tau)$ are known, all of the other coordinates are determined from

$$z^i = \frac{\dot{z}^{i+1}(\tau)}{\dot{z}^1(\tau)}$$

where the dot indicated the standard derivative with respect to the independent variable τ. Because of this property, these two coordinates are sometimes referred to as linearizing outputs for the Pfaffian system.

Example. Consider the following nonlinear system with s configuration variables and a single input,

$$\begin{array}{rcl} \dot{x}_1 & = & f_1(x_1, \ldots, x_s, u) \\ \dot{x}_2 & = & f_2(x_1, \ldots, x_s, u) \\ & \vdots & \\ \dot{x}_s & = & f_s(x_1, \ldots, x_s, u) \end{array}$$

Equivalently we can look at following Pfaffian system,

$$I = \{dx^i - f_i(x^1, \ldots, x^s, u)dt\} \ \ 1 \leq i \leq s$$

The system is of codimension 2 since we have s constraints and $s+2$ variables, namely $x^1, \ldots, x^s, u, t$. Assume that the form $\pi = dt$ satisfies the Goursat congruences. Then by Goursat's Theorem there exists a coordinate transformation $z = \Phi(x, u, t)$ such that I is generated by

$$I = \{dz^3 - z^2 dz^1, dz^4 - z^3 dz^1, \ldots, dz^{s+2} - z^{s+1} dz^1\}$$

The annihilating distribution of the above codistribution is

$$\begin{array}{rcl} \dot{z}^1 & = & v_1 \\ \dot{z}^2 & = & v_2 \\ \dot{z}^3 & = & z^2 v_1 \\ & \vdots & \\ \dot{z}^{s+2} & = & z^{s+1} v_1 \end{array}$$

which, if we set $v_1 = 1$, is clearly a linear system. If it turned out that the z^1 coordinate corresponds to time in the original coordinates, that is, $z^1 = t$, then the connection becomes even more clear. Goursat's Theorem can thus be used to linearize single-input nonlinear systems which satisfy the Goursat congruences.

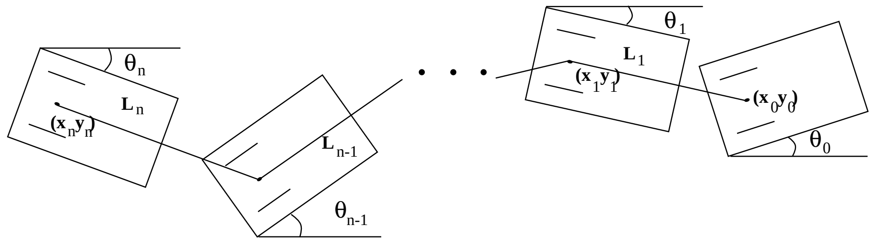

FIGURE 1. The mobile robot Hilare with n trailers.

These and other issues related to control systems will be explored more fully in Section 4. □

3.2. **The N-trailer Pfaffian system.** In this section, we will show how the system of a mobile robot towing n trailers can be represented as a Pfaffian system. As we saw in the unicycle example, the constraint that a wheel rolls without slipping can be represented as a one-form on the configuration manifold. The velocity of the N-trailer system is constrained in n directions corresponding to the n axles of wheels. A basis for this constraint codistribution (or equivalently, the Pfaffian system) is found by writing down the rolling without slipping conditions for all n axles.

3.2.1. *The system of rolling constraints and its derived flag.* Consider a single-axle mobile robot such as Hilare[5] with n trailers attached, as sketched in Figure 1. Each trailer is attached to the body in front of it by a rigid bar, and the rear set of wheels of each body is constrained to roll without slipping. The trailers are assumed to be identical, with possibly different link lengths L_i. The x, y coordinates of the midpoint between the two wheels on the i^{th} axle are referred to as (x^i, y^i) and the hitch angles (all measured with respect to the horizontal) are given by θ^i. The connections between the bodies give rise to the following relations:

$$\begin{aligned} x^{i-1} &= x^i + L_i \cos\theta^i \\ y^{i-1} &= y^i + L_i \sin\theta^i \end{aligned} \qquad i = 1, 2, \ldots, n. \tag{77}$$

Obviously, the space parameterized by the coordinates $(x^0, y^0, \theta^0, \ldots, x^n, y^n, \theta^n) \in \mathbb{R}^{2n+2} \times (S^1)^{n+1}$ is not reachable. Taking into account the connection relations (77), any one of the Cartesian positions x^i, y^i together with all the hitch angles $\theta^0, \ldots, \theta^n$ will completely represent the configuration of the system. The configuration space is thus $M = \mathbb{R}^2 \times (S^1)^{n+1}$ and has dimension $n + 3$. In any neighborhood, the configuration space can be parameterized by $\mathbb{R}^{n+3}$.

The velocity constraints on the system arise from constraining the wheels of the robot and trailers to roll without slipping; the velocity of each body in the direction perpendicular to its wheels must be zero. Each pair of wheels is modeled as a single wheel at the midpoint of the axle. Each velocity constraint can be written as a one-form,

$$\alpha^i = \sin\theta^i dx^i - \cos\theta^i dy^i \quad i = 0, \ldots, n \tag{78}$$

The one-forms $\alpha^0, \alpha^1, \ldots, \alpha^n$ represent the constraints that the wheels of the zeroth trailer (i.e. the cab), the first trailer, ..., the n^{th} trailer, respectively roll

[5]The Hilare family of mobile robots resides at LAAS in Toulouse, see for example [7, 12].

without slipping. The Pfaffian system corresponding to this mobile robot system is generated by the codistribution spanned by all of the rolling without slipping constraints:

$$I = \{\alpha^0, \dots, \alpha^n\} \tag{79}$$

and has dimension $n+1$ on a manifold of dimension $n+3$.

Before finding the derived flag associated with I, it is useful to investigate some properties of the constraints and their exterior derivatives. Notice that equation (78) can be rearranged (after a division by a cosine) to give the congruence:

$$dy^i \equiv \tan\theta^i dx^i \mod \alpha^i \tag{80}$$

This division by a cosine introduces a singularity; the resulting coordinate transformation will not be valid at points where $\theta^i = \pm\pi/2$. See Remark 3.2.2 for a brief discussion of singularities.

All of the (x^i, y^i) are related by the hitch relationships. The exterior derivatives of these relationships can be taken,

$$\begin{aligned} x^{i-1} &= x^i + L_i\cos\theta^i \\ y^{i-1} &= y^i + L_i\sin\theta^i \end{aligned} \quad \Longrightarrow \quad \begin{aligned} dx^{i-1} &= dx^i - L_i\sin\theta^i d\theta^i \\ dy^{i-1} &= dy^i + L_i\cos\theta^i d\theta^i \end{aligned}$$

and these expressions can then be substituted into the formula for α^{i-1} from (78), allowing the constraint for the $(i-1)^{st}$ axle to be rewritten as:

$$\begin{aligned} \alpha^{i-1} &= \sin\theta^{i-1}dx^{i-1} - \cos\theta^{i-1}dy^{i-1} && (81)\\ &= \sin\theta^{i-1}dx^i - \cos\theta^{i-1}dy^i - L_i\cos(\theta^{i-1}-\theta^i)d\theta^i && (82)\\ &\equiv (\sin\theta^{i-1} - \tan\theta^i\cos\theta^{i-1})dx^i - L_i\cos(\theta^{i-1}-\theta^i)d\theta^i \mod \alpha^i && (83)\\ &\equiv \sec\theta^i\sin(\theta^{i-1}-\theta^i)dx^i - L_i\cos(\theta^{i-1}-\theta^i)d\theta^i \mod \alpha^i && (84) \end{aligned}$$

after an application of the congruence (80). A rearrangement of terms and a division by cosine in equation (81) will give the congruence

$$\begin{aligned} d\theta^i &\equiv \frac{1}{L_i}\sec\theta^i\tan(\theta^{i-1}-\theta^i)dx^i \mod \alpha^i, \alpha^{i-1} && (85)\\ &\equiv f_{\theta^i}dx^i \mod \alpha^i, \alpha^{i-1} && (86) \end{aligned}$$

The exact form of the function f_{θ^i} is unimportant; what will be needed is the relationship between $d\theta^i$ and dx^i.

The first lemma relates the exterior derivatives of the x coordinates,

Lemma 42. *The exterior derivatives of any of the x variables are congruent modulo the Pfaffian system, that is: $dx^i \equiv f_{x^{i,j}}dx^j \mod I$.*

Proof. For two adjacent axles, the relationship between the x coordinates is given by the hitching,

$$\begin{aligned} x^{i-1} &= x^i + L_i\cos\theta^i && (87)\\ dx^{i-1} &= dx^i - L_i\sin\theta^i d\theta^i && (88)\\ &\equiv (1 - L_i\sin\theta^i f_{\theta^i})dx^i \mod \alpha^{i-1}, \alpha^i && (89)\\ &\equiv f_{x^{i-1,i}}dx^i \mod \alpha^{i-1}, \alpha^i && (90) \end{aligned}$$

The congruence (85) was applied. □

A complement to the Pfaffian system $I = \{\alpha^0, \dots, \alpha^n\}$ is given by

$$\{d\theta^0, dx^i\} \tag{91}$$

for any x^i, since by Lemma 42 their exterior derivatives are congruent modulo the system, and the complement is only defined modulo the system. These two one-forms, together with the codistribution I, form a basis for the space of all one-forms on the configuration manifold, or $\Omega^1(M)$.

Now consider the exterior derivative of the constraint corresponding to the i^{th} axle,

$$\begin{aligned}
\alpha^i &= \sin\theta^i dx^i - \cos\theta^i dy^i && (92)\\
d\alpha^i &= d\theta^i \wedge (\cos\theta^i dx^i + \sin\theta^i dy^i) && (93)\\
&\equiv d\theta^i \wedge dx^i(\cos\theta^i + \sin\theta^i \tan\theta^i) \quad \mod \alpha^i && (94)\\
&\equiv d\theta^i \wedge dx^i(\sec\theta^i) \quad \mod \alpha^i && (95)\\
&\equiv 0 \quad \mod \alpha^i, \alpha^{i-1} && (96)
\end{aligned}$$

using (85). Thus, the exterior derivative of the constraint corresponding to the i^{th} axle is congruent to zero modulo itself and the constraint corresponding to the axle directly in front of it. The congruences (80) and (85) were useful in deriving this result.

This is all the information that is needed to find the derived flag for the system.

Theorem 43 (Derived flag for the N-trailer Pfaffian system)**.** *Consider the Pfaffian system of the N-trailer system* (79) *with the one forms α^i defined by equations* (78)*. The one-forms α^i are adapted to the derived flag in the following sense:*

$$\begin{aligned}
I^{(0)} &= \{\alpha^0, \alpha^1, \dots, \alpha^n\} && (97)\\
I^{(1)} &= \{\alpha^1, \dots, \alpha^n\} && (98)\\
&\vdots && (99)\\
I^{(n)} &= \{\alpha^n\} && (100)\\
I^{(n+1)} &= \{0\} && (101)
\end{aligned}$$

Proof. The proof is merely a repeated application of equation (92). Noting that the exterior derivative of the i^{th} constraint is equal to zero modulo itself and the constraint corresponding to the axle directly in front of it, it is simple to check that the derived flag has the form given in equation (97). □

Note that $I^{(n+1)} = \{0\}$ implies that there is no integrable subsystem contained in the constraints which define N-trailer Pfaffian system.

3.2.2. *Conversion to Goursat normal form.* In the preceding subsection, it was shown that basis $\{\alpha^0, \dots, \alpha^n\}$ defined in equation (78) is adapted to its derived flag in the sense of (97). It remains to be checked whether the α^i satisfy the Goursat congruences and if they do, to find a transformation that puts them into the Goursat canonical form. The following theorem guarantees the existence of such a transformation.

Theorem 44 (Goursat congruences for the N-trailer system). *Consider the Pfaffian system $I = \{\alpha^0, \dots, \alpha^n\}$ associated with the N-trailer system (79) with the one-forms α^i defined by equation (78). There exists a change of basis of the one-forms α^i to $\bar{\alpha}^i$ which preserves the adapted structure, and a one-form π which satisfies the Goursat congruences for this new basis:*

$$\begin{aligned} &d\bar{\alpha}^i \equiv -\bar{\alpha}^{i-1} \wedge \pi \mod \bar{\alpha}^i, \dots, \bar{\alpha}^n \qquad i = 1, \dots, n \\ &d\bar{\alpha}^0 \not\equiv 0 \qquad\qquad\qquad\qquad\qquad\qquad \mod I. \end{aligned}$$

A one-form which satisfies these congruences is given by

$$\pi = dx^n \tag{102}$$

Proof. First of all, consider the original basis of constraints. The expression for α^i can be written in the configuration space coordinates from equation (78) together with the connection relations (77) and some bookkeeping as:

$$\alpha^i = \sin\theta^i dx^n - \cos\theta^i dy^n - \sum_{k=i+1}^{n} L_k \cos(\theta^i - \theta^k) d\theta^k \tag{103}$$

Before beginning the main part of the proof, it will be helpful to define a new basis of constraints $\tilde{\alpha}^i$, which is also adapted to the derived flag, but is somewhat simpler to work with. Each $\tilde{\alpha}^i$ will have only two terms. Although the last constraint already has only two terms, it will be scaled by a factor,

$$\tilde{\alpha}^n = \sec\theta^n \alpha^n = \tan\theta^n dx^n - dy^n \tag{104}$$

Note that a rearrangement of terms will give the congruence

$$dy^n \equiv \tan\theta^n dx^n \mod \tilde{\alpha}^n \tag{105}$$

Now consider the next to last constraint, α^{n-1}, and apply the preceding congruence:

$$\begin{aligned} \alpha^{n-1} &= \sin\theta^{n-1} dx^n - \cos\theta^{n-1} dy^n - L_n \cos(\theta^n - \theta^{n-1}) d\theta^n \\ &\equiv \sec\theta^n \sin(\theta^{n-1} - \theta^n) dx^n - L_n \cos(\theta^{n-1} - \theta^n) d\theta^n \mod \tilde{\alpha}^n \end{aligned} \tag{106}$$

Dividing once again by a cosine, the new basis element $\tilde{\alpha}^{n-1}$ is defined as

$$\tilde{\alpha}^{n-1} = \sec\theta^n \tan(\theta^{n-1} - \theta^n) dx^n - L_n d\theta^n \tag{107}$$

Thus, $\tilde{\alpha}^{n-1} \equiv f_{\alpha^{n-1}} \alpha^{n-1} \mod \alpha^n$. Also, the exterior derivative $d\theta^n$ is related to dx^n by the congruence

$$d\theta^n \equiv \frac{1}{L_n} \sec\theta^n \tan(\theta^{n-1} - \theta^n) dx^n \mod \tilde{\alpha}^{n-1} \tag{108}$$

This procedure of eliminating the terms $dy^n, d\theta^n, \dots, d\theta^i$ from α^{i+1} can be continued.

Lemma 45. *A new basis of constraints $\tilde{\alpha}^i$ of the form*

$$\begin{aligned} &\tilde{\alpha}^n = \tan\theta^n dx^n - dy^n \\ &\tilde{\alpha}^i = \sec\theta^n \sec(\theta^{n-1} - \theta^n) \dots \sec(\theta^{i+1} - \theta^{i+2}) \tan(\theta^i - \theta^{i+1}) dx^n - L_{i+1} d\theta^{i+1} \\ &\qquad i = 0, \dots, n-1 \end{aligned} \tag{109}$$

is related to the original basis of constraints α^i through the following congruences:

$$\tilde{\alpha}^i \equiv f_{\alpha^i} \alpha^i \mod \alpha^{i+1}, \dots, \alpha^n \tag{110}$$

and thus the basis $\tilde{\alpha}^i$ is also adapted to the derived flag.

Note that by the definition of $\tilde{\alpha}^i$, the exterior derivative $d\theta^{i+1}$ is related to dx^n by the congruence

$$d\theta^{i+1} \equiv \frac{1}{L_{i+1}} \sec\theta^n \sec(\theta^{n-1}-\theta^n)\ldots\sec(\theta^{i+1}-\theta^{i+2})\tan(\theta^i-\theta^{i+1})dx^n \mod \tilde{\alpha}^i \tag{111}$$

The lemma is proved by induction. It has already been shown that $\tilde{\alpha}^n = f_{\alpha^n}\alpha^n$ and $\tilde{\alpha}^{n-1} \equiv f_{\alpha^{n-1}}\alpha^{n-1} \mod \tilde{\alpha}^n$. Assume that $\tilde{\alpha}^i \equiv f_{\alpha^i}\alpha^i \mod \alpha^{i+1},\ldots,\alpha^n$ for $i = j+1,\ldots,n$. Consider $\tilde{\alpha}^j$ as defined by equation (109),

$$\tilde{\alpha}^j = \sec\theta^n \sec(\theta^{n-1}-\theta^n)\ldots\sec(\theta^{j+1}-\theta^{j+2})\tan(\theta^j-\theta^{j+1})dx^n - L_{j+1}d\theta^{j+1} \tag{112}$$

Recall from equation (103) that α^j has the form

$$\alpha^j = \sin\theta^j dx^n - \cos\theta^j dy^n - \sum_{k=j+1}^{n} L_k\cos(\theta^j-\theta^k)d\theta^k \tag{113}$$

Now, applying the congruences

$$\begin{aligned} dy^n &\equiv \tan\theta^n dx^n && \mod \tilde{\alpha}^n \\ d\theta^i &\equiv \frac{1}{L_i}\sec\theta^n\sec(\theta^{n-1}-\theta^n)\ldots\sec(\theta^i-\theta^{i+1})\tan(\theta^{i-1}-\theta^i)dx^n && \mod \tilde{\alpha}^{i-1} \end{aligned}$$

to the expression for α^j, and expanding the summation, yields

$$\begin{aligned} \alpha^j \equiv{} & \sin\theta^j dx^n - \cos\theta^j\tan\theta^n dx^n - L_{j+1}\cos(\theta^j-\theta^{j+1})d\theta^{j+1} \\ & - \cos(\theta^j-\theta^{j+2})\sec\theta^n\sec(\theta^{n-1}-\theta^n)\ldots\sec(\theta^{j+2}-\theta^{j+3})\tan(\theta^{j+1}-\theta^{j+2})dx^n \\ & - \cdots \\ & - \cos(\theta^j-\theta^{n-1})\sec\theta^n\sec(\theta^{n-1}-\theta^n)\tan(\theta^{n-2}-\theta^{n-1})dx^n \\ & - \cos(\theta^j-\theta^n)\sec\theta^n\tan(\theta^{n-1}-\theta^n)dx^n \\ & \qquad \mod \tilde{\alpha}^{j+1},\ldots,\tilde{\alpha}^{n-2},\tilde{\alpha}^{n-1},\tilde{\alpha}^n \end{aligned} \tag{114}$$

To simplify the above expression, the trigonometric identity

$$\sin a - \cos a\tan b = \sec b\sin(a-b) \tag{115}$$

is repeatedly applied. After all the terms are collected, it can be seen that the equation will read:

$$\begin{aligned} \alpha^j \equiv{} & \sin(\theta^j-\theta^{j+1})\sec\theta^n\sec(\theta^{n-1}-\theta^n)\ldots\sec(\theta^{j+1}-\theta^{j+2})dx^n \\ & - L_{j+1}\cos(\theta^j-\theta^{j+1})d\theta^{j+1} \qquad \mod \tilde{\alpha}^{j+1},\ldots,\tilde{\alpha}^{n-2},\tilde{\alpha}^{n-1},\tilde{\alpha}^n \\ \equiv{} & \cos(\theta^j-\theta^{j+1})\tilde{\alpha}^j \mod \tilde{\alpha}^{j+1},\ldots,\tilde{\alpha}^{n-2},\tilde{\alpha}^{n-1},\tilde{\alpha}^n \end{aligned} \tag{116}$$

and the lemma is proved.

The basis $\tilde{\alpha}^i$ will now be scaled to find the basis $\bar{\alpha}^i$ which will satisfy the congruences (75). Once again, the procedure will start with the last congruence, $\tilde{\alpha}^n$. The exterior derivative of $\tilde{\alpha}^n$ is given by

$$d\tilde{\alpha}^n = \sec^2\theta^n d\theta^n \wedge dx^n \tag{117}$$

Looking at the expression for $\tilde{\alpha}^{n-1}$ given in equation (107), it can be seen that π should be chosen to be some multiple of dx^n or $d\theta^n$. In fact, either $\pi = dx^n$ or $\pi = d\theta^n$ will work, although the computations are different for each case. The calculations here are for choosing $\pi = dx^n$. Choosing the new basis element $\bar{\alpha}^{n-1}$ as

$$\bar{\alpha}^{n-1} = \frac{1}{L_n} \sec^2 \theta^n \tilde{\alpha}^{n-1} \tag{118}$$

will result in the desired congruence,

$$d\bar{\alpha}^n \equiv -\bar{\alpha}^{n-1} \wedge \pi \mod \alpha^n \tag{119}$$

Now consider the exterior derivative of $\bar{\alpha}^{n-1}$,

$$\begin{aligned} d\bar{\alpha}^{n-1} &= d(\frac{1}{L_n} \sec^3 \theta^n \tan(\theta^{n-1} - \theta^n) dx^n - L_n \sec^2 \theta^n d\theta^n) \\ &\equiv \frac{1}{L_n} \sec^3 \theta^n \sec^2(\theta^{n-1} - \theta^n) d\theta^{n-1} \wedge dx^n \mod \tilde{\alpha}^{n-1} \end{aligned} \tag{120}$$

since any terms $d\theta^n \wedge dx^n$ are congruent to 0 mod $\tilde{\alpha}^{n-1}$. Thus, in order to achieve the next Goursat congruence $d\bar{\alpha}^{n-1} \equiv \bar{\alpha}^{n-2} \wedge \pi$, the new basis element $\bar{\alpha}^{n-2}$ should be chosen as

$$\bar{\alpha}^{n-2} = \frac{1}{L_n L_{n-1}} \sec^3 \theta^n \sec^2(\theta^{n-1} - \theta^n) \tilde{\alpha}^{n-2} \tag{121}$$

In general, the new basis is defined by

$$\begin{aligned} \bar{\alpha}^i \quad = \quad & \frac{1}{L_n L_{n-1} \dots L_{i+1}} \sec^{n-i+1} \theta^n \sec^{n-i}(\theta^{n-1} - \theta^n) \dots \\ & \sec^3(\theta^{i+2} - \theta^{i+3}) \sec^2(\theta^{i+1} - \theta^{i+1}) \tilde{\alpha}^i \end{aligned} \tag{122}$$

It has already been shown that the congruences hold for $i = n$ and $i = n - 1$. Assume that the congruences

$$d\bar{\alpha}^i = -\bar{\alpha}^{i-1} \wedge \pi \mod \alpha^i, \dots, \alpha^n. \tag{123}$$

hold for $i = j + 1, \dots, n$. Consider the exterior derivative of $\bar{\alpha}^j$,

$$\begin{aligned} d\bar{\alpha}^j = d\big[(&\frac{1}{L_n L_{n-1} \dots L_{j+1}} \sec^{n-j+1} \theta^n \sec^{n-j}(\theta^{n-1} - \theta^n) \dots \\ & \sec^3(\theta^{j+2} - \theta^{j+3}) \sec^2(\theta^{j+1} - \theta^{j+2})) \tilde{\alpha}^i \end{aligned}$$

Before calculating all of the terms, recall that the following congruences hold:

$$\begin{aligned} d\theta^i \wedge dx^n &\equiv 0 \mod \tilde{\alpha}^{i-1} \\ d\theta^i \wedge d\theta^k &\equiv 0 \mod \tilde{\alpha}^{i-1}, \tilde{\alpha}^{k-1} \end{aligned} \tag{124}$$

and thus the only term in $d\bar{\alpha}^j \mod \tilde{\alpha}^j, \dots, \tilde{\alpha}^n$ will be a multiple of $d\theta^j \wedge dx^n$,

$$\begin{aligned} d\bar{\alpha}^j \equiv & \frac{1}{L_n L_{n-1} \dots L_{j+1}} \sec^{n-j+2} \theta^n \sec^{n-j+1}(\theta^{n-1} - \theta^n) \dots \\ & \sec^4(\theta^{j+2} - \theta^{j+3}) \sec^3(\theta^{j+1} - \theta^{j+2}) \sec^2(\theta^j - \theta^{j+1}) \\ & \quad \mod \tilde{\alpha}^j, \dots, \tilde{\alpha}^n \\ \equiv & \bar{\alpha}^{j-1} \wedge \pi \mod \tilde{\alpha}^j, \dots, \tilde{\alpha}^n \end{aligned} \tag{125}$$

This completes the proof that the Goursat congruences are satisfied. $\square$

Since the one-forms $\bar{\alpha}^i$ do satisfy the Goursat congruences, a coordinate transformation into Goursat normal form can be found. As seen in the proof of Goursat's Theorem, the one-form α^n in the last nonzero derived system has rank 1. We can therefore use Pfaff's Theorem to find functions f_1 and f_2 which satisfy the following equations

$$\begin{aligned} d\alpha^n \wedge \alpha^n \wedge df_1 &= 0 \\ \alpha^n \wedge df_1 \wedge df_2 &= 0 \end{aligned} \qquad \text{and} \qquad \begin{aligned} \alpha^n \wedge df_1 &\neq 0 \\ df_1 \wedge df_2 &\neq 0. \end{aligned} \tag{126}$$

The constraint corresponding to the last axle is once again given by[6]

$$\alpha^n = \sin\theta^n dx^n - \cos\theta^n dy^n \tag{128}$$

and its exterior derivative has the form

$$d\alpha^n = -\cos\theta^n dx^n \wedge d\theta^n - \sin\theta^n dy^n \wedge d\theta^n, \tag{129}$$

It follows that the exterior product of these two quantities is given by

$$d\alpha^n \wedge \alpha^n = -dx^n \wedge dy^n \wedge d\theta^n. \tag{130}$$

By the first equation of (126), f_1 may be chosen to be *any* function of x^n, y^n, θ^n *exclusively.*

Two different solutions of the equations (126) are explained here.

Transformation 1. *Coordinates of the N^{th} trailer. Motivated by Sørdalen* [29]*, f_1 can be chosen to be x^n. The second equation of* (126) *then becomes*

$$\sin\theta^n dx^n \wedge dy^n \wedge df_2 = 0 \tag{131}$$

with the proviso that $df_1 \wedge df_2 \neq 0$. A non-unique *choice of f_2 is*

$$f_2 = y^n. \tag{132}$$

The change of coordinates is defined by:

$$z_1 = f_1(x) = x^n \tag{133}$$

$$z_{n+3} = f_2(x) = y^n. \tag{134}$$

The one form α^n may be written by dividing through by $\sin\theta^n$ *as*

$$\begin{aligned} \alpha^n &= dy^n - \tan\theta^n dx^n &(135)\\ &= dz_{n+3} - z_{n+2}dz_1, &(136)\end{aligned}$$

giving $z_{n+1} = \tan\theta^n$. The remaining coordinates are found by solving the equations

$$\alpha^i \equiv dz_{i+3} - z_{i+2}dz_1 \mod \alpha^{i+1}, \ldots, \alpha^n \tag{137}$$

[6]The basis that satisfies the Goursat congruences was a scaled version of the original basis, $\bar{\alpha}^n = f_{\alpha^n}\alpha^n$. However, it can be checked that

$$\begin{aligned} d\bar{\alpha}^n \wedge \bar{\alpha}^n &= (df_{\alpha^n} \wedge \alpha^n + f_{\alpha^n} d\alpha^n) \wedge f_{\alpha^n}\alpha^n \\ &= (f_{\alpha^n})^2 d\alpha^n \wedge \alpha^n \end{aligned} \tag{127}$$

and thus a function f_1 will satisfy $d\alpha^n \wedge \alpha^n \wedge df_1 = 0$ if and only if $d\bar{\alpha}^n \wedge \bar{\alpha}^n \wedge df_1 = 0$.

for $i = n-1, \ldots, 1$. In fact, because $dz_1 = \pi$ as chosen in the proof of Theorem 44, the one-forms $\bar{\alpha}^i$ already satisfy these equations,

$$\bar{\alpha}^i = \frac{1}{L_n \ldots L_{i+1}} \sec^{n-i+1}\theta^n \sec^{n-i}(\theta^{n-1}-\theta^n)\ldots\sec^3(\theta^{i+2}-\theta^{i+3})\sec^2(\theta^{i+1}-\theta^{i+1}) \Big(\sec\theta^n \sec(\theta^{n-1}-\theta^n)\ldots\sec(\theta^{i+1}-\theta^{i+2})\tan(\theta^i-\theta^{i+1})dx^n - L_{i+1}d\theta^{i+1}\Big) \tag{138}$$

and so the coordinates z_i are given by the coefficients of dx^n in the expression for the $\bar{\alpha}^i$.

Transformation 2. *Coordinates of the origin seen from the last trailer. Yet another choice for f_1 corresponds to writing the coordinates of the origin as seen from the last trailer. This is reminiscent of a transformation used by Samson* [24] *in a different context, and is given by*

$$z_1 := f_1(x) = x^n \cos\theta^n + y^n \sin\theta^n. \tag{139}$$

This has the physical interpretation of being the origin of the reference frame when viewed from a coordinate frame attached to the n^{th} trailer. It satisfies the first of the equations of (126) *simply by virtue of the fact that it is only a function of x^n, y^n, θ^n. It may be verified that a choice of f_2 given by*

$$z_{n+3} := f_2 = x^n \sin\theta^n - y^n \cos\theta^n - \theta^n z_1 \tag{140}$$

satisfies the Pfaff equation,

$$\alpha^1 \wedge df_1 \wedge df_2 = 0. \tag{141}$$

The remaining coordinates $z_2, \ldots, z_{n+2}$ corresponding to this transformation may be obtained by solving the equations

$$\alpha^i \equiv dz_{i+3} - z_{i+2}dz_1 \mod \alpha^{i+1}, \ldots, \alpha^n \tag{142}$$

for $i = n-1, \ldots, 1$. The details are tedious and are omitted.

Remark (Singularities)**.** There are two types of singularities associated with the transformation into Goursat form. At $\theta^n = \pi/2$, for example, the transformation will be singular, but this singularity can be avoided by choosing another coordinate chart at the singular point (such as by interchanging x and y, using the $SE(2)$ symmetry of the system). A singularity also occurs when the angle between two adjacent axles is equal to $\pi/2$; at this point, some of the codistributions in the derived flag will lose rank. The derived flag is not defined at these points; nor is the transformation. There are no singularities of the second type for the unicycle ($n = 0$) or for the front-wheel drive car ($n = 1$).

Once the constraints are in the Goursat normal form, paths can be found which connect any two desired configurations. See Tilbury, Murray, and Sastry [34] for details.

3.3. **The Extended Goursat Normal Form.** While the Goursat normal form is powerful, it is restricted to Pfaffian systems of codimension two. In order to study Pfaffian systems of higher codimension, we present the extended Goursat normal form. Whereas the Goursat normal form can be thought of as a single chain of integrators, the extended Goursat form consists of many chains of integrators. Consider the following definition,

Definition 40 (Extended Goursat Normal Form). *A Pfaffian system I on $\mathbb{R}^{n+m+1}$ of codimension $m+1$ is in* extended Goursat normal form *if it is generated by n constraints of the form:*

$$I = \{dz_i^j - z_{i+1}^j dz^0 : i = 1, \dots, s_j; j = 1, \dots, m\}, \tag{143}$$

This is a direct extension of the Goursat normal form, and all integral curves of (143) are determined by the $m+1$ functions $z^0(t), z_1^1(t), \dots, z_1^m(t)$ and their derivatives with respect to the parameter t. The notation has been changed slightly; the canonical constraints are now $dz_i^j - z_{i+1}^j dz^0$ whereas before they were $dz^i - z^{i-1}dz_1$. For the Goursat form, the constraint in the last nontrivial derived system was $dz^n - z^{n-1}dz^1$; in the extended Goursat normal form, it will be $dz_1^j - z_2^j dz^0$. We refer to the set of constraints with superscript j as the j^{th} tower (the reason for this name will become clear after we compute the derived flag).

Conditions for converting a Pfaffian system to extended Goursat normal form are given by the following theorem:

Theorem 46 (Extended Goursat Normal Form). *Let I be a Pfaffian system of codimension $m+1$. If (and only if) there exists a set of generators $\{\alpha_i^j : i = 1, \dots, s_j; j = 1, \dots, m\}$ for I and an* integrable *one-form π such that for all j,*

$$\begin{aligned} d\alpha_i^j &\equiv -\alpha_{i+1}^j \wedge \pi \qquad && \text{mod } I^{(s_j - i)} \quad i = 1, \dots, s_j - 1 && (144) \\ d\alpha_{s_j}^j &\not\equiv 0 && \text{mod } I && (145) \end{aligned}$$

then there exists a set of coordinates z such that I is in extended Goursat normal form,

$$I = \{dz_i^j - z_{i+1}^j dz^0 : i = 1, \dots, s_j; j = 1, \dots, m\}.$$

Proof. If the Pfaffian system is already in extended Goursat normal form, the congruences are satisfied with $\pi = dz^0$ (which is integrable) and the basis of constraints $\alpha_i^j = dz_i^j - z_{i+1}^j dz^0$.

Now assume that a basis of constraints for I has been found which satisfies the congruences (144). It is easily checked that this basis is adapted to the derived flag, that is:

$$I^{(k)} = \{\alpha_i^j : \quad i = 1, \dots, s_j - k; j = 1, \dots, m\}$$

The coordinates z which comprise the Goursat normal form can now be constructed.

Since π is integrable, any first integral of π can be used for the coordinate z^0. If necessary, the constraints α_i^j can be scaled so that the congruences (144) are satisfied with dz^0:

$$d\alpha_i^j \equiv -\alpha_{i+1}^j \wedge dz^0 \quad \text{mod } I^{(s_j - i)} \quad i = 1, \dots, s_j - 1$$

and the constraints can be renumbered so that $s_1 \geq s_2 \geq \cdots \geq s_m$.

Consider the last nontrivial derived system, $I^{(s_1 - 1)}$. The one-forms $\alpha_1^1, \dots, \alpha_1^{r_1}$ form a basis for this codistribution, where $s_1 = s_2 = \cdots = s_{r_1}$. From the fact that

$$d\alpha_1^j \equiv -\alpha_2^j \wedge dz^0 \quad \text{mod } I^{(s_1 - 1)},$$

it follows that the one-forms $\alpha_1^1, \dots, \alpha_1^{r_1}$ satisfy the Frobenius condition:

$$d\alpha_1^j \wedge \alpha_1^1 \wedge \cdots \wedge \alpha_1^{r_1} \wedge dz^0 = 0$$

and thus, by the Frobenius Theorem, coordinates $z_1^1, \ldots, z_1^{r_1}$ can be found such that

$$\begin{bmatrix} \alpha_1^1 \\ \vdots \\ \alpha_1^{r_1} \end{bmatrix} = A \begin{bmatrix} dz_1^1 \\ \vdots \\ dz_1^{r_1} \end{bmatrix} + B dz^0$$

The matrix A must be nonsingular, since the α_1^j's are a basis for $I^{(s_1-1)}$ and they are independent of dz^0. Therefore, a new basis $\bar{\alpha}_1^j$ can be defined as:

$$\begin{bmatrix} \bar{\alpha}_1^1 \\ \vdots \\ \bar{\alpha}_1^{r_1} \end{bmatrix} := A^{-1} \begin{bmatrix} \alpha_1^1 \\ \vdots \\ \alpha_1^{r_1} \end{bmatrix} = \begin{bmatrix} dz_1^1 \\ \vdots \\ dz_1^{r_1} \end{bmatrix} + (A^{-1}B) dz^0$$

and the coordinates $z_2^j := -(A^{-1}B)_j$ are defined so that the one-forms $\bar{\alpha}_1^j$ have the form

$$\bar{\alpha}_1^j = dz_1^j - z_2^j dz^0$$

for $j = 1, \ldots, r_1$. In these coordinates, the exterior derivative of $\bar{\alpha}_1^j$ is equal to

$$d\bar{\alpha}_1^j = -dz_2^j \wedge dz^0$$

If there were some coordinate z_2^k which could be expressed as a function of the other z_2^j's and z_1^j's, then there would be some linear combination of the $\bar{\alpha}_1^j$'s whose exterior derivative would be zero modulo $I^{(s_1-1)}$, which is a contradiction. Thus, this is a valid choice of coordinates.

By the proof of the standard Goursat Theorem, all of the coordinates in the j^{th} tower can be found from z_1^j and z^0. By the above procedure, all the coordinates in the first r_1 towers can be found.

To find the coordinates for the other towers, the lowest derived systems in which they appear must be considered . The coordinates for the longest towers were found first, next those for the next-longest tower(s) will be found.

Consider the smallest integer k such that dim $I^{(s_1-k)} > k r_1$; more towers will appear at this level. A basis for $I^{(s_1-k)}$ is

$$\{\bar{\alpha}_1^1, \ldots, \bar{\alpha}_k^1, \ldots, \bar{\alpha}_1^{r_1}, \ldots, \bar{\alpha}_k^{r_1}, \alpha_1^{r_1+1}, \ldots, \alpha_1^{r_1+r_2}\}$$

where $\bar{\alpha}_i^j = dz_i^j - z_{i+1}^j dz^0$ for $j = 1, \ldots, r_1$, as found in the first step, and α_1^j for $j = r_1 + 1, \ldots, r_2$ are the one-forms which satisfy the congruences (144) and are adapted to the derived flag. The lengths of these towers are $s_{r_1+1} = \cdots s_{r_1+r_2} = s_1 - k + 1$. For notational convenience, define $z_{(k)}^j := (z_1^j, \ldots, z_k^j)$ for $j = 1, \ldots, r_1$.

By the Goursat congruences, $d\alpha_1^j \equiv -\alpha_2^j \wedge dz^0 \mod I^{(s_1-k)}$ for $j = r_1 + 1, \ldots, r_1 + r_2$, thus the Frobenius condition

$$d\alpha_1^j \wedge \alpha_1^{r_1+1} \wedge \cdots \wedge \alpha_1^{r_1+r_2} \wedge dz_1^1 \wedge \cdots \wedge dz_k^1 \wedge \cdots \wedge dz_1^{r_1} \wedge \cdots \wedge dz_k^{r_1} \wedge dz^0 = 0$$

is satisfied for $j = r_1+1, \ldots, r_1+r_2$. Using the Frobenius Theorem, new coordinates $z_1^{r_1+1}, \ldots, z_1^{r_1+r_2}$ can be found such that

$$\begin{bmatrix} \alpha_1^{r_1+1} \\ \vdots \\ \alpha_1^{r_1+r_2} \end{bmatrix} = A \begin{bmatrix} dz_1^{r_1+1} \\ \vdots \\ dz_1^{r_1+r_2} \end{bmatrix} + B dz^0 + C \begin{bmatrix} dz_{(k)}^1 \\ \vdots \\ dz_{(k)}^{r_1} \end{bmatrix}$$

Since the congruences are only defined up to mod $I^{(s_1-k)}$, the last group of terms (those multiplied by the matrix C) can be eliminated by adding in the appropriate multiples of $\bar{\alpha}_i^j = dz_i^j - z_{i+1}^j dz^0$ for $j = 1, \ldots, r_1$ and $i = 1, \ldots, k$. This will change the B matrix, leaving the equation

$$\begin{bmatrix} \tilde{\alpha}_1^{r_1+1} \\ \vdots \\ \tilde{\alpha}_1^{r_1+r_2} \end{bmatrix} = A \begin{bmatrix} dz_1^{r_1+1} \\ \vdots \\ dz_1^{r_1+r_2} \end{bmatrix} + \tilde{B} dz^0$$

Again, note that A must be nonsingular because the α_1^j's are linearly independent mod $I^{(s_1-k)}$ and also independent of dz^0. Define

$$\begin{bmatrix} \bar{\alpha}_1^{r_1+1} \\ \vdots \\ \bar{\alpha}_1^{r_1+r_2} \end{bmatrix} := A^{-1} \begin{bmatrix} \tilde{\alpha}_1^{r_1+1} \\ \vdots \\ \tilde{\alpha}_1^{r_1+r_2} \end{bmatrix} = \begin{bmatrix} dz_1^{r_1+1} \\ \vdots \\ dz_1^{r_1+r_2} \end{bmatrix} + (A^{-1}\tilde{B}) dz^0$$

and then define the coordinates $z_2^j := -(A^{-1}\tilde{B})_j$ for $j = r_1 + 1, \ldots, r_1 + r_2$ so that $\bar{\alpha}_1^j = dz_1^j - z_2^j dz^0$. Again, by the standard Goursat Theorem, all of the coordinates in the towers $r_1 + 1, \ldots, r_1 + r_2$ are now defined.

The coordinates for the rest of the towers are defined in a manner exactly analogous to that of the second-longest tower. □

If the one-form π which satisfies the congruences (144) is not integrable, then the Frobenius Theorem cannot be used to find the coordinates. In the special case where $s_1 > s_2$, that is, there is one tower which is strictly longer than the others, it can be shown that if there exists *any* π which satisfies the congruences, then there also exists an *integrable* π' which also satisfies the congruences (with a rescaling of the basis forms), see [4, 21]. However, if $s_1 = s_2$, or there are at least two towers which are longest, this is no longer true. Thus, the assumption that π is integrable is necessary for the general case.

If I can be converted to extended Goursat normal form, then the derived flag of I has the structure:

$$\begin{array}{rlllllllll}
I = & \{\alpha_1^1, & \cdots, & \cdots, & \alpha_{s_1-1}^1, & \alpha_{s_1}^1, & \cdots, & \alpha_1^m, & \cdots, & \alpha_{s_m-1}^m, \quad \alpha_{s_m}^m\} \\
I^{(1)} = & \{\alpha_1^1, & \cdots, & \cdots, & \alpha_{s_1-1}^1, & & \cdots, & \alpha_1^m, & \cdots, & \alpha_{s_m-1}^m\} \\
& \vdots & & & \iddots & & & \vdots & \iddots & \\
I^{(s_m-1)} = & \{\alpha_1^1, & \cdots, & \alpha_{s_1-s_m+1}^1, & & & \cdots, & \alpha_1^m\} & & \\
& \vdots & & \iddots & & & & & & \\
I^{(s_1-2)} = & \{\alpha_1^1, & \alpha_2^1\} & & & & & & & \\
I^{(s_1-1)} = & \{\alpha_1^1\} & & & & & & & & \\
I^{(s_1)} = & \{0\} & & & & & & & &
\end{array}$$

where the forms in each level have been arranged to show the different towers. The superscripts j indicate the tower to which each form belongs, and the subscripts i index the position of the form within the j^{th} tower. There are s_j forms in the j^{th} tower.

Another version of the extended Goursat normal form Theorem is given here, which is easier to check, since it does not require finding a basis which satisfies the congruences but only one which is adapted to the derived flag. One special case of this theorem is proven in [25].

Theorem 47 (Extended Goursat Normal Form)**.** *A Pfaffian system I of codimension $m+1$ on $\mathbb{R}^{n+m+1}$ can be converted to extended Goursat normal form if and only if $I^{(N)} = \{0\}$ for some N and there exists a one-form π such that $\{I^{(k)}, \pi\}$ is integrable for $k = 0, \ldots, N-1$.*

Proof. The only if part is easily shown by taking $\pi = dz^0$ and noting that

$$\begin{aligned} I^{(k)} &= \{dz_i^j - z_{i+1}^j dz^0 : \qquad i = 1, \ldots, s_j - k; j = 1, \ldots, m\} \\ \{I^{(k)}, \pi\} &\equiv \{dz_i^j, dz^0 : \qquad i = k+1, \ldots, s_j; j = 1, \ldots, m\} \end{aligned}$$

which is integrable for every k.

Now assume that such a π exists. After the derived flag of the system, $I =: I^{(0)} \supset I^{(1)} \supset \cdots \supset I^{(s_1)} = \{0\}$, has been found, a basis which is adapted to the derived flag and which satisfies the Goursat congruences (144) can be iteratively constructed.

The lengths of each tower are determined from the dimensions of the derived flag. Indeed, the longest tower of forms has length s_1. If the dimension of $I^{(s_1-1)}$ is r_1, then there are r_1 towers which each have length s_1; and we have $s_1 = s_2 = \cdots = s_{r_1}$. Now, if the dimension of $I^{(s_1-2)}$ is $2r_1 + r_2$, then there are r_2 towers with length $s_1 - 1$, and $s_{r_1+1} = \cdots = s_{r_1+r_2} = s_1 - 1$. Each s_j is found similarly.

A π which satisfies the conditions must be in the complement of I, for if π were in I, then $\{I, \pi\}$ integrable means that I is integrable, and this contradicts the assumption that $I^{(N)} = \{0\}$ for some N.

Consider the last nontrivial derived system, $I^{(s_1-1)}$. Let $\{\alpha_1^1, \ldots, \alpha_1^{r_1}\}$ be a basis for $I^{(s_1-1)}$. The definition of the derived flag, specifically $I^{(s_1)} = \{0\}$, implies that

$$d\alpha_1^j \not\equiv 0 \mod I^{(s_1-1)} \quad j = 1, \ldots, r_1 \tag{146}$$

Also, the assumption that $\{I^{(k)}, \pi\}$ is integrable gives the congruence

$$d\alpha_1^j \equiv 0 \mod \{I^{(s_1-1)}, \pi\} \quad j = 1, \ldots, r_1 \tag{147}$$

combining equations (146) and (147), the congruence

$$d\alpha_1^j \equiv \pi \wedge \beta^j \mod I^{(s_1-1)} \quad j = 1, \ldots, r_1 \tag{148}$$

must be satisfied for some $\beta^j \not\equiv 0 \mod I^{(s_1-1)}$.

Now, from the definition of the derived flag,

$$d\alpha_1^j \equiv 0 \mod I^{(s_1-2)} \quad j = 1, \ldots, r_1$$

which combined with (148) implies that β^j is in $I^{(s_1-2)}$.

Claim. $\beta^1, \ldots, \beta^{r_1}$ are linearly independent mod $I^{(s_1-1)}$.

Proof of Claim. The proof is by contradiction. Suppose there exists some combination of the β^j's, say

$$\beta = b_1\beta^1 + \cdots + b_{r_1}\beta^{r_1} \equiv 0 \mod I^{(s_1-1)}$$

with not all of the b_j's equal to zero. Consider $\alpha = b_1\alpha_1^1 + \cdots + b_{r_1}\alpha_1^{r_1}$. This one-form $\alpha \neq 0$ because the α_1^j are a basis for $I^{(s_1-1)}$. The exterior derivative of α

can be found by the product rule,

$$\begin{aligned} d\alpha &= \sum_{j=1}^{r_1} b_j d\alpha_1^j + \sum_{j=1}^{r_1} db_j \wedge \alpha_1^j \\ &\equiv \sum_{j=1}^{r_1} b_j(\pi \wedge \beta^j) \quad \bmod I^{(s_1-1)} \\ &\equiv \pi \wedge (\sum_{j=1}^{r_1} b_j \beta^j) \quad \bmod I^{(s_1-1)} \\ &\equiv 0 \quad \bmod I^{(s_1-1)} \end{aligned}$$

which implies that α is in $I^{(s_1)}$. However, this contradicts the assumption that $I^{(s_1)} = \{0\}$. Thus the β^j's must be linearly independent mod $I^{(s_1-1)}$.

Define $\alpha_2^j := \beta^j$ for $j = 1, \ldots, r_1$. Note that these basis elements satisfy the first level of Goursat congruences, that is:

$$d\alpha_1^j \equiv -\alpha_2^j \wedge \pi \quad \bmod I^{(s_1-1)} \quad j = 1, \ldots, r_1$$

If the dimension of $I^{(s_1-2)}$ is greater than $2r_1$, then one-forms $\alpha_1^{r_1+1}, \ldots, \alpha_1^{r_1+r_2}$ are chosen such that

$$\{\alpha_1^1, \ldots, \alpha_1^{r_1}, \alpha_2^1, \ldots, \alpha_2^{r_1}, \alpha_1^{r_1+1}, \ldots, \alpha_1^{r_1+r_2}\}$$

is a basis for $I^{(s_1-2)}$.

For the induction step, assume that a basis for $I^{(i)}$ has been found,

$$\{\alpha_1^1, \ldots, \alpha_{k_1}^1, \alpha_1^2, \ldots, \alpha_{k_2}^2, \ldots, \alpha_1^c, \ldots, \alpha_{k_c}^c\}$$

which satisfies the Goursat congruences up to this level:

$$d\alpha_k^j = -\alpha_{k+1}^j \wedge \pi \quad \bmod I^{(s_j-k)} \quad k = 1, \ldots, k_j - 1; \quad j = 1, \ldots, c$$

Note c towers of forms have appeared in $I^{(i)}$. Consider only the last form in each tower that appears in $I^{(i)}$, that is $\alpha_{k_j}^j, j = 1, \ldots, c$. By the construction of this basis (or from the Goursat congruences), $\alpha_{k_j}^j$ is in $I^{(i)}$ but is not in $I^{(i+1)}$, thus

$$d\alpha_{k_j}^j \not\equiv 0 \quad \bmod I^{(i)} \quad j = 1, \ldots, c$$

The assumption that $\{I^{(i)}, \pi\}$ is integrable assures

$$d\alpha_{k_j}^j \equiv 0 \quad \bmod \{I^{(i)}, \pi\} \quad j = 1, \ldots, c$$

thus $d\alpha_{k_j}^j$ must be a multiple of π mod $I^{(i)}$,

$$d\alpha_{k_j}^j \equiv \pi \wedge \beta^j \quad \bmod I^{(i)} \quad j = 1, \ldots, c$$

for some $\beta^j \not\equiv 0 \mod I^{(i)}$. From the fact that $\alpha_{k_j}^k$ is in $I^{(i)}$ and the definition of the derived flag,

$$d\alpha_{k_j}^j \equiv 0 \quad \bmod I^{(i-1)} \quad j = 1, \ldots, c$$

which implies that $\beta^j \in I^{(i-1)}$. By a similar argument to the claim above, it can be shown that the β^j's are independent mod $I^{(i)}$. Define $\alpha_{k_j+1}^j := \beta^j$, and thus

$$\{\alpha_1^1, \ldots, \alpha_{k_1+1}^1, \alpha_1^2, \ldots, \alpha_{k_2+1}^2, \ldots, \alpha_1^c, \ldots, \alpha_{k_c+1}^c\}$$

forms part of a basis of $I^{(i-1)}$. If the dimension of $I^{(i-1)}$ is greater than $k_1 + k_2 + \cdots + k_c + c$, then complete the basis of $I^{(i-1)}$ with any linearly independent one-forms $\alpha_1^{c+1}, \ldots, \alpha_1^{c+r_c}$ such that

$$\{\alpha_1^1, \ldots, \alpha_{k_1+1}^1, \alpha_1^2, \ldots, \alpha_{k_2+1}^2, \ldots, \alpha_1^c, \ldots, \alpha_{k_c+1}^c, \alpha_1^{c+1}, \ldots, \alpha_1^{c+r_c}\}$$

is a basis for $I^{(i-1)}$.

Repeated application of this procedure will construct a basis for I which is not only adapted to the derived flag, but also satisfies the Goursat congruences.

By assumption, π is integrable mod the last nontrivial derived system, $I^{(s_1-1)}$. Looking at the congruences (144), any integrable one-form π' which is congruent to π up to a scaling factor f,

$$\pi' = dz^0 \equiv f\pi \mod I^{(s_1-1)}$$

will satisfy the same set of congruences up to a rescaling of the constraint basis by multiples of this factor f. □

3.4. **Prolongations of Pfaffian Systems.** If a Pfaffian system I of codimension k satisfies the necessary and sufficient conditions for conversion into extended Goursat form, then its integral curves are determined by k arbitrary functions. However, even if a system cannot be transformed into Goursat form, its integral curves may still have this property. If so, then I is said to be absolutely equivalent in the sense of Cartan to the trivial system (the system with no constraints) on $\mathbb{R}^k$.

Definition 41. *Two Pfaffian systems I and J are said to be* absolutely equivalent *(in the sense of Cartan) if there is a one-to-one correspondence between their solution trajectories.*

Although the concept of absolute equivalence will not be examined here in its full generality, some sufficient conditions will be given for a Pfaffian system to have a prolongation which can be converted to Goursat form, and thus the integral curves of I are determined by k independent functions. Consider a Pfaffian system I in extended Goursat normal form:

$$I = \{dz_i^j - z_{i+1}^j dz^0 : i = 1, \ldots, s_j; j = 1, \ldots, m\},$$

with independence condition dz^0. Let the Pfaffian system J be defined by:

$$J = \{dz_i^j - z_{i+1}^j dz^0 : i = 1, \ldots, s_1 + 1 \text{ and } i = 1, \ldots, s_j; j = 2, \ldots, m\},$$

The coordinate $z_{s_1+2}^1$ has been added, but the new system is also in extended Goursat normal form. It is clear that there is a one-to-one correspondence of integral curves between I and J although they are defined on manifolds of different dimensions. J is said to be a *prolongation by differentiation* (of order one) of I with respect to the independence condition dz^0.

Prolongations by differentiation can also be defined for systems which are not *a priori* in extended Goursat normal form. Let I be a Pfaffian system on a manifold M with independence condition dt, and let $d\eta$ be a one-form in the complement of I. Define the system J on $M \times \mathbb{R}$ given by

$$J = \{I, d\eta - y dt\}$$

to be a prolongation by differentiation of I, where the new coordinate y is the fiber coordinate on $\mathbb{R}$. In effect, this adds the derivative of η (with respect to the independence condition) as a state variable. As long as all solution trajectories are "smooth enough" (C^∞), there will be a one-to-one correspondence between solution trajectories of the original and the prolonged system.

In general, many of these partial prolongations by differentiation may be taken.

Definition 42 (Prolongation by differentiation)**.** *Let I be a Pfaffian system of codimension $m+1$ on $\mathbb{R}^{n+m+1}$ with coordinates (z, v, t) for which dt is an independence condition and $\{dv_1, \ldots, dv_m, dt\}$ forms a complement. Let $b_1, \ldots, b_m$ be nonnegative integers and let b denote their sum. The system I augmented by the b one-forms*

$$\begin{array}{llll} dv_1 - v_1^1 dt, & \ldots, & dv_1^{b_1-1} - v_1^{b_1} dt, & \\ dv_2 - v_2^1 dt, & \ldots, & & dv_2^{b_2-1} - v_2^{b_2} dt \\ & \vdots & & \ddots \\ dv_m - v_m^1 dt, & \ldots, & \ldots, & dv_m^{b_m-1} - v_m^{b_m} dt, \end{array}$$

is a prolongation by differentiation *of I. The augmented system is defined on $\mathbb{R}^{n+m+b+1}$.*

If a Pfaffian system I does not satisfy the necessary and sufficient conditions of Theorems 46 and 47, then I cannot be converted into extended Goursat normal form. It is possible, however, that there exists a prolongation by differentiation J of I which does satisfy the extended Goursat conditions. In this case, the prolonged system J can be put into Goursat normal form, paths can be found for the transformed system using one of the methods described in [36, 32], and these paths can be projected down onto the original Pfaffian system I to give integral curves.

Although the general problem of determining which Pfaffian systems can be converted into extended Goursat normal form after prolongation is still an open one, the following theorem gives some sufficient conditions under which such a transformation exists.

Theorem 48 (Conversion to Goursat form with prolongation by differentiation)**.** *Consider a Pfaffian system $I = \{\alpha^1, \ldots, \alpha^n\}$ on $\mathbb{R}^{n+m+1}$ with independence condition dz^0 and complement $\{dv_1, \ldots, dv_m, dz^0\}$. If there exists a list of integers $b_1, \ldots, b_m$ such that the prolonged system*

$$\begin{aligned} J \; = \{ \quad & \alpha^1, \ldots, \alpha^n, \\ & dv_1 - v_1^1 dz^0, \ldots, dv_1^{b_1-1} - v_1^{b_1} dz^0, \\ & \quad \vdots \\ & dv_m - v_m^1 dz^0, \ldots, dv_m^{b_m-1} - v_m^{b_m} dz^0 \ \} \end{aligned}$$

satisfies the condition that $\{J^{(k)}, dz^0\}$ is integrable for all k, then I can be transformed to extended Goursat normal form using a prolongation by differentiation.

Proof. The proof is by application of Theorem 47 to the prolonged system J. □

3.5. **The multi-steering trailer problem.** Previously, we discussed the system of a car-like mobile robot towing n trailers. A similar system consisting of a chain of wheeled trailers, several of which are steerable, will be considered as the main example in this section.

First, consider a system of n (passive) trailers and m (steerable) cars linked together by rigid bars, as sketched in Figure 2. It is assumed that each body (trailer or car) has only one axle, since a two-axle car is equivalent (under coordinate transformation and state feedback) to a one-axle car towing one trailer.

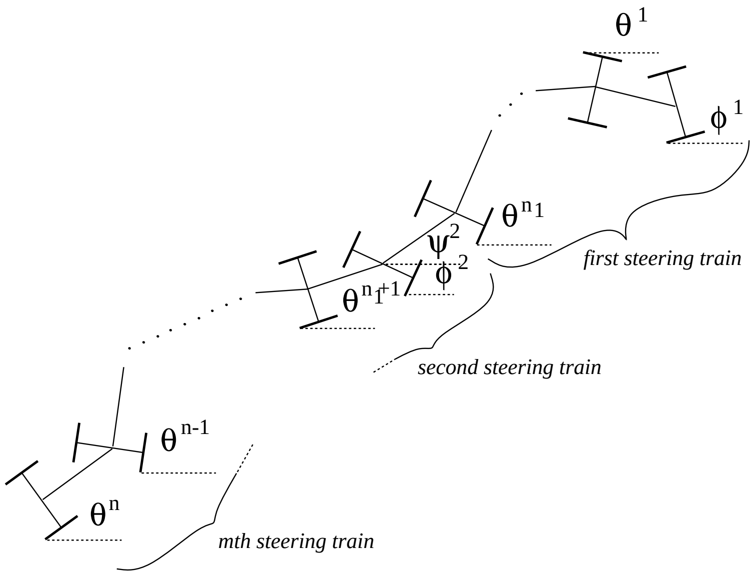

FIGURE 2. A multi-trailer system with n (passive) trailers and m (active) steering wheels.

3.5.1. *Configuration space.* The active or steering axles are numbered from front to back, starting with 1 and going up to m, and the passive axles are numbered similarly from 1 to n. There are a total of $n + m$ axles in the system. The angle of each passive axle with respect to the horizontal will be represented by θ^i where $i \in \{1, \ldots, n\}$ is the axle number. Each steerable axle together with the passive axles directly behind it will be called a *steering train.*

The steerable axles may be interspersed among the passive axles in any fashion. The indices of the passive axles which are directly in front of the steerable axles will be denoted by $n_1, \ldots, n_{m-1}$. The first axle is always assumed to be steerable, and thus $n_0 = 0$. The angle of the first axle with respect to the horizontal is denoted by ϕ^1. If there are n_1 passive trailers in the first steering train, their angles are denoted $\theta^1, \ldots, \theta^{n_1}$. The axle directly behind the first steering train is steerable, and its angle with respect to the horizontal will be ϕ^2. The (passive) axles behind the second steering wheel are thus $\theta^{n_1+1}, \ldots, \theta^{n_2}$; the angle of the third steering wheel will be ϕ^3, and so forth. For convenience of notation, let $n_m = n$, although in general the last axle will not be steerable. If the last axle is steerable, then $n_{m-1} = n_m$.

Let ψ^j denote the absolute angle (with respect to the horizontal) of the bar connecting the j^{th} steered axle to the last axle of the $(j-1)^{st}$ steering train (which may be either steered or passive). This can be considered to be the angle of the bar connecting the j^{th} steering train to the $(j-1)^{st}$ steering train. The Cartesian position (x, y) of *any one* of the axles, along with all of the angles described above, will determine the state of the system. The choice of which (x, y) will be deferred for the time being, but it is noted that only one pair is needed.

The configuration of a trailer system consisting of n trailers and m steerable cars is thus completely given by

$$\xi = [\theta^1, \ldots, \theta^n, \phi^1, \ldots, \phi^m, \psi^2, \ldots, \psi^m, x, y]^T \in (S^1)^{n+2m-1} \times \mathbb{R}^2 \ .$$

3.5.2. *Pfaffian system.* The nonholonomic constraints on the velocities, representing the fact that each axle of wheels rolls without slipping, form a codistribution of one-forms in the cotangent bundle to the configuration manifold and thus generate a Pfaffian system.

If the variables (x^i, y^i) are used to represent the Cartesian position of the i^{th} passive axle, then the constraint that the i^{th} passive axle roll without slipping can be written in these coordinates as:

$$\omega^i = \sin\theta^i dx^i - \cos\theta^i dy^i \tag{149}$$

Similarly, let (x_s^j, y_s^j) represent the Cartesian position of the j^{th} steerable axle (where the subscript s stands for steerable). The constraint that the j^{th} steerable axle roll without slipping may be written as:

$$\alpha^j = \sin\phi^j dx_s^j - \cos\phi^j dy_s^j \tag{150}$$

Of course, as noted before, only one pair of (x, y), along with all of the angles, is needed to specify the state of the system.

The Pfaffian system generated by this mobile robot system is the collection of all the nonholonomic (rolling without slipping) constraints:

$$I = \{\omega^1, \dots, \omega^n, \alpha^1, \dots, \alpha^m\}$$

Thus I has dimension $n+m$ in a space of dimension $n+2m+1$; the codimension of I is $m+1$, or one more than the number of steering angles.

Notice that from equations (149) and (150) it can be seen that:

$$dy^i \equiv \tan\theta^i dx^i \mod \omega^i \tag{151}$$

$$dy_s^j \equiv \tan\phi^j dx_s^j \mod \alpha^j \tag{152}$$

All of the (x^i, y^i)'s and (x_s^i, y_s^i)'s are related by the hitch relationships. The exterior derivatives of these relationships can be taken, yielding

$$\begin{array}{l} x^{i-1} = x^i + L_i\cos\theta^i \\ y^{i-1} = y^i + L_i\sin\theta^i \end{array} \Longrightarrow \begin{array}{l} dx^{i-1} = dx^i - L_i\sin\theta^i d\theta^i \\ dy^{i-1} = dy^i + L_i\cos\theta^i d\theta^i \end{array}$$

and substituting these quantities into the expression for ω^{i-1} from (149), the constraint for the $(i-1)^{st}$ passive axle can be rewritten as:

$$\begin{aligned} \omega^{i-1} &= \sin\theta^{i-1}dx^{i-1} - \cos\theta^{i-1}dy^{i-1} && (153)\\ &= \sin\theta^{i-1}dx^i - \cos\theta^{i-1}dy^i - L_i\cos(\theta^i - \theta^{i-1})d\theta^i && (154)\\ &\equiv (\sin\theta^{i-1} - \tan\theta^i\cos\theta^{i-1})dx^i - L_i\cos(\theta^i - \theta^{i-1})d\theta^i && \mod \omega^i\\ &\equiv \sec\theta^i\sin(\theta^{i-1} - \theta^i)dx^i - L_i\cos(\theta^i - \theta^{i-1})d\theta^i && \mod \omega^i \end{aligned}$$

where the congruence (151) has been used. Once again, a rearrangement of terms and a division by cosine in (153) will give the congruence

$$\begin{aligned} d\theta^i &\equiv \frac{1}{L_i}\sec\theta^i\tan(\theta^{i-1} - \theta^i)dx^i && \mod \omega^i, \omega^{i-1} && (155)\\ d\theta^i &\equiv f_{\theta^i}dx^i && \mod \omega^i, \omega^{i-1} && (156) \end{aligned}$$

The exact form of the function f_{θ^i} is unimportant; what will be needed is the relationship between $d\theta^i$ and dx^i.

The first lemma can now be proved,

Lemma 49. *The exterior derivatives of any of the x variables are congruent modulo the Pfaffian system, that is:* $dx^i \equiv f_{x^{i,j}} dx^j \equiv f_{x_s^{i,k}} dx_s^k \mod I$

Proof. For two passive axles, the relationship between the x coordinates is given by the hitching relationship,

$$\begin{aligned} x^{i-1} &= x^i + L_i \cos\theta^i && (157)\\ dx^{i-1} &= dx^i - L_i \sin\theta^i d\theta^i && (158)\\ &\equiv (1 - L_i \sin\theta^i f_{\theta^i})dx^i \qquad \mod \omega^{i-1}, \omega^i && (159)\\ &\equiv f_{x^{i-1}} dx^i \qquad \mod \omega^{i-1}, \omega^i && (160) \end{aligned}$$

where the congruence (155) was used.

The computations are similar when there is a steerable axle involved instead of two passive axles. If the i^{th} passive axle is located in front of the j^{th} steerable axle, then the hitch relationship and its exterior derivative are given by:

$$x^i = x_s^j + l_j \cos\psi^j \tag{161}$$

$$dx^i = dx_s^j - l_j \sin\psi^j d\psi^j \tag{162}$$

In this case, the constraint corresponding to the i^{th} passive axle has the form

$$\begin{aligned} \omega^i &= \sin\theta^i dx^i - \cos\theta^i dy^i && (163)\\ &= \sin\theta^i dx_s^j - \cos\theta^i dy_s^j - l_j \cos(\theta^i - \psi^j) d\psi^j && (164)\\ &\equiv (\sin\theta^i - \cos\theta^i \tan\phi^j) dx_s^j - l_j \cos(\theta^i - \psi^j) d\psi^j \qquad \mod \alpha^j && (165)\\ &\equiv \sec\phi^j \sin(\theta^i - \phi^j) dx_s^j - l_j \cos(\theta^i - \psi^j) d\psi^j \qquad \mod \alpha^j && (166) \end{aligned}$$

Again, the standard trick of dividing through by a cosine and rearranging terms will result in the congruence

$$\begin{aligned} d\psi^j &\equiv \frac{1}{l_j} \sec\phi^j \sin(\theta^i - \phi^j) \sec(\theta^i - \psi^j) dx_s^j \qquad \mod \alpha^j, \omega^i && (167)\\ d\psi^j &\equiv f_{\psi^j} dx_s^j \qquad \mod \alpha^j, \omega^i && (168) \end{aligned}$$

Now, combining (167) with (161), it can be seen that

$$dx^i \equiv f_{x_s^j} dx_s^j \mod \alpha^j, \omega^i$$

The case where there are two adjacent steerable axles is done exactly the same way, with different notation, and will not be written out in detail here. □

A complement to the Pfaffian system $I = \{\omega^1, \ldots, \omega^n, \alpha^1, \ldots, \alpha^m\}$ is given by

$$\{d\phi^1, \ldots, d\phi^m, dx\}$$

for any $x \in \{x^1, \ldots, x^n, x_s^1, \ldots, x_s^m\}$, since by Lemma 49 their exterior derivatives are congruent modulo the system, and the complement is only defined modulo the system. Since the derivatives $d\phi^j$ do not appear in any of the constraints, they are in the complement to I.

From the exterior derivative of the constraint corresponding to the i^{th} passive axle, it can be seen that

$$\begin{aligned}
\omega^i &= \sin\theta^i dx^i - \cos\theta^i dy^i && && (169)\\
d\omega^i &= d\theta^i \wedge (\cos\theta^i dx^i + \sin\theta^i dy^i) && && (170)\\
&\equiv d\theta^i \wedge (\cos\theta^i + \sin\theta^i \tan\theta^i) dx^i && \mod \omega^i && (171)\\
&\equiv \sec\theta^i d\theta^i \wedge dx^i && \mod \omega^i && (172)\\
&\equiv 0 && \mod \omega^i, \omega^{i-1} && (173)
\end{aligned}$$

where the congruences (151) and (155) have been used. That is, the exterior derivative of the constraint corresponding to the i^{th} passive axle is equal to zero modulo itself and the constraint which corresponds to the axle most directly in front. Without redoing the calculations, which are identical except for the notation, it can be seen that if the i^{th} passive axle is behind a steerable axle with angle ϕ^k instead of a passive axle with angle θ^{i-1}, that is, $i = n_{k-1} + 1$, then the following congruence will result:

$$d\omega^i \equiv 0 \mod \omega^i, \alpha^k \tag{174}$$

Proceeding similarly, the exterior derivatives of the constraints associated with the steerable axles can be found,

$$\begin{aligned}
\alpha^j &= \sin\phi^j dx_s^j - \cos\phi^j dy_s^j && && (175)\\
d\alpha^j &= d\phi^j \wedge (\cos\phi^j dx_s^j + \sin\phi^j dy_s^j) && && (176)\\
&\equiv d\phi^j \wedge (\cos\phi^j + \sin\phi^j \tan\phi^j) dx_s^j && \mod \alpha^j && (177)\\
&\equiv \sec\phi^j d\phi^j \wedge dx_s^j && \mod \alpha^j && (178)\\
&\not\equiv 0 && \mod I && (179)
\end{aligned}$$

and it can be seen that their exterior derivatives are nonzero modulo the Pfaffian system I.

Recalling the definition of the derived flag from Chapter 2, it is now easy to see that all of the constraints corresponding to the passive axles are in the first derived system, and none of those corresponding to the steerable axles are. That is, the first derived system is given by:

$$I^{(1)} = \{\omega^1, \ldots, \omega^n\}$$

In fact, the entire derived flag can be found just from the three congruences, (169), (174), and (175),

Lemma 50 (Derived Flag). *The derived flag associated with the m-steering, n-trailer system has the form:*

$$I^{(k)} = \{\omega^i : n_{j-1} + k \le i \le n_j, j = 1, \ldots, m\}$$

for $k = 1, \ldots, n$. In addition,

$$I^{(n+1)} = \{0\}.$$

Proof. The proof is just a one-time application of (175), to show that none of the constraints α^j corresponding to the steering axles are in $I^{(1)}$, and then a repeated application of (169) to show at which level each constraint falls out of the derived flag. □

If n_1 is the greatest of the indices n_i, the derived flag has the structure:

$$\begin{array}{rlllll}
I & = \{ & \alpha^1, \omega^1, \omega^2, \dots, \omega^{n_1}, & \alpha^2, \omega^{n_1+1}, \omega^{n_1+2}, \dots, & \alpha^m, \omega^{n_{m-1}+1}, \dots, \omega^n \} \\
I^{(1)} & = \{ & \omega^1, \omega^2, \dots, \omega^{n_1}, & \omega^{n_1+1}, \omega^{n_1+2}, \dots, & \omega^{n_{m-1}+1}, \dots, \omega^n \} \\
I^{(2)} & = \{ & \omega^2, \dots, \omega^{n_1}, & \omega^{n_1+2}, \dots, & \dots, \omega^n \} \\
\vdots & & \vdots & & \\
I^{n_1} & = \{ & \omega^{n_1} \} & & \\
I^{n_1+1} & = \{ & 0 \} & &
\end{array}$$

In the general case, the Pfaffian system I consists of the constraints corresponding to all the axles, the first derived system lacks the steerable axles, the second derived system lacks those passive axles that are directly behind steerable axles, and at every subsequent level, the constraint which is most toward the front of each steering train will drop off. Since the longest possible chain of contiguous passive axles is equal to n, the total number of passive axles that are in the chain, the $(n+1)^{st}$ derived system must be equal to $\{0\}$.

3.5.3. *Conversion to extended Goursat normal form.* In this section, it will be shown how the general multi-steering trailer system can be converted into extended Goursat normal form after prolongation. The configurations of this system which satisfy the conditions for conversion without prolongation will also be detailed. The first lemma gives a candidate choice for π, since for a system in extended Goursat normal form it must always be true that $\{I, \pi\}$ is integrable.

Lemma 51. *$\{I, dx\}$ is integrable for any $x \in \{x^1, \dots, x^n, x_s^1, \dots, x_s^m\}$.*

Proof. Each constraint in I satisfies the congruence

$$\begin{aligned}
d\omega^i &\equiv d\theta^i \wedge dx^i \mod \omega^i \\
&\equiv 0 \mod \{I, dx^i\} \\
d\alpha^j &\equiv d\phi^j \wedge dx_s^j \mod \alpha^j \\
&\equiv 0 \mod \{I, dx_s^j\}
\end{aligned}$$

(by equations (169) and (175)). Also, all of the dx^i, dx_s^j are congruent by Lemma 49. Thus, the exterior derivative of any constraint in $\{I, dx\}$ is congruent to zero mod $\{I, dx_s^j\}$, which is the condition for integrability. □

It can be shown that for the general case, there does not exist a dx (or any other one-form) which will satisfy the condition that $\{I^{(i)}, dx\}$ is integrable for every i. However, the general multi-steering system can be transformed into Goursat normal form after prolongation.

The concept of "virtual trailers" was first introduced in [36] as a type of dynamic state feedback for the multi-steering trailer system. A chain of these virtual trailers, each analogous to a physical trailer, was added in front of each actual steering wheel, and a virtual steering wheel was added at the front of each virtual chain. The sketch of this augmented system in Figure 3 helps make the concept more clear. Each virtual trailer adds one state to the system, as well as one constraint. Thus the codimension of the extended system is the same as that of the original system, $m+1$.

Theorem 52 (Converting the multi-steering system to Goursat form)**.** *The multi-steering system with n trailers and m steering wheels can be put into extended*

Goursat normal form, for any n, m and for any configuration of steerable cars and passive trailers, using a prolongation of degree less than or equal to $n_1 + \cdots + n_{m-1}$.

Proof. Consider the n-trailer, m-steering system with virtual extension as shown in Figure 3. That is, in front of each steerable axle, imagine that there are n_{j-1} virtual axles, and that only the front axle in each virtual chain is steerable. Note that with this virtual axle formulation, the actual steerable axles within the multi-trailer chain are no longer assumed to be directly steerable, but rather are controlled through the virtual steering axles and the the chains of virtual trailers.

Let ϕ_v^j represent the angle of the j^{th} virtual steering axle, where the subscript v stands for virtual. The angles of the passive axles that are added are denoted by θ_j^i, where the subscript j stands for the index of the virtual chain that they are in, and the superscript i indexes their position from the front of the virtual train.

A total of $n_1 + \cdots + n_{m-1}$ states have been added to the system, corresponding to the angles of the virtual axles. The same number of constraints have also been added. The first axle is always assumed to be steerable, and no virtual axles are added in front of the front steering wheel.

Because the constraints that were added have the same form as those in the system already, it is easy to see that they can be written in coordinates as

$$\nu_j^i = \sin\theta_j^i dx_j^i - \cos\theta_j^i dy_j^i$$

for the passive virtual axles and

$$\alpha_v^j = \sin\phi_v^j dx_v^j - \cos\theta_v^j dy_v^j$$

for the steerable virtual axles at the front of each chain. Although these constraints do not immediately appear to be of the same form as a prolongation by differentiation, it can be shown that

$$\nu_j^i \equiv d\theta_j^{i+1} - \tan\theta_j^i dx^n \mod J^{(i)}$$

for $i < n_{j-1} - 2$ and

$$\nu_j^{n_{j-1}-1} \equiv d\phi^j - \tan\theta_j^{n_{j-1}-1} dx^n \mod J^{(n_{j-1}-1)}$$

where x^n is the x-position of the last passive axle. This particular form of a prolongation by differentiation was chosen so that the constraints which were added to the system would have the same expression (in coordinates) as the physical constraints; the computations are somewhat simplified by this choice. Because of the equivalence, a standard prolongation by differentiation could have been used; it would be difficult to interpret the meaning of the added states.

The prolonged Pfaffian system is given by the collection of actual and virtual constraints,

$$J = \{\alpha^1, \dots, \alpha^m, \alpha_v^2, \dots, \alpha_v^m, \omega^1, \dots, \omega^n, \nu_j^i : j = 2, \dots, m; i = 1, \dots, n_{j-1} - 1\}$$

The derived flag corresponding to the extended Pfaffian system can now be found. First, performing a similar calculation to that in equation (169), it can be seen that

$$d\nu_j^i \equiv 0 \mod \nu_j^i, \nu_j^{i-1}$$

Then, similar to equation (175),

$$d\alpha_v^j \not\equiv 0 \mod J$$

It is also not difficult to show that

$$d\alpha^j \equiv 0 \mod \alpha^j, \nu_j^{n_{j-1}-1}$$

From these three congruences, the structure of the derived flag is seen to be

$$\begin{aligned}
J^{(1)} &= \{\alpha^2, \ldots, \alpha^m, \omega^1, \ldots, \omega^n, \nu_j^i : j = 2, \ldots, m; i = 1, \ldots, n_{j-1} - 1\} \\
J^{(2)} &= \{\alpha^2, \ldots, \alpha^m, \omega^2, \ldots, \omega^n, \nu_j^i : j = 2, \ldots, m; i = 2, \ldots, n_{j-1} - 1\} \\
&\vdots \\
J^{(k+j_k-1)} &= \{\alpha^{j_k+1}, \ldots, \alpha^m, \omega^k, \ldots, \omega^n, \nu_j^i : j = 2, \ldots, m; i = k, \ldots, n_{j-1} - 1\} \\
&\vdots \\
J^{(n+m-1)} &= \{\omega^n\} \text{ or } \{\alpha^m\} \\
J^{(n+m)} &= \{0\}
\end{aligned}$$

where j_k is defined to be the number of steerable axles that are in front of the k^{th} passive axle in the actual chain of trailers, and $J^{(n+m-1)} = \{\omega^n\}$ if the last axle in the chain is passive, and $J^{(n+m-1)} = \{\alpha^m\}$ if the last axle in the chain is steerable. In words, the (extended) Pfaffian system J consists of all the constraints corresponding to both the actual and the virtual axles. The first derived system consists of all constraints except the ones at the front of each (virtual) chain. At the second level, the constraints corresponding to the axles directly behind each virtual steering wheel fall off, and at the k^{th} level, the constraints corresponding to the axles which are k behind each virtual steering wheel fall off, until at the last level, there is only the constraint corresponding to the last axle in the chain (ω^n if it is passive, α^m if it is steerable). The $(n+m)^{th}$ derived system is trivial, which implies that the augmented system is controllable.

At each level of the derived flag, exactly one of the constraints which falls out of the flag corresponds to a real axle, and all the rest which fall out correspond to virtual axles.

The one-form π which satisfies the Goursat conditions of Theorem 47 is equal to the exterior derivative of the x coordinate of the last body in the actual multi-steering chain; dx^n (if the last axle in the chain is passive) or dx_s^m (if the last axle in the chain is steerable). The rest of the details are straightforward, although the notation is cumbersome. □

Now that it has been shown that the system with virtual trailers can always be converted into extended Goursat normal form, some special cases of the multi-steering trailer system which can be converted into extended Goursat normal form without any prolongation will be examined.

Theorem 53. *If there is only one steering train which has passive axles in it, that is, all the passive axles are contiguous, then the system can be converted into extended Goursat normal form without prolongation.*

Proof. The Pfaffian system has the form,

$$I = \{\alpha^1, \ldots, \alpha^k, \omega^1, \ldots, \omega^n, \alpha^{k+1}, \ldots, \alpha^m\}$$

where the constraints have been arranged in the order in which the axles appear in the chain. Choose $\pi = dx^n$, and note that by Lemma 51, $\{I, dx^n\}$ is integrable.

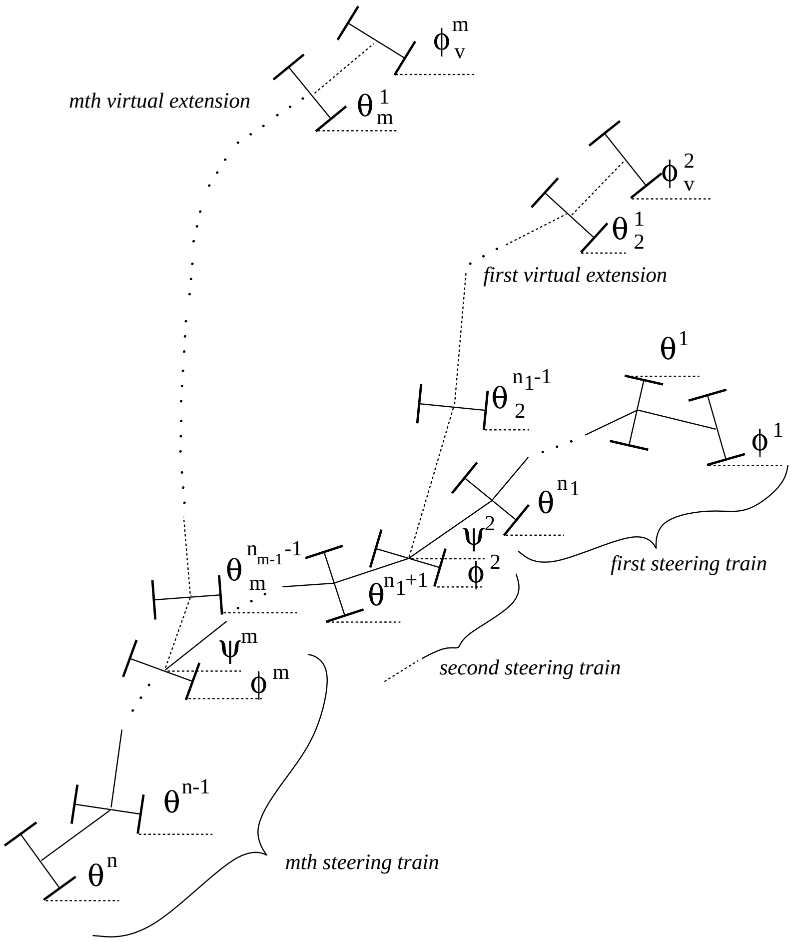

FIGURE 3. A multi-trailer system with n (passive) trailers and m (active) steering wheels, with a virtual extension of n_{j-1} virtual trailers in front of each steering wheel.

The derived flag associated to this case is simply found using either Lemma 50 or equation (169). It has the form:

$$\begin{aligned} I^{(1)} &= \{\omega^1, \omega^2, \dots, \omega^n\} \\ I^{(2)} &= \{\omega^2, \dots, \omega^n\} \\ &\vdots \\ I^{(n)} &= \{\omega^n\} \\ I^{(n+1)} &= \{0\} \end{aligned}$$

which is reminiscent of the N-trailers case from Section 3.2.

Equation (169),

$$d\omega^i \equiv d\theta^i \wedge \sec\theta^i dx^i \mod \omega^i$$

combined with equation (157),

$$dx^{i-1} \equiv f_{x^{i-1}} dx^i \mod \omega^{i-1}, \omega^i$$

gives the congruence

$$d\omega^i \equiv f_{\omega^i} d\theta^i \wedge dx^n \mod \omega^i, \omega^{i+1}, \ldots, \omega^n$$

which implies that $\{I^{(i)}, dx^n\}$ is integrable for $i = 1, \ldots, n+1$, and thus by Theorem 47, the system can be converted into extended Goursat normal form. □

The Goursat coordinates are defined by (x^n, y^n), the Cartesian position of the last passive axle, along with $\phi^1, \ldots, \phi^{m-1}$, the angles of the hitches.

Corollary 54 (Special cases). *As special cases of the general case described in Theorem 53, the following systems can be converted into Goursat form without prolongation:*

- *There is only one steering wheel, $m = 1$, which by convention is located at the front of the chain. This is the n-trailer problem of Section 3.2.*
- *There is one steering wheel at the front of the chain and another at the end of the chain, as in the firetruck example* [5, 33].
- *All the steering wheels are at the front, that is $n_1 = n_2 = \cdots = n_{m-1} = 0$.*
- *All the steering wheels are either at the front or the back of the chain, in a generalized firetruck situation.*
- *All the axles are steerable, $n = 0$.*
- *There is only one passive trailer, $n = 1$.*

The other special case which does not require prolongation to achieve Goursat normal form is slightly more complicated. The following can be shown.

Proposition 55. *If there are two sets of passive axles, separated by only one steerable axle, and the set towards the back has only one axle, then the system can be converted to extended Goursat normal form without prolongation.*

All configurations which do not satisfy either Theorem 53 or Proposition 55 require prolongation to be converted into extended Goursat normal form. The minimum dimension of the prolongation can be computed as follows. Recall that there are a total of n passive trailers and m steerable axles, and let k equal the index of the first steerable axle which has no passive trailers behind it. That is, $n_k = n_{k+1} = \cdots = n_m = n$ and $n_{k-1} < n$. There are two possible cases:

1. If $n_{k-1} = n-1$, then the minimum dimension of prolongation is $n_1+\cdots+n_{k-2}$.
2. Otherwise, a prolongation of dimension $n_1 + \cdots + n_{k-1}$ is needed to convert the system into extended Goursat normal form.

Now some specific multi-steering mobile robot systems will be considered and it will be shown how their associated Pfaffian systems satisfy the extended Goursat conditions.

Example. [Two, Three, or Four Axles] It is a simple exercise in combinatorics to check that all of the possible configurations with two or three axles and one, two or three steering wheels satisfy the conditions of Theorem 53. Note particularly that the firetruck example [5], sketched in Figure 4, satisfies these conditions with $n = 1$.

In addition, it can be shown that all except one configuration of a system with four axles will satisfy the conditions of Theorem 53. The exception is $m = 2$, two

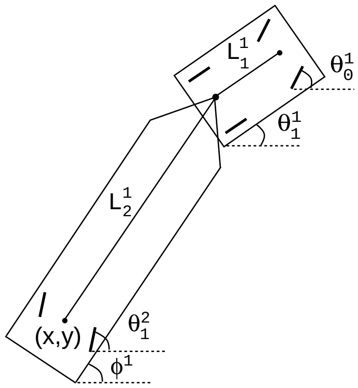

FIGURE 4. A sketch of the firetruck, with steering wheels on the front and back axles.

steerable axles, two passive axles, alternating. That is, the first and third axles are steerable, and the second and fourth axles are passive. This situation would arise if a car were towing another car and both of the cars had drivers at the steering wheels. This example satisfies Proposition 55, and thus can be converted into Goursat form without prolongation. □

The 5-axle system with two steering wheels is the lowest-dimensional case where interesting things begin to happen.

Example. [5-axle, 1-4 steering] First consider the 5-axle system with the first and fourth axles steerable, as sketched in Figure 5.

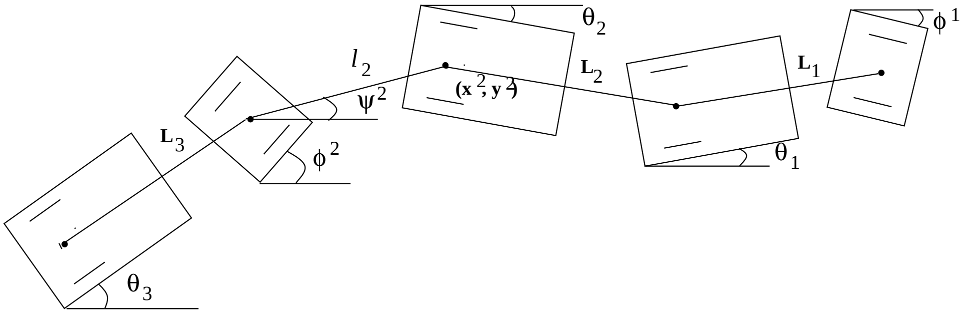

FIGURE 5. A 5-axle trailer system with the first and fourth axles steerable.

The constraints are that each axle rolls without slipping:

$$\omega^i = \sin\theta^i dx^i - \cos\theta^i dy^i \quad i = 1,2,3 \qquad \alpha^j = \sin\phi^j dx^j_s - \cos\phi^j dy^j_s \quad j = 1,2$$

The Pfaffian system is thus $I = \{\alpha^1, \omega^1, \omega^2, \alpha^2, \omega^3\}$ and a complement to this system is $\{d\phi^1, d\phi^2, dx^2\}$. This basis is adapted to the the derived flag,

$$\begin{array}{rl} I = & \{\alpha^1,\ \omega^1,\ \omega^2,\quad \alpha^2,\ \omega^3\} \\ I^{(1)} = & \qquad \{\omega^1,\ \omega^2,\qquad\quad \omega^3\} \\ I^{(2)} = & \qquad\qquad \{\omega^2\} \\ I^{(3)} = & \qquad\qquad\ \ \{0\} \end{array}$$

and it can be checked that each $\{I^{(i)}, dx^2\}$ is integrable. The coordinates which put the system into Goursat form are the following:

$$\begin{aligned} z^0 &= x^2 \\ z_1^1 &= y^2 \\ z_1^2 &= \begin{cases} \psi^2 - \tan\frac{(\psi^2-\theta^3)}{2} & L_3 = l_2 \\ \psi^2 + \frac{2L_3/l_2}{\sqrt{(\frac{L_3}{l_2})^2-1}} \arctan\left(\frac{\sqrt{(\frac{L_3}{l_2})^2-1}\tan(\frac{\psi^2-\theta^3}{2})}{1+L_3/l_2}\right) & L_3 > l_2 \\ \psi^2 + \frac{L_3/l_2}{\sqrt{1-(\frac{L_3}{l_2})^2}} \log\left(\frac{\sqrt{1-(\frac{L_3}{l_2})^2}\tan(\frac{\psi^2-\theta^3}{2})+1+L_3/l_2}{\sqrt{1-(\frac{L_3}{l_2})^2}\tan(\frac{\psi^2-\theta^3}{2})-1-L_3/l_2}\right) & L_3 < l_2 \end{cases} \end{aligned}$$

Note the dependence on the relative lengths of the hitches in the system. The remaining coordinates are defined by the relationships

$$\begin{aligned} z_k^1 &= \dot{z}_{k-1}^1/\dot{z}^0 \quad k = 2, \ldots, 4 \\ z_k^2 &= \dot{z}_{k-1}^2/\dot{z}^0 \quad k = 2, 3 \end{aligned}$$

Of course, by Theorem 48, this system can also be converted into an extended Goursat normal form using a prolongation of dimension two, and the coordinates in this case are given by:

$$\zeta^0 = x^3 \qquad \zeta_1^1 = y^3 \qquad \zeta_1^2 = \psi^2$$

together with the relations

$$\begin{aligned} \zeta_k^1 &= \dot{\zeta}_{k-1}^1/\dot{\zeta}^0 \quad k = 2, \ldots, 5 \\ \zeta_k^2 &= \dot{\zeta}_{k-1}^2/\dot{\zeta}^0 \quad k = 2, \ldots, 4 \end{aligned}$$

The two sets of coordinates (z^0, z_1^1, z_1^2) and $(\zeta^0, \zeta_1^1, \zeta_1^2)$ parameterize all integral curves for the system, in the sense that all the states and inputs to the system can be found by taking derivatives of these quantities. More differentiations will be required for the ζ coordinates.

Both coordinate transformations have two types of singularities. Because of the division by the derivative of z^0 (or ζ^0), whenever this coordinate is constant (corresponding to $\cos\theta^2$ or $\cos\theta^3$ respectively being zero), the transformations will be undefined. This type of singularity can be avoided by choosing a different coordinate chart at the singular point (interchanging x and y for example). A singularity also occurs when the angle between two adjacent axles is equal to $\pi/2$; at this point, some of the codistributions in the derived flag will lose rank. The derived flag is not defined at these points; nor is the transformation. The methods described herein will not work for controlling the multi-steering trailer system when the trailers must go through such a jack-knifed configuration. □

Example. [5-axle, 1-3 steering] The only instance of the 5-axle trailer system with two steering wheels which satisfies neither Theorem 53 or Proposition 55 has the first and third axles steerable, as shown in Figure 6.

The constraints are that each axle roll without slipping:

$$\omega^i = \sin\theta^i dx^i - \cos\theta^i dy^i \quad i = 1, 2, 3 \qquad \alpha^j = \sin\phi^j dx_s^j - \cos\phi^j dy_s^j \quad j = 1, 2$$

The Pfaffian system is $I = \{\alpha^1, \omega^1, \alpha^2, \omega^2, \omega^3\}$, and a complement to the system is given by $\{d\phi^1, d\phi^2, dx^3\}$.

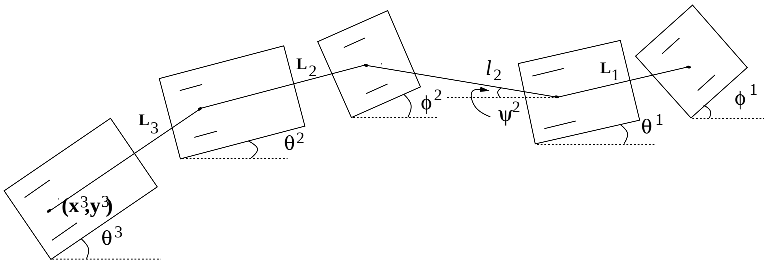

FIGURE 6. A 5-axle trailer system with the first and third axles steerable. This is the only configuration of the 5-axle system with two steering wheels which does not satisfy the conditions for converting to extended Goursat normal form without prolongation.

By Lemma 50, the derived flag has the form

$$\begin{array}{rl} I = & \{\alpha^1,\ \omega^1,\ \ \alpha^2,\ \omega^2,\ \omega^3\} \\ I^{(1)} = & \quad\ \{\omega^1,\ \qquad \omega^2,\ \omega^3\} \\ I^{(2)} = & \qquad\qquad\qquad\quad \{\omega^3\} \\ I^{(3)} = & \qquad\qquad\qquad\qquad \{0\} \end{array}$$

In order to have $\{I^{(2)}, \pi\}$ integrable, π must be $dx^3 \pmod{\omega^3}$. This will also give $\{I^0, dx^3\}$ integrable by Lemma 51, but a simple check will show that $\{I^{(1)}, dx^3\}$ is *not* integrable. Thus, as predicted by the theorems, this system does not satisfy the conditions for conversion to extended Goursat normal form without prolongation.

The system I can be prolonged by differentiation, adding the additional form $\nu = d\phi^2 - v\ dx^3$. The new coordinate v can be thought of as the tangent of the angle of the virtual axle that is added to the system in Theorem 48. The derived flag of the augmented system is:

$$\begin{array}{rl} J = & \{\alpha^1,\ \omega^1,\ \ \nu,\ \alpha^2,\ \omega^2,\ \omega^3\} \\ J^{(1)} = & \quad\ \{\omega^1,\ \qquad \alpha^2,\ \omega^2,\ \omega^3\} \\ J^{(2)} = & \qquad\qquad\qquad\quad \{\omega^2,\ \omega^3\} \\ J^{(3)} = & \qquad\qquad\qquad\qquad\quad \{\omega^3\} \\ J^{(4)} = & \qquad\qquad\qquad\qquad\qquad \{0\} \end{array}$$

and the systems $\{J^{(k)}, dx^3\}$ are integrable for all k. Thus, the prolonged system J can be converted into extended Goursat normal form. □

For the case of a 5-axle system with three steering wheels (two passive trailers), if the two passive trailers are connected we know from Theorem 53 that the system can be converted into extended Goursat normal form without prolongation. If the two passive trailers are separated by only one steerable axle, then we apply Proposition 55. The only configuration which does not satisfy one of these two conditions has the passive axles in the second and fifth positions, and this configuration will again require prolongation to convert it to extended Goursat normal form.

4. CONTROL SYSTEMS

The examples considered in Section 3, multi-body mobile robots towing trailers, required purely kinematic models. There were no drift terms considered, and no variable representing time which needed special consideration. Because of this, and the fact the velocity constraints could be represented as one-forms, exterior differential systems are particularly appropriate for their analysis.

Nonlinear control systems have traditionally been defined by distributions of vector fields on manifolds. Because of the duality between vector fields and one-forms, as seen in Section 2.1.1, a control system can also be defined as a Pfaffian system on a manifold and analyzed using techniques from exterior differential systems. In this section, we will present some results on linearization for nonlinear control systems and also examine the connections between the two different formalisms of vector fields and one-forms.

We will consider the nonlinear dynamical system:

$$\dot{x} = f(x, u) \tag{180}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and f is a smooth map [7]:

$$\begin{aligned} f : \mathbb{R}^n \times \mathbb{R}^m &\longrightarrow T\mathbb{R}^n \\ (x, u) &\longrightarrow f(x, u) \in T_x\mathbb{R}^n \end{aligned}$$

A very important special case of system (180) is the one where the input enters affinely in the dynamics:

$$\dot{x} = f(x) + g(x)u \tag{181}$$

where $g(x) = [g_1(x) \dots g_m(x)]$ and $g_i(x)$ are smooth vector fields. Most of the results presented here will be concerned with systems belonging to this class, even though some can be extended to the the more general case (180).

We would like to establish conditions under which the dynamics of (180) and (181) are adequately described by those of a linear system:

$$\dot{\hat{x}} = A\hat{x} + B\hat{u} \tag{182}$$

where $\hat{x} \in \mathbb{R}^{\hat{n}}$, $\hat{u} \in \mathbb{R}^m$, $A \in \mathbb{R}^{\hat{n} \times \hat{n}}$ and $B \in \mathbb{R}^{\hat{n} \times m}$ with $\hat{n} \geq n$.

4.1. **Static Feedback Linearization.** One of the best-studied problems in nonlinear control is that of exact linearization using static state feedback and change of coordinates. First we will present the well-known results on static feedback linearization, and then we will show how these results can be restated in terms of the Goursat normal form. In the next section, we will consider the problem of dynamic feedback linearization.

4.1.1. *Problem Statement.* Following the notation of Isidori [17], the problem of exact linearization by static state feedback and coordinate transformation can be stated as follows:

Problem 1. (State Space Exact Linearization Problem)
Given a control system of the form (180) *and an initial state x°, find, if possible, a neighborhood U of x°, a feedback function $c : U \times \mathbb{R}^m \to \mathbb{R}^m$, a coordinate*

[7] Most of the techniques presented here can be generalized to the case where the state evolves on a manifold. $\mathbb{R}^n$ will be used to simplify the calculations

transformation $z = \Phi(x)$, *all defined on* U, *and matrices* $A \in \mathbb{R}^{n\times n}$ *and* $B \in \mathbb{R}^{n\times m}$, *such that:*

$$\left[\frac{\partial \Phi}{\partial x}(f(x, c(x,0)))\right]_{x=\Phi^{-1}(z)} = Az \tag{183}$$

$$\left[\frac{\partial \Phi}{\partial x}\frac{\partial (f(x, c(x,v))}{\partial v}\right]_{x=\Phi^{-1}(z)} = B \tag{184}$$

$$rank(B \;\; AB \;\; \dots \;\; A^{n-1}B) = n \tag{185}$$

In the special case of systems affine in the inputs, the problem simplifies to:

Problem 2. *Given a control system of the form* (181) *and an initial state* x°, *find, if possible, a neighborhood* U *of* x°, *a pair of feedback functions* $a(x)$ *and* $b(x)$, *a coordinate transformation* $z = \Phi(x)$, *all defined on* U, *and matrices* $A \in \mathbb{R}^{n\times n}$ *and* $B \in \mathbb{R}^{n\times m}$, *such that:*

$$\left[\frac{\partial \Phi}{\partial x}(f(x) + g(x)a(x))\right]_{x=\Phi^{-1}(z)} = Az \tag{186}$$

$$\left[\frac{\partial \Phi}{\partial x}(g(x)b(x))\right]_{x=\Phi^{-1}(z)} = B \tag{187}$$

$$rank(B \;\; AB \;\; \dots \;\; A^{n-1}B) = n \tag{188}$$

The last condition of both problem statements allows us to assume that without loss of generality, the resulting linear system will be in Brunovsky canonical form, i.e.:

$$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & A_m \end{bmatrix} \qquad B = \begin{bmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & b_m \end{bmatrix}$$

$$A_i = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \qquad b_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

The dimensions of A_i correspond to the Kronecker indices of the pair (A, B).

4.1.2. *The Vector Field Approach.* The standard results on linearization by static state feedback and coordinate transformation concern systems which are affine in the input (181). The relevant theorems can be found in [17, 23]; we will use the notation and definitions of Isidori [17].

Theorem 56. *For the control system* (181) *define the filtration:*

$$\begin{aligned} G_0 &= span\{g_1, \dots, g_m\} \\ G_{i+1} &= G_i + \; span\{[f, G_i]\} = \; span\{ad_f^k g_j : 0 \le k \le i+1, 1 \le j \le m\} \end{aligned}$$

Suppose the distribution $G_0(x)$ *has dimension* m *at* x°. *Then, the state space exact linearization problem is solvable if and only if:*

1. *for each* $0 \le i \le n-1$ *the distribution* G_i *has constant dimension near* x°
2. *The distribution* G_{n-1} *has dimension* n
3. *for each* $0 \le i \le n-2$ *the distribution* G_i *is involutive*

Proof. In [17, 23]. □

If the system has only one input ($m = 1$), the involutivity of G_{n-2} will imply the involutivity of the other G_i. Thus, the conditions for static feedback linearization can be restated in the single-input case as follows:

Corollary 57. *The state space exact linearization problem for a control system* (181) *with a single input is solvable if and only if:*

1. *the distribution* $[g(x)\ \ ad_f g(x)\ \ldots\ ad_f^{n-1} g(x)]$ *has dimension* n *at* x°.
2. *the distribution* $G_{n-2} = \ span\{g,\ ad_f g,\ \ldots,\ ad_f^{n-2} g\}$ *is involutive near* x°.

Proof. Special case of Theorem 56 when $m = 1$. □

It should be noted that even for the multi-input case, the involutivity of certain distributions (namely those corresponding to the Kronecker indices of the resulting linear system) implies the involutivity of others. However, an equivalent statement of Theorem 56 that takes this fact into account is notationally complicated.

4.1.3. *The Pfaffian System Approach.* The problem of linearization can also be approached from the point of view of exterior differential systems. Note that any control system of the form (180) can also be thought of as a Pfaffian system of co-dimension $m + 1$ in $\mathbb{R}^{n+m+1}$. The corresponding ideal is generated by the co-distribution:

$$I = \{dx_i - f_i(x, u)dt : i = 1, \ldots, n\} \tag{189}$$

The $n+m+1$ variables for the Pfaffian system correspond to the n states, m inputs and time, t. For the special case of the affine system (181) the co-distribution becomes:

$$I = \{dx_i - (f_i(x) + \sum_{j=1}^{m} g_{ij}(x)u_j)dt : i = 1, \ldots, n\} \tag{190}$$

In this light the extended Goursat normal form looks remarkably similar to the Brunovsky normal form with Kronecker indices $s_j, j = 1, \ldots, m$. Indeed if we identify coordinates $z^0, z^j_{s_j+1}, j = 1, \ldots, m$ in the Goursat Normal Form with $t, u_j, j = 1, \ldots, m$, the Pfaffian system becomes equivalent (in vector field notation) to a collection of m chains of integrators, each one of length s_j and terminating with an input in the right hand side. With this in mind, Theorems 46 and 47, which provide conditions under which a Pfaffian system can be transformed to extended Goursat normal form, can be viewed as linearization theorems with the additional restriction that $\pi = dt$.

An equivalent formulation of the conditions of Theorem 41 involving the annihilating distributions is given by Murray [22]. The result is restricted to Pfaffian systems of co-dimension two.

Theorem 58. *Given a 2-dimensional distribution* Δ *construct two filtrations:*

$$\begin{array}{ll} E_0 = \Delta & F_0 = \Delta \\ E_{i+1} = E_i + [E_i, E_i] & F_{i+1} = F_i + [F_i, F_0] \end{array}$$

If all the distributions are of constant rank and:

$$dim\ E_i = dim\ F_i = i + 2 \quad i = 0, \ldots, n-2$$

there exists a local basis $\{\alpha^1, \ldots, \alpha^s\}$ *and a one-form* π *such that the Goursat congruences are satisfied for the differential system* $I = \Delta^\perp$.

Proof. In [22]. □

In [22] this Theorem is shown to be equivalent to Theorem 41. However, there is no known analog of Theorem 58 to the extended Goursat case covered by Theorems 46 and 47.

4.1.4. *The connection between the two approaches.* We explicitly work through the connection between the classical static feedback linearization theorem (Theorem 56) and the extended Goursat normal form theorem (Theorem 47).

Proposition 59. *The control system* (181) *satisfies the conditions of Theorem 56 if and only if the corresponding Pfaffian system* (190) *satisfies the conditions of Theorem 47 for* $\pi = dt$.

Proof. Consider control system (181) and the equivalent Pfaffian system (190). For simplicity, we will consider the case $m = 2$. The Pfaffian system $I^{(0)}$ and its annihilating distribution Δ_0 are given by:

$$\begin{aligned} I^{(0)} &= \{dx_i - (f_i(x) + g_{i1}(x)u_1 + g_{i2}(x)u_2)dt : i = 1, \ldots, n\} \\ \Delta_0 &= (I^{(0)})^\perp = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ f + g_1u_1 + g_2u_2 \end{bmatrix} \right\} \end{aligned}$$

As the notation suggests, the top three entries in each vector field in the distribution Δ_0 are scalars (corresponding to the coordinates t, u_1 and u_2) while the bottom entry is a column vector of dimension n. We will construct the derived flag $I^{(0)} \supset I^{(1)} \supset \ldots \supset I^{(N)}$ and the corresponding orthogonal filtration $\Delta_0 \subset \Delta_1 \subset \ldots \subset \Delta_N$. We will denote by $\hat{I}^{(i)} = \{I^{(i)}, dt\}$ and $\hat{\Delta}_i = (\hat{I}^{(i)})^\perp$. We will go through the conditions of Theorem 47 step by step, assuming $\pi = dt$:

Step 0: As above:

$$\begin{aligned} I^{(0)} &= \{dx_i - (f_i(x) + g_{i1}(x)u_1 + g_{i2}(x)u_2)dt : i = 1, \ldots, n\} \\ \Delta_0 &= (I^{(0)})^\perp = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ f + g_1u_1 + g_2u_2 \end{bmatrix} \right\} \\ &= \{v_1, v_2, v_3\} \end{aligned}$$

The condition of Theorem 47 requires that $\hat{I}^{(0)} = \{I^{(0)}, dt\}$ be integrable. Its annihilator is $\hat{\Delta}_0 = \{v_1, v_2\}$ which is indeed involutive since $[v_1, v_2] = 0$ are constant vector fields.

Step 1: It is easy to show that:

$$[v_1, v_2] = 0 \quad [v_1, v_3] = \begin{bmatrix} 0 \\ 0 \\ 0 \\ g_1 \end{bmatrix} \quad [v_2, v_3] = \begin{bmatrix} 0 \\ 0 \\ 0 \\ g_2 \end{bmatrix}$$

Therefore:

$$\begin{aligned}
I^{(1)} &= \{\alpha \in I^{(0)} : d\alpha \equiv 0 \bmod I^{(0)}\} \\
\Delta_1 &= (I^{(1)})^{\perp} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ f + g_1 u_1 + g_2 u_2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ g_1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ g_2 \end{bmatrix} \right\} \\
&= \{v_1, v_2, v_3, v_4, v_5\}
\end{aligned}$$

The condition of Theorem 47 requires that $\hat{I}^{(1)} = \{I^{(1)}, dt\}$ be integrable. To check this, consider its annihilator $\hat{\Delta}_1 = \{v_1, v_2, v_4, v_5\}$ being involutive. Now:

$$[v_1, v_2] = [v_1, v_4] = [v_1, v_5] = [v_2, v_4] = [v_2, v_5] = 0 \text{ and } [v_4, v_5] = \begin{bmatrix} 0 \\ 0 \\ 0 \\ [g_1, g_2] \end{bmatrix}$$

Therefore $\hat{\Delta}_1$ is involutive if and only if $[g_1, g_2]$ is in the span of $\{g_1, g_2\}$. The condition of Theorem 47 holds for the first iteration of the derived flag if and only if distribution G_0 of Theorem 56 is involutive.

Step 2: We compute the bracket of the vector fields v_3 and v_4.

$$[v_3, v_4] = \begin{bmatrix} 0 \\ 0 \\ 0 \\ ad_f g_1 - [g_1, g_2] u_2 \end{bmatrix}$$

The computation of $[v_3, v_5]$ is similar. Therefore, assuming that the conditions of Step 1 hold and in particular that $[g_1, g_2] \in \operatorname{span}\{g_1, g_2\}$:

$$\begin{aligned}
I^{(2)} &= \{\alpha \in I^{(1)} : d\alpha \equiv 0 \bmod I^{(1)}\} \\
\Delta_2 &= (I^{(2)})^{\perp} = \Delta_1 + \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ ad_f g_1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ ad_f g_2 \end{bmatrix} \right\} \\
&= \{v_i : i = 1, \ldots 7\}
\end{aligned}$$

The condition of Theorem 47 requires that $\hat{I}^{(2)}$ be integrable. This is equivalent to $\hat{\Delta}_2 = \{v_1, v_2, v_4, v_5, v_6, v_7\}$ being involutive. As before the only pairs whose involutivity needs to be verified are the ones not involving v_1 and v_2, i.e. the condition is equivalent to $\{g_1, g_2, ad_f g_1, ad_f g_2\}$ being involutive. Overall the condition of Theorem 47 holds for the the second iteration of the derived flag if and only if distribution G_1 of Theorem 56 is involutive.

Step i: Assume that:

$$\Delta_{i-1} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ f + g_1 u_1 + g_2 u_2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ ad_f^k g_1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ ad_f^k g_2 \end{bmatrix} \right\}$$

for $0 \leq k \leq i-2$. Also assume that $\hat{I}^{(k)}$, $0 \leq k \leq i-1$ are integrable, or, equivalently, that $\hat{\Delta}_k$ for $0 \leq k \leq i-1$ (which is the same as Δ_k without the third vector field) are involutive, or, equivalently, that $G_k = \{ad_f^l g_j : 0 \leq l \leq k, j = 1,2\}$ for $0 \leq k \leq i-2$ are involutive. Construct $\Delta_i = \Delta_{i-1} + [\Delta_{i-1}, \Delta_{i-1}]$. By involutivity of $\hat{\Delta}_{i-1}$ and the construction of the filtration the only terms not already in Δ_{i-1} are ones of the form:

$$\left[\begin{bmatrix} 1 \\ 0 \\ 0 \\ f + g_1 u_1 + g_2 u_2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ ad_f^{i-2} g_1 \end{bmatrix}\right] = \begin{bmatrix} 0 \\ 0 \\ 0 \\ ad_f^{i-1} g_1 + [g_1, ad_f^{i-2} g_1] u_1 + [g_2, ad_f^{i-2} g_1] u_2 \end{bmatrix}$$

and similarly for $ad_f^{i-1} g_2$. By the assumed involutivity of $\hat{\Delta}_{i-1}$ the last two terms are already in Δ_{i-1}. Therefore we can write:

$$\Delta_i = \Delta_{i-1} + \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ ad_f^{i-1} g_1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ ad_f^{i-1} g_2 \end{bmatrix} \right\}$$

The condition of Theorem 47 requires that $\hat{I}^{(i)}$ be integrable, or equivalently that $\hat{\Delta}_i$ be involutive. As before the only pairs that can cause trouble are the ones not involving v_1 and v_2. Hence the condition is equivalent to $G_{i-1} = \{ad_f^k g_j : 0 \leq k \leq i-1, j = 1,2\}$ being involutive.

By induction, the condition of Theorem 47 holds for the i^{th} iteration of the derived flag if and only if distribution G_{i-1} is involutive, i.e. if and only if condition (3) of Theorem 56 holds. In addition, note that the dimension of G_i keeps increasing by *at least one* until, for some M, $G_{M-1} = G_M$. The involutivity assumption on G_{M-1} prevents any further increase in dimension after this stage is reached. Since the dimension of G_i is necessarily bounded above by n, the number of steps until saturation is bounded by the maximum final dimension, $M \leq n$. By construction, the dimension of Δ_i is three greater than the dimension of G_{i-1}. Moreover $\Delta_M = \Delta_{M+1}$ and therefore $I^{(M)} = I^{(M+1)}$, i.e. the derived flag stops shrinking after M steps. The remaining condition of Theorem 47, namely that there exists N such that $I^{(N)} = \{0\}$ is equivalent to the existence of M such that $I^{(M)} = \{0\}$, or that Δ_M has dimension $n+3$. As noted above, this is equivalent G_{M-1} having dimension n. Since $M \leq n$, this can also be stated as G_{n-1} having dimension n, i.e. condition (2) of Theorem 56. The remaining condition of Theorem 56, namely that the dimension of G_i is constant for all $0 \leq i \leq n-1$, is taken care of by the implicit assumption that all co-distributions in the derived flag have constant dimension. $\square$

Note that a coordinate transformation in the exterior differential systems context corresponds to a coordinate transformation together with a state feedback in the vector field notation. Because the state space $\mathbb{R}^{n+m+1}$ in the forms context does not discriminate between states, inputs and time, a coordinate transformation on this larger space can make the inputs in the original coordinates functions of the state in the original coordinates and possibly time. It can be shown (see [11]) that time need not enter into the transformation at all; that is, if the conditions of Theorem 46 are satisfied, a time-invariant state feedback and coordinate change can always be found. In addition the coordinate transformation can also be chosen to be independent of both time and input.

Theorems 46 and 47 in their general form are **not** equivalent to Theorem 56. The extended Goursat Theorems allow π to be any integrable one-form and not just dt. Therefore we expect more systems to match the conditions of Theorems 46 and 47 than those of Theorem 56. However, a choice of π other than dt implies a rescaling of time as a function of the state. Even though this effect is very useful for the case of driftless systems (where the role of time is effectively played by an input), solutions for $\pi \neq dt$ are probably not very useful for linearizing control systems with drift. Because of their generality, Theorems 46 and 47 are capable of dealing with the more general case of control systems of the form (189) (or equivalently (180)), as well as drift-free systems which were investigated in Section 3. Equivalent conditions for the vector field case have not been thoroughly investigated.

Finally, Theorem 58 is a very interesting alternative to Theorems 46 and 47 since it provides a way of determining if a Pfaffian system can be converted to Goursat normal form just by looking at the annihilating distributions, without having to determine a one-form π or an appropriate basis. Unfortunately a generalization to multi-input systems (or more precisely to the extended Goursat normal form) is not easy to formulate. It should be noted that the conditions on the filtrations are very much like involutivity conditions. It is interesting to try to relate these conditions to the conditions of Theorem 47 (the connection to the conditions of Theorem 46 is provided in [22]) and see if a formulation for the extended problem can be constructed in this way.

4.2. **Dynamic Feedback Linearization.** We now consider the more general problem of linearization by dynamic state feedback. The feedback compensator is allowed to have its own dynamics, and we search for a transformation on the extended space, including the states of the original systems and the controller, into a linear form.

4.2.1. *Problem Statement.* Following the notation of Charlet, Lévine, and Marino [6], the problem of exact linearization by dynamic state feedback and coordinate transformation can be stated as follows:

Problem 3. (Dynamic Feedback Linearization Problem)[8]
Given a control system of the form (180)*, find, if possible, a dynamic feedback compensator:*

$$\begin{aligned} \dot{w} &= a(x,w) + B(x,w)v \\ u &= \alpha(x,w) + \beta(x,w)v \end{aligned}$$

where $w \in \mathbb{R}^q, v \in \mathbb{R}^m$ *and an extended state space diffeomorphism* $z = \Phi(x,w), z \in \mathbb{R}^{n+q}$ *such that the resulting system is linear and controllable (without loss of generality in Brunovsky form).*

[8] As in the case of Problem 1, all conditions may be restricted to a neighborhood U of an equilibrium point x°.

It should be noted here that the problem statement requires that both the system and the controller dynamics be rendered linear and controllable. An interesting special case of the general dynamic feedback linearization problem restricts the controller dynamics to consist of derivatives of the inputs:

Problem 4. *(***Feedback Linearization by Dynamic Extension***)*
Given a control system of the form (180)*, find, if possible, a dynamic feedback compensator of the form:*

$$\begin{aligned}
\dot{w}_j^i &= w_j^{i+1} \quad 0 \le i \le \mu_j - 1, \mu_j > 0 && (191)\\
w_j^{\mu_j} &= \alpha_j(x,w) + \sum_{\ell=1}^{m} \beta_{j,\ell}(x,w) v_\ell(t) \quad 1 \le j \le m, \mu_j > 0 && (192)\\
u_j &= w_j^0 \quad 1 \le j \le m, \mu_j > 0 && (193)\\
u_j &= \alpha_j(x,w) + \sum_{\ell=1}^{m} \beta_{j,\ell}(x,w) v_\ell(t) \quad 1 \le j \le m, \mu_j = 0 && (194)
\end{aligned}$$

for some integers $\mu_j \ge 0$ and $\beta(x,w)$ has full rank m in a neighborhood an equilibrium point in $\mathbb{R}^{n+\mu}$, $\mu = \sum_{i=1}^m \mu_j$.

For dynamic extension, chains of integrators are added in front of some of the input channels, and the new inputs are defined to be linear combinations of the resulting derivatives of the original inputs:

$$\begin{pmatrix} u_1^{(\mu_1)} \\ \vdots \\ u_m^{(\mu_m)} \end{pmatrix} = \alpha(x,w) + \beta(x,w) \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \tag{195}$$

4.2.2. *The Vector Field Approach.* The results for dynamic feedback linearization using the vector field approach are again restricted to systems of the form (181). The problem of linearization by general dynamic state feedback and coordinate transformation is still mostly open. Even for the special case of dynamic extension no necessary and sufficient conditions exist. The following results are proven by Charlet, Lévine, and Marino [6]:

Theorem 60. *If system* (181) *is locally dynamic feedback linearizable, then its Jacobian linearization at the origin is completely controllable.*

Proof. In [6]. □

Theorem 61. *If for a set of integers $\{\mu_1, \ldots, \mu_m\}$, $0 \le \mu_1 \le \ldots \le \mu_m$, $\mu = \sum_{i=1}^m \mu_j$, the distributions, up to input reordering,*

$$\begin{aligned}
\Delta_0 &= span\{g_k : \mu_k = 0\}\\
\Delta_{i+1} &= \Delta_i + ad_f \Delta_i + \; span\{g_k : \mu_k = i+1\} \quad i \ge 0
\end{aligned}$$

are such that in a neighborhood of the origin in $\mathbb{R}^n$:

1. *Δ_i is of constant rank for $0 \le i \le n + \mu_m - 1$*
2. *Δ_i is involutive for $0 \le i \le n + \mu_m - 1$*
3. *$rank \Delta_{n+\mu_m-1} = n$*
4. *$[g_j, \Delta_i] \subset \Delta_{i+1}$ for all $j, 1 \le j \le m$ such that $\mu_j \ge 1$ and all $i, 0 \le i \le n + \mu_m - 1$*

then the system is locally dynamic feedback linearizable by dynamic extension and a local diffeomorphism on a neighborhood of the extended state space $\mathbb{R}^{n+\mu}$.

Proof. In [6]. □

In [6], the necessary condition of Theorem 60 is shown not to be sufficient and the sufficient condition of Theorem 61 is shown to not be necessary by means of counterexamples. Although the conditions of Theorem 61 are not necessary, they depend only on the original vector fields of the control system.

4.2.3. *The Prolongation Approach.* Problem 4 can also be approached in the framework of Pfaffian systems by means of prolongations by differentiation, as described in Section 3.4. The following theorem can be stated:

Theorem 62. *Consider the Pfaffian system* $I = \{\alpha^1, \ldots, \alpha^n\}$ *on* $\mathbb{R}^{n+m+1}$ *with independence condition* dt *and complement* $\{dt, du_1, \ldots, du_m\}$*. If there exists a list of non-negative integers* $\{\mu_1, \ldots, \mu_m\}, \mu = \sum_{i=1}^m \mu_j$ *such that the prolonged system:*

$$\begin{aligned} J \;=\{ \quad & \alpha^1, \ldots, \alpha^n \\ & du_1 - w_1^1 dt, \ldots, dw_1^{\mu_1 - 1} - w_1^{\mu_1} dt, \\ & du_2 - w_2^1 dt, \ldots, dw_2^{\mu_2 - 1} - w_2^{\mu_2} dt, \\ & \ldots \\ & du_1 - w_m^1 dt, \ldots, dw_m^{\mu_m - 1} - w_m^{\mu_m} dt\} \end{aligned}$$

satisfies the condition that $\{J^{(k)}, dt\}$ *is integrable for all* k*, then* I *can be transformed to extended Goursat Normal Form using prolongation by differentiation.*

Proof. The result follows from Theorem 48. □

This theorem has the advantage over Theorem 61 that if the system is linearizable by a dynamic extension of order $\mu = \{\mu_1, \ldots, \mu_m\}$, then the conditions of the Theorem will be satisfied. Of course, the derived flag must be recomputed for every choice of μ.

4.2.4. *The Infinitesimal Brunovsky Form.* An altogether different approach to dynamic feedback linearization is presented in [2]. It revolves around an alternative flag construction that can be used to derive a special normal form, the Infinitesimal Brunovsky Form. An interesting fact about this construction is that *any* accessible nonlinear system can be brought into this form.

Consider the system (181) and let $\mathcal{K}$ denote the field of meromorphic functions of $x, u, \dot{u}, \ldots$, where the dot stands for the usual time differentiation. Let $\mathcal{E}$ denote the $\mathcal{K}$ vector space of one forms, spanned by:

$$\{dx_1, \ldots, dx_n, du_1, \ldots, du_m, d\dot{u}_1, \ldots, d\dot{u}_m, \ldots\} = \{dx, du, d\dot{u}, \ldots\}$$

Define the time derivative of $\omega = \sum_j \alpha_j dv_j \in \mathcal{E}$ by

$$\begin{aligned} \dot{\omega} &= \sum_j (\dot{\alpha}_j dv_j + \alpha_j d\dot{v}_j) \\ \dot{\alpha}_j &= \frac{\partial \alpha_j}{\partial x}(f(x) + g(x)u) + \sum_{j \geq 0} \frac{\partial \alpha_j}{\partial u^{(j)}} u^{(j+1)} \\ d\dot{x}_i &= \sum_{k=1}^n \left(\frac{\partial f_i}{\partial x_k} + \sum_{j=1}^m \frac{\partial g_{ij}}{\partial x_k} u_j\right) dx_k + \sum_{k=1}^m g_{ij} du_j \end{aligned}$$

The relative degree of a one form ω is defined as the smallest integer r such that $\omega^{(r)} \notin \operatorname{span}_{\mathcal{K}}\{dx\}$. If such an integer does not exist define $r = \infty$. Consider a flag defined iteratively by:

$$H_0 = \operatorname{span}_{\mathcal{K}}\{dx, du\} \tag{196}$$

$$H_k = \{\omega \in H_{k-1} : \dot{\omega} \in H_{k-1}\} \quad k > 0 \tag{197}$$

Clearly, $\mathcal{E} \supset H_0 \supset H_1 = \operatorname{span}_{\mathcal{K}}\{dx\} \supset H_2 \supset \dots$. Moreover, because the dimension of H_1 is finite (n), the flag will stop decreasing after a finite number of steps, i.e. there exists $k^* > 0$ such that $H_{k^*+1} = H_{k^*+2} = \dots \stackrel{\text{def}}{=} H_\infty$. The importance of this flag is highlighted by the following theorems:

Proposition 63. *The following statements are equivalent:*

1. *The system* (181) *satisfies the strong accessibility rank condition*
2. *Any non-zero form has finite relative degree*
3. $H_\infty = \{0\}$

Proof. In [2]. □

Theorem 64. *Suppose $H_\infty = \{0\}$. There exist a list of integers $\{r_1, \dots, r_m\}$, invariant under regular static state feedback, and m one forms $\omega_1, \dots, \omega_m$ with relative degrees $r_1, \dots, r_m$ such that:*

1. $span_{\mathcal{K}}\{\omega_i^{(j)}, 1 \le i \le m, 0 \le j \le r_i - 1\} = span_{\mathcal{K}}\{dx\}$
2. $span_{\mathcal{K}}\{\omega_i^{(j)}, 1 \le i \le m, 0 \le j \le r_i\} = span_{\mathcal{K}}\{dx, du\}$
3. *The forms $\{\omega_i^{(j)}, 1 \le i \le m, j \ge 0\}$ are linearly independent. In particular, $\sum_{i=1}^m r_i = n$.*

Proof. In [2]. □

An equivalent form of the Theorem 64 is the following:

Corollary 65. *Suppose $H_\infty = \{0\}$. Then the basis $\{\omega_{i,j}, 0 \le j \le r_i, 1 \le i \le m\}$ of $span_{\mathcal{K}}\{dx\}$ defined by $\omega_{i,j} = \omega_i^{(j-1)}$ yields:*

$$\begin{aligned}
\dot{\omega}_{i,1} &= \omega_{i,2} \\
&\vdots \\
\dot{\omega}_{i,r_i-1} &= \omega_{i,r_i} \\
\dot{\omega}_{i,r_i} &= \sum_{j=1}^{n} a_{i,j} dx_j + \sum_{j=1}^{m} b_{i,j} du_j
\end{aligned}$$

where $a_{i,j}, b_{i,j} \in \mathcal{K}$ and the matrix $[b_{i,j}]$ has an inverse in the ring of $m \times m$ matrices with entries in $\mathcal{K}$.

Proof. In [2]. □

The last representation, called the Infinitesimal Brunovsky form, highlights the similarity of this construction with the regular Brunovsky form: the two forms are identical, with scalar quantities replaced by one forms. Using this normal form the following theorems can be proved:

Theorem 66. *The system* (181) *is linearizable by static state feedback if and only if $H_\infty = \{0\}$ and for all $k = 1, \dots, k^*$, H_k is integrable.*

Proof. In [2]. □

Theorem 67. *Suppose $H_\infty = \{0\}$ and let $\Omega = (\omega_1, \dots, \omega_m)^T$. There exists a system of linearizing outputs $y = h(x, u, \dots, u^{(\nu-1)}) \in \mathbb{R}^m$ if and only if there exist an invertible polynomial operator $P \in \mathcal{K}^{m\times m}\left[\frac{d}{dt}\right]$ such that $d(P\Omega) = 0$.*

Proof. In [2]. □

4.2.5. *Connection between the approaches.* We now consider some results where dynamic feedback linearization is successful as opposed to static feedback linearization. A simple calculation reveals the following:

Proposition 68. *Consider system* (180). *An extended system obtained by adding the same number of integrators in front of each input is linearizable by static state feedback if and only if the original system is linearizable by static state feedback.*

Proof. The proof is easy to understand in the exterior differential systems framework. If k integrators are added to each input, then the k^{th} derived system of the extended system is equal to the original system. For details see [27]. □

Therefore, dynamic extension is only useful for feedback linearization if a different number of integrators are added to each input channel, and there is at least one input channel to which no integrators are added.

Corollary 69. *A single input system of the form* (181) *is feedback linearizable by dynamic extension if and only if it is static feedback linearizable.*

Proof. The result follows from the above proposition. □

In fact Corollary 69 also holds for general dynamic feedback.

Theorem 70. *The following statements are equivalent:*

1. *System* (181) *with $m = 1$ is static feedback linearizable*
2. *System* (181) *with $m = 1$ is dynamic feedback linearizable*

Proof. See [6, 2]. □

In other words, dynamic state feedback is only helpful in the case of multi-input systems or Pfaffian systems with codimension greater than two. The proof for Theorem 70 can be found in [6] for the vector field formalism and in [2] for the infinitesimal Brunovsky form formalism. It should be noted here that the second proof is extremely simple whereas the vector field proof is rather complicated.

The relation between the dynamic extension results in the vector field approach (Theorem 61) and the exterior differential systems approach (Theorem 62) can be seen from the following statement:

Proposition 71. *If there exist integers satisfying the conditions of Theorem 61, then those same integers satisfy the conditions of Theorem 62 for $\pi = dt$.*

Proof. The proof follows if we assume the system is in Brunovsky canonical form. If the conditions of Theorem 61 are satisfied this can be done without loss of generality, as the result is intrinsic and therefore independent of the chosen coordinate frame. □

The converse is not true, as illustrated by the following example:

Example. Consider the following control system, proposed by Charlet, Lévine and Marino as a counterexample to the sufficiency of Theorem 61:

$$\begin{aligned} \dot{x}_1 &= x_2 + x_3 u_2 \\ \dot{x}_2 &= x_3 + x_1 u_2 \\ \dot{x}_3 &= u_1 + x_2 u_2 \\ \dot{x}_4 &= u_2 \end{aligned} \tag{198}$$

This control system can be written in vector field notation as:

$$\dot{x} = f(x) + g_1(x)u_1 + g_2(x)u_2$$

with the drift and input vector fields given by

$$f(x) = \begin{bmatrix} x_2 \\ x_3 \\ 0 \\ 0 \end{bmatrix}, \quad g_1(x) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad g_2(x) = \begin{bmatrix} x_3 \\ x_1 \\ x_2 \\ 1 \end{bmatrix}$$

The distribution $G_0 = \{g_1, g_2\}$ is not involutive, and thus, by Theorem 56 the system is not feedback linearizable.

Of course, we could also represent the control system (198) as a Pfaffian system,

$$I = \{\alpha^1, \alpha^2, \alpha^3, \alpha^4\}$$

with the one-forms given by

$$\begin{aligned} \alpha^1 &= dx_1 - (x_2 + x_3 u_2)dt \\ \alpha^2 &= dx_2 - (x_3 + x_1 u_2)dt \\ \alpha^3 &= dx_3 - (u_1 + x_2 u_2)dt \\ \alpha^4 &= dx_4 - u_2 dt \end{aligned}$$

A complement to this Pfaffian system is $\{du_1, du_2, dt\}$. The first derived system can be shown to be

$$I^{(1)} = \{\alpha^4 - \frac{1}{x_3}\alpha^1, \alpha^2 - \frac{x_1}{x_3}\alpha^1\}$$

and since $\{I^{(1)}, dt\}$ is not integrable, the system is not feedback linearizable.

We now consider a dynamic extension of order 3 on u_2. We can represent this extended system by $J = \{\alpha^1, \alpha^2, \alpha^3, \alpha^4, \omega^1, \omega^2, \omega^3\}$ where the one-forms that are added correspond to the dynamic extension, and the new states that are added are the first three derivatives of u_2:

$$\begin{aligned} \omega^1 &= du_2 - w_2^1 dt \\ \omega^2 &= dw_2^1 - w_2^2 dt \\ \omega^3 &= dw_2^2 - w_2^3 dt \end{aligned}$$

A complement to J is $\{du_1, dw_2^3, dt\}$ (note that I and J have the same codimension). Computing the derived flag of the extended system, we find that

$$\begin{aligned} J^{(1)} &= \{\alpha^1, \alpha^2, \alpha^4, \omega^1, \omega^2\} \\ J^{(2)} &= \{\alpha, \alpha^4, \omega^1\} \\ J^{(3)} &= \{\alpha\} \\ J^{(4)} &= \{0\} \end{aligned}$$

where $\alpha = \alpha^1 - u_2\alpha^2$. Each $\{J^{(i)}, dt\}$ is integrable, as can be seen from the following equations:

$$\begin{aligned}
d\alpha &= (u_2)^2\alpha \wedge dt + ((u_2)^3 + w_2^1 - 1)\alpha^2 \wedge dt + \alpha^2 \wedge \omega^1 + (u_2 x_1 - x_3)\omega^1 \wedge dt \\
d\alpha^2 &= -u_2\alpha \wedge dt - (u_2)^2\alpha^2 \wedge dt - \alpha^3 \wedge dt - x_1\omega^1 \wedge dt \\
d\alpha^3 &= -u_2\alpha^2 \wedge dt - du_1 \wedge dt - x_2\omega^1 \wedge dt \\
d\alpha^4 &= -\omega^1 \wedge dt \\
d\omega^1 &= -\omega^2 \wedge dt \\
d\omega^2 &= -\omega^3 \wedge dt \\
d\omega^3 &= -dw_2^3 \wedge dt
\end{aligned}$$

and thus, by Theorem 62 the extended system is feedback linearizable.

As noted by Charlet, Lévine, and Marino [6], this system does not satisfy the conditions of Theorem 61. We have seen that the system is linearizable after a dynamic extension of order 3 on u_2. Following the notation of Theorem 61, we see that

$$\Delta_0 = \{g_1\} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\} \quad \Delta_1 = \{g_1, ad_f g_1\} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} \right\} \tag{199}$$

Checking condition 3 of Theorem 61, we see that for $i = 0$ and $j = 2$,

$$[g_2, g_1] = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \notin \Delta_1 \tag{200}$$

□

The relation between the standard Pfaffian system approach and the infinitesimal Brunovsky form is more involved. The following can be shown:

Proposition 72. *If the system is linearizable by static state feedback (equivalently the conditions of Theorems 46 and 47 hold for $dz^0 = dt$) then the two flag constructions are the same, modulo dt, i.e.:*

$$H_k = I^{(k-1)} \ mod\ dt$$

Proof. As both flag constructions are intrinsic we can assume, without loss of generality, that the system is already in the canonical coordinates of the Goursat normal form. Then:

$$\begin{aligned}
I^{(0)} &= \{dz_i^j - z_{i+1}^j dz^0 : i = 1, \ldots, s_j; j = 1, \ldots, m\} \\
H_1 &= \{dz_i^j : i = 1, \ldots s_j; j = 1, \ldots, m\}
\end{aligned}$$

Recall that, in the context of system (180) (equivalently (189)), z^0 plays the role of time (hence $dz^0 = dt$) and $z_{s_j+1}^j$ plays the role of u_j, $j = 1, \ldots, m$. Observe that the above co-distributions are identical if the terms in dz^0 are dropped from $I^{(0)}$.

The next iteration of the two flags yields:

$$\begin{aligned}
I^{(1)} &= \{\alpha \in I^{(0)} : d\alpha \equiv 0 \bmod I^{(0)}\} \\
&= \{dz_i^j - z_{i+1}^j dz^0 : i = 1, \dots, s_j - 1; j = 1, \dots, m\} \\
H_2 &= \{\omega \in H_1 : \dot{\omega} \in H_1\} \\
&= \{dz_i^j : i = 1, \dots, s_j - 1; j = 1, \dots, m\}
\end{aligned}$$

Note that $d(dz_{s_j}^j - z_{s_j+1}^j dz^0) = -dz_{s_j+1}^j \wedge dz^0$ which is not equal to 0 mod $I^{(0)}$. Similarly $d\dot{z}_{s_j}^j = dz_{s_j+1} = du_j \notin H_1$. Again the two constructions are the same if the terms in dz^0 are dropped from $I^{(1)}$.
In general, for the k^{th} step assume that:

$$\begin{aligned}
I^{(k-1)} &= \{dz_i^j - z_{i+1}^j dz^0 : i = 1, \dots, s_j - k + 1; j = 1, \dots, m\} \\
H_k &= \{dz_i^j : i = 1, \dots, s_j - k + 1; j = 1, \dots, m\}
\end{aligned}$$

Then:

$$\begin{aligned}
I^{(k)} &= \{\alpha \in I^{(k-1)} : d\alpha \equiv 0 \bmod I^{(k-1)}\} \\
&= \{dz_i^j - z_{i+1}^j dz^0 : i = 1, \dots, s_j - k; j = 1, \dots, m\} \\
H_{k+1} &= \{\omega \in H_k : \dot{\omega} \in H_k\} \\
&= \{dz_i^j : i = 1, \dots, s_j - k; j = 1, \dots, m\}
\end{aligned}$$

Note again that $d(dz_{s_j-k+1}^j - z_{s_j-k+2}^j dz^0) = -dz_{s_j-k+2}^j \wedge dz^0$ which will not be zero when wedged with all the one forms spanning $I^{(k-1)}$. Similarly, $d\dot{z}_{s_j-k+1}^j = dz_{s_j-k+2} \notin H_k$. Yet again the two co-distributions are identical if the terms in dz^0 are dropped from $I^{(k)}$. □

In view of Theorem 66 we make the following conjecture:

Conjecture 1. *The two flag constructions are related by:*

$$H_k = I^{(k-1)} \ \mathit{mod}\ dt$$

only if the system is linearizable by static state feedback.

Some examples illustrate this conjecture.

Example (Modified Ball and Beam). This example is inspired by the well-known system of a ball rolling on a beam [14]; the small-angle approximation has been used to eliminate the sine term which appears in the dynamic equations, and all the constants have been normalized to unity. This is a single input system which is not linearizable by static (and hence dynamic) state feedback. The simplified equations are:

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= x_1 x_4^2 - x_3 \\
\dot{x}_3 &= x_4 \\
\dot{x}_4 &= u
\end{aligned}$$

The flag associated with the infinitesimal Brunovsky form is:

$$\begin{aligned}
H_0 &= \{dx_1, dx_2, dx_3, dx_4, du\} \\
H_1 &= \{dx_1, dx_2, dx_3, dx_4\} \\
H_2 &= \{dx_1, dx_2, dx_3\} \\
H_3 &= \{dx_1, dx_2 - 2x_1x_4dx_3\} \\
H_4 &= \{(1 + 2x_2x_4 + 2x_1u)dx_1 + 2x_1x_4(dx_2 - 2x_1x_4dx_3)\} \\
H_5 &= \{0\}
\end{aligned}$$

Note that, if we let $\omega = (1+2x_2x_4+2x_1u)dx_1+2x_1x_4(dx_2-2x_1x_4dx_3)$, $d\omega\wedge\omega \not\equiv 0$ (as it will contain, among other things, the term $x_1^2x_4du \wedge dx_1 \wedge dx_2$), therefore $H_4 = \{\omega\}$ is not integrable. Hence, according to Theorem 66 and Theorem 70, the system will not be linearizable by any of the techniques considered here, as expected.

The derived flag construction for the same system leads to:

$$\begin{aligned}
I^{(0)} &= \{dx_1 - x_2dt, dx_2 - (x_1x_4^2 - x_3)dt, dx_3 - x_4dt, dx_4 - udt\} \\
I^{(1)} &= \{dx_1 - x_2dt, dx_2 - (x_1x_4^2 - x_3)dt, dx_3 - x_4dt\} \\
I^{(2)} &= \{dx_1 - x_2dt, dx_2 - 2x_1x_4dx_3 + (x_1x_4^2 + x_3)dt\} \\
I^{(3)} &= \{0\}
\end{aligned}$$

Note that the two flags are identical (neglecting the dt terms) until the fourth step where the dimension of the derived flag drops by two. According to Theorem 47, the system is not linearizable by static state feedback. □

Example (VTOL [15]**).** This example is inspired by the dynamic equations for a planar vertical take-off and landing aircraft; parasitic effects have been eliminated to simplify the analysis. This is an example of a two input system that is not linearizable by static state feedback, but is linearizable by dynamic extension. The dynamics of the system are given by the following equations:

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= -\sin x_5 u_1 \\
\dot{x}_3 &= x_4 \\
\dot{x}_4 &= \cos x_5 u_1 - 1 \\
\dot{x}_5 &= x_6 \\
\dot{x}_6 &= u_2
\end{aligned}$$

It is easy to show that the system is not feedback linearizable by static state feedback. However, if two integrators are added in front of input u_1 the resulting eight-state, two-input system is feedback linearizable.

The flag associated with the infinitesimal Brunovsky form of the original system is:

$$
\begin{aligned}
H_0 &= \{dx_1, dx_2, dx_3, dx_4, dx_5, dx_6, du_1, du_2\} \\
H_1 &= \{dx_1, dx_2, dx_3, dx_4, dx_5, dx_6\} \\
H_2 &= \{dx_1, dx_3, dx_5, \cos x_5 dx_2 + \sin x_5 dx_4\} \\
H_3 &= \{\cos x_5 dx_1 + \sin x_5 dx_3, \\
&\quad (\sin x_5)x_6 dx_1 - (\cos x_5)x_6 dx_3 + (\cos x_5 dx_2 + \sin x_5 dx_4)\} \\
H_4 &= \{0\}
\end{aligned}
$$

Letting $\omega_1 = \cos x_5 dx_1 + \sin x_5 dx_3$ and $\omega_2 = (\sin x_5)x_6 dx_1 - (\cos x_5)x_6 dx_3 + (\cos x_5 dx_2 + \sin x_5 dx_4)$ it is easy to show that $d\omega_1 \wedge \omega_1 \wedge \omega_2 \not\equiv 0$ (as it contains terms in $dx_1 \wedge dx_3 \wedge dx_4 \wedge dx_5$ among other things). Therefore, $H_3 = \operatorname{span}_{\mathcal{K}}\{\omega_1, \omega_2\}$ is not integrable and hence, according to Theorem 66 the system is not linearizable by static state feedback. Even though it is known that the system *is* linearizable by dynamic extension, there seems to be no easy way of determining the form of the invertible operator P of Theorem 67.

The derived flag for this system has the form:

$$
\begin{aligned}
I^{(0)} &= \{dx_1 - x_2 dt, dx_2 + \sin x_5 u_1 dt, dx_3 - x_4 dt, dx_4 - (\cos x_5 u_1 - 1)dt, \\
&\quad dx_5 - x_6 dt, dx_6 - u_2 dt\} \\
I^{(1)} &= \{dx_1 - x_2 dt, dx_3 - x_4 dt, dx_5 - x_6 dt, \cos x_5 dx_2 + \sin x_5 dx_4 + \sin x_5 dt\}
\end{aligned}
$$

The calculation involved in the next step of the derived flag are rather complicated. However, the pair of one forms we would expect to find in $I^{(2)}$ because of the structure of H_3, namely $\cos x_5(dx_1 - x_2 dt) + \sin x_5(dx_3 - x_4 dt)$ and $\sin x_5 x_6(dx_1 - x_2 dt) - \cos x_5 x_6(dx_3 - x_4 dt) + (\cos x_5 dx_2 + \sin x_5 dx_4 + \sin x_5 dt)$, do not satisfy the necessary conditions. Therefore the two flags diverge at this point. An interesting observation is that, if we define outputs $y_1 = x_1$ and $y_2 = x_3$ (the position of the plane), and attempt to input-output linearize the system this is exactly the step where input u_1 shows up (without u_2) and we can conclude that the linearization will fail. It would be interesting to try to relate this observation with the maximal linearizable subsystem [19] and add more substance to this observation. □

To summarize, in this section we have approached the problem of feedback linearization of control systems using techniques such as vector fields, exterior differential systems, and the infinitesimal Brunovsky form. All give equivalent conditions for the static feedback linearization problem. A sufficient condition for linearization using dynamic extension was given using vector fields; a necessary and sufficient one was given using exterior differential systems. Of the three techniques, only the infinitesimal Brunovsky form is formulated to approach the more general problem of dynamic feedback linearization, as stated in Problem 3; however, the conditions do not appear to be easy to verify.

5. Concluding Remarks & Topics of Interest

Exterior differential systems offer a different perspective on systems of differential equations. This approach is more algebraic than the standard vector field approach which is very geometric. The main advantage of looking at systems using differential forms instead of tangent vectors is precisely this algebraic power afforded by exterior systems. In this paper, we presented a number of different approaches which can be used to linearize a nonlinear system by state feedback and coordinate transformation. It was shown that they all produce comparable results in most cases, even though some are better suited to tackle certain problems than others. These techniques represent significant progress for all the problems posed here. Much work remains to be done. For example, an extension of the vector field conditions for converting the system to extended Goursat normal form may be very useful and may provide insight into many hard problems in the area of exterior differential systems.

Another direction that deserves further attention is linearization by dynamic state feedback. It should be noted that, even for the case of dynamic extension, none of the results are constructive. In particular, all the theorem statements start with an assumption of the form "if there exist ..." ("integers such that ..." or "invertible operator ..."), but provide no insight on whether those integers or operators exist or how to determine them. Only recently have upper bounds been determined on the number of dynamic extension steps required to feedback linearize a system [27]. No bound is known on the necessary dimension of a general dynamic state feedback. All of these are properties depend only on the system dynamics, it should therefore be possible to answer the above questions given only the system equations. It should also be noted that most of the literature is concerned with dynamic extension and non-singular input transformations. Of the theorems presented here only Theorem 67 claims to address the general dynamic state feedback case. Singular input transformations are briefly discussed in [32] and compared with the corresponding results using the extended Goursat normal form and prolongations. Both of these topics merit further attention.

Finally it should be noted that the conditions for feedback linearization are "closed", i.e. they essentially hold for a set of "measure zero" in the "space" of dynamical systems. It is therefore useful to know what, if anything, can be done about systems which do not satisfy these conditions, as most systems encountered in practice fall into this category. This problem was first addressed in [14] and later, more formally, in [28]. A different approach, related more to input-output linearization is taken in [18] and [6]. It would be interesting to compare the two approaches, and hopefully determine classes of systems that are better suited for one or the other.

References

1. R. Abraham, J. E. Marsden, and T. Ratiu, *Manifolds, tensor analysis, and applications*, second ed., Springer-Verlag, 1983.
2. E. Aranda-Bricaire, C. H. Moog, and J-B. Pomet, *A linear algebraic framework for dynamic feedback linearization*, IEEE Transactions on Automatic Control **40** (1995), no. 1, 127–132.
3. R. L. Bryant, S. S. Chern, R. B. Gardner, H. L. Goldschmidt, and P. A. Griffiths, *Exterior differential systems*, Springer-Verlag, 1991.
4. L. Bushnell, D. Tilbury, and S. S. Sastry, *Extended Goursat normal forms with applications to nonholonomic motion planning*, Proceedings of the IEEE Conference on Decision and Control (San Antonio, Texas), 1993, pp. 3447–3452.
5. ———, *Steering three-input chained form nonholonomic systems using sinusoids: The fire truck example*, International Journal of Robotics Research **14** (1995), no. 4, 366–381.
6. B. Charlet, J. Lévine, and R. Marino, *Sufficient conditions for dynamic state feedback linearization*, SIAM Journal of Control and Optimization **29** (1991), no. 1, 38–57.

7. R. Chatila, *Mobile robot navigation: Space modeling and decisional processes*, Robotics Research: The Third International Symposium (O. Faugeras and G. Giralt, eds.), MIT Press, 1986, pp. 373–378.
8. M. Fliess, J. Lévine, P. Martin, and P. Rouchon, *Sur les systèmes nonlinéaires diffŕentiellement plats*, C. R. Acad. Sci. Paris **315** (1992), 619–624.
9. ———, *Towards a new differential geometric setting in nonlinear control*, Proc. Internat. Geometrical Coll. (Moscow), 1993.
10. ———, *Flatness and defect of nonlinear systems: Introductory theory and examples*, International Journal of Control **61** (1995), no. 6, 1327–1361.
11. R. B. Gardner and W. F. Shadwick, *The GS algorithm for exact linearization to Brunovsky normal form*, IEEE Transactions on Automatic Control **37** (1992), no. 2, 224–230.
12. G. Giralt, R. Chatila, and M. Vaisset, *An integrated navigation and motion control system for autonomous multisensory mobile robots*, Robotics Research : The First International Symposium (Cambridge, Massachusetts) (M. Brady and R. Paul, eds.), MIT Press, Cambridge, Massachusetts, 1984, pp. 191–214.
13. Phillip Griffiths, Exterior differential systems and the calculus of variations, Birkhauser, 1982.
14. J. Hauser, S. Sastry, and P. Kokotović, *Nonlinear control via approximate input-output linearization, the ball and beam example*, IEEE Transactions on Automatic Control **37** (1992), no. 3, 392–398.
15. J. E. Hauser, *Approximate tracking for nonlinear systems with application to flight control*, Ph.D. thesis, Department of Electrical Engineering, University of California, Berkeley, California, 1989.
16. T. W. Hungerford, *Algebra*, Springer-Verlag, 1974.
17. A. Isidori, *Nonlinear control systems*, second ed., Springer-Verlag, 1989.
18. A. J. Krener, A. Isidori, and W. Respondek, *Partial and robust linearization by feedback*, Proceedings of the IEEE Conference on Decision and Control, 1983, pp. 126–130.
19. R. Marino, *On the largest feedback linearizable subsystem*, Systems and Control Letters **6** (1986), no. 5, 345–351.
20. J. R. Munkres, *Analysis on manifolds*, Addison-Wesley, 1991.
21. R. M. Murray, *Applications and extensions of Goursat normal form to control of nonlinear systems*, Proceedings of the IEEE Conference on Decision and Control, 1993, pp. 3425–3430.
22. ———, *Nilpotent bases for a class of non-integrable distributions with applications to trajectory generation for nonholonomic systems*, Mathematics of Control, Signals, and Systems: MCSS **7** (1995), no. 1, 58–75.
23. H. Nijmeijer and A. J. van der Schaft, *Nonlinear dynamical control systems*, Springer-Verlag, 1990.
24. C. Samson, *Velocity and torque feedback control of a nonholonomic cart*, International Workshop in Adaptive and Nonlinear Control: Issues in Robotics, 1990, pp. 125–151.
25. W. F. Shadwick and W. M. Sluis, *Dynamic feedback for classical geometries*, Tech. Report FI93-CT23, The Fields Institute, Ontario, Canada, 1993.
26. W. M. Sluis, *Absolute equivalence and its applications to control theory*, Ph.D. thesis, University of Waterloo, 1992.
27. W. M. Sluis and D. M. Tilbury, *A bound on the number of integrators needed to linearize a control system*, Proceedings of the IEEE Conference on Decision and Control (New Orleans), 1995, pp. 602–607.
28. W.M. Sluis, A. Banaszuk, J. Hauser, and R.M. Murray, *A homotopy algorithm for approximating geometric distributions by integrable systems*, Systems and Control Letters **27** (1996), no. 5, 285–291.
29. O. J. Sørdalen, *Conversion of the kinematics of a car with N trailers into a chained form*, Proceedings of the IEEE International Conference on Robotics and Automation (Atlanta, Georgia), 1993, pp. 382–387.
30. M. Spivak, *Calculus on manifolds*, Addison-Wesley, 1965.
31. ———, *A comprehensive introduction to differential geometry*, second ed., vol. One, Publish or Perish, Inc., Houston, 1979.
32. D. Tilbury, *Exterior differential systems and nonholonomic motion planning*, Ph.D. thesis, University of California, Berkeley, 1994, Also available as UCB/ERL Memo, number M94/90.

33. D. Tilbury and A. Chelouah, *Steering a three-input nonholonomic system using multirate controls*, Proceedings of the European Control Conference (Groningen, The Netherlands), 1993, pp. 1428–1431.
34. D. Tilbury, R. Murray, and S. Sastry, *Trajectory generation for the N-trailer problem using Goursat normal form*, IEEE Transactions on Automatic Control **40** (1995), no. 5, 802–819.
35. D. Tilbury and S. Sastry, *The multi-steering n-trailer system: A case study of Goursat normal forms and prolongations*, International Journal of Robust and Nonlinear Control **5** (1995), no. 4, 343–364.
36. D. Tilbury, O. Sørdalen, L. Bushnell, and S. Sastry, *A multi-steering trailer system: Conversion into chained form using dynamic feedback*, IEEE Transactions on Robotics and Automation **11** (1995), no. 6, 807–818.

Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720
E-mail address: gpappas@eecs.berkeley.edu

Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720
E-mail address: lygeros@eecs.berkeley.edu

Department of Mechanical Engineering and Applied Mechanics, University of Michigan, Ann Arbor, MI 48109-2125
E-mail address: tilbury@umich.edu

Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720
E-mail address: sastry@eecs.berkeley.edu