On Vessiot's Theory of Partial Differential Equations

Dissertation zur
Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
im Fachbereich Mathematik
der Universität Kassel

vorgelegt von
Dirk Fesser
von Bad Dürkheim

Kassel im November 2007

Erster Gutachter: Professor Dr. Werner Markus Seiler, Universität Kassel
Zweiter Gutachter: Professor Dr. Bernd Heinrich Matzat, Universität Heidelberg

Tag der mündlichen Prüfung: 25. Januar 2008

Contents

Chapter 1

Introduction

The object of research presented here is Vessiot's theory of partial differential equations. During the first half of the twentieth century, Ernest Vessiot [43] developed an approach for the treatment of general systems of partial differential equations which is dual to the theory of exterior systems, the Cartan-Kähler theory [4, 18, 21], in so far as it takes vector fields as its main object of consideration and which uses the Lie bracket instead of the exterior derivative. For a given system of differential equations one seeks a distribution of vector fields that is both tangent to the differential equation and appropriate with respect to the contact structure used to describe the relation between independent and dependent coordinates. Among the subdistributions in it one special kind can be interpreted as tangential approximations for the solutions to the equation. These subdistributions of vector fields then allow to regard solutions to the differential equation as their integral manifolds.

The modern presentation [25, 31, 35, 37] of the formal theory of differential equations considers differential equations as fibred submanifolds within an appropriate jet bundle with a base space of dimension, say, n and explores formal integrability and the stronger concept of involutivity of differential equations to analyze if the equations are solvable.

The modern formulation of Vessiots approach is then to construct for a given differential equation the *Vessiot distribution*, tangent to it such that it is also contained within the contact distribution of the jet bundle. Then the aim is to find n-dimensional subdistributions in it which are fibred over that base space; they are called *integral distributions* and consist of *integral elements*, which are to be glued in such a way that they define a subdistribution that is closed under the Lie bracket. This notion is called a *flat Vessiot connection*.

Vessiot's approach has not become popular. Modern treatments of his theory are restricted to special systems (like ordinary differential equations [3] or hyperbolic equations [42]), and general considerations [15, 40] lack the precision of treatment which has been developed in the more widespread Cartan-Kähler theory; in particular, the necessary prerequisites and assumptions for the solvability of an equation and for the construction of the distributions mentioned above have not been explored yet and are neglected even in Vessiot's own work.

One main result of this thesis is to have closed this gap and to provide a foundation for

Vessiot's theory which is equally rigorous. Another result is to make clear the interrelation of Vessiot's approach and the pivotal notions of formal theory (like formal integrability and involutivity of differential equations). A major point of this thesis is the formulation of conditions which are necessary and sufficient for Vessiot's approach to succeed. This proves the equivalence of Vessiot's theory and formal theory.

We show that Vessiot's step-by-step approach to the construction of the wanted distributions succeeds if, and only if, the given system is involutive. To this end, we first prove an existence theorem for integral distributions (Theorem 3.3.9). Our definition of integral elements is new (but natural as it is based on the contact map, which any jet bundle brings along), its equivalence to the classical notion is then proven (Proposition 2.4.22). Furthermore, an existence theorem for flat Vessiot connections is proven (Theorem 3.3.28). The geometrical structure of the basic theory is being analyzed and simplified as compared to other approaches (in particular the structure equations needed for the proofs of the existence theorems). The obstructions to involution of a differential equation are deduced explicitly (Lemma 2.5.8). (The representation refers to first-order systems, which does not weaken the generality but improves clarity.)

Analyzing the structure equations not only yields theoretical insight, but also renders an algorithm for the explicit determination of the coefficients of the vector fields which span the sought integral distributions. Now an implementation of Vessiot's approach in the computer algebra system `MuPAD` is possible, and is being coded by the method developed here.

Though not an aim of this thesis but now within reach through the results of this thesis (in particular now that the integral elements in the formal theory are identified and their construction is clear), is the proof of the equivalence of the formal theory and Cartan-Kähler theory, which are linked by Vessiot theory. Though generally acknowledged, it seems an explicit proof has not been published in the literature yet.

The text is organized as follows. The second chapter summarizes the main concepts of formal theory, most of them widely known, to introduce the notation. The exposition here follows Seiler [37, 38]. We introduce jet bundles and their contact structure in Section 2.1 and differential equations as fibred submanifolds within them in Section 2.2. In this thesis we consider general systems of partial differential equations; these include arbitrarily non-linear systems. But the structure of such systems, as described by the Vessiot distribution, can be represented easily or can be reduced to the one of systems with a less complicated representation: the geometric symbol, introduced in Section 2.3, is a helpful brute-force linearization of an equation; if the given system is not involutive, it can be completed to an equivalent involutive system by a finite series of operations according to the Cartan-Kuranishi theorem in Section 2.4, and a system of arbitrary order can be rewritten as a first-order system leaving the involutivity of the given system undisturbed, in a way which is outlined in Section 2.5. (The number of variables is not in general kept constant through these transformations, but what is kept constant is the Cartan characters, which are pivotal in our theory.) When a differential equation is given as a first-order system or rewritten as such, it allows a local representation in reduced Cartan normal form. This local representation is not usual in the literature, but it helps us to clarify the argumentation because it classifies the variables of a local representation in a natural way. It is particularly convenient when, as a new result, we deduce the obstructions to involution.

In the third chapter, we give a modern presentation of Vessiot theory. We define the Vessiot distribution for a differential equation in Section 3.1—though not in the usual way, but in a way analogous to the introduction of the geometrical symbol and thus more natural in our approach—and therefore arrive at a description of the geometric symbol as a subdistribution within the Vessiot distribution. As such, it is a decisive help in the construction of flat Vessiot connections: now the Vessiot distribution can be written as a direct sum of the symbol and a horizontal complement (which is not unique). The n-dimensional, involutive subdistributions which are fibred over the base space are the linear approximations for the solutions to a differential equation of Vessiot's approach. To show their existence is now possible by analyzing the structure equations of the Vessiot distribution. (In Subsection 3.1.2 we analyze the assumptions under which n-dimensional complements, in other words Vessiot connections, exist. Then in Subsection 3.1.3 we analyze under what conditions flat Vessiot connections exist.) The ansatz that is used here is so handy that the structure equations have a simple form—at least when compared to other recent approaches. We summarize these in Section 3.2 for easier comparison.

In Section 3.3 we give the two existence theorems and the accompanying proofs. The theory of distributions and exterior systems has the advantage that many of the usual methods are linear-algebraic. The approach developed here simplifies them even further. One amendment is that the classical quadratic Equations (3.24) are replaced by linear Equations (3.29); these linear Equations can explicitly be linked with the obstructions to involution and integrability conditions. On the other hand, considering general, arbitrary systems of differential equations means that their representations and the calculations in local coordinates may appear somewhat tedious, even for linear systems, and obscured by index clouds. For the calculations here we have to develop a special notation regarding block matrices which comprehend the structure equations of the Vessiot distribution. In order not to complicate the main relations, which are of a simple nature, by unwonted notation, we sketch the core algorithm by a series a figures and give some elaborate examples, and collect the technical details in several subsections. The downside is that the length of the text grows such that it fills more than a hundred pages.

The last chapter shows possible further developments. We hint at applications of Vessiot's approach for involutive systems concerning the equivalence of formal theory and Cartan-Kähler theory (as Vessiot theory is a link between them); qualitative classification of differential equations based on the Vessiot distribution; possible developments of differential Galois theory for systems of non-finite type (which are regarded as covered by systems of finite type which correspond to involutive subdistributions within the Vessiot distribution); and the study of symmetries of a differential equation based on the equation and its Vessiot distribution alone. These hints are entirely speculative and meant to show that studying the topic of this thesis is not just an end in itself but may be connected with several interesting fields.

Danksagung

Ich möchte mich sehr herzlich bedanken bei Professor Dr. Werner Markus Seiler für die Anregung zu dieser Arbeit und seine hervorragende Betreuung.

Professor Dr. Bernd Heinrich Matzat danke ich sehr herzlich für die Übernahme des Korreferates und die großzügige Unterstützung während meiner Zeit in Heidelberg.

Den Mitgliedern seiner Arbeitsgruppe, Charlotte Hardouin, Julia Hartmann, Max Albert, Sebastian Basten, Michael Dettweiler, Michael Maier, Andreas Röscheisen, Stefan Wewers, Michael Wibmer, und Michael Winckler danke sehr ich für den äußerst angenehmen Umgang.

Besonders möchte ich Marcus Hausdorf und Dirk Saller für ihre Hilfe danken.

Professor Dr. Jacques Calmet vom Institut für Algorithmen und kognitive Systeme der Universität Karlsruhe danke für seine freundliche Unterstützung.

Heidelberg, 09. November 2007

Dirk Fesser

Chapter 2

Formal Theory

The formal theory of differential equations is an approach to describe differential equations by way of using methods from differential geometry based on the formalism of jet bundles. The overview in this chapter is mainly in local coordinates with additional hints at an intrinsic approach. A more extensive presentation is given by Seiler [37, 38].

Let $n \in \mathbb{N}$. A *multi-index* μ is an n-tuple of non-zero integers $(\mu_i \colon 1 \le i \le n) \in \mathbb{N}_0^n$. Multi-indices may be added componentwise. If for all $1 \le i \le n$ and for two multi-indices μ and ν we have $\mu_i \ge \nu_i$, then $\mu - \nu$ is defined componentwise as well. Let k be some integer. Then, for $1 \le i \le n$ and any multi-index μ, set $\mu + k_i := (\mu_1, \ldots, \mu_{i-1}, \mu_i + k, \mu_{i+1}, \ldots, \mu_n)$ if $\mu_i + k \ge 0$. We call $|\mu| := \sum_{i=1}^n \mu_i$ the *order of the multi-index* μ. It is often convenient to write a multi-index $\mu = (\mu_i)$ as a list, consisting of numbers i (or terms like x^i) each of which is written μ_i times. The Einstein convention is used where it seems appropriate and the domain of summation is obvious from the context.

2.1 The Jet Bundle and its Contact Structure

Jet bundles were introduced by Ehresmann [9, 10, 11, 12, 13]. A standard textbook on the subject is by Saunders [36]. They are considered for fibred manifolds, which provides a distinction of the variables into independent and dependent ones, but in a such way that the derivatives of the dependent variables are regarded as algebraically independent coordinates for the jet bundle.

2.1.1 Jet Bundles

All manifolds which we consider are assumed to be second-countable and Hausdorff. For any such manifold M, denote the ring of smooth functions $C^\infty(M, \mathbb{R})$ by $\mathcal{F}(M)$.

For two manifolds $\mathcal{E}$ and $\mathcal{X}$ and a map $\pi \colon \mathcal{E} \to \mathcal{X}$, the triple $(\mathcal{E}, \pi, \mathcal{X})$ is called a *fibred manifold* with the *base space* $\mathcal{X}$, if π is a surjective submersion; it is called the *projection*. (For two manifolds M and N a submersion is a smooth map $f \colon M \to N$ such that for $\dim M \ge \dim N$ its rank is maximal, where the rank of f is (pointwise) defined by the rank of its tangent map $Tf \colon TM \to TN$.) For any $\mathbf{x} \in \mathcal{X}$ the subset $\pi^{-1}(\mathbf{x})$ is called the *fibre over* $\mathbf{x}$. We consider finite dimensional manifolds and denote the dimension of $\mathcal{X}$ by

n and that of $\mathcal{E}$ by $n+m$ where m is the dimension of the fibre $\mathcal{U}$. A submanifold $\mathcal{R} \subseteq \mathcal{E}$ is called *fibred (over $\mathcal{X}$)*, if the restriction $\pi|_{\mathcal{R}} \colon \mathcal{R} \to \mathcal{X}$ is a surjective submersion, too.

Let $\pi \colon \mathcal{E} \to \mathcal{X}$ be a fibred manifold. For local coordinates on $\mathcal{X}$, we write $\mathbf{x} = (x^i \colon 1 \leq i \leq n)$ and for local coordinates on the typical fibre $\mathbf{u} = (u^\alpha \colon 1 \leq \alpha \leq m)$. Instead of x^1, x^2 and so on we often write x, y, z, s, t or something similar, and in the same way u, v, w for u^1, u^2, u^3. A *local section for the fibration* π is a smooth map $\sigma \colon \mathcal{O} \to \mathcal{E}$ for some open set $\mathcal{O} \subseteq \mathcal{X}$ with $\pi \circ \sigma = \mathrm{id}_{\mathcal{O}}$. If $\mathcal{O} = \mathcal{X}$, σ is called a *global section*. In local coordinates, any local section defines a smooth function $\mathbf{s} \colon \mathcal{O} \to \mathcal{U}$ such that $\sigma(\mathbf{x}) = (\mathbf{x}, \mathbf{s}(\mathbf{x}))$. We want to keep the notation simple and therefore suppress mentioning local charts explicitly. Let $\Gamma_{\mathbf{x}}\pi$ denote the set of all local sections where $\mathbf{x} \in \mathcal{O}$ for some open neighborhood $\mathcal{O} \subseteq \mathcal{X}$, and $\Gamma_L\pi$ the sheaf of all local sections for π. Let $\Gamma\pi$ denote the set of all global sections. Let $S_q(T^*_{\mathbf{x}}\mathcal{X}) \otimes T_\xi\mathcal{E}$ be the vector space of symmetric q-linear mappings from $(T_{\mathbf{x}}\mathcal{X})^q$ to $T_\xi\mathcal{E}$.

Definition 2.1.1. A *q-jet* can be regarded as an equivalence class $[\sigma]^{(q)}_{\mathbf{x}_0}$ of local sections where two local sections σ_1 and σ_2 are considered equivalent, if they define two functions $\mathbf{s}_1, \mathbf{s}_2 \colon \mathcal{X} \to \mathcal{E}$ with the Taylor-expansions of $\mathbf{s}_1$ and $\mathbf{s}_2$ in adapted local coordinates being equal up to order q at the expansion point $\mathbf{x}_0 \in \mathcal{X}$ (that is, they have a contact of order q there). Thus we interpret a q-jet as a truncated Taylor series. Now the *jet-space of order* q is the set of all q-jets

$$J_q\pi = \{[\sigma]^{(q)}_{\mathbf{x}_0} \colon \mathbf{x}_0 \in \mathcal{X} \text{ and } \sigma \in \Gamma_{\mathbf{x}_0}\pi\} \ .$$

As fibre coordinates for the point $[\sigma]^{(q)}_{\mathbf{x}_0}$ we may use (μ denoting a multi-index) $\mathbf{u}^{(q)} := (u^\alpha_\mu \colon 1 \leq \alpha \leq m, 0 \leq |\mu| \leq q)$ and interpret u^α_μ as the value of $\partial^{|\mu|} s^\alpha / \partial \mathbf{x}^\mu$ at the expansion point $\mathbf{x}_0 \in \mathcal{X}$ for $\mathbf{s} \colon \mathcal{X} \to \mathcal{U}$ with $\mathbf{s} = (s^\alpha \colon 1 \leq \alpha \leq m)$ and a fibre $\mathcal{U}$.

Consider two points $[\sigma_1]^{(q)}_{\mathbf{x}}$ and $[\sigma_2]^{(q)}_{\mathbf{x}}$ in $J_q\pi$ from the same fibre with regard to the fibration over π^q_{q-1}, that is, $[\sigma_1]^{(q-1)}_{\mathbf{x}} = [\sigma_2]^{(q-1)}_{\mathbf{x}}$. Then $[\sigma_1]^{(q)}_{\mathbf{x}}$ and $[\sigma_2]^{(q)}_{\mathbf{x}}$ correspond to two Taylor-series which are truncated at order q and are equal up to order $q-1$. Thus their difference yields for each u^α a homogeneous polynomial of degree q. The fibre $(\pi^q_{q-1})^{-1}([\sigma_1]^{(q-1)}_{\mathbf{x}})$ has therefore as its underlying vector space

$$S_q(T^*_{\mathbf{x}}\mathcal{X}) \otimes V_\xi\pi \cong V_{[\sigma_1]^{(q)}_{\mathbf{x}}}\pi^q_{q-1}$$

where again $\xi = \sigma(\mathbf{x}) \in \mathcal{E}$ and $V_\xi\pi \subseteq T_\xi\mathcal{E}$ is the *vertical space at ξ of the fibration over* π—it is defined as the kernel of the tangent map $T_\xi\pi \colon T_\xi\mathcal{E} \to T_{\mathbf{x}}\mathcal{X}$—and $V_{[\sigma_1]^{(q)}_{\mathbf{x}}}\pi^q_{q-1}$ is the *vertical space at $[\sigma_1]^{(q-1)}_{\mathbf{x}}$ of the fibration over π^q_{q-1}*, defined as the kernel of the tangent map $T_{[\sigma_1]^{(q)}_{\mathbf{x}}}\pi^q_{q-1} \colon T_{[\sigma_1]^{(q)}_{\mathbf{x}}} J_q\pi \to T_{[\sigma_1]^{(q-1)}_{\mathbf{x}}} J_{q-1}\pi$.

Proposition 2.1.2. *For any $\rho \in J_{q-1}\pi$, the jet-fibre at ρ of order q, $(J_q\pi)_\rho$, is an affine space with underlying vector space $V^{(q)}_\xi := S_q(T^*_{\mathbf{x}}\mathcal{X}) \otimes V_\xi\pi$. The jet-space of order q is an affine bundle*

$$\pi^q_{q-1} \colon J_q\pi \to J_{q-1}\pi$$

with underlying vector bundle $S_q(T^*\mathcal{X}) \otimes_{J_{q-1}\pi} V\pi \to J_{q-1}\pi$. Its dimension is

$$\dim J_q\pi = n + m \binom{n+q}{q} \ .$$

Proof. The proof is by checking how coordinate changes in the total space $\mathcal{E}$ which map fibres into fibres influence the derivatives, which are the coordinates in the jet bundle $J_q\pi$. Let $\mathbf{x} \leftarrow \tilde{\mathbf{x}}(\mathbf{x})$ and $\mathbf{u} \leftarrow \tilde{\mathbf{u}}(\mathbf{x}, \mathbf{u})$ be such a coordinate change. Then the chain rule implies that for the derivatives of highest order the induced change of coordinates yields as the new coordinates

$$\tilde{u}^{\alpha}_{j_1\cdots j_q} = \left(\frac{\partial \tilde{u}^\alpha}{\partial u^\beta}\frac{\partial x^{i_1}}{\partial \tilde{x}^{j_1}}\cdots\frac{\partial x^{i_q}}{\partial \tilde{x}^{j_q}}\right) u^{\beta}_{i_1\cdots i_q} + R \ .$$

where summation over repeated indices is understood, $(\frac{\partial \mathbf{x}}{\partial \tilde{\mathbf{x}}})$ means the inverse of the Jacobian matrix $(\frac{\partial \tilde{\mathbf{x}}}{\partial \mathbf{x}})$ and R denotes the terms of lower order; these do not depend on derivatives of order q (they depend on the x^i and u^α_μ where $1 \le i \le n$, $1 \le \alpha \le m$ and $0 \le |\mu| < q$). This is an affine function of the derivatives $u^{\beta}_{i_1\cdots i_q}$ of order q. Therefore $J_q\pi$ is affine over $J_{q-1}\pi$. See Saunders [36], Theorem 6.2.9, for details, or Pommaret [35], Propositions 1.9.7 and 1.9.9.

For the dimension, we have $\dim J_q\pi - \dim J_{q-1}\pi = m\binom{n+q-1}{q}$, from which follows $\dim J_q\pi = n + m\binom{n+q}{q}$. □

Remark 2.1.3. Let $\pi\colon \mathcal{E} \to \mathcal{X}$ be a fibred manifold. The jet-space of order q for $q = 0$ and $q = 1$ may be interpreted in a sense which is especially suitable for our oncoming analysis. For $q = 0$ we have $J_0\pi = \mathcal{E}$ and $\pi^q_0 = \mathrm{id}_{\mathcal{E}}$ and $\pi^q := \pi$. For $q = 1$ consider a point $\xi \in \mathcal{E}$ where $\pi\colon \mathcal{E} \to \mathcal{X}$ yields $\pi(\xi) = \mathbf{x}$. Then the jet-fibre at ρ of order 1 is

$$(J_1\pi)_\xi := \{\lambda \in T^*_{\mathbf{x}}\mathcal{X} \otimes T_\xi\mathcal{E} \colon T_\xi\pi \circ \lambda = \mathrm{id}_{T_{\mathbf{x}}\mathcal{X}}\} \ , \tag{2.1}$$

the jet-space of order 1 is

$$J_1\pi := \bigcup_{\xi\in\mathcal{E}} \{\rho\} \times (J_1\pi)_\xi \ ,$$

and a point in it is a 1-jet. Now the first-order jet bundle over $\mathcal{E}$ is the affine bundle $\pi^1_0\colon J_1\pi \to \mathcal{E}$ with its fibre at the point ξ given by the affine space given in Equation (2.1) and the underlying vector space

$$T^*_{\mathbf{x}}\mathcal{X} \otimes V_\xi\pi \ .$$

Now we have two ways to look at the first-order jet bundle; therefore we have to show their equivalence.

First let be $\sigma \in \Gamma_L\pi$. Setting $\pi \circ \pi^1_0 =: \pi^1\colon J_1\pi \to \mathcal{X}$, we see that $J_1\pi$ is fibred over $\mathcal{X}$. Then define $j_1\sigma\colon \mathcal{X} \to J_1\pi$ by $j_1\sigma(\mathbf{x}) := (\sigma(\mathbf{x}), T_{\mathbf{x}}\sigma)$ for its first prolongation. As $\sigma\colon \mathcal{X} \to \mathcal{E}$, pointwise for $\sigma(\mathbf{x}) = \xi$ we have $T_{\mathbf{x}}\sigma\colon T_{\mathbf{x}}\mathcal{X} \to T_\xi\mathcal{E}$ for its linearization. Therefore $T_{\mathbf{x}}\sigma \in T^*_{\mathbf{x}}\mathcal{X} \otimes T_\xi\mathcal{E}$. Furthermore, the definition of a section, $\pi\circ\sigma = \mathrm{id}_{\mathcal{X}}$, implies $T_\xi\pi \circ T_{\mathbf{x}}\sigma = \mathrm{id}_{T_{\mathbf{x}}\mathcal{X}}$, which is the motive for the definition in Equation (2.1).

Conversely, for any $\lambda \in T^*_{\mathbf{x}}\mathcal{X} \otimes T_\xi\mathcal{E}$, for each $\mathbf{x} \in \mathcal{X}$ there is a section $\sigma \in \Gamma_L\pi$ such that $\sigma(\mathbf{x}) = \xi$ and $T_x\sigma = \lambda$. Therefore, λ corresponds to the equivalence class $[\sigma]^{(1)}_{\mathbf{x}} = \{\sigma \in \Gamma_L\pi \colon \sigma(\mathbf{x}) = \xi \text{ and } T_{\mathbf{x}}\sigma = \lambda\}$.

The jet bundle of order $q > 1$ may be introduced by iteration as a subspace of the first order jet bundle of the previous jet bundle of order $q - 1$.

Any jet bundle of order q has the property $\pi^q(\mathcal{N}) = \mathcal{X}$ and therefore is a fibred manifold $\pi^q\colon J_q\pi \to \mathcal{X}$. For $r \geq 1$, $J_{q+r}\pi$ can be regarded as a fibred submanifold of $J_r(J_q\pi)$ by the imbedding $\iota_{q,r}\colon J_{q+r}\pi \hookrightarrow J_r\pi^q = J_r(J_q\pi)$, defined through

$$\iota_{q,r} \circ j_{q+r}(\sigma) = j_r(j_q\sigma) \tag{2.2}$$

for all $\sigma \in \Gamma_L\pi$. In local coordinates on $J_r(J_q\pi)$, it is given by setting $u^\alpha_{\mu,\nu} = u^\alpha_{\bar{\mu},\bar{\nu}}$ for all $1 \leq \alpha \leq m$, $0 \leq |\mu|, |\bar{\mu}| \leq q$ and $0 \leq |\nu|, |\bar{\nu}| \leq r$ if $\mu + \nu = \bar{\mu} + \bar{\nu}$. This means that one does not distinct derivatives which are equal except for the order of their independent coordinates any more; in $J_r(J_q\pi)$, the derivatives $u_{\mu,\nu}$ and $u_{\nu,\mu}$ are regarded as different while in $J_{q+r}\pi$ there is only $u_{\mu+\nu}$, which equals $u_{\nu+\mu}$.

Let $\mathbf{x} \in \mathcal{X}$ and $\sigma \in \Gamma_{\mathbf{x}}\pi$ Let $1 \leq r < q$. Then the mappings

$$\begin{aligned} \pi^q\colon J_q\pi &\to \mathcal{X}, & [\sigma]^{(q)}_{\mathbf{x}} &\mapsto \mathbf{x}, \\ \pi^q_0\colon J_q\pi &\to \mathcal{E}, & [\sigma]^{(q)}_{\mathbf{x}} &\mapsto \sigma(\mathbf{x}), \\ \pi^q_r\colon J_q\pi &\to J_r\pi, & [\sigma]^{(q)}_{\mathbf{x}} &\mapsto [\sigma]^{(r)}_{\mathbf{x}}, \end{aligned}$$

are called *source-*, *target-* and *jet-projection* of order q (and, in the last case, order r).

2.1.2 The Contact Structure

For a total space $\mathcal{E}$ with typical fibre $\mathcal{U}$ over the base space $\mathcal{X}$, each section $\sigma\colon \mathcal{X} \to \mathcal{U}$ can be prolonged to a section $j_q\sigma\colon \mathcal{X} \to J_q(\mathcal{X},\mathcal{U})$ by

$$\mathbf{x} \mapsto [\sigma]^{(q)}_{\mathbf{x}} \ .$$

Now the question arises, which sections of π^q are prolongations of sections of π? The answer is given by the *contact structure*, which each jet bundle brings along and which can be formulated by way of the contact map, the contact distribution or the contact codistribution. All of them characterize certain n-dimensional submanifolds $\mathcal{N} \subseteq J_q\pi$, which are fibred (over $\mathcal{X}$) because they have the property $\pi^q(\mathcal{N}) = \mathcal{X}$. The description here follows Modugno [23, 32].

Definition 2.1.4. *The mapping* $\Gamma_1\colon J_1\pi \times_{\mathcal{X}} T\mathcal{X} \to T\mathcal{E}$*, defined by*

$$(\xi, \lambda_\xi, v_{\mathbf{x}}) \mapsto (\xi, \lambda_\xi(v_{\mathbf{x}})) \ ,$$

is called the contact map of order 1*; more generally, the* contact map of order q *is* $\Gamma_q\colon J_q\pi \times_{\mathcal{X}} T\mathcal{X} \to T(J_{q-1}\pi)$*, defined by*

$$\Gamma_q = \Gamma_1 \circ \iota_{q-1,1} \ .$$

Remark 2.1.5. By definition the contact map can be regarded as an evaluation map for $q = 1$; it corresponds to $(\xi, \lambda_\xi) \mapsto (v_{\mathbf{x}} \mapsto \lambda_\xi(v_{\mathbf{x}}))$. The contact map of order q is a linear fibred morphism over π^q_{q-1}. Because of the linearity (in the argument $v_{\mathbf{x}}$), it can be regarded as a map $\Gamma_q\colon J_q\pi \to T^*\mathcal{X} \otimes_{J_{q-1}\pi} T(J_{q-1}\pi)$, in local coordinates given by

$$\Gamma_q(\mathbf{x}, \mathbf{u}^{(q)}) = \left(\mathbf{x}, \mathbf{u}^{(q-1)}, dx^i \otimes \big(\partial_{x^i} + \sum_{\alpha=1}^{m} \sum_{1 \leq |\mu| < q} u^\alpha_{\mu+1_i} \partial_{u^\alpha_\mu}\big)\right) \ . \tag{2.3}$$

The following proposition uses the contact map to show which sections $\gamma \in \Gamma_L\pi^q$ are the prolongation of sections $\sigma \in \Gamma_L\pi$.

Proposition 2.1.6. *A section $\gamma \in \Gamma_L\pi^q$ is a prolongation $\gamma = j_q\sigma$ of a section $\sigma \in \Gamma_L\pi$ if, and only if, for all points $x \in \mathcal{X}$ where γ is defined, we have*

$$\operatorname{im}\Gamma_q(\gamma(x)) = T_{\gamma(x)}\pi^q_{q-1}(T_{\gamma(x)}\operatorname{im}\gamma)\ .$$

Proof. Let $q=1$. Let $\gamma \in \Gamma_L\pi^1$. Then $\sigma := \pi^1_0 \circ \gamma \in \Gamma_L\pi$. We have $T_{\gamma(x)}\pi^1_0(T_{\gamma(x)}\operatorname{im}\gamma) = T_{\sigma(x)}\operatorname{im}\sigma = \operatorname{im}T_x\sigma$. Since $\gamma(x) = (\sigma(x), \gamma_{\sigma(x)})$, it follows $\operatorname{im}\Gamma_1(\gamma(x)) = \operatorname{im}\gamma_{\sigma(x)}$. For $\gamma = \sigma_1(x)$, we have $\gamma_{\sigma(x) = T_x\sigma}$, which proves the first direction.

For the other direction, let $\gamma_{\sigma(x)} = \operatorname{im}T_x\sigma$. Then there is a local section $\hat\sigma \in \Gamma_L\pi$ such that $\hat\sigma(x) = \sigma(x)$ and $T_x\hat\sigma = \gamma_{\sigma(x)}$ where $\gamma(x) = [\hat\sigma]^{(1)}_x$. Thus $\gamma(x) = [\sigma]^{(1)}_x$. Since x is arbitrary, the claim follows for $q=1$. For an arbitrary order q the argument is analogous but more tedious. □

Remark 2.1.7. Now we see that for any section $\sigma \in \Gamma_L\pi$, we have

$$\operatorname{im}\Gamma_{q+1}(j_{q+1}\sigma(\mathbf{x})) = \operatorname{im}T_{\mathbf{x}}(j_q\sigma)\ ;$$

in particular, for $q=1$, this means that for any $\sigma \in \Gamma_L\pi$, we have $\operatorname{im}\Gamma_1(j_1\sigma(\mathbf{x})) = \operatorname{im}T_{\mathbf{x}}\sigma$.

This may be interpreted as that the jet of order $q+1$ of $\sigma \in \Gamma_L\pi$ holds the same information as the jet of order q and its tangent space.

Let $\rho \in J_q\pi$ and consider $\operatorname{im}(\Gamma_q(\rho))$, the image of the contact map of order q at ρ:

Definition 2.1.8. *Let $\rho \in J_q\pi$ with $\pi^{q+1}_q(\hat\rho) = \rho$ for points $\hat\rho$ from its fibre. Then the vector fields in $T_\rho J_q\pi$ defined by*

$$(\mathcal{C}_q)_\rho := \operatorname{im}\Gamma_{q+1}(\hat\rho)$$

are called the contact vector fields of order q at ρ. *The vector distribution*

$$\mathcal{C}_q := \bigcup_{\rho\in J_q\pi} \{\rho\} \times (\mathcal{C}_q)_\rho$$

is called the contact distribution of order q.

Proposition 2.1.9. *Locally, the contact distribution is generated by two types of vector fields, namely, for $1 \le i \le n$,*

$$C^{(q)}_i := \partial_{x^i} + \sum_{\alpha=1}^{m} \sum_{0\le|\mu|<q} u^\alpha_{\mu+1_i}\partial_{u^\alpha_\mu}\ , \tag{2.4a}$$

and, for $1 \le \alpha \le m$ and $|\mu| = q$,

$$C^\mu_\alpha := \partial_{u^\mu_\alpha}\ . \tag{2.4b}$$

We have $\dim\mathcal{C}_q = \dim\pi^q - \dim\pi^{q-1} + \dim\mathcal{X} = n + m\binom{n+q-1}{q}$. *A map $\psi\colon \mathcal{X} \to J_q(\mathcal{X},\mathcal{U})$ is a prolongation $\psi = j_q\phi$ if, and only if, $T(\operatorname{im}\psi) \subseteq \mathcal{C}_q|_{\operatorname{im}\psi}$. (Which means if, and only if,* $\operatorname{im}\psi$ *is an integral manifold of the contact distribution, $\mathcal{C}_q$.)*

Corollary 2.1.10. *The contact distribution $\mathcal{C}_q$ can be split into*

$$\mathcal{C}_q = V\pi^q_{q-1} \oplus \mathcal{H}$$

where the complement $\mathcal{H}$ has the dimension of the base space, $\dim \mathcal{X} = n$, *and is not unique. Such a complement defines a connection on π^q.*

Proof. The possibility to split in such a way is obvious from the generating vector fields (2.4). Only the vertical bundle, generated by the vector fields (2.4b), is independent of local coordinates. □

Remark 2.1.11. The contact distribution is not closed under Lie brackets: for all $|\nu| = q-1$, we have $[C^{\nu+1_i}_\alpha, C^{(q)}_i] = \partial_{u^\alpha_\nu}$. The derived contact distribution $\mathcal{C}'_q := \mathcal{C}_q + [\mathcal{C}_q, \mathcal{C}_q]$ satisfies $\mathcal{C}'_q/\mathcal{C}_q \cong V(\pi^{q-1}_{q-2})$. If a distribution is closed under Lie brackets, it is called *involutive* or *completely integrable*. The Frobenius theorem states that an involutive D-dimensional distribution $\mathcal{V}$ of vector fields X_d, $1 \leq d \leq D$, on a manifold $\mathcal{R}$ of dimension E defines a local foliation of $\mathcal{R}$ with D-dimensional leaves which are *integral manifolds* of the distribution; this means, there are $E-D$ functions $\Phi^\tau \colon \mathcal{R} \to \mathbb{R}$ such that for all $1 \leq d \leq D$ and $1 \leq \tau \leq E-D$ we have

$$X_d \Phi^\tau = 0 \ , \tag{2.5}$$

and a set of real constants $c^\tau = \Phi^\tau$ defines a family of integral manifolds for $\mathcal{V}$. (In Chapter 3, $\mathcal{R}$ denotes a differential equation and $\mathcal{V}$ a distribution at once tangential to $\mathcal{R}$ and included in the contact distribution of a jet bundle containing $\mathcal{R}$.)

Definition 2.1.12. *The distribution of one-forms annihilating the contact distribution of order q is called the* contact codistribution of order q *and is denoted by $\mathcal{C}^0_q$.*

Proposition 2.1.13. *The contact codistribution $\mathcal{C}^0_q$ is spanned by the one-forms*

$$\omega^\alpha_\mu = du^\alpha_\mu - \sum_{i=1}^n u^\alpha_{\mu+1_i} dx^i \ , \quad 0 \leq |\mu| < q \ , 1 \leq \alpha \leq m \ . \tag{2.6}$$

They are called the contact forms.

Proof. It is obvious from the expressions in local coordinates (2.6) of the contact forms and those of the contact vector fields (2.4) that the spaces they generate are dual. □

2.2 Differential Equations

We now introduce differential equations as fibred submanifolds of suitably sized jet bundles and then consider their representations in local coordinates. This geometric description of a system of differential equations enables us to regard it as a set of algebraic equations for the coordinates of the jet bundle that contains the differential equation.

2.2.1 Differential Equations as Fibred Submanifolds

To describe a differential equation, there are two kinds of variables needed: the independent variables, for which we use the local coordinates of the base space $\mathcal{X}$, and the dependent ones, for which we use the local coordinates of the typical fibre. Though independent of coordinates, the following definition fulfills this end.

Definition 2.2.1. *Let $\pi\colon \mathcal{E} \to \mathcal{X}$ be a fibred manifold. A submanifold $\mathcal{R}_q \subset J_q\pi$ fibred over the base space $\mathcal{X}$ is called a* system of partial differential equations of order q *or simply a* differential equation *or a* system*. If the dimension of the base space is one, the differential equation is called* ordinary*.*

Here we regard $\mathcal{R}_q$ as a *regular* submanifold; that is, as a subset of $J_q\pi$ such that for any point $\rho \in \mathcal{R}_q$ there is a chart (U, ϕ_U) of the jet bundle which satisfies $\phi_U(\mathcal{R}_q \cap U) = \mathbb{R}^d \times \{\mathbf{0}\} \subseteq \mathbb{R}^D$—that is, $\mathcal{R}_q \cap U = (\mathbf{x} \in \mathbb{R}^d \times \{\mathbf{0}\}\colon x^{d+1} = \cdots = x^D = 0)$—where d and D denote the Dimensions of $\mathcal{R}_q$ and $J_q\pi$. A differential equation in this sense is also an immersed submanifold, if we use the canonical imbedding $\iota\colon \mathcal{R}_q \hookrightarrow J_q\pi$, as this defines an injective immersion. (An immersion for two manifolds M and N is a smooth map $f\colon M \to N$ such that for $\dim M \leq \dim N$ its rank is maximal, where the rank of f is (pointwise) defined by the rank of its tangent map $Tf\colon TM \to TN$.)

To give a local representation for a differential equation $\mathcal{R}_q$, let $\rho \in \mathcal{R}_q$. On a neighborhood U for ρ, the differential equation is the set of solutions to a system of equations

$$\mathcal{R}_q : \left\{ \Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0 \;\; , \quad 1 \leq \tau \leq t. \right. \tag{2.7}$$

The local representation of a system of differential equations may consist of a single equation; in this case, the term *scalar equation* is usual. It may also consist of several equations; in that case, the word "equation" has two different meanings, denoting the whole system and any single equation within it. But the meaning should always be clear from the context.

Example 2.2.2. The one-dimensional wave equation is a second-order system (of one equation) which in classical notation is written $\frac{\partial^2 u(t,x)}{\partial t^2} - \frac{\partial^2 u(t,x)}{\partial x^2} = 0$. It can be regarded as defining a submanifold $\mathcal{R}_2$ in the following second-order jet bundle. For the trivial bundle $(\mathbb{R}^2 \times \mathbb{R}, \mathrm{pr}_1, \mathbb{R}^2)$ let $(x, t; u)$ be global coordinates. Then the coordinates of the corresponding first-order jet bundle $J_2\mathrm{pr}_1$ are $(x, t; u; u_x, u_t; u_{xx}, u_{xt}, u_{tt})$. Let the map $\Phi : J_2\mathrm{pr}_1 \to \mathbb{R}$ be given by

$$\Phi(j_2\gamma|_{(x,t)}) = (u_{tt} - u_{xx})(j_2\gamma|_{(x,t)}) = 0 \;.$$

The corresponding differential equation is

$$\mathcal{R}_2 = \{ j_2\gamma|_{(x,t)} \in J_2\mathrm{pr}_1 \colon \Phi(j_2\gamma|_{(x,t)}) = 0 \}$$

or

$$\mathcal{R}_2\colon \left\{ \; u_{tt} - u_{xx} = 0 \right.$$

for short.

Analogously, the two-dimensional wave equation $\frac{\partial^2 u(x,y,t)}{\partial t^2} - \frac{\partial^2 u(x,y,t)}{\partial x^2} - \frac{\partial^2 u(x,y,t)}{\partial y^2} = 0$ may be regarded as a submanifold $\mathcal{R}_2$ in the second-order jet bundle of the trivial bundle $(\mathbb{R}^3 \times \mathbb{R}, \mathrm{pr}_1, \mathbb{R}^3)$, defined by the map $\Phi\colon J_2\mathrm{pr}_1 \to \mathbb{R}$ which is given by

$$\Phi(j_2\gamma|_{(x,y,t)}) = (u_{tt} - u_{xx} - u_{yy})(j_2\gamma|_{(x,y,t)}) = 0 \ .$$

In our short-hand notation this is

$$\mathcal{R}_2\colon \left\{ \ u_{tt} - u_{xx} - u_{yy} = 0 \ . \right.$$

The local representation (2.7) of a differential equation $\mathcal{R}_q$, according to which in a neighborhood U of a point $\rho \in U \subseteq \mathcal{R}_q$ the differential equation can be described as the solution space of the function Φ^τ, naturally leads to the following consideration.

Definition 2.2.3. *Let $(\partial\Phi^\tau/\partial u^\alpha_\mu)$ denote the Jacobian matrix of the representation (2.7) with its rank constant on U. Then its rank is called the* codimension of $\mathcal{R}_q$. *The dimension of the typical fibre of $\mathcal{R}_q$ regarded as a fibred over $\mathcal{X}$, is called the* dimension of $\mathcal{R}_q$.

The rank is the number of functionally independent mappings Φ^τ. Note that the dimension of a differential equation refers to the fibre dimension only and neglects the dimension n of the base space $\mathcal{X}$.

Lemma 2.2.4. *We have* $\dim \mathcal{R}_q = m \binom{q+n}{n} - \mathrm{codim}\mathcal{R}_q$.

Proof. The fibre dimension of $J_q\pi$ is $m \binom{q+n}{n}$. □

Definition 2.2.5. *A* (local) solution *is a smooth section $\sigma \in \Gamma_L\pi$ such that its prolongation satisfies* $\mathrm{im}\, j_q\sigma \subseteq \mathcal{R}_q$.

A word of warning is in place: The description of a differential equation as a submanifold in a jet bundle $J_q\pi$ for $q < \infty$ is not equivalent to a description of its solution space—there are cases where a differential equation contains another fibred submanifold of $J_q\pi$ as a proper submanifold and shares with this other differential equation the solution space. The reason is then that the larger differential equation contains a point which does not lie on any solution. (A differential equation is called *locally solvable* if it does not behave that way.) If a section σ is defined on some neighborhood $\mathcal{O}$ of a point $\mathbf{x}_0 \in \mathcal{X}$ and is a local solution of $\mathcal{R}_q$, too, then for its equivalence class $[\sigma]^{(q)}_{\mathbf{x}_0}$, the q-jet at $\mathbf{x}_0$, it follows $[\sigma]^{(q)}_{\mathbf{x}} \in \mathcal{R}_q$ for all $\mathbf{x} \in \mathcal{O}$. On the other hand, two sections are regarded as equivalent (according to Remark 2.1.1) if their corresponding functions have a contact of order q at a single point $\mathbf{x}$. Therefore, $[\sigma]^{(q)}_{\mathbf{x}} \in \mathcal{R}_q$ only means that the section σ solves the differential equation $\mathcal{R}_q$ at $\mathbf{x}$ up to order q without further information on higher orders or other points. So an equivalence class $[\sigma]^{(q)}_{\mathbf{x}}$ may contain sections that are not solutions of $\mathcal{R}_q$. This phenomenon is analyzed in Subsection 2.3.1.

Remark 2.2.6. In local coordinates, this coincides with the classical notion of a solution: if the local section σ is a solution, then there is an open subset $\mathcal{O} \subseteq \mathcal{X}$ and a smooth function $s\colon \mathcal{O} \to \mathcal{U}$ such that for all $1 \le \tau \le t$ and any $\mathbf{x} \in \mathcal{O}$ we have $\sigma(\mathbf{x}) = (\mathbf{x}, s(\mathbf{x}))$, and the equalities $\Phi^\tau(\mathbf{x}, j_q s(\mathbf{x})) = 0$ hold. The image of the q-prolongation of such a solution is $\{(\mathbf{x}, j_q s(\mathbf{x}))\colon \mathbf{x} \in \mathcal{O}\}$ and is contained in the subvariety $\{\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0\colon \mathbf{x} \in \mathcal{O}\} \subseteq J_q\pi$, which we identify with the differential equation.

Example 2.2.7. We continue with Example 2.2.2, the one-dimensional wave equation. Here, the function $s\colon \mathbb{R}^2 \to \mathbb{R}$, defined by $s(x,t) = x^3 + 3xt^2$, is a smooth solution (actually, a global one) since

$$\{(\mathbf{x}, j_q s(\mathbf{x})) : \mathbf{x} \in \mathcal{O}\} = \{\big(x,t; x^3+3xt^2; 3x^2+3t^2, 6xt; 6x, 6t, 6x\big) : (x,t) \in \mathbb{R}^2\}$$

is because of $\frac{\partial^2 s}{\partial x^2} - \frac{\partial^2 s}{\partial t^2} = 0$ a subset of

$$\begin{aligned}\{u_{xx} - u_{tt} = 0 : \mathbf{x} \in \mathbb{R}^2\} &= \{(\Phi^1(\mathbf{x}, \mathbf{u}^{(2)})) = 0 : \mathbf{x} \in \mathbb{R}^2\} \\ &= \{\Phi(\mathbf{x}, \mathbf{u}^{(2)}) = 0 : \mathbf{x} \in \mathcal{O}\}\ ,\end{aligned}$$

which in turn is contained in $J_2\mathrm{pr}_1$. Note that here we have a scalar equation, so $\Phi^1 = \Phi$.

Remark 2.2.8. If an ordinary differential equation of first order $\mathcal{R}_1$ is locally represented in the form $\dot{u}^\alpha = \phi^\alpha(\mathbf{x}, \mathbf{u})$, then there is a local section $\gamma\colon \mathcal{E} \to J_1\pi$ such that $\mathcal{R}_1$ can be written (locally) as $\mathcal{R}_1 = \operatorname{im}\gamma$. Thus, a differential equation of this kind defines a connection on $\mathcal{E}$. If equations of order zero are present as well, they define a constraint manifold $\mathcal{C} \subseteq \mathcal{E}$, and the differential equation (locally) defined as the image of the section $\gamma\colon \mathcal{C} \to J_1\pi$ corresponds to a connection on $\mathcal{C}$. For $q > 1$, an ordinary differential equation corresponds to a connection on $J_{q-1}\pi$ or on a constraint submanifold in it in an analogous way.

2.2.2 Prolongation and Projection of a Differential Equation

Since a differential equation $\mathcal{R}_q$ is a fibred submanifold in a jet bundle, there are two operations on it which appear natural: prolongation means to transform $\mathcal{R}_q$ into a subset of a higher order jet bundle, whereas projection means to transform it into a subset in a jet bundle of lower order.

Definition 2.2.9. *Let $\Phi \in \mathcal{F}(J_q(\mathcal{X},\mathcal{U}))$ be a smooth function. Then for all $1 \le i \le n$ the mapping $D_i\colon \mathcal{F}(J_q(\mathcal{X},\mathcal{U})) \to \mathcal{F}(J_{q+1}(\mathcal{X},\mathcal{U}))$, defined by*

$$(D_i\Phi)(\mathbf{x}, \mathbf{u}^{(q+1)}) = \frac{\partial \Phi(\mathbf{x}, \mathbf{u}^{(q)})}{\partial x^i} + \sum_{0 \le |\mu| \le q} \sum_{\alpha=1}^{m} \frac{\partial \Phi(\mathbf{x}, \mathbf{u}^{(q)})}{\partial u^\alpha_\mu} u^\alpha_{\mu+1_i}\ , \tag{2.8}$$

is called the formal *or the* total derivative of Φ with regard to x^i*.*

For a multi-index $\mu = (\mu_1, \mu_2, \ldots, \mu_n) \in \mathbb{N}_0^n$ the $|\mu|$-fold application of total derivatives yields a mapping $D_\mu\colon \mathcal{F}(J_q(\mathcal{X},\mathcal{U})) \to \mathcal{F}(J_{q+|\mu|}(\mathcal{X},\mathcal{U}))$, defined by

$$(D_\mu\Phi)(\mathbf{x}, \mathbf{u}^{(q+|\mu|)}) = D_1^{\mu_1}(D_2^{\mu_2} \ldots (D_n^{\mu_n}(\Phi(\mathbf{x}, \mathbf{u}^{(q)}))) \ldots).$$

Equation (2.8) amounts to applying the chain rule from calculus to derive the function Φ with regard to x^i. Note that the formal derivative is linear in the derivatives of highest order (that is, quasilinear). Since for smooth functions Φ cross-derivatives are equal, $D_iD_j\Phi = D_jD_i\Phi$, the formal derivative D_μ is well-defined for $|\mu| > 1$.

We use the notations $(D_\mu \Phi)(\mathbf{x}, \mathbf{u}^{(q+r)}) =: D_\mu(\Phi(\mathbf{x}, \mathbf{u}^{(q)})) =: D_\mu \Phi(\mathbf{x}, \mathbf{u}^{(q)})$.

By definition, $\mathcal{R}_q$ is fibred over $\mathcal{X}$. Let $\hat{\pi}^q \colon \mathcal{R}_q \to \mathcal{X}$ denote the corresponding projection. (Then $\hat{\pi}^q = \pi^q|_{\mathcal{R}_q}$.) Consider the jet bundle of order r over the differential equation, $J_r\hat{\pi}^q$. It can be imbedded naturally into the jet bundle of order r over the base space $J_q\pi$ by the map $\bar{\iota}_{q,r} \colon J_r\hat{\pi}^q \hookrightarrow J_r\pi^q$.

Definition 2.2.10. *The inverse image under the imbedding $\iota_{q,r} \colon J_{q+r}\pi \hookrightarrow J_r\pi^q$, defined by*

$$\mathcal{R}_{q+r} := \iota_{q,r}^{-1}(\bar{\iota}_{q,r}(J_r\hat{\pi}^q) \cap \iota_{q,r}(J_{q+r}\pi))$$

is called the rth prolongation of $\mathcal{R}_q$.

In local coordinates, a representation of the rth prolongation of $\mathcal{R}_q$ is computed by repeatedly applying the formal derivative (2.8) to the equations of the representation $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$ (where $1 \le \tau \le t$) of $\mathcal{R}_q$. A local representation for the rth prolongation is then given by

$$\mathcal{R}_{q+r} \colon \begin{cases} \Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0 \ , & 1 \le \tau \le t, \\ (D_\mu \Phi^\tau)(\mathbf{x}, \mathbf{u}^{(q+r)}) = 0 \ , & 1 \le |\mu| \le r \end{cases} . \tag{2.9}$$

Example 2.2.11. Consider the first-order system given by

$$\mathcal{R}_1 \colon \begin{cases} u_t - v = 0 \\ v_t - w_x = 0 \\ u_x - w = 0 \end{cases} .$$

Then its first prolongation is represented by

$$\mathcal{R}_2 \colon \begin{cases} u_t - v = 0, & u_{xt} - v_x = 0, & u_{tt} - v_t = 0, \\ v_t - w_x = 0, & v_{xt} - w_{xx} = 0, & v_{tt} - w_{xt} = 0, \\ u_x - w = 0, & u_{xx} - w_x = 0, & u_{xt} - w_t = 0. \end{cases} \tag{2.10}$$

Note that there are two different equations which hold the term u_{xt} in the representation of $\mathcal{R}_2$. From them follows an additional equation, namely $v_x = w_t$. We discuss this feature in Subsection 2.3.1.

While prolonging a differential equation raises its order, there is another kind of operation that lowers it.

Definition 2.2.12. *For a differential equation $\mathcal{R}_q \subseteq J_q\pi$ and for $r \le q$, the subset in $J_{q-r}\pi$, defined by*

$$\mathcal{R}_{q-r}^{(r)} = \pi_{q-r}^q(\mathcal{R}_q),$$

is called the rth projection of $\mathcal{R}_q$.

In local coordinates, a representation of the rth projection can, in principle, be computed from a local representation $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$ of $\mathcal{R}_q$ by using algebraic operations to eliminate all derivatives of order greater than $q-r$. In practice this may be hard, though.

Example 2.2.13. Reconsider the second-order system given by the representation (2.10). Then from $u_{xt} - v_x = 0$ and $u_{xt} - w_t = 0$ follows that $w_t = v_x$ by elimination of the second-order derivative u_{xt}. Thus, its first projection is represented by

$$\mathcal{R}_1^{(1)} : \begin{cases} u_t - v = 0 \\ v_t - w_x = 0 \\ w_t - v_x = 0 \\ u_x - w = 0 \end{cases} .$$

Note that projecting a differential equation may lead to singularities, and the result of a prolongation need not be a submanifold either. Those differential equations where prolongations and projections always yield submanifolds are called *regular*. For convenience we assume from now on that we deal with regular differential equations. In effect this means we retreat to suitable submanifolds within the images of these operations.

2.3 Formal Integrability

The notion of formal integrability is one of two main concepts in the formal theory (the other one being involutivity). It means that any point of the differential equation lies on at least one solution to the equation. A local representation of the differential equation then does not conceal any information about the space of formal solutions to the differential equation; but in general the equations in the representation of a differential equation may imply differential relations between them (as opposed to algebraic relations).

2.3.1 Formally Integrable Systems and Formal Solutions

Example 2.2.13 shows that prolongation and projection are not inverse operations: actually, in this example $\mathcal{R}_1^{(1)}$ is a proper subset of $\mathcal{R}_1$ since the prolongation and following projection returned for its representation the additional equation $w_t = v_x$. It arose from the equality of the mixed second-order derivatives $u_{tx} = v_x$ and $u_{xt} = w_t$. Another possible source for additional conditions for a representation are equations of an order lower than q in the representation of $\mathcal{R}_q$.

Example 2.3.1. Prolonging the second-order system

$$\mathcal{R}_2 : \begin{cases} u_{tt} = 0 \\ u_x = 0 \end{cases}$$

and then projecting $\mathcal{R}_3$ yields

$$\mathcal{R}_2^{(1)} : \begin{cases} u_{tt} = 0 \\ u_{xt} = 0 \\ u_x = 0 \end{cases} ,$$

where the condition $u_{xt} = 0$ arises from prolonging the first-order equation $u_x = 0$ to $D_t u_x = 0$ which is of second order and thus remains when projecting $\mathcal{R}_3$ to $J_2\pi$.

Such additional equations represent *integrability conditions*. Differential equations for which at no order of prolongation and projection integrability conditions arise are of special interest. While the expressions for such additional equations of the representation of a differential equation are dependent on the coordinates, it is an intrinsic property of the differential equation, considered as a geometric object, if local representations of it admit such additional equations or not.

Definition 2.3.2. *A differential equation $\mathcal{R}_q$ is called* formally integrable *if for all $r \geq 0$ the equality $\mathcal{R}_{q+r} = \mathcal{R}^{(1)}_{q+r}$ is satisfied.*

This definition is independent of local coordinates. It means that at no order prolongation and projection yield a proper submanifold. If they did, the given differential equation $\mathcal{R}_q$ would contain at least one point which would not lie on a solution to $\mathcal{R}_q$, since $\mathcal{R}_{q+r}$ and $\mathcal{R}^{(1)}_{q+r}$ have the same spaces of formal solutions. This means, if a differential equation is formally integrable, a formal solution to it can be constructed by a power series ansatz which at each order yields a correct truncation of the solution as follows. Let the system (2.7) be a local representation of $\mathcal{R}_q$. Set $\mu! = \prod_{i=1}^{n} \mu_i!$ and $(\mathbf{x}-\mathbf{x}_0)^\mu = \prod_{i=1}^{n}(x^i - x^i_0)^{\mu_i}$ as usual. A formal solution to $\mathcal{R}_q$ is locally given by its Taylor series

$$\tilde{u}^\alpha(\mathbf{x}) = \sum_{|\mu|=0}^{\infty} \frac{c^\alpha_\mu}{\mu!}(\mathbf{x}-\mathbf{x}_0)^\mu$$

with coefficient functions c^α_μ which cannot be calculated uniquely without initial or boundary conditions but which have to satisfy the equations of the system which represents $\mathcal{R}_q$ and all its prolongations. (The tilde above u^α indicates that we do not assume convergence of the Taylor series, and in particular we do not assume that its limit be u^α.) To show these interrelations explicitly, we enter the power series ansatz into the representation of $\mathcal{R}_q$ and then evaluate the equations at the point $\mathbf{x} = \mathbf{x}_0$. Here we use

$$\tilde{u}^\alpha_i(\mathbf{x}) = \sum_{|\mu|=0}^{\infty} \frac{c^\alpha_{\mu,i}}{\mu!}(\mathbf{x}-\mathbf{x}_0)^\mu$$

for all $1 \leq i \leq n$ and its analogues for higher order derivatives. Thus, we obtain a set of algebraic equations

$$\Phi^\tau(\mathbf{x}_0, \mathbf{c}^{(q)}) = 0\ , \quad 1 \leq \tau \leq t, \tag{2.11}$$

for the coefficient functions $\mathbf{c}^{(q)} := (c^\alpha_\mu \colon 1 \leq \alpha \leq m, 0 \leq |\mu| \leq q)$ up to order q. To obtain the interrelations for the $\mathbf{c}^{(q+r)}$ where $r \geq 1$, we proceed order by order. If for $r-1$ the interrelations between the coefficient function up to order $q+r-1$ are given by a set of equations (see Equation (2.11) for $r = 1$), the interrelations between them and the coefficient functions of order $q+r$ are given by that set of equations and the additional set of equations

$$(D_\mu \Phi^\tau)(\mathbf{x}_0, \mathbf{c}^{(q+r)}) = 0\ , \qquad |\mu| = r, \tag{2.12}$$

for all $1 \leq \tau \leq t$, as follows from the local representation (2.9) of the prolonged equation $\mathcal{R}_{q+r}$.

As the formal derivative (2.8) is quasilinear, for $r \geq 1$ the systems (2.12) consist of inhomogeneous linear equations for the coefficients of order $q+r$. If these systems are taken into account for all $r \geq 1$, this leads to a potentially infinitely large system for potentially infinitely many coefficients. Each solution to this system is a formal solution to the original differential equation $\mathcal{R}_q$ (and, conversely, any formal solution to $\mathcal{R}_q$ may be expressed in such a way). So long as we consider only formal solutions, there is no difference between formal integrability and local solvability. But if the solutions are supposed to be analytic, local solvability is the stronger concept of the two.

The space of formal solutions can be described using *parametric coefficients*. They can be considered as the remaining coefficients after solving each (independent) equation of the system for one of the coefficients c^α_μ, which then are called *principal*. They are determined by the differential equation and its prolongations. Their choice may be arbitrary or a matter of additional information like initial conditions.

Such a classification of the coefficients c^α_μ in each step up to some order $q+r$ yields a truncation of a formal solution. But this order by order power series ansatz is only satisfying for a differential equation which is formally integrable, since otherwise the interrelations which hold between the coefficients of an order greater than $q+r$ concern the coefficients of an order up to $q+r$, too, by means of the integrability conditions.

Example 2.3.3. The representation for the differential equation $\mathcal{R}^{(1)}_1$ in Example 2.2.13 contained the integrability condition $w_t = v_x$, which we found in Example 2.2.11. Substituting the power series ansatz into this representation and evaluating at $\mathbf{x}_0$ yields the equations

$$c^u_{(0,1)} = c^v_{(0,0)}, \quad c^v_{(0,1)} = c^w_{(1,0)}, \quad c^w_{(0,1)} = c^v_{(1,0)}, \quad c^u_{(1,0)} = c^w_{(0,0)}$$

for the coefficients up to first order. This system is under-determined, and there is no obligatory choice of coefficients for which to solve these equations. (The one given here stems from the special form of the local representation.) But without the integrability condition $w_t = v_x$ in the representation of $\mathcal{R}^{(1)}_1$, the equation $c^w_{(0,1)} = c^v_{(1,0)}$ would be missing, and thus the power series ansatz would not result in an adequate first-order truncation of a formal solution. In this example, this could be noted already in the next step, when the terms of second order are considered as well. The four equations for the coefficients up to first order are then augmented by eight additional equations

$$\begin{aligned} &c^u_{(0,2)} = c^v_{(0,1)}, \quad c^v_{(0,2)} = c^w_{(1,1)}, \quad c^w_{(0,2)} = c^v_{(1,1)}, \quad c^u_{(1,1)} = c^w_{(0,1)}, \\ &c^u_{(1,1)} = c^v_{(1,0)}, \quad c^v_{(1,1)} = c^w_{(2,0)}, \quad c^w_{(1,1)} = c^v_{(2,0)}, \quad c^u_{(2,0)} = c^w_{(1,0)}, \end{aligned} \tag{2.13}$$

which arise by calculating $D_i\Phi^\tau$ for $1 \leq \tau \leq 4$ and $i \in \{x, t\}$ and again evaluating in $\mathbf{x}_0$. Gaussian elimination would now yield the ignored condition $c^w_{(0,1)} = c^v_{(1,0)}$ for two coefficients of first order, showing that our truncation after the first-order terms contained too many degrees of freedom for the first-order coefficients. Note that the combined system of twelve equations for the coefficients up to second order is under-determined, too, and there is no distinguished classification into principal and parametric coefficients. At any step in the order-by-order construction of a formal solution the principal coefficients can be chosen arbitrarily.

Remark 2.3.4. *Note that a differential equation which is not formally integrable may still have formal solutions. For such a differential equation, these formal solutions just cannot be constructed by way a the formal power series ansatz.*

2.3.2 The Geometric Symbol

The jet-space $J_q\pi$ is according to Proposition 2.1.2 an affine bundle over $J_{q-1}\pi$. The corresponding vertical bundle $V\pi^q_{q-1}$ helps us to analyze properties of a given differential equation $\mathcal{R}_q$ which follow from the equations of highest order in a local representation of $\mathcal{R}_q$. We therefore introduce the following brute-force linearization of a differential equation.

Definition 2.3.5. *Let $\mathcal{R}_q \subseteq J_q\pi$ be a differential equation and $\rho \in \mathcal{R}_q$. Let $\iota\colon \mathcal{R}_q \hookrightarrow J_q\pi$ be the inclusion map, given by $\iota(\rho) = \rho$. The vector space $(\mathcal{N}_q)_\rho$, implicitly defined by*

$$T\iota((\mathcal{N}_q)_\rho) := T\iota(T_\rho\mathcal{R}_q) \cap V_\rho\pi^q_{q-1}$$

is called the (geometric) symbol of the differential equation $\mathcal{R}_q$ at the point ρ.

The family of vector spaces

$$\mathcal{N}_q := \bigcup_{\rho\in\mathcal{R}_q} \{\rho\} \times (\mathcal{N}_q)_\rho$$

is called the (geometric) symbol of the differential equation $\mathcal{R}_q$.

Remark 2.3.6. Explicitly, we have $(\mathcal{N}_q)_\rho = V_\rho(\pi^q_{q-1}|_{\mathcal{R}_q})$ and $(\mathcal{N}_q)_\rho = T_\rho\mathcal{R}_q \cap (V_\rho\pi^q_{q-1})|_{\mathcal{R}_q}$.

This family of vector spaces may not be a bundle since the dimension of the symbol may be different for different points $\rho \in \mathcal{R}_q$. From now on, we assume the dimension to be equal for all ρ. This amounts to restrictions onto proper subsets of $\mathcal{R}_q$ where each dimension is constant and thus each symbol $\mathcal{N}_q$ is a vector-bundle over this restriction of $\mathcal{R}_q$.

Proposition 2.3.7. *If a differential equation $\mathcal{R}_q \subseteq J_q\pi$ is locally represented by a system (2.7), then the corresponding symbol at the point $\rho \in \mathcal{R}_q$ is the solution space of the following system of t linear equations:*

$$\sum_{\alpha=1}^{m} \sum_{|\mu|=q} \left.\frac{\partial \Phi^\tau}{\partial u^\alpha_\mu}\right|_\rho v^\alpha_\mu = 0 \tag{2.14}$$

where $1 \leq \tau \leq t$.

Proof. Let $\rho \in \mathcal{R}_q$, and for a neighborhood of ρ choose local coordinates $(\mathbf{x}, \mathbf{u}^{(q)})$ for $J_q\pi$. Let be $(\mathbf{x}, \mathbf{u}^{(q)}; \dot{\mathbf{x}}, \dot{\mathbf{u}}^{(q)})$ the induced coordinates for the tangent space $T_\rho J_q\pi$. Then any vector $X \in T_\rho J_q\pi$ has the form $X = \dot{x}^i\partial_{x^i} + \dot{u}^\alpha_\mu \partial_{u^\alpha_\mu}$. If it is to be tangent to $\mathcal{R}_q$ at ρ, it has to satisfy the condition $X\Phi^\tau|_\rho = 0$ for the local representation (2.7) of $\mathcal{R}_q$. Written out, this is the system of linear equations

$$\sum_{i=1}^{n} \frac{\partial \Phi^\tau}{\partial x^i}\bigg|_\rho \dot{x}^i + \sum_{\alpha=1}^{m} \sum_{1\le|\mu|\le q} \frac{\partial \Phi^\tau}{\partial u^\alpha_\mu}\bigg|_\rho \dot{u}^\alpha_\mu = 0 \tag{2.15}$$

in the unknowns $\dot{x}^i$ and $\dot{u}^\alpha_\mu =: v^\alpha_\mu$ of which there are $n+m\binom{q+n-1}{n-1}$ since the sum runs over all i, α and μ. By definition the symbol is the vertical part of the tangent space $T_\rho J_q\pi$ and thus the solution space for those equations in the system (2.15) where $\dot{\mathbf{x}} = \dot{\mathbf{u}}^{(q-1)} = 0$. This yields the system (2.14). □

The equations (2.14) are called *symbol equations*, the corresponding matrix of coefficients *symbol matrix*; the latter is denoted by $M_q(\rho)$ or simply M_q when it is clear from the context which point ρ is meant or ρ does not matter. We assume anyway the rank of the symbol matrix to be constant over all of $\mathcal{R}_q$.

Example 2.3.8. The symbol matrix for the representation of the differential equation $\mathcal{R}_1^{(1)}$ in Example 2.2.13, which was derived from turning the one-dimensional wave equation of Example 2.2.2 into a first-order system and then making that first-order system formally integrable, is:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

The columns are indexed from left to right by u_t, v_t, w_t, u_x, v_x and w_x. The rows are ordered according to τ, like the equations in the representation of $\mathcal{R}_1^{(1)}$. A basis for the symbol space at each point of the differential equation is given by the two vector fields $\partial_{v_x} + \partial_{w_t}$ and $\partial_{v_t} + \partial_{w_x}$.

In principle, it is a matter of taste how to order the rows and columns in the symbol matrix. Since we describe row transformations in our oncoming analysis, we have to decide on an order to fix the notation. We choose an order which yields the symbol matrices of the systems which we are going to study in row echelon form. For m dependent variables u^α, there are $\dim S_q(T^*\mathcal{X}) \otimes_{J_{q-1}\pi} V\pi = m\binom{q+n-1}{q}$ derivatives of order q. We label the $m\binom{q+n-1}{q}$ columns of the symbol matrix M_q by the derivatives v^α_μ, which appear in Equation (2.14) or simply by their indices (α, μ) to order them; the order of the v^α_μ and of the index pairs (α, μ) is defined as follows.

Definition 2.3.9. *For a multi-index $\mu = (\mu_1, \mu_2, \dots, \mu_n)$ the non-negative integer* $\operatorname{cl}\mu := \min\{i : \mu^i \neq 0\}$ *is called the* class *of μ. (It is the leftmost entry different from zero in $\mu = (\mu_1, \mu_2, \dots, \mu_n)$.) The multi-index containing zeros exclusively is not said to have a class.*

The class *of anything (for example jet coordinates u^α_μ, rows or columns) indexed with μ is the class of its multi-index μ.*

For a multi-index $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$ a variable x^i with $i \leq \operatorname{cl}\mu$ is called multiplicative*, otherwise it is called* non-multiplicative.

For a ring R, each multi-index μ defines naturally a set of up to n monomials $X_i^{\mu_i}$ in the ring of polynomials $R[X_1, \ldots, X_n]$. For any $r \in R \setminus \{0\}$, a polynomial of the form $rX_i^{\mu_i}$ is called a *term*.

Remark 2.3.10. The following convention defines a total order $\preceq$ on the set of jet-coordinates (see Adams and Loustaunau [1] and Calmet, Hausdorf and Seiler [5]): let α and β denote indices for the dependant coordinates, and let μ and ν denote multi-indices for marking derivatives. Derivatives of higher order are greater than derivatives of lower order: if $|\mu| < |\nu|$, then $u^\alpha_\mu \prec u^\beta_\nu$. If derivatives have the same order $|\mu| = |\nu|$, then we distinct two cases: if the leftmost non-vanishing entry in $\mu - \nu$ is positive, then $u^\alpha_\mu \prec u^\beta_\nu$; and if $\mu = \nu$ and $\alpha < \beta$, then $u^\alpha_\mu \prec u^\beta_\nu$.

This is a class-respecting order: if $|\mu| = |\nu|$ and $\operatorname{class}\mu < \operatorname{class}\nu$, then $u^\alpha_\mu \prec u^\beta_\nu$. Any set of objects indexed with pairs (α, μ) can be ordered in an analogous way. This order of the multi-indices μ and ν is called the *degree reverse lexicographic ranking*, and we generalize it in such a way that it places more weight on the multi-indices μ and ν than on the numbers α and β of the dependent variables. In the literature [1, 5], this is called the *term-over-position lift of the degree reverse lexicographic ranking* on the following grounds. If we identify u^α for $1 \leq \alpha \leq m$ with the unit vector $\mathbf{e}_\alpha = (\delta_{\alpha\beta} \colon 1 \leq \beta \leq m)$, and if for terms $rX_i^{\mu_i} =: t$ we consider the vectors $t\mathbf{e}_\alpha$, then the order defined above defines an order for such vectors which ranks the comparison of the terms over the comparison of their positions in the vector. For terms of the same degree, the only class respecting term order is the degree reverse lexicographic ranking. For details, see Seiler [37].

Definition 2.3.11. *In the representation (2.7) of a differential equation $\mathcal{R}_q$ for any τ the* class *of $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$ is the class of that variable u^α_μ which is maximal among all u^β_ν in that equation with respect to the degree reverse lexicographic ranking. (This includes the possibility that there are $1 \leq \alpha < \beta \leq m$ such that both u^α_μ and u^β_μ appear in $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$ and their common multi-index is maximal.)*

Example 2.3.12. This example is taken from Seiler [38], Example 1.3.4. For the differential equation represented by

$$\mathcal{R}_1 \colon \begin{cases} \Phi^1(x, y, z, u, u_x, u_y, u_z) = u_z + yu_x = 0 \\ \Phi^2(x, y, z, u, u_x, u_y, u_z) = u_y \qquad\quad = 0 \end{cases},$$

equation Φ^1 is of class 3, and equation Φ^2 is of class 2. The coordinates for $J_2\pi$ are in ascending order: $u_{(0,0,0)} < u_{(1,0,0)} < u_{(0,1,0)} < u_{(0,0,1)} < u_{(2,0,0)} < u_{(1,1,0)} < u_{(1,0,1)} < u_{(0,2,0)} < u_{(0,1,1)} < u_{(0,0,2)}$, or alternatively $u < u_x < u_y < u_z < u_{xx} < u_{xy} < u_{xz} < u_{yy} < u_{yz} < u_{zz}$. The first prolongation of $\mathcal{R}_1$ is represented by

$$\mathcal{R}_2 \colon \begin{cases} u_z + yu_x = 0, & u_{xz} + yu_{xx} = 0, & u_{yz} + yu_{xy} + u_x = 0, & u_{zz} + yu_{xz} = 0, \\ u_y = 0, & u_{xy} = 0, & u_{yy} = 0, & u_{yz} = 0. \end{cases} \tag{2.16}$$

Here the two equations $D_1\Phi^1$ and $D_1\Phi^2$ are of class 1, the four equations Φ^2, $D_2\Phi^1$, $D_2\Phi^2$ and $D_3\Phi^2$ are of class 2, and the two equations Φ^1 and $D_3\Phi^1$ are of class 3.

Remark 2.3.13. From now on, we use the following convention: the columns within the symbol matrix M_q are ordered descendingly according to the degree reverse lexicographic ranking for the multi-indices μ of the variables v^α_μ in Equation (2.14) and labelled by the pairs (α, μ). (It follows that, if v^α_μ and v^β_ν are such that $\operatorname{cl}\mu > \operatorname{cl}\nu$, then the column corresponding to v^α_μ is left of the column corresponding to v^β_ν.) If derivatives have the same order $|\mu| = |\nu|$, then we distinct two cases again: if the leftmost non-vanishing entry in $\mu - \nu$ is positive, then $u^\alpha_\mu \prec u^\beta_\nu$ like in the term-over-position lift of the degree reverse lexicographic ranking; and if $\mu = \nu$ and $\alpha > \beta$, then $u^\alpha_\mu \prec u^\beta_\nu$ which is *the opposite* of what the term-over-position lift of the degree reverse lexicographic ranking does. The rows are ordered in the same way with regard to the pairs (α, μ) of the variables u^α_μ which define the classes of the equations $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$. If two rows are labelled by the same pair (α, μ), it does not matter which one comes first. This order of rows and columns guarantees that the symbol matrices for the systems we are going to consider are in row echelon form.

Lemma 2.3.14. *The dimension of the symbol is*

$$\dim \mathcal{N}_q = \dim(S_q T^*\mathcal{X} \otimes_{J_{q-1}\pi} V\pi) - \operatorname{rank} M_q. \tag{2.17}$$

Proof. We have $S_q T^*\mathcal{X} \otimes_{J_{q-1}\pi} V\pi \cong V\pi^q_{q-1}$, and Equation (2.17) is the dimension formula for systems of linear equations. $\square$

The dimension of the symbol may be zero. Differential equations with this property are called *of finite type* (because their formal solutions space is of finite dimension) or *maximally over-determined* (because for any derivative of highest order, a local representation contains an equation solved for this derivative).

Remark 2.3.15. For differential equations of finite type which do not constrain $J_{q-1}\pi$ we can generalize Remark 2.2.8 since they represent images of global sections $\gamma\colon J_{q-1}\pi \to J_q\pi$ and therefore define a connection on $J_{q-1}\pi$. Thus, they can locally be represented by equations $\mathbf{u}_{(q)} = \phi(\mathbf{x}, \mathbf{u}^{(q-1)})$ where the vector $\mathbf{u}_{(q)}$ contains exactly the derivatives of order q. (Hence the name *maximally over-determined.*) In case the system contains equations of lower order, these again define a constraint manifold $\mathcal{C} \subset J_{q-1}\pi$ and the differential equation represents a connection on $\mathcal{C}$.

If the differential equation is not maximally over-determined, it may by regarded as covered by infinitely many such systems, where the symbol gives the degrees of freedom, defining the parametrization.

For $s \geq 1$ the prolonged symbols $\mathcal{N}_{q+s} \subset S_{q+s}T^*\mathcal{X} \otimes_{J_{q+s-1}\pi} V\pi$ can be derived from $\mathcal{N}_q$, without first calculating a local representation of the prolonged system $\mathcal{R}_{q+s}$. More precisely, to each row within the symbol matrix M_q there correspond n rows within the symbol matrix M_{q+1}; and to each column within the symbol matrix M_q there correspond $m \cdot \binom{n+q}{n-1}$ columns within the symbol matrix M_{q+1}. The next lemma gives the details.

Lemma 2.3.16. *Let the differential equation $\mathcal{R}_q$ be locally represented by (2.7). Then for the prolonged symbol $\mathcal{N}_{q+s}$ there are the following symbol equations:*

$$\sum_{\alpha=1}^{m}\sum_{|\mu|=q}\frac{\partial\Phi^\tau}{\partial u_\mu^\alpha}v_{\mu+\nu}^\alpha = 0 \ , \qquad 1\le\tau\le t,\ |\nu| = s \ .$$

For the special case $s=1$ regarding $\mathcal{N}_{q+1}$, the $t\cdot n$ symbol equations are:

$$\sum_{\substack{\alpha=1\\|\mu|=q}}^{m}\frac{\partial\Phi^\tau}{\partial u_\mu^\alpha}v_{\mu,i}^\alpha = 0 \ , \qquad 1\le\tau\le t,\ 1\le i\le n. \tag{2.18}$$

Proof. First let $s=1$. Then the equations for the prolonged equation are

$$\mathcal{R}_{q+1}\colon \begin{cases} \Phi^\tau(\mathbf{x},\mathbf{u}^{(q)}) = 0 \ , & 1\le\tau\le t, \\ (D_i\Phi^\tau)(\mathbf{x},\mathbf{u}^{(q+1)}) = 0 \ , & 1\le i\le n. \end{cases}$$

Thus, according to Equation (2.14), the equations for the prolonged symbol are

$$\sum_{\alpha=1}^{m}\sum_{|\mu|=q+1}\frac{\partial(D_i\Phi^\tau)(\mathbf{x},\mathbf{u}^{(q+1)})}{\partial u_\mu^\alpha}v_\mu^\alpha = 0. \tag{2.19}$$

According to the chain rule in Equation (2.8) for the formal derivative

$$(D_i\Phi^\tau)(\mathbf{x},\mathbf{u}^{(q+1)}) = \frac{\partial\Phi^\tau(\mathbf{x},\mathbf{u}^{(q)})}{\partial x^i} + \sum_{\alpha=1}^{m}\sum_{0\le|\mu|\le q}\frac{\partial\Phi^\tau(\mathbf{x},\mathbf{u}^{(q)})}{\partial u_\mu^\alpha}u_{\mu+1_i}^\alpha \ .$$

It follows for $|\mu| = q+1$ that

$$\frac{\partial(D_i\Phi^\tau)(\mathbf{x},\mathbf{u}^{(q+1)})}{\partial u_\mu^\alpha} = \frac{\partial\Phi^\tau(\mathbf{x},\mathbf{u}^{(q)})}{\partial u_{\mu-1_i}^\alpha} \ .$$

Using this equality in (2.19) yields (2.18). The claim for $s>1$ follows by an induction from applying D_i repeatedly. $\square$

Example 2.3.17. The matrix M_2 for the prolonged symbol $\mathcal{N}_2$ of the representation of $\mathcal{R}_1^{(1)}$ in Example 2.3.8 is:

$$\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0\\
0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1\\
0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0
\end{pmatrix}.$$

The columns are indexed from left to right by u_{tt}, v_{tt}, w_{tt}, u_{xt}, v_{xt}, w_{xt}, u_{xx}, v_{xx} and w_{xx}. The rows are indexed by u_{tt}, v_{tt}, w_{tt}, u_{xt}, u_{xt} (again), v_{xt}, w_{xt}, u_{xx}, in accordance with Remark 2.3.13.

In Example 2.3.3, we used the matrix of the prolonged symbol given here in the construction of a formal solution order by order using a power series ansatz. If we collect the coefficients of second order into the vector $\mathbf{c}_2$, which then has the entries c^u_{tt}, c^v_{tt}, c^w_{tt}, c^u_{xt}, c^v_{xt}, c^w_{xt}, c^u_{xx}, c^v_{xx} and c^w_{xx}, we can write the additional Equations (2.13) for the second-order coefficients in the form

$$M_2\big|_{\mathbf{x}_0}\mathbf{c}_2 = -\frac{\partial \Phi^\tau}{\partial x^i}\big|_{\mathbf{x}_0} - \sum_{\alpha=1}^{m}\frac{\partial \Phi^\tau}{\partial u^\alpha}\big|_{\mathbf{x}_0} c^\alpha_i$$

according to the definition of the total derivative (2.8). In general, when constructing a formal solution to a differential equation $\mathcal{R}_q$ order by order as explained in Subsection 2.3.1, for $r > 0$ according to the definition of the total derivatives of order $q+r$, the Equations (2.12) for the coefficients of order $q+r$ can be written in the form

$$M_{q+r}\big|_{\mathbf{x}_0}\mathbf{c}_{q+r} = -M(\mathbf{x}_0, \mathbf{c}^{(q+r-1)})$$

where the vector $\mathbf{c}_{q+r}$ contains the coefficients $c^\alpha_{\mu+\nu}$, for all $1 \le \alpha \le m$, $|\mu| = q$ and $|\nu| = r$, in the order which is also used for the columns of the symbol matrix M_{q+r}, and where the vector $M(\mathbf{x}_0, \mathbf{c}^{(q+r-1)})$ contains the remaining summands of the formal derivatives $D_{\mu+\nu}\Phi^\tau$ up to order $q+r-1$ and has entries which depend on the coefficients up to order $q+r-1$ (here, they are collected into the vector $\mathbf{c}^{(q+r-1)}$).

Within M_{q+1}, the matrix for the prolonged symbol $\mathcal{N}_{q+1}$, the columns are ordered as described in Remark 2.3.13: the variables $v^\alpha_{\mu,i}$, which appear in Equation (2.18), are ordered descendingly with regard to the degree reverse lexicographic ranking applied to the multi-indices $\mu + 1_i$ and labelled by the pairs $(\alpha, \mu + 1_i)$. If $\mu = \nu$ for $v^\alpha_{\mu,i}$ and $v^\beta_{\mu,i}$ where $\alpha > \beta$, then $v^\alpha_{\mu,i} \prec v^\beta_{\mu,i}$. Now the columns are ordered descendingly with regard to the order of their labels $(\alpha, \mu + 1_i)$. The rows within M_{q+1} stem from the prolonged equations $D_i\Phi^\tau$ and are also ordered as described in Remark 2.3.13: if the variable which defines the class of $D_i\Phi^\tau$ is $u^\alpha_{\mu+1_i}$ then the rows are ordered descendingly with regard to the order of the labels $(\alpha, \mu+1_i)$; if for $i < j$ two different rows corresponding to $D_i\Phi^\sigma$ and $D_j\Phi^\tau$ share the variable which defines their class, the one corresponding to $D_i\Phi^\sigma$ comes first, as we did in Example 2.3.17.

Remark 2.3.18. Let the differential equation $\mathcal{R}_q$ be locally represented by $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$. Then the Jacobian matrix of $\mathcal{R}_{q+1} \subseteq J_{q+1}\pi$ can be split into four blocks like this:

$$\left(\begin{array}{c|c} \dfrac{\partial D_i\Phi^\tau}{\partial u^\alpha_\mu}\,, \quad |\mu| = q+1 & \dfrac{\partial D_i\Phi^\tau}{\partial u^\alpha_\mu}\,, \quad 0 \le |\mu| \le q \\ \hline 0 & \dfrac{\partial \Phi^\tau}{\partial u^\alpha_\mu}\,, \quad 0 \le |\mu| \le q \end{array}\right). \tag{2.20}$$

The lower block matrix stems from the equations of the original system $\mathcal{R}_q$. The block above on the left is M_{q+1}, the symbol matrix of the prolonged system $\mathcal{R}_{q+1}$. If this block has full rank, then $\mathcal{R}_q^{(1)} = \mathcal{R}_q$. If it has not, then through elementary row transformations a row within the upper block can be created which has only zero entries in its left part. If the right part of this row is independent of the rows within the block below on the right side, then this means there is an integrability condition: it can be constructed from the representation $\Phi^\tau(\mathbf{x}, u^\alpha, u_i^\alpha) = 0$ through applying the same row transformations on this representation. If the right part of the row is dependent on the rows within the block below on the right side, the representation of the system is redundant. For an inhomogenic system such equations represent *compatibility conditions*.

Proposition 2.3.19. *For the vector bundle $\mathcal{N}_{q+1}$ we have:*

$$\operatorname{rank} \mathcal{R}_q^{(1)} = \operatorname{rank} \mathcal{R}_{q+1} - \operatorname{rank} M_{q+1} \ .$$

Proof. Transform the Jacobian matrix (2.20) into row echelon form. Obviously, its rank is $\operatorname{rank} \mathcal{R}_{q+1}$. The equations in $\mathcal{R}_{q+1}$ which correspond to rows with pivots in the upper left block do not enter into $\mathcal{R}_q^{(1)}$, the projection of the prolonged system. □

As a corollary, the dual equation holds, too:

Corollary 2.3.20. *For the vector bundle $\mathcal{N}_{q+1}$ we have:*

$$\dim \mathcal{R}_q^{(1)} = \dim \mathcal{R}_{q+1} - \dim \mathcal{N}_{q+1} \ .$$

Proof. By definition of the rank, we have

$$\dim \mathcal{R}_q^{(1)} = \dim J_q\pi - \operatorname{rank} \mathcal{R}_q^{(1)}.$$

According to the preceding Proposition 2.3.19, this equals

$$\dim J_q\pi - (\operatorname{rank} \mathcal{R}_{q+1} - \operatorname{rank} M_{q+1}).$$

Proposition (2.1.2) implies that this is

$$\begin{aligned} & \dim J_{q+1}\pi - \dim(S^{q+1}T^*\mathcal{X} \otimes_{J_q\pi} V\pi) - (\operatorname{rank} \mathcal{R}_{q+1} - \operatorname{rank} M_{q+1}) \\ = \ & \dim J_{q+1}\pi - \operatorname{rank} \mathcal{R}_{q+1} - \dim(S^{q+1}T^*\mathcal{X} \otimes_{J_q\pi} V\pi) + \operatorname{rank} M_{q+1}. \end{aligned}$$

Now again from the definiton of the rank and from Lemma 2.3.14, it follows that this equals

$$\dim \mathcal{R}_{q+1} - \dim \mathcal{N}_{q+1}.$$

(Equation (2.17) is applied to the case $q + 1$ instead of q.) □

Example 2.3.21. For the differential equation in Example 2.3.12 the Jacobian matrix for the prolonged system $\mathcal{R}_2$, with its entries ordered as in Equation (2.20) and according to Remark 2.3.13, is

$$
\left(\begin{array}{cccccc|cccc}
\frac{\partial D_3\Phi^1}{\partial u_{zz}} & \frac{\partial D_3\Phi^1}{\partial u_{yz}} & \frac{\partial D_3\Phi^1}{\partial u_{yy}} & \frac{\partial D_3\Phi^1}{\partial u_{xz}} & \frac{\partial D_3\Phi^1}{\partial u_{xy}} & \frac{\partial D_3\Phi^1}{\partial u_{xx}} & \frac{\partial D_3\Phi^1}{\partial u_{z}} & \frac{\partial D_3\Phi^1}{\partial u_{y}} & \frac{\partial D_3\Phi^1}{\partial u_{x}} & \frac{\partial D_3\Phi^1}{\partial u} \\
\frac{\partial D_3\Phi^2}{\partial u_{zz}} & \frac{\partial D_3\Phi^2}{\partial u_{yz}} & \frac{\partial D_3\Phi^2}{\partial u_{yy}} & \frac{\partial D_3\Phi^2}{\partial u_{xz}} & \frac{\partial D_3\Phi^2}{\partial u_{xy}} & \frac{\partial D_3\Phi^2}{\partial u_{xx}} & \frac{\partial D_3\Phi^2}{\partial u_{z}} & \frac{\partial D_3\Phi^2}{\partial u_{y}} & \frac{\partial D_3\Phi^2}{\partial u_{x}} & \frac{\partial D_3\Phi^2}{\partial u} \\
\frac{\partial D_2\Phi^1}{\partial u_{zz}} & \frac{\partial D_2\Phi^1}{\partial u_{yz}} & \frac{\partial D_2\Phi^1}{\partial u_{yy}} & \frac{\partial D_2\Phi^1}{\partial u_{xz}} & \frac{\partial D_2\Phi^1}{\partial u_{xy}} & \frac{\partial D_2\Phi^1}{\partial u_{xx}} & \frac{\partial D_2\Phi^1}{\partial u_{z}} & \frac{\partial D_2\Phi^1}{\partial u_{y}} & \frac{\partial D_2\Phi^1}{\partial u_{x}} & \frac{\partial D_2\Phi^1}{\partial u} \\
\frac{\partial D_2\Phi^2}{\partial u_{zz}} & \frac{\partial D_2\Phi^2}{\partial u_{yz}} & \frac{\partial D_2\Phi^2}{\partial u_{yy}} & \frac{\partial D_2\Phi^2}{\partial u_{xz}} & \frac{\partial D_2\Phi^2}{\partial u_{xy}} & \frac{\partial D_2\Phi^2}{\partial u_{xx}} & \frac{\partial D_2\Phi^2}{\partial u_{z}} & \frac{\partial D_2\Phi^2}{\partial u_{y}} & \frac{\partial D_2\Phi^2}{\partial u_{x}} & \frac{\partial D_2\Phi^2}{\partial u} \\
\frac{\partial D_1\Phi^1}{\partial u_{zz}} & \frac{\partial D_1\Phi^1}{\partial u_{yz}} & \frac{\partial D_1\Phi^1}{\partial u_{yy}} & \frac{\partial D_1\Phi^1}{\partial u_{xz}} & \frac{\partial D_1\Phi^1}{\partial u_{xy}} & \frac{\partial D_1\Phi^1}{\partial u_{xx}} & \frac{\partial D_1\Phi^1}{\partial u_{z}} & \frac{\partial D_1\Phi^1}{\partial u_{y}} & \frac{\partial D_1\Phi^1}{\partial u_{x}} & \frac{\partial D_1\Phi^1}{\partial u} \\
\frac{\partial D_1\Phi^2}{\partial u_{zz}} & \frac{\partial D_1\Phi^2}{\partial u_{yz}} & \frac{\partial D_1\Phi^2}{\partial u_{yy}} & \frac{\partial D_1\Phi^2}{\partial u_{xz}} & \frac{\partial D_1\Phi^2}{\partial u_{xy}} & \frac{\partial D_1\Phi^2}{\partial u_{xx}} & \frac{\partial D_1\Phi^2}{\partial u_{z}} & \frac{\partial D_1\Phi^2}{\partial u_{y}} & \frac{\partial D_1\Phi^2}{\partial u_{x}} & \frac{\partial D_1\Phi^2}{\partial u} \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & \frac{\partial \Phi^1}{\partial u_{z}} & \frac{\partial \Phi^1}{\partial u_{y}} & \frac{\partial \Phi^1}{\partial u_{x}} & \frac{\partial \Phi^1}{\partial u} \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{\partial \Phi^2}{\partial u_{z}} & \frac{\partial \Phi^2}{\partial u_{y}} & \frac{\partial \Phi^2}{\partial u_{x}} & \frac{\partial \Phi^2}{\partial u}
\end{array}\right)
$$

$$
= \left(\begin{array}{cccccc|cccc}
1 & 0 & 0 & y & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & y & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & y & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & y & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0
\end{array}\right).
$$

The block-matrix below on the right with its rightmost column (containing two zero entries) left out is M_1, the symbol matrix of the original system $\mathcal{R}_1$. The upper left block is M_2, the symbol matrix of the prolonged system $\mathcal{R}_2$. It does not have full rank since its third row is the sum of the second row and of the y-fold of the sixth row. For the corresponding rows in the representation of the prolonged system (2.16), analogously we have $D_2\Phi^1 = D_3\Phi^2 + yD_1\Phi^2$, resulting in the integrability condition $u_x = 0$. The symbol matrix M_2 has rank 5, and here $\dim S_2T^*\mathcal{X} \otimes_{J_1\pi} V\pi = 6$, thus according to Equation (2.17) it follows that $\dim \mathcal{N}_2 = 1$. Since $\dim \mathcal{R}_2 = 2$, from Corollary 2.3.20 we conclude that the number of independent integrability conditions is $\dim \mathcal{R}_1^{(1)} = \dim \mathcal{R}_2 - \dim \mathcal{N}_2 = 2 - 1 = 1$.

2.4 Involutivity

The second main concept of formal theory is involutivity, a property concerning the symbol of a differential equation. Though, as we shall see, it is the involutivity of a differential equation which is essential for the existence of solutions, and in particular for the approach using Vessiot's techniques we have in mind, it is not always treated with due care in the literature on the subject. For example, Vessiot [43] himself does not consider it. Some modern textbooks (Stormark [40]) still neglect it.

2.4.1 Involutive Symbols

If the columns within a symbol matrix M_q are ordered according to the class respecting term order given in Remark 2.3.13, and if M_q has row echelon form, then we say the symbol (or its matrix) is given in *solved form*. The class of a row is then the class of the leftmost non-vanishing entry in that row. The corresponding derivative v^α_μ is called the *leader* of that row. It is convenient to introduce a name for the leaders of class k.

Definition 2.4.1. *Let the number of rows of class k in the row echelon form of a symbol matrix M_q be denoted by $\beta_q^{(k)}$. Then the numbers $\beta_q^{(k)}$ are called the* indices *of M_q or $\mathcal{R}_q$.*

Note that the indices $\beta_q^{(k)}$ depend on the coordinates of the base space $\mathcal{X}$ and thus may be different for different representations of an equation.

Example 2.4.2. The one-dimensional wave equation (see Example 2.2.2) may be represented both by

$$\mathcal{R}_2\colon\ \left\{\ u_{tt} = u_{xx}\right.$$

and, using characteristic coordinates, by

$$\mathcal{R}_2\colon\ \left\{\ u_{xt} = 0\right. \quad .$$

For the first representation we have $\beta_2^{(1)} = 0$ and $\beta_2^{(2)} = 1$ while for the second representation $\beta_2^{(1)} = 1$ and $\beta_2^{(2)} = 0$.

Remark 2.4.3. To achieve intrinsic results, we have to use appropriate coordinates. Since the class of a row is defined using independent coordinates $(x^i\colon 1 \leq i \leq n)$ on the base space $\mathcal{X}$, we consider an arbitrary change of coordinates $\mathbf{x} \leftarrow \tilde{\mathbf{x}}(\mathbf{x})$. When analyzing the symbol matrix M_q, we only use the derivatives of order q, the transformation of which is described by the Jacobian matrix; therefore there is no loss of generality in assuming that the change of coordinates is linear: $\mathbf{x} \leftarrow \tilde{\mathbf{x}}(\mathbf{x}) = (a_{ij})\mathbf{x}$ where (a_{ij}) is the corresponding $(n \times n)$-matrix with real entries a_{ij}. We do not specify these entries but consider a generic transformation. The symbol equations are changed in that now the new coordinates $\tilde{x}^i$ appear in homogenous polynomials of order q. If we order the columns of the symbol matrix according to class, now given by the generic coordinate system $\tilde{\mathbf{x}}$, and transform the symbol matrix into row echelon form, the number of rows of class n is the maximum among all values for $\beta_q^{(n)}$ which are admitted by the differential equation $\mathcal{R}_q$ and therefore intrinsic. Denote this maximum by $\tilde{\beta}_q^{(n)}$. Now we define $\tilde{\beta}_q^{(k)}$ for $1 \leq k < n$ by choosing the a_{ij} such that the sum $\tilde{\beta}_q^{(n)} + \tilde{\beta}_q^{(n-1)}$ takes the maximal value admitted by $\mathcal{R}_q$ as well, and so does $\sum_{k=j}^n \tilde{\beta}_q^{(k)}$ for all $1 \leq j \leq n$, without changing the previous $\tilde{\beta}_q^{(k)}$ any more. This yields indices $\tilde{\beta}_q^{(k)}$ for $\mathcal{R}_q$ for which $\sum_{k=1}^n k \cdot \tilde{\beta}_q^{(k)}$ is maximal.

Definition 2.4.4. *For a given local representation of a differential equation the local coordinates $(x^i\colon 1 \leq i \leq n)$ are called* δ-regular (for the symbol M_q) with indices $\beta_q^{(k)}$ *if $\beta_q^{(k)} = \tilde{\beta}_q^{(k)}$ for all $1 \leq k \leq n$. Otherwise the coordinates are called* δ-singular.

Remark 2.4.5. Those matrices (a_{ij}) which lead to δ-singular coordinate systems define a subvariety in the space of all $(n \times n)$-matrices because a coordinate system becomes δ-singular if the a_{ij} satisfy certain algebraic equations which make entries in the symbol matrix vanish, thus reducing the class of the corresponding row. This subvariety has a dimension less than n^2 and therefore is a set of measure zero. This means, a random choice of local coordinates on $\mathcal{X}$ yields δ-regular coordinates with probability one—though not in all cases, as Example 2.4.2 shows. (On the other hand, using characteristic coordinates is usually not a random choice.)

An intrinsic definition of δ-regularity may be based on Spencer cohomology, after the exterior derivative of which it is named; see Pommaret [35], Chapter 3, and Seiler [37, 39].

In Example 2.3.21, we have $\operatorname{rank} M_{q+1} = 5$; the symbol matrix M_q contains one row of class 2 and one row of class 3, therefore $2\beta_1^{(2)} + 3\beta_1^{(3)} = 5$. In cases like this where such an equality holds, the symbol is of special interest.

Definition 2.4.6. The symbol $\mathcal{N}_q$ of a differential equation is called *involutive* if

$$\operatorname{rank} M_{q+1} = \sum_{k=1}^{n} k\beta_q^{(k)} \ . \tag{2.21}$$

For a row of class k the variables $x^1, x^2, \ldots, x^k$ are called *multiplicative* and the remaining variables from x^{k+1} up to x^n *non-multiplicative*.

If the indices of a differential equation (or, equivalently, its Cartan characters, which are introduced below in Definition 2.4.23) are defined by way of the Spencer cohomology (see Seiler [39]), Equation (2.21) is used as a criterion to examine if a symbolic system (a sequence of vector spaces which are linked to the notion of the geometric symbol and its prolongations) is involutive at a certain order. This check up on involutivity is usually called the *Cartan test*, and we use this name for Equation (2.21) from now on.

The involutivity of the symbol in the sense of Definition 2.4.6 should not be confused with the fact that the symbol, considered as a distribution of vector fields, is always involutive, being defined as the intersection of two involutive distributions.

The next proposition characterizes the involutivity of a symbol in the sense of Definition 2.4.6.

Proposition 2.4.7. *Let (2.7) be a local representation of a differential equation in solved form where the coordinates are δ-regular. Then the symbol $\mathcal{N}_q$ is involutive if, and only if, all independent equations of order $q+1$ in a local representation of the prolonged system $\mathcal{R}_{q+1}$ are algebraically dependent on those formal derivatives of the equations of order q in (2.7) where each equation of order q is derived with respect to its multiplicative variables only.*

Proof. When calculating the prolongations of each equation in the representation of a differential equation only with respect to the multiplicative variables of each equation, the new equations derived this way are independent because they have different pivots within the symbol matrix M_q.

Since there are $\beta_q^{(k)}$ equations of class k and each of them has k multiplicative variables, there are at least $\sum k\beta_q^{(k)}$ independent equations of order $q+1$ in $\mathcal{R}_{q+1}$.

If the symbol $\mathcal{N}_q$ is involutive, then all independent equations of order $q+1$ are derived this way. All other equations derived from prolonging $\mathcal{R}_q$ are thus dependent, of lower order or both. $\square$

The following lemma is needed to prove the proposition that follows it. This proposition is then used in the analysis of matrices which describe the structure of the vector field distributions in the Vessiot theory.

Lemma 2.4.8. *Let M_q be the matrix of an involutive symbol $\mathcal{N}_q$ in solved form, where the columns are ordered according to a class-respecting ranking (for instance, the one given in Remark 2.3.13). If v^α_μ is a leader of class k, then $v^\alpha_{\mu-1_k+1_j}$ is a leader for all $j>k$, too.*

Proof. As v^α_k is a leader of class k and $j>k$, the variable x^j is non-multiplicative for the corresponding equation Φ^α_μ. For the ranking which we use to order the columns in M_q, we have for any multi-index ν that if $u^\alpha_{\mu_1}<u^\alpha_{\mu_2}$, then $u^\alpha_{\mu_1+\nu}<u^\alpha_{\mu_2+\nu}$. Therefore the equation $\Phi^\alpha_{\mu+1_j}$, the prolongation of Φ^α_μ with respect to x^j, has the leader $v^\alpha_{\mu+1_j}$. According to Proposition 2.4.7, for an involutive symbol, all independent equations can be derived as combinations of prolongations with respect to multiplicative variables only. This means that $v^\alpha_{\mu+1_j}$, being of class k, can be derived using an equation Φ^α_ν of higher or equal class prolonged with respect to x^k. It follows that $\nu+1_k=\mu+1_j$, that is, $\nu=\mu+1_j-1_k$. $\square$

The announced proposition states the monotony of the indices for a system of order $q=1$.

Proposition 2.4.9. *Let $\mathcal{R}_1$ be a first-order differential equation, where the fibre-dimension is m and the dimension of the base-space is n. If its symbol $\mathcal{N}_1$ is involutive, then the indices satisfy the following chain of inequalities:*

$$0\leq\beta_1^{(1)}\leq\beta_1^{(2)}\leq\cdots\leq\beta_1^{(n)}\leq m\ .$$

There is no such analogue for systems of an order greater than 1.

Proof. Let v^α_μ be a leader of class k. Then $\mu_k=\delta_{ik}$ for $1\leq i\leq n$ because $|\mu|=q=1$. According to Lemma 2.4.8, v^α_ν with $\nu=\mu-1_k+1_{k+1}$ is a leader of the symbol matrix M_1; it is of class $k+1$ with $\nu_i=\delta_{i,k+1}$ for $1\leq i\leq n$. It follows that $\beta_1^{(k)}\leq\beta_1^{(k+1)}$. As there are are at most m pairwise different derivatives of class n even for arbitrary $q\in\mathbb{N}$, the fibre dimension m is an upper bound. A counterexample for $q>1$ is given by $u_{yy}=u_{xy}=u_{xx}=0$ where $2=\beta_2^{(1)}>\beta_2^{(2)}=1$. $\square$

Some special cases arise when the symbol matrix has full rank.

Proposition 2.4.10. *A vanishing symbol is involutive.*

Corollary 2.4.11. *The symbol of a maximally over-determined system is involutive. In particular, the symbol of an ordinary differential equation is involutive.*

Proof. By definition, a maximally over-determined system is a system with a vanishing symbol, and a vanishing symbol is involutive according to Proposition 2.4.10. An ordinary differential equation is by definition a differential equation with a one-dimensional base space $\mathcal{X}$ and a special kind of a maximally over-determined system. □

Even if a symbol is not involutive, a finite number of prolongations suffices to reach an involutive symbol. This non-trivial fact is a cornerstone of the formal theory of differential equations:

Theorem 2.4.12. *Let $\mathcal{N}_q$ be the symbol of some differential equation. Then there is a non-negative integer s such that $\mathcal{N}_{q+s}$ is involutive.*

Proof. The proof is due to Sweeney [41], Corollary 7.7, assumes that we use δ-regular coordinates and even gives an upper bound which depends on the base-space dimension, n, the dimension of the typical fibre, m, and the order of the system, q. □

It turns out that the prolongation of an involutive symbol is itself involutive.

Proposition 2.4.13. *Let $\mathcal{R}_q$ be a differential equation with an involutive symbol $\mathcal{N}_q$.*

1. *Then for all $s \geq 0$ the symbol $\mathcal{N}_{q+s}$ is involutive.*
2. *We have $(\mathcal{R}_q^{(1)})_{+s} = \mathcal{R}_{q+s}^{(1)}$.*

Proof. The proofs for both parts contains the distinction of several cases. For the first part, consider the row echelon form of the symbol matrix M_q. Its columns are ordered according to a class respecting ranking and may be labelled by the unknowns v_μ^α. According to Proposition 2.4.7 the leaders of M_{q+1} (also given in row echelon form) are $v_{\mu+1_\ell}^\alpha$ where $1 \leq \ell \leq \operatorname{cl}\mu$. To prove the first part, one must show that the prolongation of a row in M_{q+1} with regard to a non-multiplicative variable is linearly dependent on the prolongations with regard to multiplicative variables. The necessary case distinction with regard to the class can be found in Seiler [38], pages 91/92, and shows that in fact all linearly independent rows in the matrix M_{q+2} can be derived by prolonging those in M_{q+1} with respect to their multiplicative variables only. The involutivity of $\mathcal{N}_{q+1}$ now follows from Proposition 2.4.7. For $s > 1$ the claim follows by a simple induction.

For the second part, let $s = 1$. One must show that all integrability conditions which arise from the projection from order $q+2$ to order $q+1$ (referring to $\mathcal{R}_{q+1}^{(1)}$) are prolongations of integrability conditions which arise from the projection from order $q+1$ to order q (referring to $(\mathcal{R}_q^{(1)})_{+1}$). To this end, let $\Phi^\tau = 0$ be an equation of class k in a local representation of $\mathcal{R}_q$ and consider $D_\ell D_j \Phi^\tau = 0$. Assuming without loss of generality $\ell \leq j$, there are two cases, $j \leq k$ and $k < j$, in both of which the claim follows. Again, see [38], pages 91/92, for the details. □

The second part of Proposition 2.4.13 means that for a differential equation with an involutive symbol a prolongation ρ and a projection π commute, if they follow a prolongation: $\pi \circ \rho^2 = \rho \circ \pi \circ \rho$. For the calculation of $\mathcal{R}_{q+1}^{(1)}$, the twice prolonged space $\mathcal{R}_{q+2}$ is needed while for the calculation of $(\mathcal{R}_q^{(1)})_{+1}$ it is enough to use $\mathcal{R}_{q+1}$.

Example 2.4.14. For the differential equation represented by

$$\mathcal{R}_2\colon \begin{cases} u_{tt} = 0 \\ u_{xx} = 0 \end{cases}$$

we have $\beta_1^{(1)} = 1$ and $\beta_1^{(2)} = 1$. The first prolongation is represented by

$$\mathcal{R}_3\colon \begin{cases} u_{ttt} = 0 & u_{xxt} = 0 \\ u_{xtt} = 0 & u_{xxx} = 0 \end{cases} .$$

Since the symbol matrix of $\mathcal{R}_3$ is the unit matrix $\mathbb{1}_4$, we have $\operatorname{rank} M_3 = 4 \neq 3 = \sum_{k=1}^{2} k\beta_1^{(k)}$. Though the system $\mathcal{R}_2$ is formally integrable as all its right sides vanish, it is not involutive. The reason is this: The equation $u_{xx} = 0$ has a non-multiplicative prolongation, $u_{xxt} = 0$, which is not algebraically dependent on the multiplicative prolongations $u_{xxx} = 0$, $u_{xtt} = 0$ and $u_{ttt} = 0$ of the system. Thus the criterion for the involutivity of the symbol given in Proposition 2.4.7 is not met. But the system $\mathcal{R}_3$ is involutive. So the original system $\mathcal{R}_2$ becomes involutive after two prolongations.

Definition 2.4.15. *We call such an additional condition of an order greater than q which destroys the involutivity of the symbol an* obstruction to involution.

2.4.2 Involutive Systems

For a formally integrable differential equation there are formal power series solutions, which explains the name. For an analytic equation with suitable initial conditions, these series even converge. This is the Cartan-Kähler theorem, generalizing the classical Cauchy-Kovalevskaya theorem. The Cauchy-Kovalevskaya theorem states that an analytic normal (meaning there are as many equations as there are dependent coordinates) differential equation with analytic initial conditions, represented in non-characteristic coordinates, has a unique analytic solution. (For a modern formulation and accompanying proof, see Seiler [37].) The Cartan-Kähler theorem (which we formulate below as Theorem 2.4.31) extends the Cauchy-Kovalevskaya theorem to arbitrary *involutive* (but still analytic) systems, defined below. Thus, it is of advantage to work with an involutive system, and the question arises when this is possible?

To check a differential equation on formal integrability, according to the Definition 2.3.2 a countably infinite number of conditions have to be checked. The Cartan-Kuranishi theorem 2.4.19, given below, offers a procedure which only requires a finite number of steps for the subset of the formally integrable differential equations which comprises the involutive systems.

Definition 2.4.16. *A differential equation is called* involutive*, if it is formally integrable and its symbol is involutive.*

Note that involutivity both of the symbol and the differential equation is a local property depending on $\rho \in \mathcal{R}_q$. We assume from now on that the $\beta_q^{(k)}$ do not vary for different points of the differential equation by restricting to a proper subset if necessary.

The criterion of Proposition 2.4.7 fails at checking a differential equation for involutivity because it refers only to the equations of highest order q in $\mathcal{R}_q$ while integrability conditions may arise by prolonging equations of order less than q. If a differential equation is formally integrable, its symbol may or may not be involutive, and conversely, if the symbol of a differential equation is involutive, the equation may or may not be formally integrable.

Example 2.4.17. The system $\mathcal{R}_2$ in Example 2.4.14 is formally integrable, but its symbol is not involutive. The system

$$\mathcal{R}_2 \colon \begin{cases} u_{tt} = 0 \\ u_{xx} = 0 \\ u_x = 0 \end{cases}$$

is neither formally integrable (there is the integrability condition $u_{xt} = 0$) nor has it an involutive symbol. (It has the same symbol as the system in Example 2.4.14, from which it is derived by adding the equation $u_x = 0$.) The system $\mathcal{R}_1$ in Example 2.2.11 is not formally integrable, but it has an involutive symbol. The system $\mathcal{R}_3$ in Example 2.4.14 is formally integrable and has an involutive symbol, and thus it is involutive. So is the one-dimensional wave equation in Example 2.2.2.

According to Corollary 2.4.11, for ordinary differential equations, there is no difference between formal integrability and involutivity as long as there are only equations of order q.

For a differential equation with an involutive symbol, the next proposition yields a means suitable to check the equation for involutivity using only a finite number of operations.

Proposition 2.4.18. *A differential equation $\mathcal{R}_q$ is involutive if, and only if, its symbol is involutive and $\mathcal{R}_q = \mathcal{R}_q^{(1)}$.*

Proof. Let the equation's symbol be involutive, and let $\mathcal{R}_q = \mathcal{R}_q^{(1)}$. Thus, according to Proposition 2.4.13, we have $\mathcal{R}_{q+s}^{(1)} = (\mathcal{R}_q^{(1)})_{+s}$ for all non-negative integers s. Since we assume $\mathcal{R}_q^{(1)} = \mathcal{R}_q$, it follows that $(\mathcal{R}_q^{(1)})_{+s} = \mathcal{R}_{q+s}$. Thus, the differential equation is formally integrable, too. The reverse implication is trivial. □

Involutive differential equations are remarkable in that it is possible to decide in a finite number of steps when a system is involutive, and furthermore that for an arbitrary system in a finite number of steps an involutive system can be constructed which has the same space of formal solutions. This procedure is called *completion to involution*. That it succeeds for involutive systems says the Cartan-Kuranishi theorem.

Theorem 2.4.19 (Cartan, Kuranishi [28]). *For any differential equation $\mathcal{R}_q$ there are two non-negative numbers r and s such that $\mathcal{R}_{q+s}^{(r)}$, the r-th projection of the $r+s$-th prolongation of $\mathcal{R}_q$, is involutive and has the same space of formal solutions as $\mathcal{R}_q$.*

Proof. The proof is constructive and may be sketched as follows: first consider the symbol of $\mathcal{R}_q$; if it is not involutive, prolong $\mathcal{R}_q$ to $\mathcal{R}_{q+1}$. If the symbol of $\mathcal{R}_{q+1}$ is not involutive, prolong $\mathcal{R}_{q+1}$ to $\mathcal{R}_{q+2}$ and so on until an equation $\mathcal{R}_{q+s_1}$ with an involutive symbol is reached. Such a $s_1 \geq 0$ exists according to Proposition 2.4.12. Then check this system $\mathcal{R}_{q+s_1}$ with the involutive symbol for integrability conditions: if $\mathcal{R}^{(1)}_{q+s_1} = \mathcal{R}_{q+s_1}$, then $\mathcal{R}_{q+s_1}$ is involutive, and we are through. If $\mathcal{R}^{(1)}_{q+s_1} \subset \mathcal{R}_{q+s_1}$, then prolong $\mathcal{R}^{(1)}_{q+s_1}$ to a system $(\mathcal{R}^{(1)}_{q+s_1})_{+s_2}$ with an involutive symbol. Then check if $(\mathcal{R}^{(1)}_{q+s_1})^{(1)}_{+s_2} \subseteq (\mathcal{R}^{(1)}_{q+s_1})_{+s_2}$. If equality holds, then $(\mathcal{R}^{(1)}_{q+s_1})_{+s_2}$ is involutive and we are through. If the inclusion is proper, then prolong $(\mathcal{R}^{(1)}_{q+s_1})^{(1)}_{+s_2}$ to a system with an involutive symbol and so forth. This way the algorithm eventually yields an involutive system $(...(((((\mathcal{R}_{q+s_1})^{(1)})_{+s_2})^{(1)})...)_{+s_l})^{(1)}$. The termination follows from a Noetherian argument. Since during the procedure only differential equations with an involutive symbol are being projected, Proposition 2.4.13 can be applied from which follows

$$(...(((((\mathcal{R}_{q+s_1})^{(1)})_{+s_2})^{(1)})...)_{+s_l})^{(1)} = \mathcal{R}^{(l)}_{q+s_1+s_2+...s_l} \,.$$

For details, see Seiler [38]. □

Remark 2.4.20. Since any ordinary differential equation $\mathcal{R}_q$ has an involutive symbol, the algorithm simplifies for such an equation: the prolongations for the construction of a system with an involutive symbol are not needed. Furthermore, the order of the operations in the algorithm can be reversed. In the general case we always first prolong and then project; for ordinary differential equations it suffices to first project, because in a system of ordinary differential equations only one mechanism for the generation of integrability conditions exists: the system contains equations of differing orders and the prolongation of the lower order ones leads to new equations. These equations of lower order describe a constraint manifold in $\mathcal{R}_q$, and one has to prolong just this manifold. As a consequence, the algorithm in this case yields a system of the same order q.

As mentioned in the beginning of Section 2.3, in a system of partial differential equations with $\dim \mathcal{X} > 1$ a second mechanism exists, caused by the appearance of cross-derivatives: if a linear combination of prolonged equations is such that all derivatives of maximal order cancel, this gives an integrability condition of lower order. It is crucial here to first prolong and then project; otherwise one might overlook integrability conditions. Examples like $u_{zz} + yu_{xx} = 0 = u_{yy}$ (this Example is from Janet [22]) demonstrate that several prolongations may be needed to reach suitable cross-derivatives, and the integrability conditions (here $u_{xxy} = 0$ and $u_{xxxx} = 0$) may be of higher order than the original system. Involution of the symbol $\mathcal{N}_q$ is concerned with the maximal number of prolongations needed.

According to Definition 2.2.5, a solution is a section $\sigma : \mathcal{X} \to \mathcal{E}$ such that its prolongation satisfies $\operatorname{im} j_q\sigma \subseteq \mathcal{R}_q$. For formally integrable equations it is straightforward to construct order by order formal power series solutions. Otherwise it is hard to find solutions. A constitutive insight of Cartan [6] was to introduce *infinitesimal solutions* or *integral elements* at a point $\rho \in \mathcal{R}_q$ as subspaces $\mathcal{U}_\rho \subseteq T_\rho\mathcal{R}_q$ which are potentially part of the tangent space of a prolonged solution.

Definition 2.4.21. *Let $\mathcal{R}_q \subseteq J_q\pi$ be a differential equation. A linear subspace $\mathcal{U}_\rho \subseteq T_\rho\mathcal{R}_q$ is called an* integral element *at the point $\rho \in \mathcal{R}_q$, if a point $\hat{\rho} \in \mathcal{R}_{q+1}$ exists such that $\pi^{q+1}_q(\hat{\rho}) = \rho$ and $T\iota(\mathcal{U}_\rho) \subseteq \operatorname{im}\Gamma_{q+1}(\hat{\rho})$.*

This definition of an integral element is not the customary one. Usually [15, 40, 42], one considers the pull-back $\iota^*\mathcal{C}^0_q$ of the contact codistribution or more precisely the differential ideal $\mathcal{I}[\mathcal{R}_q] := \operatorname{span}\{\iota^*\mathcal{C}^0_q\}_{\text{diff}}$ generated by it. Algebraically, $\mathcal{I}[\mathcal{R}_q]$ is therefore spanned by a basis of $\iota^*\mathcal{C}^0_q$ and the exterior derivatives of the one-forms in this basis. An integral element is then a subspace on which this ideal vanishes.

Proposition 2.4.22. *Let $\mathcal{R}_q$ be a differential equation such that $\mathcal{R}^{(1)}_q = \mathcal{R}_q$. A linear subspace $\mathcal{U}_\rho \subseteq T_\rho\mathcal{R}_q$ is an integral element at $\rho \in \mathcal{R}_q$ if, and only if, $T\iota(\mathcal{U}_\rho)$ is transversal to the fibration π^q_{q-1} and every differential form $\omega \in \mathcal{I}[\mathcal{R}_q]$ vanishes on $\mathcal{U}_\rho$.*

Proof. First let $\rho \in \mathcal{R}_q$ such that $T\iota(\mathcal{U}_\rho)$ is transversal to the fibration π^q_{q-1} and every differential form $\omega \in \mathcal{I}[\mathcal{R}_q]$ vanishes on $\mathcal{U}_\rho$. Because $T\iota(\mathcal{U}_\rho)$ is transversal there is a generating set of vectors v_i for it with $T\iota(v_1) = C^{(q)}_i|_\rho + b^\alpha_{\mu,i}C^\mu_\alpha|_\rho$ for real numbers $b^\alpha_{\mu,i}$ (for now, μ, i is a just double index). The exterior derivative of a contact form $\omega^\alpha_\nu = du^\alpha_\nu - u^\alpha_{\nu+1_i}dx^i$ of order $|\nu| = q-1$ is $d\omega^\alpha_\nu = dx^i \wedge du^\alpha_{\nu+1_i}$. It satisfies $\iota^*(d\omega^\alpha_\nu)|_\rho(v_i, v_j) = d\omega^\alpha_\nu(T\iota(v_i), T\iota(v_j)) = 0$ for all $v_i, v_j \in \mathcal{U}_\rho$ according to assumption. The skewness of $d\omega^\alpha_\nu$ now implies $b^\alpha_{\nu+1_i,j} = b^\alpha_{\nu+1_j,i}$. This means, for $\mu := \nu + 1_j$, that $b^\alpha_{\mu,i} = b^\alpha_{\mu+1_i}$ (now we can interpret μ, i as a multi-index of order $q+1$). Hence, there is a local section $\sigma \in \Gamma_{\mathbf{x}}\pi$ such that $\rho = [\sigma]^{(q)}_{\mathbf{x}}$ and the vectors $T\iota(v_i)$ where $1 \leq i \leq n$ span $T_\rho(\operatorname{im} j_q\sigma)$. It follows, since $\mathcal{U}_\rho$ is spanned by a set of linear combinations of the vectors v_i, that it is an integral element.

Now for the reverse implication let $\mathcal{U}_\rho$ be an integral element for some point $\rho \in \mathcal{R}_q$. Then by definition, there is an element $\hat{\rho} \in \mathcal{R}_{q+1}$ in the fibre such that $\hat{\pi}^q_{q-1}(\hat{\rho}) = \rho$ and $T\iota(\mathcal{U}_\rho) = \operatorname{im}\Gamma_{q+1}(\hat{\rho})$. This means that $T\iota(\mathcal{U}_\rho)$ is transversal with regard to the fibration π^q_{q-1} and $T\iota(\mathcal{U}_\rho) \subseteq \operatorname{im}\Gamma_{q+1}(\hat{\rho})$, so all one-forms $\omega \in \iota^*\mathcal{C}^0_q$ annihilate $\mathcal{U}_\rho$. Now consider the pull-backs of the exterior derivatives, $\iota^* d\mathcal{C}^0_q$. We have to show that they, too, vanish on $\mathcal{U}_\rho$. We do this by constructing a distribution which for all $\tilde{\rho} \in \mathcal{R}_q$ contains $\mathcal{U}_{\tilde{\rho}}$ and is annihilated by $\iota^* d\mathcal{C}^0_q$. The restricted projection $\hat{\pi}^{q+1}_q \colon \mathcal{R}_{q+1} \to \mathcal{R}_q$ is surjective; hence there is a local section $\gamma : \mathcal{R}_q \to \mathcal{R}_{q+1}$ such that $\gamma(\rho) = \hat{\rho}$. Now define an n-dimensional distribution $\mathcal{D}$ on $\mathcal{R}_q$ by setting $T\iota(\mathcal{D}_{\tilde{\rho}}) = \operatorname{im}\Gamma_{q+1}\bigl(\gamma(\tilde{\rho})\bigr)$ for all $\tilde{\rho} \in \mathcal{R}_q$. Then by construction $\mathcal{U}_\rho \subseteq \mathcal{D}_\rho$. From the representation in local coordinates (2.3) of the contact map it follows that locally the distribution $\mathcal{D}$ is spanned by n vector fields X_i such that

$$\iota_* X_i = C^{(q)}_i + \sum_{\alpha=1}^{m} \sum_{|\mu|=q} \gamma^\alpha_{\mu+1_i} C^\mu_\alpha \ .$$

The coefficients $\gamma^\alpha_{\mu+1_i}$ are the highest-order components of the section γ, and $C^{(q)}_i$ and C^μ_α are the contact vector fields (2.4). Thus the commutator of two such vector fields satisfies

$$\iota_*[X_i, X_j] = \bigl(C^{(q)}_i(\gamma^\alpha_{\mu+1_j}) - C^{(q)}_j(\gamma^\alpha_{\mu+1_i})\bigr)C^\mu_\alpha + \gamma^\alpha_{\mu+1_j}[C^{(q)}_i, C^\mu_\alpha] - \gamma^\alpha_{\mu+1_i}[C^{(q)}_j, C^\mu_\alpha] \ .$$

For all $|\nu| = q-1$, we have $[C_i^{(q)}, C_\alpha^{\nu+1_i}] = -\partial_{u_\nu^\alpha}$ (see Remark 2.1.11); therefore a commutator on the right side vanishes when $\mu_i = 0$ or $\mu_j = 0$. Otherwise they equal $-\partial_{u^\alpha_{\mu-1_i}}$ or $-\partial_{u^\alpha_{\mu-1_j}}$. Anyway the two sums of commutators on the right side cancel and it follows that $\iota_*\big([X_i, X_j]\big) \in \mathcal{C}_q$ as all the C_α^μ are contact vector fields. Now for any contact form $\omega \in \mathcal{C}_q^0$ we have

$$\begin{aligned}\iota^*(d\omega)(X_i, X_j) &= d\omega(\iota_* X_i, \iota_* X_j)\\ &= \iota_* X_i\big(\omega(\iota_* X_j)\big) - \iota_* X_j\big(\omega(\iota_* X_i)\big) + \omega\big(\iota_*([X_i, X_j])\big)\ .\end{aligned}$$

All vector fields which appear in the second line are contact vector fields; therefore all summands vanish. It follows that any two-form $\omega \in \iota^*(d\mathcal{C}_q^0)$ vanishes on $\mathcal{D}$ and in particular on $\mathcal{U}_\rho \subseteq \mathcal{D}_\rho$. $\square$

For an involutive differential equation, $\beta_q^{(k)}$ is the number of the principal derivatives of order q and class k. In the Taylor series of the formal solution there are only finitely many of them. The size of the formal solution spaces of several equations may be compared by comparing these indices for every order.

Definition 2.4.23. *For an involutive differential equation let $\alpha_q^{(k)}$ denote the number of the remaining derivatives of order q and class k, that is, the number of the parametric derivatives. These are called the* Cartan characters *of the differential equation.*

Remark 2.4.24. The Cartan characters tell the numbers of derivatives which have to be given by initial or boundary conditions to determine a unique solution; in this sense, the Cartan characters measure the degrees of freedom for a formal solution to $\mathcal{R}_q$. For a differential equation of finite type, all Cartan characters vanish. This means there are no free Taylor coefficients of orders q and greater than q, and therefore their space of formal solutions can be parameterized by $\dim \mathcal{R}_q$ parameters.

For $1 \le k \le n$ the Cartan characters are

$$\alpha_q^{(k)} = m\binom{q+n-k-1}{q-1} - \beta_q^{(k)}$$

since $\binom{q+n-k-1}{q-1}$ is the number of derivatives of order q and class k. Note that for $q=1$ we have $\alpha_1^{(k)} + \beta_1^{(k)} = m$, the fibre dimension of the total space of π. Unlike the indices $\beta_q^{(k)}$, the Cartan characters of an involutive differential equation are intrinsic values of $\mathcal{R}_q$ in that they do not change if $\mathcal{R}_q$ is rewritten as a first-order system (see Subsection 2.5.1). They may vary over $\mathcal{R}_q$, though.

Proposition 2.4.25. *For a first-order differential equation $\mathcal{R}_1$ where the fibre-dimension is m, the Cartan characters satisfy the following chain of inequalities:*

$$m \ge \alpha_1^{(1)} \ge \alpha_1^{(2)} \ge \cdots \ge \alpha_1^{(n)} \ge 0\ . \tag{2.22}$$

There actually is analogue for systems of an order greater than 1.

Proof. For $q = 1$, according to Proposition 2.4.9, we have $0 \le \beta_1^{(1)} \le \beta_1^{(2)} \le \cdots \le \beta_1^{(n)} \le m$, which is equivalent to the chain of inequalities (2.22) because of $m - \alpha_1^{(i)} = \beta_1^{(i)}$. For $q > 1$ see Corollary 2.5.3. $\square$

Definition 2.4.26. *For a first-order differential equation $\mathcal{R}_1$ the following local representation, a special kind of solved form,*

$$u_n^\alpha = \phi_n^\alpha(\mathbf{x}, u^\beta, u_j^\gamma, u_n^\delta) \qquad \begin{cases} 1 \le \alpha \le \beta_1^{(n)} \\ 1 \le j < n \\ \beta_1^{(n)} < \delta \le m \end{cases}, \tag{2.23a}$$

$$u_{n-1}^\alpha = \phi_{n-1}^\alpha(\mathbf{x}, u^\beta, u_j^\gamma, u_{n-1}^\delta) \qquad \begin{cases} 1 \le \alpha \le \beta_1^{(n-1)} \\ 1 \le j < n-1 \\ \beta_1^{(n-1)} < \delta \le m \end{cases}, \tag{2.23b}$$

$$\vdots$$

$$u_1^\alpha = \phi_1^\alpha(\mathbf{x}, u^\beta, u_1^\delta) \qquad \begin{cases} 1 \le \alpha \le \beta_1^{(1)} \\ \beta_1^{(1)} < \delta \le m \end{cases}, \tag{2.23c}$$

$$u^\alpha = \phi_1^\alpha(\mathbf{x}, u^\beta) \qquad \begin{cases} 1 \le \alpha \le \beta_0 \\ \beta_0 < \beta \le m \end{cases}. \tag{2.23d}$$

is called its Cartan normal form. *The equations of zeroth order, $u^\alpha = \phi^\alpha(\mathbf{x}, u^\beta)$, are called* algebraic. *The functions ϕ_k^α are called the* right sides *of $\mathcal{R}_1$.*

If, for some $1 \le k \le n$, the number of equations is $\beta_1^{(k)} = m$, then the condition $\beta_1^{(k)} < \delta \le m$ is meaningless and there are no terms u_k^δ on the right sides of those equations.

Here, each equation is solved for a principal derivative of maximal class k in such a way that the corresponding right side of the equation may depend on an arbitrary subset of the independent variables, an arbitrary subset of the dependent variables u^β with $1 \le \beta \le \beta_0$, those derivatives u_j^γ for all $1 \le \gamma \le m$ which are of a class $j < k$ and those derivatives which are of the same class k but are not principal derivatives. Note that a principle derivative u_k^α may dependent on another principle derivative u_l^γ as long as $l < k$. The equations are grouped according to their class in descending order as described in Remark 2.3.13.

Example 2.4.27. The representation of the differential equation $\mathcal{R}_1^{(1)}$ in Example 2.2.13 is in Cartan normal form. There are no algebraic equations. Multiplicative variables for each of the first three rows are t and x and for the last row x. The symbol matrix, given in Example 2.3.8, is already in reduced row echelon form, and each row is solved for its principal derivative.

Remark 2.4.28. The Cartan normal form at once yields the symbol matrix in row echelon form. For $1 \le k \le n$ in the Cartan normal form of Definition 2.4.26, we set

$$\Phi_k^\alpha(\mathbf{x}, u^\beta, u_j^\gamma, u_k^\delta) := u_k^\alpha - \phi_k^\alpha(\mathbf{x}, u^\beta, u_j^\gamma, u_k^\delta)\ .$$

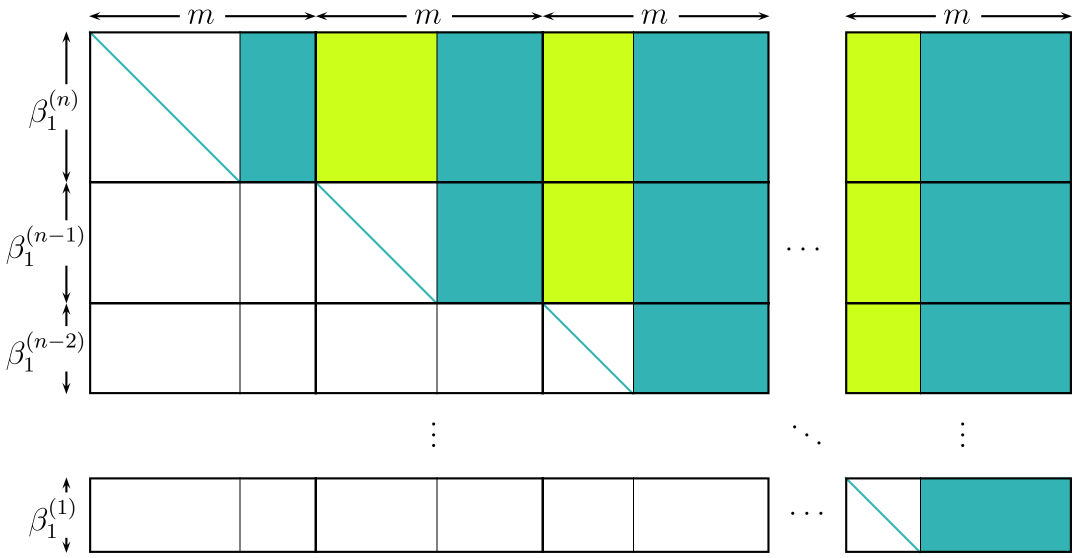


Figure 2.1: The symbol matrix for a first-order system in Cartan normal form, given in Definition 2.4.26. The columns are ordered descendingly with regard to the degree reverse lexicographic ranking. The rows are ordered descendingly according to class and within each class descendingly with respect to α. Zero entries are marked by white areas; potentially non-trivial entries are marked by light or dark shades; blocks with diagonals mark unit blocks.

When calculating the symbol matrix of this representation, for any class k and for any $1 \leq j \leq n$ we get a block of the form

$$\begin{pmatrix} \frac{\partial \Phi_k^1}{\partial u_j^1} & \frac{\partial \Phi_k^1}{\partial u_j^2} & \cdots & \frac{\partial \Phi_k^1}{\partial u_j^m} \\ \frac{\partial \Phi_k^2}{\partial u_j^1} & \frac{\partial \Phi_k^2}{\partial u_j^2} & \cdots & \frac{\partial \Phi_k^2}{\partial u_j^m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \Phi_k^{\beta_1^{(k)}}}{\partial u_j^1} & \frac{\partial \Phi_k^{\beta_1^{(k)}}}{\partial u_j^2} & \cdots & \frac{\partial \Phi_k^{\beta_1^{(k)}}}{\partial u_j^m} \end{pmatrix} =: \frac{\partial \Phi_k^\alpha}{\partial u_j^\beta} \,. \tag{2.24}$$

Since according to Remark 2.3.13 the columns in the symbol matrix are ordered decreasingly with regard to their classes (defined by the multi-indices μ in the pairs (α, μ) which label the columns) and within a class according to the indices α (such that a column with label (α, μ) is left to the column with label (β, μ) if $\alpha < \beta$), the rows in the symbol matrix which correspond to the $\beta_1^{(k)}$ rows of class k consist of the n blocks of the form (2.24) ordered decreasingly with regard to j. As any equation solved for a derivative u_k^α depends only on first-order derivatives of a class lower than k and on the u_k^δ where $\beta_1^{(k)} < \delta \leq m$, in the special cases where $k < j$ we have

$$\frac{\partial \Phi_k^\alpha}{\partial u_j^\beta} = 0_{\beta_1^{(k)} \times m} \,,$$

and for $k = j$ follows

$$\frac{\partial \Phi_k^\alpha}{\partial u_j^\beta} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & -\frac{\partial \phi_k^1}{\partial u_k^{\beta_1^{(k)}+1}} & -\frac{\partial \phi_k^1}{\partial u_k^{\beta_1^{(k)}+2}} & \dots & -\frac{\partial \phi_k^1}{\partial u_k^m} \\ 0 & 1 & & 0 & 0 & -\frac{\partial \phi_k^2}{\partial u_k^{\beta_1^{(k)}+1}} & -\frac{\partial \phi_k^2}{\partial u_k^{\beta_1^{(k)}+2}} & \dots & -\frac{\partial \phi_k^2}{\partial u_k^m} \\ \vdots & & \ddots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 & -\frac{\partial \phi_k^{\beta_1^{(k)}}}{\partial u_k^{\beta_1^{(k)}+1}} & -\frac{\partial \phi_k^{\beta_1^{(k)}}}{\partial u_k^{\beta_1^{(k)}+2}} & \dots & -\frac{\partial \phi_k^{\beta_1^{(k)}}}{\partial u_k^m} \end{pmatrix} =: \left(\mathbb{1}_{\beta_1^{(k)}}, -\frac{\partial \phi_k^\alpha}{\partial u_k^\delta}\right).$$

For each $1 \leq k \leq n$, there is a row of n blocks $\frac{\partial \Phi_k^\alpha}{\partial u_j^\beta}$. The symbol matrix consists of these n rows stacked on one another according to decreasing k. As a consequence, the rows in the symbol matrix which correspond to the $\beta_1^{(k)}$ rows of class k are

$$\left(\frac{\partial \Phi_k^\alpha}{\partial u_n^\beta}\middle|\frac{\partial \Phi_k^\alpha}{\partial u_{n-1}^\beta}\middle|\cdots\middle|\frac{\partial \Phi_k^\alpha}{\partial u_{k+1}^\beta}\middle|\left(\mathbb{1}_{\beta_1^{(k)}}, -\frac{\partial \phi_k^\alpha}{\partial u_k^\delta}\right)\middle| -\frac{\partial \phi_k^\alpha}{\partial u_{k-1}^\gamma}\middle|\cdots\middle| -\frac{\partial \phi_k^\alpha}{\partial u_1^\gamma}\right) =$$
$$\left(0_{\beta_1^{(k)}\times m}\middle|0_{\beta_1^{(k)}\times m}\middle|\cdots\middle|0_{\beta_1^{(k)}\times m}\middle|\left(\mathbb{1}_{\beta_1^{(k)}}, -\frac{\partial \phi_k^\alpha}{\partial u_k^\delta}\right)\middle| -\frac{\partial \phi_k^\alpha}{\partial u_{k-1}^\gamma}\middle|\cdots\middle| -\frac{\partial \phi_k^\alpha}{\partial u_1^\gamma}\right).$$

Thus the symbol of a representation in Cartan normal form is automatically given in row echelon form. See Figure 2.1 for a sketch of the complete symbol matrix. Now the rank of the symbol matrix is obviously $\operatorname{rank} M_1 = \sum_{k=1}^n k\beta_1^{(k)}$.

For a system given in Cartan normal form, there is the following criterion for involutivity.

Lemma 2.4.29. *Let the differential equation $\mathcal{R}_1$ be given in Cartan normal form (2.23). Then $\mathcal{R}_1$ is involutive if, and only if, all non-multiplicative prolongations of the equations (2.23a–2.23c) and all formal derivatives with respect to all the x^i of the algebraic equations (2.23d) are dependent on the equations of the system (2.23) and its multiplicative prolongations only.*

Proof. Let the differential equation $\mathcal{R}_1$ be represented by the system (2.23) in Cartan normal form. Assume that it is involutive. Then its symbol is involutive. According to Proposition 2.4.7, any second-order derivative obtained by a non-multiplicative prolongation can be linearly combined from second-order derivatives obtained by multiplicative prolongations only. Thus, there are coefficient functions A_β^{ij} such that for $1 \leq \ell < k \leq n$, when x^k is non-multiplicative for the equation $u_\ell^\alpha = \phi_\ell^\alpha$, we have

$$\frac{\partial D_k(u_\ell^\alpha - \phi_\ell^\alpha)}{\partial u_{\ell k}^\alpha} = \sum_{i=1}^{k} \sum_{\beta=1}^{\beta_1^{(i)}} \sum_{j=1}^{i} A_\beta^{ij} \frac{\partial D_j(u_i^\beta - \phi_i^\beta)}{\partial u_{ij}^\beta}.$$

Now we take into account that the formal derivative D_k applied to $u_\ell^\alpha - \phi_\ell^\alpha$ is by definition the chain rule yielding the total derivative of $u_\ell^\alpha - \phi_\ell^\alpha = 0$ with respect to x^k. This means there are additional coefficient functions B^i_β and C_β such that the formal derivative of $u^\alpha_\ell - \phi^\alpha_\ell$ is

$$\begin{aligned} D_k(u^\alpha_\ell - \phi^\alpha_\ell) &= \sum_{i=1}^{k} \sum_{\beta=1}^{\beta_1^{(i)}} \sum_{j=1}^{i} A^{ij}_\beta D_j(u^\beta_i - \phi^\beta_i) \\ &\quad + \sum_{i=1}^{k} \sum_{\beta=1}^{\beta_1^{(i)}} B^i_\beta (u^\beta_i - \phi^\beta_i) + \sum_{\beta=1}^{\beta_0} C_\beta (u^\beta_i - \phi^\beta_i) \ . \end{aligned}$$

Without the functions B^i_β and C_β, from the terms of the first part an integrability condition could arise. Nor arise integrability conditions from prolonging the algebraic conditions (2.23d), since an involutive system is formally integrable. This means, there are coefficient functions D_β such that for $1 \le k \le n$ the formal derivative of $u^\alpha_k - \phi^\alpha_k$ with respect to x^k is

$$u^\alpha_k - \frac{\partial \phi^\alpha}{\partial x^k} - \sum_{\beta=\beta_0+1}^{m} \frac{\partial \phi^\alpha}{\partial u^\beta} \phi^\beta_k \ = \ \sum_{\beta=1}^{\beta_0} D_\beta (u^\beta - \phi^\beta) \ .$$

Thus, any non-multiplicative prolongation of an equation in the system (2.23a–2.23c) and any formal derivative with respect to an x^i of an algebraic equation in (2.23d) is dependent on the equations of the system (2.23) and its multiplicative prolongations only.

The argument is also true if followed through backwards, thus showing the equivalence in the claim. □

Corollary 2.4.30. *Let $\mathcal{R}_1$ be a differential equation in one dependent coordinate u. Then its symbol $\mathcal{N}_1$ is involutive.*

Proof. By numbering the x^i adequately, we can arrange the Cartan normal form for $\mathcal{R}_1$ in such a way that it is exactly the equations of the lowest classes which are missing, if any. Let the system be represented by the equations solved for u_{x^n} to $u_{x^{n-s}}$. Its Cartan normal form is

$$\mathcal{R}_1 \colon \begin{cases} u_{x^n} = \phi_n(\mathbf{x}; u, u_{x^1}, u_{x^2}, \dots, u_{x^{n-1}}) \\ u_{x^{n-1}} = \phi_{n-1}(\mathbf{x}; u, u_{x^1}, u_{x^2}, \dots, u_{x^{n-2}}) \\ \quad \vdots \\ u_{x^{n-s}} = \phi_{n-s}(\mathbf{x}; u, u_{x^1}, u_{x^2}, \dots, u_{x^{n-s-1}}) \\ u = \phi_0(\mathbf{x}; u) \end{cases} .$$

(If $n-s=1$ then the system is maximally over-determined.) We have to show rank $M_2 = \sum_{k=1}^n k\beta_1^{(k)}$. First consider the right side of this equality. For the above representation we have for all k where $n-s \le k \le n$ the indices $\beta_1^{(k)} = 1$ and for all k where $1 \le k \le n-s-1$ the indices $\beta_1^{(k)} = 0$. It follows that $\sum_{k=1}^n k\beta_1^{(k)} = \sum_{k=n-s}^n k$. Now consider the left side. Since the differential equation $\mathcal{R}_1$ is represented by a system in row echelon form, the prolongation of any equation $u_{x^k} = \phi_k$ with respect to its multiplicative variables x^i, $1 \le i \le k$, has a different leader $u_{x^i x^k}$ than any other prolonged equation and thus its corresponding row in the symbol matrix M_2 is independent of the others. Therefore each

u_{x^k} leads to k independent leaders $u_{x^i x^k}$ in the representation of $\mathcal{R}_2$ and thus $\operatorname{rank} M_2 \geq \sum_{k=n-s}^{n} k$. For $k < \ell \leq n$, the leader $u_{x^k x^\ell}$ leads to an entry in M_2 which is in the same column as the pivot of the multiplicative prolongation $u_{x^\ell x^k}$. Therefore it does not change the rank of M_2. The prolongations of the algebraic equations render first-order equations which do not influence the matrix M_2 of the second-order symbol. Neither do the equations in the representation of the original system, as they are of first order. Thus we also have $\operatorname{rank} M_2 \leq \sum_{k=n-s}^{n} k$. □

For easier reference later, we now formulate the Cartan-Kähler theorem for an involutive first-order system in Cartan normal form without algebraic equations.

Theorem 2.4.31 (Cartan-Kähler). *Let the differential equation $\mathcal{R}_1$ be locally represented by the system (2.23a, 2.23b, 2.23c). Assume that the following initial condition are given:*

$$u^\alpha(x^1,\ldots,x^n) = f^\alpha(x^1,\ldots,x^n)\ , \qquad \beta_1^{(n)} < \alpha \leq m\ ; \tag{2.25a}$$

$$u^\alpha(x^1,\ldots,x^{n-1},0) = f^\alpha(x^1,\ldots,x^{n-1})\ , \qquad \beta_1^{(n-1)} < \alpha \leq \beta_1^{(n)}\ ; \tag{2.25b}$$

$$\vdots \qquad\qquad \vdots$$

$$u^\alpha(x^1,0,\ldots,0) = f^\alpha(x^1)\ , \qquad \beta_1^{(1)} < \alpha \leq \beta_1^{(2)}\ ; \tag{2.25c}$$

$$u^\alpha(0,\ldots,0) = f^\alpha\ , \qquad 1 \leq \alpha \leq \beta_1^{(1)}\ . \tag{2.25d}$$

Let the functions ϕ_k^α and f^α be real-analytic at the origin and let the system (2.23a, 2.23b, 2.23c) be involutive. Then this system has one and only one solution that is analytic at the origin and satisfies the initial conditions (2.25).

Proof. For the proof, see Pommaret [35] or Seiler [37]. The strategy is to split the system into subsystems according to the classes of the equations in it. The solution is constructed step by step; each step renders a normal system to which the Cauchy-Kovalevskaya theorem is applied. □

Using a transformation proposed by Drach [8], any differential Equation $\mathcal{R}_q$ may be turned into an equivalent one with only one dependent variable. If $\mathcal{R}_q$ is transformed into a first-order system before, Drach's transformation yields a differential equation of order one or two. In this sense, all partial differential equations belong to one of these two classes. The last corollary shows that for the one class, first-order equations in one dependent variable, the symbol is always involutive. See also Stormark [40], Chapter 5, for details with regard to Drach's classification. Since many kinds of differential equations (ordinary differential equations and, more generally, systems of finite type; all differential equations of first order in one dependent variable, which is one of only two classes according to Drach's classification) have an involutive symbol, for a great many differential equations there is no difference between involutivity and formal integrability. This is why the difference is sometimes overlooked.

2.5 Useful Properties of First Order Systems

It simplifies the notation in our subsequent proofs of two of the main results (Theorems 3.3.9 and 3.3.28) if we consider first-order equations $\mathcal{R}_1 \subseteq J_1\pi$. Furthermore, we will assume that any present algebraic (i. e. zeroth-order) equation has been explicitly solved and substituted into the equations Φ^τ, reducing thus the number of dependent variables. From a theoretical point of view this does not represent a restriction, as any differential equation $\mathcal{R}_q$ can be transformed into an equivalent first-order one and under some mild regularity assumptions the algebraic equations can always be solved locally. In this section we show that the crucial properties of a differential equation remain unchanged through this process.

2.5.1 Reduction to First Order

First we consider the transformation of a system of order $q > 1$ into a first-order system. There are several approaches; we follow the one demonstrated by Seiler [37] and summarize it here for the sake of completeness. Let the differential equation $\mathcal{R}_q \subseteq J_q\pi$ be locally represented by $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$. Consider the first-order jet bundle $J_1\pi^{q-1}$ with local coordinates $(\mathbf{x}, (\mathbf{u}^{(q-1)})^{(1)})$; its total space is $J_{q-1}\pi$. We identify $J_q\pi$ with a submanifold in $J_1\pi^{q-1}$ by the immersion

$$\iota_{q,1} \colon J_q\pi \to J_1\pi^{q-1} ,$$

which is given in Equation (2.2). Then it is natural to set

$$\tilde{\mathcal{R}}_1 := \iota_{q,1}(\mathcal{R}_q) \subseteq J_1\pi^{q-1} .$$

From the local representation of $\mathcal{R}_q$, we can derive a local representation for $\tilde{\mathcal{R}}_1$: equations which define $\iota_{q,1}(J_q\pi)$ as a submanifold in $J_1\pi^{q-1}$ are

$$\begin{aligned} u^\alpha_{\mu,i} &= u^\alpha_{\mu+1_i} , \quad 0 \le |\mu| < q-1 , \quad 1 \le i \le n, \\ u^\alpha_{\mu,i} &= u^\alpha_{\mu-1_k+1_i,k} , \quad |\mu| = q-1 , \quad k = \operatorname{cl}\mu < i \le n. \end{aligned}$$

It remains to derive equations $\tilde{\Phi}^\tau(\mathbf{x}, (\mathbf{u}^{(q-1)})^{(1)}) = 0$ to describe $\tilde{\mathcal{R}}_1$ within $J_1\pi^{q-1}$ from the equations $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$. We do so by expressing each derivative u^α_μ of the original representation through one of the new coordinates like this:

$$u^\alpha_\mu \mapsto \begin{cases} u^\alpha_\mu : & |\mu| \le q-1 , \\ u^\alpha_{\mu-1_k,k} : & |\mu| = q , \ \operatorname{cl}\mu = k . \end{cases}$$

The following proposition again is from Seiler [37] and says that this transformation does not change those properties of the system which are important assumptions for the results which we want to prove later.

Proposition 2.5.1. *Let the differential equation $\mathcal{R}_q$ be transformed into a first-order differential equation $\tilde{\mathcal{R}}_1$ by the procedure given above. Then*

1. *the differential equation $\tilde{\mathcal{R}}_1$ is involutive if, and only if, the differential equation $\mathcal{R}_q$ is involutive;*

2. *for all* $1 \leq k \leq n$ *the Cartan characters* $\alpha_q^{(k)}$ *of* $\mathcal{R}_q$ *and the Cartan characters* $\tilde{\alpha}_1^{(k)}$ *of* $\tilde{\mathcal{R}}_1$ *satisfy* $\alpha_q^{(k)} = \tilde{\alpha}_1^{(k)}$*; and*

3. *there is a one-to-one correspondence between the formal solutions of* $\mathcal{R}_q$ *and those of* $\tilde{\mathcal{R}}_1$.

Proof. For a straightforward proof, see Seiler [37], Appendix A. □

Example 2.5.2. The one-dimensional wave equation $u_{tt} = u_{xx}$, given in Example 2.2.2, has an involutive symbol and is formally integrable; therefore it is involutive. It may be represented as a first-order system $\tilde{\mathcal{R}}_1 \subset J_1\pi^1$, where we use x, t, u, u_x, u_t, $u_{,x}$, $u_{,t}$, $u_{x,x}$, $u_{x,t}$, $u_{t,x}$ and $u_{t,t}$ for coordinates on $J_1\pi^1$, by

$$\tilde{\mathcal{R}}_1 \colon \begin{cases} u_{t,t} - u_{x,x} = 0 \ , \\ u_{,x} = u_x \ , \\ u_{,t} = u_t \ , \\ u_{x,t} = u_{xt} \ . \end{cases}$$

For $\mathcal{R}_2$, only the first-class derivatives u_{xx} and u_{xt} are parametric, therefore we have $\alpha_2^{(1)} = 2$ and $\alpha_2^{(2)} = 0$, and since for $\tilde{\mathcal{R}}_1$ only the first-class derivatives $u_{t,x}$ and $u_{x,x}$ are parametric while $u_{,x}$, $u_{,t}$, $u_{x,t}$ and $u_{t,t}$ are principal derivatives, we have $\tilde{\alpha}_1^{(1)} = 2$ and $\tilde{\alpha}_1^{(2)} = 0$. So indeed $\alpha_q^{(k)} = \tilde{\alpha}_1^{(k)}$ for $k \in \{1, 2\}$. Note that $\beta_2^{(1)} = 0 \neq 1 = \tilde{\beta}_1^{(1)}$ and $\beta_2^{(2)} = 1 \neq 3 = \tilde{\beta}_1^{(2)}$.

To simplify the notation, set $u_t =: v$ and $u_x =: w$. Then the coordinates of $J_1\pi^1$ turn into x, t, u, w, v, u_x, u_t, w_x, w_t, v_x and v_t and may be regarded as the coordinates of the first-order jet bundle $J_1\mathrm{pr}_1$ over the trivial bundle $(\mathbb{R}^2 \times \mathbb{R}^3, \mathrm{pr}_1, \mathbb{R}^2)$ where $(t, x; u, v, w)$ are global coordinates for the total space $\mathbb{R}^2 \times \mathbb{R}^3$. The resulting system of four equations is the representation for $\mathcal{R}_1^{(1)}$ in Example 2.2.13, defining a submanifold in $J_1\mathrm{pr}_1$ by the map $\Phi : J_1\mathrm{pr}_1 \to \mathbb{R}^4$ with

$$\Phi(j_1\gamma|_{(x,t)}) := \big(\Phi^\tau(j_1\gamma|_{(x,t)}) : 1 \leq \tau \leq 4\big) \ ,$$

where the maps $\Phi^\tau : J_1\mathrm{pr}_1 \to \mathbb{R}$ are given by

$$\Phi^1 = u_t - v, \quad \Phi^2 = u_x - w, \quad \Phi^3 = v_t - w_x, \quad \Phi^4 = v_x - w_t \ .$$

We then have

$$\mathcal{R}_1^{(1)} = \{j_1\gamma|_{(x,t)} \in J_1\mathrm{pr}_1 : \Phi^\tau(j_1\gamma|_{(x,t)}) = 0\} \ .$$

Note that $\mathcal{R}_1^{(1)}$ contains the integrability condition $w_t = v_x$, which corresponds to $u_{x,t} = u_{xt}$ of $\tilde{\mathcal{R}}_1$ and was automatically produced through the transformation described above. So the system is indeed involutive, as opposed to the system of three equations in Example 2.2.11.

The section $\gamma : \mathbb{R}^2 \to \mathbb{R}^2 \times \mathbb{R}^3$, defined by $\gamma(x, t) = (x, t; x^3 + 3xt^2)$, is a global solution of the differential equation $\mathcal{R}_1$, since from

$$v = u_t = 3x^2 + 3t^2, \quad w = u_t = 6xt, \quad v_t = w_x = 6t \quad \text{and} \quad v_x = w_t = 6x$$

it follows for all $(x, t) \in \mathbb{R}^2$ that $j_1\gamma|_{(x,t)} \in \mathcal{R}_1^{(1)}$. This is the same solution as the one in Example 2.2.7 for $\mathcal{R}_2$.

We can now prove that, unlike the indices $\beta_q^{(k)}$ of a differential equation, the Cartan characters of $\mathcal{R}_q$ form a monotonous series for all $q \geq 1$.

Corollary 2.5.3. *For a differential equation $\mathcal{R}_q$, for which the dimension of the base space is n, the Cartan characters satisfy the following chain of inequalities:*

$$\alpha_q^{(1)} \geq \alpha_q^{(2)} \geq \cdots \geq \alpha_q^{(n)} \geq 0 \; . \tag{2.26}$$

Proof. According to Proposition 2.5.1, the differential equation $\mathcal{R}_q$ can be transformed into a first-order system which has the same Cartan characters $\tilde{\alpha}_1^{(i)} = \alpha_q^{(i)}$. The claim now follows from Proposition 2.4.25. □

2.5.2 Obstructions to Involution for Equations in Cartan Normal Form

For later use, we modify the Cartan normal form of a differential equation as given in Definition 2.4.26 into the *reduced Cartan normal form*. It arises by solving each equation for a derivative u_j^α, the principal derivative, and eliminating this derivative from all other equations. Again, the principal derivatives are chosen in such a manner that their classes are as great as possible. Now none of the principal derivatives appears on a right side of an equation whereas this was possible with the (non-reduced) Cartan normal form of Definition 2.4.26. All the remaining, non-principal, derivatives are called *parametric*. Ordering the obtained equations by their class, we can decompose them into subsystems:

$$u_n^\alpha = \phi_n^\alpha(\mathbf{x}, \mathbf{u}, u_j^\gamma) \qquad \begin{cases} 1 \leq \alpha \leq \beta_1^{(n)} \\ 1 \leq j \leq n \\ \beta_1^{(j)} < \gamma \leq m \end{cases} , \tag{2.27a}$$

$$u_{n-1}^\alpha = \phi_{n-1}^\alpha(\mathbf{x}, \mathbf{u}, u_j^\gamma) \qquad \begin{cases} 1 \leq \alpha \leq \beta_1^{(n-1)} \\ 1 \leq j \leq n-1 \\ \beta_1^{(j)} < \gamma \leq m \end{cases} , \tag{2.27b}$$

$$\vdots$$

$$u_1^\alpha = \phi_1^\alpha(\mathbf{x}, \mathbf{u}, u_j^\gamma) \qquad \begin{cases} 1 \leq \alpha \leq \beta_1^{(1)} \\ 1 = j \\ \beta_1^{(j)} < \gamma \leq m \end{cases} . \tag{2.27c}$$

A more compact notation for the Cartan normal form is

$$u_h^\alpha = \phi_h^\alpha(\mathbf{x}, \mathbf{u}, u_j^\gamma) \qquad \begin{cases} 1 \leq j \leq h \leq n \\ 1 \leq \alpha \leq \beta_1^{(h)} \\ \beta_1^{(j)} < \gamma \leq m \end{cases} .$$

Note that the values $\beta_1^{(k)}$ are exactly those appearing in the Cartan test (2.21), as the symbol matrix of a differential equation in Cartan normal form is automatically triangular with the principal derivatives as pivots. A sketch for the symbol matrix of a system in this reduced Cartan normal form is shown in Figure 2.1: there, the areas shaded gray in the sketch now contain only zero entries (the matrix is in reduced row echelon form then, or in Gauß-Jordan form). According to Remark 2.4.24 the Cartan characters of $\mathcal{R}_1$ are $\alpha_1^{(k)} = m - \beta_1^{(k)}$ and thus equal the number of parametric derivatives of class k.

Example 2.5.4. For $n = 3$ and $m = 5$ with $\mathbf{x} = (x, y, z)$ and $\mathbf{u} = (u^1, u^2, u^3, u^4, u^5)$, the following system is in the reduced Cartan normal form:

$$\begin{aligned}
u_z^1 &= \phi_z^1(\mathbf{x}; \mathbf{u}; u_x^3, u_x^4, u_x^5, u_y^4, u_y^5, u_z^5) \\
u_z^2 &= \phi_z^2(\mathbf{x}; \mathbf{u}; u_x^3, u_x^4, u_x^5, u_y^4, u_y^5, u_z^5) \\
u_z^3 &= \phi_z^3(\mathbf{x}; \mathbf{u}; u_x^3, u_x^4, u_x^5, u_y^4, u_y^5, u_z^5) \\
u_z^4 &= \phi_z^4(\mathbf{x}; \mathbf{u}; u_x^3, u_x^4, u_x^5, u_y^4, u_y^5, u_z^5) \\
u_y^1 &= \phi_y^1(\mathbf{x}; \mathbf{u}; u_x^3, u_x^4, u_x^5, u_y^4, u_y^5) \\
u_y^2 &= \phi_y^2(\mathbf{x}; \mathbf{u}; u_x^3, u_x^4, u_x^5, u_y^4, u_y^5) \\
u_y^3 &= \phi_y^3(\mathbf{x}; \mathbf{u}; u_x^3, u_x^4, u_x^5, u_y^4, u_y^5) \\
u_x^1 &= \phi_x^1(\mathbf{x}; \mathbf{u}; u_x^3, u_x^4, u_x^5) \\
u_x^2 &= \phi_x^2(\mathbf{x}; \mathbf{u}; u_x^3, u_x^4, u_x^5)
\end{aligned}$$

Here the dimension of the symbol is 6, the Cartan characters are $\alpha_1^{(1)} = 3$, $\alpha_1^{(2)} = 2$, and $\alpha_1^{(3)} = 1$, and the indices of the equation are $\beta_1^{(1)} = 2$, $\beta_1^{(2)} = 3$, $\beta_1^{(3)} = 4$.

For a differential equation $\mathcal{R}_1$ in Cartan normal form, it is possible to perform an involution analysis in closed form. According to Lemma 2.4.29 an effective test of involution proceeds as follows. Each equation in (2.27) is prolonged with respect to each of its non-multiplicative variables. The arising second-order equations are simplified modulo the original system and the prolongations with respect to the multiplicative variables. The symbol $\mathcal{N}_1$ is involutive if, and only if, after the simplification none of the equations is of second order any more. The differential equation $\mathcal{R}_1$ is involutive if, and only if, all new equations simplify to zero, as any remaining first-order equation would be an integrability condition. In order to apply this test, we now prove some helpful lemmata. We set $\mathcal{B} := \{(\alpha, i) \in \mathbb{N}^m \times \mathbb{N}^n : u_i^\alpha \text{ is a principal derivative}\}$, and for each $(\alpha, i) \in \mathcal{B}$ we define $\Phi_i^\alpha := u_i^\alpha - \phi_i^\alpha$. Now any prolongation of some Φ_i^α has the following explicit form.

Lemma 2.5.5. *Let the differential equation $\mathcal{R}_1$ be represented in the reduced Cartan normal form given by Equation (2.27). Then for any $(\alpha, i) \in \mathcal{B}$ and $1 \leq j \leq n$, we have*

$$D_j \Phi_i^\alpha = u_{ij}^\alpha - C_j^{(1)}(\phi_i^\alpha) - \sum_{h=1}^{i} \sum_{\gamma=\beta_1^{(h)}+1}^{m} u_{hj}^\gamma C_\gamma^h(\phi_i^\alpha) \,. \tag{2.28}$$

Proof. By straightforward calculation: according to the definition of the total derivative in Equation (2.8), which is nothing but the chain rule, and since here $q = 1$, we have

$$\begin{aligned} D_j\Phi_i^\alpha &= \frac{\partial \Phi_i^\alpha}{\partial x^j} + \sum_{0\le|\mu|\le 1}\sum_{\gamma=1}^{m} \frac{\partial \Phi_i^\alpha}{\partial u_\mu^\gamma} u_{\mu+1_j}^\gamma \\ &= \frac{\partial(u_i^\alpha - \phi_i^\alpha)}{\partial x^j} + \sum_{\gamma=1}^{m} \frac{\partial(u_i^\alpha - \phi_i^\alpha)}{\partial u^\gamma} u_j^\gamma + \sum_{h=1}^{n}\sum_{\gamma=1}^{m} \frac{\partial(u_i^\alpha - \phi_i^\alpha)}{\partial u_h^\gamma} u_{hj}^\gamma \ . \end{aligned}$$

Since the system is in reduced Cartan normal form, any right side ϕ_i^α depends at most on parametric derivatives u_h^γ where $(\gamma, h) \in \mathcal{B}$ and $1 \le h \le i$. For the special case $h = i$, we have $C_\gamma^h(\Phi_i^\alpha) = \delta_{hi} \cdot \delta_{\alpha\gamma}$ because the equation is in solved form. Therefore the double sum simplifies considerably: for $h = i$ and $\gamma = \beta_1^{(i)}$, we obtain u_{ij}^α (and then move it to the leftmost position in the formula), while otherwise only for $1 \le h \le i$ and $\beta_1^{(h)} + 1 \le \gamma \le m$ summands may be non-trivial. It follows that the prolongation may be expressed as

$$D_j\Phi_i^\alpha = u_{ij}^\alpha - \partial_{x^j}(\phi_i^\alpha) - \sum_{\gamma=1}^{m} \partial_{u^\gamma}(\phi_i^\alpha) u_j^\gamma - \sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{m} \partial_{u_h^\gamma}(\phi_i^\alpha) u_{hj}^\gamma \ .$$

Now the claim follows from the definition of the contact vector fields in Equation (2.4). □

For $j > i$, the prolongation $D_j\Phi_i^\alpha$ is non-multiplicative, otherwise it is multiplicative. Now let $j > i$, so that (2.28) is a non-multiplicative prolongation. According to our criterion in Lemma 2.4.29, the symbol $\mathcal{N}_1$ is involutive if, and only if, it is possible to eliminate on the right hand side of (2.28) all second-order derivatives by adding multiplicative prolongations, and the differential equation $\mathcal{R}_1$ is involutive if, and only if, all new equations simplify to zero, as any remaining first-order equation would be an integrability condition. The following proposition shows the explicit linear combination.

Proposition 2.5.6. *Let the differential equation $\mathcal{R}_1$ be represented in the reduced Cartan normal form given by Equation (2.27), and let $\mathcal{R}_1$ be involutive. Then for $1 \le i < j \le n$, any non-multiplicative prolongation $D_j\Phi_i^\alpha$ can be combined from multiplicative prolongations like this:*

$$D_j\Phi_i^\alpha = D_i\Phi_j^\alpha - \sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) D_h\Phi_j^\gamma \ . \tag{2.29}$$

Proof. We have to eliminate all second-order terms on the right side of Equation (2.28) by subtracting multiples of multiplicative prolongations only. To reduce the term u_{ij}^α to zero, we subtract the multiplicative prolongation $D_i\Phi_j^\alpha$, which is available for the following reason: we have $i < j$ and therefore $\beta_1^{(i)} \le \beta_1^{(j)}$. Thus, as Φ_i^α is in the system, so is Φ_j^α. The other second-order terms in $D_i\Phi_j^\alpha$ are, according to Equation (2.28), for $1 \le h \le i$ and $\beta_1^{(h)} + 1 \le \gamma \le m$ the summands $-u_{hj}^\gamma C_\gamma^h(\phi_i^\alpha)$, and we have to eliminate them as well by subtracting multiples of suitable multiplicative prolongations. For $\beta_1^{(h)} + 1 \le \gamma \le \beta_1^{(q)}$, the equation Φ_j^γ is in the system. Thus, the multiplicative prolongation $D_h\Phi_j^\gamma$ can be multiplied by $-C_\gamma^h(\phi_i^\alpha)$ and then subtracted to eliminate the term $-u_{hj}^\gamma C_\gamma^h(\phi_i^\alpha)$. Since for $\beta_1^{(q)}+1 \le \gamma \le m$, there is no equation Φ_j^γ in the system, the summands $-u_{hj}^\gamma C_\gamma^h(\phi_i^\alpha)$ cannot be eliminated by subtracting a linear combination of multiplicative prolongations and thus constitute the obstruction to involution

$$\sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(j)}+1}^{m} u_{hj}^\gamma C_\gamma^h(\phi_i^\alpha) = 0 \; .$$

The obstruction must vanish because $\mathcal{R}_1$ is involutive according to our assumption. □

Example 2.5.7. Consider the differential equation $\mathcal{R}_1$, represented by

$$\mathcal{R}_1\colon \begin{cases} v_t = \phi_t^v(x,t;u,v;v_x) \\ u_t = \phi_t^u(x,t;u,v;v_x) \\ u_x = \phi_x^u(x,t;u,v;v_x) \end{cases} ,$$

where $\Phi_t^v = v_t - \phi_t^v$, $\Phi_t^u = u_t - \phi_t^u$ and $\Phi_x^u = u_x - \phi_x^u$. The differential equation is in reduced Cartan normal form. The only non-multiplicative prolongation is

$$D_t\Phi_x^u = u_{xt} - C_t^{(1)}(\phi_x^u) - v_{xt}C_v^x(\phi_x^u) \; .$$

To eliminate u_{xt}, we subtract the multiplicative prolongation $D_x\Phi_t^u = u_{xt} - C_x^{(1)}(\phi_t^u) - v_{xx}C_v^x(\phi_t^u)$, which contains u_{xt} as well; to eliminate $v_{xt}C_v^x(\phi_x^u)$, we subtract the multiplicative prolongation $D_x\Phi_t^v = v_{xt} - C_x^{(1)}(\phi_t^v) - v_{xx}C_v^x(\phi_t^v)$, multiplied by $-C_v^x(\phi_x^u)$, which contains the term $-C_v^x(\phi_x^u)v_{xt}$. We then have

$$\begin{aligned} &D_t\Phi_x^u - D_x\Phi_t^u + C_v^x(\phi_x^u)D_x\Phi_t^v \\ &= C_x^{(1)}(\phi_t^u) - C_t^{(1)}(\phi_x^u) - C_v^x(\phi_x^u)C_x^{(1)}(\phi_t^v) - v_{xx}\left[-C_v^x(\phi_t^u) + C_v^x(\phi_x^u)C_v^x(\phi_t^v)\right] \; . \end{aligned}$$

If the differential equation $\mathcal{R}_1$ is involutive, the term on the right side vanishes, since the only non-multiplicative prolongation of the system is $D_t\Phi_x^u = D_x\Phi_t^u - C_v^x(\phi_x^u)D_x\Phi_t^v$, a linear combination of multiplicative prolongations.

Note that if the equation for v_t were missing from the system, this would leave only the equations for u_t and u_x. Instead of one, there would be two parametric derivatives, v_x and v_t, and the only non-multiplicative prolongation of the system would be

$$D_t\Phi_x^u = u_{xt} - C_t^{(1)}(\phi_x^u) - v_{xt}C_v^x(\phi_x^u) - v_{tt}C_v^t(\phi_x^u) \; ,$$

where the second-order terms are not a linear combination of the system's multiplicative prolongations $D_x\Phi_t^u$, $D_t\Phi_t^u$ and $D_x\Phi_x^u$. The obstruction to involution $v_{xt}C_v^x(\phi_x^u) = v_{tt}C_v^t(\phi_x^u)$ would arise.

If the differential equation is not involutive, Equation (2.29) is not satisfied. The difference $D_j\Phi_i^\alpha - D_i\Phi_j^\alpha + \sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{m} C_\gamma^h(\phi_h^\alpha)D_h\Phi_j^\gamma$ then does not necessarily vanish but yields an obstruction to involution for any $(\alpha,i)\in\mathcal{B}$ and $1\le i<j\le n$. The next

lemma gives all these obstructions to involution for a first-order system in reduced Cartan normal form. They appear in the proof of the existence theorem for integral distributions in Subsection 3.3.6 where the important point is that they disappear for an involutive differential equation $\mathcal{R}_1$.

Lemma 2.5.8. *For an equation in Cartan normal form, $i < j$ and α such that $(\alpha, i) \in \mathcal{B}$, we have the equality*

$$D_j\Phi_i^\alpha - D_i\Phi_j^\alpha + \sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)D_h\Phi_j^\gamma \tag{2.30}$$

$$= C_i^{(1)}(\phi_j^\alpha) - C_j^{(1)}(\phi_i^\alpha) - \sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_h^{(1)}(\phi_j^\gamma) \tag{2.31}$$

$$-\sum_{h=1}^{i-1}\sum_{\delta=\beta_1^{(h)}+1}^{m} u_{hh}^\delta\left[\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma)\right] \tag{2.32}$$

$$-\sum_{1\leq h<k<i}\left\{\sum_{\delta=\beta_1^{(h)}+1}^{\beta_1^{(k)}} u_{hk}^\delta\left[\sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma)\right]\right. \tag{2.33a}$$

$$\left.+\sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^\delta\left[\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_\delta^k(\phi_j^\gamma) + \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma)\right]\right\} \tag{2.33b}$$

$$-\sum_{h=1}^{i-1}\left\{\sum_{\delta=\beta_1^{(h)}+1}^{\beta_1^{(i)}} u_{hi}^\delta\left[-C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma)\right]\right. \tag{2.34a}$$

$$\left.+\sum_{\delta=\beta_1^{(i)}+1}^{m} u_{hi}^\delta\left[-C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma) + \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_\delta^i(\phi_j^\gamma)\right]\right\} \tag{2.34b}$$

$$-\sum_{\substack{h=1\\ i+1\leq k<j}}^{i-1}\sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^\delta\left[\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_\delta^k(\phi_j^\gamma)\right] \tag{2.35}$$

$$
-\sum_{k=i}^{j-1}\sum_{\delta=\beta_1^{(k)}+1}^{m} u^{\delta}_{ik}\left[-C^k_\delta(\phi^\alpha_j)+\sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C^i_\gamma(\phi^\alpha_i)C^k_\delta(\phi^\gamma_j)\right] \tag{2.36}
$$

$$
-\sum_{h=1}^{i-1}\sum_{\delta=\beta_1^{(j)}+1}^{m} u^{\delta}_{hj}\left[C^h_\delta(\phi^\alpha_i)+\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i)C^j_\delta(\phi^\gamma_j)\right] \tag{2.37}
$$

$$
-\sum_{\delta=\beta_1^{(j)}+1}^{m} u^{\delta}_{ij}\left[C^i_\delta(\phi^\alpha_i)-C^j_\delta(\phi^\alpha_j)+\sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C^i_\gamma(\phi^\alpha_i)C^j_\delta(\phi^\gamma_j)\right]. \tag{2.38}
$$

Proof. Line (2.31) holds all terms of lower than second order; it represents an integrability condition. In Lines (2.32–2.38) for any u^δ (and for each i with $1 \le h \le i$ and each j with $1 \le k \le j$) each of its second-order derivatives u^δ_{hk} appears exactly once. To prove the lemma, we expand our ansatz, Line (2.30), and rearrange the second-order derivatives by factoring out the u^δ_{hk}. As we shall see, for a fixed δ, each coefficient (in Lines (2.32–2.38) the terms in square brackets) belongs to one of seven classes depending on the values of h and k. In Figure (2.2) these classes are represented by the seven areas denoted

(2.32) for u^δ_{hk} where $1 \le h = k \le i-1$;
(2.33) for u^δ_{hk} where $1 \le h < k \le i-1$;
(2.34) for u^δ_{hk} where $1 \le h \le i-1$ and $k = i$;
(2.35) for u^δ_{hk} where $1 \le h \le i-1$ and $i+1 \le k \le j-1$;
(2.36) for u^δ_{hk} where $h = i$ and $i \le k \le j-1$;
(2.37) for u^δ_{hk} where $1 \le h \le i-1$ and $k = j$; and
(2.38) for u^δ_{hk} where $h = i$ and $k = j$.

h \ k	1	2	…	$i-1$	i	$i+1$	…	$j-1$	j
1	(2.32)								
2		(2.32)	(2.33)		(2.34)		(2.35)		(2.37)
⋮			⋱						
$i-1$				(2.32)					
i							(2.36)		(2.38)

Figure 2.2: Each of the terms in square brackets in Lemma 2.5.8 belongs to one of these seven classes, according to h and k in the term u^δ_{hk} before the brackets.

The numbers in the blocks shown in Figure 2.2 correspond to the line numbers in Lemma 2.5.8 where the u^δ_{hk} appear before the terms in square brackets. Expanding our ansatz in Line (2.30) yields:

$$D_j\Phi_i^\alpha - D_i\Phi_j^\alpha + \sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)D_h\Phi_j^\gamma \tag{2.30}$$

$$= u_{ij}^\alpha - C_j^{(1)}(\phi_i^\alpha) - \sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{m} u_{hj}^\gamma C_\gamma^h(\phi_i^\alpha) \tag{2.39}$$

$$- u_{ij}^\alpha + C_i^{(1)}(\phi_j^\alpha) + \sum_{k=1}^{j}\sum_{\delta=\beta_1^{(k)}+1}^{m} u_{ki}^\delta C_\delta^k(\phi_j^\alpha) \tag{2.40}$$

$$+ \sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)\left[u_{hj}^\gamma - C_h^{(1)}(\phi_j^\gamma) - \sum_{k=1}^{j}\sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^\delta C_\delta^k(\phi_j^\gamma)\right] . \tag{2.41}$$

The second-order derivatives u_{ij}^α and $-u_{ij}^\alpha$ cancel, and by dissolving the square bracket in Line (2.41), we arrive at

$$C_i^{(1)}(\phi_j^\alpha) - C_j^{(1)}(\phi_i^\alpha) \tag{2.42}$$

$$+\sum_{k=1}^{j}\sum_{\delta=\beta_1^{(k)}+1}^{m} u_{ki}^\delta C_\delta^k(\phi_j^\alpha) \tag{2.43}$$

$$-\sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{m} u_{hj}^\gamma C_\gamma^h(\phi_i^\alpha) + \sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)u_{hj}^\gamma \tag{2.44}$$

$$-\sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_h^{(1)}(\phi_j^\gamma) \tag{2.45}$$

$$+\sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)\left[-\sum_{k=1}^{j}\sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^\delta C_\delta^k(\phi_j^\gamma)\right] . \tag{2.46}$$

We combine the terms in Line (2.42) with the double sum in Line (2.45). In Line (2.43), we write two minus-signs instead of the plus-sign. Of the two double sums in Line (2.44), only the summands for $\beta_1^{(j)} + 1 \leq \gamma \leq m$ remain. This means, our ansatz (2.30) equals

$$C_i^{(1)}(\phi_j^\alpha) - C_j^{(1)}(\phi_i^\alpha) - \sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_h^{(1)}(\phi_j^\gamma) \tag{2.31}$$

$$-\sum_{k=1}^{j}\sum_{\delta=\beta_1^{(k)}+1}^{m} -u^{\delta}_{ki}C^k_\delta(\phi^\alpha_j) \tag{2.43$'$}$$

$$-\sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(j)}+1}^{m} u^{\gamma}_{hj}C^h_\gamma(\phi^\alpha_i) \tag{2.44$'$}$$

$$+\sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i)\left[-\sum_{k=1}^{j}\sum_{\delta=\beta_1^{(k)}+1}^{m} u^{\delta}_{hk}C^k_\delta(\phi^\gamma_j)\right] . \tag{2.46}$$

In the first line we now recognize the integrability condition, (2.31). In Line (2.43′) we separate the summand for $k=j$ from the rest of the sum. In Line (2.44′), we do the same with the summand for $h=i$. In Line (2.46) we rearrange the summands. We then have that our ansatz is equal to

$$(2.31)$$

$$-\sum_{\delta=\beta_1^{(j)}+1}^{m} -u^{\delta}_{ji}C^j_\delta(\phi^\alpha_j) - \sum_{k=1}^{j-1}\sum_{\delta=\beta_1^{(k)}+1}^{m} -u^{\delta}_{ki}C^k_\delta(\phi^\alpha_j) \tag{2.43$''$}$$

$$-\sum_{\gamma=\beta_1^{(j)}+1}^{m} u^{\gamma}_{ij}C^i_\gamma(\phi^\alpha_i) - \sum_{h=1}^{i-1}\sum_{\gamma=\beta_1^{(j)}+1}^{m} u^{\gamma}_{hj}C^h_\gamma(\phi^\alpha_i) \tag{2.44$''$}$$

$$+\sum_{h=1}^{i}\sum_{k=1}^{j}\sum_{\delta=\beta_1^{(k)}+1}^{m} -u^{\delta}_{hk}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i)C^k_\delta(\phi^\gamma_j) . \tag{2.46$'$}$$

Now we split (2.46′) into four parts: one for $1 \le h \le i-1$ and $1 \le k \le j-1$ giving Line (2.48) below; one for $h=i$ and $1 \le k \le j-1$ giving Line (2.49); one for $1 \le h \le i-1$ and $k=j$, to which we add the summands of Line (2.44″) for $1 \le h \le i-1$, giving Line (2.37), and one for $h=i$ and $k=j$, to which we add the last summand (where $h=i$) of Line (2.44″) and the last summand (where $k=j$) of Line (2.43″), giving Line (2.38). Thus our ansatz, Line (2.30), becomes

$$(2.31)$$

$$-\sum_{k=1}^{j-1}\sum_{\delta=\beta_1^{(k)}+1}^{m} -u^{\delta}_{ik}C^k_\delta(\phi^\alpha_j) \tag{2.47}$$

$$-\sum_{h=1}^{i-1}\sum_{k=1}^{j-1}\sum_{\delta=\beta_1^{(k)}+1}^{m} u^{\delta}_{hk}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i)C^k_\delta(\phi^\gamma_j) \tag{2.48}$$

$$-\sum_{h=1}^{i-1} \sum_{\delta=\beta_1^{(j)}+1}^{m} u_{hj}^{\delta} \left\{ \left[\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^j(\phi_j^\gamma) \right] + C_\delta^h(\phi_i^\alpha) \right\} \tag{2.37}$$

$$-\sum_{k=1}^{j-1} \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{ik}^{\delta} \left[\sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) \right] \tag{2.49}$$

$$-\sum_{\delta=\beta_1^{(j)}+1}^{m} u_{ij}^{\delta} \left\{ \left[-C_\delta^j(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^j(\phi_j^\gamma) \right] + C_\delta^i(\phi_i^\alpha) \right\} . \tag{2.38}$$

Note that now we have found Lines (2.37) and (2.38). We split the sum in Line (2.48) in two, according to whether $k \leq i$ or $k > i$, which yields Lines (2.50) and (2.35) below. The sums in Lines (2.47) and (2.49) have the same summation range; we collect the first summands (for $1 \leq k \leq i$) of both sums into Line (2.52) below and the remaining summands (for $i+1 \leq k \leq j-1$) of both sums into Line (2.51). Now our ansatz, Line (2.30), equals

$$(2.31) + (2.37) + (2.38)$$

$$-\sum_{h=1}^{i-1} \sum_{k=1}^{i} \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^{\delta} \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) \tag{2.50}$$

$$-\sum_{h=1}^{i-1} \sum_{k=i+1}^{j-1} \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^{\delta} \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) \tag{2.35}$$

$$-\sum_{k=i+1}^{j-1} \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{ik}^{\delta} \left[-C_\delta^k(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) \right] \tag{2.51}$$

$$-\sum_{k=1}^{i} \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{ik}^{\delta} \left[-C_\delta^k(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) \right] . \tag{2.52}$$

Note that we have found Line (2.35). From the sum in Line (2.50) we split its last summand (where $k = i$). Since the summation index in Line (2.52) is $1 \leq k \leq i$, we rename it h, and split the other sum according to whether $\delta \leq \beta_1^{(i)}$ or $\delta > \beta_1^{(i)}$. This yields:

$$
\begin{aligned}
&(2.31) + (2.35) + (2.37) + (2.38) + (2.51) \\
&- \sum_{h=1}^{i-1} \sum_{k=1}^{i-1} \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^{\delta} \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) && (2.53)\\
&- \sum_{h=1}^{i-1} \sum_{\delta=\beta_1^{(i)}+1}^{m} u_{hi}^{\delta} \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^i(\phi_j^\gamma) && (2.54)\\
&- \sum_{h=1}^{i} \left\{ \sum_{\delta=\beta_1^{(h)}+1}^{\beta_1^{(i)}} u_{ih}^{\delta} \left[-C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) \right] \right. && (2.55)\\
&\qquad \left. + \sum_{\delta=\beta_1^{(i)}+1}^{m} u_{ih}^{\delta} \left[-C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) \right] \right\} . && (2.56)
\end{aligned}
$$

The sum in Line (2.55) vanishes for $h = i$. From the sum in Line (2.56) we split its last summand (where $h = i$) which yields Line (2.57) below. The remaining sum and Line (2.54) are combined: we can factor out u_{ih}^{δ} in Lines (2.54) and (2.56). Thus we get Lines (2.34a–2.34b):

$$
\begin{aligned}
&(2.31) + (2.35) + (2.37) + (2.38) \\
&- \sum_{k=i+1}^{j-1} \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{ik}^{\delta} \left[-C_\delta^k(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) \right] && (2.51)\\
&- \sum_{h=1}^{i-1} \sum_{k=1}^{i-1} \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^{\delta} \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) && (2.53)\\
&- \sum_{h=1}^{i-1} \left\{ \sum_{\delta=\beta_1^{(h)}+1}^{\beta_1^{(i)}} u_{ih}^{\delta} \left[-C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) \right] \right. && (2.34a)\\
&\qquad \left. + \sum_{\delta=\beta_1^{(i)}+1}^{m} u_{ih}^{\delta} \left[-C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) + \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^i(\phi_j^\gamma) \right] \right\} && (2.34b)\\
&- \sum_{\delta=\beta_1^{(i)}+1}^{m} u_{ii}^{\delta} \left[-C_\delta^i(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^i(\phi_j^\gamma) \right] . && (2.57)
\end{aligned}
$$

The term in Line (2.57) looks like the summand for $k = i$ in the sum in Line (2.51); we combine them and thus get Line (2.36). We sort the terms in Line (2.53) according to whether $h = k$ which yields Line (2.32), $h < k$ or $h > k$, and get:

$$
\begin{aligned}
&(2.31) + (2.34a–2.34b) + (2.35) + (2.37) + (2.38)\\
&-\sum_{k=i}^{j-1} \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{ik}^{\delta} \left[-C_\delta^k(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) \right] && (2.36)\\
&-\sum_{h=1}^{i-1} \sum_{\delta=\beta_1^{(h)}+1}^{m} u_{hh}^{\delta} \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) && (2.32)\\
&-\sum_{h=1}^{i-2} \sum_{k=h+1}^{i-1} \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^{\delta} \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) && (2.58)\\
&-\sum_{k=1}^{i-2} \sum_{h=k+1}^{i-1} \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^{\delta} \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma)\ . && (2.59)
\end{aligned}
$$

We swap the summation indices k and h in Line (2.59), and combine the result with Line (2.58):

$$
\begin{aligned}
&(2.31) + (2.32) + (2.34) + (2.35) + (2.36) + (2.37) + (2.38)\\
&-\sum_{h=1}^{i-2} \sum_{k=h+1}^{i-1} \Bigg\{ \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^{\delta} \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) && (2.60a)\\
&\qquad + \sum_{\delta=\beta_1^{(h)}+1}^{m} u_{hk}^{\delta} \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) \Bigg\}\ . && (2.60b)
\end{aligned}
$$

Now we split the double sum in the curly brackets in Line (2.60b) in two according to whether $\delta \leq \beta_1^{(k)}$ or $\delta > \beta_1^{(k)}$. This makes the term within the curly brackets into

$$
\begin{aligned}
&\sum_{\delta=\beta_1^{(h)}+1}^{\beta_1^{(k)}} u_{hk}^{\delta} \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) && (2.61)\\
&+ \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^{\delta} \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) + \sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^{\delta} \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma)\ . && (2.62)
\end{aligned}
$$

Finally we factor out the second-order derivatives u^{δ}_{hk} in Line (2.62):

$$\sum_{\delta=\beta_1^{(k)}+1}^{m} u^{\delta}_{hk} \left[\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_{\gamma}(\phi^{\alpha}_i) C^k_{\delta}(\phi^{\gamma}_j) + \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^k_{\gamma}(\phi^{\alpha}_i) C^h_{\delta}(\phi^{\gamma}_j) \right] . \tag{2.62$'$}$$

Substituting (2.61+2.62′) for the term in curly brackets in Line (2.60) shows Line (2.60) equals Lines (2.33a–2.33b), and thus we see that our ansatz, Line (2.30), equals

$$(2.31) + (2.32) + (2.33) + (2.34) + (2.35) + (2.36) + (2.37) + (2.38) ,$$

which was to be shown. □

Chapter 3

Vessiot Theory

In Definition 2.4.21, integral elements were introduced; an integral element is a linearization for a local solution σ of a differential equation $\mathcal{R}_q$ in that $\operatorname{im} j_q\sigma \subseteq \mathcal{R}_q$ implies $T_\rho(\operatorname{im} j_q\sigma) \subseteq T_\rho\mathcal{R}_q$, and $T_\rho(\operatorname{im} J_q\sigma)$ is an n-dimensional integral element. Since for an integral element $\mathcal{U}_\rho \subseteq T_\rho\mathcal{R}_q$ there need not exist a local solution σ such that $\mathcal{U}_\rho \subseteq T_\rho(\operatorname{im} j_q\sigma)$, those which are tangent to a prolonged solution are of special interest: for an involutive distribution of integral elements, the Frobenius theorem guarantees the existence of a solution. Now we introduce a distribution on $\mathcal{R}_q$ which is tangent to all prolonged solutions and then analyze the assumptions needed to construct n-dimensional involutive subdistributions within it. If they exist, their integral manifolds are of the form $\operatorname{im} j_q\sigma$ for local solutions σ. The basic idea was proposed by Vessiot [43]. We show that his approach for the construction of such infinitesimal solutions step by step succeeds if, and only if, the differential equation is involutive. This proves the equivalence of Vessiot's approach to the formal theory.

3.1 The Vessiot Distribution

We can consider a differential equation $\mathcal{R}_q \subseteq J_q\pi$ as a manifold in its own right with an atlas of its own. Then we can pull back any one-form $\omega \in \Omega^1(J_q\pi)$ to the differential equation through the inclusion map $\iota\colon \mathcal{R}_q \hookrightarrow J_q\pi$ and obtain a one-form in $\Omega^1(\mathcal{R}_q)$ as follows. Let $T\iota\colon T\mathcal{R}_q \to TJ_q\pi$ be the tangent map of the inclusion map. For a description in local coordinates, let (U, ϕ_U) be a chart for $\mathcal{R}_q$ with local coordinates $\phi_U(\rho) =: \mathbf{x}$ where $\rho \in U$ and $\mathbf{x} = (x^d\colon 1 \le d \le D)$, and let (V, ψ_V) be a chart for $J_1\pi$ with local coordinates $\psi_V(\xi) =: \mathbf{y}$ where $\xi \in V$ and $\mathbf{y} = (y^g\colon 1 \le g \le G)$ and such that $\iota(U) \subseteq V$ and $y^g = \iota^g(x^1, x^2, \dots, x^D)$. The representation of ι in local coordinates is $\bar\iota := \psi_V \circ \iota \circ \phi_U^{-1}$, given by $\bar\iota(\mathbf{x}) = \mathbf{y}$. Let be $(\mathbf{x}, \dot{\mathbf{x}})$ local coordinates on $T\mathcal{R}_q$ and $(\mathbf{y}, \dot{\mathbf{y}})$ on $TJ_q\pi$. Then $T\iota\colon TU \to TV$ has in these coordinates the form

$$\mathbf{x} \mapsto \big(\bar\iota(\mathbf{x}), \frac{\partial\bar\iota(\mathbf{x})}{\partial\mathbf{x}}\dot{\mathbf{x}}\big)\ .$$

This means, $T\iota$ can locally be described by the Jacobian matrix of $\bar\iota$. For any $\rho \in \mathcal{R}_q$ and for all vectors $\vec{v} \in T_\rho \mathcal{R}_q$ set

$$(\iota^* \omega)_\rho(\vec{v}) := \omega_{\iota(\rho)}((T_\rho \iota)(\vec{v})) \ .$$

Then any one-form $\omega \in \Omega^1(J_q\pi)$ which is defined on V looks locally like $\omega = \omega_g dy^g$ for some coefficient functions $\omega_g \colon V \to \mathbb{R}$. The pull-back of ω then is in local coordinates

$$\iota^* \omega = \frac{\partial \iota^g}{\partial \overline{x}^d} \omega_g d\overline{x}^d \ .$$

Only in the trivial case the inclusion map ι is a diffeomorphism. If it is, we can push forward a vector field $\bar{X}$ on the differential equation $\mathcal{R}_q$ to make it into a vector field on the jet bundle $J_q\pi$ by setting

$$\iota_* \bar{X} := T\iota \circ \bar{X} \circ \iota^{-1}.$$

(We consider the vector field $\bar{X}$ as a section and therefore a map.) Otherwise we use that ι is always injective and generalize the concept by considering the diffeomorphism $\iota \colon \mathcal{R}_q \to \mathrm{im}(\iota)$. Now the push-forward of a vector field on $\mathcal{R}_q$ yields a vector field on $\mathrm{im}(\iota) \subset J_q\pi$, though not in general on all of $J_q\pi$. Just the same we write $\iota_* \bar{X}$ for such a vector field instead of the more cumbersome notation $T\iota(\bar{X})$.

3.1.1 Representations of the Vessiot Distribution

By Proposition 2.1.6, for any point $\mathbf{x} \in \mathcal{X}$ and any section $\sigma \in \Gamma_L\pi$ satisfying $\rho = j_q\sigma(\mathbf{x}) \in J_q\pi$, the tangent space $T_\rho(\mathrm{im}\, j_q\sigma)$ to the image of the prolonged section at the point $\rho \in J_q\pi$ is a subspace of the contact distribution $\mathcal{C}_q|_\rho$. If the section σ is a solution of $\mathcal{R}_q$, it furthermore satisfies by definition $\mathrm{im}\, j_q\sigma \subseteq \mathcal{R}_q$, and hence $T(\mathrm{im}\, j_q\sigma) \subseteq T\mathcal{R}_q$. These considerations motivate the following construction.

Definition 3.1.1. *The* Vessiot distribution *of a differential equation $\mathcal{R}_q \subseteq J_q\pi$ is the distribution $\mathcal{V}[\mathcal{R}_q] \subseteq T\mathcal{R}_q$, defined by*

$$T\iota\big(\mathcal{V}[\mathcal{R}_q]\big) = T\iota\big(T\mathcal{R}_q\big) \cap \mathcal{C}_q|_{\mathcal{R}_q} \ .$$

The Vessiot distribution is sometimes called the “Cartan distribution” in the literature (Kuperschmidt [27], Vinogradov [44]). By naming it after Vessiot, we follow Fackerell [15] and Vassiliou [42].

Example 3.1.2. We calculate the Vessiot distribution for the first-order system of Example 2.2.13. For coordinates on $J_1\pi$ choose $x, t; u, v, w; u_x, v_x, w_x, u_t, v_t, w_t$. Since $\mathcal{R}_1^{(1)}$ is represented by a system in solved form, it is natural to choose appropriate local coordinates for $\mathcal{R}_1^{(1)}$, which we bar to distinct them: $\overline{x}, \overline{t}; \overline{u}, \overline{v}, \overline{w}; \overline{v_x}, \overline{w_x}$. The contact codistribution for $J_1\pi$ is generated by the one-forms $\omega^\alpha = du^\alpha - \sum_{i=1}^n u_i^\alpha dx^i$. Written out these are

$$\omega^1 = du - u_x dx - u_t dt, \quad \omega^2 = dv - v_x dx - v_t dt, \quad \omega^3 = dw - w_x dx - w_t dt \ .$$

The tangent space $T\mathcal{R}_1$ is spanned by $\partial_{\overline{x}}, \partial_{\overline{t}}, \partial_{\overline{u}}, \partial_{\overline{v}}, \partial_{\overline{w}}, \partial_{\overline{v_x}}, \partial_{\overline{w_x}}$, and we have $T\iota T\mathcal{R}_1 = \operatorname{span}(\partial_x, \partial_t, \partial_u, \partial_v + \partial_{u_t}, \partial_w + \partial_{u_x}, \partial_{v_x} + \partial_{w_t}, \partial_{w_x} + \partial_{v_t})$, which is annihilated by

$$\omega^4 = du_x - dw, \quad \omega^5 = dv_x - dw_t \text{ (corresponding to the integrability condition)},$$
$$\omega^6 = dw_x - dv_t \quad \text{and} \quad \omega^7 = du_t - dv \ .$$

These seven one-forms annihilate the Vessiot distribution of $\mathcal{R}_1^{(1)}$, which is spanned by the four vector fields

$$\begin{aligned} X_1 &= \partial_x + u_x\partial_u + v_x\partial_v + w_x\partial_w + v_x\partial_{u_t} + w_x\partial_{u_x}, \\ X_2 &= \partial_t + u_t\partial_u + v_t\partial_v + w_t\partial_w + v_t\partial_{u_t} + w_t\partial_{u_x}, \\ X_3 &= \partial_{v_x} + \partial_{w_t} \quad \text{and} \\ X_4 &= \partial_{v_t} + \partial_{w_x}. \end{aligned}$$

Again (like the definition of integral elements), Definition 3.1.1 is not the usual definition found in the literature. The following proposition proves the equivalence to the standard approach.

Proposition 3.1.3. *The Vessiot distribution satisfies* $\mathcal{V}[\mathcal{R}_q] = (\iota^*\mathcal{C}_q^0)^0$.

Proof. Let $\omega \in \mathcal{C}_q^0$ be a contact form, and let $\bar{X}$ be a tangent vector field on $\mathcal{R}_q$. Then from basic differential calculus with push-forwards and pull-backs, it follows that

$$\iota^*\omega(\bar{X}) = \omega(\iota_*\bar{X}) \ .$$

This means that $X \in (\iota^*\mathcal{C}_q^0)^0$ if, and only if, $\iota_*X \in \mathcal{C}_q|_{\mathcal{R}_q}$. Therefore

$$T\iota((\iota^*\mathcal{C}_q^0)^0) = T\iota(T\mathcal{R}_q) \cap \mathcal{C}_q|_{\mathcal{R}_q} \ .$$

As the tangent map $T\iota\colon T\mathcal{R}_q \to TJ_q\pi$ of the inclusion map $\iota\colon \mathcal{R}_q \hookrightarrow J_q\pi$ is injective, the claim follows. ☐

The Vessiot distribution is not necessarily of constant rank along $\mathcal{R}_q$; we will restrict to the case where its rank does not vary over the differential equation.

For a differential equation which is given in explicitly solved form, the inclusion map $\iota\colon \mathcal{R}_q \to J_q\pi$ is available in closed form and can be used to calculate the pull-back of the contact forms. (The linearization of the inclusion map ι is represented by the Jacobian matrix.) This has the advantage of keeping the calculations within a space of smaller dimension.

Example 3.1.4. Reconsider Example 3.1.2: the four vector fields

$$\begin{array}{rcccl} \bar{X}_1 &=& \partial_{\overline{x}} + \overline{u_x}\partial_{\overline{u}} + \overline{v_x}\partial_{\overline{v}} + \overline{w_x}\partial_{\overline{w}} &=& \partial_{\overline{x}} + \overline{w}\partial_{\overline{u}} + \overline{v_x}\partial_{\overline{v}} + \overline{w_x}\partial_{\overline{w}}, \\ \bar{X}_2 &=& \partial_{\overline{t}} + \overline{u_t}\partial_{\overline{u}} + \overline{v_t}\partial_{\overline{v}} + \overline{w_t}\partial_{\overline{w}} &=& \partial_{\overline{t}} + \overline{v}\partial_{\overline{u}} + \overline{w_x}\partial_{\overline{v}} + \overline{v_x}\partial_{\overline{w}}, \\ \bar{X}_3 &=& \partial_{\overline{v_x}}, && \\ \bar{X}_4 &=& \partial_{\overline{w_x}} && \end{array}$$

span $\mathcal{V}[\mathcal{R}_1] \subset T\mathcal{R}_1$ and satisfy $\iota_*\bar{X}_i = X_i$ (as a simple calculation using the Jacobian matrix for $T\iota$ shows). The vector fields $\bar{X}_i$ are annihilated by the pull-backs of the contact forms, $\iota^*\omega^1 = d\overline{u} - \overline{u_x}d\overline{x} - \overline{u_t}d\overline{t}$, $\iota^*\omega^2 = d\overline{v} - \overline{x_x}d\overline{x} - \overline{x_t}d\overline{t}$ and $\iota^*\omega^3 = d\overline{w} - \overline{w_x}d\overline{x} - \overline{w_t}d\overline{t}$. (The pullbacks of the other four one-forms ω^4 to ω^7 vanish on $\mathcal{R}_1$.)

Remark 3.1.5. In Example 3.1.4 the only notable difference between the coordinate expressions of the seven pulled back one forms $\iota^*\omega^j \in \Omega^1(\mathcal{R}_1)$ and their counterparts $\omega^j \in \Omega^1(J_1\pi)$ in Example 3.1.2 is the bar, which was introduced to show that we are using coordinates on $\mathcal{R}_1^{(1)}$. Without this mark, in local coordinates the forms would look alike, and this is why in the literature the pull-back of those forms is sometimes called "restriction." This is lax. It works out fine, as long as all equations of the local representation of the differential equation $\mathcal{R}_q$ are solved for terms of order q, but if equations of lower order are present, there is a notable difference. If we added the algebraic equation $u = \phi^u(x,t)$ to the system $\mathcal{R}_1^{(1)}$ in Example 2.2.13, the one-form ω^u would look the same in local coordinates on $J_1\pi$, but its pull-back onto $\mathcal{R}_1^{(1)}$ would be $\iota^*\omega^u = (\overline{u_x} + \overline{u_t} + 1)d\overline{u} - \overline{u_x}d\overline{x} - \overline{u_t}d\overline{t}$. We shall therefore always bar the local coordinates of the differential equation under consideration to distinguish them.

Instead of regarding the Vessiot distribution as a homogeneous space, we use that it can be naturally split into two subdistributions, one of which is the symbol. This subtle difference combined with the structure of the jet bundle provides the means for our approach to the construction of linear approximations to the solutions of a differential equation and to proving the existence of solutions.

Proposition 3.1.6. *For any differential equation $\mathcal{R}_q$, its symbol is contained in the Vessiot distribution: $\mathcal{N}_q \subseteq \mathcal{V}[\mathcal{R}_q]$. The Vessiot distribution can therefore be decomposed into a direct sum*

$$\mathcal{V}[\mathcal{R}_q] = \mathcal{N}_q \oplus \mathcal{H}$$

for some complement $\mathcal{H}$ (which is not unique).

Proof. Let the differential equation $\mathcal{R}_q$ be locally represented by $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$ where $1 \le \tau \le t$. Since $T\iota(\mathcal{V}[\mathcal{R}_q])$ is defined as $T\iota(T\mathcal{R}_q) \cap \mathcal{C}_q|_{\mathcal{R}_q}$, we have $T\iota(\mathcal{V}[\mathcal{R}_q]) \subseteq \mathcal{C}_q|_{\mathcal{R}_q}$. According to Proposition 2.1.6, the vector fields (2.4) form a basis for $\mathcal{C}_q$. It follows that for any vector field $\bar{X} \in \mathcal{V}[\mathcal{R}_q]$, coordinate functions $a^i, b^\alpha_\mu \in \mathcal{F}(\mathcal{R}_q)$, where $1 \le i \le n$, $1 \le \alpha \le m$ and $|\mu| = q$, exist such that

$$\iota_*\bar{X} = a^i C_i^{(q)} + b^\alpha_\mu C^\mu_\alpha \ .$$

Since the differential equation $\mathcal{R}_q$ is locally represented by $\Phi^\tau = 0$, where $1 \le \tau \le t$, from the tangency of the vector fields in $\mathcal{V}[\mathcal{R}_q]$ follows that $d\Phi^\tau(\iota_*\bar{X}) = \iota_*\bar{X}(\Phi^\tau) = 0$. This means $(a^i C_i^{(q)} + b^\alpha_\mu C^\mu_\alpha)(\Phi^\tau) = 0$, which can be considered as a system of linear equations for the unknown coefficient functions:

$$C_i^{(q)}(\Phi^\tau)a^i + C^\mu_\alpha(\Phi^\tau)b^\alpha_\mu = 0 \tag{3.1}$$

where $1 \le \tau \le t$. The zero vector is trivially a solution for Equation (3.1). The solutions for the case where all $a^i = 0$ are those for

$$C^\mu_\alpha(\Phi^\tau)b^\alpha_\mu = 0 \ . \tag{3.2}$$

The vector fields C^μ_α span $V\pi^q_{q-1}$. Since by Equation (2.4b) $C^\mu_\alpha = \partial_{u^\mu_\alpha}$, we obtain $(\partial\Phi^\tau/\partial u^\mu_\alpha)b^\alpha_\mu = 0$, the symbol equations (2.14). It follows that the symbol $\mathcal{N}_q$ as the vertical part of $T\mathcal{R}_q$ is always contained in $\mathcal{V}[\mathcal{R}_q]$. The symbol contains all the vertical vector fields of the Vessiot distribution since a vector field of the form in Equation (3.1) with at least one non-vanishing coefficient a^i is transversal to both the fibration π^q and the fibration π^q_{q-1}. Any completion of a basis for the symbol to a basis for all of $\mathcal{V}[\mathcal{R}_q]$ defines a complement $\mathcal{H}$. □

Lychagin [29] and Lychagin and Kruglikov [26] discuss such a splitting of the Vessiot distribution into a direct sum of the symbol and a transversal complement, too. (They call the Vessiot distribution "Cartan distribution.") We are going to use such a splitting and the structure of the jet bundle for the construction of convenient bases.

When the equation is not given in solved form or readily transformable explicitly into solved form, a basis for $\mathcal{V}[\mathcal{R}_q]$ is not easily available and calculations have to be carried out within the larger space $TJ_q\pi$ containing $T\iota(\mathcal{V}[\mathcal{R}_q])$. A basis for this subspace can always be found by solving a system of linear equations. We use that we can decompose $T\iota(\mathcal{V}[\mathcal{R}_q]) = T\iota(\mathcal{N}_q) \oplus T\iota(\mathcal{H})$ for some complement $\mathcal{H}$ as noted in Proposition 3.1.6. The next proposition and its proof describe this construction of bases for $T\iota(\mathcal{N}_q)$ and $T\iota(\mathcal{H})$ and thus of a basis for $T\iota(\mathcal{V}[\mathcal{R}_q])$. Before we formulate it, we give a technical lemma concerning distributions of vector fields which form what is called *Jacobian systems*.

Lemma 3.1.7. *Let for the sake of formulating this lemma n and J be arbitrary natural numbers, and let for $1 \leq i \leq n$ and $1 \leq j \leq J$ an n-distribution of vector fields be spanned by the fields $X_i := \partial_{x^i} + \sum_{j=n+1}^{J} c^j \partial_{x^j}$. (The vector fields X_i are then in* triangular form *and form a* Jacobian system.*) Then the distribution spanned by the X_i is involutive if, and only if, for all $1 \leq h, i \leq n$ the Lie-brackets $[X_h, X_i]$ vanish.*

Proof. The proof is a straightforward calculation. To keep it simple, let—without loss of generality—the distribution be spanned by the vector fields

$$X_1 = \partial_x + f\partial_z \quad \text{and} \quad X_2 = \partial_y + g\partial_z \ .$$

Then their Lie-bracket $[X_1, X_2]$ is

$$\begin{aligned}
[\partial_x + f\partial_z, \partial_y + g\partial_z] &= [\partial_x, \partial_y] + [\partial_x, g\partial_z] + [f\partial_z, \partial_y] + [f\partial_z, g\partial_z] \\
&= \partial_x(g)\partial_z + g[\partial_x, \partial_z] - \partial_y(f)\partial_z - f[\partial_y, \partial_z] + f\partial_z(g)\partial_z + g[f\partial_z, \partial_z] \\
&= \{\partial_x(g) - \partial_y(f) + f\partial_z(g) - g\partial_z(f)\}\partial_z \ .
\end{aligned}$$

This Lie-bracket is a linear combination of X_1 and X_2 (namely the trivial one) if, and only if, it vanishes. □

We now hit upon such distributions when proving the aforementioned proposition.

Proposition 3.1.8. *Let the differential equation $\mathcal{R}_q$ be locally represented by $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$ where $1 \leq \tau \leq t$. Let $\dim V\pi^q_{q-1} - \dim T\iota(\mathcal{N}_q) =: \bar{r}$. Then there are coefficient functions $\zeta^\alpha_\mu, \xi^a_i \in \mathcal{F}(\mathcal{R}_q)$ and vector fields $W_a \in V\pi^q_{q-1} \setminus T\iota(\mathcal{N}_q)$ such that for $1 \leq i \leq n$ and*

$1 \leq k \leq r = \dim \mathcal{N}_q$ *the vector fields*

$$X_i := C_i^{(q)} + \sum_{a=1}^{\bar{r}} \xi_i^a W_a \quad \textit{and} \tag{3.3}$$

$$Y_k := \sum_{\substack{\alpha=1 \\ |\mu|=q}}^{m} \zeta_{(\alpha,\mu)}^k C_\alpha^\mu \tag{3.4}$$

form a basis for $T\iota\big(\mathcal{V}[\mathcal{R}_q]\big)$ *and additionally* $[W_a, W_b] = 0$ *and* $[Y_k, Y_l] = 0$ *for all* $1 \leq k, l \leq r$ *and* $1 \leq a, b \leq \bar{r}$.

Proof. We use a splitting $V\pi_{q-1}^q = T\iota(\mathcal{N}_q) \oplus \mathcal{W}$ with some complement $\mathcal{W}$ and construct appropriate bases for the symbol $T\iota(\mathcal{N}_q)$ and for $\mathcal{W}$. Then we construct a basis for the complement $T\iota(\mathcal{H})$ from the basis for $\mathcal{W}$ and the contact vector fields which are transversal to the fibration over $\mathcal{X}$.

In Equation (3.2), we have $1 \leq \alpha \leq m$ and $|\mu| = q$. For the coefficient matrix $(C_\alpha^\mu(\Phi^\tau))$, the pair (α, μ) indicates the column and $1 \leq \tau \leq t$ the row. The solution space is the symbol $T\iota(\mathcal{N}_q)$; let its dimension be r. Let the vectors $(b_{(\alpha,\mu)}^k)$ for $1 \leq k \leq r$ be a basis for it. Then transform the matrix with these r vectors as rows into reduced row echelon form. Let $(\zeta_{(\alpha,\mu)}^k)$ be the r rows in this new matrix. Then the vectors Y_k, defined by Equation (3.4), still form a basis for $T\iota(\mathcal{N}_q)$. We have $[Y_k, Y_l] = 0$ because the vector fields Y_k span the symbol, which by definition is the intersection of two involutive distributions and therefore itself involutive, and because according to their construction the vector fields Y_k form a Jacobian system.

To construct a basis for a complement $\mathcal{W}$, consider the matrix with the vectors $(\zeta_{(\alpha,\mu)}^k)$ as its rows. Any basis vector of the symbol, as we have constructed that basis, has coordinates $(\zeta_{(\alpha,\mu)}^k)$ with respect to the vector fields C_α^μ and therefore contains as a component some C_α^μ where (α, μ) is the column index of a pivot (which equals 1) in that matrix. It follows that those contact fields C_β^ν where (β, ν) is not the column index of any pivot in that matrix span $V\pi_{q-1}^q \setminus T\iota(\mathcal{N}_q)$. Since $C_\beta^\nu = \partial_{u_\nu^\beta}$, they are linearly independent, too. We number these vector fields from 1 to $\dim V\pi_{q-1}^q \setminus T\iota(\mathcal{N}_q)$ and set them equal to W_a according to their numbers a. The Lie bracket of any two such vector fields vanishes because obviously $[\partial_{u_\nu^\beta}, \partial_{u_\mu^\alpha}] = 0$.

To construct the vector fields X_i which are to span a complement $T\iota(\mathcal{H})$ to $T\iota(\mathcal{N}_q)$ in $T\iota(\mathcal{V}[\mathcal{R}_q])$, we use that to be tangent to $\mathcal{R}_q$, they have to satisfy

$$X_i(\Phi^\tau) = C_i^{(q)}(\Phi^\tau) + \sum_{a=1}^{\bar{r}} \xi_i^a W_a(\Phi^\tau) = 0$$

for all $1 \leq i \leq n$. This condition is equivalent to the solvability of the system of t linear equations

$$\sum_{a=1}^{\bar{r}} \xi_i^a W_a(\Phi^\tau) = -C_i^{(q)}(\Phi^\tau)$$

for the $\bar{r}$ unknown coefficient functions ξ_i^a. Any solution yields a basis of the form (3.3). $\square$

Remark 3.1.9. Prolonging a differential equation $\mathcal{R}_q$ and determining its Vessiot distribution $V[\mathcal{R}_q]$ are related operations, as the similarity of the necessary computations suggests. According to the definition of the total derivative (2.8), the prolongation of $\mathcal{R}_q$ adds for $1 \leq i \leq n$ the following equations for $\mathcal{R}_{q+1}$ to the local representation $\Phi^\tau = 0$:

$$\begin{aligned} D_i\Phi^\tau &= \frac{\partial \Phi^\tau}{\partial x^i} + \sum_{0\leq|\mu|\leq q} \sum_{\alpha=1}^{m} \frac{\partial \Phi^\tau}{\partial u^\alpha_\mu} u^\alpha_{\mu+1_i} \\ &= C_i^{(q)}(\Phi^\tau) + C^\mu_\alpha(\Phi^\tau) u^\alpha_{\mu+1_i} = 0 \ . \end{aligned}$$

These are the Equations (2.12) for $\mu = i$ from the computation of formal power series solutions where they are considered as an inhomogeneous system of linear equations for the Taylor coefficients of order $q + 1$ in the formal power series ansatz.

They resemble Equations (3.1) for the determination of the Vessiot distribution because the Vessiot distribution contains the symbol and the symbol equations are used in the construction of formal solutions order by order as remarked in Example 2.3.17.

In a manner congruent with Arnold's [3] classification of the points of an ordinary differential equation $\mathcal{R}_q$ into three kinds, the symbol matrix $(M_1)_\rho$ characterizes the class of $\rho \in \mathcal{R}_1$ as is demonstrated in Seiler [37]: if, and only if, the symbol vanishes (that is, if the matrix $(M_1)_\rho$ has full rank m), the Vessiot distribution is transversal and has dimension $n = 1$ at ρ (such a point $\rho \in \mathcal{R}_q$ is called *regular*). If, and only if, $(\operatorname{rank} M_1)_\rho < m$ but the augmented matrix $(M_1|\frac{\partial\Phi^\tau}{\partial x} + \frac{\partial\Phi^\tau}{\partial \mathbf{u}}\mathbf{u}_x)_\rho$ has full rank m, the Vessiot distribution is not transversal any more at ρ but still has dimension $n = 1$ (such a point is called *regular-singular*). And if, and only if, the augmented matrix at ρ does not have full rank, the Vessiot distribution there has a dimension greater than one (such a point is called *irregular-singular*).

This property of calculating the Vessiot distribution helps in the analysis of singular solutions; these are solutions that escape the representation for a general solution of a differential equation. They may appear as envelopes of general solutions even for ordinary differential equations (see Seiler [38], Section 4.1, for an example).

For an ordinary differential equation, an integral curve of its Vessiot distribution is called a *generalized solution*. That is, a generalized solution is a curve in $J_q\pi$, not in $\mathcal{E}$. A generalized solution may stem from a solution as it is introduced in Definition 2.2.5; then this curve projects onto the image of a section $\sigma : \mathcal{X} \to \mathcal{E}$, and the generalized solution is just the prolongation of a this solution in the usual sense. Whenever the Vessiot distribution is transversal along the generalized solution, this is the case. Otherwise, the projection of a generalized solution to $\mathcal{E}$ may be more intricate.

With regard to this relation between them, Equations (3.1) may be interpreted as a projective version of the Equations

$$D_i\Phi^\tau \ = \ C_i^{(q)}(\Phi^\tau) + C^\mu_\alpha(\Phi^\tau) u^\alpha_{\mu+1_i} \ = \ 0 \ . \tag{3.5}$$

From this standpoint, prolongation is the affine calculation which is comparable to determining the Vessiot distribution as the projective calculation.

Now what helps in the analysis of singular solutions is the fact that for $n = 1$, when we consider ordinary differential equations, this comparison is exactly the classical relation between the affine and the projective formulations. For a more thorough treatment, see Seiler [37].

In general, $\mathcal{V}[\mathcal{R}_q]$ is not involutive (an exception are formally integrable equations of finite type since the symbol of such a system vanishes), but it may contain involutive subdistributions; among these, those of dimension n which are transversal (with respect to the fibration π^q) are of special interest for us—if any exist at all.

Example 3.1.10. We take up Example 3.1.2 again. The Lie-brackets of the vector fields which span the Vessiot distribution are

$$\begin{aligned}
[X_1, X_2] &= \ v_x\partial_u - w_t\partial_u &&= (v_x - w_t)\partial_u = (u_{tx} - u_{tx})\partial_u = 0,\\
[X_1, X_3] &= \ -\partial_v - \partial_{u_t}, &&\\
[X_1, X_4] &= \ -\partial_w - \partial_{u_x}, &&\\
[X_2, X_3] &= \ -\partial_w - \partial_{u_x} &&= [X_1, X_4],\\
[X_2, X_4] &= \ -\partial_v - \partial_{u_t} &&= [X_1, X_3],\\
[X_3, X_4] &= \ 0. &&
\end{aligned}$$

Note that the bracket $[X_1, X_2]$ produces the integrability condition $w_t - v_x = 0$. The vector fields $[X_1, X_3]$ and $[X_1, X_4]$ do not belong to the Vessiot distribution, which means, $\mathcal{V}[\mathcal{R}_1]$ is not involutive. But the subdistribution spanned by X_1 and X_2 is. Its dimension is $2 = n$, and it is transversal with respect to π^q.

Setting $X_3 =: Y_1$ and $X_4 =: Y_2$ we have $Y_1 = \partial_{v_x} + \partial_{w_t}$ and $Y_2 = \partial_{v_t} + \partial_{w_x}$, which are of the form shown in Equation (3.4): here $\zeta^1_{(v,x)} = \zeta^1_{(w,t)} = \zeta^2_{(v,t)} = \zeta^2_{(w,x)} = 1$. The complement to $T\iota(\mathcal{N}_1)$ in $V\pi^1_0$ is spanned by $W_1 = \partial_{u_t}$, $W_2 = \partial_{u_x}$, $W_3 = \partial_{v_x}$ and $W_4 = \partial_{w_x}$. As

$$\begin{aligned}
X_1 &= \partial_x + u_x\partial_u + v_x\partial_v + w_x\partial_w + v_x\partial_{u_t} + w_x\partial_{u_x},\\
X_2 &= \partial_t + u_t\partial_u + v_t\partial_v + w_t\partial_w + v_t\partial_{u_t} + w_t\partial_{u_x},
\end{aligned}$$

here we have $\xi^1_1 = v_x$, $\xi^2_1 = w_x$ and $\xi^3_1 = \xi^4_1 = 0$, and $\xi^1_2 = v_t$, $\xi^2_2 = w_t$ and $\xi^3_2 = \xi^4_2 = 0$, and the vector fields are of the form given in Equation (3.3). Obviously $[Y_1, Y_2] = 0$ and $[W_a, W_b] = 0$ for all $1 \leq a, b \leq 4$.

3.1.2 Vessiot Connections

The reason why n-dimensional involutive subdistribution of the Vessiot distribution which are transversal with respect to π^q are interesting is that they can be regarded as linearizations of prolonged solutions to the differential equation and conversely.

Proposition 3.1.11. *If the local section $\sigma \in \Gamma_L\pi$ is a local solution of the equation $\mathcal{R}_q$, then the tangent bundle $T(\operatorname{im} j_q\sigma)$ is an n-dimensional transversal involutive subdistribution of $\mathcal{V}[\mathcal{R}_q]|_{\operatorname{im} j_q\sigma}$. Conversely, if $\mathcal{U} \subseteq \mathcal{V}[\mathcal{R}_q]$ is an n-dimensional transversal involutive subdistribution, then any integral manifold of $\mathcal{U}$ has locally the form $\operatorname{im} j_q\sigma$ where $\sigma \in \Gamma_L\pi$ is a local solution of $\mathcal{R}_q$.*

Proof. Let $\sigma \in \Gamma_L\pi$ be a local solution of the equation $\mathcal{R}_q$. Then it satisfies by Definition 2.2.5 $\operatorname{im} j_q\sigma \subseteq \mathcal{R}_q$ and thus $T(\operatorname{im} j_q\sigma) \subseteq T\mathcal{R}_q$. Besides, by Definition 2.1.8 of the contact distribution, for $\mathbf{x} \in \mathcal{X}$ with $j_q\sigma(\mathbf{x}) = \rho \in J_q\pi$, the tangent space $T_\rho(\operatorname{im} j_q\sigma)$ is a subspace of $\mathcal{C}_q|_\rho$. By definition of the Vessiot distribution, it follows $T_\rho(\operatorname{im} j_q\sigma) \subseteq T\iota(T_\rho\mathcal{R}_q) \cap \mathcal{C}_q|_\rho$., which proves the first claim.

Now let $\mathcal{U} \subseteq \mathcal{V}[\mathcal{R}_q]$ be an n-dimensional transversal involutive subdistribution. Then according to the Frobenius theorem $\mathcal{U}$ has n-dimensional integral manifolds. By definition, $T\iota(\mathcal{V}[\mathcal{R}_q]) \subseteq \mathcal{C}_q|_{\mathcal{R}_q}$; thus according to Proposition 2.1.9 for any integral manifold of $\mathcal{U}$ there is a local section σ such that the integral manifold is of the form $\operatorname{im} j_q\sigma$. Furthermore, the integral manifold is a subset of $\mathcal{R}_q$. Thus it corresponds to a local solution of $\mathcal{R}_q$. □

Remark 3.1.12. As we noted in Remark 2.3.15, a maximally over-determined differential equation of order q is equivalent to a connection on $J_{q-1}\pi$. The images of its prolonged solutions are the integral manifolds of the corresponding horizontal bundle. If $\mathcal{R}_q$ is not maximally over-determined, it does not define one connection; but now any n-dimensional transversal (with respect to the fibration $\hat{\pi}^q : \mathcal{R}_q \to \mathcal{X}$) involutive subdistribution of $\mathcal{V}[\mathcal{R}_q]$ represents the horizontal bundle of a flat connection for $\hat{\pi}^q$. In this sense, $\mathcal{R}_q$ is covered by infinitely many flat connections. For each of these, the Frobenius theorem then guarantees that there is an n-dimensional integral manifold; by definition of $\mathcal{V}[\mathcal{R}_q]$, they are of the form $\operatorname{im} j_q\sigma$ for a solution σ of $\mathcal{R}_q$.

This observation is the basis of Vessiot's approach to the analysis of $\mathcal{R}_q$: he proposed to search for all n-dimensional, transversal involutive subdistributions of $\mathcal{V}[\mathcal{R}_q]$. Before we do this, we show how integral elements appear in this program.

Proposition 3.1.13. *Let $\mathcal{U} \subseteq \mathcal{V}[\mathcal{R}_q]$ be a transversal subdistribution of the Vessiot distribution of constant rank k. Then the spaces $\mathcal{U}_\rho$ are k-dimensional integral elements for all points $\rho \in \mathcal{R}_q$ if, and only if, $[\mathcal{U},\mathcal{U}] \subseteq \mathcal{V}[\mathcal{R}_q]$.*

Proof. Let $\{\omega_1, \dots, \omega_r\}$ be a basis of the codistribution $\iota^*\mathcal{C}_q^0$. Then an algebraic basis of the ideal $\mathcal{I}[\mathcal{R}_q]$ is $\{\omega_1, \dots, \omega_r, d\omega_1, \dots, d\omega_r\}$. Any vector field $X \in \mathcal{U}$ trivially satisfies $\omega_i(X) = 0$ by Proposition 3.1.3. For arbitrary fields $X_1, X_2 \in \mathcal{U}$, we have $d\omega_i(X_1, X_2) = X_1\big(\omega_i(X_2)\big) - X_2\big(\omega_i(X_1)\big) + \omega_i\big([X_1, X_2]\big)$. The first two summands on the right side vanish trivially and the remaining equation implies our claim. □

Definition 3.1.14. *A subdistribution $\mathcal{U} \subseteq \mathcal{V}[\mathcal{R}_q]$ satisfying the conditions of Proposition 3.1.13 is called an* integral distribution *for the differential equation $\mathcal{R}_q$.*

In the literature (Fackerell [15], Stormark [40], Vassiliou [42]) the term "involution" is common for such a distribution which, however, may be confusing since it may *not* be involutive. For the same reason an integral distribution may be not integrable; the name only reflects the fact that it is composed of integral elements.

The construction of integral distributions is the first step in our program. The second step is the construction of involutive subdistributions within them.

Since the symbol $\mathcal{N}_q$ of the equation $\mathcal{R}_q$ is contained in the Vessiot distribution, we can split the Vessiot distribution into $\mathcal{V}[\mathcal{R}_q] = \mathcal{N}_q \oplus \mathcal{H}$ where $\mathcal{H}$ is some complement, according to Proposition 3.1.6. The decomposition of the full contact distribution into $V\pi^q_{q-1}$, spanned by the vector fields given in Equation (2.4b), and a complement defining a connection on the fibration $\pi^q\colon J_q\pi \to \mathcal{X}$, for example the complement spanned by the vector fields given in Equation (2.4a), admits an analogy, which leads naturally to connections on the fibration $\hat{\pi}^q\colon \mathcal{R}_q \to \mathcal{X}$: Provided $\dim \mathcal{H} = n$, the complement $\mathcal{H}$ to the symbol $\mathcal{N}_q$ may be considered as the horizontal bundle of a connection of the fibred manifold $\mathcal{R}_q \to \mathcal{X}$.

Definition 3.1.15. *We call any such connection a* Vessiot connection *for $\mathcal{R}_q$.*

The involutive n-dimensional subdistributions within $\mathcal{V}[\mathcal{R}_q]$ which interest us are exactly such complements $\mathcal{H}$. Since they have to be closed under Lie-brackets, our aim is to construct all *flat* Vessiot connections. The existence of n-dimensional complements (which are not necessarily involutive) follows from the absence of integrability conditions. Indeed, this assumption more than suffices; the following proposition characterizes the existence of an n-dimensional complement.

Proposition 3.1.16. *Let $\mathcal{R}_q$ be a differential equation. Then its Vessiot distribution possesses locally a decomposition with an n-dimensional complement $\mathcal{H}$ such that*

$$\mathcal{V}[\mathcal{R}_q] = \mathcal{N}_q \oplus \mathcal{H}$$

if, and only if, there are no integrability conditions which arise as the prolongations of equations of lower order in the system.

Proof. Let denote $\mathbf{u}_{(q)}$ the set of all derivatives of order q only, and let $\mathcal{R}_q$ be locally represented by

$$\mathcal{R}_q\colon \begin{cases} \Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0\ , \\ \Psi^\sigma(\mathbf{x}, \mathbf{u}^{(q-1)}) = 0\ , \end{cases}$$

such that the equations $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$ do not imply lower order equations which are independent of the equations $\Psi^\sigma(\mathbf{x}, \mathbf{u}^{(q-1)}) = 0$. Then the Jacobi matrix $(\partial\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)})/\partial\mathbf{u}_{(q)})$ has maximal rank.

The ansatz (3.1) for the determination of the Vessiot distribution yields for the above representation:

$$C^{(q)}_i(\Phi^\tau)a^i + C^\mu_\alpha(\Phi^\tau)b^\alpha_\mu = 0 \tag{3.6}$$

$$C^{(q)}_i(\Psi^\sigma)a^i = 0. \tag{3.7}$$

Here, the matrix $C^\mu_\alpha(\Phi^\tau)$ has maximal rank, too; therefore the equations $C^{(q)}_i(\Phi^\tau)a^i + C^\mu_\alpha(\Phi^\tau)b^\alpha_\mu = 0$ can be solved for a subset of the unknowns b^α_μ. And since no terms of order q are present in $\Psi^\sigma(\mathbf{x}, \mathbf{u}^{(q-1)}) = 0$, we have $C^{(q)}_i(\Psi^\sigma) = D_i\Psi^\sigma$. Because the Vessiot distribution is tangential to $\mathcal{R}_q$, we have $D_i\Psi^\sigma = 0$ on $\mathcal{R}_q$. It follows that $C^{(q)}_i(\Psi^\sigma)a^i = 0$ vanishes if, and only if, no integrability conditions arise from the prolongation of lower order equations. And if, and only if, this is the case, then (3.6) has for each $1 \leq j \leq n$ a solution where $a^j = 1$ while all other a^i are zero. The existence of such a solution is equivalent to the existence of an n-dimensional complement $\mathcal{H}$. □

From this proposition follows that for the determination of the Vessiot distribution $\mathcal{V}[\mathcal{R}_q]$, equations of order less than q in the local representation of $\mathcal{R}_q$ can be ignored if there are no integrability conditions which arise from equations of lower order.

This result softens the assumptions used in Fesser and Seiler [17], Proposition 3.5, where the following sufficient condition for the existence of an n-dimensional complement is shown.

Proposition 3.1.17. *If the differential equation $\mathcal{R}_q$ satisfies $\mathcal{R}_q^{(1)} = \mathcal{R}_q$, then its Vessiot distribution possesses locally a decomposition $\mathcal{V}[\mathcal{R}_q] = \mathcal{N}_q \oplus \mathcal{H}$ with an n-dimensional complement $\mathcal{H}$.*

Proof. The assumption $\mathcal{R}_q = \mathcal{R}_q^{(1)}$ implies that to every point $\rho \in \mathcal{R}_q$ at least one point $\hat{\rho} \in \mathcal{R}_{q+1}$ with $\pi_q^{q+1}(\hat{\rho}) = \rho$ exists. We choose such a $\hat{\rho}$ and consider $\operatorname{im}\Gamma_{q+1}(\hat{\rho}) \subset T_\rho(J_q\pi)$. By definition of the contact map Γ_{q+1}, this is an n-dimensional transversal subset of $\mathcal{C}_q|_\rho$. Thus there only remains to show that it is also tangent to $\mathcal{R}_q$, as then we can define a complement by $T\iota(\mathcal{H}_\rho) = \operatorname{im}\Gamma_{q+1}(\hat{\rho})$. But this tangency is a consequence of $\hat{\rho} \in \mathcal{R}_{q+1}$; using for example the local coordinates expression (2.3) for Γ_q and a local representation $\Phi^\tau = 0$ of $\mathcal{R}_q$, it follows that the vector $v_i = \Gamma_{q+1}(\hat{\rho}, \partial_{x^i}) \in T_\rho(J_q\pi)$ satisfies $d\Phi^\tau|_\rho(v_i) = D_i\Phi^\tau(\hat{\rho}) = 0$ and thus is tangential to $\mathcal{R}_q$.

Hence we have proven that it is possible to construct for each point $\rho \in \mathcal{R}_q$ a complement $\mathcal{H}_\rho$ such that $\mathcal{V}_\rho[\mathcal{R}_q] = (\mathcal{N}_q)_\rho \oplus \mathcal{H}_\rho$. Now we must show that these complements can be chosen in such a way that they form a distribution (which by definition is smooth). Our assumption $\mathcal{R}_q = \mathcal{R}_q^{(1)}$ implies that the restricted projection $\hat{\pi}_q^{q+1} : \mathcal{R}_{q+1} \to \mathcal{R}_q$ is a surjective submersion, which means it defines a fibred manifold. Thus if we choose a local section $\gamma : \mathcal{R}_q \to \mathcal{R}_{q+1}$ and then always take $\hat{\rho} = \gamma(\rho)$, it follows immediately that the corresponding complements $\mathcal{H}_\rho$ define a smooth distribution as required. □

Compared to the necessary and sufficient condition for the existence of suitable complements $\mathcal{H}$ given in Proposition 3.1.16, Proposition 3.1.17 only gives a sufficient condition; on the other hand its proof is based on a geometrical argument while the one for Proposition 3.1.16 does not refer to the contact map.

As we see, in local coordinates, it is the integrability conditions which arise as prolongations of lower order equations which hinder the construction of n-dimensional complements, while those which follow from the relations between cross derivatives do not influence this approach.

Example 3.1.18. Consider the differential equation $\mathcal{R}_1$ in Example 2.2.11. It is locally represented by the same equations like $\mathcal{R}_1^{(1)}$ in Example 2.2.13, except that the integrability condition $w_t = v_x$ is missing. The matrix of $T\iota$ for the system $\mathcal{R}_1$ has eleven rows and eight columns—one additional column as compared to the matrix for the system $\mathcal{R}_1^{(1)}$. The symbol $T\iota(\mathcal{N}_1)$ of $\mathcal{R}_1$ is 3-dimensional, spanned by ∂_{v_x}, ∂_{w_t} and $\partial_{w_x} + \partial_{v_t}$, while the symbol $T\iota(\mathcal{N}_1^{(1)})$ of $\mathcal{R}_1^{(1)}$ has dimension 2 and is spanned by $\partial_{v_x} + \partial_{w_t}$ and $\partial_{w_x} + \partial_{v_t}$. But the one-forms ω^u, ω^v and ω^w (and their pull-backs $\iota^*\omega^u = d\overline{u} - \overline{u_x}d\overline{x} - \overline{u_t}d\overline{t}$, $\iota^*\omega^v = d\overline{v} - \overline{v_x}d\overline{x} - \overline{v_t}d\overline{t}$ and $\iota^*\omega^w = d\overline{w} - \overline{w_x}d\overline{x} - \overline{w_t}d\overline{t}$) are the same, and therefore the coordinate expressions for the vector fields X_1 and X_2 (and their representations $\bar{X}_1 = \partial_{\overline{x}} + \overline{u_x}\partial_{\overline{u}} + \overline{v_x}\partial_{\overline{v}} + \overline{w_x}\partial_{\overline{w}}$

and $\bar{X}_2 = \partial_{\bar{t}} + \overline{u_t}\partial_{\bar{u}} + \overline{v_t}\partial_{\bar{v}} + \overline{w_t}\partial_{\bar{w}}$ in $T\mathcal{R}_1$ and $T\mathcal{R}_1^{(1)}$) look alike (see Examples 3.1.2 and 3.1.4 for their representations). The integrability condition $w_t = v_x$ does not influence the results as it stems from the equality of the cross derivatives, $u_{tx} = v_x$ and $u_{xt} = w_t$, not from the prolongation of a lower order equation.

Now consider the differential equation which is locally represented by

$$\mathcal{R}_1 : \begin{cases} u_t = v \ , \quad v_t = w_x \ , \quad w_t = v_x \ , \\ u_x = w \ , \\ u = x \ . \end{cases}$$

It arises from the system in Example 2.2.13 by adding the algebraic equation $u = x$. The matrix for $T\iota$ has eleven rows and six columns, and for the pull-backs of the contact forms we have

$$\begin{aligned} \iota^*\omega^u &= d\bar{u} - \overline{u_x}d\bar{x} - \overline{u_t}d\bar{t} = d\bar{x} - \bar{w}d\bar{x} - \bar{v}d\bar{t} \ , \\ \iota^*\omega^v &= d\bar{v} - \overline{v_x}d\bar{x} - \overline{v_t}d\bar{t} = d\bar{v} - \overline{v_x}d\bar{x} - \overline{w_x}d\bar{t} \ , \\ \iota^*\omega^w &= d\bar{w} - \overline{w_x}d\bar{x} - \overline{w_t}d\bar{t} = d\bar{w} - \overline{w_x}d\bar{x} - \overline{v_x}d\bar{t} \ . \end{aligned}$$

Solving the corresponding system of linear equations, we arrive at the Vessiot distribution

$$\mathcal{V}[\mathcal{R}_1] = \operatorname{span}\{\partial_{\overline{v_x}}, \partial_{\overline{w_x}}, \bar{v}\partial_{\bar{x}} + (1-\bar{w})\partial_{\bar{t}} + (\overline{w_x}(1-\bar{w}) + \bar{v}\overline{v_x})\partial_{\bar{v}} + (\overline{v_x}(1-\bar{w}) + \bar{v}\overline{w_x})\partial_{\bar{w}}\} \ .$$

The vector fields $\partial_{\overline{v_x}}$ and $\partial_{\overline{w_x}}$ span the symbol $\mathcal{N}_1$; any of its complements in $\mathcal{V}[\mathcal{R}_1]$ is one-dimensional—as the dimension of the base space is two, none of them can be an n-dimensional transversal subdistribution in $\mathcal{V}[\mathcal{R}_1]$, and none of them can be the horizontal space of a connection.

The reason for this is that $\mathcal{R}_1$ is not formally integrable, as the prolongation of the algebraic equation $u = x$ leads to the first order equation $u_t = 0$. Projecting the prolonged system gives:

$$\mathcal{R}_1^{(1)} : \begin{cases} u_t = v \ = \ v_t = w_x \ = \ w_t = v_x \ = \ 0, \\ u_x = w \ = \ 1 \ , \\ u = x \ . \end{cases}$$

Now the symbol vanishes, and so do the pull-backs of the contact forms: $\iota^*\omega^u = d\bar{x} - d\bar{x} = 0$, $\iota^*\omega^v = d\bar{v} = 0$, $\iota^*\omega^w = d\bar{w} = 0$. Therefore we find $\mathcal{V}[\mathcal{R}_1] = \mathcal{N}_1 \oplus \mathcal{H} = \{0\} \oplus \operatorname{span}\{\partial_{\bar{x}}, \partial_{\bar{t}}\}$. As the Lie brackets of $\partial_{\bar{x}}$ and $\partial_{\bar{t}}$ trivially vanish, $T\mathcal{R}_1 = \mathcal{V}[\mathcal{R}_1] = \mathcal{H} = \operatorname{span}\{\partial_{\bar{x}}, \partial_{\bar{t}}\}$ is a two-dimensional involutive distribution.

In Proposition 3.1.8 and its proof we described the construction of an n-dimensional complement, $\mathcal{H}$, for the symbol $T\iota(\mathcal{N}_q)$, if the differential equation is represented by a system $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$ which may be fully implicit. But if the local representation is given in solved form, the approach can be simplified. The construction of a complement $\mathcal{H}$ is easily possible for the class of differential equations which are represented in the following special form. For such differential equations, the representation of $\mathcal{R}_q$ in local coordinates suggests a certain choice of a basis for $\mathcal{H}$.

Proposition 3.1.19. *For the differential equation $\mathcal{R}_q$, let be $u^\alpha_\mu = \phi^\alpha_\mu(\mathbf{x}, \mathbf{u}, \hat{\mathbf{u}}^{(q)})$ a representation where each equation is solved for a different principal derivative u^α_μ with $|\mu| = q$, and where $\hat{\mathbf{u}}^{(q)}$ denotes the set of the remaining, thus parametric, derivatives of order q. (There are no equations of order less than q; in particular, there are no algebraic conditions.) Let $\mathcal{B}$ denote the set of pairs (α, μ) where u^α_μ is a principal derivative. Then the vector fields $\bar{X}_i = \partial_{\overline{x^i}} + \sum_{\alpha=1}^m \sum_{0\leq|\mu|<q} \overline{u^\alpha_{\mu+1_i}} \partial_{\overline{u^\alpha_\mu}}$ satisfy*

$$\iota_* \bar{X}_i = C^{(q)}_i + \sum_{(\alpha,\mu)\in\mathcal{B}} C^{(q)}_i(\phi^\alpha_\mu) C^\mu_\alpha \ , \quad 1 \leq i \leq n, \tag{3.8}$$

and generate a complement $\mathcal{H}_0$ for the symbol $\mathcal{N}_q$ such that $\mathcal{V}[\mathcal{R}_q] = \mathcal{N}_q \oplus \mathcal{H}_0$.

Proof. According to Proposition 3.1.3, we have $\mathcal{V}[\mathcal{R}_q] = (\iota^* \mathcal{C}^0_q)^0$. The contact codistribution $\mathcal{C}^0_q$ is spanned by the one-forms

$$\omega^\alpha_\mu = du^\alpha_\mu - \sum_{i=1}^n u^\alpha_{\mu+1_i} dx^i \ , \quad 0 \leq |\mu| < q \ , 1 \leq \alpha \leq m \ ,$$

as given in Equation (2.6). The pullbacks of these one-forms are, for all $1 \leq \alpha \leq m$,

$$\begin{aligned} \iota^*\omega^\alpha_\mu &= d\overline{u^\alpha_\mu} - \sum_{i=1}^n \overline{u^\alpha_{\mu+1_i}} d\overline{x^i} \ , && 0 \leq |\mu| < q-1 \ , \\ \iota^*\omega^\alpha_\mu &= d\overline{u^\alpha_\mu} - \sum_{(\alpha,\mu+1_i)\in\mathcal{B}} \phi^\alpha_{\mu+1_i} d\overline{x^i} - \sum_{(\alpha,\mu+1_i)\notin\mathcal{B}} \overline{u^\alpha_{\mu+1_i}} d\overline{x^i} \ , && |\mu| = q-1 \ . \end{aligned}$$

The space which is annihilated in $T\mathcal{R}_q$ by these one-forms is spanned by two sorts of vector fields:

$$\partial_{\overline{u^\alpha_\mu}} \ , \qquad (\alpha, \mu) \notin \mathcal{B} \ , \quad 1 \leq \alpha \leq m \ , \quad |\mu| = q \ ,$$

which are vertical, and, for all $1 \leq i \leq n$,

$$\bar{X}_i = \partial_{\overline{x^i}} + \sum_{(\alpha,\mu+1_i)\in\mathcal{B}} \phi^\alpha_{\mu+1_i} \partial \overline{u^\alpha_\mu} + \sum_{(\alpha,\mu+1_i)\notin\mathcal{B}} \overline{u^\alpha_{\mu+1_i}} \partial \overline{u^\alpha_\mu} \ , \quad 0 \leq |\mu| \leq q-1 \ ,$$

which are transversal. Therefore these two sorts of vector fields span the Vessiot distribution $(\iota^*\mathcal{C}^0_q)^0$. The push-forwards of the vertical vector fields are, for all $(\alpha, \mu) \notin \mathcal{B}$ where $1 \leq \alpha \leq m$ and $|\mu| = q$,

$$\iota_* \partial_{\overline{u^\alpha_\mu}} = \partial_{u^\alpha_\mu} + \sum_{(\beta,\nu)\in\mathcal{B}} \frac{\partial \phi^\beta_\nu}{\partial u^\alpha_\mu} \partial_{u^\beta_\nu} \ . \tag{3.9}$$

As we assume that the system is given in solved form, calculating these push-forwards is the essentially same as solving the symbol equations (2.14) because of the form of the Jacobian matrix for the inclusion map $\iota\colon \mathcal{R}_q \hookrightarrow J_q\pi$ in these local coordinates. Thus, the vector fields $\partial_{\overline{u^\alpha_\mu}}$ form a basis of the symbol $\mathcal{N}_q$, and the corresponding vector fields $\iota_*\partial_{\overline{u^\alpha_\mu}}$

a basis for the symbol $T\iota(\mathcal{N}_q)$. To calculate the push-forwards of the transversal fields $\bar{X}_i$, we first consider the respective push-forward of each type of their summands. For all $1 \leq i \leq n$, we have

$$\iota_*\partial_{\overline{x^i}} = \partial_{x^i} + \sum_{(\alpha,\mu)\in\mathcal{B}} \frac{\partial\phi^\alpha_\mu}{\partial x^i}\partial_{u^\alpha_\mu} \ ;$$

and for (α, μ), where $1 \leq \alpha \leq m$ and $0 \leq |\mu| \leq q-1$, the formula for the vector fields $\iota_*\partial_{\overline{u^\alpha_\mu}}$ looks like the one in Equation (3.9)—but mind that now the order of the multi-index μ is less than q. From the expressions for $\iota_*\partial_{\overline{x^i}}$ and $\iota_*\partial_{\overline{u^\alpha_\mu}}$, it follows that for all $1 \leq i \leq n$, we have

$$\begin{aligned}\iota_*\bar{X}_i = \partial_{x^i} + \sum_{(\alpha,\mu)\in\mathcal{B}} \frac{\partial\phi^\alpha_\mu}{\partial x^i}\partial_{u^\alpha_\mu} + \sum_{(\alpha,\mu+1_i)\in\mathcal{B}} \phi^\alpha_{\mu+1_i}(\partial_{u^\alpha_\mu} + \sum_{(\beta,\nu)\in\mathcal{B}} \frac{\partial\phi^\beta_\nu}{\partial u^\alpha_\mu}\partial_{u^\beta_\nu}) \\ + \sum_{(\alpha,\mu+1_i)\notin\mathcal{B}} u^\alpha_{\mu+1_i}(\partial_{u^\alpha_\mu} + \sum_{(\beta,\nu)\in\mathcal{B}} \frac{\partial\phi^\beta_\nu}{\partial u^\alpha_\mu}\partial_{u^\beta_\nu}) \ .\end{aligned}$$

What follows now is a straightforward calculation. Rearranging the summands makes

$$\begin{aligned}&\partial_{x^i} + \sum_{(\alpha,\mu+1_i)\in\mathcal{B}} \phi^\alpha_{\mu+1_i}\partial_{u^\alpha_\mu} + \sum_{(\alpha,\mu+1_i)\notin\mathcal{B}} u^\alpha_{\mu+1_i}\partial_{u^\alpha_\mu} + \sum_{(\alpha,\mu)\in\mathcal{B}} \frac{\partial\phi^\alpha_\mu}{\partial x^i}\partial_{u^\alpha_\mu} \\ &+ \sum_{(\beta,\nu)\in\mathcal{B}}\sum_{(\alpha,\mu+1_i)\in\mathcal{B}} \phi^\alpha_{\mu+1_i}\frac{\partial\phi^\beta_\nu}{\partial u^\alpha_\mu}\partial_{u^\beta_\nu} + \sum_{(\beta,\nu)\in\mathcal{B}}\sum_{(\alpha,\mu+1_i)\notin\mathcal{B}} u^\alpha_{\mu+1_i}\frac{\partial\phi^\beta_\nu}{\partial u^\alpha_\mu}\partial_{u^\beta_\nu} \ .\end{aligned}$$

We factor out in the second line. This yields

$$\begin{aligned}&\partial_{x^i} + \sum_{(\alpha,\mu+1_i)\in\mathcal{B}} \phi^\alpha_{\mu+1_i}\partial_{u^\alpha_\mu} + \sum_{(\alpha,\mu+1_i)\notin\mathcal{B}} u^\alpha_{\mu+1_i}\partial_{u^\alpha_\mu} + \sum_{(\alpha,\mu)\in\mathcal{B}} \partial_{x^i}(\phi^\alpha_\mu)\partial_{u^\alpha_\mu} \\ &+ \sum_{(\beta,\nu)\in\mathcal{B}}\left\{\sum_{(\alpha,\mu+1_i)\in\mathcal{B}} \phi^\alpha_{\mu+1_i}\partial_{u^\alpha_\mu}(\phi^\beta_\nu)\partial_{u^\beta_\nu} + \sum_{(\alpha,\mu+1_i)\notin\mathcal{B}} u^\alpha_{\mu+1_i}\partial_{u^\alpha_\mu}(\phi^\beta_\nu)\partial_{u^\beta_\nu}\right\} \ .\end{aligned}$$

Now we use that on the differential equation $\mathcal{R}_q$ we have $u^\alpha_\mu = \phi^\alpha_\mu$ for all $(\alpha, \mu) \in \mathcal{B}$. This enables us to collect parametric and principal derivatives with summation indices of the same kind into one sum for each kind. We arrive at

$$\begin{aligned}&\partial_{x^i} + \sum_{\alpha=1}^m\sum_{0\leq|\mu|<q} u^\alpha_{\mu+1_i}\partial_{u^\alpha_\mu} + \sum_{(\alpha,\mu)\in\mathcal{B}}\left\{\partial_{x^i} + \sum_{\beta=1}^m\sum_{0\leq|\beta|<q} u^\beta_{\nu+1_i}\partial_{u^\beta_\nu}\right\}(\phi^\alpha_\mu)\partial_{u^\alpha_\mu} \\ &= C^{(q)}_i + \sum_{(\alpha,\mu)\in\mathcal{B}} C^{(q)}_i(\phi^\alpha_\mu)C^\mu_\alpha \ .\end{aligned}$$

The last equality follows from Equation (2.4a), applied twice. □

When proving the existence theorems for integral distributions and flat Vessiot connections (Theorems 3.3.9 and 3.3.28), we reduce technical nuisances by transforming differential equations into equivalent systems of first order. For these, the results of Proposition 3.1.19 take the following form.

Remark 3.1.20. A differential equation $\mathcal{R}_1$ represented in the reduced Cartan normal form (2.27) satisfies the assumptions of Proposition 3.1.19. In this case, from Equation (3.8) follows that a reference complement $\mathcal{H}_0$ is spanned by the n vector fields

$$\bar{X}_i = \partial_{\overline{x^i}} + \sum_{(\alpha,i)\in\mathcal{B}} \phi_i^\alpha \partial_{\overline{u^\alpha}} + \sum_{(\alpha,i)\notin\mathcal{B}} \overline{u_i^\alpha} \partial_{\overline{u^\alpha}} \ , \tag{3.10}$$

for which the push-forwards are

$$\iota_* \bar{X}_i = C_i^{(1)} + \sum_{(\alpha,j)\in\mathcal{B}} C_i^{(1)}(\phi_j^\alpha) C_\alpha^j \ , \quad 1 \le i \le n. \tag{3.11}$$

3.1.3 Flat Vessiot Connections

Now we return to the general case where the differential equation $\mathcal{R}_q$ is represented in the form (2.7) and the equations of the system may be fully implicit. Any n-dimensional complement $\mathcal{H}$ is obviously a transversal subdistribution of $\mathcal{V}[\mathcal{R}_q]$, but not necessarily involutive. Conversely, any n-dimensional subdistribution $\mathcal{H}$ of $\mathcal{V}[\mathcal{R}_q]$ is a possible choice for a complement, and Vessiot's goal is the construction of all flat Vessiot connections. If we choose a reference complement $\mathcal{H}_0$ with a basis $(X_i\colon 1 \le i \le n)$, then a basis for any other complement $\mathcal{H}$ arises by adding some vertical fields to the vectors X_i. We follow this approach in Section 3.3. For the remainder of this section we turn our attention to the choice of a convenient basis of $\mathcal{V}[\mathcal{R}_q]$ that facilitates our computations.

As we already noticed in Example 3.1.10, the Vessiot distribution $\mathcal{V}[\mathcal{R}_q]$ may not be involutive. Therefore it is understandable that its structure equations are important. Therefore we begin our considerations by examining the structure equations of $\mathcal{V}[\mathcal{R}_q]$. For $\mathcal{V}[\mathcal{R}_q] = \mathcal{N}_q \oplus \mathcal{H}$, the complement $\mathcal{H}$ may be not involutive, causing the Vessiot distribution to be not involutive. Let

$$[\mathcal{V}[\mathcal{R}_q], \mathcal{V}[\mathcal{R}_q]] =: \mathcal{V}'[\mathcal{R}_q]$$

be the derived Vessiot distribution. Then, because $T\mathcal{R}_q$ is involutive and therefore $T\iota(T\mathcal{R}_q)$, too, we have $T\iota(\mathcal{V}'[\mathcal{R}_q]) \subseteq T\iota(T\mathcal{R}_q) \cap \mathcal{C}_q'|_{\mathcal{R}_q}$. Since the only non-vanishing Lie brackets of contact fields in $\mathcal{C}_q$ are

$$[C_\alpha^{\nu+1_i}, C_i^{(q)}] = \partial_{u_\nu^\alpha} \ , \qquad |\nu| = q-1 \ , \tag{3.12}$$

it follows that, in local coordinates, we may extend a basis of $T\iota(\mathcal{V}[\mathcal{R}_q])$ to a basis of the derived Vessiot distribution $T\iota(\mathcal{V}'[\mathcal{R}_q])$ by adding vector fields Z_c in $T\iota(T\mathcal{R}_q) \cap \mathcal{C}_q'|_{\mathcal{R}_q}$, $1 \le c \le C = \dim \mathcal{V}'[\mathcal{R}_q] - \dim \mathcal{V}[\mathcal{R}_q]$. Using Equation (3.12), for each c we may linearly combine the vector fields Z_c from the vector fields $\partial_{u_\nu^\alpha}$ where $|\nu| = q-1$, which means there are coefficient functions $\kappa_{c\nu}^\alpha \in \mathcal{F}(\mathcal{R}_q)$ such that

$$Z_c = \kappa_{c\nu}^\alpha \partial_{u_\nu^\alpha} \qquad |\nu| = q-1 \ ,$$

and the derived Vessiot distribution now is

$$T\iota(\mathcal{V}'[\mathcal{R}_q]) = T\iota(\mathcal{V}[\mathcal{R}_q]) \oplus \operatorname{span}\{Z_c\colon 1 \le c \le C\} \ . \tag{3.13}$$

The vector fields Z_c may span a proper subspace in $\operatorname{span}\{\partial_{u^\alpha_\nu} : 1 \le \alpha \le m, |\nu| = q-1\}$. (For a formally integrable system of finite type $\mathcal{V}[\mathcal{R}_q]$ is involutive and thus $C = 0$.) To analyze the construction of flat Vessiot connections, we have to examine the structure equations for vector fields in $T\iota(\mathcal{V}[\mathcal{R}_q])$. We exploit that they can be represented uniquely as linear combinations of vector fields from bases for $T\iota(\mathcal{N}_q)$ and $T\iota(\mathcal{H})$ and first consider the structure equations for these.

Lemma 3.1.21. *Let the vector fields $(\check{Y}_k : 1 \le k \le r)$ be a basis for the symbol $T\iota(\mathcal{N}_q)$, let $T\iota(\mathcal{V}[\mathcal{R}_q]) = T\iota(\mathcal{N}_q) \oplus T\iota(\mathcal{H})$ with some complement $T\iota(\mathcal{H})$, and let $(\check{X}_i : 1 \le i \le n)$ be a basis for $T\iota(\mathcal{H})$. Let $T\iota(\mathcal{V}'[\mathcal{R}_q])$ be given as in Equation (3.13). Then coefficient functions $\hat{A}^h_{ij}$, $\hat{B}^p_{ij}$, Θ^c_{ij}, $\tilde{A}^h_{ik}$, $\tilde{B}^p_{ik}$, Ξ^c_{ik}, $B^p_{kl} \in \mathcal{F}(\mathcal{R}_q)$ exist such that the structure equations for the generators of the Vessiot distribution are*

$$[\check{X}_i, \check{X}_j] = \hat{A}^h_{ij}\check{X}_h + \hat{B}^p_{ij}\check{Y}_p + \Theta^c_{ij}Z_c\ , \qquad 1 \le i,j \le n\ , \tag{3.14a}$$
$$[\check{X}_i, \check{Y}_k] = \tilde{A}^h_{ik}\check{X}_h + \tilde{B}^p_{ik}\check{Y}_p + \Xi^c_{ik}Z_c\ , \qquad 1 \le i \le n,\ 1 \le k \le r\ , \tag{3.14b}$$
$$[\check{Y}_k, \check{Y}_l] = B^p_{kl}\check{Y}_p\ , \qquad 1 \le k,l \le r\ . \tag{3.14c}$$

Proof. Since the symbol $T\iota(\mathcal{N}_q)$ is defined as the intersection $V\pi^q_{q-1} \cap T\iota(T\mathcal{R}_q)$ of two involutive distributions and thus itself an involutive distribution, there exists a basis $(\check{Y}_1, \check{Y}_2, \dots, \check{Y}_r)$ for it which is closed under Lie brackets. Thus, there exist coefficient functions $B^p_{kl} \in \mathcal{F}(\mathcal{R}_q)$ such that (3.14c) is satisfied. Since the vector fields $[\check{X}_i, \check{X}_j]$ and $[\check{X}_i, \check{Y}_k]$ are in $T\iota(\mathcal{V}'[\mathcal{R}_q])$, the equalities (3.14a) and (3.14b) follow from equalities (3.13) and $T\iota(\mathcal{V}[\mathcal{R}_q]) = T\iota(\mathcal{N}_q) \oplus T\iota(\mathcal{H})$. □

In principle, in a concrete application it suffices to analyze the structure equations in this form. Approaches so far (Fackerell [15], Stormark [40], Vassiliou [42]) are based on Vessiot's classical procedure [43] in that they do not use the decomposition $T\iota(\mathcal{V}[\mathcal{R}_q]) = T\iota(\mathcal{N}_q) \oplus T\iota(\mathcal{H})$ and thus have to analyze nonlinear equations (Section 3.2 shows the details). Our ansatz yields linear equations (see Section 3.3), and they can be simplified even further by way of applying Proposition 3.1.8 whenever its assumptions are met. Even if they are not, it is always possible to choose appropriate bases in triangular form by transforming the given bases into Jacobian systems by way of Gaussian elimination. This makes our considerations in what follows easier.

Lemma 3.1.22. *Under the assumptions of Lemma 3.1.21, there is a basis $(Y_k : 1 \le k \le r)$ for the symbol $T\iota(\mathcal{N}_q)$ and a basis $(X_i : 1 \le i \le n)$ for the complement $T\iota(\mathcal{H})$ such that*

$$[X_i, X_j] = \Theta^c_{ij}Z_c\ , \qquad 1 \le i,j \le n\ , \tag{3.14a$'$}$$
$$[X_i, Y_k] = \Xi^c_{ik}Z_c\ , \qquad 1 \le i \le n,\ 1 \le k \le r\ , \tag{3.14b$'$}$$
$$[Y_k, Y_l] = 0\ , \qquad 1 \le k,l \le r\ . \tag{3.14c$'$}$$

Proof. Define the vector fields Y_k by Equation (3.4). Then they form a Jacobian system as was shown in the proof of Proposition 3.1.8. Since the symbol is involutive (as a distribution of vector fields), it follows that $[Y_k, Y_l] = 0$ for all $1 \le k,l \le r$. Transforming the vector fields $\check{X}_i$ into a Jacobian system yields vector fields X_i for which the coefficient functions $\hat{A}^h_{ij} \equiv \tilde{A}^h_{ik} \equiv 0$. (Note that the X_i as they are given in Equation (3.3) already form a Jacobian system.) □

Remark 3.1.23. So far we have considered differential equations $\mathcal{R}_q$ which are represented by a system of equations which may be given in fully implicit form $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$. Our approach simplifies further, if in some δ-regular local coordinates $(\mathbf{x}, \mathbf{u}^{(q)})$ on $J_q\pi$ each of the equations in the representation (2.7) of $\mathcal{R}_q$ can be solved for another principal derivative. This induces a coordinate chart $(\overline{\mathbf{x}}, \overline{\mathbf{u}}^{(q)})$ on $\mathcal{R}_q$, and these local coordinates on $\mathcal{R}_q$ simplify the representation of the vector fields and distributions under consideration, since then a particularly convenient choice for the fields Y_k and W_a exists. We can choose for any $1 \leq k \leq r$ a parametric derivative $\overline{u^\alpha_\mu}$ of order $|\mu| = q$, that is, $(\alpha, \mu) \notin \mathcal{B}$, and set

$$\bar{Y}_k := \bar{Y}^\alpha_\mu := \partial_{\overline{u^\alpha_\mu}} \, . \tag{3.15}$$

Then Equation (3.14c′) is satisfied, because trivially $[\partial_{\overline{u}^\alpha_\mu}, \partial_{\overline{u}^\beta_\nu}] = 0$. Additionally, for any $1 \leq a \leq s$ (where s is the dimension of a complement $\mathcal{W}$ for $T\iota(\mathcal{N}_q)$ in $V\pi^q_{q-1}$) we can choose a principal derivative u^β_ν, that is, $(\beta, \nu) \in \mathcal{B}$, such that

$$W_a := W^\beta_\nu := \partial_{u^\beta_\nu} \, .$$

The vector fields $\iota_* \bar{Y}_k = Y_k$ and W_a combined then yield a basis for $(V\pi^q_{q-1})|_{\mathcal{R}_q}$. Therefore the vertical bundle $(V\pi^q_{q-1})|_{\mathcal{R}_q}$, as an involutive distribution, can be decomposed into

$$(V\pi^q_{q-1})|_{\mathcal{R}_q} = T\iota(\mathcal{N}_q) \oplus \mathcal{W} \, , \tag{3.16}$$

and $\mathcal{W}$ is again an involutive distribution. The distribution $\mathcal{W}$ is spanned by vector fields $\{W_a \colon 1 \leq a \leq s\}$, where $s = \sum_{k=1}^n \beta^{(k)}_q$ equals the number of principal derivatives in the representation of $\mathcal{R}_q$ (using δ-regular coordinates). Since the vector fields W_a are chosen in triangular form (as in Proposition 3.1.8), we have $[W_a, W_b] = 0$ for all $1 \leq a, b \leq s$.

Now the reference complement $\mathcal{H}_0$ for $\mathcal{N}_q$ within $\mathcal{V}[\mathcal{R}_q]$ can be chosen as follows. (It need not be involutive.) Any basis of it must consist of n vector fields which are mapped into contact fields by ι_* and which are transversal with respect to the fibration $\hat{\pi}^q \colon \mathcal{R}_q \to \mathcal{X}$. Among the generating vector fields (2.4), only the vector fields $C^{(q)}_i$ are transversal. And since the vector fields C^μ_α are vertical, for any complement $\mathcal{H}$ we can construct a basis $(\tilde{X}_i \colon 1 \leq i \leq n)$ of the form $\iota_* \tilde{X}_i = C^{(q)}_i + \xi^\alpha_{i\mu} C^\mu_\alpha$ with some coefficient functions $\xi^\alpha_{i\mu} \in \mathcal{F}(\mathcal{R}_q)$ chosen such that $\tilde{X}_i$ is tangent to $\mathcal{R}_q$. The vector fields C^μ_α also span the vertical bundle $V\pi^q_{q-1}$, and hence we may use the decomposition (3.16) to further simplify the basis of the complement $\mathcal{H}$. By subtracting from each $\tilde{X}_i$ a suitable linear combination of the fields Y_k spanning the symbol $\mathcal{N}_q$, we arrive at a basis $(\bar{X}_i \colon 1 \leq i \leq n)$ for a coordinate-dependent reference complement $\mathcal{H}_0$ where

$$\iota_* \bar{X}_i = C^{(q)}_i + \xi^a_i W_a \, . \tag{3.17}$$

This construction of a basis $(\bar{X}_1, \ldots, \bar{X}_n, \bar{Y}_1, \ldots, \bar{Y}_r)$ for the Vessiot distribution $\mathcal{V}[\mathcal{R}_q]$ means, the non-vanishing structure equations of $\mathcal{V}[\mathcal{R}_q]$ now take the simple form (3.14a′) and (3.14b′) with smooth functions Θ^c_{ij} and Ξ^c_{ik} in $\mathcal{F}(\mathcal{R}_q)$.

Example 3.1.24. We continue Example 3.1.4 and compare the system's representation in local coordinates on $\mathcal{R}^{(1)}_1$ there with the one in local coordinates on $J_q\pi$ given in Example

3.1.2. The Lie-brackets for the vector fields X_i and Y_k are given in Example 3.1.10. Now the Lie brackets for $\bar{X}_i$ and $\bar{Y}_k$ are

$$\begin{aligned}
[\bar{X}_1,\bar{X}_2] &= \overline{v_x}\partial_{\overline{u}} - \overline{v_x}\partial_{\overline{u}} &&= 0,\\
[\bar{X}_1,\bar{X}_3] &= -\partial_{\overline{v}},\\
[\bar{X}_1,\bar{X}_4] &= -\partial_{\overline{w}},\\
[\bar{X}_2,\bar{X}_3] &= -\partial_{\overline{w}} &&= [\bar{X}_1,\bar{X}_4],\\
[\bar{X}_2,\bar{X}_4] &= -\partial_{\overline{v}} &&= [\bar{X}_1,\bar{X}_3],\\
[\bar{X}_3,\bar{X}_4] &= 0.
\end{aligned}$$

We have $\iota_*\partial_{\overline{v}} = \partial_v + \partial_{u_t}$ and $\iota_*\partial_{\overline{w}} = \partial_w + \partial_{u_x}$. This expresses the fact that $\iota_*[\bar{X}_i,\bar{X}_j] = [\iota_*\bar{X}_i,\iota_*\bar{X}_j]$ for arbitrary vector fields $\bar{X}_i$ and $\bar{X}_j$ in $T\mathcal{R}_q$.

The dimension of the base space $\mathcal{X}$ is $n=2$. The two fields $\bar{X}_1$ and $\bar{X}_2$ are transversal (with regard to the fibration $\hat{\pi}^q\colon \mathcal{R}_q \to \mathcal{X}$). Setting $\bar{X}_3 =: \bar{Y}_1$ and $\bar{X}_4 =: \bar{Y}_2$ we have $\bar{Y}_1 = \partial_{\overline{v_x}}$ and $\bar{Y}_2 = \partial_{\overline{w_x}}$ as proposed in Equation (3.15). They form a basis for $\mathcal{N}_1$ and satisfy $\iota_*\bar{Y}_1 = \partial_{v_x} + \partial_{w_t}$ and $\iota_*\bar{Y}_2 = \partial_{w_x} + \partial_{v_t}$, which form the basis for $T\iota(\mathcal{N}_1)$ in Example 3.1.10. For the four principal derivatives we have the vector fields $W_1 = \partial_{u_t}$, $W_2 = \partial_{v_t}$, $W_3 = \partial_{w_t}$ and $W_4 = \partial_{u_x}$, spanning $\mathcal{W}$. We saw in Example 3.1.4 that

$$\begin{aligned}
\iota_*\bar{X}_1 = X_1 = \partial_x + u_x\partial_u + v_x\partial_v + w_x\partial_w + v_x\partial_{u_t} + w_x\partial_{u_x},\\
\iota_*\bar{X}_2 = X_2 = \partial_t + u_t\partial_u + v_t\partial_v + w_t\partial_w + v_t\partial_{u_t} + w_t\partial_{u_x},
\end{aligned}$$

which are indeed of the form given in Equation (3.17) and form the basis for the reference complement $\mathcal{H}_0$ (the coefficients ξ_i^a are as in Example 3.1.10).

Remark 3.1.25. For a first-order equation $\mathcal{R}_1$ in reduced Cartan normal form (2.27) there are no integrability conditions which arise as prolongations of equations of lower order as there are no algebraic conditions. Thus it satisfies the assumptions of Proposition 3.1.17, and it is possible to exploit the above considerations for the construction of flat Vessiot connections explicitly. We choose as a reference complement $\mathcal{H}_0$ the linear span of the vector fields $\bar{X}_i$ given in Equation (3.11). According to Proposition 3.1.19, this is a valid choice. Furthermore, the $\iota_*\bar{X}_i$ have the form given in Equation (3.17). Using the vector fields given in Equation (3.15) as a basis for the symbol, we can explicitly evaluate the Lie brackets (3.14) in the simplified form (3.14′) on $\mathcal{R}_1$. Since the vector fields Z_c, which are given by Equation (3.13) and appear on the right sides of the structure equations (3.14′), for $q=1$ may span a proper subspace in $\operatorname{span}\{\partial_{u^\alpha}\colon 1\leq\alpha\leq m\}$, about the exact size of which we know nothing, we write them as linear combinations $Z_c =: \kappa_c^\alpha\partial_{u^\alpha}$. The structure equations (3.14a′, 3.14b′) then become

$$[X_i,X_j] = \varTheta_{ij}^c\kappa_c^\alpha\partial_{u^\alpha} =: \varTheta_{ij}^\alpha\partial_{u^\alpha}\ , \qquad 1\leq i,j\leq n\ , \tag{3.14a$''$}$$

$$[X_i,Y_k] = \varXi_{ik}^c\kappa_c^\alpha\partial_{u^\alpha} =: \varXi_{ik}^\alpha\partial_{u^\alpha}\ , \qquad 1\leq i\leq n,\ 1\leq k\leq r\ . \tag{3.14b$''$}$$

Knowing the larger sets of coefficients $\varTheta_{ij}^\alpha$, $\varXi_{ik}^\alpha$, we can reconstruct the true structure coefficients $\varTheta_{ij}^c$, $\varXi_{ik}^c$ by solving the over-determined systems of linear equations

$$\varTheta_{ij}^c\kappa_c^\alpha = \varTheta_{ij}^\alpha \quad\text{and}\quad \varXi_{ik}^c\kappa_c^\alpha = \varXi_{ik}^\alpha\ .$$

This is always possible since the fields Z_c are assumed to be part of a basis for the derived Vessiot distribution $\mathcal{V}'[\mathcal{R}_1]$ and therefore linearly independent. Thus there exist some coefficient functions κ_α^c such that

$$\Theta_{ij}^c = \Theta_{ij}^\alpha \kappa_\alpha^c \quad \text{and} \quad \Xi_{ik}^c = \Xi_{ik}^\alpha \kappa_\alpha^c \ .$$

When we calculate in local coordinates on $\mathcal{R}_q$, we can use that for a system in reduced Cartan normal form $\iota_* \partial_{\overline{u^\alpha}} = \partial_{u^\alpha}$ for all $1 \leq \alpha \leq m$ and set $\iota_* \bar{Z}_c := Z_c$.

In Section 3.3 we have to analyze certain matrices with the coefficients Θ_{ij}^c and Ξ_{ik}^c as their entries. It turns out that this analysis becomes simpler, if we use the extended sets of coefficients Θ_{ij}^α and Ξ_{ik}^α instead.

Lemma 3.1.26. *If $i < j$, then we obtain for the extended set of structure coefficients Θ_{ij}^α in local coordinates on $\mathcal{R}_1$ the following results:*

$$\Theta_{ij}^\alpha = \begin{cases} C_i^{(1)}(\phi_j^\alpha) - C_j^{(1)}(\phi_i^\alpha) : (\alpha, i) \in \mathcal{B} \text{ and } (\alpha, j) \in \mathcal{B} \ , \\ C_i^{(1)}(\phi_j^\alpha) \qquad\qquad\qquad : (\alpha, i) \notin \mathcal{B} \text{ and } (\alpha, j) \in \mathcal{B} \ , \\ 0 \qquad\qquad\qquad\qquad\quad\ : (\alpha, i) \notin \mathcal{B} \text{ and } (\alpha, j) \notin \mathcal{B} \ . \end{cases} \tag{3.18}$$

Proof. According to Equation (3.10), the Lie-brackets of the vector fields $\bar{X}_i$ and $\bar{X}_j$ are

$$[\bar{X}_i, \bar{X}_j] = [\partial_{\overline{x^i}} + \sum_{\gamma=1}^{\beta_1^{(i)}} \phi_i^\gamma \partial_{\overline{u^\gamma}} + \sum_{\gamma=\beta_1^{(i)}+1}^{m} \overline{u_i^\gamma} \partial_{\overline{u^\gamma}} \ , \ \partial_{\overline{x^j}} + \sum_{\delta=1}^{\beta_1^{(j)}} \phi_j^\delta \partial_{\overline{u^\delta}} + \sum_{\delta=\beta_1^{(j)}+1}^{m} \overline{u_j^\delta} \partial_{\overline{u^\delta}}] \ .$$

We now use that for $1 \leq i, j \leq n$ and $1 \leq \gamma, \delta \leq m$ the terms $\partial_{\overline{x^i}}(\overline{u_j^\delta})$, $\phi_i^\gamma \partial_{\overline{u^\gamma}}(\overline{u_j^\delta}) \partial_{\overline{u^\delta}}$, $\overline{u_i^\gamma} \partial_{\overline{u^\gamma}}(\overline{u_j^\delta}) \partial_{\overline{u^\delta}}$ and $[\partial_{\overline{u^\gamma}}, \partial_{\overline{u^\delta}}]$ all vanish. Therefore what remains of the Lie-brackets is

$$\begin{aligned} [\bar{X}_i, \bar{X}_j] = & \sum_{\delta=1}^{\beta_1^{(j)}} \left\{ \partial_{\overline{x^i}} + \sum_{\gamma=1}^{\beta_1^{(i)}} \phi_i^\gamma \partial_{\overline{u^\gamma}} + \sum_{\gamma=\beta_1^{(i)}+1}^{m} \overline{u_i^\gamma} \partial_{\overline{u^\gamma}} \right\} (\phi_j^\delta) \partial_{\overline{u^\delta}} \\ & - \sum_{\gamma=1}^{\beta_1^{(i)}} \left\{ \partial_{\overline{x^j}} + \sum_{\delta=1}^{\beta_1^{(j)}} \phi_j^\delta \partial_{\overline{u^\delta}} + \sum_{\delta=\beta_1^{(j)}+1}^{m} \overline{u_j^\delta} \partial_{\overline{u^\delta}} \right\} (\phi_i^\gamma) \partial_{\overline{u^\gamma}} \ . \end{aligned}$$

In the curly brackets, we recognize $C_i^{(1)}$ and $C_j^{(1)}$. Now let $i < j$. Then $\beta_1^{(i)} \leq \beta_1^{(j)}$, and we have

$$\begin{aligned} [\bar{X}_i, \bar{X}_j] = & \sum_{\delta=1}^{\beta_1^{(j)}} C_i^{(1)}(\phi_j^\delta) \partial_{\overline{u^\delta}} - \sum_{\gamma=1}^{\beta_1^{(i)}} C_j^{(1)}(\phi_i^\gamma) \partial_{\overline{u^\gamma}} \\ = & \sum_{\gamma=1}^{\beta_1^{(i)}} \left\{ C_i^{(1)}(\phi_j^\gamma) - C_j^{(1)}(\phi_i^\gamma) \right\} \partial_{\overline{u^\gamma}} + \sum_{\delta=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_i^{(1)}(\phi_j^\delta) \partial_{\overline{u^\delta}} \ . \end{aligned}$$

The coefficients of the first sum are those for which both (γ, i) and (γ, j) are in $\mathcal{B}$, while for the coefficients of the second sum $(\delta, i) \notin \mathcal{B}$ and only $(\delta, j) \in \mathcal{B}$. Thus, the coefficients are as given in Equation (3.18). $\square$

We collect these coefficients into vectors Θ_{ij}which have m rows each where the entries are ordered according to increasing α.

Lemma 3.1.27. *If we set $\bar{Y}_k = \bar{Y}_h^\beta$, then the extended set of structure coefficients Ξ_{ik}^α in local coordinates on $\mathcal{R}_1$ are*

$$\Xi_{ik}^\alpha = \begin{cases} -C_\beta^h(\phi_i^\alpha) : (\alpha, i) \in \mathcal{B} \ , \\ -1 \qquad\qquad : (\alpha, i) \notin \mathcal{B} \text{ and } (\alpha, i) = (\beta, h) \ , \\ 0 \qquad\qquad\quad : (\alpha, i) \notin \mathcal{B} \text{ and } (\alpha, i) \neq (\beta, h) \ . \end{cases} \tag{3.19}$$

Proof. According to Equation (3.10) and Equation (3.15) for $\mu = h$ and $\bar{Y}_k = \bar{Y}_h^\beta$, the Lie-brackets of $\bar{X}_i$ and $\bar{Y}_k$ are

$$[\bar{X}_i, \bar{Y}_k] = [\partial_{\overline{x^i}} + \sum_{\alpha=1}^{\beta_1^{(i)}} \phi_i^\alpha \partial_{\overline{u^\alpha}} + \sum_{\alpha=\beta_1^{(i)}+1}^{m} \overline{u_i^\alpha} \partial_{\overline{u^\alpha}} \ , \ \partial_{\overline{u_h^\beta}}] \ .$$

Since for $1 \leq h, i \leq n$ and $1 \leq \alpha, \beta \leq m$ the Lie-brackets $[\partial_{\overline{u^\alpha}}, \partial_{\overline{u_h^\beta}}]$ vanish, and because $\partial_{\overline{u_h^\beta}}(\overline{u_i^\alpha})\partial_{\overline{u^\alpha}} = \delta_{hi}\delta_{\alpha\beta}\partial_{\overline{u^\alpha}}$, what remains of the Lie-brackets is

$$\begin{aligned} [\bar{X}_i, \bar{Y}_h^\beta] &= -\sum_{\alpha=1}^{\beta_1^{(i)}} \partial_{\overline{u_h^\beta}}(\phi_i^\alpha)\partial_{\overline{u^\alpha}} - \sum_{\alpha=\beta_1^{(i)}+1}^{m} \delta_{hi}\delta_{\alpha\beta}\partial_{\overline{u^\alpha}} \\ &= -\sum_{\alpha=1}^{\beta_1^{(i)}} C_\beta^h(\phi_i^\alpha)\partial_{\overline{u^\alpha}} - \partial_{\overline{u^\beta}} \ . \end{aligned}$$

Since for $1 \leq \alpha \leq \beta_1^{(i)}$ we have $(\alpha, i) \in \mathcal{B}$ and of the coefficients for $\beta_1^{(i)} + 1 \leq \alpha \leq m$ only the one for $(\alpha, i) = (\beta, h)$ remains and equals -1, the coefficients are as given in Equation (3.19). $\square$

Remark 3.1.28. Some of the $-C_\beta^h(\phi_i^\alpha)$ where $(\alpha, i) \in \mathcal{B}$ vanish, too: all of the parametric derivatives on the right side of an equation Φ_i^α in the reduced Cartan normal form (2.27) are of a class lower than that of the equation's left side as otherwise we would solve this equation for the derivative of highest class. This means $-C_\beta^h(\phi_i^\alpha) = 0$ whenever $i = \text{class}(u_i^\alpha) < \text{class}(u_h^\beta) = h$.

We collect the coefficients Ξ_{ik}^α into matrices Ξ_i using i as the number of the matrix to which the entry Ξ_{ik}^α belongs, α as the row index of the entry and k as its column index. These matrices have m rows each, ordered according to increasing α, and which have

$r = \dim \mathcal{N}_1$ columns each of which can be labelled by pairs $(\beta, h) \notin \mathcal{B}$ or the symbol fields $\bar{Y}_k = \bar{Y}_h^\beta$. More precisely, for $1 \leq h \leq n$, we set

$$\begin{pmatrix} -C^h_{\beta_1^{(h)}+1}(\phi_i^1) & -C^h_{\beta_1^{(h)}+2}(\phi_i^1) & \cdots & -C^h_m(\phi_i^1) \\ -C^h_{\beta_1^{(h)}+1}(\phi_i^2) & -C^h_{\beta_1^{(h)}+2}(\phi_i^2) & \cdots & -C^h_m(\phi_i^2) \\ \vdots & \vdots & \ddots & \vdots \\ -C^h_{\beta_1^{(h)}+1}(\phi_i^{\beta_1^{(i)}}) & -C^h_{\beta_1^{(h)}+2}(\phi_i^{\beta_1^{(i)}}) & \cdots & -C^h_m(\phi_i^{\beta_1^{(i)}}) \end{pmatrix} =: [\Xi_i]^h \ . \tag{3.20}$$

Such a matrix with an upper index h collects all those Ξ^α_{ik} into a block where $(\alpha, i) \in \mathcal{B}$. For any $1 \leq i \leq n$, we have $m - \beta_1^{(i)} = \alpha_1^{(i)}$, so such a matrix has $\beta_1^{(i)}$ rows and $\alpha_1^{(i)}$ columns. Since, for any h with $i < h$ and for all $\beta_1^{(h)} + 1 \leq \beta \leq m$, we have $-C^h_\beta(\phi_i^\alpha) = 0$, such matrices $[\Xi_i]^h$ where $i < h$ are zero. Furthermore, for $1 \leq h \leq n$, we assemble the remaining terms Ξ^α_{ik} (which are those where $(\alpha, i) \notin \mathcal{B}$) in a matrix. As above, let $\bar{Y}_k = \bar{Y}_h^\beta$, and denote any entry Ξ^α_{ik} accordingly, for the sake of introducing the following matrix, by ${}^h_\beta\Xi^\alpha_i$. Now set

$$\begin{pmatrix} {}^{\;\;h}_{\beta_1^{(h)}+1}\Xi_i^{\beta_1^{(i)}+1} & {}^{\;\;h}_{\beta_1^{(h)}+2}\Xi_i^{\beta_1^{(i)}+1} & \cdots & {}^h_m\Xi_i^{\beta_1^{(i)}+1} \\ {}^{\;\;h}_{\beta_1^{(h)}+1}\Xi_i^{\beta_1^{(i)}+2} & {}^{\;\;h}_{\beta_1^{(h)}+2}\Xi_i^{\beta_1^{(i)}+2} & \cdots & {}^h_m\Xi_i^{\beta_1^{(i)}+2} \\ \vdots & \vdots & \ddots & \vdots \\ {}^{\;\;h}_{\beta_1^{(h)}+1}\Xi_i^m & {}^{\;\;h}_{\beta_1^{(h)}+2}\Xi_i^m & \cdots & {}^h_m\Xi_i^m \end{pmatrix} =: [\Xi_i]_h \ .$$

For any $1 \leq h \leq n$, such a matrix with the index h written below has $\alpha_1^{(i)}$ rows and $\alpha_1^{(h)}$ columns. Let for any natural numbers a and b denote $0_{a \times b}$ the $a \times b$ zero matrix. According to equation (3.19), we have

$$[\Xi_i]_h = \begin{cases} -\mathbb{1}_{\alpha_1^{(i)}} & : h = i \ , \\ 0_{\alpha_1^{(i)} \times \alpha_1^{(h)}} & : h \neq i \ . \end{cases}$$

Using the matrices $[\Xi_i]^h$ and $[\Xi_i]_h$ as blocks we now build the matrix

$$\Xi_i = \begin{pmatrix} [\Xi_i]^1 & [\Xi_i]^2 \cdots & [\Xi_i]^n \\ [\Xi_i]_1 & [\Xi_i]_2 \cdots & [\Xi_i]_n \end{pmatrix} \ .$$

Taking into account what we have noted on its entries this means

$$\Xi_i = \begin{pmatrix} [\Xi_i]^1 & \cdots & [\Xi_i]^{i-1} & [\Xi_i]^i & 0 \cdots 0 \\ 0 & \cdots & 0 & -\mathbb{1}_{\alpha_1^{(i)}} & 0 \cdots 0 \end{pmatrix} \ . \tag{3.21}$$

A sketch is of such a matrix Ξ_i where the entries which may be different from zero are marked as shaded areas and $-\mathbb{1}_{\alpha_1^{(i)}}$ as a diagonal line is given in Figure (3.1).

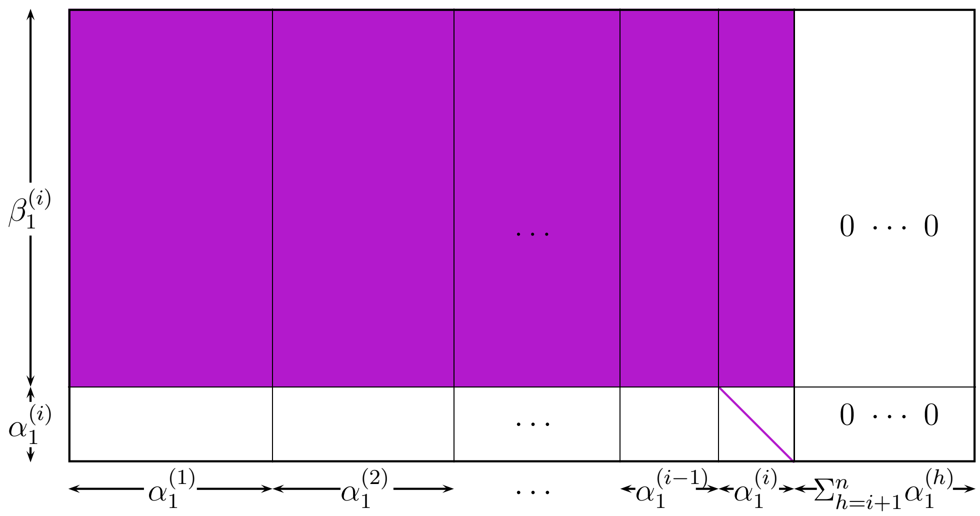


Figure 3.1: A sketch for Ξ_i in Equation (3.21).

For all h where $1 \leq h \leq n$, we call the block $[\Xi_i]^h$ in Ξ_i stacked upon the block $[\Xi_i]_h$ in Ξ_i *the hth block of columns in Ξ_i*. For those h with $\beta_1^{(h)} = m$ the hth block of columns is empty. Now the symbol fields $\bar{Y}_h^{\beta}$, or equivalently the pairs $(\beta, h) \notin \mathcal{B}$, are used to order the $\dim \mathcal{N}_1 = r$ columns of Ξ_i, according to increasing h into n blocks (empty for those h with $\alpha_1^{(h)} = 0$) and within each block according to increasing β (with $\beta_1^{(j)} + 1 \leq \beta \leq m$).

Remark 3.1.29. This means, the columns in Ξ_i are ordered increasingly with respect to the term-over-position lift of the degree reverse lexicographic ranking. Therefore, the entry $-C^h_{\beta_1^{(h)}+\gamma}(\phi_i^\alpha)$ stands in the matrix Ξ_i in line α, in the hth block of columns of which it is the γth one from the left. Entries different from zero and from -1 may appear in Ξ_i only in a $[\Xi_i]^h$ for $h \leq i$. To be exact, for any class i, the matrix Ξ_i has $\alpha_1^{(i)}$ rows where all entries are zero with only one exception: for each $1 \leq \ell_i < \alpha_1^{(i)}$ we have

$$\Xi_{i\,k}^{\beta_1^{(i)}+\ell_i} = -\delta_{\ell\, k}$$

where $\ell := \sum_{h=1}^{i-1} \alpha_1^{(h)} + \ell_i$. The entries in the remaining upper $\beta_1^{(i)}$ rows are $-C^h_\beta(\phi_i^\alpha)$. The potentially non-trivial ones of them are marked as shaded areas in Figure (3.1).

Note that for a differential equation with constant coefficients all vectors Θ_{ij} vanish and for a maximally over-determined equation there are no matrices Ξ_i.

The unit block of $\alpha_1^{(i)}$ rows, $\mathbb{1}_{\alpha_1^{(i)}}$, leads immediately to the estimate

$$\alpha_1^{(i)} \leq \operatorname{rank} \Xi_i \leq \min\{m, \sum_{h=1}^{i} \alpha_1^{(h)}\} \,.$$

Considerations of the rank for such matrices Ξ_i are at the heart of the proof of the existence theorem for Vessiot connections, Theorem 3.3.9.

3.2 Constructing Flat Vessiot Connections I: Recent Approaches

For easier comparison, we summarize in this subsection approaches from the modern literature (Fackerell [15], Stormark [40], Vassiliou [42]) on the subject of constructing the Vessiot distribution. Let $\mathcal{R}_q \subseteq J_q\pi$ be a differential equation of order q and dimension E, and let $\mathcal{V}[\mathcal{R}_q]$ denote its Vessiot distribution. Let $\mathcal{V}[\mathcal{R}_q]$ be D-dimensional, generated by the vector fields X_d where $1 \leq d \leq D$. If $\mathcal{V}[\mathcal{R}_q]$ is an involutive distribution, then as mentioned in Remark 2.1.11, according to the Frobenius theorem it has D-dimensional integral manifolds. Therefore there are $E - D$ functions $\Phi^\tau \colon \mathcal{R}_q \to \mathbb{R}$ such that

$$X_d \Phi^\tau = 0, \quad 1 \leq d \leq D,\ 1 \leq \tau \leq E - D\ ,$$

and the real constants $c_b = \Phi^\tau$ define a family of integral manifolds for the Vessiot distribution. If, on the other hand, $\mathcal{V}[\mathcal{R}_q]$ is not an involutive distribution, the aim is to construct involutive subdistributions $\mathcal{U}$ of greatest possible dimension $n < D$ within $\mathcal{V}[\mathcal{R}_q]$. Let $\mathcal{U}$ be spanned by the vector fields

$$U_i := \xi_i^d X_d, \quad 1 \leq d \leq D,\ 1 \leq i \leq n\ , \tag{3.22}$$

where $\xi_i^d \in \mathcal{F}(\mathcal{R}_q)$. Let Z_c, $1 \leq c \leq C$, be vector fields such that $\mathcal{V}'[\mathcal{R}_q] = \mathcal{V}[\mathcal{R}_q] \oplus \operatorname{span}\{Z_c \colon 1 \leq c \leq C\}$. Then the structure equations for the vector fields X_d generating $\mathcal{V}[\mathcal{R}_q]$ are

$$[X_d, X_e] = \Delta_{de}^f X_f + \Theta_{de}^c Z_c \tag{3.23}$$

for certain functions $\Delta_{de}^f, \Theta_{de}^c \in \mathcal{F}(\mathcal{R}_q)$. (Note that the conditions (3.23) are more complicated than those given in Equations (3.14′) since here the generating vector fields X_d do not form a Jacobian system. Stormark [40] does consider a system in triangular form and therefore has $[X_d, X_e] = \Theta_{de}^c Z_c$.) It follows that the structure equations for the vector fields U_i generating $\mathcal{U} \subset \mathcal{V}[\mathcal{R}_q]$ are

$$[U_i, U_j] = \frac{1}{2}(\xi_i^d \xi_j^e - \xi_j^d \xi_i^e)\Theta_{de}^c Z_c + \Big\{U_i(\xi_j^f) - U_j(\xi_i^f) + \xi_i^d \xi_j^e \Delta_{de}^f\Big\} X_f$$

where $1 \leq f \leq D$ and $1 \leq i, j \leq n$. (We have $\frac{1}{2}(\xi_i^d \xi_j^e - \xi_j^d \xi_i^e) = \xi_i^d \xi_j^e$.) The vector field $[U_i, U_j]$ lies in the Vessiot distribution if, and only if, for $1 \leq c \leq C$ the algebraic constraints

$$(\xi_i^d \xi_j^e - \xi_j^d \xi_i^e)\Theta_{de}^c = 0 \tag{3.24}$$

are satisfied. It lies in $\mathcal{U}$ if, and only if, additionally there are functions $\nu_{ij}^h, \xi_h^f \in \mathcal{F}(\mathcal{R}_q)$ such that for $[U_i, U_j] = \nu_{ij}^h U_h = \nu_{ij}^h \xi_h^f X_f$ the differential constraints

$$U_i(\xi_j^f) - U_j(\xi_i^f) + \xi_i^d \xi_j^e \Delta_{de}^f = \nu_{ij}^h \xi_h^f \tag{3.25}$$

are satisfied. (Turning the U_i into triangular form makes the terms ν_{ij}^h and $\xi_i^d \xi_j^e \Delta_{de}^f$ vanish, as is done in Stormark [40].)

The algebraic constraints (3.24) are satisfied if there are functions $\eta_{ij}^d \in \mathcal{F}(\mathcal{R}_q)$ such that $[U_i, U_j] = \eta_{ij}^d X_d$. Another way to write this is

$$[U_i, U_j] \equiv 0 \mod \mathcal{V}[\mathcal{R}_q] \ .$$

In the literature [15, 40, 42], two vector fields $U, V \in \mathcal{V}[\mathcal{R}_q]$ are said to be "in involution," if $[U, V] \equiv 0 \mod \mathcal{V}[\mathcal{R}_q]$. A set of linearly independent vector fields $\{U_i \in \mathcal{V}[\mathcal{R}_q] : 1 \leq i \leq I\}$ which are pairwise in involution is called an "involution of degree I". The most general involution of degree I is called an "involution of order I." Vessiot's approach to construct the involution of order I in $\mathcal{V}[\mathcal{R}_q]$ is to consider the vector fields which are to be pairwise in involution,

$$\begin{array}{llll}
[U_1, U_2] \equiv 0 & & & \mod \mathcal{V}[\mathcal{R}_q] \ , \\
[U_1, U_3] \equiv 0 \ , & [U_2, U_3] \equiv 0 & & \mod \mathcal{V}[\mathcal{R}_q] \ , \\
 & \vdots & & \\
[U_1, U_I] \equiv 0 \ , & [U_2, U_I] \equiv 0 \ , \ \ldots, & [U_{I-1}, U_I] & \mod \mathcal{V}[\mathcal{R}_q] \ ,
\end{array}$$

and solve this series of congruences successively. First set $U_1 := \xi_1^d X_d$ with coefficient functions $\xi_1^d \in \mathcal{F}(\mathcal{R}_q)$ for the general representation of a vector field in the Vessiot distribution (which means, the functions ξ_1^d are not constrained). Then set $U := \xi^d X_d \in \mathcal{V}[\mathcal{R}_q]$ for yet unknown coefficient functions $\xi^d \in \mathcal{F}(\mathcal{R}_q)$ which yield for

$$[U_1, U] \equiv 0 \mod \mathcal{V}[\mathcal{R}_q]$$

the most general vector field U as a solution. If $U = \mathbb{R}U_1$ is the only solution, then the subdistribution of maximal dimension $\mathcal{U} \subset \mathcal{V}[\mathcal{R}_q]$ is an involution of order 1. A one-dimensional subdistribution is trivially involutive. In this case $[\mathcal{U}, \mathcal{U}] \subseteq \mathcal{U}$, so there is no need to consider the differential conditions (3.25). Otherwise set $U =: U_2$. Then U_1 and U_2 make an involution of order 2. Next construct $U = \xi^d X_d$ as the most general solution of the congruence

$$[U_1, U] \equiv 0, \quad [U_2, U] \equiv 0 \mod \mathcal{V}[\mathcal{R}_q].$$

There already are the two solutions: $U = U_1$ and $U = U_2$; if there is another, linearly independent, solution, then set $U =: U_3$, and $\mathcal{V}[\mathcal{R}_q]$ contains an involution of order at least 3. Proceeding this way, one constructs from an involution of order $i-1$ an involution of order i by calculating U_i as the most general vector field U to solve the congruence

$$[U_1, U] \equiv 0, \quad [U_2, U] \equiv 0, \ldots, [U_{i-1}, U] \equiv 0 \mod \mathcal{V}[\mathcal{R}_q]$$

if there is such a U_i which is not already an element of the involution of order $i-1$. If there is no such U_i, then $\mathcal{V}[\mathcal{R}_q]$ contains an involution of order $i-1$ only.

For a subdistribution $\mathcal{U}$ of dimension at least 2 satisfying $[\mathcal{U}, \mathcal{U}] \subseteq \mathcal{V}[\mathcal{R}_q]$, the differential conditions (3.25) have to be satisfied, too, if $\mathcal{U}$ is to satisfy $[\mathcal{U}, \mathcal{U}] \subseteq \mathcal{U}$ as well. Vessiot proves the existence of the necessary functions ν_{ij}^h and ξ_h^f to solve this system of differential equations using the Cauchy-Kovalevskaya theorem. It follows that, if the algebraic conditions (3.24) are met, for analytic equations (3.25) there are i-dimensional subdistributions $\mathcal{U}$ which are closed under the Lie bracket. Then from the Frobenius theorem follows that there are i-dimensional integral manifolds which correspond to solutions of $\mathcal{R}_q$.

For all $1 \leq i \leq I$, the set of algebraic conditions

$$[U_1, U] \equiv 0, \quad [U_2, U] \equiv 0, \ldots, [U_i, U] \equiv 0 \mod \mathcal{V}[\mathcal{R}_q] \tag{3.26}$$

forms a system of linear equations in the unknowns ξ^d where $1 \leq d \leq D$. Let r_i denote the rank of this system. As the system for determining U in step i is contained in the system for determining U in step j if $i < j$, we have the chain of inequalities

$$r_1 \leq r_2 \leq \cdots \leq r_i$$

for all $1 \leq i \leq I$. For all i, the difference $r_i - r_{i-1}$ shows the increase in the rank of the system combined from all systems (3.26) up to i and are called "characters of $\mathcal{V}[\mathcal{R}_q]$". We shall see in Section 3.3 that these differences are indeed the Cartan characters $\alpha_q^{(i)}$ of the differential equation $\mathcal{R}_q$, if the symbol $\mathcal{N}_q$ is involutive. The sum $\sum_k^i \alpha_q^{(k)}$ then gives the rank of the combined system up to step i and is called an "index of $\mathcal{V}[\mathcal{R}_q]$" in the literature [15, 40, 42], but this does not correspond to our notion of an index introduced in Definition 2.4.23.

3.3 Constructing Flat Vessiot Connections II: A New Approach

In this section, we develop an approach for constructing flat Vessiot connections which improves the recent approaches in so far as we exploit the splitting of $\mathcal{V}[\mathcal{R}_q]$ suggested by Proposition 3.1.6 to introduce convenient bases for integral distributions which yield structure equations that are simpler than those in Equation (3.25), which are derived by a procedure where no such splitting is used. As a consequence, we know from the beginning the maximally possible dimension of an integral distribution (it is $\dim \mathcal{X} + \dim \mathcal{N}_q$) whereas in recent approaches the dimension is assumed to be unknown at the outset. We give necessary and sufficient conditions for Vessiot's approach to succeed.

We discuss the general case when the system is not necessarily represented in solved form (as is the case with recent approaches; Fackerell [15] and Vassiliou [42] consider special kinds of systems, like hyperbolic equations, where such a solved form naturally appears). In particular, we distinguish throughout whether the additional vector fields Z_c for the derived Vessiot distribution are in $T\mathcal{R}_q$ or in $T\iota(T\mathcal{R}_q) \cap \mathcal{C}'_q|_{\mathcal{R}_q}$; this distinction is not necessary in these recent approaches since for a system given in solved form coordinates on $\mathcal{R}_q$ are used.

3.3.1 Structure Equations for the Vessiot Distribution

Let the differential equation $\mathcal{R}_q$ locally be represented by the system $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$ for $1 \leq \tau \leq t$. Our goal is the construction of all n-dimensional transversal involutive subdistributions $\mathcal{U}$ within the Vessiot distribution $\mathcal{V}[\mathcal{R}_q]$. Taking some basis (X_i, Y_k) for $T\iota(\mathcal{V}[\mathcal{R}_q]) = T\iota(\mathcal{N}_q) \oplus T\iota(\mathcal{H})$, such that the vector fields Y_k are a basis for the symbol $T\iota(\mathcal{N}_q)$ and the vector fields X_i are a basis for some complement $T\iota(\mathcal{H})$, we make for the basis $(U_i\colon 1 \leq i \leq n)$ of such a distribution $\mathcal{U}$ the ansatz

$$U_i = X_i + \zeta_i^k Y_k \tag{3.27}$$

with yet undetermined coefficient functions $\zeta_i^k \in \mathcal{F}(\mathcal{R}_q)$. This ansatz follows naturally from our considerations in Proposition 3.1.8 and Lemma 3.1.22, as the fields X_i are transversal to the fibration over $\mathcal{X}$ and in $\mathcal{N}_q$ and all fields Y_k are vertical.

Lemma 3.3.1. *Let the functions $\hat{A}_{ij}^h$, $\tilde{A}_{i\ell}^h$, $\hat{B}_{ij}^p$, $\tilde{B}_{i\ell}^p$, $B_{\ell k}^p$, Θ_{ij}^c and $\Xi_{i\ell}^c$ be given as in the structure equations (3.14) for the Vessiot distribution. Set*

$$\hat{A}_{ij}^h + \zeta_j^\ell \tilde{A}_{i\ell}^h - \zeta_i^k \tilde{A}_{jk}^h =: \Gamma_{ij}^h \ .$$

Then for the n vector fields U_i given by equation (3.27), the structure equations are:

$$\begin{aligned}
[U_i, U_j] = {}& \Gamma_{ij}^h U_h \\
& + (\hat{B}_{ij}^p + \zeta_j^\ell \tilde{B}_{i\ell}^p - \zeta_i^k \tilde{B}_{jk}^p - \zeta_j^\ell \zeta_i^k B_{\ell k}^p + U_i(\zeta_j^p) - U_j(\zeta_i^p) - \zeta_h^p \Gamma_{ij}^h) Y_p \\
& + (\Theta_{ij}^c + \zeta_j^\ell \Xi_{i\ell}^c - \zeta_i^k \Xi_{jk}^c) Z_c \ .
\end{aligned} \tag{3.28}$$

Proof. The proof is a straightforward calculation. We give it just because the result is not obvious at first glance. Using $U_i = X_i + \zeta_i^k Y_k$ we have

$$\begin{aligned}
[U_i, U_j] = {}& [X_i, X_j] + \zeta_j^\ell [X_i, Y_\ell] + X_i(\zeta_j^\ell) Y_\ell - \zeta_i^k [X_j, Y_k] - X_j(\zeta_i^k) Y_k \\
& - \zeta_j^\ell \zeta_i^k [Y_\ell, Y_k] - \zeta_j^\ell Y_\ell(\zeta_i^k) Y_k + \zeta_i^k Y_k(\zeta_j^\ell) Y_\ell \ .
\end{aligned}$$

From the structure equations (3.14) follows that this equals

$$\begin{aligned}
& \hat{A}_{ij}^h X_h + \hat{B}_{ij}^p Y_p + \Theta_{ij}^c Z_c \\
& + \zeta_j^\ell \tilde{A}_{i\ell}^h X_h + \zeta_j^\ell \tilde{B}_{i\ell}^p Y_p + \zeta_j^\ell \Xi_{i\ell}^c Z_c + X_i(\zeta_j^\ell) Y_\ell \\
& - \zeta_i^k \tilde{A}_{jk}^h X_h - \zeta_i^k \tilde{B}_{jk}^p Y_p - \zeta_i^k \Xi_{jk}^c Z_c - X_j(\zeta_i^k) Y_k \\
& - \zeta_j^\ell \zeta_i^k B_{\ell k}^p Y_p - \zeta_j^\ell Y_\ell(\zeta_i^k) Y_k + \zeta_i^k Y_k(\zeta_j^\ell) Y_\ell \ .
\end{aligned}$$

Factoring out, changing summation indices where necessary and using (3.27), we arrive at

$$\begin{aligned}
& (\hat{A}_{ij}^h + \zeta_j^\ell \tilde{A}_{i\ell}^h - \zeta_i^k \tilde{A}_{jk}^h) X_h \\
& + (\hat{B}_{ij}^p + \zeta_j^\ell \tilde{B}_{i\ell}^p - \zeta_i^k \tilde{B}_{jk}^p - \zeta_j^\ell \zeta_i^k B_{\ell k}^p + U_i(\zeta_j^p) - U_j(\zeta_i^p)) Y_p \\
& + (\Theta_{ij}^c + \zeta_j^\ell \Xi_{i\ell}^c - \zeta_i^k \Xi_{jk}^c) Z_c \ .
\end{aligned}$$

Now the claim follows from $X_h = U_h - \zeta_h^p Y_p$ and the definition of Γ_{ij}^h. □

Now we can give a criterion for a distribution $\mathcal{U}$ to be an integral distribution or even integrable.

Corollary 3.3.2. *Let $\mathcal{U} \in \mathcal{V}[\mathcal{R}_q]$ be a subdistribution with a basis of the form (3.27). Then $\mathcal{U}$ is an integral distribution if, and only if, the vector fields U_i satisfy the algebraic conditions*

$$G^c_{ij} := \Theta^c_{ij} - \Xi^c_{jk}\zeta^k_i + \Xi^c_{ik}\zeta^k_j = 0\,, \qquad \begin{cases} 1 \le c \le C\,, \\ 1 \le i < j \le n\,. \end{cases} \tag{3.29}$$

The subdistribution $\mathcal{U}$ is integrable if, and only if, it additionally satisfies the differential conditions

$$\begin{aligned} H^p_{ij} := {} & U_i(\zeta^p_j) - U_j(\zeta^p_i) \\ & + \hat{B}^p_{ij} + \zeta^\ell_j\tilde{B}^p_{i\ell} - \zeta^k_i\tilde{B}^p_{jk} - \zeta^\ell_j\zeta^k_i B^p_{\ell k} - \zeta^p_h\Gamma^h_{ij} = 0\,, \end{aligned} \qquad \begin{cases} 1 \le p \le r\,, \\ 1 \le i < j \le n\,. \end{cases} \tag{3.30}$$

Proof. If, and only if, the algebraic constraints (3.29) are satisfied, from the structure equations (3.28) and $\mathcal{N}_q \subseteq \mathcal{V}[\mathcal{R}_q]$ follows that $[U_i, U_j] \in \mathcal{V}[\mathcal{R}_q]$. Now the claim for the first part follows from Proposition 3.1.13.

And if, and only if, additionally the differential constraints (3.30) are satisfied, from the structure equations (3.28) follows that even $[U_i, U_j] \in \mathcal{U}$. By definition of the Y_p and Z_c, these fields are linearly independent, so their coefficients in Equation (3.28) must vanish for $\mathcal{U}$ to be involutive. □

Whenever the vector fields U_i can be constructed in triangular form—which is the case for a system given in the reduced Cartan normal form (2.27)—their structure equations become especially simple. Then the criterion of Corollary 3.3.2 simplifies as well.

Corollary 3.3.3. *Let $\mathcal{U} \in \mathcal{V}[\mathcal{R}_q]$ be a subdistribution with a basis of vector fields (3.27) where the X_i and Y_k are in triangular form—for example as they are given in Equations (3.3, 3.4). Then $\mathcal{U}$ is an integral distribution if, and only if, the vector fields U_i satisfy the algebraic conditions (3.29), and integrable if, and only if, they additionally satisfy*

$$U_i(\zeta^p_j) - U_j(\zeta^p_i) = 0\,, \qquad \begin{cases} 1 \le p \le r\,, \\ 1 \le i < j \le n\,. \end{cases} \tag{3.30$'$}$$

Proof. The first equivalence (for $\mathcal{U}$ being an integral distribution) has been shown in Corollary 3.3.2. Now let the algebraic conditions (3.29) be met and let $\mathcal{U}$ be integrable. Since the fields X_i are in triangular form, so are the U_i, and therefore the distribution $\mathcal{U}$ is involutive if, and only if, all Lie brackets $[U_i, U_j]$ vanish. What remains of Equation (3.28) is

$$0 = \Gamma^h_{ij}U_h + (\hat{B}^p_{ij} + \zeta^\ell_j\tilde{B}^p_{i\ell} - \zeta^k_i\tilde{B}^p_{jk} - \zeta^\ell_j\zeta^k_i B^p_{\ell k} + U_i(\zeta^p_j) - U_j(\zeta^p_i) - \zeta^p_h\Gamma^h_{ij})Y_p\,.$$

Since the X_i and Y_k are in triangular form, the structure equations (3.14′) can be applied. This means that the functions Γ^h_{ij}, $\hat{B}^p_{ij}$, $\tilde{B}^p_{i\ell}$ and $B^p_{\ell k}$ in (3.28) vanish.

Conversely, if the algebraic conditions are met and all $U_i(\zeta^p_j) - U_j(\zeta^p_i) = 0$, then what remains of Equation (3.28) contains only terms which vanish according to the structure equations (3.14′), and it follows that all $[U_i, U_j] = 0$. □

This form, (3.30′), of the differential conditions is simpler than those in some recent approaches [15, 42], given in Equation (3.25), where the advantage of having a representation of the differential equation in solved form was not used to construct the generating vector fields as a Jacobian system.

Example 3.3.4. We continue Example 3.1.10 which goes back to Example 2.2.2, the wave equation, which was rewritten as an involutive system of first order. Here $n = 2$. The two fields X_1 and X_2 are transversal (with regard to the fibration over $\mathcal{X}$). Setting

$$U_1 := X_1 + \zeta_1^1 Y_1 + \zeta_1^2 Y_2 \quad \text{and} \quad U_2 := X_2 + \zeta_2^1 Y_1 + \zeta_2^2 Y_2$$

according to our ansatz in Equation (3.27), we have

$$\begin{aligned}[U_1, U_2] = {} & (v_x - w_t)\partial_u \\ & + (\zeta_1^2 - \zeta_2^1)(\partial_v + \partial_{u_t}) + (\zeta_1^1 - \zeta_2^2)(\partial_w + \partial_{u_x}) \\ & + (U_1(\zeta_2^1) - U_2(\zeta_1^1))Y_1 + (U_1(\zeta_2^2) - U_2(\zeta_1^2))Y_2 \,.\end{aligned}$$

This yields the algebraic conditions $\zeta_1^2 - \zeta_2^1$ and $\zeta_1^1 - \zeta_2^2$ and the differential conditions $U_1(\zeta_2^1) - U_2(\zeta_1^1)$ and $U_1(\zeta_2^2) - U_2(\zeta_1^2)$. It also yields the integrability condition $w_t = v_x$.

In the algebraic conditions (3.29) the true structure coefficients Θ_{ij}^c, Ξ_{jk}^c appear. For our subsequent analysis we follow Remark 3.1.25 and replace them by the extended set of coefficients Θ_{ij}^α, Ξ_{jk}^α. This corresponds to replacing (3.29) by an equivalent but larger linear system of equations which is simpler to analyze.

The vector fields Y_k lie in the Vessiot distribution $\mathcal{V}[\mathcal{R}_1]$. Thus, according to Proposition 3.1.13, $\mathcal{U}$ is an integral distribution if, and only if, the coefficients ζ_k^i satisfy the algebraic conditions (3.29). This observation permits us immediately to reduce the number of unknowns ζ_k^i in our ansatz.

Lemma 3.3.5. Let $1 \leq i < j \leq n$, and let

$$\iota_* \bar{Y}_{(\alpha,\kappa)} = \partial_{u_\kappa^\alpha} + \sum_{\substack{(\beta,\mu)\in\mathcal{B} \\ |\mu|=q}} \xi_{(\alpha,\kappa)}^{(\beta,\mu)} \partial_{u_\mu^\beta} \quad \text{and} \quad \iota_* \bar{Y}_{(\alpha,\lambda)} = \partial_{u_\lambda^\alpha} + \sum_{\substack{(\beta,\mu)\in\mathcal{B} \\ |\mu|=q}} \xi_{(\alpha,\lambda)}^{(\beta,\mu)} \partial_{u_\mu^\beta} \tag{3.31}$$

be vector fields from the symbol of the differential equation $\mathcal{R}_q$ such that $\kappa + 1_j = \lambda + 1_i$. Then the coefficients in the algebraic conditions (3.29) are interrelated by $\zeta_j^{(\alpha,\kappa)} = \zeta_i^{(\alpha,\lambda)}$.

Proof. Let σ be a local section with $\sigma = (\mathbf{x}, \mathbf{s}(\mathbf{x}))$ for some appropriate function $\mathbf{s} = (s^\alpha \colon 1 \leq \alpha \leq m)$. Let the Taylor-expansion of s^α at $\mathbf{x}_0 \in \mathcal{X}$ be $s^\alpha(\mathbf{x}) = c_\mu^\alpha(\mathbf{x} - \mathbf{x}_0)^\mu/\mu!$. If σ is a solution to $\mathcal{R}_q$, then $\operatorname{im} j_q\sigma$ is an integral manifold of an n-dimensional transversal involutive subdistribution $\mathcal{U} \subseteq \mathcal{V}[\mathcal{R}_q]$. This means, at $\mathbf{x}_0 \in \mathcal{X}$ with $\rho := j_q\sigma(\mathbf{x}_0)$, we have $T_\rho(\operatorname{im} j_q\sigma) = \mathcal{U}_q$. Now we use the relation between the tangent space of the qth prolongation of σ and the prolongation of order $q+1$ which is given by the contact map (see Remark 2.1.7): let $\hat{\rho} := j_{q+1}\sigma(\mathbf{x}_0)$. Then $T\iota(\mathcal{U}_\rho) = \operatorname{im}\Gamma_{q+1}(\hat{\rho})$.

According to Remark 3.1.23 we can choose vector fields $\bar{Y}_k$ and $\bar{X}_i$ which generate the symbol and an n-dimensional reference complement and satisfy

$$\iota_* \bar{X}_i = C_i^{(q)} + \sum_{\substack{(\beta,\nu)\in\mathcal{B}\\ |\nu|=q}} \xi_i^{(\beta,\nu)} \partial_{u_\nu^\beta} \quad \text{and} \quad \iota_* \bar{Y}_k = Y_{(\alpha,\mu)} = \partial_{u_\mu^\alpha} + \sum_{\substack{(\beta,\nu)\in\mathcal{B}\\ |\nu|=q}} \xi_{(\alpha,\mu)}^{(\beta,\nu)} \partial_{u_\nu^\beta}$$

for $1 \le i \le n$ and $(\alpha, \mu) \notin \mathcal{B}$ with $|\mu| = q$ such that the vector fields

$$U_i = C_i^{(q)} + \xi_i^a W_a + \zeta_i^k Y_k \tag{3.32}$$

are a basis for $\mathcal{U}$. (We have again $W_a = \partial_{u_\nu^\beta}$ for all $(\beta, \nu) \in \mathcal{B}$ where $|\nu| = q$.) This means $U_i|_\rho = \Gamma_{q+1}(\hat{\rho}, \partial_{x^i})$ according to the interpretation of Γ_{q+1} as an evaluation map. It follows from the coordinate form (2.3) of the contact map

$$U_i|_\rho = \partial_{x^i}|_\rho + \sum_{\alpha=1}^{m} \sum_{0\le|\mu|<q+1} u_{\mu+1_i}^\alpha\big|_{\mathbf{x}_0} \partial_{u_\mu^\alpha}|_\rho = C_i^{(q)}|_\rho + \sum_{\alpha=1}^{m} \sum_{|\mu|=q} c_{\mu+1_i}^\alpha \partial_{u_\mu^\alpha}|_\rho \ .$$

Here we use that $u_{\mu+1_i}^\alpha = \partial^{|\mu+1_i|} s^\alpha(\mathbf{x}) / \partial \mathbf{x}^{\mu+1_i}$ for the prolongation $j_{q+1}\sigma\colon \mathcal{X} \to J_{q+1}\pi$, which equals $c_{\mu+1_i}^\alpha$ when evaluated in $\mathbf{x}_0$. Using $W_a = W_{(\alpha,\mu)} = \partial_{u_\mu^\alpha}$ for $(\alpha, \mu) \in \mathcal{B}$, the comparison with U_i as given in Equation (3.32) yields for the summands of order q

$$\begin{aligned}
&\sum_{\substack{(\alpha,\mu)\in\mathcal{B}\\ |\mu|=q}} c_{\mu+1_i}^\alpha \partial_{u_\mu^\alpha} + \sum_{\substack{(\alpha,\mu)\notin\mathcal{B}\\ |\mu|=q}} c_{\mu+1_i}^\alpha \partial_{u_\mu^\alpha} \\
&= \sum_{\substack{(\alpha,\mu)\in\mathcal{B}\\ |\mu|=q}} \xi_i^{(\alpha,\mu)} W_{(\alpha,\mu)} + \sum_{\substack{(\alpha,\mu)\notin\mathcal{B}\\ |\mu|=q}} \zeta_i^{(\alpha,\mu)} \Biggl(\partial_{u_\mu^\alpha} + \sum_{\substack{(\beta,\nu)\in\mathcal{B}\\ |\nu|=q}} \xi_{(\alpha,\mu)}^{(\beta,\nu)} W_{(\beta,\nu)} \Biggr) \\
&= \sum_{\substack{(\beta,\nu)\in\mathcal{B}\\ |\nu|=q}} \Biggl(\xi_i^{(\beta,\nu)} + \sum_{\substack{(\alpha,\mu)\notin\mathcal{B}\\ |\mu|=q}} \zeta_i^{(\alpha,\mu)} \xi_{(\alpha,\mu)}^{(\beta,\nu)} \Biggr) W_{(\beta,\nu)} + \sum_{\substack{(\alpha,\mu)\notin\mathcal{B}\\ |\mu|=q}} \zeta_i^{(\alpha,\mu)} \partial_{u_\mu^\alpha} \ .
\end{aligned}$$

It follows that if $(\beta, \nu) \in \mathcal{B}$, then

$$c_{\nu+1_i}^\beta = \xi_i^{(\beta,\nu)} + \sum_{\substack{(\alpha,\mu)\notin\mathcal{B}\\ |\mu|=q}} \zeta_i^{(\alpha,\mu)} \xi_{(\alpha,\mu)}^{(\beta,\nu)} \ ,$$

and if $(\alpha, \mu) \notin \mathcal{B}$, then

$$c_{\mu+1_i}^\alpha = \zeta_i^{(\alpha,\mu)} \ .$$

We conclude that if $\iota_* \bar{Y}_{(\alpha,\kappa)}$ and $\iota_* \bar{Y}_{(\alpha,\lambda)}$ are as in Equation (3.31) and if $\kappa + 1_j = \lambda + 1_i$, too, then the coefficients are interrelated by $\zeta_j^{(\alpha,\kappa)} = c_{\kappa+1_j}^\alpha = c_{\lambda+1_i}^\alpha = \zeta_i^{(\alpha,\lambda)}$. □

For the special case when $q = 1$ and therefore $|\kappa| = |\lambda| = 1$, this means the following. Assume that we have values $1 \le i < j \le n$ and $1 \le \alpha \le m$ such that both (α, i) and (α, j) are not contained in $\mathcal{B}$, which means both u_i^α and u_j^α are parametric derivatives (and thus the second-order derivative u^α_{ij}, too). Then there exist two symbol fields $\iota_*(\partial_{\overline{u^\alpha_i}}) = Y_{(\alpha,i)}$ and $\iota_*(\partial_{\overline{u^\alpha_j}}) = Y_{(\alpha,j)}$. In the notation of Lemma 3.3.5, we have $i = \kappa$ and $j = \lambda$ if, and only if, $\kappa + 1_j = \lambda + 1_i$ holds. If we set $Y_k = Y_{(\alpha,i)}$ and $Y_l = Y_{(\alpha,j)}$ and assume $\kappa + 1_j = \lambda + 1_i$, then according to Lemma 3.3.5 their coefficients satisfy

$$\zeta^k_j = \zeta^l_i \quad \text{or, equivalently,} \quad \zeta^{(\alpha,i)}_j = \zeta^{(\alpha,j)}_i \ . \tag{3.33}$$

Now it follows that $\mathcal{U}$, spanned by the vector fields $U_h = X_h + \zeta^k_h Y_k$ for $1 \le h \le n$, can be an integral distribution if, and only if, $\zeta^k_j = \zeta^l_i$ for all $1 \le i < j \le n$ and $1 \le k, l \le r$.

Remark 3.3.6. The matrix for the system of linear equations which forms the algebraic conditions (3.29) is

$$\begin{pmatrix}
\Xi_1 & \Xi_2 & & & & -\Theta_{12} \\
\Xi_1 & \Xi_3 & & & & -\Theta_{13} \\
\Xi_2 & & \Xi_3 & & & -\Theta_{23} \\
\Xi_1 & \Xi_4 & & & & -\Theta_{14} \\
\Xi_2 & & \Xi_4 & & & -\Theta_{24} \\
\Xi_3 & & & \Xi_4 & & -\Theta_{34} \\
\vdots & & \vdots & & & \vdots \\
\Xi_1 & \Xi_n & & & & -\Theta_{1n} \\
\Xi_2 & & \Xi_n & & & -\Theta_{2n} \\
\vdots & & & \ddots & & \vdots \\
\Xi_{n-1} & & & & \Xi_n & -\Theta_{n-1n}
\end{pmatrix} . \tag{3.34}$$

(Empty spaces mark zero entries.) For all $1 \le i < j \le n$, the columns of the matrices Ξ_j in Equation (3.29) can be labelled by the indices $1 \le l \le r$ of the unknowns ζ^l_i. We use this to label the columns in the matrix (3.34), except for the first r columns, which make up the stack of matrices from Ξ_1 to Ξ_{n-1}, and the last column, which contains the entries of the vectors $-\Theta_{ij}$, as follows. If an entry X of a column which we want to label is non-trivial, it appears in some matrix Ξ_j and has a label ζ^l_i indicating its column in Ξ_j by l. This matrix Ξ_j is right to some Ξ_i in the matrix (3.34) with $i < j$, and this i is the same for all non-trivial entries in the whole column in the matrix (3.34) which contains the entry X. Therefore it is possible to label that column by ζ^l_i. So far, the order q has been arbitrary. Now let $q = 1$. Then the label ζ^l_i is equivalent to a pair (α, ij) where (α, j) corresponds to l and ij is a multi-index of second order. Another way to write this label is by using the symbol field Y^α_{ij} from the representation of the prolonged system $\mathcal{R}_2$.

Now the identification shown in Equation (3.33) means that in the matrix (3.34) several pairs of columns share their label and have to be combined into one (those which are labelled by $\zeta^{(\alpha,i)}_j$ and $\zeta^{(\alpha,j)}_i$ or, equivalently, by Y^α_{ij} and Y^α_{ji}). This leads to a *contraction* of the matrix (3.34). We introduce now *contracted matrices* $\hat{\Xi}_h$ which arise as follows: whenever $\zeta^k_j = \zeta^l_i$, then the corresponding columns of the matrix (3.34) are added. If the column labelled l is left to the one labelled k, we enter the resulting column instead of the one labelled ζ^l_i and cancel the column labelled ζ^k_j. The matrix which is made in this way from a matrix Ξ_h which appears as a block in the matrix (3.34) is denoted $\hat{\Xi}_h$. Similarly, we introduce reduced vectors $\hat{\zeta}_h$ where the redundant components are left out. From now on we always understand that in the equations above this reduction has been performed. Mind that the notation $\hat{\Xi}_h$ just shows that the matrix Ξ_h has changed because we now take the equalities (3.33) into account; it may denote different matrices, as Ξ_h may appear as a block of the matrix (3.34) several times and therefore may be affected in different ways by different contractions. We analyze this point in detail in Subsection 3.3.4.

According to these considerations, the existence of flat Vessiot connections is equivalent to the solvability of the combined system of conditions (3.29, 3.30). The conditions themselves are easily derived for a differential equation $\mathcal{R}_q$, and in our approach the algebraic part (3.29) is an inhomogeneous linear system in the unknowns ζ_i^k and therefore poses no problem. The differential conditions (3.30) form a quasi-linear differential system for the functions ζ_i^k, with an inhomogeneous term consisting of linear and mixed-quadratic expressions for the functions ζ_i^k which vanishes for our special ansatz of Lemma 3.1.22 when the differential conditions simplify to (3.30′).

3.3.2 The Existence Theorem for Integral Distributions

Now the question arises, when the combined system (3.29, 3.30) has solutions. We begin by analyzing the algebraic part (3.29). The solvability of the differential conditions is discussed in Subsection 3.3.7. Given the linear system (3.29) for the vectors ζ_i and ζ_j, we now seek to build a solution step by step with j increasing and $1 \leq i < j$ for each j; this step-by-step approach is Vessiot's [43] original proposal, and we are going to examine the necessary and sufficient assumptions for it to succeed. According to Lemma 3.3.5, we may replace ζ_2 by $\hat{\zeta}_2$ since for the entries $\zeta_2^{(\beta,1)}$ where $\beta_1^{(2)} + 1 \leq \beta \leq m$ we know already that $\zeta_2^{(\beta,1)} = \zeta_1^{(\beta,2)}$. Thus we begin the construction of the integral distribution $\mathcal{U}$ by first choosing an arbitrary vector field U_1 and then aiming for another vector field U_2 such that $[U_1, U_2] \in T\iota(\mathcal{V}[\mathcal{R}_q])$. During the construction of U_2 we regard the components of the vector $\zeta_1 = \hat{\zeta}_1$ as given parameters and the components of $\hat{\zeta}_2$ as the only unknowns of the system

$$\hat{\Xi}_1 \hat{\zeta}_2 = \hat{\Xi}_2 \hat{\zeta}_1 - \Theta_{12} \,. \tag{3.35}$$

Since the components of $\hat{\zeta}_1$ are not considered as unknowns, the system (3.35) must not lead to any restrictions for the coefficients $\hat{\zeta}_1^k$. Obviously, this is the case if, and only if,

$$\operatorname{rank} \hat{\Xi}_1 = \operatorname{rank} \begin{pmatrix} \hat{\Xi}_1 & \hat{\Xi}_2 \end{pmatrix} . \tag{3.36}$$

Assuming that (3.36) holds, the system (3.35) is solvable if, and only if, it satisfies the augmented rank condition

$$\operatorname{rank} \hat{\Xi}_1 = \operatorname{rank} \begin{pmatrix} \hat{\Xi}_1 & \hat{\Xi}_2 & -\Theta_{12} \end{pmatrix} . \tag{3.37}$$

Assuming we have succeeded in constructing U_2, the next step is to seek yet another vector field U_3 such that $[U_1, U_3] \in T\iota(\mathcal{V}[\mathcal{R}_q])$ and $[U_2, U_3] \in T\iota(\mathcal{V}[\mathcal{R}_q])$. Now the components of both vectors $\hat{\zeta}_1$ and $\hat{\zeta}_2$ are regarded as given, and the components of $\hat{\zeta}_3$ are regarded as the unknowns of the system

$$\hat{\Xi}_1\hat{\zeta}_3 = \hat{\Xi}_3\hat{\zeta}_1 - \Theta_{13}\ , \qquad \hat{\Xi}_2\hat{\zeta}_3 = \hat{\Xi}_3\hat{\zeta}_2 - \Theta_{23}\ . \tag{3.38}$$

Now this system is not to restrict the components of both $\hat{\zeta}_1$ and $\hat{\zeta}_2$ any further; again, the conditions on the $\hat{\zeta}_i$ following from the condition (3.33) for the existence of integral distributions, following from Lemma 3.3.5, is taken care of by contracting Ξ_3 into $\hat{\Xi}_3$. This implies that now the rank condition

$$\operatorname{rank}\begin{pmatrix}\hat{\Xi}_1\\ \hat{\Xi}_2\end{pmatrix} = \operatorname{rank}\begin{pmatrix}\hat{\Xi}_1 & \hat{\Xi}_3 & 0\\ \hat{\Xi}_2 & 0 & \hat{\Xi}_3\end{pmatrix} \tag{3.39}$$

has to be satisfied. If it is, then for $1 \leq c \leq C = \dim \mathcal{V}'[\mathcal{R}_q] - \dim \mathcal{V}[\mathcal{R}_q]$ the system

$$\Theta^c_{13} - \hat{\Xi}^c_{3k}\zeta^k_1 + \hat{\Xi}^c_{1k}\zeta^k_3 = 0\ , \qquad \Theta^c_{23} - \hat{\Xi}^c_{3k}\zeta^k_2 + \hat{\Xi}^c_{2k}\zeta^k_3 = 0$$

is solvable if, and only if, the augmented rank condition

$$\operatorname{rank}\begin{pmatrix}\hat{\Xi}_1\\ \hat{\Xi}_2\end{pmatrix} = \operatorname{rank}\begin{pmatrix}\hat{\Xi}_1 & \hat{\Xi}_3 & 0 & -\Theta_{13}\\ \hat{\Xi}_2 & 0 & \hat{\Xi}_3 & -\Theta_{23}\end{pmatrix}$$

holds. Now we proceed by iteration. Given $j-1$ vector fields $U_1, U_2, \dots, U_{j-1}$ of the required form spanning an involutive subdistribution of $T\iota(\mathcal{V}[\mathcal{R}_1])$, we construct the next vector field U_i by solving the system

$$\begin{aligned} \hat{\Xi}_1\hat{\zeta}_j &= \hat{\Xi}_j\hat{\zeta}_1 - \Theta_{1j}\\ &\ \ \vdots\\ \hat{\Xi}_{j-1}\hat{\zeta}_j &= \hat{\Xi}_j\hat{\zeta}_{j-1} - \Theta_{j-1,j}\ . \end{aligned} \tag{3.40}$$

Again we consider only the components of the vector $\hat{\zeta}_j$ as unknowns, and the system (3.40) must not imply any further restrictions on the components of the vectors $\hat{\zeta}_i$ for $1 \leq i < j$. The corresponding rank condition is

$$\operatorname{rank}\begin{pmatrix}\hat{\Xi}_1\\ \hat{\Xi}_2\\ \vdots\\ \hat{\Xi}_{j-1}\end{pmatrix} = \operatorname{rank}\begin{pmatrix}\hat{\Xi}_1 & \hat{\Xi}_j & & & \\ \hat{\Xi}_2 & & \hat{\Xi}_j & & 0\\ \vdots & 0 & & \ddots & \\ \hat{\Xi}_{j-1} & & & & \hat{\Xi}_j\end{pmatrix}. \tag{3.41}$$

Assuming that it holds, the equations (3.40) are solvable and yield solutions for the components of $\hat{\zeta}_j$ if, and only if, it satisfies the augmented rank condition

$$\operatorname{rank}\begin{pmatrix}\hat{\Xi}_1\\ \hat{\Xi}_2\\ \vdots\\ \hat{\Xi}_{j-1}\end{pmatrix} = \operatorname{rank}\begin{pmatrix}\hat{\Xi}_1 & \hat{\Xi}_j & & & & -\Theta_{1j}\\ \hat{\Xi}_2 & & \hat{\Xi}_j & & 0 & -\Theta_{2j}\\ \vdots & 0 & & \ddots & & \vdots\\ \hat{\Xi}_{j-1} & & & & \hat{\Xi}_j & -\Theta_{j-1,j}\end{pmatrix}. \tag{3.42}$$

Remark 3.3.7. Again, if each equation of the local representation of the differential equation $\mathcal{R}_q$ is solvable for some different principal derivative and satisfies the assumptions of Proposition 3.1.19, we can use the local coordinates on $\mathcal{R}_q$, thus simplifying our calculations. According to Proposition 3.1.8, for the vector fields W_a we may choose the contact vector fields C^{μ}_{α} where $(\alpha, \mu) \in \mathcal{B}$. For the generators of the symbol $\mathcal{N}_q$ we may choose for $(\alpha, \mu) \notin \mathcal{B}$ the vector fields $\bar{Y}^{\alpha}_{\mu}$, given in Equation (3.15). And for a basis for the complement $\mathcal{H} \subseteq \mathcal{V}[\mathcal{R}_q]$, we can take the vector fields $\bar{X}_i$ given in Equation (3.8). The further procedure, solving first the structure equations, which now take the simple form (3.14′), to find the generators of $\mathcal{V}'(\mathcal{R}_q)$ and then solving Equation (3.29) for the coefficient functions ζ_j, is the same as in the case where the representation is not in solved form.

Another prerequisite of any step-by-step approach for the construction of integral distributions concerns the chosen local coordinates.

Remark 3.3.8. Whenever some kind of Cartan test is used, the notion of δ-regularity comes to the fore. Thus, to construct the respective generators U_i, given in Equations (3.22) or (3.27), through a step-by-step approach, we have to use δ-regular coordinates. A usual evasion maneuver in all approaches—using vector fields or, dually, an exterior system (like Hartley and Tucker [18])—whenever the choice of coordinates turns out to be δ-singular and, as a consequence, the step-by-step construction does not work out, is to introduce some linear combination of the generators which, in effect, means using some random transformation of the coefficients.

Now within the formal theory, δ-regularity is well understood in that a systematic method to analyze if the problem of δ-singular coordinates arises is available, and if it does arise, to introduce δ-regular coordinates using a deterministic procedure which avoids expensive random transformations. See Hausdorf, Mehdi and Seiler [20] for details.

The following theorem links the satisfaction of the rank conditions (3.41) and (3.42), and thus the solvability of the algebraic system (3.29) by the above described step-by-step process, with intrinsic properties of the differential equation $\mathcal{R}_q$ and its symbol $\mathcal{N}_q$.

Theorem 3.3.9. *Assume that δ-regular coordinates have been chosen for the differential equation $\mathcal{R}_q$. Then the rank condition (3.41) is satisfied for all $1 \leq j \leq n$ if, and only if, the symbol $\mathcal{N}_q$ is involutive. The augmented rank condition (3.42) holds for all $1 \leq j \leq n$ if, and only if, the differential equation $\mathcal{R}_q$ is involutive.*

The proof of Theorem 3.3.9 requires some technical considerations concerning the transformation of the matrices (3.41) and (3.42) into row echelon form, working out their contractions and analyzing the interrelation between these operations. These considerations are amplified in the following three subsections to keep the main argument clear. Since any regular differential equation can be transformed into a first-order system with a representation in the reduced Cartan normal form (2.27), there is no loss of generality if we assume the differential equation is represented that way; the advantage of this assumption is that the calculations in the oncoming considerations are simplified considerably. The proof of Theorem 3.3.9 is then given in Subsection 3.3.6.

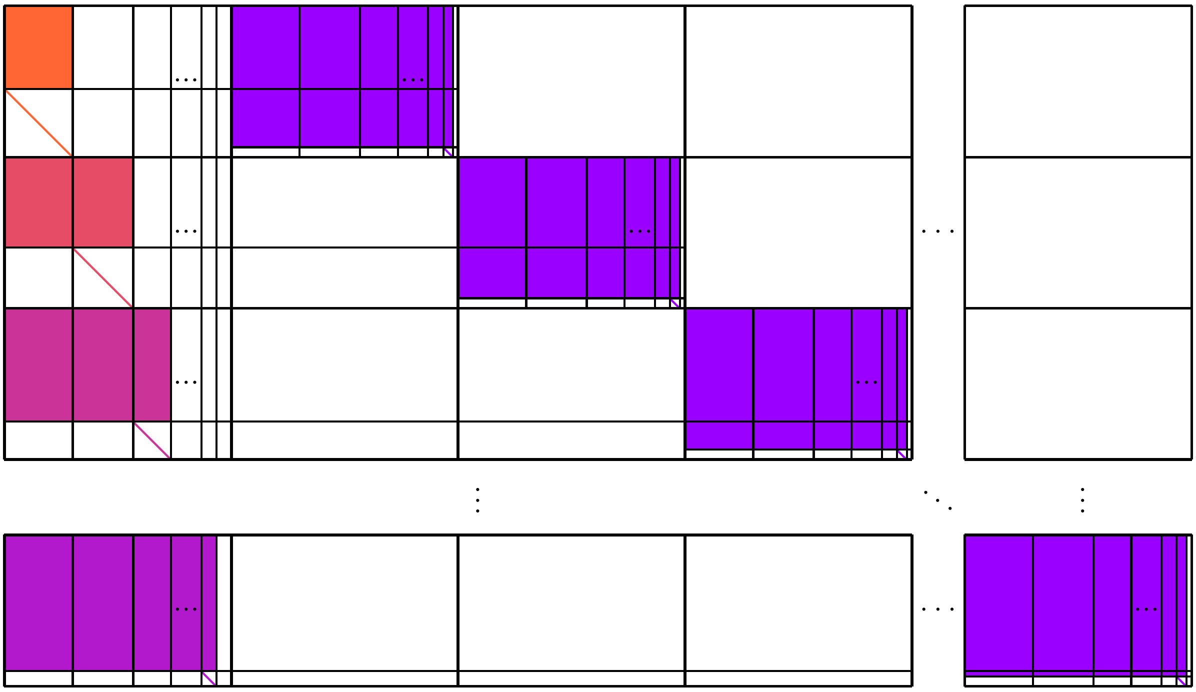


Figure 3.2: The complete matrix at step j as given in Definition 3.3.10.

3.3.3 Technical Details I: Structure Matrices

For the proof of the rank conditions (3.41) and (3.42), we need some more notation. First we introduce names for the matrices we need.

Definition 3.3.10. *Let Ξ^α_{ik} and Θ^α_{ij} be given as in (3.19) and (3.18). Then the* complete matrix (at the step j) *is defined as*

$$\begin{pmatrix} \Xi_1 & \Xi_j & & & \\ \Xi_2 & & \Xi_j & & 0 \\ \vdots & 0 & & \ddots & \\ \Xi_{j-1} & & & & \Xi_j \end{pmatrix} .$$

If we introduce the transpose of the vector $(\Theta^\alpha_{ij} \colon 1 \le i \le j-1)$ as another column to the right of the complete matrix, this yields the augmented complete matrix (at the step j)*.*

See Figure (3.2) for a sketch which shows the distribution of the potentially non-trivial entries. As we did in the sketch of the matrix Ξ_j, Figure (3.1), in the sketch of the complete matrix we mark those entries that possibly do not vanish as shaded areas, while diagonal lines denote negative unit blocks. The matrix (3.34) shows all the complete matrices stacked upon one another, augmented by an additional column for the entries of the vectors $-\Theta_{ij}$.

The complete matrix at the jth step is built from $(j-1)j$ blocks: the stack of $j-1$ matrices Ξ_i, $1 \le i \le j-1$, on the left, with each of the Ξ_i having another $j-1$ blocks to its right, the ith of which being Ξ_j and all the others being zero.

For easier reference, let, for $1 \leq i, k \leq j-1$, be $[i,k]$ the kth block right of Ξ_i. Then, for all $1 \leq i, k \leq j-1$, we have

$$[i,k] = \begin{cases} \Xi_j \colon i = k \\ 0_{m\times r} \colon i \neq k \end{cases} .$$

For convenience, we set $[i,0] := \Xi_i$. Matrix multiplication is denoted by a dot, while writing two matrices A and B next to one another and putting brackets around them, $[AB]$, means the matrix made by combining the entries of A and B into one matrix in the obvious way.

We call the block matrix $[[i,0][i,1]\cdots[i,j-1]]$ within the complete matrix at step j the *ith shift (of the complete matrix at step j)*. The block matrix within the complete matrix at step j which is made, for all $1 \leq i < j$, from the blocks $[i,k]$ being stacked upon one another is called the *kth shaft (of the complete matrix at step j)* and denoted by $[*,k]$.

Let, for $1 \leq g \leq h \leq n$,

$$[\Xi_i]^{g\ldots h} \tag{3.43}$$

denote the matrix that results from writing the block matrices $[\Xi_i]^g$, $[\Xi_i]^{g+1}$, $\ldots$, $[\Xi_i]^h$ (from left to right) next to each other.

Any matrix $[i,k]$, $1 \leq k \leq n$, contains as m rows and r columns. As a first step, we group the columns in all the $[i,k]$ into blocks the way we did for the matrices Ξ_i, labelling them, too, by the symbol fields Y_h^α (or u_h^α, the corresponding derivatives, or by the pairs $(\alpha,h) \notin \mathcal{B}$), namely according to increasing h into n blocks (empty for those h with $\alpha_1^{(h)} = 0$) and within each block according to increasing α (with $\beta_1^{(j)} + 1 \leq \alpha \leq m$). (This means, we order them ascendingly according to the term-over-position lift of the degree reverse lexicographic ranking applied to their labels $(\alpha,h) \notin \mathcal{B}$.) A second step is required to distinguish columns within the complete matrix at step j which are labelled by the same symbol field, say Y_h^α, but have entries in different blocks $[i,k] \neq [i,\ell]$. For this reason we label these columns of the complete matrix by the fields Y_{hk}^α and $Y_{h\ell}^\alpha$ (or their corresponding derivatives u_{hk}^α and $u_{h\ell}^\alpha$, or by the pairs (α,hk)). This is the analogue of labelling the columns in the matrix (3.34) in Remark 3.3.6. As a consequence, two columns within the complete matrix now may be labelled by, say, Y_{hk}^α and Y_{kh}^α, which in fact are equal since $u_{hk}^\alpha = u_{kh}^\alpha$. Exactly such pairs of columns are the ones to be added when making the matrices Ξ_i into the $\hat{\Xi}_i$ through contraction according to the identities (3.33).

For second-order derivatives u_{hk}^δ where k is the number of $[i,k]$, h denotes the number of the block of columns in $[i,k]$, and for $\delta = \beta_1^{(k)} + t$ the index t is the number of the column (its *label* being δ) within the hth block of columns.

Let $[i,k]^h$ be the hth block of columns in $[i,k]$, and let, for $1 \leq g \leq h \leq n$, in analogy to the shorthand (3.43),

$$[i,k]^{g\ldots h}$$

denote the matrix that results from writing the matrices $[i,k]^g$, $[i,k]^{g+1}$, $\ldots$, $[i,k]^h$ (from left to right) next to each other.

If M is any $a \times b$-matrix, then let ${}_c^d[M]$ be the matrix made from the *rows* with indices (that is, labels) c to d, $[M]_e^f$ the matrix made from the *columns* with indices e to f and ${}_c^d[M]_e^f$ the matrix made from the entries in the rows with indices c to d and in the columns with indices from e to f. If the columns of M are grouped into blocks and g denotes which blocks are meant, then we write ${}_1^c[M]^g$ (with g up right) to show that the first c upper rows are being selected, and we write ${}_d^b[M]_g$ (with g below right) to show that the last $b-d+1$ lower rows are being selected. For the block matrix made from M by selecting the rows indexed c to d within the block of columns labelled g, we write ${}_c^d[M]g$. This notation is redundant in that the position of g does not give new information, but in the calculations to come it increases readability.

The block $[i,k]^h$ stacked above the block $[i,k]_h$ yields a matrix of m rows and $\alpha_1^{(h)}$ columns; let it be denoted by $[i,k]h$. Let $[*,k]h$ denote all the $[i,k]h$ stacked upon one another for $1 \le i \le j-1$; this block is made up of the columns labelled from $u_{hk}^{\beta_1^{(h)}+1}$ to u_{hk}^m in the complete matrix at step j.

3.3.4 Technical Details II: Contractions

The next lemma shows which of the columns are to be contracted in step j. For a sketch, see Figure (3.3).

Lemma 3.3.11. *In step j, the number of pairs to be contracted is $\sum_{i=2}^{j-1}(i-1)\alpha_1^{(i)}$. For any column within a block*

$$[*,k]h, \qquad 1 \le k < h < j\ , \tag{3.44}$$

there is exactly one column which is not within any of the blocks (3.44) and which is to be contracted with the given column only. Any contraction concerns a pair of columns of which one is in one of the blocks (3.44) while the other column is not.

Proof. For $1 \le j \le n$ fix j. We show that for any given column within one of the blocks (3.44) there is a different unique column not within any of these blocks which is to be added to the given column only. Then we show that for any column not within such a block no contraction is possible.

In step j, for fixed h and k with $1 \le k < h < j$, consider the block $[*,k]h$. The $\alpha_1^{(h)}$ columns within $[*,k]h$ are labelled by

$$u_{hk}^{\beta_1^{(h)}+1}, u_{hk}^{\beta_1^{(h)}+2}, \ldots, u_{hk}^m\ .$$

Since $k < h$ according to assumption, we have $\beta_1^{(k)} \le \beta_1^{(h)}$, and so $\alpha_1^{(h)} \le \alpha_1^{(k)}$. There are $\alpha_1^{(k)}$ columns in the block $[*,h]k$, labelled by

$$u_{kh}^{\beta_1^{(k)}+1}, u_{kh}^{\beta_1^{(k)}+2}, \ldots, u_{kh}^m\ .$$

Since $\beta_1^{(k)} \le \beta_1^{(h)}$, among them are those labelled by

$$u_{kh}^{\beta_1^{(h)}+1}, u_{kh}^{\beta_1^{(h)}+2}, \ldots, u_{kh}^m\ .$$

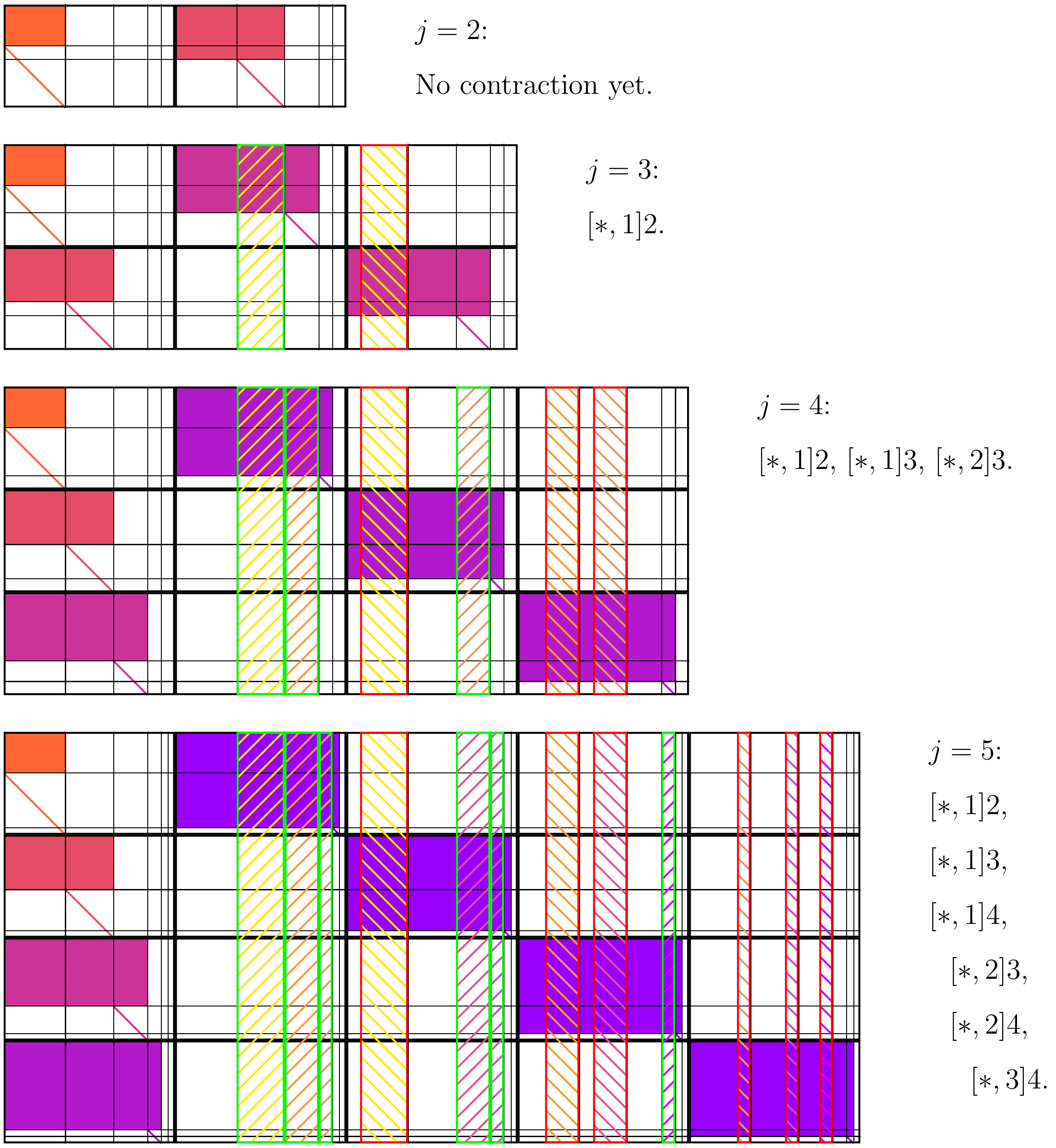


Figure 3.3: Contracting pairs of columns up to step $j = 5$: to each column in a block shown green-framed and hatched another column from one of the areas hatched the other way around is to be added. This is done blockwise: If $[*, h]k$ is green-framed and hatched, it is to be contracted with $[[*, k]h]^m_{\beta_1^{(k+1)}+1}$, shown red-framed and hatched the other way around in the same hue. For each step the blocks both green-framed and hatched, concerned by the contraction, are listed next to the matrix. There are $\sum_{i=1}^{j-2} i$ of them in step j. The contracted matrices arise through adding to the green-framed, hatched blocks their corresponding differently hatched counterparts and then cancelling those. See Figure 3.11 for a sketch of the contracted complete matrix in row echelon form in step $j = 5$.

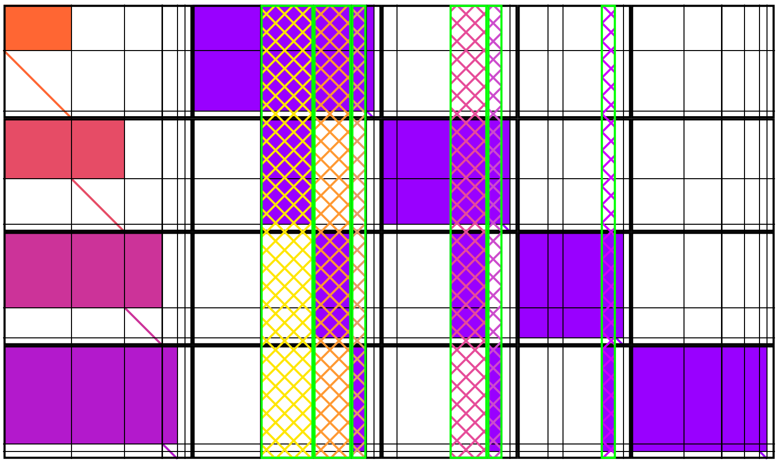


Figure 3.4: Complete matrix for step $j = 5$ after contraction: blocks of columns concerned by contraction are shown green-framed and hatched crosswise.

This means for any of the labels $u_{hk}^{\beta_1^{(h)}+t}$ with $1 \leq t \leq \alpha_1^{(h)}$, there is one matching label $u_{kh}^{\beta_1^{(h)}+t}$ for a column in $[*, h]k$. This column is unique because any other column is labelled by a second-order derivative different from $u_{hk}^{\beta_1^{(h)}+t} = u_{kh}^{\beta_1^{(h)}+t}$. But as the column labelled $u_{kh}^{\beta_1^{(h)}+t}$ is in $[*, h]k$ and $k < h$, it cannot be in $[*, h]k$ with $h < k$ as it had to if it were to be among the blocks in (3.44).

On the other hand, for a column which is not contained in one of the blocks (3.44) there is no other column outside the blocks (3.44) which is to be added to it: any such column is indexed by some $u_{hk}^{\beta_1^{(h)}+t}$; for $h = k$ the column is not contracted at all (because hk and kh do not yield two different labels), and for $h \neq k$ we have either $k < h$ which means $u_{hk}^{\beta_1^{(h)}+t}$ is contained in (3.44) while $u_{kh}^{\beta_1^{(h)}+t}$ is not, or we have $h < k$ in which case $u_{hk}^{\beta_1^{(h)}+t}$ is not contained in (3.44) while $u_{kh}^{\beta_1^{(h)}+t}$ is. For the same reason no two columns contained in the blocks (3.44) are to be contracted. This means that contraction of columns can be described blockwise: for $1 \leq k < h < j$, the block $[*, h]k$ is to be contracted with the block $[[*, k]h]^m_{\beta_1^{(k+1)}+1}$ in such a way that the column in $[[*, k]h]^m_{\beta_1^{(k+1)}+1}$ with index β is to be added to the column in $[*, h]k$ with that same index β.

The number of blocks like (3.44) which are concerned by contractions is $\sum_{i=1}^{j-2} i$ for step j. This is because for each u^δ and each k, h with $1 \leq k < h < j$, there are $\sum_{i=1}^{j-2} i$ second-order derivatives u_{hk}^δ so that $\sum_{i=1}^{j-2} i$ such derivatives u_{hk}^δ appear as labels in step j; since for each k and h the derivative u_{hk}^δ appears exactly once a block of columns $[*, k]h$, the number of such columns equals the number of these derivatives. $\square$

Now we give an explicit expression for the terms affected by contraction.

Lemma 3.3.12. *Contraction changes the entries of the complete matrix of step j for $1 \le k < i < j$ by the substitutions*

$$ {}^{\beta_1^{(j)}}_{1}[i,k]i \quad \leftarrow \quad {}^{\beta_1^{(j)}}_{1}[i,k]i + {}^{\beta_1^{(j)}}_{1}[[i,i]k]^m_{\beta_1^{(i)}+1} \tag{3.45} $$

and by cancelling the columns $[[*,h]k]^m_{\beta_1^{(h)}+1}$ for $1 \le k < h < j$. All other entries remain unchanged.

Proof. We just have to consider the special structure of the complete matrix at step j as given in Definition 3.3.10. Fix $1 < j \le n$. Let $[\,]$ denote a column that is cancelled. Then according to the proof of Lemma 3.3.11, for all $1 \le k < h < j$ contraction means to transform the blocks of columns in the following way:

$$ [*,k]h \quad \leftarrow \quad [*,k]h + [[*,h]k]^m_{\beta_1^{(h)}+1} \,, \tag{3.46a} $$

$$ [[*,h]k]^m_{\beta_1^{(h)}+1} \quad \leftarrow \quad [\,] \,. \tag{3.46b} $$

(We assume $k < h$ because we place two contracted blocks of columns in the complete matrix where the left block was and cancel the right one.) For $1 \le i \le j-1$ fix i and consider the effect of contraction on the row of block matrices

$$ [[i,1][i,2]\cdots[i,j-1]] \,. \tag{3.47} $$

For these rows, contraction (3.46a) becomes

$$ [i,k]h \quad \leftarrow \quad [i,k]h + [[i,h]k]^m_{\beta_1^{(h)}+1} \,. \tag{3.48} $$

Since $k < h$, for $i \le k$, according to the structure of the complete matrix, given in Definition 3.3.10, $[i,h]$ is a zero matrix. Thus for $i \le k$ the contraction (3.48) leaves $[i,k]h$ unchanged, and the only changes in lines (3.47) may appear in blocks $[i,k]h$ where $k < i$. We now specify where exactly and thus assume $k < i$. Non-trivial entries in (3.47) may only be found in $[i,i] = \Xi_j$. So contraction (3.48) reduces to

$$ [i,k]i \quad \leftarrow \quad [i,k]i + [[i,i]k]^m_{\beta_1^{(i)}+1} \,, \tag{3.49} $$

while for $h \ne i$ the $[i,k]h$ remain unchanged zero blocks.

According to (3.21), for $k < i$, the entries in the last $\alpha_1^{(j)}$ rows of $[i,i]$ are zero. So (3.49) further simplifies to

$$ {}^{\beta_1^{(j)}}_{1}[i,k]i \quad \leftarrow \quad {}^{\beta_1^{(j)}}_{1}[i,k]i + {}^{\beta_1^{(j)}}_{1}[[i,i]k]^m_{\beta_1^{(i)}+1} \,. \tag{3.50} $$

For fixed i, these are the only entries which are affected by contraction in that they may become non-trivial. To sum up: for $1 \le k < j$, in the hth block of columns in the kth shaft the rightmost columns (those with labels $\beta_1^{(k)} + 1 \le \beta \le m$) are to be deleted, if, and only if, $1 \le h < k$, and the hth block of columns in the kth shaft receives added columns if, and only if $k < h < j$. $\square$

Corollary 3.3.13. *For all $1 \leq i, k < j$, let $\widehat{[i,k]}$ be $[i,k]$ after contraction. Then for $k < i$ non-vanishing entries may appear in the blocks ${}_{1}^{\beta_1^{(j)}}\widehat{[i,k]}i$, and for $k = i$ in the block ${}_{1}^{\beta_1^{(j)}}\widehat{[i,i]}^{1\dots j}$ and on the main diagonal of ${}^{\;\;m}_{\beta_1^{(j)}+1}\widehat{[i,i]}_j$, but nowhere else in the ith shift of the contracted complete matrix (that is, right of Ξ_i).*

Proof. According to Lemma 3.3.12, for all $k < i$, we have

$$ {}_{1}^{\beta_1^{(j)}}\widehat{[i,k]}i = {}_{1}^{\beta_1^{(j)}}[[i,i]k]^{m}_{\beta_1^{(i)}+1} \;, $$

and according to the structure of ${}_{1}^{\beta_1^{(j)}}[i,i]k = [\Xi_j]^k$ as shown in Equation (3.20), it follows

$$ {}_{1}^{\beta_1^{(j)}}[[i,i]k]^{m}_{\beta_1^{(i)}+1} = \begin{pmatrix} -C^k_{\beta_1^{(i)}+1}(\phi^1_j) & -C^k_{\beta_1^{(i)}+2}(\phi^1_j) & \cdots & -C^k_m(\phi^1_j) \\ -C^k_{\beta_1^{(i)}+1}(\phi^2_j) & -C^k_{\beta_1^{(i)}+2}(\phi^2_j) & \cdots & -C^k_m(\phi^2_j) \\ \vdots & \vdots & \ddots & \vdots \\ -C^k_{\beta_1^{(i)}+1}(\phi^{\beta_1^{(j)}}_j) & -C^k_{\beta_1^{(i)}+2}(\phi^{\beta_1^{(j)}}_j) & \cdots & -C^k_m(\phi^{\beta_1^{(j)}}_j) \end{pmatrix} . \tag{3.51a} $$

For $k = i$ we have to specify the possibly non-vanishing entries in $\widehat{[i,i]}$. This block is made through contraction from $[i,i]$ according to substitution (3.46b) by cancelling for all $1 \leq h < i$ the columns in the blocks $[[i,i]h]^{m}_{\beta_1^{(i)}+1}$. This yields for all $1 \leq h < i$

$$ {}_{1}^{\beta_1^{(j)}}[\widehat{[i,i]}h] = \begin{pmatrix} -C^h_{\beta_1^{(h)}+1}(\phi^1_j) & -C^h_{\beta_1^{(h)}+2}(\phi^1_j) & \cdots & -C^h_{\beta_1^{(i)}}(\phi^1_j) \\ -C^h_{\beta_1^{(h)}+1}(\phi^2_j) & -C^h_{\beta_1^{(h)}+2}(\phi^2_j) & \cdots & -C^h_{\beta_1^{(i)}}(\phi^2_j) \\ \vdots & \vdots & \ddots & \vdots \\ -C^h_{\beta_1^{(h)}+1}(\phi^{\beta_1^{(j)}}_j) & -C^h_{\beta_1^{(h)}+2}(\phi^{\beta_1^{(j)}}_j) & \cdots & -C^h_{\beta_1^{(i)}}(\phi^{\beta_1^{(j)}}_j) \end{pmatrix} \tag{3.51b} $$

while for $i \leq h \leq j$, we have

$$ {}_{1}^{\beta_1^{(j)}}[\widehat{[i,i]}h] = {}_{1}^{\beta_1^{(j)}}[[i,i]h] \; . \tag{3.51c} $$

Since $\widehat{[i,i]}j = [i,i]j = [\Xi_j]j$, we have

$$ {}^{\;\;m}_{\beta_1^{(j)}+1}\widehat{[i,i]}_j = -\mathbb{1}_{\alpha_1^{(j)}} \; . \tag{3.51d} $$

All other entries remain zero throughout the process. □

Remark 3.3.14. There is the following relation between the negative unit blocks in the complete matrix and the blocks to be contracted: whenever the block $[*,k]h$ in the step $j-1$ has a negative unit block as its last rows, in step j the columns of $[*,k]h$ join the blocks (3.44) of columns to be contracted. (See Figure 3.3 for a sketch.) Since in step j, there are $j-1$ negative unit blocks $-\mathbb{1}_{\alpha_1^{(j)}}$, the number of pairs to be contracted is $\sum_{i=2}^{j-1}(i-1)\alpha_1^{(i)}$.

The unit blocks $-\mathbb{1}_{\alpha_1^{(j)}}$ within the complete matrix at step j comprehend the information about the cross relations among the entries of ζ_j and those of ζ_k for $1 \leq k \leq j-1$ given by Equation (3.33). The oncoming Corollary 3.3.23 gives the complete set of relations between all the coefficients ζ_k.

Lemma 3.3.15. *The columns of the complete matrix after contraction are ordered ascendingly with respect to the term-over-position lift of the degree reverse lexicographic ranking applied to their labels* (α, hk) *where* $(\alpha, h) \notin \mathcal{B}$.

Proof. For $1 \leq k < j$, the columns of the kth shaft of the complete matrix before contraction are labelled (from left to right)

$$\begin{aligned}
&u_{1k}^{\beta_1^{(1)}+1}, u_{1k}^{\beta_1^{(1)}+2}, \ldots, u_{1k}^{m},\\
&\quad u_{2k}^{\beta_1^{(2)}+1}, u_{2k}^{\beta_1^{(2)}+2}, \ldots, u_{2k}^{m},\\
&\qquad \ldots, u_{nk}^{\beta_1^{(n)}+1}, u_{nk}^{\beta_1^{(n)}+2}, \ldots, u_{nk}^{m}.
\end{aligned}$$

Here, for $1 \leq h \leq n$, the indices $u_{hk}^{\beta_1^{(h)}+1}$, $u_{hk}^{\beta_1^{(h)}+2}$, $\ldots$, u_{hk}^{m} label the hth block of columns in the kth shaft. (The columns of the kth shaft are ordered ascendingly with respect to the term-over-position lift of the degree reverse lexicographic ranking applied to their labels $(\alpha, h) \notin \mathcal{B}$.)

As mentioned at the end of the proof of Lemma 3.3.12, for $1 \leq k < j$, the columns of the kth shaft of the complete matrix after contraction (made of the blocks $\widehat{[i,k]}$ stacked upon one another) are labelled (from left to right)

$$\begin{aligned}
&u_{1k}^{\beta_1^{(1)}+1}, u_{1k}^{\beta_1^{(1)}+2}, \ldots, u_{1k}^{\beta_1^{(k)}},\\
&\quad u_{2k}^{\beta_1^{(2)}+1}, u_{2k}^{\beta_1^{(2)}+2}, \ldots, u_{2k}^{\beta_1^{(k)}},\\
&\qquad \ldots, u_{k-1,k}^{\beta_1^{(k-1)}+1}, u_{k-1,k}^{\beta_1^{(k-1)}+2}, \ldots, u_{k-1,k}^{\beta_1^{(k)}},\\
&\qquad\quad u_{kk}^{\beta_1^{(k)}+1}, u_{kk}^{\beta_1^{(k)}+2}, \ldots, u_{kk}^{m},\\
&\qquad\qquad u_{k+1,k}^{\beta_1^{(k+1)}+1}, u_{k+1,k}^{\beta_1^{(k+1)}+2}, \ldots, u_{k+1,k}^{m},\\
&\qquad\qquad\quad \ldots, u_{nk}^{\beta_1^{(n)}+1}, u_{nk}^{\beta_1^{(n)}+2}, \ldots, u_{nk}^{m}.
\end{aligned}$$

This order is according to the term-over-position lift of the degree reverse lexicographic ranking applied to the labels $(\alpha, hk) \notin \mathcal{B}$ where $(\alpha, h) \notin \mathcal{B}$. ☐

3.3.5 Technical Details III: Row Transformations

In order to prove the rank condition (3.41), we transform the matrices into row echelon form. Since each matrix Ξ_i contains a unit block, there is an obvious way to do this. We describe the procedure using the notation for subblocks, $[\Xi_i]^h$ and $[\Xi_i]_h$, introduced in Remark 3.1.28. As we shall see in Subsection 3.3.6 where we prove the existence theorem for integral distributions, the relevant entries in this row echelon form are the coefficients of the second-order derivatives u^δ_{hk} which appear in Lemma 2.5.8 and therefore their vanishing is equivalent to involution of the symbol $\mathcal{N}_1$.

Example 3.3.16. Before proving the general case in Subsection 3.3.6, we consider the case $i = 1$ and $j = 2$, that is, Equation (3.36), as an example to show the method while the calculations involved are comparatively short. (The proof for the general case is independent of this consideration.) Since Ξ_1 consists of a negative unit matrix of $\alpha_1^{(1)}$ rows with a $\beta_1^{(1)} \times \alpha_1^{(1)}$-matrix stacked upon it and only zeros for all other entries, we have $\operatorname{rank}(\Xi_1) = \alpha_1$. Next, we transform the matrix $(\Xi_1\ \Xi_2)$ into row echelon form using the special structure of the matrices Ξ_i as given in Equation (3.21); the blocks are replaced in this way:

$$[\Xi_1]^1 \leftarrow\ [\Xi_1]^1 + [\Xi_1]^1 \cdot [\Xi_1]_1\ , \tag{3.52a}$$

$${}^{\beta_1^{(1)}}_{1}[\Xi_2]^1 \leftarrow\ {}^{\beta_1^{(1)}}_{1}[\Xi_2]^1 + [\Xi_1]^1 \cdot {}^{m}_{\beta_1^{(1)}+1}[\Xi_2]_1\ , \tag{3.52b}$$

$${}^{\beta_1^{(1)}}_{1}[\Xi_2]^2 \leftarrow\ {}^{\beta_1^{(1)}}_{1}[\Xi_2]^2 + [\Xi_1]^1 \cdot {}^{m}_{\beta_1^{(1)}+1}[\Xi_2]_2\ . \tag{3.52c}$$

If, for the sake of simplicity, we use the same names for the changed blocks, then we have

$$[\Xi_1]^1 = 0_{\beta_1^{(1)} \times \alpha_1^{(1)}} \tag{3.53a}$$

$${}^{\beta_1^{(1)}}_{1}[\Xi_2]^1 = \left(-C^1_\delta(\phi^\alpha_2) + \sum_{\gamma=\beta_1^{(1)}+1}^{\beta_1^{(2)}} C^1_\gamma(\phi^\alpha_1) C^1_\delta(\phi^\gamma_2) \right)_{\substack{1 \le \alpha \le \beta_1^{(1)} \\ \beta_1^{(1)}+1 \le \delta \le m}}\ , \tag{3.53b}$$

$${}^{\beta_1^{(1)}}_{1}[\Xi_2]^2 = \left(C^1_\delta(\phi^\alpha_1) - C^2_\delta(\phi^\alpha_2) + \sum_{\gamma=\beta_1^{(1)}+1}^{\beta_1^{(2)}} C^1_\gamma(\phi^\alpha_1) C^2_\delta(\phi^\gamma_2) \right)_{\substack{1 \le \alpha \le \beta_1^{(1)} \\ \beta_1^{(2)}+1 \le \delta \le m}}\ . \tag{3.53c}$$

According to Lemma 2.5.8 all these entries vanish for an involutive system, since for $i = 1$, $j = 2$ the formula yields for any $1 \le \alpha \le \beta_1^{(1)}$ and $\beta_1^{(1)} + 1, \beta_1^{(2)} + 1 \le \delta \le m$:

$$0 = D_2\Phi^\alpha_1 - D_1\Phi^\alpha_2 + \sum_{\gamma=\beta_1^{(1)}+1}^{\beta_1^{(2)}} C^1_\gamma(\phi^\alpha_1) D_1\Phi^\gamma_2 \tag{2.30}$$

$$= C^{(1)}_1(\phi^\alpha_2) - C^{(1)}_2(\phi^\alpha_1) - \sum_{\gamma=\beta_1^{(1)}+1}^{\beta_1^{(2)}} C^1_\gamma(\phi^\alpha_1) C^{(1)}_1(\phi^\gamma_2) \tag{2.31}$$

$$- \sum_{\delta=\beta_1^{(1)}+1}^{m} u_{11}^{\delta} \left[-C_\delta^1(\phi_2^\alpha) + \sum_{\gamma=\beta_1^{(1)}+1}^{\beta_1^{(2)}} C_\gamma^1(\phi_1^\alpha) C_\delta^1(\phi_2^\gamma) \right] \tag{2.36}$$

$$- \sum_{\delta=\beta_1^{(2)}+1}^{m} u_{12}^{\delta} \left[C_\delta^1(\phi_1^\alpha) - C_\delta^2(\phi_2^\alpha) + \sum_{\gamma=\beta_1^{(1)}+1}^{\beta_1^{(2)}} C_\gamma^1(\phi_1^\alpha) C_\delta^2(\phi_2^\gamma) \right] . \tag{2.38}$$

For $i = 1$ and $j = 2$ all other lines vanish. The coefficients in Line (2.36) are the new entries of block matrix (3.53b) while those in Line (2.38) are the new entries of block matrix (3.53c). It follows that now the first $\beta_1^{(1)}$ rows of $(\Xi_1 \Xi_2)$ have become zero, leaving only the last $\alpha_1^{(1)}$ rows non-trivial; these last rows begin with the block $-\mathbb{1}_{\alpha_1^{(1)}}$, thus $\operatorname{rank}(\Xi_1 \Xi_2) = \alpha_1^{(1)} = \operatorname{rank} \Xi_1$. In this case ($i = 1$ and $j = 2$) contracting columns does not change the matrices, and it follows that $\operatorname{rank}(\hat{\Xi}_1 \hat{\Xi}_2) = \alpha_1^{(1)} = \operatorname{rank} \hat{\Xi}_1$.

So whenever the symbol is involutive, we can use Lemma 2.5.8 to construct the prolongations needed for the transformation into row echelon form. And whenever we can transform the matrix into row echelon form such that the rank condition is satisfied, we have found the coefficients to write any non-multiplicative prolongation of the system as a linear combination of multiplicative ones, which means the symbol is involutive. Thus we may conclude that the rank condition (3.36) holds if, and only if, no non-multiplicative prolongation $D_2 \Phi_1^a$ leads to an obstruction of involution.

The claim for the augmented condition (3.37) follows from the explicit expression (3.18) for the entries Θ_{12}^α. Performing the same computations as 3.52 with the augmented system (having an additional column for the entries Θ_{12}^α) yields as additional relevant entries exactly the integrability conditions arising from Lemma 2.5.8 applied for $i = 1$ and $j = 2$. They are collected in Line (2.31). Hence (3.37) holds if, and only if, no non-multiplicative prolongation $D_2 \Phi_1^a$ yields an integrability condition.

Remark 3.3.17. We could combine the three parts (3.52) of the substitution into

$$_{1}^{\beta_1^{(1)}}[\Xi_1[1,1]] \leftarrow {}_{1}^{\beta_1^{(1)}}[\Xi_1[1,1]] + [\Xi_1]^1 \cdot {}_{\beta_1^{(1)}+1}^{m}[\Xi_1[1,1]] \, . \tag{3.52$'$}$$

This would shorten the notation, especially for greater j. Figure (3.6) shows a sketch of this variant. Though both variants give the same result, in (3.52′) all zero columns of the complete matrix would, quite unnecessarily, be introduced into the calculation. Thus, for implementation or the calculation with concrete examples, variant (3.52) is to be preferred.

On the other hand, in the oncoming proof of the general case a concise notation is of greater advantage, and the symbolic computations involved do not become more complicated by these vanishing terms. Thus, variant (3.52′) is preferred there; sketches for the cases up to $j = 4$ are shown in Figures (3.6) to (3.9).

Another technical point concerns the order of the operations used: while transforming the complete matrix into row echelon form we only use row transformations; contraction

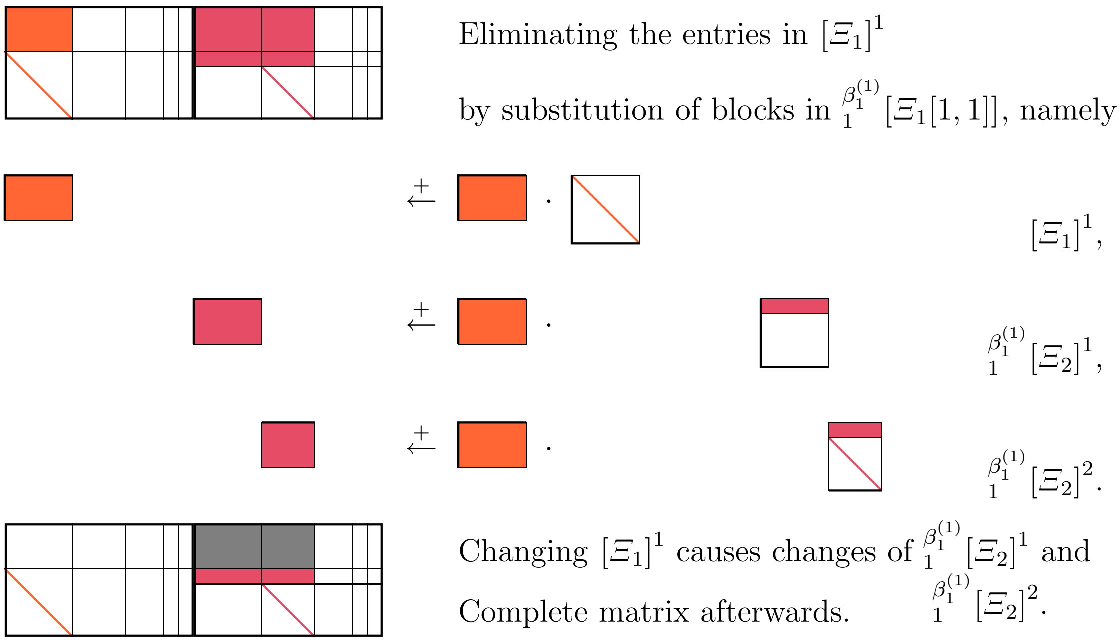


Figure 3.5: Turning the complete matrix in Definition 3.3.10 into row echelon form, step $j = 2$, by applying (3.52). Critical changed entries as given in (3.53b-c) are shaded gray. Complete matrix not contracted as no contraction is needed for $i = 1$ and $j = 2$.

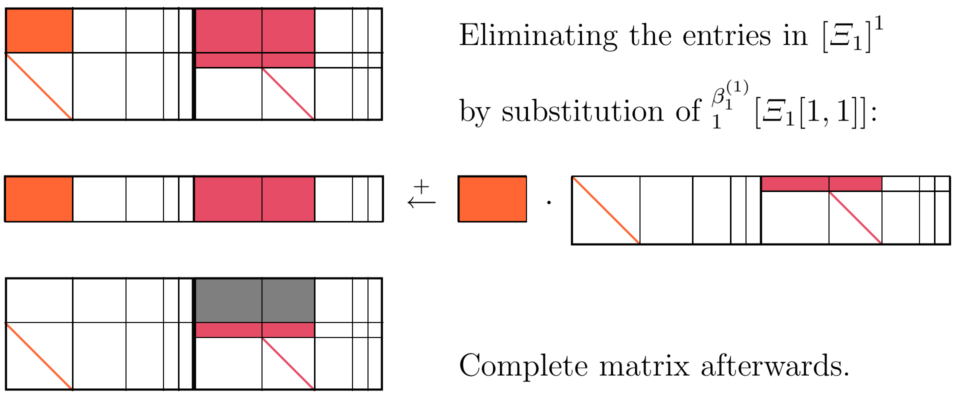


Figure 3.6: Turning the complete matrix in Definition 3.3.10 into row echelon form, step $j = 2$, by applying (3.52′). Critical changed entries as given in (3.53b-c) are shaded gray. Complete matrix not contracted as no contraction is needed for $i = 1$ and $j = 2$.

of columns involves adding pairs of columns. Both operations can be described by multiplying the complete matrix by elementary matrices, and the associative law guaranties that the outcome does not depend on the order of these operations. (Contraction also involves projections described by leaving out one column of each pair after adding it to the other; as addition comes before projection it does not matter whether we first transform into row echelon form and then add and project or the other way around.) For easier reference, we formulate this as a lemma.

Lemma 3.3.18. *Transforming the complete matrix into contracted row echelon form yields the same result independent of the order of operations—whether the matrix is first transformed into row echelon form and then contracted or the other way around.*

Proof. The proof is elementary. Row transformations correspond to multiplicating the complete matrix with elementary matrices from the left, while contracting columns corresponds to multiplications with elementary matrices (for the addition of two columns) and the unit matrix with one column left out (for discarding a column) from the right, and matrix multiplication is associative. □

As one might expect from the above considerations for $i = 1$ and $j = 2$, the analysis of the rank condition (3.41) for general $1 \leq i < j \leq n$ requires the non-multiplicative prolongations $D_j\Phi^\alpha_1, D_j\Phi^\alpha_2, \dots, D_j\Phi^\alpha_{j-1}$. It follows trivially from the block form (3.21) of the matrices Ξ_j that the rank of the matrix on the left sides of (3.41) is $\sum_{k=1}^{j-1}\alpha_1^{(k)}$.

For the general case, we follow the same steps as in the case $i = 1$ and $j = 2$. The transformation of the matrix on the right hand side of (3.41) can be described using block matrices, and the resulting matrix in row echelon form has as its entries in the rows where no unit block appears the coefficients of the second-order derivatives in Lemma 2.5.8. Thus we may conclude again that satisfaction of (3.41) is equivalent to the fact that in the non-multiplicative prolongations $D_j\Phi^\alpha_i$, where $1 \leq i < j$, no obstructions to involution arise. In the case of the augmented conditions (3.42), it follows again from the explicit expression (3.18) for the entries Θ^α_{ij} that the additional relevant entries are identical with the potential integrability conditions produced by the non-multiplicative prolongations $D_j\Phi^\alpha_i$.

At this point it becomes apparent why we had to introduce the contracted matrices $\hat{\Xi}_i$. As all functions are assumed to be smooth, partial derivatives commute: $u^\alpha_{ij} = u^\alpha_{ji}$. In Lemma 2.5.8 each obstruction to involution corresponding to these partial derivatives actually consists of two parts: one arises as coefficient of u^α_{ij}, the other one as coefficient of u^α_{ji}. While this decomposition does not show in Lemma 2.5.8 because both derivatives are collected into one term, the two parts appear in different columns of the matrices Ξ_i and the rank condition (3.41) will not hold in general, if we replace the contracted matrices $\hat{\Xi}_i$ by the original matrices Ξ_i (see Example 3.3.29). The effect of the contraction is to combine the two parts in order to obtain the right rank.

Example 3.3.19. Consider the system in Cartan normal form as given in Example 2.5.4. Then the dimension of the symbol is 6 and, using $*$ as a shorthand for all terms of the

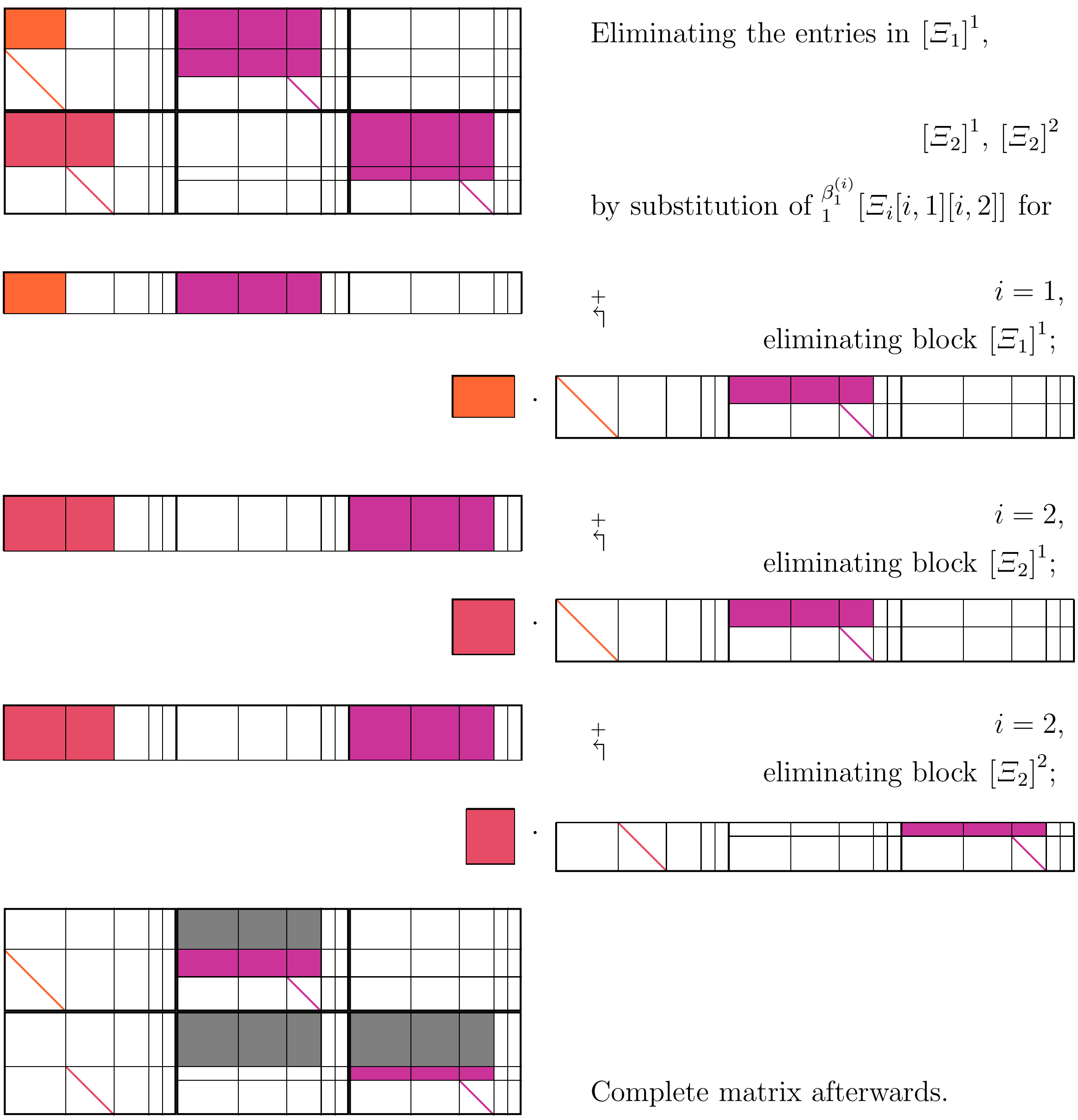


Figure 3.7: Turning the complete matrix in Definition 3.3.10 into row echelon form, step $j = 3$. Critical changed entries are shaded gray. Complete matrix not yet contracted.

form $-C^h_{\beta_1^{(h)}+\gamma}(\phi_i^\alpha)$, we get for $j = 2$:

$$\begin{pmatrix} \Xi_1 & [1,1] \end{pmatrix} = \begin{pmatrix} \Xi_1 & \Xi_2 \end{pmatrix} = \begin{pmatrix} * \; * \; * & 0\;0 & 0 & * \; * \; * & * \; * & 0 \\ * \; * \; * & 0\;0 & 0 & * \; * \; * & * \; * & 0 \\ -1\,0\,0 & 0\;0 & 0 & * \; * \; * & * \; * & 0 \\ 0-1\;0 & 0\;0 & 0 & 0\;0\;0 & -1\,0 & 0 \\ 0\;0-1 & 0\;0 & 0 & 0\;0\;0 & 0-1 & 0 \end{pmatrix}.$$

The first six columns are labelled from left to right by u^3_x, u^4_x, u^5_x, u^4_y, u^5_y and u^5_z; the following six columns are labelled u^3_{xx}, u^4_{xx}, u^5_{xx}, u^4_{yx}, u^5_{yx} and u^5_{zx}.

In order to transform the complete matrix for $j = 2$ into row echelon form, we transform the upper $\beta_1^{(1)} = 2$ rows of the complete matrix. These are rows within Ξ_1 and next to it, in the block matrix $[1,1]$. As the necessary substitutions for the entries in the block matrix Ξ_1 we have

$$[\Xi_1]^1 \leftarrow \; [\Xi_1]^1 + [\Xi_1]^1 \cdot [\Xi_1]_1 \, , \tag{3.54a}$$

according to Line (3.52a). According to Lines (3.52b, 3.52c) the necessary substitutions for the entries in the block matrix $[1,1] = \Xi_2$ are

$$\begin{smallmatrix}\beta_1^{(1)}\\1\end{smallmatrix}[1,1]^k \leftarrow \; \begin{smallmatrix}\beta_1^{(1)}\\1\end{smallmatrix}[1,1]^k + [\Xi_1]^1 \cdot {}^{m}_{\beta_1^{(1)}+1}[1,1]_k \, , \tag{3.54b}$$

where $1 \le k \le j$, that is, $k = 1$ and $k = 2$. The substitution (3.54a) changes the entries in those columns of the complete matrix which are labelled from u^3_x to u^5_x, while the substitution (3.54b) changes the entries in the columns of the complete matrix which are labelled from u^3_{xx} to u^5_{yx}. All the other entries in the first $\beta_1^{(1)}$ rows remain unchanged. Note that for $j = 2$ there are no contractions, as we already saw in Example 3.3.16. Therefore the transformed complete matrix for $j = 2$ looks like this:

$$\begin{pmatrix} 0\;0\;0 & 0\;0 & 0 & \Sigma\,\Sigma\,\Sigma & \Sigma\,\Sigma & 0 \\ 0\;0\;0 & 0\;0 & 0 & \Sigma\,\Sigma\,\Sigma & \Sigma\,\Sigma & 0 \\ -1\,0\,0 & 0\;0 & 0 & * \; * \; * & * \; * & 0 \\ 0-1\;0 & 0\;0 & 0 & 0\;0\;0 & -1\,0 & 0 \\ 0\;0-1 & 0\;0 & 0 & 0\;0\;0 & 0-1 & 0 \end{pmatrix}.$$

In the above sketch, unchanged entries of the form $-C^h_{\beta_1^{(h)}+\gamma}(\phi_i^\alpha)$ appear as stars again, while changed entries in the blocks $\begin{smallmatrix}\beta_1^{(1)}\\1\end{smallmatrix}[1,1]^1$ and $\begin{smallmatrix}\beta_1^{(1)}\\1\end{smallmatrix}[1,1]^2$ are represented by a summation sign each as they equal the matrix entries in Equations (3.53b, 3.53c), which are sums. See Example 3.3.16 for the exact calculations for the general case, where still $j = 2$ but the numbers of rows and columns are arbitrary. The rank condition for $j = 2$ is satisfied if all the entries denoted Σ vanish. That they do so indeed for an involutive system follows from Lemma 2.5.8 and was shown in Example 3.3.16, too. Figure 3.5 shows a sketch of these operations and of the complete matrix for $j = 2$ before and after.

Now for the next step, where $j = 3$. Since $\dim \mathcal{X} = n = 3$, too, the algorithm already terminates with this step. The complete matrix now looks like this:

$$
\begin{pmatrix} \Xi_1 & [1,1] & [1,2] \\ \Xi_2 & [2,1] & [2,2] \end{pmatrix} = \begin{pmatrix} \Xi_1 & \Xi_3 & 0_{5\times 6} \\ \Xi_2 & 0_{5\times 6} & \Xi_3 \end{pmatrix}
$$
$$
= \left(\begin{array}{cccccc|cccccc|cccccc}
* & * & * & 0 & 0 & 0 & * & * & * & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & 0 & 0 & 0 & * & * & * & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & * & * & * & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & * & * & * & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * & * & * \\
* & * & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * & * & * \\
* & * & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * & * & * \\
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * & * & * \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1
\end{array}\right) .
$$

The first six columns are labelled from left to right by u^3_x, u^4_x, u^5_x, u^4_y, u^5_y and u^5_z;
the next six columns are labelled u^3_{xx}, u^4_{xx}, u^5_{xx}, u^4_{yx}, u^5_{yx} and u^5_{zx};
and the rightmost six columns by u^3_{xy}, u^4_{xy}, u^5_{xy}, u^4_{yy}, u^5_{yy} and u^5_{zy}.

In order to transform the complete matrix for $j = 3$ into row echelon form, we use the negative unit block $[\Xi_1]_1$ to first transform the upper $\beta_1^{(1)} = 2$ rows of the complete matrix (the first shift, $i = 1$). These are rows within and next to Ξ_1. The necessary substitutions are:

$$
[\Xi_1]^1 \leftarrow \ [\Xi_1]^1 + [\Xi_1]^1 \cdot [\Xi_1]_1 \ , \tag{3.55a}
$$

and, for $1 \le k \le j$:

$$
{}_{1}^{\beta_1^{(1)}}[1,1]^k \leftarrow \ {}_{1}^{\beta_1^{(1)}}[1,1]^k + [\Xi_1]^1 \cdot {}^{m}_{\beta_1^{(1)}+1}[1,1]_k \ . \tag{3.55b}
$$

(Actually, as here $j = n = 3$, the indices k in substitution (3.55b) may as well be left out. In fact, this may be done in general, simplifying the notation; but, since for $j < k \le n$ the columns have only zero entries, this would increase the number of calculations unnecessarily, see Remark 3.3.17.) The substitution (3.55a) changes the entries in those columns of the complete matrix which are labelled from u^3_x to u^5_x, while the substitution (3.55b) changes the entries in those columns of the complete matrix which are labelled from u^3_{xx} to u^5_{zx}. All other entries in the first $\beta_1^{(1)}$ rows remain unchanged.

Next we consider the lower $m = 5$ rows in the complete matrix (the second shift, $i = 2$). The first three of those are to be eliminated. They are rows within and next to Ξ_2, labelled from 1 to $\beta_1^{(2)} = 3$. The necessary substitutions are:

$$
[\Xi_2]^1 \leftarrow \ [\Xi_2]^1 + [\Xi_2]^1 \cdot [\Xi_1]_1 \ , \tag{3.56a}
$$
$$
[\Xi_2]^2 \leftarrow \ [\Xi_2]^2 + [\Xi_2]^2 \cdot [\Xi_2]_2 \ , \tag{3.56b}
$$
$$
{}_{1}^{\beta_1^{(2)}}[2,1]^k \leftarrow \ {}_{1}^{\beta_1^{(2)}}[2,1]^k + [\Xi_2]^1 \cdot {}^{m}_{\beta_1^{(1)}+1}[1,1]_k \ , \tag{3.56c}
$$
$$
{}_{1}^{\beta_1^{(2)}}[2,2]^k \leftarrow \ {}_{1}^{\beta_1^{(2)}}[2,2]^k + [\Xi_2]^2 \cdot {}^{m}_{\beta_1^{(2)}+1}[2,2]_k \ , \tag{3.56d}
$$

where again $1 \le k \le j$. The substitution (3.56a) changes the entries in those columns of the complete matrix which are labelled from u^3_x to u^5_x, the substitution (3.56b) those in the columns which are labelled u^4_y and u^5_y, the substitution (3.56c) those in the columns which are labelled from u^3_{xx} to u^5_{zx} and the substitution (3.56d) changes those in the columns which are labelled from u^3_{xy} to u^5_{zy}. All the other entries in the last five rows remain unchanged. The complete matrix in after these row transformations is:

$$\begin{array}{c}\left(\begin{array}{ccc ccc cccccc cccccc}
0&0&0&0&0&0&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&0&0&0&0&0&0\\
0&0&0&0&0&0&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&0&0&0&0&0&0\\
-1&0&0&0&0&0&*&*&*&*&*&*&0&0&0&0&0&0\\
0&-1&0&0&0&0&*&*&*&*&*&*&0&0&0&0&0&0\\
0&0&-1&0&0&0&0&0&0&0&0&-1&0&0&0&0&0&0\\
0&0&0&0&0&0&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma\\
0&0&0&0&0&0&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma\\
0&0&0&0&0&0&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma\\
0&0&0&-1&0&0&0&0&0&0&0&0&*&*&*&*&*&*\\
0&0&0&0&-1&0&0&0&0&0&0&0&0&0&0&0&0&-1
\end{array}\right)\\
\begin{array}{ccc ccc cccccc cccccc}
&&&&&&&&&\Uparrow&\uparrow&&&&&\Uparrow&\uparrow&
\end{array}\end{array}.$$

Again, unchanged entries of the form $-C^h_{\beta_1^{(h)}+\gamma}(\phi^\alpha_i)$ appear as stars, while changed entries are represented by a summation sign each, on the ground that they equal the sums in square brackets in Lemma 2.5.8 as we shall see when proving the existence theorem for integral distributions in Subsection 3.3.6. See Figure 3.7 for a sketch of the complete matrix before and after these operations.

Both operations, elementary row transformations and contraction as combining two columns into one, commute as the elementary Lemma 3.3.18 says, so their order does not change the result.

There are two occurrences where labels for columns appear twice, so two contractions are in order: the columns marked with double arrows, labelled u^4_{yx} (which is the tenth from the left) and u^4_{xy} (the fourteenth from the left), are to be added, and so are the columns marked with single arrows, labelled u^5_{yx} (the eleventh column) and u^5_{xy} (the fifteenth one). This gives the contracted complete matrix of $2m = 10$ rows and 16 columns, that is:

$$\begin{aligned}\begin{pmatrix}\Xi_1 & [1,1] & [1,2]\\ \Xi_2 & [2,1] & [2,2]\end{pmatrix} &\rightsquigarrow \begin{pmatrix}\hat{\Xi}_1 & \hat{\Xi}_3 & 0_{5\times 4}\\ \hat{\Xi}_2 & \widehat{[2,1]} & \hat{\Xi}_3\end{pmatrix}\\
&= \left(\begin{array}{ccc ccc cccccc cccc}
0&0&0&0&0&0&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&0&0&0&0\\
0&0&0&0&0&0&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma&0&0&0&0\\
-1&0&0&0&0&0&*&*&*&*&*&*&0&0&0&0\\
0&-1&0&0&0&0&*&*&*&*&*&*&0&0&0&0\\
0&0&-1&0&0&0&0&0&0&0&0&-1&0&0&0&0\\
0&0&0&0&0&0&\Sigma&\Sigma&\Sigma&\Sigma'&\Sigma'&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma\\
0&0&0&0&0&0&\Sigma&\Sigma&\Sigma&\Sigma'&\Sigma'&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma\\
0&0&0&0&0&0&\Sigma&\Sigma&\Sigma&\Sigma'&\Sigma'&\Sigma&\Sigma&\Sigma&\Sigma&\Sigma\\
0&0&0&-1&0&0&0&0&0&*&*&0&*&*&*&*\\
0&0&0&0&-1&0&0&0&0&0&0&0&0&0&0&-1
\end{array}\right).\end{aligned}$$

We have again, for convenience, kept the notation for matrices which have only been changed into row echelon form, while the hats above the blocks indicate that the complete matrix has been changed through contraction. The entries which have changed through contraction and which are of importance for the rank are primed. The rank condition for $j = 3$ is satisfied if all the entries denoted Σ or Σ' vanish. That they do so indeed for an involutive system follows from Lemma 2.5.8 as these terms appear there in the square brackets. As an explicit example on how, we consider the new tenth column (of the contracted matrix in row echelon form) which is labelled by u^4_{xy}: its entries are, from top to bottom:

$-C_4^2(\phi_3^1) + C_3^1(\phi_1^1)C_4^2(\phi_3^3) + C_4^1(\phi_1^1)C_4^2(\phi_3^4),$
$-C_4^2(\phi_3^2) + C_3^1(\phi_1^2)C_4^2(\phi_3^3) + C_4^1(\phi_1^2)C_4^2(\phi_3^4),$
$*,$
$*,$
$0,$
$-C_4^1(\phi_3^1) + C_4^2(\phi_2^1)C_4^1(\phi_3^4) + C_3^1(\phi_2^1)C_4^2(\phi_3^3) + C_4^1(\phi_2^1)C_4^2(\phi_3^4),$
$-C_4^1(\phi_3^2) + C_4^2(\phi_2^2)C_4^1(\phi_3^4) + C_3^1(\phi_2^2)C_4^2(\phi_3^3) + C_4^1(\phi_2^2)C_4^2(\phi_3^4),$
$-C_4^1(\phi_3^3) + C_4^2(\phi_2^3)C_4^1(\phi_3^4) + C_3^1(\phi_2^3)C_4^2(\phi_3^3) + C_4^1(\phi_2^3)C_4^2(\phi_3^4),$
$*,$
$0.$

The first $m = 5$ entries belong to the first shift: they stand in the complete matrix right of Ξ_1. For $i = 1$ and $j = 3$ the coefficient for the term u^4_{xy} (in this example we have $\delta = 4$) according to Lemma 2.5.8 is to be found in Line (2.34b) where for all α with $(\alpha, i) \in \mathcal{B}$ (here these are $\alpha = 1, 2$) we have

$$-u^4_{12}[(-C_4^2(\phi_3^\alpha) + C_3^1(\phi_1^\alpha)C_4^2(\phi_3^3) + (C_4^1(\phi_1^\alpha)C_4^2(\phi_3^4))] \ .$$

So the terms Σ among the first five entries vanish. The last $m = 5$ entries belong to the second shift: they stand in the complete matrix right of Ξ_2. For $i = 2$ and $j = 3$ the coefficient for the term u^4_{xy} (still, $\delta = 4$) according to Lemma 2.5.8, Line (2.33b), for all α with $(\alpha, i) \in \mathcal{B}$ (in this case $\alpha = 1, 2, 3$) is

$$-u^4_{12}[(-C_4^1(\phi_3^\alpha) + C_4^2(\phi_2^\alpha)C_4^1(\phi_3^4)) + (C_3^1(\phi_2^\alpha)C_4^2(\phi_3^3) + C_4^1(\phi_2^\alpha)C_4^2(\phi_3^4))] \ .$$

So the terms Σ' among the last five entries vanish. The sums for the terms Σ are shorter than those for the terms Σ' because contraction only adds zeros to them.

To prove the rank conditions (3.41) and (3.42) and thus the existence theorem for integral distributions, we now turn these considerations into a general lemma concerning the calculation with block matrices. Though it appears technical, it uses only linear algebra and is very simple. We give it here for the sake of completeness and to clarify notation used later in Subsection 3.3.6.

Lemma 3.3.20. *Let A, R, S, T be natural numbers or zero. Let $\mathbb{1}_S$ denote the $S \times S$ unit matrix. Consider the matrices*

$$\begin{array}{ll} \big(a_{\alpha r}\colon 1 \le \alpha \le A, 1 \le r \le R\big)\,, & \big(b_{\alpha s}\colon 1 \le \alpha \le A, 1 \le s \le S\big)\,, \\ \big(c_{\alpha t}\colon 1 \le \alpha \le A, 1 \le t \le T\big)\,, & \big(d_{st}\colon 1 \le s \le S, 1 \le t \le T\big)\,, \\ 0_{S\times R}\ , & -\,\mathbb{1}_S \end{array}$$

and the matrix

$$\begin{pmatrix} a_{\alpha r} & b_{\alpha s} & c_{\alpha t} \\ 0_{S\times R} & -\mathbb{1}_S & d_{st} \end{pmatrix} =: \ (g_{ij}\colon 1 \le i \le A+S, 1 \le j \le R+S+T) \tag{3.57}$$

built from these blocks. (Here the index r is just some index and not supposed to be the dimension of any symbol.) Then the substitution

$$\begin{pmatrix} a_{\alpha r} & b_{\alpha s} & c_{\alpha t} \end{pmatrix} \leftarrow \begin{pmatrix} a_{\alpha r} & b_{\alpha s} & c_{\alpha t} \end{pmatrix} + (b_{\alpha s}) \cdot \begin{pmatrix} 0_{S\times R} & -\mathbb{1}_S & d_{st} \end{pmatrix} \tag{3.58}$$

transforms the matrix (g_{ij}) *into*

$$\begin{pmatrix} a_{\alpha r} & 0_{A\times S} & c_{\alpha t}+\sum_{s=1}^{S} b_{\alpha s} d_{st} \\ 0_{S\times R} & -\mathbb{1}_S & d_{st} \end{pmatrix} . \tag{3.59}$$

Proof. The simple matrix calculation

$$(b_{\alpha s})\cdot\begin{pmatrix}0_{S\times R} & -\mathbb{1}_S & d_{st}\end{pmatrix} = \begin{pmatrix}0_{S\times R} & -b_{\alpha s} & b_{\alpha s}\cdot d_{st}\end{pmatrix} = \begin{pmatrix}0_{S\times R} & -b_{\alpha s} & \sum_{s=1}^{S} b_{\alpha s} d_{st}\end{pmatrix} \tag{3.60}$$

yields the result. □

We are going to use this lemma to eliminate entries in matrices built from blocks of the type (3.59) or, similarly, of the type

$$\begin{pmatrix} 0_{S\times R} & -\mathbb{1}_S & d_{st} \\ a_{\alpha r} & b_{\alpha s} & c_{\alpha t} \end{pmatrix} =: (g_{ij}\colon 1\le i\le A+S, 1\le j\le R+S+T) . \tag{3.61}$$

The matrix calculation in the lemma emulates the following obvious way to eliminate the entries $b_{\alpha s}$ for all $1\le\alpha\le A$ and $1\le s\le S$: Consider the entry $b_{\alpha' s'}$. Then $g_{A+s',s'}=-1$, and adding $b_{\alpha' s'}$ times the row $(g_{A+s',j}\colon 1\le j\le R+S+T)$ to the row $(g_{\alpha' j}\colon 1\le j\le R+S+T)$ eliminates $b_{\alpha' s'}$ and changes, for all $1\le t\le T$, the entries $c_{\alpha' t}$ into $c_{\alpha' t}+b_{\alpha' s'}d_{s't}$. Doing this for all $1\le s\le S$ eliminates all the $b_{\alpha' s}$ in the row $(b_{\alpha' s}\colon 1\le s\le S)$ and changes, for all $1\le t\le T$, the entries $c_{\alpha' t}$ into $c_{\alpha' t}+\sum_{s=1}^{S} b_{\alpha' s}d_{st}$. Doing this for all $1\le\alpha'\le A$ makes the matrix $\begin{pmatrix}a_{\alpha r} & b_{\alpha s} & c_{\alpha t}\end{pmatrix}$ into the same matrix,

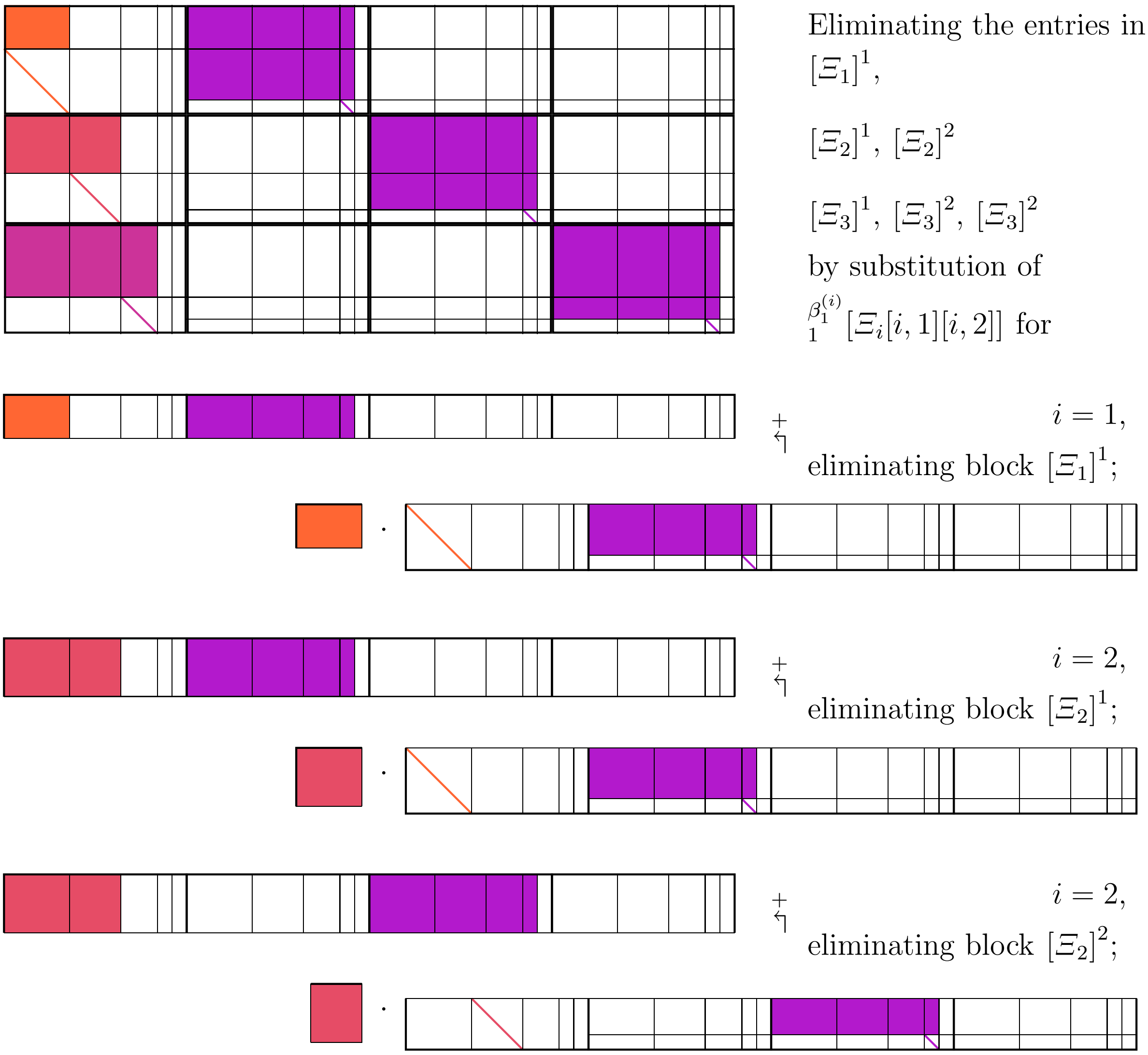
$$\begin{pmatrix}a_{\alpha r} & 0_{A\times S} & c_{\alpha t}+\sum_{s=1}^{S} b_{\alpha s}d_{st}\end{pmatrix} , \tag{3.62}$$

as in the lemma. We give this wordy explanation because it is exactly this kind of row transformations that we use in the proof of the existence theorem for integral distributions and which we reduce to block matrix calculations of the kind described here.

3.3.6 The Proof of the Existence Theorem for Integral Distributions

With the technical means which we collected in the last subsections now at hand, we can tackle the proof of Theorem 3.3.9, the existence theorem for integral distributions. It is in principle by straightforward matrix calculation and involves a tedious distinction of several cases and subcases, since we have to compare the entries in the augmented complete matrix for step j after turning it into row echelon form and contracting its columns with the integrability conditions and the obstructions to involution as they are given in Lemma 2.5.8 for an arbitrary shift i, where $1\le i<j$, and all the shafts of the complete matrix.

Since for a differential equation with an involutive symbol the obstructions to involution vanish and for an involutive differential equation the integrability conditions vanish, too, it follows that the augmented rank condition, stated in Equation (3.42), is equivalent to the differential equation being involutive, which in turn is the case if, and only if, the



(to be continued in Figure 3.9)

Figure 3.8: Turning the complete matrix in Definition 3.3.10 into row echelon form, step $j = 4$ for $i = 1, 2$. Complete matrix not yet contracted.

(continued from Figure 3.8)

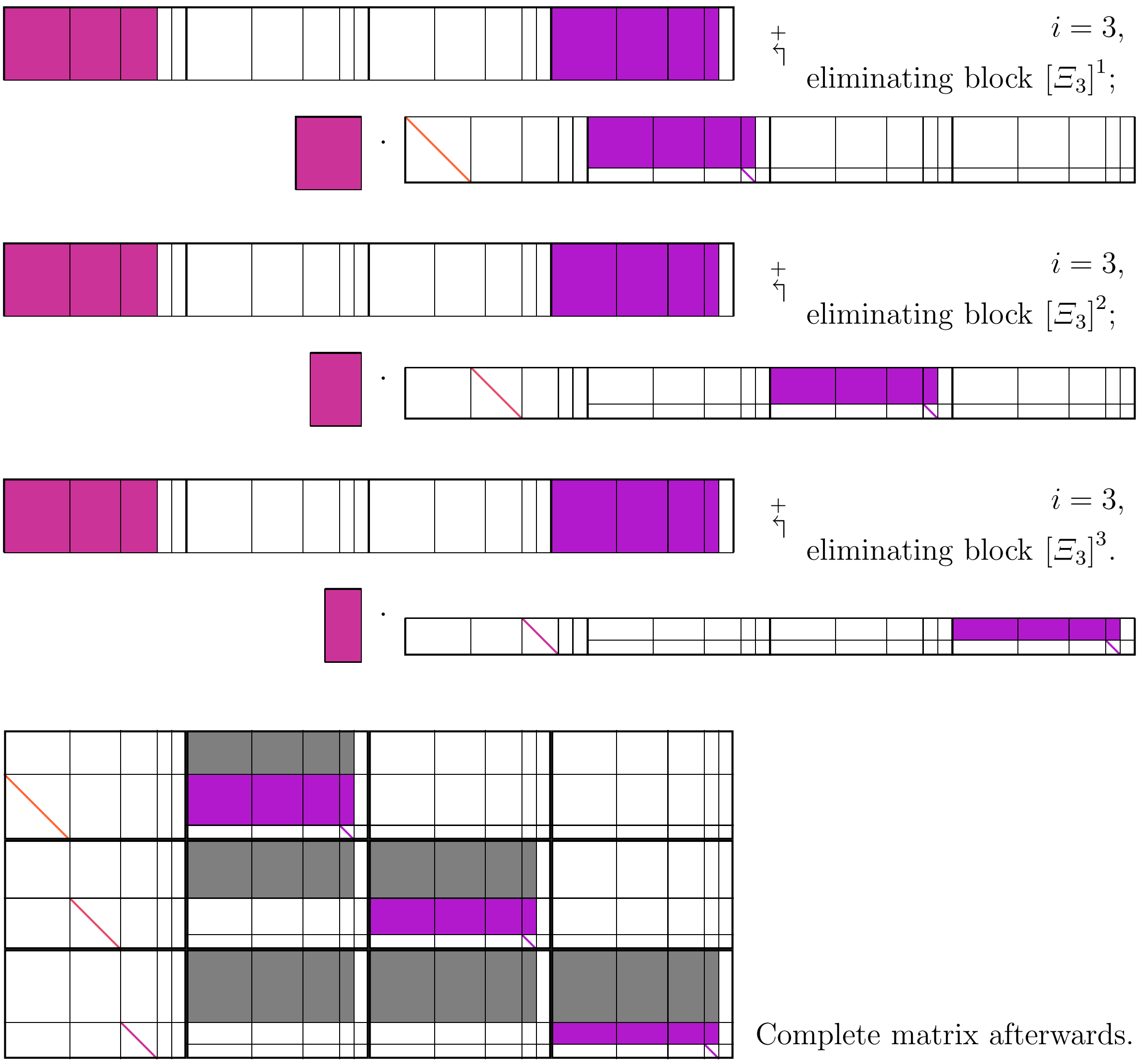


Figure 3.9: Turning the complete matrix in Definition 3.3.10 into row echelon form, step $j = 4$ for $i = 3$. Critical changed entries are shaded gray. Complete matrix not yet contracted.

algebraic conditions (3.29) are satisfied, which is necessary and sufficient for the existence of integral distributions within the Vessiot distribution.

Figures 3.6, 3.7, 3.8 and 3.9 show sketches for the transformations used in the proof for the steps up to $j = 5$.

Proof of Theorem 3.3.9. If the order of the differential equation is $q > 1$, transform it into a first-order equation by the procedure described in Subsection 2.5.1. According to Proposition 2.5.1, if $\mathcal{R}_q$ is involutive, then so is $\mathcal{R}_1$. We assume this first-order equation is represented in reduced Cartan normal form (2.27). We first prove the rank condition (3.41) for the homogeneous system. (The proof for the augmented rank condition (3.42) follows.) We proceed in two steps: We first turn the complete matrix, given in Definition 3.3.10, into row echelon form; then we contract columns and consider the effect.

For the transformation into row echelon form, let $1 < j \leq n$ be given. Choose $1 \leq i < j$. Now we eliminate all the non-trivial entries in the rows

$$\overset{\beta_1^{(i)}}{\underset{1}{}}[\Xi_i\,[i,1]\,[i,2]\cdots[i,j-1]]$$

by using the block matrix calculation as described in Lemma 3.3.20 on the blocks of columns from left to right. For sketches of the procedure up to $j = 4$, see Figures (3.6) to (3.9). We turn to the block $\overset{\beta_1^{(i)}}{\underset{1}{}}[\Xi_i]$ first. Non-trivial entries therein may appear according to Equation (3.21) only for $1 \leq h \leq i$ in the blocks $\overset{\beta_1^{(i)}}{\underset{1}{}}[\Xi_i]^h$. So we fix $1 \leq h \leq i$ and consider the block $\overset{\beta_1^{(i)}}{\underset{1}{}}[\Xi_i]^h$. To apply Lemma 3.3.20, set

$$
\begin{aligned}
(a_{\alpha r}) &= \overset{\beta_1^{(i)}}{\underset{1}{}}[\Xi_i]^{1\dots h-1}, &\quad (3.63a)\\
(b_{\alpha s}) &= \overset{\beta_1^{(i)}}{\underset{1}{}}[\Xi_i]^{h}, &\quad (3.63b)\\
(c_{\alpha t}) &= \overset{\beta_1^{(i)}}{\underset{1}{}}[[\Xi_i]^{h+1\dots n}[i,1][i,2]\dots[i,j-1]]\ , &\quad (3.63c)\\
(d_{st}) &= \overset{m}{\underset{\beta_1^{(h)}+1}{}}[[\Xi_h]_{h+1\dots n}[h,1][h,2]\dots[h,j-1]]\ , &\quad (3.63d)\\
-\mathbb{1}_S &= -\mathbb{1}_{\alpha_1^{(h)}}\ ,\\
0_{S\times R} &= 0_{\alpha_1^{(h)}\times\sum_{l=1}^{h-1}\alpha_1^{(l)}}\ .
\end{aligned}
$$

Then it follows that the substitution (3.58) leaves $(a_{\alpha r})$ unchanged, turns (b_{as}) into zero as required and makes $(c_{\alpha t})$ into $(c_{\alpha t} + \sum_{s=1}^{S} b_{\alpha s}d_{st})$. According to (3.63d) the entries $\sum_{s=1}^{S} b_{\alpha s}d_{st}$ here are of two types: from $\overset{m}{\underset{\beta_1^{(h)}+1}{}}[\Xi_h]_{h+1\dots n}$, the left part of (d_{st}), we have the entries in $(b_{\alpha s})\cdot \overset{m}{\underset{\beta_1^{(h)}+1}{}}[\Xi_h]_{h+1\dots n}$, and from $\overset{m}{\underset{\beta_1^{(h)}+1}{}}[[h,1][h,2]\dots[h,j-1]]$, the right part of (d_{st}), we have those of $(b_{\alpha s})\cdot \overset{m}{\underset{\beta_1^{(h)}+1}{}}[h,l]$ where $1 \leq l \leq j-1$.

Consider the $c_{\alpha t} + \sum_{s=1}^{S} b_{\alpha s}d_{st}$ where the factors d_{st} are from the left part of (d_{st}): according to (3.21), $\overset{m}{\underset{\beta_1^{(h)}+1}{}}[\Xi_h]_{h+1\dots n} = 0$, so

$$(b_{\alpha s})\cdot \overset{m}{\underset{\beta_1^{(h)}+1}{}}[\Xi_h]_{h+1\dots n} = 0\ .$$

It follows $c_{\alpha t} + \sum_{s=1}^{S} b_{\alpha s} d_{st} = c_{\alpha t}$, so that all entries in ${}_{1}^{\beta_1^{(i)}}[\Xi_i]^{h+1...n}$, the left part of (c_{st}), remain unchanged through the elimination of $(b_{\alpha s}) = {}_{1}^{\beta_1^{(i)}}[\Xi_i]^{h}$.

For the rest of the proof, we consider the remaining entries $c_{\alpha t} + \sum_{s=1}^{S} b_{\alpha s} d_{st}$,where the factors d_{st} are from ${}_{\beta_1^{(h)}+1}^{m}[[h,1][h,2]...[h,j-1]]$, the right part of (d_{st}). Fix $1 \le l \le j-1$ and consider $(b_{\alpha s}) \cdot {}_{\beta_1^{(h)}+1}^{m}[h,l]$.

There are two cases: $h = l$ and $h \neq l$. For $h \neq l$ we have $[h,l] = 0$ according to Definition 3.3.10 of the complete matrix, thus ${}_{\beta_1^{(h)}+1}^{m}[h,l] = 0$ and so

$$(b_{\alpha s}) \cdot {}_{\beta_1^{(h)}+1}^{m}[h,l] = 0 \ .$$

Again it follows $c_{\alpha t} + \sum_{s=1}^{S} b_{\alpha s} d_{st} = c_{\alpha t}$, so that for $h \neq l$ all entries in

$${}_{1}^{\beta_1^{(i)}}[[i,1][i,2]...[i,j-1]] \ ,$$

the right part of $(c_{\alpha t})$, remain unchanged through the elimination of $(b_{\alpha s}) = {}_{1}^{\beta_1^{(i)}}[\Xi_i]^{h}$, too.

For $h = l$, we have $[h,h] = \Xi_j$. The structure of Ξ_j, which is (3.21) with i replaced by j, implies that non-vanishing entries are possible in the blocks ${}_{\beta_1^{(h)}+1}^{m}[h,h]_k$ where $1 \le k \le j$. In fact

$$\begin{aligned} {}_{\beta_1^{(j)}+1}^{m}[h,h]_k &= 0 \quad \text{for } 1 \le k \le j-1 \text{ and} && \text{(3.64a)} \\ {}_{\beta_1^{(j)}+1}^{m}[h,h]_j &= -\mathbb{1}_{\alpha_1^{(j)}} . && \text{(3.64b)} \end{aligned}$$

The entries $\sum_{s=1}^{S} b_{\alpha s} d_{st}$ for those d_{st} within ${}_{\beta_1^{(h)}+1}^{m}[h,h]_k$ are

$$\begin{aligned} (b_{\alpha s}) \cdot {}_{\beta_1^{(h)}+1}^{m}[h,h]_k &= {}_{1}^{\beta_1^{(i)}}[\Xi_i]^{h} \cdot {}_{\beta_1^{(h)}+1}^{m}[\Xi_j]_k \\ &= \left(\sum_{s=1}^{S} C^h_{\beta_1^{(h)}+s}(\phi_i^\alpha) C^k_{\beta_1^{(k)}+t}(\phi_j^{\beta_1^{(h)}+s}) \right) \ . \end{aligned} \tag{3.65}$$

This matrix has $A = \beta_1^{(i)}$ rows and $T = \alpha_1^{(k)}$ columns. We consider its entry in row α and column t. Setting $\gamma := \beta_1^{(h)} + s$, $\delta := \beta_1^{(k)} + t$ and using $S = \alpha_1^{(h)}$ in (3.65), this entry is

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(h)}+\alpha_1^{(h)}} C^h_\gamma(\phi_i^\alpha) C^k_\delta(\phi_j^\gamma) \tag{3.66}$$

for all $1 \le k \le j$. Some of the $\beta_1^{(h)} + \alpha_1^{(h)} = m$ summands vanish because of the special entries (3.64). As a consequence, from $\beta_1^{(j)} + 1$ on, of all the summands $C^h_\gamma(\phi_i^\alpha) C^k_\delta(\phi_j^\gamma)$ in (3.66) at most one remains: none for $k \neq j$, exactly one for $k = j$, namely the one for $\gamma = \beta_1^{(j)} + t = \delta$. Using the Kronecker-delta, for all $\beta_1^{(j)} + 1 \le \delta \le m$ we have

$$C^h_\delta(\phi_i^\alpha) C^j_\delta(\phi_j^\delta) = \delta_{kj} \cdot C^h_\delta(\phi_i^\alpha).$$

Now (3.66) becomes for all $1 \le \alpha \le \beta_1^{(i)}$ and all $\beta_1^{(k)} + 1 \le \delta \le m$

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) + \delta_{kj} \cdot C_\delta^h(\phi_i^\alpha) \; . \tag{3.66′}$$

These are the terms $\sum_{s=1}^{S} b_{\alpha s} d_{st}$ in the case $h = l$ when $1 \le k \le j$. Now for the subcase $j+1 \le k \le n$: these blocks ${}^{\;\;m}_{\beta_1^{(h)}+1}[h,h]_k = {}^{\;\;m}_{\beta_1^{(h)}+1}[\Xi_j]_k$, again on account of the structure of Ξ_j, which is (3.21) with i replaced by j, are zero (if they exist at all; they do not for k with $\alpha_1^{(k)} = 0$). So in this case we have

$$(b_{\alpha s}) \cdot {}^{\;\;m}_{\beta_1^{(h)}+1}[\Xi_j]_{j+1\ldots n} = 0 \; ,$$

which contains the terms $\sum_{s=1}^{S} b_{\alpha s} d_{st} = 0$. As a consequence, here, in the case $h = l$, the terms $c_{\alpha t} + \sum_{s=1}^{S} b_{\alpha s} d_{st}$ are of the following form. The $c_{\alpha t}$ of interest in Line (3.63c) are those in the block ${}^{\beta_1^{(i)}}_{1}[i,h]$. For $h \ne i$, we have $[i,h] = 0$, thus ${}^{\beta_1^{(i)}}_{1}[i,h] = 0$. Non-trivial entries $c_{\alpha t}$ are possible only for $h = i$; in that case $[i,i] = \Xi_j$, thus ${}^{\beta_1^{(i)}}_{1}[i,h] = {}^{\beta_1^{(i)}}_{1}[\Xi_j]$. According to the structure of Ξ_j, which is (3.21) with i replaced by j, the only entries $c_{\alpha t}$ within the block ${}^{\beta_1^{(i)}}_{1}[\Xi_j]$ which may not vanish are, for $1 \le k \le j$, those of the form $-C_\delta^k(\phi_j^\alpha)$. They make up the block ${}^{\beta_1^{(i)}}_{1}[\Xi_j]^{1\ldots j} = {}^{\beta_1^{(i)}}_{1}[i,i]^{1\ldots j}$.

So for any row index α and any column index $\delta = \beta_1^{(k)} + t$ we have $c_{\alpha t} = -\delta_{hi} C_\delta^k(\phi_j^\alpha)$ as the most general form of an entry.

Since (3.66′) is $\sum_{s=1}^{S} b_{\alpha s} d_{st}$, it follows that $c_{\alpha t} + \sum_{s=1}^{S} b_{\alpha s} d_{st}$ is

$$-\delta_{hi} C_\delta^k(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) + \delta_{kj} C_\delta^h(\phi_i^\alpha) \; . \tag{3.67}$$

Since $1 \le \alpha \le \beta_1^{(i)}$ and $\beta_1^{(k)} + 1 \le \delta \le m$, any term in the row echelon form of the complete matrix without a -1 in a negative unit block somewhere to its left has this form or is a zero in ${}^{\beta_1^{(i)}}_{1}[\Xi_i]^{i+1\ldots n}$ for some $i < j$ (in which case it remains zero throughout the elementary row transformations and so does not influence the rank; the elementary row transformations in step j for the shift i do not change the transformed shifts a for $a < i$ any more). This means, if all these expressions vanish (when contracted), the rank condition is satisfied. To show that they do vanish (when contracted) for a system with an involutive symbol, we consider them as the new entries in ${}^{\beta_1^{(i)}}_{1}[i,h]$ with $1 \le h \le j$. Now with regard to the relation between h and i there are three cases: $h > i$, $h = i$ and $h < i$. We consider them in that order.

1. Let $h > i$. Then according to the structure of Ξ_i, (3.21), all the $C_\gamma^h(\phi_i^\alpha)$ and $C_\delta^h(\phi_i^\alpha)$ in (3.67) vanish. Since $h \ne i$, $\delta_{hi} = 0$, thus all of (3.67) vanishes.

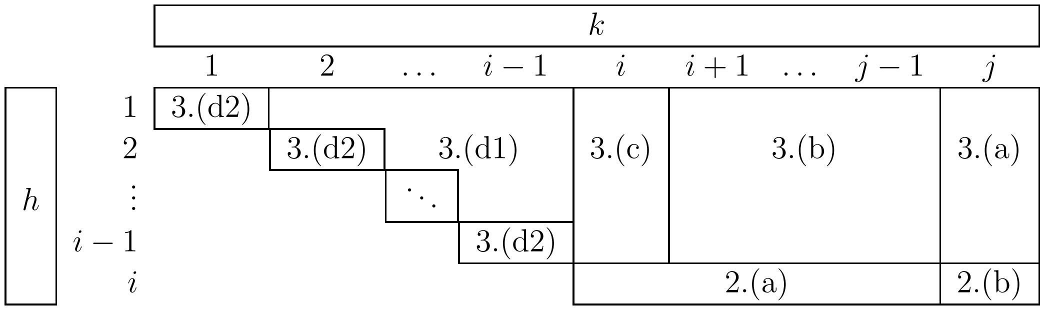


Figure 3.10: Possible combinations of h and k in the terms u^{δ}_{hk} used as labels for columns in the contracted matrix, each corresponding to one subcase in the consideration of cases 2. $h = i$ and 3. $h < i$.

2. Let $h = i$. We shall consider several subcases for $h = i$, and for $h < i$ after that, which may be labelled by second-order derivatives u^{δ}_{hk} according to Figure 3.10 (which is the analogue of Figure 2.2). For fixed δ, any u^{δ}_{hk} belongs to exactly one of these blocks, according to its indices h and k.

   It turns out that (3.67) is the common form of all the sums that appear in the squared brackets of Lemma 2.5.8 as the coefficients of second-order derivatives u^{δ}_{hk}. Not only can the case distinctions of the following argument, which are sketched above, be labelled by these u^{δ}_{hk}, but in fact the case labelled u^{δ}_{hk} is dealt with by using the fact that according to Lemma 2.5.8 the coefficient of that same u^{δ}_{hk} vanishes for an involutive system. This is why sketches (2.2) and (3.10) look alike.

   The cases $h = i$ with subcase $k < i$ and $h < i$ with subcase $k < i$ need not be considered because the columns of the complete matrix are to be contracted when proving the rank condition. This contraction concerns those columns of the complete matrix which are labelled by the same second-order derivative, and each second-order derivative is used exactly once in the argument.

   (a) For the first subcase of $h = i$ let $i \le k < j$. Then (3.67) becomes

   $$-C^{k}_{\delta}(\phi^{\alpha}_{j}) + \sum_{\gamma=\beta^{(i)}_{1}+1}^{\beta^{(j)}_{1}} C^{i}_{\gamma}(\phi^{\alpha}_{i}) C^{k}_{\delta}(\phi^{\gamma}_{j})$$

   where $\beta^{(k)}_{1} + 1 \le \delta \le m$. This vanishes for a system with an involutive symbol according to Lemma 2.5.8, Line (2.36).

   (b) For the second subcase of $h = i$ choose $k = j$. Then (3.67) becomes

   $$-C^{j}_{\delta}(\phi^{\alpha}_{j}) + \sum_{\gamma=\beta^{(i)}_{1}+1}^{\beta^{(j)}_{1}} C^{i}_{\gamma}(\phi^{\alpha}_{i}) C^{j}_{\delta}(\phi^{\gamma}_{j}) + C^{i}_{\delta}(\phi^{\alpha}_{i})$$

where $\beta_1^{(j)} + 1 = \delta$. According to Lemma 2.5.8, Line (2.38), this vanishes for a system with an involutive symbol.

3. Let $h < i$. We consider several subcases with regard to the relation between $h < i$, j and k.

(a) First choose $k = j$. Then (3.67) becomes

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^j(\phi_j^\gamma) + C_\delta^h(\phi_i^\alpha)$$

where $\beta_1^{(j)} + 1 \leq \delta \leq m$. According to Lemma 2.5.8, line (10), this vanishes for a system with an involutive symbol.

(b) For the second subcase choose $i < k < j$. Then (3.67) becomes

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma)$$

where $\beta_1^{(k)} + 1 \leq \delta \leq m$. According to Lemma 2.5.8, line (8), this vanishes for a system with an involutive symbol.

(c) For the third subcase choose $k = i$. Then (3.67) becomes

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^i(\phi_j^\gamma) = \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(i)}} C_\gamma^h(\phi_i^\alpha) C_\delta^i(\phi_j^\gamma) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^i(\phi_j^\gamma) \tag{3.68}$$

where $\beta_1^{(k)} + 1 \leq \delta \leq m$. Since we have $h < k = i$, for any such δ the cross derivative $u_{ih}^\delta = u_{hi}^\delta$ labels two columns in the complete matrix: the one with label δ in ${}_{1}^{\beta_1^{(i)}}[i,h]^i$ and the one with label δ in ${}_{1}^{\beta_1^{(i)}}[i,i]^h$ which according to (3.67) has the new entries

$$-\delta_{ki} C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) + \delta_{hj} C_\delta^k(\phi_i^\alpha) \tag{3.69}$$

where $\beta_1^{(h)} \leq \delta \leq m$. In the current subcase, (3.69) becomes

$$-C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) \tag{3.70}$$

where $\beta_1^{(h)} \leq \delta \leq m$. Since the columns of both (3.68) and (3.70) are labelled by the same second-order derivatives, we have to contract them, which means adding their new entries. For all $\beta_1^{(h)} + 1 \leq \delta \leq \beta_1^{(i)}$ this yields

$$-C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) ,$$

which, for a system with an involutive symbol, vanishes according to Lemma 2.5.8, Line (2.34a), and for all $\beta_1^{(i)} + 1 \leq \delta \leq m$

$$-C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) + \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^i(\phi_j^\gamma) ,$$

which, for a system with an involutive symbol, again vanishes according to Lemma 2.5.8, Line (2.34b).

(d) For the fourth subcase choose $k < i$. Under this assumption, we have to distinct two further subcases: $k = h$ and $k < h$. The subcase $k > h$ need not be considered since $u_{kh}^\delta = u_{hk}^\delta$ and any second-order derivative is used only once to label a column in the contracted matrix.

(d1) First consider $k < i$ and $k = h$. Then still $h < i$, furthermore $k < j$, and (3.67) becomes

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma)$$

where $\beta_1^{(h)} + 1 \leq \delta \leq m$. According to Lemma 2.5.8, Line (2.32), this vanishes for a system with an involutive symbol.

(d2) At last consider $h < k < i$. Then $k < j$, and (3.67) becomes

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) = \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(k)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) + \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) \tag{3.71}$$

where $\beta_1^{(k)} + 1 \leq \delta \leq m$. Since we have $h < k < i$, for any such δ the cross derivative $u_{kh}^\delta = u_{hk}^\delta$ labels two columns in the complete matrix: the one with label δ in ${}_{1}^{\beta_1^{(i)}}[i,h]^k$ and the one with label δ in ${}_{1}^{\beta_1^{(i)}}[i,k]^h$ which according to (3.67) has the new entries

$$-\delta_{ki} C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) + \delta_{hj} C_\delta^k(\phi_i^\alpha) \tag{3.72}$$

where $\beta_1^{(h)} \leq \delta \leq m$. In the current subcase, (3.72) becomes

$$\sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) \tag{3.73}$$

where $\beta_1^{(h)} \leq \delta \leq m$. Since the columns of both (3.71) and (3.73) are labelled by the same second-order derivatives, we have to contract them, which means adding their new entries. This yields

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(k)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) + \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) + \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma)$$

where, for a system with an involutive symbol, the right term vanishes according to Lemma 2.5.8, Line (2.33a), and the sum of the left and the middle terms vanishes according to the same lemma, Line (2.33b).

What we have shown is that in the contracted matrix all the entries without some entry -1 in a negative unit block to their left may be eliminated by elementary row transformations if the equation has an involutive symbol. Thus under this assumption the rank condition (3.41) is satisfied.

Now for the augmented rank condition, Equation (3.42). To transform the augmented complete matrix into row echelon form, we use for each j and i the same procedure as for the transformation of the non-augmented complete matrix, except that now the matrices $(c_{\alpha t})$ and (d_{st}) are augmented by one more column each as follows. Fix $1 < j \leq n$. Let $1 \leq i < j$. Then, according to the structure of the augmented complete matrix given in Definition 3.3.10, for its transformation into row echelon form we have to eliminate for $1 \leq h \leq i$ the entries in the blocks $[\Xi_i]^h$, like we did for the non-augmented complete matrix. We have to consider the effect of these transformations on the additional entries which make up the rightmost column in the augmented complete matrix. These are $-\Theta_{ij}^1$, $-\Theta_{ij}^2$, …, $-\Theta_{ij}^m$, given in Equation (3.18). Of these entries, only $-\Theta_{ij}^1$, $-\Theta_{ij}^2$, …, $-\Theta_{ij}^{\beta_1^{(i)}}$ are affected (since we eliminate the entries which are in rows 1 to $\beta_1^{(i)}$ of the matrices $[\Xi_i]^h$). We add them as the rightmost column in the augmented matrix $(c_{\alpha t})$. Now fix $1 \leq h \leq i$. Then augment the matrix (d_{st}), used in the process of eliminating the entries in $[\Xi_i]^h$, by adding the entries $-\Theta_{hj}^{\beta_1^{(h)}+1}$, $-\Theta_{hj}^{\beta_1^{(h)}+2}$, …, $-\Theta_{hj}^m$ as its rightmost column, in accordance with the structure of the augmented complete matrix as given in Definition 3.3.10. Then the substitution (3.58) yields as the transformed entries $c_{\alpha t} + \sum_{s=1}^S b_{\alpha s} d_{st}$ the same as for the transformation of the non-augmented matrix except, of course, for the new last column. Now let denote t the index of this last column. Fix some row index $1 \leq \alpha \leq \beta_1^{(i)}$. Then the entry $c_{\alpha t} = -\Theta_{ij}^\alpha$ transforms as follows: For all $1 \leq h \leq i$ we have to add to it

$$\sum_{s=1}^S b_{\alpha s} d_{st} = \sum_{\gamma=\beta_1^{(h)}+1}^m b_{\alpha\gamma} d_{\gamma t} \ .$$

Since here the matrix $(b_{\alpha\gamma}) = [\Xi_i]^h$, this is the product of the row with index α in the matrix $[\Xi_i]^h$ and the transpose of the vector $(-\Theta_{hj}^{\beta_1^{(h)}+1}, -\Theta_{hj}^{\beta_1^{(h)}+2}, \dots, -\Theta_{hj}^m)$. According to the structure of $[\Xi_i]^h$ as defined in Equation (3.20), this equals

$$\sum_{\gamma=\beta_1^{(h)}+1}^{m} C_\gamma^h(\phi_i^\alpha)\Theta_{hj}^\gamma \ .$$

The entries Θ_{hj}^γ can be taken from Equation (3.18). Since $\beta_1^{(h)} + 1 \leq \gamma \leq m$, we have $\Theta_{hj}^\gamma = C_h^{(1)}(\phi_j^\alpha)$ if $\beta_1^{(h)} + 1 \leq \gamma \leq \beta_1^{(j)}$ and $\Theta_{hj}^\gamma = 0$ if $\beta_1^{(j)} + 1 \leq \gamma \leq m$. Therefore we get

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_h^{(1)}(\phi_j^\alpha)$$

as the summand for each h to be added to $c_{\alpha t} = -\Theta_{ij}^\alpha$. Since $1 \leq \alpha \leq \beta_1^{(i)}$, we have $-\Theta_{ij}^\alpha = -C_i^{(1)}(\phi_j^\alpha) + C_j^{(1)}(\phi_i^\alpha)$. Therefore the entry $c_{\alpha t}$ transforms into

$$-C_i^{(1)}(\phi_j^\alpha) + C_j^{(1)}(\phi_i^\alpha) + \sum_{h=1}^{i} \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_h^{(1)}(\phi_j^\gamma) \ ,$$

which is the integrability condition in Line (2.31) of Lemma 2.5.8, except for the sign. The difference in sign comes from the fact that the augmented complete matrix describes the system of equations (3.40), where the entries from the row with index α in the matrix Ξ_i are on the opposite side from the entries from the row with index α in the matrix Ξ_j and the inhomogeneous term $-\Theta_{ij}^\alpha$, while in Lemma 2.5.8 for each i, j and α the corresponding equation is set to zero, if the differential equation is involutiv.

This means that the augmented rank condition holds for all $1 \leq i < j \leq n$ if, and only if, the equation has an involutive symbol and is formally integrable. □

As a first corollary, we assert our announcement in Section 3.2 concerning Cartan characters and the rank of the matrices used in the step-by-step construction of an integral manifold.

Corollary 3.3.21. *Let $\mathcal{R}_1$ be a differential equation with an involutive symbol $\mathcal{N}_1$. Then for a local representation in δ-regular coordinates, for all $1 \leq i < j \leq n$, the Cartan characters $\alpha_1^{(i)}$ satisfy*

$$\sum_{i=1}^{j-1} \alpha_1^{(i)} = \operatorname{rank} \begin{pmatrix} \hat{\Xi}_1 & \hat{\Xi}_j & & & \\ \hat{\Xi}_2 & & \hat{\Xi}_j & & 0 \\ \vdots & 0 & & \ddots & \\ \hat{\Xi}_{j-1} & & & & \hat{\Xi}_j \end{pmatrix} .$$

For $q > 1$, the Cartan characters $\alpha_q^{(i)}$ may be used on the left side instead of the $\alpha_1^{(i)}$ if the complete matrix at step j on the right side is that of the first-order system which is deduced from the representation of $\mathcal{R}_q$ according to the procedure in Subsection 2.5.1.

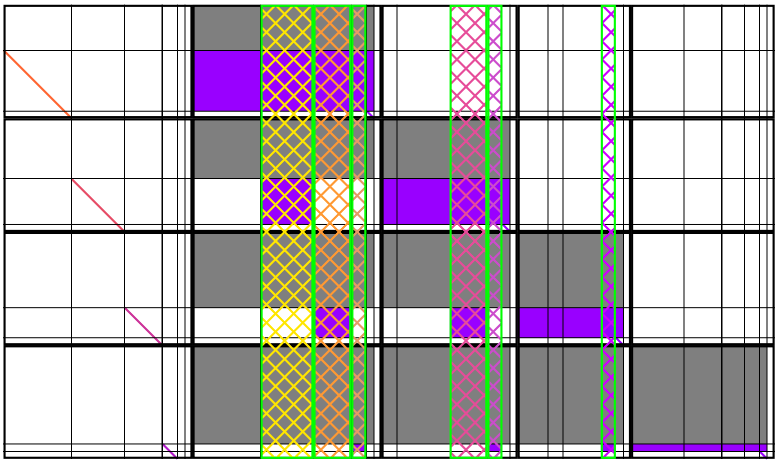


Figure 3.11: Complete matrix for step $j = 5$ after contraction (as shown in Figure 3.4), now in row echelon form. Blocks of columns shown green-framed and hatched are those which were concerned by contraction. Areas shaded gray contain the coefficients of Lemma 2.5.8 for $j = 5$ and $1 \leq i \leq 4$ as entries. Areas shaded violet are not changed through the process but may have moved through contraction.

Proof. Under the above assumptions, the contracted complete matrix in row echelon form has non-trivial entries only in the negative unit blocks $-\mathbb{1}_{\alpha_1^{(i)}}$ and rows which have their leftmost entries in such a block. The assertion on systems of order q follows from Proposition 2.5.1, second item. $\square$

Example 3.3.22. We take up Example 3.3.19, where $n = 3$, $m = 5$, $\beta_1^{(1)} = 2$, $\beta_1^{(2)} = 3$ and $\beta_1^{(3)} = 4$. For $j = 2$, only $i = 1$ is possible, and the last column in the augmented complete matrix before transforming it into row echelon form has the five entries $-\Theta_{12}^1 = -C_1^{(1)}(\phi_2^1) + C_2^{(1)}(\phi_1^1)$, $-\Theta_{12}^2 = -C_1^{(1)}(\phi_2^2) + C_2^{(1)}(\phi_1^2)$, $-\Theta_{12}^3 = -C_1^{(1)}(\phi_2^3)$, $-\Theta_{12}^4 = 0$ and $-\Theta_{12}^5 = 0$. Afterwards it contains

$$\begin{aligned}
-\Theta_{12}^1 &\leftarrow -C_1^{(1)}(\phi_2^1) + C_2^{(1)}(\phi_1^1) + C_3^1(\phi_1^1)C_1^{(1)}(\phi_2^1)\ , \\
-\Theta_{12}^2 &\leftarrow -C_1^{(1)}(\phi_2^2) + C_2^{(1)}(\phi_1^2) + C_3^1(\phi_1^2)C_1^{(1)}(\phi_2^1)\ , \\
-\Theta_{12}^3 &\leftarrow -C_1^{(1)}(\phi_2^3)\ , \\
-\Theta_{12}^4 &\leftarrow 0 \quad \text{and} \\
-\Theta_{12}^5 &\leftarrow 0\ .
\end{aligned}$$

The first two substitutions produce the integrability conditions for $j = 2$ and $i = 1$, while the other three entries remain unchanged.

For $j = 3$, we have the ten entries: five for $i = 1$, which are $-\Theta_{13}^1 = -C_1^{(1)}(\phi_3^1) + C_3^{(1)}(\phi_1^1)$, $-\Theta_{13}^2 = -C_1^{(1)}(\phi_3^2) + C_3^{(1)}(\phi_1^2)$, $-\Theta_{13}^3 = -C_1^{(1)}(\phi_3^3)$, $-\Theta_{13}^4 = -C_1^{(1)}(\phi_3^4)$, $-\Theta_{13}^5 = 0$, and five for $i = 2$, which are $-\Theta^1_{23} = -C^{(1)}_2(\phi^1_3) + C^{(1)}_3(\phi^1_2)$, $-\Theta^2_{23} = -C^{(1)}_2(\phi^2_3) + C^{(1)}_3(\phi^2_2)$, $-\Theta^3_{23} = -C^{(1)}_2(\phi^3_3)+C^{(1)}_3(\phi^3_2)$, $-\Theta^4_{23} = -C^{(1)}_2(\phi^4_3)$ and $-\Theta^5_{23} = 0$ before the transformation. The substitutions are for the first shift, $i = 1$:

$$
\begin{aligned}
&-\Theta^1_{13} \leftarrow -C^{(1)}_1(\phi^1_3) + C^{(1)}_3(\phi^1_1) + C^1_3(\phi^1_1)C^{(1)}_1(\phi^3_3) + C^1_4(\phi^1_1)C^{(1)}_1(\phi^4_3)\ ,\\
&-\Theta^2_{13} \leftarrow -C^{(1)}_1(\phi^2_3) + C^{(1)}_3(\phi^2_1) + C^1_3(\phi^2_1)C^{(1)}_1(\phi^3_3) + C^1_4(\phi^2_1)C^{(1)}_1(\phi^4_3)\ ,\\
&-\Theta^3_{13} \leftarrow -C^{(1)}_1(\phi^3_3)\ ,\\
&-\Theta^4_{13} \leftarrow -C^{(1)}_1(\phi^4_3)\ ,\\
&-\Theta^5_{13} \leftarrow 0\ ,
\end{aligned}
$$

and for the second shift, $i = 2$:

$$
\begin{aligned}
&-\Theta^1_{23} \leftarrow -C^{(1)}_2(\phi^1_3) + C^{(1)}_3(\phi^1_2) + C^1_3(\phi^1_2)C^{(1)}_1(\phi^3_3) + C^1_4(\phi^1_2)C^{(1)}_1(\phi^4_3) + C^2_4(\phi^1_2)C^{(1)}_2(\phi^4_3)\ ,\\
&-\Theta^2_{23} \leftarrow -C^{(1)}_2(\phi^2_3) + C^{(1)}_3(\phi^2_2) + C^1_3(\phi^2_2)C^{(1)}_1(\phi^3_3) + C^1_4(\phi^2_2)C^{(1)}_1(\phi^4_3) + C^2_4(\phi^2_2)C^{(1)}_2(\phi^4_3)\ ,\\
&-\Theta^3_{23} \leftarrow -C^{(1)}_2(\phi^3_3) + C^{(1)}_3(\phi^3_2) + C^1_3(\phi^3_2)C^{(1)}_1(\phi^3_3) + C^1_4(\phi^3_2)C^{(1)}_1(\phi^4_3) + C^2_4(\phi^3_2)C^{(1)}_2(\phi^4_3)\ ,\\
&-\Theta^4_{23} \leftarrow -C^{(1)}_2(\phi^4_3) \quad \text{and}\\
&-\Theta^5_{23} \leftarrow 0\ .
\end{aligned}
$$

Five entries remain unchanged (three for $i = 1$ and two for $i = 2$), while the other five are the integrability conditions which have to be satisfied in the step $j = 3$ for the system to be formally integrable.

In the proof of the Theorem 3.3.9 we produced an explicit row echelon form of the system matrix given in Definition 3.3.10, which we now exploit to solve that system of linear equations: we express some of the unknowns ζ^p_ℓ as linear combinations of the remaining ones and thus parametrize the solution space. Then we eliminate those unknowns ζ^p_ℓ from the differential conditions (3.30) and analyze the arising differential equation. The following lemma shows the exact relations between these unknowns.

Corollary 3.3.23. *Let $1 \leq i < j \leq n$. For all $1 \leq \ell \leq n$ the entries of the vector $\zeta_\ell = (\zeta^p_\ell\colon 1 \leq p \leq r) = (\zeta^{(\alpha,h)}_\ell\colon (\alpha, h) \notin \mathcal{B})$ are grouped into n blocks of which the hth one is the transpose of*

$$(\zeta_\ell^{(\beta^{(h)}_1+1,h)}, \zeta_\ell^{(\beta^{(h)}_1+2,h)}, \ldots, \zeta_\ell^{(m,h)})\ .$$

Then the relations between the entries of ζ_i and ζ_j are for all (α, i) where $\beta^{(j)}_1+1 \leq \alpha \leq m$:

$$\zeta_j^{(\alpha,i)} = \zeta_i^{(\alpha,j)}\ ; \tag{3.74}$$

and for all (α, i) where $\beta^{(i)}_1 + 1 \leq \alpha \leq \beta^{(j)}_1$:

$$\zeta_j^{(\alpha,i)} = \sum_{k=1}^{j} \sum_{\gamma=\beta^{(k)}_1+1}^{m} C^k_\gamma(\phi^\alpha_j)\zeta_i^{(\gamma,k)} + C^{(1)}_i(\phi^\alpha_j)\ . \tag{3.75}$$

Proof. Fix j. Let be $i < j$. The ith block of ζ_j is $(\zeta_j^{(\alpha,i)} \colon \beta_1^{(i)} + 1 \le \alpha \le m)$. According to the structure of the complete contracted matrix in row echelon form, the entries of $(\zeta_j^{(\alpha,i)} \colon \beta_1^{(i)} + 1 \le \alpha \le m)$ are related to those of the vectors ζ_k for $1 \le k < j$ according to the rows

$$ {}^{m}_{\beta_1^{(i)}+1}[\widehat{[i,0]}\widehat{[i,1]}\widehat{[i,2]}\cdots\widehat{[i,j-1]}] \ . \tag{3.76} $$

There are two sorts of such rows: one for $\beta_1^{(i)} + 1 \le \alpha \le \beta_1^{(j)}$ and one for $\beta_1^{(j)} + 1 \le \alpha \le m$.

Let $\beta_1^{(j)} + 1 \le \alpha \le m$. Then according to Corollary 3.3.13, the following entries of the row with index α in the block matrix (3.76) are of interest: in $[i,0]_i = -\mathbb{1}_{\alpha_1^{(i)}}$ the -1 which is the coefficient of $\zeta_j^{(\alpha,i)}$ and therefore appears in the column indexed by Y_{ij}^{α}; and the entries of the row with index α in ${}^{m}_{\beta_1^{(j)}+1}\widehat{[i,i]}_j$. This row contains the coefficients of $\zeta_i^{(\beta_1^{(j)}+1,j)}$ to $\zeta_i^{(m,j)}$, and according to Equation (3.51d), we have

$$ \zeta_j^{(\alpha,i)} = \sum_{\gamma=\beta_1^{(j)}+1}^{m} C_\gamma^j(\phi_j^\alpha)\zeta_i^{(\gamma,j)} = \zeta_i^{(\alpha,j)} \ , \tag{3.77} $$

which proves Equation (3.74).

Now let $\beta_1^{(i)} + 1 \le \alpha \le \beta_1^{(j)}$. Again according to Corollary 3.3.13, in the row with index α of the block matrix (3.76) the -1 in the block $[i,0]_i = -\mathbb{1}_{\alpha_1^{(i)}}$ is the coefficient of $\zeta_j^{(\alpha,i)}$. The only further entries of that row with index α in the block matrix (3.76) that possibly do not vanish are, according to Corollary 3.3.13, for $k < i$ the entries of the row with index α in the block matrix ${}^{\beta_1^{(j)}}_{1}\widehat{[i,k]}i$, and for $k = i$ the entries of the row with index α in the block matrix ${}^{\beta_1^{(j)}}_{1}\widehat{[i,i]}^{1\dots j}$. (For $i < k \le j-1$ the entries vanish according to the structure of the contracted complete matrix.) For $k < i$, they can be read off from Equation (3.51a), which gives for all $1 \le k < i$ the coefficients for the entries in the vectors $\hat{\zeta}_k$: they are

$$ -C_\gamma^k(\phi_j^\alpha) \ , \quad \beta_1^{(i)} + 1 \le \gamma \le m \ . $$

For $k = i$, Equation (3.51b) shows for all $1 \le h < i$ the coefficients of the entries in $\hat{\zeta}_i$; they are

$$ -C_\gamma^h(\phi_j^\alpha) \ , \quad \beta_1^{(h)} + 1 \le \gamma \le \beta_1^{(i)} \ . $$

For $k = i$ and all $i \le h \le j$, Equation (3.51c) gives as the coefficients of the entries in $\hat{\zeta}_i$

$$ -C_\gamma^h(\phi_j^\alpha) \ , \quad \beta_1^{(h)} + 1 \le \gamma \le m \ . $$

The rightmost column of the augmented complete matrix contains the entries of the vectors $-\Theta_{ij}$, of which the one in row number α is $-\Theta_{ij}^\alpha$. According to our considerations at the end of the proof of Theorem 3.3.9 (on page 118) this entry remains unchanged throughout contraction and transformation into row echelon form. Equation (3.18) gives it as $-\Theta^{\alpha}_{ij} = -C^{(1)}_i(\phi^{\alpha}_j)$. Summing up, we arrive at

$$\zeta^{(\alpha,i)}_j = \sum_{k=1}^{i-1} \sum_{\gamma=\beta^{(i)}_1+1}^{m} C^k_\gamma(\phi^\alpha_j)\zeta^{(\gamma,i)}_k \tag{3.78a}$$

$$+\sum_{h=1}^{i-1} \sum_{\gamma=\beta^{(h)}_1+1}^{\beta^{(i)}_1} C^h_\gamma(\phi^\alpha_j)\zeta^{(\gamma,h)}_i \tag{3.78b}$$

$$+\sum_{h=i}^{j} \sum_{\gamma=\beta^{(h)}_1+1}^{m} C^h_\gamma(\phi^\alpha_j)\zeta^{(\gamma,h)}_i + C^{(1)}_i(\phi^\alpha_j)\ . \tag{3.78c}$$

In Lines (3.78b) and (3.78c) we change the summation index h to k. Since we have already shown that Equation (3.77) holds for all $i < j$, it holds in particular for $1 \le k \le i$, so we have $\zeta^{(\gamma,i)}_k = \zeta^{(\gamma,k)}_i$ and thus can combine the right term in (3.78a) and (3.78b) into $\sum_{k=1}^{i-1}\sum_{\gamma=\beta^{(k)}_1+1}^{m} C^k_\gamma(\phi^\alpha_j)\zeta^{(\gamma,k)}_i$. Adding (3.78c) to it, we arrive at (3.75). □

Remark 3.3.24. This calculation shows that we would have obtained the same Equations (3.74) and (3.75) for the interrelations between the entries of the vectors ζ_i and ζ_j, if we had used the non-contracted complete matrices without transforming them into row echelon form. Transformation into row echelon form does not influence these interrelations, because the rows which are used to describe them are not changed by the row transformations at all. Contraction does not influence them, because of the two coefficients in such a row being added through contraction, one is certainly zero. Therefore, contraction only means to take into account that $\zeta^{(\alpha,i)}_j = \zeta^{(\alpha,j)}_i$ whenever $i < j$ and $\beta^{(j)}_1 + 1 \le \alpha \le m$, as Equation (3.74) shows it. In other words, it means to take into account that in Equation (3.75) and sums like it the terms $\zeta^{(\alpha,i)}_j$ and $\zeta^{(\alpha,j)}_i$ are interchangeable because both are equal and in the non-contracted complete matrix one of them has a zero coefficient while the other one has not.

Transforming the complete matrix and the augmented complete matrix is not necessary to exhibit the interrelations between $\zeta^{(\alpha,i)}_j$ and $\zeta^{(\alpha,j)}_i$, but to prove the rank conditions (3.41) and (3.42) for an involutive system.

This remark and Corollary 3.3.23 before it imply another interrelation between the entries of the vectors ζ_i and ζ_j which helps us prove the existence theorem for flat Vessiot connections in Subsection 3.3.7.

Corollary 3.3.25. *Let $1 \le i < j \le n$. Then for all β with $1 \le \beta \le \beta^{(i)}_1$, the entries of the vector ζ_i and those of vector the ζ_j satisfy*

$$\sum_{a=1}^{i} \sum_{\gamma=\beta^{(a)}_1+1}^{m} C^a_\gamma(\phi^\beta_i)\zeta^{(\gamma,a)}_j - \sum_{b=1}^{j} \sum_{\delta=\beta^{(b)}_1+1}^{m} C^b_\delta(\phi^\beta_j)\zeta^{(\delta,b)}_i = C^{(1)}_i(\phi^\beta_j) - C^{(1)}_j(\phi^\beta_i)\ . \tag{3.79}$$

Proof. Since $i < j$, we have $\beta_1^{(i)} \leq \beta_1^{(j)}$. Thus, for $1 \leq \beta \leq \beta_1^{(i)}$ both $(\beta, i) \in \mathcal{B}$ and $(\beta, j) \in \mathcal{B}$. Now it follows from Equation (3.18) that $-\Theta_{ij}^{\beta} = -C_i^{(1)}(\phi_k^{\beta}) + C_k^{(1)}(\phi_i^{\beta})$. If contraction of columns could be ignored, the equality (3.79) would follow from Definition 3.3.10 of the augmented complete, non-contracted, matrix and the distribution of its entries (3.18) and (3.19) in the row with index β. But as noted in Remark 3.3.24, the obtained interrelations between the entries $\zeta_j^{(\gamma,a)}$ and $\zeta_i^{(\delta,b)}$ are the same, whether we use the contracted complete matrix or the non-contracted one. □

Example 3.3.26. As a counterexample for Vessiot's step-by-step approach consider the differential equation of second order locally represented by

$$\mathcal{R}_2\colon \begin{cases} u_{tt} = au \\ u_{xx} = bu \end{cases} ,$$

where a and b are real constants. It is taken from Seiler [37]. For $a = b = 0$, we arrive at the system considered in Example 2.4.14 where we saw that its symbol is not involutive. The symbol remains the same for arbitrary real numbers a and b and hence is not involutive in the general case, too. Transforming it into a first order equation according to the procedure given in Subsection 2.5.1 would still yield a system with a symbol that is not involutive. This is why this differential equation does not meet the necessary requirements of Theorem 3.3.9. As with the special case of this system that is treated in Example 2.4.14, this system is formally integrable.

The pull-backs of the three generating contact forms are

$$\begin{aligned} \iota^*\omega &= d\overline{u} - \overline{u_x}d\overline{x} - \overline{u_t}d\overline{t} \\ &= d\overline{u} - \overline{u_x}d\overline{x} - \overline{u_t}d\overline{t}\ , \\ \iota^*\omega_x &= d\overline{u_x} - \overline{u_{xx}}d\overline{x} - \overline{u_{xt}}d\overline{t} \\ &= d\overline{u_x} - \overline{bu}d\overline{x} - \overline{u_{xt}}d\overline{t}\ , \\ \iota^*\omega_t &= d\overline{u_t} - \overline{u_{xt}}d\overline{x} - \overline{u_{tt}}d\overline{t} \\ &= d\overline{u_t} - \overline{u_{xt}}d\overline{x} - \overline{au}d\overline{t}\ . \end{aligned}$$

They annihilate the subdistribution $\mathcal{V}[\mathcal{R}_2] \subset T\mathcal{R}_2 = \operatorname{span}\{\partial_{\overline{x}}, \partial_{\overline{t}}, \partial_{\overline{u}}, \partial_{\overline{u_x}}, \partial_{\overline{u_t}}, \partial_{\overline{u_{xt}}}\}$ which is spanned by the vector fields

$$\begin{aligned} \bar{X}_1 &= \partial_{\overline{x}} + \overline{u_x}\partial_{\overline{u}} + b\overline{u}\partial_{\overline{u_x}} + \overline{u_{xt}}\partial_{\overline{u_t}}\ , \\ \bar{X}_2 &= \partial_{\overline{t}} + \overline{u_t}\partial_{\overline{u}} + \overline{u_{xt}}\partial_{\overline{u_x}} + a\overline{u}\partial_{\overline{u_t}}\ , \\ \bar{Y} &= \partial_{\overline{xt}}\ . \end{aligned}$$

Their non-trivial Lie brackets are

$$[\bar{X}_1, \bar{X}_2] = a\overline{u_x}\partial_{\overline{u_t}} - b\overline{u_t}\partial_{\overline{u_x}}\ , \qquad [\bar{X}_1, \bar{Y}] = -\partial_{\overline{u_t}} \quad \text{and} \quad [\bar{X}_2, \bar{Y}] = -\partial_{\overline{u_x}}\ .$$

To construct $\bar{U}_1 := \bar{X}_1 + \zeta_1\bar{Y}$, choose an arbitrary function $\zeta_1 \in \mathcal{F}(\mathcal{R}_2)$. Then the next step in Vessiot's approach is to choose a second function $\zeta_2 \in \mathcal{F}(\mathcal{R}_2)$ such that for

$\bar{U}_2 := \bar{X}_2 + \zeta_2\bar{Y}$ we have a two-dimensional integral distribution $\mathcal{U} := \operatorname{span}\{\bar{U}_1, \bar{U}_2\}$. To achieve this, the Lie bracket

$$[\bar{U}_1, \bar{U}_2] \equiv (a\overline{u_x} - \zeta_2)\partial_{\overline{u_t}} - (b\overline{u_t} - \zeta_1)\partial_{\overline{u_x}} \mod \mathcal{V}[\mathcal{R}_2]$$

has to satisfy the condition $[\bar{U}_1, \bar{U}_2] \equiv 0 \mod \mathcal{V}[\mathcal{R}_2]$ as $\bar{U}_1$ and $\bar{U}_2$ are in triangular form. It follows that the two conditions $a\overline{u_x} = \zeta_2$ and $b\overline{u_t} = \zeta_1$ are to be met, fixing both functions. The second of these conditions concerns ζ_1 though in a step-by-step process, it should not be restricted by conditions arising during the second step.

Nevertheless, the combined system of algebraic equations is solvable, and therefore the vector fields

$$\bar{U}_1 := \bar{X}_1 + b\overline{u_t}\bar{Y} \quad \text{and} \quad \bar{U}_2 := \bar{X}_2 + a\overline{u_x}\bar{Y}$$

span a two-dimensional subdistribution in $\mathcal{V}[\mathcal{R}_2]$ which is also transversal. It is even involutive:

$$\begin{aligned}
[\bar{U}_1, \bar{U}_2] &= [\bar{X}_1 + b\overline{u_t}\bar{Y}, \bar{X}_2 + a\overline{u_x}\bar{Y}] \\
&= a\overline{u_x}\partial_{\overline{u_t}} - b\overline{u_t}\partial_{\overline{u_x}} + a\overline{u_x}[\bar{X}_1, \bar{Y}] + \bar{X}_1(a\overline{u_x})\bar{Y} \\
&\quad - \bar{X}_2(b\overline{u_t})\bar{Y} - b\overline{u_t}[\bar{X}_2, \bar{Y}] + a\overline{u_x}[b\overline{u_t}\bar{Y}, \bar{Y}] + b\overline{u_t}\bar{Y}(a\overline{u_x})\bar{Y} \\
&= a\overline{u_x}\partial_{\overline{u_t}} - b\overline{u_t}\partial_{\overline{u_x}} + a\overline{u_x}(-\partial_{\overline{u_t}}) + \bar{X}_1(a\overline{u_x})\bar{Y} \\
&\quad - \bar{X}_2(b\overline{u_t})\bar{Y} + b\overline{u_t}\partial_{\overline{u_x}} - a\overline{u_x}\bar{Y}(b\overline{u_t})\bar{Y} + b\overline{u_t}\bar{Y}(a\overline{u_x})\bar{Y} \\
&= \bar{X}_1(a\overline{u_x})\bar{Y} - \bar{X}_2(b\overline{u_t})\bar{Y} - a\overline{u_x}\bar{Y}(b\overline{u_t})\bar{Y} + b\overline{u_t}\bar{Y}(a\overline{u_x})\bar{Y} \\
&= b\overline{u}a\bar{Y} - a\overline{u}b\bar{Y} = 0\ .
\end{aligned}$$

This means, for $\mathcal{R}_2$ there is an n-dimensional transversal involutive subdistribution in $\mathcal{V}[\mathcal{R}_2]$—defining a flat Vessiot connection—which can not be constructed step by step because the symbol of the system is not involutive.

In Remark 2.3.4 we ran into a similar problem concerning the step-by-step construction of formal power series solutions to a differential equation that is not formally integrable: there may be formal power series solutions for such a differential equation but they cannot be constructed step by step.

3.3.7 The Existence Theorem for Flat Vessiot Connections

At this point, we have proven that integral distributions within the Vessiot distribution exist if, and only if, the algebraic conditions (3.29) are solvable, and that this is equivalent to the augmented rank condition (3.41) being satisfied. This in turn is the case exactly if the differential equation is involutive. Now we have characterized the existence of Vessiot connections for Vessiot's [43] step-by-step approach.

There remains to analyze the solvability, if we add the differential system (3.30). Its solvability is equivalent to the existence of flat Vessiot connections in that each flat Vessiot connection of $\mathcal{R}_1$ corresponds to a solution of the combined system (3.29, 3.30). We first note that the set of differential conditions (3.30) alone is again an involutive system.

Proposition 3.3.27. *The differential conditions (3.30) represent an involutive differential equation of first order.*

Proof. For all $1 \leq i \leq n$ and $1 \leq p \leq r$, the independent variables of the functions $\zeta_i^p \in \mathcal{F}(\mathcal{R}_1)$ are the coordinates on $\mathcal{R}_1$, which are $\mathbf{x}, \mathbf{u}$ and all u_h^α such that $(\alpha, h) \notin \mathcal{B}$. To apply a vector field $U_j = \partial_{x^j} + \cdots$ to a function ζ_i^p includes a derivation with respect to x^j. We order the independent variables such that if $j > i$, then x^j is greater than x^i, and each x^i is greater than all the variables u^α and u_h^α where $(\alpha, h) \notin \mathcal{B}$. For any equation H_{ij}^p within the system (3.30), the application of the vector field $U_j = \partial_{x^j} + \cdots$ to ζ_i^p yields $\partial\zeta_i^p/\partial x^j$ as the leader of that equation; therefore equation H_{ij}^p is of class j, and the equations of maximal class are H_{in}^p; the equations of second highest class in the system are H_{in-1}^p and so on. There are only equations H_{ij}^p of a class indicated by some index $2 \leq j \leq n$.

From the Jacobi identity for vector fields U_i, U_j and U_k where $1 \leq i < j < k \leq n$, we have

$$[U_i, [U_j, U_k]] + [U_j, [U_k, U_i]] + [U_k, [U_i, U_j]] = 0 \ .$$

The structure equations (3.28) for the vector fields U_h and the definitions of G_{ij}^c and H_{ij}^p in Equations (3.29) and (3.30) imply that this is

$$\begin{aligned}
0 &= [U_i, \Gamma_{jk}^h U_h + G_{jk}^c Z_c + H_{jk}^p Y_p] \\
&\quad + [U_j, \Gamma_{ki}^h U_h + G_{ki}^c Z_c + H_{ki}^p Y_p] \\
&\quad + [U_k, \Gamma_{ij}^h U_h + G_{ij}^c Z_c + H_{ij}^p Y_p] \\
&= [U_i, \Gamma_{jk}^h U_h] + [U_i, G_{jk}^c Z_c] + [U_i, H_{jk}^p Y_p] \\
&\quad + [U_j, \Gamma_{ki}^h U_h] + [U_j, G_{ki}^c Z_c] + [U_j, H_{ki}^p Y_p] \\
&\quad + [U_k, \Gamma_{ij}^h U_h] + [U_k, G_{ij}^c Z_c] + [U_k, H_{ij}^p Y_p] \\
&= \Gamma_{jk}^h [U_i, U_h] + U_i(\Gamma_{jk}^h) U_h + G_{jk}^c [U_i, Z_c] + U_i(G_{jk}^c) Z_c + H_{jk}^p [U_i, Y_p] + U_i(H_{jk}^p) Y_p \\
&\quad + \Gamma_{ki}^h [U_j, U_h] + U_j(\Gamma_{ki}^h) U_h + G_{ki}^c [U_j, Z_c] + U_j(G_{ki}^c) Z_c + H_{ki}^p [U_j, Y_p] + U_j(H_{ki}^p) Y_p \\
&\quad + \Gamma_{ij}^h [U_k, U_h] + U_k(\Gamma_{ij}^h) U_h + G_{ij}^c [U_k, Z_c] + U_k(G_{ij}^c) Z_c + H_{ij}^p [U_k, Y_p] + U_k(H_{ij}^p) Y_p \ .
\end{aligned}$$

The combined system (3.29, 3.30) means that all $G_{ab}^c = 0$ and all $H_{ab}^p = 0$ which implies that $\mathcal{U}$ is involutive, which it is, being in triangular form, exactly if all $[U_a, U_b] = 0$. This leaves only

$$\begin{aligned}
0 \ &= \{U_i(\Gamma_{jk}^h) + U_j(\Gamma_{ki}^h) + U_k(\Gamma_{ij}^h)\} \, U_h \\
&\quad + \{U_i(G_{jk}^c) + U_j(G_{ki}^c) + U_k(G_{ij}^c)\} \, Z_c \\
&\quad + \{U_i(H_{jk}^p) + U_j(H_{ki}^p) + U_k(H_{ij}^p)\} \, Y_p \ .
\end{aligned}$$

As part of a basis for $\mathcal{V}'[\mathcal{R}_q]$, the vector fields U_h, Z_c and Y_p are linearly independent, which means their coefficients must vanish individually. So in particular

$$U_i(H_{jk}^p) + U_j(H_{ki}^p) + U_k(H_{ij}^p) = 0 \ .$$

Under the assumption $i < j < k$, the term $U_k(H^p_{ij})$ contains derivations with respect to x_k of $U_i(\zeta^p_j)$ and $U_j(\zeta^p_i)$. Thus, according to our order, this is a non-multiplicative prolongation, and the remaining terms are multiplicative prolongations. But since any non-multiplicative prolongation within the system (3.30) must be of such a form, it is a linear combination of multiplicative prolongations. Therefore, no integrability conditions arise from cross-derivatives (and none arise from a prolongation of lower order equations since all equations of the system are of first order).

If we set

$$\hat{B}^p_{ij} + \zeta^\ell_j \tilde{B}^p_{i\ell} - \zeta^k_i \tilde{B}^p_{jk} - \zeta^\ell_j \zeta^k_i B^p_{\ell k} - \zeta^p_h \Gamma^h_{ij} =: \Delta^p_{ij}$$

for the inhomogeneous parts of the system (3.30), $\partial\zeta^p_i/\partial x^j =: (\zeta^p_i)_j$ for the leaders and

$$\tilde{U}_j(\zeta^p_i) := U_j(\zeta^p_i) - (\zeta^p_i)_j + \Delta^p_{ij}$$

and solve each equation of the system (3.30) for its leader, then it takes the form

$$\begin{aligned}
(\zeta^p_1)_n &= U_1(\zeta^p_n) - \tilde{U}_n(\zeta^p_1) + \Delta^p_{1n}\ ,\\
(\zeta^p_2)_n &= U_2(\zeta^p_n) - \tilde{U}_n(\zeta^p_2) + \Delta^p_{2n}\ ,\\
&\ \vdots\\
(\zeta^p_{n-1})_n &= U_{n-1}(\zeta^p_n) - \tilde{U}_n(\zeta^p_{n-1}) + \Delta^p_{n-1,n}\ ,\\
(\zeta^p_1)_{n-1} &= U_1(\zeta^p_{n-1}) - \tilde{U}_{n-1}(\zeta^p_1) + \Delta^p_{1,n-1}\ ,\\
(\zeta^p_2)_{n-1} &= U_2(\zeta^p_{n-1}) - \tilde{U}_{n-1}(\zeta^p_2) + \Delta^p_{2,n-1}\ ,\\
&\ \vdots\\
(\zeta^p_{n-2})_{n-1} &= U_{n-2}(\zeta^p_{n-1}) - \tilde{U}_{n-1}(\zeta^p_{n-2}) + \Delta^p_{n-2,n-1}\ ,\\
&\ \vdots\\
(\zeta^p_2)_3 &= U_2(\zeta^p_3) - \tilde{U}_3(\zeta^p_2) + \Delta^p_{23}\ ,\\
(\zeta^p_1)_3 &= U_1(\zeta^p_3) - \tilde{U}_3(\zeta^p_1) + \Delta^p_{13}\ ,\\
(\zeta^p_1)_2 &= U_1(\zeta^p_2) - \tilde{U}_2(\zeta^p_1) + \Delta^p_{12}\ ;
\end{aligned} \tag{3.30*}$$

here for each line $1 \leq p \leq r$. Therefore the system (3.30*) is in Cartan normal form given in Definition 2.4.26. From Lemma 2.4.29 now follows that the system is involutive. □

If the original equation $\mathcal{R}_1$ is analytic, then the quasi-linear system (3.30) is analytic, too. Thus we may apply the Cartan-Kähler theorem 2.4.31 to it, which guarantees the existence of solutions.

The problem remains that the combined system (3.29, 3.30) is in general not involutive, as the prolongation of the algebraic equations (3.29) leads to additional differential equations. Instead of analyzing the effect of these integrability conditions, we proceed as follows. If we assume that $\mathcal{R}_1$ is involutive, then we know from Theorem 3.3.9 that the algebraic equations (3.29) are solvable. Now we use the interrelation between the unknowns ζ^p_ℓ as shown in Corollary 3.3.23 to eliminate in (3.30) some of the unknowns

ζ_ℓ^p by expressing them as linear combinations of the remaining ones; that is, we plug the algebraic conditions into the differential conditions and thus get rid of them. We can then prove the following existence theorem for flat Vessiot connections.

Theorem 3.3.28. *Assume that δ-regular coordinates have been chosen for the differential equation $\mathcal{R}_1$ and that $\mathcal{R}_1$ is analytic. Then the combined system (3.29, 3.30) is solvable.*

Proof. First we transform the generators of $\mathcal{U}$ into a triangular system. Then the inhomogeneous terms $\Delta_{ij}^p := \hat{B}_{ij}^p + \zeta_j^\ell \tilde{B}_{i\ell}^p - \zeta_i^k \tilde{B}_{jk}^p - \zeta_j^\ell \zeta_i^k B_{\ell k}^p - \zeta_h^p \Gamma_{ij}^h$ in the differential conditions

$$U_i(\zeta_j^p) - U_j(\zeta_i^p) + \Delta_{ij}^p = 0 \tag{3.30}$$

vanish, and it suffices to consider the system (3.29, 3.30′) instead of (3.29, 3.30). We follow the strategy outlined above and eliminate some of the unknowns ζ_ℓ^p. As we consider each of the equations of (3.30) as being solved for its derivative $\partial\zeta_i^{(\beta,h)}/\partial x^j$ of highest class j, as given in Equation (3.30*), we must take a closer look only at those equations where this leading derivative is of one of the unknowns we eliminate. The structure of the vectors ζ_i, given in Corollary 3.3.23, shows which ones these are. Let k be such that $2 \leq k \leq n$. Then for the subsystem of the equations of class k in the system (3.30), the equations which hold the following terms are concerned:

$$\begin{aligned}
&U_k(\zeta_2^{(\beta,1)})\,,\\
&U_k(\zeta_3^{(\beta,1)})\,,\ U_k(\zeta_3^{(\beta,2)})\,,\\
&\quad\vdots\\
&U_k(\zeta_{k-1}^{(\beta,1)})\,,\ U_k(\zeta_{k-1}^{(\beta,2)})\,,\dots,\ U_k(\zeta_{k-1}^{(\beta,k-2)})\ ;
\end{aligned}$$

here, for any $U_k(\zeta_i^{(\beta,h)})$, we have $\beta_1^{(h)} + 1 \leq \beta \leq m$. We now show that these equations vanish. The proof is by straightforward calculation, though tedious and requiring a case distinction. Let $1 < i < k$. Fix some $\beta_1^{(h)} + 1 \leq \beta \leq m$. Consider the equation

$$U_i(\zeta_k^{(\beta,h)}) = U_k(\zeta_i^{(\beta,h)})\ . \tag{3.80}$$

Then $h < i < k$. According to the structure of the vector ζ_i, the entries of which in its hth block are of two kinds, there are two cases.

1. The interrelation for $\zeta_i^{(\beta,h)}$ is an equality: $\zeta_i^{(\beta,h)} = \zeta_h^{(\beta,i)}$. This is so if, and only if, $\beta_1^{(i)} + 1 \leq \beta \leq m$ according to the structure of ζ_i. Now there arise two subcases.

   (a) The other interrelation is an equality, too: $\zeta_k^{(\beta,h)} = \zeta_h^{(\beta,k)}$. This is so if, and only if, $\beta_1^{(k)} + 1 \leq \beta \leq m$ according to the structure of ζ_k. In this subcase, Equation (3.80) becomes

$$U_i(\zeta_h^{(\beta,k)}) = U_k(\zeta_h^{(\beta,i)})\ . \tag{3.81}$$

Since the system (3.30) contains the equalities $U_i(\zeta_h^{(\beta,k)}) = U_h(\zeta_i^{(\beta,k)})$ and $U_k(\zeta_h^{(\beta,i)}) = U_h(\zeta_k^{(\beta,i)})$, Equation (3.81) becomes

$$U_h(\zeta_i^{(\beta,k)}) = U_h(\zeta_k^{(\beta,i)}) \ . \tag{3.82}$$

Since $i < k$ and $\beta_1^{(k)}+1 \leq \beta \leq m$, from the structure of ζ_k follows $\zeta_i^{(\beta,k)} = \zeta_k^{(\beta,i)}$. Thus, Equation (3.80) vanishes.

(b) The other interrelation is an affine-linear combination:

$$\zeta_k^{(\beta,h)} = \sum_{a=1}^{k} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_k^\beta)\zeta_h^{(\gamma,a)} + C_h^{(1)}(\phi_k^\beta) \ .$$

This is so if, and only if, $\beta_1^{(i)} + 1 \leq \beta \leq \beta_1^{(k)}$ according to the structure of ζ_k. In this subcase, the term $U_i(\zeta_k^{(\beta,h)})$ in Equation (3.80) becomes

$$U_i(\zeta_k^{(\beta,h)}) = \sum_{a=1}^{k} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_k^\beta)U_i(\zeta_h^{(\gamma,a)}) \tag{3.83a}$$

$$+\sum_{a=1}^{k} \sum_{\gamma=\beta_1^{(a)}+1}^{m} U_i(C_\gamma^a(\phi_k^\beta))\zeta_h^{(\gamma,a)} + U_i(C_h^{(1)}(\phi_k^\beta)) \ . \tag{3.83b}$$

The term $U_k(\zeta_i^{(\beta,h)})$ in Equation (3.80) becomes

$$\begin{aligned} U_k(\zeta_i^{(\beta,h)}) &= U_k(\zeta_h^{(\beta,i)}) \\ &= U_h(\zeta_k^{(\beta,i)}) \\ &= U_h\Biggl(\sum_{a=1}^{k} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_k^\beta)\zeta_i^{(\gamma,a)} + C_i^{(1)}(\phi_k^\beta)\Biggr) \\ &= \sum_{a=1}^{k} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_k^\beta)U_h(\zeta_i^{(\gamma,a)}) \end{aligned} \tag{3.84a}$$

$$+\sum_{a=1}^{k} \sum_{\gamma=\beta_1^{(a)}+1}^{m} U_h(C_\gamma^a(\phi_k^\beta))\zeta_i^{(\gamma,a)} + U_h(C_i^{(1)}(\phi_k^\beta)) \ ; \tag{3.84b}$$

here we have the first equality because we are considering the first main case, the second equality because of the structure of the system (3.30) and the third equality according to the structure of ζ_k, since $i < k$ and because $\beta_1^{(i)} + 1 \leq \beta \leq \beta_1^{(k)}$. Substituting (3.83) and (3.84) in Equation (3.80) and factoring out, we get

$$0 = \sum_{a=1}^{k} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_k^\beta)\big\{U_i(\zeta_h^{(\gamma,a)}) - U_h(\zeta_i^{(\gamma,a)})\big\} \tag{3.85a}$$

$$+ \sum_{a=1}^{k} \sum_{\gamma=\beta_1^{(a)}+1}^{m} \big\{U_i(C_\gamma^a(\phi_k^\beta))\zeta_h^{(\gamma,a)} - U_h(C_\gamma^a(\phi_k^\beta))\zeta_i^{(\gamma,a)}\big\} \tag{3.85b}$$

$$+ U_i(C_h^{(1)}(\phi_k^\beta)) - U_h(C_i^{(1)}(\phi_k^\beta)) \ . \tag{3.85c}$$

Line (3.85c) contains the Lie bracket $[U_i, C_h^{(1)}](\phi_k^\beta)$. According to the structure of the system (3.30), the term (3.85a) vanishes. If the terms (3.85b) and (3.85c) vanish, too, then so does equation (3.30). Otherwise they form a new algebraic condition for (3.30), which can be solved for some function $\zeta_h^{(\beta,a)}$. Substituting this function in (3.30) does not change the classes or the numbers of the single equations therein. Thus, equation (3.80) vanishes.

2. The interrelation for $\zeta_i^{(\beta,h)}$ is an affine-linear combination:

$$\zeta_i^{(\beta,h)} = \sum_{a=1}^{i} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_i^\beta)\zeta_h^{(\gamma,a)} + C_h^{(1)}(\phi_i^\beta) \ .$$

This is so if, and only if, $\beta_1^{(h)} + 1 \leq \beta \leq \beta_1^{(i)}$ according to the structure of ζ_i. Since we have $h < i < k$ and $\beta_1^{(i)} \leq \beta_1^{(k)}$, according to the structure of ζ_k the other interrelation is an affine-linear combination, too:

$$\zeta_k^{(\beta,h)} = \sum_{b=1}^{i} \sum_{\delta=\beta_1^{(b)}+1}^{m} C_\delta^b(\phi_k^\beta)\zeta_h^{(\delta,b)} + C_h^{(1)}(\phi_k^\beta) \ .$$

Thus, Equation (3.80) becomes

$$0 = \sum_{a=1}^{i} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_i^\beta)U_k(\zeta_h^{(\gamma,a)}) - \sum_{b=1}^{k} \sum_{\delta=\beta_1^{(b)}+1}^{m} C_\delta^b(\phi_k^\beta)U_i(\zeta_h^{(\delta,b)}) \tag{3.86a}$$

$$+ \sum_{a=1}^{i} \sum_{\gamma=\beta_1^{(a)}+1}^{m} U_k(C_\gamma^a(\phi_i^\beta))\zeta_h^{(\gamma,a)} - \sum_{b=1}^{k} \sum_{\delta=\beta_1^{(b)}+1}^{m} U_i(C_\delta^b(\phi_k^\beta))\zeta_h^{(\delta,b)} \tag{3.86b}$$

$$+ U_k(C_h^{(1)}(\phi_i^\beta)) - U_i(C_h^{(1)}(\phi_k^\beta)) \ . \tag{3.86c}$$

In part (3.86a), the terms $U_k(\zeta_h^{(\gamma,a)})$ and $U_i(\zeta_h^{(\delta,b)})$ are equal to $U_h(\zeta_k^{(\gamma,a)})$ and $U_h(\zeta_i^{(\delta,b)})$ according to the structure of the system (3.30). Thus, Equation (3.86) becomes

$$0 = \sum_{a=1}^{i} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_i^\beta)U_h(\zeta_k^{(\gamma,a)}) - \sum_{b=1}^{k} \sum_{\delta=\beta_1^{(b)}+1}^{m} C_\delta^b(\phi_k^\beta)U_h(\zeta_i^{(\delta,b)}) \tag{3.86a'}$$

$$+ (3.86b) + (3.86c) \ .$$

Factoring out the vector field U_h in part (3.86a′), this equals

$$0 = U_h\Bigl(\sum_{a=1}^{i}\sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_i^\beta)\zeta_k^{(\gamma,a)} - \sum_{b=1}^{k}\sum_{\delta=\beta_1^{(b)}+1}^{m} C_\delta^b(\phi_k^\beta)\zeta_i^{(\delta,b)}\Bigr) \tag{3.87a}$$

$$- \Bigl(\sum_{a=1}^{i}\sum_{\gamma=\beta_1^{(a)}+1}^{m} U_h(C_\gamma^a(\phi_i^\beta))\zeta_k^{(\gamma,a)} - \sum_{b=1}^{k}\sum_{\delta=\beta_1^{(b)}+1}^{m} U_h(C_\delta^b(\phi_k^\beta))\zeta_i^{(\delta,b)}\Bigr) \tag{3.87b}$$

$$+ (3.86\text{b}) + (3.86\text{c})\ . \tag{3.87c}$$

According to Corollary 3.3.25 (for $j = k$), the term (3.87a) equals

$$U_h\bigl(C_i^{(1)}(\phi_k^\beta) - C_k^{(1)}(\phi_i^\beta)\bigr),$$

which does not contain any $\zeta_k^{(\gamma,a)}$ or $\zeta_i^{(\delta,b)}$ any more; it is an algebraic expression instead of the differential expression that it seems to be when written in the form (3.87a). The other terms, (3.87b) and (3.87c), are algebraic, too. So all of Equation (3.80) has shown to be an algebraic condition when the interrelations between the entries of the vectors ζ_h, ζ_i and ζ_k, as noted in Corollary 3.3.23, are taken into account.

If this new algebraic condition for the system (3.30) vanishes, Equation (3.80) vanishes. Otherwise, this new algebraic condition given in Equation (3.87) now appears as

$$0 = U_h\bigl(C_i^{(1)}(\phi_k^\beta) - C_k^{(1)}(\phi_i^\beta)\bigr) \tag{3.87a′}$$

$$- \Bigl(\sum_{a=1}^{i}\sum_{\gamma=\beta_1^{(a)}+1}^{m} U_h(C_\gamma^a(\phi_i^\beta))\zeta_k^{(\gamma,a)} - \sum_{b=1}^{k}\sum_{\delta=\beta_1^{(b)}+1}^{m} U_h(C_\delta^b(\phi_k^\beta))\zeta_i^{(\delta,b)}\Bigr) \tag{3.87b}$$

$$+ \sum_{a=1}^{i}\sum_{\gamma=\beta_1^{(a)}+1}^{m} U_k(C_\gamma^a(\phi_i^\beta))\zeta_h^{(\gamma,a)} - \sum_{b=1}^{k}\sum_{\delta=\beta_1^{(b)}+1}^{m} U_i(C_\delta^b(\phi_k^\beta))\zeta_h^{(\delta,b)} \tag{3.86b}$$

$$+ U_k(C_h^{(1)}(\phi_i^\beta)) - U_i(C_h^{(1)}(\phi_k^\beta))\ . \tag{3.86c}$$

Collecting terms in lines (3.87a′) and (3.86c), this yields

$$\begin{aligned} 0 &= (3.86\text{b}) + (3.87\text{b}) \\ &\quad + U_h(C_i^{(1)}(\phi_k^\beta)) - U_i(C_h^{(1)}(\phi_k^\beta)) + U_k(C_h^{(1)}(\phi_i^\beta)) - U_h(C_k^{(1)}(\phi_i^\beta))\ . \end{aligned}$$

The lower line contains the Lie brackets $[U_h, C_i^{(1)}](\phi_k^\beta)$ and $[U_k, C_h^{(1)}](\phi_i^\beta)$. There must be some non-vanishing summand containing a factor $\zeta_k^{(\gamma,a)}$, $\zeta_i^{(\delta,b)}$, $\zeta_h^{(\gamma,a)}$ or $\zeta_h^{(\delta,a)}$. As we did in case 1. (b), we solve (3.86) for this non-vanishing factor and substitute it into the system (3.30), which does not change the class of any equation therein. Therefore Equation (3.80) drops out from the system (3.30).

Now we have shown that all those equations vanish where the leading derivative is subject to being substituted through the interrelations concerning the coefficient function $\zeta_k^{(\beta,i)}$. In the system (3.30*), these are the equations with the leaders

$$\begin{aligned}
&(\zeta_2^{(\beta,1)})_k \;, \\
&(\zeta_3^{(\beta,1)})_k, (\zeta_3^{(\beta,2)})_k \;, \\
&\quad\vdots \\
&(\zeta_{k-1}^{(\beta,1)})_k, (\zeta_{k-1}^{(\beta,2)})_k, \dots, (\zeta_{k-1}^{(\beta,k-2)})_k \;;
\end{aligned}$$

here $2 \leq k \leq n$ and $\beta_1^{(h)} + 1 \leq \beta \leq m$. The remaining equations still form an involutive system (we may numerate the remaining ζ_i^p in such a way that no gaps appear) as the considerations for the system (3.30) in Proposition 3.3.27 apply likewise. Thus we eventually arrive at an analytic involutive differential equation for the coefficient functions ζ_i^k which is solvable according to the Cartan-Kähler theorem 2.4.31. □

Example 3.3.29. Consider the first-order equation

$$\mathcal{R}_1 \colon \begin{cases} u_t = v_t = w_t = u_s = 0 \;, & v_s = 2u_x + 4u_y \;, \\ w_s = -u_x - 3u_y \;, & u_z = v_x + 2w_x + 3v_y + 4w_y \;. \end{cases}$$

It is formally integrable, and its symbol is involutive with $\dim \mathcal{N}_1 = 8$. Thus $\mathcal{R}_1$ is an involutive equation. For the matrices Ξ_i, all of which have three rows and eight columns, we find

$$\Xi_1 = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \Xi_2 = \begin{pmatrix} 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{pmatrix},$$

$$\Xi_3 = \begin{pmatrix} 0 & -1 & -2 & 0 & -3 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}, \quad \Xi_4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & -4 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\Xi_5 = 0_{3\times 8} \;.$$

For the first two steps in the construction of the fields U_i, the rank conditions are trivially satisfied even for the non-contracted matrices. But not so in the third step where we have in the row echelon form of the arising 9×32-matrix in the 7th row zero entries throughout except in the 12th column (where we have -2) and in the 17th column (where we have 2). As a consequence, we obtain the equality $\zeta_1^4 = \zeta_2^1$ and the rank condition for this step does not hold. However, since both u_x and u_y are parametric derivatives and in our ordering $Y_1 = \iota_*(\partial_{u_x})$ and $Y_4 = \iota_*(\partial_{u_y})$, this equality is already taken into account in our reduced ansatz and for the matrices $\hat{\Xi}_i$ the rank condition is satisfied.

Note that the rank condition is first violated when the rank reaches the symbol dimension. From then on, the rank of the left matrix in Equation (3.41) for the rank condition stagnates at $\dim \mathcal{N}_1$ while the rank of the augmented matrix may rise further. The entries breaking the rank condition differ by their sign, while their corresponding coefficients in Lemma 2.5.8 are collected into one sum and thus vanish.

Remark 3.3.30. One of Vessiot's [43] aims was to give a proof of his own for the Cartan-Kähler theorem 2.4.31. Within this same argumentation, he developed his step-by-step approach for the construction of transversal involutive subdistributions within the Vessiot distribution, thus striving for two separate objectives at once. Since, in contrast, we split both parts of this argument by basing the Vessiot theory on the formal theory, we need not prove the Cartan-Kähler theorem while analyzing the step-by-step construction at the same time.

As with Vessiot's original approach, our strategy guarantees only the existence of an analytical solution. But we have an advantage: if for some system under consideration we can show that the differential conditions (3.30) are solvable without turning to the Cartan-Kähler theorem—for the special equation at hand some other argument may be available—then our approach yields the existence of a unique solution in the smooth category because of the remarkable setting: to prove the existence of the necessary involutive subdistributions within the Vessiot distribution, we need the Cartan-Kähler theorem. But once we have them, we use the Frobenius theorem (as is outlined in Remark 3.1.12). Now the Frobenius theorem refers to smooth solutions as apposed to analytical ones: here the integral manifolds are C^∞-manifolds. Therefore, by dividing both parts of the argument in two—proving the Cartan-Kähler theorem and developing the step-by-step approach—for those differential equations where the differential conditions (3.30) can be shown to be solvable without recourse to the Cartan-Kähler theorem, it is possible to sidestep the Cartan-Kähler theorem altogether, gaining a stronger existence theorem.

Chapter 4

Possible Further Developments

In this closing chapter we hint at some possible further developments. Apart from an interest in Vessiot theory on aesthetic grounds there is the question of how it can be applied or linked to other ideas.

Cartan-Kähler theory Now that we have related the approach of Vessiot's theory with the central concepts of the formal theory by proving that Vessiot's construction succeeds if, and only if, it is applied to an involutive differential equation, the relation to the Cartan-Kähler theory of exterior differential systems may be considered. As Cartan-Kähler theory is dual to Vessiot's approach, it seems natural to assume that the formal theory and the Cartan-Kähler theory are equivalent, too. Malgrange [31] discusses this point in an appendix. Here, equivalence means, the pull-back of the contact codistribution is involutive in the sense of Cartan-Kähler theory if, and only if, the differential equation is involutive in the sense of formal theory. As Vessiot theory is intermediate between Cartan-Kähler theory and formal theory, it should facilitate an explicit equivalence proof.

To prove the equivalence of formal theory and Cartan-Kähler theory one can use now Proposition 2.4.22 on how the integral elements fit in the formal theory as the new definition of integral elements based on the contact map makes the relations between the formal theory and the Cartan-Kähler theory more transparent. One should clarify the following point: Matsushima [33] considers the link between Cartan's test in the version of the theory of exterior systems [4, 21] and Cartan's test as it appears in the formal theory based on intrinsic concepts of Spencer cohomology as given by Seiler [37, 39]. Both times the comparison of Cartan characters is linked to the involutive symbols of the formal theory, and the intention should be to formulate the role of the symbol and the Spencer cohomology clearly.

Qualitative Investigations One also could try to develop methods from Vessiot theory into an algorithm for the completion of differential equations to involutive systems. But as it turns out, the method of formal theory is at least as simple. Then again, completion to involution is not an end in itself. Indeed, it is at the beginning in the analysis of over- and under-determined systems. One could proceed to qualitative investigations: can the geometrical structure of the differential equation—its Vessiot distribution and the involutive subdistributions in it—be used to classify differential equations? Can it distinguish, for example, an elliptic system from a hyperbolic system? There are such considerations for exterior systems. Vassiliou [42] studies for hyperbolic equations algebraic structures within the Vessiot distribution: here, the Vessiot distribution is constructed as a direct sum of two subdistributions which correspond to the characteristics.

As Vessiot theory is dual to the theory of exterior systems one can translate the methods of Cartan-Kähler theory into methods of Vessiot theory. Not to replace one approach with the other; but if a question is posed in the language of formal theory, it should be an advantage to stick to it throughout, without the need to reformulate it as a question regarding exterior systems, just like when a question arises naturally within the context of exterior systems one does not reformulate it as a partial differential equation. Now Vessiot theory offers the development of analogues. This would avoid the indefiniteness in the transcription, too. Though there is a manifest way to transform a problem concerning a differential equation into a problem concerning an exterior system (pulling back the contact codistribution), there are several other approaches as well: the Korteweg-de-Vries equation $u_{xxx} = 6uu_x - u_y$ rewritten as an exterior system the way Estabrook and Wahlquist [14] do it is not a Pfaffian system but contains higher order forms.

Vessiot theory may help to clarify phenomena concerning singularities as Arnold [3] studies them for ordinary differential equations; this is indicated in Remark 3.1.9.

We have assumed throughout that the Vessiot distribution has constant rank; a more general consideration would be to omit this restriction. For a non-linear differential equation, the Vessiot distribution is not a distribution in the strict sense any more as its rank then may vary for different points on the differential equation.

The algorithm for the completion to involution then does not complete one single differential equation but has to work out a tree of subsystems, where the terminal knots of which, the leaves, correspond to different systems of differential equations which are regular in that the Vessiot distribution has constant rank on each of them. The question is being studied by Marcus Hausdorf [19].

Differential Galois Theory Now that we know the Vessiot approach is successful for an involutive system, we can regard the differential equation as covered with systems of finite type, as any involutive subdistribution within the Vessiot distribution corresponds to a system of finite type. In differential Galois theory systems of finite type are considered; now if a differential equation of non-finite type is given, and one regards it as covered by systems of finite type, this should offer a way to confer the methods of differential Galois theory to such general systems.

Malgrange [30] proposes a Galois theory for nonlinear differential equations where his starting-point is the projection of the Vessiot distribution from $TJ_1\pi$ to $T\mathcal{E}$. The interest of this theory concentrates on the case when there are singularities of the kind that arise at those points where the Vessiot distribution is not transversal any more (as opposed to the points where the Vessiot distribution is of different rank). Those points of the jet bundle are projected onto the zero vector, and the projected distribution is not of constant rank any more. An example is provided by the Clairaut equation

$u = x\dot{u} + f(\dot{u})$, the symbol matrix of which is $(x + f'(\dot{u}))$. Among the solutions, there is an envelope which is not represented within a general solution (a family of straight lines, $u(x) = cx + f(c)$). The singular points constitute a fibred submanifold which locally represents a differential equation that is over-determined to the extent that it admits only one solution (a parabola). To analyze these singularities, one has to prefer the Vessiot distribution on the level of $J_1\pi$ to its projected version in $T\mathcal{E}$.

Symmetries Another point for further research is to analyze symmetries using the Vessiot distribution. Lie's classical approach is to analyze contact transformations of the jet bundle (where a diffeomorphism $\phi\colon J_q\pi \to J_q\pi$ is called a *contact transformation*, if it satisfies $T\phi(\mathcal{C}_q) = \mathcal{C}_q$) to check which of them leave the differential equation invariant: if $\phi(\mathcal{R}_q) \subseteq \mathcal{R}_q$, then ϕ is called an *external symmetry*. But then one uses too much of the contact distribution which is irrelevant because it is not tangential to the differential equation, and thus some symmetries may be overlooked in some cases. An *inner symmetry* is a diffeomorphism $\phi\colon \mathcal{R}_q \to \mathcal{R}_q$ with $\phi^*(\iota^*\mathcal{C}_q^0) \subseteq \iota^*\mathcal{C}_q^0$ and $\phi(\mathcal{R}_q) \subseteq \mathcal{R}_q$. Cartan [6] computed inner symmetries. Olver [34] and Kamran [24] and both of them with Anderson [2] distinguish inner, external and generalized symmetries and compute them for special differential equations. For the Hilbert-Cartan equation $v_x = (u_{xx})^2$ they find that the Lie algebra of external symmetries is six-dimensional and contained in the Lie algebra of inner symmetries, which equals the Lie algebra of generalized symmetries and has dimension fourteen.

Now another way to compute inner symmetries of partial differential equations is this: instead of regarding a differential equation as a fibred submanifold in a jet bundle consider it as a manifold with a distribution $\mathcal{V}[\mathcal{R}_q]$. (Now this distribution encodes the information of imbedding $\mathcal{R}_q$ into $J_q\pi$.) Then analyze the symmetries of this distribution. This yields inner symmetries without further ado, while external symmetries are avoided. This would be the natural geometric approach to the analysis of symmetries. When is there a difference between inner and external symmetries? For systems of Cauchy-Kovalevskaya type (well-determined equations) there are none according to Anderson, Kamran and Olver [2]: any inner symmetry can be extended to a external symmetry. As the different kinds of symmetries were studied separately until fairly recently, general criteria to characterize those systems where there is a difference between them are not yet available. (Ordinary and under-determined differential equations like the Hilbert-Cartan equation allow such a difference.) To develop a geometrical insight instead of several technical criteria for special types of differential equations, Vessiot's theory may be helpful because there, internal symmetries appear as an intrinsic concept.

Field Theories In [16] an intrinsic definition of Hamiltonian differential equations as fibred submanifolds of a first order jet bundle with a one-dimensional base space for the description of explicitly time-dependent systems is given. A natural extension of this theory is to field theories where the base space may have arbitrary dimension. In [7], de Léon and others summarize the various generalizations of the tangent and cotangent structures and bundles that are used in the Lagrangian and Hamiltonian formulations of classical mechanics and classify them into two categories: one where the geometric structure of the bundles is being generalized, which results in several axiomatic systems such as k-symplectic and k-tangent structures; and one where the bundles themselves are being extended and the geometric properties of these extensions are studied, which results in the multi-symplectic geometry on jet and cojet bundles and the n-symplectic geometry on frame bundles. These theories study several distributions of vector fields; their relation to the Vessiot distribution of the differential equation which describes a dynamical system is an open question.

Chapter 5

Deutsche Zusammenfassung

Forschungsgegenstand der Arbeit ist Vessiots Theorie partieller Differentialgleichungen: Zu einem gegebenen System von Differentialgleichungen sucht man eine Distribution von Vektorfeldern, von deren Unterdistributionen manche aufgefasst werden können als tangentiale Näherungen an Lösungen der Gleichung.

Ernest Vessiot [43] verfolgte in den zwanziger bis vierziger Jahren des 20. Jahrhunderts einen Ansatz zur Behandlung allgemeiner Systeme partieller Differentialgleichungen, der dual ist zur Theorie äußerer Systeme, der Cartan-Kähler-Theorie [4, 21], insofern, als er Vektorfelder zum zentralen Gegenstand der Betrachtung macht und die äußere Ableitung ersetzt durch die Lie-Klammer von Vektorfeldern. Gewisse Distributionen von Vektorfeldern erlauben dann, Lösungen einer Differentialgleichung als Integralmannigfaltigkeiten dieser Distributionen aufzufassen. Vessiots Ansatz besteht darin, zu einer gegebenen Differentialgleichung eine Distribution zu konstruieren, die tangential ist zur Differentialgleichung und zudem in der Kontaktdistribution des *Jet*-Bündels enthalten. Dann sucht man darin nach n-dimensionalen, zu der Basismannigfaltigkeit transversalen Teildistributionen, den *Integraldistributionen*. Diese bestehen aus *Integralelementen*, und diese wiederum sollen so aneinandergepasst werden, dass sie eine unter der Lie-Klammer schließende Unterdistribution bilden. Man spricht dann von einem *flachen Vessiot-Zusammenhang*.

Vessiots Ansatz ist nicht populär geworden. Modern formulierte Darstellungen seiner Theorie beschränken sich auf spezielle Systeme (wie hyperbolische Gleichungen, siehe [42]), und allgemeinen Betrachtungen [15, 40] fehlt die Rigorosität der Betrachtungen, wie sie in der verbreiteteren Cartan-Kähler-Theorie ausgearbeitet worden ist; speziell die nötigen Voraussetzungen zur Lösbarkeit einer Gleichung und zur Konstruktion oben erwähnter Distributionen sind noch nicht erforscht worden und werden selbst in Vessiots eigenen Arbeiten vernachlässigt. Ein Ergebnis dieser Arbeit ist, diese Lücke geschlossen und für Vessiots Theorie ebenso rigorose Grundlagen geschaffen zu haben. Zudem wird der Zusammenhang zwischen Vessiots Ansatz und den zentralen Begriffen der formalen Theorie (wie formale Integrierbarkeit und Involution von Differentialgleichungen) herausgearbeitet. Ein Hauptergebnis der Arbeit ist zu zeigen, unter welchen Bedingungen Vessiots Ansatz gelingen kann.

Im ersten Teil der Arbeit gebe ich eine aktuelle Übersicht über die formale Theorie partieller Differentialgleichungen. Ich folge in der Darstellung Seiler [37, 38] und Pommaret [35]. Die moderne Beschreibung der formalen Theorie von Differentialgleichung betrachtet Differentialgleichungen als gefaserte Untermannigfaltigkeiten in einem geeigneten *Jet*-Bündel und untersucht formale Integrierbarkeit und das stärkere Konzept der Involutivität von Differentialgleichungen zur Analyse ihrer Lösbarkeit.

In dieser Arbeit werden allgemeine Systeme partieller Differentialgleichungen betrachtet; dies schließt beliebig komplizierte nicht-lineare Systeme ein. Die Struktur dieser Systeme, wie sie durch die Vessiot-Distribution beschrieben wird, lässt sich trotzdem leicht darstellen oder auf einfach darstellbare Systeme zurückführen: Zunächst lässt sich jedes System beliebig hoher Ordnung umschreiben in ein System erster Ordnung. Ist das System nicht involutiv, lässt es sich durch endlich viele Operationen vervollständigen zu einem involutiven System; dies besagt der Satz von Cartan-Kuranishi. Ist das Ausgangssystem involutiv, so gibt es auch ein äquivalentes System erster Ordnung, das involutiv ist. (Die Zahl der Veränderlichen bleibt dabei natürlich nicht unbedingt erhalten. Erhalten bleibt aber die für die Theorie wichtige Zahl an Cartan-Charakteren.) Nachdem ein beliebiges System umgeschrieben ist in ein System erster Ordnung, lässt es sich (lokal) darstellen in der reduzierten Cartan-Normalform. Diese ist in der Literatur nicht üblich, hilft aber in dieser Arbeit wesentlich, die Argumentation zu vereinfachen, da sie die Veränderlichen der lokalen Darstellung der Differentialgleichung auf naheliegende Weise klassifiziert. Dies führt zu einer natürlichen Beschreibung des zugehörigen geometrischen Symboles. Dieses wieder ist als Unterdistribution der Vessiot-Distribution eine entscheidende Hilfe bei der Konstruktion flacher Vessiot-Zusammenhänge: Die Vessiot-Distribution lässt sich nun zerlegen in die direkte Summe des Symboles und eines (nicht eindeutigen) horizontalen Komplementes. Die n-dimensionalen, unter der Lie-Klammer geschlossenen, zu der Basismannigfaltigkeit transversalen Unterdistributionen sind die gesuchten tangentialen Näherungen an die Lösungen der Differentialgleichung. Ihre Existenz zu zeigen, ist nun möglich durch Analyse der Strukturgleichungen. Der hier verwendete Ansatz ist so geschickt, dass diese eine sehr einfache (oder jedenfalls verglichen mit bisherigen Ansätzen einfache) Form haben.

Im zweiten Teil der Arbeit wird gezeigt, dass Vessiots Ansatz zur schrittweisen Konstruktion der gewünschten Distributionen genau dann gelingt, wenn das gegebene System involutiv ist. Bewiesen wird zunächst ein Existenzsatz für Integraldistributionen. Weiter wird ein Existenzsatz für flache Vessiot-Zusammenhänge bewiesen. Die differentialgeometrische Struktur der zugrundeliegenden Systeme wird analysiert und gegenüber anderen Ansätzen vereinfacht (speziell die für die Beweise der Existenzsätze betrachteten Strukturgleichungen). Die möglichen Obstruktionen zur Involution einer Differentialgleichung werden explizit hergeleitet. (Die Darstellung bezieht sich auf Systeme erster Ordnung, was die Allgemeinheit nicht einschränkt und die Übersicht erhöht.)

Die Analyse der Strukturgleichungen liefert nicht nur theoretische Einsichten, sondern auch ein Verfahren, mit dem die Koeffizienten der Vektorfelder, welche die gesuchten Integraldistributionen aufspannen, explizit bestimmt werden können. Dadurch ist eine Implementierung des Verfahrens in dem Computeralgebrasystem `MuPAD` möglich geworden, an der zur Zeit gearbeitet wird.

Bibliography

[1] William W. Adams and Philippe Loustaunau. *An Introduction to Gröbner Bases.* Graduate Studies in Mathematics Volume 3. AMS, sine loco, 1994.

[2] Ian M. Anderson, Niky Kamran, and Peter J. Olver. "Internal, external and generalzed Symmetries". *Adv. in Math.*, 100:53–100, 1993.

[3] Vladimir Igorevich Arnold. *Geometrical Methods in the Theory of Ordinary Differential Equations*, volume 250 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, Berlin, Heidelberg, New York, second edition, 1988.

[4] Robert L. Bryant, Shiing-shen Chern, Robert B. Gardner, Hubert L. Goldschmidt, and Phillip A. Griffiths. *Exterior Differential Systems.* Mathematical Sciences Research Institute Publications 18. Springer-Verlag, New York, 1991.

[5] Jacques Calmet, Marcus Hausdorf, and Werner M. Seiler. "A constructive Introduction to Involution". In R. Akerkar, editor, *Proc. Int. Symp. Applications of Computer Algebra — ISACA 2000*, pages 33–50. Allied Publishers, New Delhi, 2001.

[6] Élie Joseph Cartan. *Les Systèmes Différentiels Extérieurs et leurs Applications Géométriques.* Hermann;, Paris, 1945.

[7] Manuel de Léon, Michael McLean, Larry K. Norris, Angel Rey Roca, and Modesto Salgado. *Geometric Structures in Field Theory.* arXiv:math-ph/0208036, 26.08.2002.

[8] Jules Joseph Drach. "Sur les systèmes complètement orthogonaux dans l'espace à n dimensions et sur la réduction des systèmes différentielle les plus généraux". *Comp. Rend. Acad. Sci.*, 125:598–601, 1897.

[9] Charles Ehresmann. "Sur les Espaces Fibrés Associés à une Variété Differentiable". *Comp. Rend. Acad. Sci.*, 216:628–630, 1943.

[10] Charles Ehresmann. "Les Connections Infinitésimales dans un Espace Fibré Différentiable". In *Colloque de Topologie (Espaces Fibrés)*, pages 29–55. Thone, Liège, 1950.

[11] Charles Ehresmann. "Les Prolongements d'une Variété Différentiable". *Comp. Rend. Acad. Sci.*, 233:598–600, 777–779, 1081–1083, 1951.

[12] Charles Ehresmann. “Les Prolongements d’une Variété Différentiable”. *Comp. Rend. Acad. Sci.*, 234:1028–1030, 1424–1425, 1951.

[13] Charles Ehresmann. “Introduction à la Théorie des Structures Infinitésimals et des Pseudo-Groupes de Lie”. In *Géométrie Différentielle*, Colloq. Intern. Centre Nat. Rech. Sci. 52, pages 97–110. CNRS, Paris, 1953.

[14] Frank B. Estabrook and Hugo D. Wahlquist. “Prolongation Structures of Nonlinear Evolution Equations”. *J. Math. Phys.*, 16:1–7, 1975.

[15] Edward J. Fackerell. “Isovectors and Prolongation Structures by Vessiot’s Vector Field Formulation of Partial Differential Equations”. In R. Martini, editor, *Geometric Aspects of the Einstein Equations and integrable Systems*, Lecture Notes in Physics 239, pages 303–321. Springer-Verlag, Berlin, Heidelberg, 1985.

[16] Dirk Fesser, Dirk Saller, and Werner M. Seiler. “A Differential Equations Approach to Hamiltonian Systems”. *Reports on Mathematical Physics*, 49(2/3):345–359, 2002.

[17] Dirk Fesser and Werner M. Seiler. “Vessiot Connections of Partial Differential Equations”. In Jacques Calmet, Werner M. Seiler, and Robin W. Tucker, editors, *Global Integrability of Field Theories*, Proceedings of GIFT 2006, pages 111–134. Universitätsverlag Karlsruhe, Karlsruhe, 2006.

[18] David Hartley and Robin W. Tucker. “A Constructive Implementation of the Cartan-Kähler Theory of Exterior Differential Systems”. *Journal of Symbolic Computation*, Volume 12:655–667, 1991.

[19] Marcus Hausdorf. *Konstruktive Theorie algebraischer Differentialgleichungen.* In preparation; doctoral thesis to appear at Universität Kassel.

[20] Marcus Hausdorf, Mehdi Sahbi, and Werner M. Seiler. “Delta- and Quasi-regularity for Polynomial Ideals”. In Jacques Calmet, Werner M. Seiler, and Robin W. Tucker, editors, *Global Integrability of Field Theories*, Proceedings of GIFT 2006, pages 179–200. Universitätsverlag Karlsruhe, Karlsruhe, 2006.

[21] Thomas A. Ivey and Joseph M. Landsberg. *Cartan for Beginners: Differential Geometry via Moving Frames and Exterior Differential Systems.* Graduate Studies in Mathematics 61. American Mathematical Society, Providence, 2003.

[22] Maurice Janet. “Sur les Systèmes d’Équations aux Dérivées Partielles”. *J. Math. Pure Appl.*, 3:65–151, 1920.

[23] Josef Janyška and Marco Modugno. “Relations between linear Connections on the Tangent Bundle and Connections on the Jet Bundle of a fibred Manifold”. *Archivum Mathematicum*, 32:281–288, 1996.

[24] Niky Kamran. “Selected Topics in the Geomatrical Study of Differential Equations”. Regional Conference Series in Mathematics Number 96. American Mathematical Society, 2002.

[25] Iosef S. Krasilshchik, Valentin V. Lychagin, and Alexander M. Vinogradov. *Geometry of Jet Spaces and Nonlinear Partial Differential Equations.* Gordon & Breach, New York, 1986.

[26] Boris Kruglikov and Valentin V. Lychagin. "Geometry of Differential Equations". January 2007.

[27] Boris A. Kuperschmidt. *Geometry of Jet Bundles and the Structure of Lagrangian and Hamiltonian Formalisms.* Lecture Notes in Physics 775. Springer-Verlag, Berlin, Heidelberg, 1980.

[28] Masatake Kuranishi. "On E. Cartan's Prolongation Theorem of Exterior Differential Systems". *Amer. J. Math.*, 79:1–47, 1957.

[29] Valentin V. Lychagin. "Homogeneous Geometric Structures and Homogeneous Differential Equations". In Valentin V. Lychagin, editor, *The Interplay between Differential Geometry and Differential Equations*, Americ. Math. Soc. Transl. 167, pages 143–164. American Mathematical Society, Providence, 1995.

[30] Bernard Malgrange. "On Nonlinear Differential Galois Theory". *Chinese Ann. Math, Ser. B*, 23(2):219–226, 2002.

[31] Bernard Malgrange. *Systèmes Différentiels Involutifs.* Panoramas et Synthèses 19. Societé Mathématique de France, Paris, 2005.

[32] Luigi Mangiarotti and Marco Modugno. "Fibered Spaces, Jet Spaces and Connections for Field Theories". In *Proceedings of the International Meeting "Geometry and Physics" in Florence*, pages 135–165. Pitagora Editrice, 1983.

[33] Yozô Matsushima. "On a Theorem Concerning the Prolongation of a Differential System". *Nagoya Math. J.*, 6:1–16, 1953.

[34] Peter J. Olver. *Applications of Lie Groups to Differential Equations.* Graduate Texts in Mathematics 107. Springer-Verlag, New York, Berlin, Heidelberg, 1986.

[35] Jean-François Pommaret. *Systems of Partial Differential Equations and Lie Pseudogroups.* Gordon and Breach, New York, London, Paris, 1978.

[36] David J. Saunders. *The Geometry of Jet Bundles.* Cambridge University Press, Cambridge, 1989.

[37] Werner M. Seiler. *Involution. The formal theory of differential equations and its applications in computer algebra.* In preparation; manuscript accepted for publication by Springer-Verlag.

[38] Werner M. Seiler. *Involution. The formal theory of differential equations and its applications in computer algebra and numerical analysis.* Habilitationschrift. Universität Mannheim, Mannheim, 2002.

[39] Werner M. Seiler. “Spencer Cohomology, Differential Equations, and Pommaret Bases”. In M. Rosenkranz and D.M. Wang, editors, *Gröbner Bases in Symbolic Analysis*, Radon Series on Computational and Applied Mathematics 2, pages 171–219. de Gruyter, Berlin, 2007.

[40] Olle Stormark. *Lie's Structural Approach to PDE Systems*. Cambridge University Press, Cambridge, 2000.

[41] William John Sweeney. “The D-Neumann Problem”. *Acta Mathematicae*, (120):223–251, 1968.

[42] Peter J. Vassiliou. “Coupled Systems of Nonlinear Wave Equations and Finite Dimensional Lie Algebras I”. *Acta Applicandae Mathematicae*, Volume 8(2):107–147, February 1987.

[43] Ernest Vessiot. “Sur une Théorie nouvelle des Problèmes généraux d'Intégration”. *Bull. Soc. Math. Fr.*, Volume 52:336–395, 1924.

[44] Alexander M. Vinogradov. “Geometry of Nonlinear Differential Equations”. *English translation in: Journal Sov. Math.*, Volume 17:1624–1649, 1981.

Index

Hiermit versichere ich, dass ich die vorliegende Dissertation selbständig und ohne unerlaubte Hilfe angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen sind, habe ich als solche kenntlich gemacht. Kein Teil dieser Arbeit ist in einem anderen Promotions- oder Habilitationsverfahren verwendet worden.

Heidelberg, 09. November 2007