Existence and Construction of Vessiot Connections*

Dirk FESSER † *and Werner M. SEILER* ‡

† *IWR, Universität Heidelberg, INF 368, 69120 Heidelberg, Germany*
E-mail: *dirk.fesser@iwr.uni-heidelberg.de*

‡ *AG "Computational Mathematics", Universität Kassel, 34132 Kassel, Germany*
E-mail: *seiler@mathematik.uni-kassel.de*
URL: `http://www.mathematik.uni-kassel.de/~seiler/`

Received May 05, 2009, in final form September 14, 2009; Published online September 25, 2009
doi:10.3842/SIGMA.2009.092

Abstract. A rigorous formulation of Vessiot's vector field approach to the analysis of general systems of partial differential equations is provided. It is shown that this approach is equivalent to the formal theory of differential equations and that it can be carried through if, and only if, the given system is involutive. As a by-product, we provide a novel characterisation of transversal integral elements via the contact map.

Key words: formal integrability; integral element; involution; partial differential equation; Vessiot connection; Vessiot distribution

2000 Mathematics Subject Classification: 35A07; 35A30; 35N99; 58A20

Contents

*This paper is a contribution to the Special Issue "Élie Cartan and Differential Geometry". The full collection is available at http://www.emis.de/journals/SIGMA/Cartan.html

1 Introduction

Constructing solutions for (systems of) partial differential equations is obviously difficult – in particular for non-linear systems. Élie Cartan [4] proposed to construct first infinitesimal solutions or integral elements. These are possible tangent spaces to (prolonged) solutions. Thus they always lead to a linearisation of the problem and their explicit construction requires essentially only straightforward linear algebra. In the Cartan–Kähler theory [3, 12], differential equations are represented by exterior differential systems and integral elements consist of tangent vectors pointwise annihilated by differential forms.

Vessiot [28] proposed in the 1920s a dual approach which does not require the use of exterior differential systems. Instead of individual integral elements, it always considers distributions of them generated by vector fields and their Lie brackets replace the exterior derivatives of differential forms. This approach takes an intermediate position between the formal theory of differential equations [13, 19, 23] and the Cartan–Kähler theory of exterior differential systems. Thus it allows for the transfer of many techniques from the latter to the former one, although this point will not be studied here.

Vessiot's approach may be considered a generalisation of the Frobenius theorem. Indeed, if one applies his theory to a differential equation of finite type, then one obtains an involutive distribution such that its integral manifolds are in a one-to-one correspondence with the smooth solutions of the equation. For more general equations, Vessiot proposed to "cover" the equation with infinitely many involutive distributions such that any smooth solution corresponds to an integral manifold of at least one of them.

Vessiot's theory has not attracted much attention: presentations in a more modern language are contained in [5, 24]; applications have mainly appeared in the context of the Darboux method for solving hyperbolic equations, see for example [27]. While a number of textbooks provide a very rigorous analysis of the Cartan–Kähler theory, the above mentioned references (including Vessiot's original work [28]) are somewhat lacking in this respect. In particular, the question of under what assumptions Vessiot's construction succeeds has been ignored.

The purpose of the present article is to close this gap and at the same time to relate the Vessiot theory with the key concepts of the formal theory like formal integrability and involution (we will not develop it as a dual form of the Cartan–Kähler theory, but from scratch within the formal theory). We will show that Vessiot's construction succeeds if, and only if, it is applied to an involutive system of differential equations. This result is of course not surprising, given the well-known fact that the formal theory and the Cartan–Kähler theory are equivalent. However, to our knowledge an explicit proof has never been given. As a by-product, we will provide a new characterisation of integral elements based on the contact map, making also the relations between the formal theory and the Cartan–Kähler theory more transparent. Furthermore, we simplify the construction of the integral distributions. Up to now, quadratic equations had to be considered (under some assumptions, their solution can be obtained via a sequence of linear systems. We will show how the natural geometry of the jet bundle hierarchy can be exploited for always obtaining a linear system of equations.

This article contains the main results of the first author's doctoral thesis [6] (a short summary of the results has already appeared in [7]). It is organised as follows. The next three sections recall the needed elements from the formal theory of differential equations and provide our new characterisation of transversal integral elements. The following two sections introduce the key concepts of Vessiot's approach: the Vessiot distribution, integral distributions and flat Vessiot connections. Two further sections discuss existence theorems for the latter two based on a step-by-step approach already proposed by Vessiot. As the proofs of some results are fairly technical, the main text contains only an outline of the underlying ideas and full details are given in three appendices. The Einstein convention is sometimes used to indicate summation.

2 The contact structure

Before we outline the formal theory of partial differential equations, we briefly review its underlying geometry: the jet bundle and its contact structure. Many different ways exist to introduce these geometric constructions, see for example [9, 18, 20]. Furthermore, they are discussed in any book on the formal theory (see the references in the next section).

Let $\pi : \mathcal{E} \to \mathcal{X}$ be a smooth fibred manifold. We call coordinates $\mathbf{x} = (x^i : 1 \le i \le n)$ of $\mathcal{X}$ *independent variables* and fibre coordinates $\mathbf{u} = (u^\alpha : 1 \le \alpha \le m)$ in $\mathcal{E}$ *dependent variables.* Sections $\sigma : \mathcal{X} \to \mathcal{E}$ correspond locally to functions $\mathbf{u} = \mathbf{s}(\mathbf{x})$. We will use throughout a "global" notation in order to avoid the introduction of many local neighbourhoods even though we mostly consider local sections.

Derivatives are written in the form $u^\alpha_\mu = \partial^{|\mu|}u^\alpha/\partial x_1^{\mu_1}\cdots\partial x_n^{\mu_n}$ where $\mu = (\mu_1, \dots, \mu_n)$ is a multi-index. The set of derivatives u^α_μ up to order q is denoted by $\mathbf{u}^{(q)}$; it defines a local coordinate system for the q-th order jet bundle $J_q\pi$, which may be regarded as the space of truncated Taylor expansions of functions $\mathbf{s}$.

The hierarchy of jet bundles $J_q\pi$ with $q = 0, 1, 2, \dots$ possesses many natural fibrations which correspond to "forgetting" higher-order derivatives. For us particularly important are $\pi^q_{q-1} : J_q\pi \to J_{q-1}\pi$ and $\pi^q : J_q\pi \to \mathcal{X}$. To each section $\sigma : \mathcal{X} \to \mathcal{E}$, locally defined by $\sigma(\mathbf{x}) = \bigl(\mathbf{x}, \mathbf{s}(\mathbf{x})\bigr)$, we may associate its *prolongation* $j_q\sigma : \mathcal{X} \to J_q\pi$, a section of the fibration π^q locally given by $j_q\sigma(\mathbf{x}) = \bigl(\mathbf{x}, \mathbf{s}(\mathbf{x}), \partial_{\mathbf{x}}\mathbf{s}(\mathbf{x}), \partial_{\mathbf{xx}}\mathbf{s}(\mathbf{x}), \dots\bigr)$.

The geometry of the jet bundle $J_q\pi$ is to a large extent determined by its *contact structure.* It can be introduced in various ways. For our purposes, three different approaches are convenient. First, we adopt the *contact codistribution* $\mathcal{C}^0_q \subseteq T^*(J_q\pi)$; it consists of all one-forms such that their pull-back by a prolonged section vanishes. Locally, it is spanned by the *contact forms*

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

Dually, we may consider the *contact distribution* $\mathcal{C}_q \subseteq T(J_q\pi)$ consisting of all vector fields annihilated by $\mathcal{C}^0_q$. A straightforward calculation shows that it is generated by the *contact fields*

$$\begin{aligned} C^{(q)}_i &= \partial_i + \sum_{\alpha=1}^m \sum_{0\le|\mu|<q} u^\alpha_{\mu+1_i}\partial_{u^\alpha_\mu}, \qquad 1 \le i \le n, \\ C^\mu_\alpha &= \partial_{u^\alpha_\mu}, \qquad |\mu| = q, \quad 1 \le \alpha \le m. \end{aligned} \tag{1}$$

Note that the latter fields, C^μ_α, span the vertical bundle $V\pi^q_{q-1}$ of the fibration π^q_{q-1}. Thus the contact distribution can be split into $\mathcal{C}_q = V\pi^q_{q-1} \oplus \mathcal{H}$. Here the complement $\mathcal{H}$ is an n-dimensional transversal subbundle of $T(J_q\pi)$ and obviously not uniquely determined (though any local coordinate chart induces via the span of the vectors $C^{(q)}_i$ one possible choice). Any such complement $\mathcal{H}$ may be considered the horizontal bundle of a connection on the fibred manifold $\pi^q : J_q\pi \to \mathcal{X}$ (but not for the fibration π^q_{q-1}). Following Fackerell [5], we call any connection on π^q the horizontal bundle of which consists of contact fields a *Vessiot connection* (in the literature the terminology *Cartan connection* is also common, see for example [15]).

For later use, we note the structure equations of the contact distribution. The only non-vanishing Lie brackets of the vector fields (1) are

$$\bigl[C^{\nu+1_i}_\alpha, C^{(q)}_i\bigr] = \partial_{u^\alpha_\nu}, \qquad |\nu| = q - 1. \tag{2}$$

Note that this observation implies that the vertical bundle $V\pi^q_{q-1}$ is involutive.

As a third approach to the contact structure we consider, following Modugno [17], the *contact map* (see also [6, 23]). It is the unique map $\Gamma_q : J_q\pi \underset{\mathcal{X}}{\times} T\mathcal{X} \to T(J_{q-1}\pi)$ such that the diagram

$$\begin{tikzcd}
J_q\pi \underset{\mathcal{X}}{\times} T\mathcal{X} \arrow[rr, "\Gamma_q"] & & T(J_{q-1}\pi) \\
& T\mathcal{X} \arrow[ul, "((j_q\sigma)\circ\tau_{\mathcal{X}})\times\mathrm{id}_{T\mathcal{X}}"] \arrow[ur, "T(j_{q-1}\sigma)"'] &
\end{tikzcd}$$

commutes for any section σ. Because of its linearity over π^q_{q-1}, we may also consider it a map $\Gamma_q : J_q\pi \to T^*\mathcal{X} \underset{J_{q-1}\pi}{\otimes} T(J_{q-1}\pi)$ with the local coordinate form

$$\Gamma_q :\ (\mathbf{x}, \mathbf{u}^{(q)}) \mapsto \big(\mathbf{x}, \mathbf{u}^{(q-1)}; dx^i \otimes (\partial_{x^i} + u^\alpha_{\mu+1_i}\partial_{u^\alpha_\mu})\big). \tag{3}$$

Obviously, $\Gamma_q(\rho, \partial_{x^i}) = T_\rho\pi^q_{q-1}(C^{(q)}_i)$ and hence $(\mathcal{C}_q)_\rho = (T_\rho\pi^q_{q-1})^{-1}\big(\operatorname{im}\Gamma_q(\rho)\big)$ for any point $\rho \in J_q\pi$. Note that all vectors in the image of $\Gamma_q(\rho)$ are transversal to the fibration π^{q-1}_{q-2}.

One of the main applications of the contact structure is given by the following proposition (for a proof, see [6, Proposition 2.1.6] or [23, Proposition 2.2.7]). It characterises those sections of the jet bundle $\pi^q\colon J_q\pi \to \mathcal{X}$ which are prolongations of sections of the underlying fibred manifold $\pi\colon \mathcal{E} \to \mathcal{X}$.

Proposition 1. *A section $\gamma : \mathcal{X} \to J_q\pi$ is of the form $\gamma = j_q\sigma$ for a section $\sigma : \mathcal{X} \to \mathcal{E}$ if, and only if, $\operatorname{im}\Gamma_q\big(\gamma(x)\big) = T_{\gamma(x)}\pi^q_{q-1}\big(T_{\gamma(x)}\operatorname{im}\gamma\big)$ for all points $x \in \mathcal{X}$ where γ is defined.*

Thus for any section $\sigma : \mathcal{X} \to \mathcal{E}$ the equality $\operatorname{im}\Gamma_{q+1}\big(j_{q+1}\sigma(x)\big) = \operatorname{im} T_x(j_q\sigma)$ holds and we may say that knowing the $(q+1)$-jet $j_{q+1}\sigma(x)$ of a section σ at some $x \in \mathcal{X}$ is equivalent to knowing its q-jet $\rho = j_q\sigma(x)$ at x as well as the tangent space $T_\rho(\operatorname{im} j_q\sigma)$ at this point. This observation will later be the key for the Vessiot theory.

3 The formal theory of differential equations

We are now going to outline the formal theory of partial differential equations to introduce the basic notation. Our presentation follows [23]; other general references are [13, 14, 19].

Definition 1. A *differential equation* of order q is a fibred submanifold $\mathcal{R}_q \subseteq J_q\pi$ locally described as the zero set of some smooth functions on $J_q\pi$:

$$\mathcal{R}_q :\ \begin{cases} \Phi^\tau\big(\mathbf{x}, \mathbf{u}^{(q)}\big) = 0, \\ (\tau = 1, \dots, t). \end{cases} \tag{4}$$

Note that we do not distinguish between scalar equations and systems.

We denote by $\iota : \mathcal{R}_q \hookrightarrow J_q\pi$ the canonical inclusion map. Differentiating every equation in the local representation (4) leads to the *prolonged equation* $\mathcal{R}_{q+1} \subseteq J_{q+1}\pi$ defined by the equations $\Phi^\tau = 0$ and $D_i\Phi^\tau = 0$ where the formal derivative D_i is given by

$$D_i\Phi^\tau\big(\mathbf{x}, \mathbf{u}^{(q+1)}\big) = \frac{\partial\Phi^\tau}{\partial x^i}\big(\mathbf{x}, \mathbf{u}^{(q)}\big) + \sum_{0\le|\mu|\le q}\sum_{\alpha=1}^{m} \frac{\partial\Phi^\tau}{\partial u^\alpha_\mu}\big(\mathbf{x}, \mathbf{u}^{(q)}\big)u^\alpha_{\mu+1_i}. \tag{5}$$

Iteration of this process gives the higher prolongations $\mathcal{R}_{q+r} \subseteq J_{q+r}\pi$. A subsequent projection leads to $\mathcal{R}^{(1)}_q = \pi^{q+1}_q(\mathcal{R}_{q+1}) \subseteq \mathcal{R}_q$, which is a proper submanifold whenever integrability conditions appear.

Definition 2. A differential equation $\mathcal{R}_q$ is *formally integrable* if at any prolongation order $r > 0$ the equality $\mathcal{R}^{(1)}_{q+r} = \mathcal{R}_{q+r}$ holds.

In local coordinates, the following definition coincides with the usual notion of a solution.

Definition 3. 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 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 4. Let $\mathcal{R}_q \subseteq J_q\pi$ be a differential equation and $\iota : \mathcal{R}_q \to J_q\pi$ the canonical inclusion map. Let $\mathcal{I}[\mathcal{R}_q] = \langle \iota^*\mathcal{C}^0_q\rangle_{\text{diff}}$ be the differential ideal generated by the pull-back of the contact codistribution on $\mathcal{R}_q$ (algebraically, $\mathcal{I}[\mathcal{R}_q]$ is then spanned by a basis of $\iota^*\mathcal{C}^0_q$ and the exterior derivatives of the forms in this basis). A linear subspace $\mathcal{U}_\rho \subseteq T_\rho\mathcal{R}_q$ is an *integral element* at the point $\rho \in \mathcal{R}_q$, if all forms in $(\mathcal{I}[\mathcal{R}_q])_\rho$ vanish on it.

The following result provides an alternative characterisation of *transversal* integral elements via the contact map. It requires that the projection $\pi^{q+1}_q : \mathcal{R}_{q+1} \to \mathcal{R}_q$ is surjective.

Proposition 2. *Let $\mathcal{R}_q$ be a differential equation such that $\mathcal{R}^{(1)}_q = \mathcal{R}_q$. A linear subspace $\mathcal{U}_\rho \subseteq T_\rho\mathcal{R}_q$ such that $T_\rho\iota(\mathcal{U}_\rho)$ lies transversal to the fibration π^q_{q-1} is an integral element at the point $\rho \in \mathcal{R}_q$ if, and only if, a point $\hat\rho \in \mathcal{R}_{q+1}$ exists on the prolonged equation $\mathcal{R}_{q+1}$ such that $\pi^{q+1}_q(\hat\rho) = \rho$ and $T_\rho\iota(\mathcal{U}_\rho) \subseteq \operatorname{im}\Gamma_{q+1}(\hat\rho)$.*

Proof. Assume first that $\mathcal{U}_\rho$ satisfies the given conditions. It follows immediately from the coordinate form of the contact map that then firstly $T_\rho\iota(\mathcal{U}_\rho)$ is transversal to π^q_{q-1} and secondly that every one-form $\omega \in \iota^*\mathcal{C}^0_q$ vanishes on $\mathcal{U}_\rho$, as $\operatorname{im}\Gamma_{q+1}(\hat\rho) \subset (\mathcal{C}_q)_\rho$. Thus there only remains to show that the same is true for the two-forms $d\omega \in \iota^*(d\mathcal{C}^0_q)$.

Choose a section $\gamma : \mathcal{R}_q \to \mathcal{R}_{q+1}$ such that $\gamma(\rho) = \hat\rho$ and define a distribution $\mathcal{D}$ of rank n on $\mathcal{R}_q$ by setting $T\iota(\mathcal{D}_{\tilde\rho}) = \operatorname{im}\Gamma_{q+1}\bigl(\gamma(\tilde\rho)\bigr)$ for any point $\tilde\rho \in \mathcal{R}_q$. Obviously, by construction $\mathcal{U}_\rho \subseteq \mathcal{D}_\rho$. It follows now from the coordinate form (3) of the contact map that locally the distribution $\mathcal{D}$ is spanned by n vector fields X_i such that $\iota_*X_i = C^{(q)}_i + \gamma^\alpha_{\mu+1_i}C^\mu_\alpha$ where the coefficients γ^α_ν are the highest-order components of the section γ. Thus the commutator of two such vector fields satisfies

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

The commutators on the right hand side vanish whenever $\mu_i = 0$ or $\mu_j = 0$, respectively. Otherwise we obtain $-\partial_{u^\alpha_{\mu-1_i}}$ and $-\partial_{u^\alpha_{\mu-1_j}}$, respectively. But this fact implies that the two sums on the right hand side cancel each other and we find that $\iota_*\bigl([X_i, X_j]\bigr) \in \mathcal{C}_q$. Thus we find for any contact form $\omega \in \mathcal{C}^0_q$ that

$$\iota^*(d\omega)(X_i, X_j) = d\omega(\iota_*X_i, \iota_*X_j) = \iota_*X_i\bigl(\omega(\iota_*X_j)\bigr) - \iota_*X_j\bigl(\omega(\iota_*X_i)\bigr) + \omega\bigl(\iota_*([X_i, X_j])\bigr).$$

Each summand in the last expression vanishes, as all appearing fields are contact fields. Hence any form $\omega \in \iota^*(d\mathcal{C}^0_q)$ vanishes on $\mathcal{D}$ and in particular on $\mathcal{U}_\rho \subseteq \mathcal{D}_\rho$.

For the converse, note that any transversal integral element $\mathcal{U}_\rho \subseteq T_\rho\mathcal{R}_q$ is spanned by linear combinations of vectors v_i such that $T_\rho\iota(v_i) = C^{(q)}_i|_\rho + \gamma^\alpha_{\mu,i}C^\mu_\alpha|_\rho$ where $\gamma^\alpha_{\mu,i}$ are real coefficients. Now consider a contact form ω^α_ν with $|\nu| = q - 1$. Then $d\omega^\alpha_\nu = dx^i \wedge du^\alpha_{\nu+1_i}$. Evaluating the condition $\iota^*(d\omega^\alpha_\nu)|_\rho(v_i, v_j) = d\omega\bigl(T_\rho\iota(v_i), T_\rho\iota(v_j)\bigr) = 0$ yields the equation $\gamma^\alpha_{\nu+1_i,j} = \gamma^\alpha_{\nu+1_j,i}$. Hence the coefficients are of the form $\gamma^\alpha_{\mu,i} = \gamma^\alpha_{\mu+1_i}$ and a section σ exists such that $\rho = j_q\sigma(x)$ and $T_\rho(\operatorname{im} j_q\sigma)$ is spanned by the vectors $T_\rho\iota(v_1), \dots, T_\rho\iota(v_n)$. This observation implies that $\mathcal{U}_\rho$ satisfies the given conditions. ∎

For many purposes the purely geometric notion of formal integrability is not sufficient, and one needs the stronger algebraic concept of involution. This concerns in particular the derivation of uniqueness results but also the numerical integration of overdetermined systems [21]. An intrinsic definition of involution is possible using the Spencer cohomology (see for example [22] and references therein for a discussion). We apply here a simpler approach requiring that one works in "good", more precisely: δ-regular, coordinates $\mathbf{x}$. This assumption represents a mild restriction, as generic coordinates are δ-regular and it is possible to construct systematically "good" coordinates – see [11]. Furthermore, it will turn out that the use of δ-regular coordinates is essential for Vessiot's approach.

Definition 5. The (*geometric*) *symbol* of a differential equation $\mathcal{R}_q$ is $\mathcal{N}_q = V\pi^q_{q-1}|_{\mathcal{R}_q} \cap T\mathcal{R}_q$.

Thus, the symbol is the vertical part of the tangent space to $\mathcal{R}_q$. Locally, $\mathcal{N}_q$ consists of all vertical vector fields $\sum_{\alpha=1}^m \sum_{|\mu|=q} v^\alpha_\mu \partial_{u^\alpha_\mu}$ where the coefficients v^α_μ satisfy the following linear system of algebraic equations:

$$\mathcal{N}_q\colon \left\{ \sum_{\alpha=1}^{m} \sum_{|\mu|=q} \left(\frac{\partial \Phi^\tau}{\partial u^\alpha_\mu} \right) v^\alpha_\mu = 0 . \right. \tag{6}$$

The matrix of this system is called the *symbol matrix* M_q. The *prolonged symbols* $\mathcal{N}_{q+r}$ are the symbols of the prolonged equations $\mathcal{R}_{q+r}$ with corresponding symbol matrices M_{q+r}.

The *class* of a multi-index $\mu = (\mu_1, \dots, \mu_n)$, denoted $\operatorname{cls}\mu$, is the smallest k for which μ_k is different from zero. The columns of the symbol matrix (6) are labelled by the v^α_μ. We order them as follows. Let α and β denote indices for the dependent 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 distinguish 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 $\operatorname{cls}\mu < \operatorname{cls}\nu$, 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 generalise 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. This is called the *term-over-position lift of the degree reverse lexicographic ranking.*

Now the columns within the symbol matrix are ordered descendingly according to the degree reverse lexicographic ranking for the multi-indices μ of the variables v^α_μ in equation (6) and labelled by the pairs (α, μ). (It follows that, if v^α_μ and v^β_ν are such that $\operatorname{cls}\mu > \operatorname{cls}\nu$, then the column corresponding to v^α_μ is left of the column corresponding to v^β_ν.) 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.

We compute now a row echelon form of the symbol matrix. We denote the number of rows where the pivot is of class k by $\beta^{(k)}_q$, the *indices* of the symbol $\mathcal{N}_q$, and associate with each such row its *multiplicative variables* $x^1, \dots, x^k$. Prolonging each equation only with respect to its multiplicative variables yields independent equations of order $q+1$, as each has a different leading term.

Definition 6. If prolongation with respect to the non-multiplicative variables does not lead to additional independent equations of order $q+1$, in other words if

$$\operatorname{rank} M_{q+1} = \sum_{k=1}^{n} k\beta^{(k)}_q , \tag{7}$$

then the symbol $\mathcal{N}_q$ is *involutive*. The differential equation $\mathcal{R}_q$ is called *involutive*, if it is formally integrable and its symbol is involutive.

The criterion (7) is also known as *Cartan's test*, as it is analogous to a similar test in the Cartan–Kähler theory of exterior differential systems. We stress again that it is valid only in δ-regular coordinates (in fact, in other coordinate systems it will always fail).

4 The Cartan normal form

For notational simplicity, we will consider in our subsequent analysis almost exclusively first-order equations $\mathcal{R}_1 \subseteq J_1\pi$. At least from a theoretical standpoint, this is not a restriction, as any higher-order differential equation $\mathcal{R}_q$ can be transformed into an equivalent first-order one (see for example [23, Appendix A.3]). For these we now introduce a convenient local representation.

Definition 7. For a first-order differential equation $\mathcal{R}_1$ the following local representation, a special kind of solved form,

$$u_n^\alpha = \phi_n^\alpha\bigl(\mathbf{x}, u^\beta, u_j^\gamma, u_n^\delta\bigr) \qquad \begin{cases} 1 \le \alpha \le \beta_1^{(n)}, \\ 1 \le j < n, \\ \beta_1^{(n)} < \delta \le m, \end{cases} \tag{8a}$$

$$u_{n-1}^\alpha = \phi_{n-1}^\alpha\bigl(\mathbf{x}, u^\beta, u_j^\gamma, u_{n-1}^\delta\bigr) \qquad \begin{cases} 1 \le \alpha \le \beta_1^{(n-1)}, \\ 1 \le j < n-1, \\ \beta_1^{(n-1)} < \delta \le m, \end{cases} \tag{8b}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$u_1^\alpha = \phi_1^\alpha\bigl(\mathbf{x}, u^\beta, u_1^\delta\bigr) \qquad \begin{cases} 1 \le \alpha \le \beta_1^{(1)}, \\ \beta_1^{(1)} < \delta \le m, \end{cases} \tag{8c}$$

$$u^\alpha = \phi^\alpha\bigl(\mathbf{x}, u^\beta\bigr) \qquad \begin{cases} 1 \le \alpha \le \beta_0, \\ \beta_0 < \beta \le m, \end{cases} \tag{8d}$$

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 empty and no terms u_k^δ appear 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 depend on another principle derivative u_l^γ as long as $l < k$. The equations are grouped according to their class in descending order.

Theorem 1 (Cartan–Kähler). *Let the involutive differential equation $\mathcal{R}_1$ be locally represented in δ-regular coordinates by the system* (8a), (8b), (8c)*. Assume that the following initial conditions are given:*

$$u^\alpha(x^1, \ldots, x^n) = f^\alpha(x^1, \ldots, x^n), \qquad \beta_1^{(n)} < \alpha \le m; \tag{9a}$$

$$u^\alpha(x^1, \ldots, x^{n-1}, 0) = f^\alpha(x^1, \ldots, x^{n-1}), \qquad \beta_1^{(n-1)} < \alpha \le \beta_1^{(n)}; \tag{9b}$$

$$\cdots\cdots\cdots \qquad \cdots\cdots\cdots$$

$$u^\alpha(x^1, 0, \dots, 0) = f^\alpha(x^1), \qquad \beta_1^{(1)} < \alpha \leq \beta_1^{(2)}; \tag{9c}$$

$$u^\alpha(0, \dots, 0) = f^\alpha, \qquad 1 \leq \alpha \leq \beta_1^{(1)}. \tag{9d}$$

If the functions ϕ_k^α *and* f^α *are* (*real-*)*analytic at the origin, then this system has one and only one solution that is analytic at the origin and satisfies the initial conditions* (9)*.*

Proof. For the proof, see [19] or [23, Section 9.4] and references therein. The strategy is to split the system into subsystems according to the classes of the equations in it (see below). The solution is constructed step by step; each step renders a normal system in fewer independent variables to which the Cauchy–Kovalevskaya theorem is applied. Finally, the condition that $\mathcal{R}_1$ is involutive leads to further normal systems ensuring that the constructed functions are indeed solutions of the full system with respect to all independent variables. ■

Under some mild regularity assumptions the algebraic equations can always be solved locally. From now on, we will assume that any present algebraic equation has been explicitly solved, reducing thus the number of dependent variables. We simplify the Cartan normal form of a differential equation as given in Definition 7 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 no principal derivative appears on a right side of an equation (whereas this was possible with the non-reduced Cartan normal form of Definition 7). All the remaining, non-principal, derivatives are called *parametric*. Ordering the obtained equations by their class, we again can decompose them into subsystems:

$$u_k^\alpha = \phi_k^\alpha\big(\mathbf{x}, \mathbf{u}, u_j^\gamma\big) \qquad \begin{cases} 1 \leq j \leq k \leq n, \\ 1 \leq \alpha \leq \beta_1^{(k)}, \\ \beta_1^{(j)} < \gamma \leq m. \end{cases} \tag{10}$$

Note that the values $\beta_1^{(k)}$ are exactly those appearing in the Cartan test (7), as the symbol matrix of a differential equation in Cartan normal form is automatically triangular with the principal derivatives as pivots.

Definition 8. The *Cartan characters* of $\mathcal{R}_1$ are defined as $\alpha_1^{(k)} = m - \beta_1^{(k)}$ and thus equal the number of parametric derivatives of class k and order 1.

Provided that δ-regular coordinates are chosen, it is possible to perform a closed form involution analysis for a differential equation $\mathcal{R}_1$ in reduced Cartan normal form. We remark that an effective test of involution proceeds as follows (see for example [23, Remark 7.2.10]). Each equation in (10) 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 prolonged equations is of second-order any more. The equation $\mathcal{R}_1$ is involutive if, and only if, all new equations even simplify to zero, as any remaining first-order equation would be an integrability condition.

In order to apply this test, we now prove two helpful lemmata. We introduce the set $\mathcal{B} := \big\{(\alpha, i) \in \mathbb{N} \times \mathbb{N} \colon u_i^\alpha \text{ is a principal derivative}\big\}$, and for each $(\alpha, i) \in \mathcal{B}$ we define $\Phi_i^\alpha := u_i^\alpha - \phi_i^\alpha$. Using the contact fields (1), any prolongation of some Φ_i^α can be expressed in the following form.

Lemma 1. *Let the differential equation* $\mathcal{R}_1$ *be represented in the reduced Cartan normal form given by equation* (10)*. Then for any* $(\alpha, i) \in \mathcal{B}$ *and* $1 \leq j \leq n$*, we have*

$$D_j \Phi_i^\alpha = u_{ij}^\alpha - C_j^{(1)}(\phi_i^\alpha) - \sum_{h=1}^{i} \sum_{\gamma=\beta_1^{(h)}+1}^{m} u_{hj}^\gamma C_\gamma^h(\phi_i^\alpha). \tag{11}$$

Proof. By straightforward calculation; see [6, Lemma 2.5.5]. ■

For $j > i$, the prolongation $D_j\Phi_i^\alpha$ is non-multiplicative, otherwise it is multiplicative. Now let $j > i$, so that equation (11) shows a non-multiplicative prolongation, and assume that we are using δ-regular coordinates. According to our test, the symbol $\mathcal{N}_1$ is involutive if, and only if, it is possible to eliminate on the right hand side of (11) all second-order derivatives by adding multiplicative prolongations.

If the differential equation is not involutive, then the difference

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

does not necessarily vanish but may yield an obstruction to involution for any $(\alpha, i) \in \mathcal{B}$ and any $i < j \le n$. The next lemma gives all these obstructions to involution for a first-order system in reduced Cartan normal form.

Lemma 2. *Assume that δ-regular coordinates are used. For an equation in Cartan normal form and indices $i < j$ and α such that $(\alpha, i) \in \mathcal{B}$, we have the equality*

$$\begin{aligned}
&D_j\Phi_i^\alpha - D_i\Phi_j^\alpha + \sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)D_h\Phi_j^\gamma \\
&\quad = C_i^{(1)}(\phi_j^\alpha) - C_j^{(1)}(\phi_i^\alpha) - \sum_{h=1}^{i}\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_h^{(1)}(\phi_j^\gamma) && (12)\\
&\qquad - \sum_{h=1}^{i-1}\sum_{\delta=\beta_1^{(h)}+1}^{m} u_{hh}^\delta\left[\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma)\right] && (13)\\
&\qquad - \sum_{1\le h<k<i}\left\{\sum_{\delta=\beta_1^{(h)}+1}^{\beta_1^{(k)}} u_{hk}^\delta\left[\sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma)\right]\right. && (14a)\\
&\qquad + \left.\sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^\delta\left[\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_\delta^k(\phi_j^\gamma) + \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma)\right]\right\} && (14b)\\
&\qquad - \sum_{h=1}^{i-1}\left\{\sum_{\delta=\beta_1^{(h)}+1}^{\beta_1^{(i)}} u_{hi}^\delta\left[-C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma)\right]\right. && (15a)\\
&\qquad + \left.\sum_{\delta=\beta_1^{(i)}+1}^{m} u_{hi}^\delta\left[-C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma) + \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_\delta^i(\phi_j^\gamma)\right]\right\} && (15b)\\
&\qquad - \sum_{\substack{h=1\\ i+1\le k<j}}^{i-1}\sum_{\delta=\beta_1^{(k)}+1}^{m} u_{hk}^\delta\left[\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_\delta^k(\phi_j^\gamma)\right] && (16)
\end{aligned}$$

$$-\sum_{k=i}^{j-1}\sum_{\delta=\beta_1^{(k)}+1}^{m} u^\delta_{ik}\left[-C^k_\delta(\phi^\alpha_j)+\sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C^i_\gamma(\phi^\alpha_i)C^k_\delta(\phi^\gamma_j)\right] \tag{17}$$

$$-\sum_{h=1}^{i-1}\sum_{\delta=\beta_1^{(j)}+1}^{m} u^\delta_{hj}\left[C^h_\delta(\phi^\alpha_i)+\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i)C^j_\delta(\phi^\gamma_j)\right] \tag{18}$$

$$-\sum_{\delta=\beta_1^{(j)}+1}^{m} u^\delta_{ij}\left[C^i_\delta(\phi^\alpha_i)-C^j_\delta(\phi^\alpha_j)+\sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C^i_\gamma(\phi^\alpha_i)C^j_\delta(\phi^\gamma_j)\right]. \tag{19}$$

Proof. By a tedious but fairly straightforward computation, see [6, pages 48–55]. ■

In line (12) we have collected all terms which are of lower than second order. Furthermore, none of the appearing second-order derivatives is of a form that it could be eliminated by adding some multiplicative prolongation. Hence, under our assumption of δ-regular coordinates, the symbol $\mathcal{N}_1$ is involutive if, and only if, all the expressions in square brackets vanish. The differential equation $\mathcal{R}_1$ is involutive if, and only if, in addition line (12) vanishes, as it represents an integrability condition. Thus Lemma 2 gives us all obstructions to involution for $\mathcal{R}_1$ in explicit form. They will reappear in the proof of the existence theorem for integral distributions in Section 7.

5 The Vessiot distribution

By Proposition 1, the tangent spaces $T_\rho(\operatorname{im} j_q\sigma)$ of prolonged sections at points $\rho \in J_q\pi$ are always subspaces of the contact distribution $\mathcal{C}_q|_\rho$. If the section σ is a solution of the differential equation $\mathcal{R}_q$, then by definition it furthermore satisfies $\operatorname{im} j_q\sigma \subseteq \mathcal{R}_q$, and therefore $T(\operatorname{im} j_q\sigma) \subseteq T\mathcal{R}_q$. Hence, the following construction suggests itself.

Definition 9. 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

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

Note that the Vessiot distribution is not necessarily of constant rank along $\mathcal{R}_q$ (just like the symbol $\mathcal{N}_q$); for simplicity, we will almost always assume here that this is the case. This definition of the Vessiot distribution is not the one usually found in the literature. But the equivalence to the standard approach is an elementary exercise in computing with pull-backs.

Proposition 3. *The Vessiot distribution satisfies* $\mathcal{V}[\mathcal{R}_q] = (\iota^*\mathcal{C}^0_q)^0$.

For a differential equation 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. This has the advantage of keeping the calculations within a space of smaller dimension, namely the submanifold $\mathcal{R}_q$. Thereby regarding the differential equation as a manifold in its own right, we bar its coordinates to distinguish them from those of the jet bundle.

Example 1. Consider the first-order system given by the representation $\mathcal{R}_1\colon u_t - v = v_t - w_x = u_x - w = 0$. Then from the prolongations $u_{xt} - v_x = 0$ and $u_{xt} - w_t = 0$ follows the integrability condition $w_t = v_x$ by elimination of the second-order derivative u_{xt}. Thus, the first projection of $\mathcal{R}_1$ is (in Cartan normal form) represented by

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

It is not difficult to verify that the projected equation $\mathcal{R}_1^{(1)}$ is involutive. 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 distinguish 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

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

The tangent space $T\mathcal{R}_1^{(1)}$ 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 $T\iota(T\mathcal{R}_1^{(1)})$ therefore by the fields ∂_x, ∂_t, ∂_u, $\partial_v + \partial_{u_t}$, $\partial_w + \partial_{u_x}$, $\partial_{v_x} + \partial_{w_t}$ and $\partial_{w_x} + \partial_{v_t}$. This space is annihilated by

$$\omega^4 = du_x - dw, \qquad \omega^5 = dv_x - dw_t, \qquad \omega^6 = dw_x - dv_t, \qquad \omega^7 = du_t - dv.$$

These seven one-forms $\omega^1, \dots, \omega^7$ annihilate the Vessiot distribution $\mathcal{V}[\mathcal{R}_1^{(1)}]$, which is spanned by the four vector fields

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

In local coordinates on $\mathcal{R}_1^{(1)}$, these four vector fields become

$$\begin{aligned} \bar{X}_1 &= \partial_{\overline{x}} + \overline{w}\partial_{\overline{u}} + \overline{v_x}\partial_{\overline{v}} + \overline{w_x}\partial_{\overline{w}}, & \bar{X}_3 &= \partial_{\overline{v_x}},\\ \bar{X}_2 &= \partial_{\overline{t}} + \overline{v}\partial_{\overline{u}} + \overline{w_x}\partial_{\overline{v}} + \overline{v_x}\partial_{\overline{w}}, & \bar{X}_4 &= \partial_{\overline{w_x}}. \end{aligned}$$

They 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 remaining four one-forms $\omega^4, \dots, \omega^7$ trivially vanish on $\mathcal{R}_1^{(1)}$).

For a fully nonlinear differential equation $\mathcal{R}_q$, in particular an implicit one, this approach to compute its Vessiot distribution $\mathcal{V}[\mathcal{R}_q]$ via a pull-back is in general not effectively feasible. However, applying directly our definition of $\mathcal{V}[\mathcal{R}_q]$, it is easily possible even for such equations to determine effectively $T\iota(\mathcal{V}[\mathcal{R}_q])$, in other words: to realize it as a subbundle of $T(J_q\pi)|_{\mathcal{R}_q}$. The contact fields (1) form a basis for $\mathcal{C}_q$. It follows that for any vector field $\bar{X} \in \mathcal{V}[\mathcal{R}_q]$, coefficients $a^i, b^\alpha_\mu \in \mathcal{F}(\mathcal{R}_q)$, where $1 \le i \le n$, $1 \le \alpha \le m$ and $|\mu| = q$, exist such that

$$\iota_*\bar{X} = a^i C_i^{(q)} + b^\alpha_\mu C^\mu_\alpha. \tag{20}$$

If the differential equation $\mathcal{R}_q$ is locally represented by $\Phi^\tau = 0$, where $1 \le \tau \le t$, it follows from the tangency of the vector fields in $\mathcal{V}[\mathcal{R}_q]$ that $d\Phi^\tau(\iota_*\bar{X}) = \iota_*\bar{X}(\Phi^\tau) = 0$ and thus the coefficient functions must satisfy the following system of linear equations:

$$C_i^{(q)}(\Phi^\tau)a^i + C^\mu_\alpha(\Phi^\tau)b^\alpha_\mu = 0, \tag{21}$$

where $1 \le \tau \le t$. Note that this approach to determine the Vessiot distribution is closely related to prolonging the differential equation $\mathcal{R}_q$ and requires essentially the same computations. Indeed, the formal derivative (5) can be written in the form

$$D_i\Phi^\tau = C_i^{(q)}(\Phi^\tau) + C^\mu_\alpha(\Phi^\tau)u^\alpha_{\mu+1_i} = 0 \tag{22}$$

and in the context of the order-by-order determination of formal power series solutions (see for example [23, Section 2.3]) these equations are considered an inhomogeneous system for the Taylor coefficients of order $q+1$ depending on the lower order coefficients. Taking this point of view, we may call (21) the “projective” version of (22). In fact for $n = 1$, that is, for ordinary differential equations, this is even true in a rigorous sense.

Example 2. We consider the fully nonlinear first-order ordinary differential equation $\mathcal{R}_1$ locally defined by $(u')^2 + u^2 + x^2 = 1$. The contact distribution $\mathcal{C}_1$ is spanned by the two vector fields $X_1 = \partial_x + u'\partial_u$ and $X_2 = \partial_{u'}$ and the Vessiot distribution $T\iota(\mathcal{V}[\mathcal{R}_q])$ consists of all linear combinations of these two fields which are tangent to $\mathcal{R}_1$. Setting $\omega = xdx + udu + u'du'$, we thus have to solve the linear equation $\omega(aX_1 + bX_2) = 0$ in order to determine $T\iota(\mathcal{V}[\mathcal{R}_q])$. Its solution requires a case distinction (which is typical for fully nonlinear differential equations). If $u' \neq 0$, then we find

$$T\iota(\mathcal{V}[\mathcal{R}_q]) = \langle u'\partial_x + (u')^2\partial_u - (x + u'u)\partial_{u'}\rangle.$$

For $u' = 0$ and $x \neq 0$, the Vessiot distribution is spanned by the vertical contact field X_2. Finally, for $x = u' = 0$ the rank of the Vessiot distribution jumps, as at these points the whole contact plane is contained in it.

The definition of the symbol $\mathcal{N}_q$ and of the Vessiot distribution $\mathcal{V}[\mathcal{R}_q]$, respectively, of a differential equation $\mathcal{R}_q \subseteq J_q\pi$ immediately imply the following generalisation of the above discussed splitting of the contact distribution $\mathcal{C}_q = V\pi^q_{q-1} \oplus \mathcal{H}$ (such a splitting of the Vessiot distribution is also discussed by Lychagin and Kruglikov [14, 15] where the Vessiot distribution is called "Cartan distribution").

Proposition 4. *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*

$$\mathcal{V}[\mathcal{R}_q] = \mathcal{N}_q \oplus \mathcal{H} \tag{23}$$

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

Such a complement $\mathcal{H}$ is necessarily transversal to the fibration $\mathcal{R}_q \to \mathcal{X}$ and thus leads naturally to connections: provided $\dim \mathcal{H} = n$, it may be considered the horizontal bundle of a connection of the fibred manifold $\mathcal{R}_q \to \mathcal{X}$.

Definition 10. Any such connection is called a *Vessiot connection* for $\mathcal{R}_q$.

In general, the Vessiot distribution $\mathcal{V}[\mathcal{R}_q]$ is not involutive (that is, closed under the Lie bracket; an exception are formally integrable equations of finite type [23, Remark 9.5.8]), but it may contain involutive subdistributions. If these are furthermore transversal (to the fibration $\mathcal{R}_q \to \mathcal{X}$) and of dimension n, then they define a *flat* Vessiot connection.

Lemma 3. *If the section $\sigma : \mathcal{X} \to \mathcal{E}$ is a 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 $\operatorname{im} j_q\sigma$ for a solution σ of $\mathcal{R}_q$.*

Proof. Let σ be a local solution of the equation $\mathcal{R}_q$. Then it satisfies by Definition 3 $\operatorname{im} j_q\sigma \subseteq \mathcal{R}_q$ and thus $T(\operatorname{im} j_q\sigma) \subseteq T\mathcal{R}_q$. Besides, by the definition 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, $T\iota(\mathcal{V}[\mathcal{R}_q]) \subseteq \mathcal{C}_q|_{\mathcal{R}_q}$; this characterises prolonged sections. Hence, 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 $\mathcal{R}_q$. ■

This simple observation forms the basis of Vessiot's approach to the analysis of $\mathcal{R}_q$: he proposed to construct all flat Vessiot connections. Before we do this, we first show how integral elements appear in this program.

Proposition 5. *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. 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 hand side vanish trivially and the remaining equation implies our claim. ■

We call a subdistribution $\mathcal{U} \subseteq \mathcal{V}[\mathcal{R}_q]$ satisfying the conditions of Proposition 5 an *integral distribution* for the differential equation $\mathcal{R}_q$. In the literature [24], the terminology "involution" is common for such distributions which, however, is confusing. Note that generally an integral distribution is *not* integrable; the name only reflects the fact that it consists of integral elements.

A general differential equations $\mathcal{R}_q$ does not necessarily possess any Vessiot connection (not even a non-flat one). Their existence is linked to the absence of integrability conditions. More precisely, we obtain the following characterisation.

Proposition 6. *Let $\mathcal{R}_q$ be a differential equation. Then its Vessiot distribution possesses locally a direct 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 prolongation of equations of lower order in the system.*

Proof. Let $\mathcal{R}_q$ be locally represented by

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

such that the equations $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$ do not imply lower-order equations which are independent of the equations $\Psi^\sigma(\mathbf{x}, \mathbf{u}^{(q-1)}) = 0$. Let $\mathbf{u}_{(q)}$ denote the subset of all derivatives of order q only; then the Jacobi matrix $\big(\partial\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)})/\partial\mathbf{u}_{(q)}\big)$ has maximal rank. If we proceed as in the last proof, then the ansatz (20) for the determination of the Vessiot distribution yields for the above representation the linear system

$$\begin{aligned} &C_i^{(q)}(\Phi^\tau)a^i + C_\alpha^\mu(\Phi^\tau)b_\mu^\alpha = 0, \\ &C_i^{(q)}(\Psi^\sigma)a^i = 0. \end{aligned} \tag{24}$$

Here, the matrix $C_\alpha^\mu(\Phi^\tau)$ has maximal rank, too; thus the equations $C_i^{(q)}(\Phi^\tau)a^i + C_\alpha^\mu(\Phi^\tau)b_\mu^\alpha = 0$ can be solved for a subset of the unknowns b_μ^α. And since no terms of order q are present in $\Psi^\sigma(\mathbf{x}, \mathbf{u}^{(q-1)}) = 0$, we have $C_i^{(q)}(\Psi^\sigma) = D_i\Psi^\sigma$. Recall that we consider the Vessiot distribution, and thus the linear system (24), only on $\mathcal{R}_q$. It follows that the subsystem $C_i^{(q)}(\Psi^\sigma)a^i = 0$ becomes trivial 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 (24) has for each $1 \le j \le n$ 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 transversal complement $\mathcal{H}$. ■

From the proof of 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. It is the integrability conditions which arise as prolongations of lower order equations that hinder the construction of n-dimensional complements, while those which follow from the relations between cross derivatives do not influence this approach.

Remark 1. If one does not care about the distinction between different kinds of integrability conditions and simply requires that $\mathcal{R}_q = \mathcal{R}_q^{(1)}$ (meaning that no integrability conditions at all appear in the first prolongation of $\mathcal{R}_q$), then one can provide a more geometric proof for the existence of an n-dimensional complement (of course, in contrast to Proposition 6, the converse is not true then).

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 point 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 tangential to $\mathcal{R}_q$, as then we can define a complement by $T_\rho\iota(H_\rho) = \operatorname{im}\Gamma_{q+1}(\hat{\rho})$. But this tangency is a trivial consequence of $\hat{\rho} \in \mathcal{R}_{q+1}$; using for example the local coordinates expression (3) for the contact map and a local representation $\Phi^\tau = 0$ of $\mathcal{R}_q$, one immediately sees 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 it is possible to construct for each point $\rho \in \mathcal{R}_q$ a transversal complement $\mathcal{H}_\rho$ such that $\mathcal{V}_\rho[\mathcal{R}_q] = (\mathcal{N}_q)_\rho \oplus \mathcal{H}_\rho$. There remains to show that these complements can be chosen so that they form a smooth distribution. 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, that is, it defines a fibred manifold. Thus if we choose a 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.

Example 3. Consider again the differential equation $\mathcal{R}_1$ in Example 1. It is locally represented by the same equations as $\mathcal{R}_1^{(1)}$, 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 column more than the symbol 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^1, ω^2 and ω^3 (and their pull-backs $\iota^*\omega^1 = d\overline{u} - \overline{u_x}d\overline{x} - \overline{u_t}d\overline{t}$, $\iota^*\omega^2 = d\overline{v} - \overline{v_x}d\overline{x} - \overline{v_t}d\overline{t}$ and $\iota^*\omega^3 = d\overline{w} - \overline{w_x}d\overline{x} - \overline{w_t}d\overline{t}$) 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 $\bar{X}_2 = \partial_{\overline{t}} + \overline{u_t}\partial_{\overline{u}} + \overline{v_t}\partial_{\overline{v}} + \overline{w_t}\partial_{\overline{w}}$ in $T\mathcal{R}_1$ and $T\mathcal{R}_1^{(1)}$) look alike (see Example 1 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 $u_{xt} = w_t$, and not from the prolongation of a lower order equation.

Now consider for comparison the differential equation which is locally represented by

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

It arises from the system $\mathcal{R}_1^{(1)}$ in Example 1 by adding the algebraic equation $u = x$. Proceeding as in Example 1, we find that the Vessiot distribution $\mathcal{V}[\tilde{\mathcal{R}}_1]$ is spanned by the three vector fields

$$\partial_{\overline{v_x}}, \qquad \partial_{\overline{w_x}} \qquad \text{and} \qquad \overline{v}\partial_{\overline{x}} + (1-\overline{w})\partial_{\overline{t}} + \big(\overline{w_x}(1-\overline{w}) + \overline{v}\overline{v_x}\big)\partial_{\overline{v}} + \big(\overline{v_x}(1-\overline{w}) + \overline{v}\overline{w_x}\big)\partial_{\overline{w}}.$$

The vector fields $\partial_{\overline{v_x}}$ and $\partial_{\overline{w_x}}$ generate the symbol $\tilde{\mathcal{N}}_1$; any of its complements in $\mathcal{V}[\tilde{\mathcal{R}}_1]$ is one-dimensional and, as the dimension of the base space is two, none of them has the right dimension to be the horizontal space of a connection.

The reason for this is that $\tilde{\mathcal{R}}_1$ is not formally integrable, as the prolongation of the algebraic equation $u = x$ leads to the integrability conditions $u_x = 1$ and $u_t = 0$. Projecting the prolonged equation gives

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

Now the symbol vanishes, and so do the pull-backs of the contact forms: $\iota^*\omega^1 = d\overline{x} - d\overline{x} = 0$, $\iota^*\omega^2 = d\overline{v} = 0$, $\iota^*\omega^3 = d\overline{w} = 0$. Therefore we find $\mathcal{V}[\tilde{\mathcal{R}}_1^{(1)}] = \tilde{\mathcal{N}}_1^{(1)} \oplus \mathcal{H} = \{0\} \oplus \langle \partial_{\overline{x}}, \partial_{\overline{t}} \rangle$. As the Lie brackets of $\partial_{\overline{x}}$ and $\partial_{\overline{t}}$ trivially vanish, $T\tilde{\mathcal{R}}_1^{(1)} = \mathcal{V}[\tilde{\mathcal{R}}_1^{(1)}] = \mathcal{H} = \langle \partial_{\overline{x}}, \partial_{\overline{t}} \rangle$ is a two-dimensional involutive distribution.

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 as a complement. Choosing a "reference" complement $\mathcal{H}_0$ with a basis $(X_i \colon 1 \leq i \leq n)$, a basis for any other complement $\mathcal{H}$ arises by adding some vertical fields to the vectors X_i. We will follow this approach in the next section. For the remainder of this section we turn our attention to the choice of a convenient basis of $\mathcal{V}[\mathcal{R}_q]$ that will facilitate our computations.

Let $r := \dim \mathcal{N}_q$. As an intersection of two involutive distributions, the symbol $\mathcal{N}_q$ is an involutive distribution, too. Hence, there exists a basis $(Y_k \colon 1 \leq k \leq r)$ for it whose Lie brackets vanish: $[Y_k, Y_\ell] = 0$ for all $1 \leq k, \ell \leq r$. Since the vertical bundle $V\pi^q_{q-1}$ is also involutive, we can decompose it into

$$V\pi^q_{q-1} = \mathcal{N}_q \oplus \mathcal{W},$$

where $\mathcal{W}$ is again an involutive distribution. It can be spanned by vector fields $W_1, W_2, \dots, W_s$ (where $s = \sum_{k=1}^n \beta_q^{(k)}$ equals the number of principal derivatives) which are chosen such that we have $[W_a, W_b] = 0$ for all $1 \leq a, b \leq s$. In local coordinates, a particularly convenient choice for the fields Y_k and W_a exists. We first choose for any $1 \leq k \leq r$ a *parametric* derivative u^α_μ, that is $(\alpha, \mu) \notin \mathcal{B}$, and set $Y_k := Y^\alpha_\mu := \iota_*(\partial_{\overline{u^\alpha_\mu}})$; then we choose for any $1 \leq a \leq s$ a *principal* derivative u^α_μ, that is $(\alpha, \mu) \in \mathcal{B}$, and set $W_a := W^\alpha_\mu := \partial_{u^\alpha_\mu}$.

The reference complement $\mathcal{H}_0$ is chosen as follows. Any basis of it must consist of n transversal contact fields. Since the fields C^μ_α are vertical, we can always use a basis $(\tilde{X}_1 \colon 1 \leq i \leq n)$ of the form

$$\tilde{X}_i = C_i^{(q)} + \xi^\alpha_{i\mu} C^\mu_\alpha$$

with some coefficient functions $\xi^\alpha_{i\mu}$ chosen such that $\tilde{X}_i$ is tangential to $\mathcal{R}_q$. The fields C^μ_α also span the vertical bundle $V\pi^q_{q-1}$ and hence we may exploit the above decomposition for a further simplification of the basis. 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 $(X_i \colon 1 \leq i \leq n)$ where

$$X_i = C_i^{(q)} + \xi^a_i W_a. \tag{25}$$

As already mentioned above, the Vessiot distribution $\mathcal{V}[\mathcal{R}_q]$ is not necessarily involutive. So it is not surprising that its structure equations are going to be important later. We may extend the above chosen basis (X_i, Y_k) of $\mathcal{V}[\mathcal{R}_q]$ to a basis of the derived Vessiot distribution, $\mathcal{V}'[\mathcal{R}_q]$, by adding vector fields Z_c, $1 \leq c \leq C := \dim \mathcal{V}'[\mathcal{R}_q] - \dim \mathcal{V}[\mathcal{R}_q]$, where, using (2), for each c we have some coefficients $\kappa^\alpha_{c\nu}$ such that $Z_c = \kappa^\alpha_{c\nu} \partial_{u^\alpha_\nu}$ with $|\nu| = q - 1$. By construction, the non-vanishing structure equations of $\mathcal{V}[\mathcal{R}_q]$ take now the form

$$[X_i, X_j] = \Theta^c_{ij} Z_c \qquad \text{and} \qquad [X_i, Y_k] = \Xi^c_{ik} Z_c \tag{26}$$

for $1 \leq i, j \leq n$ and $1 \leq k \leq r$, with smooth functions Θ^c_{ij} and Ξ^c_{ik}. (For the complete set of structure equations, we have to add $[Y_k, Y_l] = 0$ for $1 \leq k, l \leq r$.)

Remark 2. Since the vector fields Z_c, which appear on the right sides of the structure equations (26), may, for $q = 1$, span a proper subspace in $\langle \partial_{u^\alpha} : 1 \leq \alpha \leq m \rangle$, about the exact size of which we know nothing, we write them as linear combinations $Z_c =: \kappa^\alpha_c \partial_{u^\alpha}$. The structure equations then become

$$[X_i, X_j] = \Theta^c_{ij} \kappa^\alpha_c \partial_{u^\alpha} =: \Theta^\alpha_{ij} \partial_{u^\alpha}, \qquad 1 \leq i, j \leq n, \tag{26a$'$}$$

$$[X_i, Y_k] = \Xi^c_{ik} \kappa^\alpha_c \partial_{u^\alpha} =: \Xi^\alpha_{ik} \partial_{u^\alpha}, \qquad 1 \leq i \leq n, \quad 1 \leq k \leq r. \tag{26b$'$}$$

Knowing the larger sets of coefficients Θ^α_{ij}, Ξ^α_{ik}, we can reconstruct the true structure coefficients Θ^c_{ij}, Ξ^c_{ik} by solving the overdetermined systems of linear equations

$$\Theta^c_{ij} \kappa^\alpha_c = \Theta^\alpha_{ij} \qquad \text{and} \qquad \Xi^c_{ik} \kappa^\alpha_c = \Xi^\alpha_{ik}.$$

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^c_α such that

$$\Theta^c_{ij} = \Theta^\alpha_{ij} \kappa^c_\alpha \qquad \text{and} \qquad \Xi^c_{ik} = \Xi^\alpha_{ik} \kappa^c_\alpha.$$

For our later proof of an existence theorem for integral distributions, we will have to analyse certain matrices with the coefficients Θ^c_{ij} and Ξ^c_{ik} as their entries. It turns out that this analysis becomes simpler, if we use the extended sets of coefficients Θ^α_{ij} and Ξ^α_{ik} instead.

For a first-order equation $\mathcal{R}_1$ with Cartan normal form (8) satisfying the assumptions of Proposition 6 it is possible to perform this process explicitly. We choose as a reference complement $\mathcal{H}_0$ the linear span of the vector fields

$$\iota_* \bar{X}_i = C^{(q)}_i + \sum_{(\alpha,\mu) \in \mathcal{B}} C^{(q)}_i (\phi^\alpha_\mu) C^\mu_\alpha \tag{27}$$

for $1 \leq i \leq n$. One verifies in a straightforward computation that (27) represents a valid choice (see [6, Proposition 3.1.19]). Using this reference complement, we can explicitly evaluate the Lie brackets (26) on $\mathcal{R}_1$. As we are not able to determine a simple expression for the derived Vessiot distribution, we follow the approach of Remark 2 and consider the equations (26$'$) instead.

Lemma 4. *Let $i < j$, without loss of generality. Then we obtain for the extended set of structure coefficients Θ^α_{ij} in local coordinates on $\mathcal{R}_1$ the following results:*

$$\Theta^\alpha_{ij} = \begin{cases} C^{(1)}_i(\phi^\alpha_j) - C^{(1)}_j(\phi^\alpha_i), & (\alpha, i) \in \mathcal{B} \text{ and } (\alpha, j) \in \mathcal{B}, \\ C^{(1)}_i(\phi^\alpha_j), & (\alpha, i) \notin \mathcal{B} \text{ and } (\alpha, j) \in \mathcal{B}, \\ 0, & (\alpha, i) \notin \mathcal{B} \text{ and } (\alpha, j) \notin \mathcal{B}. \end{cases} \tag{28}$$

Proof. By a rather straightforward computation; see [6, pages 75 and 76]. ∎

We collect these coefficients into vectors Θ_{ij} which have m rows each where the entries are ordered according to increasing α. It is useful to denote the symbol fields $Y_k = \iota_*(\partial_{u^\beta_h})$ by using double indices: $Y_k = Y^\beta_h$ for any $(\beta, h) \notin \mathcal{B}$.

Lemma 5. *Set $\bar{Y}_k =: \bar{Y}^\beta_h$, then the extended set of structure coefficients Ξ^α_{ik} in local coordinates on $\mathcal{R}_1$ are*

$$\Xi^\alpha_{ik} = \begin{cases} -C^h_\beta(\phi^\alpha_i), & (\alpha,i)\in\mathcal{B},\\ -1, & (\alpha,i)\notin\mathcal{B} \text{ and } (\alpha,i)=(\beta,h),\\ 0, & (\alpha,i)\notin\mathcal{B} \text{ and } (\alpha,i)\neq(\beta,h). \end{cases} \tag{29}$$

Proof. Again by a rather straightforward computation; see [6, page 76]. ■

Example 4. We calculate the structure equations for Example 1. Here, $\beta^{(1)}_1 = 1$ and $\beta^{(2)}_1 = 3$. We set $\bar{X}_3 =: \bar{Y}_1 = \bar{Y}^2_1$ and $\bar{X}_4 =: \bar{Y}_2 = \bar{Y}^3_1$. Then besides the trivial structure equations we get:

$$[\bar{X}_1,\bar{X}_2] = 0, \qquad [\bar{X}_1,\bar{Y}_1] = [\bar{X}_2,\bar{Y}_2] = -\partial_{\overline{v}}, \qquad [\bar{X}_2,\bar{Y}_1] = [\bar{X}_1,\bar{Y}_2] = -\partial_{\overline{w}}.$$

Noting that here the set $\mathcal{B}$ contains the pairs $(1,1)\equiv u_x$, $(1,2)\equiv u_t$, $(2,2)\equiv v_t$ and $(3,2)\equiv w_t$ while $(2,1)\equiv v_x$ and $(3,1)\equiv w_x$ are not in $\mathcal{B}$, one easily verifies that all coefficients are as given in equations (28) and (29).

We end this section with two technical remarks on how these coefficients are collected into matrices Ξ_i. The examination of the ranks of these matrices Ξ_i is basic for the proof of the existence Theorem 2 for integral distributions.

Remark 3. Some of the terms $-C^h_\beta(\phi^\alpha_i)$ where $(\alpha,i)\in\mathcal{B}$ vanish, too: all of the parametric derivatives on the right side of an equation Φ^α_i in the reduced Cartan normal form (10) 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 $-C^h_\beta(\phi^\alpha_i) = 0$ whenever $i = \operatorname{cls}(u^\alpha_i) < \operatorname{cls}(u^\beta_h) = h$.

We collect the coefficients Ξ^α_{ik} into matrices Ξ_i using i as the number of the matrix to which the entry Ξ^α_{ik} 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 they have $r = \dim\mathcal{N}_1$ columns each of which can be labelled by pairs $(\beta,h)\notin\mathcal{B}$ or the symbol fields $\bar{Y}_k = \bar{Y}^\beta_h$. More precisely, for $1\le h\le n$, we set

$$\begin{pmatrix} -C^h_{\beta^{(h)}_1+1}(\phi^1_i) & -C^h_{\beta^{(h)}_1+2}(\phi^1_i) & \cdots & -C^h_m(\phi^1_i)\\ -C^h_{\beta^{(h)}_1+1}(\phi^2_i) & -C^h_{\beta^{(h)}_1+2}(\phi^2_i) & \cdots & -C^h_m(\phi^2_i)\\ \vdots & \vdots & \ddots & \vdots\\ -C^h_{\beta^{(h)}_1+1}(\phi^{\beta^{(i)}_1}_i) & -C^h_{\beta^{(h)}_1+2}(\phi^{\beta^{(i)}_1}_i) & \cdots & -C^h_m(\phi^{\beta^{(i)}_1}_i) \end{pmatrix} =: [\Xi_i]^h. \tag{30}$$

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 $1\le i\le n$, we have $m-\beta^{(i)}_1 = \alpha^{(i)}_1$, so such a matrix has $\beta^{(i)}_1$ rows and $\alpha^{(i)}_1$ columns. Since, for any h with $i<h$ and for all $\beta^{(h)}_1+1\le\beta\le m$, we have $-C^h_\beta(\phi^\alpha_i)=0$, such matrices $[\Xi_i]^h$ where $i<h$ are zero. Furthermore, for $1\le h\le n$, we assemble the remaining terms Ξ^α_{ik} (which are those where $(\alpha,i)\notin\mathcal{B}$) in a matrix. As above, let $\bar{Y}_k = \bar{Y}^\beta_h$, and denote any entry Ξ^α_{ik} accordingly, for the sake of introducing the following matrix, by ${}^h_\beta\Xi^\alpha_i$. Now set

$$\begin{pmatrix} {}^{\;\;h}_{\beta^{(h)}_1+1}\Xi^{\beta^{(i)}_1+1}_i & {}^{\;\;h}_{\beta^{(h)}_1+2}\Xi^{\beta^{(i)}_1+1}_i & \cdots & {}^h_m\Xi^{\beta^{(i)}_1+1}_i\\ {}^{\;\;h}_{\beta^{(h)}_1+1}\Xi^{\beta^{(i)}_1+2}_i & {}^{\;\;h}_{\beta^{(h)}_1+2}\Xi^{\beta^{(i)}_1+2}_i & \cdots & {}^h_m\Xi^{\beta^{(i)}_1+2}_i\\ \vdots & \vdots & \ddots & \vdots\\ {}^{\;\;h}_{\beta^{(h)}_1+1}\Xi^m_i & {}^{\;\;h}_{\beta^{(h)}_1+2}\Xi^m_i & \cdots & {}^h_m\Xi^m_i \end{pmatrix} =: [\Xi_i]_h.$$

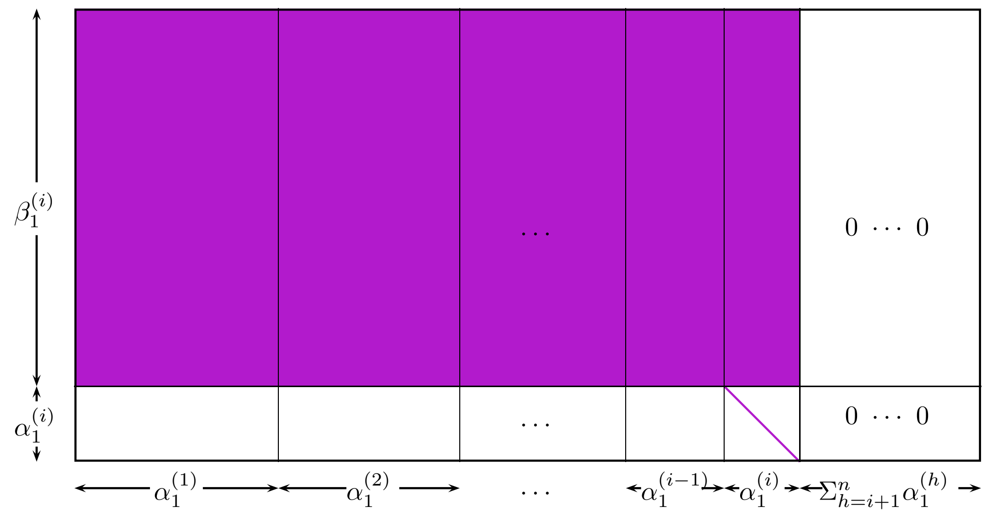

Figure 1. A sketch for the matrix Ξ_i in equation (31).

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 (29), 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

$$\Xi_i = \begin{pmatrix} [\Xi_i]^1 & \cdots & [\Xi_i]^{i-1} & [\Xi_i]^i & 0 \cdots 0 \\ 0 & \cdots & 0 & -\mathbb{1}_{\alpha_1^{(i)}} & 0 \cdots 0 \end{pmatrix}. \tag{31}$$

A sketch 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 Fig. 1.

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 4. 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 $-C_\beta^h(\phi_i^\alpha)$. The potentially non-trivial ones of them are marked as shaded areas in Fig. 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

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

Example 5. For Example 1, we have $n = 2$ and therefore two matrices Ξ_1 and Ξ_2. Their entries follow immediately from our results in Example 4. For $i = 1$, we have

$$\Xi_1 = \begin{pmatrix} 0 & 0 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The first line is $[\Xi_1]^1 = (0, 0)$. The unit block below it is $[\Xi_1]_1$. (We have $\beta_1^{(1)} = 1$ and therefore $\alpha_1^{(1)} = 2$. There is only $h = 1$ to consider, as both parametric derivatives v_x and w_x are with respect to x only.) Because of the very simple nature of our system, we find here by accident that $\Xi_1 = \Xi_2$. However, $[\Xi_2]^1 = \Xi_2$ is the whole matrix while there is no $[\Xi_2]_1$ because $\beta_1^{(2)} = m = 3$ and therefore $\alpha_1^{(2)} = 0$. Finally, we obtain $\Theta_{12} = (v_x - w_t, 0, 0)^t$. (The t top right marks the transpose.) Note that its first entry is in fact the integrability condition.

6 Flat Vessiot connections

In this section, we develop an approach for constructing flat Vessiot connections which improves recent approaches [5, 24, 27]: we exploit the splitting of $\mathcal{V}[\mathcal{R}_q]$ suggested by (23) to introduce convenient bases for integral distributions which yield structure equations that are particularly simple. Finally, we give necessary and sufficient conditions for Vessiot's approach to succeed.

Let $\mathcal{R}_q$ locally be represented by the system $\Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0$ where $1 \leq \tau \leq t$. Our goal is the construction of all n-dimensional transversal involutive subdistributions $\mathcal{U}$ within the Vessiot distribution $\mathcal{V}[\mathcal{R}_q]$. Taking for the Vessiot distribution the basis (X_i, Y_k) where the vector fields Y_k are the above mentioned basis of the symbol $T\iota(\mathcal{N}_q)$ with vanishing Lie brackets and the vector fields X_i are given in (25), we make for the basis $(U_i\colon 1 \leq i \leq n)$ of such a subdistribution $\mathcal{U}$ the ansatz

$$U_i = X_i + \zeta_i^k Y_k$$

with yet undetermined coefficient functions $\zeta_i^k \in \mathcal{F}(\mathcal{R}_q)$. This ansatz follows naturally from our considerations so far, as the fields X_i are transversal to the fibration over $\mathcal{X}$ and span a reference complement to the symbol and in $\mathcal{N}_q$ all fields Y_k are vertical.

Since the fields U_i are in triangular form, the distribution $\mathcal{U}$ is involutive if, and only if, their Lie brackets vanish, and using the structure equations (26) this means:

$$\begin{aligned}[U_i, U_j] &= [X_i, X_j] + \zeta_i^k [Y_k, X_j] + \zeta_j^k [X_i, Y_k] + \big(U_i(\zeta_j^k) - U_j(\zeta_i^k)\big) Y_k \\ &= \big(\Theta_{ij}^c - \Xi_{jk}^c \zeta_i^k + \Xi_{ik}^c \zeta_j^k\big) Z_c + \big(U_i(\zeta_j^k) - U_j(\zeta_i^k)\big) Y_k = 0. \end{aligned} \tag{32}$$

It follows from the definition of the fields Y_k and Z_c that they are linearly independent and so their coefficients must vanish for $\mathcal{U}$ to be involutive. Thus the Lie bracket (32) yields two sets of conditions for the coefficient functions ζ_i^k: a system of algebraic equations

$$G_{ij}^c := \Theta_{ij}^c - \Xi_{jk}^c \zeta_i^k + \Xi_{ik}^c \zeta_j^k = 0, \qquad \begin{cases} 1 \leq c \leq C, \\ 1 \leq i < j \leq n, \end{cases} \tag{33}$$

and a system of differential equations

$$H_{ij}^p := U_i(\zeta_j^p) - U_j(\zeta_i^p) = 0, \qquad \begin{cases} 1 \leq p \leq r, \\ 1 \leq i < j \leq n. \end{cases} \tag{34}$$

In the algebraic system (33) the true structure coefficients Θ_{ij}^c, Ξ_{jk}^c appear. For our subsequent analysis we follow Remark 2 and replace them by the extended set of coefficients Θ_{ij}^α, Ξ_{jk}^α. This corresponds to replacing (33) by an equivalent but larger linear system of equations which is, however, simpler to analyse.

Obviously, (33) is an inhomogeneous *linear* system to which any solution method for linear systems can be applied. We shall see in the next section that its structure (induced by the structure equations of the Vessiot distribution) allows us to decompose it into simpler subsystems which are considered step by step. Many papers on the Vessiot theory (see for example [5, 24]) study at this stage a *quadratic* system which *only* by such a step-by-step approach can be reduced to a series of linear problems. The linearity of (33) is a simple consequence of our choice of a basis for $\mathcal{V}[\mathcal{R}_q]$ which in turn exploits the natural splitting $\mathcal{V}[\mathcal{R}_q] = \mathcal{N}_q \oplus \mathcal{H}$.

Remark 5. The vector fields Y_k lie in the Vessiot distribution $\mathcal{V}[\mathcal{R}_1]$. Thus, according to Proposition 5, the subdistribution $\mathcal{U}$ is an integral distribution if, and only if, the coefficients ζ_k^i satisfy the algebraic conditions (33). This observation permits us immediately to reduce the number of unknowns in our ansatz. Assume that we have values $1 \leq i, j \leq n$ and $1 \leq \alpha \leq m$ such that both (α, i) and (α, j) are not contained in $\mathcal{B}$, that is, u_i^α and u_j^α are both parametric derivatives (and thus obviously the second-order derivative u_{ij}^α, too). Then there exist two symbol fields $Y_k = \iota_*(\partial_{u_i^\alpha})$ and $Y_l = \iota_*(\partial_{u_j^\alpha})$. Now it follows from the coordinate form (3) of the contact map that the subdistribution $\mathcal{U}$, spanned by the vector fields $U_h = X_h + \zeta_h^k Y_k$ for $1 \leq h \leq n$, can be an integral distribution if, and only if,

$$\zeta_j^k = \zeta_i^l \qquad \text{or, equivalently,} \qquad \zeta_j^{(\alpha,i)} = \zeta_i^{(\alpha,j)}. \tag{35}$$

for all $1 \leq i < j \leq n$ and $1 \leq k, l \leq r$.

As the unknowns ζ_j^k may be understood as labels for the columns of the matrices Ξ_h, this identification leads to a contraction of these matrices. The contracted matrices, denoted by $\hat{\Xi}_h$, arise as follows: whenever $\zeta_j^k = \zeta_i^l$ then the corresponding columns of Ξ_h are added; their sum replaces the first of these columns, while the second column is dropped. 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. Otherwise some results would not be correct (see Example 10 below).

For a more thorough outline of these technical details, see [6, Lemma 3.3.5 and Remark 3.3.6].

Example 6. We consider a first-order equation in one dependent variable

$$\mathcal{R}_1\colon \quad \begin{cases} u_n = \phi_n(\mathbf{x}, u, u_1, \ldots, u_{n-r-1}), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ u_{n-r} = \phi_{n-r}(\mathbf{x}, u, u_1, \ldots, u_{n-r-1}). \end{cases}$$

It is easy to show that the symbol of such an equation is always involutive and that the used coordinates are δ-regular. In order to simplify the notation, we use the following convention for the indices: $1 \leq k, \ell < n - r$ and $n - r \leq a, b \leq n$. Thus we may express our system in the concise form $u_a = \phi_a(\mathbf{x}, u, u_k)$ and local coordinates on the submanifold $\mathcal{R}_1$ are $(\overline{\mathbf{x}}, \overline{u}, \overline{u}_k)$.

The pull-back of the contact form $\omega = du - u_i dx^i$ generating the contact codistribution $\mathcal{C}_1^0$ is $\iota^*\omega = d\overline{u} - \overline{u}_k d\overline{x}^k - \phi_a d\overline{x}^a$ and $\mathcal{V}[\mathcal{R}_1] = \langle \bar{X}_1, \dots, \bar{X}_n, \bar{Y}_1, \dots, \bar{Y}_{n-r-1} \rangle$ where

$$\bar{X}_k = \partial_{\overline{x}^k} + \overline{u}_k \partial_{\overline{u}}, \qquad \bar{X}_a = \partial_{\overline{x}^a} + \phi_a \partial_{\overline{u}}, \qquad \bar{Y}_k = \partial_{\overline{u}_k}.$$

The fields $\bar{Y}_k$ span the symbol $\mathcal{N}_1$ and the fields $\bar{X}_i$ our choice of a reference complement $\mathcal{H}_0$. Setting $\bar{Z} = \partial_{\overline{u}}$, the structure equations of $\mathcal{V}[\mathcal{R}_1]$ are

$$\begin{aligned} &[\bar{X}_k, \bar{X}_\ell] = 0, && [\bar{Y}_k, \bar{Y}_\ell] = 0, \\ &[\bar{X}_k, \bar{X}_a] = \bar{X}_k(\phi_a)\bar{Z}, && [\bar{X}_a, \bar{X}_b] = \big(\bar{X}_a(\phi_b) - \bar{X}_b(\phi_a)\big)\bar{Z}, \\ &[\bar{X}_k, \bar{Y}_\ell] = -\delta_{k\ell}\bar{Z}, && [\bar{X}_a, \bar{Y}_k] = -\bar{Y}_k(\phi_a)\bar{Z}. \end{aligned}$$

Now we make the above discussed ansatz $U_i = X_i + \zeta_i^k Y_k$ for the generators of a transversal complement $\mathcal{H}$. Modulo the Vessiot distribution $\mathcal{V}[\mathcal{R}_1]$ we obtain for their Lie brackets

$$[U_k, U_\ell] \equiv (\zeta_k^\ell - \zeta_\ell^k) Z \mod \mathcal{V}[\mathcal{R}_1], \tag{36a}$$

$$[U_a, U_k] \equiv \big(\zeta_a^k - \zeta_k^\ell Y_\ell(\phi_a) - X_k(\phi_a)\big) Z \mod \mathcal{V}[\mathcal{R}_1], \tag{36b}$$

$$[U_a, U_b] \equiv \big(\zeta_a^k Y_k(\phi_b) - \zeta_b^\ell Y_\ell(\phi_a) + X_a(\phi_b) - X_b(\phi_a)\big) Z \mod \mathcal{V}[\mathcal{R}_1]. \tag{36c}$$

The algebraic system (33) is now obtained by requiring that all the expressions in parentheses on the right hand sides vanish. Its solution is straightforward. The first subsystem (36a) implies the equalities $\zeta_k^\ell = \zeta_\ell^k$. This result was to be expected by the discussion in Remark 5: both u_k and u_ℓ are parametric derivatives for $\mathcal{R}_1$ and thus we could have made this identification already in our ansatz for the complement. The second subsystem (36b) yields that $\zeta_a^k = \zeta_k^\ell Y_\ell(\phi_a) + X_k(\phi_a)$. If we enter these results into the third subsystem (36c), then all unknowns ζ drop out and the solvability condition

$$X_a(\phi_b) - X_b(\phi_a) + X_k(\phi_a) Y_k(\phi_b) - X_k(\phi_b) Y_k(\phi_a) = 0$$

arises. Thus in this example the algebraic system (33) has a solution if, and only if, this condition is satisfied.

Comparing with the classical theory of such systems, one easily verifies that this solvability condition is equivalent to the vanishing of the Mayer or Jacobi bracket $[u_a - \phi_a, u_b - \phi_b]$ on the submanifold $\mathcal{R}_1$, which in turn is a necessary and sufficient condition for the formal integrability of the differential equation $\mathcal{R}_1$. Thus we may conclude that $\mathcal{R}_1$ possesses n-dimensional integral distributions if, and only if, it is formally integrable (which in our case is also equivalent to $\mathcal{R}_1$ being involutive).

Thus here involution can be decided solely on the basis of the algebraic system (33). We will show in the next section that this observation does not represent a special property of a very particular class of differential equations but a general feature of the Vessiot theory.

7 The existence theorem for integral distributions

Now the question arises, when the combined system (33), (34) has solutions? We begin by analysing the algebraic part (33). We use for this analysis a step-by-step approach originally proposed by Vessiot [28]. Our analysis will automatically reveal necessary and sufficient assumptions for it to succeed. As outlined in Remark 5, we 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 from equation (35) 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 the field 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

$$\hat{\Xi}_1\hat{\zeta}_2 = \hat{\Xi}_2\hat{\zeta}_1 - \Theta_{12}. \tag{37}$$

Since the components of $\hat{\zeta}_1$ are not considered unknowns, the system (37) must not lead to any restrictions for the coefficients $\hat{\zeta}_1^k$. Obviously, this is the case if, and only if,

$$\operatorname{rank}\hat{\Xi}_1 = \operatorname{rank}\begin{pmatrix}\hat{\Xi}_1 & \hat{\Xi}_2\end{pmatrix}. \tag{38}$$

Assuming that (38) holds, the system (37) is solvable if, and only if, it satisfies the augmented rank condition

$$\operatorname{rank}\hat{\Xi}_1 = \operatorname{rank}\begin{pmatrix}\hat{\Xi}_1 & \hat{\Xi}_2 & -\Theta_{12}\end{pmatrix}. \tag{39}$$

When we have succeeded in constructing U_2, the next step is to seek a further 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

$$\hat{\Xi}_1\hat{\zeta}_3 = \hat{\Xi}_3\hat{\zeta}_1 - \Theta_{13}, \qquad \hat{\Xi}_2\hat{\zeta}_3 = \hat{\Xi}_3\hat{\zeta}_2 - \Theta_{23}.$$

Now this system is not to restrict the components of both $\hat{\zeta}_1$ and $\hat{\zeta}_2$ any further. The interrelations between the $\hat{\zeta}_i$ following from the condition (35) for the existence of integral distributions, given in Remark 5, are taken care of by contracting Ξ_3 into $\hat{\Xi}_3$. This implies that now the rank condition

$$\operatorname{rank}\begin{pmatrix}\hat{\Xi}_1\\ \hat{\Xi}_2\end{pmatrix} = \operatorname{rank}\begin{pmatrix}\hat{\Xi}_1 & \hat{\Xi}_3 & 0\\ \hat{\Xi}_2 & 0 & \hat{\Xi}_3\end{pmatrix}$$

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

$$\Theta_{13}^c - \hat{\Xi}_{3k}^c\zeta_1^k + \hat{\Xi}_{1k}^c\zeta_3^k = 0, \qquad \Theta_{23}^c - \hat{\Xi}_{3k}^c\zeta_2^k + \hat{\Xi}_{2k}^c\zeta_3^k = 0$$

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_j by solving the system

$$\begin{aligned}
&\hat{\Xi}_1\hat{\zeta}_j = \hat{\Xi}_j\hat{\zeta}_1 - \Theta_{1j},\\
&\dots\dots\dots\dots\dots\dots\dots\\
&\hat{\Xi}_{j-1}\hat{\zeta}_j = \hat{\Xi}_j\hat{\zeta}_{j-1} - \Theta_{j-1,j}.
\end{aligned} \tag{40}$$

Again we consider only the components of the vector $\hat{\zeta}_j$ unknowns, and the system (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

$$\operatorname{rank}\begin{pmatrix}\hat{\Xi}_1\\ \hat{\Xi}_2\\ \vdots\\ \hat{\Xi}_{j-1}\end{pmatrix} = \operatorname{rank}\begin{pmatrix}\hat{\Xi}_1 & \hat{\Xi}_j & & \\ \hat{\Xi}_2 & & \hat{\Xi}_j & 0\\ \vdots & 0 & & \ddots & \\ \hat{\Xi}_{j-1} & & & & \hat{\Xi}_j\end{pmatrix}. \tag{41}$$

Assuming that it holds, the equations (40) are solvable and yield solutions for the components of $\hat{\zeta}_j$ if, and only if, the system satisfies the augmented rank condition

$$\operatorname{rank}\begin{pmatrix}\hat{\Xi}_1\\ \hat{\Xi}_2\\ \vdots\\ \hat{\Xi}_{j-1}\end{pmatrix} = \operatorname{rank}\begin{pmatrix}\hat{\Xi}_1 & \hat{\Xi}_j & & & -\Theta_{1j}\\ \hat{\Xi}_2 & & \hat{\Xi}_j & 0 & -\Theta_{2j}\\ \vdots & 0 & & \ddots & \vdots\\ \hat{\Xi}_{j-1} & & & \hat{\Xi}_j & -\Theta_{j-1,j}\end{pmatrix}. \tag{42}$$

Remark 6. We can simplify our calculations for a differential equation $\mathcal{R}_q$, represented by $u^\alpha_\mu = \phi^\alpha_\mu(\mathbf{x}, \mathbf{u}, \hat{\mathbf{u}}^{(q)})$ where each equation is solved for a principal derivative (all pairwise different) u^α_μ with $|\mu| = q$, and where $\hat{\mathbf{u}}^{(q)}$ denotes the set of the remaining, thus parametric, derivatives of order q and less. (Example 1 discusses a system of that kind.) Now we can use the local coordinates on $\mathcal{R}_q$. For the generators of the symbol $\mathcal{N}_q$ we may choose for $(\alpha, \mu) \notin \mathcal{B}$ the vector fields $\bar{Y}^\alpha_\mu = \partial_{\overline{u^\alpha_\mu}}$; as a basis for the complement $\mathcal{H} \subseteq \mathcal{V}[\mathcal{R}_q]$, we can take the vector fields

$$\bar{X}_i = \partial_{\overline{x^i}} + \sum_{\alpha=1}^{m} \sum_{0\le|\mu|<q} \overline{u^\alpha_{\mu+1_i}}\,\partial_{\overline{u^\alpha_\mu}},$$

which satisfy equation (27). For the vector fields W_a which appear in equation (25) we may choose the contact vector fields C^μ_α where $(\alpha, \mu) \in \mathcal{B}$. The further procedure, solving first the structure equations, which now take the simple form (26), to find the generators of $\mathcal{V}'(\mathcal{R}_q)$ and then solving equation (33) for the coefficient functions ζ_j, is the same as in the case where the representation is not in the above solved form.

Example 7. Before we proceed with our theoretical analysis, let us demonstrate with a concrete differential equation that in general we cannot expect that the above outlined step-by-step construction of integral distributions works. In order to keep the size of the example reasonably small, we use the second-order equation[1] $\mathcal{R}_2$ defined by the system $u_{xx} = \alpha u$ and $u_{yy} = \beta u$ with two real constants α, β. Its symbol $\mathcal{N}_2$ is *not* involutive. However, one easily proves that $\mathcal{R}_2$ is formally integrable for arbitrary choices of the constants α, β. One readily computes that the Vessiot distribution $\mathcal{V}[\mathcal{R}_2]$ is generated by the following three vector fields on $\mathcal{R}_2$:

$$X_1 = \partial_{\overline{x}} + \overline{u_x}\partial_{\overline{u}} + \alpha\overline{u}\partial_{\overline{u_x}} + \overline{u_{xy}}\partial_{\overline{u_y}}, \qquad X_2 = \partial_{\overline{y}} + \overline{u_y}\partial_{\overline{u}} + \overline{u_{xy}}\partial_{\overline{u_x}} + \beta\overline{u}\partial_{\overline{u_y}}, \qquad Y_1 = \partial_{\overline{u_{xy}}}.$$

They yield as structure equations for $\mathcal{V}[\mathcal{R}_2]$:

$$[X_1, X_2] = \beta\overline{u_x}\partial_{\overline{u_y}} - \alpha\overline{u_y}\partial_{\overline{u_x}}, \qquad [X_1, Y_1] = -\partial_{\overline{u_y}}, \qquad [X_2, Y_1] = -\partial_{\overline{u_x}}.$$

For the construction of a two-dimensional integral distribution $\mathcal{U} \subset \mathcal{V}[\mathcal{R}_2]$ we make as above the ansatz $U_i = X_i + \zeta^1_i Y_1$ with two coefficients ζ^1_i. As we want to perform a step-by-step construction, we assume that we have chosen some fixed value for ζ^1_1 and try now to determine ζ^1_2 such that $[U_1, U_2] \equiv 0 \bmod \mathcal{V}[\mathcal{R}_2]$. Evaluation of the Lie bracket yields the equation

$$[U_1, U_2] \equiv (\beta\overline{u_x} - \zeta^1_2)\partial_{\overline{u_y}} - (\alpha\overline{u_y} - \zeta^1_1)\partial_{\overline{u_x}} \mod \mathcal{V}[\mathcal{R}_2]. \tag{43}$$

A necessary condition for the vanishing of the right side is that $\zeta^1_1 = \alpha\overline{u_y}$. Hence one cannot choose this coefficient arbitrarily, as we assumed, but (43) determines *both* functions ζ^1_i uniquely. Note that the conditions on the coefficients ζ^1_i imposed by (43) are trivially solvable and thus $\mathcal{R}_2$ possesses an integral distribution (which is even involutive). The problem is that it could not be constructed systematically with the above outlined *step-by-step* process.

[1] It could always be rewritten as a first-order one satisfying the assumptions made above, and the phenomenon we want to discuss is independent of this transformation.

The following theorem links the satisfaction of the rank conditions (41) and (42), and thus the solvability of the algebraic system (33) by the above described step-by-step process, with intrinsic properties of the differential equation $\mathcal{R}_q$ and its symbol $\mathcal{N}_q$. It represents an existence theorem for integral distributions.

Theorem 2. *Assume that δ-regular coordinates have been chosen for the local representation of the differential equation $\mathcal{R}_q$. Then the rank condition* (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* (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 2 is given in Appendix A, as parts of it are rather technical. First, we transform the differential equation into an equivalent first-order system with a representation in the reduced Cartan normal form (10) and the same Cartan characters. This is always possible; see, for example, [23, Proposition A.3.1]. The basic idea of the proof is then that the rank condition (41) is equivalent to the vanishing of the obstructions to involution of the symbol in Lemma 2, and the augmented rank condition (42) is equivalent to the vanishing of the remaining integrability conditions (recall that Lemma 2 gives all obstructions to involution only if the used coordinates are δ-regular so that this assumption is necessary in Theorem 2). The concrete implementation of this idea requires some technical considerations concerning the transformation of the matrices (41) and (42) into row echelon form, working out their contractions and analysing the interrelation between these operations.

Example 8. We demonstrate the role of δ-regularity of the coordinates with the wave equation $u_{xy} = 0$ in characteristic coordinates which are not δ-regular. A straightforward computation yields as generators for the Vessiot distribution the fields $X_1 = C_1^{(2)}$, $X_2 = C_2^{(2)}$, $Y_1 = \partial_{u_{xx}}$ and $Y_2 = \partial_{u_{yy}}$. Following our approach, we make the ansatz $U_1 = X_1 + \zeta_1^1 Y_1 + \zeta_1^2 Y_2$ and $U_2 = X_2 + \zeta_2^1 Y_1 + \zeta_2^2 Y_2$. Evaluation of the Lie bracket $[U_1, U_2]$ yields now the algebraic equations $\zeta_1^2 = 0$ and $\zeta_2^1 = 0$ and the differential equations $U_1(\zeta_2^1) - U_2(\zeta_1^1) = 0$ and $U_1(\zeta_2^2) - U_2(\zeta_1^2) = 0$. As in Example 7, we see that the step-by-step process is broken, since we obtain a condition on the coefficient $\zeta_1^2 = 0$ which we should be able to consider a parameter here.

At this point, we have proven that integral distributions within the Vessiot distribution exist if, and only if, the algebraic conditions (33) are solvable, and that this is equivalent to the augmented rank condition (41) being satisfied. This in turn is the case precisely if the differential equation is involutive. Thus we have characterised the existence of Vessiot connections for Vessiot's [28] step-by-step approach.

8 The existence theorem for flat Vessiot connections

There remains to analyse the solvability, if we add the differential system (34). 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 (33), (34). We first note that the set of differential conditions (34) alone is again an involutive system.

Proposition 7. *The differential conditions* (34) *alone represent an involutive differential equation of first order.*

The proof, which is somewhat technical, is given in Appendix B. It is based on an evaluation of the Jacobi identity for the Lie brackets of the fields U_i.

If the original equation $\mathcal{R}_1$ is analytic, then the quasi-linear system (34) is analytic, too. Thus we may apply the Cartan–Kähler theorem (Theorem 1) to it which guarantees the existence of solutions. The problem is that the combined system (33), (34) is not necessarily involutive, as the prolongation of the algebraic equations (33) may lead to additional differential equations. Instead of analysing the effect of these integrability conditions, we proceed as follows. If we assume that $\mathcal{R}_1$ is involutive, then we know from Theorem 2 that the algebraic equations (33) are solvable. Actually, the proof of the Theorem 2 produces an explicit row echelon form of the system matrix given on the right hand side of equation (41). Now we use the interrelations between the unknowns $\zeta_\ell^{(\alpha,h)}$, where $(\alpha,h) \notin \mathcal{B}$, to eliminate in the differential conditions (34) some of them by expressing them as linear combinations of the remaining ones. Let $1 \le i < j \le n$; then for all (α,i) where $\beta_1^{(i)} + 1 \le \alpha \le \beta_1^{(j)}$ we find

$$\zeta_j^{(\alpha,i)} = \sum_{k=1}^{j} \sum_{\gamma=\beta_1^{(k)}+1}^{m} C_\gamma^k(\phi_j^\alpha)\zeta_i^{(\gamma,k)} + C_i^{(1)}(\phi_j^\alpha), \tag{44}$$

while for all (α,i) where $\beta_1^{(j)} + 1 \le \alpha \le m$ the correspondence (35), given in Remark 5, holds again (see [6, Corollary 3.3.23] for the detailed calculation). In this way, we plug the algebraic conditions into the differential conditions and thus get rid of them.

Example 9. For Example 1, where $i = 1$ and $j = 2$, we deduce for

$$U_1 = X_1 + \zeta_1^{v_x}\partial_{v_x} + \zeta_1^{w_x}\partial_{w_x} \qquad \text{and} \qquad U_2 = X_2 + \zeta_2^{v_x}\partial_{v_x} + \zeta_2^{w_x}\partial_{w_x}$$

from the algebraic condition (37) that $-\zeta_2^{v_x} = -\zeta_1^{v_x}$ and $-\zeta_2^{w_x} = -\zeta_1^{w_x}$. This is equation (44) for $(\alpha,i) \equiv v_x$ and $(\alpha,i) \equiv w_x$. The rank conditions (38) and (39) are satisfied because $v_x - w_t = 0$, so that the matrix $(\Xi_1\ \Xi_2\ \ -\Theta_{12})$ has a vanishing first row. (There is no contraction of matrices for $j = 2$.) There are no interrelations like those in equation (35) because there is no index value α for which both $\iota_*(\partial_{u_1^\alpha})$ and $\iota_*(\partial_{u_2^\alpha})$ would be symbol fields. The differential conditions (34) are

$$U_1(\zeta_2^{v_x}) - U_2(\zeta_1^{v_x}) = 0, \qquad U_1(\zeta_2^{w_x}) - U_2(\zeta_1^{w_x}) = 0,$$

and by substituting $\zeta_2^{v_x} = \zeta_1^{v_x}$ and $\zeta_2^{w_x} = \zeta_1^{w_x}$, we can drop the algebraic conditions from the system. Since $n = 2$, there is no further procedure.

We can now prove the following existence theorem for flat Vessiot connections.

Theorem 3. *Assume that δ-regular coordinates have been chosen for the local representation of the analytic differential equation $\mathcal{R}_1$. Then the combined system* (33), (34) *is solvable.*

The idea of the proof is this: following the strategy we have just outlined, we eliminate some of the unknowns $\hat{\zeta}_i^k$. Because of the simple structure of (34), it turns out that we must take a closer look only at those equations where the leading derivative is of one of the unknowns we eliminate. A somewhat lengthy but straightforward computation shows that these equations actually vanish. The remaining equations still form an involutive system. Thus we eventually arrive at an analytic involutive differential equation for the coefficient functions $\hat{\zeta}_i^k$ which is solvable according to the Cartan–Kähler theorem. The details are worked out in Appendix C.

Example 10. Consider the first-order equation

$$\mathcal{R}_1\colon \begin{cases} u_t = v_t = w_t = u_s = 0, & v_s = 2u_x + 4u_y, \\ w_s = -u_x - 3u_y, & u_z = v_x + 2w_x + 3v_y + 4w_y \end{cases}$$

in the five independent variables x, y, z, s, t and the three dependent variables u, v, w. 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 are 3×8-matrices, we find

$$\Xi_1 = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \Xi_2 = \begin{pmatrix} 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{pmatrix},$$

$$\Xi_3 = \left(\begin{smallmatrix} 0 & -1 & -2 & 0 & -3 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{smallmatrix}\right), \qquad \Xi_4 = \left(\begin{smallmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & -4 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \end{smallmatrix}\right), \qquad \Xi_5 = 0_{3\times 8}.$$

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 (which is 8). From then on, the rank of the left matrix in (41) stagnates at $\dim \mathcal{N}_1$ while the rank of the augmented matrix rises further. The entries breaking the rank condition differ by their sign, while their corresponding coefficients in Lemma 2 are collected into one sum and thus vanish. Here it shows that contracting the matrices, as explained in Remark 5, is necessary.

9 Conclusions

Vessiot's [28] original motivation for the introduction of his theory was to provide an alternative proof of the Cartan–Kähler theorem (the same holds for Stormark's presentation [24] of the theory). If one takes this point of view, then one may say that our proof of Theorem 3 is a "cheat", as it uses the Cartan–Kähler theorem instead of proving it. However, in our opinion, this point of view is the main reason why the corresponding proofs in [24, 28] are so difficult to read (and actually incomplete as they neglect that the involutivity of the system is a necessary condition). The direct proof of the Cartan–Kähler theorem (for differential equations) given in textbooks on the formal theory like [19, 23] is much simpler and more transparent; in particular, it makes clear where involution (as opposed to mere formal integrability) is needed. Thus, if the only goal consisted of proving such an existence and uniqueness theorem, then there would be no need to bother with Vessiot's theory.

We believe that the real value of Vessiot's theory lies in the fact that it provides via the distribution $\mathcal{V}[\mathcal{R}_q]$ an additional geometric structure on an involutive differential equation $\mathcal{R}_q$ which is very useful for the *further* analysis of the equation, that is, after its solvability has been established. One possible application is the investigation of certain forms of singular behaviour of solutions. In fact, the classical works of Arnold and collaborators (see [2] for an elementary introduction) on implicit ordinary differential equations are based on the Vessiot distribution (without using this terminology); for some further works also in the context of partial differential equations or numerical analysis see, for example, [16, 25, 26]. The basic idea here is that such behaviour mainly stems from the fact that at some points the considered involutive distributions cease to be transversal to the fibration $\mathcal{R}_q \to \mathcal{X}$ so that integral manifolds can no longer be interpreted as prolonged solutions in the classical sense. However, it often makes sense to consider them a generalised form of (potentially multi-valued) solutions.

Another application concerns symmetry theory. In classical symmetry theory one always considers the differential equation $\mathcal{R}_q$ as a submanifold of $J_q\pi$ and looks then for diffeomorphisms of $J_q\pi$ which are (i) compatible with the contact structure of $J_q\pi$ and (ii) leave $\mathcal{R}_q$ invariant. It was only fairly late realized [1] that this approach yields only what is now called external symmetries and that in general further useful symmetries may exist. Using Vessiot's theory, we may represent a differential equation as the pair $(\mathcal{R}_q, \mathcal{V}[\mathcal{R}_q])$, that is, as a manifold together with a distribution on it – without any recourse to an ambient space (all relevant properties of $J_q\pi$ are captured in the Vessiot distribution $\mathcal{V}[\mathcal{R}_q]$). This point of view yields automatically all internal symmetries (and is equivalent to the symmetry theory of exterior differential systems).

We mentioned already in the Introduction that Vessiot's theory takes an intermediate position between the formal theory and the theory of exterior differential systems. It allows for the transfer of ideas from the latter to the former one without the need to rewrite a differential equation as an exterior system. For example, [23, Section 9.5] contains a formulation of prolongation structures without exterior forms using the Vessiot distribution instead[2]. It should also be possible to give an *explicit* proof of the equivalence of the formal theory and the Cartan–Kähler theory on the basis of the results presented here.

Finally, we comment on the practical side of the Vessiot theory. Throughout this article, we made a number of assumptions on the treated differential equation: we assumed that it is first rewritten as a first-order system, then that all present algebraic equations are explicitly solved and finally that each equation is solved for its principal derivative. While, from a theoretical point of view, these operations are always possible under fairly mild regularity assumptions, usually they cannot be performed effectively. However, the made assumptions are only needed in order to be able to prove results like Lemma 2 on the explicit form of the obstructions to involution. These expressions are already bad enough with our simplifying assumptions; for more general systems they would hardly be manageable.

A concrete differential equation need not be transformed to such a special normal form in order to apply the Vessiot theory. Even for fully implicit equations of arbitrary order, the determination of all integral distributions requires only linear algebra and is easily implemented in a computer algebra system. In fact, already Fackerell [5] reported of an implementation of Vessiot's theory. Within a recent project thesis [8], Globke implemented large parts of the theory in MuPAD. The problems and limitations are here exactly the same as in implementations of the Cartan–Kähler theory (see, for example, [10]).

A Proof of Theorem 2

In this section of the appendix, we prove Theorem 2, the existence theorem for integral distributions. It is in principle by straightforward matrix calculation and involves a tedious distinction of several cases and subcases: we have to compare the entries in the matrix on the right hand side of equation (42) for step j after turning it into row echelon form with the integrability conditions and the obstructions to involution as they are given in Lemma 2. For this purpose, we fix $1 \leq i < j$ and consider the block $\hat{\Xi}_i$ and the entries to its right in 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 (42), is equivalent to the differential equation being involutive, which in turn is the case if, and only if, the algebraic conditions (33) are satisfied, which is necessary and sufficient for the existence of integral distributions within the Vessiot distribution.

If the order of the differential equation is $q > 1$, transform it into an equivalent first-order equation; for the details of this procedure, see [6, Subsection 2.5.1] or [23]. If $\mathcal{R}_q$ is involutive, then so is $\mathcal{R}_1$ (see [23] for a straightforward proof of this). We assume this first-order equation is represented in reduced Cartan normal form (10). We first prove the rank condition (41) for the homogeneous system. (The proof for the augmented rank condition (42) follows.) We proceed in two steps: We consider the *complete matrix* at the jth step,

$$\begin{pmatrix} \Xi_1 & \Xi_j & & & \\ \Xi_2 & & \Xi_j & & 0 \\ \vdots & 0 & & \ddots & \\ \Xi_{j-1} & & & & \Xi_j \end{pmatrix}. \tag{45}$$

[2]Both symmetry theory and prolongation structures were also discussed by Fackerell [5] via Vessiot theory.

(It differs from the matrix on the right hand side of equation (41) in that it is not yet contracted.) 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 $1 \le i, k \le j-1$, be $[i,k]$ the kth block right of a Ξ_i. Then, for all $1 \le i, k \le j-1$, we have

$$[i,k] = \begin{cases} \Xi_j, & i = k, \\ 0_{m\times r}, & i \ne k. \end{cases}$$

For convenience, we set $[i,0] := \Xi_i$. Let, for $1 \le g \le h \le n$,

$$[\Xi_i]^{g\ldots h} \tag{46}$$

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. Let $[i,k]^h$ be the hth block of columns in $[i,k]$, and let, for $1 \le g \le h \le n$, in analogy to the shorthand (46),

$$[i,k]^{g\ldots h}$$

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 ${}^d_c[M]$ be the matrix made from the *rows* with indices (that is, labels) c to d, $[M]^f_e$ the matrix made from the *columns* with indices e to f and ${}^d_c[M]^f_e$ 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 ${}^c_1[M]^g$ (with g up right) to show that the first c upper rows are being selected, and we write ${}^b_d[M]_g$ (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 ${}^d_c[M]g$. This notation is redundant in that the position of g does not give new information, but in the calculations to come it increases readability.

We first turn the complete matrix into row echelon form. Then we contract columns and consider the effect.

The matrices Ξ_i are shown in equation (31) and have a block-structure of the form

$$\begin{pmatrix} a_{\alpha r} & b_{\alpha s} & c_{\alpha t} \\ 0_{S\times R} & -\mathbb{1}_S & d_{st} \end{pmatrix} =: (g_{ij}\colon 1 \le i \le A+S, 1 \le j \le R+S+T),$$

where A, R, S, T are adequate natural numbers or zero, and the blocks are

$$\begin{aligned} &\left(a_{\alpha r}\colon 1 \le \alpha \le A, 1 \le r \le R\right), && \left(b_{\alpha s}\colon 1 \le \alpha \le A, 1 \le s \le S\right), \\ &\left(c_{\alpha t}\colon 1 \le \alpha \le A, 1 \le t \le T\right), && \left(d_{st}\colon 1 \le s \le S, 1 \le t \le T\right), \end{aligned}$$

the $S \times S$ unit matrix $\mathbb{1}_S$ or the zero matrix $0_{S\times R}$. (Here the index r is just some index and not supposed to be the dimension of any symbol.) Then the substitution

$$\begin{pmatrix} a_{\alpha r} & b_{\alpha s} & c_{\alpha t} \end{pmatrix} \leftarrow \begin{pmatrix} a_{\alpha r} & b_{\alpha s} & c_{\alpha t} \end{pmatrix} + (b_{\alpha s}) \cdot \begin{pmatrix} 0_{S\times R} & -\mathbb{1}_S & d_{st} \end{pmatrix} \tag{47}$$

transforms the matrix (g_{ij}) into

$$\begin{pmatrix} a_{\alpha r} & 0_{A\times S} & c_{\alpha t} + \sum_{s=1}^{S} b_{\alpha s} d_{st} \\ 0_{S\times R} & -\mathbb{1}_S & d_{st} \end{pmatrix}.$$

For the transformation of the complete matrix into row echelon form, we use this obvious method. Let $1 < j \le n$ be given. Choose $1 \le i < j$. Now we eliminate all the non-trivial entries in the rows

$$ {}^{\beta_1^{(i)}}_{1}\left[\Xi_i \, [i,1] \, [i,2] \dots [i,j-1]\right] \ . $$

We turn to the block ${}^{\beta_1^{(i)}}_{1}[\Xi_i]$ first. Non-trivial entries therein may appear according to equation (31) only for $1 \le h \le i$ in the blocks ${}^{\beta_1^{(i)}}_{1}[\Xi_i]^h$. So we fix $1 \le h \le i$ and consider the block ${}^{\beta_1^{(i)}}_{1}[\Xi_i]^h$. Now set

$$
\begin{aligned}
(a_{\alpha r}) &= {}^{\beta_1^{(i)}}_{1}[\Xi_i]^{1\dots h-1}, \\
(b_{\alpha s}) &= {}^{\beta_1^{(i)}}_{1}[\Xi_i]^{h}, \\
(c_{\alpha t}) &= {}^{\beta_1^{(i)}}_{1}\left[[\Xi_i]^{h+1\dots n}[i,1][i,2]\dots[i,j-1]\right], && \text{(48a)} \\
(d_{st}) &= {}^{m}_{\beta_1^{(h)}+1}\left[[\Xi_h]_{h+1\dots n}[h,1][h,2]\dots[h,j-1]\right], && \text{(48b)} \\
-\mathbb{1}_S &= -\mathbb{1}_{\alpha_1^{(h)}}, \\
0_{S\times R} &= 0_{\alpha_1^{(h)} \times \sum_{l=1}^{h-1}\alpha_1^{(l)}}.
\end{aligned}
$$

Then it follows that the substitution (47) 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 (48b) the entries $\sum_{s=1}^{S} b_{\alpha s} d_{st}$ here are of two types: from ${}^{m}_{\beta_1^{(h)}+1}[\Xi_h]_{h+1\dots n}$, the left part of (d_{st}), we have the entries in $(b_{\alpha s}) \cdot {}^{m}_{\beta_1^{(h)}+1}[\Xi_h]_{h+1\dots n}$, and from ${}^{m}_{\beta_1^{(h)}+1}[[h,1][h,2]\dots[h,j-1]]$, the right part of (d_{st}), we have those of $(b_{\alpha s}) \cdot {}^{m}_{\beta_1^{(h)}+1}[h,l]$ where $1 \le l \le j-1$.

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 equation (31), ${}^{m}_{\beta_1^{(h)}+1}[\Xi_h]_{h+1\dots n} = 0$, so

$$ (b_{\alpha s}) \cdot {}^{m}_{\beta_1^{(h)}+1}[\Xi_h]_{h+1\dots n} = 0. $$

It follows $c_{\alpha t} + \sum_{s=1}^{S} b_{\alpha s} d_{st} = c_{\alpha t}$, so that all entries in ${}^{\beta_1^{(i)}}_{1}[\Xi_i]^{h+1\dots n}$, the left part of (c_{st}), remain unchanged through the elimination of $(b_{\alpha s}) = {}^{\beta_1^{(i)}}_{1}[\Xi_i]^h$.

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 ${}^{m}_{\beta_1^{(h)}+1}[[h,1][h,2]\dots[h,j-1]]$, the right part of (d_{st}). Fix $1 \le l \le j-1$ and consider $(b_{\alpha s}) \cdot {}^{m}_{\beta_1^{(h)}+1}[h,l]$.

There are two cases: $h = l$ and $h \ne l$. For $h \ne l$ we have $[h,l] = 0$ according to the definition of the complete matrix (45), thus ${}^{m}_{\beta_1^{(h)}+1}[h,l] = 0$ and so

$$ (b_{\alpha s}) \cdot {}^{m}_{\beta_1^{(h)}+1}[h,l] = 0. $$

Again it follows $c_{\alpha t} + \sum_{s=1}^{S} b_{\alpha s} d_{st} = c_{\alpha t}$, so that for $h \ne l$ all entries in

$$ {}^{\beta_1^{(i)}}_{1}\left[[i,1][i,2]\dots[i,j-1]\right], $$

the right part of $(c_{\alpha t})$, remain unchanged through the elimination of $(b_{\alpha s}) = {}^{\beta_1^{(i)}}_{1}[\Xi_i]^h$, too.

For $h = l$, we have $[h, h] = \Xi_j$. The structure of Ξ_j, which is (31) with i replaced by j, implies that non-vanishing entries are possible in the blocks ${}^{\;\;m}_{\beta_1^{(h)}+1}[h, h]_k$ where $1 \leq k \leq j$. In fact

$$
\begin{aligned}
&{}^{\;\;m}_{\beta_1^{(j)}+1}[h, h]_k = 0 \qquad \text{for} \;\; 1 \leq k \leq j-1 \quad \text{and} && \text{(49a)}\\
&{}^{\;\;m}_{\beta_1^{(j)}+1}[h, h]_j = -\mathbb{1}_{\alpha_1^{(j)}}. && \text{(49b)}
\end{aligned}
$$

The entries $\sum_{s=1}^{S} b_{\alpha s} d_{st}$ for those d_{st} within ${}^{\;\;m}_{\beta_1^{(h)}+1}[h, h]_k$ are

$$
(b_{\alpha s}) \cdot {}^{\;\;m}_{\beta_1^{(h)}+1}[h, h]_k = {}^{\beta_1^{(i)}}_{1}[\Xi_i]^h \cdot {}^{\;\;m}_{\beta_1^{(h)}+1}[\Xi_j]_k = \left(\sum_{s=1}^{S} C^h_{\beta_1^{(h)}+s}(\phi_i^\alpha) C^k_{\beta_1^{(k)}+t}(\phi_j^{\beta_1^{(h)}+s}) \right). \tag{50}
$$

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 (50), this entry is

$$
\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(h)}+\alpha_1^{(h)}} C^h_\gamma(\phi_i^\alpha) C^k_\delta(\phi_j^\gamma) \tag{51}
$$

for all $1 \leq k \leq j$. Some of the $\beta_1^{(h)} + \alpha_1^{(h)} = m$ summands vanish because of the special entries (49). 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 (51) 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 $\beta_1^{(j)} + 1 \leq \delta \leq m$ we have

$$
C^h_\delta(\phi_i^\alpha) C^j_\delta(\phi_j^\delta) = \delta_{kj} \cdot C^h_\delta(\phi_i^\alpha).
$$

Now (51) becomes for all $1 \leq \alpha \leq \beta_1^{(i)}$ and all $\beta_1^{(k)} + 1 \leq \delta \leq m$

$$
\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi_i^\alpha) C^k_\delta(\phi_j^\gamma) + \delta_{kj} \cdot C^h_\delta(\phi_i^\alpha). \tag{51$'$}
$$

These are the terms $\sum_{s=1}^{S} b_{\alpha s} d_{st}$ in the case $h = l$ when $1 \leq k \leq j$. Now for the subcase $j + 1 \leq k \leq n$: 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 (31) 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 (48a) are those in the block ${}^{\beta_1^{(i)}}_{1}[i, h]$. For $h \neq i$, 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 (31) 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 $1 \leq k \leq j$, those of the form $-C^k_\delta(\phi_j^\alpha)$. They make up the block ${}^{\beta_1^{(i)}}_{1}[\Xi_j]^{1\ldots j} = {}^{\beta_1^{(i)}}_{1}[i, i]^{1\ldots j}$.

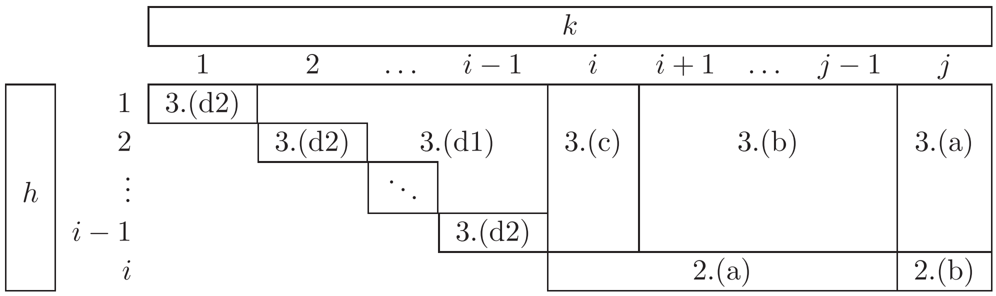

Figure 2. Possible combinations of h and k in the terms u^{δ}_{hk} used as labels for columns in the contracted matrix. Each term u^{δ}_{hk} corresponds to one subcase in the consideration of cases 2. $h = i$ and 3. $h < i$. Each of these subcases refers to one of the terms in square brackets in Lemma 2 as follows: 2.(a) – line (17), 2.(b) – line (19), 3.(a) – line (18), 3.(b) – line (16), 3.(c) – line (15), 3.(d1) – line (14), 3.(d2) – line (13).

So for any row index α and any column index $\delta = \beta_1^{(k)} + t$ we have $c_{\alpha t} = -\delta_{hi} C^k_\delta(\phi^\alpha_j)$ as the most general form of an entry.

Since $(51')$ 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

$$-\delta_{hi} C^k_\delta(\phi^\alpha_j) + \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i) C^k_\delta(\phi^\gamma_j) + \delta_{kj} C^h_\delta(\phi^\alpha_i). \tag{52}$$

Since $1 \leq \alpha \leq \beta_1^{(i)}$ and $\beta_1^{(k)} + 1 \leq \delta \leq m$, 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 ${}_{1}^{\beta_1^{(i)}}[\Xi_i]^{i+1\ldots n}$ for some $i < j$ (in which case it remains zero throughout the elementary row transformations and so does not influence the rank). 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 ${}_{1}^{\beta_1^{(i)}}[i, h]$ with $1 \leq h \leq j$. 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, (31), all the $C^h_\gamma(\phi^\alpha_i)$ and $C^h_\delta(\phi^\alpha_i)$ in (52) vanish. Since $h \neq i$, $\delta_{hi} = 0$, thus all of (52) vanishes.
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} where h and k are combined as shown in Fig. 2. For fixed δ, any u^δ_{hk} belongs, according to its indices h and k, to exactly one of the blocks marked by the case denominations 2.(a), 2.(b) and 3.(a) to 3.(d2).

   It turns out that (52) is the common form of all the sums that appear in the squared brackets of Lemma 2 as the coefficients of second-order derivatives u^δ_{hk}. Not only can the case distinctions of the following argument, 2.(a), 2.(b) and 3.(a) to 3.(d2), be labelled by these u^δ_{hk}, but in fact a case labelled u^δ_{hk} is dealt with by using the fact that according to Lemma 2 the coefficient in square brackets of that same u^δ_{hk} vanishes for an involutive system. The correspondence of the cases and the terms u^δ_{hk} is given in the caption of Fig. 2.

   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 (52) becomes

$$-C^k_\delta(\phi^\alpha_j) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C^i_\gamma(\phi^\alpha_i) C^k_\delta(\phi^\gamma_j)$$

where $\beta_1^{(k)} + 1 \le \delta \le m$. This vanishes for a system with an involutive symbol according to Lemma 2, line (17).

(b) For the second subcase of $h = i$ choose $k = j$. Then (52) becomes

$$-C^j_\delta(\phi^\alpha_j) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C^i_\gamma(\phi^\alpha_i) C^j_\delta(\phi^\gamma_j) + C^i_\delta(\phi^\alpha_i)$$

where $\beta_1^{(j)} + 1 \le \delta \le m$. According to Lemma 2, line (19), 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 (52) becomes

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i) C^j_\delta(\phi^\gamma_j) + C^h_\delta(\phi^\alpha_i)$$

where $\beta_1^{(j)} + 1 \le \delta \le m$. According to Lemma 2, line (18), this vanishes for a system with an involutive symbol.

(b) For the second subcase choose $i < k < j$. Then (52) becomes

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i) C^k_\delta(\phi^\gamma_j)$$

where $\beta_1^{(k)} + 1 \le \delta \le m$. According to Lemma 2, line (16), this vanishes for a system with an involutive symbol.

(c) For the third subcase choose $k = i$. Then (52) becomes

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i) C^i_\delta(\phi^\gamma_j) \tag{53}$$

where $\beta_1^{(k)} + 1 \le \delta \le m$. Since we have $h < k = i$, for any such δ the cross derivative $u^\delta_{ih} = u^\delta_{hi}$ labels two columns in the complete matrix: the one with label δ in ${}^{\beta_1^{(i)}}_{1}[i,h]^i$ and the one with label δ in ${}^{\beta_1^{(i)}}_{1}[i,i]^h$ which according to (52) has the new entries

$$-\delta_{ki} C^h_\delta(\phi^\alpha_j) + \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C^k_\gamma(\phi^\alpha_i) C^h_\delta(\phi^\gamma_j) + \delta_{hj} C^k_\delta(\phi^\alpha_i) \tag{54}$$

where $\beta_1^{(h)} + 1 \leq \delta \leq m$. In the current subcase, (54) becomes

$$-C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C_\gamma^i(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) \tag{55}$$

where $\beta_1^{(h)} + 1 \leq \delta \leq m$. Since the columns of both (53) and (55) 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, line (15a), 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, line (15b).

(d) For the fourth subcase choose $k < i$. Under this assumption, we have to distinguish 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 (52) 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, line (13), this vanishes for a system with an involutive symbol.

(d2) At last consider $h < k < i$. Then $k < j$, and (52) becomes

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^k(\phi_j^\gamma) \tag{56}$$

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 (52) has the new entries

$$-\delta_{ki} C_\delta^h(\phi_j^\alpha) + \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) + \delta_{hj} C_\delta^k(\phi_i^\alpha) \tag{57}$$

where $\beta_1^{(h)}+1 \le \delta \le m$. In the current subcase, (57) becomes

$$\sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma) \tag{58}$$

where $\beta_1^{(h)}+1 \le \delta \le m$. Since the columns of both (56) and (58) are labelled by the same second-order derivatives, we have to contract them, which means adding their new entries. This yields

$$\sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma)$$

for $\beta_1^{(h)}+1 \le \delta \le \beta_1^{(k)}$ and vanishes according to Lemma 2, line (14a), and

$$\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha)C_\delta^k(\phi_j^\gamma) + \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha)C_\delta^h(\phi_j^\gamma)$$

for $\beta_1^{(k)}+1 \le \delta \le m$ and vanishes according to Lemma 2, line (14b).

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 (41) is satisfied.

Now for the augmented rank condition, equation (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 $1 < j \le n$. Let $1 \le i < j$. Then, according to the structure of the augmented complete matrix, for its transformation into row echelon form we have to eliminate for $1 \le h \le i$ the entries in the blocks $[\Xi_i]^h$, as we did for the non-augmented complete matrix given in equation (45). 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 $-\Theta_{ij}^1, -\Theta_{ij}^2, \dots, -\Theta_{ij}^m$, given in equation (28). Of these entries, only $-\Theta_{ij}^1, -\Theta_{ij}^2, \dots, -\Theta_{ij}^{\beta_1^{(i)}}$ 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 $1 \le h \le i$. Then augment the matrix (d_{st}), used in the process of eliminating the entries in $[\Xi_i]^h$, by adding the entries $-\Theta_{hj}^{\beta_1^{(h)}+1}, -\Theta_{hj}^{\beta_1^{(h)}+2}, \dots, -\Theta_{hj}^m$ as its rightmost column, in accordance with the structure of the augmented complete matrix. Then the substitution (47) 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 $1 \le \alpha \le \beta_1^{(i)}$. Then the entry $c_{\alpha t} = -\Theta_{ij}^\alpha$ transforms as follows: For all $1 \le h \le i$ 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 $(-\Theta_{hj}^{\beta_1^{(h)}+1}, -\Theta_{hj}^{\beta_1^{(h)}+2}, \ldots, -\Theta_{hj}^m)$. According to the structure of $[\Xi_i]^h$ as defined in equation (30), 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 (28). 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 (12) of Lemma 2, except for the sign. The difference in sign comes from the fact that the augmented complete matrix describes the system of equations (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 for each i, j and α the corresponding equation is set to zero, if the differential equation is involutive.

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.

B Proof of Proposition 7

We are going to show that the system (34) is involutive. Consider the basis $(U_i\colon 1 \leq i \leq n)$ of the distribution $\mathcal{U}$ given by $U_i = X_i + \zeta_i^p Y_p$ with yet undetermined coefficient functions $\zeta_i^p \in \mathcal{F}(\mathcal{R}_1)$. For all $1 \leq i \leq n$ and $1 \leq p \leq r$, the independent variables of the functions ζ_i^p 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 (34), 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 (32) for the vector fields U_h and the definitions of G^c_{ij} and H^p_{ij} in equations (33) and (34) imply that this is

$$\begin{aligned}
0 &= [U_i, G^c_{jk} Z_c + H^p_{jk} Y_p] + [U_j, G^c_{ki} Z_c + H^p_{ki} Y_p] + [U_k, G^c_{ij} Z_c + H^p_{ij} Y_p] \\
&= G^c_{jk}[U_i, Z_c] + U_i(G^c_{jk}) Z_c + H^p_{jk}[U_i, Y_p] + U_i(H^p_{jk}) Y_p \\
&\quad + G^c_{ki}[U_j, Z_c] + U_j(G^c_{ki}) Z_c + H^p_{ki}[U_j, Y_p] + U_j(H^p_{ki}) Y_p \\
&\quad + G^c_{ij}[U_k, Z_c] + U_k(G^c_{ij}) Z_c + H^p_{ij}[U_k, Y_p] + U_k(H^p_{ij}) Y_p.
\end{aligned}$$

The combined system (33, 34) means that all $G^c_{ab} = 0$ and all $H^p_{ab} = 0$ 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

$$0 = \{U_i(G^c_{jk}) + U_j(G^c_{ki}) + U_k(G^c_{ij})\} Z_c + \{U_i(H^p_{jk}) + U_j(H^p_{ki}) + U_k(H^p_{ij})\} Y_p.$$

As part of a basis for $\mathcal{V}'[\mathcal{R}_q]$, the vector fields Z_c and Y_p are linearly independent, which means their coefficients must vanish individually. So in particular

$$U_i(H^p_{jk}) + U_j(H^p_{ki}) + U_k(H^p_{ij}) = 0.$$

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 (34) 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 $\partial\zeta^p_i/\partial x^j =: (\zeta^p_i)_j$ for the leaders and

$$\tilde{U}_j(\zeta^p_i) := U_j(\zeta^p_i) - (\zeta^p_i)_j$$

and solve each equation of the system (34) for its leader, then it takes the form

$$\begin{aligned}
&(\zeta^p_1)_n = U_1(\zeta^p_n) - \tilde{U}_n(\zeta^p_1), \\
&(\zeta^p_2)_n = U_2(\zeta^p_n) - \tilde{U}_n(\zeta^p_2), \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
&(\zeta^p_{n-1})_n = U_{n-1}(\zeta^p_n) - \tilde{U}_n(\zeta^p_{n-1}), \\
&(\zeta^p_1)_{n-1} = U_1(\zeta^p_{n-1}) - \tilde{U}_{n-1}(\zeta^p_1), \\
&(\zeta^p_2)_{n-1} = U_2(\zeta^p_{n-1}) - \tilde{U}_{n-1}(\zeta^p_2), \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
&(\zeta^p_{n-2})_{n-1} = U_{n-2}(\zeta^p_{n-1}) - \tilde{U}_{n-1}(\zeta^p_{n-2}), \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
&(\zeta^p_2)_3 = U_2(\zeta^p_3) - \tilde{U}_3(\zeta^p_2), \\
&(\zeta^p_1)_3 = U_1(\zeta^p_3) - \tilde{U}_3(\zeta^p_1), \\
&(\zeta^p_1)_2 = U_1(\zeta^p_2) - \tilde{U}_2(\zeta^p_1);
\end{aligned} \tag{34*}$$

here for each line $1 \leq p \leq r$. Therefore the system (34*) is in Cartan normal form given in Definition 7.

Now one can prove (see [6, Lemma 2.4.29]) that such a differential equation $\mathcal{R}_1$ is involutive if, and only if, all non-multiplicative prolongations of the equations (8a)–(8c) and all formal derivatives with respect to all the x^i of the algebraic equations (8d) are dependent on the equations of the system (8) and its multiplicative prolongations only. Therefore, it follows that the system (34) is involutive.

C Proof of Theorem 3

We consider the combined system of algebraic and differential conditions (33), (34) and want to show that it has a solution. We follow the strategy outlined above and eliminate some of the unknowns ζ_ℓ^p. As we consider each of the equations of (34) as being solved for its derivative $\partial\zeta_i^{(\beta,h)}/\partial x^j$ of highest class j, as given in equation (34*), 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 equations (35) and (44), 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 (34), the equations which hold the following terms are concerned:

$$
\begin{aligned}
&U_k\big(\zeta_2^{(\beta,1)}\big),\\
&U_k\big(\zeta_3^{(\beta,1)}\big),\ U_k\big(\zeta_3^{(\beta,2)}\big),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&U_k\big(\zeta_{k-1}^{(\beta,1)}\big),\ U_k\big(\zeta_{k-1}^{(\beta,2)}\big),\ \dots,\ U_k\big(\zeta_{k-1}^{(\beta,k-2)}\big);
\end{aligned}
$$

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

$$
U_i\big(\zeta_k^{(\beta,h)}\big) = U_k\big(\zeta_i^{(\beta,h)}\big). \tag{59}
$$

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 (59) becomes

   $$
   U_i\big(\zeta_h^{(\beta,k)}\big) = U_k\big(\zeta_h^{(\beta,i)}\big). \tag{60}
   $$

   Since the system (34) contains the equalities $U_i\big(\zeta_h^{(\beta,k)}\big) = U_h\big(\zeta_i^{(\beta,k)}\big)$ and $U_k\big(\zeta_h^{(\beta,i)}\big) = U_h\big(\zeta_k^{(\beta,i)}\big)$, equation (60) becomes

   $$
   U_h\big(\zeta_i^{(\beta,k)}\big) = U_h\big(\zeta_k^{(\beta,i)}\big).
   $$

   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 (59) vanishes.

   (b) The other interrelation is an affine-linear combination:

   $$
   \zeta_k^{(\beta,h)} = \sum_{a=1}^{k} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a\big(\phi_k^\beta\big)\zeta_h^{(\gamma,a)} + C_h^{(1)}\big(\phi_k^\beta\big).
   $$

   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 (59) becomes

   $$
   U_i\big(\zeta_k^{(\beta,h)}\big) = \sum_{a=1}^{k} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_k^\beta)U_i\big(\zeta_h^{(\gamma,a)}\big) \tag{61a}
   $$

$$+\sum_{a=1}^{k}\sum_{\gamma=\beta_1^{(a)}+1}^{m} U_i\big(C_\gamma^a(\phi_k^\beta)\big)\zeta_h^{(\gamma,a)} + U_i\big(C_h^{(1)}(\phi_k^\beta)\big). \tag{61b}$$

The term $U_k(\zeta_i^{(\beta,h)})$ in equation (59) becomes

$$\begin{aligned}
U_k\big(\zeta_i^{(\beta,h)}\big) &= U_k\big(\zeta_h^{(\beta,i)}\big) = U_h\big(\zeta_k^{(\beta,i)}\big)\\
&= U_h\left(\sum_{a=1}^{k}\sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_k^\beta)\zeta_i^{(\gamma,a)} + C_i^{(1)}(\phi_k^\beta)\right)\\
&= \sum_{a=1}^{k}\sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_k^\beta)U_h\big(\zeta_i^{(\gamma,a)}\big) && \text{(62a)}\\
&\quad + \sum_{a=1}^{k}\sum_{\gamma=\beta_1^{(a)}+1}^{m} U_h\big(C_\gamma^a(\phi_k^\beta)\big)\zeta_i^{(\gamma,a)} + U_h\big(C_i^{(1)}(\phi_k^\beta)\big); && \text{(62b)}
\end{aligned}$$

here we have the first equality because we are considering the first main case, the second equality because of the structure of the system (34) 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 (61) and (62) in equation (59) and factoring out, we get

$$\begin{aligned}
0 &= \sum_{a=1}^{k}\sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a(\phi_k^\beta)\big\{U_i\big(\zeta_h^{(\gamma,a)}\big) - U_h\big(\zeta_i^{(\gamma,a)}\big)\big\} && \text{(63a)}\\
&\quad + \sum_{a=1}^{k}\sum_{\gamma=\beta_1^{(a)}+1}^{m} \big\{U_i\big(C_\gamma^a(\phi_k^\beta)\big)\zeta_h^{(\gamma,a)} - U_h\big(C_\gamma^a(\phi_k^\beta)\big)\zeta_i^{(\gamma,a)}\big\} && \text{(63b)}\\
&\quad + U_i\big(C_h^{(1)}(\phi_k^\beta)\big) - U_h\big(C_i^{(1)}(\phi_k^\beta)\big). && \text{(63c)}
\end{aligned}$$

Line (63c) contains the Lie bracket $[U_i, C_h^{(1)}](\phi_k^\beta)$. According to the structure of the system (34), the term (63a) vanishes. If the terms (63b) and (63c) vanish, too, then so does equation (34). Otherwise they form a new algebraic condition for (34), which can be solved for some function $\zeta_h^{(\beta,a)}$. Substituting this function in (34) does not change the classes or the numbers of the single equations therein. Thus, equation (59) 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 (59) becomes

$$
\begin{aligned}
0 = &\sum_{a=1}^{i} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a\big(\phi_i^\beta\big) U_k\big(\zeta_h^{(\gamma,a)}\big) - \sum_{b=1}^{k} \sum_{\delta=\beta_1^{(b)}+1}^{m} C_\delta^b\big(\phi_k^\beta\big) U_i\big(\zeta_h^{(\delta,b)}\big) && \text{(64a)}\\
&+ \sum_{a=1}^{i} \sum_{\gamma=\beta_1^{(a)}+1}^{m} U_k\big(C_\gamma^a\big(\phi_i^\beta\big)\big)\zeta_h^{(\gamma,a)} - \sum_{b=1}^{k} \sum_{\delta=\beta_1^{(b)}+1}^{m} U_i\big(C_\delta^b\big(\phi_k^\beta\big)\big)\zeta_h^{(\delta,b)} && \text{(64b)}\\
&+ U_k\big(C_h^{(1)}\big(\phi_i^\beta\big)\big) - U_i\big(C_h^{(1)}\big(\phi_k^\beta\big)\big). && \text{(64c)}
\end{aligned}
$$

In part (64a), 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 (34). Thus, equation (64) becomes

$$
\begin{aligned}
0 = &\sum_{a=1}^{i} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a\big(\phi_i^\beta\big) U_h\big(\zeta_k^{(\gamma,a)}\big) - \sum_{b=1}^{k} \sum_{\delta=\beta_1^{(b)}+1}^{m} C_\delta^b\big(\phi_k^\beta\big) U_h\big(\zeta_i^{(\delta,b)}\big) && \text{(64a}'\text{)}\\
&+ \text{(64b)} + \text{(64c)}.
\end{aligned}
$$

Factoring out the vector field U_h in part (64a′), this equals

$$
\begin{aligned}
0 = &\,U_h\Bigg(\sum_{a=1}^{i} \sum_{\gamma=\beta_1^{(a)}+1}^{m} C_\gamma^a\big(\phi_i^\beta\big) \zeta_k^{(\gamma,a)} - \sum_{b=1}^{k} \sum_{\delta=\beta_1^{(b)}+1}^{m} C_\delta^b\big(\phi_k^\beta\big) \zeta_i^{(\delta,b)}\Bigg) && \text{(65a)}\\
&- \Bigg(\sum_{a=1}^{i} \sum_{\gamma=\beta_1^{(a)}+1}^{m} U_h\big(C_\gamma^a\big(\phi_i^\beta\big)\big)\zeta_k^{(\gamma,a)} - \sum_{b=1}^{k} \sum_{\delta=\beta_1^{(b)}+1}^{m} U_h\big(C_\delta^b\big(\phi_k^\beta\big)\big)\zeta_i^{(\delta,b)}\Bigg) && \text{(65b)}\\
&+ \text{(64b)} + \text{(64c)}. && \text{(65c)}
\end{aligned}
$$

Now one can show (see [6, Corollary 3.3.25 (for $j = k$)]) that the term (65a) equals

$$
U_h\big(C_i^{(1)}\big(\phi_k^\beta\big) - C_k^{(1)}\big(\phi_i^\beta\big)\big),
$$

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 (65a). The other terms, (65b) and (65c), are algebraic, too. So all of equation (59) 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 equations (35) and (44), are taken into account.

If this new algebraic condition for the system (34) vanishes, equation (59) vanishes. Otherwise, this new algebraic condition given in equation (65) now appears as

$$
\begin{aligned}
0 = &\,U_h\big(C_i^{(1)}\big(\phi_k^\beta\big) - C_k^{(1)}\big(\phi_i^\beta\big)\big) && \text{(65a}'\text{)}\\
&- \Bigg(\sum_{a=1}^{i} \sum_{\gamma=\beta_1^{(a)}+1}^{m} U_h\big(C_\gamma^a\big(\phi_i^\beta\big)\big)\zeta_k^{(\gamma,a)} - \sum_{b=1}^{k} \sum_{\delta=\beta_1^{(b)}+1}^{m} U_h\big(C_\delta^b\big(\phi_k^\beta\big)\big)\zeta_i^{(\delta,b)}\Bigg) && \text{(65b)}\\
&+ \sum_{a=1}^{i} \sum_{\gamma=\beta_1^{(a)}+1}^{m} U_k\big(C_\gamma^a\big(\phi_i^\beta\big)\big)\zeta_h^{(\gamma,a)} - \sum_{b=1}^{k} \sum_{\delta=\beta_1^{(b)}+1}^{m} U_i\big(C_\delta^b\big(\phi_k^\beta\big)\big)\zeta_h^{(\delta,b)} && \text{(64b)}\\
&+ U_k\big(C_h^{(1)}\big(\phi_i^\beta\big)\big) - U_i\big(C_h^{(1)}\big(\phi_k^\beta\big)\big). && \text{(64c)}
\end{aligned}
$$

Collecting terms in lines (65a′) and (64c), this yields

$$\begin{aligned}0 &= (64b) + (65b)\\ &\quad + U_h\big(C_i^{(1)}(\phi_k^\beta)\big) - U_i\big(C_h^{(1)}(\phi_k^\beta)\big) + U_k\big(C_h^{(1)}(\phi_i^\beta)\big) - U_h\big(C_k^{(1)}(\phi_i^\beta)\big).\end{aligned}$$

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 (64) for this non-vanishing factor and substitute it into the system (34), which does not change the class of any equation therein. Therefore equation (59) drops out from the system (34).

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 (34*), these are the equations with the leaders

$$\begin{aligned}&\big(\zeta_2^{(\beta,1)}\big)_k,\\ &\big(\zeta_3^{(\beta,1)}\big)_k,\ \big(\zeta_3^{(\beta,2)}\big)_k,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ &\big(\zeta_{k-1}^{(\beta,1)}\big)_k,\ \big(\zeta_{k-1}^{(\beta,2)}\big)_k,\ \ldots,\ \big(\zeta_{k-1}^{(\beta,k-2)}\big)_k;\end{aligned}$$

here $2 \le k \le n$ and $\beta_1^{(h)} + 1 \le \beta \le m$. The remaining equations still form an involutive system (we may enumerate the remaining ζ_i^p in such a way that no gaps appear) as the considerations for the system (34) in Proposition 7 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 (Theorem 1).

Acknowledgements

This work received partial financial support by the European NEST-Adventure grant 5006, *Global Integrability of Field Theories*.

References

[1] Anderson I., Kamran N., Olver P.J., Internal, external, and generalized symmetries, *Adv. Math.* **100** (1993), 53–100.

[2] Arnold V.I., Geometrical methods in the theory of ordinary differential equations, *Grundlehren der Mathematischen Wissenschaften*, Vol. 250, Springer-Verlag, New York, 1988.

[3] Bryant R.L., Chern S.S., Gardner R.B., Goldschmidt H.L., Griffiths P.A., Exterior differential systems, *Mathematical Sciences Research Institute Publications*, Vol. 18, Springer-Verlag, New York, 1991.

[4] Cartan É., Les Systèmes différentiels extérieurs et leurs applications géométriques, *Actualités Sci. Ind.*, no. 994, Hermann et Cie., Paris, 1945.

[5] Fackerell E.D., Isovectors and prolongation structures by vessiot's vector field formulation of partial differential equations, in Geometric Aspects of the Einstein Equations and Integrable Systems (Scheveningen, 1984), Editor R. Martini, *Lecture Notes in Phys.*, Vol. 239, Springer-Verlag, Berlin, 1985, 303–321.

[6] Fesser D., On Vessiot's theory of partial differential equations. A geometric approach to constructing infinitesimal solutions, PhD Thesis, Saarbrücken, 2008.

[7] Fesser D., Seiler W.M., Vessiot connections of partial differential equations, in Global Integrability of Field Theories – GIFT 2006, Editors J. Calmet, W.M. Seiler and R.W. Tucker, Universitätsverlag Karlsruhe. Karlsruhe, 2006, 111–134.

[8] Globke W., Eine Objektorientierte Programmierumgebung für Differentialgeometrische Berechnungen in MuPAD, Project Thesis, Fakultät für Informatik, Universität Karlsruhe, 2006.

[9] Golubitsky M., Guillemin V.W., Stable mappings and their singularities, *Graduate Texts in Mathematics*, Vol. 14, Springer-Verlag, New York – Heidelberg, 1973.

[10] Hartley D.H., Tucker R.W., A constructive implementation of the Cartan–Kähler theory of exterior differential systems, *J. Symbolic Comput.* **12** (1991), 655–667.

[11] Hausdorf M., Seiler W.M., An efficient algebraic algorithm for the geometric completion to involution, *Appl. Algebra Engrg. Comm. Comput.* **13** (2002), 163–207.

[12] Ivey T.A., Landsberg J.M., Cartan for beginners: differential geometry via moving frames and exterior differential systems, *Graduate Studies in Mathematics*, Vol. 61, American Mathematical Society, Providence, RI, 2003.

[13] Krasilshchik I.S., Lychagin V.V., Vinogradov A.M., Geometry of jet spaces and nonlinear partial differential equations, *Advanced Studies in Contemporary Mathematics*, Vol. 1, Gordon and Breach Science Publishers, New York, 1986.

[14] Kruglikov B.S., Lychagin V.V., Geometry of differential equations, in Handbook of Global Analysis, Editors D. Krupka and D. Saunders, Elsevier, Amsterdam, 2008, 725–771.

[15] Lychagin V.V., Homogeneous geometric structures and homogeneous differential equations, in The Interplay between Differential Geometry and Differential Equations, Editor V.V. Lychagin, *Amer. Math. Soc. Transl. Ser. 2*, Vol. 167, Amer. Math. Soc., Providence, RI, 1995, 143–164.

[16] Lychagin V.V., Singularities of multivalued solutions of nonlinear differential equations, and nonlinear phenomena, *Acta Appl. Math.* **3** (1985), 135–173.

[17] Modugno M., Covariant quantum mechanics, Unpublished manuscript, Dept. of Mathematics, University of Florence, 1999.

[18] Olver P.J., Applications of Lie groups to differential equations, *Graduate Texts in Mathematics*, Vol. 107, Springer-Verlag, New York, 1986.

[19] Pommaret J.F., Systems of partial differential equations and Lie pseudogroups, *Mathematics and its Applications*, Vol. 14, Gordon & Breach Science Publishers, New York, 1978.

[20] Saunders D.J., The geometry of jet bundles, *London Mathematical Society Lecture Note Series*, Vol. 142, Cambridge University Press, Cambridge, 1989.

[21] Seiler W.M., Completion to involution and semi-discretisations, *Appl. Numer. Math.* **42** (2002), 437–451.

[22] Seiler W.M., Spencer cohomology, differential equations, and Pommaret bases, in Gröbner Bases and Symbolic Analysis, Editors M. Rosenkranz and D.M. Wang, *Radon Ser. Comput. Appl. Math.*, Vol. 2, Walter de Gruyter, Berlin, 2007, 169–216.

[23] Seiler W.M., Involution – the formal theory of differential equations and its applications in computer algebra, *Algorithms and Computation in Mathematics*, Vol. 24, Springer-Verlag, Berlin – Heidelberg, 2009.

[24] Stormark O., Lie's structural approach to PDE systems, *Encyclopedia of Mathematics and its Applications*, Vol. 80, Cambridge University Press, Cambridge, 2000.

[25] Tuomela J., On singular points of quasilinear differential and differential-algebraic equations, *BIT* **37** (1997), 968–977.

[26] Tuomela J., On the resolution of singularities of ordinary differential equations, *Numer. Algorithms* **19** (1998), 247–259.

[27] Vassiliou P.J., Vessiot structure for manifolds of (p, q)-hyperbolic type: Darboux integrability and symmetry, *Trans. Amer. Math. Soc.* **353** (2001), 1705–1739.

[28] Vessiot E., Sur une théorie nouvelle des problèmes généraux d'intégration, *Bull. Soc. Math. France* **52** (1924), 336–395.