# Geometry and classification of contact systems: applications to control of nonholomic mechanical systems

Shunjie Li

Institut National des Sciences Appliquées de Rouen

# Mémoire de thèse

pour obtenir le titre de

## Docteur de L'Insa de Rouen

présenté par

**LI SHUNJIE**

Discipline: Mathématiques appliquées
Spécialité: Théorie du contrôle

# GÉOMÉTRIE ET CLASSIFICATION DES SYSTÈMES DE CONTACT: APPLICATIONS AU CONTRÔLE DES SYSTÈMES MÉCANIQUES NON HOLONOMES

Soutenue le 16 février 2010

Composition du Jury:

Président:

| | |
|---|---|
| **Fernand PELLETIER** | Professeur à l'Université de Savoie |

Rapporteurs:

| | |
|---|---|
| **Frédéric JEAN** | Enseignant-chercheur à l'ENSTA |
| **Pierre ROUCHON** | Professeur à l'Ecole des Mines de Paris |

Examinateurs:

| | |
|---|---|
| **Claude MOOG** | Directeur de recherche au CNRS (IRCCyN) |
| **Jean-Baptiste POMET** | Chargé de recherche à l'INRIA (Sophia Antipolis) |

Directeur de thèse:

| | |
|---|---|
| **Witold RESPONDEK** | Professeur à l'INSA de Rouen |

# Remerciements

Je voudrais remercier d'abord de tout mon cœur mon responsable de thèse Witold Respondek qui est à origine de ce travail. Il a dirigé mon mémoire de Master 2 avec gentillesse et ensuite accepté de me prendre en thèse. C'est un honneur pour moi de travailler avec lui et je ne peux qu'admirer son talent. Il a dirigé ma thèse avec beaucoup de patience et il a consacré beaucoup de temps à mon travail en étant toujours très disponible inconditionnellement à toutes les heures du jour et de la nuit, ce qui m'a énormément encouragé. Je le remercie d'avoir lu et corrigé très sérieusement beaucoup de versions préliminaires de mes travaux et ses redactions exceptionnelles ont donné plus de clarté aux résultats de ce mémoire. Je le remercie aussi, ainsi que son épouse Eva, de m'avoir beaucoup aidé durant ma présence à Rouen.

Je remercie Frédéric Jean et Pierre Rouchon, qui m'ont fait l'honneur d'accepter de rapporter ma thèse et de faire partie du jury. Je remercie également Claude Moog, Fernand Pelletier et Jean-Baptiste Pomet pour l'honneur qu'ils ont accepté d'examiner et juger favorablement ma thèse.

J'adresse mes remerciements à tous ceux qui m'ont entouré, permanents ou thésards du LMI pour les bons moments partagés ensemble. En particulier, je tiens à saluer Sandra Tamion, Arnaud Knippel, Khaled et Lamia qui m'ont donné beaucoup d'aide pendant ces années de thèse.

Je remercie finalement ma femme dont le soutien a été très important pendant mes études de la thèse en France.

# Contents

# Résumé


Les problèmes suivants ont été étudiés et résolus dans nos travaux :

1. **Caractérisation des sorties plates des systèmes avec deux contrôles** : D'abord, nous avons trouvé la relation entre la platitude d'un système avec deux contrôles et la structure de Goursat. Ensuite, nous avons caractérisé complètement toutes les $x$-sorties plates des systèmes avec deux contrôles qui sont équivalent au système chaînée (système de contact sur $J^n(\mathbb{R}^1, \mathbb{R}^1)$) et aussi décrit leur lieux singuliers. Enfin, nous avons appliqué nos résultats au système du robot mobile avec des remorques.

2. **Le système à $n$-barres dans** $\mathbb{R}^{m+1}$ : Un nouveau modèle cinématique pour le système à $n$-barres dans $\mathbb{R}^{m+1}$ a été présenté, qui généralise le système du robot mobile avec des remorques dans $\mathbb{R}^2$. En utilisant ce modèle, nous avons donné une caractérisation complète du système à $n$-barres dans $\mathbb{R}^{m+1}$ et aussi de ses lieux singuliers. Ensuite, la propriété de la platitude de ce système a été analysée et ses sorties plates ont été déterminées.

3. **Caractérisation de la distribution de Cartan** $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$ : Nous avons donné des conditions nécessaires et suffisantes vérifiables pour qu'une distribution soit équivalente à la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$. Nous avons aussi répondu à la question : quand une distribution $\mathcal{D}$ de rang $k + mk$ ($m \geq 3$), sur une variété de dimension $m + k + mk$, contient une sous-distribution involutive de corang $k$ dans $\mathcal{D}$.

4. **Linéarisation par bouclage orbital pour un système avec multi-contrôles défini dans** $\mathbb{R}^{(n+1)m+1}$ : Nous avons étudié le système de contrôle sous la forme $\Sigma : \dot{x} = f(x) + \sum_{i=1}^{m} g_i(x)u_i$, où $m \geq 2$, $x \in X = \mathbb{R}^{(n+1)m+1}$, et $f(x_0) \neq 0$ pour $x_0 \in X$. Nous avons obtenu des conditions nécessaires et suffisantes pour que ce système soit, localement, linéarisable par bouclage orbital. Toutes ces conditions peuvent être vérifiées directement sur le système original et une construction de la fonction $\gamma$ qui décrit le changement de temps a été donnée.

# Introduction

Le contrôle des systèmes non linéaires constitue un domaine très actif de recherche en automatique et mathématique. Un système de contrôle non linéaire est un système d'équations non linéaires, décrivant l'évolution temporelle des variables du système sous l'action d'un nombre fini de variables indépendantes appelées contrôles de ce système. Un système de contrôle non linéaire peut être écrit généralement sous la forme :

$$\Xi : \quad \dot{x} = f(x, u), \tag{0.0.1}$$

avec $x \in X$, l'état du système dans une variété differentielle et $u$ est le contrôle à valeur dans $U$, une variété de dimension $m$ appelée l'espace du contôle. Dans cette thèse, tous les systèmes que nous étudions sont des *systèmes affines*, i.e., des systèmes qui admettent la forme suivante

$$\Sigma_{\text{aff}} : \quad \dot{x} = f(x) + \sum_{i=1}^{m} g_i(x) u_i = f(x) + g(x)u,$$

ou des systèmes linéaires par rapport aux contrôles ($f \equiv 0$) de la forme suivante

$$\Sigma_{\text{lin}} : \quad \dot{x} = \sum_{i=1}^{m} g_i(x) u_i = g(x)u.$$

## Equivalence des systèmes par bouclage

Soit $\tilde{\Sigma}_{\text{aff}}$ un autre système de contrôle défini par

$$\tilde{\Sigma}_{\text{aff}} : \quad \dot{\tilde{x}} = \tilde{f}(\tilde{x}) + \sum_{i=1}^{m} \tilde{g}_i(\tilde{x}) \tilde{u}_i = \tilde{f}(\tilde{x}) + \tilde{g}(\tilde{x})\tilde{u},$$

où $\tilde{x} \in \tilde{X}$ et $\tilde{u} \in \tilde{U}$. Alors on dit que $\Sigma_{\text{aff}}$ et $\tilde{\Sigma}_{\text{aff}}$ sont équivalents par bouclage s'il existe un difféomorphisme $\varphi : X \to \tilde{X}$ et un bouclage $u = \alpha(x) + \beta(x)\tilde{u}$, où $\beta(x)$ est une matrice $m \times m$ inversible et $\tilde{u} = (\tilde{u}_1, \ldots, \tilde{u}_m)^\top$, tels que

$$\tilde{f} = \varphi_*(f + g\alpha) \quad \text{et} \quad \tilde{g} = \varphi_*(g\beta).$$

Pour les systèmes $\Sigma_{\text{lin}}$ la relation entre $\tilde{f}$ et $f$ est évidemment absente. La distribution associée au système $\Sigma_{\text{aff}}$, notée par $\mathcal{D}$, est la distribution engendrée par les champs de vecteurs $f, g_1, \ldots, g_m$, i.e.,

$$\mathcal{D} = \operatorname{span}\{f, g_1, \ldots, g_m\} \subset TX.$$

Il est important de remarquer que pour les systèmes $\Sigma_{\text{lin}}$, l'équivalence locale par bouclage coïncide avec l'équivalence des distributions $\mathcal{D}$ engendrée par les champs de vecteurs $g_1, \ldots, g_m$.

La linéarisation par bouclage est un outil très important dans la recherche des systèmes non linéaires dans le fait qu'on peut ainsi appliquer des propriétés d'un système linéaire à un système non linéaire. Le problème de la linéarisation par bouclage d'un système avec un contrôle a été résolu par Brockett [5]. Jakubczyk et Respondek [29], Hunt et Su [22] ont donné des conditions nécessaires et suffisantes de la linéarisation par bouclage pour un système affine avec multi-contrôles.

## Système de contact canonique

Le but de cette thèse est d'étudier la géométrie et la structure des systèmes de contrôle non linéaires, en particulier les systèmes équivalents aux systèmes de contact canonique qui est une classe de systèmes non holonomes. Un système non holonome est un système soumis aux contraintes non holonomes qui ne sont pas intégrables. On rencontre ce type de contraintes par exemple dans les phénomènes de roulement sans glissement apparaissant pour les systèmes du type du robot mobile.

Considérons $J^n(\mathbb{R}^k, \mathbb{R}^m)$, l'espace de $n$-jets des applications de classe $C^\infty$ de $\mathbb{R}^k$ dans $\mathbb{R}^m$ avec les coordonnées canoniques données par

$$(x_1, \ldots, x_k, y_1, \ldots, y_m, p_j^\sigma, \quad 1 \leq j \leq m, \quad 1 \leq |\sigma| \leq n),$$

où $x_i$, pour $1 \leq i \leq k$, représentent les variables indépendantes et $y_j$, pour $1 \leq j \leq m$, représentent les variables dépendantes; $\sigma = (\sigma_1, \ldots, \sigma_k)$ est un vecteur de nombres entiers non-négatifs tel que $|\sigma| = \sigma_1 + \cdots + \sigma_k \leq n$ et $p_j^\sigma$, pour $1 \leq j \leq m$, correspondent aux derivées partielles $\dfrac{\partial^{|\sigma|} y_j}{\partial x_\sigma}$. Notons $p_j^0 = y_j$, pour $1 \leq j \leq m$. Chaque application lisse $\varphi = (\varphi_1, \ldots, \varphi_m)$ de $\mathbb{R}^k$ à $\mathbb{R}^m$ définit une sous-variété dans $J^n(\mathbb{R}^k, \mathbb{R}^m)$ par

$$p_j^\sigma = \frac{\partial^{|\sigma|} \varphi_j}{\partial x_\sigma}(x_1, \ldots, x_k),$$

pour $1 \leq j \leq m$, $0 \leq |\sigma| \leq n$ et on l'appelle le $n$-graphe de $\varphi$. Tous les $n$-graphes sont sous-variétés intégrales, de dimension $k$, d'une distribution appelée la *distribution de Cartan* sur $J^n(\mathbb{R}^k, \mathbb{R}^m)$, notée par $\mathcal{CC}^n(\mathbb{R}^k, \mathbb{R}^m)$, qui est annulée par les formes différentielles suivantes:

$$\mathrm{d}p_j^\sigma - \sum_{i=1}^{k} p_j^{\sigma+1_i} dx_i = 0, \quad 0 \leq j \leq m,\ 1 \leq |\sigma| \leq n-1,$$

où $\sigma + 1_i = (\sigma_1, \ldots, \sigma_i + 1, \ldots, \sigma_k)$. Il est clair que la distribution de Cartan est de rang constant. Nous appelons un système linéaire par rapport aux contrôles un *système de contact canonique* sur $J^n(\mathbb{R}^k, \mathbb{R}^m)$ telle que sa distribution associée est engendrée par tous les champs de vecteurs de la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^k, \mathbb{R}^m)$. Les applications des systèmes de contact canonique ou des distributions de Cartan sont de ce fait extrêmement nombreuses dans des domaines très différents tels que la mécanique, l'automatique, la thermodynamique, et l'électromagnétisme, etc. Par exemple, quand $k = m = 1$, le système de contact canonique sur $J^n(\mathbb{R}^1, \mathbb{R}^1)$ peut s'écrire sous la forme du système de contrôle suivant

$$\begin{array}{rcl}
\dot{x}_0^0 & = & u_0 \\
\dot{x}_1^0 & = & x_1^1 u_0 \\
\dot{x}_1^1 & = & x_1^2 u_0 \\
 & \vdots & \\
\dot{x}_1^{n-1} & = & x_1^n u_0 \\
\dot{x}_1^n & = & u_1.
\end{array}$$

On obtient un système sous la forme chaînée - un système non holonome connu qui est extrêmement important en automatique et a été étudié par de nombreux scientifiques tels que Murray et Sastry [46], Samson [65], Sørdalen [73], etc. La distribution associée au système de contact canonique sur $J^n(\mathbb{R}^1, \mathbb{R}^1)$ (la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^1, \mathbb{R}^1)$), est appelé aussi la forme normale de Goursat.

Nous étudions les systèmes de contact canoniques dans des cas différents afin de comprendre les différentes notions d'équivalence (par bouclage statique, par bouclage dynamique, par bouclage orbital, etc.) correspondant aux objets géometriques attachés aux systèmes de contrôles tels que les distributions, les distributions affines, les distributions caractéristiques, etc.

## Platitude

La notion de la platitude différentielle a été introduite dans la théorie du contrôle par Fliess, Lévine, Martin et Rouchon [12], [13], [14] (voir aussi [2], [25], [26], [56], [70], [77]). Un système plat est un système pour lequel il est possible de trouver un système linéaire équivalent par bouclage endogène.

Considérons le système $\Xi$, défini par (0.0.1), nous lui associons $l$-prolongation $\Xi^l$ donnée par

$$\Xi^l : \begin{array}{rcl} \dot{x} & = & f(x, u^0) \\ \dot{u}^0 & = & u^1 \\ & \vdots & \\ \dot{u}^l & = & u^{l+1}, \end{array}$$

pour un certain entier $l$, qui peut être considéré comme un système de contrôle sur $X \times U \times \mathbb{R}^{ml}$. Ses variables d'état sont $(x, u^0, u^1, \ldots, u^l)$ et ses $m$ contrôles sont les $m$ composantes de $u^{l+1}$. Notons $\bar{u}^l = (u^0, u^1, \ldots, u^l)$.

**Definition 0.0.1** Le système $\Xi$, défini par (0.0.1), est plat en $(x_0, \bar{u}_0^l) \in X \times U \times \mathbb{R}^{ml}$, où $l$ est un nombre entier non négatif, s'il existe un voisinage $\mathcal{O}^l$ de $(x_0, \bar{u}_0^l)$ et $m$ fonctions lisses, définies dans $\mathcal{O}^l$,

$$h_i = h_i(x, u, \dot{u}, \ldots, u^{(l)}), \quad 1 \leq i \leq m,$$

et s'il existe un entier $s$, des fonctions lisses $\gamma_i$, pour $1 \leq i \leq n$, et des fonctions lisses $\delta_i$, pour $1 \leq i \leq m$, tels que

$$\begin{array}{rcl} x_i & = & \gamma_i(h, \dot{h}, \ldots, h^{(s)}), \quad 1 \leq i \leq n, \\ u_i & = & \delta_i(h, \dot{h}, \ldots, h^{(s)}), \quad 1 \leq i \leq m, \end{array}$$

où $h = (h_1, \ldots, h_m)^\top$ le long de chaque trajectoire $x(t)$ définie par un contrôle $u(t)$ tels que $(x(t), u(t), \dot{u}(t), \ldots, u^{(l)}(t)) \in \mathcal{O}^l$. Le vecteur $h = (h_1, \ldots, h_m)$ est appelé *sortie plate* de ce système. Si $h_i = h_i(x, u^0, u^1, \ldots, u^r)$, $r \leq l$, le système est appelé $(x, u, \ldots, u^r)$-plat et, en particulier, $x$-plat si $h_i = h_i(x)$.

D'après la définition, le comportement du système plat est complètement caractérisé par la sortie plate et ses dérivées successives sous forme de relations algébriques. Ainsi le problème de la planification de trajectoire pour un système plat peut être résolu facilement grâce à la propriété de platitude. Il est connu qu'un système plat est toujours localement contrôlable. Notons qu'il est facile de montrer que le système de contact sur $J^n(\mathbb{R}^1, \mathbb{R}^1)$, i.e., le système chaîné, est plat avec une sortie plate $h = (h_1, h_2) = (x_0^0, x_1^0)$ en tous les points tels que $u_0 \neq 0$.

Rappelons les définitions suivantes :

Le drapeau des systèmes dérivés d'une distribution $\mathcal{D}$ est défini par la suite de distributions

$$\mathcal{D}^{(0)} = \mathcal{D} \quad \text{et} \quad \mathcal{D}^{(i+1)} = \mathcal{D}^{(i)} + [\mathcal{D}^{(i)}, \mathcal{D}^{(i)}], \quad \text{pour} \quad i \geq 0.$$

Le drapeau de Lie est défini par une autre suite de distributions

$$\mathcal{D}_0 = \mathcal{D} \quad \text{et} \quad \mathcal{D}_{i+1} = \mathcal{D}_i + [\mathcal{D}_0, \mathcal{D}_i], \quad \text{pour} \quad i \geq 0.$$

En parallèle, on peut aussi définir le système dérivé pour un système de Pfaff $\mathcal{I} = \operatorname{span}\{\omega_1, \dots, \omega_s\}$, engendré par $s$ formes différentielles de degré 1 qui sont indépendantes partout. Son système dérivé est défini par

$$\mathcal{I}^{(1)} = \operatorname{span}\{\omega \in \mathcal{I} : \ \mathrm{d}\omega \wedge \omega_1 \wedge \cdots \wedge \omega_s = 0\}.$$

Le drapeau des systèmes dérivés de $\mathcal{I}$ est la suite de systèmes de Pfaff

$$\mathcal{I}^{(0)} \supset \mathcal{I}^{(1)} \supset \cdots \supset \mathcal{I}^{(i)} \supset \cdots,$$

qui satisfait les relations $\mathcal{I}^{(0)} = \mathcal{I}$ et $\mathcal{I}^{(i+1)} = (\mathcal{I}^{(i)})^{(1)}$. Si toutes les distributions du système dérivé sont de rang constant et $\mathcal{D} = \mathcal{I}^{\perp}$, alors on a $\mathcal{D}^{(i)} = (\mathcal{I}^{(i)})^{\perp}$, for $i \geq 0$.

Nous allons maintenant présenter chapitre par chapitre les résultats contenus dans cette thèse.

**Chapitre** 1. Ce chapitre est dédié à la caractérisation complète de toutes les sorties plates pour un système avec deux contrôles qui est équivalent au système chaîné (le système de contact canonique sur $J^n(\mathbb{R}^1, \mathbb{R}^1)$). Ce problème vient d'un phénomène que nous avons observé pour le système de la voiture (robot mobile avec une remorque)

$$\Sigma_{\text{car}} : \quad \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta}_0 \\ \dot{\theta}_1 \end{pmatrix} = \begin{pmatrix} \cos(\theta_1 - \theta_0)\cos\theta_0 \\ \cos(\theta_1 - \theta_0)\sin\theta_0 \\ \sin(\theta_1 - \theta_0) \\ 0 \end{pmatrix} u_1 + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} u_2,$$

où $(x, y)$ sont les coordonnées caractésiennes du milieu de l'axe des roues arrières, et notons par $\theta_0$ et $\theta_1$, respectivement, les orientations de l'axe principal de la voiture et les roues avant par rapport à $x$-axe. Rappelons que la platitude du système du robot mobile avec remorques roulant sans glissement a été montrée par Fliess, Lévine, Martin et Rouchon [13] (voir aussi [26]) et donc le système $\Sigma_{\text{car}}$ est $x$-plat. En effet, en chaque point de l'espace de configuration $\mathbb{R}^2 \times S^1 \times S^1$, ce système peut être transformé au système chaîné. En conséquence, autour de chaque point dans $\mathbb{R}^2 \times S^1 \times S^1$, il existe une $x$-sortie plate. Par contre, nous observons que *chaque $x$-sortie plate possède un lieu singulier où il faut la remplacer par une autre $x$-sortie plate.* Plus précisément, pour le système $\Sigma_{\text{car}}$, le couple $(h_1, h_2) = (x, y)$ est une $x$-sortie plate pour les points où $\{q = (x, y, \theta_1, \theta_2) \mid \theta_1 - \theta_0 \neq \pm\frac{\pi}{2}\}$ et donc $\theta_1 - \theta_0 = \pm\frac{\pi}{2}$ détermine le lieu singulier de la $x$-sortie plate $(x, y)$. Cependant, l'équation $\theta_1 - \theta_0 = \pm\frac{\pi}{2}$ ne définit pas les points de singularité pour le système $\Sigma_{\text{car}}$. En effet, en ces points où $\theta_1 - \theta_0 = \pm\frac{\pi}{2}$, ce système est encore $x$-plat et la $x$-sortie plate peut être donnée par le couple $(\bar{h}_1, \bar{h}_2) = (\theta_0, x\sin\theta_0 - y\cos\theta_0)$. Notons que cette deuxième $x$-sortie plate possède aussi un lieu singulier défini par $\{q = (x, y, \theta_1, \theta_2) | \ \theta_1 - \theta_0 = 0\,, \pi\}$.

Toutes ces dernières analyses mènent naturellement aux questions suivantes : Combien de $x$-sorties plates possède le système? Peut-on les caractériser complètement?

Plus généralement, peut-on donner une caractérisation de toutes les $x$-sorties plates pour un système avec deux contrôles qui est équivalent au système chaîné (système de contact canonique sur $J^n(\mathbb{R}^1, \mathbb{R}^1)$)?

Le problème de la platitude du système avec deux contrôles a été étudié et résolu par Martin et Rouchon [40]. Ils ont montré qu'un système avec deux contrôles est plat si et seulement si sa distribution associée $\mathcal{D}$ satisfait, dans un ouvert dense de $M$, la condition

$$\operatorname{rang} \mathcal{D}^{(i)} = i + 2, \quad 0 \leq i \leq n. \tag{0.0.2}$$

*Une structure de Goursat* est une distribution de rang deux qui vérifie la condition (0.0.2) en tout point $x \in M$ . Il est connu (voir von Weber [78], Cartan [9], Goursat [18]) que sur un ouvert dense de $M$, la condition (0.0.2) implique que $\mathcal{D}$ peut être transformée à la forme normal de Goursat (autrement dit, le système associé est localement équivalent au système chaîné). Kumpera, Ruiz et Giaro [17] ont découvert l'existence des points singuliers pour le problème de transformer une distribution de rang deux à la forme normale de Goursat. Murray [45] a montré qu'un système avec deux contrôles est équivalent au système chaîné au point $x_0$, autrement dit, la distribution associée est équivalente à la forme de Goursat, si et seulement si sa distribution associée satisfait la condition (0.0.2) et aussi la condition de régularité:

$$\dim \mathcal{D}^{(i)}(x_0) = \dim \mathcal{D}_i(x_0), \quad 0 \leq i \leq n. \tag{0.0.3}$$

Alors, une question naturelle se pose : est-ce qu'un système dont la distribution associée $\mathcal{D}$ est une structure de Goursat est plat en les points où la condition de régularité (0.0.3) n'est pas satisfaite?

**Theorem 0.0.2** *Considérons un système $\Sigma$ avec deux contrôles,*

$$\Sigma: \quad \dot{x} = f_1(x)u_1 + f_2(x)u_2,$$

*où $x \in M$, une variété de dimension $n+2$, pour $n \geq 1$. Supposons que la distribution associée $\mathcal{D} = \operatorname{span}\{f_1, f_2\}$ est une structure de Goursat, i.e., $\operatorname{rang} \mathcal{D}^{(i)} = i + 2$, pour $0 \leq i \leq n$. Alors les conditions suivantes sont équivalentes au point $x_0 \in M$*

(i) *$\Sigma$ est $x$-plat en $(x_0, \bar{u}_0^l) \in M \times \mathbb{R}^{2(l+1)}$, pour certain $l \geq 0$;*
(ii) *$\Sigma$ est $x$-plat en $(x_0, u_0) \in M \times \mathbb{R}^2$;*
(iii) *$\dim \mathcal{D}^{(i)}(x_0) = \dim \mathcal{D}_i(x_0)$, pour $0 \leq i \leq n$ et $u_0 \notin U_{\text{sing}}(x_0)$;*
(iv) *$\Sigma$ est localement, autour du point $x_0$, équivalent au système chaîné et $u_0 \notin U_{\text{sing}}(x_0)$.*

**Remarque**. Théorème 0.0.2 montre qu'un système tel que la distribution associée est une structure de Goursat est $x$-plat seulement aux points où la condition de régularité est satisfaite, i.e., $\dim \mathcal{D}^{(i)}(x_0) = \dim \mathcal{D}_i(x_0)$, pour $0 \leq i \leq n$ et $u_0 \notin U_{\text{sing}}(x_0)$ (voir chapitre 3 pour la définition de $U_{\text{sing}}$). Donc notre résultat donne une réponse négative pour la question posée par Martin et Rouchon [39] (voir aussi [40]) (concernant la $x$-platitude). Toute structure de Goursat peut être transformé à la forme normale de Kumpera-Ruiz ([33], [42], [52]). Alors notre résultat montre aussi que aucune forme normale de Kumpera-Ruiz singulière n'est $x$-plate.

Théorème 0.0.2 montre que les seules structures de Goursat $x$-plates sont celles qui satisfaitent la condition de régularitée (le système associé est équivalent au système chaîné). Alors nous allons caractériser toutes les $x$-sorties plates pour ce type de système avec deux contrôles.

**Theorem 0.0.3 (Caractérisation des sorties plates)**(Première version)
*Considérons un système $\Sigma$ avec 2 contrôles, défini sur $M$, une variété de dimension $n+2$, tel que sa distribution associée $\mathcal{D}$ satisfait $\operatorname{rang} \mathcal{D}^{(i)} = \operatorname{rang} \mathcal{D}_i = i+2$, pour $0 \le i \le n$. Fixons un point $x_0 \in M$ et supposons que $g_1$ est un champ de vecteurs arbitraire dans $\mathcal{D}$ tel que $g_1(x_0) \notin \mathcal{C}_{n-1}(x_0)$, où $\mathcal{C}_{n-1}$ est la distribution caractéristique de $\mathcal{D}^{(n-1)}$, et $\varphi_1, \varphi_2$ sont deux fonctions lisses définies dans un voisinage ouvert $\mathcal{M}$ du point $x_0$. Alors $(\varphi_1, \varphi_2)$ est une $x$-sortie plate du système $\Sigma$ en $(x_0, u_0)$, où $u_0 \notin U_{\text{sing}}(x_0)$, si et seulement si les conditions suivantes sont satisfaites :*

(i) $\mathrm{d}\varphi_1(x_0) \wedge \mathrm{d}\varphi_2(x_0) \neq 0$, *i.e.*, $\mathrm{d}\varphi_1$ *et* $\mathrm{d}\varphi_2$ *sont indépendantes en* $x_0$;

(ii) $(L_{g_1}\varphi_1(x_0), L_{g_1}\varphi_2(x_0)) \neq (0,0)$;

(iii) $L_c\varphi_1 \equiv L_c\varphi_2 \equiv L_c(\frac{L_{g_1}\varphi_2}{L_{g_1}\varphi_1}) \equiv 0$, $\forall\, c \in \mathcal{C}_{n-1}$, *où les fonctions* $\varphi_1$ *et* $\varphi_2$ *sont permutées telles que* $L_{g_1}\varphi_1 \neq 0$. *C'est toujours possible grâce à la condition* (ii).

*De plus, si $(\varphi_1, \varphi_2)$ satisfait* (i) *partout dans $\mathcal{M}$ et est une $x$-sortie plate en $(x, u)$ pour chaque point $x \in \tilde{\mathcal{M}}$ et un certain $u = u(x)$, où $\tilde{\mathcal{M}}$ est un ouvert dense dans $\mathcal{M}$, alors*

$$Sing(\varphi_1, \varphi_2) = \{x \in \mathcal{M} : (L_g\varphi_1(x), L_g\varphi_2(x)) = (0,0)\}.$$

Le Théorème 0.0.3 nous permet de vérifier si une paire de fonctions donnée $(\varphi_1, \varphi_2)$ est une $x$-sortie plate pour un système linéaire par rapport aux contrôles avec deux contrôles. Le théorème suivant va répondre à la question : y-a-t-il beaucoup de paires $(\varphi_1, \varphi_2)$ qui satisfont les conditions du Théorème 0.0.3?

**Theorem 0.0.4** *Considérons un système $\Sigma$ avec 2 contrôles, défini sur $M$, une variété de dimension $n+2$, tel que sa distribution associée $\mathcal{D}$ satisfait $\operatorname{rang} \mathcal{D}^{(i)} = \operatorname{rang} \mathcal{D}_i = i+2$, pour $0 \le i \le n$. Fixons un point $x_0 \in M$ et soit $g_1$ un champ de vecteurs arbitraire dans $\mathcal{D}$ tel que $g_1(x_0) \notin \mathcal{C}_{n-1}(x_0)$. Alors, etant donné une fonction lisse $\varphi_1$ telle que $L_c\varphi_1 = 0$, $\forall\, c \in \mathcal{C}_{n-1}$, et $L_{g_1}\varphi_1(x_0) \neq 0$, il existe toujours une fonction $\varphi_2$ telle que $(\varphi_1, \varphi_2)$ forme une $x$-sortie plate de $\Sigma$ en $(x_0, u_0)$, où $u_0 \notin U_{\text{sing}}(x_0)$. De plus, si les deux paires $(\varphi_1, \varphi_2)$ et $(\varphi_1, \tilde{\varphi}_2)$ sont des $x$-sorties plates de $\Sigma$, alors on a*

$$\operatorname{span}\{\mathrm{d}\varphi_1(x), \mathrm{d}\varphi_2(x)\} = \operatorname{span}\{\mathrm{d}\varphi_1(x), \mathrm{d}\tilde{\varphi}_2(x)\},$$

*pour tous les points $x$ dans un voisinage du point $x_0$.*

L'importance du Théorème 0.0.3 est qu'il nous permet non seulement de vérifier si une paire de fonctions donnée forme une $x$-sortie plate, mais aussi, avec Théorème 0.0.4, d'exprimer explicitement un système d'équations aux dérivées partielles du $1^{\text{er}}$ ordre dans le but de calculer toutes les $x$-sorties plates pour un système $\Sigma$ avec deux contrôles (voir Section 1.3.2 pour les détails). En fin, nous appliquons nos résultats au système du disque roulant et du robot mobile avec remorques afin de montre comment calculer toutes les $x$-sorties plates en résolvant des équations aux dérivées partielles du $1^{\text{er}}$ ordre.

**Chapitre** 2. Ce chapitre est consacré à la modélisation, à la caractérisation, aux singularités et à la propriété de la platitude du système à $n$-barres dans l'espace $\mathbb{R}^{m+1}$, qui généralise le système du robot mobile avec $n$ remorques dans $\mathbb{R}^2$. Le système du type robot mobile avec contraintes de roulement sans glissement, comme un cas typique de système mécanique non holonome, est très important en automatique et a été étudié par de nombreux mathématiciens et automaticiens depuis les dernières vingtaine d'années (voir les livres [35] et [36]; les articles [13], [27], [26], [34], [46], [53], [73]). En 1991, Laumond [34] a présenté un modèle cinématique du robot mobile avec remorques et a montré la controllabilité pour ce modèle. En 1994, Tilbury, Murray et Sastry [75] ont donné une caractérisation complète du système du robot mobile avec $n$ remorques : autour des points réguliers, ce système peut être transformé en un système chaîné, autrement dit, il est localement équivalent au système de contact sur $J^n(\mathbb{R}^1, \mathbb{R}^1)$. Les points réguliers sont caractérisés par la condition $\dim \mathcal{D}^{(i)} = \dim \mathcal{D}_i$, pour $0 \leq i \leq n$, où $\mathcal{D}$ désigne la distribution associée à ce système. Respondek et Pasillas-Lépine [52] ont montré qu'en dehors des points réguliers, le système du robot mobile avec $n$ remorques peut être transformé à la forme normale de Kumpera et Ruiz. Une caractérisation complète du lieu singulier du système du robot mobile avec $n$-remorques a été donnée par Jean [27].

Notre objectif dans ce chapitre est de généraliser le système du robot mobile avec remorques dans l'espace $\mathbb{R}^{m+1}$ (le système à $n$-barres), pour $m \geq 1$, et de donner une caractérisation de ce système. Récemment, ce système, sa géométrie et, notamment, son lieu singulier ont été étudiés par Slayaman et Pelletier ([71], [72]). Notre résultat est basé sur une nouvelle modélisation cinématique efficace pour ce système. Considérons le système à $n$-barres dans $\mathbb{R}^{m+1}$ tel que:

(1) La fin de la barre précédente coïncide avec la source de la barre suivante;
(2) La longueur de chaque barre est de 1;
(3) Chaque barre bouge sans glisser latéralement et la vitesse instantanée du point $P_i$ est parallèle au vecteur $\overrightarrow{P_iP_{i+1}}$, pour $0 \leq i \leq n-1$.

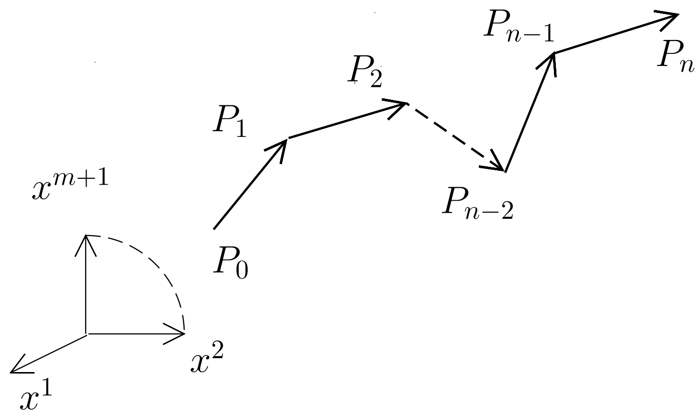


Figure 1: le système à $n$-barres dans $\mathbb{R}^{m+1}$

Supposons que les coordonnées des points $P_i$, pour $0 \leq i \leq n$, sont données par $P_i = (x_i^1, x_i^2, \ldots, x_i^{m+1})$. Alors l'espace de configuration de ce système peut être décrit complètement par les $(n+1)(m+1)$ coordonnées $(x_0^1, \ldots, x_0^{m+1}, x_1^1, \ldots, x_1^{m+1}, \ldots, x_n^1, \ldots, x_n^{m+1})$ dans $X = \mathbb{R}^{(n+1)(m+1)}$. Les hypothèses $|\overrightarrow{P_iP_{i+1}}| = 1$, pour $0 \leq i \leq n-1$, nous donnent les contraintes holonomes suivantes

$$\Psi(x) = 0,$$

où $\Psi = \{\Psi_1, \ldots, \Psi_n\} : X = \mathbb{R}^{(n+1)(m+1)} \to \mathbb{R}^n$ est définie par

$$\left\{ \begin{array}{lcl} \Psi_1(x) & = & (x_1^1 - x_0^1)^2 + (x_1^2 - x_0^2)^2 + \cdots + (x_1^{m+1} - x_0^{m+1})^2 - 1 \\ \Psi_2(x) & = & (x_2^1 - x_1^1)^2 + (x_2^2 - x_1^2)^2 + \cdots + (x_2^{m+1} - x_1^{m+1})^2 - 1 \\ & \vdots & \\ \Psi_n(x) & = & (x_n^1 - x_{n-1}^1)^2 + (x_n^2 - x_{n-1}^2)^2 + \cdots + (x_n^{m+1} - x_{n-1}^{m+1})^2 - 1. \end{array} \right. \quad (0.0.4)$$

D'après ces $n$ contraintes holonomes, le vrai espace de configuration du système à $n$-barres devient la sous-variété $Q = \mathbb{R}^{m+1} \times (S^m)^n \subset X$ définie par $Q = \{x \in X : \Psi(x) = 0\}$.

En effet, le modèl $\Gamma$ du système à $n$-barres

$$\Gamma : \quad \dot{q} = \sum_{i=0}^{m} f_i(q) u_i, \quad q \in \mathbb{R}^{m+1} \times (S^m)^n,$$

peut être défini par le système non holonome sans dérive $\Sigma$ dans l'espace $X = \mathbb{R}^{(n+1)(m+1)}$

$$\Delta : \quad \dot{x} = \sum_{i=1}^{n+m+1} g_i(x) v_i, \quad x \in X, \quad (0.0.5)$$

où $g_1, \ldots, g_{n+m+1}$ sont donnés par

$$
\begin{aligned}
g_1 &= (x_1^1 - x_0^1)\frac{\partial}{\partial x_0^1} + (x_1^2 - x_0^2)\frac{\partial}{\partial x_0^2} + \cdots + (x_1^{m+1} - x_0^{m+1})\frac{\partial}{\partial x_0^{m+1}} \\
g_2 &= (x_2^1 - x_1^1)\frac{\partial}{\partial x_1^1} + (x_2^2 - x_1^2)\frac{\partial}{\partial x_1^2} + \cdots + (x_2^{m+1} - x_1^{m+1})\frac{\partial}{\partial x_1^{m+1}} \\
&\vdots \\
g_n &= (x_n^1 - x_{n-1}^1)\frac{\partial}{\partial x_{n-1}^1} + (x_n^2 - x_{n-1}^2)\frac{\partial}{\partial x_{n-1}^2} + \cdots + (x_n^{m+1} - x_{n-1}^{m+1})\frac{\partial}{\partial x_{n-1}^{m+1}} \\
g_{n+i} &= \frac{\partial}{\partial x_n^i}, \quad 1 \leq i \leq m+1,
\end{aligned}
$$

sous les contraintes holonomes $\Psi(x) = (\Psi_1(x), \ldots, \Psi_n(x)) = 0$ définies par (0.0.4) (voir Section 2.2.2 pour plus de détails). Notons que nous n'avons pas donné l'expression explicite des champs de vecteurs $f_i$, $0 \leq i \leq m$, ni celle de la distribution associée $\mathcal{D}_\Gamma$. Toutes ses propriétés seront formulées et analysées en utilisant la distribution associée $\mathcal{E} = \text{span}\{g_1, \ldots, g_{n+m+1}\}$ du système $\Delta$ et les contraintes holonomes $\Psi(x) = 0$. Observons aussi que notre modèle n'utilise pas de variables d'angle et ce caractère nous permet de pouvoir étudier les objets géométriques attachés au système à $n$-barres tels que la distribution, la distribution caractéristique, le système de Pfaff, etc. Quand $m = 1$, le système à $n$-barres retrouve une modélisation du célèbre système du robot mobile avec $n-1$ remorques.

Un système de contrôle $\Sigma : \ \dot{x} = \sum_{i=0}^m u_i f_i(x)$, défini sur $\mathbb{R}^{(n+1)m+1}$, est dit sous la forme $m$-chaînée s'il est representé par

$$
\begin{array}{ccccccccccc}
\dot{x}_0^0 & = & u_0 & \dot{x}_1^0 & = & x_1^1 u_0 & \cdots & \dot{x}_m^0 & = & x_m^1 u_0 \\
 & & & & \cdots & & \cdots & & \cdots & \\
 & & & \dot{x}_1^{n-1} & = & x_1^n u_0 & \cdots & \dot{x}_m^{n-1} & = & x_m^n u_0 \\
 & & & \dot{x}_1^n & = & u_1 & \cdots & \dot{x}_m^n & = & u_m.
\end{array}
$$

Evidemment, la forme $m$-chaînée coïncide justement avec le système de contact canonique sur $J^n(\mathbb{R}^1, \mathbb{R}^m)$. D'autre part, des conditions nécessaires et suffisantes vérifiables pour qu'une distribution soit équivalente à la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^1, \mathbb{R}^m)$ ont été présentées par Respondek et Pasillas-Lépine [50] : Une distribution $\mathcal{D}$ de rang $m+1$, définie sur une variété $M$ de dimension $(n+1)m+1$, est localement, autour d'un point $p$, équivalente à la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^1, \mathbb{R}^m)$ si et seulement si les conditions suivantes sont satisfaites : (i) $\mathcal{D}^{(n)} = TM$, (ii) $\mathcal{D}^{(n-1)}$ est de rang constant $nm+1$ et contient une sous-distribution involutive $\mathcal{L}$ dont le corang dans $\mathcal{D}^{(n-1)}$ est constant et égal à 1, (iii) $\mathcal{D}(p)$ n'est pas contenue dans $\mathcal{L}(p)$. Parmi les trois conditions, la condition (iii) distingue les points singuliers et réguliers.

La caractérisation du système à $n$-barres dans $\mathbb{R}^{m+1}$ peut être exprimée par le théorème suivant :

**Theorem 0.0.5** *Le système à $n$-barres $\Gamma$ dans $\mathbb{R}^{m+1}$, pour $m \geq 1$, est localement équivalent au système $m$-chaîné en chaque point $x \in X = \mathbb{R}^{(n+1)(m+1)}$ qui satisfait $\Psi(x) = 0$ (c'est à dire $x$ correspond à un point $q \in Q$) et*

$$\text{(R1)} \quad \sum_{j=1}^{m+1} (x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0, \quad \text{pour } 1 \leq i \leq n-1, \quad \text{si } m \geq 2,$$

$$\text{(R2)} \quad \sum_{j=1}^{m+1} (x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0, \quad \text{pour } 2 \leq i \leq n-1, \quad \text{si } m = 1.$$

Les conditions $\Psi(x) = 0$ impliquent que $x$ est un point dans $Q$ et les conditions (R1) et (R2) décrivent les points réguliers du système $\Gamma$ dans $Q$. La condition $\sum_{j=1}^{m+1}(x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0$ est appelée *la condition de régularité* pour le système à $n$-barres. Notons $\theta_i$, pour $1 \leq i \leq n-1$, l'angle orienté entre le vecteur $\overrightarrow{P_{i-1}P_i}$ et $\overrightarrow{P_iP_{i+1}}$. Alors les conditions de régularité sont équivalentes à $\theta_i \neq \pm\frac{\pi}{2}$. Notre théorème montre que l'angle $\theta_1 = \pm\frac{\pi}{2}$ (i.e., la première barre est perpendiculaire à la deuxième bar) est une singularité pour le cas $m \geq 2$ mais pas pour le cas $m = 1$. Il est intéressant de remarquer que la propriété de la contrôlabilité peut être obtenue comme un sous-résultat du Théorèm 0.0.5 : *le système à $n$-barres $\Gamma$ est globalement contrôlable dans $Q = \mathbb{R}^{m+1} \times (S^m)^n$.*

La propriété de la platitude du système à $n$-barres dans $\mathbb{R}^{m+1}$ est aussi analysée dans ce chapitre. Dans le cas $m = 1$, la platitude du système du robot mobile avec remorques a été résolue par Fliess, Lévine, Martin et Rouchon [13] (voir aussi [26]). Dans le cas $m \geq 2$, nous donnons le résultat suivant :

**Theorem 0.0.6** (***Platitude du système à $n$-barre***) *Pour le système à $n$-barres $\Gamma$ dans $\mathbb{R}^{m+1}$, $m \geq 2$, nous avons*

- (i) $\Gamma$ *est $x$-plat au point $(q_0, u^0) \in Q \times \mathbb{R}^{m+1}$ qui satisfait*
  - (1) $\Psi(x) = 0$ *et* $\sum_{j=1}^{m+1}(x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0$, *où $q_0 \in Q$ est identifié avec un point $x \in \mathbb{R}^{(n+1)(m+1)}$ satisfaisant $\Psi(x) = 0$;*
  - (2) $u^0$ *est le contrôle tel que la vitesse $\dot{P}_0$ du point $P_0$ ne s'annule pas (et donc les vitesses $\dot{P}_i$, $0 \leq i \leq n-1$ sont non-nulles).*
- (ii) $P_0 = (x_0^1, x_0^2, \ldots, x_0^{m+1})$ *est une $x$-sortie plate minimale de $\Gamma$ en tout $(q_0, u^0)$ comme ci-dessus.*
- (iii) *Si $h = (h_0, \ldots, h_m)$ est une $x$-sortie plate minimale en $(q_0, u^0)$, alors localement autour de $q_0$ on a*

$$\mathrm{span}\{\mathrm{d}h_0, \ldots, \mathrm{d}h_m\} = \mathrm{span}\{\mathrm{d}x_0^1, \mathrm{d}x_0^2, \ldots, \mathrm{d}x_0^{m+1}\}.$$

**Remarque**. Il est important de remarquer la différence pour les $x$-sorties plates entre deux cas différents: dans le cas $m \geq 2$, les coordonnées du point $P_0 = (x_0^1, x_0^2, \ldots, x_0^{m+1})$ forment la seule $x$-sortie plate minimale de $\Gamma$. Contrairement, au cas $m = 1$, la $x$-sortie plate minimale n'est pas unique.

**Chapitre** 3. Le but de ce chapitre est de trouver des conditions nécessaires et suffisantes vérifiables pour qu'une distribution soit équivalente à la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$. En géométrie différentielle, la caractérisation intrinsèque de la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^k, \mathbb{R}^m)$ est encore un problème ouvert même si des solutions ont été obtenues dans certains cas particuliers. Le cas $n = 1$ et $m = 1$ a été résolu par Darboux [10] dans son théorème célèbre en généralisant les résulats de Pfaff [55] et Frobenius [15]; le cas $n = 2$, $m = 1$ et $k = 1$ par Engel [11]; le cas $n > 2$, $m = 1$ et $k = 1$ par E.von Weber, Cartan et Goursat (en point régulier), et par Libermann [37], Kumpera et Ruiz [33], et Murray [45](en point quelconque); le cas $n = 1$ par Bryant [7]; le cas général par Yamaguchi [79]; le cas $k = 1$ par Gardner et Shadwick [16], Murray [45], Tilbury et Sastry [76], Aranda-Bricaire et Pomet [3], Mormul [43], Respondek et Pasillas-Lépine [54].

Avant de présenter nos résultats, nous rappelons les concepts du rang d'Engel et du rang de Cartan. Le rang d'Engel, en point $p$, du système de Pfaff $\mathcal{I} = \operatorname{span}\{\omega_1, \ldots, \omega_s\}$ est l'entier le plus large $\rho$ tel qu'il existe une 1-forme différentielle $\alpha$ dans $\mathcal{I}$ qui satisfait la condition suivante

$$((d\alpha)^\rho \wedge \omega_1 \wedge \cdots \wedge \omega_s)(p) \neq 0.$$

Le *rang de Cartan* de $\mathcal{I}$ est l'entier le plus petit $k$ tel qu'il existe $\pi^1, \ldots, \pi^k \in \Lambda^1(M)$ tels que

$$\pi^1 \wedge \cdots \wedge \pi^k \wedge \omega^1 \wedge \cdots \wedge \omega^m \neq 0,$$

et

$$\mathrm{d}\omega \wedge \pi^1 \wedge \cdots \wedge \pi^k \equiv 0 \bmod \mathcal{I},\ \forall\, \omega \in \mathcal{I}.$$

**Theorem 0.0.7** *Soit $m \geq 3$ et $\mathcal{D}$ une distribution de rang $m(n+1)+2$ définie sur une variété $M$ de dimension $\frac{m}{2}(n+1)(n+2)+2$. Alors $\mathcal{D}$ est localement, autour du point $q \in M$, équivalente à la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$ si et seulement si les conditions suivantes sont satisfaites :*

- (i) *Chacune de ses distributions dérivées $\mathcal{D}^{(i)}$, pour $0 \leq i \leq n$, est de rang constant $\frac{m}{2}(2n+2-i)(i+1)+2$.*
- (ii) *Chacune de ses distributions dérivées $\mathcal{D}^{(i)}$, pour $0 \leq i \leq n-2$, contient la distribution caractéristique $\mathcal{C}_{i+1}$ de $\mathcal{D}^{(i+1)}$ dont le corang dans $\mathcal{D}^{(i)}$ est de deux. De plus, la distribution caractéristique de $\mathcal{D}$ est réduite à $\{0\}$.*
- (iii) *Le rang d'Engel de $(\mathcal{D}^{(n-1)})^\perp$ est constant et égal à deux.*
- (iv) $\dim(\mathcal{L} + \mathcal{D}^{(i)})(q) = \dim \mathcal{L}(q) + 2$, *pour $0 \leq i \leq n-2$, où $\mathcal{L}$ est une sous-distribution involutive de corang deux dans $\mathcal{D}^{(n-1)}$. L'existence de $\mathcal{L}$ est garantie par la condition* (iii).

Toutes ces conditions (i) $-$ (iv) sont vérifiables directement sur la distribution $\mathcal{D}$. Premièrement, le rang d'Engel peut être vérifié directement d'après la définition. Deuxièmement, si les conditions (i) $-$ (iii) sont satisfaites, la condition (iv) peut être vérifiée algébriquement parce que la sous-distribution $\mathcal{L}$, si elle existe, est unique et caclulable (voir Chapitre 3 pour les détails).

Le theorème suivant montre une relation entre le rang d'Engel et l'existence d'une sous-distribution involutive de corang $k$ et répond à la question : quand une distribution $\mathcal{D}$ contient une sous-distribution involutive $\mathcal{L}$ qui soit de corang $k$ dans $\mathcal{D}$?

**Theorem 0.0.8** *Soit $m \geq 3$ et $\mathcal{D}$ une distribution de rang $k + mk$, définie sur une variété $M$ de dimension $m + k + mk$, telle que $\mathcal{D}^{(1)} = TM$. Si la distribution caractéristique de $\mathcal{D}$ est réduite à $\{0\}$, alors les conditions suivantes sont équivalentes:*

(i) *Le rang de Cartan de $\mathcal{D}^{\perp}$ est constant et égal à $k$;*
(ii) *Le rang d'Engel de $\mathcal{D}^{\perp}$ est constant et égal à $k$;*
(iii) *Il existe une sous-distribution $\mathcal{F}$ de $\mathcal{D}$ de corang $k$ telle que $[\mathcal{F}, \mathcal{F}] \subset \mathcal{D}$ et il n'existe aucune autre sous-distribution de $\mathcal{D}$ vérifiant cette propriété et de corang plus petit que $k$.*
(iv) *Il existe une sous-distribution involutive $\mathcal{L}$ qui est de corang $k$ dans $\mathcal{D}$ et il n'existe aucune autre sous-distribution involutive $\mathcal{D}$ de corang plus petit que $k$.*

*De plus, la sous-distribution involutive $\mathcal{L}$ de $\mathcal{D}$ de corang $k$, si elle existe, est unique et calculable.*

Afin de démontrer Theorème 0.0.7, nous proposons aussi une forme normale de Bryant étendue qui généralise la caractérisation de la distribution de Cartan $\mathcal{CC}^1(\mathbb{R}^k, \mathbb{R}^m)$.

**Chapitre** 4. Cette dernière partie est consacrée à la linéarisation d'un système avec multi-contrôles par le bouclage orbital. La linéarisation par bouclage est un outil très important dans l'étude des systèmes non linéaires dans le fait qu'on peut ainsi appliquer les propriétés d'un système linéaire à un système non linéaire. Le problème de la linéarisation par bouclage d'un système avec un seul contrôle a été résolu par Brockett [5]. Ensuite Jakubczyk et Respondek [29] ainsi que Hunt et Su [22] ont donné des conditions nécessaires et suffisantes de la linéarisation par bouclage pour un système affine avec multi-contrôles. Pour les systèmes qui ne peuvent pas être linéarisés par bouclage, il est possible qu'ils soient linéarisés par bouclage avec un changement de temps qui s'appelle linéarisation par bouclage orbital. Ce problème a été proposé et étudié premièrement par Sampei et Furuta [67]. Malheureusement leurs conditions necessitent la resolution d'un système d'EDP non linéaire afin d'obtenir la transformation du changement de temps et donc elles ne peuvent pas être verifiées pour le système original. En 1998, Respondek [62] a présenté des contitions nécessaires et suffisantes pour la linéarisation par bouclage orbital d'un système à un seul contrôle et ses conditions sont vérifiables pour le système original. Après son travail, Guay [19] a donné en 2000 des conditions équivalentes avec une approche basée sur les formes différentielles et les a développées en 2001 à la linéarisation par bouclage orbital d'un système à multi-contrôles [20]. Pour vérifier les conditions de Guay, on a besoin de chercher des générateurs convenables du système de Pfaff associés au système original.

Considérons le système suivant, pour $m \geq 2$,

$$\Sigma : \quad \frac{dx}{dt} = f(x) + \sum_{i=1}^{m} g_i(x)u_i, \;\; f(x_0) \neq 0, \tag{0.0.6}$$

où $x \in X = \mathbb{R}^{(n+1)m+1}$.

Dans cette thèse, le système $\Sigma$ est dit linéarisable par bouclage orbital s'il est localement équivalent au système linéaire

$$\Lambda_t : \quad \begin{array}{rcl} \dot{x}^0 & = & 1 \\ \dot{\bar{x}} & = & \bar{A}\bar{x} + \sum_{i=1}^{m} \bar{b}_i \bar{u}_i. \end{array}$$

Il est clair que le système $\Lambda_t$ n'est jamais controllable mais il peut être accessible (voir [24], [47] pour la définition) si et seulement si le sous-système $(\bar{A}, \bar{B})$ est controllable, où $\bar{B} = (b_1, \ldots, b_m)$. Par la suite nous appelerons les indices de contrôlabilité et la forme canonique de Brunovsky de $\Lambda_t$ comme étant celles du sous-système $(\bar{A}, \bar{B})$.

Définissons les distributions suivantes

$$\begin{array}{rcl} \mathcal{G} & = & \operatorname{span}\{g_1, \ldots, g_m\}, \\ \mathcal{G}_f^j & = & \operatorname{span}\{f, g_i, \operatorname{ad}_f g_i, \ldots, \operatorname{ad}_f^{j-1} g_i, \quad 1 \leq i \leq m\}, \end{array}$$

pour $1 \leq j \leq n+1$. Supposons que $\dim \mathcal{G}_f^{n+1}(x) = (n+1)m + 1$, alors nous avons $\dim \mathcal{G}_f^n(x) = nm + 1$. En conséquence, il existe $m$ formes différentielles $\omega^1, \ldots, \omega^m \in \Lambda(X)$ définies uniquement par

$$\begin{array}{rcl} < \omega^j, g > & = & 0, \quad \forall\, g \in \mathcal{G}_f^n \\ < \omega^j, ad_f^n g_i > & = & \delta_i^j, \end{array}$$

pour $1 \leq i, j \leq m$. Grâce aux $m$ formes differentielles, les fonctions suivantes peuvent être introduites :

$$\begin{array}{rcl} T_{i,j}^{k,l} & = & < \omega^k, [\operatorname{ad}_f^{n-1} g_i, \operatorname{ad}_f^l g_j] > \\ & = & \omega^k([\operatorname{ad}_f^{n-1} g_i, \operatorname{ad}_f^l g_j]). \end{array}$$

Le théorème suivant décrit des conditions nécessaires et suffisantes du problème de la linéarisation par bouclage orbital pour le système (0.0.6). Toutes ses conditions peuvent être vérifiées directement sur le système original.

**Theorem 0.0.9** *Le système non linéaire $\Sigma$, $m \geq 2$, défini par (0.0.6), est localement équivalent par bouclage orbital au système linéaire $\Lambda_t$ dont tous les indices de contrôlabilité égalent $n+1$ si et seulement s'il satisfait, autour du point $p$, les conditions suivantes :*

(OPL1) $\dim \mathcal{G}_f^{n+1}(x) = (n+1)m + 1$

(OPL2) $[\mathcal{G}_f^j, \mathcal{G}_f^j] \subset \mathcal{G}_f^{j+1}, \quad 1 \leq j \leq n,$

(OPL3) $[\mathcal{G}, \mathcal{G}_f^2] \subset \mathcal{G}_f^2,$

(OPL4) *Les fonctions $\mathrm{T}_{i,j}^{k,l}$ satisfont les conditions suivantes:*

(i) $\mathrm{T}_{i,j}^{k,l} = 0$, pour $\begin{cases} 1 \leq k \neq i \leq m, \quad 1 \leq j \leq m, \quad \text{si} \quad l < n-1 \\ 1 \leq i \neq k \neq j \leq m \quad \text{si} \quad l = n-1 \end{cases}$

(ii) $\mathrm{T}_{i,j}^{i,l} = T_{k,j}^{k,l}$, pour $1 \leq l \leq n-1$, $1 \leq i, j, k \leq m$ t.q $j \neq i, k$ si $l = n-1$,

(iii) *Quand $m = 2$, la distribution définie par*

$$\mathcal{B} = \operatorname{span}\{g_i, \operatorname{ad}_f^l g_i + b_i^l f, \quad i = 1, 2 \text{ et } 1 \leq l \leq n-1\}$$

*doit être involutive, où*

$$\begin{aligned} &b_i^l = \mathrm{T}_{1,i}^{1,l} = \mathrm{T}_{2,i}^{2,l}, \text{ pour } 1 \leq i \leq 2, 1 \leq l \leq n-2 \\ &b_1^{n-1} = \mathrm{T}_{1,2}^{1,n-l}, \quad b_2^{n-1} = \mathrm{T}_{2,1}^{2,n-l}. \end{aligned}$$

Il est important de remarquer que la fonction $\gamma$ qui décrit le changement de l'échelle de temps, par $\dfrac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma(x(t))$, peut être construite directement à partir de la distribution $\mathcal{B}$ que l'on a définie dans ce théorème (voir les détails dans Chapitre 4).

Notre résultat implique aussi une relation intéressante entre la linéarisation par bouclage orbital du système (0.0.6) et la caractérisation de la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^1, \mathbb{R}^m)$. Il est connu que les systèmes de type $\Lambda_t$, dont tous les indices de contrôlabilité sont constants et égaux à $n+1$, peuvent être transformés par un changement de coordonnées et un bouclage à la $t$-etendue forme normale de Brunovsky ($t$-augmented Brunovsky canonical form)

$$\Lambda_t^{Br} : \quad \begin{array}{lllllll} \dot{x}_0^0 & = & 1 & & & & \\ \dot{x}_1^0 & = & x_1^1 & \cdots & \dot{x}_m^0 & = & x_m^1 \\ & \vdots & & \vdots & & \vdots & \\ \dot{x}_1^{n-1} & = & x_1^n & \cdots & \dot{x}_m^{n-1} & = & x_m^n \\ \dot{x}_1^n & = & u_1 & \cdots & \dot{x}_m^n & = & u_m. \end{array}$$

A chaque système de contrôle défini par (0.0.6), il correspond à une distribution affine $\mathcal{A} = f + \mathcal{G}$ qui définit en chaque point $x \in X$ un sous-espace affine $\mathcal{A}(x) = f(x) + \mathcal{G}(x)$ de l'espace tangent $T_xX$. Considérons deux systèmes affines $\Sigma$ et $\tilde{\Sigma}$ tels que les distributions de contrôle $\mathcal{G}$ et $\tilde{\mathcal{G}}$ soient de rang constant. Alors ils sont équivalents par bouclage si et seulement s'il existe un diffeomorphisme $\varphi$ tel que

$$\tilde{\mathcal{A}} = \varphi_* \mathcal{A},$$

et sont équivalents par bouclage orbital s'il existe un diffeomorphisme $\varphi$ et une fonction $\gamma(\cdot) \neq 0$ tels que

$$\tilde{\mathcal{A}} = \varphi_*(\gamma \mathcal{A}),$$

où $\gamma\mathcal{A} = \gamma f + \mathcal{G}$. Soit $\mathcal{D}$ est une distribution de rang constant sur $X$. Pour chaque distribution affine $\mathcal{A} \subset \mathcal{D}$, où $\mathcal{A} = f + \mathcal{G}$ est telle que $\operatorname{corang}(\mathcal{G} \subset \mathcal{D}) = 1$ et $f(x) \notin \mathcal{G}(x)$, nous l'associons au système de contrôle $\Sigma_{\mathcal{A}}$. Nous avons les deux résultats suivants qui décrivent la relation entre la linéarisation par bouclage orbital du système (0.0.6) et la caractérisation de la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^1, \mathbb{R}^m)$.

**Proposition 0.0.10** *Soit $\mathcal{D}$ une distribution de rang constant dans laquelle on choisit une distribution affine $\mathcal{A} = f + \mathcal{G} \subset \mathcal{D}$ telle que* $\operatorname{corang}(\mathcal{G} \subset \mathcal{D}) = 1$ *et $f(x) \notin \mathcal{G}(x)$. Si $\mathcal{D}$ satisfait les conditions suivantes*

(D1) *$\mathcal{D}$ est localement équivalente à la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$;*
(D2) *$\mathcal{C}(\mathcal{D}^{(1)}) = \mathcal{G}$,*

*alors le système affine associé $\Sigma_{\mathcal{A}}$ est localement équivalent par bouclage orbital au système $\Lambda_t^{Br}$.*

**Proposition 0.0.11** *Un système de contrôle $\Sigma$, défini par (0.0.6), est localement équivalent par bouclage orbital au système $\Lambda_t^{Br}$ dont tous les indices de contrôlabilité sont constants et égaux à $n+1$ si et seulement si la distribution associée $\mathcal{D}_\Sigma$ satisfait les conditions suivantes:*

(C1) *$\mathcal{D}_\Sigma$ est localement équivalente à la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$;*
(C2) *$\mathcal{C}(\mathcal{D}_\Sigma^{(1)}) = \mathcal{G}$.*

Grâce aux deux Propositions, nous présentons une version complète pour Théorème 0.0.9 dans laquelle nous donnons des conditions équivalentes aux conditions (OPL1) − (OPL4) (voir Chapitre 4 pour les détails).

# Chapter 1

# Flat outputs of driftless two-input control systems

## 1.1 Introduction

The notion of flatness has been introduced by Fliess, Lévine, Martin and Rouchon in [12], [13], [14] in order to describe the class of control systems, whose set of trajectories can be parameterized by a finite number of functions and their time-derivatives. More formally, a system with $m$ controls is flat if we can find $m$ functions (of the state and control variables and their time-derivatives), called *flat outputs*, such that the evolution in time of the state and control can be expressed in terms of flat outputs and their time derivatives (see Section 1.2 for a precise definition and references).

As an introductory example, consider the nonholonomic car or, equivalently, a unicycle-like robot towing a trailer (see, e.g., [34]). Denote by $(x, y) \in \mathbb{R}^2$ the position of the mid-point of the rear wheels, and by $\theta_0$ and $\theta_1$, respectively, the angles between the rear and front wheels and the $x$-axis. The controls $u_1$ and $u_2$ allow to move (forward and backward) the car and to turn. The car is subject to two nonholonomic constraints: the wheels are not allowed to slide. This leads to the following model given by a driftless, i.e., control-linear, system on $\mathbb{R}^2 \times S^1 \times S^1$:

$$\Sigma_{\mathrm{car}} : \quad \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta}_0 \\ \dot{\theta}_1 \end{pmatrix} = \begin{pmatrix} \cos(\theta_1 - \theta_0)\cos\theta_0 \\ \cos(\theta_1 - \theta_0)\sin\theta_0 \\ \sin(\theta_1 - \theta_0) \\ 0 \end{pmatrix} u_1 + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} u_2.$$

It is well known, as proved by Fliess, Lévine, Martin and Rouchon [13] (see also [26]) that the nonholonomic car is flat and that the position of the mid-point of the rear

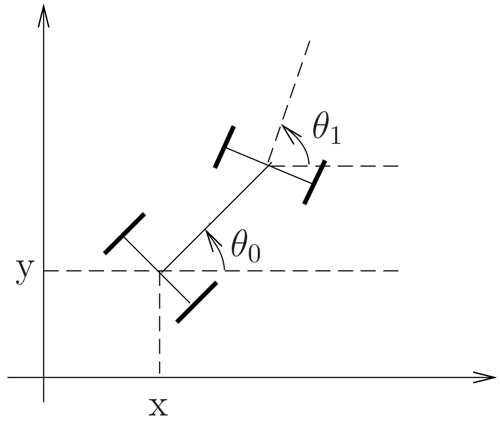


Figure 1.1: nonholonomic car

wheels is a flat output. Indeed, the following coordinates change

$$
\begin{aligned}
x_1 &= x \\
x_2 &= y \\
x_3 &= \tan\theta_0 \\
x_4 &= \tan(\theta_1 - \theta_0)\sec^3\theta_0
\end{aligned}
$$

for $\theta_0 \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and $\theta_1 - \theta_0 \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, followed by the invertible feedback

$$
\begin{aligned}
v_1 &= u_1\cos(\theta_1 - \theta_0)\cos\theta_0 \\
v_2 &= u_1\sec^3\theta_0\sin(\theta_1 - \theta_0)\left(3\tan(\theta_1 - \theta_0)\tan\theta_0 - \sec^2(\theta_1 - \theta_0)\right) + u_2\sec^3\theta_0\sec^2(\theta_1 - \theta_0),
\end{aligned}
$$

brings the system $\Sigma_{\text{car}}$ into the *chained form*:

$$
\begin{cases}
\dot{x_1} = v_1 \\
\dot{x_2} = x_3 v_1 \\
\dot{x_3} = x_4 v_1 \\
\dot{x_4} = v_2.
\end{cases}
$$

It is easy to see that the pair of functions $h = (h_1, h_2) = (x_1, x_2) = (x, y)$ are flat outputs for the chained form. Indeed, we have $x_1 = h_1, x_2 = h_2, x_3 = \frac{\dot{h}_2}{\dot{h}_1}, x_4 = \frac{1}{\dot{h}_1}\frac{d}{dt}(\frac{\dot{h}_2}{\dot{h}_1})$, $v_1 = \dot{h}_1$ and $v_2 = \frac{1}{\dot{h}_1}(\dot{x}_4)$. The applied transformation, consisting of a change of coordinates and feedback, is invertible which proves that, indeed, $h = (x, y)$ are flat outputs of the nonholonomic car.

The presented procedure, to express the state and control in terms of $h$ and its time-derivatives, exhibits three singularities: two in the state space (at $\theta_0 = \pm\frac{\pi}{2}$ and

$\theta_1 - \theta_0 = \pm\frac{\pi}{2}$) and one in the control space (when $u_1(t) = v_1(t) = 0$). Let us analyze those singularities.

First, the singularity at $u_1(t) = 0$ (at least at $u_1(t) \equiv 0$) seems to be intrinsic: we cannot see how the angles $\theta_0(t)$ and $\theta_1(t)$ evolve if the observed point $(x(t), y(t))$ does not move.

Secondly, $\theta_0 = \pm\frac{\pi}{2}$ is not a singularity of the flat output $h = (x, y)$. Indeed, around a point such that $\theta_0 = \pm\frac{\pi}{2}$, we can choose the following coordinates change:

$$
\begin{aligned}
x_1 &= y \\
x_2 &= x \\
x_3 &= \cot\theta_0 \\
x_4 &= -\tan(\theta_1 - \theta_0)\csc^3\theta_0,
\end{aligned}
$$

for $\theta_0 \in (0, \pi)$ including $\theta_0 = -\frac{\pi}{2}$ (or $\theta_0 \in (-\pi, 0)$, including $\theta_0 = -\frac{\pi}{2}$) and $\theta_1 - \theta_0 \in (-\frac{\pi}{2}, \frac{\pi}{2})$, followed by a suitable feedback transformation, which brings the system $\Sigma_{\text{car}}$ into the chained form around $\theta_0 = \pm\frac{\pi}{2}$. Therefore, the nonholonomic car is flat, with $(x, y)$ as a flat output, at any $q = (x, y, \theta_0, \theta_1)$ where $\theta_1 - \theta_0 \neq \pm\frac{\pi}{2}$. Actually, the singularities at $\theta_0 = \pm\frac{\pi}{2}$ are related with the domain of inversion: when calculating $\theta_0$ in terms of $x$, $y$, $\dot{x} = u_1\cos(\theta_1 - \theta_0)\cos\theta_0$ and $\dot{y} = u_1\cos(\theta_1 - \theta_0)\sin\theta_0$, we have to invert either $\tan\theta_0 = \frac{\dot{y}}{\dot{x}}$ or $\cot\theta_0 = \frac{\dot{x}}{\dot{y}}$.

Thirdly, and most interestingly, the nature of the singularity at $\theta_1 - \theta_0 = \pm\frac{\pi}{2}$ is completely different. It is an *intrinsic* singularity of the flat output $h = (x, y)$, see Section 1.3.1 for details, nevertheless the nonholonomic car is flat at $\theta_1 - \theta_0 = \pm\frac{\pi}{2}$ if we choose another flat output! To see this, define two functions

$$
\begin{aligned}
\bar{h}_1(x, y, \theta_0) &= x\sin\theta_0 - y\cos\theta_0 \\
\bar{h}_2(x, y, \theta_0) &= x\cos\theta_0 + y\sin\theta_0
\end{aligned}
$$

and consider the change of coordinates

$$
\begin{aligned}
y_1 &= \theta_0 \\
y_2 &= \bar{h}_1(x, y, \theta_0) \\
y_3 &= \bar{h}_2(x, y, \theta_0) \\
y_4 &= \cot(\theta_1 - \theta_0) - \bar{h}_1(x, y, \theta_0)
\end{aligned}
$$

followed by the invertible feedback

$$
\begin{aligned}
v_1 &= u_1\sin(\theta_1 - \theta_0) \\
v_2 &= u_1(\csc^2(\theta_1 - \theta_0) - \bar{h}_2(x, y, \theta_0)) - u_2\csc^2(\theta_1 - \theta_0),
\end{aligned}
$$

which also brings $\Sigma_{\text{car}}$ into *chained form*, but this time around $\theta_1 - \theta_0 = \pm\frac{\pi}{2}$,

$$
\begin{cases}
\dot{y}_1 = v_1 \\
\dot{y}_2 = y_3 v_1 \\
\dot{y}_3 = y_4 v_1 \\
\dot{y}_4 = v_2.
\end{cases}
$$

Thus $(\bar{h}_1, \bar{h}_2) = (\theta_0, x\sin\theta_0 - y\cos\theta_0)$ is another flat output of the nonholonomic car, valid around $\theta_1 - \theta_0 = \pm\frac{\pi}{2}$, but singular at $\theta_1 - \theta_0 = 0, \pm\pi$ (notice that the same singularity $v_1(t) = u_1(t) = 0$, as previously, occurs in the control space).

A series of natural questions arises: are there other flat outputs of the nonholonomic car and, if so, how many and how to describe them? More generally, how to characterize all flat outputs of any 2-input driftless control system and how to describe their singular loci and singular controls? The aim of this chapter is to give complete answers to those questions.

This chapter is organized as follows. In Section 1.2, we define the crucial notion of flatness and recall a description of flat driftless 2-input systems. In section 1.3, we give our main results. We characterize all flat outputs of driftless 2-input systems and give a way of parameterizing them: it turns out that all flat outputs can be parameterized by an arbitrary function of intrinsically defined three variables. We also construct a system of 1st order PDE's whose solutions are flat outputs of a given system. Still in Section 1.3, we illustrate our results by describing all flat outputs of some examples: nonholonomic car (1-trailer system) which we have just discussed and then the $n$-trailer system. We prove our results in Section 1.4.

## 1.2 Flatness of driftless two-input control systems

Throughout this chapter, the word smooth will always mean $C^\infty$-smooth. Consider a smooth nonlinear control system

$$\Xi : \quad \dot{x} = f(x, u),$$

where $x \in X$, an $n$-dimensional manifold and $u \in U$, an $m$-dimensional manifold. Given any integer $l$, we associate to $\Xi$ its $l$-prolongation $\Xi^l$ given by

$$\Xi^l : \quad \begin{array}{rcl} \dot{x} &=& f(x, u^0) \\ \dot{u}^0 &=& u^1 \\ &\vdots& \\ \dot{u}^l &=& u^{l+1}, \end{array}$$

which can be considered as a control system on $X^l = X \times U \times \mathbb{R}^{ml}$, whose state variables are $(x, u^0, u^1, \ldots, u^l)$ and whose $m$ controls are the $m$ components of $u^{l+1}$. Denote $\bar{u}^l = (u^0, u^1, \ldots, u^l)$.

**Definition 1.2.1** The system $\Xi$ is called flat at a point $(x_0, \bar{u}_0^l) \in X^l = X \times U \times \mathbb{R}^{ml}$, for some $l \geq 0$, if there exist a neighborhood $\mathcal{O}^l$ of $(x_0, \bar{u}_0^l)$ and $m$ smooth functions

$$h_i = h_i(x, u^0, u^1, \ldots, u^l), \quad 1 \leq i \leq m,$$

called *flat outputs*, defined in $\mathcal{O}^l$, having the following property: there exist an integer $s$ and smooth functions $\gamma_i$, $1 \le i \le n$, and $\delta_i$, $1 \le i \le m$, such that we have

$$\begin{aligned} x_i &= \gamma_i(h, \dot{h}, \ldots, h^{(s)}) \\ u_i &= \delta_i(h, \dot{h}, \ldots, h^{(s)}), \end{aligned}$$

where $h = (h_1, \ldots, h_m)^\top$, along any trajectory $x(t)$ given by a control $u(t)$ that satisfy $(x(t), u(t), \dot{u}(t), \ldots, u^{(l)}(t)) \in \mathcal{O}^l$.

The compositions $\gamma_i(h, \dot{h}, \ldots, h^{(s)})$ and $\delta_i(h, \dot{h}, \ldots, h^{(s)})$ are, a priori, defined in an open set $\mathcal{O}^{s+l} \subset X^{s+l} = X \times U \times \mathbb{R}^{m(s+l)}$. The above definition requires that $\pi(\mathcal{O}^{s+l}) \supset \mathcal{O}^l$, where $\pi(x, \bar{u}^{s+l}) = (x, \bar{u}^l)$, and that for all such $(x, \bar{u}^{s+l})$, the compositions yield, respectively, $x_i$ and $u_i$.

If $h_i = h_i(x, u^0, u^1, \ldots, u^r)$, $r \le l$, we will say that the system is $(x, u, \ldots, u^r)$-flat and, in particular, $x$-flat if $h_i = h_i(x)$. In the case $h_i = h_i(x, u^0, u^1, \ldots, u^r)$, we will assume that they are defined on $\mathcal{O}^r \subset X^r = X \times U \times \mathbb{R}^{mr}$, where $\pi^{-1}(\mathcal{O}^r) \supset \mathcal{O}^l$ and $\pi$ stands for the projection $\pi(x, u^0, \ldots, u^r, \ldots, u^l) = (x, u^0, \ldots, u^r)$.

The notion of flatness has been introduced in control theory by Fliess, Lévine, Martin and Rouchon [12], [13], [14] (see also [2], [25], [26], [56], [70], [77]), and has attracted a lot of attention because of its extensive applications in constructive controllability and trajectory tracking, compare [38] and references therein. A similar notion (of underdetermined systems of differential equations that are integrable without integration) has already been studied by Cartan [9] and Hilbert [21].

In this chapter, we deal only with two-input driftless (equivalently, control-linear) systems of the form

$$\Sigma: \quad \dot{x} = f_1(x)u_1 + f_2(x)u_2,$$

on an $(n+2)$-dimensional manifold $M$, where $f_1$ and $f_2$ are $C^\infty$-smooth vector fields independent everywhere on $M$ and $u = (u_1, u_2)^\top \in \mathbb{R}^2$. To this system, we associate the distribution $\mathcal{D}$ spanned by the vector fields $f_1, f_2$, which will be denoted by $\mathcal{D} = \operatorname{span}\{f_1, f_2\}$. Consider another 2-input driftless system

$$\tilde{\Sigma}: \quad \dot{\tilde{x}} = \tilde{f}_1(\tilde{x})\tilde{u}_1 + \tilde{f}_2(\tilde{x})\tilde{u}_2,$$

where $\tilde{f}_1$ and $\tilde{f}_2$ are $C^\infty$-smooth vector fields on $\tilde{M}$. Form the matrices $f(x) = (f_1(x), f_2(x))$ and $\tilde{f}(\tilde{x}) = (\tilde{f}_1(\tilde{x}), \tilde{f}_2(\tilde{x}))$. The systems $\Sigma$ and $\tilde{\Sigma}$ are feedback equivalent if there exist an invertible $2 \times 2$-matrix $\beta$, whose entries $\beta_{ij}$, $1 \le i, j \le 2$, are $C^\infty$-smooth functions on $M$, and a diffeomorphism $\varphi : M \to \tilde{M}$ such that

$$\mathrm{D}\varphi(x) \cdot f(x) \cdot \beta(x) = \tilde{f}(\varphi(x)).$$

It is easily seen that $\Sigma$ and $\tilde{\Sigma}$ are locally feedback equivalent if and only if the associated distributions $\mathcal{D} = \operatorname{span}\{f_1, f_2\}$ and $\tilde{\mathcal{D}} = \operatorname{span}\{\tilde{f}_1, \tilde{f}_2\}$ are locally equivalent via $\varphi$, i.e.,

$$\mathrm{D}\varphi \cdot \mathcal{D}(x) = \tilde{\mathcal{D}}(\varphi(x)).$$

The *derived flag* of a distribution $\mathcal{D}$ is the sequence of modules of vector fields $\mathcal{D}^{(0)} \subset \mathcal{D}^{(1)} \subset \cdots$ defined inductively by

$$\mathcal{D}^{(0)} = \mathcal{D} \quad \text{and} \quad \mathcal{D}^{(i+1)} = \mathcal{D}^{(i)} + [\mathcal{D}^{(i)}, \mathcal{D}^{(i)}], \quad \text{for} \quad i \geq 0.$$

The *Lie flag* of $\mathcal{D}$ is the sequence of modules of vector fields $\mathcal{D}_0 \subset \mathcal{D}_1 \subset \cdots$ defined inductively by

$$\mathcal{D}_0 = \mathcal{D} \quad \text{and} \quad \mathcal{D}_{i+1} = \mathcal{D}_i + [\mathcal{D}_0, \mathcal{D}_i], \quad \text{for} \quad i \geq 0.$$

In general, the derived and Lie flags are different though for any point $x$, the inclusion $\mathcal{D}_i(x) \subset \mathcal{D}^{(i)}(x)$ holds, for $i \geq 0$.

A characteristic vector field of a distribution $\mathcal{D}$ is a vector field $f$ that belongs to $\mathcal{D}$ and satisfies $[f, \mathcal{D}] \subset \mathcal{D}$. The characteristic distribution of $\mathcal{D}$, which will be denoted by $\mathcal{C}$, is the subdistribution spanned by all its characteristic vector fields. It follows directly from the Jacobi identity that the characteristic distribution is always involutive but, in general, it need not be of constant rank.

The problem of flatness of driftless 2-input systems has been studied and solved by Martin and Rouchon [39] (see also [40] and a related work of Cartan [9]). Their important result proves that a system is flat if and only if its associated distribution $\mathcal{D}$ satisfies, on an open and dense subset $M'$ of $M$, the conditions

$$\operatorname{rank} \mathcal{D}^{(i)} = i + 2, \quad 0 \leq i \leq n. \tag{1.2.1}$$

A distribution $\mathcal{D}$ is called a Goursat structure (also a "système en drapeau" in [33] and a Goursat flag in [42]) if it satisfies the conditions (1.2.1) at any point $x \in M$. It is known since the work of von Weber [78], Cartan [9] and Goursat [18] that the conditions (1.2.1) imply that on an open and dense subset $M''$ of $M$, the distribution $\mathcal{D}$ can be brought to the *Goursat normal form*, or equivalently, the corresponding control system is feedback equivalent to the *chained form*:

$$\Sigma_{\text{chain}} : \quad \begin{cases} \dot{z}_1 & = & v_1 \\ \dot{z}_2 & = & z_3 v_1 \\ \dot{z}_3 & = & z_4 v_1 \\ & \vdots & \\ \dot{z}_{n+1} & = & z_{n+2} v_1 \\ \dot{z}_{n+2} & = & v_2. \end{cases} \tag{1.2.2}$$

It is easy to see that $\Sigma_{\text{chain}}$ is $x$-flat with $x$-flat outputs chosen as $h = (h_1, h_2) = (z_1, z_2)$ and provided that the control $v_1 \neq 0$ (compare Introduction, where we brought the nonholonomic car to the chained form for $\dim M = 4$). Giaro, Kumpera and Ruiz [17] were the first to observe the existence of singular points in the problem of transforming a distribution of rank two into the Goursat normal form. Murray [45] proved that the feedback equivalence of $\Sigma$ to the chained form $\Sigma_{\text{chain}}$ (or, in other words, equivalence of the associated distribution to the Goursat normal form), around an arbitrary point $x_0$ requires, in addition to (1.2.1), the regularity condition (see Theorem 1.2.2 below)

$$\dim \mathcal{D}^{(i)}(x_0) = \dim \mathcal{D}_i(x_0), \quad 0 \le i \le n. \tag{1.2.3}$$

A natural question arises: can $\Sigma$ be locally flat at a singular point of $\mathcal{D}$, i.e., at a point not satisfying the regularity condition (1.2.3)? In other words, can a driftless 2-input system be flat without being locally equivalent to the chained form? Theorem 1.2.2 answers this question (in what concerns $x$-flatness).

Let $\mathcal{D}$ be any distribution of rank two such that $\operatorname{rank} \mathcal{D}^{(1)} = 3$ and $\operatorname{rank} \mathcal{D}^{(2)} = 4$. Then there exists a distribution $\mathcal{C}_1 \subset \mathcal{D}$ of corank one which is characteristic for $\mathcal{D}^{(1)}$, i.e., $[\mathcal{C}_1, \mathcal{D}^{(1)}] \subset \mathcal{D}^{(1)}$. Indeed, the above rank assumptions imply that (after permuting $f_1$ and $f_2$, if necessary) there exists a smooth function $\alpha$ such that

$$[f_2, [f_1, f_2]] = \alpha [f_1, [f_1, f_2]] \mod \mathcal{D}^{(1)}.$$

It follows that $[f_2 - \alpha f_1, [f_1, f_2]] = 0 \mod \mathcal{D}^{(1)}$ and hence $\mathcal{C}_1 = \operatorname{span}\{f_2 - \alpha f_1\}$. Let $U_{\text{sing}}(x)$ be the 1-dimensional subspace of $\mathbb{R}^2$ such that for any feedback control $(u_1(x), u_2(x))^\top = u(x) \in U_{\text{sing}}(x)$, we have $f_1(x)u_1(x) + f_2(x)u_2(x) \in \mathcal{C}_1(x)$ (clearly, $U_{\text{sing}}(x)$ is spanned by $(\alpha(x), -1)^\top$). Any control $u(t) \in U_{\text{sing}}(x(t))$ will be called *singular* and the trajectories of the system governed by a singular control remain *tangent* to the characteristic subdistribution $\mathcal{C}_1$. We have just given the definition of $U_{\text{sing}}(x)$ for $\dim M \ge 4$ (since we have used $\operatorname{rank} \mathcal{D}^{(2)} = 4$). If $\dim M = 3$, we define $U_{\text{sing}}(x) = 0 \in \mathbb{R}^2$. Note that if $l = 0$, we will denote a fixed control value by $u_0$ (instead of more complicated $u_0^0$).

**Theorem 1.2.2** *Consider a 2-input driftless control system*

$$\Sigma: \quad \dot{x} = f_1(x)u_1 + f_2(x)u_2,$$

*where $x \in M$, an $(n+2)$-dimensional manifold, $n \ge 1$. Assume that the distribution $\mathcal{D} = \operatorname{span}\{f_1, f_2\}$ associated to $\Sigma$ is a Goursat structure, that is, satisfies $\operatorname{rank} \mathcal{D}^{(i)} = i + 2$, for $0 \le i \le n$, everywhere on $M$. Then the following conditions are equivalent:*

(i) *$\Sigma$ is $x$-flat at $(x_0, \bar{u}_0^l) \in M \times \mathbb{R}^{2(l+1)}$, for a certain $l \ge 0$;*
(ii) *$\Sigma$ is $x$-flat at $(x_0, u_0) \in M \times \mathbb{R}^2$;*
(iii) *$\dim \mathcal{D}^{(i)}(x_0) = \dim \mathcal{D}_i(x_0)$ for $0 \le i \le n$ and $u_0 \notin U_{\text{sing}}(x_0)$;*
(iv) *$\Sigma$ is locally, around $x_0$, feedback equivalent to the chained form $\Sigma_{\text{chain}}$ and $u_0 \notin U_{\text{sing}}(x_0)$.*

We assume that $\mathcal{D}$ satisfies $\operatorname{rank} \mathcal{D}^{(i)} = i+2$, for $0 \leq i \leq n$, so the characteristic distribution $\mathcal{C}_1$ and the set of singular controls $U_{\text{sing}}$ are well defined. The above theorem implies that a driftless 2-input system is never flat at $(x_0, u_0)$ such that $u_0 \in U_{\text{sing}}(x_0)$. Therefore any $x$-flat outputs $(\varphi_1, \varphi_2)$ become singular in the control space (at $u_0 \in U_{\text{sing}}$) but they may also exhibit singularities in the state space $M$. To formalize this, assume that a pair of functions $(\varphi_1, \varphi_2)$ defined in an open set $\mathcal{M} \subset M$ are $x$-flat outputs at a point $(x_0, u_0) \in M \times \mathbb{R}^2$, that is, there exists a neighborhood $\mathcal{O}^0 \subset M \times \mathbb{R}^2$, satisfying $\mathcal{O}^0 \subset \pi^{-1}(\mathcal{M})$, where $\pi(x, u) = x$, in which the conditions of Definition 1.2.1 hold. If a pair of functions $(\varphi_1, \varphi_2)$, defined in $\mathcal{M} \subset M$, is an $x$-flat output at $(x, u)$ for any $x \in \tilde{\mathcal{M}}$ and certain $u = u(x)$, where $\tilde{\mathcal{M}}$ is open and dense in $\mathcal{M}$, then by the singular locus of $(\varphi_1, \varphi_2)$, denoted by $Sing(\varphi_1, \varphi_2)$, we will mean the set of points $x \in \mathcal{M}$ such that $(\varphi_1, \varphi_2)$ are not $x$-flat outputs at $(x, u)$ for any $u \in \mathbb{R}^2$.

The interest of the above theorem is two-fold. First, together with its proof, it will allow us to characterize all $x$-flat outputs of driftless 2-input systems (see Section 1.3). Secondly, it shows that a Goursat structure is $x$-flat at points $x_0$ satisfying $\dim \mathcal{D}^{(i)}(x_0) = \dim \mathcal{D}_i(x_0)$, for $0 \leq i \leq n$, only, that is, at regular points of $\mathcal{D}$. Martin and Rouchon asked in [39] (see also [40]) whether a Goursat structure $\mathcal{D}$ is flat (dynamically linearizable) at points that do not satisfy $\dim \mathcal{D}^{(i)}(x_0) = \dim \mathcal{D}_i(x_0)$. So our result gives a negative answer to their question (for $x$-flatness). Any Goursat structure can be brought to a generalization of the Goursat normal form, called Kumpera-Ruiz normal form (see [33], [42], [52]). It follows that none of Kumpera-Ruiz normal forms is $x$-flat (except for the regular Kumpera-Ruiz normal form, that is, Goursat normal form). In particular, the system

$$\begin{cases} \dot{x}_1 &= x_5 u_1 \\ \dot{x}_2 &= x_3 x_5 u_1 \\ \dot{x}_3 &= x_4 x_5 u_1 \\ \dot{x}_4 &= u_1 \\ \dot{x}_5 &= u_2 \end{cases}$$

which is historically the first discovered Kumpera-Ruiz normal form [17], is not $x$-flat at any point of its singular locus $\{x \in \mathbb{R}^5 : x = 0\}$. This answers negatively another question of [39].

It is known (see [27], [34], [42], [52]) that the model of $n$-trailer system is a Goursat structure at any configuration point but is equivalent to the chained form out of the singular locus only, that is, if all angles $\theta_{i+1} - \theta_i$ between two consecutive wheels are not $\pm\frac{\pi}{2}$ (except for $\theta_1 - \theta_0$ which is the most far from the top of the train). Therefore our theorem implies that at any singular configuration $\theta_{i+1} - \theta_i = \pm\frac{\pi}{2}$, $i \geq 1$, the $n$-trailer system is not $x$-flat.

The proof of Theorem 1.2.2 is given in Section 1.4.2 and is based on normal forms at singular points (introduced in [33], [42], and [52] and called in the latter the Kumpera-Ruiz normal forms) and on the following result which is of independent interest (as is also proved in Section 1.4.2). It turns out that flat outputs and the chained form $\Sigma_{\rm chain}$ are compatible: in fact, for any given pair of flat outputs $(\varphi_1, \varphi_2)$ of a system feedback equivalent to $\Sigma_{\rm chain}$, we can bring the system to the chained form $\Sigma_{\rm chain}$ for which $\varphi_1$ and $\varphi_2$ serve as the two top variables.

**Proposition 1.2.3** *Consider a driftless 2-input smooth control system $\Sigma$, defined on a manifold $M$ of dimension $n+2$, whose associated distribution $\mathcal{D}$ satisfies* $\operatorname{rank} \mathcal{D}^{(i)} = \operatorname{rank} \mathcal{D}_i = i + 2$, *for $0 \le i \le n$. Given any pair $(\varphi_1, \varphi_2)$ of flat outputs at $(x_0, u_0) \in M \times \mathbb{R}^2$, there exists a feedback transformation $(\Psi, \beta)$ around $x_0$ bringing the system $\Sigma$ into the chained form $\Sigma_{\rm chain}$, given by (1.2.2), such that $\varphi_1 = z_1$ and $\varphi_2 = z_2$.*

# 1.3 Characterization of flat outputs

## 1.3.1 Main Theorems

Recall a useful result due to Cartan [9] whose proof can be found in, e.g., [40] and [51].

**Lemma 1.3.1 (E. Cartan)** *Consider a rank two distribution $\mathcal{D}$ defined on a manifold $M$ of dimension $n + 2$, for $n \ge 2$. If $\mathcal{D}$ satisfies* $\operatorname{rank} \mathcal{D}^{(i)} = i + 2$, *for $0 \le i \le n$, everywhere on $M$, then each distribution $\mathcal{D}^{(i)}$, for $0 \le i \le n - 2$, contains a unique involutive subdistribution $\mathcal{C}_{i+1}$ that is characteristic for $\mathcal{D}^{(i+1)}$ and has constant corank one in $\mathcal{D}^{(i)}$.*

Theorem 1.2.2 implies that the only Goursat structures that are $x$-flat are those equivalent to the chained form (equivalently, whose associated distribution $\mathcal{D}$ is equivalent to the Goursat normal form). For this reason, we will consider in two theorems below such distributions only. Moreover, any distribution equivalent to the Goursat normal form obviously satisfies the assumptions of Lemma 1.3.1 and defines the involutive distribution $\mathcal{C}_{n-1}$ that is characteristic distribution for $\mathcal{D}^{(n-1)}$ and of corank one in $\mathcal{D}^{(n-2)}$.

**Theorem 1.3.2 (Characterization of flat outputs, first version)** *Consider a driftless 2-input smooth control system $\Sigma$ defined on a manifold $M$ of dimension $n+2$ whose associated distribution $\mathcal{D}$ satisfies* $\operatorname{rank} \mathcal{D}^{(i)} = \operatorname{rank} \mathcal{D}_i = i + 2$, *for $0 \le i \le n$. Fix $x_0 \in M$ and let $g$ be an arbitrary vector field in $\mathcal{D}$ such that $g(x_0) \notin \mathcal{C}_{n-1}(x_0)$ and $\varphi_1, \varphi_2$ be two smooth functions defined in a neighborhood $\mathcal{M}$ of $x_0$. Then $(\varphi_1, \varphi_2)$ is an $x$-flat output of $\Sigma$ at $(x_0, u_0)$, $u_0 \notin U_{\rm sing}(x_0)$, if and only if the following conditions hold:*

(i) $\mathrm{d}\varphi_1(x_0) \wedge \mathrm{d}\varphi_2(x_0) \neq 0$, *i.e.,* $\mathrm{d}\varphi_1$ *and* $\mathrm{d}\varphi_2$ *are independent at $x_0$;*

(ii) $L_c\varphi_1 \equiv L_c\varphi_2 \equiv L_c(\dfrac{L_g\varphi_2}{L_g\varphi_1}) \equiv 0$, *for any $c \in \mathcal{C}_{n-1}$, where the functions $\varphi_1$ and $\varphi_2$ are ordered such that $L_g\varphi_1(x_0) \neq 0$ which is always possible due to item* (iii) *below*;

(iii) $(L_g\varphi_1(x_0), L_g\varphi_2(x_0)) \neq (0,0)$;

*Moreover, if a pair of functions* $(\varphi_1, \varphi_2)$ *satisfies* (i) *everywhere in* $\mathcal{M}$ *and forms an* $x$*-flat output at* $(x,u)$ *for any* $x \in \tilde{\mathcal{M}}$ *and certain* $u = u(x)$*, where* $\tilde{\mathcal{M}}$ *is open and dense in* $\mathcal{M}$*, then*

$$Sing(\varphi_1, \varphi_2) = \{x \in \mathcal{M} : (L_g\varphi_1(x), L_g\varphi_2(x)) = (0,0)\}.$$

**Theorem 1.3.3 (Characterization of flat outputs, second version)** *Consider a driftless* 2*-input smooth control system* $\Sigma$ *defined on a manifold* $M$ *of dimension* $n+2$ *whose associated distribution* $\mathcal{D}$ *satisfies* $\operatorname{rank}\mathcal{D}^{(i)} = \operatorname{rank}\mathcal{D}_i = i+2$*, for* $0 \leq i \leq n$*. Fix* $x_0 \in M$ *and let* $\varphi_1, \varphi_2$ *be two smooth functions defined in a neighborhood* $\mathcal{M}$ *of* $x_0$*. Then* $(\varphi_1, \varphi_2)$ *is an* $x$*-flat output of* $\Sigma$ *at* $(x_0, u_0)$*,* $u_0 \notin U_{\mathrm{sing}}(x_0)$*, if and only if the following conditions hold:*

(i)$'$ $\mathrm{d}\varphi_1(x_0) \wedge \mathrm{d}\varphi_2(x_0) \neq 0$*, i.e.,* $\mathrm{d}\varphi_1$ *and* $\mathrm{d}\varphi_2$ *are independent at* $x_0$*;*

(ii)$'$ $\mathcal{L} = (\operatorname{span}\{\mathrm{d}\varphi_1, \mathrm{d}\varphi_2\})^{\perp} \subset \mathcal{D}^{n-1}$ *in* $\mathcal{M}$*;*

(iii)$'$ $\mathcal{D}(x_0)$ *is not contained in* $\mathcal{L}(x_0)$*.*

*Moreover, if a pair of functions* $(\varphi_1, \varphi_2)$ *satisfies* (i)$'$ *everywhere in* $\mathcal{M}$ *and forms an* $x$*-flat output at* $(x,u)$ *for any* $x \in \tilde{\mathcal{M}}$ *and certain* $u = u(x)$*, where* $\tilde{\mathcal{M}}$ *is open and dense in* $\mathcal{M}$*, then*

$$Sing(\varphi_1, \varphi_2) = \{x \in \mathcal{M} : \mathcal{D}(x) \subset \mathcal{L}(x)\}.$$

**Remark 1**. Notice that Theorem 1.3.3 is valid for any $n \geq 1$ (i.e., $\dim M \geq 3$; if $\dim M = 3$, then (iii)$'$ is satisfied automatically) while Theorem 1.3.2 is true for $n \geq 2$ only (i.e., $\dim M \geq 4$). In fact, in Theorem 1.3.2 we use the characteristic distribution $\mathcal{C}_{n-1}$ of $\mathcal{D}^{(n-1)}$ but if $\dim M = 3$, such a distribution does not exist and therefore Theorem 1.3.2 can not applied in that case.

**Remark 2**. The two items (iii) and (iii)$'$ describing the singular locus of an $x$-flat output $(\varphi_1, \varphi_2)$ are equivalent (which will be shown in the proofs of the two theorems) under the condition $\operatorname{rank}\mathcal{D}^{(i)} = \operatorname{rank}\mathcal{D}_i = i+2$, for $0 \leq i \leq n$, i.e.,

$$\{x \in \mathcal{M} : (L_g\varphi_1(x), L_g\varphi_2(x)) = (0,0)\} = \{x \in \mathcal{M} : \mathcal{D}(x) \subset \mathcal{L}(x)\}.$$

**Remark 3**. The conditions of both theorems are verifiable, i.e., given a pair of functions $(\varphi_1, \varphi_2)$ in a neighborhood of a point $x_0$, we can easily verify whether $(\varphi_1, \varphi_2)$ forms an $x$-flat output of a control system under considerations and verification involves derivations and algebraic operations only (without solving PDE's or bringing the system to a normal form). Moreover, the theorems allow us to find the singular locus of a given flat output $(\varphi_1, \varphi_2)$.

A natural question to ask is if there is a lot of pairs $(\varphi_1, \varphi_2)$ which satisfy the conditions of Theorem 1.3.2 or 1.3.3? In other words, is there a lot of pairs $(\varphi_1, \varphi_2)$ which are $x$-flat outputs for a 2-input driftless control system? This question has an elegant answer given by the following theorem. Recall that $\mathcal{C}_{n-1}$ denotes the characteristic distribution of $\mathcal{D}^{(n-1)}$.

**Theorem 1.3.4 (Uniqueness of $x$-flat outputs)** *Consider a driftless 2-input smooth control system $\Sigma$ whose associated distribution $\mathcal{D}$ satisfies* $\operatorname{rank} \mathcal{D}^{(i)} = \operatorname{rank} \mathcal{D}_i = i+2$, *for $0 \le i \le n$, locally around a point $x_0 \in M$, an $(n+2)$ dimensional manifold. Let $g$ be an arbitrary vector field in $\mathcal{D}$ such that $g(x_0) \notin \mathcal{C}_{n-1}(x_0)$. Then for a given arbitrary smooth function $\varphi_1$ such that $L_c\varphi_1 = 0$, for any $c \in \mathcal{C}_{n-1}$, and $L_g\varphi_1(x_0) \neq 0$, there always exists a function $\varphi_2$ such that $(\varphi_1, \varphi_2)$ is an $x$-flat output of $\Sigma$ at $(x_0, u_0)$, $u_0 \notin U_{\mathrm{sing}}(x_0)$. Moreover, if for a given function $\varphi_1$ as above, the pairs $(\varphi_1, \varphi_2)$ and $(\varphi_1, \tilde{\varphi}_2)$ are both $x$-flat outputs of $\Sigma$ at $(x_0, u_0)$, then*

$$\operatorname{span}\{\mathrm{d}\varphi_1, \mathrm{d}\varphi_2\}(x) = \operatorname{span}\{\mathrm{d}\varphi_1, \mathrm{d}\tilde{\varphi}_2\}(x),$$

*for any $x$ in a neighborhood of $x_0$.*

**Remark**. Observe that $x$-flat outputs $(h_1, \ldots, h_m)$ and $(\tilde{h}_1, \ldots, \tilde{h}_m)$ of a system with $m$ controls such that

$$\operatorname{span}\{\mathrm{d}h_1, \ldots, \mathrm{d}h_m\} = \operatorname{span}\{\mathrm{d}\tilde{h}_1, \ldots, \mathrm{d}\tilde{h}_m\}$$

can be considered as statically equivalent. Indeed, in that case there exist smooth functions $H_i$ and $\tilde{H}_i$ of $m$ variables such that $h_i = H_i(\tilde{h}_1, \ldots, \tilde{h}_m)$ and $\tilde{h}_i = \tilde{H}_i(h_1, \ldots, h_m)$. It thus follows from Theorem 1.3.4 that for a given arbitrary $\varphi_1$ (satisfying the assumptions of the theorem), the choice of $\varphi_2$ is unique in the sense that all functions $\varphi_2$ giving $x$-flat outputs $(\varphi_1, \varphi_2)$ yield, actually, statically equivalent $x$-flat outputs.

### 1.3.2 Finding $x$-flat outputs

The importance of Theorem 1.3.2 is that it not only allows to check whether a given pair of functions forms an $x$-flat output but also, together with Theorem 1.3.4, to express explicitly a system of $1^{\text{st}}$ order PDE's to be solved in order to calculate all $x$-flat outputs for a given 2-input driftless system.

To this aim, choose $n-1$ vector fields $c_1, \ldots, c_{n-1}$ spanning the characteristic distribution $\mathcal{C}_{n-1}$ of $\mathcal{D}^{(n-1)}$. Recall that $\mathcal{C}_{n-1}$ can be easily calculated as (see, e.g., [8])

$$\mathcal{C}_{n-1} = \{f \in \mathcal{D}^{(n-1)} : f \lrcorner \mathrm{d}\omega \in (\mathcal{D}^{(n-1)})^{\perp}\},$$

where $\omega$ is any non-zero differential 1-form annihilating $\mathcal{D}^{(n-1)}$. Fix any vector field $g$ of $\mathcal{D}$, independent at $x_0$ with the characteristic distribution $\mathcal{C}_{n-1}$ of $\mathcal{D}^{(1)}$. According to Theorem 1.3.2 and Theorem 1.3.4, in order to find $\varphi_1$, we have to solve the following system of $1^{\text{st}}$ order PDE's

$$\begin{aligned} L_{c_i}\varphi_1 &= 0, \quad 1 \le i \le n-1, \\ L_g\varphi_1(x_0) &\neq 0. \end{aligned}$$

The above system possesses solutions (since $\mathcal{C}_{n-1}$ is involutive) and the space of solutions is that of functions of three variables (since corank $\mathcal{C}_{n-1} \subset TM) = 3$).

Now we will establish a system of equations for $\varphi_2$. According to Theorem 1.3.2, it is given by

$$\begin{aligned} L_{c_i}\varphi_2 &= 0, \quad 1 \le i \le n-1 \\ L_{c_i}\left(\frac{L_g\varphi_2}{L_g\varphi_1}\right) &= 0, \quad 1 \le i \le n-1. \end{aligned}$$

The last $n-1$ equations are equivalent to

$$(L_g\varphi_1)L_{c_i}L_g\varphi_2 - (L_g\varphi_2)L_{c_i}L_g\varphi_1 = 0, \ 1 \le i \le n-1.$$

Applying $L_{[c_i,g]}\psi = L_{c_i}L_g\psi - L_gL_{c_i}\psi$ and taking into account that $L_{c_i}\varphi_1 = L_{c_i}\varphi_2 = 0$, we get

$$(L_g\varphi_1)L_{[c_i,g]}\varphi_2 - (L_{[c_i,g]}\varphi_1)L_g\varphi_2 = 0,$$

which we rewrite as

$$L_{v_i}\varphi_2 = 0, \quad 1 \le i \le n-1,$$

where the vector fields $v_1, \ldots, v_{n-1}$ are given by

$$v_i = (L_g\varphi_1)[c_i, g] - (L_{[c_i,g]}\varphi_1)g, \quad 1 \le i \le n-1.$$

We want to emphasize that the vector fields $v_i$ are easily calculable in terms of the vector fields $g, c_1, \ldots, c_{n-1}$ and the chosen solution $\varphi_1$. So finally, we have to solve the system $L_{c_i}\varphi_2 = L_{v_i}\varphi_2 = 0$ which, surprisingly, consists of $2(n-1)$ $1^{\text{st}}$ order PDE's on an $(n+2)$-dimensional manifold. We will show below that this system reduces, actually, to $n$ equations.

### 1.3.3 Reducing equations for $\varphi_2$

By Cartan's Lemma 1.3.1 we have $\mathcal{C}_1 \subset \mathcal{C}_2 \subset \cdots \subset \mathcal{C}_{n-1}$ and $\operatorname{corank}(\mathcal{C}_i \subset \mathcal{D}^{(i-1)}) = 1$, for $1 \le i \le n-1$. Thus we can always choose vector fields $c_1, \ldots, c_{n-1}$ such that $\mathcal{C}_{n-1} = \operatorname{span}\{c_1, \cdots, c_{n-1}\}$ and $c_{n-1}(x_0) \notin \mathcal{C}_{n-2}(x_0)$, i.e., $\mathcal{C}_{n-2} = \operatorname{span}\{c_1, \cdots, c_{n-2}\}$. Fix a vector field $g \in \mathcal{D}$ such that $g(x_0) \notin \mathcal{C}_{n-1}(x_0)$ and clearly we have

$$\mathcal{D}^{(n-2)} = \operatorname{span}\{c_1, \cdots, c_{n-2}, c_{n-1}, g\}.$$

We claim that the system

$$\begin{aligned} L_{c_i}\varphi_2 &= 0, \quad 1 \le i \le n-1 \\ L_{v_i}\varphi_2 &= 0, \quad 1 \le i \le n-1, \end{aligned} \tag{1.3.1}$$

where $v_i = (L_g\varphi_1)[c_i, g] - (L_{[c_i, g]}\varphi_1)g$, can be reduced to a system of $n$ equations. Since $\mathcal{C}_{n-2}$ is the characteristic distribution of $\mathcal{D}^{(n-2)}$, we get

$$[c_i, g] = \beta_i g \mod \mathcal{C}_{n-1}, \quad 1 \le i \le n-2 \tag{1.3.2}$$

where $\beta_i$, for $1 \le i \le n-2$, are smooth functions defined in a neighborhood of $x_0$. The equation (1.3.2) implies that (recall that $L_c\varphi_1 = 0$, for any $c \in \mathcal{C}_{n-1}$)

$$L_{[c_i, g]}\varphi_1 = L_{\beta_i g}\varphi_1 = \beta_i L_g\varphi_1. \tag{1.3.3}$$

Therefore, for $1 \le i \le n-2$,

$$\begin{aligned} v_i &= (L_g\varphi_1)[c_i, g] - (L_{[c_i,g]}\varphi_1)g \\ &= (L_g\varphi_1)\beta_i g - \beta_i(L_g\varphi_1)g \mod \mathcal{C}_{n-1} \\ &= 0 \mod \mathcal{C}_{n-1}. \end{aligned}$$

It follows that the equations $L_{c_i}\varphi_2 = 0$, for $c_i \in \mathcal{C}_{n-1}$, imply $L_{v_i}\varphi_2 = 0$, $1 \le i \le n-2$, and therefore the system (1.3.1) is equivalent to the following system of $n$ equations

$$\begin{aligned} L_{c_i}\varphi_2 &= 0, \quad 1 \le i \le n-1 \\ L_{v_{n-1}}\varphi_2 &= 0, \end{aligned} \tag{1.3.4}$$

where $v_{n-1} = (L_g\varphi_1)[c_{n-1}, g] - (L_{[c_{n-1},g]}\varphi_1)g$. Notice that $\varphi_1$ solves the system (1.3.4). We are looking for a solution $\varphi_2$ of (1.3.4), independent with $\varphi_1$, and by Frobenious theorem the system (1.3.4) possesses two independent solutions if and only if the distribution $\mathcal{L} = \text{span}\{c_1, \ldots, c_{n-1}, v_{n-1}\} = \mathcal{C}_{n-1} \oplus \text{span}\{v_{n-1}\}$ is involutive.

Below we will show that $\mathcal{L}$ is, indeed, involutive and to this end it is sufficient to show that $[c_i, v_{n-1}] \in \mathcal{L}$, for any $1 \le i \le n-1$. Since $\mathcal{D}^{(n-2)} = \text{span}\{c_1, \ldots, c_{n-1}, g\}$, we have $\mathcal{D}^{(n-1)} = \mathcal{D}^{(n-2)} + [\mathcal{D}^{(n-2)}, \mathcal{D}^{(n-2)}] = \text{span}\{c_1, \ldots, c_{n-1}, g, [c_{n-1}, g]\}$. The fact that $\mathcal{C}_{n-1} = \text{span}\{c_1, \ldots, c_{n-1}\}$ is the characteristic distribution of $\mathcal{D}^{(n-1)}$ implies that

$$[c_i, [c_{n-1}, g]] = \xi_i g + \tau_i[c_{n-1}, g] \mod \mathcal{C}_{n-1}, \quad 1 \le i \le n-1, \tag{1.3.5}$$

where $\xi_i$, $\tau_i$ are smooth functions defined in a neighborhood of $x_0$, and thus

$$L_{[c_i,[c_{n-1},g]]}\varphi_1 = \xi_i L_g\varphi_1 + \tau_i L_{[c_{n-1},g]}\varphi_1, \quad 1 \le i \le n-1. \tag{1.3.6}$$

Observing that

$$\begin{aligned} L_{[c_i,g]}\varphi_1 &= L_{c_i}L_g\varphi_1 - L_gL_{c_i}\varphi_1 = L_{c_i}L_g\varphi_1 \\ L_{[c_i,[c_{n-1},g]]}\varphi_1 &= L_{c_i}L_{[c_{n-1},g]}\varphi_1 - L_{[c_{n-1},g]}L_{c_i}\varphi_1 = L_{c_i}L_{[c_{n-1},g]}\varphi_1 \end{aligned} \tag{1.3.7}$$

and applying the relations (1.3.2)-(1.3.6), we have, for $1 \le i \le n-2$,

$$\begin{aligned}
[c_i, v_{n-1}] &= [c_i, (L_g\varphi_1)[c_{n-1}, g]] - [c_i, (L_{[c_{n-1},g]}\varphi_1)g] \\
&= (L_g\varphi_1)[c_i, [c_{n-1}, g]] + (L_{c_i}L_g\varphi_1)[c_{n-1}, g] - (L_{[c_{n-1},g]}\varphi_1)[c_i, g] - (L_{c_i}L_{[c_{n-1},g]}\varphi_1)g \\
&= L_g\varphi_1\,(\xi_i g + \tau_i[c_{n-1},\, g]) + \beta_i(L_g\varphi_1)[c_{n-1}, g] \\
&\quad -(L_{[c_{n-1},g]}\varphi_1)\beta_i g - \big(\xi_i(L_g\varphi_1)g + \tau_i(L_{[c_{n-1},\,g]}\varphi_1)g\big) \mod \mathcal{C}_{n-1} \\
&= (\tau_i + \beta_i)v_{n-1} \mod \mathcal{C}_{n-1}
\end{aligned}$$

which implies that $[c_i, v_{n-1}] \in \mathcal{L}$ , for $1 \le i \le n-2$. Now consider the case $i = n-1$, that is, the Lie bracket $[c_{n-1}, v_{n-1}]$. Applying (1.3.5) and (1.3.6) for $i = n-1$, we get

$$\begin{aligned}
[c_{n-1}, v_{n-1}] &= [c_{n-1}, (L_g\varphi_1)[c_{n-1}, g]] - [c_{n-1}, (L_{[c_{n-1},g]}\varphi_1)g] \\
&= (L_g\varphi_1)[c_{n-1}, [c_{n-1}, g]] + (L_{c_{n-1}}L_g\varphi_1)[c_{n-1}, g] \\
&\quad -(L_{[c_{n-1},g]}\varphi_1)[c_{n-1}, g] - (L_{c_{n-1}}L_{[c_{n-1},g]}\varphi_1)g \\
&= (L_g\varphi_1)[c_{n-1}, [c_{n-1}, g]] - (L_{c_{n-1}}L_{[c_{n-1},g]}\varphi_1)g \\
&= \xi_{n-1}(L_g\varphi_1)g + \tau_{n-1}(L_g\varphi_1)[c_{n-1}, g] \\
&\quad -\big(\xi_{n-1}(L_g\varphi_1)g + \tau_{n-1}(L_{[c_{n-1},g]})g\big) \mod \mathcal{C}_{n-1} \\
&= \tau_{n-1}v_{n-1} \mod \mathcal{C}_{n-1},
\end{aligned}$$

which implies that $[c_{n-1}, v_{n-1}] \in \mathcal{L}$. In conclusion, the distribution $\mathcal{L}$ is involutive and hence the system (1.3.4) is solvable. Together with the analysis of Section 1.3.2, we get the following theorem.

**Theorem 1.3.5** *Assume that a control system $\Sigma$ is $x$-flat at $(x_0, u_0)$, $u_0 \notin U_{\rm sing}(x_0)$, that is, the associated distribution $\mathcal{D}$ is, locally at $x_0$, equivalent to the Goursat normal form on an $(n+2)$-dimensional manifold $M$. Let $\mathcal{C}_{n-1} = \operatorname{span}\{c_1, \ldots, c_{n-1}\}$ be the characteristic distribution of $\mathcal{D}^{(n-1)}$ such that $c_{n-1}(x_0) \notin \mathcal{C}_{n-2}(x_0)$ and $g$ any vector field in $\mathcal{D}$ such that $g(x_0) \notin \mathcal{C}_{n-1}(x_0)$. Then*

(i) *For any smooth function $\varphi_1$ such that*

$$\text{(Flat 1)} \qquad \begin{aligned} L_{c_i}\varphi_1 &= 0, \quad 1 \le i \le n-1, \\ L_g\varphi_1(x_0) &\ne 0, \end{aligned}$$

*the distribution $\mathcal{L} = \operatorname{span}\{c_1, \ldots, c_{n-1}, v\}$ is involutive, where*

$$v = (L_g\varphi_1)[c_{n-1}, g] - (L_{[c_{n-1},g]}\varphi_1)g.$$

(ii) *A pair of functions $(\varphi_1, \varphi_2)$ forms an $x$-flat output of $\Sigma$ at $(x_0, u_0)$, $u_0 \notin U_{\rm sing}(x_0)$, if and only if after a permutation (if necessary) $\varphi_1$ satisfies* (Flat 1), $d\varphi_1(x_0) \wedge d\varphi_2(x_0) \ne 0$, *and $\varphi_2$ satisfies*

$$\text{(Flat 2)} \qquad \begin{aligned} L_{c_i}\varphi_2 &= 0, \quad 1 \le i \le n-1, \\ L_v\varphi_2 &= 0. \end{aligned}$$

**Remark**. In (ii) only one implication may need permuting $\varphi_1$ and $\varphi_2$. Indeed, if $(\varphi_1, \varphi_2)$ satisfies (Flat 1) and (Flat 2), then it is an $x$-flat output (and no permutation is needed). If $(\varphi_1, \varphi_2)$ is an $x$-flat output, then at least one $\varphi_i$, $1 \le i \le 2$, satisfies $L_g\varphi_i(x_0) \neq 0$ and we choose $\varphi_1$ such that $L_g\varphi_1(x_0) \neq 0$.

**Example 1.3.6** To illustrate the above-presented procedure of finding flat outputs, we will consider the case of 2-input system

$$\dot{x} = f_1(x)u_1 + f_2(x)u_2$$

on a 4-dimensional manifold $M$. Assume that the system is $x$-flat, that is, the associated distributions $\mathcal{D} = \operatorname{span}\{f_1, f_2\}$ satisfies the conditions of Theorem 1.3.5. Choose a vector field $c \in \mathcal{C}_1$ characteristic for $\mathcal{D}^{(1)}$ and $g \in \mathcal{D}$ such that $g(x_0) \wedge c(x_0) \neq 0$.

According to above procedure we take as $\varphi_1$ an arbitrary solution of

$$\begin{aligned} L_c\varphi_1 &= 0 \\ L_g\varphi_1(x_0) &\neq 0 \end{aligned}$$

and, in order to find $\varphi_2$, we have to solve

$$\begin{aligned} L_c\varphi_2 &= 0 \\ L_v\varphi_2 &= 0 \end{aligned}$$

where $v = (L_g\varphi_1)[c, g] - (L_{[c,\,g]}\varphi_1)g$. Notice that the above system of three 1$^{\text{st}}$ order PDE's contains a fourth one; indeed we have

$$L_v\varphi_1 = (L_g\varphi_1)L_{[c,g]}\varphi_1 - (L_{[c,g]}\varphi_1)L_g\varphi_1 = 0.$$

The system

$$L_c\varphi_i = L_v\varphi_i = 0, \quad 1 \le i \le 2, \tag{1.3.8}$$

admits two independent functions $\varphi_1$ and $\varphi_2$ as solutions if and only if the distribution $\operatorname{span}\{c, v\}$ is integrable. A direct calculation shows that this is the case. All becomes clear: the involutive distribution $\operatorname{span}\{c, v\}$ is just the distribution $\mathcal{L}$ of Theorem 1.3.3 while $\varphi_1$ and $\varphi_2$ satisfying (1.3.8) are $x$-flat outputs since their differentials span $\mathcal{L}^\perp$. We also see that $\mathcal{L}$ is not unique: different choices of $\varphi_1$ lead to different vector fields $v$ which, in turn, give different distributions $\mathcal{L} = \operatorname{span}\{c, v\}$, although all of them are involutive and thus define (via $\operatorname{span}\{\mathrm{d}\varphi_1, \mathrm{d}\varphi_2\} = \mathcal{L}^\perp$) non equivalent flat outputs. This is in a perfect accordance with Theorem 1.3.4.

**Example 1.3.7** (Vertical rolling disk) Consider a vertical disk of radius $R$ rolling without slipping on a horizontal plane. Denote by $(x, y)$ the position of the contact point in the $xy$-plane, and by $\theta$ and $\phi$, respectively, the rotation angle of the disk and

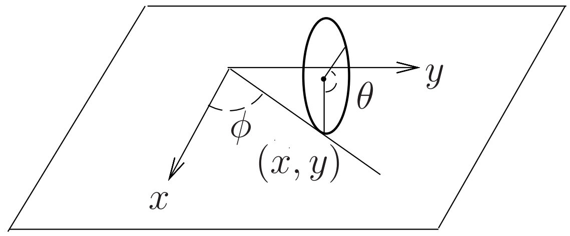


Figure 1.2: the rolling disk

the orientation of the disk. The controls $u_1$ and $u_2$ allow the disk to rotate and turn. This leads to the following model given by a driftless system on $Q = \mathbb{R}^2 \times S^1 \times S^1$:

$$\Sigma_{\text{disk}} : \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{\phi} \end{pmatrix} = \begin{pmatrix} R\cos\phi \\ R\sin\phi \\ 1 \\ 0 \end{pmatrix} u_1 + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} u_2 = f_1 u_1 + f_2 u_2.$$

A direct computation shows that $\operatorname{rank}\mathcal{D}^{(i)} = \operatorname{rank}\mathcal{D}_i = i+2$, for $0 \le i \le 2$, and $\mathcal{C}_1 = \operatorname{span}\{f_1\}$. Therefore by Theorem 1.2.2, the model $\Sigma_{\text{disk}}$ is $x$-flat at any point of its configuration space $Q$. Moreover, it satisfies the hypothesis of Theorem 1.3.2, 1.3.3 and 1.3.4 and $U_{\text{sing}}$ is given by $U_{\text{sing}} = \{u = (u_1, u_2)^\top : u_2 = 0\}$. Thus the singular control corresponds to rolling the disk along a straight line. Now let us calculate all its $x$-flat outputs by using the procedure given in the Section 1.3.2. We choose

$$c = f_1 = R\cos\phi\frac{\partial}{\partial x} + R\sin\phi\frac{\partial}{\partial y} + \frac{\partial}{\partial\theta},$$

and take $g = f_2 = \frac{\partial}{\partial\phi}$. Then as a first flat output we can take any function $\varphi_1$ satisfying the following system of equations

$$\begin{aligned} L_c\varphi_1 &= R\cos\phi\frac{\partial\varphi_1}{\partial x} + R\sin\phi\frac{\partial\varphi_1}{\partial y} + \frac{\partial\varphi_1}{\partial\theta} \equiv 0 \\ L_g\varphi_1(q_0) &\neq 0. \end{aligned}$$

Solving this system of equations, we get that $\varphi_1$ is any function of the form

$$\varphi_1 = \varphi_1(\phi, x - R\theta\cos\phi, y - R\theta\sin\phi)$$

satisfying $\frac{\partial\varphi_1}{\partial\phi}(q_0) \neq 0$. Choose one such $\varphi_1$ and then $\varphi_2$ is any function independent with $\varphi_1$ that satisfies $L_c\varphi_2 = L_v\varphi_2 = 0$, where the vector field $v$ is given by

$$v = (L_g\varphi_1)[c, g] - (L_{[c,g]}\varphi_1)g.$$

To illustrate this, choose the function $\varphi_1 = x - R\theta\cos\phi$ around a point $q_0$ such that $L_g\varphi_1(q_0) = R\theta\sin\phi \neq 0$ and then

$$v = R^2\theta\sin^2\phi\frac{\partial}{\partial x} - R^2\theta\sin\phi\cos\phi\frac{\partial}{\partial y} - R\sin\phi\frac{\partial}{\partial\phi}.$$

Solving the system of equations $L_c\varphi_2 = L_v\varphi_2 = 0$, we get

$$\varphi_2 = \varphi_2(x - R\theta\cos\phi, y - R\theta\sin\phi)$$

satisfying $(\mathrm{d}\varphi_1 \wedge \mathrm{d}\varphi_2)(q_0) \neq 0$. All such functions satisfy $\mathrm{span}\,\{\mathrm{d}\varphi_1, \mathrm{d}\varphi_2\} = \mathrm{span}\,\{\mathrm{d}\varphi_1, \mathrm{d}\tilde{\varphi}_2\}$ and we can take, for instance, $\varphi_2 = y - R\theta\sin\phi$. Moreover, the singular locus of the $x$-flat output $(x - R\theta\cos\phi,\ y - R\theta\sin\phi)$ is given by

$$\begin{aligned} Sing(\varphi_1, \varphi_2) &= \{q \in Q : (L_g\varphi_1(q), L_g\varphi_2(q)) = (0,0)\} \\ &= \{q \in Q : \theta = 0\}. \end{aligned}$$

To see that $\sin\phi = 0$ is, indeed, not a singularity, we just permute $\varphi_1$ and $\varphi_2$.

To consider another possibility, we choose $\varphi_1 = \phi$ and then we have

$$v = R\sin\phi\frac{\partial}{\partial x} - R\cos\phi\frac{\partial}{\partial y}.$$

Solving the system of equations $L_c\varphi_2 = L_v\varphi_2 = 0$, we get

$$\varphi_2 = \varphi_2(\phi, R\theta - x\cos\phi - y\sin\phi),$$

satisfying $(\mathrm{d}\varphi_1 \wedge \mathrm{d}\varphi_2)(q_0) \neq 0$. We can take, for instance, $\varphi_2 = R\theta - x\cos\phi - y\sin\phi$ and a simple calculation shows that there does not exist singular point of the $x$-flat output $(\phi,\ R\theta - x\cos\phi - y\sin\phi)$ in the state space $Q$. In other words, $(\phi,\ R\theta - x\cos\phi - y\sin\phi)$ is an $x$-flat output at any point $(q,u) \in Q \times \mathbb{R}^2$ provided that $u \notin U_{\mathrm{sing}}(q)$.

For various choices of functions, our result allows to eliminate them as candidates for $x$-flat outputs. For example, we can conclude that if $(\varphi_1, \varphi_2)$ is an $x$-flat output, then $L_c\varphi_i \equiv 0$, for $i = 1, 2$, where $c = R\cos\phi\frac{\partial}{\partial x} + R\sin\phi\frac{\partial}{\partial y} + \frac{\partial}{\partial\theta}$ is a characteristic vector field of $\mathcal{D}^{(1)}$. It follows that independently of the choice of $\varphi_2$, neither $(x, \varphi_2)$, nor $(y, \varphi_2)$, nor $(\theta, \varphi_2)$ can serve as an $x$-flat output.

### 1.3.4 A complete description of $x$-flat outputs for the nonholonomic car system

Come back to the example of the nonholonomic car $\Sigma_{\mathrm{car}}$ that we analyzed in Section 1.1.

$$\Sigma_{\mathrm{car}} : \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta}_0 \\ \dot{\theta}_1 \end{pmatrix} = \begin{pmatrix} \cos(\theta_1 - \theta_0)\cos\theta_0 \\ \cos(\theta_1 - \theta_0)\sin\theta_0 \\ \sin(\theta_1 - \theta_0) \\ 0 \end{pmatrix} u_1 + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} u_2 \qquad ,$$

We choose as a characteristic vector field $c = \dfrac{\partial}{\partial \theta_1}$ and take

$$g = \cos(\theta_1 - \theta_0)\cos\theta_0 \frac{\partial}{\partial x} + \cos(\theta_1 - \theta_0)\sin\theta_0 \frac{\partial}{\partial y} + \sin(\theta_1 - \theta_0)\frac{\partial}{\partial \theta_0}.$$

As a first $x$-flat output we can take any function $\varphi_1$ satisfying $L_c\varphi_1 = \dfrac{\partial \varphi_1}{\partial \theta_1} \equiv 0$ and $L_g\varphi_1(q) \neq 0$, that is any function $\varphi_1 = \varphi_1(x, y, \theta_0)$ such that $L_g\varphi_1(q) \neq 0$. Let us choose one such $\varphi_1$ then $\varphi_2$ satisfies $L_c\varphi_2 = L_v\varphi_2 = 0$, where the vector field $v$ is given by

$$\begin{aligned} v &= (L_g\varphi_1)[c, g] - (L_{[c,g]}\varphi_1)g \\ &= -\frac{\partial \varphi_1}{\partial \theta_0}\cos\theta_0 \frac{\partial}{\partial x} - \frac{\partial \varphi_1}{\partial \theta_0}\sin\theta_0 \frac{\partial}{\partial y} + (\frac{\partial \varphi_1}{\partial x}\cos\theta_0 + \frac{\partial \varphi_1}{\partial y}\sin\theta_0)\frac{\partial}{\partial \theta_0}. \end{aligned}$$

Therefore $\varphi_2$ can be taken as any functions $\varphi_2(x, y, \theta_0)$ satisfying $L_v\varphi_2 = 0$ and $(\mathrm{d}\varphi_1 \wedge \mathrm{d}\varphi_2)(q) \neq 0$. Given $\varphi_1$ as above, the space of solutions for $\varphi_2$ is thus parameterized by one function of two variables but any two solutions $\varphi_2$ and $\tilde{\varphi}_2$ give statically equivalent flat outputs, that is $\mathrm{span}\,\{\mathrm{d}\varphi_1, \mathrm{d}\varphi_2\} = \mathrm{span}\,\{\mathrm{d}\varphi_1, \mathrm{d}\tilde{\varphi}_2\}$. On the other hand, different choices of $\varphi_1$ will lead to nonequivalent pairs $(\varphi_1, \varphi_2)$ of $x$-flat outputs.

To illustrate this, take $\varphi_1 = x$, then $v = \cos\theta_0 \dfrac{\partial}{\partial \theta_0}$ and $L_c\varphi_2 = L_v\varphi_2 = 0$ imply that $\varphi_2$ is any function of the form $\varphi_2(x, y)$ satisfying $\dfrac{\partial \varphi_2}{\partial y}(q) \neq 0$ (because of $(\mathrm{d}\varphi_1 \wedge \mathrm{d}\varphi_2)(q) \neq 0$). All such functions satisfy $\mathrm{span}\,\{\mathrm{d}x, \mathrm{d}\varphi_2\} = \mathrm{span}\,\{\mathrm{d}x, \mathrm{d}\tilde{\varphi}_2\}$ and we can take, for instance, $\varphi_2 = y$. This gives the well-known $x$-flat output $(x, y)$.

To see another choice, take $\varphi_1 = \theta_0$, then $v = -\cos\theta_0 \dfrac{\partial}{\partial x} - \sin\theta_0 \dfrac{\partial}{\partial y}$ and the general solution of $L_c\varphi_2 = L_v\varphi_2 = 0$ is $\varphi_2 = \varphi_2(\theta_0, x\sin\theta_0 - y\cos\theta_0)$, which gives as an $x$-flat output $(\theta_0, x\sin\theta_0 - y\cos\theta_0)$. Notice that the singular loci of the two choices of $x$-flat outputs are different. In fact, $Sing(x, y) = \{\theta_1 - \theta_0 = \pm\frac{\pi}{2}\}$ and $Sing(\theta_0, x\sin\theta_0 - y\cos\theta_0) = \{\theta_1 - \theta_0 = 0, \pm\pi\}$.

Now take $\varphi_1 = x + \theta_0$ around $\cos\theta_0 \neq 0$, then $v = -\cos\theta_0 \dfrac{\partial}{\partial x} - \sin\theta_0 \dfrac{\partial}{\partial y} + \cos\theta_0 \dfrac{\partial}{\partial \theta_0}$. Thus the general solution of $L_c\varphi_2 = L_v\varphi_2 = 0$ is $\varphi_2 = \varphi_2(x + \theta_0, y - \ln|\cos\theta_0|)$. We can take, for instance, $\varphi_2 = y - \ln|\cos\theta_0|$ which gives a third $x$-flat output $(x + \theta_0, y - \ln|\cos\theta_0|)$ of $\Sigma_{\mathrm{car}}$ and its singular locus is defined by

$$Sing(x + \theta_0, y - \ln|\cos\theta_0|) = \{\cos\theta_0 = 0\} \cup \{\cos(\theta_1 - \theta_0)\cos\theta_0 + \sin(\theta_1 - \theta_0) = 0\}.$$

### 1.3.5 A complete description of $x$-flat outputs for the nonholonomic $n$-trailer system

Consider the kinematic model for a unicycle-like mobile robot towing $n$ trailers such that the towing hook of each trailer is located at the center of its unique axle (with the assumption that the distances between any two consecutive trailers are equal). The $n$-trailer system is subject to nonholonomic constraints: it is assumed that the wheels of each individual trailer are aligned with the body and are not allowed to slip [34] . This model and its control properties have attracted a lot of attention (see the books [35] and [36]; and the papers [13], [27], [46], [52], [73]). We use here the following description introduced in [52].

Consider the $n$-trailer system $\Sigma_{\rm tr}^{\rm n}$ defined on $\mathbb{R}^2 \times (S^1)^{n+1}$, for $n \geq 0$,

$$\Sigma_{\rm tr}^{\rm n}: \quad \dot{q} = f_1(q)u_1 + f_2(q)u_2, \quad q \in \mathbb{R}^2 \times (S^1)^{n+1},$$

where the vector fields $f_1$ and $f_2$ are given by

$$\begin{aligned} f_1 &= \pi_0 \cos\theta_0 \frac{\partial}{\partial x} + \pi_0 \sin\theta_0 \frac{\partial}{\partial y} + \sum_{i=0}^{n-1} \pi_{i+1} \sin(\theta_{i+1} - \theta_i) \frac{\partial}{\partial \theta_i} \\ f_2 &= \frac{\partial}{\partial \theta_n} \end{aligned}$$

with $\pi_i = \prod_{j=i+1}^{n} \cos(\theta_j - \theta_{j-1})$ and $\pi_n = 1$. The configuration of this system is described by $q = (x, y, \theta_0, \ldots, \theta_n) \in \mathbb{R}^2 \times (S^1)^{n+1}$, where $(x, y)$ denotes the position of the last trailer while $\theta_0, \ldots, \theta_n$ represent the angles between each trailer's axle and the $x$-axis. According to Theorem 1.2.2, the $n$-trailer system is locally $x$-flat at any $(q_0, u_0)$ such that $\dim \mathcal{D}^{(i)}(q_0) = \dim \mathcal{D}_i(q_0)$, $0 \leq i \leq n$ (equivalently $\mathcal{D}$ is equivalent to the Goursat normal form at $q_0$) and $u_0 \notin U_{\rm sing}(q_0)$. The former condition yields $\cos(\theta_{i,0} - \theta_{i-1,0}) \neq 0$ and the latter means that $u_0 = (u_{10}, u_{20})$ satisfies $u_{10} \neq 0$. In other words, the $n$-trailer is $x$-flat along a trajectory $q(t)$ besides those time instances $t_0$ at which the angle between two consecutive trailers becomes $\pm\frac{\pi}{2}$ (besides the angle between the last trailer and one before the last which can be any) or those instance at which the velocity $\dot{q}(t_0)$ becomes tangent to $\frac{\partial}{\partial \theta_n}$, that is, the whole $n$-trailer movement stops.

Let $\mathcal{D}$ be the distribution associated to the $n$-trailer system $\Sigma_{\rm tr}^{\rm n}$, that is $\mathcal{D} = \operatorname{span}\{f_1, f_2\}$. It is easy to check that $\operatorname{rank} \mathcal{D}^{(i)} = i + 2$, for $0 \leq i \leq n + 1$ and that, see e.g., Lemma 4.10 in [52], the characteristic distribution $\mathcal{C}_i$ of $\mathcal{D}^{(i)}$, for $1 \leq i \leq n$, is given by

$$\mathcal{C}_i = \operatorname{span}\{c_1, \ldots, c_i\} = \operatorname{span}\{\frac{\partial}{\partial \theta_n}, \ldots, \frac{\partial}{\partial \theta_{n-i+1}}\}.$$

Taking

$$g = f_1 = \pi_0 \cos\theta_0 \frac{\partial}{\partial x} + \pi_0 \sin\theta_0 \frac{\partial}{\partial y} + \sum_{i=0}^{n-1} \pi_{i+1} \sin(\theta_{i+1} - \theta_i) \frac{\partial}{\partial \theta_i},$$

we have clearly $\mathcal{D}^{(i)} = \operatorname{span}\{c_1, \ldots, c_{i+1}, g\}$, for $1 \leq i \leq n - 1$, around any point $q_0$ such that $\cos(\theta_{i,0} - \theta_{i-1,0}) \neq 0$. According to the analysis performed in Section 1.3.2, and summarized in Theorem 1.3.5, as a first flat output around a given point $q_0 \in \mathbb{R}^2 \times (S^1)^{n+1}$, we can take any function $\varphi_1$ satisfying $L_{c_i}\varphi_1 = \frac{\partial \varphi_1}{\partial \theta_i} \equiv 0$, for

$1 \le i \le n$, and $L_g\varphi_1(q_0) \neq 0$, that is any function $\varphi_1 = \varphi_1(x, y, \theta_0)$ such that $L_g\varphi_1(q_0) \neq 0$. Let us choose one such $\varphi_1$ and then $\varphi_2$ has to satisfy

$$\begin{aligned} L_{c_i}\varphi_2 &= 0, \quad 1 \le i \le n \\ L_{v_n}\varphi_2 &= 0, \end{aligned} \tag{1.3.9}$$

where $v_n = (L_g\varphi_1)[c_n, g] - (L_{[c_n,g]}\varphi_1)g$ (notice that the dimension of the state space is $n+3$). The conditions $L_{c_i}\varphi_2 = 0$, $1 \le i \le n$, imply that $\varphi_2 = \varphi_2(x, y, \theta_0)$ and now we consider the equation $L_{v_n}\varphi_2 = 0$. We have

$$g = \eta\tilde{g} + \sum_{i=1}^{n-1} \pi_{i+1} \sin(\theta_{i+1} - \theta_i) c_{n-i+1},$$

where $\eta = \prod_{i=1}^{n-1} \cos(\theta_{i+1} - \theta_i)$ and

$$\tilde{g} = \cos(\theta_1 - \theta_0)\cos\theta_0 \frac{\partial}{\partial x} + \cos(\theta_1 - \theta_0)\sin\theta_0 \frac{\partial}{\partial y} + \sin(\theta_1 - \theta_0)\frac{\partial}{\partial \theta_0}.$$

Recall that around $q_0$ under consideration $\theta_{i+1,0} - \theta_{i,0} \neq \pm\frac{\pi}{2}$. So $\eta \neq 0$ and by a direct calculation we check that $L_{v_n}\varphi_2 = 0$ if and only if $L_{\tilde{v}_n}\varphi_2 = 0$, where

$$\begin{aligned} \tilde{v}_n &= (L_{\tilde{g}}\varphi_1)[c_n, \tilde{g}] - (L_{[c_n,\tilde{g}]}\varphi_1)\tilde{g} \\ &= -\frac{\partial \varphi_1}{\partial \theta_0}\cos\theta_0 \frac{\partial}{\partial x} - \frac{\partial \varphi_1}{\partial \theta_0}\sin\theta_0 \frac{\partial}{\partial y} + \left(\frac{\partial \varphi_1}{\partial x}\cos\theta_0 + \frac{\partial \varphi_1}{\partial y}\sin\theta_0\right)\frac{\partial}{\partial \theta_0}. \end{aligned}$$

Given any $\varphi_1 = \varphi_1(x, y, \theta_0)$ such that $L_g\varphi_1(q_0) \neq 0$, the space of solution of $L_{\tilde{v}_n}\varphi_2 = 0$ is clearly parameterized by one function of two variables. Moreover any two solutions $\varphi_2$ and $\tilde{\varphi}_2$ give equivalent $x$-flat outputs $\mathrm{span}\,\{\mathrm{d}\varphi_1, \mathrm{d}\varphi_2\} = \mathrm{span}\,\{\mathrm{d}\varphi_1, \mathrm{d}\tilde{\varphi}_2\} = \mathcal{L}^\perp$, where $\mathcal{L} = \mathrm{span}\,\{c_1, \ldots, c_n, v\} = \mathrm{span}\,\{c_1, \ldots, c_n, \tilde{v}\}$ (as we have already discussed in Section 1.3.3). This equivalence is very easy to proved and so we omit it here. Therefore $\varphi_2$ can be taken as any functions $\varphi_2 = \varphi_2(x, y, \theta_0)$ satisfying $L_{\tilde{v}_n}\varphi_2 = 0$ and $(\mathrm{d}\varphi_1 \wedge \mathrm{d}\varphi_2)(x_0) \neq 0$. The space of solutions is thus defined on $\{q : \eta^2(q) \neq 0\}$, i.e., $\{q : \theta_{i+1} - \theta_i \neq \pm\frac{\pi}{2}\}$.

Take $\varphi_1 = x$, then $\tilde{v}_n = \cos\theta_0 \frac{\partial}{\partial \theta_0}$ and $L_{c_i}\varphi_2 = L_{\tilde{v}_n}\varphi_2 = 0$, for $1 \le i \le n$, imply that $\varphi_2$ is any function of the form $\varphi_2(x, y)$ satisfying $\frac{\partial \varphi_2}{\partial y}(q_0) \neq 0$ (because of $(\mathrm{d}\varphi_1 \wedge \mathrm{d}\varphi_2)(q_0) \neq 0$). All such functions satisfy $\mathrm{span}\,\{\mathrm{d}x, \mathrm{d}\varphi_2\} = \mathrm{span}\,\{\mathrm{d}x, \mathrm{d}\tilde{\varphi}_2\}$ and we can take, for instance, $\varphi_2 = y$.

To see the other choice, take $\varphi_1 = \theta_0$, then

$$\tilde{v}_n = -\cos\theta_0 \frac{\partial}{\partial x} - \sin\theta_0 \frac{\partial}{\partial y}$$

and the general solution of $L_{c_i}\varphi_2 = L_{\tilde{v}_n}\varphi_2 = 0$ is $\varphi_2 = \varphi_2(\theta_0, x\sin\theta_0 - y\cos\theta_0)$.

It is interesting to notice that the family of all $x$-flat outputs of the $n$-trailer system coincides with the family of all $x$-flat outputs of the nonholonomic car, more precisely of the car defined as the tail consisting of the last and one-before-the-last trailers (those indexed , respectively, by $i = 0$ and $i = 1$) controlled by towing (forward and backward) and rotating the one-before-the-last trailer. Indeed, the family of $x$-flat outputs coincide because they are given by any $\varphi_1 = \varphi_1(x, y, \theta_0)$ and any $\varphi_2(x, y, \theta_0)$ satisfying $L_{\tilde{v}_n}\varphi_2 = 0$ (for the $n$-trailer) and $L_{v_1}\varphi_2 = 0$ (for the car) but, clearly, these two equations are equivalent (compare Section 1.3.4).

## 1.4 Proof of main theorems

### 1.4.1 Useful results

In this section we give a series of results that we will use in the subsequent sections when proving our theorems. We start with a weaker version of Proposition 1.2.3 proving that the statement of the latter holds on open and dense subset.

**Lemma 1.4.1** *Consider a driftless 2-input smooth control system $\Sigma$ defined on a manifold $M$ of dimension $n+2$. Let $\varphi_1$, $\varphi_2$ be two functions defined on $M$. If $(\varphi_1, \varphi_2)$ is an $x$-flat output at $(x_0, u_0)$, then there exists an open neighborhood $\mathcal{M}$ of $x_0$ and an open and dense subset $\tilde{\mathcal{M}}$ such that around any point $q \in \tilde{\mathcal{M}}$ there exist coordinates $(z_1, \ldots, z_{n+2})$ in which $\Sigma$ is locally feedback equivalent to the chained form $\Sigma_{\text{chain}}$, given by (1.2.2) such that $\varphi_1 = z_1$ and $\varphi_2 = z_2$.*

**Proof:** Let $(\varphi_1, \varphi_2)$ be an $x$-flat output of $\Sigma$ at $(x_0, u_0)$. There exists an open neighborhood $\mathcal{M}$ of $x_0$ such that $(\varphi_1, \varphi_2)$ is an $x$-flat output at $(x, u)$ for any $x \in \mathcal{M}$ and $u = u(x)$. It is known (see, e.g., [14], [26], [56]) that the differentials of flat output are independent at $x_0$ and thus we put $x_1 = \varphi_1$, $x_2 = \varphi_2$, and complete them to a coordinate system $\xi = (x_1, x_2, \ldots, x_{n+2})$. Consider the $(2 \times 2)$-matrix $\mathrm{D} = (\mathrm{D}_{ij})$ given by $\mathrm{D}_{ij} = L_{g_j}\varphi_i$, $1 \le i, j \le 2$. It is immediate to see that $\operatorname{rank}\mathrm{D}(q) \le 1$, for any $q \in \mathcal{M}$. Indeed, if the rank were two then by a suitable invertible feedback $u = \beta(x)v$ we would get

$$\begin{aligned} \varphi_1 &= x_1, \qquad \dot{x}_1 = v_1 \\ \varphi_2 &= x_2, \qquad \dot{x}_2 = v_2 \end{aligned}$$

which contradicts the flatness assumption because $\varphi_i^{(j)} = v_i^{(j-1)}$, for $1 \le i \le 2$, and any $j \ge 1$, and thus the coordinates $x_3, \ldots, x_{n+2}$ cannot be represented as functions of $\varphi_i^{(j)}$, $j \ge 0$. Therefore on an open and dense subset $\mathcal{M}'$ of $\mathcal{M}$, $\operatorname{rank}\mathrm{D}(q) = 1$, for $q \in \mathcal{M}'$. After applying around any $q \in \mathcal{M}$ a suitable invertible feedback $u = \beta(x)v$, the system becomes

$$\dot{x} = g_1(x)v_1 + g_2(x)v_2,$$

where the flat outputs and their derivatives are

$$\begin{array}{ll}\varphi_1 = x_1, & \dot{x}_1 = v_1 \\ \varphi_2 = x_2, & \dot{x}_2 = \psi(x) v_1\end{array}$$

with $\psi(x)$ being a smooth function. Consider the vector fields

$$\begin{aligned} g_1 &= \frac{\partial}{\partial x_1} + \psi \frac{\partial}{\partial x_2} + \sum_{i=3}^{n+2} g_{1i}(x) \frac{\partial}{\partial x_i} \\ g_2 &= \sum_{i=3}^{n+2} g_{2i}(x) \frac{\partial}{\partial x_i}. \end{aligned}$$

We claim that for any $q$ in an open and dense subset $\mathcal{M}''$ of $\mathcal{M}'$ there exists $\rho$ such that

$$\begin{aligned} L_{g_2} L_{g_1}^{\mu} \psi &\equiv 0, \quad \text{for } \ 0 \le \mu \le \rho - 2, \\ L_{g_2} L_{g_1}^{\rho-1} \psi(q) &\neq 0, \end{aligned}$$

and, moreover, that $\rho = n$ for any $q \in \mathcal{M}''$. In other words, $\rho = n$ is the relative degree of the single-input system $\dot{x} = f + vg$, where $f = g_1, g = g_2$ and $v = v_2$.

To prove our claim, first, observe that $\rho$ exists on an open and dense subset $\mathcal{M}''$ of $\mathcal{M}'$. If not, on an open set in $\mathcal{M}'$, we would have $L_{g_2} L_{g_1}^{\mu} \psi \equiv 0$ for any $\mu \ge 0$, which contradicts the flatness of $\Sigma$ since $v_2$ could not be expressed in terms of $\varphi_i^{(j)}$, $i = 1, 2,\ j \ge 0$. Thus $\rho$ exists and is locally constant on an open and dense subset $\mathcal{M}'''$ of $\mathcal{M}''$, with a priori different constant values on different connected components of $\mathcal{M}'''$. We claim that on each connected component the constant value of $\rho$ is $n$. On one hand, we have $\rho \le n$ since $L_{g_2}\varphi_1 = L_{g_2}\varphi_2 = 0$. On the other hand, if $\rho < n$, then we put $\tilde{v}_2 = (L_{g_2} L_{g_1}^{\rho-1}\psi) v_2 + (L_{g_1}^{\rho}\psi) v_1$ which is an invertible feedback, because of the definition of $\rho$. Now knowing $\varphi_1$ and $\varphi_2$, we can obtain $v_1 = \dot{\varphi}_1, \psi = \dfrac{\dot{\varphi}_2}{\dot{\varphi}_1}$, by differentiation $\rho - 1$ functions $L_{g_1}^{\rho}\psi = \psi^{(\mu)}$, $1 \le \mu \le \rho - 1$, and finally the control $\tilde{v}_2$. This gives 2 controls and $\rho + 2 < n + 2$ functions, so one function among $x_3, \ldots, x_{n+2}$ is missing. This contradicts the flatness assumption. We thus have proved that on an open and dense subset $\mathcal{M}'''$ of $\mathcal{M}'$, the relative degree $\rho$ is well defined and equals $n$.

Fix an arbitrary $q \in \mathcal{M}'''$, we can assume that $\psi(q) = g_{1i}(q) = 0$, for $3 \le i \le n+2$ (if not, we replace $x_i$ by $x_i - k_i x_1$, for $2 \le i \le n+2$, where $k_2 = \psi(q)$ and $k_i = g_{1i}(q)$, for $3 \le i \le n+2$). We claim that the differentials of the functions $x_1, x_2, \psi, L_f\psi, \ldots, L_f^{n-1}\psi$ are independent at $q \in \mathcal{M}'''$. To this end, we will use the following result (see, e.g., Isidori [24]): if two vector fields $f$ and $g$ and a function $\phi$ satisfy

$$L_g\phi = L_g L_f \phi = \cdots = L_g L_f^{k-2}\phi = 0, \quad L_g L_f^{k-1}\phi = \lambda,$$

where $\lambda$ is function, then for any $1 \le j \le k - 1$,

$$L_{\mathrm{ad}_f^j g}\phi = L_{\mathrm{ad}_f^j g}(L_f\phi) = \cdots = L_{\mathrm{ad}_f^j g}(L_f^{k-j-2}\phi) = 0, \quad L_{\mathrm{ad}_f^j g}(L_f^{k-j-1}\phi) = -\lambda.$$

We apply this result to $f = g_1, g = g_2, \phi = \psi$ and $k = n$ so, in particular, $\lambda(q) \neq 0$. Since, $L_{g_2}\varphi_1 = 0$ and $L_{g_1}\varphi_1 = 1$, it follows that

$$L_{\mathrm{ad}_{g_1} g_2}\varphi_1 = L_{[g_1,g_2]}\varphi_1 = L_{g_1}L_{g_2}\varphi_1 - L_{g_2}L_{g_1}\varphi_1 = 0$$

and by induction we prove easily that $L_{\mathrm{ad}^j_{g_1} g_2}\varphi_1 = 0$. Moreover, $\psi = L_{g_1}\varphi_2$ implies that

$$\begin{aligned}
&L_{\mathrm{ad}_{g_1} g_2}\varphi_2 = \cdots = L_{\mathrm{ad}^{n-1}_{g_1} g_2}\varphi_2 = 0\\
&L_{\mathrm{ad}^n_{g_1} g_2}\varphi_2(q) = (-1)^n L_{g_2}L^n_{g_1}\varphi_2(q) = (-1)^n L_{g_2}L^{n-1}_{g_1}\psi(q) \neq 0
\end{aligned}$$

Putting $L_{\mathrm{ad}^n_{g_1} g_2}\varphi_2 = \tilde{\lambda}$ and evaluating the differential forms $\mathrm{d}\varphi_i, \mathrm{d}L^j_{g_1}\psi$, $0 \leq j \leq n-1$ on the vector fields $g_2, \mathrm{ad}_{g_1}g_2, \ldots, \mathrm{ad}^n_{g_1}g_2, g_1$ we get (notice that $L_{g_1}\varphi_2 = \psi$)

$$\begin{pmatrix} \mathrm{d}\varphi_1 \\ \mathrm{d}\varphi_2 \\ \mathrm{d}\psi \\ \mathrm{d}L_{g_1}\psi \\ \vdots \\ \mathrm{d}L^{n-1}_{g_1}\psi \end{pmatrix} \left(g_2, \mathrm{ad}_{g_1}g_2, \ldots, \mathrm{ad}^n_{g_1}g_2, g_1\right) = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & \tilde{\lambda} & \psi \\ \cdots & \cdots & \cdots & * & * \\ 0 & (-1)^{n-1}\tilde{\lambda} & \cdots & \cdots & \cdots \\ (-1)^n\tilde{\lambda} & * & \cdots & \cdots & * \end{pmatrix}$$

The determinant of the above matrix is nonzero since $\tilde{\lambda}(q) \neq 0$, and therefore the functions $\varphi_1, \varphi_2, \psi, L_{g_1}\psi, \ldots, L^{n-1}_{g_1}\psi$ are independent. It follows that

$$\begin{aligned}
z_1 &= \varphi_1\\
z_2 &= \varphi_2\\
z_3 &= \psi\\
z_4 &= L_{g_1}\psi\\
&\vdots\\
z_{n+2} &= L^{n-1}_{g_1}\psi
\end{aligned}$$

is a valid local change of coordinates in a neighborhood of any $q \in \tilde{\mathcal{M}} = \mathcal{M}'''$, in which the system after applying feedback $\tilde{v}_1 = v_1$, $\tilde{v}_2 = (L^n_{g_1}\psi)v_1 + (L_{g_2}L^{n-1}_{g_1}\psi)v_2$ (invertible since $(L_{g_2}L^{n-1}_{g_1}\psi)(q) \neq 0$) takes the chained form $\Sigma_{\mathrm{chain}}$ (1.2.2). $\square$

Although the above lemma was proved at generic points only, it implies the following result in a whole neighborhood of the point $x_0$ under consideration. Recall that for any Goursat structure $\mathcal{D}$ we denote by $\mathcal{C}_{n-1}$ the characteristic distribution of $\mathcal{D}^{(n-1)}$ (see Lemma 1.3.1).

**Corollary 1.4.2** *Consider a Goursat structure $\mathcal{D}$ on $M$ of dimension $n+2$, that is,* $\mathrm{rank}\,\mathcal{D}^{(i)} = i+2$, *for $0 \leq i \leq n$, hold everywhere on $M$. If the associated control system $\Sigma$ is $x$-flat at $(x_0, u_0) \in M \times \mathbb{R}^2$, $u_0 \notin U_{\mathrm{sing}}$, then for any $x$-flat output $(\varphi_1, \varphi_2)$ at $(x_0, u_0)$, there exists an open neighborhood $\mathcal{M}$ of $x_0$ in which we have $L_c\varphi_i = 0$, for $i = 1, 2$ and any $c \in \mathcal{C}_{n-1}$.*

**Proof:** Let $(\varphi_1, \varphi_2)$ be an $x$-flat output of $\Sigma$ at $(x_0, u_0)$. By Lemma 1.4.1, there exists an open neighborhood $\mathcal{M}$ of $x_0$ and an open and dense subset $\tilde{\mathcal{M}}$ of $\mathcal{M}$ such that around any $q \in \tilde{\mathcal{M}}$, the system $\Sigma$ is feedback equivalent to the chained form $\Sigma_{\text{chain}}$, given by (1.2.2), with $\varphi_1 = z_1$ and $\varphi_2 = z_2$. We have $\mathcal{C}_{n-1} = \operatorname{span}\{\frac{\partial}{\partial z_4}, \ldots, \frac{\partial}{\partial z_{n+2}}\}$ and hence $L_c\varphi_i = 0$, for $i = 1, 2$ and any $c \in \mathcal{C}_{n-1}$ on $\tilde{\mathcal{M}}$ and hence on $\mathcal{M}$ (since $\tilde{\mathcal{M}}$ is dense and the functions $\varphi_i$ as well as the distribution $\mathcal{C}_{n-1}$ are well defined on the whole $\mathcal{M}$). $\square$

### 1.4.2 Proof of Theorem 1.2.2

We will show the implications (iii) $\Rightarrow$ (iv) $\Rightarrow$ (ii) $\Rightarrow$ (i) $\Rightarrow$ (iii).

(iii) $\Rightarrow$ (iv): It is a well known result (proved by Murray in [45]) that $\dim \mathcal{D}^{(i)}(x_0) = \dim \mathcal{D}_i(x_0)$, $0 \le i \le n$, are necessary and sufficient for local feedback equivalent to the chained form.

(iv) $\Rightarrow$ (ii): It is obvious for a system the chained form $\Sigma_{\text{chain}}$ given by (1.2.2), that $\varphi_1 = z_1$ and $\varphi_2 = z_2$ yield flatness for $v_1 \neq 0$ and the latter means that we can take in (ii) any $u_0 \notin U_{\text{sing}}(x_0)$.

(ii) $\Rightarrow$ (i): Obvious.

(i) $\Rightarrow$ (iii): This is the only difficult implication. Its proof will be based on Corollary 1.4.2 and on the result that assures that any Goursat structure can be brought to the following polynomial normal form, called Kumpera-Ruiz normal, as proved by Pasillas-Lépine and Respondek [52] (see also Mormul [42]):

**Theorem 1.4.3 (Kumpera-Ruiz normal form)** *Assume that $n \ge 2$. Any Goursat structure defined on a manifold $M$ of dimension $n+2$ is locally equivalent, at any point $x_0$ in $M$, to a distribution spanned in a small neighborhood of zero by a pair of vector fields that has the following form:*

$$f_1 = \sum_{i=0}^{m} \Big(\prod_{j=0}^{i-1} x^j_{k_j}\Big)\Big(\sum_{j=1}^{k_i-1} (x^i_j + c^i_j)\frac{\partial}{\partial x^i_{j+1}} + \frac{\partial}{\partial x^{i+1}_1}\Big), \quad f_2 = \frac{\partial}{\partial x^0_1} \tag{1.4.1}$$

*where the coordinates $x^i_j$, for $0 \le i \le m+1$ and $1 \le j \le k_i$, are centered at $x_0$ (that means that $x(x_0) = 0$); the integer $m$ is such that $0 \le m \le n-2$; and $k_i$, for $0 \le i \le m-1$, satisfy $k_0 \ge 1, \ldots, k_{m-1} \ge 1$, $k_m \ge 3$, $k_{m+1} = 1$ and $\sum_{i=0}^{m+1} k_i = n+2$. The constants $c^i_j$, for $1 \le j \le k_i - 1$, are real constants.*

**Remark**. In the above normal form, the integer $m$ gives the number of singularities of the Kumpera-Ruiz normal form. When $m = 0$, the Kumpera-Ruiz normal form coincides with the Goursat normal form (since in this case all constant $c^i_j$ can be eliminated).

In order to prove (i) $\Rightarrow$ (iii), assume that $\Sigma$ is $x$-flat at $(x_0, \bar{u}_0^l)$, for a certain $l \geq 0$, and that there exists an integer $2 \leq i \leq n$ such that $\dim \mathcal{D}^{(i)}(x_0) \neq \dim \mathcal{D}_i(x_0)$. Since $\mathcal{D}$ is a Goursat structure and $\dim \mathcal{D}^{(i)}(x_0) \neq \dim \mathcal{D}_i(x_0)$, for certain $2 \leq i \leq n$, by Theorem 1.4.3, there exists a new coordinate system $(x_1^0, \ldots, x_{k_0}^0, \ldots, x_1^m, \ldots, x_{k_m}^m, x_1^{m+1})$ in which $\mathcal{D}$ takes, in a small neighborhood of zero, the Kumpera-Ruiz normal form with $m \geq 1$ and $\sum_{i=0}^m k_i = n+1$ , i.e., $\mathcal{D} = \mathrm{span}\,\{f_1, f_2\}$ where $f_1$ and $f_2$ are given by (1.4.1). A direct calculation shows that

$$\mathcal{D}^{(j)} = \mathrm{span}\,\{\frac{\partial}{\partial x_1^0}, \ldots, \frac{\partial}{\partial x_{j+1}^0}, f_1\}, \quad 1 \leq j \leq k_0 - 1,$$

and $\mathcal{C}_j = \mathrm{span}\,\{\frac{\partial}{\partial x_1^0}, \ldots, \frac{\partial}{\partial x_j^0}\}$, $1 \leq j \leq k_0 - 1$. Observing that

$$\Big[\frac{\partial}{\partial x_{k_0}^0}, f_1\Big] = \sum_{i=1}^{m} \Big(\prod_{j=1}^{i-1} x_{k_j}^j\Big)\Big(\sum_{j=1}^{k_i - 1} (x_j^i + c_j^i)\frac{\partial}{\partial x_{j+1}^i} + \frac{\partial}{\partial x_1^{i+1}}\Big)$$

and

$$\frac{\partial}{\partial x_1^1} = f_1 - x_{k_0}^0 \Big[\frac{\partial}{\partial x_{k_0}^0}, f_1\Big] - \sum_{j=1}^{k_0-1} (x_j^0 + c_j^0)\frac{\partial}{\partial x_{i+1}^0},$$

we then get

$$\begin{aligned}
\mathcal{D}^{(k_0)} &= \mathcal{D}^{(k_0-1)} + [\mathcal{D}^{(k_0-1)}, \mathcal{D}^{(k_0-1)}] \\
&= \mathrm{span}\,\Big\{\frac{\partial}{\partial x_1^0}, \ldots, \frac{\partial}{\partial x_{k_0}^0}, \frac{\partial}{\partial x_1^1}, \sum_{i=1}^{m}\Big(\prod_{j=1}^{i-1} x_{k_j}^j\Big)\Big(\sum_{j=1}^{k_i-1}(x_j^i + c_j^i)\frac{\partial}{\partial x_{j+1}^i} + \frac{\partial}{\partial x_1^{i+1}}\Big)\Big\},
\end{aligned}$$

and $\mathcal{C}_{k_0} = \mathrm{span}\,\{\frac{\partial}{\partial x_1^0}, \ldots, \frac{\partial}{\partial x_{k_0}^0}\}$. In the same way, we obtain that

$$\mathcal{C}_{k_i + j} = \mathcal{C}_{k_i} \oplus \mathrm{span}\,\{\frac{\partial}{\partial x_1^{i+1}}, \ldots, \frac{\partial}{\partial x_j^{i+1}}\},$$

for $0 \leq i \leq m$, $1 \leq j \leq k_i$ and $k_m + j \leq n - 1$. Therefore, finally, the characteristic distribution $\mathcal{C}_{n-1}$ of $\mathcal{D}^{(n-1)}$ is given by

$$\mathcal{C}_{n-1} = \mathrm{span}\,\{\frac{\partial}{\partial x_1^0}, \ldots, \frac{\partial}{\partial x_{k_0}^0}, \ldots, \frac{\partial}{\partial x_1^m}, \ldots, \frac{\partial}{\partial x_{k_m-2}^m}\} = (\mathrm{span}\,\{\mathrm{d}x_{k_m-1}^m, \mathrm{d}x_{k_m}^m, \mathrm{d}x_1^{m+1}\})^{\perp}.$$

For simplicity, we denote the coordinates $x_{k_m-1}^m, x_{k_m}^m, x_1^{m+1}$ by $y_1, y_2, y_3$, respectively, and denote the remaining $n-1$ coordinates $x_j^i$ by $x_1, \ldots, x_{n-1}$, that is,

$$(x, y) = (x_1, \ldots, x_{n-1}, y_1, y_2, y_3) = (x_1^0, \ldots, x_{k_0}^0, \ldots, x_1^m, \ldots, x_{k_m}^{m-1}, x_{k_m}^m, x_1^{m+1}).$$

In $(x, y)$-coordinates, the control system $\Sigma$ associated to the Kumpera-Ruiz normal form (1.4.1), reads as

$$\Sigma_{\mathrm{KR}}: \quad \begin{cases} \dot{x}_1 & = & u_1 \\ \dot{x}_2 & = & \gamma_2(x,y)u_2 \\ & \vdots & \\ \dot{x}_{n-1} & = & \gamma_{n-1}(x,y)u_2 \\ \dot{y}_1 & = & x_r\beta_1(x,y)u_2 \\ \dot{y}_2 & = & x_r\beta_2(x,y)u_2 \\ \dot{y}_3 & = & x_r\beta_3(x,y)u_2 \end{cases}$$

where $\beta_1$, $\beta_2$, $\beta_3$ and $\gamma_j$, for $2 \leq j \leq n-1$, are smooth functions defined in a neighborhood of $0 \in \mathbb{R}^{n+2}$, $1 \leq r \leq n-1$ is an integer, and the characteristic distribution $\mathcal{C}_{n-1}$ is given by

$$\mathcal{C}_{n-1} = \mathrm{span}\left\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_{n-1}}\right\} = (\mathrm{span}\left\{\mathrm{d}y_1, \mathrm{d}y_2, \mathrm{d}y_3\right\})^{\perp}.$$

By our assumption, $\Sigma_{\mathrm{KR}}$ is $x$-flat at $(0, \bar{u}_0^l) \in \mathbb{R}^{n+2} \times \mathbb{R}^{2(l+1)}$ and let $(\varphi_1, \varphi_2)$ be an $x$-flat output defined in a neighborhood $\mathcal{O}$ of $0 \in \mathbb{R}^{n+2}$. Being a Goursat structure, $\Sigma_{\mathrm{KR}}$ satisfies $\dim \mathcal{D}^{(i)}(z) = \dim \mathcal{D}_i(z)$, $0 \leq i \leq n$, for any $z = (x, y)$ in an open and dense subset $\mathcal{O}'$ of $\mathcal{O}$ and by Corollary 1.4.2 and the form of $\mathcal{C}_{n-1}$, we conclude that

$$\frac{\partial \varphi_1}{\partial x_i} = \frac{\partial \varphi_2}{\partial x_i} = 0,\ 1 \leq i \leq n-1,$$

holds in $\mathcal{O}'$ and since $\mathcal{O}'$ is dense in $\mathcal{O}$, also in $\mathcal{O}$. It follows that $\varphi_i = \varphi_i(y_1, y_2, y_3)$ for $i = 1, 2$. Moreover, the fact that $(\varphi_1, \varphi_2)$ is an $x$-flat output at $(0, \bar{u}_0^l)$ implies that $\varphi_1, \varphi_2, \dot{\varphi}_1, \dot{\varphi}_2$ must be independent at $(0, u_0) \in \mathbb{R}^{n+2} \times \mathbb{R}^2$, where $u_0 = (u_{10}, u_{20})$, (see, e.g.,[14] and [56]). Calculating the derivatives $\dot{\varphi}_i$, for $i = 1, 2$, we get

$$\dot{\varphi}_i = \sum_{j=1}^{3} x_r \beta_j u_2 \frac{\partial \varphi_i}{\partial y_j} = F_i(x, y_1, y_2, y_3, u_2), \quad i = 1, 2.$$

Then

$$\mathrm{d}\dot{\varphi}_i = \mathrm{d}F_i = \sum_{j=1}^{n-1} \frac{\partial F_i}{\partial x_j} \cdot \mathrm{d}x_j + \sum_{j=1}^{3} \frac{\partial F_i}{\partial y_j} \cdot \mathrm{d}y_j + \frac{\partial F_i}{\partial u_2} \cdot \mathrm{d}u_2, \quad i = 1, 2$$

and at $(0, u_0)$,

$$\mathrm{d}\dot{\varphi}_i(0, u_0) = \mathrm{d}F_i(0, u_0) = \frac{\partial F_i}{\partial x_r}(0, u_0) \cdot \mathrm{d}x_r, \quad i = 1, 2,$$

which implies that

$$(\mathrm{d}\varphi_1 \wedge \mathrm{d}\varphi_2 \wedge \mathrm{d}\dot{\varphi}_1 \wedge \mathrm{d}\dot{\varphi}_2)(0, u_{20}) = 0,$$

independently of the value of $u_{20}$, which gives a contradiction. Therefore if a system associated to a Goursat structure is $x$-flat at $(x, \bar{u}_0^l)$, for some $\bar{u}_0^l$, then we have $\dim \mathcal{D}^{(i)}(x_0) = \dim \mathcal{D}_i(x_0)$, $0 \leq i \leq n$.

### 1.4.3 Proof of Theorem 1.3.2

**Proof:** *Sufficiency:* Take any 2-input system whose associated distribution satisfies $\dim \mathcal{D}^{(i)}(x) = \dim \mathcal{D}_i(x) = i+2$ everywhere in a neighborhood of $x_0$ and choose two functions fulfilling (i) – (iii). We can bring $\Sigma$ to the chained form $\Sigma_{\text{chain}}$, given by (1.2.2), in coordinates $(z_1, \ldots, z_{n+2})$, transforming $x_0$ into $0 \in \mathbb{R}^{n+2}$, and its associated distribution is given by $\mathcal{D} = \operatorname{span}\{g_1, g_2\}$, where

$$
\begin{aligned}
g_1 &= \frac{\partial}{\partial z_1} + z_3 \frac{\partial}{\partial z_2} + \cdots + z_{n+2} \frac{\partial}{\partial z_{n+1}} \\
g_2 &= \frac{\partial}{\partial z_{n+2}}.
\end{aligned}
$$

The characteristic distribution $\mathcal{C}_{n-1}$ is given by

$$
\mathcal{C}_{n-1} = \operatorname{span}\{\frac{\partial}{\partial z_4}, \ldots, \frac{\partial}{\partial z_{n+2}}\},
$$

and the condition $L_c \varphi_i = 0$, for any $c \in \mathcal{C}_{n-1}$, of item (ii) implies that $\varphi_i = \varphi_i(z_1, z_2, z_3)$, for $i = 1, 2$. Item (iii) implies that there exists $\varphi_i$ (say $\varphi_1$, if not we permute) such that $L_{g_1}\varphi_1(0) = \frac{\partial \varphi_1}{\partial z_1}(0) \neq 0$. Due to (i) we can complete $\varphi_1$ and $\varphi_2$ by a function $\varphi_3(z_1, z_2, z_3)$ such that $(\mathrm{d}\varphi_1 \wedge \mathrm{d}\varphi_2 \wedge \mathrm{d}\varphi_3)(0) \neq 0$ to introduce new coordinates $\tilde{z}_i = \varphi_i(z_1, z_2, z_3)$, $1 \leq i \leq 3$, followed by $\tilde{z}_i = z_i$, $4 \leq i \leq n$. We have

$$
\dot{\tilde{z}}_i = \psi_i(z_1, z_2, z_3, z_4) u_1, \quad 1 \leq i \leq 3,
$$

where $\psi_i = L_{g_1}\varphi_i$. Since $\psi_1(0) = L_{g_1}\varphi_1(0) \neq 0$, we apply invertible feedback $\tilde{u}_1 = \psi_1 u_1$ to get

$$
\Sigma : \begin{cases}
\dot{\tilde{z}}_1 &= \tilde{u}_1 \\
\dot{\tilde{z}}_2 &= \tilde{\psi}_2 \tilde{u}_1 \\
\dot{\tilde{z}}_3 &= \tilde{\psi}_3 \tilde{u}_1 \\
\dot{\tilde{z}}_i &= \tilde{z}_{i+1} \dfrac{\tilde{u}_1}{\psi_1}, \quad \text{for } \ 4 \leq i \leq n+1 \\
\dot{\tilde{z}}_{n+2} &= u_2.
\end{cases}
$$

Notice that the characteristic distribution $\mathcal{C}_{n-1} = \operatorname{span}\{\frac{\partial}{\partial \tilde{z}_4}, \ldots, \frac{\partial}{\partial \tilde{z}_{n+2}}\}$ (since $\tilde{z}_i = \psi_i(z_1, z_2, z_3)$, for $1 \leq i \leq 3$). It thus follows from $L_c \left( \frac{L_{g_1}\varphi_2}{L_{g_1}\varphi_1} \right) = 0$, $c \in \mathcal{C}_{n-1}$, that $\tilde{\psi}_2 = \frac{\psi_2(z_1, z_2, z_3, z_4)}{\psi_1(z_1, z_2, z_3, z_4)}$ is actually a function of $\tilde{z}_1, \tilde{z}_2, \tilde{z}_3$ only. Moreover, the condition $\operatorname{rank} \mathcal{D}^{(n)} = \operatorname{rank} \mathcal{D}_n = n+2$ implies that $\frac{\partial \tilde{\psi}_3}{\partial \tilde{z}_4}(0) \neq 0$ and a direct calculation shows that $L_{\tilde{g}_2} L_{\tilde{g}_1}^{\mu} \tilde{\psi}_2 \equiv 0$ for $0 \leq \mu \leq n-2$ and $L_{\tilde{g}_2} L_{\tilde{g}_1}^{n-1} \tilde{\psi}_2(0) \neq 0$, where

$$
\begin{aligned}
\tilde{g}_1 &= \frac{\partial}{\partial \tilde{z}_1} + \tilde{\psi}_2 \frac{\partial}{\partial \tilde{z}_2} + \tilde{\psi}_3 \frac{\partial}{\partial \tilde{z}_3} + \frac{\tilde{z}_5}{\psi_1} \frac{\partial}{\partial \tilde{z}_4} + \cdots + \frac{\tilde{z}_{n+2}}{\psi_1} \frac{\partial}{\partial \tilde{z}_{n+1}} \\
\tilde{g}_2 &= \frac{\partial}{\partial \tilde{z}_{n+2}}.
\end{aligned}
$$

Therefore the function $\tilde{\psi}_2$ satisfies the condition of $\psi$ from the proof of Lemma 1.4.1, and following that proof we can thus bring the system to the chained form (1.2.2), with $(\varphi_1, \varphi_2) = (\tilde{z}_1, \tilde{z}_2)$ which proves that $(\varphi_1, \varphi_2)$ is indeed an $x$-flat output at $(x_0, u_0)$, $u_0 \notin U_{\text{sing}}(x_0)$.

*Necessity*: Assume that $\Sigma$ is $x$-flat at $(x_0, u_0)$, $u_0 \notin U_{\text{sing}}(x_0)$, and let $(\varphi_1, \varphi_2)$ be its $x$-flat output defined in a neighborhood $\mathcal{M}$ of $x_0$. It is well known (see [14], [26], [56]) that $\mathrm{d}\varphi_1(x_0) \wedge \mathrm{d}\varphi_2(x_0) \neq 0$. By Lemma 1.4.1, we can bring $\Sigma$, around any point $q \in \tilde{\mathcal{M}}$ (open and dense in $\mathcal{M}$), into the chained form $\Sigma_{\text{chain}}$, given by (1.2.2), with $\varphi_1 = z_1$, $\varphi_2 = z_2$, and $x_0$ is transformed into $z_0 = 0 \in \mathbb{R}^{n+2}$. We have $\mathcal{C}_{n-1} = \operatorname{span}\{\frac{\partial}{\partial z_4}, \ldots, \frac{\partial}{\partial z_{n+2}}\}$. By a direct calculation we get that $L_g\varphi_1(0) \neq 0$ for any $g \in \mathcal{D}$ such that $g(0) \notin \mathcal{C}_{n-1}(0)$, and that $L_c\varphi_1 \equiv L_c\varphi_2 \equiv L_c(\frac{L_g\varphi_2}{L_g\varphi_1}) \equiv 0$, for any $c \in \mathcal{C}_{n-1}$ and $g$ as above, which gives the item (ii) on $\tilde{\mathcal{M}}$. Now observe that the flat outputs $\varphi_1$, $\varphi_2$ are well defined in $\mathcal{M}$ and so is the characteristic distribution $\mathcal{C}_{n-1}$ (since the distribution $\mathcal{D}$ associated to $\Sigma$ satisfies $\operatorname{rank} \mathcal{D}^{(i)} = i+2$, $0 \leq i \leq n$ everywhere in $\mathcal{M}$). It follows by continuity that $L_c\varphi_1 \equiv L_c\varphi_2 \equiv (L_g\varphi_1)(L_cL_g\varphi_2) - (L_g\varphi_2)(L_cL_g\varphi_1) \equiv 0$ holds also on $\mathcal{M}$ thus implying (ii) on $\mathcal{M}$.

It remains to prove (iii). Bring $\Sigma$, locally around $x_0 \in \mathcal{M}$, into the chained form $\Sigma_{\text{chain}}$, given by (1.2.2) (which is always possible by the assumption of theorem). Then item (ii), which we have just proved on $\mathcal{M}$, implies that $\varphi_i = \varphi_i(z_1, z_2, z_3)$, for $1 \leq i \leq 2$ and $g = \frac{\partial}{\partial z_1} + z_3\frac{\partial}{\partial z_2} + z_4\frac{\partial}{\partial z_3} \mod \mathcal{C}_{n-1}$. If for $i = 1$ or 2, we have $L_g\varphi_i(0) \neq 0$, then (iii) holds. So assume that $L_g\varphi_1(0) = L_g\varphi_2(0) = 0$ implying $\frac{\partial\varphi_1}{\partial z_1}(0) = \frac{\partial\varphi_2}{\partial z_1}(0) = 0$. We have

$$\begin{aligned}
\dot{\varphi}_1 &= (\frac{\partial\varphi_1}{\partial z_1} + z_3\frac{\partial\varphi_1}{\partial z_2} + z_4\frac{\partial\varphi_1}{\partial z_3})u_1 &&= (a + bz_4)u_1 \\
\dot{\varphi}_2 &= (\frac{\partial\varphi_2}{\partial z_1} + z_3\frac{\partial\varphi_2}{\partial z_2} + z_4\frac{\partial\varphi_2}{\partial z_3})u_1 &&= (c + dz_4)u_1
\end{aligned}$$

and thus $\dot{\varphi}_1(0, u_0) = \dot{\varphi}_2(0, u_0) = 0$, where $u_0$ stands for the value of the nominal control. Defin $\tilde{\mathcal{M}} = \{z \in \mathcal{M} : (L_g\varphi_1(z) = L_g\varphi_2(z) = (0, 0)\}$. On $\tilde{\mathcal{M}}$ we, clearly, have $z_i = \tilde{\gamma}_i(\varphi_1, \varphi_2, \dot{\varphi}_1, \dot{\varphi}_2)$, $1 \leq i \leq 3$. We claim that around 0 there must exist smooth functions $\gamma_i$, for $1 \leq i \leq 3$, such that $z_i = \gamma_i(\varphi_1, \varphi_2, \dot{\varphi}_1, \dot{\varphi}_2)$ that coincide with $\tilde{\gamma}_i$ on $\tilde{\mathcal{M}}$. Indeed, from the definition of flatness it follows that $z_i = \gamma_i(\bar{\varphi}_1^s, \bar{\varphi}_2^s)$, for some $s \geq 0$. But if $s > 1$, then there exists nontrivial relations $\gamma_i = \tilde{\gamma}_i$ on $\tilde{\mathcal{M}}$ contradicting differential independence of $\varphi_1$ and $\varphi_2$ (see [14], [26], [56]). We have $\varphi_1 = \varphi_1(z_1, z_2, z_3)$, $\varphi_2 = \varphi_2(z_1, z_2, z_3)$ and since $\frac{\partial\varphi_1}{\partial z_1}(0) = \frac{\partial\varphi_2}{\partial z_1}(0) = 0$, it follows that $z_1 = \gamma_1(\varphi_1, \varphi_2, \dot{\varphi}_1, \dot{\varphi}_2)$, where $\gamma_1$ depends explicitly on $\dot{\varphi}_i$, but the composition $\gamma_1(\varphi_1, \varphi_2, \dot{\varphi}_1, \dot{\varphi}_2)(z_1, z_2, z_3, z_4, u_1)$ depends actually on $z_1, z_2, z_3$ only. Notice that $\dot{\varphi}_1$ and $\dot{\varphi}_2$ are linear with respect to $u_1$ and affine with respect to $z_4$ which implies that $(ad - bc) \equiv 0$ (in order that it does not depend on $z_4$), and that $\gamma_1$ is a function of $\frac{a + bz_4}{c + dz_4}$ (in order that it does not depend on $u_1$). However neither $\frac{a + bz_4}{c + dz_4}$ nor $\frac{c + dz_4}{a + bz_4}$

is smooth at 0. Therefore $L_g\varphi_i(0) \neq 0$ for at least one $1 \leq i \leq 2$ and the item (iii) holds as well. The above analysis also shows that at any point $x \in \mathcal{M}$ that satisfies $L_g\varphi_1(x) = L_g\varphi_2(x) = 0$, the pair $(\varphi_1, \varphi_2)$ fails to be an $x$-flat output. In other words,

$$Sing(\varphi_1, \varphi_2) = \{x \in \mathcal{M} : (L_g\varphi_1(x), L_g\varphi_2(x)) = (0, 0)\}.$$

□

Now we will show how Proposition 1.2.3 follows from Theorem 1.3.2.

**Proof: (of Proposition 1.2.3)** Let $(\varphi_1, \varphi_2)$ be a pair of $x$-flat outputs at $(x_0, u_0)$, $u_0 \notin U_{\text{sing}}(x_0)$. Then $(\varphi_1, \varphi_2)$ satisfy the items (i) – (iii) of Theorem 1.3.2 and we can follow the procedure described in the sufficiency part of that theorem in order to bring $\Sigma$ into the chained form $\Sigma_{\text{chain}}$ such that $\varphi_1 = z_1$ and $\varphi_2 = z_2$. □

### 1.4.4 Proof of Theorem 1.3.3

**Proof:** *Sufficiency*: We will prove separately the cases $\dim M = 3$ and $\dim M \geq 4$.

Case (**I**): $\dim M = 3$

Let $\varphi_1, \varphi_2$ be any functions satisfying (i)$'$ and (ii)$'$. Introduce coordinates $x_1 = \varphi_1$, $x_2 = \varphi_2$ and complete them to a coordinate system $(x_1, x_2, x_3)$, centered at $0 \in \mathbb{R}^3$. Choose two vector fields $\tilde{g}_1, g_2$ such that $\mathcal{D} = \text{span}\,\{\tilde{g}_1, g_2\}$, $\mathcal{L} = \text{span}\,\{g_2\}$. We have $\tilde{g}_1(0) \notin \mathcal{L}(0)$ and thus there exits a function $\varphi_i$, for $i = 1, 2$, such that $L_{\tilde{g}_1}\varphi_i(x_0) \neq 0$, say $L_{\tilde{g}_1}\varphi_1(x_0) \neq 0$. Define

$$g_1 = \frac{1}{L_{\tilde{g}_1}\varphi_1}\tilde{g}.$$

The associated control system $\dot{x} = u_1 g_1(x) + u_2 g_2(x)$ is

$$\begin{aligned} \dot{x}_1 &= u_1 \\ \dot{x}_2 &= \psi(x)u_1 \\ \dot{x}_3 &= \eta(x)u_2 \end{aligned}$$

where $\psi$ and $\eta$ are smooth functions. We have $\eta(0) \neq 0$ and we can suppose that $\psi(0) = 0$ (if not, replace $x_2$ by $x_2 - \psi(0)x_1$). The condition $\dim \mathcal{D}^{(1)}(0) = 3$ implies that $g_1$, $g_2$ and $[g_1, g_2]$ are independent at $0 \in \mathbb{R}^3$ and hence $\dfrac{\partial\psi}{\partial x_3}(0) \neq 0$. Replacing $x_3$ by $\psi$ and applying feedback to normalize $\dot{\psi}$, we get

$$\begin{aligned} \dot{x}_1 &= u_1 \\ \dot{x}_2 &= x_3 u_1 \\ \dot{x}_3 &= u_2 \end{aligned}$$

for which $(\varphi_1, \varphi_2) = (x_1, x_2)$ is an $x$-flat output at $(0, u_0)$, where $u_0 = (u_{10}, u_{20})^\top$ such that $u_{10} \neq 0$. To see that $u_{10} = 0$ is not a singular control, introduce the new coordinate $\tilde{x}_2 = x_2 - x_1 x_3$ to get

$$\begin{array}{rcl} \dot{x}_1 & = & u_1 \\ \dot{\tilde{x}}_2 & = & -x_1 u_2 \\ \dot{x}_3 & = & u_2 \end{array}$$

for which $(\tilde{\varphi}_1, \tilde{\varphi}_2) = (\tilde{x}_2, x_3)$ is an $x$-flat output at $(0, u_0)$ such that $u_{20} \neq 0$. If follows that the singular control is $U_{\text{sing}} = \{(0,0)\}$ only.

Case (**II**): $\dim M \geq 4$
Take any 2-input system whose associated distribution satisfies $\dim \mathcal{D}^{(i)}(x) = \dim \mathcal{D}_i(x) = i+2$ everywhere in a neighborhood of $x_0$ and choose two functions fulfilling $(\text{i})' - (\text{iii})'$. We can bring $\Sigma$ to the chained form (1.2.2) in coordinates $(z_1, \ldots, z_{n+2})$, transforming $x_0$ into $0 \in \mathbb{R}^{n+2}$, and its associated distribution is $\mathcal{D} = \text{span}\,\{g_1, g_2\}$, where

$$\begin{aligned} g_1 &= \frac{\partial}{\partial z_1} + z_3 \frac{\partial}{\partial z_2} + \cdots + z_{n+2} \frac{\partial}{\partial z_{n+1}} \\ g_2 &= \frac{\partial}{\partial z_{n+2}}. \end{aligned}$$

The characteristic distribution $\mathcal{C}_{n-1}$ is given by

$$\mathcal{C}_{n-1} = \text{span}\,\{\frac{\partial}{\partial z_4}, \ldots, \frac{\partial}{\partial z_{n+2}}\}.$$

Item $(\text{ii})'$ implies that $\mathcal{C}_{n-1} \subset \mathcal{L} = (\text{span}\,\{\text{d}\varphi_1, \text{d}\varphi_2\})^{\perp}$. Indeed, if there was a vector field $f \in \mathcal{C}_{n-1}$ such that $f \notin \mathcal{L}$, then $\mathcal{D}^{(n-1)} = \mathcal{L} + \text{span}\{f\}$ and hence

$$\mathcal{D}^{(n)} = \mathcal{D}^{(n-1)} + [\mathcal{D}^{(n-1)}, \mathcal{D}^{(n-1)}] = \mathcal{L} + \text{span}\{f\} + [f, \mathcal{L}] = \mathcal{L} + \text{span}\{f\} = \mathcal{D}^{(n-1)}$$

which contradicts the condition $\text{rank}\,\mathcal{D}^{(n)} = n+2$ . Therefore $\mathcal{C}_{n-1} \subset \mathcal{L}$ holds indeed. Consequently we have $L_c \varphi_i \equiv 0$, for $i = 1, 2$ and any $c \in \mathcal{C}_{n-1}$, which implies that $\varphi_i = \varphi_i(z_1, z_2, z_3)$ for $i = 1, 2$. Moreover observing that $g_2 \in \mathcal{C}_{n-1} \subset \mathcal{L}$, by item $(\text{iii})'$ we conclude, $g_1 \notin \mathcal{L}$, which implies that there exist $\varphi_i$, $1 \leq i \leq 2$, such that $L_{g_1} \varphi_i(0) \neq 0$, says $i = 1$. In other words, we have $L_{g_1} \varphi_1(0) = \dfrac{\partial \varphi_1}{\partial z_1}(0) \neq 0$. By $(\text{i})'$ we can complete $\varphi_1$ and $\varphi_2$ by a function $\varphi_3(z_1, z_2, z_3)$ such that $(\text{d}\varphi_1 \wedge \text{d}\varphi_2 \wedge \text{d}\varphi_3)(0) \neq 0$ to introduce new coordinates $\tilde{z}_i = \varphi_i(z_1, z_2, z_3)$, $1 \leq i \leq 3$, followed by $\tilde{z}_i = z_i$, $4 \leq i \leq n$. The remaining part of the proof follows the same line as that of Theorem 1.3.2.

*Necessity*: Assume that $\Sigma$ is $x$-flat at $(x_0, u_0)$, $u_0 \notin U_{\text{sing}}(x_0)$, and $(\varphi_1, \varphi_2)$ is an $x$-flat output at $(x_0, u_0)$, defined in a neighborhood $\mathcal{M}$ of $x_0$. It is well known (see [14], [26], [56]) that $\text{d}\varphi_1(x_0) \wedge \text{d}\varphi_2(x_0) \neq 0$. Now we will prove the item $(\text{ii})'$. Clearly, Lemma 1.4.1 applies and thus there exists an open and dense subset $\tilde{\mathcal{M}} \subset \mathcal{M}$ with the properties claimed by the lemma. Around any $x \in \tilde{\mathcal{M}}$, there exists a local coordinate system $(z_1, \ldots, z_{n+2})$ such that $z_1 = \varphi_1$, $z_2 = \varphi_2$ in which $\Sigma$ takes the chained form

$\Sigma_{\text{chain}}$ (1.2.2) and $x$ is transformed into $0 \in \mathbb{R}^{n+2}$. Then by a simple computation we get

$$\mathcal{D}^{(n-1)} = \text{span}\,\{\frac{\partial}{\partial z_{n+2}}, \frac{\partial}{\partial z_{n+1}}, \ldots, \frac{\partial}{\partial z_3}, \frac{\partial}{\partial z_1} + z_3\frac{\partial}{\partial x_2}\}.$$

Hence

$$\mathcal{L} = (\text{span}\,\{\mathrm{d}\varphi_1, \mathrm{d}\varphi_2\})^{\perp} = (\text{span}\,\{\mathrm{d}z_1, \mathrm{d}z_2\})^{\perp} = \text{span}\,\{\frac{\partial}{\partial z_{n+2}}, \frac{\partial}{\partial z_{n+1}}, \ldots, \frac{\partial}{\partial z_3}\} \subset \mathcal{D}^{(n-1)}.$$

Now consider a point $x \in \mathcal{M} \setminus \tilde{\mathcal{M}}$. The distributions $\mathcal{L} = (\text{span}\,\{\mathrm{d}\varphi_1, \mathrm{d}\varphi_2\})^{\perp}$ and $\mathcal{D}^{(n-1)}$ are of constant rank and they are well defined at any point of $\mathcal{M}$. Assume that $\mathcal{L}(x) = (\text{span}\,\{\mathrm{d}\varphi_1(x), \mathrm{d}\varphi_2(x)\})^{\perp} \not\subset \mathcal{D}^{(n-1)}(x)$, then, because of constant ranks of $\mathcal{L}$ and $\mathcal{D}^{(n-1)}$, the inclusion does not hold at any $\tilde{x}$ in an open neighborhood $\tilde{\mathcal{O}}$ of $x$. Clearly, $\tilde{\mathcal{M}} \cap \tilde{\mathcal{O}} \neq \emptyset$ which gives a contradiction. In conclusion, we have $\mathcal{L}(x) = \text{span}\,\{\mathrm{d}\varphi_1(x), \mathrm{d}\varphi_2(x)\}^{\perp} \subset \mathcal{D}^{(n-1)}(x)$ for any point $x$ in $\mathcal{M}$. Observe that $\mathcal{D}(x_0) \not\subset \mathcal{L}(x_0)$ holds if and only if there exists a vector field $g_1 \in \mathcal{D}$ such that

$$(L_{g_1}\varphi_1(x_0), L_{g_1}\varphi_2(x_0)) \neq (0, 0),$$

which is just the item (iii) of Theorem 1.3.2. This shows the equivalence of the two singular loci defined in Theorem 1.3.2 and Theorem 1.3.3 and proves, due to Theorem 1.3.2, the necessity of (iii)$'$. □

### 1.4.5 Proof of Theorem 1.3.4

**Proof:** The results of Section 1.3.2 and Section 1.3.3 show that for a given arbitrary smooth function $\varphi_1$ such that $L_c\varphi_1 = 0$, for any $c \in \mathcal{C}_{n-1}$, and $L_g\varphi_1(x_0) \neq 0$, there always exists a function $\varphi_2$, independent with $\varphi_1$, such that $(\varphi_1, \varphi_2)$ is an $x$-flat output of $\Sigma$ at $(x_0, u_0)$, $u_0 \notin U_{\text{sing}}(x_0)$. By Proposition 1.2.3, we can introduce new coordinates by $z_1 = \varphi_1$, $z_2 = \varphi_2$ and complete them to a coordinate system $(z_1, \ldots, z_{n+2})$ in which our original system $\Sigma$ takes, via a feedback transformation, the chained form

$$\Sigma_{\text{chain}}: \begin{cases} \dot{z}_1 &= v_1 \\ \dot{z}_2 &= z_3 \cdot v_1 \\ \dot{z}_3 &= z_4 \cdot v_1 \\ &\vdots \\ \dot{z}_{n+1} &= z_{n+2} \cdot v_1 \\ \dot{z}_{n+2} &= v_2 \end{cases}$$

Suppose that there exists another function $\tilde{\varphi}_2$ such that $(\varphi_1, \tilde{\varphi}_2) = (z_1, \tilde{\varphi}_2)$ is an $x$-flat output of $\Sigma$ at $(x_0, u_0)$, $u_0 \notin U_{\text{sing}}(x_0)$. Clearly, $(\varphi_1, \tilde{\varphi}_2)$ is also an $x$-flat output of $\Sigma_{\text{chain}}$ at $(0, v_0)$, $v_0 \notin U_{\text{sing}}(0)$. Take $g = \frac{\partial}{\partial z_1} + z_3\frac{\partial}{\partial z_2} + \cdots + z_{n+2}\frac{\partial}{\partial z_{n+1}}$ and let $\mathcal{C}_{n-1}$ be the characteristic distribution of $\mathcal{D}^{(n-1)}$, where $\mathcal{D}$ is the associated distribution of

$\Sigma_{\text{chain}}$. By the result of Section 1.3.2 and Section 1.3.3, the function $\tilde{\varphi}_2$ must satisfy the equations

$$\begin{aligned} L_c\tilde{\varphi}_2 &= 0, \quad \forall\, c \in \mathcal{C}_{n-1}, \\ L_v\tilde{\varphi}_2 &= 0. \end{aligned}$$

where $v = (L_g\varphi_1)[c_{n-1}, g] - (L_{[c_{n-1},g]}\varphi_1)g$ and $c_{n-1} = \dfrac{\partial}{\partial z_4}$. Solving the above system of equations, we get $\tilde{\varphi}_2 = \{\tilde{\varphi}_2(z_1, z_2) : \dfrac{\partial\tilde{\varphi}_2}{\partial z_2}(0) \neq 0\}$. Thus

$$\begin{aligned} \operatorname{span}\{\mathrm{d}\varphi_1, \mathrm{d}\varphi_2\}(z) &= \operatorname{span}\{\mathrm{d}z_1, \mathrm{d}z_2\}(z) \\ &= \operatorname{span}\{\mathrm{d}z_1, \frac{\partial\tilde{\varphi}_2}{\partial z_1}(z)\mathrm{d}z_1 + \frac{\partial\tilde{\varphi}_2}{\partial z_2}(z)\mathrm{d}z_2\}(z) \\ &= \operatorname{span}\{\mathrm{d}\varphi_1, \mathrm{d}\tilde{\varphi}_2\}(z), \end{aligned}$$

for any $z$ in a neighborhood of 0. Correspondingly, for the original system $\Sigma$, the above equality is true for any point in a neighborhood of $x_0$. □

Let $Flat(\varphi_1)$ be the codistribution spanned by the differentials of all $x$-flat outputs determined by function $\varphi_1$, i.e., $Flat(\varphi_1) = \operatorname{span}\{\mathrm{d}\varphi_1, \mathrm{d}\varphi_2\}$, where $\varphi_2$ is any function such that $(\varphi_1, \varphi_2)$ is an $x$-flat output. Clearly, $Flat(\varphi_1)$ is well defined because of Theorem 1.3.4.

**Corollary 1.4.4** *If* $Flat(\varphi_1) = \operatorname{span}\{\mathrm{d}\varphi_1, \mathrm{d}\varphi_2\}$, *then*

$$Flat(\varphi_1) = Flat(\varphi_2)$$

This corollary is easy to prove and here we omit its proof.

# Chapter 2

# The geometry and flatness property of the $n$-bar system

## 2.1 Introduction

The well known $n$-trailer system was proposed by [34] to model a unicycle-like robot towing trailers. This nonholonomic model has attracted a lot of attention and has been a source of inspiration to study its various properties: controllability ([34]), structure ([27], [42], [52], [73]), flatness ([13], [26]), motion planning and tracking ([35], [46], [53]), optimal control ([35]), etc. In this chapter we propose its generalization, which we call the $n$-bar system, consisting of a "train" of $n$ rigid bars subject to nonholonomic constraints (see a detailed description in Section 2.2.2 below). We study the geometry of the model of the $n$-bar system and prove that around any regular configuration (that is, none of the angles between two consecutive bars is $\pm\frac{\pi}{2}$), the control system associated to the $n$-bar system is feedback equivalent to the $m$-chained form. This implies that the $n$-bar system is flat around any regular configuration and we show that the cartesian position of the source point of the last (from the top) bar is a flat output. We show also that all other minimal flat outputs are equivalent to that one. This is in contrast with the $n$-trailer system for which the position of the last trailer is also a flat output but there is a whole family of not equivalent flat outputs (parameterized by one function of three variables, see Chapter 1). As a by-product of our consideration we deduce the global controllability of the $n$-bar system since it is accessible at any (regular or not) configurations. We send the reader to [71] and [72] for another, although similar, model for the $n$-bar system (called there an articulated arm) and for a detailed analysis of singular configurations.

This chapter is organized as follows. We define our model of the $n$-bar system in Section 2.2. We provide geometric notions and recall a characterization of Cartan distributions $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$ ($m$-chained forms) and given our first main result, equivalence of the $n$-bar system in $\mathbb{R}^{m+1}$ to the $m$-chained form, in Section 2.3 and its proofs in Section 2.4. The flatness property of the $n$-bar system is analyzed in Section 2.5. In Section 2.6, we give a proof of technical Lemma 2.4.1 which is used in Section 2.4.

## 2.2 $n$-bar system in $\mathbb{R}^{m+1}$

In this section, a model of the $n$-bar system will be proposed with which we will work throughout this chapter. In order to do that, we analyze first the model of a rigid bar moving in $\mathbb{R}^3$ and then we will extend it to the general case of $n$-bar system moving in $\mathbb{R}^{m+1}$.

### 2.2.1 Model of a rigid bar moving in $\mathbb{R}^3$

Consider a rigid bar moving in $\mathbb{R}^3$ which is described by a vector $\overrightarrow{P_0P_1}$, where $P_i = (x_i^1, x_i^2, x_i^3)$ for $i = 0, 1$. It it assumed that the endpoint $P_1$ of the bar can rotate freely when the source point $P_0$ remains fixed and that the bar moves in such a way that the direction of the instantaneous velocity of $P_0$ is parallel to the direction of the bar $\overrightarrow{P_0P_1}$. For simplicity, the length of the bar is assumed to be one.

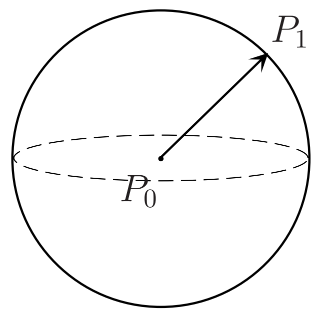


Figure 2.1: the rigid bar in $\mathbb{R}^3$

No matter how the bar moves, its position can always be determined uniquely by the source point $P_0$ and the endpoint $P_1$. Therefore the configuration of the bar system can be described completely by $x = (x_0^1, x_0^2, x_0^3, x_1^1, x_1^2, x_1^3) \in \mathbb{R}^6$. At the same time, the assumption of $|\overrightarrow{P_0P_1}| = 1$ implies that these six variables must satisfy the following holonomic constraint

$$\psi(x) = (x_1^1 - x_0^1)^2 + (x_1^2 - x_0^2)^2 + (x_1^3 - x_0^3)^2 - 1 = 0 \qquad (2.2.1)$$

which is the equation of the unit sphere $S^2$ in $\mathbb{R}^3$ centered at $P_0$. The holonomic constraint (2.2.1) reduces the dimension of the configuration space of the bar system to five. In fact, if we consider only the rotation of the endpoint $P_1$ with respect to the source point $P_0$, the former belongs always to the sphere $S^2$ with center $P_0$ and radius 1. Hence the configuration of the bar can be described completely by the position of the point $P_0 \in \mathbb{R}^3$ together with the position of the point $P_1$ in a unite sphere $S^2$. In other words, the true configuration space is the cartesian product $Q = \mathbb{R}^3 \times S^2$, which is a submanifold of $\mathbb{R}^6$ given by

$$Q = \{x \in \mathbb{R}^6 : \psi(x) = 0\}.$$

Notice that the equation (2.2.1) implies that there exists an index $j$, for $1 \le j \le 3$, such that $x_1^j - x_0^j \neq 0$, say $j = 1$. Then the assumption that the direction of the instantaneous velocity $(\dot{x}_0^1, \dot{x}_0^2, \dot{x}_0^3)$ of $P_0$ is always parallel to $\overrightarrow{P_0P_1}$ introduces two nonholonomic constraints:

$$\begin{cases} (x_1^2 - x_0^2)\dot{x}_0^1 - (x_1^1 - x_0^1)\dot{x}_0^2 = 0 \\ (x_1^3 - x_0^3)\dot{x}_0^1 - (x_1^1 - x_0^1)\dot{x}_0^3 = 0 \end{cases}$$

which geometrically means that the velocity $(\dot{x}_0^1, \dot{x}_0^2, \dot{x}_0^3, \dot{x}_1^1, \dot{x}_1^2, \dot{x}_1^3)^\top$ is always annihilated by the two following differential 1-forms:

$$\begin{aligned} &(x_1^2 - x_0^2)\mathrm{d}x_0^1 - (x_1^1 - x_0^1)\mathrm{d}x_0^2 \\ &(x_1^3 - x_0^3)\mathrm{d}x_0^1 - (x_1^1 - x_0^1)\mathrm{d}x_0^3. \end{aligned}$$

The distribution $\mathcal{E}$ on $\mathbb{R}^6$, annihilated by the above two differential forms, is given by

$$\mathcal{E} = \mathrm{span}\,\{g_1, \cdots, g_4\},$$

where

$$\begin{aligned} g_1 &= (x_1^1 - x_0^1)\frac{\partial}{\partial x_0^1} + (x_1^2 - x_0^2)\frac{\partial}{\partial x_0^2} + (x_1^3 - x_0^3)\frac{\partial}{\partial x_0^3} \\ g_2 &= \frac{\partial}{\partial x_1^1}, \quad g_3 = \frac{\partial}{\partial x_1^2}, \quad g_4 = \frac{\partial}{\partial x_1^3}, \end{aligned}$$

and it defines the control-linear (driftless, in other words) system on $\mathbb{R}^6$

$$\Delta_{\mathrm{bar}}: \quad \dot{x} = \sum_{i=1}^{4} g_i(x)v_i, \quad x \in \mathbb{R}^6,$$

where $u_1, u_2, u_3, u_4$ are four arbitrary functions of time can be interpreted as controls (and will do so throughout). In order to obtain a kinematic model of the rigid bar we have to constrain the system $\Delta_{\mathrm{bar}}$ to its true configuration space $Q$. The crucial observation that the intersection $TQ \cap \mathcal{E}$ defines a distribution $\mathcal{D}$ of constant rank equal to 3 on $Q$ and thus gives rise to the kinematic model of the rigid bar moving in $\mathbb{R}^3$,

$$\Gamma_{\mathrm{bar}}: \quad \dot{q}(t) \in \mathcal{D}(q(t)), \quad q \in \mathbb{R}^3 \times S^2.$$

Locally, we can choose three independent vector fields $f_1, f_2, f_3$ on $Q$ such that $\mathcal{D} = \operatorname{span}\{f_1, f_2, f_3\}$ which yields a local representation of the bar system

$$\Gamma_{\text{bar}}: \quad \dot{q} = f_1(q)u_1 + f_2(q)u_2 + f_3(q)u_3, \quad q \in \mathbb{R}^3 \times S^2.$$

To summarize, the model of the rigid bar system $\Gamma_{\text{bar}}$ is defined by the system $\Delta_{\text{bar}}$, together with the holonomic constraint

$$\psi(x) = (x_1^1 - x_0^1)^2 + (x_1^2 - x_0^2)^2 + (x_1^3 - x_0^3)^2 - 1 = 0.$$

Notice that we do not express the distribution $\mathcal{D}$ explicitly, however, all its properties can be deduced by means of the distribution $\mathcal{E}$ and the holonomic constraint (2.2.1).

The configuration space $\mathbb{R}^3 \times S^2$ of the rigid bar system $\Gamma_{\text{bar}}$ can also be interpreted in a natural way through the classical rigid body theory. To analyze the motion of a rigid body, we choose an inertial reference frame $(O, x, y, z)$ and a body reference frame $(O', x', y', z')$ that is fixed to move with the body (thus the point $O'$ coincides with the source point $P_0$ of the rigid bar system). Then the position of the body is specified by the vector $r = \overrightarrow{O'O} \in \mathbb{R}^3$, along with the orientation of the orthogonal frame $\{x', y', z'\}$ relative to $\{x, y, z\}$, which is determined by a $(3 \times 3)$ orthogonal matrix $R$ belonging to the group of special orthogonal matrices

$$SO(3, \mathbb{R}) = \{R \in Gl(3, \mathbb{R}) \mid RR^\top = \mathrm{I}, \quad \det R = 1\}.$$

Now consider the rigid bar moving in $\mathbb{R}^3$ and let the origin $O'$ be the source point $P_0$. Notice that the bar $\overrightarrow{P_0P_1}$ can be looked at as such a thin rigid body that its rotations are reduced to the rotations of the unit vector $\overrightarrow{O'z'}$ of the $z'$-axis, which is $\overrightarrow{P_0P_1} = \overrightarrow{O'z'}$. The configuration space is thus the quotient $SO(3)/SO(2) \cong S^2$ with $SO(2)$ corresponding to the rotations preserving the plane $x'O'y'$. We find again that the configuration of the bar system is described by the position of the point $O' = P_0 \in \mathbb{R}^3$ together with the position of a point of a unit sphere $S^2$, that is, the configuration space is $\mathbb{R}^3 \times S^2$.

### 2.2.2 Model of the $n$-bar system moving in $\mathbb{R}^{m+1}$

In this section we will consider the $n$-bar system moving in $\mathbb{R}^{m+1}$, as shown on Figure 2.2, and derive a kinematic model for it. It is assumed that all $n$ components of the $n$-bar system are attached in such a way that $P_i$ is the source point of the $(i+1)$-th bar and simultaneously the endpoint of the $i$-th bar and that the instantaneous velocity of the point $P_i$ is parallel to the vector $\overrightarrow{P_iP_{i+1}}$, for $0 \le i \le n-1$. Furthermore, each rigid bar is assumed to have length one. The coordinates of $P_i$ in $\mathbb{R}^{m+1}$ are given by

$$P_i = (x_i^1, x_i^2, \cdots, x_i^{m+1}), \quad 0 \le i \le n.$$

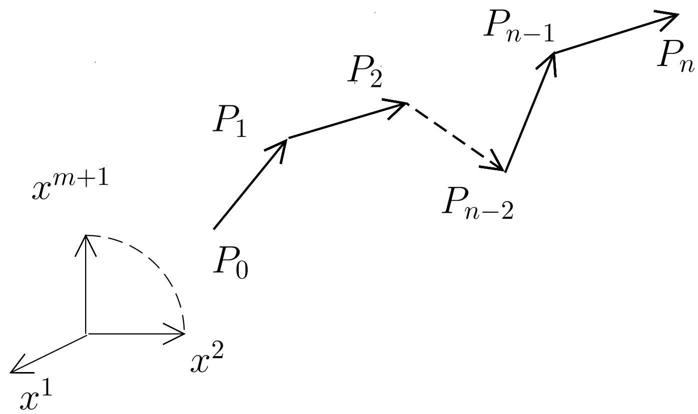


Figure 2.2: the $n$-bar system in $\mathbb{R}^{m+1}$

Clearly, the configuration of the $n$-bar system can be described completely by the $(n+1)(m+1)$ coordinates

$$(x_0^1, \cdots, x_0^{m+1}, x_1^1, \cdots, x_1^{m+1}, \cdots, x_n^1, \cdots, x_n^{m+1}) \in X = \mathbb{R}^{(n+1)(m+1)}.$$

Due to the assumption $|\overrightarrow{P_iP_{i+1}}| = 1$, for $0 \leq i \leq n-1$, we have the following holonomic constraints

$$\Psi(x) = 0,$$

where $\Psi = \{\Psi_1, \cdots, \Psi_n\} : X = \mathbb{R}^{(n+1)(m+1)} \to \mathbb{R}^n$ is given by

$$\left\{ \begin{array}{lcl} \Psi_1(x) & = & (x_1^1 - x_0^1)^2 + (x_1^2 - x_0^2)^2 + \cdots + (x_1^{m+1} - x_0^{m+1})^2 - 1 \\ \Psi_2(x) & = & (x_2^1 - x_1^1)^2 + (x_2^2 - x_1^2)^2 + \cdots + (x_2^{m+1} - x_1^{m+1})^2 - 1 \\ & \vdots & \\ \Psi_n(x) & = & (x_n^1 - x_{n-1}^1)^2 + (x_n^2 - x_{n-1}^2)^2 + \cdots + (x_n^{m+1} - x_{n-1}^{m+1})^2 - 1. \end{array} \right. \tag{2.2.2}$$

Under these $n$ holonomic constraints, the true configuration space of the $n$-bar system becomes the regular embedded submanifold $Q = \mathbb{R}^{m+1} \times (S^m)^n \subset X$ defined by

$$Q = \{x \in X : \Psi(x) = 0\}.$$

Moreover, the equation $\Psi(x) = 0$ implies that for any $1 \leq i \leq n$, there always exists $1 \leq \sigma(i) \leq m+1$, such that

$$x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)} \neq 0.$$

Now the assumption that the instantaneous velocity of the point $P_i$ is parallel to the vector $\overrightarrow{P_iP_{i+1}}$, for $0 \leq i \leq n-1$, imposes the following nonholonomic constraints on the $n$-bar system: the velocity of the system along any trajectory is annihilated by the following differential 1-forms

$$\Omega_i^j = (x_i^j - x_{i-1}^j)\mathrm{d}x_{i-1}^{\sigma(i)} - (x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})\mathrm{d}x_{i-1}^j,$$

for $1 \leq i \leq n$, $1 \leq j \leq m+1$ and $j \neq \sigma(i)$. The distribution $\mathcal{E}$ annihilated by all forms $\Omega_i^j$ is given by

$$\mathcal{E} = \bigcap_{i,j} \ker \Omega_i^j = \operatorname{span}\{g_1, \ldots, g_{n+m+1}\},$$

where

$$\begin{aligned}
g_1 &= (x_1^1 - x_0^1)\frac{\partial}{\partial x_0^1} + \cdots + (x_1^{m+1} - x_0^{m+1})\frac{\partial}{\partial x_0^{m+1}} \\
g_2 &= (x_2^1 - x_1^1)\frac{\partial}{\partial x_1^1} + \cdots + (x_2^{m+1} - x_1^{m+1})\frac{\partial}{\partial x_1^{m+1}} \\
&\vdots \\
g_n &= (x_n^1 - x_{n-1}^1)\frac{\partial}{\partial x_{n-1}^1} + \cdots + (x_n^{m+1} - x_{n-1}^{m+1})\frac{\partial}{\partial x_{n-1}^{m+1}} \\
g_{n+i} &= \frac{\partial}{\partial x_n^i}, \quad 1 \leq i \leq m+1,
\end{aligned} \tag{2.2.3}$$

which defines the control-linear system on $X = \mathbb{R}^{(n+1)(m+1)}$

$$\Delta : \quad \dot{x} = \sum_{i=1}^{n+m+1} g_i(x) v_i, \quad x \in X. \tag{2.2.4}$$

To obtain a kinematic model of the $n$-bar system we have to constrain the system $\Delta$ to the regular submanifold $Q \subset X$. Consider the embedding $\Phi : Q \to X$ such that $\Phi(q) = q$, for any $q \in Q$. Let $\mathcal{J}$ be the codistribution spanned by all differential forms $\Omega_i^j$, i.e.,

$$\mathcal{J} = \operatorname{span}\{\Omega_i^j, \quad 1 \leq i \leq n, \ 1 \leq j \leq m+1, \ j \neq \sigma(i)\}. \tag{2.2.5}$$

Clearly $\mathcal{J} = \mathcal{E}^{\perp}$ and the pull back $\Phi^*$ maps $\mathcal{J}$ into a codistribution $\mathcal{I}$ on $Q$,

$$\mathcal{I} = \Phi^* \mathcal{J} = \operatorname{span}\{\omega_i^j, \quad 1 \leq i \leq n, \ 1 \leq j \leq m+1, \ j \neq \sigma(i)\}, \tag{2.2.6}$$

where $\omega_i^j = \Phi^* \Omega_i^j$. Define a distribution $\mathcal{D}$ by $\mathcal{D} = \mathcal{I}^{\perp}$. Notice that $\mathcal{D}$ is just the intersection $TQ \cap \mathcal{E}$ and is of constant rank equal to $m+1$ (see Section 2.4.2 for details) and thus gives rise to a control-linear system

$$\Gamma : \quad \dot{q} = \sum_{i=0}^{m} f_i(q) u_i, \quad q \in \mathbb{R}^{m+1} \times (S^m)^n, \tag{2.2.7}$$

where locally $\mathcal{D} = \operatorname{span}\{f_0, \ldots, f_m\}$, which describes completely the $n$-bar system moving in $\mathbb{R}^{m+1}$. To summarize, the model $\Gamma$ describing the $n$-bar moving in $\mathbb{R}^{m+1}$, for $m \geq 1$, is defined by the control-linear system $\triangle$, given by (2.2.3)-(2.2.4), together with the following holonomic constraint

$$\Psi(x) = (\Psi_1(x), \ldots, \Psi_n(x)) = 0,$$

where $\Psi_i$, $1 \leq i \leq n$, are defined by (2.2.2). Notice that, like in the case of $\mathbb{R}^3$ described in Section 2.2.1, we give explicit expression neither for the distribution $\mathcal{D}$ nor for the vector fields $f_i$, for $0 \leq i \leq m$. All their properties will be formulated and analyzed in terms of the distribution $\mathcal{E}$ and the holonomic constraint $\Psi(x) = 0$.

Another, although similar model for the $n$-bar system (called articulated arm) has been very recently introduced and studied in [71], and [72]; in the former a detailed analysis of the singular locus (see also section 2.3.2) has been performed.

The presented model of the $n$-bar system in $\mathbb{R}^{m+1}$ is a natural generalization of the well known $n$-trailer system on $\mathbb{R}^2$. The latter is a model for a unicycle-like mobile robot towing $n$-trailers such that the tow hook of each trailer is located at the center of its unique axle (with the assumption that the distances between any two consecutive trailers are equal). The $n$-trailer system is subject to nonholonomic constraints: it is assumed that the wheels of each individual trailer are aligned with the body and are not allowed to slip [34]. This model and its control properties have attracted a lot of attention (see the books [35] and [36]; and the papers [13], [27], [46], [52], [73]).

Clearly the nonholonomic constraint that the wheel cannot slip on the plane $\mathbb{R}^2$ can be equivalently rephrased that the instantaneous velocity of the middle point of the $i$-th trailer axle, say point $P_i$, is parallel to the vector $\overrightarrow{P_iP_{i+1}}$ joining the two consecutive axles. These are exactly our nonholonomic constraints imposed for the $n$-bar system, the only difference being to allow the vectors $\overrightarrow{P_iP_{i+1}}$ to move in $\mathbb{R}^{m+1}$ and not on the plane $\mathbb{R}^2$.

## 2.3 Equivalence of the $n$-bar system to the $m$-chained form

### 2.3.1 Characterization of Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$

Consider an arbitrary distribution $\mathcal{D}$. The *derived flag* of $\mathcal{D}$ is the sequence of modules of vector fields $\mathcal{D}^{(0)} \subset \mathcal{D}^{(1)} \subset \cdots$ defined inductively by

$$\mathcal{D}^{(0)} = \mathcal{D} \ \text{ and } \ \mathcal{D}^{(i+1)} = \mathcal{D}^{(i)} + [\mathcal{D}^{(i)}, \mathcal{D}^{(i)}], \text{ for } i \geq 0.$$

The *Lie flag* of $\mathcal{D}$ is the sequence of modules of vector fields $\mathcal{D}_0 \subset \mathcal{D}_1 \subset \cdots$ defined inductively by

$$\mathcal{D}_0 = \mathcal{D} \ \text{ and } \mathcal{D}_{i+1} = \mathcal{D}_i + [\mathcal{D}_0, \mathcal{D}_i], \quad \text{for } i \geq 0.$$

In general, the derived and Lie flags are different; though for any point $p$ in the underlying manifold the inclusion $\mathcal{D}_i(p) \subset \mathcal{D}^{(i)}(p)$ holds, for $i \geq 0$.

Two distributions $D$ and $\tilde{\mathcal{D}}$ defined on two manifolds $M$ and $\tilde{M}$, respectively, are *equivalent* if there exists a smooth diffeomorphism $\varphi$ between $M$ and $\tilde{M}$ such that $(\varphi_*\mathcal{D})(\tilde{p}) = \tilde{\mathcal{D}}(\tilde{p})$, for each point $\tilde{p}$ in $\tilde{M}$.

An alternative description of the above definitions can also be given using the dual language of differential forms. A *codistribution* $\mathcal{I}$ of rank $s$ on a smooth manifold $M$ (or a *Pfaffian system*) is a map that assigns smoothly to each point $p$ in $M$ a linear subspace $\mathcal{I}(p) \subset T_p^*M$ of dimension $s$. Such a field of cotangent $s$-planes is spanned locally by $s$ pointwise linearly independent smooth differential 1-forms $\omega_1, \ldots, \omega_s$ on $M$, which will be denoted by $\mathcal{I} = \operatorname{span}\{\omega_1, \ldots, \omega_s\}$. In the case of constant rank we can (locally) identity codistribution and Pfaffian system which we will do throughout this chapter. Two codistributions (Pfaffian systems) $\mathcal{I}$ and $\tilde{\mathcal{I}}$ defined on two manifolds $M$ and $\tilde{M}$, respectively, are equivalent if there exists a smooth diffeomorphism $\varphi$ between $M$ and $\tilde{M}$ such that $\mathcal{I}(p) = (\varphi^*\tilde{\mathcal{I}})(p)$ for each point $p$ in $M$.

For a codistribution $\mathcal{I}$, its *derived flag* $\mathcal{I}^{(0)} \supset \mathcal{I}^{(1)} \supset \cdots$ can be defined by

$$\mathcal{I}^{(0)} = \mathcal{I} \ \text{ and } \ \mathcal{I}^{(i+1)} = \{\omega \in \mathcal{I}^{(i)} : \mathrm{d}\omega \equiv 0 \bmod \mathcal{I}^{(i)}\}, \ \text{for } i \geq 0,$$

provided that each element $\mathcal{I}^{(i)}$ of this sequence has constant rank. In this case, it is immediate to see that the derived flag of the distribution $\mathcal{D} = \mathcal{I}^{\perp}$ coincides with the sequence of distributions that annihilate the elements of the derived flag of $\mathcal{I}$, that is

$$\mathcal{D}^{(i)} = (\mathcal{I}^{(i)})^{\perp}, \ \text{for } i \geq 0.$$

Consider $J^n(\mathbb{R}, \mathbb{R}^m)$, the space of $n$-jets of smooth maps from $\mathbb{R}$ into $\mathbb{R}^m$ and denote its canonical coordinates by

$$x_0^0, x_1^0, \ldots, x_m^0, x_1^1, \ldots, x_m^1, \ldots, x_1^n, \ldots, x_m^n,$$

where $x_0^0$ represents the independent variable and $x_i^0$ for $1 \leq i \leq m$, represent the dependent variables, and $x_i^j$, for $1 \leq i \leq m$ and $1 \leq j \leq n$, correspond to the ordinary derivatives $\dfrac{\mathrm{d}^j x_i^0}{\mathrm{d}(x_0^0)^j}$. Any smooth map $\varphi$ from $\mathbb{R}$ into $\mathbb{R}^m$ defines a submanifold in $J^n(\mathbb{R}, \mathbb{R}^m) \cong \mathbb{R}^{(n+1)m+1}$ by the relation $x_i^j = \varphi_i^{(j)}(x_0^0)$, for $1 \leq i \leq m$ and $0 \leq j \leq n$, where $\varphi_i^{(j)}(x_0^0)$ denotes the $j$-th derivative with respect to $x_0^0$ of the $i$-th component $\varphi_i$ of $\varphi$. This submanifold is called the $n$-graph of $\varphi$.

The *Cartan distribution* on $J^n(\mathbb{R}, \mathbb{R}^m)$, denote by $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$, is the completely nonholonomic distribution spanned by the following family of vector fields

$$\frac{\partial}{\partial x_0^0} + \sum_{j=0}^{n-1} \sum_{i=1}^{m} x_i^{j+1} \frac{\partial}{\partial x_i^j}, \ \frac{\partial}{\partial x_1^n}, \ \ldots, \ \frac{\partial}{\partial x_m^n}$$

or, equivalently, annihilated by the following family of differential forms:

$$\mathrm{d}x_i^j - x_i^{j+1}\mathrm{d}x_0^0, \quad 0 \leq j \leq n-1, \quad 1 \leq i \leq m.$$

It turns out that all $n$ -graphs are integral submanifolds of dimension 1, that is integral curves, of the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$.

The problem of characterizing distributions that are locally equivalent to the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$ has been studied and solved in the following way by Pasillas-Lépine and Respondek [54] (see also [42], [51], [79]). Recall that the *Engel rank*, at a point $p$, of a codistribution $\mathcal{I} = \operatorname{span}\{\omega_1, \cdots, \omega_s\}$ is the largest integer $\rho$ such that there exists a 1-form $\alpha$ in $\mathcal{I}$ for which $((d\alpha)^\rho \wedge \omega_1 \wedge \cdots \wedge \omega_s)(p) \neq 0$.

**Theorem 2.3.1** *A distribution $\mathcal{D}$ of rank $m+1$, with $m \geq 2$, on a manifold $M$ of dimension $(n+1)m+1$ is equivalent, in a small enough neighborhood of a point $p$ in $M$, to the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$ if and only if the following conditions hold:*

(i) $\mathcal{D}^{(n)}(p) = T_pM$;
(ii) $\mathcal{D}^{(n-1)}$ *is of constant rank* $nm+1$ *and contains an involutive subdistribution* $\mathcal{L}_{n-1}$ *that has constant corank one in* $\mathcal{D}^{(n-1)}$;
(iii) $\mathcal{D}(p)$ *is not contained in* $\mathcal{L}_{n-1}(p)$.

*Moreover, if $m \geq 3$, $\mathcal{L}_{n-1}$ exists if and only if the Engel rank of $(\mathcal{D}^{(n-1)})^\perp$ equals one and* $\operatorname{rank} \mathcal{C}(\mathcal{D}^{(n-1)}) = (n-1)m$ *and is given as*

$$\mathcal{L}_{n-1} = \mathcal{F}_1 + \cdots + \mathcal{F}_m,$$

*where $\mathcal{F}_i = \{f \in \mathcal{D}^{(n-1)} : f \lrcorner \mathrm{d}\omega_i \in (\mathcal{D}^{(n-1)})^\perp\}$ and $\omega_i$'s are any differential 1-forms such that*

$$\mathcal{I}^{(n-1)} = (\mathcal{D}^{(n-1)})^\perp = \operatorname{span}\{\omega_1, \ldots, \omega_m\}.$$

**Remarks 1**. The involutive subdistribution $\mathcal{L}_{n-1}$, whose existence is claimed by (ii), is unique (if it exists) and it will be called the canonical involutive subdistribution in $\mathcal{D}^{(n-1)}$. The uniqueness and involutivity of $\mathcal{L}_{n-1}$ follow from a result of Bryant [7] and have been shown in [54].
**Remarks 2**. Item (i) and (ii) describe the essential geometric property of distributions equivalent to the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$ while the condition (iii) distinguishes regular points $p$ at which $\mathcal{D}(p) \not\subset \mathcal{L}_{n-1}(p)$ form singular points, where this last condition is violated.

The case $m = 1$ is excluded from Theorem 2.3.1 because if an involutive subdistribution of corank one $\mathcal{L}_{n-1} \subset \mathcal{D}^{(n-1)}$ exists it cannot be unique and therefore there is no a canonical one. However, a "non-canonical" version of Theorem 2.3.1 holds for $m = 1$ as well, as proved in [54]: a rank-two distribution is equivalent to the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R})$, called also a chained normal form or Goursat normal form if and only if there exists a distribution $\mathcal{L}_{n-1}$ satisfying the conditions (i), (ii) and (iii) of Theorem 2.3.1.

Let a distribution $\mathcal{D}$ of rank $m+1$, with $m \geq 1$, satisfy the items (i) and (ii) of Theorem 2.3.1. The *regular locus* of $\mathcal{D}$, denoted by *Reg*$(\mathcal{D})$, is the subset of $M$ consisting of points at which $\mathcal{D}$ is equivalent to the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$ at

$0 \in \mathbb{R}^{(n+1)(m+1)}$. It can be proved that $Reg(\mathcal{D})$ is an open and dense subset of $M$; its complement, called *the singular locus* of $\mathcal{D}$, will be denoted by $Sing(\mathcal{D})$. In the case $m \geq 2$, since $\mathcal{L}_{n-1}$ is unique we clearly have

$$Reg(\mathcal{D}) = \{p \in M : \mathcal{D}(p) \not\subset \mathcal{L}_{n-1}(p)\}$$

and

$$Sing(\mathcal{D}) = \{p \in M : \mathcal{D}(p) \subset \mathcal{L}_{n-1}(p)\}.$$

### 2.3.2 Main result: equivalence of the $n$-bar system to the $m$-chained form

Consider two driftless systems

$$\Sigma : \quad \dot{x} = \sum_{i=0}^{m} f_i(x)u_i = f(x)u, \quad x \in M,$$

and

$$\tilde{\Sigma} : \quad \dot{\tilde{x}} = \sum_{i=0}^{m} \tilde{f}_i(\tilde{x})\tilde{u}_i = \tilde{f}(\tilde{x})\tilde{u}, \quad \tilde{x} \in \tilde{M},$$

where $u = (u_0, \ldots, u_m)^\top \in \mathbb{R}^{m+1}$, $\tilde{u} = (\tilde{u}_0, \ldots, \tilde{u}_m)^\top \in \mathbb{R}^{m+1}$ and the rows $f = (f_0, \ldots, f_m)$ and $\tilde{f} = (\tilde{f}_0, \ldots, \tilde{f}_m)$ are formed by $C^\infty$-smooth vector fields $f_i$ and $\tilde{f}_i$, $0 \leq i \leq m$, on $M$ and $\tilde{M}$, respectively. We say that $\Sigma$ and $\tilde{\Sigma}$ are feedback equivalent if there exists a diffeomorphism $\varphi : M \to \tilde{M}, \tilde{x} = \varphi(x)$, and a feedback transformation $u = \beta(x)\tilde{u}$ where $\beta(x)$ is an invertible $C^\infty$-smooth $(m+1) \times (m+1)$-matrix such that

$$\mathrm{D}\varphi(x) \cdot f(x)\beta(x) = \tilde{f}(\varphi(x)).$$

**Definition 2.3.2** An $(m+1)$-input driftless control system $\Sigma : \ \dot{x} = \sum_{i=0}^{m} u_i f_i(x)$, defined on $\mathbb{R}^{(n+1)m+1}$, is said to be in the $m$-chained form if it is represented by

$$\begin{array}{lllllllll}
\dot{x}_0^0 & = & u_0 & \dot{x}_1^0 & = & x_1^1 u_0 & \cdots & \dot{x}_m^0 & = & x_m^1 u_0 \\
 & & & & \cdots & & \cdots & & \cdots & \\
 & & & \dot{x}_1^{n-1} & = & x_1^n u_0 & \cdots & \dot{x}_m^{n-1} & = & x_m^n u_0 \\
 & & & \dot{x}_1^n & = & u_1 & \cdots & \dot{x}_m^n & = & u_m.
\end{array}$$

A system in the $m$-chained form is also called the *canonical contact system* on $J^n(\mathbb{R}, \mathbb{R}^m)$. In fact, the vector fields $f_0, \ldots, f_m$ of the $m$-chained form coincide with those generating the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$ given in Section 2.3.1.

To any control-linear system $\Sigma$, we associate the distribution spanned by all its vector fields $\mathcal{D}_\Sigma = \operatorname{span}\{f_0, \ldots, f_m\}$. The (local) feedback equivalence of $\Sigma$ and $\tilde{\Sigma}$ coincides with the (local) equivalence of the associated distributions $\mathcal{D}_\Sigma$ and $\mathcal{D}_{\tilde{\Sigma}}$.

Therefore the statement that a control system $\Sigma$ is locally feedback equivalent to the $m$-chained form (equivalently, to the canonical contact system on $J^n(\mathbb{R}, \mathbb{R}^m)$) will always mean that the associated distribution $\mathcal{D}_\Sigma$ is locally equivalent to the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$.

In this section we will formulate our first main result. See [71] and [72] for another approach to the problem of equivalence of the $n$-bar system (called there the articulated arm system) to the Cartan distribution (and, more generally, to the multi-flag system).

**Theorem 2.3.3** *The $n$-bar system $\Gamma$ moving in $\mathbb{R}^{m+1}$, for $m \geq 1$, defined by* (2.2.7), *is locally feedback equivalent to the $m$-chained form at any point $x \in X = \mathbb{R}^{(n+1)(m+1)}$ satisfying $\Psi(x) = 0$ (that is, at $x$ corresponding to a point $q \in Q$) such that*

(R1) $\displaystyle\sum_{j=1}^{m+1}(x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0$, for $1 \leq i \leq n-1$ , if $m \geq 2$;

(R2) $\displaystyle\sum_{j=1}^{m+1}(x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0$, for $2 \leq i \leq n-1$ , if $m = 1$.

*Moreover, at any point $q \in Q$ (equivalently, at any point $x \in X = \mathbb{R}^{(n+1)(m+1)}$ satisfying $\Psi(x) = 0$), the $n$-bar system satisfies the condition* (i) *and* (ii) *of Theorem 2.3.1.*

**Remark 1**. Let $\mathcal{D}_\Gamma$ be the distribution associated to the $n$-bar system $\Gamma$. Define the regular locus $Reg(\Gamma)$ of $\Gamma$ as $Reg(\Gamma) = Reg(\mathcal{D}_\Gamma)$. Then Theorem 2.3.3 implies that the regular locus of $\Gamma$ is given by

$$Reg(\Gamma)_{m\geq 2} = \{x \in X : \Psi(x) = 0, \sum_{j=1}^{m+1}(x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0, \quad 1 \leq i \leq n-1\}$$

$$Reg(\Gamma)_{m=1} = \{x \in X : \Psi(x) = 0, \sum_{j=1}^{m+1}(x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0, \quad 2 \leq i \leq n-1\}.$$

The condition $\Psi(x) = 0$ implies that $x$ is a point of $Q$ while the conditions (R1) and (R2) identify the regular points of $\Gamma$ in $Q$. It is obvious that $Reg(\Gamma)$ is open and dense in the configuration space $Q$ for both $m \geq 2$ and $m = 1$.

**Remark 2**. The regularity condition $\sum_{j=1}^{m+1}(x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0$ has a clear interpretation for the $n$-bar system. Let $\theta_i$, for $1 \leq i \leq n-1$, denote the angles of the $(i+1)$-th bar with respect to the $i$-th bar, i.e., the angles between the vectors $\overrightarrow{P_{i-1}P_i}$ and $\overrightarrow{P_iP_{i+1}}$ on the plane spanned by them. Then clearly the regularity conditions mean that $\theta_i$ are different from $\pm\frac{\pi}{2}$, in other words, the $i$-th bar is not perpendicular to the $(i+1)$-th one. Using the angles $\theta_i$, the regular locus can be rewritten as

$$Reg(\Gamma)_{m\geq 2} = \{q \in Q : \theta_i \neq \pm\frac{\pi}{2}, \quad 1 \leq i \leq n-1\}$$

$$Reg(\Gamma)_{m=1} = \{q \in Q : \theta_i \neq \pm\frac{\pi}{2}, \quad 2 \leq i \leq n-1\}.$$

It is interesting to observe the difference between the planar ($m = 1$) and all other cases ($m \geq 2$). Namely, the angle $\pm\frac{\pi}{2}$ between the last two bars (the two most far from the controlled one) is a singularity for $m \geq 2$ but is not for the planar case. The latter implies, in particular, that the 2-bar system in $\mathbb{R}^2$ is transformable into the chained form even if the bars are perpendicular. This is not true any longer if we consider the 2-bar system in the space $\mathbb{R}^{m+1}$, $m \geq 2$ (in $\mathbb{R}^3$, for instance). Of course, the 2-bar system in $\mathbb{R}^2$ is just the 1-trailer system (a unicycle-like mobile robot towing one trailer or, equivalently, a nonholonomic car) and it is well known that the system can be brought into the chained form even if the axles are perpendicular. In other words, the rank 2 distributions on 4-dimensional manifolds with the growth vector $(2, 3, 4)$ have no singularities, a result that goes back to [11].
**Remark 3**. The property of controllability of the $n$-bar system can also be obtained from Theorem 2.3.3. Indeed, we have the following corollary.

**Corollary 2.3.4** *The $n$-bar system $\Gamma$ is globally controllable on $Q = \mathbb{R}^{m+1} \times (S^m)^n$.*

## 2.4 Proof of Theorem 2.3.3

It is well known that (see, e.g., [27], [46], [52], [73] and Chapter 1 of this thesis) that the $n$-trailer system (that is, the $n$-bar system on $\mathbb{R}^{m+1}$ for $m = 1$) is locally feedback equivalent to the chained form (that is, its associated distribution is equivalent to the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R})$) if and only if the angles between the consecutive trailers are not $\pm\pi/2$, except for the angle between the last two ones (equivalently, except for the angle between $\overrightarrow{P_0P_1}$ and $\overrightarrow{P_1P_2}$). Therefore, we will prove only the case $m \geq 2$ here. The proof of Theorem 2.3.3, for $m \geq 2$, will use Theorem 2.3.1. First, the conditions (i) and (ii) of Theorem 2.3.1 will be verified for the $n$-bar system $\Gamma$. Secondly, we will prove that the regular locus $Reg(\Gamma)_{m\geq 2}$ is just the set of the points which satisfy the condition (R1) of Theorem 2.3.3. For simplicity, we denote the associated distribution $\mathcal{D}_\Gamma$ by $\mathcal{D}$ and clearly we have $\mathcal{D}^\perp = \mathcal{I}$ where $\mathcal{I}$ is defined by (2.2.6).

### 2.4.1 Notations

For convenience, we express $\mathcal{J}$, the codistribution given by (2.2.5), in the following form

$$\mathcal{J} = \text{span}\,\{\Omega_i^j,\ 1 \leq i \leq n, 1 \leq j \leq m+1, j \neq \sigma(i)\} = \text{span}\,\{\Omega_1, \Omega_2, \ldots, \Omega_n\},$$

where $\Omega_i = (\Omega_i^1, \ldots, \Omega_i^{\sigma(i)-1}, \Omega_i^{\sigma(i)+1}, \ldots, \Omega_i^{m+1})$, for $1 \leq i \leq n$. Similarly, the codistribution $\mathcal{I}$, given by (1.6), is expressed as

$$\mathcal{I} = \text{span}\,\{\omega_i^j,\ 1 \leq i \leq n, 1 \leq j \leq m+1, j \neq \sigma(i)\} = \text{span}\,\{\omega_1, \omega_2, \ldots, \omega_n\},$$

where $\omega_i = (\omega_i^1, \ldots, \omega_i^{\sigma(i)-1}, \omega_i^{\sigma(i)+1}, \ldots, \omega_i^{m+1})$, for $1 \le i \le n$.
The following symbols will be used in this chapter:

$$\begin{array}{lll}
(1)\ \ \mathrm{d}X_i & = & (\mathrm{d}x_i^1, \ldots, \mathrm{d}x_i^{m+1}) \\
(2)\ \ \wedge\mathrm{d}X_i & = & \mathrm{d}x_i^1 \wedge \cdots \wedge \mathrm{d}x_i^{m+1} \\
(3)\ \ \wedge\mathrm{d}\check{X}_i^{\sigma(i)} & = & \mathrm{d}x_i^1 \wedge \cdots \wedge \mathrm{d}x_i^{\sigma(i)-1} \wedge \mathrm{d}x_i^{\sigma(i)+1} \cdots \wedge \mathrm{d}x_i^{m+1} \\
(4)\ \ \wedge\Omega_i & = & \Omega_i^1 \wedge \cdots \wedge \Omega_i^{\sigma(i)-1} \wedge \Omega_i^{\sigma(i)+1} \wedge \cdots \wedge \Omega_i^{m+1} \\
(5)\ \ \wedge\omega_i & = & \omega_i^1 \wedge \cdots \wedge \omega_i^{\sigma(i)-1} \wedge \omega_i^{\sigma(i)+1} \wedge \cdots \wedge \omega_i^{m+1}.
\end{array}$$

### 2.4.2 Condition (i): $\mathcal{D}^{(n)} = TQ$

Let $\Lambda^k(Q)$ denote the space of smooth differential k-forms on $Q$ and put $\Lambda(Q) = \oplus\Lambda^k(Q)$. An ideal of $\Lambda(Q)$ is a subspace $V$ of $\Lambda^k(Q)$ such that if $\alpha \in V$, then $\alpha \wedge \beta \in V$ for any $\beta \in \Lambda(Q)$.

**Lemma 2.4.1** *Consider the codistribution $\mathcal{I}$ defined by (2.2.6) and let $(\Phi^*\mathrm{d}X_i)$ be the ideal in $\Lambda(Q)$ generated by $\Phi^*\mathrm{d}x_i^1, \ldots, \Phi^*\mathrm{d}x_i^{m+1}$, then we have:*

(a) $\mathrm{d}\omega_i^j \wedge \omega_i \wedge \omega_{i+1} = 0$, *for* $1 \le i \le n-1$, $1 \le j \le m+1$, *and* $j \neq \sigma(i)$*;*

(b) $\wedge\omega_i = (-1)^m (x_i^{\sigma(i)} - x_i^{\sigma(i-1)})^{m-2} (\wedge\Phi^*\mathrm{d}\check{X}_{i-1}^{\sigma(i)})$ mod $(\Phi^*\mathrm{d}X_i)$, *for* $1 \le i \le n-1$*;*

(c) $\mathrm{d}\omega_i^j \wedge \omega_i \wedge \cdots \wedge \omega_1 \neq 0$, *for* $1 \le i \le n$, $1 \le j \le m+1$ *and* $j \neq \sigma(i)$ *.*

The proof of Lemma 2.4.1 is quite tedious and is given in Section 2.6. Below we will show how Lemma 2.4.1 implies the condition $\mathcal{D}^{(n)} = TQ$.

If $i = n$, the condition (c) shows that $\mathrm{d}\omega_n^j \wedge \omega_1 \wedge \cdots \wedge \omega_n \neq 0$, for $1 \le j \le m+1$ and $j \neq \sigma(n)$, which imply that all differential forms $\omega_1, \omega_2, \ldots, \omega_n$ are independent. Recall that each $\omega_i$ consists of $m$ differential 1-forms. Therefore $\mathcal{I}$ is of constant rank $nm$ and thus

$$\operatorname{rank}\mathcal{D} = \dim Q - \operatorname{rank}\mathcal{I} = (n+1)m + 1 - nm = m + 1.$$

Observe that

$$\begin{aligned}
\mathrm{d}\omega_n^j &= \Phi^*\mathrm{d}\Omega_n^j \\
&= \Phi^*\mathrm{d}x_n^j \wedge \Phi^*\mathrm{d}x_{n-1}^{\sigma(n)} - 2\Phi^*\mathrm{d}x_{n-1}^j \wedge \Phi^*\mathrm{d}x_{n-1}^{\sigma(n)} - \Phi^*\mathrm{d}x_n^{\sigma(n)} \wedge \Phi^*\mathrm{d}x_{n-1}^j.
\end{aligned}$$

It is obvious that the differential forms $\mathrm{d}\omega_n^1, \ldots, \mathrm{d}\omega_n^{\sigma(n)-1}, \mathrm{d}\omega_n^{\sigma(n)+1}, \ldots, \mathrm{d}\omega_n^{m+1}$ are linearly independent on $Q$. Consequently, all the differential formes

$$\mathrm{d}\omega_n^j \wedge \omega_n \wedge \cdots \wedge \omega_1,$$

for $1 \leq j \leq m+1$ and $j \neq \sigma(n)$, are also linearly independent on $Q$. From condition (a), we get easily that

$$\mathrm{d}\omega_i^j \wedge \omega_1 \wedge \cdots \wedge \omega_n = 0, \quad 1 \leq i \leq n-1,\, 1 \leq j \leq m+1, \quad \text{and} \quad j \neq \sigma(i). \quad (2.4.1)$$

Recall that the derived codistribution $\mathcal{I}^{(1)}$ is given by

$$\mathcal{I}^{(1)} = \{\omega \in \mathcal{I} : \mathrm{d}\omega \equiv 0 \bmod \mathcal{I}\}.$$

For any $\omega \in \mathcal{I}$, we have

$$\omega = \sum_{i=1}^{n} \sum_{j=1, j\neq\sigma(i)}^{m+1} \alpha_j^i \cdot \omega_i^j,$$

where $\alpha_j^i \in C^\infty(Q)$. Applying (2.4.1), we deduce

$$\begin{aligned} \mathrm{d}\omega \wedge \omega_1 \wedge \cdots \wedge \omega_n &= \mathrm{d}\left(\sum_{i=1}^{n} \sum_{j=1, j\neq\sigma(i)}^{m+1} \alpha_j^i \cdot \omega_i^j\right) \wedge \omega_1 \wedge \cdots \wedge \omega_n \\ &= \sum_{j=1,\, j\neq\sigma(n)}^{m+1} \alpha_j^n \cdot \mathrm{d}\omega_n^j \wedge \omega_1 \wedge \cdots \wedge \omega_n. \end{aligned}$$

The differential forms $\mathrm{d}\omega_n^j \wedge \omega_n \wedge \cdots \wedge \omega_1$, for $1 \leq j \leq m+1$ and $j \neq \sigma(n)$, are linearly independent and thus $\mathrm{d}\omega \wedge \omega_1 \wedge \cdots \wedge \omega_n = 0$ implies that all coefficients $\alpha_j^n$ must vanish, i.e., $\alpha_j^n = 0$, for $1 \leq j \leq m+1$, and $j \neq \sigma(n)$. This shows that, indeed, any $\omega \in \mathcal{I}^{(1)}$ can be always expressed by a linear combination of the differential forms $\omega_1, \omega_2, \ldots, \omega_{n-1}$. In other words, we have

$$\mathcal{I}^{(1)} = \operatorname{span}\{\omega_1, \omega_2, \ldots, \omega_{n-1}\}.$$

In the same way, it can be shown that

$$\begin{aligned} \mathcal{I}^{(2)} &= \operatorname{span}\{\omega_1, \omega_2, \ldots, \omega_{n-2}\} \\ &\ \vdots \\ \mathcal{I}^{(n-1)} &= \operatorname{span}\{\omega_1\} \\ \mathcal{I}^{(n)} &= 0. \end{aligned}$$

The expressions of the derived flag $\mathcal{I}^{(0)} \supset \cdots \supset \mathcal{I}^{(n)}$ show immediately that $\operatorname{rank} \mathcal{I}^{(i)} = (n-i)m$, and hence $\operatorname{rank} \mathcal{D}^{(i)} = (i+1)m+1$. In particular, $\operatorname{rank} \mathcal{D}^{(n)} = (n+1)m+1$ which implies that $\mathcal{D}^{(n)} = TQ$.

### 2.4.3 Condition (ii) : $\mathcal{D}^{(n-1)}$ contains a corank one involutive subdistribution $\mathcal{L}_{n-1} \subset \mathcal{D}^{(n-1)}$

Define a distribution $\mathcal{K}$ on $X = \mathbb{R}^{(n+1)(m+1)}$ by

$$\mathcal{K} = \operatorname{span}\{\frac{\partial}{\partial x_1^1}, \ldots, \frac{\partial}{\partial x_1^{m+1}}, \ldots, \frac{\partial}{\partial x_n^1}, \ldots, \frac{\partial}{\partial x_n^{m+1}}\} \subset TX.$$

Let $\mathcal{L}$ be the distribution on $Q$ defined by $\mathcal{L}^{\perp} = \Phi^*(\mathcal{K}^{\perp})$ , i.e.,

$$\mathcal{K} \longmapsto (\mathcal{K})^{\perp} \overset{\Phi^*}{\longmapsto} \mathcal{L}^{\perp} \longmapsto \mathcal{L}$$

**Lemma 2.4.2** *The distribution $\mathcal{L}$ is involutive and contained in $\mathcal{D}^{(n-1)}$.*

**Proof:** Since $\mathcal{K}$ is involutive and $\Phi^*(\mathcal{K}^{\perp}) = \mathcal{L}^{\perp}$, the involutivity of $\mathcal{L}$ is evident. Let $\mathcal{J}_{n-1}$ be the codistribution defined on $X = \mathbb{R}^{(n+1)(m+1)}$ by

$$\mathcal{J}_{n-1} = \operatorname{span}\{\Omega_1^j, \quad 1 \leq j \leq m+1,\ j \neq \sigma(1)\},$$

where

$$\Omega_1^j = (x_1^j - x_0^j)\mathrm{d}x_0^{\sigma(1)} - (x_1^{\sigma(1)} - x_0^{\sigma(1)})\mathrm{d}x_0^j.$$

Since $\mathcal{K}^{\perp} = \operatorname{span}\{\mathrm{d}x_0^1, \ldots, \mathrm{d}x_0^{m+1}\}$, it is obvious that $\mathcal{J}_{n-1} \subset \mathcal{K}^{\perp}$. Recall that we showed in Section 2.4.2

$$\mathcal{I}^{(n-1)} = \operatorname{span}\{\omega_1^j, \quad 1 \leq j \leq m+1,\ j \neq \sigma(1)\} = \Phi^*\mathcal{J}_{n-1}.$$

Then clearly we have $\Phi^*(\mathcal{J}_{n-1}) \subset \Phi^*(\mathcal{K}^{\perp})$ and therefore $\mathcal{I}^{(n-1)} \subset \mathcal{L}^{\perp}$ which implies that $\mathcal{L} \subset \mathcal{D}^{(n-1)}$. $\square$

**Lemma 2.4.3** *$\mathcal{L}$ is of constant corank one in $\mathcal{D}^{(n-1)}$.*

**Proof:** Observe that

$$\mathcal{L}^{\perp} = \Phi^*(\mathcal{K}^{\perp}) = \operatorname{span}\{\Phi^*\mathrm{d}x_0^1, \ldots, \Phi^*\mathrm{d}x_0^{m+1}\}.$$

Since $\Phi^*\mathrm{d}x_0^1 \wedge \cdots \wedge \Phi^*\mathrm{d}x_0^{m+1} \neq 0$, we have $\operatorname{rank}\mathcal{L}^{\perp} = m+1$ and thus

$$\operatorname{rank}\mathcal{L} = \dim Q - \operatorname{rank}\mathcal{L}^{\perp} = (n+1)m + 1 - (m+1) = nm.$$

In Subsection 2.4.2, we proved $\operatorname{rank}\mathcal{D}^{(n-1)} = nm+1$ which yields

$$\operatorname{corank}(\mathcal{L} \subset \mathcal{D}^{(n-1)}) = 1.$$

Define $\mathcal{L}_{n-1} = \mathcal{L}$, then the condition (ii) of Theorem 2.3.1 holds. $\square$

### 2.4.4 The regular locus of the $n$-bar system in $\mathbb{R}^{m+1}$

**Proposition 2.4.4** *The regular locus of the $n$-bar system $\Gamma$ in $\mathbb{R}^{m+1}$, for $m \geq 2$, is given by*

$$Reg(\Gamma)_{m\geq 2} = \{x \in X : \Psi(x) = 0, \quad \sum_{j=1}^{m+1}(x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0, \quad 1 \leq i \leq n-1\}.$$

For simplicity of exposition, we assume throughout this section that $\sigma(i) = m+1$, for $1 \le i \le n$, i.e.,

$$x_i^{m+1} - x_{i-1}^{m+1} \neq 0, \quad 1 \le i \le n.$$

We can always achieve this by a suitable choice of coordinates. Under this assumption, the codistribution $\mathcal{J} = \mathcal{E}^{\perp}$ on $\mathbb{R}^{(n+1)(m+1)}$ is $\mathcal{J} = \operatorname{span}\{\Omega_i^j,\ 1 \le i \le n,\ 1 \le j \le m\}$, where

$$\Omega_i^j = (x_i^{m+1} - x_{i-1}^{m+1})\,\mathrm{d}x_{i-1}^j - (x_i^j - x_{i-1}^j)\,\mathrm{d}x_{i-1}^{m+1}.$$

In Section 2.4.3, the subdistribution $\mathcal{L}_{n-1}$ satisfying the conditions (i) and (ii) of Theorem 2.3.1 was constructed. Since $m \ge 2$, $\mathcal{L}_{n-1}$ is unique (see [50]). Therefore the singular locus of the system $\Gamma$ is defined by

$$Sing(\Gamma) = \{q \in Q : \mathcal{D}(q) \subset \mathcal{L}_{n-1}(q)\}.$$

From the relation $(\mathcal{L}_{n-1})^{\perp} = \Phi^*\mathcal{K}^{\perp}$ (defining $\mathcal{L}_{n-1} = \mathcal{L}$, see Section 2.4.3) and $\mathcal{I} = \mathcal{D}^{\perp} = \Phi^*\mathcal{E}^{\perp} = \Phi^*\mathcal{J}$ (see Section 2.2.2), we conclude that the singular locus of the $n$-bar system $\Gamma$ can be expressed as

$$\begin{aligned} Sing(\Gamma) &= \{q \in Q : \mathcal{D}(q) \subset \mathcal{L}_{n-1}(q)\} \\ &= \{q \in Q : (\mathcal{L}_{n-1})^{\perp}(q) \subset \mathcal{I}(q)\} \\ &= \{q \in Q : \Phi^*\mathcal{K}^{\perp}(q) \subset \Phi^*\mathcal{J}(q)\}. \end{aligned}$$

Notice that $\Phi^*\mathcal{K}^{\perp} = \operatorname{span}\{\Phi^*\mathrm{d}x_0^1, \dots, \Phi^*\mathrm{d}x_0^{m+1}\}$, and

$$\Phi^*\mathcal{J} = \operatorname{span}\{\Phi^*\Omega_i^j, \quad 1 \le i \le n, \quad 1 \le j \le m\},$$

where $\Phi^*\Omega_i^j = (x_i^{m+1} - x_{i-1}^{m+1})\Phi^*\mathrm{d}x_{i-1}^j - (x_i^j - x_{i-1}^j)\Phi^*\mathrm{d}x_{i-1}^{m+1}$. Since $\mathcal{L}_{n-1}$ and $\mathcal{I}$ are both of constant rank, we have immediately that the relation $\Phi^*\mathcal{K}^{\perp}(q) \subset \Phi^*\mathcal{J}(q)$ holds if and only if

$$\Phi^*\mathrm{d}x_0^1 \wedge \cdots \wedge \Phi^*\mathrm{d}x_0^{m+1} \wedge \Phi^*\Omega_1 \wedge \cdots \wedge \Phi^*\Omega_n = 0. \tag{2.4.2}$$

By a straightforward calculation, it is easy to see that (2.4.2) is equivalent to

$$\Phi^*\mathrm{d}x_0^1 \wedge \cdots \wedge \Phi^*\mathrm{d}x_0^{m+1} \wedge \Phi^*\Omega_2 \wedge \cdots \wedge \Phi^*\Omega_n = 0.$$

Denoting $\mathfrak{S} = \Phi^*\mathrm{d}x_0^1 \wedge \cdots \wedge \Phi^*\mathrm{d}x_0^{m+1} \wedge \Phi^*\Omega_2 \wedge \cdots \wedge \Phi^*\Omega_n$, we thus have

$$Sing(\Gamma) = \{q \in Q : \mathfrak{S}(q) = 0\}.$$

**Lemma 2.4.5** *Define the differential form* $\Theta_k$ *by*

$$\Theta_k = \Phi^*\mathrm{d}x_k^1 \wedge \cdots \wedge \Phi^*\mathrm{d}x_k^{m+1} \wedge \Phi^*\Omega_{k+2},$$

*then we have*

$$\Theta_k = (-1)^m (x_{k+2}^{m+1} - x_{k+1}^{m+1})^{m-1} \cdot \lambda_{k+1} \cdot (\wedge \Phi^*\mathrm{d}\breve{X}_k^{m+1}) \wedge \Phi^*\mathrm{d}x_{k+1}^1 \wedge \cdots \wedge \Phi^*\mathrm{d}x_{k+1}^{m+1},$$

*for* $0 \le k \le n-2$, *where* $\wedge\Phi^*\mathrm{d}\breve{X}_k^{m+1} = \Phi^*\mathrm{d}x_k^1 \wedge \cdots \wedge \Phi^*\mathrm{d}x_k^m$ *and*

$$\lambda_{k+1} = (x_{k+2}^{m+1} - x_{k+1}^{m+1}) + \frac{x_{k+1}^1 - x_k^1}{x_{k+1}^{m+1} - x_k^{m+1}}(x_{k+2}^1 - x_{k+1}^1) + \cdots + \frac{x_{k+1}^m - x_k^m}{x_{k+1}^{m+1} - x_k^{m+1}}(x_{k+2}^m - x_{k+1}^m).$$

**Proof:** Differentiating the relation

$$\Psi_{k+1} = (x_{k+1}^1 - x_k^1)^2 + (x_{k+1}^2 - x_k^2)^2 + \cdots + (x_{k+1}^{m+1} - x_k^{m+1})^2 - 1 = 0$$

we obtain

$$\mathrm{d}x_k^{m+1} = \mathrm{d}x_{k+1}^{m+1} + \frac{x_{k+1}^1 - x_k^1}{x_{k+1}^{m+1} - x_k^{m+1}}(\mathrm{d}x_{k+1}^1 - \mathrm{d}x_k^1) + \cdots + \frac{x_{k+1}^m - x_k^m}{x_{k+1}^{m+1} - x_k^{m+1}}(\mathrm{d}x_{k+1}^m - \mathrm{d}x_k^m).$$

Therefore, on the configuration space $Q$, we have

$$\Phi^*\mathrm{d}x_k^{m+1} = \Phi^*\mathrm{d}x_{k+1}^{m+1} + \frac{x_{k+1}^1 - x_k^1}{x_{k+1}^{m+1} - x_k^{m+1}}(\Phi^*\mathrm{d}x_{k+1}^1 - \Phi^*\mathrm{d}x_k^1) + \cdots + \frac{x_{k+1}^m - x_k^m}{x_{k+1}^{m+1} - x_k^{m+1}}(\Phi^*\mathrm{d}x_{k+1}^m - \Phi^*\mathrm{d}x_k^m).$$

Recall that

$$\Omega_{k+2}^j = (x_{k+2}^{m+1} - x_{k+1}^{m+1})\mathrm{d}x_{k+1}^j - (x_{k+2}^j - x_{k+1}^j)\mathrm{d}x_{k+1}^{m+1},$$

for $1 \le j \le m$, and substituting the latter into $\Theta_k$, we get

$$\begin{aligned}
\Theta_k &= \Phi^*\mathrm{d}x_k^1 \wedge \cdots \wedge \Phi^*\mathrm{d}x_k^{m+1} \wedge \Phi^*\Omega_{k+2} \\
&= \Phi^*\mathrm{d}x_k^1 \wedge \cdots \wedge \Phi^*\mathrm{d}x_k^m \wedge (\Phi^*\mathrm{d}x_{k+1}^{m+1} + \frac{x_{k+1}^1 - x_k^1}{x_{k+1}^{m+1} - x_k^{m+1}}\,\Phi^*\mathrm{d}x_{k+1}^1 + \cdots \\
&\quad + \frac{x_{k+1}^m - x_k^m}{x_{k+1}^{m+1} - x_k^{m+1}}\,\Phi^*\mathrm{d}x_{k+1}^m) \wedge \Phi^*\Omega_{k+2}^1 \wedge \cdots \wedge \Phi^*\Omega_{k+2}^m \\
&= (x_{k+2}^{m+1} - x_{k+1}^{m+1})^m \cdot (\wedge\mathrm{d}\breve{X}_k^{m+1}) \wedge \Phi^*\mathrm{d}x_{k+1}^{m+1} \wedge \Phi^*\mathrm{d}x_{k+1}^1 \wedge \cdots \wedge \Phi^*\mathrm{d}x_{k+1}^m \\
&\quad -(x_{k+2}^{m+1} - x_{k+1}^{m+1})^{m-1}(x_{k+2}^1 - x_{k+1}^1)\frac{x_{k+1}^1 - x_k^1}{x_{k+1}^{m+1} - x_k^{m+1}} \cdot (\wedge\mathrm{d}\breve{X}_k^{m+1}) \\
&\quad \wedge\Phi^*\mathrm{d}x_{k+1}^1 \wedge \Phi^*\mathrm{d}x_{k+1}^{m+1} \wedge \Phi^*\mathrm{d}x_{k+1}^2 \wedge \cdots \wedge \Phi^*\mathrm{d}x_{k+1}^m \\
&\quad -(x_{k+2}^{m+1} - x_{k+1}^{m+1})^{m-1}(x_{k+2}^2 - x_{k+1}^2)\frac{x_{k+1}^2 - x_k^2}{x_{k+1}^{m+1} - x_k^{m+1}} \cdot (\wedge\mathrm{d}\breve{X}_k^{m+1}) \\
&\quad \wedge\Phi^*\mathrm{d}x_{k+1}^2 \wedge \Phi^*\mathrm{d}x_{k+1}^1 \wedge \Phi^*\mathrm{d}x_{k+1}^{m+1} \wedge \Phi^*\mathrm{d}x_{k+1}^3 \wedge \cdots \Phi^*\mathrm{d}x_{k+1}^{m-1} \wedge \Phi^*\mathrm{d}x_{k+1}^m \\
&\quad - \cdots - (x_{k+2}^{m+1} - x_{k+1}^{m+1})^{m-1}(x_{k+2}^m - x_{k+1}^m)\frac{x_{k+1}^m - x_k^m}{x_{k+1}^{m+1} - x_k^{m+1}} \cdot (\wedge\mathrm{d}\breve{X}_k^{m+1}) \\
&\quad \wedge\Phi^*\mathrm{d}x_{k+1}^m \wedge \Phi^*\mathrm{d}x_{k+1}^1 \wedge \cdots \wedge \Phi^*\mathrm{d}x_{k+1}^{m-1}\Phi^*\mathrm{d}x_{k+1}^{m+1} \\
&= (-1)^m(x_{k+2}^{m+1} - x_{k+1}^{m+1})^{m-1} \cdot \lambda_{k+1} \cdot (\wedge\Phi^*\mathrm{d}\breve{X}_k^{m+1}) \wedge \Phi^*\mathrm{d}x_{k+1}^1 \wedge \cdots \wedge \Phi^*\mathrm{d}x_{k+1}^{m+1}
\end{aligned}$$

where

$$\lambda_{k+1} = (x_{k+2}^{m+1} - x_{k+1}^{m+1}) + \frac{x_{k+1}^1 - x_k^1}{x_{k+1}^{m+1} - x_k^{m+1}}(x_{k+2}^1 - x_{k+1}^1) + \cdots + \frac{x_{k+1}^m - x_k^m}{x_{k+1}^{m+1} - x_k^{m+1}}(x_{k+2}^m - x_{k+1}^m).$$

$\square$

**Proof of Proposition 2.4.4**: Applying lemma 2.4.5, we obtain

$$
\begin{aligned}
\mathfrak{S} &= \Phi^* dx_0^1 \wedge \cdots \wedge \Phi^* dx_0^{m+1} \wedge \Phi^*\Omega_2 \wedge \cdots \wedge \Phi^*\Omega_n \\
&= \Theta_0 \wedge \Phi^*\Omega_3 \wedge \cdots \wedge \Phi^*\Omega_n \\
&= (-1)^m (x_2^{m+1} - x_1^{m+1})^{m-1} \cdot \lambda_1 \cdot (\wedge \Phi^* d\breve{X}_0^{m+1}) \wedge \Phi^* dx_1^1 \wedge \cdots \wedge \Phi^* dx_1^m \wedge \Phi^* dx_1^{m+1} \\
&\quad \wedge \Phi^*\Omega_3 \wedge \cdots \wedge \Phi^*\Omega_n \\
&= (-1)^m (x_2^{m+1} - x_1^{m+1})^{m-1} \cdot \lambda_1 \cdot (\wedge \Phi^* d\breve{X}_0^{m+1}) \wedge \Theta_1 \wedge \Phi^*\Omega_4 \wedge \cdots \wedge \Phi^*\Omega_n \\
&= (-1)^{2m} \left( \prod_{i=1}^{2} (x_{i+1}^{m+1} - x_i^{m+1})^{m-1} \right) \cdot (\lambda_1 \cdot \lambda_2) \cdot (\wedge \Phi^* d\breve{X}_0^{m+1}) \wedge \Phi^* d\breve{X}_1^{m+1} \\
&\quad \wedge \Phi^* dx_2^1 \wedge \cdots \wedge \Phi^* dx_2^m \wedge \Phi^* dx_2^{m+1} \wedge \Phi^*\Omega_4 \wedge \cdots \wedge \Phi^*\Omega_n \\
&= (-1)^{2m} \left( \prod_{i=1}^{2} (x_{i+1}^{m+1} - x_i^{m+1})^{m-1} \right) \cdot (\lambda_1 \cdot \lambda_2) \cdot (\wedge \Phi^* d\breve{X}_0^{m+1}) \wedge \Phi^* d\breve{X}_1^{m+1} \\
&\quad \wedge \Theta_2 \wedge \Phi^*\Omega_5 \wedge \cdots \wedge \Phi^*\Omega_n \\
&= \cdots \\
&= (-1)^{(n-1)m} \left( \prod_{i=1}^{n-1} (x_{i+1}^{m+1} - x_i^{m+1})^{m-1} \right) \cdot \left( \prod_{k=1}^{n-1} \lambda_k \right) \cdot (\wedge \Phi^* d\breve{X}_0^{m+1}) \\
&\quad \wedge \cdots \wedge \Phi^* d\breve{X}_{n-2}^{m+1} \wedge \Phi^* dx_{n-1}^1 \wedge \cdots \Phi^* dx_{n-1}^{m+1}.
\end{aligned}
$$

Notice that the differential forms $\Phi^* d\breve{X}_0^{m+1}, \ldots, \Phi^* d\breve{X}_{n-2}^{m+1}, \Phi^* dx_{n-1}^1, \ldots, \Phi^* dx_{n-1}^{m+1}$ are always independent on $Q$ and $\prod_{i=1}^{n-1}(x_{i+1}^{m+1} - x_i^{m+1})^{m-1} \neq 0$, so it can be concluded that $\mathfrak{S}(q) = 0$ if and only if there exists $1 \leq k \leq n-1$ such that $\lambda_k = 0$. The latter holds if and only if

$$\sum_{j=1}^{m+1} (x_k^j - x_{k-1}^j)(x_{k+1}^j - x_k^j) = 0$$

which, in fact, is equivalent to the orthogonality $\overrightarrow{P_{k-1}P_k} \perp \overrightarrow{P_k P_{k+1}}$. Therefore $Sing(\Gamma) = \{q \in Q : \mathfrak{S}(q) = 0\}$ are equivalent to

$$Sing(\Gamma) = \{x \in X : \Psi(x) = 0, \exists\, 1 \leq i \leq n-1 \quad s.t. \quad \sum_{j=1}^{m+1} (x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) = 0\},$$

and

$$Reg(\Gamma) = \{x \in X : \Psi(x) = 0, \sum_{j=1}^{m+1} (x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0, \quad 1 \leq i \leq n-1\}.$$

## 2.5 Flatness of the $n$-bar system in $\mathbb{R}^{m+1}$

Consider a smooth nonlinear control system

$$\Xi : \quad \dot{x} = f(x,u)$$

where $x \in X$, an $n$-dimensional manifold and $u \in U$, an $m$-dimensional manifold. Given any integer $l$, we associate to $\Xi$ its $l$-prolongation $\Xi^l$ given by

$$\Xi^l : \quad \begin{array}{rcl} \dot{x} & = & f(x,u^0) \\ \dot{u}^0 & = & u^1 \\ & \vdots & \\ \dot{u}^l & = & u^{l+1} \end{array}$$

which can be considered as a control system on $X^l = X \times U \times \mathbb{R}^{ml}$, whose state variables are $(x, u^0, u^1, \ldots, u^l)$ and whose $m$ controls are the $m$ components of $u^{l+1}$. Denote $\bar{u}^l = (u^0, u^1, \ldots, u^l)$.

**Definition 2.5.1** The system $\Xi$ is called flat at a point $(x_0, \bar{u}_0^l) \in X^l = X \times U \times \mathbb{R}^{ml}$, for some $l \geq 0$, if there exist a neighborhood $\mathcal{O}^l$ of $(x_0, \bar{u}_0^l)$ and $m$ smooth functions

$$h_i = h_i(x, u^0, u^1, \ldots, u^l), \quad 1 \leq i \leq m,$$

called *flat outputs*, defined in $\mathcal{O}^l$, having the following property: there exist an integer $s$ and smooth functions $\gamma_i$, $1 \leq i \leq n$, and $\delta_i$, $1 \leq i \leq m$, such that we have

$$\begin{array}{rcl} x_i & = & \gamma_i(h, \dot{h}, \ldots, h^{(s)}), \quad 1 \leq i \leq n \\ u_i & = & \delta_i(h, \dot{h}, \ldots, h^{(s)}), \quad 1 \leq i \leq m, \end{array}$$

where $h = (h_1, \ldots, h_m)^\top$, along any trajectory $x(t)$ given by a control $u(t)$ that satisfy $(x(t), u(t), \dot{u}(t), \ldots, u^{(l)}(t)) \in \mathcal{O}^l$.

The compositions $\gamma_i(h, \dot{h}, \ldots, h^{(s)})$ and $\delta_i(h, \dot{h}, \ldots, h^{(s)})$ are, a priori, defined in an open set $\mathcal{O}^{s+l} \subset X^{s+l} = X \times U \times \mathbb{R}^{m(s+l)}$. The above definition requires that $\pi(\mathcal{O}^{s+l}) \supset \mathcal{O}^l$, where $\pi(x, \bar{u}^{s+l}) = (x, \bar{u}^l)$, and that for all such $(x, \bar{u}^{s+l})$, the compositions yield, respectively, $x_i$ and $u_i$.

If $h_i = h_i(x, u^0, u^1, \ldots, u^r)$, $r \leq l$, we will say that the system is $(x, u, \ldots, u^r)$-flat and, in particular, $x$-flat if $h_i = h_i(x)$. In the case $h_i = h_i(x, u^0, u^1, \ldots, u^r)$, we will assume that they are defined on $\mathcal{O}^r \subset X^r = X \times U \times \mathbb{R}^{mr}$, where $\pi^{-1}(\mathcal{O}^r) \supset \mathcal{O}^l$ and $\pi$ stands for the projection $\pi(x, u^0, \ldots, u^r, \ldots, u^l) = (x, u^0, \ldots, u^r)$.

Let $h_1, \ldots, h_m$ be flat outputs of a system $\Xi$. It has been observed in [57] that there exist integers $k_1, \ldots, k_m$ such that

$$\mathrm{span}\,\{\mathrm{d}x_1, \ldots, \mathrm{d}x_n, \mathrm{d}u_1, \ldots, \mathrm{d}u_m\} \subset \mathrm{span}\,\{\mathrm{d}h_i^{(j)}, 1 \leq i \leq m,\ 0 \leq j \leq k_i\},$$

and if at the same time

$$\mathrm{span}\,\{\mathrm{d}x_1,\ldots,\mathrm{d}x_n,\mathrm{d}u_1,\ldots,\mathrm{d}u_m\}\subset \mathrm{span}\,\{\mathrm{d}h_i^{(j)},1\le i\le m,\ 0\le j\le \mu_i\},$$

then $k_i \le \mu_i$, for $1 \le i \le m$. The $m$-tuple $(k_1,\ldots,k_m)$ will be called the *differential $m$-weight* of $h=(h_1,\ldots,h_m)$ and $k=\sum_{i=1}^{m} k_i$ will be called the *differential weight* of $h$.

**Definition 2.5.2** Flat outputs of $\Xi$ at $(x_0,\bar{u}_0^l)$ are called minimal if their differential weight is the lowest among all flat outputs of $\Xi$ at $(x_0,\bar{u}_0^l)$.

Let $U_{\mathrm{sing}}(x)$ be the $m$-dimensional subspace of $\mathbb{R}^{m+1}$ such that for any feedback control $(u_0(x),\ldots,u_m(x))^\top = u(x)\in U_{\mathrm{sing}}(x)$ we have $\sum_{i=0}^m f_i(x)u_i(x)\in\mathcal{C}_1(x)$, where $\mathcal{C}_1\subset\mathcal{D}=\mathrm{span}\,\{f_0,\ldots,f_m\}$ is the characteristic subdistribution of $\mathcal{D}^{(1)}$. Any control $u(t)\in U_{sing}(x(t))$ will be called *singular* and the trajectories of the system governed by a singular control remain *tangent* to the characteristic distribution $\mathcal{C}_1$.

The following theorem, given in [57], characterizes the minimal flat outputs for systems that are feedback equivalent to the $m$-chained form (i.e., the canonical contact system on $J^n(\mathbb{R},\mathbb{R}^m)$), with $m\ge 2$.

**Theorem 2.5.3** *Consider the driftless control-linear system*

$$\Sigma:\quad \dot{x}=\sum_{i=0}^{m} f_i(x)u_i,$$

*defined on a manifold $X$ and let $\mathcal{D}=\mathrm{span}\,\{f_0,\ldots,f_m\}$ be the distribution associated to $\Sigma$. If $\Sigma$ is locally feedback equivalent, at $x_0\in X$, to the $m$-chained form, with $m\ge 2$, then the following conditions are equivalent:*

(i) $(\mathcal{L}_{n-1})^\perp = \mathrm{span}\,\{\mathrm{d}h_0,\ldots,\mathrm{d}h_m\}$ *around $x_0$, where $\mathcal{L}_{n-1}$ denotes the subdistribution that is involutive and of corank one in $\mathcal{D}^{(n-1)}$;*

(ii) $h_0,\ldots,h_m$ *are minimal flat outputs of $\Sigma$ at $(x_0,u^0)$, where $u^0\notin U_{\mathrm{sing}}(x_0)$.*

Recall that the $n$-bar system, as defined in Section 2.2.2, evolves on its configuration manifold $Q=\mathbb{R}^{m+1}\times(S^m)^n$, whose points $q\in Q$ correspond to those $x\in\mathbb{R}^{(n+1)(m+1)}$ that satisfy $\Psi(x)=(\Psi_1(x),\ldots,\Psi_n(x))=0$, where $\Psi_i$, for $1\le i\le n$, are given by (2.2.2). The following theorem describes the flatness property of the $n$-bar system $\Gamma$.

**Theorem 2.5.4** (Flatness of the $n$-bar system) *For the $n$-bar system $\Gamma$ moving in $\mathbb{R}^{m+1}$, where $m\ge 2$, we have*

(i) $\Gamma$ *is $x$-flat at any $(q_0,u^0)\in Q\times\mathbb{R}^{m+1}$ satisfying*

  (1) $\Psi(x)=0$ *and* $\sum_{j=1}^{m+1}(x_i^j-x_{i-1}^j)(x_{i+1}^j-x_i^j)\neq 0$, *where $q_0\in Q$ is identified with a point $x\in\mathbb{R}^{(n+1)(m+1)}$ satisfying $\Psi(x)=0$;*

(2) *$u^0$ is such that the instantaneous velocity $\dot{P}_0$ of the point $P_0$ is nonzero (and thus the instantaneous velocities $\dot{P}_i$, $0 \le i \le n-1$, are nonzero).*

(ii) *The coordinates of the source point of the last bar $P_0 = (x_0^1, x_0^2, \ldots, x_0^{m+1})$ are minimal $x$-flat outputs of $\Gamma$ at any $(q_0, u^0)$ as above.*

(iii) *If $h_0, \ldots, h_m$ are any minimal $x$-flat outputs at $(q_0, u^0)$, then locally around $q_0$ we have*

$$\operatorname{span}\{\mathrm{d}h_0, \ldots, \mathrm{d}h_m\} = \operatorname{span}\{\mathrm{d}x_0^1, \mathrm{d}x_0^2, \ldots, \mathrm{d}x_0^{m+1}\}.$$

The item (iii) is in contrast with the planar case $m = 1$, where minimal flat outputs are not unique and their totality is actually parameterized by an arbitrary function of three variables (See a detailed analysis in Chapter 1). Another difference between the planar case ($m = 1$) and higher-dimensional case ($m \geq 2$) is that in the latter all angles $\pm\frac{\pi}{2}$ between two consecutive bars form singularities while in the former the angle $\pm\frac{\pi}{2}$ between the last and one before the last bars (trailers) $\overrightarrow{P_0P_1}$ and $\overrightarrow{P_1P_2}$ is not a singularity.

**Proof:** The items (i), (ii) and (iii) are natural consequences of Theorem 2.3.3 and Theorem 2.5.3. Theorem 2.3.3 assures that for $m \geq 2$, the $n$-bar system $\Gamma$ is locally feedback equivalent to the $m$-chained form at any point $q_0$ that corresponds to $x \in \mathbb{R}^{(n+1)(m+1)}$ satisfying $\Psi(x) = 0$ and $\sum_{j=1}^{m+1}(x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0$, for $1 \leq i \leq n-1$. Moreover, it was proved in Section 2.4.3 that around $q_0$, the subdistribution $\mathcal{L}_{n-1}$, which is involutive and of corank one in $\mathcal{D}^{(n-1)}$, is given by

$$\begin{aligned}(\mathcal{L}_{n-1})^{\perp} &= \operatorname{span}\{\Phi^*\mathrm{d}x_0^1, \ldots, \Phi^*\mathrm{d}x_0^{m+1}\} \\ &= \operatorname{span}\{\mathrm{d}\Phi^*x_0^1, \ldots, \mathrm{d}\Phi^*x_0^{m+1}\}.\end{aligned}$$

Notice that on the configuration space $Q$, we have always that $\Phi^*x_0^j = x_0^j$, for $1 \leq j \leq m+1$. Thus according to Theorem 2.5.3, the coordinates of $P_0 = (x_0^1, x_0^2, \ldots, x_0^{m+1})$ are minimal $x$-flat outputs of $\Gamma$ around $q_0$ which implies immediately that $\Gamma$ is $x$-flat at $(q_0, u^0)$ for some control $u^0$. Before we characterize the control $u^0$, notice that Theorem 2.5.3 implies that for control systems that are feedback equivalent to the $m$-chained form, for $n \geq 2, m \geq 2$, the minimal flat outputs are equivalent in the sense that for any two families of minimal flat outputs $(h_0, \ldots, h_m)$ and $(\tilde{h}_0, \ldots, \tilde{h}_m)$, we have $\operatorname{span}\{\mathrm{d}h_0, \ldots, \mathrm{d}h_m\} = \operatorname{span}\{\mathrm{d}\tilde{h}_0, \ldots, \mathrm{d}\tilde{h}_m\}$. In view of this and the item (ii) of Theorem 2.5.4, any minimal flat outputs $(h_0, \ldots, h_m)$ of the $n$-bar system satisfy $\operatorname{span}\{\mathrm{d}h_0, \ldots, \mathrm{d}h_m\} = \operatorname{span}\{\mathrm{d}x_0^1, \mathrm{d}x_0^2, \ldots, \mathrm{d}x_0^{m+1}\}$ in $\mathbb{R}^{m+1}$ for $n \geq 2, m \geq 2$. This proves (iii).

Now we are going to characterize the control $u^0$. According to the definition of the flat output, the entire state and all the controls should be functions of the coordinates

$x_0^1, x_0^2, \dots, x_0^{m+1}$ and their derivatives. Recall the system $\Delta$ given be (2.2.3) and (2.2.4)

$$\Delta : \quad \begin{cases} \dot{x}_0^j &= v_1(x_1^j - x_0^j) \\ &\vdots \\ \dot{x}_{n-1}^j &= v_n(x_n^j - x_{n-1}^j) \\ \dot{x}_n^j &= v_{n+j} \end{cases}, \quad 1 \le j \le m+1,$$

and consider the system of equation for the $x_0^j$-variables

$$\begin{cases} \dot{x}_0^1 &= v_1(x_1^1 - x_0^1) \\ \dot{x}_0^2 &= v_1(x_1^2 - x_0^2) \\ &\vdots \\ \dot{x}_0^{m+1} &= v_1(x_1^{m+1} - x_0^{m+1}) \\ \Psi_1(x) &= \sum_{j=1}^{m+1}(x_1^j - x_0^j)^2 - 1 = 0. \end{cases} \tag{2.5.1}$$

A direct computation shows that

$$v_1 = \left((\dot{x}_0^1)^2 + \cdots + (\dot{x}_0^{m+1})^2\right)^{\frac{1}{2}} = \eta_1(P_0, \dot{P}_0), \tag{2.5.2}$$

for some function $\eta_1$. Substituting (2.5.2) into (2.5.1), we get

$$\begin{aligned} x_1^1 &= x_0^1 + \frac{\dot{x}_0^1}{v_1} &&= \varphi_1^1(P_0, \dot{P}_0) \\ x_1^2 &= x_0^2 + \frac{\dot{x}_0^2}{v_1} &&= \varphi_1^2(P_0, \dot{P}_0) \\ &\vdots && \vdots \\ x_1^{m+1} &= x_0^{m+1} + \frac{\dot{x}_0^{m+1}}{v_1} &&= \varphi_1^{m+1}(P_0, \dot{P}_0), \end{aligned}$$

for some functions $\varphi_1^i$, for $1 \le i \le m+1$. Put $\varphi_1 = (\varphi_1^1, \dots, \varphi_1^{m+1})$, we thus have

$$\begin{aligned} P_1 &= (x_1^1, \dots, x_1^{m+1}) \\ &= (\varphi_1^1(P_0, \dot{P}_0), \dots, \varphi_1^{m+1}(P_0, \dot{P}_0)) \\ &= \varphi_1(P_0, \dot{P}_0). \end{aligned}$$

In the same way, we obtain that, for $2 \le i \le n$,

$$v_i = \eta_i(P_{i-1}, \dot{P}_{i-1}) = \tilde{\eta}_i(P_0, \dot{P}_0, P_0^{(2)}, \dots, P_0^{(i)})$$

and

$$P_i = \varphi_i(P_{i-1}, \dot{P}_{i-1}) = \tilde{\varphi}_i(P_0, \dot{P}_0, P_0^{(2)}, \dots, P_0^{(i)}),$$

for some functions $\tilde{\eta}_i$ and $\tilde{\varphi}_i$. Finally, the controls $v_{n+j}$, for $1 \le j \le m+1$, can be expressed by

$$v_{n+j} = \dot{x}_n^j$$

$$
\begin{aligned}
&= \frac{\mathrm{d}}{\mathrm{d}t}(\tilde{\varphi}_n^j(P_0, \dot{P}_0, P_0^{(2)}, \ldots, P_0^{(n)})) \\
&= \tau_j(P_0, \dot{P}_0, P_0^{(2)}, \ldots, P_0^{(n+1)}),
\end{aligned}
$$

for some functions $\tau_j$. In this way, the entire state and all controls $v_i$, for $1 \leq i \leq n+m+1$, are expressed as functions of the coordinates of $P_0 = (x_0^1, x_0^2, \ldots, x_0^{m+1})$ and their derivatives up to order $n+1$. The $n$-bar system $\Gamma$ has $m+1$ controls while the system $\Delta$ has $n+m+1$ controls. So there must be $n$ relations between controls of $\Delta$ when restricted to $Q = \{x \in X : \Psi(x) = 0\}$. We will see this below and at the same time we will clarify the problem of singularities. Clearly, in order that the above computations hold, all the controls $v_i$, for $1 \leq i \leq n$, cannot vanish. It is sufficient, however, to assume that the control $v_1$ is nonzero since around any point $x$ satisfying $\Psi(x) = 0$ and the regularity condition $\sum_{j=1}^{m+1}(x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j) \neq 0$, the condition $v_1 \neq 0$ implies that $v_i \neq 0$, for $2 \leq i \leq n$. In fact, differentiating the constraint

$$
\Psi_1(x) = (x_1^1 - x_0^1)^2 + (x_1^2 - x_0^2)^2 + \cdots + (x_1^{m+1} - x_0^{m+1})^2 - 1 = 0,
$$

we get

$$
\sum_{j=1}^{m+1} 2\left((x_1^j - x_0^j)\dot{x}_1^j - (x_1^j - x_0^j)\dot{x}_0^j\right) = 0.
$$

Substituting $\dot{x}_0^j = v_1(x_1^j - x_0^j)$ and $\dot{x}_1^j = v_2(x_2^j - x_1^j)$, for $1 \leq j \leq m+1$, into the above equation, by a simple calculation we get

$$
v_1 = v_2 \sum_{j=1}^{m+1}(x_1^j - x_0^j)(x_2^j - x_1^j).
$$

Recall that around any regular point $q_0$, we have always that $\sum_{j=1}^{m+1}(x_1^j - x_0^j)(x_2^j - x_1^j) \neq 0$. Therefore, the condition $v_1 \neq 0$ implies that $v_2 \neq 0$ and similarly, it can be shown that

$$
v_i = v_{i+1} \sum_{j=1}^{m+1}(x_i^j - x_{i-1}^j)(x_{i+1}^j - x_i^j),
$$

for $1 \leq i \leq n-1$ and

$$
v_n = \sum_{j=1}^{m+1}(x_n^j - x_{n-1}^j)v_{n+j}.
$$

The above equations show, first, that $v_1 \neq 0$ is equivalent to $v_i \neq 0$, $1 \leq i \leq n$. Secondly, they show that there exist $n$ relations between the controls $v_i$, $1 \leq i \leq n+m+1$, of $\Delta$ implying that the $n$-bar system possesses, indeed, $m+1$ controls. Moreover, the condition $v_1 \neq 0$ (which yields $v_i \neq 0$, $1 \leq i \leq n$) implies that the instantaneous velocity $\dot{P}_0$ of the point $P_0$ is nonzero (and, consequently, the instantaneous velocities $\dot{P}_i$ of the points $P_i$, $0 \leq i \leq n-1$, are nonzero). Therefore the condition (2) holds. □

## 2.6 Proof of Lemma 2.4.1

(a). $\mathrm{d}\omega_i^j \wedge \omega_i \wedge \omega_{i+1} = 0$, for $1 \le i \le n-1$, $1 \le j \le m+1$, and $j \ne \sigma(i)$.

**Proof:** For simplicity, we assume that $j < \sigma(i)$. The case of $j > \sigma(i)$ can be proved in the same way. Since $\Omega_i^j = (x_i^j - x_{i-1}^j)\mathrm{d}x_{i-1}^{\sigma(i)} - (x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})\mathrm{d}x_{i-1}^j$, we have

$$\mathrm{d}\Omega_i^j = \mathrm{d}x_i^j \wedge \mathrm{d}x_{i-1}^{\sigma(i)} - 2\mathrm{d}x_{i-1}^j \wedge \mathrm{d}x_{i-1}^{\sigma(i)} - \mathrm{d}x_i^{\sigma(i)} \wedge \mathrm{d}x_{i-1}^j.$$

Therefore

$$\begin{aligned}
\mathrm{d}\Omega_i^j \wedge \Omega_i &= \mathrm{d}\Omega_i^j \wedge \Omega_i^1 \wedge \cdots \wedge \Omega_i^{m+1} \\
&= (-1)^{\sigma(i)}(x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})^m(\wedge \mathrm{d}X_{i-1}) \wedge \mathrm{d}x_i^j \\
&\quad +(-1)^{\sigma(i)+1}(x_i^j - x_{i-1}^j)(x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})^{m-1}(\wedge \mathrm{d}X_{i-1}) \wedge \mathrm{d}x_i^{\sigma(i)} \\
&= \delta_1 + \delta_2,
\end{aligned}$$

where

$$\begin{aligned}
\delta_1 &= (-1)^{\sigma(i)}(x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})^m(\wedge \mathrm{d}X_{i-1}) \wedge \mathrm{d}x_i^j \\
\delta_2 &= +(-1)^{\sigma(i)+1}(x_i^j - x_{i-1}^j)(x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})^{m-1}(\wedge \mathrm{d}X_{i-1}) \wedge \mathrm{d}x_i^{\sigma(i)}.
\end{aligned}$$

Suppose that $\sigma(i) \le \sigma(i+1)$, a direct computation shows that

$$\begin{aligned}
\delta_1 \wedge \Omega_{i+1} &= (-1)^{\sigma(i)+m-1}(x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})^m(x_{i+1}^j - x_i^j)(x_{i+1}^{\sigma(i+1)} - x_i^{\sigma(i+1)})^{m-1} \\
&\quad (\wedge \mathrm{d}X_{i-1}) \wedge x_i^j \wedge x_i^1 \wedge \cdots \wedge x_i^{j-1} \wedge x_i^{\sigma(i+1)} \wedge x_i^{j+1} \wedge \cdots \\
&\quad x_i^{\sigma(i+1)-1} \wedge x_i^{\sigma(i+1)+1} \wedge \cdots \wedge x_i^{m+1} \\
&= (-1)^{m+\sigma(i)+\sigma(i+1)-1}(x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})^m(x_{i+1}^j - x_i^j)(x_{i+1}^{\sigma(i+1)} - x_i^{\sigma(i+1)})^{m-1} \\
&\quad (\wedge \mathrm{d}X_{i-1}) \wedge \mathrm{d}X_i,
\end{aligned}$$

and

$$\begin{aligned}
\delta_2 \wedge \Omega_{i+1} &= (-1)^{\sigma(i)+m}(x_i^j - x_{i-1}^j)(x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})^{m-1}(x_{i+1}^j - x_i^j) \\
&\quad (x_{i+1}^{\sigma(i+1)} - x_i^{\sigma(i+1)})^{m-1}(x_{i+1}^{\sigma(i)} - x_i^{\sigma(i)}) \\
&\quad (\wedge \mathrm{d}X_{i-1}) \wedge \mathrm{d}x_i^{\sigma(i)} \wedge \mathrm{d}x_i^1 \wedge \cdots \wedge \mathrm{d}x_i^{\sigma(i)-1} \wedge \mathrm{d}x_i^{\sigma(i+1)} \wedge \mathrm{d}x_i^{\sigma(i)+1} \\
&\quad \wedge \cdots \wedge \mathrm{d}x_i^{\sigma(i+1)-1} \wedge \mathrm{d}x_i^{\sigma(i+1)+1} \wedge \cdots \wedge \mathrm{d}x_i^{m+1} \\
&= (-1)^{m+\sigma(i)+\sigma(i+1)}(x_i^j - x_{i-1}^j)(x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})^{m-1}(x_{i+1}^j - x_i^j) \\
&\quad (x_{i+1}^{\sigma(i+1)} - x_i^{\sigma(i+1)})^{m-1}(x_{i+1}^{\sigma(i)} - x_i^{\sigma(i)})(\wedge \mathrm{d}X_{i-1}) \wedge \mathrm{d}X_i.
\end{aligned}$$

Differentiating the constraint

$$\Psi_i = (x_i^1 - x_{i-1}^1)^2 + (x_i^2 - x_{i-1}^2)^2 + \cdots + (x_i^{m+1} - x_{i-1}^{m+1})^2 - 1 = 0,$$

we get

$$(x_i^1-x_{i-1}^1)\mathrm{d}x_i^1-(x_i^1-x_{i-1}^1)\mathrm{d}x_{i-1}^1+\cdots+(x_i^{m+1}-x_{i-1}^{m+1})\mathrm{d}x_i^{m+1}-(x_i^{m+1}-x_{i-1}^{m+1})\mathrm{d}x_{i-1}^{m+1}=0,$$

which implies that $\Phi^*\mathrm{d}X_{i-1}$ and $\Phi^*\mathrm{d}X_i$ are dependent on the configuration space $Q$. Then

$$\begin{aligned}
\mathrm{d}\omega_i^j\wedge\omega_i\wedge\omega_{i+1} &= \Phi^*(\mathrm{d}\Omega_i^j\wedge\Omega_i\wedge\Omega_{i+1})\\
&= \Phi^*(\delta_1\wedge\Omega_{i+1}+\delta_2\wedge\Omega_{i+1})\\
&= (-1)^{\sigma(i)+m+\sigma(i+1)}\cdot[(x_i^j-x_{i-1}^j)(x_i^{\sigma(i)}-x_{i-1}^{\sigma(i)})^{m-1}(x_{i+1}^j-x_i^j)\\
&\quad (x_{i+1}^{\sigma(i+1)}-x_i^{\sigma(i+1)})^{m-1}(x_{i+1}^{\sigma(i)}-x_i^{\sigma(i)})-(x_i^{\sigma(i)}-x_{i-1}^{\sigma(i)})^m\\
&\quad (x_{i+1}^j-x_i^j)(x_{i+1}^{\sigma(i+1)}-x_i^{\sigma(i+1)})^{m-1}]\cdot(\wedge\Phi^*\mathrm{d}X_{i-1})\wedge\Phi^*\mathrm{d}X_i\\
&= 0.
\end{aligned}$$

The case of $\sigma(i)>\sigma(i+1)$ can be treated in a similar way. $\square$

(b). $\quad \wedge\,\omega_i=(-1)^m(x_i^{\sigma(i)}-x_{i-1}^{\sigma(i)})^{m-2}(\wedge\Phi^*\mathrm{d}\check{X}_{i-1}^{\sigma(i)}) \mod (\Phi^*\mathrm{d}X_i)$,

for $1\le i\le n-1$, where $(\Phi^*\mathrm{d}X_i)$ denotes the ideal generated by $\Phi^*\mathrm{d}x_i^1,\ldots,\Phi^*\mathrm{d}x_i^{m-1}$.

**Proof:** Consider the case $i=1$ and denote $\sigma(1)=k$, then

$$\Omega_1^j=(x_1^j-x_0^j)\mathrm{d}x_0^k-(x_1^k-x_0^k)\mathrm{d}x_0^j,$$

for $1\le j\le m+1$ and $j\ne k$. Differentiating the constraint

$$\Psi_1=(x_1^1-x_0^1)^2+(x_1^2-x_0^2)^2+\cdots+(x_1^{m+1}-x_0^{m+1})^2-1=0,$$

we obtain

$$\mathrm{d}x_0^k=-\sum_{j=1,j\ne k}^{m+1}\frac{x_1^j-x_0^j}{x_1^k-x_0^k}\mathrm{d}x_0^j \mod (\mathrm{d}X_1),$$

and hence

$$\begin{aligned}
\omega_1^j &= \Phi^*(\Omega_1^j)\\
&= -\Big(\frac{(x_1^j-x_0^j)^2}{x_1^k-x_0^k}+(x_1^j-x_0^j)\Big)\Phi^*\mathrm{d}x_0^j-\sum_{l=1,l\ne k,l\ne j}^{m+1}\frac{x_1^l-x_0^l}{x_1^k-x_0^k}(x_1^j-x_0^j)\Phi^*\mathrm{d}x_0^1 \mod (\Phi^*\mathrm{d}X_1).
\end{aligned}$$

Since $\Phi^*\mathrm{d}x_0^1,\ldots,\Phi^*\mathrm{d}x_0^{k-1},\Phi^*\mathrm{d}x_0^{k+1},\ldots,\Phi^*\mathrm{d}x_0^{m+1}$ are independent on $Q$, comparing the coefficients of $\Phi^*\mathrm{d}x_0^j$ it is easy to see that

$$\begin{aligned}
\wedge\omega_1 &= \omega_1^1\wedge\cdots\omega_1^{k-1}\wedge\omega_1^{k+1}\wedge\cdots\wedge\omega_1^{m+1}\\
&= (-1)^m\frac{1}{(x_1^k-x_0^k)^m}\cdot|A|\,\Phi^*\mathrm{d}x_0^1\wedge\cdots\Phi^*\mathrm{d}x_0^{k-1}\wedge\Phi^*\mathrm{d}x_0^{k+1}\wedge\cdots\wedge\Phi^*\mathrm{d}x_0^{m+1},
\end{aligned}$$

where $|A|$ is given by, in fact, a special case of the following determinant $J_m$,

$$J_m = \begin{vmatrix} a_1^2 + a_k^2 & a_1 a_2 & \cdots & a_1 a_{k-1} & a_1 a_{k+1} & \cdots & a_1 a_{m+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{k-1} a_1 & a_{k-1} a_2 & \cdots & a_{k-1}^2 + a_k^2 & a_{k-1} a_{k+1} & \cdots & a_{k-1} a_{m+1} \\ a_{k+1} a_1 & a_{k+1} a_2 & \cdots & a_{k+1} a_{k-1} & a_{k+1}^2 + a_k^2 & \cdots & a_{k+1} a_{m+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m+1} a_1 & a_{m+1} a_2 & \cdots & a_{m+1} a_{k-1} & a_{m+1} a_{k+1} & \cdots & a_{m+1}^2 + a_k^2 \end{vmatrix}$$

with $a_i = x_1^i - x_0^i$, where $a_k \neq 0$ and $\displaystyle\sum_{i=1}^{m+1} a_i^2 = 1$. By a straightforward calculation, we obtain that

$$J_m = a_k^{2m-1} \cdot (a_1^2 + \cdots + a_{m+1}^2) = a_k^{2(m-1)},$$

and thus

$$|A| = (x_1^k - x_0^k)^{2(m-1)}.$$

Therefore, we have

$$\begin{aligned} \wedge\omega_1 &= (-1)^m (x_1^k - x_0^k)^{m-2} (\wedge \Phi^* \mathrm{d}\check{X}_0^k) \mod (\Phi^* \mathrm{d}X_1) \\ &= (-1)^m (x_1^{\sigma(1)} - x_0^{\sigma(1)})^{m-2} (\wedge \Phi^* \mathrm{d}\check{X}_0^{\sigma(1)}) \mod (\Phi^* \mathrm{d}X_1). \end{aligned}$$

Similarly, it can be proved, for $1 \leq i \leq n-1$, that

$$\wedge\omega_i = (-1)^m (x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})^{m-2} (\wedge \Phi^* \mathrm{d}\check{X}_{i-1}^{\sigma(i)}) \mod (\Phi^* \mathrm{d}X_i).$$

□

(c). $\quad \mathrm{d}\omega_i^j \wedge \omega_i \wedge \cdots \wedge \omega_1 \neq 0, \quad 1 \leq j \leq m+1, j \neq \sigma(i)$

**Proof:** We give here the proof for $i = n$ and the other cases can be shown in the same way. For simplicity, we assume that $j < \sigma(n)$. A direct computation shows that

$$\begin{aligned} \mathrm{d}\Omega_n^j \wedge \Omega_n &= (-1)^{m+\sigma(n)-1} (x_n^{\sigma(n)} - x_{n-1}^{\sigma(n)})^m \mathrm{d}x_n^j \wedge \mathrm{d}X_{n-1} \\ &\quad + (-1)^{m+\sigma(n)} (x_n^j - x_{n-1}^j)(x_n^{\sigma(n)} - x_{n-1}^{\sigma(n)})^{m-1} \mathrm{d}x_n^{\sigma(n)} \wedge \mathrm{d}X_{n-1}. \end{aligned}$$

Moreover, differentiating the constraint

$$\Psi_n = (x_n^1 - x_{n-1}^1)^2 + (x_n^2 - x_{n-1}^2)^2 + \cdots + (x_n^{m+1} - x_{n-1}^{m+1})^2 - 1 = 0$$

we get

$$\mathrm{d}x_n^{\sigma(n)} = - \sum_{l=1, l\neq\sigma(n)}^{m+1} \frac{x_n^l - x_{n-1}^l}{x_n^{\sigma(n)} - x_{n-1}^{\sigma(n)}} \mathrm{d}x_n^l \mod (\mathrm{d}X_{n-1}).$$

Therefore

$$\mathrm{d}\omega_n^j \wedge \omega_n = \Phi^* (\mathrm{d}\Omega_n^j \wedge \Omega_n)$$

$$
\begin{aligned}
&= (-1)^{m+\sigma(n)-1}[(x_n^{\sigma(n)} - x_{n-1}^{\sigma(n)})^m \, \Phi^*\mathrm{d}x_n^j + (x_n^j - x_{n-1}^j)^2 (x_n^{\sigma(n)} - x_{n-1}^{\sigma(n)})^{m-2} \, \Phi^*\mathrm{d}x_n^j \\
&\quad + \sum_{l=2, l\neq\sigma(n), l\neq j}^{m+1} (x_n^{\sigma(n)} - x_{n-1}^{\sigma(n)})^{m-2} (x_n^j - x_{n-1}^j)(x_n^l - x_{n-1}^l) \, \Phi^*\mathrm{d}x_n^l] \wedge \Phi^*\mathrm{d}X_{n-1} \\
&= (-1)^{m+\sigma(n)-1} \, \eta_j \cdot \, \Phi^*\mathrm{d}x_n^j \wedge \Phi^*\mathrm{d}X_{n-1} \\
&\quad + (-1)^{m+\sigma(n)-1} \sum_{l=2, l\neq\sigma(n), l\neq j}^{m+1} \eta_l \cdot \, \Phi^*\mathrm{d}x_n^l \wedge \Phi^*\mathrm{d}X_{n-1},
\end{aligned}
$$

where

$$
\begin{aligned}
\eta_j &= (x_n^{\sigma(n)} - x_{n-1}^{\sigma(n)})^m + (x_n^j - x_{n-1}^j)^2 (x_n^{\sigma(n)} - x_{n-1}^{\sigma(n)})^{m-2}, \\
\eta_l &= (x_n^{\sigma(n)} - x_{n-1}^{\sigma(n)})^{m-2} (x_n^j - x_{n-1}^j)(x_n^l - x_{n-1}^l).
\end{aligned}
$$

Since $\Phi^*\mathrm{d}x_n^1, \ldots, \Phi^*\mathrm{d}x_n^{\sigma(n)-1}, \Phi^*\mathrm{d}x_n^{\sigma(n)+1}, \ldots, \Phi^*\mathrm{d}x_n^{m+1}$ are independent on $Q$, all differential forms $\Phi^*\mathrm{d}x_n^i \wedge \Phi^*\mathrm{d}X_{n-1} \wedge \omega_{n-1} \wedge \cdots \wedge \omega_1$, for $1 \le i \le m+1$ and $i \neq \sigma(n)$, are also independent. Therefore, it is enough to prove

$$
\eta_j \, \cdot \Phi^*\mathrm{d}x_n^j \wedge \Phi^*\mathrm{d}X_{n-1} \wedge \omega_{n-1} \wedge \cdots \wedge \omega_1 \neq 0, \tag{2.6.1}
$$

to conclude that

$$
\begin{aligned}
\mathrm{d}\omega_n^j \wedge \omega_n \wedge \cdots \wedge \omega_1 &= (-1)^{m+\sigma(n)-1} \, \eta_j \, \cdot \Phi^*\mathrm{d}x_n^j \wedge \Phi^*\mathrm{d}X_{n-1} \wedge \omega_{n-1} \wedge \cdots \wedge \omega_1 \\
&\quad + (-1)^{m+\sigma(n)-1} \sum_{l=2, l\neq\sigma(n), j}^{m+1} \eta_l \cdot \, \Phi^*\mathrm{d}x_n^l \wedge \Phi^*\mathrm{d}X_{n-1} \wedge \omega_{n-1} \wedge \cdots \wedge \omega_1 \neq 0.
\end{aligned}
$$

Notice that $\eta_j = (x_n^{\sigma(n)} - x_{n-1}^{\sigma(n)})^m + (x_n^j - x_{n-1}^j)^2 (x_n^{\sigma(n)} - x_{n-1}^{\sigma(n)})^{m-2} \neq 0$ is obvious since $x_n^{\sigma(n)} - x_{n-1}^{\sigma(n)} \neq 0$. Hence in order to prove (2.6.1), it is sufficient to show that

$$
\Phi^*\mathrm{d}x_n^j \wedge \Phi^*\mathrm{d}X_{n-1} \wedge \omega_{n-1} \wedge \cdots \wedge \omega_1 \neq 0.
$$

Differentiating the constraint

$$
\Psi_{n-1} = (x_{n-1}^1 - x_{n-2}^1)^2 + (x_{n-1}^2 - x_{n-2}^2)^2 + \cdots + (x_{n-1}^{m+1} - x_{n-2}^{m+1})^2 - 1 = 0
$$

implies the following relation

$$
\mathrm{d}x_{n-1}^{\sigma(n-1)} = \mathrm{d}x_{n-2}^{\sigma(n-1)} \mod (\mathrm{d}x_{n-1}^j, \mathrm{d}x_{n-2}^j, \quad 1 \le j \le m+1, \quad j \neq \sigma(n-1)).
$$

Recall the condition (b) of Lemma 2.4.1,

$$
\wedge\omega_i = (-1)^m (x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})^{m-2} \, (\wedge \Phi^*\mathrm{d}\check{X}_{i-1}^{\sigma(i)}) \mod (\Phi^*\mathrm{d}X_i),
$$

and a straightforward calculation shows that

$$
\Phi^*\mathrm{d}x_n^j \wedge \Phi^*\mathrm{d}X_{n-1} \wedge \omega_{n-1}
$$

$$
\begin{aligned}
&= (-1)^m (x_{n-1}^{\sigma(n-1)} - x_{n-2}^{\sigma(n-1)})^{m-2}\, \Phi^* \mathrm{d}x_n^j \wedge \Phi^* \mathrm{d}X_{n-1} \wedge \Phi^* \mathrm{d}\check{X}_{n-2}^{\sigma(n-1)} \\
&= (-1)^{2m+1} (x_{n-1}^{\sigma(n-1)} - x_{n-2}^{\sigma(n-1)})^{m-2}\, \Phi^* \mathrm{d}x_n^j \wedge \Phi^* \mathrm{d}\check{X}_{n-1}^{\sigma(n-1)} \wedge \Phi^* \mathrm{d}X_{n-2}.
\end{aligned}
$$

By a similar computation as above, we get

$$
\begin{aligned}
&\Phi^* \mathrm{d}x_n^j \wedge \Phi^* \mathrm{d}X_{n-1} \wedge \omega_{n-1} \wedge \cdots \wedge \omega_1 = \\
&\quad (-1)^{n(m+1)-1} \prod_{i=1}^{n-1} (x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)})^{m-2}\, \Phi^* \mathrm{d}x_n^j \wedge \Phi^* \mathrm{d}\check{X}_{n-1}^{\sigma(n-1)} \wedge \cdots \Phi^* \mathrm{d}\check{X}_1^{\sigma(1)} \wedge \Phi^* \mathrm{d}X_0 \neq 0
\end{aligned}
$$

since $x_i^{\sigma(i)} - x_{i-1}^{\sigma(i)} \neq 0$, for $1 \leq i \leq n-1$, and $\Phi^* \mathrm{d}x_n^j, \Phi^* \mathrm{d}\check{X}_{n-1}^{\sigma(n-1)}, \ldots, \Phi^* \mathrm{d}\check{X}_1^{\sigma(1)}, \Phi^* \mathrm{d}X_0$ are independent on the configuration space $Q$. Therefore, (2.6.1) is true and item (c) holds. □

# Chapter 3

# Cartan distribution for surfaces

## 3.1 Introduction

Consider $J^n(\mathbb{R}^k,\mathbb{R}^m) \cong \mathbb{R}^N$, for a suitable $N$, the space of $n$-jets of smooth maps from $\mathbb{R}^k$ into $\mathbb{R}^m$ and denote its canonical coordinates by

$$(x_1,\ldots,x_k,y_1,\ldots,y_m,p_j^\sigma,\quad 1\leq j\leq m,\quad 1\leq|\sigma|\leq n),$$

where $x_i$, for $1\leq i\leq k$, represent independent variables and $y_j$, for $1\leq j\leq m$, represent dependent variables; the vector of non-negative integers $\sigma=(\sigma_1,\ldots,\sigma_k)$ is a multi-index such that $|\sigma|=\sigma_1+\cdots+\sigma_k\leq n$ and $p_j^\sigma$, for $1\leq j\leq m$, correspond to the partial derivatives $\dfrac{\partial^{|\sigma|}y_j}{\partial x^\sigma}$, where $\partial x^\sigma$ stands for $\partial^{\sigma_1}x_1\cdots\partial^{\sigma_k}x_k$. Denote $p_j^0=y_j$, for $1\leq j\leq m$. Any smooth map $\varphi=(\varphi_1,\ldots,\varphi_m)$ from $\mathbb{R}^k$ into $\mathbb{R}^m$ defines a submanifold in $J^n(\mathbb{R}^k,\mathbb{R}^m)$ by the relation

$$p_j^\sigma=\frac{\partial^{|\sigma|}\varphi_j}{\partial x^\sigma}(x_1,\ldots,x_k),$$

for $1\leq j\leq m$ and $0\leq|\sigma|\leq n$. This submanifold is called the $n$-graph of $\varphi$. It turns out that all $n$-graphs are integral submanifolds, of dimension $k$, of a distribution called the *Cartan distribution* on $J^n(\mathbb{R}^k,\mathbb{R}^m)$ and denoted by $\mathcal{CC}^n(\mathbb{R}^k,\mathbb{R}^m)$. The Pfaffian system that annihilates $\mathcal{CC}^n(\mathbb{R}^k,\mathbb{R}^m)$, which is called the *canonical contact Pfaffian system on* $J^n(\mathbb{R}^k,\mathbb{R}^m)$, is given in the canonical coordinates of $J^n(\mathbb{R}^k,\mathbb{R}^m)$ by

$$\mathrm{d}p_j^\sigma-\sum_{i=1}^k p_j^{\sigma+1_i}dx_i=0,\quad \text{for}\ \ 1\leq j\leq m,\ \ \text{and}\ \ 0\leq|\sigma|\leq n-1,$$

where $\sigma+1_i=(\sigma_1,\ldots,\sigma_i+1,\ldots,\sigma_k)$.

A natural problem is to characterize those distributions which are (locally) equivalent to a Cartan distribution. This problem was posed by Pfaff [55] in 1814 and is still open in its full generality although many important particular solutions have been obtained.

In the case of $n = m = 1$ and an arbitrary $k$, the solution has been obtained by Darboux [10] in his famous theorem generalizing earlier results of Pfaff [55] and Frobenius [15]. The case $n = 2$, $m = 1$ and $k = 1$ was solved by Engel in [11]. The case $n \geq 2$, $m = 1$ and $k = 1$ was solved by E. von Weber [78], Cartan [9] and Goursat [18] (at generic points) and by Libermann [37], Kumpera and Ruiz [33], and Murray [45] (at an arbitrary point). The case $n = 1$ with $k$ and $m$ arbitrary has bee studied by Gardner and solved by Bryant [7], [8]. The case $k = 1$, with $n$ and $m$ arbitrary (that is, characterization of Cartan distributions for curves) has been studied by Gardner et Shadwick [16], Murray [45], Tilbury et Sastry [76], Aranda-Bricaire and Pomet [3], Mormul [43], Respondek and Pasillas-Lépine [51], [54]. The general case of arbitrary $n$,$m$, and $k$ has been studied by Yamaguchi [79].

This chapter is devoted to the problem of when a given distribution is locally equivalent to the Cartan distribution in the case $k = 2$, that is, Cartan distributions for surfaces, and is organized as follows. In Section 3.2, we define the Cartan distribution for surfaces. In Section 3.3, we give our main results. In Section 3.4, we recall the Bryant theorem (describing the case $n = 1$, with $k$ and $m$ arbitrary) and propose its extension that will be useful in the proof of our main result. Finally, we prove our results in Section 3.5.

## 3.2 Cartan distributions for surfaces

A *distribution* $\mathcal{D}$ on a $C^\infty$-smooth manifold $M$ is a map that assigns smoothly to each point $q \in M$ a linear subspace $\mathcal{D}(q) \subset T_qM$. Given a distribution $\mathcal{D}$, we will denote by $\Gamma(\mathcal{D})$ the submodule of the module $V^\infty(M)$ of $C^\infty$-smooth vector fields consisting of $C^\infty$-smooth sections of $TM$ with values in $\mathcal{D}$, i.e.,

$$\Gamma(\mathcal{D}) = \{f \in V^\infty(M) : f(q) \in \mathcal{D}(q)\},$$

for any $q \in M$. We will say that a distribution $\mathcal{D}$ is $C^\infty$-smooth if $\mathcal{D}(q) = \Gamma(\mathcal{D})(q)$ for any $q \in M$. $\mathcal{D}$ will be called *nonsingular* if $\dim \mathcal{D}(q)$ is constant and in this case we will say that $\mathcal{D}$ is of rank $k$, where $\dim \mathcal{D}(q) = k$ for any $q \in M$.

The *derived flag* of a distribution $\mathcal{D}$ is the sequence of modules of vector fields $\mathcal{D}^{(0)} \subset \mathcal{D}^{(1)} \subset \cdots$ defined inductively by

$$\mathcal{D}^{(0)} = \Gamma(\mathcal{D}) \quad \text{and} \quad \mathcal{D}^{(i+1)} = \Gamma(\mathcal{D}^{(i)}) + [\Gamma(\mathcal{D}^{(i)}), \Gamma(\mathcal{D}^{(i)})], \quad \text{for} \quad i \geq 0.$$

The *Lie flag* is the sequence of modules of vector fields $\mathcal{D}_0 \subset \mathcal{D}_1 \subset \cdots$ defined inductively by

$$\mathcal{D}_0 = \Gamma(\mathcal{D}) \quad \text{and} \quad \mathcal{D}_{i+1} = \Gamma(\mathcal{D}_i) + [\Gamma(\mathcal{D}_0), \Gamma(\mathcal{D}_i)], \quad \text{for} \quad i \geq 0.$$

In general, the derived and Lie flags are different; though for any point $q$ in the underlying manifold the inclusion $\mathcal{D}_i(q) \subset \mathcal{D}^{(i)}(q)$ holds, for $i \geq 0$.

A Pfaffian system on a smooth manifold $M$ is a differential system $\omega^1 = \cdots = \omega^m = 0$, where $\omega^i$ are differential 1-forms on $M$. We associate to it the codistribution $\mathcal{I}$ by putting

$$\mathcal{I}(q) = \mathrm{vect}_{\mathbb{R}} \{\omega^1(q), \ldots, \omega^m(q)\} \subset T_q^*M.$$

Conversely, let $\mathcal{I}$ be a smooth codistribution that assigns smoothly to each point $q$ in $M$ a linear subspace $\mathcal{I}(q) \subset T_q^*M$ of dimension $m$. Such a field of cotangent $m$-planes is spanned locally by $m$ pointwise linearly independent smooth differential 1-forms $\omega^1, \ldots, \omega^m$ on $M$, which will be denoted by $\mathcal{I} = \mathrm{span}\{\omega^1, \ldots, \omega^m\}$. So in the case of constant rank we can (locally) identity codistributions and Pfaffian systems which we will do throughout this paper.

For a Pfaffian system $\mathcal{I}$, we can define its *derived flag* $\mathcal{I}^{(0)} \supset \mathcal{I}^{(1)} \supset \cdots$ by

$$\mathcal{I}^{(0)} = \mathcal{I} \quad \text{and} \quad \mathcal{I}^{(i+1)} = \{\omega \in \mathcal{I}^{(i)} : \mathrm{d}\omega \equiv 0 \mod \mathcal{I}^{(i)}\}, \quad \text{for } i \geq 0,$$

provided that each element $\mathcal{I}^{(i)}$ of this sequence has constant rank. In this case, it is immediate to see that the derived flag of the distribution $\mathcal{D} = \mathcal{I}^{\perp}$ coincides with the sequence of distributions that annihilate the elements of the derived flag of $\mathcal{I}$, that is $\mathcal{D}^{(i)} = (\mathcal{I}^{(i)})^{\perp}$, for $i \geq 0$.

**Remark**: If $\mathcal{D}$ is a smooth nonsingular distribution, then with no confusion we can identify $\mathcal{D}$ and $\Gamma(\mathcal{D})$ and will do it throughout this chapter. If, however, $\mathcal{D}$ is singular or non-smooth, then the distinction between $\mathcal{D}$ and $\Gamma(\mathcal{D})$ is essential. Notice that if $\mathcal{D}$ is $C^\infty$-smooth, then the function $q \mapsto \dim \mathcal{D}(q)$ is lower semi-continuous (in other words, it can drop at a point but not jump up) because if vector fields belonging to $\mathcal{D}$ are independent at a point so they are in a neighborhood. In the paper, however, we will have to deal with non-smooth distributions $\mathcal{D}$ for which $q \to \dim \mathcal{D}(q)$ jumps up at a point (and thus is not lower semi-continuous). This will be the case if $\mathcal{I}$ is a smooth codistribution of a non-constant rank and thus $\mathcal{D} = \mathcal{I}^{\perp}$, defined as $\mathcal{D}(q) = (\mathcal{I}(q))^{\perp}$ is non-smooth for which $q \to \dim\mathcal{D}(q)$ is actually upper semi-continuous. Indeed, $\dim \mathcal{D}(q) = n - \dim \mathcal{I}(q)$, where $q \to \dim \mathcal{I}(q)$ is lower semi-continuous because $\mathcal{I}$ was assumed to be smooth.

Given a distribution $\mathcal{D}$ of constant rank $k$ on a $n$-dimensional manifold $M$, choose locally $m = n - k$ differential 1-forms $\omega^1, \ldots, \omega^m$ such that

$$\mathcal{I} = \mathcal{D}^{\perp} = \mathrm{span}\{\omega^1, \ldots, \omega^m\}.$$

We define the characteristic distribution $\mathcal{C}(\mathcal{D})$ of $\mathcal{D}$ pointwise by

$$\mathcal{C}(\mathcal{D})(q) = \{v \in \mathcal{D}(q) : v \lrcorner \, \mathrm{d}\omega(q) \in \mathcal{I}(q), \quad \forall \, \omega \in \mathcal{I}\}.$$

Notice that, in general, $\mathcal{C}(\mathcal{D})$ may not be a smooth distribution and $\dim \mathcal{C}(\mathcal{D})(q)$ may jump up at a point. If $\mathcal{C}(\mathcal{D})$ is nonsingular (in other words, of constant rank), then it can be equivalently defined as the module of its sections:

$$\Gamma(\mathcal{C}(\mathcal{D})) = \{g \in \Gamma(\mathcal{D}) : [g, f] \in \Gamma(\mathcal{D}), \quad \forall\, f \in \Gamma(\mathcal{D})\}.$$

Denote by $(x_1, x_2)$ coordinates on $\mathbb{R}^2$ and by $(y_1, \ldots, y_m)$ coordinates on $\mathbb{R}^m$. According to Introduction, we denote by

$$(x_1, x_2, p_j^\sigma, \quad 1 \leq j \leq m, \quad 0 \leq |\sigma| \leq n),$$

the canonical coordinates on $J^n(\mathbb{R}^2, \mathbb{R}^m)$, where $\sigma = (\sigma_1, \sigma_2)$ such that $\sigma_1, \sigma_2 \geq 0$ and $|\sigma| = \sigma_1 + \sigma_2 \leq n$ and (by definition) $p_j^0 = y_j$, for $1 \leq j \leq m$. As we already explained in the Introduction, the surface $\{y_j = \varphi_j(x_1, x_2)\}$, $1 \leq j \leq m$, defined by any smooth map $\varphi$ from $\mathbb{R}^2$ into $\mathbb{R}^m$, can be lifted to the surface

$$\left\{ p_j^\sigma = \frac{\partial^{|\sigma|}\varphi_j}{\partial x^\sigma}(x_1, x_2) \right\},$$

in $J^n(\mathbb{R}^2, \mathbb{R}^m)$, called the $n$-graph of the surface $\{y_j = \varphi_j(x_1, x_2)\}$. As we pointed out in the Introduction, we can endow $J^n(\mathbb{R}^2, \mathbb{R}^m)$ with a (nonintegrable) distribution. Namely, a surface in $J^n(\mathbb{R}^2, \mathbb{R}^m)$, that projects to $\mathbb{R}^2$, is the lift of a surface in $\mathbb{R}^m$ if and only if it is an integral submanifold of the Cartan distribution $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$ on $J^n(\mathbb{R}^2, \mathbb{R}^m)$, called also the *Cartan distribution for surfaces*, spanned by the following vector fields

$$\frac{\partial}{\partial x_1} + \sum_{|\sigma|=0}^{n-1} \sum_{j=1}^{m} p_j^{\sigma+1_1} \frac{\partial}{\partial p_j^\sigma}, \quad \frac{\partial}{\partial x_2} + \sum_{|\sigma|=0}^{n-1} \sum_{j=1}^{m} p_j^{\sigma+1_2} \frac{\partial}{\partial p_j^\sigma}, \quad \frac{\partial}{\partial p_j^\sigma}, \quad 1 \leq j \leq m, \quad |\sigma| = n.$$

The annihilator of $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$ is the Pfaffian system, called the *canonical contact Pfaffian system for surfaces*, spanned by

$$\mathrm{d}p_j^\sigma - p_j^{\sigma+1_1}\mathrm{d}x_1 - p_j^{\sigma+1_2}\mathrm{d}x_2, \quad 1 \leq j \leq m, \ 0 \leq |\sigma| \leq n-1.$$

We have $\dim J^n(\mathbb{R}^2, \mathbb{R}^m) = \frac{m}{2}(n+1)(n+2) + 2$; the Cartan distribution $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$ and its annihilator are, respectively, of rank $m(n+1) + 2$ and $\frac{m}{2}n(n+1)$.

The aim of this chapter is to identify all distributions $\mathcal{D}$ of rank $m(n+1)+2$ on $\mathbb{R}^N$ where $N = \frac{m}{2}(n+1)(n+2) + 2$, that are locally equivalent to the Cartan distribution $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$, that is, for which there exists a local diffeomorphism $\varphi : \mathbb{R}^N \to J^n(\mathbb{R}^2, \mathbb{R}^m)$ such that

$$\varphi_*\mathcal{D} = \mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m).$$

## 3.3 Characterization of Cartan distributions for surfaces

### 3.3.1 Main result

The crucial notion will be that of the Engel rank.

**Definition 3.3.1** The *Engel rank*, at a point $q$, of a Pfaffian system $\mathcal{I} = \operatorname{span}\{\omega^1, \ldots, \omega^m\}$, denoted by $\operatorname{rank}_E \mathcal{I}(q)$, is the largest integer $k$ such that there exists a 1-form $\alpha$ in $\mathcal{I}$ for which

$$((d\alpha)^k \wedge \omega^1 \wedge \cdots \wedge \omega^m)(q) \neq 0.$$

In Section 3.3.2, we will describe relations of the Engel rank with Cartan rank and, in particular, with the existence of involutive subdistributions of corank $k$ in $\mathcal{D} = \mathcal{I}^\perp$.

**Theorem 3.3.2** *Assume that $m \geq 3$. A distribution $\mathcal{D}$ of rank $m(n+1)+2$ on a manifold $M$ of dimension $\frac{m}{2}(n+1)(n+2)+2$ is locally, around a point $q$ of $M$, equivalent to the Cartan distribution for surfaces $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$ if and only if the following conditions hold:*

(i) *Each element $\mathcal{D}^{(i)}$, for $0 \leq i \leq n$, has constant rank equal $\frac{m}{2}(2n+2-i)(i+1)+2$.*

(ii) *Each element $\mathcal{D}^{(i)}$, for $0 \leq i \leq n-2$, contains the characteristic distribution $\mathcal{C}_{i+1}$ of $\mathcal{D}^{(i+1)}$ that is of constant corank two in $\mathcal{D}^{(i)}$. Moreover, the characteristic distribution of $\mathcal{D}$ is empty.*

(iii) *The Engel rank of $(\mathcal{D}^{(n-1)})^\perp$ is constant equal two.*

(iv) $\dim(\mathcal{L}_{n-1} + \mathcal{D}^{(i)})(q) = \dim \mathcal{L}_{n-1}(q) + 2$, *for $0 \leq i \leq n-2$, where $\mathcal{L}_{n-1}$ is an involutive subdistribution of corank two in $\mathcal{D}^{(n-1)}$ that exists due to* (iii).

**Remark 1:** The condition (iii) is the essential condition identifying Cartan distribution for surfaces. We want to underline two of its features. First, it is easily verifiable in terms of a given distribution $\mathcal{D}$ (see the definition of the Engel rank). Secondly, it has an elegant geometric interpretation: it is equivalent to the existence of an involutive subdistribution, which we denote by $\mathcal{L}_{n-1}$, of corank two in $\mathcal{D}^{(n-1)}$. We will discuss the geometric interpretation of the condition (iii) in Section 3.3.2

**Remark 2:** The condition (ii) implies that each $\mathcal{D}^{(i)}$, $0 \leq i \leq n-2$, also contains an involutive subdistribution $\mathcal{L}_i$ of corank two in $\mathcal{D}^{(i)}$. Checking the existence of such an involutive distribution does not require, however, calculating the Engel rank of $\mathcal{D}^{(i)}$ because the condition (ii) assures that, indeed, $\mathcal{L}_i = \mathcal{C}_{i+1} = \mathcal{C}(\mathcal{D}^{(i+1)})$. In view of this, the particularity of $\mathcal{L}_{n-1}$ is seen: in fact, $\mathcal{L}_{n-1}$ is not a characteristic distribution of $\mathcal{D}^{(n)} = TM$ and therefore its existence has to be verified by calculating the Engel rank of $\mathcal{D}^{(n-1)}$.

**Remark 3:** The condition (iv) is a sort of regularity condition. If the conditions (i) – (iii) are satisfied everywhere on $M$, then (iv) holds in an open and dense subset $\tilde{M}$ of $M$. In this sense, (i) – (iii) are structural conditions that guarantee equivalence to a Cartan distribution for surfaces at generic points and (iv) distinguishes regular points from singular ones. Given a distribution and a fixed point $q \in M$, the condition (iv) can be checked algebraically because the involutive subdistribution $\mathcal{L}_{n-1}$ is unique (if it exists) and can be calculated; we will discuss that in Section 3.3.2 below.

The proof of the above theorem will be given in Section 3.5. It is based on the Bryant normal form, which we will discuss in Section 3.4.

### 3.3.2 Involutive subdistributions of corank $k$

The condition (iii) of Theorem 3.3.2, saying that the Engel rank of $\mathcal{D}^{(n-1)}$ equals two, is essential for identifying Cartan distributions for surfaces. In this section, we will relate that condition to that expressed via the Cartan rank and to other geometric properties, in particular, to the existence of an involutive subdistribution of corank two.

**Definition 3.3.3** The *Cartan rank* of a Pfaffian system $\mathcal{I} = \operatorname{span}\{\omega^1, \dots, \omega^m\}$ at $q \in M$, denoted by $\operatorname{rank}_C \mathcal{I}(q)$, is the smallest integer $k$ for which there exist $\pi^1, \dots, \pi^k \in \Lambda^1(M)$ such that

$$\pi^1 \wedge \cdots \wedge \pi^k \wedge \omega^1 \wedge \cdots \wedge \omega^m(q) \neq 0,$$

and

$$\mathrm{d}\omega \wedge \pi^1 \wedge \cdots \wedge \pi^k \equiv 0 \bmod \mathcal{I}, \text{ for any } \omega \in \mathcal{I}.$$

The following theorem answers the question: when does a given distribution $\mathcal{D}$ of constant rank $k + mk$, for $m \geq 3$, contain an involutive subdistribution $\mathcal{L} \subset \mathcal{D}$ that has constant corank $k$ in $\mathcal{D}$?

**Theorem 3.3.4** *Assume that $m \geq 3$. Let $\mathcal{D}$ be a distribution of constant rank $k+mk$ defined on a manifold $M$ of dimension $m + k + mk$ such that $\mathcal{D}^{(1)} = TM$. If the characteristic distribution of $\mathcal{D}$ satisfies $\mathcal{C}(\mathcal{D}) = 0$, then the following conditions are equivalent:*

- (i) *The Cartan rank of $\mathcal{D}^{\perp}$ is constant and equals $k$;*
- (ii) *The Engel rank of $\mathcal{D}^{\perp}$ is constant and equals $k$;*
- (iii) *There exists a subdistribution $\mathcal{F}$ of $\mathcal{D}$ of corank $k$ such that $[\mathcal{F}, \mathcal{F}] \subset \mathcal{D}$ and there does not exist any subdistribution of $\mathcal{D}$ with that property of a corank smaller than $k$;*
- (iv) *There exists an involutive subdistribution $\mathcal{L}$ that is of corank $k$ in $\mathcal{D}$ and there does not exist any involutive subdistribution of $\mathcal{D}$ of a corank smaller than $k$.*

*Moreover, if any of the equivalent conditions* (i)-(iv) *is satisfied, then the involutive subdistribution $\mathcal{L}$ of corank $k$ in $\mathcal{D}$ is unique and can be calculated as*

$$\mathcal{L} = \tilde{\mathcal{F}}_1 + \cdots \tilde{\mathcal{F}}_m,$$

*where $\tilde{\mathcal{F}}_i = \operatorname{span}\{f \in \mathcal{D} : f \lrcorner \, d\omega^i \in \mathcal{D}^{\perp}\}$ and $\omega^i$ are any differential 1-forms such that $\mathcal{D}^{\perp} = \operatorname{span}\{\omega^1, \dots, \omega^m\}$; in fact, it is enough to take in the above sum defining $\mathcal{L}$ only two terms corresponding to any $1 \leq i \neq j \leq m$.*

Notice that, in general, $\operatorname{rank}_C \mathcal{I}(q) \neq \operatorname{rank}_E \mathcal{I}(q)$ even if both ranks are constant (see [8] for examples). The equality of both ranks is a consequences of the assumptions $\mathcal{D}^{(1)} = TM$ and $\mathcal{C}(\mathcal{D}) = 0$ and was proved by Bryant [7]. The proof of the above theorem is given in Section 3.5.3.

## 3.4 Extended Bryant normal form

The Bryant normal form is the canonical form of the Cartan distribution $\mathcal{CC}^1(\mathbb{R}^k, \mathbb{R}^m)$ (that is, for 1-jets). Our proof of Theorem 3.3.2 is based on an extension of the Bryant normal form. In this section we will recall the Bryant normal form (and for completeness we will provide its alternative proof), then we will give its extension and, finally, provide a dual version of that extension which we will use directly in the proof of our main result, Theorem 3.3.2.

### 3.4.1 Bryant normal form

The question of which Pfaffian systems are locally equivalent to the canonical contact Pfaffian system on $J^1(\mathbb{R}^k, \mathbb{R}^m)$ (equivalently, the problem to characterize the Cartan distributions $\mathcal{CC}^1(\mathbb{R}^k, \mathbb{R}^m)$) has been studied by Gardner since 1972 and Bryant. In this subsection, we will recall the solution to this problem given by Bryant [7] in 1979.

If for a constant rank Pfaffian system $\mathcal{I}$, the characteristic distribution $\mathcal{C}(\mathcal{D})$ of the distribution $\mathcal{D} = \mathcal{I}^{\perp}$ is of constant rank, then its annihilator $(\mathcal{C}(\mathcal{D}))^{\perp}$ is a Pfaffian system, called the *retracting codistribution* of $\mathcal{I}$ and denoted by $\mathcal{R}(\mathcal{I})$, in other words, $\mathcal{R}(\mathcal{I}) = (\mathcal{C}(\mathcal{D}))^{\perp}$. The following result (see, e.g., [8]) interprets the rank of $\mathcal{R}(\mathcal{I})$ in terms of the minimal number of coordinate variables needed to express $\mathcal{I}$.

**Lemma 3.4.1** *Let $\mathcal{I}$ be a Pfaffian system on $M$ such that the rank of $\mathcal{I}$ and of its retracting codistribution $\mathcal{R}(\mathcal{I})$ are constant. If $\operatorname{rank} \mathcal{R}(\mathcal{I}) = s$, then around any $q \in M$ there are coordinates $(x^1, \ldots, x^r, y_1, \ldots, y_s)$, where $r+s = n$ such that $\mathcal{I}$ is generated by 1-forms that are expressed via the coordinates $y_1, \ldots, y_s$ only. Moreover, $s = \operatorname{rank} \mathcal{R}(\mathcal{I})$ is the minimal integer with that property.*

This Lemma shows that given a Pfaffian system, the rank of its retracting system equals, indeed, the minimal number of variables needed to describe the system (see [8]).

The following characterization of the canonical contact Pfaffian system on $J^1(\mathbb{R}^k, \mathbb{R}^m)$ was proved by Bryant in [7] (see also [8]).

**Theorem 3.4.2 (Bryant normal form)** *Assume that $m \neq 2$. A Pfaffian system*

$$\mathcal{I} = \mathrm{span}\,\{\omega^1, \ldots, \omega^m\}$$

*of constant rank $m$, defined on a manifold of dimension $m + k + mk$ and such that $\mathcal{I}^{(1)} = 0$, is locally equivalent to the canonical contact Pfaffian system on $J^1(\mathbb{R}^k, \mathbb{R}^m)$, called also the Bryant normal form,*

$$\begin{array}{lcl} \mathrm{d}p_1^0 - p_1^1 \mathrm{d}x_1 - & \cdots & -p_1^k \mathrm{d}x_k = 0 \\ & \vdots & \\ \mathrm{d}p_m^0 - p_m^1 \mathrm{d}x_1 - & \cdots & -p_m^k \mathrm{d}x_k = 0 \end{array} \tag{3.4.1}$$

*if and only if its Engel rank is constant and equals $k$, and the rank of its retracting codistribution $\mathcal{R}(\mathcal{I})$ is constant and equals $m + k + mk$.*

Above, the variables $p_j^i$, where $1 \le i \le k$ and $1 \le j \le m$, stand for $p_j^{1_i}$, where $1_i = (0, \ldots, 1, \ldots, 0)$, with the $i$-th component being equal to one.

For completeness, we will give an alternative proof of Theorem 3.4.2 (Bryant normal form for $m \ge 3$) based on Theorem 3.3.4, and on Lemma 3.5.3.

**Proof:** Let $\mathcal{D}$ be the distribution defined by $\mathcal{D} = \mathcal{I}^{\perp}$. Obviously, the condition that the rank of the retracting system $\mathcal{R}(\mathcal{I})$ equals $m + k + mk$, implies that $\mathcal{C}(\mathcal{D}) = 0$. Since $m \ge 3$ and the Engel rank of $\mathcal{D}^{\perp}$ equals $k$, it follows by Theorem 3.3.4 that $\mathcal{D}$ contains an involutive subdistribution $\mathcal{L}$ that has corank $k$ in $\mathcal{D}$. Therefore, there exists a coordinate system $(x_1, \ldots, x_k, p_j^i)$, for $0 \le i \le k$ and $1 \le j \le m$, such that

$$\mathcal{D} = \mathcal{L} \oplus \mathrm{span}\,\{f_1, \ldots, f_k\},$$

where

$$\mathcal{L} = \mathrm{span}\,\{\frac{\partial}{\partial p_j^i}, \quad 1 \le i \le k, \quad 1 \le j \le m\},$$

and

$$f_i = \frac{\partial}{\partial x_i} + a_{i1} \frac{\partial}{\partial p_1^0} + \cdots + a_{im} \frac{\partial}{\partial p_m^0}, \quad \text{for } 1 \le i \le k,$$

for some functions $a_{ij}$. We can assume that all functions $a_{ij}$ satisfy $a_{ij}(q_0) = 0$ replacing, if necessary, $p_j^0$ by $p_j^0 - \sum_{i=1}^k a_{ij}(q_0) x_i$.

The condition $\mathcal{C}(\mathcal{D}) = 0$ implies that $\mathcal{D}$ satisfies the assumptions of Lemma 3.5.3 (see the next section), with $r = km$ and $l = k$. Therefore

$$\mathrm{rank}\, \frac{\partial a}{\partial p}(q) = km,$$

for any $q$ in a neighborhood of $q_0$, where

$$a \;=\; (a_{11}, \ldots, a_{1m}, a_{21}, \ldots, a_{km})^{\top} \;\text{ and}$$

$$p = (p_1^1, \dots, p_m^1, p_1^2, \dots, p_m^k)^\top.$$

It follows that the map defined by

$$\begin{aligned}
\tilde{x}_i &= x_i, \quad 1 \le i \le k, \\
\tilde{p}_j^0 &= p_j^0, \quad 1 \le i \le k, \\
\tilde{p}_j^i &= a_{ij}, \quad 1 \le i \le k, \ 1 \le j \le m,
\end{aligned}$$

is a local diffeomorphism transforming $q_0$ into 0. We have

$$\begin{aligned}
\varphi_* \mathcal{L} &= \operatorname{span}\{\frac{\partial}{\partial \tilde{p}_j^i}, \quad 1 \le i \le k, \quad 1 \le j \le m\} \text{ and} \\
\varphi_* f_i &= \frac{\partial}{\partial \tilde{x}_i} + \sum_{j=1}^{m} \tilde{p}_j^i \frac{\partial}{\partial \tilde{p}_j^0} \bmod \mathcal{L}.
\end{aligned}$$

It follows that (we skip the tildes)

$$\varphi_* \mathcal{D} = \operatorname{span}\{\frac{\partial}{\partial \tilde{x}_i} + \sum_{s=1}^{m} \tilde{p}_s^i \frac{\partial}{\partial \tilde{p}_s^0}, \quad \frac{\partial}{\partial \tilde{p}_j^i}, \quad 1 \le i \le k, \ 1 \le j \le m\}$$

thus proving that $\varphi_* \mathcal{D}$ is the Cartan distribution $\mathcal{CC}^1(\mathbb{R}^k, \mathbb{R}^m)$ expressed in $z = (\tilde{x}, \tilde{p})$-coordinates. $\square$

### 3.4.2 Extended Bryant normal form

We give in this section, a generalization of Bryant's Theorem 3.4.2 and then provide a dual version of that generalization.

**Theorem 3.4.3 (Extended Bryant normal form, first version)** *Assume that $m \neq 2$. Let $\mathcal{I} = \operatorname{span}\{\omega^1, \dots, \omega^m\}$ be a Pfaffian system defined on a manifold $M$ of dimension $m + k + mk + l$ such that $\mathcal{I}^{(1)} = 0$. If $\mathcal{I}$ satisfies, in a neighborhood of a point $q \in M$, the following conditions:*

- (i) *The Engel rank of $\mathcal{I}$ is constant and equals $k$;*
- (ii) *The rank of the retracting codistribution $\mathcal{R}(\mathcal{I})$ is constant and equals $m+k+mk$,*

*then there exist local coordinates $(z, y) = (x, p, y) = (x_i, p_j^i, y_1, \dots, y_l)$, for $0 \le i \le k$, $1 \le j \le m$, around $q$ such that $\mathcal{I}$ takes in $(z, y) = (x, p, y)$-coordinates the Bryant normal form* (3.4.1).

**Proof:** Recall that the characteristic distribution is always involutive. Since rank $\mathcal{R}(\mathcal{I})$ equals $m + k + mk$, there exists, by Lemma 3.4.1, a local coordinate system $(z, y) = (z_1, \dots, z_{m+k+mk}, y_1, \dots, y_l)$, centered at $(0, 0) \in \mathbb{R}^{m(k+1)+k} \times \mathbb{R}^l$ such that

$$\omega^i = \sum_{j=1}^{m+k+mk} \omega_j^i(z) dz_j, \quad 1 \le i \le m.$$

Therefore $\omega^1, \ldots, \omega^m$ can be considered locally as differential 1-forms on $\mathbb{R}^{m+k+mk}$, the latter equipped with $z$-coordinates. Define the codistribution (Pfaffian system) $\tilde{\mathcal{I}}$ spanned by them locally around $0 \in \mathbb{R}^{m+k+mk}$. Assume for a moment that $\tilde{\mathcal{I}}$ satisfies the conditions of Theorem 3.4.2, then there exists a local change of coordinates $z = \psi(\tilde{z})$ such that the codistribution $\psi^*\tilde{\mathcal{I}}$ is locally given, in $\tilde{z} = (x, p)$-coordinates, by the Bryant normal form (3.4.1). Define the local coordinates change

$$\Psi : \qquad \begin{cases} z = \psi(\tilde{z}) \\ y = \tilde{y} \end{cases}$$

and we have immediately that the pull-back $\Psi^*$ takes $\mathcal{I}$ into the canonical form (3.4.1) in $(\tilde{z}, \tilde{y}) = (x, p, y)$-coordinates. So in order to prove Theorem 3.4.3, it remains to verify that $\tilde{\mathcal{I}}$ satisfies the conditions of Theorem 3.4.2, that is, (1): $\tilde{\mathcal{I}}^{(1)} = 0$, (2): $\operatorname{rank}_E \tilde{\mathcal{I}} = k$ and (3): $\operatorname{rank} \mathcal{R}(\tilde{\mathcal{I}}) = m + k + mk$.

**(1)**. For any $\alpha \in \tilde{\mathcal{I}}$, we have

$$\alpha = \sum_{i=1}^{m} \alpha_i(z)\omega^i \in \mathcal{I},$$

which gives (since $\mathcal{I}^{(1)} = 0$)

$$\mathrm{d}\alpha \wedge \omega^1 \wedge \cdots \omega^m = 0,$$

thus implying that $\tilde{\mathcal{I}}^{(1)} = 0$.

**(2)**. The condition $\operatorname{rank}_E \tilde{\mathcal{I}} = k$ holds if and only if

(1) $(\mathrm{d}\alpha)^{k+1} \wedge \omega^1 \wedge \cdots \wedge \omega^m = 0$, for any $\alpha \in \tilde{\mathcal{I}}$;
(2) There exists $\alpha \in \tilde{\mathcal{I}}$ such that $(\mathrm{d}\alpha)^k \wedge \omega^1 \wedge \cdots \wedge \omega^m \neq 0$.

On one hand, since $\operatorname{rank}_E \mathcal{I} = k$, for any $\Omega = \sum_{i=1}^m \eta_i(z, y)\omega^i \in \mathcal{I}$, we have

$$(\mathrm{d}\,\Omega)^k \wedge \omega^1 \wedge \cdots \omega^m = 0,$$

which implies that for any $\alpha = \sum_{i=1}^m \alpha_i(z)\omega^i \in \tilde{\mathcal{I}}$, we have $(\mathrm{d}\alpha)^{k+1} \wedge \omega^1 \wedge \cdots \omega^m = 0$. On the other hand, the condition $\operatorname{rank}_E \mathcal{I} = k$ implies that there exists a 1-form $\Omega = \sum_{i=1}^m \eta_i(z, y)\omega^i \in \mathcal{I}$ such that

$$(\mathrm{d}\,\Omega)^k \wedge \omega^1 \wedge \cdots \omega^m(0, 0) \neq 0$$

(recall that the coordinates are centered at zero). Define a differential 1-form $\tilde{\Omega} \in \tilde{\mathcal{I}}$ by

$$\tilde{\Omega} = \sum_{i=1}^{m} c_i\omega^i$$

where the constant coefficients $c_i$ are given by $c_i = \eta(0,0)$. We have $\mathrm{d}\tilde{\Omega} = \sum_{i=1}^m c_i \mathrm{d}\omega^i$ and thus

$$(\mathrm{d}\,\tilde{\Omega})^k \wedge \omega^1 \wedge \cdots \omega^m(0) = (\mathrm{d}\,\Omega)^k \wedge \omega^1 \wedge \cdots \omega^m(0,0) \neq 0$$

which implies that $\mathrm{rank}_E \tilde{\mathcal{I}} = k$.

**(3)**. If in a neighborhood of $0 \in \mathbb{R}^{m(k+1)+k}$ there exists a point $\tilde{q}$ such that $\mathrm{rank}\,\mathcal{R}(\tilde{\mathcal{I}})(\tilde{q}) = m + k + mk - s$ where $s > 0$, it follows that $\mathrm{rank}\,\mathcal{C}(\tilde{\mathcal{I}}^\perp)(\tilde{q}) = s > 0$ implying that there exists $0 \neq \tilde{v} \in \tilde{\mathcal{I}}^\perp(\tilde{q})$ such that $\tilde{v} \lrcorner \mathrm{d}\omega(\tilde{q}) \in \tilde{\mathcal{I}}(\tilde{q})$ for any $\omega \in \tilde{\mathcal{I}}$. It follows that $v = (\tilde{v}, 0)$, where 0 stands for the $y$-components of $v$, satisfies $v \in \mathcal{I}^\perp(q)$ and $v \in \mathcal{C}(\mathcal{I}^\perp)(q)$ at any $q = (\tilde{q}, \bar{q})$ for any $\bar{q}$ corresponding to the $y$-components, in particular at $q = (\tilde{q}, 0)$. This contradicts the assumption $\mathrm{rank}\,\mathcal{C}(\mathcal{I}^\perp)(q) = 0$ in a neighborhood of $q_0 = 0$. Therefore we have $\mathrm{rank}\,\mathcal{R}(\tilde{\mathcal{I}}) = m + k + mk$. □

In the dual language of vector fields and distributions, in particular, using $\mathcal{D}$ instead of $\mathcal{I}$, as well as $\mathcal{C}(\mathcal{D})$ instead of $\mathcal{R}(\mathcal{I})$, and calculating vector fields annihilated by the differential forms of (3.4.1), we obtain the following result:

**Theorem 3.4.4 (Extended Byrant normal form, dual version)** *Assume that $m \neq 2$. Let $\mathcal{D}$ be a distribution of constant rank $k + mk + l$ defined on a manifold $M$ of dimension $m + k + mk + l$ such that $\mathcal{D}^{(1)} = TM$. If $\mathcal{D}$ satisfies, in a neighborhood of a point $q \in M$, the following conditions:*

(i) *the Engel rank of $\mathcal{D}^\perp$ is constant and equals $k$,*
(ii) *the rank of the characteristic distribution $\mathcal{C}(\mathcal{D})$ of $\mathcal{D}$ is constant and equals $l$,*

*then there exist local coordinates $(z, y) = (x, p, y) = (x_1, \ldots, x_k, p_1^0, \ldots, p_1^k, p_2^0, \ldots, p_2^k, \ldots, p_m^0, \ldots, p_m^k, y_1, \ldots, y_l)$ around $q$ such that*

$$\mathcal{D} = \tilde{\mathcal{D}} \oplus \mathcal{C},$$

*where $\tilde{\mathcal{D}}$ is the Cartan distribution $\mathcal{CC}^1(\mathbb{R}^k, \mathbb{R}^m)$ expressed in $z = (x,p)$-coordinates, that is,*

$$\tilde{D} = \mathrm{span}\,\{\frac{\partial}{\partial x_i} + \sum_{s=1}^m p_s^i \frac{\partial}{\partial p_s^0}, \quad \frac{\partial}{\partial p_j^i}, \quad 1 \leq i \leq k, \ 1 \leq j \leq m\}$$

*and*

$$\mathcal{C} = \mathcal{C}(\mathcal{D}) = \mathrm{span}\,\{\frac{\partial}{\partial y}\} = \mathrm{span}\,\{\frac{\partial}{\partial y_1}, \ldots, \frac{\partial}{\partial y_l}\}$$

*is the characteristic distribution of $\mathcal{D}$.*

## 3.5 Proofs

### 3.5.1 Useful results

In this section we give a series of results that we will use in the subsequent sections when proving our theorems.

**Lemma 3.5.1** *Let $\mathcal{D}$ be a constant rank distribution on a manifold $M$. Assume that both $\mathcal{D}$ and its derived distribution $\mathcal{D}^{(1)}$ are of constant rank and let $\mathcal{C}_0$ and $\mathcal{C}_1$ denote, respectively, their characteristic distributions. Then $\mathcal{C}_0 \subset \mathcal{C}_1$.*

**Proof:** In order to prove this lemma, it is sufficient to show that for any $g \in \mathcal{C}_0$, we have also $g \in \mathcal{C}_1$. Assume that $f$ is an arbitrary vector field of $\mathcal{D}^{(1)}$. Since

$$f \in \mathcal{D}^{(1)} = \mathcal{D} + [\mathcal{D}, \mathcal{D}],$$

we have either $f \in \mathcal{D}$ or $f \in [\mathcal{D}, \mathcal{D}]$. If $f \in \mathcal{D}$, then $[g, f] \in \mathcal{D} \subset \mathcal{D}^{(1)}$ because $g \in C_0$. If $f \in [\mathcal{D}, \mathcal{D}]$, then there exist $f_1, f_2 \in \mathcal{D}$ such that $f = [f_1, f_2]$. Therefore

$$[g, f] = [g, [f_1, f_2]] = [[g, f_1], f_2] + [f_1, [g, f_2]].$$

The condition $g \in C_0$ implies that $[g, f_1] \in \mathcal{D}$ and $[g, f_2] \in \mathcal{D}$. Thus we have

$$[[g, f_1], f_2] \in [\mathcal{D}, \mathcal{D}] \subset \mathcal{D}^{(1)}$$

$$[f_1, [g, f_2]] \in [\mathcal{D}, \mathcal{D}] \subset \mathcal{D}^{(1)},$$

which implies that $[g, f] \in \mathcal{D}^{(1)}$. $\square$

**Lemma 3.5.2** *Let $\mathcal{E}$ and $\mathcal{F}$ be two involutive distributions on a manifold $M$. If they satisfy the following conditions*

(i) *$\mathcal{F}$ has constant corank $r_2$;*
(ii) *$\mathcal{E}$ has constant corank $r_1 + r_2$;*
(iii) *$\mathcal{E} \subset \mathcal{F}$,*

*then there exist local coordinates $(x_1, \ldots, x_{n-r_1-r_2}, p_1^1, \ldots, p_1^{r_1}, p_2^1, \ldots, p_2^{r_2})$, in a neighborhood of any point in $M$, such that*

$$\begin{aligned} \mathcal{F}^{\perp} &= \{\mathrm{d}p_2^1, \ldots, \mathrm{d}p_2^{r_2}\} \\ \mathcal{E}^{\perp} &= \{\mathrm{d}p_1^1, \ldots, \mathrm{d}p_1^{r_1}, \mathrm{d}p_2^1, \ldots, \mathrm{d}p_2^{r_2}\}. \end{aligned}$$

The proof of this lemma (simultaneous rectification of involutive distributions) can be found, e.g., in [29]. For completeness we give below another proof.

**Proof:** Since $\mathcal{F}$ is involutive and has constant rank $n - r_2$, there exist, by Frobenius theorem, local coordinates $(x_1, \dots, x_{n-r_1-r_2}, p_1^1, \dots, p_1^{r_1}, p_2^1, \dots, p_2^{r_2})$, in a neighborhood $\mathcal{O}$ of $q \in M$, such that

$$\mathcal{F}^{\perp} = \mathrm{span}\,\{\mathrm{d}p_2^1, \dots, \mathrm{d}p_2^{r_2}\}.$$

Moreover, the condition that $\mathcal{E}$ is involutive and has constant corank $r_1 + r_2$ implies that there exist $r_1 + r_2$ smooth functions $\varphi_1, \dots, \varphi_{r_1+r_2}$ in $\mathcal{O}$ such that

$$\mathcal{E}^{\perp} = \mathrm{span}\,\{\mathrm{d}\varphi_1, \dots, \mathrm{d}\varphi_{r_1+r_2}\},$$

The relation $\mathcal{E} \subset \mathcal{F}$ yields immediately that $\mathcal{F}^{\perp} \subset \mathcal{E}^{\perp}$, i.e.,

$$\mathrm{span}\,\{\mathrm{d}p_2^1, \dots, \mathrm{d}p_2^{r_2}\} \subset \mathrm{span}\,\{\mathrm{d}\varphi_1, \dots, \mathrm{d}\varphi_{r_1+r_2}\}.$$

which implies that there exist $r_1$ functions among the $\varphi_i$'s (say $\varphi_1, \dots, \varphi_{r_1}$, after a reordering) such that

$$\mathrm{span}\,\{\mathrm{d}\varphi_1, \dots, \mathrm{d}\varphi_{r_1}, \mathrm{d}p_2^1, \dots, \mathrm{d}p_2^{r_2}\} = \mathrm{span}\,\{\mathrm{d}\varphi_1, \dots, \mathrm{d}\varphi_{r_1+r_2}\}.$$

Introduce new coordinates by replacing $p_1^i$ by $\varphi_i$, for $1 \le i \le r_1$, and in new coordinates we have clearly

$$\begin{aligned} \mathcal{F}^{\perp} &= \{\mathrm{d}p_2^1, \dots, \mathrm{d}p_2^{r_2}\} \\ \mathcal{E}^{\perp} &= \{\mathrm{d}p_1^1, \dots, \mathrm{d}p_1^{r_1}, \mathrm{d}p_2^1, \dots, \mathrm{d}p_2^{r_2}\}. \end{aligned}$$

□

Recall that we have defined the characteristic distribution $\mathcal{C}(\mathcal{E})$ of a distribution $\mathcal{E}$ pointwise as

$$\mathcal{C}(\mathcal{E})(q) = \{v \in \mathcal{E}(q) : v \lrcorner \, \mathrm{d}\omega(q) \in \mathcal{I}(q), \quad \forall\, \omega \in \mathcal{I}\},$$

where $\mathcal{I} = \mathcal{E}^{\perp}$.

**Lemma 3.5.3** *Let $\mathcal{E}$ be a distribution of rank $r + l$ defined on a smooth manifold $M$ of dimension $r + s$, for $l \le s$. Suppose that at any point $q \in M$, the characteristic distribution $\mathcal{C}(\mathcal{E})$ of $\mathcal{E}$ satisfies $\mathcal{C}(\mathcal{E})(q) = 0$. Assume that in a local coordinate system $(p_1, \dots, p_r, y_1, \dots, y_s)$ in a neighborhood $\mathcal{O}(q_0)$ of a point $q_0 \in M$, the distribution $\mathcal{E}$ is spanned by*

$$\mathcal{E} = \mathrm{span}\,\{\frac{\partial}{\partial p_1}, \dots, \frac{\partial}{\partial p_r}, f_1, \dots, f_l\},$$

*where*

$$\begin{aligned} f_1 &= f_1^1 \frac{\partial}{\partial y_1} + f_1^2 \frac{\partial}{\partial y_2} + \cdots + f_1^s \frac{\partial}{\partial y_s} \\ &\vdots \\ f_l &= f_l^1 \frac{\partial}{\partial y_1} + f_l^2 \frac{\partial}{\partial y_2} + \cdots + f_l^s \frac{\partial}{\partial y_s} \end{aligned} \tag{3.5.1}$$

*for $1 \leq i \leq m$ and $1 \leq j \leq s$. Then*

$$\operatorname{rank}\left(\frac{\partial f}{\partial p}\right)(q) = r,$$

*for any $q \in \mathcal{O}(q_0)$, where $p = (p_1, \ldots, p_r)$ and the map $f : \mathcal{O}(q_0) \to \mathbb{R}^{ls}$ is formed by*

$$f = (f_1^1, \ldots, f_1^s, f_2^1, \ldots, f_l^1, \ldots, f_l^s)^\top.$$

**Proof:** Assume that at a point $q \in \mathcal{O}(q_0)$, we have $\operatorname{rank}\left(\frac{\partial f}{\partial p}\right)(q) = r' < r$ and, without loss of generality, that the first $r'$ columns of $\left(\frac{\partial f}{\partial p}\right)(q)$ are independent. For any $r' + 1 \leq i \leq r$, we have

$$\frac{\partial f}{\partial p_i}(q) = \sum_{j=1}^{r'} c_i^j \frac{\partial f}{\partial p_j}(q),$$

for some constants $c_i^j \in \mathbb{R}$. Define the vector fields

$$g_i = \frac{\partial}{\partial p_i} - \sum_{j=1}^{r'} c_i^j \frac{\partial}{\partial p_j}, \quad r' + 1 \leq i \leq r.$$

We have, $[g_i,\ f_t](q) = 0$, for $r'+1 \leq i \leq r$ and any $f_t$, $1 \leq t \leq l$, which implies that for any $\omega_j \in \mathcal{I} = \mathcal{E}^\perp$, we have $(g_i \lrcorner \, \mathrm{d}\omega_j)(q) = 0$ and hence $g_i(q) \in \mathcal{C}(\mathcal{E})(q)$, for $r'+1 \leq i \leq r$, thus contradicting the assumption $\mathcal{C}(\mathcal{E})(q) = 0$. Therefore, $\operatorname{rank}\left(\frac{\partial f}{\partial p}\right)(q) = r$ for any $q \in \mathcal{O}(q_0)$. $\square$

### 3.5.2 Proof of Theorem 3.3.2

**Proof:** For simplicity, we will write $p^{\sigma_1,\sigma_2}$ instead of $p^{(\sigma_1,\sigma_2)}$ and $p^0$ instead of $p^{(0,0)}$. If $\mathcal{D}$ is equivalent to the Cartan distribution for surfaces $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$, then by a direct calculation, we can easily check that the conditions (i) – (iv) are satisfied.

Consider now a distribution that satisfies, at a point $q \in M$, the conditions (i) – (iv). To start with, observe that the distribution $\mathcal{D}^{(n-1)}$ (of rank $\frac{m}{2}(n+1)(n+2) - m + 2$, by (ii)) satisfies the conditions of Theorem (3.4.4) with $k = 2$ and $l = \frac{m}{2}(n+1)(n+2) - 3m$. Indeed, (iii) implies that the Engel rank of $(\mathcal{D}^{(n-1)})^\perp$ equals two and by (ii) its characteristic distribution $\mathcal{C}_{n-1}$ is of constant rank given by

$$\operatorname{rank} \mathcal{C}_{n-1} = \operatorname{rank} \mathcal{D}^{(n-2)} - 2 = \frac{m}{2}(n+1)(n+2) - 3m = c_{n-1},$$

(we used $\operatorname{rank} \mathcal{D}^{(n-2)} = \frac{m}{2}(n+1)(n+2) - 3m + 2$ following from (i)). Therefore, by Theorem 3.4.4, $\mathcal{D}^{(n-1)}$ is the direct sum of its characteristic distribution $\mathcal{C}_{n-1}$ and a distribution locally equivalent to the Cartan distribution $\mathcal{CC}^1(\mathbb{R}^2,\mathbb{R}^m)$. It follows that there exists a local coordinate system, in a neighborhood $\mathcal{O}(q)$ of $q$, denoted by $(x_1,x_2,p_j^\sigma)$, for $1\leq j\leq m$, $0\leq|\sigma|\leq n$, such that

$$\mathcal{D}^{(n-1)}=\operatorname{span}\Big\{\frac{\partial}{\partial p_j^\sigma}\Big|_{2\leq|\sigma|\leq n},\frac{\partial}{\partial p_j^{1,0}},\frac{\partial}{\partial p_j^{0,1}},f^{n-1},g^{n-1},\quad 1\leq j\leq m\Big\},$$

where the vector fields $f^{n-1}$ and $g^{n-1}$ are given by

$$f^{n-1}=\frac{\partial}{\partial x_1}+\sum_{j=1}^m p_j^{1,0}\frac{\partial}{\partial p_j^0}$$
$$g^{n-1}=\frac{\partial}{\partial x_2}+\sum_{j=1}^m p_j^{0,1}\frac{\partial}{\partial p_j^0}.$$

Moreover, the characteristic distribution $\mathcal{C}_{n-1}$ is given by

$$\mathcal{C}_{n-1}=\operatorname{span}\Big\{\frac{\partial}{\partial p_j^\sigma},\quad 1\leq j\leq m,\quad 2\leq|\sigma|\leq n\Big\}.$$

Observe that the coordinates $y_i$, $1\leq i\leq l=c_{n-1}$ of Theorem 3.4.4 correspond to $p_j^\sigma$, for $1\leq j\leq m$, $2\leq|\sigma|\leq n$ (notice that the number of $p_j^\sigma$'s is, indeed, $c_{n-1}$).

Clearly, the involutive subdistribution $\mathcal{L}_{n-1}$ of corank two in $\mathcal{D}^{n-1}$ is given by

$$\mathcal{L}_{n-1}=\operatorname{span}\Big\{\frac{\partial}{\partial p_j^\sigma},\quad 1\leq j\leq m,\quad 1\leq|\sigma|\leq n\Big\}.$$

For a notational convenience we put $\mathcal{C}_n=\mathcal{L}_{n-1}$. We have

$$\mathcal{C}_1\subset\cdots\subset\mathcal{C}_{n-1}\subset\mathcal{C}_n=\mathcal{L}_{n-1},$$

(see Lemma 3.5.1) and thus the coordinates $p_j^\sigma$ can be chosen, due to a successive application of Lemma 3.5.2, such that we have, for $1\leq j\leq m$,

$$\mathcal{C}_i=\operatorname{span}\Big\{\frac{\partial}{\partial p_j^\sigma},\quad 1\leq j\leq m,\quad n-i+1\leq|\sigma|\leq n\Big\}.$$

Now the proof follows from Lemma 3.5.4 given below. Notice that the hypothesis of Lemma 3.5.4 holds for $k=1$. Therefore a successive applying the lemma $(n-1)$-times (for $k$ from 1 up to $n-1$) yields local coordinates $(x_1,x_2,p_j^\sigma)$, for $1\leq j\leq m$, $0\leq|\sigma|\leq n$, in which

$$\mathcal{D}=\mathcal{D}^{(0)}=\mathcal{C}_1\oplus\operatorname{span}\{f^0,g^0\},$$

where

$$\mathcal{C}_1=\operatorname{span}\{\frac{\partial}{\partial p_j^\sigma},\quad 1\leq j\leq m,\quad |\sigma|=n\},$$

and

$$f^0 = \frac{\partial}{\partial x_1} + \sum_{|\sigma|=0}^{n-1} \sum_{j=1}^{m} p_j^{\sigma+1_1} \frac{\partial}{\partial p_j^\sigma}$$

$$g^0 = \frac{\partial}{\partial x_2} + \sum_{|\sigma|=0}^{n-1} \sum_{j=1}^{m} p_j^{\sigma+1_2} \frac{\partial}{\partial p_j^\sigma},$$

which shows that $\mathcal{D}$ is locally equivalent to the Cartan distribution for surfaces $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$. $\square$

**Lemma 3.5.4** *Let a distribution $\mathcal{D}$ satisfy the conditions* (i)-(iv) *of Theorem 3.3.2. Assume that for a certain $1 \leq k \leq n$, there exist coordinates $(x_1, x_2, p_j^\sigma)$, for $1 \leq j \leq m$, $0 \leq |\sigma| \leq n$, centered at $0 \in \mathbb{R}^N$, where $N = \frac{m}{2}(n+1)(n+2)+2$, such that*

$$\begin{aligned} \mathcal{L}_{n-1} &= \operatorname{span}\left\{\frac{\partial}{\partial p_j^\sigma}, \quad 1 \leq j \leq m, \quad 1 \leq |\sigma| \leq n\right\} \\ \mathcal{C}_i &= \operatorname{span}\left\{\frac{\partial}{\partial p_j^\sigma}, \quad 1 \leq j \leq m, \quad n-i+1 \leq |\sigma| \leq n\right\} \end{aligned}$$

*for $1 \leq i \leq n-1$ and $\mathcal{D}^{(n-k)} = \mathcal{C}_{n-k+1} \oplus \operatorname{span}\{f^{n-k}, g^{n-k}\}$ where*

$$f^{n-k} = \frac{\partial}{\partial x_1} + \sum_{|\sigma|=0}^{k-1} \sum_{j=1}^{m} p_j^{\sigma+1_1} \frac{\partial}{\partial p_j^\sigma}$$

$$g^{n-k} = \frac{\partial}{\partial x_2} + \sum_{|\sigma|=0}^{k-1} \sum_{j=1}^{m} p_j^{\sigma+1_2} \frac{\partial}{\partial p_j^\sigma}.$$

*(Recall that for a notational convenience we put, for $k=1$, $\mathcal{C}_{n-k+1} = \mathcal{C}_n = \mathcal{L}_{n-1}$). Then we can modify the coordinates in such a way that in new coordinates $(\tilde{x}_1, \tilde{x}_2, \tilde{p}_j^\sigma)$, for $1 \leq j \leq m$, $0 \leq |\sigma| \leq n$, centered at $0 \in \mathbb{R}^N$, we have*

$$\begin{aligned} \mathcal{L}_{n-1} &= \operatorname{span}\left\{\frac{\partial}{\partial \tilde{p}_j^\sigma}, \quad 1 \leq j \leq m, \quad 1 \leq |\sigma| \leq n\right\} \\ \mathcal{C}_i &= \operatorname{span}\left\{\frac{\partial}{\partial \tilde{p}_j^\sigma}, \quad 1 \leq j \leq m, \quad n-i+1 \leq |\sigma| \leq n\right\} \end{aligned}$$

*for $1 \leq i \leq n-1$ and $\mathcal{D}^{(n-k-1)} = \mathcal{C}_{n-k} \oplus \operatorname{span}\{f^{n-k-1}, g^{n-k-1}\}$ where*

$$f^{n-k-1} = \frac{\partial}{\partial \tilde{x}_1} + \sum_{|\sigma|=0}^{k} \sum_{j=1}^{m} \tilde{p}_j^{\sigma+1_1} \frac{\partial}{\partial \tilde{p}_j^\sigma}$$

$$g^{n-k-1} = \frac{\partial}{\partial \tilde{x}_2} + \sum_{|\sigma|=0}^{k} \sum_{j=1}^{m} \tilde{p}_j^{\sigma+1_2} \frac{\partial}{\partial \tilde{p}_j^\sigma}.$$

**Proof:** (**of Lemma 3.5.4**) To explain the idea of the proof, we will show the lemma for $k = 1$ and then, following similar arguments, for an arbitrary $1 \le k \le n$.

Suppose that the assumptions of the lemma hold for $k = 1$. The condition (ii) of Theorem 3.3.2 implies that $\mathrm{corank}\,(\mathcal{C}_{n-1} \subset \mathcal{D}^{(n-2)}) = 2$ and thus there exist two vector fields $f^{n-2}$ and $g^{n-2}$ such that locally

$$\begin{aligned}\mathcal{D}^{(n-2)} &= \mathcal{C}_{n-1} \oplus \mathrm{span}\,\{f^{n-2},\, g^{n-2}\} \\ &= \mathcal{C}_{n-2} \oplus \mathrm{span}\,\Big\{\frac{\partial}{\partial p_j^{\sigma}},\;\; 1 \le j \le m,\;\; |\sigma| = 2\Big\} \oplus \mathrm{span}\,\{f^{n-2},\, g^{n-2}\}.\end{aligned}$$

We have

$$\mathcal{D}^{(n-1)} = \mathcal{C}_{n-1} \oplus \mathrm{span}\,\Big\{\frac{\partial}{\partial p_j^{\sigma}},\;\; 1 \le j \le m,\;\; |\sigma| = 1\Big\} \oplus \mathrm{span}\,\{f^{n-1}, g^{n-1}\}, \tag{3.5.2}$$

then the relation $f^{n-2}, g^{n-2} \in \mathcal{D}^{(n-2)} \subset \mathcal{D}^{(n-1)}$ implies immediately that

$$\begin{aligned} f^{n-2} &= s_1 f^{n-1} + s_2 g^{n-1} + \sum_{j=1}^{m} (A_j^1 \frac{\partial}{\partial p_j^{1,0}} + A_j^2 \frac{\partial}{\partial p_j^{0,1}}) \mod \mathcal{C}_{n-1} \\ g^{n-2} &= t_1 f^{n-1} + t_2 g^{n-1} + \sum_{j=1}^{m} (B_j^1 \frac{\partial}{\partial p_j^{1,0}} + B_j^2 \frac{\partial}{\partial p_j^{0,1}}) \mod \mathcal{C}_{n-1} \end{aligned}$$

where $s_1, s_2, t_1, t_2, A_j^i, B_j^i$, for $1 \le i \le 2$, $1 \le j \le m$, are smooth functions defined in a neighborhood $\mathcal{O}$ of $0 \in \mathbb{R}^N$. The condition (iv) of Theorem 3.3.2, applied for $i = n-2$, implies that $f^{n-2}$ and $g^{n-2}$ are independent at $0 \in \mathbb{R}^N$, modulo $\mathcal{L}_{n-1}$. Hence $(s_1 t_2 - s_2 t_1)(0) \ne 0$ and thus by a suitable choice of $f^{n-2}$ and $g^{n-2}$ (and by renaming $A_j^i$ and $B_j^i$ as well as taking a smaller $\mathcal{O}$, if necessary), we can assume that

$$\begin{aligned} f^{n-2} &= f^{n-1} + \sum_{j=1}^{m} (A_j^1 \frac{\partial}{\partial p_j^{1,0}} + A_j^2 \frac{\partial}{\partial p_j^{0,1}}) \\ g^{n-2} &= g^{n-1} + \sum_{j=1}^{m} (B_j^1 \frac{\partial}{\partial p_j^{1,0}} + B_j^2 \frac{\partial}{\partial p_j^{0,1}}). \end{aligned}$$

We will now analyze the functions $A_j^i$ and $B_j^i$, for $1 \le i \le 2$ and $1 \le j \le m$. First, we replace $p_j^{1,0}$ by $p_j^{1,0} - A_j^1(0)x_1 - B_j^1(0)x_2$ and $p_j^{0,1}$ by $p_j^{0,1} - A_j^2(0)x_1 - B_j^2(0)x_2$ in order to obtain $A_j^1(0) = A_j^2(0) = B_j^1(0) = B_j^2(0) = 0$. This change of coordinates preserves $\mathcal{L}_{n-1}$ and $\mathcal{C}_i$, for $1 \le i \le n-1$. Secondly, the characteristic distribution of $\mathcal{D}^{(n-2)}$ is $\mathcal{C}_{n-2} = \mathrm{span}\,\big\{\frac{\partial}{\partial p_j^{\sigma}},\;\; 1 \le j \le m,\;\; 3 \le |\sigma| \le n\big\}$. It follows that, for any $1 \le j \le m$ and $3 \le |\sigma| \le n$,

$$\Big[\frac{\partial}{\partial p_j^{\sigma}}, f^{n-2}\Big] = \Big[\frac{\partial}{\partial p_j^{\sigma}}, g^{n-2}\Big] = 0$$

and hence

$$A_j^i = A_j^i(x_1, x_2, p^{\sigma}), \quad B_j^i = B_j^i(x_1, x_2, p^{\sigma}), \quad 0 \le |\sigma| \le 2, \tag{3.5.3}$$

for $i = 1, 2$ and $1 \leq j \leq m$, where $p^\sigma = (p_1^\sigma, \dots, p_m^\sigma)$. Thirdly, a direct calculation gives

$$[f^{n-2}, g^{n-2}] = (A_1^2 - B_1^1)\frac{\partial}{\partial p_1^0} + (A_2^2 - B_2^1)\frac{\partial}{\partial p_2^0} + \cdots + (A_m^2 - B_m^1)\frac{\partial}{\partial p_m^0} \mod \mathcal{L}_{n-1}.$$

Since $[f^{n-2}, g^{n-2}] \in \mathcal{D}^{(n-1)}$, it is easy to conclude that $B_j^1 = A_j^2$, for $1 \leq j \leq m$. Now consider the distribution

$$\begin{aligned} \mathcal{E}^{n-2} \;&=\; \operatorname{span}\Big\{\frac{\partial}{\partial x_1} + \sum_{j=1}^{m}\Big(p_j^{1,0}\frac{\partial}{\partial p_j^0} + A_j^1\frac{\partial}{\partial p_j^{1,0}} + A_j^2\frac{\partial}{\partial p_j^{0,1}}\Big), \\ &\qquad \frac{\partial}{\partial x_2} + \sum_{j=1}^{m}\Big(p_j^{0,1}\frac{\partial}{\partial p_j^0} + B_j^1\frac{\partial}{\partial p_j^{1,0}} + B_j^2\frac{\partial}{\partial p_j^{0,1}}\Big),\quad \frac{\partial}{\partial p_j^\sigma},\quad 1 \leq j \leq m,\quad |\sigma| = 2\Big\} \end{aligned}$$

around $0 \in \mathbb{R}^{6m+2}$, equipped with coordinates $(x_1, x_2, p_j^\sigma)$, for $1 \leq j \leq m$ and $0 \leq |\sigma| \leq 2$. It is well defined (because of (3.5.3)) and it satisfies the assumptions of Lemma 3.5.3 with $r = 3m$, $l = 2$ and $s = 3m + 2$. Indeed $\mathcal{C}(\mathcal{E}^{n-2}) = 0$ because $\mathcal{C}(\mathcal{D}^{(n-2)}) = \mathcal{C}_{n-2}$ and, moreover, the coordinates $(p_i)$ for $1 \leq i \leq r$, correspond to $(p_j^\sigma)$, for $1 \leq j \leq m$ and $|\sigma| = 2$, and the coordinates $y_i$, $1 \leq i \leq s$, to $(x_1, x_2, p_j^\sigma)$, $1 \leq j \leq m$ and $0 \leq |\sigma| \leq 1$. It follows by Lemma 3.5.3 that the map given by

$$\begin{array}{lcll} \tilde{x}_i &=& x_i, & i = 1, 2 \\ \tilde{p}_j^\sigma &=& p_j^\sigma, & 1 \leq j \leq m,\quad |\sigma| \neq 2 \\ \tilde{p}_j^{2,0} &=& A_j^1, & 1 \leq j \leq m \\ \tilde{p}_j^{1,1} &=& A_j^2 = B_j^1, & 1 \leq j \leq m \\ \tilde{p}_j^{0,2} &=& B_j^2, & 1 \leq j \leq m \end{array}$$

is a local diffeomorphism, denoted by $\varphi$, in a neighborhood of $0 \in \mathbb{R}^{\frac{m}{2}(n+1)(n+2)+2}$. It is easy to see that $\varphi$ preserves the nested sequence of distributions

$$\mathcal{C}_1 \subset \cdots \subset \mathcal{C}_{n-1} \subset \mathcal{C}_n = \mathcal{L}_{n-1},$$

that is,

$$\varphi_*\mathcal{C}_i = \operatorname{span}\Big\{\frac{\partial}{\partial \tilde{p}_j^\sigma},\quad 1 \leq j \leq m,\quad n - i + 1 \leq |\sigma| \leq n\Big\},$$

for $1 \leq i \leq n$, and that the vector fields $f^{n-2}$ and $g^{n-2}$ become, respectively,

$$\varphi_* f^{n-2} = \frac{\partial}{\partial \tilde{x}_1} + \sum_{j=1}^{m}\Big(\tilde{p}_j^{1,0}\frac{\partial}{\partial \tilde{p}_j^0} + \tilde{p}_j^{2,0}\frac{\partial}{\partial \tilde{p}_j^{1,0}} + \tilde{p}_j^{1,1}\frac{\partial}{\partial \tilde{p}_j^{0,1}}\Big)$$

$$\varphi_* g^{n-2} = \frac{\partial}{\partial \tilde{x}_2} + \sum_{j=1}^{m}\Big(\tilde{p}_j^{0,1}\frac{\partial}{\partial \tilde{p}_j^0} + \tilde{p}_j^{1,1}\frac{\partial}{\partial \tilde{p}_j^{1,0}} + \tilde{p}_j^{0,2}\frac{\partial}{\partial \tilde{p}_j^{0,1}}\Big).$$

To prove the general case, which follows the same line, assume that in a coordinate system $(x_1, x_2, p_j^\sigma)$, $1 \le j \le m$, $0 \le |\sigma| \le n$, centered at $0 \in \mathbb{R}^N$, we have

$$\mathcal{D}^{(n-k)} = \mathcal{C}_{n-k+1} \oplus \mathrm{span}\,\{f^{n-k}, g^{n-k}\}$$

and

$$\mathcal{C}_i = \mathrm{span}\,\Big\{\frac{\partial}{\partial p_j^\sigma}, \quad 1 \le j \le m, \quad n-i+1 \le |\sigma| \le n\Big\},$$

for $1 \le i \le n$ (recall that we denoted $\mathcal{L}_{n-1} = \mathcal{C}_n$), where

$$f^{n-k} = \frac{\partial}{\partial x_1} + \sum_{|\sigma|=0}^{k-1}\sum_{j=1}^{m} p_j^{\sigma+1_1}\frac{\partial}{\partial p_j^\sigma}$$

$$g^{n-k} = \frac{\partial}{\partial x_2} + \sum_{|\sigma|=0}^{k-1}\sum_{j=1}^{m} p_j^{\sigma+1_2}\frac{\partial}{\partial p_j^\sigma}.$$

By the condition (ii), we have $\mathrm{corank}\,(\mathcal{C}_{n-k} \subset \mathcal{D}^{(n-k-1)}) = 2$ which implies that there exist two vector fields $f^{n-k-1}, g^{n-k-1}$ such that

$$\begin{aligned}
\mathcal{D}^{(n-k-1)} &= \mathcal{C}_{n-k} \oplus \mathrm{span}\,\{f^{n-k-1}, g^{n-k-1}\}\\
&= \mathcal{C}_{n-k-1} \oplus \mathrm{span}\,\{\frac{\partial}{\partial p^\sigma}\Big|_{|\sigma|=k+1}\} \oplus \mathrm{span}\,\{f^{n-k-1}, g^{n-k-1}\}.
\end{aligned}$$

Since $f^{n-k-1}$, $g^{n-k-1} \in \mathcal{D}^{(n-k-1)} \subset \mathcal{D}^{(n-k)}$, they can be expressed as

$$f^{n-k-1} = s_1 f^{n-k} + s_2 g^{n-k} + \sum_{j=1}^{m}(A_j^1\frac{\partial}{\partial p_j^{k,0}} + A_j^2\frac{\partial}{\partial p_j^{k-1,1}} + \cdots + A_j^{k+1}\frac{\partial}{\partial p_j^{0,k}}) \mod \mathcal{C}_{n-k}$$

$$g^{n-k-1} = t_1 f^{n-k} + t_2 g^{n-k} + \sum_{j=1}^{m}(B_j^1\frac{\partial}{\partial p_j^{k,0}} + B_j^2\frac{\partial}{\partial p_j^{k-1,1}} + \cdots + B_j^{k+1}\frac{\partial}{\partial p_j^{0,k}}) \mod \mathcal{C}_{n-k},$$

where $s_1, s_2, t_1, t_2, A_j^i, B_j^i$, for $1 \le i \le k+1$ and $1 \le j \le m$, are smooth functions defined in a neighborhood $\mathcal{O}$ of $0 \in \mathbb{R}^N$. The condition (iv), applied for $i = n-k-1$, implies that $f^{n-k-1}$ and $g^{n-k-1}$ are independent at $0 \in \mathbb{R}^N$, modulo $\mathcal{L}_{n-1}$. Hence $(s_1t_2 - s_2t_1)(0) \ne 0$ and thus by a suitable choice of $f^{n-k-1}$ and $g^{n-k-1}$, we can assume that

$$f^{n-k-1} = f^{n-k} + \sum_{j=1}^{m}(A_j^1\frac{\partial}{\partial p_j^{k,0}} + A_j^2\frac{\partial}{\partial p_j^{k-1,1}} + \cdots + A_j^{k+1}\frac{\partial}{\partial p_j^{0,k}})$$

$$g^{n-k-1} = g^{n-k} + \sum_{j=1}^{m}(B_j^1\frac{\partial}{\partial p_j^{k,0}} + B_j^2\frac{\partial}{\partial p_j^{k-1,1}} + \cdots + B_j^{k+1}\frac{\partial}{\partial p_j^{0,k}}).$$

We will now analyze the functions $A_j^i, B_j^i$, for $1 \le i \le k+1$ and $1 \le j \le m$. First, for $0 \le i \le k$ and $1 \le j \le m$, we will replace $p^{i,k-i}$ by $p^{i,k-i} - A_j^{i+1}(0)x_1 - B_j^{i+1}(0)x_2$

in order to obtain $A_j^{i+1}(0) = B_j^{i+1}(0) = 0$. This change of coordinates preserves $\mathcal{L}_{n-1}$ and $\mathcal{C}_i$, for $1 \le i \le n-1$. Secondly, the characteristic distribution of $\mathcal{D}^{(n-k-1)}$ is $\mathcal{C}_{n-k-1} = \operatorname{span}\{\frac{\partial}{\partial p_j^\sigma}, \quad 1 \le j \le m, \quad k+2 \le |\sigma| \le n\}$. It follows that, for any $1 \le j \le m$ and $k+2 \le |\sigma| \le n$,

$$[\frac{\partial}{\partial p_j^\sigma}, f^{n-k-1}] = [\frac{\partial}{\partial p_j^\sigma}, g^{n-k-1}] = 0$$

and hence

$$A_j^i = A_j^i(x_1, x_2, p^\sigma), \quad B_j^i = B_j^i(x_1, x_2, p^\sigma), \quad 0 \le |\sigma| \le k+1, \tag{3.5.4}$$

for $1 \le i \le k+1$, $1 \le j \le m$, where $p^\sigma = (p_1^\sigma, \ldots, p_m^\sigma)$. Thirdly, a direct calculation shows that

$$\begin{aligned}[f^{n-k-1}, g^{n-k-1}] \quad = \quad & \sum_{j=1}^m \Big[(A_j^2 - B_j^1)\frac{\partial}{\partial p_j^{k-1,0}} + (A_j^3 - B_j^2)\frac{\partial}{\partial p_j^{k-2,1}} + \cdots \\ & +(A_j^{k+1} - B_j^k)\frac{\partial}{\partial p_j^{0,k-1}}\Big] \quad \operatorname{mod} \quad \mathcal{C}_{n-k+1}.\end{aligned}$$

On the other hand, $[f^{n-k-1}, g^{n-k-1}] \in \mathcal{D}^{n-k} = \mathcal{C}_{n-k+1} \oplus \operatorname{span}\{f^{n-k-1}, \, g^{n-k-1}\}$, which implies that $B_j^i = A_j^{i+1}$, for $1 \le i \le k$ and $1 \le j \le m$. Now consider the distribution

$$\begin{aligned}\mathcal{E}^{n-k-1} \quad = \quad & \operatorname{span}\Big\{\frac{\partial}{\partial x_1} + \sum_{j=1}^m \Big(\sum_{|\sigma|=0}^{k-1} p_j^{\sigma+1_1}\frac{\partial}{\partial p_j^\sigma} + A_j^1\frac{\partial}{\partial p_j^{k,0}} + A_j^2\frac{\partial}{\partial p_j^{k-1,1}} + \cdots + A_j^{k+1}\frac{\partial}{\partial p_j^{0,k}}\Big), \\ & \frac{\partial}{\partial x_2} + \sum_{j=1}^m \Big(\sum_{|\sigma|=0}^{k-1} p_j^{\sigma+1_2}\frac{\partial}{\partial p_j^\sigma} + B_j^1\frac{\partial}{\partial p_j^{k,0}} + B_j^2\frac{\partial}{\partial p_j^{k-1,1}} + \cdots + B_j^{k+1}\frac{\partial}{\partial p_j^{0,k}}\Big), \\ & \frac{\partial}{\partial p_j^\sigma}, \quad 1 \le j \le m, \quad |\sigma| = k+1\Big\}.\end{aligned}$$

around $0 \in \mathbb{R}^{\frac{m}{2}(k+2)(k+3)+2}$, equipped with the coordinates $(x_1, x_2, p_j^\sigma)$, for $1 \le j \le m$ and $0 \le |\sigma| \le k+1$. The distribution $\mathcal{E}^{n-k-1}$ is well defined (because of (3.5.4)) and, moreover, it satisfies the assumption of Lemma 3.5.3 with $r = (k+2)m$, $l = 2$ and $s = \frac{m}{2}(k+1)(k+2) + 2$. Indeed, $\mathcal{C}(\mathcal{E}^{n-k-1}) = 0$ because $\mathcal{C}(\mathcal{D}^{(n-k-1)}) = \mathcal{C}_{n-2}$ and, moreover, the coordinates $p_i$, $1 \le i \le r$, correspond to $p_j^\sigma$, $1 \le j \le m$, $|\sigma| = k+1$, (notice that $\sharp\{\sigma : |\sigma| = k+1\} = k+2$), and the coordinates $y_i$, $1 \le i \le s$, to $(x_1, x_2, p_j^\sigma)$, $1 \le j \le m$ and $0 \le |\sigma| \le k$. It follows by Lemma 3.5.3 that the map $\varphi$

given by

$$\begin{array}{lcll}
\tilde{x}_i & = & x_i, & i = 1, 2 \\
\tilde{p}_j^\sigma & = & p_j^\sigma, & 1 \le j \le m, \quad |\sigma| \ne k+1 \\
\tilde{p}_j^{k+1,0} & = & A_j^1, & 1 \le j \le m \\
\tilde{p}_j^{k,1} & = & A_j^2 = B_j^1, & 1 \le j \le m \\
 & \vdots & & \\
\tilde{p}_j^{1,k} & = & A_j^{k+1} = B_j^k, & 1 \le j \le m \\
\tilde{p}_j^{0,k+1} & = & B_j^{k+1}, & 1 \le j \le m,
\end{array}$$

is a local diffeomorphism, denoted by $\varphi$, in a neighborhood of $0 \in \mathbb{R}^{\frac{m}{2}(n+1)(n+2)+2}$. It is easy to see that the local diffeomorphism $\varphi$ preserves the nested sequence of distributions

$$\mathcal{C}_1 \subset \cdots \subset \mathcal{C}_{n-1} \subset \mathcal{C}_n = \mathcal{L}_{n-1},$$

that is,

$$\varphi_* \mathcal{C}_i = \operatorname{span}\left\{\frac{\partial}{\partial \tilde{p}_j^\sigma}, \quad 1 \le j \le m, \quad n-i+1 \le |\sigma| \le n\right\},$$

for $1 \le i \le n$, and that the vector fields $f^{n-k-1}$ and $g^{n-k-1}$ become, respectively,

$$\begin{aligned}
\varphi_* f^{n-k-1} &= \frac{\partial}{\partial \tilde{x}_1} + \sum_{|\sigma|=0}^{k} \sum_{j=1}^{m} \tilde{p}_j^{\sigma+1_1} \frac{\partial}{\partial \tilde{p}_j^\sigma} \\
\varphi_* g^{n-k-1} &= \frac{\partial}{\partial \tilde{x}_2} + \sum_{|\sigma|=0}^{k} \sum_{j=1}^{m} \tilde{p}_j^{\sigma+1_2} \frac{\partial}{\partial \tilde{p}_j^\sigma}.
\end{aligned}$$

□

### 3.5.3 Proof of Theorem 3.3.4

**Proof:** Let $\mathcal{I}$ be a codistribution defined by $\mathcal{I} = \mathcal{D}^\perp$ and locally spanned by differential 1-forms $\omega^1, \ldots, \omega^m$, that is, $\mathcal{I} = \operatorname{span}\{\omega^1, \ldots, \omega^m\}$.

(i) $\Rightarrow$ (iv): Since $\operatorname{rank}_C \mathcal{I} = k$, there exist $\pi^1, \ldots, \pi^k \in \Lambda^1(M)$ such that

$$\pi^1 \wedge \cdots \wedge \pi^k \wedge \omega^1 \wedge \cdots \wedge \omega^m \ne 0,$$

and

$$\mathrm{d}\omega \wedge \pi^1 \wedge \cdots \wedge \pi^k \equiv 0 \mod \mathcal{I} \tag{3.5.5}$$

for all $\omega \in \mathcal{I}$. The relation (3.5.5) can be rewritten, for the $\omega^i$'s that span $\mathcal{I}$, as

$$\mathrm{d}\omega^i = \sum_{j=1}^{k} \eta_j^i \wedge \pi^j \mod \mathcal{I}. \tag{3.5.6}$$

The condition $\mathcal{C}(\mathcal{D}) = 0$ implies that the differential forms $\omega^i, \pi^j, \eta^i_j$ are independent everywhere. To see this, assume that there exists a point $q \in M$ such that the differential forms $\omega^i, \pi^j, \eta^i_j$, for $1 \le i \le m$ and $1 \le j \le k$ are dependent at $q$. Since $\dim M = m + k + mk$, there exists a vector $0 \neq v \in T_qM$ such that $< \eta^i_j(q), v >=< \pi^l(q), v >=< \omega^r(q), v >= 0$. It follows that $v \in \mathcal{I}^\perp(q) = \mathcal{D}(q)$ and that $\mathrm{d}\omega^i(v, w) = 0$, for any $w \in \mathcal{D}(q)$, thus implying that $0 \neq v \in \mathcal{C}(\mathcal{D})(q)$ and contradicting the assumption $\mathcal{C}(\mathcal{D})(q) = 0$ for any $q \in M$.

Now taking the exterior derivative of both sides of (3.5.6), we get

$$\sum_{j=1}^{k} \eta^i_j \wedge \mathrm{d}\pi^j = 0 \mod \left(\mathcal{I}, \pi^1, \ldots, \pi^k\right)$$

which implies for any $1 \le j \le k$,

$$\mathrm{d}\pi^j \equiv 0 \mod (\mathcal{I}, \pi^1, \ldots, \pi^k, \eta^i_1, \ldots, \eta^i_k), \tag{3.5.7}$$

where $(\cdot)$ stands for the ideal in the exterior algebra. Notice that the above system is formed by $m$ relations, obtained for $1 \le i \le m$ by the corresponding differential forms $\omega^1, \ldots, \omega^m$. Let $\mathcal{F}$ be a distribution defined by

$$\begin{aligned} \mathcal{F} &= \left(\mathrm{span}\,\{\omega^1, \ldots, \omega^m, \pi^1, \ldots, \pi^k\}\right)^\perp \\ &= \mathrm{span}\,\{f \in \mathcal{D} : \pi^j(f) = 0, \quad 1 \le j \le k\}. \end{aligned}$$

Obviously, $\mathcal{F}$ satisfies $\mathrm{corank}\,(\mathcal{F} \subset \mathcal{D}) = k$ and now we will show that $\mathcal{F}$ is involutive.

Define the distributions $\mathcal{F}_i$, for $1 \le i \le m$, as

$$\begin{aligned} \mathcal{F}_i &= \left(\mathrm{span}\,\{\omega^1, \ldots, \omega^m, \pi^1, \ldots, \pi^k, \eta^i_1, \ldots, \eta^i_k\}\right)^\perp \\ &= \mathrm{span}\,\{f \in \mathcal{D} : \pi^j(f) = 0, \quad \eta^i_j(f) = 0, \quad 1 \le j \le k\}. \end{aligned}$$

The relation (3.5.6) implies that for any $1 \le s \le m$,

$$\mathrm{d}\omega^s = \sum_{j=1}^{k} \eta^s_j \wedge \pi^j + \sum_{i=1}^{m} A^s_i \wedge \omega^i. \tag{3.5.8}$$

On one hand, for any $f, g \in \mathcal{F}_i$ and $1 \le s \le m$,

$$\mathrm{d}\omega^s(f, g) = \sum_{j=1}^{k} \eta^s_j \wedge \pi^j(f, g) + \sum_{i=1}^{m} A^s_i \wedge \omega^i(f, g) = 0,$$

on the other hand,

$$\mathrm{d}\omega^s(f, g) = L_f \omega^s(g) - L_g \omega^s(f) - \omega^s([f, g]) = -\omega^s([f, g]).$$

Therefore $\omega^s([f,g]) = 0$ for $1 \leq s \leq m$ and any $f, g \in \mathcal{F}_i$. Moreover, since $\mathrm{d}\pi^j \equiv 0 \mod (\mathcal{I}, \pi^1, \ldots, \pi^k, \eta^i_1, \ldots, \eta^i_k)$, by an analogous argument as above, we obtain that $\mathrm{d}\pi^j(f,g) = -\pi^j([f,g]) = 0$, for $1 \leq j \leq k$. Therefore it can be concluded that for any $f, g \in \mathcal{F}_i$, we always have $[f,g] \in \mathcal{F}$ which implies that

$$[\mathcal{F}_i, \mathcal{F}_i] \subset \mathcal{F}, \quad 1 \leq i \leq m.$$

We claim that $\mathcal{F} = \sum_{i=1}^m \mathcal{F}_i$. The relation $\sum_{i=1}^m \mathcal{F}_i \subset \mathcal{F}$ is obvious. Recall that the differential forms $\omega^1, \ldots, \omega^m, \pi^1, \ldots, \pi^k, \eta^i_j$, for $1 \leq i \leq m$ and $1 \leq j \leq k$, are independent everywhere so they generate $\Lambda^1(M)$ over $C^\infty(M)$. Choose vector fields $f^1_1, \ldots, f^1_k, \ldots, f^m_1, \ldots, f^m_k$, such that $f^i_j$, for $1 \leq i \leq m$ and $1 \leq j \leq k$, satisfy

$$\eta^l_s(f^i_j) = \begin{cases} 1 & i = l, \ j = s \\ 0 & \text{otherwise.} \end{cases}$$

and, moreover,

$$\mathcal{F} = \operatorname{span}\{f^1_1, \ldots, f^1_k, \ldots, f^m_1, \ldots, f^m_k\}.$$

Then the subdistributions $\mathcal{F}_i$, for $1 \leq i \leq m$, are given by

$$\mathcal{F}_i = \operatorname{span}\{f^1_1, \ldots, f^1_k, \ldots, f^{i-1}_1, \ldots, f^{i-1}_k, f^{i+1}_1, \ldots, f^{i+1}_k, \ldots, f^m_1, \ldots, f^m_k\}.$$

Clearly, for any $f^i_j \in \mathcal{F}$, where $1 \leq i \leq m$ and $1 \leq j \leq k$, we have $f^i_j \in \mathcal{F}_s$, for any $s \neq i$ and $1 \leq s \leq m$. This implies that $\mathcal{F} \subset \sum_{i=1}^m \mathcal{F}_i$ and thus we get

$$\mathcal{F} = \mathcal{F}_1 + \cdots + \mathcal{F}_m.$$

Moreover, it is easy to see that it is enough to take in the above sum only two terms corresponding to any $1 \leq i \neq j \leq m$.

In order to prove involutivity of $\mathcal{F}$, consider any two vector fields $f^l_j, f^r_s$ in $\mathcal{F} = \sum_{i=1}^m \mathcal{F}_i$. Since $m \geq 3$, there always exists $\mathcal{F}_i$ such that $i \neq l$ and $i \neq r$ and thus $f^l_j, f^r_s \in \mathcal{F}_i$. Recall that $[\mathcal{F}_i, \mathcal{F}_i] \subset \mathcal{F}$ and thus we get

$$[f^l_j, f^r_s] \in \mathcal{F},$$

proving that $[\mathcal{F}, \mathcal{F}] \subset \mathcal{F}$, i.e., $\mathcal{F}$ is involutive. Put $\mathcal{L} = \mathcal{F}$, then $\mathcal{L}$ is involutive and has corank k in $\mathcal{D}$.

Now we will prove the uniqueness of $\mathcal{L}$. Suppose that there exists another involutive subdistribution $\tilde{\mathcal{L}}$ in $\mathcal{D}$ of corank $k$ and $\mathcal{L} \neq \tilde{\mathcal{L}}$. Observe that $\operatorname{rank}(\mathcal{L} \cap \tilde{\mathcal{L}})$ is locally constant on an open and dense submanifold $\tilde{M}$ of $M$ and constant on each connected component of $\tilde{M}$. Choose one such component. Clearly, we have

$mk-k \le \operatorname{rank}(\mathcal{L}\cap\tilde{\mathcal{L}}) \le mk-1$. If $\operatorname{rank}(\mathcal{L}\cap\tilde{\mathcal{L}}) = mk-k$, then there exist vector fields $h_1,\ldots,h_{mk-k}, f_1,\ldots,f_k, g_1,\ldots,g_k$ such that

$$\begin{aligned}\mathcal{L} &= \operatorname{span}\{h_1,\ldots,h_{mk-k}, f_1,\ldots,f_k\}\\ \tilde{\mathcal{L}} &= \operatorname{span}\{h_1,\ldots,h_{mk-k}, g_1,\ldots,g_k\}.\end{aligned}$$

On one hand,

$$\mathcal{L}\cup\tilde{\mathcal{L}} = \operatorname{span}\{h_1,\ldots,h_{mk-k}, f_1,\ldots,f_k, g_1,\ldots,g_k\}\subset\mathcal{D}.$$

On the other hand, we have

$$\operatorname{rank}(\mathcal{L}\cup\tilde{\mathcal{L}}) = \operatorname{rank}\mathcal{L}+\operatorname{rank}\tilde{\mathcal{L}}-\operatorname{rank}(\mathcal{L}\cap\tilde{\mathcal{L}}) = k+mk = \operatorname{rank}\mathcal{D}.$$

This implies that $\mathcal{D} = \operatorname{span}\{h_1,\ldots,h_{mk-k}, f_1,\ldots,f_k, g_1,\ldots,g_k\}$ and, moreover, $h_i$ for $1\le i\le mk-k$, are characteristic vector fields of $\mathcal{D}$ which contradicts the condition $\mathcal{C}(\mathcal{D})=0$.

Assume now that $\operatorname{rank}(\mathcal{L}\cap\tilde{\mathcal{L}}) = mk-r$ where $1\le r<k$. Then there exist vector fields $h_1,\ldots,h_{mk-r}, f_1,\ldots,f_r, g_1,\ldots,g_r$ such that

$$\begin{aligned}\mathcal{L} &= \operatorname{span}\{h_1,\ldots,h_{mk-r}, f_1,\ldots,f_r\}\\ \tilde{\mathcal{L}} &= \operatorname{span}\{h_1,\ldots,h_{mk-r}, g_1,\ldots,g_r\}.\end{aligned}$$

The involutivity of $\mathcal{L}$ implies that, without loss of generality, we can choose the $h_i$'s and $f_j$'s that commute. Therefore there exist local coordinates $(x,y)=(x_1,\ldots,x_{mk}, y_1,\ldots,y_{m+k})$, in a neighborhood of any point $q\in\tilde{M}$, such that for $1\le i\le mk-r$ and $1\le j\le r$,

$$\begin{aligned}&h_i = \frac{\partial}{\partial x_i},\quad f_j = \frac{\partial}{\partial x_{mk-r+j}},\\ &g_j = \frac{\partial}{\partial y_j} + \sum_{l=r+1}^{k+m} g_j^l(x,y)\frac{\partial}{\partial y_l} \mod \mathcal{L}\end{aligned}$$

(the form of $g_1,\ldots,g_r$ follows from their independence). Since $\tilde{\mathcal{L}}$ is involutive, it is easy to see that $\dfrac{\partial g_j^l}{\partial x_i}\equiv 0$, for $1\le i\le mk-r$, $1\le j\le r$ and $r+1\le l\le k+m$ which implies that $g_j^l = g_j^l(x_{mk-r+1},\ldots,x_{mk},y)$. Clearly, $\mathcal{L}\cup\tilde{\mathcal{L}} = \operatorname{span}\{h_1,\ldots,h_{mk-r}, f_1,\ldots,f_r, g_1,\ldots,g_r\}\subset\mathcal{D}$, and thus there exist $k-r$ vector fields $g_{r+1},\ldots,g_k$ such that

$$\begin{aligned}\mathcal{D} &= \operatorname{span}\{h_1,\ldots,h_{mk-r}, f_1,\ldots,f_r, g_1,\ldots,g_r, g_{r+1},\ldots,g_k\}\\ &= \operatorname{span}\{\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_{mk}}, g_1,\ldots,g_r, g_{r+1},\ldots,g_k\}\end{aligned}$$

with $g_j = \dfrac{\partial}{\partial y_j} + \displaystyle\sum_{l=k+1}^{k+m} g_j^l(x,y)\dfrac{\partial}{\partial y_l}, \quad r+1 \le j \le k$. Since $\mathcal{C}(\mathcal{D}) = 0$, clearly $\mathcal{D}$ satisfies the assumption of Lemma 3.5.3, with $r = mk$, $l = k$ and $s = m+k$. Therefore, at any $q \in \tilde{M}$, we must have rank $\left(\dfrac{\partial G}{\partial x}\right)(q) = mk$, where

$$G = \left(g_1^{r+1}, \ldots, g_1^{k+m}, \ldots, g_r^{r+1}, \ldots, g_r^{k+m}, g_{r+1}^{k+1}, \ldots, g_{r+1}^{k+m}, \ldots, g_k^{k+1}, \ldots, g_k^{k+m}\right)^\top.$$

A direct computation gives

$$\frac{\partial G}{\partial x} = \begin{pmatrix} 0 & \cdots & 0 & \dfrac{\partial g_1^{r+1}}{\partial x_{mk-r+1}} & \cdots & \dfrac{\partial g_1^{r+1}}{\partial x_{mk}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & \dfrac{\partial g_r^{k+m}}{\partial x_{mk-r+1}} & \cdots & \dfrac{\partial g_r^{k+m}}{\partial x_{mk}} \\ \dfrac{\partial g_{r+1}^{k+1}}{\partial x_1} & \cdots & \dfrac{\partial g_{r+1}^{k+1}}{\partial x_{mk-r}} & \dfrac{\partial g_{r+1}^{k+1}}{\partial x_{mk-r+1}} & \cdots & \dfrac{\partial g_{r+1}^{k+1}}{\partial x_{mk}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial g_k^{k+m}}{\partial x_1} & \cdots & \dfrac{\partial g_k^{k+m}}{\partial x_{mk-r}} & \dfrac{\partial g_k^{k+m}}{\partial x_{mk-r+1}} & \cdots & \dfrac{\partial g_k^{k+m}}{\partial x_{mk}} \end{pmatrix} = \begin{pmatrix} 0 & G_{12} \\ G_{21} & G_{22} \end{pmatrix}.$$

Since $m \ge 3$, we have rank $\left(\dfrac{\partial G}{\partial x}\right) \le \operatorname{rank} G_{21} + \operatorname{rank} G_{12} \le m(k-r) + r < mk$, which contradicts rank $\left(\dfrac{\partial G}{\partial x}\right)(q) = mk$. Therefore $\mathcal{L}$ and $\tilde{\mathcal{L}}$ coincide on $\tilde{M}$, which is open and dense, and hence (being of constant rank) coincide everywhere on $M$.

(iv) $\Rightarrow$ (iii) is obvious.

(iii) $\Rightarrow$ (i): Choose, locally, $m$ differential 1-forms $\omega^1, \ldots, \omega^m$ and $k$ differential 1-forms $\pi^1, \ldots, \pi^k$ such that $\pi^1 \wedge \cdots \wedge \pi^k \wedge \omega^1 \wedge \cdots \wedge \omega^m \neq 0$, and

$$\begin{aligned} \mathcal{F}^\perp &= \operatorname{span}\{\pi^1, \ldots, \pi^k, \omega^1, \ldots, \omega^m\} \\ \mathcal{D}^\perp &= \operatorname{span}\{\omega^1, \ldots, \omega^m\} \end{aligned}$$

Since $\dim M = k + m + mk$, locally there exist $mk$ differential 1-forms $\eta^1, \ldots, \eta^{mk}$ such that $\pi^1, \ldots, \pi^k, \omega^1, \ldots, \omega^m, \eta^1, \ldots, \eta^{mk}$ are independent everywhere. Then we can choose $m + k + mk$ vector fields $g_1, \ldots, g_k, h_1, \ldots, h_m, f_1, \ldots, f_{km}$, satisfying

$$\pi^i(g_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \quad 1 \le i, j \le k,$$

$$\omega^i(h_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \quad 1 \le i, j \le m,$$

and

$$\eta^i(f_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \quad 1 \le i, j \le mk,$$

such that $\mathcal{D} = \text{span}\{g_1, \ldots, g_k, f_1, \ldots, f_{km}\}$ and $\mathcal{F} = \text{span}\{f_1, \ldots, f_{km}\}$. For simplicity, we denote

$$\{\Omega^1, \ldots, \Omega^{k+m}\} = \{\pi^1, \ldots, \pi^k, \omega^1, \ldots, \omega^m\}.$$

Then, for any $1 \le i \le m$, the 2-form $\mathrm{d}\omega^i$ can be expressed as

$$\mathrm{d}\omega^i = \sum_{j=1}^{k+m} A^i_j \wedge \Omega^j + \sum_{l<s,} B^i_{ls} \eta^l \wedge \eta^s, \tag{3.5.9}$$

for some differential 1-forms $A^i_j$ and some functions $B^i_{ls}$. On one hand, for any vector fields $f_l, f_s \in \mathcal{F}$, $1 \le l < s \le km$, we have

$$\mathrm{d}\omega^i(f_l, f_s) = B^i_{ls}.$$

On the other hand,

$$\begin{aligned} \mathrm{d}\omega^i(f_l, f_s) &= L_{f_l}\omega^i(f_s) - L_{f_s}\omega^i(f_l) - \omega^i([f_l, f_s]) \\ &= \omega^i([f_l, f_s]) \\ &= 0. \end{aligned}$$

Therefore we get $B^i_{ls} \equiv 0$ and thus

$$\mathrm{d}\omega^i = \sum_{j=1}^{k+m} A^i_j \wedge \Omega^j = \sum_{j=1}^{m} \alpha^i_j \wedge \omega^j + \sum_{s=1}^{k} \beta^i_s \wedge \pi^s,$$

for some differential 1-forms $\alpha^i_j$ and $\beta^i_s$, which gives immediately that

$$\mathrm{d}\omega \wedge \pi^1 \wedge \cdots \wedge \pi^k \wedge \omega^1 \wedge \cdots \wedge \omega^m \equiv 0$$

for any $\omega \in \mathcal{I} = \mathcal{D}^\perp$. By the definition of the Cartan rank, it can be conclude that $\text{rank}_C \mathcal{D} = k$.

(ii) $\Leftrightarrow$ (i) has been proved by Bryant in [7].

In order to prove $\mathcal{L} = \tilde{\mathcal{F}}_1 + \cdots + \tilde{\mathcal{F}}_m$, it is enough to show that $\tilde{\mathcal{F}}_i = \mathcal{F}_i$, for $1 \le i \le m$, where $\mathcal{F}_i$ have been defined when proving the implication (i) $\Leftrightarrow$ (iv). Recall the equations (3.5.8)

$$\mathrm{d}\omega^i = \sum_{j=1}^{k} \eta^i_j \wedge \pi^j + \sum_{s=1}^{k} A^i_s \wedge \omega^s.$$

According to the definition of the distribution $\mathcal{F}_i$, for any $f \in \mathcal{F}_i$, we have

$$\omega^i(f) = 0, \quad \pi^j(f) = 0, \quad \eta^i_j(f) = 0$$

for $1 \le i \le m$ and $1 \le j \le k$. Therefore

$$f \lrcorner d\omega^i = \sum_{j=1}^{k} \left(\eta^i_j(f) \cdot \pi^j - \pi^j(f) \cdot \eta^i_j\right) + \sum_{s=1}^{m} \left(A^i_s(f) \cdot \omega^s - \omega^s(f) \cdot A^i_s\right)$$

$$= \sum_{s=1}^{m} A_s^i(f) \cdot \omega^i \in \mathcal{I} = \mathcal{D}^{\perp},$$

which implies that $f \in \tilde{\mathcal{F}}_i$ and thus $\mathcal{F}_i \subset \tilde{\mathcal{F}}_i$. On the other hand, for any $f \in \tilde{\mathcal{F}}_i \subset \mathcal{D}$, we have clearly $\omega^s(f) = 0$, for $1 \leq s \leq m$ and

$$\begin{aligned} f \lrcorner d\omega^i &= \sum_{j=1}^{k} \left(\eta_j^i(f) \cdot \pi^j - \pi^j(f) \cdot \eta_j^i\right) + \sum_{s=1}^{m} \left(A_s^i(f) \cdot \omega^s - \omega^s(f) \cdot A_s^i\right) \\ &= \sum_{j=1}^{k} \left(\eta_j^i(f) \cdot \pi^j - \pi^j(f) \cdot \eta_j^i\right) + \sum_{s=1}^{m} A_s^i(f) \cdot \omega^s. \end{aligned}$$

The above expression belongs to $\mathcal{D}^{\perp}$ (since $f \lrcorner d\omega^i \in \mathcal{D}^{\perp}$) and therefore $\eta_j^i(f) = 0$ and $\pi^j(f) = 0$ because the differential forms $\eta_j^i$, $\pi^j$, and $\omega^s$ are independent. This implies that $f \in \mathcal{F}_i$ and thus $\tilde{\mathcal{F}}_i \subset \mathcal{F}_i$ implying that $\tilde{\mathcal{F}}_i = \mathcal{F}_i$. $\square$

# Chapter 4

# Orbital feedback linearization for multi-input control systems

## 4.1 Introduction

Feedback linearization is a powerful tool for nonlinear control systems and has attracted a lot of research in recent years. Following the work of Brockett [5] who solved the state feedback linearization for single-input systems under a restricted feedback, Jakubczyk and Respondek [29] and Hunt and Su [22] gave geometric necessary and sufficient conditions for linearizing multi-input affine control systems under change of coordinates and general feedback which modifies both the drift and the control vector fields.

In the theory of dynamical systems two natural equivalence relations are considered: equivalence under diffeomorphisms and orbital equivalence. In the former case, the diffeomorphism establishing the equivalence maps trajectories into trajectories, understood as parameterized curves. In the latter case, equivalent systems have also the same trajectories but considered as non-parameterized curves (that is, the same trajectories up to a diffeomorphism and a time re-scaling). Observe that any two feedback equivalent systems have the same families of trajectories, up to a diffeomorphism, although their parametrization with respect to controls is different. It is then natural to ask when two control systems have the same families of trajectories, up to a reparametrization with respect to controls and a time re-scaling (that is, when they are orbital feedback equivalent). The linearization of a control system under feedback and time re-scaling was studied for the first time by Sampei and Furuta [67]. In their approach, solving a system of PDE's is necessary in order to verify feedback linearizability conditions and to find the time re-scaling function. Respondek developed in [62], for the orbital feedback linearization of a single-input system, necessary and sufficient conditions, which can be conveniently verified upon the original system.

Independently, Guay adapted an approach using exterior differential systems to derive necessary and sufficient conditions for the orbital feedback linearization of single-input systems [19] and then of multi-input systems [20]. Verifying Guay's conditions requires the search of suitable generators for the associated Pfaffian system.

This chapter considers the orbital feedback linearization for specific multi-input affine control systems. Both necessary and sufficient conditions are established so that they can be verified in terms of the original system. In addition, the time re-scaling function can be constructed using those conditions.

The chapter is organized as follows. The definitions and notations are given in Section 4.2. The main result is presented in Section 4.3. In Section 4.4, a complete version of the main theorem is derived due to an analysis of the distributions associated to any control-affine system. Two illustrating examples are proposed in Section 4.5. All proofs are given in Section 4.6.

## 4.2 Orbital feedback equivalence

Consider a control-affine system, with $m$ controls, of the following form

$$\Sigma : \quad \frac{\mathrm{d}x}{\mathrm{d}t} = f(x) + \sum_{i=1}^{m} g_i(x) u_i, \tag{4.2.1}$$

where $x \in X$, a $C^\infty$-manifold of dimension $d$, and $f$ and $g_i$, for $1 \leq i \leq m$, are $C^\infty$-vector fields on $X$. Throughout this chapter, smooth means $C^\infty$-smooth.

Define a new time scale $\tau$ such that

$$\frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma(x(t)),$$

where $\gamma$ is a smooth $\mathbb{R}$-valued function on $X$ satisfying $\gamma(\cdot) \neq 0$. Then in the new time scale $\tau$, the system $\Sigma$ can be rewritten as

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \frac{\mathrm{d}x}{\mathrm{d}t} \cdot \frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma(x) f(x) + \sum_{i=1}^{m} \gamma(x) g_i(x) u_i = \gamma(x) f(x) + \sum_{i=1}^{m} g_i(x) v_i,$$

where $v_i = \gamma(x) u_i$, $1 \leq i \leq m$, are feedback modified controls. A time re-scaling allows, as we have just seen, to multiply the drift $f$ by an arbitrary nonvanishing function while its action on control vector fields $g_i$ can be compensated by a suitable feedback transformation $u_i(t) \mapsto \gamma(x(t)) u_i(t) = v_i(t)$.

Consider another control system of the same form evolving on a manifold $\tilde{X}$

$$\tilde{\Sigma} : \quad \frac{\mathrm{d}\tilde{x}}{\mathrm{d}t} = \tilde{f}(\tilde{x}) + \sum_{i=1}^{m} \tilde{g}_i(\tilde{x}) \tilde{u}_i.$$

The two systems $\Sigma$ and $\tilde{\Sigma}$ are said to be *orbital feedback equivalent* if there exists a diffeomorphism $\varphi : X \to \tilde{X}$, an $\mathbb{R}^m$-valued smooth function $\alpha = (\alpha_1, \ldots, \alpha_m)$, a smooth matrix $\beta = (\beta_i^j)$, $1 \le i, j \le m$, invertible everywhere, and a smooth function $\gamma$, satisfying $\gamma(\cdot) \ne 0$, such that

$$\tilde{f} = \varphi_*(\gamma f + \sum_{i=1}^{m} \alpha_i g_i) \tag{4.2.2}$$

and for $1 \le i \le m$,

$$\tilde{g}_i = \varphi_*(\sum_{j=1}^{m} \beta_i^j g_j), \tag{4.2.3}$$

where $\varphi_*$ denotes the tangent map of $\varphi$, i.e.,

$$(\varphi_* f)(\tilde{x}) = \mathrm{D}\varphi(\varphi^{-1}(\tilde{x})) \cdot f(\varphi^{-1}(\tilde{x})).$$

The system $\Sigma$ defined by (4.2.1) is said to be *orbital feedback linearizable* if it is orbital feedback equivalent to a linear system of the form

$$\Lambda : \quad \dot{\tilde{x}} = A\tilde{x} + \sum_{i=1}^{m} b_i \tilde{u}_i + c. \tag{4.2.4}$$

There are two reasons to add the extra constant vector field $c$ to the dynamics of the linear control system

$$\dot{x} = Ax + \sum_{i=1}^{m} b_i u_i.$$

The first is to consider it around a non-equilibrium point $x_0$. Then in coordinates $\tilde{x} = x - x_0$, the system becomes

$$\dot{\tilde{x}} = A\tilde{x} + \sum_{i=1}^{m} b_i u_i + c,$$

where $c = Ax_0$ (of course $c$ can be incorporated into $Ax$ if the system is controllable and is considered globally on $\mathbb{R}^d$, but this may become impossible if we work locally only). The second reason occurs when dealing with linear time-varying systems of the form

$$\dot{x} = A(t)x + \sum_{i=1}^{m} b_i(t) u_i.$$

Add time as a new variable $x^0 = t$, that is, $\dot{x}^0 = 1$. If all controllability indices are time-invariant then, as it is well known ([6], [31], [44], [68]), we can bring the system, via a linear change of coordinates and feedback (both time-varying) to the form

$$\Lambda_t : \quad \begin{array}{rcl} \dot{x}^0 & = & 1 \\ \dot{\bar{x}} & = & \bar{A}\bar{x} + \sum_{i=1}^{m} \bar{b}_i \bar{u}_i, \end{array} \tag{4.2.5}$$

which in the augmented state space $\tilde{x} = (x^0, \bar{x})$ is of the form (4.2.4) with $c = (1, 0, \ldots, 0)^\top$. We will call the system $\Lambda_t$ time-augmented linear system. Throughout this chapter, whenever we speak about orbital feedback linearization, we will always mean orbital feedback equivalence to the time-augmented linear control system $\Lambda_t$. Of course, the system $\Lambda_t$ is never controllable but it can be accessible [24], [47] (the latter if and only if the linear subsystem $(\bar{A}, \bar{B})$ is controllable where $\bar{B}$ is the matrix whose columns are $(\bar{b}_1, \ldots, \bar{b}_m)$). We will speak about controllability indices and Brunovsky canonical form of $\Lambda_t$ meaning the respective objects of $(\bar{A}, \bar{B})$.

Notice that we have defined time re-scaling using any non vanishing function $\gamma$. Of course in many problems, it would be more natural to use exclusively positive functions $\gamma$ which preserve not only unparameterized system's trajectories but also the direction of the time arrow along the trajectories. It is actually easy to observe that if $\Sigma$ is orbital feedback equivalent to a time-augmented linear system $\Lambda_t$ via a negative time re-scaling function, it is so via a positive time re-scaling one. To see it, assume that $\Sigma$ is orbital feedback equivalent to a time-augmented form (4.2.5) via a feedback transformation and a time re-scaling function given by $\gamma < 0$. Then the same feedback transformation completed by the time re-scaling function defined by $-\gamma > 0$ brings $\Sigma$ into

$$\begin{aligned} \dot{x}^0 &= -1 \\ \dot{\bar{x}} &= -\bar{A}\bar{x} - \sum_{i=1}^{m} \bar{b}_i \bar{u}_i, \end{aligned}$$

and it is enough to apply the linear isomorphism replacing $x^0$ by $-x^0$. Notice that, in general, the pairs $(\bar{A}, \bar{B})$ and $(-\bar{A}, -\bar{B})$ are not conjugated via a linear isomorphism but are feedback equivalent and can be brought to the same $t$-augmented Brunovsky canonical form $\Lambda_t^{Br}$ (see the definition at the beginning of Section 4.4). More precisely, we have

**Proposition 4.2.1** *The following are equivalent, locally around $x_0 \in X$:*

(i) $\Sigma$ *is orbital feedback equivalent to a $t$-augmented linear system;*
(ii) $\Sigma$ *is orbital feedback equivalent, with $\gamma < 0$, to the $t$-augmented Brunovsky canonical form $\Lambda_t^{Br}$;*
(iii) $\Sigma$ *is orbital feedback equivalent, with $\gamma > 0$, to the $t$-augmented Brunovsky canonical form $\Lambda_t^{Br}$;*

The aim of this chapter is to find checkable geometric conditions for the problem: when is the multi-input, i.e., $m \geq 2$, system

$$\Sigma : \quad \frac{\mathrm{d}x}{\mathrm{d}t} = f(x) + \sum_{i=1}^{m} g_i(x) u_i, \quad f(x_0) \neq 0 \tag{4.2.6}$$

where $x \in X$, locally orbital feedback equivalent to a time-augmented linear system $\Lambda_t$ with equal controllability indices? The statement of the problem requires that

$f(x_0) \neq 0$, where $x_0$ is a given point around which we work, and that $\dim X = (n+1)m+1$, where all controllability indices (which coincide, by assumption) equal $n+1$. We will assume those two conditions throughout.

## 4.3 Main result

Put $\mathrm{ad}_f^0 g_i = g_i$ and inductively $\mathrm{ad}_f^{j+1} g_i = [f, \mathrm{ad}_f^j g_i]$, for $1 \leq i \leq m$ and $j \geq 0$. We define the following distributions

$$\begin{array}{rcl} \mathcal{G} & = & \mathrm{span}\,\{g_1, \ldots, g_m\}, \\ \mathcal{G}_f^j & = & \mathrm{span}\,\{f, g_i, \mathrm{ad}_f g_i, \ldots, \mathrm{ad}_f^{j-1} g_i, \quad 1 \leq i \leq m\}, \end{array}$$

for $1 \leq j \leq n+1$.

If we suppose $\dim \mathcal{G}_f^{n+1}(x) = (n+1)m+1$ in a neighborhood of $x_0$ (which we will, indeed, assume below), then we have $\dim \mathcal{G}_f^j(x) = jm+1$, for $0 \leq j \leq n$. It follows that $m$ differential 1-forms $\omega^1, \ldots, \omega^m \in \Lambda(X)$ are defined uniquely by

$$\begin{array}{rcl} \omega^j(h) & = & 0, \quad \text{for any } h \in \mathcal{G}_f^n, \\ \omega^j(ad_f^n g_i) & = & \delta_i^j \end{array} \tag{4.3.1}$$

for $1 \leq i, j \leq m$. Using those differential forms, we introduce the following $\mathbb{R}$-valued functions:

$$\mathrm{T}_{i,j}^{k,l} = \omega^k([\mathrm{ad}_f^{n-1} g_i, \mathrm{ad}_f^l g_j]) \tag{4.3.2}$$

for any $0 \leq l \leq n-1$ and any $1 \leq i, j, k \leq m$ such that $i \neq j$ in the case $l = n-1$. Of course the functions $\mathrm{T}_{i,i}^{k,n-1}$ can also be defined (actually, they vanish), but we exclude them to simplify formulations of results. We have obviously

$$\mathrm{T}_{i,j}^{k,n-1} = -\mathrm{T}_{j,i}^{k,n-1}, \tag{4.3.3}$$

for any $1 \leq i, j, k \leq m$. Our main result on orbital feedback linearization is as follows:

**Theorem 4.3.1** *The multi-input control-affine system $\Sigma$, defined by (4.2.6) for $m \geq 2$, is locally orbital feedback equivalent to a time-augmented linear system $\Lambda_t$, with all controllability indices equal to $n+1$, if and only if it satisfies, in a neighborhood of $x_0$,*

(OPL1) $\dim \mathcal{G}_f^{n+1}(x) = (n+1)m+1$;
(OPL2) $[\mathcal{G}_f^j, \mathcal{G}_f^j] \subset \mathcal{G}_f^{j+1}$, *for* $1 \leq j \leq n$;
(OPL3) $[\mathcal{G}, \mathcal{G}_f^2] \subset \mathcal{G}_f^2$;
(OPL4) *The functions* $\mathrm{T}_{i,j}^{k,l}$ *satisfy the following conditions*:

(i) $\mathrm{T}_{i,j}^{k,l} = 0$, for $\begin{cases} 1 \leq k \neq i \leq m, \quad 1 \leq j \leq m, & \text{if} \quad l < n-1 \\ 1 \leq i \neq k \neq j \leq m, & \text{if} \quad l = n-1, \end{cases}$

(ii) $\mathrm{T}_{i,j}^{i,l} = T_{k,j}^{k,l}$, for $1 \le l \le n-1$, $1 \le i,j,k \le m$ s.t $j \ne i,k$ if $l = n-1$,

(iii) *when $m = 2$, then additionally, the distribution defined by*

$$\mathcal{B} = \mathrm{span}\,\{g_i, \mathrm{ad}_f^l g_i + b_i^l f, \quad i = 1,2 \text{ and } 1 \le l \le n-1\}$$

*must be involutive, where*

$$b_i^l = \mathrm{T}_{1,i}^{1,l} = \mathrm{T}_{2,i}^{2,l}, \text{ for } 1 \le i \le 2, 1 \le l \le n-2$$
$$b_1^{n-1} = \mathrm{T}_{1,2}^{1,n-l}, \quad b_2^{n-1} = \mathrm{T}_{2,1}^{2,n-l}.$$

Theorem 4.3.1 is actually the implication (i) $\Leftrightarrow$ (ii) of the more general Theorem 4.4.6 (where other characterizations of orbital feedback linearization are given) and therefore the proof of Theorem 4.3.1 follows form the proof of Theorem 4.4.6 given in Section 4.6.
**Remark 1**. The following generalization of (OFL3) is necessary for orbital feedback linearization (see Lemma 4.6.1 in Section 4.6.1)

$$(\mathrm{OFL3})' \quad [\mathcal{G}, \mathcal{G}_f^j] \subset \mathcal{G}_f^j, \quad 2 \le j \le n+1.$$

But it can be proved that (OFL2) and (OFL3) imply (OFL3)$'$ for any $j \ge 3$.

**Remark 2**. We would like to emphasize two important features of Theorem 4.3.1. First, all its conditions are easily checkable in terms of the original system. Secondly, if the conditions (OFL1) $-$ (OFL4) hold for $\Sigma$, the time re-scaling function $\gamma$ can be constructed as follows:

(1). Put $\mathcal{B} = \mathrm{span}\,\{g_i, \mathrm{ad}_f^l g_i + b_i^l f : 1 \le i \le m, 1 \le l \le n-1\}$, where the functions $b_i^l$ are given by

$$b_i^l = \mathrm{T}_{k,i}^{k,l},$$

where $1 \le k \le m$ is any integer such that $k \ne i$ when $l = n-1$ (actually, $\mathrm{T}_{k,i}^{k,l}$ is the same function for any $k$ because of (OFL4)(ii)). We will prove below (see Lemma 4.6.3 and Proposition 4.6.4) that for $m \ge 3$, $\mathcal{B}$ is an involutive subdistribution of $\mathcal{G}_f^n$ of corank one and, for $m = 2$, we assume that in (OFL4)(iii).

(2). Choose $m+1$ smooth functions $\phi_0, \phi_1, \ldots, \phi_m$ such that

$$\mathrm{span}\,\{\mathrm{d}\phi_0, \mathrm{d}\phi_1, \ldots, \mathrm{d}\phi_m\} = \mathcal{B}^\perp,$$

where the codistribution $\mathcal{B}^\perp$ denotes the annihilator of $\mathcal{B}$.

(3). The construction of $\mathcal{B}$ implies immediately that $f(x_0) \notin \mathcal{B}(x_0)$. Thus there exists a function $\phi_i$, for a certain $0 \le i \le m$, such that $\mathrm{L}_f\phi_i(x_0) \ne 0$ where $\mathrm{L}_f\phi_i$ denotes the Lie derivative of $\phi_i$ along the vector field $f$.

(4). The time re-scaling function is given by $\gamma = \dfrac{1}{\mathrm{L}_f\phi_i}$.

Note that the time re-scaling is not unique. Indeed, take any smooth function $\tilde{\phi}$ such that $\mathrm{d}\tilde{\phi} \in \mathcal{B}^{\perp}$ and $\mathrm{L}_f\tilde{\phi}(x_0) \neq 0$. Then $\tilde{\gamma} = \frac{1}{\mathrm{L}_f\tilde{\phi}}$ defines a time re-scaling function that renders the system $\dot{x} = \tilde{f}(x) + \sum_{i=1}^{m} g_i(x)\tilde{u}_i$, with $\tilde{f} = \tilde{\gamma} f$, feedback linearizable and, conversely, any $\tilde{\gamma}$ achieving that goal is of the form $\tilde{\gamma} = \frac{1}{\mathrm{L}_f\tilde{\phi}}$, where $\tilde{\phi}$ is as above.

**Remark 3**. Obviously, a control system that is feedback linearizable is always orbital feedback linearizable. To see how it is reflected in conditions (OFL1) – (OFL4), put $\mathcal{G}_j = \operatorname{span}\{g_i, \mathrm{ad}_f g_i, \ldots, \mathrm{ad}_f^{j-1} g_i, 1 \leq i \leq m\}$ and observe that $\Sigma$, defined on $X$ of dimension $(n+1)m+1$, is locally feedback equivalent to a $t$-augmented linear system $\Lambda_t$ if and only if it satisfies the following conditions

(FL1) $\dim \mathcal{G}_{n+1}(x) = (n+1)m$;
(FL2) $\dim \mathcal{G}_{n+2}(x) = (n+1)m$;
(FL3) $[\mathcal{G}_j, \mathcal{G}_j] \subset \mathcal{G}_j$, $1 \leq j \leq n$.

Indeed, the condition (FL1) implies that $\Sigma$ decomposes into 1-dimensional system $\dot{x}^0 = 1$ followed by a $(n+1)m$-dimensional subsystem that is feedback linearizable due to the standard linearizability conditions (FL2) and (FL3) (for the latter see [24], [29],[47]). Now it is immediate to see that (FL3) implies (OFL2), (OFL3) and (OFL4) (with all $\mathrm{T}_{i,j}^{k,l} = 0$), while (FL1) and (FL2) are related with (OFL1) via $f(x_0) \notin \mathcal{G}_f^n(x_0)$, which reflects the time-augmented form of $\Lambda_t$.

## 4.4 From distributions to control-affine systems and back

In the previous section we characterized control-affine systems that are orbital feedback equivalent to a time-augmented linear system $\Lambda_t$, with all controllability indices equal. Any such system $\Lambda_t$ (with the common value of the controllability indices being $n+1$) can be transformed, via a linear change of coordinates and linear feedback, to the $t$-augmented Brunovsky canonical form

$$
\Lambda_t^{Br}: \quad
\begin{array}{rclcrcl}
\dot{x}_0^0 & = & 1 & & & & \\
\dot{x}_1^0 & = & x_1^1 & \cdots & \dot{x}_m^0 & = & x_m^1 \\
 & \vdots & & & & \vdots & \\
\dot{x}_1^{n-1} & = & x_1^n & \cdots & \dot{x}_m^{n-1} & = & x_m^n \\
\dot{x}_1^n & = & u_1 & \cdots & \dot{x}_m^n & = & u_m.
\end{array}
$$

The drift and the control vector fields of the above form are, respectively,

$$
\begin{aligned}
f^{cc} &= \frac{\partial}{\partial x_0^0} + \sum_{j=0}^{n-1}\sum_{i=1}^{m} x_i^{j+1} \frac{\partial}{\partial x_i^j}, \\
g_1^{cc} &= \frac{\partial}{\partial x_1^n}, \quad \cdots, \quad g_m^{cc} = \frac{\partial}{\partial x_m^n},
\end{aligned}
\tag{4.4.1}
$$

and define on $\mathbb{R}^{(n+1)m+1}$ a distribution of rank $m+1$

$$\mathcal{CC}^n(\mathbb{R},\mathbb{R}^m) = \mathrm{span}\,\{f^{cc}, g_1^{cc}, \ldots, g_m^{cc}\},$$

called the Cartan distribution. Formal similarity of the $t$-augmented Brunovsky canonical form $\Lambda_t^{Br}$ and of the vector fields spanning $\mathcal{CC}^n(\mathbb{R},\mathbb{R}^m)$ rise a natural question: what are relations between control systems that are orbital feedback equivalent to $\Lambda_t^{Br}$ and distributions equivalent to $\mathcal{CC}^n(\mathbb{R},\mathbb{R}^m)$. This section is devoted to various aspects of that question.

Consider an arbitrary distribution $\mathcal{D}$. The *derived flag* of $\mathcal{D}$ is the sequence of modules of vector fields $\mathcal{D}^{(0)} \subset \mathcal{D}^{(1)} \subset \cdots$ defined inductively by

$$\mathcal{D}^{(0)} = \mathcal{D} \quad \text{and} \quad \mathcal{D}^{(i+1)} = \mathcal{D}^{(i)} + [\mathcal{D}^{(i)}, \mathcal{D}^{(i)}], \quad \text{for} \quad i \geq 0.$$

The *Lie flag* of $\mathcal{D}$ is the sequence of modules of vector fields $\mathcal{D}_0 \subset \mathcal{D}_1 \subset \cdots$ defined inductively by

$$\mathcal{D}_0 = \mathcal{D} \quad \text{and} \quad \mathcal{D}_{i+1} = \mathcal{D}_i + [\mathcal{D}_0, \mathcal{D}_i], \quad \text{for} \quad i \geq 0.$$

In general, the derived and Lie flags are different, though for any point $x$ in the underlying manifold the inclusion $\mathcal{D}_i(x) \subset \mathcal{D}^{(i)}(x)$ holds, for $i \geq 0$.

A *characteristic vector field* of a distribution $\mathcal{D}$ is a vector field $f$ that belongs to $\mathcal{D}$ and satisfies $[f, \mathcal{D}] \subset \mathcal{D}$. The *characteristic distribution* of $\mathcal{D}$, which will be denoted by $\mathcal{C}(\mathcal{D})$, is the module spanned by all its characteristic vector fields. It follows directly from the Jacobi identity that the characteristic distribution is always involutive but, in general, it need not be of constant rank. Let $\mathcal{L}$ be a subdistribution contained in $\mathcal{D}$, we denote by $\mathrm{corank}\,(\mathcal{L} \subset \mathcal{D})$ the corank of $\mathcal{L}$ in $\mathcal{D}$, that is the value of $\mathrm{rank}\,\mathcal{D} - \mathrm{rank}\,\mathcal{L}$.

The following theorem, given by Pasillas-Lépine and Respondek [51], [54], characterizes distributions which are equivalent to the Cartan distribution $\mathcal{CC}^n(\mathbb{R},\mathbb{R}^m)$.

**Theorem 4.4.1** *A rank $m+1$ distribution $\mathcal{D}$, with $m \geq 2$, on a manifold $X$ of dimension $(n+1)m+1$ is locally, at $x_0 \in X$, equivalent to the Cartan distribution $\mathcal{CC}^n(\mathbb{R},\mathbb{R}^m)$ if and only if the following conditions hold around $x_0$, for $0 \leq j \leq n$.*

- (B1) *Each distribution $\mathcal{D}^{(j)}$ is of rank $(j+1)m+1$ and contains an involutive subdistribution $\mathcal{L}_j$ such that* $\mathrm{corank}\,(\mathcal{L}_j \subset \mathcal{D}^{(j)}) = 1$*;*
- (B2) $\mathrm{rank}\,\mathcal{D}^{(j)} = \mathrm{rank}\,\mathcal{D}_j$.

**Remark**. One can prove that, for $0 \leq j \leq n-2$, we have $\mathcal{L}_j = \mathcal{C}(\mathcal{D}^{(j+1)})$, the characteristic distribution of $\mathcal{D}^{(j+1)}$. It follows (see [54] for details) that the condition (B1) can be weaken to

(B1)′ $\mathcal{D}^{(n)} = TX$;
(B1)″ rank $\mathcal{D}^{(n-1)} = nm + 1$ and there exists an involutive subdistribution $\mathcal{L}_{n-1}$ of $\mathcal{D}^{(n-1)}$ such that corank $(\mathcal{L}_{n-1} \subset \mathcal{D}^{(n-1)}) = 1$.

### 4.4.1 From distributions to control-affine systems

Recall that $\mathcal{G} = \operatorname{span}\{g_1, \ldots, g_m\}$ denotes the control distribution spanned by the control vector fields. Observe that the set of all admissible velocities of the control-affine system $\Sigma$ at $x \in X$ is the affine subspace $\mathcal{A}(x) = f(x) + \mathcal{G}(x)$ of the tangent space $T_xX$. Therefore to any control affine system there corresponds an affine distribution $\mathcal{A} = f + \mathcal{G}$ which assigns to any $x \in X$ the affine subspace $\mathcal{A}(x)$ of $T_xX$. Feedback equivalence can be rephrased, using the notion of affine distributions, in an invariant way that does not depend on particular control parameterizations of $\Sigma$. Indeed, two control-affine systems $\Sigma$ and $\tilde{\Sigma}$ are feedback equivalent if and only if there exists a diffeomorphism $\varphi$ such that

$$\tilde{\mathcal{A}} = \varphi_* \mathcal{A},$$

provided that the control distributions $\mathcal{G}$ and $\tilde{\mathcal{G}}$ are of constant rank. Similarly, they are orbital feedback equivalent if there exists a diffeomorphism $\varphi$ and a function $\gamma(\cdot) \neq 0$ such that

$$\tilde{\mathcal{A}} = \varphi_*(\gamma \mathcal{A}).$$

Here $\gamma\mathcal{A}$ means that all vector fields of $\mathcal{A}$ are multiplied by $\gamma$, i.e.,

$$\gamma \mathcal{A} = \gamma f + \mathcal{G}.$$

Consider an arbitrary constant rank distribution $\mathcal{D}$ on $X$. To any affine distribution $\mathcal{A} \subset \mathcal{D}$, where $\mathcal{A} = f + \mathcal{G}$ such that corank $(\mathcal{G} \subset \mathcal{D}) = 1$ and $f(x) \notin \mathcal{G}(x)$, we attach the control system $\Sigma_{\mathcal{A}}$, given up to orbital feedback.

**Proposition 4.4.2** *Consider a constant rank distribution $\mathcal{D}$ in which we choose an affine distribution $\mathcal{A} = f + \mathcal{G} \subset \mathcal{D}$ such that corank$(\mathcal{G} \subset \mathcal{D}) = 1$ and $f(x) \notin \mathcal{G}(x)$. If the distribution $\mathcal{D}$ satisfies*

(D1) *$\mathcal{D}$ is locally equivalent to the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$;*
(D2) *$\mathcal{C}(\mathcal{D}^{(1)}) = \mathcal{G}$,*

*then the associated control-affine system $\Sigma_{\mathcal{A}}$ is locally orbital feedback equivalent to the t-augmented Brunovsky canonical form $\Lambda_t^{Br}$.*

**Remark**. It seems that the assertion of that theorem may a priori depend on the choice of $f$ but actually it does not. Indeed, if for one vector field $f$, the corresponding system $\Sigma_{\mathcal{A}}$ is orbital feedback linearizable it is so for any $\tilde{f}$ such that $\operatorname{span}\{f\} + \mathcal{G} = \operatorname{span}\{\tilde{f}\} + \mathcal{G} = \mathcal{D}$, where $\mathcal{D}$ is equivalent to the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$. The proof of this result follows immediately form that of Proposition 4.4.5, given in the next subsection.

### 4.4.2 From control-affine systems to distributions

To any control-affine system $\Sigma$ we will attach a distribution spanned by all vector fields corresponding to all controls:

**Definition 4.4.3** Consider a given control-affine system $\Sigma$ defined by (4.2.6). The distribution $\mathcal{D}_\Sigma$ *associated with* $\Sigma$ is defined as

$$\mathcal{D}_\Sigma = \operatorname{span}\{f, g_1, \dots, g_m\}.$$

Let $\mathcal{D}_\Sigma$ and $\tilde{\mathcal{D}}_{\tilde{\Sigma}}$ be the distributions associated to $\Sigma$ and $\tilde{\Sigma}$, respectively, i.e., $\mathcal{D}_\Sigma = \operatorname{span}\{f\} + \mathcal{G}_\Sigma$ and $\tilde{\mathcal{D}}_{\tilde{\Sigma}} = \operatorname{span}\{\tilde{f}\} + \tilde{\mathcal{G}}_{\tilde{\Sigma}}$. Clearly the orbital feedback equivalence of $\Sigma$ and $\tilde{\Sigma}$ implies the equivalence of the distributions $\mathcal{D}_\Sigma$ and $\tilde{\mathcal{D}}_{\tilde{\Sigma}}$. But the former is apparently stronger because it also implies the equivalence of the control distributions $\mathcal{G}_\Sigma$ and $\tilde{\mathcal{G}}_{\tilde{\Sigma}}$.

The pairs $(\mathcal{D}, \mathcal{G})$ and $(\tilde{\mathcal{D}}, \tilde{\mathcal{G}})$ of distributions on $X$ and $\tilde{X}$, respectively, are called equivalent if there exists a diffeomorphism $\varphi : X \to \tilde{X}$ such that

$$\varphi_* \mathcal{D} = \tilde{\mathcal{D}} \quad \text{and} \quad \varphi_* \mathcal{G} = \tilde{\mathcal{G}}.$$

We have the following characterization of orbital feedback equivalence.

**Proposition 4.4.4** *Assume that for control-affine systems $\Sigma$ and $\tilde{\Sigma}$, the control distributions $\mathcal{G}_\Sigma$ and $\tilde{\mathcal{G}}_{\tilde{\Sigma}}$ are of constant rank and $f(x) \notin \mathcal{G}_\Sigma(x)$, for any $x \in X$ and $\tilde{f}(\tilde{x}) \notin \tilde{\mathcal{G}}_{\tilde{\Sigma}}(\tilde{x})$, for any $\tilde{x} \in \tilde{X}$. Then $\Sigma$ and $\tilde{\Sigma}$ are orbital feedback equivalent if and only if the pairs $(\mathcal{D}_\Sigma, \mathcal{G}_\Sigma)$ and $(\tilde{\mathcal{D}}_{\tilde{\Sigma}}, \tilde{\mathcal{G}}_{\tilde{\Sigma}})$ are equivalent.*

**Proof:** As we have already observed, necessity is obvious. To prove sufficiency, denote by $\varphi$ a diffeomorphism that establish the equivalence of $(\mathcal{D}_\Sigma, \mathcal{G}_\Sigma)$ and $(\tilde{\mathcal{D}}_{\tilde{\Sigma}}, \tilde{\mathcal{G}}_{\tilde{\Sigma}})$. The assumption $\varphi_* \mathcal{G}_\Sigma = \tilde{\mathcal{G}}_{\tilde{\Sigma}}$ implies, because of the constant rank of $\mathcal{G}_\Sigma$ and $\tilde{\mathcal{G}}_{\tilde{\Sigma}}$, the existence of functions $\beta_i^j$ such that (4.2.3) holds.

Choose any vector fields $f \in \mathcal{A}$ and $\tilde{f} \in \tilde{\mathcal{A}}$ such that $f(x) \notin \mathcal{G}_\Sigma(x)$ and $\tilde{f}(\tilde{x}) \notin \tilde{\mathcal{G}}_{\tilde{\Sigma}}(\tilde{x})$. It follows that $\varphi_*(\operatorname{span}\{f\}) = \operatorname{span}\{\tilde{f}\} \bmod \tilde{\mathcal{G}}_{\tilde{\Sigma}}$, and thus there exists a nonvanishing function $\gamma$ such that $\varphi_*(\gamma f) = \tilde{f} \bmod \tilde{\mathcal{G}}_{\tilde{\Sigma}}$. Hence (4.2.2) holds which proves feedback equivalence of $\Sigma$ and $\tilde{\Sigma}$. $\square$

This leads to a new, with respect to Theorem 4.3.1, characterization of orbital feedback linearizable systems.

**Proposition 4.4.5** *Consider a control-affine system $\Sigma = (f, \mathcal{G}_\Sigma)$ and let $\mathcal{D}_\Sigma$ be its associated distribution. Then $\Sigma$ is locally orbital feedback equivalent to the $t$-augmented Brunovsky canonical form $\Lambda_t^{Br}$, with all controllability indices equal $n+1$, if and only if $\mathcal{D}_\Sigma$ satisfies the following conditions:*

(C1) $\mathcal{D}_\Sigma$ *is locally equivalent to the Cartan distribution* $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$*;*
(C2) $\mathcal{C}(\mathcal{D}_\Sigma^{(1)}) = \mathcal{G}_\Sigma$.

**Proof:** *Necessity.* Without loss of generality, we can assume that

$$\Sigma: \quad \dot{x} = f(x) + \sum_{i=1}^m g_i(x) u_i$$

is locally orbital feedback equivalent to the $t$-augmented Brunovsky canonical form

$$\dot{\tilde{x}} = f^{cc}(\tilde{x}) + \sum_{i=1}^m g_i^{cc}(\tilde{x}) \tilde{u}_i,$$

where the vector fields $f^{cc}$ and $g_i^{cc}$, for $1 \le i \le m$, are given by (4.4.1). This proves that the distribution $\mathcal{D}_\Sigma$ is locally equivalent to $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$.

*Sufficiency.* Since $\mathcal{D}_\Sigma$ is locally equivalent to $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m) = \tilde{\mathcal{D}}$, there exists a local diffeomorphism $\varphi : X \to \tilde{X}$ such that

$$\varphi_* \mathcal{D}_\Sigma = \tilde{\mathcal{D}} = \operatorname{span}\{f^{cc}, g_1^{cc}, \ldots, g_m^{cc}\}.$$

In other words, there exists an everywhere invertible matrix-valued smooth function $M = (\mu_i^j)$, $0 \le i, j \le m$, such that

$$\begin{aligned} \varphi_*(\mu_0^0 f + \sum_{j=1}^m \mu_0^j g_j) &= f^{cc} \\ \varphi_*(\mu_i^0 f + \sum_{j=1}^m \mu_i^j g_j) &= g_i^{cc}. \end{aligned}$$

From the structure of $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m) = \tilde{\mathcal{D}} = \operatorname{span}\{f^{cc}, g_1^{cc}, \ldots, g_m^{cc}\}$, it is easy to see that $\tilde{\mathcal{G}} = \operatorname{span}\{g_1^{cc}, \ldots, g_m^{cc}\}$ is the characteristic distribution of $\tilde{\mathcal{D}}^{(1)}$. Therefore the characteristic distribution of $\mathcal{D}_\Sigma^{(1)}$ is

$$\mathcal{C}(\mathcal{D}_\Sigma^{(1)}) = \operatorname{span}\{\mu_i^0 f + \sum_{j=1}^m \mu_i^j g_j, \quad 1 \le i \le m\}.$$

The condition (C2) gives $\mathcal{C}(\mathcal{D}_\Sigma^{(1)}) = \mathcal{G}_\Sigma = \operatorname{span}\{g_1, \ldots, g_m\}$ which implies

$$\mu_1^0 = \cdots = \mu_m^0 = 0.$$

Define a smooth matrix $\beta = (\beta_i^j)$ by $\beta_i^j = \mu_i^j$, for $1 \le i, j \le m$. Clearly the function $\mu_0^0$ cannot be zero at $x_0$ and $\beta$ is invertible because the matrix $M(x)$ is invertible everywhere.

Put $\gamma = \mu_0^0$ and $\alpha_j = \mu_0^j$, for $1 \leq j \leq m$, then we get

$$\begin{aligned}
\varphi_*(\gamma f + \sum_{j=1}^{m} \alpha_j g_j) &= f^{cc} \\
\varphi_*(\sum_{j=1}^{m} \beta_i^j g_j) &= g^{cc},
\end{aligned}$$

which implies that $\Sigma$ is orbital feedback equivalent to the $t$-augmented Brunovsky canonical form with time re-scaling function $\gamma = \mu_0^0$. $\square$

All considerations presented in this section, together with Theorem 4.3.1 lead to the following complete characterization of the orbital feedback linearizability of the control-affine system $\Sigma$:

**Theorem 4.4.6** *Consider the control-affine system $\Sigma$, defined by (4.2.6). Let $\mathcal{D}_\Sigma$ be the distribution associated to $\Sigma$ and $\mathcal{G}_\Sigma = \operatorname{span}\{g_1, \ldots, g_m\}$ be the control distribution spanned by the control vectors. Then the following conditions are equivalent, locally around an arbitrary point $x_0 \in X$:*

(i) *$\Sigma$ is orbital feedback equivalent to a $t$-augmented linear system $\Lambda_t$, with all controllability indices equal to $n+1$.*
(ii) *$\Sigma$ satisfies the conditions* (OFL1) – (OFL4) *of Theorem* 4.3.1.
(iii) *The distribution $\mathcal{D}_\Sigma$ associated to $\Sigma$ satisfies the conditions* (C1) – (C2) *of Proposition* 4.4.5.
(iv) *The distribution $\mathcal{D}_\Sigma$ associated to $\Sigma$ satisfies the following conditions:*

(C1)′ $\mathcal{D}_\Sigma^{(n)} = TX$;

(C2)′ $\operatorname{rank} \mathcal{D}_\Sigma^{(n-1)} = nm + 1$ *and there exists an involutive subdistribution $\mathcal{L}_{n-1}$ of $\mathcal{D}_\Sigma^{(n-1)}$ such that* $\operatorname{corank}(\mathcal{L}_{n-1} \subset \mathcal{D}_\Sigma^{(n-1)}) = 1$;

(C3)′ $[\mathcal{G}_\Sigma, \mathcal{D}_\Sigma^{(1)}] \subset \mathcal{D}_\Sigma^{(1)}$;

(C4)′ $\mathcal{D}_\Sigma(x_0) \not\subset \mathcal{L}_{n-1}(x_0)$.

(v) *The distribution $\mathcal{D}_\Sigma$ satisfies the following conditions, for $0 \leq j \leq n$,*

(C1)″ $\operatorname{rank} \mathcal{D}_\Sigma^{(j)} = (j+1)m+1$ *and each element $\mathcal{D}_\Sigma^{(j)}$ contains an involutive subdistribution $\mathcal{L}_j$ such that* $\operatorname{corank}(\mathcal{L}_j \subset \mathcal{D}_\Sigma^{(j)}) = 1$;

(C2)″ $\operatorname{rank}(\mathcal{D}_\Sigma)_j = (j+1)m + 1$, *where $(\mathcal{D}_\Sigma)_j$ stands for the $j$-th element of the Lie flag of $\mathcal{D}_\Sigma$;*

(C3)″ $[\mathcal{G}_\Sigma, \mathcal{D}_\Sigma^{(1)}] \subset \mathcal{D}_\Sigma^{(1)}$.

The equivalence (i) $\Leftrightarrow$ (ii) is just Theorem 4.3.1 whereas (i) $\Leftrightarrow$ (iii) is just Proposition 4.4.5. The conditions (OFL1) – (OFL4) are expressed directly in terms of the control system $\Sigma$. On the other hand, the conditions of (iii), (iv) and (v) are given in terms of the distribution $\mathcal{D}_\Sigma$ associated to $\Sigma$. Equivalence of $\mathcal{D}_\Sigma$ to the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$ can be tested using either the conditions (B1) – (B2) of Theorem 4.4.1 or the conditions (B1)$'$, (B1)$''$ and (B2) of Remark following it. The latter lead to the conditions (C1)$'$ – (C4)$'$ and the former to (C1)$''$ – (C3)$''$.

## 4.5 Examples

In this section we will illustrate our results on orbital feedback linearizability by two examples.

**Example 4.5.1** Consider a control system described on $\mathbb{R}^7$, equipped with the coordinates $(x, y_1, y_2, z_1, z_2, w_1, w_2)$ by

$$\Sigma : \quad \begin{cases} \dot{x} = 1 \\ \dot{y}_1 = z_1 + w_1 \\ \dot{y}_2 = z_2 + y_1 w_2 \\ \dot{z}_1 = w_1 \\ \dot{z}_2 = w_2 \\ \dot{w}_1 = u_1 \\ \dot{w}_2 = u_2 \end{cases}$$

where $(u_1, u_2)^\top \in \mathbb{R}^2$ is the control. This system is not feedback linearizable because it does not satisfy the conditions (FL1) – (FL3) given in Section 4.3 (see the calculation below). We will show that it fulfils however the conditions of Theorem 4.3.1 almost everywhere and thus it is locally orbital feedback linearizable on an open and dense subset of $\mathbb{R}^7$. We have

$$f = \frac{\partial}{\partial x} + (z_1 + w_1)\frac{\partial}{\partial y_1} + (z_2 + y_1 w_2)\frac{\partial}{\partial y_2} + w_1 \frac{\partial}{\partial z_1} + w_2 \frac{\partial}{\partial z_2},$$

$$g_1 = \frac{\partial}{\partial w_1}, \quad g_2 = \frac{\partial}{\partial w_2}.$$

A direct calculation gives

$$\begin{aligned} \mathrm{ad}_f g_1 &= -\frac{\partial}{\partial y_1} - \frac{\partial}{\partial z_1} \\ \mathrm{ad}_f g_2 &= -y_1 \frac{\partial}{\partial y_2} - \frac{\partial}{\partial z_2} \\ \mathrm{ad}_f^2 g_1 &= \frac{\partial}{\partial y_1} + w_2 \frac{\partial}{\partial y_2} \\ \mathrm{ad}_f^2 g_2 &= (1 - z_1 - w_1)\frac{\partial}{\partial y_2}. \end{aligned}$$

Then it is easily seen that the distribution $\mathcal{G}_2 = \mathrm{span}\,\{g_1, g_2, \mathrm{ad}_f g_1, \mathrm{ad}_f g_2\}$ is not involutive and hence $\Sigma$ is not feedback linearizable. On the other hand, the distributions

$$\begin{aligned} \mathcal{G} &= \mathrm{span}\,\{g_1, g_2\}, \\ \mathcal{G}_f^j &= \mathrm{span}\,\{f, g_1, g_2, \ldots, \mathrm{ad}_f^{j-1} g_1, \mathrm{ad}_f^{j-1} g_2\}, \end{aligned}$$

for $1 \leq j \leq 3$, satisfy the conditions (OFL1) – (OFL3) of Theorem 4.3.1 around any point in $\mathbb{R}^7$ such that $1 - z_1 - w_1 \neq 0$. Now we examine the condition (OFL4) of Theorem 4.3.1. Define the differential forms $\omega_1, \omega_2$ by

$$\begin{array}{rcl} \omega^j(h) & = & 0, \quad \text{for any} \quad h \in \mathcal{G}_f^2, \\ \omega^j(ad_f^2 g_i) & = & \delta_i^j, \end{array}$$

for $1 \leq i, j \leq 2$. Solving the above equations, we obtain

$$\begin{array}{rcl} \omega_1 & = & -z_1 \mathrm{d}x + \mathrm{d}y_1 - \mathrm{d}z_1 \\ \omega_2 & = & \dfrac{1}{1 - z_1 - w_1}((z_1 w_2 - z_2)dx - w_2 \mathrm{d}y_1 + \mathrm{d}y_2 + w_2 \mathrm{d}z_1 - y_1 \mathrm{d}z_2). \end{array}$$

Taking into account that $[\mathrm{ad}_f g_1, \mathrm{ad}_f g_2] = \dfrac{\partial}{\partial y_2}$, we get

$$\mathrm{T}_{1,2}^{1,1} = \omega_1([\mathrm{ad}_f g_1, \mathrm{ad}_f g_2]) = 0,$$
$$\mathrm{T}_{1,2}^{2,1} = \omega_2([\mathrm{ad}_f g_1, \mathrm{ad}_f g_2]) = \frac{1}{1 - z_1 - w_1}.$$

Notice that for the system $\Sigma$, the conditions (i) and (ii) of (OFL4) are trivially satisfied and thus it remains to verify the condition (iii) of (OFL4). That is, the distribution

$$\mathcal{B} = \mathrm{span}\,\{g_1, g_2, \mathrm{ad}_f g_1 + b_1^1 f, \mathrm{ad}_f g_2 + b_2^1 f\}$$

should be involutive, where the functions $b_1^1$ and $b_2^1$ are given, respectively, by

$$b_1^1 = \mathrm{T}_{2,1}^{2,1} = -\mathrm{T}_{1,2}^{2,1} = \frac{1}{z_1 + w_1 - 1}$$
$$b_2^1 = \mathrm{T}_{1,2}^{1,1} = 0.$$

Substituting $b_1^1$ and $b_2^1$ into $\mathcal{B}$, we get

$$\begin{array}{rcl} \mathcal{B} & = & \mathrm{span}\,\{g_1, g_2, \mathrm{ad}_f g_1 + b_1^1 f, \mathrm{ad}_f g_2 + b_2^1 f\} \\ & = & \mathrm{span}\,\{g_1, g_2, f + (z_1 + w_1 - 1)\mathrm{ad}_f g_1, \mathrm{ad}_f g_2\} \\ & = & \mathrm{span}\,\{g_1, g_2, h_1, h_2\}, \end{array}$$

where

$$h_1 = f + (z_1 + w_1 - 1)\mathrm{ad}_f g_1 = \frac{\partial}{\partial x} + \frac{\partial}{\partial y_1} + (z_2 + w_2 y_1)\frac{\partial}{\partial y_2} + (1 - z_1)\frac{\partial}{\partial z_1} + w_2 \frac{\partial}{\partial z_2}$$
$$h_2 = \mathrm{ad}_f g_2 = -y_1 \frac{\partial}{\partial y_2} - \frac{\partial}{\partial z_2}.$$

Computing the Lie brackets

$$[g_1, g_2] = 0, [g_1, h_1] = 0, [g_1, h_2] = 0,$$

$$[g_2, h_1] = -h_2, [g_2, h_2] = 0, [h_1, h_2] = 0,$$

we conclude that the distribution $\mathcal{B}$ is indeed involutive and therefore the system $\Sigma$ is orbital feedback linearizable. The time re-scaling function $\gamma$ is constructible. To this end, choose three functions $\phi_1 = x - y_1, \phi_2 = y_1 z_2 - y_2, \phi_3 = e^x(z_1 - 1)$ so that they satisfy

$$\mathcal{B}^\perp = \text{span}\,\{\mathrm{d}\phi_1, \mathrm{d}\phi_2, \mathrm{d}\phi_3\}.$$

Since $L_f\phi_1 = (1 - z_1 - w_1) \neq 0$ out of the set $\{z_1 + w_1 = 1\}$, a time re-scaling function $\gamma$ can be taken as

$$\gamma = \frac{1}{1 - z_1 - w_1}.$$

By a direct calculation we can verify that the system $\Sigma$ with respect to the new time scale $\mathrm{d}t = \gamma \mathrm{d}\tau$, i.e., the system $\Sigma_1$, which is $\Sigma$ with $f$ replaced by $\gamma f$, satisfies the feedback linearizability conditions (FL1) – (FL3) (see Remark 3 following Theorem 4.3.1). Notice that we can also choose $\gamma_2 = \dfrac{1}{L_f\phi_2}$, in $\{1 - z_1 - w_1 \neq 0\} \cap \{z_2 \neq 0\}$ or $\gamma_3 = \dfrac{1}{L_f\phi_3}$, in $\{1 - z_1 - w_1 \neq 0\}$ and the corresponding control systems $\Sigma_2$, with $f$ replaced by $\gamma_2 f$, as well as $\Sigma_3$, with $f$ replaced by $\gamma_3 f$, become feedback linearizable. This illustrates an interesting phenomenon that the linearizing time re-scaling function is not unique and may lead to different control systems (in fact, $\Sigma_1$, $\Sigma_2$ and $\Sigma_3$ are not mutually locally equivalent via a diffeomorphism $\phi$ in the state space) which are, of course, feedback equivalent to each other.

**Example 4.5.2** (Rigid bar moving in $\mathbb{R}^3$)
Consider a rigid bar moving in $\mathbb{R}^3$ such that the instantaneous velocity of the bar is parallel to its direction. Let $(x, y, z) \in \mathbb{R}^3$ denotes the coordinate position of the source point of the bar, $\varphi$ denotes the angle between the bar axis and the plane $XOY$ and $\theta$ denotes the angle between the $x$-axis and the projection of the bar on the $XOY$-plane. We assume that we control both angular velocities and that the bar moves forward only with a velocity of norm one.

$$\Sigma_{\text{bar}} : \quad \begin{cases} \dot{x} = \cos\varphi\cos\theta \\ \dot{y} = \cos\varphi\sin\theta \\ \dot{z} = \sin\varphi \\ \dot{\theta} = u_1 \\ \dot{\varphi} = u_2 \end{cases}$$

where $\theta \in S^1$ and $-\frac{\pi}{2} < \varphi < \frac{\pi}{2}$. Note that this model is not defined at the point $\varphi = \pm\frac{\pi}{2}$. Therefore the state is $(x, y, z, \varphi, \theta) \in X = \mathbb{R}^3 \times (-\frac{\pi}{2}, \frac{\pi}{2}) \times S^1$ and the drift and the control vector fields are, respectively,

$$f = \cos\varphi\cos\theta\frac{\partial}{\partial x} + \cos\varphi\sin\theta\frac{\partial}{\partial y} + \sin\varphi\frac{\partial}{\partial z},$$

$$g_1 = \frac{\partial}{\partial\theta}, \quad g_2 = \frac{\partial}{\partial\varphi}.$$

A direct computation gives

$$\begin{aligned} \mathrm{ad}_f g_1 &= \cos\varphi\sin\theta\frac{\partial}{\partial x} - \cos\varphi\cos\theta\frac{\partial}{\partial y} \\ \mathrm{ad}_f g_2 &= \sin\varphi\cos\theta\frac{\partial}{\partial x} + \sin\varphi\sin\theta\frac{\partial}{\partial y} - \cos\varphi\frac{\partial}{\partial z}. \end{aligned}$$

It is easy to find that the distribution $\mathcal{G}_2 = \mathrm{span}\,\{g_1, g_2, \mathrm{ad}_f g_1, \mathrm{ad}_f g_2\}$ is not involutive and thus $\Sigma_{\mathrm{bar}}$ is not feedback linearizable because it does not satisfy the conditions (FL1) – (FL3) given in Section 4.3. However it is locally orbital feedback linearizable. Clearly, the distributions

$$\begin{aligned} \mathcal{G} &= \mathrm{span}\,\{g_1, g_2\}, \\ \mathcal{G}_f^j &= \mathrm{span}\,\{f, g_1, g_2, \ldots, \mathrm{ad}_f^{j-1} g_1, \mathrm{ad}_f^{j-1} g_2\}, \end{aligned}$$

for $1 \leq j \leq 2$, satisfy the conditions (OFL1) – (OFL3) of Theorem 4.3.1 around any point in $X$. Moreover the condition (OPL4) is trivially satisfied because all the functions $\mathrm{T}_{i,j}^{k,l}$ vanish identically. Indeed, recall that $m = 2$ and $n = 1$ so $l = 0$ since $l \leq n - 1$. Then we have,

$$\begin{aligned} \mathrm{T}_{i,j}^{k,l} &= \omega^k([\mathrm{ad}_f^{n-1} g_i, \mathrm{ad}_f^{n-1} g_j]) \\ &= \omega^k([g_i, g_j]) \\ &= 0 \end{aligned}$$

for $1 \leq i, j, k \leq 2$, since $[g_1, g_2] = [\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\varphi}] \equiv 0$. Consequently, the involutive distribution $\mathcal{B}$ coincides with the distribution $\mathcal{G} = \mathrm{span}\,\{g_1, g_2\}$. In order to find a suitable time re-scaling function, we choose $\phi_1 = x, \phi_2 = y, \phi_3 = z$ such that

$$\mathcal{B}^\perp = \mathrm{span}\,\{\mathrm{d}\phi_1, \mathrm{d}\phi_2, \mathrm{d}\phi_3\}.$$

If $\theta \neq \pm\frac{\pi}{2}$, then $L_f\phi_1 = L_f x = \cos\varphi\cos\theta > 0$. Therefore a time re-scaling function is given by

$$\gamma = \frac{1}{\cos\varphi\cos\theta}.$$

Multiplying the right hand side of $\Sigma_{\mathrm{bar}}$ by $\frac{1}{\cos\varphi\cos\theta}$ and putting $v = \tan\theta$, $w = \frac{\tan\varphi}{\cos\theta}$, we obtain finally the linear system by a suitable feedback

$$\begin{cases} \dot{x} = 1 \\ \dot{y} = v \\ \dot{z} = w \\ \dot{v} = \tilde{u}_1 \\ \dot{w} = \tilde{u}_2. \end{cases}$$

In the case $\theta = \pm\frac{\pi}{2}$, the orbital feedback linearization can be realized in a similar way with the time re-scaling function given by

$$\gamma = \frac{1}{\cos\varphi\sin\theta},$$

since $L_f\phi_2 = L_f y = \cos\varphi \sin\theta \neq 0$. As we have explained in Remark 2 following Theorem 4.3.1, the choice of the time re-scaling function $\gamma$ is not unique. In fact, we can choose any $\varphi$ such that $\mathrm{d}\varphi \in \mathcal{B}^{\perp}$, that is any $\varphi = \varphi(x, y, z)$, such that $L_f\varphi(x_0) \neq 0$ and re-scale the time via $\gamma = \dfrac{1}{L_f\varphi}$.

## 4.6 Proof of Theorem 4.4.6

Clearly, both Theorem 4.3.1 and Proposition 4.4.5 are contained in Theorem 4.4.6 which gives a complete characterization of the orbital feedback linearization of the system $\Sigma$ defined by (4.2.6). So it is enough to prove Theorem 4.4.6 which we will do in this section. The proof will be based on some lemmata given in Subsection 4.6.1.

### 4.6.1 Useful results

**Lemma 4.6.1** *If the system $\Sigma$ satisfies the conditions* (OFL1) – (OFL3) *of* Theorem 4.3.1*, then we have*

$$[\mathcal{G}, \mathcal{G}_f^j] \subset \mathcal{G}_f^j, \quad 3 \leq j \leq n+1.$$

**Proof:** We proceed by induction. If $j = 3$ we have

$$\mathcal{G}_f^3 = \mathrm{span}\,\{f, g_i, \mathrm{ad}_f g_i, \mathrm{ad}_f^2 g_i, \quad 1 \leq i \leq m\}.$$

Given an arbitrary $g \in \mathcal{G}$, in order to show that $[g, \mathcal{G}_f^3] \in \mathcal{G}_f^3$, it is enough to prove

$$[g, \mathrm{ad}_f^2 g_i] \in \mathcal{G}_f^3, \qquad 1 \leq i \leq m.$$

Computing that Lie bracket, we have, by the Jacobi identity,

$$\begin{aligned} [g, \mathrm{ad}_f^2 g_i] &= [g, [f, \mathrm{ad}_f g_i]] \\ &= [[g, f], \mathrm{ad}_f g_i] + [f, [g, \mathrm{ad}_f g_i]] \\ &= -[\mathrm{ad}_f g, \mathrm{ad}_f g_i] + [f, [g, \mathrm{ad}_f g_i]]. \end{aligned}$$

The condition $[\mathcal{G}, \mathcal{G}_f^2] \subset \mathcal{G}_f^2$ implies that $[g, \mathrm{ad}_f g_i] \in \mathcal{G}_f^2$. Moreover, since the vector fields $\mathrm{ad}_f g$, $\mathrm{ad}_f g_i$ and $f$ belong to $\mathcal{G}_f^2$, it follows from the condition (OFL2) that

$$[\mathrm{ad}_f g, \mathrm{ad}_f g_i] \in \mathcal{G}_f^3, \quad [f, [g, \mathrm{ad}_f g_i]] \in \mathcal{G}_f^3,$$

and thus $[g, \mathrm{ad}_f^2 g_i] \in \mathcal{G}_f^3$.

Assume now that the statement is true for $j = k$, i.e., $[\mathcal{G}, \mathcal{G}_f^k] \subset \mathcal{G}_f^k$. Consider

$$\mathcal{G}_f^{k+1} \;=\; \mathcal{G}_f^k \oplus \mathrm{span}\,\{\mathrm{ad}_f^k g_i, \quad 1 \leq i \leq m\}$$

$$= \quad \mathrm{span}\,\{f, g_i, \mathrm{ad}_f g_i, \ldots, \mathrm{ad}_f^{k-1} g_i, \mathrm{ad}_f^k g_i, \quad 1 \le i \le m\}.$$

For any $g \in \mathcal{G}$, by the Jacobi identity we have, for $1 \le i \le m$,

$$\begin{aligned}
[g, \mathrm{ad}_f^k g_i] &= [g, [f, \mathrm{ad}_f^{k-1} g_i]] \\
&= [[g, f], \mathrm{ad}_f^{k-1} g_i] + [f, [g, \mathrm{ad}_f^{k-1} g_i]] \\
&= -[\mathrm{ad}_f g, \mathrm{ad}_f^{k-1} g_i] + [f, [g, \mathrm{ad}_f^{k-1} g_i]].
\end{aligned}$$

The induction assumption $[\mathcal{G}, \mathcal{G}_f^k] \subset \mathcal{G}_f^k$ implies that $[g, \mathrm{ad}_f^{k-1} g_i] \in \mathcal{G}_f^k$, for $1 \le i \le m$. Moreover, since $\mathrm{ad}_f g$, $\mathrm{ad}_f^{k-1} g_i$ and $f$ belong to $\mathcal{G}_f^k$, the condition (OFL2) implies that

$$[\mathrm{ad}_f g, \mathrm{ad}_f^{k-1} g_i] \in \mathcal{G}_f^{k+1}, \quad [f, [g, \mathrm{ad}_f^{k-1} g_i]] \in \mathcal{G}_f^{k+1}.$$

Thus $[g, \mathrm{ad}_f^k g_i] \in \mathcal{G}_f^{k+1}$, for $1 \le i \le m$, which yields $[\mathcal{G}, \mathcal{G}_f^{k+1}] \subset \mathcal{G}_f^{k+1}$ and the lemma follows by an induction argument. □

Recall that in Section 4.3 we have defined, under (OFL1), the differential forms $\omega^1, \ldots, \omega^m$ and with their help, the functions $T_{i,j}^{k,l}$ were also well defined for any $0 \le l \le n-1$ and any $1 \le i, j, k \le m$ such that $i \ne j$ in the case $l = n-1$.

**Lemma 4.6.2** *If the control-affine system* $\Sigma$, *defined by* (4.2.6) *satisfies the conditions* (OFL1) – (OFL3) *of* Theorem 4.3.1 *and, moreover, there exists a subdistribution* $\mathcal{B}$ *of* $\mathcal{G}_f^n$ *such that* $\mathrm{corank}\,(\mathcal{B} \subset \mathcal{G}_f^n) = 1$ *and* $[\mathcal{B}, \mathcal{B}] \subset \mathcal{G}_f^n$, *then*

$$\mathcal{B} = \mathcal{G} \oplus \mathcal{H},$$

*where* $\mathcal{G} = \mathrm{span}\,\{g_1, \ldots, g_m\}$ *and* $\mathcal{H} = \mathrm{span}\,\{\mathrm{ad}_f^j g_i + b_i^j f, \quad 1 \le i \le m, 1 \le j \le n-1\}$ *for some suitable* $C^\infty$*-functions* $b_i^j$ *defined on* $X$.

**Proof:** To begin with, we will show that $\mathcal{G} = \mathrm{span}\,\{g_1, \ldots, g_m\} \subset \mathcal{B}$ locally around $x_0$. Assume that there exist $g \in \mathcal{G}$ such that $g(x_0) \notin \mathcal{B}(x_0)$, since $\mathcal{B}$ is of constant rank, we have $g(x) \notin \mathcal{B}(x)$, for any $x$ in a neighborhood $\mathcal{O}$ of $(x_0)$. Then the the condition $\mathrm{corank}\,(\mathcal{B} \subset \mathcal{G}_f^n) = 1$ implies that $\mathcal{G}_f^n = \mathcal{B} \oplus \mathrm{span}\,\{g\}$ in $\mathcal{O}$. Moreover, Lemma 4.6.1 shows that $g$ is a characteristic vector field of $\mathcal{G}_f^n$ and therefore, locally around $x_0$,

$$[\mathcal{G}_f^n, \mathcal{G}_f^n] = [\mathcal{B}, \mathcal{B}] + [\mathcal{B}, g] \subset \mathcal{G}_f^n,$$

which implies that $\mathcal{G}_f^n$ is involutive. This gives a contradiction and thus $\mathcal{G} \subset \mathcal{B}$.

Since $\mathrm{rank}\,\mathcal{B} = \mathrm{rank}\,\mathcal{G}_f^n - 1 = nm$ , there exists a rank $(n-1)m$ distribution

$$\mathcal{H} = \mathrm{span}\,\{h_i^1, \ldots, h_i^{n-1}, 1 \le i \le m\} \subset \mathcal{B}$$

such that $\mathcal{B} = \mathcal{G} \oplus \mathcal{H}$. Forming the following row-vectors whose elements are vector fields,

$$\mathrm{H} = (h_1^1, \ldots, h_m^1, \ldots, h_1^{n-1}, \ldots, h_m^{n-1}),$$

$$\mathrm{A} = (\mathrm{ad}_f g_i, \ldots, \mathrm{ad}_f^{n-1} g_i, f, \quad 1 \le i \le m),$$

and taking into account the structure of $\mathcal{G}$, we obtain that

$$\mathrm{H} = \mathrm{A} \cdot \mathrm{B},$$

where B is a matrix-valued function of rank $(n-1)m$ with $(n-1)m+1$ rows and $(n-1)m$ columns. Let $\mathrm{B}_1$ be the matrix formed by the first $(n-1)m$ rows,

$$\mathrm{B} = \begin{pmatrix} \mathrm{B}_1 \\ * \end{pmatrix}.$$

Assume for a moment $\operatorname{rank} \mathrm{B}_1 = (n-1)m$ and then putting

$$\tilde{\mathrm{B}} = \mathrm{B} \cdot \mathrm{B}_1^{-1} = \begin{pmatrix} \mathrm{Id} \\ * \end{pmatrix}$$

we have $\operatorname{Im} \tilde{\mathrm{B}} = \operatorname{Im} \mathrm{B}$. Now, define a new row vector

$$\begin{aligned} \tilde{\mathrm{H}} &= \mathrm{A} \cdot \tilde{\mathrm{B}} \\ &= (\mathrm{ad}_f g_i, \ldots, \mathrm{ad}_f^{n-1} g_i, f) \begin{pmatrix} \mathrm{Id} \\ * \end{pmatrix} \\ &= (\mathrm{ad}_f^1 g_i + b_i^1 f, \ldots, \mathrm{ad}_f^{n-1} g_i + b_i^{n-1} f, \quad 1 \le i \le m), \end{aligned}$$

where $(b_1^1, \ldots, b_m^1, \ldots, b_1^{n-1}, \ldots, b_m^{n-1})$ coincides with the $((n-1)m+1)$-th row of $\tilde{\mathrm{B}}$. Let $\tilde{\mathcal{H}}$ be the distribution spanned by all the vector fields of $\tilde{\mathrm{H}}$. Then the relation $\operatorname{Im} \tilde{\mathrm{B}} = \operatorname{Im} \mathrm{B}$ implies that

$$\mathcal{H} = \tilde{\mathcal{H}} = \operatorname{span} \{\mathrm{ad}_f^1 g_i + b_i^1 f, \ldots, \mathrm{ad}_f^{n-1} g_i + b_i^{n-1} f, \quad 1 \le i \le m\},$$

and hence the statement of Lemma 4.6.2 follows. In order to finish the proof, it remains to show that $\operatorname{rank} \mathrm{B}_1$ can not be less than $(n-1)m$.

Suppose that $\operatorname{rank} \mathrm{B}_1(x_0) < (n-1)m$, then clearly we have $\operatorname{rank} \mathrm{B}_1(x_0) = (n-1)m - 1$. We can assume, for simplicity, that the matrix $\mathrm{B}_2$ obtained by neglecting the first row of B is invertible, i.e.,

$$\mathrm{B} = \begin{pmatrix} * \\ \mathrm{B}_2 \end{pmatrix}, \quad \operatorname{rank} \mathrm{B}_2 = (n-1)m.$$

Let $\bar{\mathrm{B}}$ be the matrix defined by

$$\bar{\mathrm{B}} = \mathrm{B} \cdot \mathrm{B}_2^{-1} = \begin{pmatrix} * \\ \mathrm{Id} \end{pmatrix},$$

and clearly we have $\operatorname{Im}\bar{\mathrm{B}} = \operatorname{Im}\mathrm{B}$. Let $\bar{h}_2^1, \dots, \bar{h}_m^1, \bar{h}_i^2, \dots, \bar{h}_i^{n-1}, \bar{h}_0$ , for $1 \le i \le m$, be the vector fields given by

$$\begin{aligned}(\bar{h}_2^1, \dots, \bar{h}_m^1, \bar{h}_i^2, \dots, \bar{h}_i^{n-1}, \bar{h}_0) &= (\mathrm{ad}_f g_i, \dots, \mathrm{ad}_f^{n-1} g_i, f)\bar{B} \\ &= (\mathrm{ad}_f g_i, \dots, \mathrm{ad}_f^{n-1} g_i, f)\begin{pmatrix} * \\ \mathrm{Id} \end{pmatrix}.\end{aligned}$$

A straightforward calculation shows that

$$\begin{aligned}&\bar{h}_i^j = \mathrm{ad}_f^j g_i + \bar{b}_i^j \cdot \mathrm{ad}_f g_1, \quad (i,j) \ne (1,1) \\ &\bar{h}_0 = f + \bar{b}_0 \cdot \mathrm{ad}_f g_1,\end{aligned}$$

where $(\bar{b}_2^1, \dots, \bar{b}_m^1, \bar{b}_1^2, \dots, \bar{b}_m^2, \dots, \bar{b}_1^{n-1}, \dots, \bar{b}_m^{n-1}, \bar{b}_0)$ coincides with the first row of $\bar{\mathrm{B}}$. Now we discuss the value of the function $\bar{b}_0$ at point $x_0$. If $\bar{b}_0(x_0) \ne 0$, notice that $\bar{b}_0$ is just the element that lies in the first row and $(n-1)m$-th column of the matrix $\bar{\mathrm{B}}$ and then, clearly, the first $(n-1)m$ rows of $\bar{\mathrm{B}}$ are linearly independent. In other words,

$$\operatorname{rank}\mathrm{B}_1\mathrm{B}_2^{-1} = (n-1)m,$$

which contradicts the assumption that $\operatorname{rank}\mathrm{B}_1 = (n-1)m - 1$. If $\bar{b}_0(x_0) = 0$, let us consider the vector fields $\bar{h}_0$ and $\bar{h}_i^{n-1}$, for $1 \le i \le m$. On one hand, since $\bar{h}_0 \subset \mathcal{H} \subset \mathcal{B}$ and $\bar{h}_i^{n-1} \subset \mathcal{H} \subset \mathcal{B}$, we have, for $1 \le i \le m$

$$h(x_0) = [\bar{h}_0, \bar{h}_i^{n-1}](x_0) \subset [\mathcal{B}, \mathcal{B}](x_0) \subset \mathcal{G}_f^n(x_0). \tag{4.6.1}$$

On the other hand, calculating the Lie bracket of $\bar{h}_0$ and $\bar{h}_i^{n-1}$, $1 \le i \le m$, we have

$$\begin{aligned}h(x_0) &= [\bar{h}_0, \bar{h}_i^{n-1}](x_0) \\ &= [f + \bar{b}_0 \cdot \mathrm{ad}_f g_1,\ \mathrm{ad}_f^{n-1} g_i + \bar{b}_i^{n-1} \cdot \mathrm{ad}_f g_1](x_0) \\ &= \mathrm{ad}_f^n g_i(x_0) + \bar{b}_i^{n-1}(x_0) \cdot \mathrm{ad}_f^2 g_1(x_0) + (L_f \bar{b}_i^{n-1})(x_0) \cdot \mathrm{ad}_f g_1(x_0) \\ &\quad + \bar{b}_0(x_0)[\mathrm{ad}_f g_1, \mathrm{ad}_f^{n-1} g_i](x_0) - (L_{\mathrm{ad}_f^{n-1} g_i}\bar{b}_0)(x_0) \cdot \mathrm{ad}_f g_1(x_0) \\ &\quad + (\bar{b}_0(L_{\mathrm{ad}_f g_1}\bar{b}_i^{n-1}) - \bar{b}_i^{n-1}(L_{\mathrm{ad}_f g_1}\bar{b}_0))(x_0) \cdot \mathrm{ad}_f g_1(x_0) \\ &= \mathrm{ad}_f^n g_i(x_0) + \bar{b}_i^{n-1}(x_0) \cdot \mathrm{ad}_f^2 g_1(x_0) + \eta(x_0) \cdot \mathrm{ad}_f g_1(x_0),\end{aligned}$$

where $\eta = L_f \bar{b}_i^{n-1} - L_{\mathrm{ad}_f^{n-1} g_i}\bar{b}_0 + \bar{b}_0(L_{\mathrm{ad}_f g_1}\bar{b}_i^{n-1}) - \bar{b}_i^{n-1}(L_{\mathrm{ad}_f g_1}\bar{b}_0)$. By the condition (OFL1), the vectors $\mathrm{ad}_f g_1(x_0)$, $\mathrm{ad}_f^2 g_1(x_0)$ and $\mathrm{ad}_f^n g_i(x_0)$ are linearly independent. Then the fact $\mathrm{ad}_f^n g_i \notin \mathcal{G}_f^n$ implies that $h(x_0) \notin \mathcal{G}_f^n(x_0)$ which contradicts the relation (4.6.1). Therefore $\operatorname{rank}\mathrm{B}_1 = (n-1)m$ and Lemma 4.6.2 holds. $\square$

The following version of a result of Bryant [7] was given by Pasillas-Lépine and Respondek in [51], [54] in order to study the Cartan distribution $\mathcal{CC}^n(\mathbb{R}, \mathbb{R}^m)$.

**Lemma 4.6.3** *Let $\mathcal{D}$ be a distribution such that $\mathcal{D}$ and $\mathcal{D}^{(1)}$ have constant rank $d_0$ and $d_1$, respectively. Put $r_0 = d_1 - d_0$ and assume that $\mathcal{D}$ contains a subdistribution $\mathcal{B} \subset \mathcal{D}$ that has constant corank one in $\mathcal{D}$ and satisfies $[\mathcal{B}, \mathcal{B}] \subset \mathcal{D}$.*

(i) *If $r_0 \geq 2$, then $\mathcal{B}$ is unique.*
(ii) *If $r_0 \geq 3$, then $\mathcal{B}$ is involutive.*

In the case of $\mathcal{D} = \mathcal{G}_f^n$ the existence of $\mathcal{B}$ as above is described by the following

**Lemma 4.6.4** *If the control-affine system $\Sigma$, defined by* (4.2.6)*, for $m \geq 2$, satisfies the conditions* (OFL1) – (OFL3) *of Theorem 4.3.1, then the following conditions are equivalent:*

(i) *There exists a subdistribution $\mathcal{B}$ of $\mathcal{G}_f^n$ such that* $\operatorname{corank}(\mathcal{B} \subset \mathcal{G}_f^n) = 1$ *and* $[\mathcal{B}, \mathcal{B}] \subset \mathcal{G}_f^n$*;*
(ii) *The functions $\mathrm{T}_{i,j}^{k,l}$, defined by (4.3.2), satisfy the following conditions:*

$$\text{(OFL4)}^{\text{(i)}} \quad \mathrm{T}_{i,j}^{k,l} = 0, \text{ for } \begin{cases} 1 \leq k \neq i \leq m, \quad 1 \leq j \leq m, \quad \text{if} \quad l < n-1 \\ 1 \leq i \neq k \neq j \leq m, \quad \text{if} \quad l = n-1 \end{cases},$$

$$\text{(OFL4)}^{\text{(ii)}} \quad \mathrm{T}_{i,j}^{i,l} = T_{k,j}^{k,l}, \text{ for } 1 \leq l \leq n-1,\ 1 \leq i, j, k \leq m \text{ s.t } i \neq j,\ k \neq j \text{ if } l = n-1$$

**Remark**. The condition (OFL1) assumes that the distributions $\mathcal{G}_f^{n+1}$ and $\mathcal{G}_f^n$ are of constant rank and $\operatorname{rank} \mathcal{G}_f^{n+1} - \operatorname{rank} \mathcal{G}_f^n = m$. Then if $m \geq 3$, by Lemma 4.6.3, the subdistribution $\mathcal{B} \subset \mathcal{G}_f^n$ given in item (i) of Lemma 4.6.4 is involutive and unique. Therefore Lemma 4.6.4 implies that if $\Sigma$, for $m \geq 3$, satisfies the conditions (OFL1)−(OFL3) of Theorem 4.3.1, then the condition (OFL4) holds if and only if there exists an involutive subdistribution $\mathcal{B}$ of $\mathcal{G}_f^n$ which satisfies $\operatorname{corank}(\mathcal{B} \subset \mathcal{G}_f^n) = 1$.

**Proof:** (i) $\Longrightarrow$ (ii): Applying Lemma 4.6.2 we have

$$\begin{aligned} \mathcal{B} &= \mathcal{G} \oplus \mathcal{H} \\ &= \operatorname{span}\{g_1, \ldots, g_m\} \oplus \operatorname{span}\{h_i^j, \quad 1 \leq i \leq m, 1 \leq j \leq n-1\} \end{aligned}$$

with $h_i^j = \operatorname{ad}_f^j g_i + b_i^j f$ where $b_i^j$ are $C^\infty$-functions defined on $X$. Compute the following Lie brackets, for any $1 \leq i, j \leq m$ and $1 \leq l \leq n-1$,

$$\begin{aligned} [h_i^{n-1}, h_j^l] &= [\operatorname{ad}_f^{n-1} g_i + b_i^{n-1} f,\ \operatorname{ad}_f^l g_j + b_j^l f] \\ &= [\operatorname{ad}_f^{n-1} g_i, \operatorname{ad}_f^l g_j] - b_j^l \cdot \operatorname{ad}_f^n g_i + (L_{\operatorname{ad}_f^{n-1} g_i} b_j^l) \cdot f \\ &\quad + b_i^{n-1} \cdot \operatorname{ad}_f^{l+1} g_j - (L_{\operatorname{ad}_f^l g_j} b_i^{n-1}) \cdot f + (b_i^{n-1} L_f b_j^l - b_j^l L_{b_i^{n-1}}) f \\ &= [\operatorname{ad}_f^{n-1} g_i, \operatorname{ad}_f^l g_j] - b_j^l \cdot \operatorname{ad}_f^n g_i + b_i^{n-1} \cdot \operatorname{ad}_f^{l+1} g_j + \eta \cdot f \end{aligned} \tag{4.6.2}$$

where $\eta = L_{\operatorname{ad}_f^{n-1} g_i} b_j^l - L_{\operatorname{ad}_f^l g_j} b_i^{n-1} + b_i^{n-1} L_f b_j^l - b_j^l L_{b_i^{n-1}}$. On one hand, the definition (4.3.1) of the differential forms $\omega^1, \ldots, \omega^m$ implies that

$$\omega^k([h_i^{n-1}, h_j^l]) \equiv 0, \quad 1 \leq k \leq m, \quad 1 \leq l \leq n-1,$$

since $[h_i^{n-1}, h_j^l] \in [\mathcal{B}, \mathcal{B}] \subset \mathcal{G}_f^n$. On the other hand, we have

$$\begin{aligned}\omega^k([h_i^{n-1}, h_j^l]) &= \omega^k([\mathrm{ad}_f^{n-1} g_i, \mathrm{ad}_f^l g_j]) - b_j^l \cdot \omega^k(\mathrm{ad}_f^n g_i) \\ &\quad + b_i^{n-1} \cdot \omega^k(\mathrm{ad}_f^{l+1} g_j) + \eta \cdot \omega^k(f) \\ &= T_{i,j}^{k,l} - b_j^l \cdot \omega^k(\mathrm{ad}_f^n g_i) + b_i^{n-1} \cdot \omega^k(\mathrm{ad}_f^{l+1} g_j).\end{aligned}$$

It follows that

$$T_{i,j}^{k,l} = b_j^l \cdot \omega^k(\mathrm{ad}_f^n g_i) - b_i^{n-1} \cdot \omega^k(\mathrm{ad}_f^{l+1} g_j). \tag{4.6.3}$$

We will consider separately the case $l < n-1$ and $l = n-1$.

If $l < n-1$, then obviously $\mathrm{ad}_f^{l+1} g_j \in \mathcal{G}_f^n$ and so $\omega^k(\mathrm{ad}_f^{l+1} g_j) = 0$, for any $1 \le k \le m$. If $k \ne i$, by the definition of $\omega^k$, we have $\omega^k(\mathrm{ad}_f^n g_i) = 0$. Thus the equation (4.6.3) gives

$$\omega^k([h_i^{n-1}, h_j^l]) = \mathrm{T}_{i,j}^{k,l} = 0.$$

If $k = i$, since $\omega^k(\mathrm{ad}_f^n g_i) = \omega^k(\mathrm{ad}_f^n g_k) = 1$, then the equation (4.6.3) gives

$$\omega^k([h_k^{n-1}, h_j^l]) = \mathrm{T}_{k,j}^{k,l} - b_j^l = 0,$$

and thus

$$T_{k,j}^{k,l} = b_j^l,$$

for any $1 \le k \le m$ and hence we conclude that

$$\mathrm{T}_{r,j}^{r,l} = b_j^l = \mathrm{T}_{k,j}^{k,l}, \quad \text{for any} \quad 1 \le r, j, k \le m. \tag{4.6.4}$$

In the case $l = n-1$, if $i \ne k \ne j$ then clearly $\omega^k(\mathrm{ad}_f^n g_i) = \omega^k(\mathrm{ad}_f^n g_j) = 0$ and from the equation (4.6.3) we get

$$\mathrm{T}_{i,j}^{k,n-1} = 0.$$

If $k = i$ and $i \ne j$, then clearly $\omega^k(\mathrm{ad}_f^n g_i) = \omega^k(\mathrm{ad}_f^n g_k) = 1$ and $\omega^k(\mathrm{ad}_f^n g_j) = 0$. Therefore it follows from the equation (4.6.3) that

$$\mathrm{T}_{i,j}^{k,n-1} = \mathrm{T}_{k,j}^{k,n-1} = b_j^{n-1}. \tag{4.6.5}$$

Since the equation (4.6.5) does not depend on $k$, we get

$$\mathrm{T}_{r,j}^{r,n-1} = b_j^{n-1} = \mathrm{T}_{k,j}^{k,n-1}, \ \ 1 \le j \ne r, j \ne k \le m. \tag{4.6.6}$$

If $k = j$ and $i \ne j$, then $\mathrm{T}_{i,j}^{k,n-1} = \mathrm{T}_{i,k}^{k,n-1} = -\mathrm{T}_{k,i}^{k,n-1}$ and the result follows by (4.6.5).

The relations (4.6.4) and (4.6.6) show that we always have

$$\mathrm{T}_{i,j}^{i,l} = \mathrm{T}_{k,j}^{k,l},$$

for any $1 \leq l \leq n-1$ and any $1 \leq i, j, k \leq m$ such that $j \neq i$ and $j \neq k$ in the case $l = n-1$.

(ii) $\Longrightarrow$ (i): Assume that a system $\Sigma$ satisfies (OFL1) – (OFL3) and that the functions $\mathrm{T}_{i,j}^{k,l}$, $1 \leq i, j, k \leq m$, $1 \leq l \leq n-1$, satisfy the conditions (OFL4)(i) – (OFL4)(ii). For any $1 \leq i \leq m$ and $1 \leq l \leq n-1$, define the functions

$$b_i^l = \mathrm{T}_{k,i}^{k,l}, \tag{4.6.7}$$

where we take any $1 \leq k \leq m$ if $l < n-1$ and any $k \neq i$ if $l = n-1$ (actually (OFL4)(ii) implies that the functions $\mathrm{T}_{k,i}^{k,l}$, do not depend on $k$). Define a distribution $\mathcal{B}$ by

$$\mathcal{B} \;\; = \;\; \operatorname{span}\{g_i, \operatorname{ad}_f^l g_i + b_i^l f, \;\; 1 \leq i \leq m, \;\; 1 \leq l \leq n-1\}.$$

We claim that $\mathcal{B}$ satisfies item (i) of Lemma 4.6.4. The relations $\mathcal{B} \subset \mathcal{G}_f^n$ and $\operatorname{corank}(\mathcal{B} \subset \mathcal{G}_f^n) = 1$ are obvious. Therefore it remains to prove that $[\mathcal{B}, \mathcal{B}] \subset \mathcal{G}_f^n$. Denote $h_i^l = \operatorname{ad}_f^l g_i + b_i^l f$. Lemma 4.6.1 says that $\mathcal{G}$ is contained in the characteristic distribution of $\mathcal{G}_f^n$ and thus any $g \in \mathcal{G} \subset \mathcal{B}$ and any $h_j^l$, $1 \leq j \leq m$, $1 \leq l \leq n-1$, satisfy

$$[g, h_j^l] \in \mathcal{G}_f^n.$$

So we need only to prove that for any vector fields $h_i^r$ and $h_j^l$, for $1 \leq i, j \leq m$, $1 \leq l, r \leq n-1$, we have

$$[h_i^r, h_j^l] \in \mathcal{G}_f^n, \tag{4.6.8}$$

which is equivalent to

$$\omega^k([h_i^r, h_j^l]) = 0, \quad 1 \leq k \leq m.$$

Notice that in the case $r < n-1$ and $l < n-1$, the relation (4.6.8) is obvious since $h_i^r \in \mathcal{G}_f^{n-1}$, $h_j^l \in \mathcal{G}_f^{n-1}$ and then the condition (OFL2) yields directly

$$[h_i^r, h_j^l] \in [\mathcal{G}_f^{n-1}, \mathcal{G}_f^{n-1}] \subset \mathcal{G}_f^n.$$

Now consider the case $r = n-1$. Following the same calculation as in (4.6.2), we get

$$\begin{aligned}
\omega^k([h_i^r, h_j^l]) &= \omega^k([h_i^{n-1}, h_j^l]) \\
&= \omega^k([\operatorname{ad}_f^{n-1} g_i, \operatorname{ad}_f^l g_j]) - b_j^l \cdot \omega^k(\operatorname{ad}_f^n g_i) + b_i^{n-1} \cdot \omega^k(\operatorname{ad}_f^{l+1} g_j) + \eta \cdot \omega^k(f) \\
&= \mathrm{T}_{i,j}^{k,l} - b_j^l \cdot \omega^k(\operatorname{ad}_f^n g_i) + b_i^{n-1} \cdot \omega^k(\operatorname{ad}_f^{l+1} g_j).
\end{aligned}$$

We will consider separately the case $l < n-1$ and $l = n-1$. If $l < n-1$, notice that $\operatorname{ad}_f^{l+1} g_j \in \mathcal{G}_f^n$ and thus we have $\omega^k(\operatorname{ad}_f^{l+1} g_j) = 0$, for any $1 \leq k \leq m$. If $k \neq i$, the definition of $\omega^k$ gives that $\omega^k(\operatorname{ad}_f^n g_i) = 0$ and moreover, by the condition $\mathrm{T}_{i,j}^{k,l} = 0$, we get

$$\omega^k([h_i^{n-1}, h_j^l]) = 0.$$

If $k = i$, the definition of $\omega^k$ gives that $\omega^k(\mathrm{ad}_f^n g_i) = \omega^k(\mathrm{ad}_f^n g_k) = 1$. Then the relation (4.6.7) implies that

$$\omega^k([h_i^{n-1}, h_j^l]) = \mathrm{T}_{i,j}^{k,l} - b_j^l = 0.$$

Now assume that $l = n-1$. If $k \neq i, j$, from the definition of $\omega^k$ and the assumption (OFL4)(ii), we find $\omega^k(\mathrm{ad}_f^n g_i) = 0$, $\omega^k(\mathrm{ad}_f^n g_j) = 0$ and $\mathrm{T}_{i,j}^{k,n-1} = 0$ which imply

$$\omega^k([h_i^{n-1}, h_j^{n-1}]) = \mathrm{T}_{i,j}^{k,n-1} - b_j^{n-1} \cdot \omega^k(\mathrm{ad}_f^n g_i) + b_i^{n-1} \cdot \omega^k(\mathrm{ad}_f^n g_j) = 0.$$

If $k = i$ and $i \neq j$, we have $\omega^k(\mathrm{ad}_f^n g_i) = \omega^k(\mathrm{ad}_f^n g_k) = 1$ and $\omega^k(\mathrm{ad}_f^n g_j) = 0$. Taking into account the equality $\mathrm{T}_{k,j}^{k,n-1} = b_j^{n-1}$, we get

$$\omega^k([h_i^{n-1}, h_j^{n-1}]) = \omega^k([h_k^{n-1}, h_j^{n-1}]) = \mathrm{T}_{k,j}^{k,n-1} - b_j^{n-1} = 0.$$

The case $k = j$ and $i \neq j$ follows from the previous one by permuting $i$ and $j$. Therefore the relation (4.6.8) holds for $r = n-1$, any $1 \leq i, j \leq m$ and any $1 \leq l < n-1$. Finally the case $l = n-1$ follows from that for $r = n-1$ just by permuting $l$ and $r$. Therefore the relation (4.6.8) is always true and implies $[\mathcal{B}, \mathcal{B}] \subset \mathcal{G}_f^n$. $\square$

### 4.6.2 Proof of Theorem 4.4.6

The equivalence (i) $\Longleftrightarrow$ (iii) $\Longleftrightarrow$ (iv) $\Longleftrightarrow$ (v) is a direct consequence of Proposition 4.4.5 and the results given by Pasillas-Lépine and Respondek in [54] (summarized in Theorem 4.4.1 and Remarks following it). To prove Theorem 4.4.6, it is enough to show that (v) $\Longrightarrow$ (ii) $\Longrightarrow$ (iv). For simplicity, we will denote the distribution $\mathcal{D}_\Sigma$ associated to $\Sigma$ by $\mathcal{D}$.

**Proof:** (**v**) $\Longrightarrow$ (**ii**) Note that (C1)$''$ and (C2)$''$ imply that $\mathcal{D}^{(i)} = \mathcal{D}_i$ for $i \geq 0$. We start by proving, by induction, that $\mathcal{G}_f^{i+1} = \mathcal{D}_i$, for $1 \leq i \leq n$. Obviously, we have

$$\mathcal{G}_f^1 = \mathcal{D}_0 = \mathrm{span}\,\{f, g_1, \ldots, g_m\}.$$

Assume that

$$\mathcal{G}_f^i = \mathrm{span}\,\{f, g_j, \mathrm{ad}_f g_j, \ldots, \mathrm{ad}_f^{i-1} g_j, \quad 1 \leq j \leq m\} = \mathcal{D}_{i-1}.$$

The condition (C3)$''$ implies that $[\mathcal{G}, \mathcal{D}_i] \subset \mathcal{D}_i$, for $1 \leq i \leq m$. Indeed, (C3)$''$ gives $[\mathcal{G}, \mathcal{D}_0] \subset \mathcal{D}_0$ from which we conclude easily (via Jacobi identity) that $[\mathcal{G}, \mathcal{D}^{(i)}] \subset \mathcal{D}^{(i)}$, but $\mathcal{D}^{(i)} = \mathcal{D}_i$. We have

$$\mathcal{D}_{i+1} = \mathcal{D}_i + [\mathcal{D}_0, \mathcal{D}_i] = [f, \mathcal{D}_i] + [\mathcal{D}_0, \mathcal{D}_i],$$

which implies

$$\mathcal{D}_{i+1} = \mathcal{G}_f^{i+2} = \mathrm{span}\,\{f, g_j, \mathrm{ad}_f g_j, \ldots, \mathrm{ad}_f^{i+1} g_j, \quad 1 \leq j \leq m\}.$$

We then have $\mathcal{G}_f^{i+1} = \mathcal{D}_i = \mathcal{D}^{(i)}$ and, by definition, $\mathcal{D}^{(i)} = \mathcal{D}^{(i-1)} + [\mathcal{D}^{(i-1)}, \mathcal{D}^{(i-1)}]$ which yields $[\mathcal{G}_f^i, \mathcal{G}_f^i] \subset \mathcal{G}_f^{i+1}$. This proves the conditions (OFL1) and (OFL2). The condition (OFL3) coincides with (C3)$''$ while (OFL4) follows from Lemma 4.6.4, where we take $\mathcal{B} = \mathcal{L}_{n-1}$.

(**ii**) $\Longrightarrow$ (**iv**) The condition (OFL1) implies that $\operatorname{rank} \mathcal{G}_f^{i+1} = (i+1)m + 1$, for $1 \le i \le n$, in a neighborhood of $x_0$. We will show $\mathcal{D}^{(i)} = \mathcal{G}_f^{i+1}$ by induction. The relations $\mathcal{D}^{(0)} = \mathcal{G}_f^1$ is obvious. Assume that

$$\mathcal{G}_f^i = \operatorname{span}\{f, g_j, \operatorname{ad}_f g_j, \ldots, \operatorname{ad}_f^{i-1} g_j, \quad 1 \le j \le m\} = \mathcal{D}^{(i-1)},$$

then

$$\begin{aligned} \mathcal{D}^{(i)} &= \mathcal{D}^{(i-1)} + [\mathcal{D}^{(i-1)}, \mathcal{D}^{(i-1)}] \\ &= \mathcal{G}_f^i + [\mathcal{G}_f^i, \mathcal{G}_f^i] \\ &= \mathcal{G}_f^{i+1} + \operatorname{span}\{[\operatorname{ad}_f^l g_j, \operatorname{ad}_f^s g_k], \quad 0 \le l, s \le i-1, 1 \le j, k \le m\}. \end{aligned}$$

Now (OFL2) implies that for $0 \le l, s \le i-1,\ 1 \le j, k \le m$,

$$[\operatorname{ad}_f^l g_j, \operatorname{ad}_f^s g_k] \in \mathcal{G}_f^{i+1},$$

and thus $\mathcal{D}^{(i)} \subset \mathcal{G}_f^{i+1}$. The inclusion $\mathcal{G}_f^{i+1} \subset \mathcal{D}^{(i)}$ is obvious and thus by an inductive argument, $\mathcal{G}_f^{i+1} \subset \mathcal{D}^{(i)}$ for any $i \ge 0$. Therefore the conditions (C1)$'$ and (C4)$'$ hold and $\operatorname{rank} \mathcal{D}^{(n-1)} = nm + 1$. Moreover, the condition (C2)$'$ follows from Lemma 4.6.4. Finally, the condition (C3)$'$ can be verified directly due to the structure of $\mathcal{L}_{n-1}$ given in the proof of Lemma 4.6.4. In fact, by Lemma 4.6.3, the involutive subdistribution $\mathcal{L}_{n-1}$ is unique and therefore it is given by

$$\mathcal{L}_{n-1} = \mathcal{B} = \operatorname{span}\{g_j, \operatorname{ad}_f^1 g_j + b_j^1 f, \ldots, \operatorname{ad}_f^{n-1} g_j + b_j^{n-1} f, \quad 1 \le j \le m\}.$$

It is immediate to see that

$$\mathcal{D}_0 = \operatorname{span}\{f, g_1, \ldots, g_m\} \not\subset \mathcal{L}_{n-1},$$

and hence the condition (C3)$'$ holds. $\square$

# Conclusions

Nous avons étudié dans ce mémoire la géométrie et la structure des systèmes de contrôle non linéaires qui sont équivalents aux systèmes de contact - une classe de systèmes non holonomes. Nous résumons dans cette partie nos résultats principaux .

Dans le premier chapitre, nous avons étudié la propriété de la platitude des systèmes avec deux contrôles. Premièrement, nous avons montré qu'une structure de Goursat est $x$-plate en $(x_0, u_0)$, $u_0 \notin U_{\mathrm{sing}}(x_0)$, si et seulement si les conditions de régularité $\dim \mathcal{D}^{(i)}(x_0) = \dim \mathcal{D}_i(x_0)$, pour $0 \leq i \leq n$, sont satisfaites. Ceci répond à la question posée par Martin et Rouchon [39] (voir aussi [40]) pour la $x$-platitude. Ensuite, nous avons donné des conditions nécessaires et suffisantes vérifiables (Theorème 1.3.2 et 1.3.3) pour qu'une paire de fonctions données $(\varphi_1, \varphi_2)$ forme une $x$-sortie plate du système $\Sigma$ avec deux contrôles qui est équivalent au système chaîné (système de contact canonique sur $J^n(\mathbb{R}^1, \mathbb{R}^1)$). Nous avons aussi décrit le lieu singulier de cette $x$-sorties plate. Notons $\mathcal{C}_{n-1}$ la distribution caractéristique de $\mathcal{D}^{(n-1)}$ et $g$ un champ de vecteurs dans $\mathcal{D}$ tel que $g(x_0) \notin \mathcal{C}_{n-1}(x_0)$. Nous avons montré qu'étant donné une fonction lisse $\varphi_1$ telle que $L_c\varphi_1 = 0$, $\forall\, c \in \mathcal{C}_{n-1}$, et $L_g\varphi_1(x_0) \neq 0$, il existe toujours une fonction $\varphi_2$ telle que $(\varphi_1, \varphi_2)$ frome une $x$-sortie plate de $\Sigma$ en $(x_0, u_0)$. De plus, $\varphi_2$ est unique au sens de $\mathrm{span}\,\{\mathrm{d}\varphi_1(x), \mathrm{d}\varphi_2(x)\} = \mathrm{span}\,\{\mathrm{d}\varphi_1(x), \mathrm{d}\tilde{\varphi}_2(x)\}$, pour tous les points $x$ dans un voisinage du point $x_0$, où $\tilde{\varphi}_2$ est une autre fonction telle que $(\varphi_1, \tilde{\varphi}_2)$ est aussi une $x$-sortie plate. Ces derniers résultats nous permettent de calculer toutes les $x$-sorties plates pour un système qui est équivalent au système chaîné en résolvant un système d'équations aux dérivées partielles du 1^er^ ordre. En fin, nous avons appliqué nos résultats au système du robot mobile avec remorques pour décrire toutes ses $x$-sorties plates.

Dans le deuxième chapitre, nous avons étudié le système à $n$-barres dans l'espace $\mathbb{R}^{m+1}$ qui généralise le système du robot mobile avec remorques sur le plan. Nous avons introduit un modèle cinématique de ce système sans utiliser les variables d'angle. Les résultats principaux sont les suivants: (1) Le système à $n$-barres dans $\mathbb{R}^{m+1}$ est localement équivalent au système $m$-chaîné (i.e., système de contact canonique sur $J^n(\mathbb{R}^1, \mathbb{R}^m)$) en tous les points réguliers que nous avons caractérisés. Nous avons aussi montré que les lieux singuliers sont essentiellement différents entre les deux cas $m = 1$ et $m \geq 2$. De plus, le système à $n$-barres est contrôllable globalement dans l'espace de configuration $Q = \mathbb{R}^{m+1} \times (S^m)^n$. (2) Le système à $n$-barres $\mathbb{R}^{m+1}$, pour $m \geq 2$, est $x$-plat en tous les points réguliers. De plus, les coordonnées cartésiennes du point source $P_0 = (x_0^1, x_0^2, \cdots, x_0^{m+1})$ de la première barre $\overrightarrow{P_0P_1}$ forment la seule $x$-sortie plate minimale. Contrairement au cas $m = 1$, où la sortie plate minimale n'est pas unique.

Dans le troisième chapitre, nous avons donné des conditions nécessaires et suffisantes pour qu'une distribution $\mathcal{D}$ soit équivalente à la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$. Toutes ces conditions sont vérifiables sur la distribution $\mathcal{D}$. Pour montrer notre théorème, nous avons proposé une forme normale de Bryant étendue qui généralise la caractérisation de la distribution de Cartan $\mathcal{CC}^1(\mathbb{R}^k, \mathbb{R}^m)$. Nous avons aussi étudié le problème suivant : quand une distribution $\mathcal{D}$ contient une sous-distribution involutive qui est de corang $k$ dans $\mathcal{D}$. Nous avons obtenu le résultat suivant : si $\mathcal{D}$ est une distribution de rang $k + mk$ définie sur une variété $M$ de dimension $m + k + mk$, pour $m \geq 3$, telle que $\mathcal{D}^{(1)} = TM$ et $\mathcal{C}(\mathcal{D}) = 0$, alors $\mathcal{D}$ contient une sous-distribution involutive de corang $k$ dans $\mathcal{D}$ si et seulement si le rang d'Engel de $\mathcal{D}^\perp$ est constant et égal à $k$.

Dans le quatrième chapitre, nous avons étudié le système de contrôle sous la forme $\Sigma : \dot{x} = f(x) + \sum_{i=1}^m g_i(x)u_i$, où $x \in X = \mathbb{R}^{(n+1)m+1}$, en supposant que $f(x_0) \neq 0$ où $x_0$ est le point autour duquel nous travaillons. Nous avons obtenu des conditions nécessaires et suffisantes vérifiables pour que le système $\Sigma$ soit linéarisable par bouclage orbital. Toutes ces conditions peuvent être vérifiées directement sur le système original. De plus, nous avons aussi donné une construction de la fonction $\gamma$ qui décrit le changement de temps, i.e., la transformation orbitale. Il est aussi important de remarquer que notre résultat implique une relation intéressante entre la linéarisation orbitale par bouclage du système $\Sigma$ et la caractérisation du système de contact sur $J^n(\mathbb{R}^1, \mathbb{R}^m)$.

# Perspectives

A l'issue de cette thèse, certains problèmes liés à nos études demeurent encore non-résolus et indiquent de possibles directions de recherche pour l'avenir :

(1). Dans le premier chapitre, nous avons montré qu'une structure de Goursat est $x$-plate en $(x_0, u_0)$, $u_0, \notin U_{\mathrm{sing}}(x_0)$, si et seulement si les conditions de regularité $\dim \mathcal{D}^{(i)}(x_0) = \dim \mathcal{D}_i(x_0)$, pour $0 \leq i \leq n$, sont satisfaites. Deux problèmes intéressants se posent : (i) Est-ce que l'hypothèse que la distribution associée $\mathcal{D}$ soit une structure de Goursat est nécessaire? Autrement dit, si un système avec deux contrôles est $x$-plat au point $x_0$, est-ce qu'il est toujours équivalent au système chaîné en $x_0$? (ii) Est-ce que ce résultat est aussi vrai si on remplace la $x$-platitude par la platitude? Autrement dit, est-ce qu'une structure de Goursat est plate en $(x_0, \bar{u}_0^k)$ si et seulement si les conditions de regularité sont satisfaites en $x_0$?

(2). Dans le chapitre 3, nous avons donné des conditions nécessaires et suffisantes pour qu'une distribution soit équivalente à la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^2, \mathbb{R}^m)$. Deux problèmes intéressants relatifs aux chapitre 2 et 3 sont : (i) Le premier problème consiste à considérer un système à $n$-plans qui bouge dans l'espace $\mathbb{R}^{m+1}$ et de voir comment construire un modèle cinématique pour ce système? Est-ce qu'il est localement équivalent au système de contact canonique sur $J^n(\mathbb{R}^2, \mathbb{R}^m)$? Est-ce que nous pouvons le montrer en utilisant nos résultats dans chapitre 3, caractériser son lieu singulier, étudier la propriété de la platitude et determiner ses $x$-sorties plates? (ii) Le deuxième problème semble plus difficile à étudier mais il est aussi naturel : comment caractériser la distribution de Cartan dans le cas general? Autrement dit, nous voulons trouver des conditions nécessaires et suffisantes vérifiables pour qu'une distribution soit equivalente à la distribution de Cartan $\mathcal{CC}^n(\mathbb{R}^k, \mathbb{R}^m)$, pour $k > 2$.

(3). Considérons le système de contrôle sous la forme $\Sigma : \dot{x} = f(x) + \sum_{i=1}^m g_i(x)u_i$, où $x \in X = \mathbb{R}^N$, pour un certain entier $N$ et $x_0$ est un point de $X$. Les problèmes liés au chapitre 4 sont les suivants : (i) Si $f(x_0) \neq 0$, quand $\Sigma$ est localement equivalent par bouclage orbital à la forme de Goursat généralisée (Extended Goursat normal form [63]), avec des indices de contrôlabilité differents? (ii) Si $f(x_0) = 0$, i.e., au point d'équilibre, quand $\Sigma$ est localement linéarisable par bouclage orbital? Il serait souhaitable que toutes conditions puissent être vérifiables directement sur le système original.

# References


[1] C. Altafini. Some properties of the general n-Trailer. *International Journal of Control*, vol. 74, n. 4, pp. 409–424, March 2001.

[2] E. Aranda-Bricaire, C.H. Moog and J.-B. Pomet. Infinitesimal Brunovsky form for nonlinear systems with applications to Dynamic Linearization. *Geometry in nonlinear control and differential inclusions, Banach Center Publications*, vol. 32, Warszawa, pp. 19–33, 1995.

[3] E. Aranda-Bricaire and J.-B. Pomet. Some explicit conditions for a control system to be feedback equivalent to extended Goursat normal form. In *Proceedings of the IFAC Nonlinear Control Systems Design Symposium*, Tahoe City (California), 1995.

[4] V. I. Arnold. *Geometrical Methods in the Theory of Ordinary Differential Equations*, Second Edition, Springer-Verlag, 1988.

[5] R. W. Brockett. Feedback invariants for nonlinear systems. *IFAC Congress* **6**, Helsinki, pp. 1115–1120, 1978.

[6] P. Brunovský. A classification of linear controllable systems. *Kybernetica* **6** pp. 173–188, 1970.

[7] R. Bryant. *Some aspects of the local and global theory of Pfaffian systems*. Ph.D. thesis, University of North Carolina, Chapel Hill, 1979.

[8] R. Bryant, S-S. Chern, R. Gardner, H. Goldschmidt, and P. Griffiths. *Exterior Differential Systems*. Mathematical Sciences Research Institute Publications. Springer-Verlag, New York, 1991.

[9] E. Cartan. Sur l'équivalence absolue de certains systèmes d'équations différentielles et sur certaines familles de courbes. *Bulletin de la Société Mathématique de France*, 42:12–48, 1914. *Œuvres complètes*, Part. II, Vol. 2, Gauthiers-Villars, Paris.

[10] G. Darboux. Sur le problème de Pfaff. *Bulletin des Sciences mathématiques*, 2(6):14–36,49–68, 1882.

[11] F. Engel. Zur Invariantentheorie der Systeme Pfaff'scher Gleichungen. *Berichte Verhandlungen der Koniglich Sachsischen Gesellshaft der Wissenshaften Mathematisch-Physikalische Klasse, Leipzig*, 41,42:157–176;192–207, 1889,1890.

[12] M. Fliess, J. Lévine, P. Martin and P. Rouchon. Sur les systèmes non linéaires différentiellement plats. *Comptes Rendus de l'Académie des Sciences*, 315:619–624, 1992.

[13] M. Fliess, J. Lévine, P. Martin and P. Rouchon. Flatness and defect of nonlinear systems: Introductory theory and examples. *International Journal of Control*, 61(6):1327–1361, 1995.

[14] M. Fliess, J. Lévine, P. Martin and P. Rouchon. A Lie-Bäcklund approach to equivalence and flatness of nonlinear systems. *IEEE Trans. Automat. Control*, 61:1327–1361, 1999.

[15] G. Frobenius. Über das Pfaffsche problem, *Journal für die reine und angewandte Mathematik*. 82, pp. 230–315, 1877

[16] R. Gardner and W. Shadwick. The GS algorithm for exact linearization to Brunovsky normal form. *IEEE Transactions on Automatic Control*, 37(2):224–230, 1992.

[17] A. Giaro, A. Kumpera and C. Ruiz. Sur la lecture correcte d'un resultat d'Élie Cartan. *Comptes Rendus de l'Académie des Sciences de Paris*, 287:241–244, 1978.

[18] E. Goursat. *Leçons sur le problème de Pfaff*. Hermann, Paris, 1923.

[19] M. Guay. An algorithm for orbital feedback linearization of single-input control affine systems. *Systems & Control Letters*, 38, pp, 271–281 1999.

[20] M. Guay. Orbital feedback linearization of multi-input control affine systems. *Proceeding of the American Control Conference*, Arlington, 2001.

[21] D. Hilbert. Über den Begriff der Klasse von Differentialgleichungen. *Math. Ann.*, 73(1):95–108, 1912.

[22] L. R. Hunt and R. Su. Linear equivalents of nonlinear time varying systems. *Proc. of the MTNS*, Santa Monica, pp. 119–123, 1981.

[23] R. Hunt, R. Su, G. Meyer. Design for multi-input nonlinear systems. *Differential Geometric Control Theory*, Birkhauser, Boston, MA, pp. 268–298, 1982.

[24] A. Isidori. *Nonlinear Control Systems*. 3rd edition, Springer-Verlag, London, 1995.

[25] A. Isidori, C. H. Moog, and A. de Luca. A sufficient condition for full linearization via dynamic state feedback. *In Proc. 25th IEEE Conf. on Decision & Control*, Athens, pp. 203–207, 1986.

[26] B. Jakubczyk. Invariants of dynamic feedback and free systems. In *Proceedings of the European Control Conference*, Groningen, 1993.

[27] F. Jean. The car with n trailers: Characterisation of the singular configurations. *Control, Optimisaton and Calculus of Variations*, Vol. 1, pp. 241–266, 1998.

[28] Z-P. Jiang and H. Nijmeijer. A recursive technique for tracking control of nonholonomic systems in chained form. *IEEE Transaction on Automatic control*, 44(2):265–279, 1999.

[29] B. Jakubczyk and W. Respondek. On linearization of control systems. *Bull. Acad. Polonaise Sci., Ser. Sci. Math.*, vol.28 pp. 517–522, 1980.

[30] V. Jurdjevic. *Geometric Control Theory*. Studies in Advanced Math., 52, Cambridge University Press, N.Y., 1997.

[31] T. Kailath. *Linear systems*. Prentice-Hall, 1980.

[32] A. J. Krener. On the equivalence of control systems and linearization of nonlinear systems. *SIAM J. Contr.* vol. 11, pp. 670–676, 1973.

[33] A. Kumpera and C. Ruiz. Sur l'équivalence locale des systèmes de Pfaff en drapeau. In F. Gherardelli, editor, *Monge-Ampère equations and related topics*, pp. 201–247. Instituto Nazionale di Alta Matematica Francesco Severi, Rome, 1982.

[34] J.P. Laumond. Controllability of a multibody robot. *IEEE Trans. Robotics and Automation*, 9:755–763, 1991.

[35] J. P. Laumond, *Robot Motion Planning and Control*. Lecture Notes on Control and Information Sciences. Springer-Verkag, 229. Springer, ISBN 3-540-76219-1, 1998.

[36] Z. Li and J. F. Canny, editors. *Nonholonomic Motion Plannging*. Internqtional Series in Engineering and Computer Sciences. Kluwer, Dordrecht, 1992

[37] P. Libermann. Sur le problème d'équivalence des systèmes de Pfaff non complètement intégrables. *Publications Paris VII*, 3:73–110, 1977.

[38] P. Martin, R. M. Murray, and P. Rouchon. Flat systems. *In Mathematical control theory,* Part 1, 2 (Trieste, 2001), ICTP Lect. Notes, VIII, pp. 705–768 (electronic). Abdus Salam Int. Cent. Theoret. Phys., Trieste, 2002.

[39] P. Martin and P. Rouchon. Feedback linearization and driftless systems. *CAS internal report* No. 446, 1993.

[40] P. Martin and P. Rouchon. Feedback linearization and driftless systems. *Mathematics of Control, Signals, and Systems*, 7:235–254, 1994.

[41] R. Montgomery and M. Zhitomirskiĭ. Geometric approach to Goursat flags. Annales de l'Institut Henri Poincaré. Analyse non linaire vol. 18, no4, pp. 459–493, 2001.

[42] P. Mormul. Goursat flags: classification of codimension-one singularities. *Journal of dynamical and control systems* vol.6, pp. 311–330, 2000.

[43] P. Mormul. Multi-dimensional Cartan prolongation and special k-flags, *in Geometric Singularity Theory, Banach Center Publications 65 of Math. Editors H. Hironaka, S. Janeczko and S. Lojasiewicz,* Vol. 65, Polish Acad. Sci., Warsaw, pp.157–178, 2004.

[44] A. S. Morse and L. M. Silverman. Structure of index-invariant systems, *SIAM J. Control*, 11(2):215–225, 1973.

[45] R. Murray. Nilpotent bases for a class of nonintegrable distributions with applications to trajectory generation for nonholonomic systems. *Mathematics of Control, Signals, and Systems*, 7:58–75. 1994

[46] R. Murray, S. Sastry. Nonholonomic motion planning: Steering using sinusoids. *IEEE Transactions on Automatic Control*, Vol. 38, pp. 700-716, 1993.

[47] H. Nijmeijer and A. J. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York, 1990.

[48] P. Olver. *Equivalence, Invariants, and Symmetry.* Cambridge University Press, 1995.

[49] W. Pasillas-Lépine and W. Respondek. Conversion of the n-trailer into Kumpera-Ruiz normal form and motion planning through the singular locus. *Proceeding of the IFAC Nonlinear Control Systems Design Symposium*, Enschede, Neherlands, pp. 789–794, 1998.

[50] W. Pasillas-Lépine and W. Respondek. On the geometry of control systems equivalent to canonical contact systems: regular points, singular points and flatness. *Proceedings of the* 39*th IEEE Conference of Decision and Control,* Sydney, Australia, pp. 5151–5156, 2000.

[51] W. Pasillas-Lépine and W. Respondek. Contact systems and corank one involutive subdistributions. *Acta Applicandae Mathematicae* 69, pp. 105–128, 2001.

[52] W. Pasillas-Lépine and W. Respondek. On the geometry of Goursat structures. *ESAIM Control, Optimisation, and Calculus of Variations*, 6:119–181, 2001.

[53] W. Pasillas-Lépine and W. Respondek. Nilpotentization of the kinematics of the n-trailer system at singular points and motion planning through the singular locus. *International Journal of Control 74*, pp. 628–637, 2001.

[54] W. Pasillas-Lépine, W. Respondek. Canonical Contact Systems for Curves: A Survey. In : *Contemporary Trends in Geometric Control Theory and Applications*, A. Anzaldo, B. Bonnard, J.P. Gauthier et F. Monroy (eds), World Scientific, Singapore, pp. 77–112. 2001.

[55] J-F. Pfaff. Methodus generalis, aequationes differentiarum partialum, nec non aequationes differentiales vulgares, utrasque primi ordinis, inter quotcunque variabiles, completi integrandi. *Abhandlungen der Königlich-Preußischen Akademie der Wissenschaften zu Berlin, Mathematische Klasse*, pp. 76–136, 1814-1815.

[56] J.-B. Pomet. A differential geometric setting for dynamic equivalence and dynamic linearization. In B. Jakubczyk, W. Respondek, and T. Rzezuchowski, editors, *Geometry in Nonlinear Control and Differential Inclusions*, volume 32, pp. 319–339. Banach Center Publications, Warszawa, 1995.

[57] W. Respondek, Symmetries and minimal flat outputs of nonlinear control systems, *Symposium of New Trends in Nonlinear Dynamics and Control and their Applications*, Monterey, USA, Octobre 2002.

[58] W. Respondek, Geometric methods in linearization of control systems. in *Mathematical Control Theory*, Cz. Olech, B. Jakubczyk, and J. Zabczyk (eds.), Banach Center Publications, vol. 14, PWN-Polish Scientific Publishers, Warsaw, pp. 453–467, 1985.

[59] W. Respondek. Disturbance decoupling via dynamic feedback. in: *Analysis of Controlled Dynamical Systems*, B. Bonnard, B. Bride, J.P. Gauthier, I. Kupka (eds), Birkhäuser, Boston, pp. 347–357, 1991.

[60] W. Respondek. Transforming nonholonomic control system into the canonical contact form, *P*roc. of the $40^{th}$ IEEE Conf. on Decision and Control, Orlando, Florida, pp. 1781–1786, 2001.

[61] W. Respondek. Dynamic input-output linearizaton and decoupling of nonlinear systems. *Proc. 2nd European Control Conference*, Groningen, pp. 1523-1527, 1993.

[62] W. Respondek. Orbital Feedback Linearization of Single-input Nonliear Control Systems. *Proceedings of IFAC NOLCOS'*98, Enschede, The Netherlands, pp. 499–504, 1998.

[63] W. Respondek, W. Pasillas-Lépine. Extended Goursat Normal Form: a Geometric Characterization. in *Lecture Notes in Control and Inform. Sci.* 259 pp. 323–338, Springer, London 2001.

[64] P. Rouchon, M. Fliess, J. Lévine, and P. Martin. Flateness and motion planning: the car with $n$ traliers. Proc. of the European Control Conference ECC'93, Groningen, pp. 1518-1522, 1993.

[65] C. Samon. Control of chained system: application to path-following and time-varying point stabilization of mobile robots. *IEEE Trans. on Automatic Control*, 40, pp. 64-77, 1995.

[66] S. Sastry. *Nonlinear systems: analysis, stability, and control*. Springer Verlag, 1999.

[67] M. Sampei and K. Furuta. On time scaling for nonlinear systems: applications to linearization. *IEEE Trans. Automat. Control* AC-31, pp. 459-462, 1986.

[68] L. M. Silverman and H. E. Meadows. Controllability and observability in time-variable linear systems, *SIAM J. Control*, 5:64–73, 1967.

[69] P. S. Pereira da Silva. Flatness of nonlinear control systems and exterior differential systems. Organizado por A. Isidori, F. Lamnabh Lagarrigue, W. Respondek. *Nonlinear Control in the Year 2000*. Springer Verlag, Berlin, Vol.2, pp.205–227, 2001.

[70] P. S. Pereira da Silva, C. Corrêa Filho. Relative Flatness and Flatness of Implicit Systems. *SIAM Journal on Control and Optimization*, Vol.39, n.6, pp.1929–1951, 2001.

[71] M. Slayaman. *Bras articulé et distributions multi-drapeaux*, Université de Savoie, Laboratoire de Mathématiques (LAMA), Ph.D. thesis 2008.

[72] M. Slayaman, F. Pelletier. Articulated arm and special multi-flags (priprint)

[73] O. J. Sørdalen. Conversion of the kinematics of a car with n trailers into a chained form. *Proceeding of* 1993 *International Conference on Robotics and Automation*, atlanta, CA, pp. 382–387, 1993.

[74] O. J. Sørdalen. On the global degree of nonholonomy of a car with n trailers. *In Proceeding of the IFAC Symposium on Robot Control*, Capri, Italy, pp. 343–348, 1994.

[75] D. Tilbury, R. Murray and S. Sastry. Trajectroy generation for the n-trailer problem using Goursat normal form. *IEEE Transactions on Automatic Control*, 40(5):802–819, 1995.

[76] 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(4):343–364, 1995.

[77] M. van Nieuwstadt, M. Rathinam, and R. M. Murray. Differential Flatness and Absolute Equivalence of Nonlinear Control Systems *SIAM J. Control Optim.* Vol 36, Issue 4, pp. 1225–1239, 1998.

[78] E. von Weber. Zur Invariantentheorie der Systeme Pfaff'scher Gleichungen. *Berichte Verhandlungen der Koniglich Sachsischen Gesellshaft der Wissenshaften Mathematisch-Physikalische Klasse, Leipzig*, 50:207–229, 1898.

[79] K. Yamaguchi. Contact geometry of higher order. *Japanese J. of Math*, 8: 109–176, 1982.

**GÉOMÉTRIE ET CLASSIFICATION DES SYSTÈMES DE CONTACT: APPLICATIONS AU CONTRÔLE DES SYSTÈMES MÉCANIQUES NON HOLONOMES**

**RÉSUMÉ:** Dans la première partie de cette thèse, nous caractérisons complètement toutes les sorties plates et leurs lieux singuliers pour un système avec deux contrôles qui est équivalent au système chaîné. Nous appliquons aussi ce résultat au système du robot mobile avec des remorques pour calculer toutes ses sorties plates. Dans la deuxième partie, nous présentons un nouveau modèle pour le système à $n$-barres dans l'espace de dimension $m+1$. Nous montrons que ce système est localement équivalent au système $m$-chaîné (système de contact sur $J^n(\mathbb{R}, \mathbb{R}^m)$) et caractérisons aussi ses lieux singuliers. Ensuite, nous analysons sa propriété de platitude et donnons ses sorties plates minimales. Dans la troisième partie, nous donnons des conditions nécessaires et suffisantes pour qu'une distribution soit équivalente à la distribution de Cartan pour des surfaces. Finalement, dans la quatrième partie, nous donnons des conditions nécessaires et suffisantes vérifiables pour qu'un système multi-entrées soit linéarisable par bouclage orbital.
**Mots-clés:** Système de contrôle, système nonholonome, platitude, sortie plate, système de contact, distributon de Cartan, système à $n$-barres, linearisation par bouclage orbital

**GEOMETRY AND CLASSIFICATION OF CONTACT SYSTEMS: APPLICATIONS TO CONTROL OF NONHOLONOMIC MECHANICAL SYSTEMS**

**ABSTRACT:** In the first part of this Ph.D. thesis, we characterize all flat outputs and their singular loci of any 2-input driftless control system which is equivalent to the chained form. Then we apply that result to the $n$-trailer system in order to calculate all its flat outputs. In the second part, we establish a new model of the $n$-bar system in $(m+1)$-dimensional space. With the help of this model, we show that the system is locally equivalent to the $m$-chained form (canonical contact system on $J^n(\mathbb{R}, \mathbb{R}^m)$) and also describe its singular locus. Furthermore, we analyze its flatness property and determine its minimal flat outputs. In the third part, we give necessary and sufficient conditions for a distribution to be locally equivalent to the Cartan distribution for surfaces. Finally, in the fourth part, we give necessary and sufficient verifiable conditions for a multi-input affine control system to be orbital feedback linearizable.
**Key-words:** Control system, nonholonomic system, flatness, flat output, contact system, Cartan distributon, $n$-bar system, orbital feedback linearization