Invariant Source Forms, Conservation Laws, and the Inverse Problem of the Calculus of Variations

Vor der Fakultät für Mathematik und Naturwissenschaften der Carl von Ossietzky Universität Oldenburg zur Erlangung des Grades und Titels eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

angenommene Dissertation

von Herrn Markus Dafinger
geboren am 5. Januar 1983 in Erding

Gutachter: Prof. Dr. Daniel Grieser
Zweitgutachter: Prof. Dr. Ian M. Anderson

Tag der Disputation: 12. Juni 2018

Abstract

We prove that under certain assumptions a partial differential equation (PDE) is necessarily variational, i.e. that it can be written as the Euler-Lagrange equation for some Lagrangian. Instead of investigating the PDE directly, we consider the so-called source form which is the natural object in this problem. We show that if a source form satisfies certain symmetries and corresponding conservation laws then it must necessarily be variational. This question was first formulated by Floris Takens in a paper in 1977 and was then generalized by Ian M. Anderson, Juha Pohjanpelto and others. The source forms and all other objects are defined on so-called jet spaces J^kE which are constructed over fiber bundles $\pi : E \to M$. Symmetries are described by projectable vector fields on E which are then prolonged to obtain vector fields on the jet space. In the case of ordinary differential equations (ODEs), a conservation law can be considered as a first integral, i.e. a quantity which is constant along solutions of the differential equation. Conservation laws describe important properties of physical differential equations, like the conservation of energy, momentum, angular-momentum, mass, charge and so on. It is well-known from Noether's first and second theorems that symmetries of a variational functional lead to conservation laws. In some sense, we want to reverse this statement and prove that symmetries and corresponding conservation laws lead to a variational functional. The main motivation comes from physics, where the fundamental theories, like gravity and the standard model, are formulated as variational differential equations. We want to get a deeper understanding of how this formulation possibly follows from required symmetries, like Lorentz invariance and, conservation laws, like mass-conservation.

Abstract

In dieser Arbeit wird bewiesen, dass eine partielle Differentialgleichung (PDG) unter gewissen Voraussetzungen notwendigerweise variationell sein muss, d.h., dass sie sich als Euler-Lagrange Gleichung schreiben lässt. Anstatt die PDG direkt zu untersuchen, wird die sogenannte source form betrachtet, welche das natürliche Objekt in dieser Problemstellung ist. Wir zeigen, dass wenn eine source form gewisse Symmetrien und dazugehörige Erhaltungsgleichungen erfüllt, dann muss sie notwendigerweise variationell sein. Diese Fragestellung wurde zuerst von Floris Takens in einer Arbeit im Jahre 1977 formuliert und später von Ian M. Anderson, Juha Pohjanpelto und anderen verallgemeinert. Die source form und alle weiteren Objekte werden auf sogenannten Jet Räumen J^kE definiert, welche über einem Faserbündel $\pi : E \to M$ konstruiert werden. Symmetrien werden durch projezierbare Vektorfelder auf E beschrieben, welche anschließend prolongiert werden, um Vektorfelder auf dem Jet Raum zu erhalten. Im Falle von gewöhnlichen Differentialgleichungen (GDG) kann eine Erhaltungsgleichung als erstes Integral betrachtet werden, d.h. eine Größe welche entlang der Lösungen der Differentialgleichung konstant ist. Erhaltungsgleichungen beschreiben wichtige Eigenschaften von physikalischen Differentialgleichungen, wie z.B. Energie-, Impuls-, Drehimpuls-, Massen-, Ladungserhaltung usw. Wie aus dem ersten und zweiten Noetherschem Theorem bekannt ist, führen Symmetrien des variationellen Funktionals auf Erhaltungsgleichungen. In einem gewissen Sinne wollen wir diese Aussage umdrehen und beweisen, dass Symmetrien und Erhaltungsgleichungen zu einem variatonellen Funktional führen. Die Hauptmotivation kommt dabei aus der Physik, in der die fundamentalen Theorien, wie Gravitation und das Standardmodell, als variationelle Differentialgleichungen formuliert werden. Wir wollen ein tieferes Verständnis dafür erlangen, warum diese Formulierungen möglicherweise aus notwendigen Symmetrien, wie der Lorentz Invarianz, und Erhaltungsgleichungen, wie der Massenerhaltung, folgen.

Contents

1. Introduction

The main results of this dissertation are the proofs two theorems. The general structure of both theorems is as follows (the exact formulation can be found in Theorem 1.0.2 and Theorem 1.0.3 below):

Theorem 1.0.1. *Let $f = f(x,u(x),u_x(x),u_{xx}(x)) = 0$ be a differential equation for the unknown function $u = u(x)$. If the differential equation satisfies certain symmetries and conservation laws, then f must be variational.*

Variational means that f can be written as $f = \frac{\partial L}{\partial u} - \frac{d}{dx}\frac{\partial L}{\partial u_x}$ for some function $L = L(x,u(x),u_x(x))$ which is called *Lagrange function* (we also call it Lagrangian). In other words, variational means that the differential equation can be written as Euler-Lagrange equation. Note that such theorems can also be formulated for higher order differential equations $f(x,u,u_x,u_{xx},u_{xxx},...)$ and also what it means that f is variational (we will explain it later). A symmetry could be, for example, $f = f(u(x),u_x(x),...)$, where f is translation invariant with respect to the x-coordinate. A conservation law is, e.g. $\frac{1}{2}(u_x^2 + u^2) = c$, where $c \in \mathbb{R}$ is a constant. In this case we think of some kind of energy conservation. Note that Theorem 1.0.1, as it is formulated above, is not true in general, but it provides the idea of what we want to prove.

For example, let us consider the differential equation $f = u_{xx} + u = 0$ which describes oscillations. This differential equation is translation invariant with respect to x. This symmetry can also be described by the Lie derivative of some vector field. In this case, the Lie derivative is $\frac{\partial}{\partial x}$ and we get

$$\frac{\partial}{\partial x} f = \frac{\partial}{\partial x}(u_{xx} + u) = 0,$$

where $(u,u_x,...)$ now are considered as variables and not as functions depending on x. Furthermore, f satisfies a conservation law of the form

$$u_x f = u_x(u_{xx} + u) = \frac{1}{2}\frac{d}{dx}(u_x^2 + u^2) = 0 \quad \Leftrightarrow \quad \frac{1}{2}(u_x^2 + u^2) = \text{constant}, \qquad (1.1)$$

where the equations in (1.1) are satisfied for every solution $u = u(x)$ of the differential equation $f = 0$. The equation is variational and a Lagrangian is $L = \frac{1}{2}(u^2 - u_x^2)$, where one can easily check the identity $f = \frac{\partial L}{\partial u} - \frac{d}{dx}\frac{\partial L}{\partial u_x}$.

We want to mention here that symmetries and conservation laws are somehow connected and it is better to speak about symmetries and *corresponding* conservation laws. For the above example this means that it is well-known from Noether's first theorem that invariance in time (what we called x here) leads to energy conservation. But in order to apply Noether's theorem, we already need a variational formulation from the beginning. In what we want to prove in this dissertation, we do not assume a variational formulation from the beginning, we rather want to prove such a formulation. This kind of problem can be considered as inverting the statement in Noether's theorem in some sense. In the next section, we will give some more motivation why this is an interesting problem.

Here, we will only explain the result of this dissertation a little bit more precisely. As we saw in the above example, when talking about symmetries and conservation laws, sometimes we consider $(u,u_x,...)$ as variables or coordinates and sometimes we consider them as functions $(u(x),u_x(x),...)$ depending on x. The first part of Chapter 2 will give a precise definition of these two pictures. In this section, we simply say which formulation we mean if it is essential to distinguish both. As one usually defines a differential equation $f = f(x,u,u_x)$ of order one, $f = f(x,u,u_x,u_{xx})$ of order two and so on, we define the order of the coordinates $(x,u,u_x,...)$ in the same way and we write $(x,u,u_x,...,u_{(k)})$ for k-th order coordinates.

Let g_s, $s \in \mathbb{R}$, be a 1-parameter group acting on the coordinates (x,u). For example, $g_s(x,u) = (x+s,u+s)$ which describes translations with respect to the parameter $s \in \mathbb{R}$ in x and u direction. These kinds of transformations induce transformations on the coordinates $(x,u,u_x,...)$ in a certain way and the induced transformation is written as $\mathrm{pr}\, g_s$, where pr stands for *prolongation*. A differential equation $f = 0$, or better a differential expression f, is invariant with respect to some prolonged group action g_s if $f(\mathrm{pr}\, g_s(x,u,...)) = f(x,u,...)$ for all $s \in \mathbb{R}$. It is well-known that this is equivalent to

$$\frac{d}{ds} f(\mathrm{pr}\, g_s(x,u,u_x,...))|_{s=0} = 0$$

and this equation can be read as applying a vector field $V := \frac{d}{ds}\mathrm{pr}\, g_s|_{s=0}$ to f (for translations in x direction we get $g_s(x,u) = (x+s,u)$ with corresponding vector field $\frac{d}{ds} g_s|_{s=0} = \frac{\partial}{\partial x}$).

Actually, in this dissertation we want to consider relatively arbitrary systems of partial differential equations (PDEs) and therefore we need to write (x^i,u^α), $i = 1,2,...n$ and $\alpha = 1,2,...,m$ instead of (x,u). The coordinates (x^i,u^α) belong to the space E which is a fiber bundle $\pi : E \to M$ and we have certain projections, like $\pi(x^i,u^\alpha) = (x^i)$, where (x^i) are coordinates on the base manifold M. The coordinates $(x^i,u^\alpha,\frac{\partial u^\alpha}{\partial x^i},...)$ belong to a space J^kE, called the k-th order jet space of E, which is also a fiber bundle with different kinds of projections. Furthermore, we will always write $\dim M = n$, and m is the dimension of the fibers of E, i.e. $\dim E = n + m$.

Let us also briefly introduce the so-called *source form*, defined as

$$\Delta := f_\alpha du^\alpha \wedge dx^1 \wedge ... \wedge dx^n,$$

which is a differential $(n+1)$-form on J^kE. The source form is used to describe the differential equation $f_\alpha = 0$. The brief explanation why we use Δ instead of f_α is that in the calculus of variations the functional

$$I(u) := \int L(x,u(x),u_x(x)...)dx$$

plays a fundamental role and there we can see that the so-called Lagrange form $\lambda := Ldx$ is fundamental, as well. When computing the first variation δI, we then get the source form Δ in a very natural way. Roughly speaking, if we considered the Lagrangian L instead of the Lagrange form $\lambda = Ldx$, then we would not get an object which transforms correctly under local coordinate transformations. The same would happen if we considered f_α instead of Δ. This will be explained later in more detail and is a technical detail at this point.

There is one last definition we would like to present before we give a precise formulation of the main results in this dissertation. The space $\mathcal{V}$ is the space of vector fields V on the fiber bundle E which describes symmetries of Δ. Moreover, these vector fields have the additional property of being so-called *projectable vector fields* (in short: $\pi_* V$ exists and is a vector field on M, this will also be explained later). The short notation will be

$$\mathcal{V} := \{V \in \mathfrak{X}(E) : V \text{ is projectable with respect to } \pi \text{ and } \mathcal{L}_{\mathrm{pr}V}\Delta = 0\},$$

where $\mathcal{L}_{\mathrm{pr}V}$ is the Lie derivative with respect to the prolonged vector field V. For vector fields on E we usually write $V = V^i\partial_{x^i} + V^\alpha\partial_{u^\alpha}$ or $V = V^x\partial_x + V^u\partial_u$ in local coordinates. The set of symmetry vector fields forms a vector space over $\mathbb{R}$. In general, it will not be sufficient to consider only one symmetry vector field (modulo vector space structure) and we will need the additional condition that

$$\mathrm{span}_{\mathbb{R}}\{V_p : V \in \mathcal{V}\} = T_pE \quad \text{for all } p \in E. \tag{1.2}$$

The main results of this dissertation are:

Theorem 1.0.2. *Let $n,m \in \mathbb{N}$ be arbitrary and let $\pi : E \to M$ be a fiber bundle of fiber dimension m and base dimension n. Furthermore, let $\Delta = f_\alpha du^\alpha \wedge dx^1 \wedge ... \wedge dx^n$ be a second order source form defined on J^2E. Assume:*
i) The set $\mathcal{V}$ of symmetries of Δ satisfies (1.2).
ii) Each $V \in \mathcal{V}$ generates a conservation law of the from $Q_V^\alpha f_\alpha = D_i C_V^i$, where $Q_V^\alpha = V^\alpha - u_i^\alpha V^i$ are the characteristics.
Then Δ must be locally variational.

Theorem 1.0.3. *Let $\pi : E \to M$ be a fiber bundle of fiber dimension one and base dimension one. Furthermore, let $\Delta = f du \wedge dx$ be a 4-th order source form defined on J^4E. Assume:*
i) The set $\mathcal{V}$ of symmetries of Δ satisfies (1.2).
ii) Each $V \in \mathcal{V}$ generates a conservation law of the from $Q_V f = D_x C_V$, where $Q_V = V^u - u_x V^x$ are the characteristics.
Then Δ must be locally variational.

We conjecture that Theorem 1.0.3 is also true with the same assumptions, but arbitrary base dimension n.

The functions C_V^i are different kinds of conserved quantities and, in the case of ordinary differential equations (ODEs), a first integral for the differential equation $f_\alpha = 0$.

Note that Theorem 1.0.2 is also true when the source form Δ is only defined on open subsets $\mathcal{R}^2 \subset J^2E$, whereas Theorem 1.0.3 is no longer true in general for open subsets $\mathcal{R}^4 \subset J^4E$. For example, $f = \frac{1}{u_x}$ is not defined on J^4E, but it can be defined on some open subset $\mathcal{R}^4 \subset J^4E$. In which cases Theorem 1.0.3 is also true when allowing such topological obstructions depends on the subset $\mathcal{R}^4 \subset J^4E$ and needs further investigation. In other words, Theorem 1.0.3 is only true for non-singular source forms and it is no longer true for singular source forms (singular in the sense that there is no smooth continuation of Δ from $\mathcal{R}^4$ to J^4E). We will discuss this in detail in Subsection 3.8.1.

Similar theorems have already been proven by others and we only want to mention two of them which are strongly connected to Theorem 1.0.2 and 1.0.3. The first theorem is

Theorem 1.0.4. *Let (x^i), $i = 1,2,...,n$ be coordinates on $\mathbb{R}^n$. A classical field theory is described by a 1-form $A = A_i dx^i$, where (A_i) is the field. Assume the source form $\Delta = f^i dA_i \wedge dx^1 \wedge ... \wedge dx^n$ is translation and gauge invariant, and has corresponding conservation laws of the form*

$$\begin{aligned} A_{j,i} f^j &= D_k C_i^k, \quad \textit{for some functions } C_i^k \\ D_i f^i &= 0. \end{aligned}$$

i) Then Δ is locally variational if $n = 2$ and f^i is of third order or $n \geq 3$ and f^i is of second order.
ii) Then Δ is locally variational if the functions f^i are polynomials of degree at most n in the field variables A_i and their derivatives.

This theorem can be found in (AP96, p.370). There is a similar theorem which holds also for non-Abelian gauge transformations, see (MPV08, p.4). For the second theorem we need the condition

$$\operatorname{span}_{\mathbb{R}}\{\pi_* V_p : V \in \mathcal{V}\} = T_{\pi(p)}M \quad \text{for all } p \in E, \tag{1.3}$$

which is similar to (1.2), but is a actually a weaker assumption. The second theorem is

Theorem 1.0.5. *Let $n \in \mathbb{N}$ be arbitrary and $m = 1$. If a second order, non-singular source form $\Delta = f du \wedge dx^1 \wedge ... \wedge dx^n$ is invariant under the prolonged vector fields $V \in \mathcal{V}$, such that* (1.3) *is satisfied and Δ has corresponding conservation laws of the form $Q_V^u f = D_i C_V^i$, where $Q_V^u = V^u - u_i V^i$ are the characteristics, then Δ must be locally variational.*

This theorem can be found in (AP94, p.213). Note that for $m > 1$, the stronger assumption (1.2) is necessary otherwise the theorem is no longer true (see 4th counter example in Section 4.1.). Notice to the reader: Sometimes, when citing such theorems or definition we will slightly change the notation compared to the original papers or books, to get a consistent notation in this thesis. In the papers (AP12, MPV08, Poh95, AP96, AP95, AP94) and the original work (Tak77), different versions and proofs of Takens' problem can be found. As far as we know, this should be a relatively complete list of such theorems. In the next section, we will discuss in more detail what kinds of theorems have already been proven and what kinds of new results we derive in this thesis.

I want to thank the Professors:

D. Grieser for advising me during my Ph.D. studies and supporting me in my research. Especially, I want to thank him for giving me a lot of freedom in my research and supporting my interests.

I.M. Anderson for a very nice stay at the Utah State University and for helping me to solve a version of Takens' problem.

P.J. Olver for giving me the opportunity to work at the University of Minnesota and helping me in my research.

D. Krupka, especially for the beginning phase of my Ph.D. studies and for long discussions about basics on jet bundles and concepts in the calculus of variations.

I also want to thank my colleagues I. Shestakov, A. Beyer and M. Behr for a very nice collaboration at the University of Oldenburg.

1.1. Literature Review and New Results

There is quite a long history of questions similar to Takens' problem. For example, in the paper (LL74, p.1694) from 1974 we can find the following conjecture: *"For* *metric theories of gravity, the existence of a conserved energy-momentum ... is equivalent to the existence of a Lagrangian formulation.*". In 1977, Takens (Tak77) was able to answer a similar question for second order source forms. In 2012, Ian M. Anderson and Juha Pohjanpelto (AP12) answered such a question for third order source forms. We could even go back to the paper (Noe18) of Emmy Noether in 1918 and set this as the starting point of these kinds of questions. It is no surprise that such problems came up at that time and later, since Albert Einstein discovered the special and general relativity at about the same time and the investigation of symmetries became a very important factor. Einstein had different versions of the nowadays so-called Einstein's field equations. Some of them did not satisfy necessary local energy-momentum conservation laws and they were also not variational (VK15, p.1). Finally, Einstein (and Hilbert) found an equation which satisfies these conservation laws and it turned out to be variational. Nowadays, the variational formulation is given by the so called Einstein-Hilbert functional.

Now, we will start a more detailed discussion of the paper (Tak77) of Takens from 1977, since, there, source forms are used (probably for the first time) and they are necessary to give a precise formulation of the problem. Takens formulates three different theorems and we will only describe the main features in informal notation:

Theorem 5,2 (Tak77, p.599): Let $\pi : E \to M$ be a linear fiber bundle, i.e. which has a linear structure on the fibers. Furthermore, let Δ be a linear source form, i.e. which depends linearly on $(u^\alpha, u^\alpha_i, ...)$ (PDE of arbitrary order). If we have at least one symmetry vector field $V \in \mathcal{V}$ and corresponding conservation law, such that $\pi_* V \neq 0$, then Δ is variational.

This theorem is special in this regard as the order is arbitrary. The proof can be found in Appendix A., but we have to introduce more notation to be able to understand it later. Basically all other theorems which we will present need to make assumptions about the order (or polynomial degree). The proof is actually very simple and probably the simplest proof compared to the proofs of all other theorems which give an answer to Takens' question. The restriction to linear source forms makes the proof relatively easy. More generally, restricting to polynomial source forms of degree two or three has a similar effect and simplifies the proof. This will be clear later, when we discuss polynomial structures in certain equations, especially in the so-called *equations of conservation laws and symmetries* (ECS). The ECS is the equation $\mathcal{L}_{\mathrm{pr}V}\Delta = 0$, where the corresponding conservation law condition is eliminated (see Section 2.9. for further details).

Theorem 5,3 (Tak77, p.600): Let $\pi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ be a linear fiber bundle of fiber dimension one and base dimension n, i.e. $E = \mathbb{R}^n \times \mathbb{R}$ and $M = \mathbb{R}^n$. Furthermore, let $\Delta = f du \wedge dx^1 \wedge ... \wedge dx^n$ be a second order source form defined on J^2E. Assume:
i) Δ is invariant under the symmetries $\mathcal{V}$.
ii) For each $\partial_{x^i} \in \mathfrak{X}(M)$, $i = 1,2,...,n$, there exists a $V \in \mathcal{V}$ such that $\pi_* V = \partial_{x^i}$.

iii) Each $V \in \mathcal{V}$ generates a conservation law.
Then Δ is variational.

Note that we slightly changed *Theorem 5,3* compared to the original work. The condition in *ii*) is similar to (1.3). However, it is different in this regard that the vector fields $\{\partial_{x^i}\}$ are constant on M and this set is an Abelian Lie algebra on M, but $\mathcal{V}$ is not necessarily an Abelian Lie algebra on E. This theorem is in some sense similar to Theorem 1.0.3, at least concerning the proof and the importance of investigating singularities.

We will not formulate *Theorem 5,4* from Takens' paper (Tak77, p.601), since it needs more notation. Roughly speaking, the theorem is about metric field theories of second order. There occur two complications for us here: First, we have to define what we mean by metric field theory. Second, the symmetries are given in a certain way and they involve so-called arbitrary functions used in Noether's second theorem. This also dictates how the corresponding conservation laws have to look like. In fact, these symmetry vector fields generate an infinite dimensional Lie algebra which is not the case in the context of Theorem 1.0.2 and 1.0.3. Infinite dimensional Lie algebras usually lead to stronger conditions than (1.2) (for example, $\mathrm{span}\{(\mathrm{pr}^1 V)_p : V \in \mathcal{V}\} = T_p J^1 E$ for all $p \in J^1 E$).

We could say that many theorems, proven by others, and proven in this thesis, are generalizations of these three theorems in one way or the other. However, significantly more notation has to be used and sometimes new techniques have to be introduced, as well. We summarize these theorems in the following table:

article	year	dimension	order	pol. deg.	sym. $V \in \mathcal{V}$	cons. laws
(AP94, p.213)	1994	$m = 1$, any n	2	-	(1.3) is satisfied	Noet. 1st
(AP95, p.629)	1995	any n,m	any	$\leq r$	$\dim\{\pi_* V_p\} = r$	Noet. 1st
(Poh95, p.352)	1995	any $n = m$	1	-	translations ∂_{x^i}	Noet. 1st
(AP96, p.370) (classical field th.)	1996	$n = m = 2$	3	-	(1.2) is satisfied gauge $\phi_{x^i}\partial_{u_i}$ translations ∂_{x^i}	$D_i f^i = 0$ Noet. 1st
(AP96, p.370) (classical field th.)	1996	any $n = m$	2	-	(1.2) is satisfied gauge $\phi_{x^i}\partial_{u_i}$ translations ∂_{x^i}	$D_i f^i = 0$ Noet. 1st
(AP96, p.370) (classical field th.)	1996	any $n = m$	any	$\leq n$	(1.2) is satisfied gauge $\phi_{x^i}\partial_{u_i}$ translations ∂_{x^i}	$D_i f^i = 0$ Noet. 1st
(MPV08, p.4) (gauge field th.)	2008	any $n = m$	2	-	(1.2) is satisfied, $(\varphi^\alpha_{x^i} + c^\alpha_{\beta\gamma} u^\beta_i \varphi^\gamma)\partial_{u^\alpha_i}$ translations ∂_{x^i}	$\nabla_i f^i_\alpha = 0$ Noet. 1st
(MPV08, p.4) (gauge field th.)	2008	any $n = m$	any	$\leq n$	(1.2) is satisfied $(\varphi^\alpha_{x^i} + c^\alpha_{\beta\gamma} u^\beta_i \varphi^\gamma)\partial_{u^\alpha_i}$ translations ∂_{x^i}	$\nabla_i f^i_\alpha = 0$ Noet. 1st
(AP12, p.4) (metrics)	2012	any n, $m = \frac{n(n+1)}{2}$	3	-	$(*)$ $\xi^i \partial_{x^i} - 2\xi^i_{,(k} g_{l)i} \partial_{g_{kl}}$	$\nabla_i f^{ji} = 0$
this thesis	2018	any n,m	2	-	(1.2) is satisfied	Noet. 1st
this thesis	2018	n,$m = 1$	4	-	(1.2) is satisfied	Noet. 1st

The short notation *Noet. 1st* stands for *Noether's 1st theorem* and means that the corresponding conservation laws are given in the form

$$Q_V f = D_i C_V^i. \tag{1.4}$$

The other conservation laws, i.e. $D_i f^i = 0$, $\nabla_i f_\alpha^i = 0$ and $\nabla_i f^{ji} = 0$, correspond to Noether's second theorem. Note that we sometimes write f^i,u^i instead of f_α,u^α and, u_i does not always have the meaning of partial derivative $\frac{\partial u}{\partial x^i}$, rather describes the field u^α, but with lower indices i (this is a technical detail at this point and we refer to the original papers and the notation there). The functions ϕ, φ^α and ξ^i in the symmetry vector fields are arbitrary functions and therefore we always have an infinite dimensional Lie algebra in these cases. This can also be seen by the corresponding conservation laws, as we mentioned, since they correspond to Noether's second theorem. The star $(*)$ in the table above means that (1.2) is satisfied for $n = 2$ and probably also for $n > 2$, but we did not check it in detail (maybe the symmetries even satisfy $\mathrm{span}\{(\mathrm{pr}^1 V)_p : V \in \mathcal{V}\} = T_p J^1 E$ for all $p \in J^1 E$). Note that the conservation laws (or differential identities) corresponding to Noether's second theorem can always be brought into the form (1.4) of Noether's first theorem. This is mainly important when investigating applications and the physical meaning of conservation laws and how to reformulate these equations.

Notice to the reader: We will always try to explain, as well as possible, all theorems, definitions, calculations and so on. However, some aspects of the following discussion are probably only understandable for the experts with advanced knowledge of Takens' problem. The reader could skip this short part and continue with Section 1.2 and come back to this later.

In the following, we compare the new results in this dissertation with the theorems in the above table:

New results in Theorem 1.0.2:

Dimension of Lie algebra: The papers from 1996 to 2012 consider infinite dimensional Lie algebras, where ϕ, φ^α and ξ^i are arbitrary functions. Roughly speaking, this means that the amount of symmetries leads to strong restrictions. In more detail, and concerning the proofs: The zeroth order Helmholtz expression (in our notation $H_{\alpha\beta}$) in (AP96, p.378) immediately vanishes, because ϕ_{x^i} can be chosen arbitrarily and independently of $\phi_{x^i x^j}$ and $\phi_{x^i x^j x^k}$. Nearly similarly in (MPV08, p.12), the expression with the $\varphi^\beta_{x^j}$-coefficient immediately vanishes, since it can be chosen independently of φ^β and $\varphi^\beta_{x^j x^k}$. Our proof of Theorem 1.0.2 in Section 3.6 only requires a finite dimensional set of vector fields (Lie algebra) and the zeroth order Helmholtz expression $H_{\alpha\beta}$ vanishes at the very end of the proof in Step 7 (see Section 3.6). But more importantly, during the proof, $H_{\alpha\beta}$ will be needed in Step 4 and Step 5. The so-called *Helmholtz dependencies* will play a fundamental role there and the ECS is also needed to deduce properties of $H_{\alpha\beta}$. The proofs of Theorem 1.0.2, Theorem 1.0.3 and similar ones can be divided into different steps. The vanishing of the zeroth order Helmholtz expression in (AP96) at the very beginning does not need so many different steps in the proof in some sense. It also seems that the zeroth order Helmholtz expression does not occur once again in the proof in (MPV08) and equation (32) in (MPV08, p.12) is sufficient to complete the proof. The Helmholtz dependencies are also not used in (MPV08), but in (AP96).

New algorithm: In general, the three theorems proven by Takens in 1977 have similarities to all other theorems and provide useful techniques how to solve the problem. For (AP96, MPV08) and Theorem 1.0.2, these techniques are basically to discuss successively the order of jet coordinates in the equations forced by the symmetry and conservation law assumptions (we call these equations the ECS). Usually, the discussion starts with the highest order jet coordinates.[1] However, the problem is the complexity and that there is no straightforward way telling us at which point in the proof we have to use which of these equations and conditions. To find the way (or the algorithm) in which these equations and conditions can be used, are the main parts of these proofs. Beside the technical problems of how to handle a tremendous amount of indices. Therefore, that $H_{\alpha\beta}$ does not immediately vanish in our proofs and that it will be needed in different steps also means that we will find a new way in some sense (a new algorithm) how to prove a version of Takens' problem and this way seems to be different from all the previous versions proven by others.

New induction method: Step 3 in Section 3.6 can probably be proven with the so-called d-fold operator, used in (AP96, p.379) and (MPV08, p.12). We prove Step 3 with a kind of induction which is an alternative method and which has not been used in this form before. The proof of Step 6 in Section 3.6 is also the same kind of induction and these inductions work fairly well in different situations. This new kind of induction has the advantage that it provides information in every induction-step and, more importantly, every induction-step is relatively easy to understand. On the other hand, applying the d-fold operator does not split the problem into smaller pieces, it rather solves the problem in one relatively complicated step. The application of the d-fold operator is a transformation of an equation into a new equation and the new equation will be discussed instead. If the original equation is already relatively complicated, then the transformed equation will be again relatively complicated and it is, in some sense, hard to understand what exactly happens during the transformation. A lot of hard calculations are hidden behind notation. Sometimes, hiding complicated calculations behind notation is reasonable, but sometimes it also makes sense to do a calculation where we can directly follow every reformulation. The operator $\partial^b_{\beta,k,X}$ in (MPV08, p.11) *replaces* k-th order coordinates by $X_{i_1}...X_{i_k}$, whereas in the induction we *eliminate* k-th order coordinates by applying

[1] For (AP94) and Theorem 1.0.3 we also have to solve differential equations and the situation is slightly different there.

partial derivatives, like $\partial_{u^\gamma_{JJ}}$. Eliminating almost all of the expressions reduces a big problem into a small problem, whereas the replacement method just transforms a problem into a new problem. However, we will also use this kind of technique in Section 3.6 because in some situations it is a very good method and sometimes there are simply no other techniques available. The proofs with the d-fold operator in (AP96, MPV08) need *density arguments*, whereas the induction does not need such arguments. Therefore, and for the other reasons mentioned previously, the induction method is a different and new method. Note that, although the proof of Step 3 in Subsection 3.6.1 looks much longer than similar proofs in (AP96, MPV08), it can actually be brought in a much shorter version. It could probably be written on one and a half pages, but we show the longer version in order to explain every detail of the proof.

Many theorems make assumptions stronger than (1.2)**:** For many theorems, condition (1.2) is satisfied by the assumptions. In our Theorem 1.0.2 in this thesis, we only need this condition and no specific symmetries, like translation- or gauge symmetries or the symmetries for metrics. We could even say that if the theorems in (AP96, MPV08, AP12) can be brought in such a form, then they are just corollaries of the more general version we will prove here. For (AP12), this is of course only true when restricting to second order.[2] However, it is fair to say the the more complicated fields, especially the metric field, probably need their own formulations and some work has to be done there. It could be future work to check in which way (AP96, MPV08, AP12) can be considered as corollaries or not.

Arbitrary n,m**:** We can prove Theorem 1.0.2 for any n,m. The condition $n = m$ in (AP96, MPV08), or $m = \frac{n(n+1)}{2}$ in (AP12), since the metric g_{ij} is symmetric in i,j, seems not to be a crucial factor in these proofs. Especially, the discussion of the formulas (3.8), (3.9), (3.10) and (3.11) in (AP96, p.378) makes explicitly use of the property $n = m$.

Polynomial structure: To derive polynomial structure in some of the jet coordinates in f_α is crucial in almost all proofs of Takens' problem. In (AP96, MPV08), polynomial structure in second order jet coordinates is derived with the help of the equations $D_i f^i = 0$ and $\nabla_i f^i_\alpha = 0$. We prove it with the help of Step 2 in Section 3.6. instead. Again, this means that we find an algorithm how to prove the theorem and this method differs from others. Although the statement in the Local Simplification Lemma 3.4.1 is almost trivial, the use of this lemma in Step 1 in Section 3.6 should not be underestimated and there is no similar transformation in (AP96, MPV08). This actually allows us to derive the Hyperjacobian structure in Step 2.

Polynomial degree and conservation laws: The proof of Step 6 in Subsection 3.6.2 is a new result in the following sense: It is known (or easy to see) that n equations of the form $u^\beta_{ik} H^k_{\alpha\beta} = 0$, for $i = 1,2,...,n$, have non-trivial solu-

[2] For third order source forms, the assumption $\mathrm{span}\{(\mathrm{pr}^1 V)_p : V \in \mathcal{V}\} = T_p J^1 E$ for all $p \in J^1 E$ is probably needed to prove Takens' problem. I want to thank Ian M. Anderson for a fruitful discussion concerning this.

tions $H^k_{\alpha\beta}$ of polynomial degree n in the second order coordinates u^β_{ij}. There do not exist non-trivial solutions of degree $\leq n-1$. Therefore, the number n is the critical number here. We will prove that $(O^\beta_{ik}(1) - u^\beta_{ik})H^k_{\alpha\beta} = 0$ for $i = 1,2,...,n$ only allows the trivial solution under the assumption $H^i_{\alpha\beta,u^\gamma_{ii}} = 0$, but it is still assumed that $H^i_{\alpha\beta}$ is a polynomial of degree $\leq n$. We show this with a kind of induction which demonstrates the use of such methods once again. The analogous equation $(A^\gamma_{j,kp} + c^\gamma_{\beta\zeta}A^\zeta_j A^\beta_{k,p})H^{ij,k}_{\alpha\gamma} = 0$ in (MPV08, p.13) assumes polynomial degree $\leq n-1$ in second order coordinates of $H^{ij,k}_{\alpha\gamma}$. But the induction is not the only difference in these proofs. In order to use the condition $H^i_{\alpha\beta,u^\gamma_{ii}} = 0$ in Step 5 in Section 3.6, we need to find a way (or algorithm) how to use the symmetry and conservation law assumptions, as we mentioned above. The condition $\nabla_i f^i_\alpha = 0$ in (MPV08, p.10) leads to polynomial degree at most $n-1$ in the source form and then to polynomial degree at most $n-1$ in $H^{ij,k}_{\alpha\gamma}$. In contrast, we allow source forms of polynomial degree n and we do not require a conservation law (or differential identity) of the special form $\nabla_i f^i_\alpha = 0$. All the conservation laws $D_i f^i = 0$, $\nabla_i f^i_\alpha = 0$ and $\nabla_i f^{ji} = 0$ in (AP96, MPV08, AP12) are strong restrictions from the very beginning (for $n = 1$, $D_i f^i = 0$ even means that $f = f^i$ must be constant). Note that we could also derive that $H^i_{\alpha\beta}$ is polynomial degree at most $n-1$ in Step 5 in Section 3.6, but the stronger assumption is not needed to complete the proof and the source form is still allowed to have degree n (for example, we allow the variational Monge-Ampere expression $f = u_{xx}u_{yy} - u^2_{xy}$ of degree $n = 2$). In any case, we derive this condition differently and we use so-called Helmholtz dependencies there.

<u>New results in Theorem 1.0.3:</u> As far as we know, and according to the above table, a theorem for 4-th order source forms has not been proven (except if we make the additional assumption of source forms which are polynomial in the fields u^α and derivatives of u^α, see (AP95) and Theorem 5,2 in (Tak77)). We find a very interesting structure in the ECS, also for higher order. In more detail, we find integrating factors such that the ECS can be written as a total derivative and this leads to a first integral. After eliminating some of the unknowns we repeat this and we find the general solution of the ECS under the assumption of (1.2). We show that the characteristics Q^α are integrating factors for the ECS for arbitrary n,m and arbitrary order and we indicate how this can possibly be used more generally. The proof of Theorem 1.0.3 shows how the theorem for 4-th order, $m = 1$, and arbitrary n can probably be proven. This proof differs from a pure order discussion of jet coordinates and the investigation of singularities becomes important.

Above we mentioned some of the most important new results and new methods. Note that it is almost impossible to discuss all proofs and differences between (AP12, MPV08, Poh95, AP96, AP95, AP94, Tak77) and our work here. This would be a topic of its own. In the following, we only mention some notational innovations of this thesis.

We use the notation of O which describes lower order terms, i.e. non-important terms (note that O does not mean the Landau symbol here). When the symmetries are not given explicitly, as in (AP96, MPV08, AP12), then it is reasonable to work with this notation. This would also simplify some of the explicit expressions in (MPV08) and it clarifies the structure of such proofs in general, in our opinion. We distinguish explicitly between Helmholtz conditions and Helmholtz dependencies, which also clarifies what the actual problem is and how we can try to solve it. We also give the equation $\mathcal{L}_{\mathrm{pr}V}\Delta = 0$ a name, after eliminating the conservation law assumptions, we call it the ECS. Finally, we formulate an algorithm of the proof and work out the main steps. This can help to find differences in the proofs of Takens' problem and it explains the main ideas in these proofs.

1.2. Motivation and Physical Background

Our motivation for this thesis mainly comes from the following considerations:

- Since the fundamental laws of physics, described by differential equations, should hold everywhere in the universe, the differential equations should be translation invariant, rotation invariant and so on. In general, they do not (explicitly) depend on the (t,x,y,z)-coordinates. Furthermore, when we think of special relativity, then we assume Lorentz or Poincaré invariance.

- Since in physics (we observe in experiments that) we have energy-, angular-momentum-, momentum-, charge-, and so on conservation, the solutions of these differential equations must satisfy conservation laws, or rather, they must allow such conservation laws.

Therefore, in physics we consider only a subset of the set of all differential equations, satisfying the above and even more conditions. Now, the question is: Can we describe this subset more precisely? For example, are all such differential equations variational? That is, does a Lagrangian L exist such that the differential equation can be written as the Euler-Lagrange equation for L? When $n,m = 1$ and $L = L(x,u,u_x)$ this means

$$\text{Euler-Lagrange equation} = \frac{\partial L}{\partial u} - \frac{d}{dx}\frac{\partial L}{\partial u_x}, \tag{1.5}$$

as we already mentioned above. The Euler-Lagrange equations are the differential equations which must be satisfied by extremal values of the variational functional

$$I(u) = \int L(x,u(x),u_x(x))dx,$$

The main result of this thesis is that the answer to this question is yes under certain conditions. In the paper by Floris Takens in 1977 (Tak77), we can find this question (maybe for the first time in this form). Therefore, we call it **Takens' problem**. Other people also investigated this problem, for example, Ian M. Anderson, Juha Pohjanpelto, Gianni Manno and Raffaele Vitolo, to name a few.

From a quantum mechanics point of view it could also be important to understand if a variational formulation exists. With the help of Feynman's path integral formalism, variational equations can be quantized. In fact, all equations in the standard model and Einstein's field equations are variational. The variational functional for Einstein's field equations is the famous Einstein-Hilbert functional (Wan94)

$$I_{EH}(g) = \frac{1}{16\pi G}\int (R + 16\pi G L_{\text{matter}})\sqrt{-g}d^4x$$

and, as an example regarding the standard model, the functional for quantum electrodynamics (QED) is given by (PS95, p.303)

$$I_{\text{QED}}(\psi,A) = \int [\bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - e\bar{\psi}\gamma^\mu\psi A_\mu]d^4x.$$

Not only in physics, also in mathematics, the corresponding functionals of differential equations are of great interest and can help to solve problems. We only want to mention the famous functional

$$\mathcal{F}(g,f) = \int (R + |\nabla f|^2)e^{-f}dV$$

used by G. Perelman in (Per02, p.5). This functional can be considered as a modification of the Einstein-Hilbert functional. However, in his paper (Per02, p.6) G. Perelman says that the functional $\mathcal{F}$ can be found in the literature on string theory, for example, in (DEF$^+$99, p.911). These functionals can also be used for proving existence of asymptotic stationary solutions and stability analysis of stationary solutions. In this case, they are called Lyapunov functionals and, usually, the corresponding differential equation is described by gradient flow.

Even if a differential equation is not variational, we can still ask if we can transform it equivalently into a variational one. For example, one of these transformations is known as the problem of finding a variational multiplier and was investigated by different people (Dou41, AT92). In this sense, we could extend the question of Takens and ask if a differential equation, which satisfies certain symmetries and conservation laws, always allows a variational multiplier and similar questions are thinkable, as well. For example, Maxwell's equations are not variational when using the electric and magnetic fields $\boldsymbol{E}$,$\boldsymbol{B}$. But they are variational when using the vector potential A_μ. The formulation with the vector potential A_μ also describes the relativistic properties better and it is quite surprising that specifically this version is variational. Switching from the fields $\boldsymbol{E}$,$\boldsymbol{B}$ to A_μ can be considered as a sort of transformation.[3]

1.3. Classical Calculus of Variations

Let us briefly recall the classical calculus of variations. We want to find a function $u : \Omega \to \mathbb{R}^m$, $\Omega \subset \mathbb{R}^n$ open, such that the functional

$$I(u) := \int_\Omega L(x,u(x),u_i(x))dx$$

is minimal or maximal, i.e. extremal ($u_i(x) = \frac{du(x)}{dx^i}$). Usually, the closure $\bar{\Omega}$ is a compact subset of $\mathbb{R}^n$. The function u is chosen from a set S, for example,

$$S := \{u \in C^\infty(\bar{\Omega},\mathbb{R}^m) : u|_{\partial\Omega} = g\}, \tag{1.6}$$

where g is a fixed function on the boundary. The condition $u \in C^\infty$ could be weakened and we could choose $u \in C^2$ or u in some Sobolev space. The regularity of suitable spaces is dictated by the corresponding differential equation and can be very different for different problems. Also, specific boundary conditions for u are not necessary. For example, in classical mechanics we usually have initial conditions

$$S := \{u \in C^2([a,b],\mathbb{R}^3) : u(a) = u_0, u_x(a) = v_0\}, \quad a,b \in \mathbb{R}, \quad u_0,v_0 \in \mathbb{R}^3,$$

where $u : [a,b] \to \mathbb{R}^3$ describes the position of a particle at time x with initial position u_0 and velocity v_0. The calculus of variations can be applied in all these cases. The only important assumption we have to make is that the test function φ has to have compact support in Ω (here in (a,b)), since then partial integration works without getting a boundary term and this will be crucial (this will be discussed below in more detail). Recall that a **test function** φ is a function $\varphi \in C_c^\infty(\Omega)$, where c denotes compact support in Ω, and Ω is open.

[3]Maxwell's equations are $\nabla\boldsymbol{E} = \rho$, $\nabla \times B = \boldsymbol{j} + \partial_t\boldsymbol{E}$, $\nabla\boldsymbol{B} = 0$ and $\nabla \times \boldsymbol{E} = -\partial_t\boldsymbol{B}$. Two of them can be solved immediately, namely, $\nabla\boldsymbol{B} = 0$ leads to $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ and then $\nabla \times \boldsymbol{E} = -\partial_t\boldsymbol{B}$ can be written as $\nabla \times (\boldsymbol{E} + \partial_t\boldsymbol{A}) = 0$, which leads to $\boldsymbol{E} + \partial_t\boldsymbol{A} = -\nabla\phi$ (at least locally). Now, we can plug the homogenous solutions $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ and $\boldsymbol{E} = -\nabla\phi - \partial_t\boldsymbol{A}$ into the inhomogeneous equations and solve them for the vector potential $(A_\mu) = (\phi,\boldsymbol{A})$. Using the vector potential A_μ requires only the four inhomogeneous equations and the homogeneous ones are automatically satisfied or already solved. Therefore, we have four unknowns A_μ and four (inhomogeneous) equations, which have a chance of being variational. Six unknowns $\boldsymbol{E}$,$\boldsymbol{B}$ and eight equations cannot be variational (if they are independent equations). Note that using the vector potential A_μ and the definition of field strength tensor $F^{\mu\nu}$, the so-called dual equation $\partial_\mu\tilde{F}^{\mu\nu} = 0$ is automatically satisfied and we could say that the variational formulation only describes the four inhomogeneous equations.

The classical calculus of variations does also not deliver a global minimum or maximum, just a necessary condition for an extremal value, which can be derived when using compactly supported test functions and doing perturbation up to first order. Let us consider $u : \mathbb{R}^n \to \mathbb{R}$. A perturbation of u, which can be written as $u + \epsilon\varphi$, where φ is a test function and ϵ is a small parameter (such that $u + \epsilon\varphi$ is again in S when having boundary conditions, but it also works for other conditions), leads to the necessary condition

$$\begin{aligned}\delta I(u;\varphi) =& \frac{d}{d\epsilon} I(u+\epsilon\varphi)|_{\epsilon=0} = \\ =& \frac{d}{d\epsilon}\int_\Omega L(x,u(x)+\epsilon\varphi(x),u_i(x)+\epsilon\varphi_i(x))dx = \\ =& \int_\Omega (\frac{\partial L}{\partial u}\varphi + \frac{\partial L}{\partial u_i}\varphi_i)(x)dx = \\ =& \int_\Omega (\frac{\partial L}{\partial u} - \frac{d}{dx^i}\frac{\partial L}{\partial u_i})(x)\varphi(x)dx + \int_{\partial\Omega} \frac{d}{dx^i}(\frac{\partial L}{\partial u_i}\varphi)(x)dS(x) = \\ =& \int_\Omega (\frac{\partial L}{\partial u} - \frac{d}{dx^i}\frac{\partial L}{\partial u_i})(x)\varphi(x)dx \overset{!}{=} 0 \quad \text{for all } \varphi \in C_c^\infty(\Omega),\end{aligned}$$

i.e. the first variation $\delta I(u;\varphi)$ must vanish. Using Du Bois-Reymond's lemma (also called fundamental lemma of the calculus of variations), we get

$$(L_u - \frac{d}{dx^i}L_{u_i})(x) = 0$$

as a necessary condition and this equation is called Euler-Lagrange equation. As we can see here, a variational equation comes together with a weak formulation. This will be important later. An expansion in ϵ of

$$I(u+\epsilon\varphi) = I(u) + \delta I(u;\varphi)\epsilon + \frac{1}{2}\delta^2 I(u;\varphi)\epsilon^2 + ...$$

leads to the first variation δI, second variation $\delta^2 I$ and so on. The functional I has a minimum at u if $\delta^2 I > 0$ and a maximum if $\delta^2 I < 0$. A necessary condition for an extremal value is $\delta I = 0$. This is completely analog to curve sketching in $\mathbb{R}^n$. Note that $\mathbb{R}^n$ is to distinguish from the base manifold M when we talk about curve sketching and analogies.

Let us briefly consider the analogous case in $\mathbb{R}^n$, i.e. curve sketching in $\mathbb{R}^n$. Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a function (instead of a functional I) and v a vector in $\mathbb{R}^n$ (instead of a test function φ). Then we get the expansion

$$\phi(x+\epsilon v) = \phi(x) + < \nabla\phi(x), v > \epsilon + \frac{1}{2} < v, \text{Hess}_\phi(x) v > \epsilon^2 + ...,$$

where ∇ is the gradient, Hess_ϕ is the Hessian of ϕ and $< .,. >$ is the Euclidian scalar product in $\mathbb{R}^n$. Finding conditions under which a differential equation is variational is analog to finding conditions under which a vector field V in $\mathbb{R}^n$ is a gradient field $\nabla\phi$. It is well-known that when the Jacobian of a vector field V

$$J_V = \begin{pmatrix} V_{1x^1} & V_{1x^2} & ... & V_{1x^n} \\ V_{2x^1} & V_{2x^2} & ... & V_{2x^n} \\ \vdots & & & \\ V_{nx^1} & V_{nx^2} & ... & V_{nx^n} \end{pmatrix}$$

is symmetric, then V is locally a gradient field $\nabla\phi$. It is clear that the Jacobian of a gradient field $\nabla\phi$ is the Hessian Hess_ϕ and this matrix is always symmetric (by Schwarz's theorem). The symmetry of the Jacobian can also be written as

$$\frac{\partial V_k}{\partial x^l} - \frac{\partial V_l}{\partial x^k} = 0, \quad \text{for all } k,l = 1,2,...,n.$$

These conditions are called *integrability conditions* and, in a similar way, we get conditions for a differential equation to be variational. In the case of differential equations these conditions are called *Helmholtz conditions* and they will be explained in detail later. The Helmholtz conditions seem to be formulated for the first time in (Hel87) and the original version of the paper (StbM12). Later, several authors investigated this problem. For example, see the paper (Ton84) and references therein.

Finding the right analogies in $\mathbb{R}^n$ solves a lot of problems in the calculus of variations and we will refer to analogies in $\mathbb{R}^n$ several times. The ideas are sometimes quite simple, but the notation can get very complicated, especially when considering relatively arbitrary manifolds and the calculus of variations with higher order differential expressions.

For example, the following problem occurs when considering more general manifolds. Let p be a point in a manifold and γ_ϵ a flow of some vector field V. The expansion in ϵ of

$$\phi(\gamma_\epsilon(p)) = \phi(p) + d\phi_p(V)\epsilon + ...$$

has terms in ϵ^2, which includes objects which are not vector fields, since there occurs a second derivative on γ_ϵ. Therefore, it is not clear at this point how to define a Hessian of ϕ on general manifolds. We also have no canonical scalar product and we cannot say what it means that the Jacobian of some vector field is symmetric. However, we will solve all these problems later. For example, instead of the gradient ∇, we will use the differential $d\phi$, instead of the scalar product, we will use the interior product ι and then most of the calculations are straight forward.

1.4. Outline of this Dissertation

In Chapter 2, we introduce more notation to clarify what we are doing later. We also explain Takens' problem in detail, show what the above mentioned Helmholtz conditions are and how to derive them. Furthermore, in Chapter 2, we formulate known results from the literature in order to solve and understand Takens' problem later on. The reader familiar with Takens' problem, jet spaces, prolongations, Euler-Lagrange mapping, Helmholtz mapping, Helmholtz conditions and locally exact sequences in the calculus of variations can continue with Chapter 3, where the actual results of this dissertation can be found. These results are mainly formulated in Section 3.6 and Section 3.8. In the Sections 3.1-3.5 and Section 3.7 we explain necessary information and techniques which will be needed for the proofs in Section 3.6 and Section 3.8. Chapter 4 provides further information about Takens' problem and investigates the question of applications and if Theorems 1.0.2 and 1.0.3 are sharp. Finally, in Chapter 5 we formulate open problems and give a conclusion of our research.

1.5. Notation

The standard definitions of manifolds, fiber bundles and related definitions in differential geometry can be found in, for example, (Lee13). All of our manifolds are smooth manifolds. Vector fields, sections and all maps between manifolds are also smooth, unless stated otherwise. Furthermore, our notation is the following:

- M is a manifold (base manifold), $\dim M = n$, local coordinates are (x^i), where $i = 1,2,...,n$.
- E is a fiber bundle with projection π and base M, we also write $\pi : E \to M$. Local coordinates on E are (x^i,u^α), where $i = 1,2,...,n$ and $\alpha = 1,2,...,m$. Furthermore, x^i are pull-backs of coordinates on M and $\dim E = n + m$.
- J^kE is the jet space of order k of sections of $\pi : E \to M$ and has local coordinates $(x^i,u^\alpha,u_i^\beta,u_{ij}^\gamma,...,u_I^\delta)$, where I is a multi-index of length k. The jet space will be introduced more precisely in Definition 2.1.4. Note that we will allow unordered multi-indices I, see Section 2.12.
- We write $f_{u_I^\beta} := \frac{\partial f}{\partial u_I^\beta}$ for partial derivative, as well as $\frac{\partial f_\beta}{\partial u_I^\alpha} := f_{\beta,u_I^\alpha}$, (where I is a multi-index). Here, f and f_β are functions on J^kE.
- $D_i := \frac{\partial}{\partial x^i} + u_i^\alpha \frac{\partial}{\partial u^\alpha} + u_{ij}^\alpha \frac{\partial}{\partial u_j^\alpha} + ...$ is the total derivative. The meaning of D_i is that

it treats functions $g(x^i,u^\alpha,u_i^\alpha,...)$ on J^kE as functions $g(x^i,u^\alpha(x^j),u_i^\alpha(x^j),...)$ on M and $(D_ig)(x^i,u^\alpha(x^j),u_i^\alpha(x^j),...) = \partial_{x^i}[g(x^i,u^\alpha(x^j),u_i^\alpha(x^j),...)]$ for all sections $(x^i,u^\alpha(x^i))$ and for all functions g on J^kE. See Subsection 2.2.3 for further details.

- $D_I := D_{i_1}D_{i_2}...D_{i_k}$, where $|I| = k$.
- By definition $u_i^\beta = D_iu^\beta$, but u_i^β is also considered as a local coordinate on J^kE. This is consistent by definition of D_i and coordinates in J^kE.
- $f_{\alpha,x^i} := \frac{\partial f_\alpha}{\partial x^i}$, as well as $f_{\alpha,i} := \frac{\partial f_\alpha}{\partial x^i}$.
- $f_{\alpha;x^i} := D_if_\alpha$, as well as $f_{\alpha;i} = D_if_\alpha$.
- We use Einstein summation. For example, $A^\alpha B_\alpha = \sum_{\alpha=1}^m A^\alpha B_\alpha$ and $A^iB_i = \sum_{i=1}^n A^iB_i$.
- $\Omega^l(M), \Omega^l(E)$ and $\Omega^l(J^kE)$ are sets of differential l-forms on M,E,J^kE.
- $\mathcal{L}_V$ denotes the Lie derivative with respect to the vector field V.
- $[.,.]$ denotes the Lie bracket.
- π^* is the pull-back under the projection π.
- π_* is the push-forward under the projection π.
- $\mathfrak{X}(M)$,$\mathfrak{X}(E)$,$\mathfrak{X}(J^kE)$ are the sets of smooth vector fields on M,E,J^kE.
- $\Gamma(E)$ denotes the set of sections σ of π.
- ι denotes the interior product of vector fields and differential forms.

2. Definition of the Basic Objects and Ideas Behind Them

2.1. Jet Spaces

For explaining the following ideas, we start mostly with ordinary differential equations (ODEs) and later we will consider partial differential equations (PDEs) in more detail. All ideas can be understood when discussing ODEs, where the notation is much easier and we can refer to the literature for generalizations. For example, additional notation is needed for PDEs when using multi-indices I and when we have to say if this multi-index is ordered or not and we do not want to bother the reader with that at this point.

A system of ODEs can be written as

$$f_\alpha(x,u^\beta,u^\beta_x,\ldots) = 0; \quad \alpha = 1,2,\ldots,m; \quad \beta = 1,2,\ldots,m;$$

where f_α is smooth in all of the coordinates $(x,u^\beta,u^\beta_x,\ldots)$. To write such equations in a coordinate independent way, we need the definition of a fiber bundle and jet space. The coordinates (x,u^β) and $(x,u^\beta,u^\beta_x,u^\beta_{xx},\ldots)$ are fiber bundle coordinates. Roughly speaking, we need functions u^β, depending on x, and derivatives of these functions, which are written as u^β_x,u^β_{xx} and so on. To clarify for which functions (or better, coordinates pulled back by sections) we can compute derivatives, we need the language of **dependent** and **independent coordinates** and this is the motivation for fiber bundles and jet spaces. We call x the independent and u^β the dependent coordinate and we think of $u^\beta = u^\beta(x)$. Jet spaces describe the concept of dependent and independent coordinates in an appropriate way.

The rough idea of fiber bundle is the following: The coordinates (x,u^β) correspond to a space E and we have a projection π, which maps $\pi(x,u^\beta) = x$. This structure allows us to speak about dependent and independent coordinates. For example, if (x,u) are coordinates on the standard 2-dimensional torus in $\mathbb{R}^3$ then we have to clarify what it means to talk about functions and derivatives $u(x),u_x(x),\ldots$, or if we want to talk about functions and derivatives $x(u),x_u(u),\ldots$. The following definition describes the necessary structure.

Definition 2.1.1. *Let M and F be topological spaces. A* ***fiber bundle over*** *M* ***with model fiber*** *F is a topological space E together with a surjective continuous* *map $\pi : E \to M$ with the property that for each $q \in M$, there exists a neighborhood U^0 of q in M and a homeomorphism $\Phi : \pi^{-1}(U^0) \to U^0 \times F$, called a* ***local trivialization of** E **over** U^0, such that the following diagram commutes:*

$$\begin{tikzcd}
\pi^{-1}(U^0) \arrow[r, "\Phi"] \arrow[d, "\pi"'] & U^0 \times F \arrow[dl, "\mathrm{proj}"] \\
U^0 &
\end{tikzcd}$$

.

The space E is called the ***total space of the bundle**, M is its* ***base**, and π is its* ***projection**. Furthermore,* proj *is the canonical projection to the first factor. If E,M and F are smooth manifolds, π is a smooth map, and the local trivializations can be chosen to be diffeomorphisms, then it is called a* ***smooth fiber bundle**.*

This definition can be found in (Lee13, p.268). Note that we are always considering smooth fiber bundles here and we simply call them fiber bundles.

Since we work locally most of the time, we can avoid the relatively complicated definition of fiber bundle in some sense (especially for the reader who is not familiar with it). We can reformulate the definition as follows: Locally, a fiber bundle looks like $\tilde{\pi} : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$ with canonical projection $\tilde{\pi}$ and it is usually used without further comment when talking about differential equations, dependent and independent coordinates. We want to discuss local coordinates on fiber bundles in more detail. Let $p \in E$ and $\pi(p) = q \in M$. Near $p \in E$, there exists a local fiber bundle chart $\varphi : U \to \Omega$, where $U \subset E$ and $\Omega \subset \mathbb{R} \times \mathbb{R}^m$ open, such that $\varphi(p) = (x,u^\beta) \in \Omega$. We also write $(x,u^\beta) = (x(p),u^\beta(p))$ and we will identify the point $p \in E$ with the local coordinates (x,u^β) without mentioning it every time and, when possible, we do not use the map φ, we rather consider $x(p)$ and $u^\beta(p)$ as maps. Near $q \in M$, there exists also a chart $\varphi^0 : U^0 \to \Omega^0$, where $U^0 \subset M$, $U^0 = \pi(U)$ and $\Omega^0 \subset \mathbb{R}$ open, such that $\varphi^0(q) = x \in \Omega^0$. Furthermore, for the canonical projection

$$\tilde{\pi} : \begin{cases} \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}, \\ (x,u^\beta) \mapsto x, \end{cases}$$

the diagram

$$\begin{tikzcd}
U \subset E \arrow[r, "\varphi"] \arrow[d, "\pi"'] & \Omega \subset \mathbb{R} \times \mathbb{R}^m \arrow[d, "\tilde{\pi}"] \\
U^0 \subset M \arrow[r, "\varphi^0"] & \Omega^0 \subset \mathbb{R}
\end{tikzcd}$$

commutes. This structure allows us to work locally with the fiber bundle $\tilde{\pi} : \Omega \to \Omega^0$ and we can use the well understood and nice structure of $\mathbb{R}^n$. Local coordinate transformations (the transition function) on fiber bundles have to be of a special form, namely

$$(x,u^\beta) \longrightarrow (y,v^\gamma) = (y(x),v^\gamma(x,u^\beta)). \tag{2.1}$$

This is called a **fiber preserving coordinate transformation**. This special transformation property will have an impact on locally defined objects and we will refer to it several times. Note that a more general transformation (not on fiber bundles) would be of the form

$$(x,u^\beta) \longrightarrow (y,v^\gamma) = (y(x,u^\beta),v^\gamma(x,u^\beta)),$$

but for fiber bundles, only $y(x)$ dependencies are allowed. The next definition is a generalization of a function $u : \mathbb{R}^n \to \mathbb{R}$, defined on more general manifolds, namely on fiber bundles.

Definition 2.1.2. *Let $\pi : E \to M$ be a fiber bundle. A* ***section of*** *E is a section of the map π, that is, a continuous map $\sigma : M \to E$ satisfying $\pi \circ \sigma = id_M$.*

This definition can be found in (Lee13, p.88, p.255). Again, let us discuss local coordinates of sections. Let $q \in M$, then $\varphi(\sigma(q)) = (x(q),u^\beta(x(q)))$ and therefore, locally, we can identify the section σ with $(x,u^\beta(x))$. Or even simpler, we can identify σ with the function $u^\beta(x)$.

Usually, $u^\beta(x)$ describes some physical quantity, like the position of a particle at time x, or the electromagnetic field at some point x.

Definition 2.1.3. *Two locally defined sections σ_1 and σ_2 are called* ***k-equivalent*** *at a point $q \in M$, if $\sigma_1(q) = \sigma_2(q)$ and if there exist local charts φ,φ^0 such that*

$$\partial_x^l(\varphi \circ \sigma_1 \circ (\varphi^0)^{-1})(x) = \partial_x^l(\varphi \circ \sigma_2 \circ (\varphi^0)^{-1})(x), \quad \textit{for all } 1 \le l \le k, \tag{2.2}$$

where $x = \varphi^0(q)$.

Note that ∂_x^l means $\partial_x ... \partial_x$ l-times. We took this definition from (Kru97b, p.29) and changed it slightly. Also note that instead of writing (2.2), we can also write

$$\partial_x^l u_1^\beta(x) = \partial_x^l u_2^\beta(x), \quad \text{for all } 1 \le l \le k, \text{for all } \beta = 1,2,...,m,$$

where $u_1^\beta(x)$ is identified with the section σ_1 and $u_2^\beta(x)$ with σ_2. The difference is that (2.2) is the same as

$$\partial_x^l(x,u_1^\beta(x)) = \partial_x^l(x,u_2^\beta(x)), \quad \text{for all } 1 \le l \le k, \text{for all } \beta = 1,2,...,m,$$

where the vector $(x,u^\beta(x))$ can be differentiated componentwise. If two sections are k-equivalent with respect to one coordinate system, then it is possible to prove that they are equivalent in every coordinate system. All k-equivalent sections at a point x (or $q \in M$) define an equivalence class, which is denoted by $[\sigma]_k(q)$ and we write $\sigma_1 \sim \sigma_2$ if σ_1 and σ_2 are equivalent (up to some order k).

Definition 2.1.4. *We define $J^k_q E$ as the set of all equivalence classes of local sections at a point $q \in M$ up to order k and*

$$J^k E := \bigcup_{q \in M} J^k_q E$$

and it is called the ***k-th order jet space of*** *E.*

We took this definition from (Kru97b, p.29). Therefore, points in $J^k E$ are given by equivalence classes of sections up to a certain order, but locally, we think of points as coordinates $(x,u^\beta,u^\gamma_x,u^\delta_{xx},\dots,u^\epsilon_{(k)})$, since these coordinates define the equivalence class at x. We will also use the short notation

$$(x,u^\beta_{[k]}) := (x,u^\beta,u^\gamma_x,u^\delta_{xx},\dots,u^\epsilon_{(k)}) \tag{2.3}$$

for local coordinates on $J^k E$. The coordinates (x,u^β) are called 0-th order coordinates, since there are zero derivatives on u^β, we call (x,u^β,u^β_x) the first order coordinates, since there is one derivative on u^β and, in general, we call (2.3) the **k-th order jet coordinates**. Moreover, a function $g = g(x,u^\beta_{[k]})$ is called a **function of order k**. A similar definition holds for tensors, vector fields, differential forms and so on. One can show that $J^k E$ has the structure of a fibre bundle with projections

$$\begin{aligned} \pi^k :&\quad J^k E \to M, \\ \pi^{k,0} :&\quad J^k E \to E, \\ \pi^{k,l} :&\quad J^k E \to J^l E, \quad k > l. \end{aligned}$$

Sometimes we refer to the infinite jet bundle, which is denoted by JE or $J^\infty E$. For example, we can use it if the order is not important primarily and it will simplify some of the notation then.

Definition 2.1.5. *The* ***infinite jet bundle*** *$J^\infty E$ is defined as follows. The inverse sequence of topological spaces $\{J^k(E),\pi^{l,k}\}$ determine an inverse limit space $J^\infty(E)$ together with projection maps*

$$\begin{aligned} \pi^{\infty,k} &: J^\infty(E) \to J^k(E), \\ \pi^{\infty,0} &: J^\infty(E) \to E, \\ \pi^{\infty} &: J^\infty(E) \to M. \end{aligned}$$

The definition can be found in (And89, p.3) and also see (Tak79).

If we have a section $(x,u^\alpha(x))$ on E, we can consider $(x,u^\alpha(x),u^\alpha_x(x),\dots,u^\alpha_{(k)}(x))$ and this will give us a section of π^k.

Definition 2.1.6. *If σ is a local section of π then the mapping $q \mapsto [\sigma]_k(q)$ is a section of π^k and it is called the* ***k-jet prolongation of the section*** *σ and is denoted by $\mathrm{pr}^k\sigma$ (sometimes written as $j^k\sigma$, where j denotes jet).*[1]

The definition can be found in (Kru97b, p.30). Most of the time we simply call it prolongation of σ. Technically, it is a lift from E to J^kE in a certain way. We will also use the notation

$$(x,u^\alpha_{[k]}(x)) = (x,u^\beta(x),u^\beta_x(x),...,u^\beta_{(k)}(x)) \quad \text{or} \quad \mathrm{pr}^k\sigma(q).$$

Note that we have to distinguish between the two expressions

$$\begin{aligned} &f_\alpha(x,u^\beta_{[k]}) \quad \text{and} \\ &f_\alpha(x,u^\beta_{[k]}(x)) = [(\mathrm{pr}^k\sigma)^* f_\alpha](x). \end{aligned}$$

The first expression is defined on J^kE and the second on M. A differential equation is an equation for sections $u^\beta(x)$, of the form

$$f_\alpha(x,u^\beta_{[k]}(x)) = 0. \tag{2.4}$$

Before we continue with differential equations and the question under which conditions they are variational, we introduce a bit more notation in the next section. This notation will be needed to solve Takens' problem.

2.2. Vector Fields and Differential Forms

In local coordinates, a general vector field $V \in \mathfrak{X}(E)$ can be written as

$$V = V^x(x,u^\alpha)\partial_x + V^\beta(x,u^\alpha)\partial_{u^\beta}.$$

Usually, we write ϕ_t for the flow of V and we have

$$\frac{d}{dt}\phi_t|_{t=0} = V.$$

Note that for PDEs, we write $V = V^i\partial_{x^i} + V^\alpha\partial_{u^\alpha}$ and when explicitly writing, for example, V^1, we use the notation $V^{x,1}$ or $V^{u,1}$, to distinguish V^i from V^β. The flow ϕ_t transforms points $p \in E$ to new points $\phi_t(p) \in E$ and we will use it, for example, to describe symmetries of differential equations. Since we can identify the point p

[1] It seems that most of the time j is used in the literature instead of pr. However, in his book (Olv86, p.98) Peter Olver uses the notation of pr and we will use it, as well. For the reader who is not familiar with prolongations, it may be easier to remember what it stands for and i,j will also be used as indices later.

with (x,u^β), we also write $\phi_t(x,u^\beta)$ for the transformation of local coordinates and we will use the same ϕ_t if there is no danger of confusion.

Roughly speaking, a symmetry of $f = f(x,u^\beta)$ is a flow ϕ_t such that

$$f(x,u^\beta) = f(\phi_t(x,u^\beta)) \quad \text{for all } t \in \mathbb{R} \text{ (or where it is defined) and for all } (x,u^\beta).$$

When we want to describe symmetries of differential equations $f(x,u^\beta_{[k]}(x)) = 0$, then, intuitively, solutions should be mapped to solutions. Since solutions are section of π, the map ϕ_t should map sections to sections. But this is not always the case. For example, the vector field

$$V = V^x(x,u)\partial_x + V^u(x,u)\partial_u = -u\partial_x + x\partial_u \in \mathfrak{X}(E),$$

defined on $E = \mathbb{R} \times \mathbb{R}$, describes rotations in the (x,u)-plane and the corresponding flow is

$$\phi_t(x,u) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix}, \quad \text{for all } t \in \mathbb{R}.$$

When we rotate by $t = \frac{\pi}{2}$, then a section is not mapped to a section. The property that sections are mapped to sections is not only useful if we want to investigate symmetries of differential equations. It is a special case of a bundle map and preserves the structure of the fiber bundle.

Definition 2.2.1. *A vector field $V \in \mathfrak{X}(E)$ is called π-**projectable** if there exists a vector field $V^0 \in \mathfrak{X}(M)$ such that $\pi_* V = V^0$. Let $l < k$. A vector field $Z \in \mathfrak{X}(J^k E)$ is called $\pi^{k,l}$-**projectable** if there exists a vector field $W \in \mathfrak{X}(J^l E)$ such that $\pi^{k,l}_* Z = W$.*

We took the definition from (Kru97b, p.31) and changed it slightly. In local coordinates, π-projectable vector fields can be written as

$$V = V^x(x)\partial_x + V^\beta(x,u^\alpha)\partial_{u^\beta}.$$

There is another property of vector fields, which is quite important and which can naturally be defined on fiber bundles.

Definition 2.2.2. *A π-projectable vector field $V \in \mathfrak{X}(E)$ is called π-**vertical** if $\pi_* V = 0$. Let $l < k$. A $\pi^{k,l}$-projectable vector field $Z \in \mathfrak{X}(J^k E)$ is called $\pi^{k,l}$-**vertical** if $\pi^{k,l}_* Z = 0$.*

The definition can be found in (Kru97b, p.31). Note that a π-vertical vector field on E can be written in local coordinates as

$$V = V^\alpha(x,u^\beta)\partial_{u^\alpha}$$

and local coordinate transformations are of the form

$$V = V^\alpha \partial_{u^\alpha} = V^\alpha \left(\frac{\partial y}{\partial u^\alpha} \partial_y + \frac{\partial v^\gamma}{\partial u^\alpha} \partial_{u^\gamma} \right), \quad \text{where} \quad \frac{\partial y}{\partial u^\alpha} = 0.$$

See (2.1) and Proposition 2.4.3 for local coordinate transformations. Also note that, at this point, there is no natural way how to define horizontal vector fields (although we will do it later, but we will need further notation). For example,

$$V = V^x(x,u^\alpha)\partial_x$$

is not defined invariantly, since it transforms as

$$V = V^x \partial_x = V^x \left(\frac{\partial y}{\partial x} \partial_y + \frac{\partial v^\gamma}{\partial x} \partial_{v^\gamma} \right)$$

and $\frac{\partial v^\beta}{\partial x} \neq 0$ in general, again see (2.1).

Since differential forms are the "dual" objects to vector fields, the concept of vertical vector fields can be transferred to differential forms, where we have the concept of horizontal forms.

Definition 2.2.3. *Let V be a vector field. Then the **annihilator** of V is the set of all differential forms ω such that $\iota_V \omega = 0$.*

Definition 2.2.4. *Let $l = 0,1,2,...$ be any integer. A form $\omega \in \Omega^l(J^k E)$ is called $\pi^{r,s}$**-horizontal** if $\iota_V \omega = 0$ for every $\pi^{r,s}$-vertical vector field $V \in \mathfrak{X}(J^k E)$.*

We took Definition 2.2.4 from (Kru97b, p.33) and changed it slightly.

2.2.1. Prolongation of Vector Fields

We already defined the prolongation of sections on E, see Definition 2.1.6. Now we also want to define the prolongation of vector fields on E. For example, symmetries of differential equations are described by prolonged vector fields. Since a differential equation is usually not of the form $f_\alpha(x,u^\beta) = 0$, but $f_\alpha(x,u^\beta,u^\beta_x,...) = 0$, we have to find out how to transform derivatives u^β_x,u^β_{xx} and so on. The idea is the following: When we transform a section on E and get a new section on E, then this transformation will have an impact on the derivatives of this sections at a point x. Therefore, the coordinates $(u_x,u_{xx},...)$ will also be transformed. This induced transformation is exactly described by prolonged vector fields.

Now let us explain how this prolongation works. First, we only want to prolong vertical vector fields on E. The idea is the following: Let

$$V = V^\alpha(x,u^\beta)\partial_{u^\alpha}$$

be a vertical vector field on E and ϕ_t its flow. Let $p \in E$, then we can find a local section $\sigma \in \Gamma(E)$ such that

$$p = \sigma(q), \quad \text{where } q = \pi(p) \in M.$$

At a fixed point $p \in E$, we can write

$$V(p) = V(\sigma(q)) \quad \text{and} \quad \phi_t(p) = \phi_t(\sigma(q)).$$

Then, for fixed $t \in \mathbb{R}$, we can show that $\phi_t \circ \sigma(q)$ is a local section on E, because $\pi \circ \phi_t \circ \sigma(q) = q$ (since $q \in M$ will not be transformed by vertical vector fields), and we can prolong this section. We compute

$$\frac{d}{dt}\mathrm{pr}^k[\phi_t \circ \sigma(q)]|_{t=0} =: \mathrm{pr}^k V(...) \in \mathfrak{X}(J^k E). \tag{2.5}$$

The bracket (...) will be explained below. Since $\mathrm{pr}^k[\phi_t \circ \sigma(q)]$ is a curve (or section of π^k) in $J^k E$, the expression

$$\frac{d}{dt}\mathrm{pr}^k[\phi_t \circ \sigma(q)]|_{t=0}$$

will give us a vector field in $J^k E$ which is what we want. The remaining question is: At which point does it define which vector? To understand this: Let γ_t be a curve in a manifold. Then $(\frac{d}{dt}\gamma_t)|_{t=0} = V(\gamma_{t=0})$ defines a vector V at the point $\gamma_{t=0}$. The same happens with prolongations of sections, where (2.5) defines a vector at the point $\mathrm{pr}^k[\phi_t \circ \sigma(q)]|_{t=0} = \mathrm{pr}^k \sigma(q)$. This expression just looks complicated, but there is no new concept behind it. Therefore, we get

$$\frac{d}{dt}\mathrm{pr}^k[\phi_t \circ \sigma(q)]|_{t=0} = \mathrm{pr}^k V(\mathrm{pr}^k[\phi_t \circ \sigma(q)]|_{t=0}) = \mathrm{pr}^k V(\mathrm{pr}^k \sigma(q)).$$

Let us consider the local coordinate expression of prolonged vector fields. We use the same ϕ there and $\phi_t(x,u^\alpha) = (x,u^\alpha_t)$ are the coordinates at time $t \in \mathbb{R}$:

$$\begin{aligned}
&\frac{d}{dt}\mathrm{pr}^k[\phi_t(x,u^\alpha(x))]|_{t=0} = \\
=&\frac{d}{dt}\begin{pmatrix} x \\ u^\alpha_t(x) \\ u^\alpha_{t,x}(x) \\ u^\alpha_{t,xx}(x) \\ \vdots \\ u^\alpha_{t,(k)}(x) \end{pmatrix}\Bigg|_{t=0} = \begin{pmatrix} \frac{d}{dt}x \\ \frac{d}{dt}u^\alpha_t(x) \\ D_x\frac{d}{dt}u^\alpha_t(x) \\ D^2_x\frac{d}{dt}u^\alpha_t(x) \\ \vdots \\ D^k_x\frac{d}{dt}u^\alpha_t(x) \end{pmatrix}\Bigg|_{t=0} = \begin{pmatrix} 0 \\ V^\alpha(x,u^\beta(x)) \\ D_xV^\alpha(x,u^\beta(x)) \\ D^2_xV^\alpha(x,u^\beta(x)) \\ \vdots \\ D^k_xV^\alpha(x,u^\beta(x)) \end{pmatrix} = \\
=&V^\alpha(x,u^\beta(x))\partial_{u^\alpha} + [D_xV^\alpha(x,u^\beta(x))]\partial_{u^\alpha_x} + ... + [D^k_xV^\alpha(x,u^\beta(x))]\partial_{u^\alpha_{(k)}}.
\end{aligned} \tag{2.6}$$

In (2.6), we used

$$\frac{d}{dt}D_x = D_x\frac{d}{dt}.$$

We get the following definition:

Definition 2.2.5. *Let $P \in J^kE$ and $V \in \mathfrak{X}(E)$ a vertical vector field. There exists a local section σ of π such that $pr^k\sigma(q) = P$ (by definition of points in J^kE). We define the* ***prolongation of*** V ***at a point*** $P \in J^kE$ *as the unique vector field pr^kV at P as*

$$pr^kV(P) := \frac{d}{dt}pr^k[\phi_t \circ \sigma](q)|_{t=0} \quad \in \mathfrak{X}(J^kE).$$

The definition can be found in (Kru97b, p.32), but we changed it slightly. It can be seen that the right hand side depends only on the k-jet of σ.

Now let us investigate how to prolong projectable vector fields. For projectable vector fields V on E, in general,

$$\pi \circ \phi_t \circ \sigma(q) \neq q$$

and therefore the definition is a bit more complicated. In this case, we also need to consider the flow ϕ^0 of π_*V, to get

$$\pi[\phi_t \circ \sigma \circ \phi^0_{-t}] = id_M.$$

For projectable vector fields, the definition is the following:

Definition 2.2.6. *Let V be a projectable vector field on E with flow ϕ. Furthermore, let ϕ^0 be the flow of π_*V and let $P \in J^kE$. There exists a local section σ of π such that $pr^k\sigma(q) = P$ (by definition of points in J^kE). We define the* ***prolongation of*** V ***at a point*** $P \in J^kE$ *as*

$$pr^kV(P) := \frac{d}{dt}\{pr^k[\phi_t \circ \sigma \circ \phi^0_{-t}](\phi^0_t(q))\}|_{t=0} \quad \in \mathfrak{X}(J^kE).$$

Again, this definition can be found in (Kru97b, p.32). The local coordinate expressions for projectable vector fields are

$$\mathrm{pr}^kV = V^x\partial_x + V^\alpha\partial_{u^\alpha} + (D_xV^\alpha - u^\alpha_xV^x_x)\partial_{u^\alpha_x} + \sum_{l=2}^{k}\xi^\alpha_l\partial_{u^\alpha_{(l)}},$$

where

$$\begin{aligned} \xi^\alpha_x = \xi^\alpha_1 &:= D_xV^\alpha - u^\alpha_xV^x_x, \\ \xi^\alpha_l &:= D_x\xi^\alpha_{l-1} - u^\alpha_{(l)}V^x_x, \quad \text{for } l \geq 2. \end{aligned}$$

In Appendix B, we will derive this expression in detail. One can also find it in (Kru97b, p.32).

The prolongation pr^kV can be written in a different way, which will also be important later. For $l \geq 1$, let us write ξ_l^α as

$$\begin{aligned}
\xi_l^\alpha &= D_x \xi_{l-1}^\alpha - u_{(l)}^\alpha V_x^x = \\
&= D_x(\xi_{l-1}^\alpha - u_{(l)}^\alpha V^x) + u_{(l+1)}^\alpha V^x = \\
&= D_x[(D_x \xi_{l-2}^\alpha - u_{(l-1)}^\alpha V_x^x) - u_{(l)}^\alpha V^x] + u_{(l+1)}^\alpha V^x = \\
&= D_x[D_x(\xi_{l-2}^\alpha - u_{(l-1)}^\alpha V^x) + u_{(l)}^\alpha V^x - u_{(l)}^\alpha V^x] + u_{(l+1)}^\alpha V^x = \\
&= D_x^2(\xi_{l-2}^\alpha - u_{(l-1)}^\alpha V^x) + u_{(l+1)}^\alpha V^x = \\
&= ... = \\
&= D_x^l(V^\alpha - u_x^\alpha V^x) + u_{(l+1)}^\alpha V^x.
\end{aligned}$$

Then we can write

$$\begin{aligned}
\mathrm{pr}^k V &= V^x \partial_x + V^\alpha \partial_{u^\alpha} + \sum_{l=1}^k D_x^l (V^\alpha - u_x^\alpha V^x) \partial_{u_{(l)}^\alpha} + V^x \sum_{l=1}^k u_{(l+1)}^\alpha \partial_{u_l^\alpha} = \\
&= V^x \partial_x + (V^\alpha - u_x^\alpha V^x) \partial_{u^\alpha} + \sum_{l=1}^k D_x^l (V^\alpha - u_x^\alpha V^x) \partial_{u_{(l)}^\alpha} + V^x \sum_{l=0}^k u_{(l+1)}^\alpha \partial_{u_l^\alpha} = \\
&= V^x D_x + \sum_{l=0}^k D_x^l (V^\alpha - u_x^\alpha V^x) \partial_{u_{(l)}^\alpha},
\end{aligned}$$

where

$$D_x = \partial_x + \sum_{l=0}^k u_{(l+1)}^\alpha \partial_{u_l^\alpha} \tag{2.7}$$

and $Q^\alpha := V^\alpha - u_x^\alpha V^x$ are called **characteristics of the vector field** V. It is elaborate to write the sum $\sum_{l=0}^k$ and therefore we will also use the short notation

$$\mathrm{pr}^k V = V^x D_x + (D_x^l Q^\alpha) \partial_{u_{(l)}^\alpha}. \tag{2.8}$$

Note that the decomposition (2.8) cannot be done in $J^k E$, but in $J^{k+1} E$, since the coefficients in (2.7) are defined in $J^{k+1} E$. We will come back to this later, when we define horizontal vector fields. Further information on prolongation of vector fields and characteristics can be found in (Olv86, And89).

Earlier, we already defined vertical vector fields and, the dual objects, horizontal forms. Now we want to define **total vector fields** and **contact forms**. Total vector fields can be considered as the horizontal vector fields and contact forms are the dual objects, the vertical forms. These are just different names for one and the same thing. With the help of these definitions, we would like to uniquely decompose vector fields and differential forms into these parts. But it turns out that this cannot be done in the same space $J^k E$, but in $J^{k+1} E$. We can also do the decomposition in the infinite jet bundle.

2.2.2. Contact Forms

Definition 2.2.7. *A differential form ω on a jet manifold J^kE is called* ***contact form****, if for every local section $\sigma : U^0 \to E$, $U^0 \subset M$, we have $(pr^k\sigma)^*\omega = 0$.*

This definition can be found in (KS08, p.1044). Since sections are in some sense naturally defined on fiber bundles, contact forms are as well. The contact one-forms

$$\Theta_l^\alpha = du_{(l)}^\alpha - u_{(l+1)}^\alpha dx, \quad 0 \leq l \leq k-1.$$

are a set of differential generators of contact forms on J^kE, that is, these forms and their exterior derivatives are algebraic generators. See Proposition 2.4.3 below, where we investigate local coordinate transformations of these basis elements. For low order forms, sometimes we write $\Theta_1^\alpha = \Theta_x^\alpha$, $\Theta_2^\alpha = \Theta_{xx}^\alpha$ and so on. For $m = 1$, we also write $\Theta_l^\alpha = \Theta_l^u$ or $\Theta_l^\alpha = \Theta_l$. For example, the form $du^\alpha \wedge dx$ is a contact form and can be written as $du^\alpha \wedge dx = \Theta^\alpha \wedge dx$. Note that the symbol Θ is probably motivated by $0 \in \mathbb{R}$, which satisfies $0 \cdot a = 0$ for every $a \in \mathbb{R}$ and similarly with contact forms, where $\Theta \wedge \omega$ is a contact form for every form ω. In the language of algebra this is called an ideal, i.e. the set $\{\text{contact forms}\} \subset \Omega^s(J^kE)$ is an ideal. Moreover, it is a **differential ideal**, i.e.

$$\begin{aligned} \Theta_l^\alpha \wedge \eta &= \text{contact form}, \quad \text{for every form } \eta \\ d\Theta_l^\alpha &= \text{contact form} \end{aligned}$$

and for prolonged vector fields $V \in \mathfrak{X}(E)$, the Lie derivative $\mathcal{L}_{\mathrm{pr}V}\Theta$ is a contact form, as well.

Proposition 2.2.8. *Let V be a projectable vector field on E, then $\mathcal{L}_{prV}\Theta_l^\alpha$ is a contact form. More generally, $\mathcal{L}_{prV}\eta$ is a contact form if η is a contact form. Furthermore, $d\Theta_{(l)}^\alpha = -\Theta_{(l+1)}^\alpha \wedge dx$.*

Proof: We simply prove it in local coordinates. Let V be a projectable vector field on E. In the following, we will write (see (2.8))

$$\mathrm{pr}^kV = V^xD_x + (D^lQ^\alpha)\partial_{u_l^\alpha}, \quad (0 \leq l \leq k).$$

For any function g of order k we can do the standard decomposition (of horizontal and contact forms)

$$\begin{aligned} dg =& g_xdx + g_{u^\alpha}du^\alpha + g_{u_x^\alpha}du_x^\alpha + ... + g_{u_{(k)}^\alpha}du_{(k)}^\alpha = \\ =& g_xdx + g_{u^\alpha}(du^\alpha - u_x^\alpha dx) + g_{u_x^\alpha}(du_x^\alpha - u_{xx}^\alpha dx) + ... + g_{u_{(k)}^\alpha}(du_{(k)}^\alpha - u_{(k+1)}^\alpha dx) + \\ &+ u_x^\alpha g_{u^\alpha}dx + u_{xx}^\alpha g_{u_x^\alpha}dx + ... + u_{(k+1)}^\alpha g_{u_{(k)}^\alpha}dx = \\ =& (D_xg)dx + g_{u^\alpha}\Theta^\alpha + g_{u_x^\alpha}\Theta_x^\alpha + ... + g_{u_{(k)}^\alpha}\Theta_{(k)}^\alpha. \end{aligned}$$

Then we can write

$$\begin{aligned}
\mathcal{L}_{\mathrm{pr}V}\Theta^\alpha_{(l)} &= (\iota_{\mathrm{pr}V}d + d\iota_{\mathrm{pr}V})(du^\alpha_{(l)} - u^\alpha_{(l+1)}dx) = \\
&= -\iota_{\mathrm{pr}V}(du^\alpha_{(l+1)} \wedge dx) + d(D^l_x Q^\alpha) = \\
&= -((D^{l+1}_x Q^\alpha) + V^x u^\alpha_{(l+2)})dx + V^x du^\alpha_{(l+1)} + d(D^l_x Q^\alpha) = \\
&= -(D^{l+1}_x Q^\alpha)dx + V^x \Theta^\alpha_{(l+1)} + d(D^l_x Q^\alpha) = \\
&= V^x \Theta^\alpha_{(l+1)} + \sum_{r=0}^{l+1} [\partial_{u^\beta_{(r)}}(D^l_x Q^\alpha)]\Theta^\beta_{(r)} = \\
&= V^x \Theta^\alpha_{(l+1)} + \sum_{r=0}^{l} [\partial_{u^\beta_{(r)}}(D^l_x Q^\alpha)]\Theta^\beta_{(r)} - V^x\Theta^\beta_{(l+1)} = \\
&= \sum_{r=0}^{l} [\partial_{u^\beta_{(r)}}(D^l_x Q^\alpha)]\Theta^\beta_{(r)} = \text{contact},
\end{aligned}$$

where we consider $D^l_x Q^\alpha$ as the function g in the fourth line and applied the standard decomposition. For an arbitrary differential form η we can do a similar calculation. Furthermore, we get

$$d\Theta^\alpha_{(l)} = d(du^\alpha_{(l)} - u^\alpha_{(l+1)}dx) = -du^\alpha_{(l+1)} \wedge dx = -\Theta^\alpha_{(l+1)} \wedge dx,$$

which proves Proposition 2.2.8. □

Note that a basis of differential forms on $\Omega^s(J^k E)$ is given by

$$dx, \Theta^\alpha, \Theta^\alpha_x, ..., \Theta^\alpha_{(k-1)}, du^\alpha_{(k)}$$

and the horizontal and contact forms

$$dx, \Theta^\alpha, \Theta^\alpha_x, ..., \Theta^\alpha_{(k-1)}$$

are not a basis. We cannot use

$$dx, \Theta^\alpha, \Theta^\alpha_x, ..., \Theta^\alpha_{(k-1)}, \Theta^\alpha_k$$

as a basis, since Θ^α_k is not a form on $J^k E$, rather on $J^{k+1}E$. But it would be nice to have a unique decomposition of differential forms into horizontal and contact forms. In the following, we consider the two spaces $J^{k-1}E$ and $J^k E$ instead of $J^k E$ and $J^{k+1}E$. Therefore, we lift the differential form $\omega \in \Omega^s(J^{k-1}E)$ to a form $\pi^{k,k-1,*}\omega \in \Omega^s(J^k E)$ and then we can do the decomposition in $J^k E$. Let us consider two simple examples. The first example is to apply the decomposition to one-forms $\omega \in \Omega^1(J^{k-1}E)$, where $n,m = 1$. We can write

$$\begin{aligned}
\omega &= \omega^x dx + \omega^u du + \omega^u_x du_x + ... + \omega^u_{k-1} du_{(k-1)} = \\
&= \omega^x dx + \omega^u(du - u_x dx) + \omega^u_x(du_x - u_{xx}dx) + ... + \omega^u_{k-1}(du_{(k-1)} - u_{(k)}dx) + \\
&\quad + \omega^u u_x dx + \omega^u_x u_{xx} dx + ... + \omega^u_{(k-1)} u_{(k)} dx = \\
&= (\omega^x + \omega^u u_x + ... + \omega^u_{(k-1)} u_{(k)})dx + \omega^u \Theta^u + ... + \omega^u_{(k-1)}\Theta^u_{(k-1)}.
\end{aligned}$$

The second example is to decompose a certain two-form $\omega \in \Omega^2(J^{k-1}E)$, where $n = 1$ and m is arbitrary. That is, let us consider

$$\begin{aligned}\omega &= A_\alpha dx \wedge du^\alpha + B_{\alpha\beta} du^\alpha \wedge du^\beta = \\ &= A_\alpha dx \wedge (du^\alpha - u^\alpha_x dx) + B_{\alpha\beta}(du^\alpha - u^\alpha_x dx) \wedge du^\beta + B_{\alpha\beta} u^\alpha_x dx \wedge du^\beta = \\ &= A_\alpha dx \wedge \Theta^\alpha + B_{\alpha\beta}\Theta^\alpha \wedge du^\beta + B_{\alpha\beta} u^\alpha_x dx \wedge \Theta^\beta = \\ &= A_\alpha dx \wedge \Theta^\alpha + B_{\alpha\beta}\Theta^\alpha \wedge \Theta^\beta + B_{\alpha\beta} u^\beta_x \Theta^\alpha \wedge dx + B_{\alpha\beta} u^\alpha_x dx \wedge \Theta^\beta. \end{aligned} \tag{2.9}$$

Similar decompositions can be done for any form on $J^{k-1}E$. We do not formulate a general proof for that, since the idea is clear. More generally, if a differential form ω is defined on J^kE and if it is $\pi^{k,k-1}$-horizontal, then we can do the same decomposition and we do not have to lift it to $J^{k+1}E$ to do the decomposition. In this case, the forms have coefficients in J^kE, but are generated by basis elements dx,du^α,...,$du^\alpha_{(k-1)}$ in $J^{k-1}E$. In any case, we either lift a form to J^kE, or if it is $\pi^{k,k-1}$-horizontal, we can do these kinds of decompositions. Further information and a decomposition theorem can be found in (Kru97b, p.35). On the infinite jet bundle, this decomposition can be done without lifting the forms, or requiring that they have to be $\pi^{k,k-1}$ horizontal. This is probably one of the motivations to define the infinite jet bundles (see (And89)). We can also use the notation of r-horizontal and s-vertical forms

$$\tilde{\Omega}^l(J^kE) = \bigoplus_{r+s=l} \Omega^{r,s}(J^kE),$$

where $\tilde{\Omega}^l(J^kE)$ is the set of $\pi^{k,k-1}$-horizontal forms in $\Omega^l(J^kE)$. Here, r counts the number of dx^i-forms (in PDE case, in ODE case $r = 0$ or $r = 1$) and s the number of contact forms Θ^α_I. For example, if we consider the forms in (2.9), then we get

$$\begin{aligned} A_\alpha dx \wedge \Theta^\alpha &\in \Omega^{1,1}(J^kE), \\ B_{\alpha\beta}\Theta^\alpha \wedge \Theta^\beta &\in \Omega^{0,2}(J^kE). \end{aligned}$$

Although it is a very nice notation, instead of using $\Omega^{r,s}$, we will mostly use the notation of horizontal and s-contact forms, since it seems that this is mainly used in the finite variational sequence (Ku04, Kru97a) and we will sometimes refer to it.

Definition 2.2.9. *Let $\omega \in \Omega^l(J^kE)$ be a $\pi^{k,k-1}$-horizontal l-form. We say that ω is a* ***1-contact l-form*** *if it is a contact form and if for every π^k-vertical vector field $V \in \mathfrak{X}(J^kE)$, the contraction $\iota_V\omega$ is π^k-horizontal.*

Note that for 1-forms, we only need the language of contact and horizontal forms. Inductively we define

Definition 2.2.10. *Let $\omega \in \Omega^l(J^kE)$ be a $\pi^{k,k-1}$-horizontal l-form. We say that ω is a* ***s-contact l-form*** *if it is a contact form and if for every π^k-vertical vector field $V \in \mathfrak{X}(J^kE)$, the contraction $\iota_V\omega$ is $(s-1)$-contact.*

We took Definition 2.2.9 and 2.2.10 from (KS08, p.1045) and changed them slightly.

Lemma 2.2.11. *Let $\omega \in \Omega^l(J^kE)$ be a $\pi^{k,k-1}$-horizontal l-form, then it can be uniquely decomposed into a sum of s-contact forms.*

See Proposition 2.6.3 in (Kru97b, p.35).

2.2.3. Total- and Generalized Vector Fields

In this subsection, we want to define the dual objects to contact forms. The contact forms can be considered as the vertical forms (these are just different names). We already defined vertical vector fields and now we want to define horizontal vector fields. The finite and infinite jet space are different in this regard and in the latter we have total vector fields as the dual objects to contact forms. In the finite jet space, the Cartan distribution describes the dual vector fields to contact forms. We mostly work with the finite jet space, and therefore now we want to investigate the Cartan distribution in more detail. More precisely, we want to consider a special part of the Cartan distribution, namely, the part which is very similar to the total derivative. To understand this in more detail, let us first define the Cartan distribution and then we continue with the discussion.

Definition 2.2.12. *A map δ on J^kE is called* ***distribution*** *if it assigns at every point $p \in J^kE$ a vector sub-scpace of T_pJ^kE.*

We took this definition from (Kru97b, p.21) and changed it slightly. Note that δ is to distinguish from the Dirac-δ-distribution here.

The annihilator space of the contact one-forms is the so-called **Cartan distribution**

$$\mathcal{C}_k := \operatorname{span}\{\partial_x + \sum_{0\leq l\leq k-1} u^\alpha_{(l+1)}\partial_{u^\alpha_l}, \partial_{u^\alpha_{(k)}}\}$$

and span means the span over smooth functions.

Now we want to define total vector fields on the finite dimensional jet spaces. Actually, these are not vector fields in the classical sense, as we will see below, rather vector fields along a map. The idea is to define the total derivative operator

$$D_x = \partial_x + u^\alpha_x\partial_{u^\alpha} + u^\alpha_{xx}\partial_{u^\alpha_x} + ... + u^\alpha_{(k+1)}\partial_{u^\alpha_{(k)}}$$

as a vector field. Local coordinate transformations lead to

$$\frac{\partial y}{\partial x}D_y = \frac{\partial y}{\partial x}(\partial_y + v^\alpha_y\partial_{v^\alpha} + v^\alpha_{yy}\partial_{v^\alpha_y} + ... + v^\alpha_{(k+1)}\partial_{v^\alpha_{(k)}}),$$

see Proposition (2.4.3). Usually, total derivatives are written together with horizontal forms, for example,

$$\omega = (D_x A)dx = \frac{\partial y}{\partial x}(D_y A)\frac{\partial x}{\partial y}dy = (D_y A)dy,$$

such that we get coordinate independent expressions (we will come back to this later). More generally, any operator of the form

$$V = V^x(x,u^\beta_{[k+1]})(\partial_x + u^\alpha_x \partial_{u^\alpha} + u^\alpha_{xx}\partial_{u^\alpha_x} + ...u^\alpha_{(k+1)}\partial_{u^\alpha_{(k)}}),$$

is called total vector field on J^kE (see the definition below). It is clear that this V cannot be a (classical) vector field on J^kE, since it includes the coordinates $u^\alpha_{(k+1)}$, even when $V^x(x,u^\beta_{[k+1]}) = V^x(x,u^\beta_{[k]})$ or when $V^x(x,u^\beta_{[k+1]}) = V^x(x)$. Therefore, we are confronted with the problem to define V in some space, different from J^kE. To solve this problem, we need the definition of a vector field along a map. This definition will also be needed when defining generalized vector fields.

Definition 2.2.13. *Let P,Q be finite dimensional manifolds and $\phi : P \to Q$ a smooth map. A **vector field along** ϕ is a smooth map $Z : P \to TQ$ such that for all $p \in P$, $Z(p)$ is a tangent vector to Q at the point $\phi(p)$.*

We took this definition from (And89, p.8) (and changed it slightly). Note that in the standard definition for vector fields we consider a smooth map $\phi : Q \to Q$, where $P = Q$ and $\phi = id_Q$. In the case where $Q = J^kE$, a vector field is a (smooth) map from $J^kE \to TJ^kE$. In a similar way, we define it for differential forms. We only define it for one-forms, the definition is straight forward for l-forms

Definition 2.2.14. *Let P,Q be finite dimensional manifolds and $\phi : P \to Q$ a smooth map. A **differential 1-form along** ϕ is a smooth map $Z : P \to T^*Q$ such that for all $p \in P$, $Z(p)$ is a cotangent vector to Q at the point $\phi(p)$.*

Finally, we are able to define total vector fields in a geometric way:

Definition 2.2.15. *A **total vector field** on J^kE is a vector field Z along $\pi^{k+1,k}$ which annihilates all contact 1-forms on $J^{k+1}E$ (or along $\pi^{k+1,k}$), that is, $\iota_Z\Theta^\alpha_{(l)} = 0$ for all $\alpha = 1,2,...,m$ and for all $0 \leq l \leq k$.*

See (And89, p.27) or (KS08, p.1046) for further details. Note that total vector fields in (And89, p.27) have a slightly different meaning. **Total derivatives** V are of the form

$$V = V^x(x)(\partial_x + u^\alpha_x \partial_{u^\alpha} + u^\alpha_{xx}\partial_{u^\alpha_x} + ...u^\alpha_{(k+1)}\partial_{u^\alpha_{(k)}}),$$

and this form holds in every local coordinate system (see Proposition 2.4.3). They are a special case of total vector fields, where $V^x(x,u^\beta_{[k+1]}) = V^x(x)$. Beside that they are total vector fields, they have the property that they are π^{k+1}-projectable, i.e. $\pi^{k+1}_*(V^x(x)D_x) = V^x(x)\partial_x$, which means that $V^x(x)D_x$ projects to a vector field on M.

Note that in a fixed coordinate system, the total derivative of a function g, written as $D_x g$, can also be characterised by the equation

$$(D_x g)(x,u^\alpha_{[k+1]}(x)) = \partial_x[g(x,u^\alpha_{[k]}(x)] \tag{2.10}$$

for every section $(x,u^\alpha(x))$ (see (Olv86, p.112)). This equation explains the name total derivative, since every coordinate is considered to depend on x when applying ∂_x.

Definition 2.2.16. *A **generalized vector field** Z **on** J^kE is a vector field along the map $\pi^{l,k}$ for some $l > k$, i.e. Z is a smooth map*

$$Z: \ J^lE \to TJ^kE$$

such that for all $p \in J^lE$, $Z(p) \in T_{\pi^{l,k}(p)}(J^kE)$.

We took this definition from (And89, p.8). A similar definition can be found in (Olv86, p295). A generalized vector field means that the coefficients are defined in J^lE, but the basis elements $\partial_x,\partial_{u^\alpha},\ldots,\partial_{u^\alpha_{(k)}}$ are from J^kE. For example,

$$V = \partial_x + u_x\partial_u$$

is a generalized vector on E. Note that total vector fields do not have a flow.

The prolongation of generalized vector fields can be found in (Olv86) in Section 5.1, or in (And89), especially Proposition 1.12 therein. Let V be a generalized vector field on E, written in local coordinates as

$$V = V^x(x,u^\alpha_{[k]})\partial_x + V^\beta(x,u^\alpha_{[k]})\partial_{u^\beta}.$$

Roughly speaking, we define the l-th prolongation of V as the generalized vector field (see the proposition below)

$$\mathrm{pr}^lV = V^xD_x + D^r_x(V^\beta - u^\beta_xV^x)\partial_{u^\beta_{(r)}}, \quad 0 \le r \le l. \tag{2.11}$$

Intrinsically, we have the following proposition, which is also a definition:

Proposition 2.2.17. *Let l,k be integers and $N = k + l$. Furthermore, let V be a vector field along $\pi^{k,0}$, i.e. a generalized vector field on E. Then there exists a unique vector field Z along $\pi^{N,l}$, i.e. a generalized vector field on J^lE, such that*
i) Z projects to V, that is, $\pi^{N,0}Z = V$, and
ii) Z preserves the contact ideal, that is, $\mathcal{L}_Z\omega$ must be a contact form on J^NE whenever ω is a contact form on J^NE.
*We call $Z = pr^lV$ the l**-th prolongation of** V.*

We took this proposition from (And89, p.23), where we can also find the proof. Note that in (And89), the definition is for the infinite jet bundle, but we can also use the definition for the finite jet bundle.

Proof: Let Z be the generalized vector field

$$Z = Z^x \partial_x + Z_r^\beta \partial_{u^\beta_{(r)}}, \quad \text{where } 0 \le r \le l,$$

which preserves the contact ideal. It is sufficient to apply the Lie derivative on the basis contact forms $\Theta^\alpha_{(s)}$, where we get

$$\begin{aligned}
\mathcal{L}_Z \Theta^\alpha_{(s)} =& \mathcal{L}_Z (du^\alpha_{(s)} - u^\alpha_{(s+1)} dx) = \\
=& dZ^\alpha_s - (\iota_Z d + d\iota_Z)(u^\alpha_{(s+1)} dx) = \\
=& dZ^\alpha_s - [\iota_Z du^\alpha_{(s+1)} \wedge dx + d(u^\alpha_{(s+1)} Z^x)] = \\
=& dZ^\alpha_s - [Z^\alpha_{s+1} dx - Z^x du^\alpha_{(s+1)} + du^\alpha_{(s+1)} Z^x + u^\alpha_{(s+1)} dZ^x] = \\
=& dZ^\alpha_s - Z^\alpha_{s+1} dx - u^\alpha_{(s+1)} D_x Z^x dx + \text{contact form} = \\
=& (D_x Z^\alpha_s - Z^\alpha_{s+1} - u^\alpha_{(s+1)} D_x Z^x) dx + \text{contact form} \stackrel{!}{=} \text{contact form}.
\end{aligned}$$

Then, inductively, we get the equation

$$\begin{aligned}
Z^\alpha_{s+1} &= D_x Z^\alpha_s - u^\alpha_{(s+1)} D_x Z^x = \\
&= D_x (Z^\alpha_s - u^\alpha_{(s+1)} Z^x) + u^\alpha_{(s+2)} Z^x = \\
&= D_x [(D_x Z^\alpha_{s-1} - u^\alpha_{(s)} D_x Z^x) - u^\alpha_{(s+1)} Z^x)] + u^\alpha_{(s+2)} Z^x = \\
&= D_x^2 (Z^\alpha_{s-1} - u^\alpha_{(s)} Z^x) + u^\alpha_{(s+2)} Z^x = \\
&= \ldots = \\
&= D_x^{s+1} (Z^\alpha - u^\alpha_x Z^x) + u^\alpha_{(s+2)} Z^x,
\end{aligned}$$

which uniquely determines the coefficients Z^α_s for all $0 \le s \le l-1$. That is, we get

$$\begin{aligned}
Z &= Z^x \partial_x + Z^\alpha \partial_{u^\alpha} + [D_x^{s+1}(Z^\alpha - u^\alpha_x Z^x) + u^\alpha_{(s+2)} Z^x] \partial_{u^\alpha_{(s+1)}} = \\
&= Z^x D_x + (Z^\alpha - u^\alpha_x Z^x) \partial_{u^\alpha} + D_x^{s+1}(Z^\alpha - u^\alpha_x Z^x) \partial_{u^\alpha_{(s+1)}}.
\end{aligned}$$

If Z projects to V, i.e. $\pi^{N,0} Z = V$ then we get $Z^x = V^x$ and $Z^\alpha = V^\alpha$, which leads to the expression in (2.11). $\square$

A generalized vector field on E which is π-vertical is also called an **evolutionary vector field**. This definition can be found in (And89, p.25).

2.3. Differential Equations, Weak Formulations and Source Forms

In this section, we try to motivate the definition of source form, which will be a fundamental object later. It seems that the formal definition can be found in many books and articles, but a good explanation and motivation is hard to find. Therefore, we find our own motivation in the following. In the paper (Tak77) of Takens in 1977 we can probably find the definition of source form for the first time.

Before we start with the motivation, recall that a variational equation always comes via a weak formulation (see Section 1.3). Therefore, let us first define what we mean by a weak formulation precisely.

If we multiply a differential equation

$$f_\alpha(x,u^\beta_{[k]}(x)) = 0 \tag{2.12}$$

by a test function $\varphi = (\varphi^\alpha) \in C_0^\infty(U^0)$, where $U^0 \subset M$ is open and $\bar{U}^0$ is the closure of U^0, such that $\bar{U}^0$ is compact, and let Ω^0 be the corresponding subset of $\mathbb{R}^n$ in local coordinates, then we can integrate over this expression, to get a new equation of the form

$$K(\sigma,\varphi) :=< f_\alpha,\varphi^\alpha >_{L^2(U^0)}= \int_{\Omega^0} f_\alpha(x,u^\beta_{[k]}(x))\varphi^\alpha(x)dx = 0. \tag{2.13}$$

If we require this equation for all φ^α then we get a **weak formulation** of the problem (2.12). Note that in (2.13) the expression $< f_\alpha,\varphi^\alpha >_{L^2(U^0)}$ is a short notation, which explains the structure, however, f_α has to be pull-backed by a section $(x,u^\beta_{[k]}(x))$, otherwise integration does not make sense and the corresponding section of $(x,u^\beta(x))$ is σ.

A variational equation is an equation of the form

$$\frac{d}{dt}I(\sigma_t)|_{t=0} = \delta I(\sigma, \frac{d}{dt}\sigma_t|_{t=0}) \int_{\Omega^0} (\mathcal{E}_\alpha L)(x,u^\beta_{[k]}(x))\varphi^\alpha(x)dx = 0, \tag{2.14}$$

where $f_\alpha = \mathcal{E}_\alpha L$ is the Euler-Lagrange expression (see Section 1.3). After applying the Du-Bois Reymond lemma, we get the Euler-Lagrange equation

$$(\mathcal{E}_\alpha L)(x,u^\beta_{[k]}(x)) = 0.$$

Again, let $(x,u^\beta_t(x))$ be the local coordinate section corresponding to the 1-parameter section σ_t, $t \in \mathbb{R}$. The following observation is crucial: For variational equations, like (2.14), the test function φ^α is of the form

$$\frac{d}{dt}u^\alpha_t(x)|_{t=0} = \varphi^\alpha(x)$$

and $u_t^\alpha(x)$ has the expansion

$$u_t^\alpha(x) = u^\alpha(x) + t\varphi^\alpha(x) + O(t^2).$$

That is, the test function φ^α is not an arbitrary function, rather a perturbation of u^α and we have certain transformation rules for u^α, and therefore also for φ^α. More precisely, the test function φ^α can be identified with a vertical vector field $V = V^\alpha \partial_{u^\alpha}$, where $\varphi^\alpha = V^\alpha$ after pull-back by a section. Furthermore, V^α has support in U^0 after pull-back by a section. For vector fields $V = V^x \partial_x \in \mathfrak{X}(M)$, or $V = V^x \partial_x + V^\alpha \partial_{u^\alpha} \in \mathfrak{X}(E)$, support in U^0, or in $\pi^{-1}(U^0)$, means that V and all derivatives of V^x, or (V^x,V^α), must vanish at ∂U^0, or $\pi^{-1}(\partial U^0)$ (note that the vector fields must be smooth on M or E).

Now we want to consider the expression in (2.14) without writing the integral. Let us define the differential form $\eta_V \in \Omega^1(J^k E)$, which is defined as

$$\eta_V := f_\alpha(x,u^\beta_{[k]}) V^\alpha(x,u^\beta) dx$$

and let us also consider the form

$$f_\alpha(x,u^\beta_{[k]}(x))\varphi^\alpha(x)dx,$$

which is defined on M. We pull-back η_V by a prolonged section σ on E and we get

$$\mathrm{pr}^k\sigma^*[f_\alpha(x,u^\beta_{[k]})V^\alpha(x,u^\beta)dx] = f_\alpha(x,u^\beta_{[k]}(x))\varphi^\alpha(x)dx,$$

where we can observe the relation between the vertical vector field V and the test function φ^α. This is exactly the expression in (2.14), when writing $\mathcal{E}_\alpha L = f_\alpha$. Let us check if the differential form η_V is well-defined. Since dx transforms (because of the fiber preserving local coordinate transformations) as

$$dx = \frac{\partial x}{\partial y}dy + \frac{\partial x}{\partial v^\beta}dv^\beta = \frac{\partial x}{\partial y}dy,$$

and

$$V = V^\alpha \partial_{u^\alpha} = V^\alpha \left(\frac{\partial y}{\partial u^\alpha}\partial_y + \frac{\partial v^\beta}{\partial u^\alpha}\partial_{v^\beta} \right) = V^\alpha \frac{\partial v^\beta}{\partial u^\alpha}\partial_{v^\beta},$$

we get

$$\eta_V = f_\alpha V^\alpha dx = f_\beta V^\alpha \frac{\partial v^\beta}{\partial u^\alpha}\frac{\partial x}{\partial y}dy. \tag{2.15}$$

Therefore, the set of forms η_V is in the sense well-defined under local coordinate transformations, as we do not get $dv^\alpha, dv^\alpha_x, \ldots$-basis elements after coordinate transformation. Intrinsically, η_V is a π^k-horizontal 1-form on $J^k E$. It turns out that η_V can be written as

$$\eta_V = \iota_W[f_\alpha(x,u^\beta_{[k]})du^\alpha \wedge dx], \tag{2.16}$$

where $W \in \mathfrak{X}(J^kE)$ is a $\pi^{k,0}$-projectable vector field such that

$$\pi^{k,0}_* W = V^\alpha \partial_{u^\alpha}.$$

The vector field W is not uniquely determined. For example,

$$W_1 = V^\alpha \partial_{u^\alpha} = V^\alpha \partial_{u^\alpha} + 0 \cdot \partial u^\alpha_x + ... + 0 \cdot \partial u^\alpha_{(k)}$$

and

$$W_2 = \mathrm{pr}^k(V^\alpha \partial_{u^\alpha}) = V^\alpha \partial_{u^\alpha} + (D_x V^\alpha)\partial_{u^\alpha_x} + ... + (D^k_x V^\alpha)\partial_{u^\alpha_{(k)}} \tag{2.17}$$

define the same η_V. Most of the time we will use $W = \mathrm{pr}^k V$. This is because for the variational weak formulation we apply $\mathrm{pr}^k V$ to L and only the component V^α (or φ^α) occurs in (2.14) because of partial integration.

Equation (2.16) motivates the definition of the differential form

$$\Delta := f_\alpha du^\alpha \wedge dx \quad \in \Omega^2(J^kE), \tag{2.18}$$

which is called a source form. In the case of PDEs, we define

$$\Delta := f_\alpha du^\alpha \wedge dx^1 \wedge ... \wedge dx^n \quad \in \Omega^{n+1}(J^kE). \tag{2.19}$$

Definition 2.3.1. *Any $(n+1)$-form on J^kE (here $n = 1$) which is $\pi^{k,0}$-horizontal and 1-contact is called* ***source form*** *and usually we write Δ for it.*

For example, see (Kru97b, p.37), where it is called dynamical form (in case of classical mechanics) or see (AP94, p.197). In local coordinates, source forms can be written as we did in (2.18) or as

$$\Delta = f_\alpha \Theta^\alpha \wedge dx.$$

Note that it is not sufficient to define source forms as elements in $\Omega^{1,1}(J^kE)$, since then, for example, $f_\alpha \Theta^\alpha_x \wedge dx$ would also be allowed, but this form is not $\pi^{k,0}$-horizontal. Roughly speaking, source forms represent differential equations via weak formulations. One can easily check that the local coordinate representation is well defined, since the transformation of coordinates is

$$\begin{aligned}\Delta = f_\alpha du^\alpha \wedge dx = f_\alpha \left(\frac{\partial u^\alpha}{\partial y} dy + \frac{\partial u^\alpha}{\partial v^\beta} dv^\beta \right) \wedge \left(\frac{\partial x}{\partial y} dy + \frac{\partial x}{\partial v^\gamma} dv^\gamma \right) = \\ = f_\alpha \frac{\partial u^\alpha}{\partial v^\beta} \frac{\partial x}{\partial y} dv^\beta \wedge dy = \tilde{f}_\beta dv^\beta \wedge dy,\end{aligned} \tag{2.20}$$

where $\frac{\partial x}{\partial v^\gamma} = 0$ because of fiber preserving local coordinate transformations. Now we are able to give a reformulation of equation (2.12), which can be written as

$$(\mathrm{pr}^k \sigma)^*(\iota_{\mathrm{pr} V} \Delta) = 0$$

for all projectable and π-vertical vector fields V on E. In this notation, the section σ is the solution of the differential equation.

Proposition 2.3.2. *The differential equation* $f_\alpha(x,u^\beta_{[k]}(x)) = 0$ *is equivalent to*
i) $pr\sigma^*[\iota_{prV}(f_\alpha du^\alpha \wedge dx)] = 0$ *for all* π*-vertical* $V \in \mathfrak{X}(E)$*,*
ii) $pr\sigma^*[\iota_W(f_\alpha du^\alpha \wedge dx)] = 0$ *for all* π^k*-vertical* $W \in \mathfrak{X}(J^kE)$*.*

Let us say a few more words about the definition of source form, since it is a very fundamental object in Theorem 1.0.2 and 1.0.3. As we explained, source forms represent differential equations via weak formulations. But, even more, the source form implies a transformation property for the differential equation $f_\alpha = 0$. When we start with a source form and extract the equation $f_\alpha = 0$ in local coordinates, then this is clear, since in another local coordinate system we get the equation $f_\alpha \frac{\partial u^\alpha}{\partial v^\beta}\frac{\partial x}{\partial y} = 0$. However, when we start with an equation $f_\alpha = 0$, then it is problematic to assign a corresponding source form. Usually, a differential equations is written down in local coordinates as an equation of the form

$$f_\alpha(x,u^\beta_{[k]}(x)) = 0$$

and usually it means that f_α are scalar equations (they do not transform under the indices α). The definition of source form however tells us exactly how to transform the components of f_α (this is also partially indicated by the lower index α) and f_α transforms like

$$f_\alpha \longrightarrow f_\alpha \frac{\partial u^\alpha}{\partial v^\beta}\frac{\partial x}{\partial y}, \tag{2.21}$$

as we saw in (2.20). Therefore, when we assign to an arbitrary system of differential equations $f_\alpha = 0$ a source form Δ, then we force a special transformation property of that f_α.

Note that only for differential equations, where $\alpha = 1,2,...,m$, we can assign a source form. For example, in the classical formulation of Maxwell's equations, we have the electromagnetic field $\boldsymbol{E},\boldsymbol{B}$. Maxwell's equations are eight equations with six unknowns $\boldsymbol{E},\boldsymbol{B}$. Therefore, we cannot assign a source form to Maxwell's equations (so easily). But, if we consider the formulation with the vector potential A_μ, then we can assign a source form and even more, the equations are variational.

Note that assigning a source form to a given differential equation is not problematic from the solution point of view. Since $f_\alpha = 0$ and

$$f_\alpha \frac{\partial u^\alpha}{\partial v^\beta}\frac{\partial x}{\partial y} = 0 \tag{2.22}$$

have the same solutions, because the fiber coordinate transformation is a diffeomorphism (otherwise Proposition 2.3.2 would not make sense). Solutions are sections in E and they are defined independently of local coordinates. But, the symmetries of m scalar equations $f_\alpha = 0$ and the equations we get from source forms Δ, are different in general. We will come back to this later, when we have precisely defined what we mean by a symmetry.

Note that there is no unique way how to assign a source form to a given differential equation. For example, let

$$f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = 0$$

be a system of differential equations. Then we can assign

$$\Delta_1 := f_1 du^1 \wedge dx + f_2 du^2 \wedge dx$$

or

$$\Delta_2 := f_2 du^1 \wedge dx + f_1 du^2 \wedge dx.$$

The source form Δ_1 might be variational, but the source form Δ_2 might not be variational (see Proposition 2.4.2 for the definition of variational source forms).

In the following, we always have the situation where we have a source form from the beginning or where we assigned a source form to a given system of m equations. We do not have to say how we actually assigned it. The assumptions in Theorem 1.0.2 and 1.0.3 do not rely on this information. These theorems are statements about a given source form. If one assigns another source form to a given system of m equations then it is a statement about this other source form.

2.4. Variational Equations and the Variational Sequence

A variational functional is a map $I : S \to \mathbb{R}$,

$$I(u) := \int_{\Omega^0} L(x, u^\alpha_{[k]}(x))dx, \tag{2.23}$$

where S is a suitable space for the functions u (or sections σ). For example, $S = \{u \in C^\infty(\Omega^0) : u|_{\partial\Omega^0} = 0\}$. The Lagrangian L is given and usually one wants to minimize or maximise the functional I, that is, to find the extremals u. Note that since I is a functional, which depends non-locally on u and also on derivatives of u, sometimes it is better to write $I[u]$ instead of $I(u)$ to indicate the more complicated dependency. In the following, we also write

$$I(\sigma) := \int_{U^0} \mathrm{pr}^k \sigma^*(L dx) \tag{2.24}$$

and it is the same as (2.23), when identifying the section σ with $u(x)$. We also write $U^0 \subset M$ for the corresponding $\Omega^0 \subset \mathbb{R}$ in local coordinates.

As we saw in the introduction, extremizing the functional I leads to the Euler-Lagrange equations. For simplicity, at the beginning we only consider first order Lagrangians $L = L(x,u^\alpha,u^\alpha_x)$ and corresponding variational equations of the form

$$f_\alpha = (\partial_{u^\alpha} - D_x\partial_{u^\alpha_x})L = 0. \tag{2.25}$$

The operator $\partial_{u^\alpha} - D_x\partial_{u^\alpha_x}$ will be needed quite often. For higher order expressions we define

$$\mathcal{E}_\alpha := \partial_{u^\alpha} - D_x\partial_{u^\alpha_x} + D_x^2\partial_{u^\alpha_{xx}} \pm ... + (-1)^k D_x^k\partial_{u^\alpha_{(k)}} \tag{2.26}$$

and we call it the **Euler-Lagrange operator of order** k. The coordinate invariance of this operator will be shown below. When we consider differential equations, we are in the field of functional analysis, which is dealing with ∞-dimensional spaces. We say ∞-dimensional spaces, since the space S is ∞-dimensional.

In the following, we want to understand the geometric meaning of variational equations and how they differ from general differential equations. For example, in the case of 0-th order Lagrangian $L = L(x,u^\alpha)$, (2.25) becomes

$$f_\alpha = L_{u^\alpha} = (\nabla L)_\alpha, \tag{2.27}$$

where $\nabla = (\partial_{u^1},...,\partial_{u^m})$ is the gradient of the u^α-coordinates (the vertical coordinates). At least in this case, variational equations are exactly described by vector fields f_α on E (given in local coordinates in the components f_α), which can be written as gradients. Or in the language of differential forms, they correspond to 1-forms, which are (locally) exact. This picture also holds for higher order jet coordinates and we want to investigate it in more detail. To understand this picture, we first consider finite dimensional spaces.

Finite dimensional space $\mathbb{R}^n$: Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a function (instead of a functional I). To find the extremals of ϕ, we have to consider the equation

$$\frac{d}{dt}\phi(x + tv) =< \nabla\phi(x),v >= 0,$$

where $t \in \mathbb{R}$, $v \in \mathbb{R}^n$ and $< .,. >$ denotes the Euclidian scalar product. More generally, when we do not have a vector space structure and when we are not able to write $x + tv$: Let γ_t be any smooth curve in $\mathbb{R}^n$ (or some other manifold) such that $\gamma_{t=0} = x$ and $(\frac{d}{dt}\gamma_t)(x)|_{t=0} = v(x)$, then we can write

$$(\mathcal{L}_v\phi)(x) =< \nabla\phi(x),v(x) >= d\phi_x(v) = 0, \quad \text{for all } v \in \mathfrak{X}(\mathbb{R}^n),$$

where $\mathcal{L}_v$ denotes the Lie derivative with respect to the vector field v.

A vector field w in $\mathbb{R}^n$ can be written as a gradient field $\nabla\phi$ for some function ϕ in $\mathbb{R}^n$ if

$$(\mathcal{L}_v\phi)(x) =< \nabla\phi(x),v(x) >=< w(x),v(x) >,$$

for all vector fields v in $\mathbb{R}^n$ and for all $x \in \mathbb{R}^n$ (this is just a reformulation, for example, we take $v = e_l$, where e_l are the canonical unit vectors). On a general manifold we do not have a gradient operator ∇ or canonical scalar product, but we do have the concept of differential $d\phi$ and interior product ι. Therefore, more generally we can say: A differential 1-form $\omega \in \Omega^1$ can be written as $d\phi$ for some function ϕ if

$$\iota_v \omega = \mathcal{L}_v \phi = (d\phi)(v), \tag{2.28}$$

for all vector fields v on that manifold.

Infinite dimensional analog in the calculus of variations: Let σ be a section in E and γ_t a 1-parameter family of sections such that $\gamma_{t=0} = \sigma$. Let us assume that $\gamma_t = \phi_t \circ \sigma$ for some flow ϕ_t and $(\frac{d}{dt}\phi_t)|_{t=0} = V$, where V is a vertical vector field on E and $\mathrm{supp}V \subset \pi^{-1}(U^0)$. Then we define

$$\begin{aligned}(\mathcal{L}_{\mathrm{pr}V} I)(\sigma) &:= \frac{d}{dt} I(\gamma_t)|_{t=0} = \delta I(\sigma; V) = \\ &=< (\mathcal{E}_\alpha L), V^\alpha >_{L^2(U^0)} = \int_{U^0} \mathrm{pr}^k \sigma^* [(\mathcal{E}_\alpha L) V^\alpha dx]. \end{aligned} \tag{2.29}$$

Note that the expressions in the scalar product $< .,. >$ in (2.29) have to be pull-backed by a section σ, otherwise the integral does not make sense. We did not write the pull-back for simplicity and because it is more important to understand the structure here. Let us further consider a weak formulation

$$K(\sigma; V) := \int_{U^0} \mathrm{pr}^k \sigma^* (f_\alpha V^\alpha dx) =< f_\alpha, V^\alpha >_{L^2(U^0)} .$$

Then f_α (or K) is **variational** if there exists a functional I such that

$$K(\sigma; V) =< f_\alpha, V^\alpha >_{L^2(U^0)} =< (\mathcal{E}_\alpha L), V^\alpha >_{L^2(U^0)} = \delta I(\sigma; V) \tag{2.30}$$

for all vertical vector fields V on E and $\mathrm{supp}V \subset \pi^{-1}(U^0)$.

Local differential geometry on $J^k E$ approach: Instead of working with the integral in (2.24), we also want to work with the differential form

$$\lambda = L(x, u^\alpha_{[k]}) dx, \tag{2.31}$$

which is called **Lagrange form**. It can be invariantly characterized in the following way:

Definition 2.4.1. *Any n-form (here $n = 1$) on $J^k E$ which is π^k-horizontal is called **Lagrange form of order** k.*

This definition can be found in (Kru97b, p.37). In local coordinates, such forms are exactly described by the expression in (2.31) and they only include the dx-basis elements (no $du^\alpha_{(l)}$-terms are allowed). Since L is not a function and transforms more complicated, we will call it **Lagrangian**.

What we should be careful about the local diff. geom. on J^kE approach: We would probably assume, or let us say this is what we hope, that

$$\mathcal{L}_{\mathrm{pr}V}(Ldx) = (\mathcal{E}_\alpha L)V^\alpha dx \ ? \tag{2.32}$$

or even

$$d(Ldx) = (\mathcal{E}_\alpha L)du^\alpha \wedge dx \ ? \tag{2.33}$$

but this is wrong for at least two reasons:

- The integral in (2.29) has a non-trivial kernel.
- We have to pull-back (2.32) and (2.33) by a section when we want to get (2.29) and the pull-back also has a non-trivial kernel, namely contact forms.

To understand that (2.32) and (2.33) is not satisfied and how it should be correctly, let us investigate the following: Let $L = L(x,u,u_x)$ and V a projectable vector field:

$$\begin{aligned}
&\mathcal{L}_{\mathrm{pr}V}(Ldx) = \\
&= (\mathcal{L}_{\mathrm{pr}V}L)dx + L\mathcal{L}_{\mathrm{pr}V}dx = \\
&= (L_xV^x + L_uV^u + L_{u_x}D_xV^u)dx + LV^x_x dx = \\
&= (L_uV^u + L_{u_x}D_xV^u)dx + D_x(LV^x)dx = \\
&= (L_u - D_xL_{u_x})V^u dx + D_x\underbrace{(L_{u_x}V^u + LV^x)}_{=:A} dx = \\
&= (\mathcal{E}L)V^u dx + (A_x + u_xA_u + u_{xx}A_{u_x})dx = \\
&= (\mathcal{E}L)V^u dx + A_x dx + A_u(u_x dx - du) + A_{u_x}(u_{xx}dx - du_x) + A_u du + A_{u_x}du_x = \\
&= (\mathcal{E}L)V^u dx + A_u\Theta + A_{u_x}\Theta_x + dA.
\end{aligned} \tag{2.34}$$

It is clear that the contact forms $A_u\Theta$ and $A_{u_x}\Theta_x$ vanish if we pull-back them by a section and it is also clear that

$$\int_{U^0} \mathrm{pr}^k\sigma^*(dA) = 0$$

for all projectable vector fields $V \in \mathfrak{X}(E)$ and $\mathrm{supp}V \subset \pi^{-1}(U^0)$, since A, and therefore also dA, depend linearly on V^x,V^u and derivatives of V^x,V^u. These are the two points which we mentioned above. We can do a similar calculation with the left hand side of (2.33) and we get

$$\begin{aligned}
d(Ldx) &= (L_u du + L_{u_x} du_x) \wedge dx = \\
&= L_u du \wedge dx + L_{u_x} d(u_x dx) = \\
&= L_u du \wedge dx + L_{u_x} d(-du + u_x dx) = \\
&= L_u du \wedge dx - L_{u_x} d\Theta^u = \\
&= L_u du \wedge dx - d(L_{u_x}\Theta^u) + (dL_{u_x}) \wedge \Theta^u = \\
&= L_u du \wedge dx - d(L_{u_x}\Theta^u) + (D_x L_{u_x} dx + L_{uu_x}\Theta^u + L_{u_x u_x}\Theta^u_x) \wedge \Theta^u = \\
&= (\mathcal{E}L) du \wedge dx + d(\text{1-contact}) + (\text{2-contact}). \qquad (2.35)
\end{aligned}$$

Therefore, d maps the Lagrange form Ldx to the Euler-Lagrange source form

$$(\mathcal{E}L)du \wedge dx$$

modulo d(1-contact) and 2-contact forms. In the quotient spaces of the variational sequence these terms are considered to be zero. Very roughly speaking, d maps Ldx to the Euler-Lagrange source form $(\mathcal{E}L)du \wedge dx$. See (Ku04, Kru97a) for further details.

Now we want to understand the meaning of the identities (2.34) and (2.35) without doing the calculation. Let us consider the first variation, which can be written as

$$\begin{aligned}
\delta I(\sigma; V) &= \frac{d}{dt}\int_{U^0} \mathrm{pr}^k \gamma_t^*(Ldx)|_{t=0} = \int_{U^0} \mathrm{pr}^k \sigma^*[(L_u V^u + L_{u_x} D_x V^u)dx] = \\
&= \int_{U^0} \mathrm{pr}^k \sigma^*[(\mathcal{E}L)V^u dx]. \qquad (2.36)
\end{aligned}$$

Here γ_t is a 1-parameter family of sections such that $\gamma_{t=0} = \sigma$, $\frac{d}{dt}\gamma_t|_{t=0} = V$ and $V \in \mathfrak{X}(E)$ is vertical and $\mathrm{supp}V \subset \pi^{-1}(U^0)$. Without writing the integral and pull-back by a section in (2.36) we get basically two expressions

$$\begin{aligned}
(L_u V^u + L_{u_x} D_x V^u)dx &= \iota_{\mathrm{pr}V}[(L_u du + L_{u_x} du_x) \wedge dx] \quad \text{and} \\
(\mathcal{E}L)V^u dx &= \iota_{\mathrm{pr}V}[(\mathcal{E}L)du \wedge dx],
\end{aligned}$$

which both can be identified with the the first variation δI. They only differ by a total derivative, which vanishes when integrated, and a contact form, which vanishes when pull-backed by a section. Neglecting $\iota_{\mathrm{pr}V}$, these two expressions are

$$(L_u du + L_{u_x} du_x) \wedge dx \quad \text{and} \qquad (2.37)$$

$$(\mathcal{E}L) du \wedge dx \qquad (2.38)$$

and they should be equivalent in some sense. Indeed, (2.37) and (2.38) are two different representations in the same equivalence class in the variational sequence and we can write

$$[(L_u du + L_{u_x} du_x) \wedge dx] = [(\mathcal{E}L)du \wedge dx],$$

where the brackets [...] denote equivalence class. We can also write it as

$$[d(Ldx)] = [(\mathcal{E}L)du \wedge dx],$$

what we roughly mentioned above. This is one example, which impressively shows, that it makes a lot of sense to use equivalence classes and develop objects like the variational sequence.

Actually, there are at least two ways how to define the Euler-Lagrange mapping $\mathcal{E}$. The first way is to use equivalence classes and the operator d as we discussed above. A second way is to use interior Euler-operators, especially when considering infinite jet spaces and the Bicomplex in (And89). The interior Euler-operator is a projective operator. Roughly speaking, the interior Euler-operator chooses the representative (2.38), which is in some sense special (VU13, p.1) (that is why it is usually identified with the first variation). We will neither need the variational sequence or Bicomplex essentially to solve Takens' problem and one can find further information in (Ku04, Kru97a, And89). We would need more space to introduce these objects here in detail and for us the ideas are more important.

Without referring to the variational sequence or Bicomplex, the question for us, how to define variational expressions f_α invariantly, without the integral in (2.30), is still open and we will solve it now. Intuitively, we would of course define it as: f_α is (locally) **variational** if there exists a Lagrangian L such that

$$f_\alpha = \mathcal{E}_\alpha L.$$

We basically have to show that this is well defined and independent of the choice of local coordinates.

Note that to write the first variation δI in the form $\int(\mathcal{E}_\alpha L)V^\alpha dx$ makes sense and it is the way one usually writes it. But, for example, for the second variation $\delta^2 I$ we cannot find a similar expression in the following sense: We cannot partially integrate such that there are no $D_x V^\alpha$,$D_x^2 V^\alpha$,...-terms. That is, we cannot shift all total derivatives on V^α to other terms. For example, let us consider $L = \frac{1}{2}(u^2 + u_x^2)$, then

$$\delta^2 I(\sigma; V) = \int_{U^0} \mathrm{pr}^k \sigma^* \{\frac{1}{2}[(V^u)^2 + (D_x V^u)^2]dx\}$$

and there is no way to write $\delta^2 I$ without total derivatives D_x on some of the V^u-terms. It seems that partial integration technique can be used very powerful for the first variation, but it is not such powerful for higher variations when we have non-linear expressions in V^u.

Now let us define in local coordinates what we mean by a variational source form.

The definition is straightforward and we only have to check that the definition is independent of the choice of local coordinates. The following proposition is also a definition:

Proposition 2.4.2 (Definition). *Let $\Delta = f_\alpha du^\alpha \wedge dx$ be a source form on J^kE. We say that Δ is* ***locally variational*** *if for each $p \in E$ there exist a neighbourhood $U \subset E$, a Lagrange form $\lambda = Ldx$ on $(\pi^{k,0})^{-1}(U) \subset J^kE$ of order $l \leq k$, and local coordinates on $(\pi^{k,0})^{-1}(U) \subset J^kE$, such that we can write*

$$\Delta = [(\partial_{u^\alpha} - D_x \partial_{u^\alpha_x} + D_x^2 \partial_{u^\alpha_{xx}} \pm ...(-1)^l D_x^l \partial_{u^\alpha_{(l)}})L] du^\alpha \wedge dx \quad \text{on } (\pi^{k,0})^{-1}(U). \tag{2.39}$$

The expression in (2.39) *is independent of the choice of local coordinates. If there exists a globally defined Lagrange form $\lambda = Ldx$ on J^kE such that Δ can be written as* (2.39)*, then we say Δ is* ***globally variational****.*

If Δ is variational then we also write $\Delta = E(\lambda)$, where E is the (formal) Euler-Lagrange operator which leads to the expression in (2.39). Whether there exists a Lagrangian L of the same order as f_α, or lower order, or any order, is a different question and we do not want to discuss it here in detail. Let us only note that we can always lift a Lagrangian L defined on J^lE to a Lagrangian defined on J^kE such that Ldx and Δ are defined in the same space J^kE. When the order of Ldx is greater than Δ then we can lift Δ such that Ldx and Δ are defined in the same jet space. Later, we will construct a Lagrangian with the help of homotopy formulas such that Ldx and Δ have the same order. We will also show that we can always add a term $D_x\Lambda$ to L which leads to the same expression f_α. By adding this term, the order of the equivalent Lagrangian $\tilde{L} := L + D_x\Lambda$ can always be increased to arbitrary order, but this is a technical detail at this point.

The invariance of variational expressions (or source forms) can be understood in two ways. First, doing local coordinate transformations of the corresponding functional I. Then the extremals are again described by (the same) functional in these new local coordinates. When extremizing a functional, this always leads to to Euler-Lagrange equations in the corresponding local coordinates. Then it is also clear that the Lagrangian L transforms according to Ldx. More precisely, let us consider

$$I(\sigma) = \int_{\Omega^0} L(x,u(x),u_x(x))dx = \int_{\tilde{\Omega}^0} \tilde{L}(y,v(y),v_y(y))dy,$$

then the section $\sigma \in \Gamma(E)$ extremizes I if and only if the corresponding sections in local coordinates $u(x)$ and $v(y)$ satisfy the corresponding Euler-Lagrange equation. Since we are defining *variational* through source forms (pure local differential geometry on J^kE) and not through functionals I, the question basically is: How are source forms related to weak formulations of differential equations and one has to understand this in detail, what we did in (2.30) and what is completely in analogy to (2.28).

The second way how to show the invariance is simply to consider local coordinate transformations. We will do that below. Note that even if the combinations $D_x\partial_{u_x}, D_x^2\partial_{u_{xx}}$ and so on in (2.39) look pretty specific and somehow invariant (in x), it is actually non-trivial to show it, since we also need to consider the specific plus-minus-combinations. Without mentioning the relation to the variational functional I, it would probably be quite surprising to find out that the coordinate expression of variational source forms is invariant. Note that, for example, $f_x(x)dx$ is invariant since it is the differential df at the point x. But $L_u(x,u,u_x)du \wedge dx$ is not invariantly defined on J^1E, even the combination $(\partial_u L)du$ looks invariant. See Proposition 2.4.3 below and how to transform ∂_u on J^kE. The expression $L_u(x,u)du \wedge dx$ is invariantly defined on E or on J^1E.

Note that if a source form is locally variational then it is not necessarily globally variational and the Lagrangian L may not exist globally (see Section 4.1 and 6th counter example).

This is a good point where we can discuss local coordinate transformations on fiber bundles in more detail. These transformations are important in many situations and we get the following proposition:

Proposition 2.4.3. *Let $\pi : E \to M$ be a fiber bundle, where $n,m = 1$. Coordinate transformations of the coordinates (x,u,u_x,u_{xx}), vector fields $\partial_x,\partial_u,\partial_{u_x},\partial_{u_{xx}}$, total derivative D_x, differential forms dx,du,du_x,du_{xx} and contact forms $\Theta, \Theta_x,\Theta_{xx}$ are as follows:*

$$\begin{pmatrix} y \\ v \\ v_y \\ v_{yy} \end{pmatrix} = \begin{pmatrix} y(x) \\ v(x,u) \\ \frac{\partial x}{\partial y}\left(\frac{\partial v}{\partial x} + u_x\frac{\partial v}{\partial u}\right) \\ \frac{\partial^2 x}{\partial y^2}\left(\frac{\partial v}{\partial x} + u_x\frac{\partial v}{\partial u}\right) + \left(\frac{\partial x}{\partial y}\right)^2\left(\frac{\partial^2 v}{\partial x^2} + 2u_x\frac{\partial^2 v}{\partial x\partial u} + u_x^2\frac{\partial^2 v}{\partial u^2} + u_{xx}\frac{\partial v}{\partial u}\right) \end{pmatrix} =$$

$$= \begin{pmatrix} y \\ v \\ \frac{\partial x}{\partial y}D_x v \\ \frac{\partial x}{\partial y}D_x\left(\frac{\partial x}{\partial y}D_x v\right) \end{pmatrix},$$

$$\begin{pmatrix} \partial_x \\ \partial_u \\ \partial_{u_x} \\ \partial_{u_{xx}} \end{pmatrix} = \begin{pmatrix} \frac{\partial y}{\partial x} & \frac{\partial v}{\partial x} & \frac{\partial v_y}{\partial x} & \frac{\partial v_{yy}}{\partial x} \\ 0 & \frac{\partial v}{\partial u} & \frac{\partial v_y}{\partial u} & \frac{\partial v_{yy}}{\partial u} \\ 0 & 0 & \frac{\partial v_y}{\partial u_x} & \frac{\partial v_{yy}}{\partial u_x} \\ 0 & 0 & 0 & \frac{\partial v_{yy}}{\partial u_{xx}} \end{pmatrix} \begin{pmatrix} \partial_y \\ \partial_v \\ \partial_{v_y} \\ \partial_{v_{yy}} \end{pmatrix}, \quad D_x = \frac{\partial y}{\partial x}D_y,$$

$$
\begin{pmatrix} dx \\ du \\ du_x \\ du_{xx} \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial y} & 0 & 0 & 0 \\ \frac{\partial u}{\partial y} & \frac{\partial u}{\partial v} & 0 & 0 \\ \frac{\partial u_x}{\partial y} & \frac{\partial u_x}{\partial v} & \frac{\partial u_x}{\partial v_y} & 0 \\ \frac{\partial u_{xx}}{\partial y} & \frac{\partial u_{xx}}{\partial v} & \frac{\partial u_{xx}}{\partial v_y} & \frac{\partial u_{xx}}{\partial v_{yy}} \end{pmatrix} \begin{pmatrix} dy \\ dv \\ dv_y \\ dv_{yy} \end{pmatrix},
$$
$$
\begin{pmatrix} \Theta \\ \Theta_x \\ \Theta_{xx} \end{pmatrix} = \begin{pmatrix} \frac{\partial u}{\partial v} & 0 & 0 \\ \frac{\partial u_x}{\partial v} & \frac{\partial u_x}{\partial v_y} & 0 \\ \frac{\partial u_{xx}}{\partial v} & \frac{\partial u_{xx}}{\partial v_y} & \frac{\partial u_{xx}}{\partial v_{yy}} \end{pmatrix} \begin{pmatrix} \tilde{\Theta} \\ \tilde{\Theta}_y \\ \tilde{\Theta}_{yy} \end{pmatrix},
$$

where $\frac{\partial y}{\partial x}, \frac{\partial v}{\partial u}, \frac{\partial v_y}{\partial u_x}, \frac{\partial v_{yy}}{\partial u_{xx}} \neq 0$ *and furthermore*

$$
\frac{\partial v_y}{\partial u_x} = \frac{\partial x}{\partial y}\frac{\partial v}{\partial u}, \quad \frac{\partial v_{yy}}{\partial u_{xx}} = \left(\frac{\partial x}{\partial y}\right)^2 \frac{\partial v}{\partial u}. \tag{2.40}
$$

We see here that the transformations for v_{yy} and higher order coordinates can get very complicated and usually we should avoid these explicit expression. However, the highest order coordinate dependencies are quite simple and described by (2.40) and similar expressions for higher order. In the highest order coordinates, here in u_{xx},v_{yy}, we have affine linear dependencies. Furthermore, we have a polynomial structure for lower order coordinate dependencies, except for (x,u) and (y,v).

Let $(x,u,u_x,...,u_{(k)})$ and $(y,v,v_y,...,v_{(k)})$ be two local coordinate systems on J^kE. It follows that the dependencies are

$$
x(y), u(y,v), u_x(y,v,v_y),..., u_{(k)}(y,v,v_y,...,v_{(k)}),
$$

simply by the fact that J^lE are jet spaces with projections $\pi^{k,l}$,$\pi^{k,0}$,π^k and that we need fiber preserving coordinate transformations for every $0 \leq l < k$ and every k. Moreover, the transformations on J^kE are induced by the transformations on E. The charts for higher order coordinates are also called associated charts (Kru97b, p.30), since we can construct these charts from the charts on E.

In the following, sometimes it is helpful to distinguish between the coordinates (x,u), (y,v) and the mappings between these coordinates. Let ψ and ϕ be the maps between these coordinates, i.e.

$$
\begin{aligned}
&\begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} \psi_1(y) \\ \psi_2(y,v) \end{pmatrix}, && \begin{pmatrix} y \\ v \end{pmatrix} = \begin{pmatrix} \phi_1(x) \\ \phi_2(x,u) \end{pmatrix}, \\
&x = \psi_1(y) = \psi_1(\phi_1(x)), && u = \psi_2(y,v) = \psi_2(\phi_1(x),\phi_2(x,u)), \\
&y = \phi_1(x) = \phi_1(\psi_1(y)), && v = \phi_2(x,u) = \phi_2(\psi_1(y),\psi_2(y,v)), \\
&\psi \circ \phi = id_{(x,u)}, && \phi \circ \psi = id_{(y,v)}.
\end{aligned} \tag{2.41}
$$

When there is no danger of confusion then we also use the shorter notation

$$x = \psi_1(y) = x(y), \quad \frac{\partial \psi_1}{\partial y} = \frac{\partial x}{\partial y}$$

and so on. We also do not write where they are evaluated, that is, we do not write $\frac{\partial y(x)}{\partial x}$, we only write $\frac{\partial y}{\partial x}$ (all expressions are clearly defined by this notation).

Locally, the structure of a fibre bundle is one to one coded in the coordinate transformations (2.41). For global properties we have to know all charts.

The transformation on E is completely described by (2.41) and there is nothing more to say about that. However, for J^1E, J^2E and so on, there are induced transformations (associated charts) and we want to investigate them.

Proof of Proposition 2.4.3: A point in J^1E, written as $[\sigma]_1(q)$, is by definition an equivalence class of local sections at $q \in M$ (up to first order). Let $\sigma \in \Gamma(E)$ be a representative of that point and $(x,u(x))$ the corresponding local coordinate section. We also write (x_0,u,u_{x_0}) for the point $[\sigma]_1(q)$ and $\varphi^0(q) = x_0$. We fix this (local) section $u(x)$ and vary x in a small neighborhood of x_0. For any x in this small neighbourhood, we get

$$\begin{pmatrix} \phi_1(x) \\ \phi_2(x,u(x)) \end{pmatrix} = \begin{pmatrix} y \\ \phi_2(\psi_1(y),u(\psi_1(y))) \end{pmatrix} = \begin{pmatrix} y \\ v(y) \end{pmatrix},$$

that is, a local section in the (y,v)-coordinate system. Then we get the coordinate v_y as follows (since v_y defines the equivalence class of sections in the (y,v)-coordinate system up to first order)

$$\begin{aligned} v_y(x,u(x),u_x(x)) &= \partial_y \phi_2(x,u(x)) = \\ &= \frac{\partial \psi_1}{\partial y}\left(\frac{\partial \phi_2}{\partial \psi_1} + \frac{\partial u}{\partial \psi_1}\frac{\partial \phi_2}{\partial u}\right) = \frac{\partial x}{\partial y}\left(\frac{\partial v}{\partial x} + u_x \frac{\partial v}{\partial u}\right) = \\ &= \frac{\partial x}{\partial y}(v_x + u_x v_u) = \frac{\partial x}{\partial y} D_x v, \end{aligned}$$

where we have to evaluate the expression at $x = x_0$. Note that we actually do not have to use D_x in the last expression to write down the coordinate transformations, but this notation will be helpful later and it is reasonable to use it here ($D_x v$ actually means $(D_x v)(x)$, and as we said above, we do not always write the evaluation at x). This can be done for all sections $u(x)$ in the same equivalence class, since they agree up to first order and deliver the same point (x_0,u,u_{x_0}) at $\varphi^0(q) = x_0$. Therefore, we get the transformation for the coordinates $(x,u,u_x) \to (y,v,v_y)$. The whole transformation is induced by the local coordinate transformation on E. Similar calculations can be done for higher order coordinates. For v_{yy} we get

$$\begin{aligned}
&v_{yy}(x,u(x),u_x(x),u_{xx}(x)) = \partial_y\partial_y\phi_2(x,u(x)) = \\
=&D_y\left[\frac{\partial x}{\partial y}\left(v_x + u_x v_u\right)\right] = \\
=&\frac{\partial^2 x}{\partial y^2}(v_x + u_x v_u) + \left(\frac{\partial x}{\partial y}\right)^2 D_x(v_x + u_x v_u) = \\
=&\frac{\partial^2 x}{\partial y^2}(v_x + u_x v_u) + \left(\frac{\partial x}{\partial y}\right)^2 (v_{xx} + 2u_x v_{xu} + u_x^2 v_{uu} + u_{xx} v_u) = \\
=&\frac{\partial^2 x}{\partial y^2}\left(\frac{\partial v}{\partial x} + u_x\frac{\partial v}{\partial u}\right) + \left(\frac{\partial x}{\partial y}\right)^2\left(\frac{\partial^2 v}{\partial x^2} + 2u_x\frac{\partial^2 v}{\partial x\partial u} + u_x^2\frac{\partial^2 v}{\partial u^2} + u_{xx}\frac{\partial v}{\partial u}\right).
\end{aligned}$$

Higher order coordinates are probably best describe by the total derivative operator D_x, see property (2.10). For example, $v_{yy} = \frac{\partial x}{\partial y}D_x(\frac{\partial x}{\partial y}D_x v)$ and v is considered to be a function depending on x and u and D_x increases the order by one. For $v_{(k)}$ we have to write

$$v_{(k)} = \underbrace{\frac{\partial x}{\partial y}D_x\frac{\partial x}{\partial y}D_x ... \frac{\partial x}{\partial y}D_x}_{k\text{-times}}\, v.$$

Then the partial derivatives $\partial_x,\partial_u,\partial_{u_x},\partial_{u_{xx}}$ transform as

$$\begin{aligned}
\partial_x &= \frac{\partial}{\partial x} = \frac{\partial y}{\partial x}\partial_y + \frac{\partial v}{\partial x}\partial_v + \frac{\partial v_y}{\partial x}\partial_{v_y} + \frac{\partial v_{yy}}{\partial x}\partial_{v_{yy}}, \\
\partial_u &= \frac{\partial}{\partial u} = \underbrace{\frac{\partial y}{\partial u}\partial_y}_{=0} + \frac{\partial v}{\partial u}\partial_v + \frac{\partial v_y}{\partial u}\partial_{v_y} + \frac{\partial v_{yy}}{\partial u}\partial_{v_{yy}}, \\
\partial_{u_x} &= \frac{\partial}{\partial u_x} = \underbrace{\frac{\partial y}{\partial u_x}\partial_y + \frac{\partial v}{\partial u_x}\partial_v}_{=0} + \frac{\partial v_y}{\partial u_x}\partial_{v_y} + \frac{\partial v_{yy}}{\partial u_x}\partial_{v_{yy}}, \\
\partial_{u_{xx}} &= \frac{\partial}{\partial u_{xx}} = \underbrace{\frac{\partial y}{\partial u_{xx}}\partial_y + \frac{\partial v}{\partial u_{xx}}\partial_v + \frac{\partial v_y}{\partial u_{xx}}\partial_{v_y}}_{=0} + \frac{\partial v_{yy}}{\partial u_{xx}}\partial_{v_{yy}}
\end{aligned}$$

and similar for higher order partial derivatives. Note that we do not write down the explicit expressions for $\frac{\partial v_{yy}}{\partial x}$ and so on, since they are complicated. Now we use the transformations for the coordinates and $\partial_x,\partial_u,\partial_{u_x},...$ and we can write

$$\begin{aligned}
D_x =&\partial_x + u_x\partial_u + u_{xx}\partial_{u_x} = \\
=&\left(\frac{\partial y}{\partial x}\partial_y + \frac{\partial v}{\partial x}\partial_v + \frac{\partial v_y}{\partial x}\partial_{v_y}\right) + u_x\left(\frac{\partial v}{\partial u}\partial_v + \frac{\partial v_y}{\partial u}\partial v_y\right) + u_{xx}\frac{\partial v_y}{\partial u_x}\frac{\partial}{\partial v_y} = \\
=&\frac{\partial y}{\partial x}\partial_y + \underbrace{\left(\frac{\partial v}{\partial x} + u_x\frac{\partial v}{\partial u}\right)}_{=\frac{\partial y}{\partial x}v_y}\partial_v + \underbrace{\left(\frac{\partial v_y}{\partial x} + u_x\frac{\partial v_y}{\partial u} + u_{xx}\frac{\partial v_y}{\partial u_x}\right)}_{=D_x v_y = \frac{\partial y}{\partial x}D_y v_y = \frac{\partial y}{\partial x}v_{yy} \quad (*)}\partial_{v_y} = \frac{\partial y}{\partial x}D_y,
\end{aligned}$$

where $(*)$ can be seen by induction (especially for higher order coordinates). The calculation for dx,du,... can be done in a similar way and we get

$$\begin{aligned}
dx &= \frac{\partial x}{\partial y}dy + \underbrace{\frac{\partial x}{\partial v}dv + \frac{\partial x}{\partial v_y}dv_y + \dots}_{=0}\,, \\
du &= \frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial v}dv + \underbrace{\frac{\partial u}{\partial v_y}dv_y + \dots}_{=0}\,, \\
du_x &= \frac{\partial u_x}{\partial y}dy + \frac{\partial u_x}{\partial v}dv + \frac{\partial u_x}{\partial v_y}dv_y + \underbrace{\frac{\partial u_x}{\partial v_{yy}}dv_{yy} + \dots}_{=0}\,.
\end{aligned}$$

For the contact forms we get

$$\begin{aligned}
\Theta = du - u_x dx &= \left(\frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial v}dv\right) - \frac{\partial y}{\partial x}\left(\frac{\partial u}{\partial y} + v_y\frac{\partial u}{\partial v}\right)\frac{\partial x}{\partial y}dy = \\
&= \frac{\partial u}{\partial v}\left(dv - v_y dy\right) = \frac{\partial u}{\partial v}\tilde{\Theta}.
\end{aligned}$$

To compute the transformation for Θ_x we first consider

$$\begin{aligned}
u_{xx} = D_x u_x &= \frac{\partial y}{\partial x}D_y u_x = \\
&= \frac{\partial y}{\partial x}\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_x}{\partial v}v_y + \frac{\partial u_x}{\partial v_y}v_{yy}\right)
\end{aligned}$$

and then

$$\begin{aligned}
\Theta_x =& du_x - u_{xx}dx = \\
=& \left(\frac{\partial u_x}{\partial y}dy + \frac{\partial u_x}{\partial v}dv + \frac{\partial u_x}{\partial v_y}dv_y\right) - \frac{\partial y}{\partial x}\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_x}{\partial v}v_y + \frac{\partial u_x}{\partial v_y}v_{yy}\right)\frac{\partial x}{\partial y}dy = \\
=& \frac{\partial u_x}{\partial v}(dv - v_y dy) + \frac{\partial u_x}{\partial v_y}(dv_y - v_{yy}dy) \\
=& \frac{\partial u_x}{\partial v_y}\tilde{\Theta}_y + \frac{\partial u_x}{\partial v}\tilde{\Theta}. \qquad (2.42)
\end{aligned}$$

This means that the form Θ_x is not invariantly defined, but the combinations in (2.42) are. Higher order contact forms can be done in a similar and we only write down Θ_{xx}. First, we write

$$\begin{aligned}
u_{(3)} = D_x u_{xx} &= \frac{\partial y}{\partial x}D_y u_{xx} = \\
&= \frac{\partial y}{\partial x}\left(\frac{\partial u_{xx}}{\partial y} + \frac{\partial u_{xx}}{\partial v}v_y + \frac{\partial u_{xx}}{\partial v_y}v_{yy} + \frac{\partial u_{xx}}{\partial v_{yy}}v_{(3)}\right)
\end{aligned}$$

and then

$$
\begin{aligned}
\Theta_{xx} =& du_{xx} - u_{(3)}dx = \\
=& \left(\frac{\partial u_{xx}}{\partial y} dy + \frac{\partial u_{xx}}{\partial v} dv + \frac{\partial u_{xx}}{\partial v_y} dv_y + \frac{\partial u_{xx}}{\partial v_{yy}} dv_{yy} \right) - \\
& - \frac{\partial y}{\partial x} \left(\frac{\partial u_{xx}}{\partial y} + \frac{\partial u_{xx}}{\partial v} v_y + \frac{\partial u_{xx}}{\partial v_y} v_{yy} + \frac{\partial u_{xx}}{\partial v_{yy}} v_{(3)} \right) \frac{\partial x}{\partial y} dy = \\
=& \frac{\partial u_{xx}}{\partial v}(dv - v_y dy) + \frac{\partial u_{xx}}{\partial v_y}(dv_y - v_{yy}dy) + \frac{\partial u_{xx}}{\partial v_{yy}}(dv_{yy} - v_{(3)}dy) \\
=& \frac{\partial u_{xx}}{\partial v_{yy}} \tilde{\Theta}_{yy} + \frac{\partial u_{xx}}{\partial v_y} \tilde{\Theta}_y + \frac{\partial u_{xx}}{\partial v} \tilde{\Theta}.
\end{aligned}
$$

The fact that

$$\frac{\partial y}{\partial x}, \frac{\partial v}{\partial u}, \frac{\partial v_y}{\partial u_x}, \frac{\partial v_{yy}}{\partial u_{xx}} \neq 0$$

simply follows by the requirement that $\frac{\partial y}{\partial x} \neq 0$ and $\frac{\partial v}{\partial u} \neq 0$, since the transformation on E is a diffeomorphism, the corresponding Jacobian matrix must be invertible and $\frac{\partial y}{\partial u} = 0$, since we consider fiber bundles. Higher order, like $\frac{\partial v_y}{\partial u_x} = \frac{\partial x}{\partial y}\frac{\partial v}{\partial u} \neq 0$, can be reduced to these expressions (as we have shown above). □

Note that the transformation $D_x = \frac{\partial y}{\partial x} D_y$ was probably what we expected, since D_x imitates derivatives of functions which only depend on x (or y), and then the transformation is given by the standard transformation for derivatives.

Proof of Proposition 2.4.2 (for first order Lagrangians): We simply consider local coordinate transformations and show that the definition is invariant. According to Proposition 2.4.3, where we can find the transformation of local coordinates, we can write

$$
\begin{aligned}
\partial_u - D_x \partial_{u_x} &= \left(\frac{\partial v}{\partial u} \partial_v + \frac{\partial v_y}{\partial u} \partial_{v_y} \right) - \frac{\partial y}{\partial x} D_y \left(\frac{\partial v_y}{\partial u_x} \partial_{v_y} \right) = \\
&= \frac{\partial v}{\partial u} \partial_v + \frac{\partial v_y}{\partial u} \partial_{v_y} - \left(D_x \frac{\partial v_y}{\partial u_x} \right) \partial_{v_y} - \frac{\partial y}{\partial x} \underbrace{\frac{\partial v_y}{\partial u_x}}_{=\frac{\partial x}{\partial y}\frac{\partial v}{\partial u}} D_y \partial_{v_y} = \\
&= \underbrace{\left(\frac{\partial v_y}{\partial u} - D_x \frac{\partial v_y}{\partial u_x} \right)}_{=:(*)} \partial_{v_y} + \frac{\partial v}{\partial u}(\partial_v - D_y \partial_{v_y}) = \\
&= \frac{\partial v}{\partial u} \left[-\frac{\partial y}{\partial x} \frac{\partial^2 x}{\partial y^2} \partial_{v_y} + (\partial_v - D_y \partial_{v_y}) \right] = \\
&= \frac{\partial v}{\partial u} \frac{\partial y}{\partial x} (\partial_v - D_y \partial_{v_y}) \frac{\partial x}{\partial y}. \qquad (2.43)
\end{aligned}
$$

In the second and last line in (2.43) we used $\frac{\partial y}{\partial x}\frac{\partial x}{\partial y} = 1$ (for example, this is not true for $\frac{\partial u}{\partial y}\frac{\partial y}{\partial u} = 0$ and therefore one has to be careful with that). Before we continue, let us prove $(*)$:

$$\begin{aligned}
\frac{\partial v_y}{\partial u} - D_x\frac{\partial v_y}{\partial u_x} &= \frac{\partial v_y}{\partial u} - D_x\left(\frac{\partial x}{\partial y}\frac{\partial v}{\partial u}\right) = \\
&= \frac{\partial v_y}{\partial u} - \left(D_x\frac{\partial x}{\partial y}\right)\frac{\partial v}{\partial u} - \frac{\partial x}{\partial y}D_x\frac{\partial v}{\partial u} = \\
&= \frac{\partial v_y}{\partial u} - \left(\frac{\partial y}{\partial x}D_y\frac{\partial x}{\partial y}\right)\frac{\partial v}{\partial u} - \frac{\partial x}{\partial y}\frac{\partial y}{\partial x}D_y\frac{\partial v}{\partial u} = \\
&= \frac{\partial v_y}{\partial u} - \frac{\partial y}{\partial x}\frac{\partial^2 x}{\partial y^2}\frac{\partial v}{\partial u} - \partial_u D_y v = \\
&= -\frac{\partial y}{\partial x}\frac{\partial^2 x}{\partial y^2}\frac{\partial v}{\partial u}.
\end{aligned} \tag{2.44}$$

Now we use (2.43) and we can write

$$\begin{aligned}
[(\partial_u - D_x\partial_{u_x^\alpha})L]du \wedge dx &= \left[\frac{\partial v}{\partial u}\frac{\partial y}{\partial x}(\partial_v - D_y\partial_{v_y})(\frac{\partial x}{\partial y}L)\right] du \wedge dx = \\
&= [(\partial_v - D_y\partial_{v_y})(\frac{\partial x}{\partial y}L)]dv \wedge dy
\end{aligned}$$

and we have proven Proposition (2.4.2) (for first order L and $n = m = 1$). □

We want to finish this part of Section 2.4 with a table of corresponding objects:

classical analysis in $\mathbb{R}^n$	∞-dim, calculus of variations	differential geometry
$\phi : \mathbb{R}^n \to \mathbb{R}$	$I = \int L dx : S \to \mathbb{R}$	$\lambda = L dx \in \Omega(J^k E)$
$x \in \mathbb{R}^n$ finite dim.	$u \in S = \{u \in C^\infty : ...\}$ ∞-dim.	$\sigma \in \Gamma(E)$ section
v, VF in $\mathbb{R}^n$	φ^α, test function	V, vertical VF
$d\phi_x(v)$ differential	$\delta I(u; \varphi)$ first variation	$\mathcal{L}_{\mathrm{pr}V}\lambda$
d, exterior derivative	δ or $\mathcal{E}_\alpha$, Euler-Lagrange op.	E
$\omega \in \Omega^1(\mathbb{R}^n)$	f_α, differential expression	Δ, source form
$\omega = d\phi$?	$f_\alpha = \mathcal{E}_\alpha L$?	$\Delta = E(\lambda)$?

In the above table, VF denotes vector field.

2.4.1. Trivial Lagrangians

Let us start this subsection with the following question: What kind of functions are in the kernel of the gradient ∇ (or the differential d)? These are the constant functions. What kind of integral functionals are in the kernel of the first variation δ? These are definitely the constant functionals, but are there more than that. The answer is yes and this will be discussed in this subsection. Let us also consider the following question: What kind of Lagrangians are in the kernel of the Euler-Lagrange operator $\mathcal{E}_\alpha$? These are the constant Lagrangians, but are there more than that? The answer is again yes. We find immediately out that local coordinate transformation lead to

$$Ldx = L\frac{\partial x}{\partial y}dy$$

and this means that every Lagrangian $L = L(x)$, where L only depends on x, is in the kernel of the Euler-Lagrange operator (which is also easy to see without doing local coordinate transformation). But this does not completely describe the kernel. Let us first make a definition, before we further discuss this problem.

Definition 2.4.4. *Let $\lambda = Ldx$ be a Lagrange form on J^kE. If $\mathcal{E}_\alpha L = 0$ in some local coordinates and at all points in J^kE, then we call L a* ***trivial Lagrangian*** *(sometimes called null Lagrangian).*

The fact that this definition is independent of the choice of local coordinates can be seen from (2.48) below, $\mathcal{E}_\alpha D_x = 0$ and locally exact sequences later, or one proves it directly when transforming the Euler-Lagrange operator $\mathcal{E}_\alpha$ and L in local coordinates. For the transformation of $\mathcal{E}_\alpha$ see (2.43) and L transforms as $\tilde{L} = \frac{\partial x}{\partial y}L$.

How can we find out more about trivial Lagrangians and how they can be described? According to the relation between Lagrangians L and variational functionals I, we expect that trivial Lagrangians should be described by trivial first variations

$$\delta I(\sigma; V) = \int_{U^0} \mathrm{pr}^k\sigma^*[\mathcal{L}_{\mathrm{pr}V}(Ldx)] = 0, \tag{2.45}$$

where (2.45) is satisfied for all sections σ and all vertical vector fields V on E and $\mathrm{supp}V \subset \pi^{-1}(U^0)$.

<u>What we should be careful about the local diff. geom. on J^kE approach:</u> We would probably assume that trivial Lagrangians or Lagrange forms are described by the equation

$$\mathcal{L}_{\mathrm{pr}V}(Ldx) = 0 \quad \text{or} \quad \mathrm{pr}^k\sigma^*[\mathcal{L}_{\mathrm{pr}V}(Ldx)] = 0 \tag{2.46}$$

for all points in J^kE, or for all sections σ in E, and all projectable vector fields $V \in \mathfrak{X}(E)$, $\mathrm{supp}V \subset \pi^{-1}(U^0)$. But this is not the case. We already see that compact support of V is somehow unnecessary in this local description. Note that the second equation in (2.46) is the weaker condition, since it implies

$$\mathcal{L}_{\mathrm{pr}V}(Ldx) = \text{contact form}.$$

However, the first or the second equation in (2.46) is not equivalent to $\delta I(\sigma;V)=0$, since the integral in (2.45) is missing and it has a non-trivial kernel. If

$$\mathcal{L}_{\mathrm{pr}V}(Ldx)=d\eta$$

for all points in J^kE and all projectable $V\in\mathfrak{X}(E)$, supp$V\subset\pi^{-1}(U^0)$ and for some function η, then $d\eta$ must depend linearly on pr$V$ and this means that (2.45) is satisfied. Now, we would probably assume that the equation

$$\mathcal{L}_{\mathrm{pr}V}(Ldx)=d\eta+\text{contact form} \tag{2.47}$$

describes trivial Lagrangians and this is actually true (at least locally). Later, in Lemma 2.7.1, we will construct a homotopy operator which will completely describe locally trivial Lagrangians. Here instead we now want to discuss further aspects of this problem.

Let Λ be a function on J^kE and we set $L=D_x\Lambda$ as Lagrangian. The corresponding Lagrange form is $\lambda=(D_x\Lambda)dx$. This Lagrange form λ is invariantly defined, since coordinate transformations lead to

$$L=(D_x\Lambda)dx=\frac{\partial y}{\partial x}(D_y\Lambda)\frac{\partial x}{\partial y}dy=(D_y\Lambda)dy. \tag{2.48}$$

See Proposition 2.4.3 for the transformation of D_x and dx. Note that Λ is a different function in different local coordinates. More precisely, we should write

$$\Lambda(x,u,...)=\Lambda(x(y),u(y,v),...)=\tilde{\Lambda}(y,v,v_y,...),$$

and therefore $(D_x\Lambda)dx=(D_y\tilde{\Lambda})dy$. However, when there is no danger of confusion then we simply write $(D_x\Lambda)dx=(D_y\Lambda)dy$.

Let $u(x)\in S$ be a section in local coordinates in an admissible set S in the calculus of variations and $\sigma\in\tilde{S}\subset\Gamma(E)$ the corresponding section on E in the corresponding set $\tilde{S}$ (also see (1.6)). In general,

$$I(\sigma)=\int_{U^0}\mathrm{pr}^k\sigma^*[(D_x\Lambda)dx]=(\mathrm{pr}^k\sigma^*\Lambda)|_{\partial U^0}\neq\text{constant},\quad\text{for all }\sigma\in\tilde{S}, \tag{2.49}$$

since the section σ and corresponding section $u(x)$ in local coordinates and derivatives $u_x(x)$,$u_{xx}(x)$ are not fixed values at the boundary ∂U^0 (for different $u\in S$). As we mentioned in Section 1.3, in the calculus of variations, perturbations are described by test functions (or vertical vector fields) with compact support in some set $U^0\subset M$ (or in the set $(\pi)^{-1}U^0$). But there are no further assumptions on the set S in general. Especially, we do not require fixed boundary conditions for $u\in S$ and derivatives of u. However, for any sections γ and σ in the set $\tilde{S}$, where $\mathrm{pr}^k\gamma|_{\partial U^0}=\mathrm{pr}^k\sigma|_{\partial U^0}$, we get that $I(\gamma)=I(\sigma)$ if the Lagrangian is a total derivative $D_x\Lambda$ (see (2.49)). Furthermore, let $\gamma_t\in\tilde{S}$ be a 1-parameter family of sections, such that $\gamma_{t=0} = \sigma$ and, $\mathrm{pr}^k\gamma_t|_{\partial U^0} = \mathrm{pr}^k\sigma|_{\partial U^0}$ for all $t \in \mathbb{R}$. Then $I(\gamma_t) = I(\sigma)$ for all $t \in \mathbb{R}$, if the Lagrangian is a total derivative and the first variation $\frac{d}{dt}I(\gamma_t)|_{t=0}$ must vanish. Therefore, $D_x\Lambda$ is a trivial Lagrangian. Note that if the derivative of a function $\phi : \mathbb{R} \to \mathbb{R}$ vanishes everywhere, then the function must be constant on $\mathbb{R}$. In contrast here, if the first variation vanishes everywhere on S, then I does not have to be constant on the set S. In this sense, functions ϕ defined on $\mathbb{R}$, and integral functionals I defined on S behave quite differently. We could say that the first variation only allows perturbations in certain directions of S, whereas the derivative (or gradient ∇) of a function ϕ describes perturbations in all directions of $\mathbb{R}$ (or $\mathbb{R}^n$).

Let us give a second explanation why $D_x\Lambda$ is a trivial Lagrangian and let us go back to the expression (2.47). First, we compute (with the help of the standard decomposition)

$$\begin{aligned}
&(D_x\Lambda)dx = \\
=&(\Lambda_x + u_x\Lambda_u + u_{xx}\Lambda_{u_x} + ... + u_{(k+1)}\Lambda_{u_{(k)}})dx = \\
=&\Lambda_x dx + \Lambda_u(u_x dx - du) + \Lambda_{u_x}(u_{xx}dx - du_x) + ... + \Lambda_{u_{(k)}}(u_{(k+1)}dx - du_{(k)}) + \\
&+ (\Lambda_u du + \Lambda_{u_x} du_x + ... + \Lambda_{u_{(k)}} du_{(k)}) = \\
=&d\Lambda - \Lambda_u\Theta^u - \Lambda_{u_x}\Theta^u_x - ... - \Lambda_{u_{(k)}}\Theta^u_{(k)} = \\
=&d\Lambda - \sum_{l=0}^{k} \Lambda_{u_{(l)}}\Theta^u_{(l)}.
\end{aligned}$$

Second, we want to apply the Lie derivative $\mathcal{L}_{\mathrm{pr}V}$ on this special Lagrange form. Note that the Lie derivative $\mathcal{L}_{\mathrm{pr}V}$ commutes with the exterior derivative d, since by Cartan's formula we get

$$\begin{aligned}
\mathcal{L}_{\mathrm{pr}V} d &= (\iota_{\mathrm{pr}V} d + d\iota_{\mathrm{pr}V})d = \\
&= d\iota_{\mathrm{pr}V} d = \\
&= d(\iota_{\mathrm{pr}V} d + d\iota_{\mathrm{pr}V}) = d\mathcal{L}_{\mathrm{pr}V}.
\end{aligned}$$

Furthermore, for projectable vector fields $V \in \mathfrak{X}(E)$, the Lie derivative $\mathcal{L}_{\mathrm{pr}V}$ applied to a contact form will again be a contact form, see Proposition 2.2.8. Therefore,

$$\mathcal{L}_{\mathrm{pr}V}[(D_x\Lambda)dx] = d(\mathcal{L}_{\mathrm{pr}V}\Lambda) + \text{contact form}$$

and this indeed means that $(D_x\Lambda)$ is a trivial Lagrangian, which satisfies (2.45). Using the Euler-Lagrange operator $\mathcal{E}_\alpha$, we get

$$\mathcal{E}_\alpha(D_x\Lambda) = 0$$

for all points in J^kE and for all functions Λ. This is equivalent to saying that the operator identity $\mathcal{E}_\alpha D_x = 0$ holds. It turns out that this exactly describes the kernel of the Euler-Lagrange operator $\mathcal{E}_\alpha$.

Lemma 2.4.5. *Total derivatives are in the kernel of the Euler-Lagrange operator, that is, $\mathcal{E}_\alpha D_x = 0$ and $L = D_x\Lambda$ is a trivial Lagrangian.*

To prove Lemma 2.4.5, we need to prove Lemma 2.4.6 first, which might be anyway of interest in other situations (for example, to prove the Helmholtz conditions later or to show that the Cartan distribution is non-integrable).

Lemma 2.4.6. *We have the following commutator identities for vector fields and differential operators on J^kE*

$$\begin{aligned}
&[\partial_x,\partial_{u^\alpha_{(k)}}] = 0, \quad [\partial_{u^\alpha_{(k)}},\partial_{u^\beta_{(l)}}] = 0, \quad [\partial_x,D_x] = 0, \quad [\partial_{u^\alpha},D_x] = 0,\\
&[\partial_{u^\alpha_x},D_x] = \partial_{u^\alpha}, \quad [\partial_{u^\alpha_{xx}},D_x] = \partial_{u^\alpha_x}, \quad [\partial_{u^\alpha_{(k)}},D_x] = \partial_{u^\alpha_{(k-1)}},\\
&[\partial_{u^\alpha_x},D_x^2] = 2D_x\partial_{u^\alpha}, \quad [\partial_{u^\alpha_{xx}},D_x^2] = \partial_{u^\alpha} + 2D_x\partial_{u^\alpha_x}.
\end{aligned} \tag{2.50}$$

<u>Proof of Lemma 2.4.6:</u> It is clear that the partial derivatives ∂_x and $\partial_{u^\alpha_{(l)}}$ commute for all $0 \le l \le k$. Since the coefficients of D_x do not depend on x or u^α, it is also clear that ∂_x,D_x and ∂_{u^α},D_x commute. That is, we have proven the first line in (2.50). Then we are considering the identities in the second line in (2.50) and we get

$$\begin{aligned}
[\partial_{u^\alpha_x},D_x] &= [\partial_{u^\alpha_x},\, \partial_x + u^\beta_x\partial_{u^\beta} + u^\beta_{xx}\partial_{u^\beta_x} + ... + u^\beta_{(k+1)}\partial_{u^\beta_k}] =\\
&= [\partial_{u^\alpha_x},\, u^\beta_x\partial_{u^\beta}] + \underbrace{[\partial_{u^\alpha_x},\, \partial_x + u^\beta_{xx}\partial_{u^\beta_x} + ... + u^\beta_{(k+1)}\partial_{u^\beta_k}]}_{=0} =\\
&= \partial_{u^\alpha_x}(u^\beta_x\partial_{u^\beta}) - u^\beta_x\partial_{u^\beta}\partial_{u^\alpha_x} = \partial_{u^\alpha}.
\end{aligned}$$

The rest of the identities in the second line in (2.50) are proven in a similar way. For the third line in (2.50) we can already use the identities in the second line and we get

$$\begin{aligned}
[\partial_{u^\alpha_x},D_x^2] &= \partial_{u^\alpha_x}D_xD_x - D_x^2\partial_{u^\alpha_x} =\\
&= (\partial_{u^\alpha} + D_x\partial_{u^\alpha_x})D_x - D_x^2\partial_{u^\alpha_x} =\\
&= \partial_{u^\alpha}D_x + D_x(\partial_{u^\alpha} + D_x\partial_{u^\alpha_x}) - D_x^2\partial_{u^\alpha_x} = 2D_x\partial_{u^\alpha}
\end{aligned}$$

and

$$\begin{aligned}
[\partial_{u^\alpha_{xx}},D_x^2] &= \partial_{u^\alpha_{xx}}D_x^2 - D_x^2\partial_{u^\alpha_{xx}} =\\
&= (\partial_{u^\alpha_x} + D_x\partial_{u^\alpha_{xx}})D_x - D_x^2\partial_{u^\alpha_{xx}} =\\
&= \partial_{u^\alpha_x}D_x + D_x(\partial_{u^\alpha_x} + D_x\partial_{u^\alpha_{xx}}) - D_x^2\partial_{u^\alpha_{xx}} =\\
&= \partial_{u^\alpha_x}D_x + D_x\partial_{u^\alpha_x} =\\
&= (\partial_{u^\alpha} + D_x\partial_{u^\alpha_x}) + D_x\partial_{u^\alpha_x} = \partial_{u^\alpha} + 2D_x\partial_{u^\alpha_x}.
\end{aligned}$$

Note that more general identities, for example $[\partial_{u^\alpha_{(r)}}, D^s_x]$ for arbitrary r,s, are proven in a similar way and by induction. $\square$

Proof of Lemma 2.4.5: We use the commutator identities in Lemma 2.4.6 and we want to bring the operator D_x to the left side of all expressions. Then we get a telescoping sum of the following form

$$\begin{aligned}\mathcal{E}_\alpha D_x &= [\partial_{u^\alpha} - D_x\partial_{u^\alpha_x} + D_x^2\partial_{u^\alpha_{xx}} \pm ... + (-1)^k D_x^k \partial_{u^\alpha_{(k)}}]D_x = \\ &= D_x\partial_{u^\alpha} - D_x(\partial_{u^\alpha} + D_x\partial_{u^\alpha_x}) + D_x^2(\partial_{u^\alpha_x} + D_x\partial_{u^\alpha_{xx}}) \pm ...+ \\ &\qquad + (-1)^k D_x^k(\partial_{u^\alpha_{(k-1)}} + D_x\partial_{u^\alpha_{(k)}}) = \\ &= (-1)^k D^{k+1}\partial_{u^\alpha_{(k)}} = 0\end{aligned}$$

The last line vanishes, since the operator $\mathcal{E}_\alpha$ is applied to Lagrangians of order k and therefore Λ must be of order $(k-1)$, since D_x increases the order by one. $\square$

Similar telescoping sums also occur in the proof of the Helmholtz conditions later.

Example: The equation $f = u_{xx} = 0$ is variational and usually one uses $L = -\frac{1}{2}u_x^2$ as Lagrangian. The Lagrangians

$$\begin{aligned}L &= -\frac{1}{2}u_x^2, \\ \tilde{L} &= -\frac{1}{2}u_x^2 + u_x u_{xx}\end{aligned}$$

are equivalent and lead to the same differential equation, since the term $u_x u_{xx} = \frac{1}{2}D_x(u_x)^2$ is a total derivative.

In general, any two Lagrangians $L \sim \tilde{L}$ if $L - \tilde{L} = D_x\Lambda$ for some function Λ. This shows that it is reasonable to consider equivalence classes of Lagrangians denoted by $[L]$. In a more general theory, one uses Lagrange forms $\lambda = Ldx + d\Lambda +$ (1-contact) and quotient mappings, where we do not have such trivial Lagrangians, since the trivial ones are considered to be zero in the quotient space and we get $[\lambda] = [Ldx] = [Ldx + d\Lambda + \text{(1-contact)}]$. This shows once again that it is reasonable to construct such quotient spaces. Let us consider trivial Lagrange forms (and use the standard decomposition)

$$\begin{aligned}\lambda &= (D_x\Lambda)dx = \\ &= (\Lambda_x + u_x\Lambda_u + u_{xx}\Lambda_{u_x} + ... + u_{(k+1)}\Lambda_{u_k})dx = \\ &= \Lambda_x dx + (-\Lambda_u\Theta - \Lambda_{u_x}\Theta_x - ... - \Lambda_{u_{(k)}}\Theta_{(k)}) + (\Lambda_u du + ... + \Lambda_{u_{(k)}}du_{(k)}) = \\ &= d\Lambda + \text{(1-contact)}.\end{aligned}$$

We observe that $(D_x\Lambda)dx$ and $d\Lambda$+1-contact are in the same trivial equivalence class.

Let us explain the variational sequence of quotient spaces a bit more detailed. First, the naive de-Rham sequence in $\mathbb{R}^3$ can be written as

$$0 \to \mathbb{R} \to C^\infty(\mathbb{R}^3) \xrightarrow{\nabla} \mathfrak{X}(\mathbb{R}^3) \xrightarrow{\text{curl}} \mathfrak{X}(\mathbb{R}^3) \xrightarrow{\text{div}} C^\infty(\mathbb{R}^3) \to 0.$$

The actual de-Rham sequence in $\mathbb{R}^3$ is

$$0 \to \mathbb{R} \to C^\infty(\mathbb{R}^3) \xrightarrow{d} \Omega^1(\mathbb{R}^3) \xrightarrow{d} \Omega^2(\mathbb{R}^3) \xrightarrow{d} \Omega^3(\mathbb{R}^3) \to 0.$$

In analogy, in the calculus of variations, we get the naive variational sequence

$$0 \to \mathbb{R} \to C^\infty(JE) \xrightarrow{D_x} \{L\} \xrightarrow{\mathcal{E}_\alpha} \{f_\alpha\} \longrightarrow ...$$

and the actual variational sequence (for $n = 1$) is the sequence of quotient spaces

$$0 \to \mathbb{R} \to C^\infty(JE) \xrightarrow{d} \Omega^1(JE)/(dC^\infty + 1\text{-contact}) \xrightarrow{d} \Omega^2/... \longrightarrow ...,$$

where we wrote JE instead of J^kE (but actually the sequence is defined on the finite jet space). Note that instead of d one usually writes E^l for these mappings. Further information can be found in (Ku04, Kru97a). Also note that a similar sequence can be defined for the infinite jet space, but there without quotient spaces, see (And89).

2.5. The Order of Jet Coordinates, Part I

To understand how we can use the order of jet coordinates is very important when we want to solve Takens' problem and it can also be used in other situations. In this section, we discuss two examples, where the order of jet coordinates can be used. These examples connect with the previous section, where we investigated variational source forms and trivial Lagrangians.

Not every differential equation is variational. For example,

$$f = u_x = 0$$

cannot be written as an Euler-Lagrange equation and here is the idea of the proof: Let us assume there is a first order Lagrangian $L = L(x,u,u_x)$ such that

$$\begin{aligned} f &= L_u - D_x L_{u_x} = \\ &= L_u - L_{xu_x} - u_x L_{uu_x} - u_{xx} L_{u_x u_x} \stackrel{!}{=} u_x. \end{aligned} \tag{2.51}$$

Equation (2.51) must hold for every point (x,u,u_x,u_{xx}) in J^2E. Since L and therefore L_u, L_{xu_x}, L_{uu_x}, $L_{u_xu_x}$ only depend on (x,u,u_x), and on the right hand side of (2.51) there does not occur the u_{xx}-coordinate, the term

$$L_{u_xu_x} = 0$$

must vanish. The vanishing of this term forces $L = A(x,u) + u_x B(x,u)$ for some functions A,B, simply by integrating the condition two times. Now we plug this L again in equation (2.51) and we get

$$\begin{aligned} f = L_u - D_x L_{u_x} &= L_u - L_{xu_x} - u_x L_{uu_x} = \\ &= A_u + u_x B_u - B_x - u_x B_u = \\ &= A_u - B_x \overset{!}{=} u_x. \end{aligned} \tag{2.52}$$

Since the left hand side of (2.52) does not depend on u_x, this equation cannot be satisfied, and therefore the equation is not variational. If we want to do a complete proof, we also have to show that there are no higher order Lagrangians which satisfy (2.51) and this is not very complicated (one can do it as an exercise).

The main feature in the proof was basically to discuss the coordinates (u_x,u_{xx}) in equation (2.51). That is, we discussed that there cannot be second order u_{xx}-coordinates in this equation and later that there cannot be u_x-coordinates. We call these kinds of discussions the order discussion and it is also a very fundamental technique when solving Takens' problem. The fact that $f = u_x$ is not variational can also be proven in a some sense different way, namely, to show that $f = u_x$ does not satisfy the Helmholtz conditions which will be shown later.

We want to present another very simple example, where the discussion of order can be used. Not every Lagrangian is a trivial Lagrangian. For example, $L = u_x^2$ is not a trivial Lagrangian. In this small paragraph, we assume that every trivial Lagrangian can be written as a total derivative $D_x\Lambda$ for some function Λ (the exactness of the variational sequence will be proven later, we could also say that not every function is a total derivative). The first way how to prove this is that $\mathcal{E}u_x^2 \neq 0$. The second way is that we consider the equation

$$L = u_x^2 = D_x\Lambda$$

and we show that it has no solution. Let Λ be of order k then

$$L = u_x^2 \overset{!}{=} D_x\Lambda = \Lambda_x + u_x\Lambda_u + u_{xx}\Lambda_{u_x} + ... + u_{(k+1)}\Lambda_{u_{(k)}}. \tag{2.53}$$

Since the left hand side in (2.53) does not depend on the $u_{(k+1)}$-coordinate, it must vanish on the right hand side and this means

$$\Lambda_{u_{(k)}} = 0.$$

This equation tells us that Λ is not of order k, but of order $k-1$. Then we can repeat the argument until we get Λ is of order zero, i.e. $\Lambda = \Lambda(x,u)$. Again, we plug this Λ into (2.53) and we get

$$L = u_x^2 \overset{!}{=} D_x\Lambda = \Lambda_x(x,u) + u_x\Lambda_u(x,u),$$

but this equation cannot be satisfied, since the right hand side is affine linear in u_x, but the left hand side depends quadratically on u_x. Therefore, $L = u_x^2$ cannot be written as a total derivative and therefore is not a trivial Lagrangian.

In Takens' problem, we will use this sort of technique in a more complicated situation, but the idea is the same as here. Also integrating simple differential equations, like $\Lambda_{u_{(k)}} = 0$, or $L_{u_x u_x} = 0$, get a result, and plug the result again into the initial equation is used, when solving Takens' problem.

2.6. Integrability- and Helmholtz Conditions

Let us start this section with the following question: Is there a general condition which shows if an equation, or better a source form, is variational, or is not variational? We did the calculation above for the example $f = u_x$, but here we want to find a second approach, in a more general situation. We already explained the correspondence between variational equations in infinite dimensional spaces and vector fields in $\mathbb{R}^n$, which can be written as gradients, or 1-forms, which are (locally) exact. Now this will be investigated in more detail and we will find a general condition which gives the answer to the above question.

In the first part of this section (in Subsection 2.6.1), we motivate the Helmholtz conditions and we also deliver a proof why these conditions have to be satisfied for variational source forms. In the second part (in Subsection 2.6.2), we deliver an alternative proof which is more or less a straightforward calculation without using all the ideas from Subsection 2.6.1.[2]

2.6.1. Motivation and Proof of the Helmholtz Conditions

Again, let us first start with the finite dimensional analog in $\mathbb{R}^n$. More precisely, let us consider a manifold N of dimension n. Let $\omega \in \Omega^1(N)$ and we want to find out: Under which conditions is is ω exact, that is, $\omega = d\phi$ for some function ϕ? A necessary and locally sufficient condition is $d\omega = 0$. Let us write ω as

$$\omega = \omega_1 dx^1 + \omega_2 dx^2 + ... + \omega_n dx^n,$$

then $d\omega = 0$ is equivalent to

$$\frac{\partial \omega_k}{\partial x^l} - \frac{\partial \omega_l}{\partial x^k} = 0, \quad \text{for all } k,l = 1,2,...,n \tag{2.54}$$

[2] As I found out later, the approach in Subsection 2.6.1 is probably the best way to prove the Helmholtz conditions, since the method which is explained there can also be applied for PDE's of higher order, whereas the calculations in Subsection 2.6.2 can get very complicated.

and usually these are known as the **integrability conditions**. On infinite dimensional manifolds, like in the calculus of variations, we do not have an exterior derivative operator d yet. But we have the concept of Lie derivative and interior product and the Lie derivative is connected with d through Cartan's formula

$$\mathcal{L}_V = \iota_V d + d\iota_V.$$

We can use this identity to construct an operator d. To construct such an operator d on infinite dimensional spaces, we need the following characterization of d:

Proposition 2.6.1. *Let $\omega \in \Omega^1(N)$ and V,W vector fields on N. Then:*
i) $\mathcal{L}_V(\iota_W\omega) = \iota_W\mathcal{L}_V\omega + \iota_{[V,W]}\omega$.
ii) $\iota_W\iota_V d\omega = (\mathcal{L}_V\iota_W - \mathcal{L}_W\iota_V)\omega - \iota_{[V,W]}\omega$.
iii) $\iota_W\iota_V d\omega = (\mathcal{L}_V\iota_W - \iota_V\mathcal{L}_W)\omega$.

For the reader who is more familiar with vector fields than differential forms: The use of the identities in Proposition 2.6.1 can be understood very easily, when rewriting (2.54) slightly differently. Let us consider $N = \mathbb{R}^n$ for simplicity and $\tilde{\omega}$ is a vector field on N, i.e.

$$\tilde{\omega} = \tilde{\omega}_1 e_1 + \tilde{\omega}_2 e_2 + ... + \tilde{\omega}_n e_n,$$

where e_k are the canonical unit vectors (vector fields) in $\mathbb{R}^n$, and $\tilde{\omega}_k$ are the coefficients of the vector field. We can use the Euclidian scalar product and we can write

$$\begin{aligned} 0 = \frac{\partial\tilde{\omega}_k}{\partial x^l} - \frac{\partial\tilde{\omega}_l}{\partial x^k} &= \partial_{x^l} < \tilde{\omega},e_k > -\partial_{x^k} < \tilde{\omega},e_l >= \\ &=< \partial_{x^l}\tilde{\omega},e_k > - < \partial_{x^k}\tilde{\omega},e_l >= \\ &=< \partial_{x^l}\tilde{\omega},e_k > -\partial_{x^k} < \tilde{\omega},e_l >= \\ &= \mathcal{L}_{e_l} < \tilde{\omega},e_k > -\mathcal{L}_{e_k} < \tilde{\omega},e_l > , \quad \text{for all } k,l = 1,2,...,n, \end{aligned}$$

where the commutator $[e_k,e_l] = 0$, for constant vector fields e_k,e_l. Note that $\tilde{\omega}$ is the corresponding vector field to ω, that is, $\tilde{\omega} = \omega^\sharp$, where $\sharp$ is the sharp isomorphism. The general formula for non-commuting vector fields is

$$0 = \mathcal{L}_V < \tilde{\omega},W > -\mathcal{L}_W < \tilde{\omega},V > - < \tilde{\omega},[V,W] > , \quad \forall \text{ vector fields } V,W, \tag{2.55}$$

which can be easily checked and this is basically one of the identities in Proposition 2.6.1. More generally, instead of the Euclidian scalar product in $\mathbb{R}^n$, we have the interior product ι.

Proof of Proposition 2.6.1: Statement i) follows by direct computation in local coordinates and one can use Cartan's formula $\mathcal{L}_V = \iota_V d + d\iota_V$ for 1-forms. We only prove it for dim$N = 2$, where $\omega = \omega_i dx^i$, $V = V^i\partial_{x^i}$, $W = W^i\partial_{x^i}$. We get

$$\begin{aligned} \mathcal{L}_V(\iota_W\omega) &= \mathcal{L}_V(\omega_i W^i) = \\ &= (\mathcal{L}_V\omega_i)W^i + \omega_i(\mathcal{L}_V W^i) \end{aligned} \tag{2.56}$$

and on the other hand

$$
\begin{aligned}
&\iota_W \mathcal{L}_V \omega = \iota_W(\iota_V d + d\iota_V)\omega = \\
=&\iota_W[\iota_V(\omega_{1,x^2} dx^2 \wedge dx^1 + \omega_{2,x^1} dx^1 \wedge dx^2) + d(\omega_i V^i)] = \\
=&\iota_W[\iota_V(\omega_{2,x^1} - \omega_{1,x^2}) dx^1 \wedge dx^2] + \mathcal{L}_W(\omega_i V^i) = \\
=&\iota_W[(\omega_{2,x^1} - \omega_{1,x^2})(V^1 dx^2 - V^2 dx^1)] + \mathcal{L}_W(\omega_i V^i) = \\
=&(\omega_{2,x^1} - \omega_{1,x^2})(V^1 W^2 - V^2 W^1) + \mathcal{L}_W(\omega_1 V^1 + \omega_2 V^2) = \\
=&\omega_{2,x^1} V^1 W^2 - \omega_{2,x^1} V^2 W^1 - \omega_{1,x^2} V^1 W^2 + \omega_{1,x^2} V^2 W^1 + \\
+&\omega_{1,x^1} W^1 V^1 + \omega_{1,x^2} W^2 V^1 + \omega_{2,x^1} W^1 V^2 + \omega_{2,x^2} W^2 V^2 + \omega_1 \mathcal{L}_W V^1 + \omega_2 \mathcal{L}_W V^2 = \\
=&\omega_{2,x^1} V^1 W^2 + \omega_{1,x^2} V^2 W^1 + \\
+&\omega_{1,x^1} W^1 V^1 + \omega_{2,x^2} W^2 V^2 + \omega_1 \mathcal{L}_W V^1 + \omega_2 \mathcal{L}_W V^2 = \\
=&W^2 \mathcal{L}_V \omega_2 + W^1 \mathcal{L}_V \omega_1 + \omega_1 \mathcal{L}_W V^1 + \omega_2 \mathcal{L}_W V^2 = \\
=&W^i \mathcal{L}_V \omega_i + \omega_i \mathcal{L}_W V^i. \qquad (2.57)
\end{aligned}
$$

When we add the term

$$
\begin{aligned}
\iota_{[V,W]}\omega =&\iota_{(\mathcal{L}_V W^i)\partial_{x^i} - (\mathcal{L}_W V^i)\partial_{x^i}} \omega_i dx^i = \\
=&(\mathcal{L}_V W^i)\omega_i - (\mathcal{L}_W V^i)\omega_i
\end{aligned}
$$

to (2.57) then we get the expression (2.56). The statement *ii*) follows by *i*) and Cartan's formula, i.e.

$$
\begin{aligned}
\iota_W \iota_V d\omega &= \iota_W(\mathcal{L}_V - d\iota_V)\omega = \\
&= \mathcal{L}_V(\iota_W \omega) - \iota_{[V,W]} - \iota_W d(\iota_V \omega) = \\
&= \mathcal{L}_V(\iota_W \omega) - \iota_{[V,W]} - \mathcal{L}_W(\iota_V \omega).
\end{aligned}
$$

The statement *iii*) follows by *i*) and *ii*), i.e.

$$
\begin{aligned}
(\mathcal{L}_V \iota_W - \mathcal{L}_W \iota_V)\omega - \iota_{[V,W]}\omega &= [\mathcal{L}_V \iota_W - (\iota_V \mathcal{L}_W + \iota_{[W,V]})]\omega - \iota_{[V,W]}\omega = \\
&= (\mathcal{L}_V \iota_W - \iota_V \mathcal{L}_W)\omega.
\end{aligned}
$$

Therefore, we have proven Proposition 2.6.1. □

We also need the following proposition:

Proposition 2.6.2. *Let $\omega \in \Omega^1(N)$. The following statements are equivalent:*
i) $d\omega = 0$.
ii) $\iota_W \iota_V d\omega = 0$ *for all* $V, W \in \mathfrak{X}(N)$.
iii) $(\mathcal{L}_V \iota_W - \mathcal{L}_W \iota_V)\omega - \iota_{[V,W]}\omega = 0$ *for all* $V, W \in \mathfrak{X}(N)$.
iv) $(\mathcal{L}_V \iota_W - \iota_V \mathcal{L}_W)\omega = 0$ *for all* $V, W \in \mathfrak{X}(N)$.

Proof: By definition of differential forms, the equivalence in $i)$ and $ii)$ is clear. The rest follows by Proposition 2.6.1. $\square$

Now we want to investigate the infinite dimensional analog in the calculus of variations. Proposition 2.6.1 and 2.6.2 will be needed there. In the infinite dimensional case, we have the concept of flow and Lie derivative. This allows us to find similar conditions, as the integrability conditions, without having the operator d, or we could even define an operator d if we want. Further information can be found in (And89, p.68).

The infinite dimensional analog in calculus of variations: To be able to formulate the integrability conditions for the infinite dimensional case, we need a definition for Lie derivative of integral functionals. The Lie derivative uses the concept of flow of a vector field. Since integral functionals are integrals over sets $U^0 \subset M$, we have to define how to transform such sets by flows. In the following, $U^0 \subset M$ is always an open set such that the closure $\bar{U}^0 \subset M$ is compact. Compact support in U^0, or $\pi^{-1}(U^0)$, means for vector fields $V = V^x\partial_x \in \mathfrak{X}(M)$, or $V = V^x\partial_x + V^\alpha\partial_{u^\alpha} \in \mathfrak{X}(E)$, that V and all derivatives of V^x, or (V^x,V^α), must vanish at ∂U^0, or $\pi^{-1}(\partial U^0)$ (note that the vector fields must be smooth on M or E).

Definition 2.6.3. *Let V be a projectable vector field on E and ϕ_t the corresponding flow. Furthermore, let ϕ_t^0 be the flow of π_*V. Then we define $\phi_t^0 U^0 := \{q \in M : q = \phi_t^0(\tilde{q}), for\ all\ \tilde{q} \in U^0\}$.*

Note that the first variation $\delta I(\sigma; V)$ can be considered as a Lie derivative of I with respect to the (vertical) vector field V at a point σ and it is reasonable to write it as $\mathcal{L}_{\mathrm{pr}V} I$. The more general definition for projectable vector fields is the following:

Definition 2.6.4. *Let $I = \int Ldx$ be an integral functional, $\sigma \in \Gamma(E)$ a section and $V \in \mathfrak{X}(E)$ a projectable vector field. Furthermore, let ϕ_t be the flow of V, ϕ_t^0 the flow of π_*V and $U^0 \subset M$ (the flow is always defined for small $t \in \mathbb{R}$). Then we define*

$$\phi_t[I(\sigma)] = (\phi_t I)(\phi_t \circ \sigma \circ \phi_{-t}^0) := \int_{\phi_t^0 U^0} pr^k(\phi_t \circ \sigma \circ \phi_{-t}^0)^*(Ldx),$$

$$(\mathcal{L}_{prV} I)(\sigma) := \frac{d}{dt}\int_{\phi_t^0 U^0} pr^k(\phi_t \circ \sigma \circ \phi_{-t}^0)^*(Ldx)|_{t=0}. \qquad (2.58)$$

Note that (2.58) can be written differently as

$$
\begin{aligned}
&(\mathcal{L}_{\mathrm{pr}V} I)(\sigma) = \\
=&\frac{d}{dt}\int_{\phi_t^0 U^0} \mathrm{pr}^k(\phi_t \circ \sigma \circ \phi_{-t}^0)^*(Ldx)|_{t=0} = \frac{d}{dt}\int_{U^0} \mathrm{pr}^k(\phi_t \circ \sigma)^*(Ldx)|_{t=0} = \\
=&\int_{U^0} \frac{d}{dt}\mathrm{pr}^k(\phi_t \circ \sigma)^*(Ldx)|_{t=0} = \int_{U^0} \mathrm{pr}^k \sigma^*[\mathcal{L}_{\mathrm{pr}V}(Ldx)].
\end{aligned}
$$

Definition 2.6.4 and this calculation can be found in (Kru97b, p.42).

Note that this definition makes sense for the following reason: We think of the integral as a finite sum of functions depending on (different) points in $J^k E$ and all of these points are transformed by ϕ_t. Then we do the limit process of the sum to get the integral. It is basically the only way how to define the transformation ϕ_t applied to an integral functional in a natural way.

In the case of vertical vector fields $V \in \mathfrak{X}(E)$, $\mathrm{supp}V \subset \pi^{-1}(U^0)$, this definition delivers exactly the first variation. We can also define the symmetry of a weak formulation. Let V be a vertical vector field on E, $\mathrm{supp}V \subset \pi^{-1}(U^0)$, and W a projectable vector field on E, then:

$$
\begin{aligned}
\mathcal{L}_{\mathrm{pr}V} I(\sigma) &= \delta I(\sigma; V), \quad \text{(first variation)} \\
\mathcal{L}_{\mathrm{pr}W} \int_{U^0} \mathrm{pr}^k \sigma^*(f_\alpha V^\alpha dx) &= \int_{U^0} \mathrm{pr}^k \sigma^*[\mathcal{L}_{\mathrm{pr}W}(f_\alpha V^\alpha dx)], \quad \text{(sym. of weak formulation).}
\end{aligned}
$$

Our goal is to find the integrability conditions for variational equations, where we now want to continue. The ∞-dimensional analog of (2.54) or (2.55) is

$$
\mathcal{L}_{\mathrm{pr}V} < f_\alpha, W^\alpha >_{L^2(U^0)} - \mathcal{L}_{\mathrm{pr}W} < f_\alpha, V^\alpha >_{L^2(U^0)} - < f_\alpha, [V,W]^\alpha >_{L^2(U^0)} = 0, \quad (2.59)
$$

where V,W are vertical vector fields on E and $\mathrm{supp}V,W \subset \pi^{-1}(U^0)$. Here $< .,. >_{L^2}$ denotes the scalar product of L^2-functions. Note that the L^2-scalar product implies that we have to pull-back all expressions by sections, otherwise integration does not make sense, but we did not write the pull-back, to keep the notation simple and to see the main structure.

Lemma 2.6.5. *If the source form $f_\alpha du^\alpha \wedge dx$ is variational then* (2.59) *must be satisfied.*

Proof: For variational f_α there exists a functional I, such that

$$
< f_\alpha, V^\alpha >_{L^2(U^0)} = \mathcal{L}_{\mathrm{pr}V} I, \quad \text{for all vertical } V \in \mathfrak{X}(E),\ \mathrm{supp}V \subset \pi^{-1}(U^0).
$$

Let W be another vertical vector field on E and $\mathrm{supp}W \subset \pi^{-1}(U^0)$. Then we get

$$
\mathcal{L}_{\mathrm{pr}V}(\mathcal{L}_{\mathrm{pr}W} I) - \mathcal{L}_{\mathrm{pr}W}(\mathcal{L}_{\mathrm{pr}V} I) - \mathcal{L}_{\mathrm{pr}[V,W]} I = 0.
$$

This follows by definition of Lie derivative applied to such integral functionals. □

With the help of (2.59), we can also define a sort of exterior derivative d for weak formulations through the equation

$$\iota_{\mathrm{pr}V}\iota_{\mathrm{pr}W}dK := \mathcal{L}_{\mathrm{pr}V} < f_\alpha,W^\alpha >_{L^2} -\mathcal{L}_{\mathrm{pr}W} < f_\alpha,V^\alpha >_{L^2} - < f_\alpha,[V,W]^\alpha >_{L^2},$$

for all vertical vector fields V,W on E, supp$V$,$W \subset \pi^{-1}(U^0)$. Furthermore $K = K_\Delta$ is the operator defined through

$$\iota_{\mathrm{pr}V}K := \int_{U^0} \mathrm{pr}^k\sigma(f_\alpha V^\alpha dx) = \int_{U^0} \mathrm{pr}^k\sigma^*[\iota_{\mathrm{pr}V}(f_\alpha du^\alpha \wedge dx)],$$

where $V \in \mathfrak{X}(E)$ is vertical, supp$V \subset \pi^{-1}(U^0)$. If the weak formulation K is variational then $dK = 0$. Note that if K is variational then dK is not the second variation $\delta^2 I$.

In the next paragraph we want to answer the following question: How can we find conditions for f_α, or the corresponding source form, to be variational, without writing the L^2-integral in (2.59). That is, we want to find pure local conditions on J^kE.

<u>What we should be careful about the local diff. geom. on J^kE approach:</u> For the source form we would probably expect that the condition

$$\mathcal{L}_{\mathrm{pr}V}(\iota_{\mathrm{pr}W}\Delta) - \iota_{\mathrm{pr}V}(\mathcal{L}_{\mathrm{pr}W}\Delta) = 0 \quad \forall \text{ vertical } V,W \text{ on E, supp}V,W \subset \pi^{-1}(U^0) \tag{2.60}$$

is the right analog to (2.59) (also in the general case of non-commuting vector fields, see Proposition 2.6.2). But this is not true, since the integral from the L^2-scalar product in (2.59) has a non-trivial kernel, as well as the pull-back by a section. We had a similar discussion at the beginning of Section 2.4.

More precisely, any form which can be written as $\mathrm{pr}\,\sigma^*(d\eta)$, where η depends linearly on pr$V$,pr$W$ vanishes when integrated over U^0, since V,W are assumed to have support $\pi^{-1}(U^0)$. Therefore, we get in any case the weaker condition

$$\mathrm{pr}\,\sigma^*[\mathcal{L}_{\mathrm{pr}V}(\iota_{\mathrm{pr}W}\Delta) - \iota_{\mathrm{pr}V}(\mathcal{L}_{\mathrm{pr}W}\Delta)] = \mathrm{pr}\,\sigma^*(d\eta) \quad \forall \text{ vertical } V,W \text{ on } E, \quad \text{supp}V,W \subset \pi^{-1}(U^0). \tag{2.61}$$

Note that for $n = 1$, we do not have to require that η depends linearly on pr$V$,pr$W$ in (2.61), since this is clear, at least up to constants, form the left hand side in this equation, and constants vanish when d is applied. For $n \geq 2$ we have a similar effect and the kernel of d has to be considered.

To see that the condition (2.60) cannot be the right one, we consider the following simple example of a variational expression $f = u_{xx}$ and corresponding source form

$\Delta = u_{xx} du \wedge dx$. Then we compute

$$\begin{aligned}
&\mathcal{L}_{\mathrm{pr}V}(\iota_{\mathrm{pr}W}\Delta) - \iota_{\mathrm{pr}V}(\mathcal{L}_{\mathrm{pr}W}\Delta) = \\
&= \mathcal{L}_{\mathrm{pr}V}(u_{xx}W^u dx) - \iota_{\mathrm{pr}V}[(D_x^2 W^u) du \wedge dx + u_{xx} W_u^u du \wedge dx] = \\
&= [(D_x^2 V^u) W^u + u_{xx} W_u^u V^u] dx - [(D_x^2 W^u) V^u + u_{xx} W_u^u V^u] dx = \\
&= [(D_x^2 V^u) W^u - (D_x^2 W^u) V^u] dx = \\
&= D_x[(D_x V^u) W^u - (D_x W^u) V^u] dx
\end{aligned}$$

and this form is non-zero in general, but Δ is variational. However, it can be written as $\mathrm{pr}\,\sigma^*(d\eta)$, since (using the standard decomposition)

$$\begin{aligned}
&D_x[(D_x V^u) W^u - (D_x W^u) V^u] dx = D_x \underbrace{[\iota_{\mathrm{pr}W}\iota_{\mathrm{pr}V}(du_x \wedge du)]}_{=:A} dx = \\
&= (A_x + u_x A_u + u_{xx} A_{u_x}) dx = \\
&= A_x dx + A_u(u_x dx - du) + A_{u_x}(u_{xx} dx - du_x) + A_u du + A_{u_x} du_x = \\
&= dA + A_u \Theta^u + A_{u_x} \Theta_x^u
\end{aligned}$$

and when applying $\mathrm{pr}\,\sigma^*$ to contact forms they vanish and we get the expression $\mathrm{pr}\,\sigma^*(d\eta)$, where $\eta = A$. We also get that A and also dA depend linearly on $\mathrm{pr}V$ and $\mathrm{pr}W$.

Now we give some more ideas and explanations which lead to the mappings in the variational sequence. However, we cannot discuss all the details here and we focus more on the main ideas. Further information can be found in (Kru97a, Ku04). Let us assume that (2.61) holds for every $U^0 \subset M$, since we are considering pure local conditions on $J^k E$, where the precise set U^0 should not be important. Then without pull-backing the form by a section, we get

$$\iota_{\mathrm{pr}V}\iota_{\mathrm{pr}W} d\Delta = \mathcal{L}_{\mathrm{pr}V}(\iota_{\mathrm{pr}W}\Delta) - \iota_{\mathrm{pr}V}(\mathcal{L}_{\mathrm{pr}W}\Delta) = (d\eta)_{\mathrm{pr}V,\mathrm{pr}W} + (\text{1-contact})_{\mathrm{pr}V,\mathrm{pr}W}$$

for all vertical V,W on E, for variational source forms. Omitting the vector fields $\mathrm{pr}V$,$\mathrm{pr}W$, this equation is equivalent to

$$d\Delta = d(\text{2-contact}) + (\text{3-contact}). \tag{2.62}$$

We only explain briefly why (2.62) should be correct. We consider

$$\iota_{\mathrm{pr}V}\iota_{\mathrm{pr}W} d\Delta = \iota_{\mathrm{pr}V}\iota_{\mathrm{pr}W}[d(\text{2-contact}) + (\text{3-contact})]$$

and we want to find out how we can rewrite the right hand side. It is clear that $\iota_{\mathrm{pr}V}\iota_{\mathrm{pr}W}(\text{3-contact}) = (\text{1-contact})_{\mathrm{pr}V,\mathrm{pr}W}$. Next, we use Cartan's formula and we rewrite

$$\begin{aligned}
\iota_{\mathrm{pr}V}\iota_{\mathrm{pr}W} d(\text{2-contact}) &= \iota_{\mathrm{pr}V}[(\mathcal{L}_{\mathrm{pr}W} - d\iota_{\mathrm{pr}W})(\text{2-contact})] = \\
&= \iota_{\mathrm{pr}V}[(\text{2-contact})_{\mathrm{pr}W} - d(\text{1-contact})_{\mathrm{pr}W}] = \\
&= (\text{1-contact})_{\mathrm{pr}V,\mathrm{pr}W} - (\mathcal{L}_{\mathrm{pr}V} - d\iota_{\mathrm{pr}V})(\text{1-contact})_{\mathrm{pr}W} = \\
&= (\text{1-contact})_{\mathrm{pr}V,\mathrm{pr}W} + (d\eta)_{\mathrm{pr}V,\mathrm{pr}W},
\end{aligned}$$

where we used Proposition 2.2.8, that is, $\mathcal{L}$ contact = contact. If one wants to do a strict proof of (2.62), then one has to do the calculation in local coordinates, which is also not very complicated and one will probably need the identity $d\Theta^\alpha_{(l)} = -\Theta^\alpha_{(l+1)} \wedge dx$.

In the quotient spaces of the variational sequence ($n = 1$ for ODEs), where

$$[\Delta] \in \Omega^{n+1}(J^k E)/[d(\text{1-contact}) + (\text{2-contact})],$$

we get that $[\Delta]$ is closed, i.e.

$$[d\Delta] \in \Omega^{n+2}(J^k E)/[d(\text{2-contact}) + (\text{3-contact})]$$

is zero in the equivalence class if it is of the form (2.62) and then it can be shown that it is locally exact. We will not prove it here, for further details see (Kru97a, Ku04). The operator d between these quotient spaces is called (generalized) Euler operator and usually written as E^{n+1},E^{n+2} and so on (to distinguish from the fiber bundle E). It is defined as $E^l([\omega]) = [d\omega]$ for every differential form ω. As we mentioned earlier, there is another method, using interior Euler operators and the exterior derivatives d_H,d_V, see (And89).

In the following, we want to develop further methods how to formulate integrability conditions for source forms. In this case, these conditions are called Helmholtz conditions. To derive them needs some straight forward, but computational heavy work. But we need these conditions for at least two reasons:

- We need to construct a homotopy operator later and there we will need these conditions.
- We need them to solve Takens' problem.

In the next subsection, we will prove the Helmohltz conditions in detail and here we have to understand the idea of the problem. To derive these conditions in local coordinates, we use equation (2.59). Let

$$\begin{aligned} V &= V^\alpha \partial_{u^\alpha} \quad &\in \mathfrak{X}(E), \\ W &= W^\alpha \partial_{u^\alpha} \quad &\in \mathfrak{X}(E), \end{aligned}$$

be vertical vector fields on E and supp$V$,$W \subset \pi^{-1}(U^0)$. Then

$$\begin{aligned} 0 &= \mathcal{L}_{\text{pr}V} < f_\alpha, W^\alpha >_{L^2} - \mathcal{L}_{\text{pr}W} < f_\alpha, V^\alpha >_{L^2} - < f_\alpha, [V,W]^\alpha >_{L^2} = \\ &= \int_{U^0} \text{pr}\, \sigma^*[W^\alpha(\mathcal{L}_{\text{pr}V} f_\alpha) - (\mathcal{L}_{\text{pr}W} f_\alpha) V^\alpha] dx. \end{aligned} \tag{2.63}$$

Now we want to investigate equation (2.63) in more detail. More precisely, we want to derive conditions for f_α without having the components of the vector fields V,W

in it. In the following, we benefit from the assumption that we are only considering vertical vector fields, which makes some calculations easier. The integrand in (2.63) can be rewritten as

$$\begin{aligned} W^\alpha(\mathcal{L}_{\mathrm{pr}V} f_\alpha) - (\mathcal{L}_{\mathrm{pr}W} f_\alpha)V^\alpha =& W^\alpha(f_{\alpha,u^\beta} V^\beta + f_{\alpha,u^\beta_x} D_x V^\beta + f_{\alpha,u^\beta_{xx}} D_x^2 V^\beta) - \\ &- V^\beta(f_{\beta,u^\alpha} W^\alpha + f_{\beta,u^\alpha_x} D_x W^\alpha + f_{\beta,u^\alpha_{xx}} D_x^2 W^\alpha). \end{aligned} \tag{2.64}$$

We want to order all terms with respect to $W^\alpha V^\beta$, $W^\alpha D_x V^\beta$, $W^\alpha D_x^2 V^\beta$ and a total derivative of something. That is, we want all derivatives on V^β, no derivatives on W^α and a total derivative of something. This can be done with a sort of partial integration technique. Before we continue with rewriting (2.64), we consider the following identities: Let a,b,c be functions on $J^k E$, then the first identity is of the form

$$abD_x c = D_x(abc) - c(aD_x b + bD_x a).$$

However, what we actually want is (with summation over α,β)

$$\begin{aligned} V^\beta f_{\beta,u^\alpha_x} D_x W^\alpha &= D_x(V^\beta W^\alpha f_{\beta,u^\alpha_x}) - W^\alpha D_x(V^\beta f_{\beta,u^\alpha_x}) = \\ &= D_x(V^\beta W^\alpha f_{\beta,u^\alpha_x}) - W^\alpha[(D_x V^\beta) f_{\beta,u^\alpha_x} + V^\beta(D_x f_{\beta,u^\alpha_x})]. \end{aligned} \tag{2.65}$$

The second identity is of the form

$$\begin{aligned} abD_x^2 c &= D_x(abD_x c) - (D_x c)D_x(ab) = \\ &= D_x(abD_x c) - D_x[cD_x(ab)] + cD_x^2(ab) = \\ &= D_x[abD_x c - cD_x(ab)] + c[bD_x^2 a + 2(D_x a)(D_x b) + aD_x^2 b] \end{aligned}$$

and what we actually want is (with summation over α,β)

$$\begin{aligned} &V^\beta f_{\beta,u^\alpha_{xx}} D_x^2 W^\alpha = \\ =& D_x[f_{\beta,u^\alpha_{xx}} V^\beta (D_x W^\alpha)] - (D_x W^\alpha) D_x(f_{\beta,u^\alpha_{xx}} V^\beta) = \\ =& D_x[f_{\beta,u^\alpha_{xx}} V^\beta (D_x W^\alpha)] - D_x[W^\alpha D_x(f_{\beta,u^\alpha_{xx}} V^\beta)] + W^\alpha D_x^2(f_{\beta,u^\alpha_{xx}} V^\beta) = \\ =& D_x[f_{\beta,u^\alpha_{xx}} V^\beta (D_x W^\alpha) - W^\alpha D_x(f_{\beta,u^\alpha_{xx}} V^\beta)] + \\ &+ W^\alpha[(D_x^2 V^\beta) f_{\beta,u^\alpha_{xx}} + 2(D_x V^\beta)(D_x f_{\beta,u^\alpha_{xx}}) + V^\beta(D_x^2 f_{\beta,u^\alpha_{xx}})]. \end{aligned} \tag{2.66}$$

With the help of the identities (2.65) and (2.66), equation (2.64) can be written as

$$
\begin{aligned}
& W^\alpha(\mathcal{L}_{\mathrm{pr}V} f_\alpha) - (\mathcal{L}_{\mathrm{pr}W} f_\alpha)V^\alpha = \\
\\
= & - D_x[V^\beta W^\alpha f_{\beta,u_x^\alpha} + f_{\beta,u_{xx}^\alpha} V^\beta (D_x W^\alpha) - W^\alpha D_x(f_{\beta,u_{xx}^\alpha} V^\beta)] + \\
& + (f_{\alpha,u^\beta} - f_{\beta,u^\alpha}) W^\alpha V^\beta + \\
& + W^\alpha (f_{\alpha,u_x^\beta} D_x V^\beta + f_{\alpha,u_{xx}^\beta} D_x^2 V^\beta) + \\
& + W^\alpha[(D_x V^\beta) f_{\beta,u_x^\alpha} + V^\beta (D_x f_{\beta,u_x^\alpha})] - \\
& - W^\alpha[(D_x^2 V^\beta) f_{\beta,u_{xx}^\alpha} + 2(D_x V^\beta)(D_x f_{\beta,u_{xx}^\alpha}) + V^\beta (D_x^2 f_{\beta,u_{xx}^\alpha})] = \\
\\
= & - D_x[V^\beta W^\alpha f_{\beta,u_x^\alpha} + f_{\beta,u_{xx}^\alpha} V^\beta (D_x W^\alpha) - W^\alpha D_x(f_{\beta,u_{xx}^\alpha} V^\beta)] + \\
& + (f_{\alpha,u^\beta} - f_{\beta,u^\alpha} + D_x f_{\beta,u_x^\alpha} - D_x^2 f_{\beta,u_{xx}^\alpha}) W^\alpha V^\beta + \\
& + (f_{\alpha,u_x^\beta} + f_{\beta,u_x^\alpha} - 2D_x f_{\beta,u_{xx}^\alpha}) W^\alpha (D_x V^\beta) + \\
& + (f_{\alpha,u_{xx}^\beta} - f_{\beta,u_{xx}^\alpha}) W^\alpha (D_x^2 V^\beta) = \\
\\
= & - D_x[V^\beta W^\alpha f_{\beta,u_x^\alpha} + f_{\beta,u_{xx}^\alpha} V^\beta (D_x W^\alpha) - W^\alpha D_x(f_{\beta,u_{xx}^\alpha} V^\beta)] + \\
& + H_{\alpha\beta} W^\alpha V^\beta + H_{\alpha\beta}^x W^\alpha D_x V^\beta + H_{\alpha\beta}^{xx} W^\alpha D_x^2 V^\beta,
\end{aligned}
\tag{2.67}
$$

where we define

$$
\begin{aligned}
H_{\alpha\beta} &:= f_{\alpha,u^\beta} - f_{\beta,u^\alpha} + D_x f_{\beta,u_x^\alpha} - D_x^2 f_{\beta,u_{xx}^\alpha}, \\
H_{\alpha\beta}^x &:= f_{\alpha,u_x^\beta} + f_{\beta,u_x^\alpha} - 2D_x f_{\beta,u_{xx}^\alpha}, \\
H_{\alpha\beta}^{xx} &:= f_{\alpha,u_{xx}^\beta} - f_{\beta,u_{xx}^\alpha}
\end{aligned}
\tag{2.68}
$$

as the **Helmholtz expressions** (for second order ODEs). With these identities we can show that if (2.63) vanishes for all vertical V,W on E, supp$V$,$W \subset \pi^{-1}(U^0)$, then

$$
H_{\alpha\beta} = 0, \quad H_{\alpha\beta}^x = 0, \quad H_{\alpha\beta}^{xx} = 0.
$$

These are called the **Helmholtz conditions**.

Proposition 2.6.6. *i) If (2.63) is satisfied for all vertical vector fields V,W on E, supp$V$,$W \subset \pi^{-1}(U^0)$, then the Helmholtz conditions are satisfied.*
ii) If the Helmholtz conditions are satisfied, then (2.63) is satisfied for all vertical vector fields V,W on E and supp$V$,$W \subset \pi^{-1}(U^0)$.

Proof of *i*): For the proof, we need the Du Bois-Reymond lemma, or a version of it.

Equation (2.63) can be written as

$$0 = -\int_{U^0} \operatorname{pr}\sigma^* \left\{ D_x[V^\beta W^\alpha f_{\beta,u^\alpha_x} + f_{\beta,u^\alpha_{xx}} V^\beta (D_x W^\alpha) - W^\alpha D_x(f_{\beta,u^\alpha_{xx}} V^\beta)]dx \right\} +$$
$$+ \int_{U^0} \operatorname{pr}\sigma^* \left\{ [H_{\alpha\beta} W^\alpha V^\beta + H^x_{\alpha\beta} W^\alpha (D_x V^\beta) + H^{xx}_{\alpha\beta} W^\alpha (D^2_x V^\beta)]dx \right\},$$

according to (2.67). Then we get

$$0 = \int_{U^0} \operatorname{pr}\sigma^* \left\{ D_x[V^\beta W^\alpha f_{\beta,u^\alpha_x} + f_{\beta,u^\alpha_{xx}} V^\beta (D_x W^\alpha) - W^\alpha D_x(f_{\beta,u^\alpha_{xx}} V^\beta)]dx \right\}, \quad (2.69)$$

since the integrand is a total derivative of something and depends linearly on V,W and their derivatives (by fundamental theorem of integration). It remains the integral

$$0 = \int_{U^0} \operatorname{pr}\sigma^* \left\{ [H_{\alpha\beta} W^\alpha V^\beta + H^x_{\alpha\beta} W^\alpha (D_x V^\beta) + H^{xx}_{\alpha\beta} W^\alpha (D^2_x V^\beta)]dx \right\}$$

for all vertical $V,W \in \mathfrak{X}(E)$ and supp$V,W \subset \pi^{-1}(U^0)$. Let us consider

$$\begin{aligned} \operatorname{pr}\sigma^*(W^\alpha V^\beta) &=: \varphi^\alpha(x)\psi^\beta(x), \\ \operatorname{pr}\sigma^*(W^\alpha D_x V^\beta) &=: \varphi^\alpha(x)\psi^\beta_x(x), \\ \operatorname{pr}\sigma^*(W^\alpha D^2_x V^\beta) &=: \varphi^\alpha(x)\psi^\beta_{xx}(x) \end{aligned}$$

as functions of x (after pull-back by a section) and we consider $\varphi\psi,\varphi\psi_x,\varphi\psi_{xx}$ as test functions. Since these test functions are in some sense independent, we get $H_{\alpha\beta} = H^x_{\alpha\beta} = H^{xx}_{\alpha\beta} = 0$. Now we formulate this statement more precisely. Why are the functions $\varphi\psi,\varphi\psi_x,\varphi\psi_{xx}$ independent? Because we choose φ,ψ such that suppφ $\subset$ suppψ and such that $\psi(x) \equiv 1$ in the support of $\varphi(x)$. Then we get

$$\begin{aligned} \varphi(x)\psi(x) &= \varphi(x), \\ \varphi(x)\psi_x(x) &= 0, \\ \varphi(x)\psi_{xx}(x) &= 0 \end{aligned}$$

and by Du Bois-Reymond's lemma we get $\mathrm{pr}\sigma^* H_{\alpha\beta} = 0$ for all sections σ and therefore for all points in J^kE. Second, we can repeat this argument and we get $H^x_{\alpha\beta} = H^{xx}_{\alpha\beta} = 0$.

Proof of *ii)*: This direction is easy and follows immediately by (2.67) and the fundamental theorem of integration. □

In the next subsection, we show that

$$H_{\alpha\beta} = H^x_{\alpha\beta} = H^{xx}_{\alpha\beta} = 0 \quad (2.70)$$

for all second order variational f_α, without using Du Bois-Reymond's lemma and integration. Therefore, it would have been sufficient to define the Helmholtz conditions (invariantly) and then show that they are satisfied. But we wanted to provide some ideas how to find them and show the correspondence to the (standard) integrability conditions for vector fields and differential forms in finite dimensional spaces, like $\mathbb{R}^n$.

2.6.2. Alternative Proof of the Helmholtz Conditions and Summary

The task in this subsection is to deliver an alternative proof for the Helmholtz conditions.

Lemma 2.6.7. *If a second order source form $\Delta = f_\alpha du^\alpha \wedge dx$ is (locally) variational, then:*
i) The Helmholtz conditions

$$\begin{aligned} H_{\alpha\beta}(f) &= f_{\alpha,u^\beta} - f_{\beta,u^\alpha} + D_x f_{\beta,u_x^\alpha} - D_x^2 f_{\beta,u_{xx}^\alpha} = 0, \\ H^x_{\alpha\beta}(f) &= f_{\alpha,u_x^\beta} + f_{\beta,u_x^\alpha} - 2D_x f_{\beta,u_{xx}^\alpha} = 0, \\ H^{xx}_{\alpha\beta}(f) &= f_{\alpha,u_{xx}^\beta} - f_{\beta,u_{xx}^\alpha} = 0 \end{aligned}$$

are satisfied (necessary conditions).
ii) The Helmholtz conditions hold in every local coordinate system and the transformation of f_α is given by the source form Δ.

Proof of i): We only prove it when L is of second order (the more general case is similar). Then, by definition, variational f_α can be written as

$$f_\alpha = (\partial_{u^\alpha} - D_x \partial_{u_x^\alpha} + D_x^2 \partial_{u_{xx}^\alpha})L. \tag{2.71}$$

Let us consider the leading order term of f_α, that is, the term $D_x^2 \partial_{u_{xx}^\alpha} L$. First, if $\partial_{u_{xx}^\alpha} L$ depends on the u_{xx}^γ-coordinates, then D_x^2 will generate $u_{(4)}^\gamma$-coordinates. But we assumed that f_α is of second order, and therefore this cannot be the case. Second, if $(-\partial_{u_x^\alpha} + D_x \partial_{u_{xx}^\alpha})L$ depends on the u_{xx}^γ-coordinates, then D_x will generate $u_{(3)}^\gamma$-coordinates. But we assumed that f_α is of second order, and therefore this also cannot be the case. Together,

$$\partial_{u_{xx}^\gamma}(\partial_{u_{xx}^\alpha} L) = 0, \tag{2.72}$$
$$\partial_{u_{xx}^\gamma}[(-\partial_{u_x^\alpha} + D_x \partial_{u_{xx}^\alpha})L] = 0. \tag{2.73}$$

Using the identities in Lemma 2.4.6, we can rewrite equation (2.73) as

$$\begin{aligned} 0 &= \partial_{u_{xx}^\gamma}[(-\partial_{u_x^\alpha} + D_x \partial_{u_{xx}^\alpha})L] = [-\partial_{u_{xx}^\gamma}\partial_{u_x^\alpha} + (\partial_{u_x^\gamma} + D_x \underbrace{\partial_{u_{xx}^\gamma})\partial_{u_{xx}^\alpha}}_{=0\ (2.72)}]L = \\ &= (-\partial_{u_{xx}^\gamma}\partial_{u_x^\alpha} + \partial_{u_x^\gamma}\partial_{u_{xx}^\alpha})L. \end{aligned} \tag{2.74}$$

Then we use again the commutator identities in Lemma 2.4.6 and we can show that

$$\begin{aligned}
H^{xx}_{\alpha\beta} =&[\partial_{u^\beta_{xx}}(\partial_{u^\alpha} - D_x\partial_{u^\alpha_x} + D_x^2\partial_{u^\alpha_{xx}}) - \partial_{u^\alpha_{xx}}(\partial_{u^\beta} - D_x\partial_{u^\beta_x} + D_x^2\partial_{u^\beta_{xx}})]L = \\
=&\{\partial_{u^\beta_{xx}}\partial_{u^\alpha} - (\partial_{u^\beta_x} + D_x\partial_{u^\beta_{xx}})\partial_{u^\alpha_x} + (\partial_{u^\beta} + 2D_x\partial_{u^\beta_x} + D_x^2\partial_{u^\beta_{xx}})\partial_{u^\alpha_{xx}} - \\
&- [\partial_{u^\alpha_{xx}}\partial_{u^\beta} - (\partial_{u^\alpha_x} + D_x\partial_{u^\alpha_{xx}})\partial_{u^\beta_x} + (\partial_{u^\alpha} + 2D_x\partial_{u^\alpha_x} + D_x^2\partial_{u^\alpha_{xx}})\partial_{u^\beta_{xx}}]\}L = \\
=&[-3D_x\partial_{u^\beta_{xx}}\partial_{u^\alpha_x} + 3D_x\partial_{u^\alpha_{xx}}\partial_{u^\beta_x}]L = 3D_x(-\partial_{u^\beta_{xx}}\partial_{u^\alpha_x} + \partial_{u^\alpha_{xx}}\partial_{u^\beta_x})L = 0,
\end{aligned} \tag{2.75}$$

simply by ordering all terms with zero D_x, one D_x and a D_x^2-term and using (2.74) in the last line. In a similar way we prove it for $H_{\alpha\beta}$ and $H^x_{\alpha\beta}$.

Proof of *ii)*: Since we have already proven in Proposition 2.4.2 that variational f_α can be written as we did in (2.71) in every local coordinate system of J^kE, the calculation in (2.75) holds in every local coordinate system and also for $H^x_{\alpha\beta}$ and $H_{\alpha\beta}$. □

In the next section, we show that the Helmholtz conditions are also locally sufficient for the existence of a local Lagrangian. Let us define the **(naive) Helmholtz operator**

$$\mathcal{H}^\gamma_{\alpha\beta} := \begin{cases} \mathcal{H}^{0,\gamma}_{\alpha\beta} = \partial_{u^\beta}\delta^\gamma_\alpha - \partial_{u^\alpha}\delta^\gamma_\beta + D_x\partial_{u^\alpha_x}\delta^\gamma_\beta - D_x^2\partial_{u^\alpha_{xx}}\delta^\gamma_\beta, \\ \mathcal{H}^{x,\gamma}_{\alpha\beta} = \partial_{u^\beta_x}\delta^\gamma_\alpha + \partial_{u^\alpha_x}\delta^\gamma_\beta - 2D_x\partial_{u^\alpha_{xx}}\delta^\gamma_\beta, \\ \mathcal{H}^{xx,\gamma}_{\alpha\beta} = \partial_{u^\beta_{xx}}\delta^\gamma_\alpha - \partial_{u^\alpha_{xx}}\delta^\gamma_\beta, \end{cases}$$

where δ^α_γ is the Kronecker-delta. Above, we have proven that $\mathcal{H}^\gamma_{\alpha\beta}\mathcal{E}_\gamma = 0$, where $\mathcal{E}_\gamma$ is the Euler-Lagrange operator defined in (2.26). Also note that

$$\mathcal{H}^\gamma_{\alpha\beta}f_\gamma = \begin{cases} H_{\alpha\beta} = \partial_{u^\beta}f_\alpha - \partial_{u^\alpha}f_\beta + D_x\partial_{u^\alpha_x}f_\beta - D_x^2\partial_{u^\alpha_{xx}}f_\beta, \\ H^x_{\alpha\beta} = \partial_{u^\beta_x}f_\alpha + \partial_{u^\alpha_x}f_\beta - 2D_x\partial_{u^\alpha_{xx}}f_\beta, \\ H^{xx}_{\alpha\beta} = \partial_{u^\beta_{xx}}f_\alpha - \partial_{u^\alpha_{xx}}f_\beta. \end{cases}$$

Later, we can construct a (naive) locally exact sequence of the form (also see (And89))

$$0 \to \mathbb{R} \to C^\infty(J^kE) \xrightarrow{D_x} \{L\} \xrightarrow{\mathcal{E}_\alpha} \{f_\alpha\} \xrightarrow{\mathcal{H}^\gamma_{\alpha\beta}} \{H_{\alpha\beta}, H^x_{\alpha\beta}, H^{xx}_{\alpha\beta}\} \longrightarrow ....$$

As the sequence is written here, it is not very precise, especially it is not coordinate invariant, but it provides a good understanding. The idea is that we want to construct a sequence, where the morphisms are some kind of exterior derivatives d, the sets are sets of differential forms, and it looks something like

$$0 \to \mathbb{R} \to C^\infty(J^kE) \xrightarrow{d} \Omega^1(J^kE) \xrightarrow{d} \Omega^2(J^kE) \longrightarrow ..., \tag{2.76}$$

which is pretty similar to the De Rham sequence. However, this is just the idea and we would need more time and space to introduce it precisely. We would need to introduce the *Variational Bicomplex* or the *Variational Sequence* (And89, Ku04, Kru97a). As we already partially mentioned earlier, the Variational Sequence (for $n = 1$) is the sequence of quotient spaces

$$0 \to \mathbb{R} \to C^\infty(J^kE) \xrightarrow{d} \Omega^1(J^kE)/(dC^\infty + 1\text{-contact}) \xrightarrow{d}$$
$$\xrightarrow{d} \Omega^2/(d(1\text{-contact}) + 2\text{-contact}) \xrightarrow{d} \Omega^3/(d(2\text{-contact}) + 3\text{-contact}) \longrightarrow ...$$

Similar as the Lagrange and source form is defined, the Helmholtz form is defined as

$$H = H(\Delta) := \frac{1}{2}(H_{\alpha\beta}\Theta^\beta \wedge \Theta^\alpha \wedge dx + H^x_{\alpha\beta}\Theta^\beta_x \wedge \Theta^\alpha \wedge dx + H^{xx}_{\alpha\beta}\Theta^\beta_{xx} \wedge \Theta^\alpha \wedge dx),$$

where $H_{\alpha\beta}$,$H^x_{\alpha\beta}$,$H^{xx}_{\alpha\beta}$ are the Helmholtz expressions in (2.68) (see (VU13, p.13)). It turns out that

$$\iota_{\mathrm{pr}V}\iota_{\mathrm{pr}W}H = [H_{\alpha\beta}W^\beta V^\alpha + H^x_{\alpha\beta}W^\beta D_xV^\alpha + H^{xx}_{\alpha\beta}W^\beta D^2_xV^\alpha + D_x(...)]dx.$$

See Appendix C. (it also turns out that we need the factor $\frac{1}{2}$ in the Helmholtz form). The Helmholtz form is important to understand where the Helmholtz conditions come from, when working with the variational sequence.

Example: Earlier, we have already shown that $f = u_x$ is not variational. There we did a more or less straight forward calculation and we used order discussion of jet coordinates. Now we can also show it with the help of the Helmholtz conditions. The Helmholtz conditions for single, second order ODEs are (see (2.68), $\alpha = \beta = 1$ and we do not write these indices)

$$0 = D_x(f_{u_x} - D_xf_{u_{xx}}) = H_{\alpha\beta} = H, \tag{2.77}$$
$$0 = 2(f_{u_x} - D_xf_{u_{xx}}) = H^x_{\alpha\beta} = H^x, \tag{2.78}$$
$$0 = H^{xx}_{\alpha\beta} = H^{xx}. \tag{2.79}$$

Since $f_{u_x} - D_xf_{u_{xx}} = 1 \neq 0$, condition (2.78) is not satisfied and the expression $f = u_x$ is not variational.

We can see here that (2.77) and (2.78) are not independent, because (2.77) is a half the total derivative of (2.78). Actually, this sort of dependencies allow us to solve Takens' problem, as we will see later. See Section 3.8, where we discuss this in detail.

At the end of this section we want to summarize what we have found out so far with the help of the following table:

classical analysis in $\mathbb{R}^n$	∞-dim, calculus of variations	differential geometry on J^kE
$\phi : \mathbb{R}^n \to \mathbb{R}$	$I = \int L dx : S \to \mathbb{R}$	$\lambda = L dx \in \Omega^n(J^kE)$
$x \in \mathbb{R}^n$, finite dim.	$u^\beta(x) \in S$, ∞-dim.	$\sigma \in \Gamma(E)$, section
x_t, 1-parameter point	$u^\beta_t(x)$, 1-parameter function	σ_t, 1-parameter section
$\frac{d}{dt} x_t\vert_{t=0} = v$	$\frac{d}{dt} u^\beta_t(x)\vert_{t=0} = \varphi^\beta(x)$	$\sigma_t = \phi_t \circ \sigma$, $\frac{d}{dt}\phi_t\vert_{t=0} = V$
v, vector in $\mathbb{R}^n$	$\varphi^\beta(x)$, test function	$V = V^\beta \partial_{u^\beta}$, vertical VF on E
$d\phi_x(v)$, differential	$\delta I(u;\varphi)$, first variation	$\mathcal{L}_{\mathrm{pr}V} \lambda$
d, exterior derivative	δ or $\mathcal{E}_\alpha$	E^{n+1}
$d\, const. = 0$	$\mathcal{E}_\alpha D_x \Lambda = 0$	$E^{n+1}(d\eta + 1\text{-contact}) = 0$
$\omega \in \Omega^1(\mathbb{R}^n)$	f_α	Δ, source form
$\omega = d\phi$?	$f_\alpha = \mathcal{E}_\alpha L$?	$\Delta = E^{n+1}(\lambda)$?, or $[\Delta] = [d\lambda]$?
$(\mathcal{L}_{\mathrm{pr}V}\iota_{\mathrm{pr}W} - \iota_{\mathrm{pr}V}\mathcal{L}_{\mathrm{pr}W})\omega = 0$	$(\mathcal{L}_{\mathrm{pr}V}\iota_{\mathrm{pr}W} - \iota_{\mathrm{pr}V}\mathcal{L}_{\mathrm{pr}W})K = 0$	$(\mathcal{L}_{\mathrm{pr}V}\iota_{\mathrm{pr}W} - \iota_{\mathrm{pr}V}\mathcal{L}_{\mathrm{pr}W})\Delta = 0$
$d\omega = 0$, necessary cond.	$dK = 0$, necessary cond.	$E^{n+2}(\Delta) = 0$, or $[d\Delta] = 0$
$dd = 0$	$\mathcal{H}^\gamma_{\alpha\beta}\mathcal{E}_\gamma = 0$, or $d\delta I = 0$	$E^{n+2}E^{n+1} = 0$

In the above table, VF denotes vector field. Note that $d\delta I$ is not the second variation of I which is denoted by $\delta^2 I$.

2.7. Homotopy Formula and Locally Exact Sequences

In this section, we want to show that the (naive) variational sequence

$$0 \to \mathbb{R} \to C^\infty(J^kE) \xrightarrow{D_x} \{L\} \xrightarrow{\mathcal{E}_\alpha} \{f_\alpha\} \xrightarrow{\mathcal{H}^\gamma_{\alpha\beta}} \{H_{\alpha\beta}, H^x_{\alpha\beta}, H^{xx}_{\alpha\beta}\} \longrightarrow ...$$

is locally exact.

2.7.1. The Inverse Euler-Lagrange Mapping

We want to prove the following lemma:

Lemma 2.7.1. *Let $n = 1$, $m \in \mathbb{N}$ and L be a Lagrangian on J^kE. If $\mathcal{E}_\alpha L = 0$ then there exists locally a function $\Lambda \in C^\infty(J^{k-1}E)$ such that $L = D_x\Lambda$.*

Before we start with the proof, let us explain the idea. We start with the simplest case of first order Lagrangian $L = L(x,u,u_x)$ and $n,m = 1$. We have to construct the function $\Lambda = \Lambda(x,u)$ somehow. The idea is to integrate along the x-coordinate, since formally, when $L = D_x\Lambda$, then

$$\int_\gamma L dx = \int_{x_0}^{x} \mathrm{pr}^1 \sigma^*[(D_x\Lambda)dx] = \Lambda(x,u(x)) + \text{constant}, \tag{2.80}$$

where γ is the curve along the section $\mathrm{pr}^1 \sigma$ in J^1E with some initial and endpoint in J^1E. It is clear that constants are in the kernel of D_x, as we have wirtten in (2.80).

The problem with the formula in (2.80) is that $\Lambda(x,u(x))$ should be a function which depends explicitly on (x,u) and in the form $\Lambda(x,u(x)) = (\mathrm{pr}\sigma^*\Lambda)(x)$ it only depends on x and the explicit dependency on u might be not obvious. In other words, $(\mathrm{pr}\,\sigma^*\Lambda)(x)$ and $\Lambda(x,u)$ are different functions, one is locally defined on M and the other one on E. We want to get the function Λ on E (or more generally on $J^{k-1}E$) without pull-back by a section $\sigma \in \Gamma(E)$.

Since it is not clear at this point how to integrate (2.80) such that we get Λ on E (or on $J^{k-1}E$), we are inverting the operator $\partial_u - D_x\partial_{u_x}$ by hand and we try to find out if we get a better understanding thereby. Indeed, we will find the right formula. We can solve relatively easily the following equation (here $\mathcal{E}_u = \mathcal{E}_\alpha$, where $\alpha = 1$)

$$\begin{aligned} \mathcal{E}_u L = L_u - D_x L_{u_x} = \\ = L_u - L_{xu_x} - u_x L_{uu_x} - u_{xx} L_{u_x u_x} = 0. \end{aligned} \tag{2.81}$$

Since $L_u, L_{xu_x}, L_{u_x u_x}$ only depend on the coordinates (x,u,u_x), we get $L_{u_x u_x} = 0$ (since there is no second order coordinate u_{xx} on the right hand side in (2.81)). The equation $L_{u_x u_x} = 0$ can easily be solved and we get

$$L(x,u,u_x) = A(x,u) + u_x B(x,u), \tag{2.82}$$

where A,B are no more specified functions on E at this point. Therefore, we constructed an approximative L, approximation in highest order coordinate u_x, and we have to find out more about A and B. For this purpose, we plug this L again in equation (2.81) and we get

$$(A_u + u_x B_u) - (B_x + u_x B_u) = A_u - B_x = \begin{pmatrix} \partial u \\ \partial x \end{pmatrix} \times \begin{pmatrix} B \\ A \end{pmatrix} = 0. \tag{2.83}$$

This is basically the first time where we have a more or less non-trivial partial differential equation to solve. Locally, the solution is

$$\begin{pmatrix} B \\ A \end{pmatrix} = \begin{pmatrix} \partial_u \\ \partial_x \end{pmatrix} \phi \tag{2.84}$$

for some function ϕ on E. As it is well-known, the function ϕ can be obtained by the homotopy formula

$$\phi = \int_\gamma \left\langle \begin{pmatrix} A \\ B \end{pmatrix}, \begin{pmatrix} dx \\ du \end{pmatrix} \right\rangle = \int_0^1 \left\langle \begin{pmatrix} A(\gamma(t)) \\ B(\gamma(t)) \end{pmatrix}, \frac{d}{dt}\gamma(t) \right\rangle dt, \tag{2.85}$$

using the standard Poincaré lemma, where $\gamma(t)$ is a 1-parameter curve in the (x,u)-plane and $\gamma(t = 1) = (x,u)$. Therefore, we know the function ϕ and with the help of (2.82) and (2.84) we get

$$L(x,u,u_x) = \phi_x(x,u) + u_x\phi_u(x,u) = D_x\phi(x,u),$$

and therefore $\Lambda = \phi$. Now the question is: How can we explicitly construct Λ for higher order with the help of L and without doing the intermediate step, where we computed A and B. The answer is: With the help of equation (2.82) and (2.85). First, we write B as

$$B = L_{u_x},$$

and second, we write A as

$$A = L - u_x L_{u_x}.$$

Then we use (2.85) and we get

$$\Lambda(x,u) = \int_\gamma \left\langle \begin{pmatrix} L - u_x L_{u_x} \\ L_{u_x} \end{pmatrix}, \begin{pmatrix} dx \\ du \end{pmatrix} \right\rangle. \tag{2.86}$$

If we pull-back this formula by a section $\sigma \in \Gamma(E)$, where $(x,u(x))$ is the corresponding local coordinate section, then we get

$$\begin{aligned} &\Lambda(x,u(x)) = \\ &= \int_{x_0}^{x} \mathrm{pr}^1\,\sigma^* \left\langle \begin{pmatrix} L - u_x L_{u_x} \\ L_{u_x} \end{pmatrix}, \begin{pmatrix} dx \\ du \end{pmatrix} \right\rangle = \int_{x_0}^{x} \left\langle \mathrm{pr}^1\,\sigma^* \begin{pmatrix} L - u_x L_{u_x} \\ L_{u_x} \end{pmatrix}, \begin{pmatrix} dx \\ u_x(x)dx \end{pmatrix} \right\rangle = \\ &= \int_{x_0}^{x} \left\langle \mathrm{pr}^1\,\sigma^* \begin{pmatrix} L - u_x L_{u_x} \\ L_{u_x} \end{pmatrix}, \begin{pmatrix} 1 \\ u_x(x) \end{pmatrix} \right\rangle dx = \int_{x_0}^{x} \mathrm{pr}^1\,\sigma^*(L dx), \end{aligned}$$

which is exactly formula (2.80) from the beginning (modulo the constant). Formula (2.86) has two advantages compared to (2.80):

- The formula in (2.86) tells us now how to integrate without pull-backing by a section. We can integrate along any curve γ in J^1E, not necessarily a prolonged section (see Lemma 2.7.2 below).

- The formula in (2.86) can also be used when, for example, $L = u_x u_x$, where we cannot write $L = A + u_x B$, then of course $L \neq D_x\Lambda$. But we could use it to measure the **extent of not being a total derivative** by writing $L = (L - D_x\Lambda) + D_x\Lambda$, where $(L - D_x\Lambda)$ could be defined as the extent of not being a total derivative (we would have to check that this is coordinate independent). In a similar way we can define the extend of not being a variational equation.

Lemma 2.7.2. *The integral in* (2.86) *is locally independent of the choice of a curve* γ *and only depends on the initial and endpoint of* γ*.*

Proof: In other words, we have to show that the vector field

$$\begin{pmatrix} L - u_x L_{u_x} \\ L_{u_x} \end{pmatrix}$$

is conservative (see (Mat98)) and this can simply be shown by

$$\begin{pmatrix} \partial_x \\ \partial_u \end{pmatrix} \times \begin{pmatrix} L - u_x L_{u_x} \\ L_{u_x} \end{pmatrix} = L_{xu_x} - (L_u - u_x L_{uu_x}) = D_x L_{u_x} - L_u = -\mathcal{E}_u L = 0.$$

The equality before last holds since $u_{xx} L_{u_x u_x} = 0$ (order dicussion). □

It will be important soon that the integral is independent of the choice of a section, or any curve, and only depends on initial and endpoint. We could also say that the differential form $(L - u_x L_{u_x})dx + L_{u_x} du$ is closed and locally exact. This differential form is also called Poincaré Cartan form and can be found in the literature, for example, see (Kru97b, p.48).

Formula (2.86) can also be written with the help of contact forms as

$$\Lambda(x,u) = \int_\gamma (Ldx + L_{u_x} \Theta^u)$$

and this may help to find a formula for higher order Lagrangians (however, we could also do a calculation for higher order with the methods from above). Let us explain how we construct such differential forms $Ldx + L_{u_x}\Theta^u$. We are looking for a differential one form $\omega \in \Omega^1(J^k E)$, such that $\mathrm{pr}\sigma^* \omega = \mathrm{pr}\sigma^*(Ldx)$ and $d\omega = 0$. Then

$$\int_\gamma \omega = \Lambda$$

for any curve γ in $J^k E$, since ω is locally exact, and therefore the integral depends only on the initial and endpoint of γ. Furthermore,

$$\int_{x_0}^{x} \mathrm{pr}\sigma^* \omega = \int_{x_0}^{x} \mathrm{pr}\,\sigma^*(Ldx) = \Lambda(x,u(x),u_x(x),...).$$

The question is if we can always find such a differential form ω which satisfies both conditions. Let us try to do this for second order $L = L(x,u,u_x,u_{xx})$. In the following, we will write $\Theta^u = \Theta$ for simplicity. We try to determine A and A^x in

$$\omega = Ldx + A\Theta + A^x \Theta_x,$$

such that the conditions from above are satisfied. The condition $\mathrm{pr}\sigma^*\omega = \mathrm{pr}\sigma^*(Ldx)$

is of course satisfied. Furthermore, (we use the standard decomposition)

$$\begin{aligned}
0 = d\omega =& L_u du \wedge dx + L_{u_x} du_x \wedge dx + L_{u_{xx}} du_{xx} \wedge dx + \\
& + (dA) \wedge \Theta + A d\Theta + (dA^x) \wedge \Theta_x + A^x d\Theta_x = \\
\\
=& L_u \Theta \wedge dx + L_{u_x} \Theta_x \wedge dx + L_{u_{xx}} \Theta_{xx} \wedge dx + \\
& + (dA) \wedge \Theta - A\Theta_x \wedge dx + (dA^x) \wedge \Theta_x - A^x \Theta_{xx} \wedge dx = \\
\\
=& L_u \Theta \wedge dx + L_{u_x} \Theta_x \wedge dx + L_{u_{xx}} \Theta_{xx} \wedge dx + \\
& + (D_x A) dx \wedge \Theta - A\Theta_x \wedge dx + (D_x A^x) dx \wedge \Theta_x - A^x \Theta_{xx} \wedge dx + \\
& + A_{u_x} \Theta_x \wedge \Theta + A_{u_{xx}} \Theta_{xx} \wedge \Theta + A^x_u \Theta \wedge \Theta_x + A^x_{u_{xx}} \Theta_{xx} \wedge \Theta_x = \\
\\
=& (L_u - D_x A) \Theta \wedge dx + (L_{u_x} - A - D_x A^x) \Theta_x \wedge dx + (L_{u_{xx}} - A^x) \Theta_{xx} \wedge dx + \\
& + (A_{u_x} - A^x_u) \Theta_x \wedge \Theta + A_{u_{xx}} \Theta_{xx} \wedge \Theta + A^x_{u_{xx}} \Theta_{xx} \wedge \Theta_x. \qquad (2.87)
\end{aligned}$$

In the last line in (2.87) we find out that A,A^x must be of first order, i.e.

$$A_{u_{xx}} = 0, \quad A^x_{u_{xx}} = 0.$$

In the last expression of the second last line in (2.87) we find out that

$$A^x = L_{u_{xx}}. \qquad (2.88)$$

Then, the vanishing of $L_{u_x} - A - D_x A^x$ leads to

$$A = L_{u_x} - D_x A^x = L_{u_x} - D_x L_{u_{xx}}. \qquad (2.89)$$

Then it is clear that the first term in (2.87) vanishes, since

$$L_u - D_x A = L_u - D_x(L_{u_x} - D_x L_{u_{xx}}) = \mathcal{E}_u L = 0,$$

by assumption of Lemma 2.7.1. To show that we can also choose A,A^x such that $A_{u_x} - A^x_u$ vanishes, is a bit more complicated. We use the identities in Lemma 2.4.6, to commute ∂_{u_x},$\partial_{u_{xx}}$ with D_x and then we get

$$\begin{aligned}
A_{u_x} - A^x_u &= \partial_{u_x}(L_{u_x} - D_x L_{u_{xx}}) - \partial_u L_{u_{xx}} = \\
&= L_{u_x u_x} - (D_x \partial_{u_x} + \partial_u) L_{u_{xx}} - \partial_u L_{u_{xx}} = \\
&= L_{u_x u_x} - D_x L_{u_x u_{xx}} - 2L_{u u_{xx}} = \\
&= L_{u_x u_x} - D_x \partial_{u_{xx}} (A + D_x L_{u_{xx}}) - 2L_{u u_{xx}} = \\
&= L_{u_x u_x} - D_x \partial_{u_{xx}} (D_x L_{u_{xx}}) - 2L_{u u_{xx}} = \\
&= L_{u_x u_x} - (\partial_{u_{xx}} D_x - \partial_{u_x})(D_x L_{u_{xx}}) - 2L_{u u_{xx}} = \\
&= L_{u_x u_x} - 2L_{u u_{xx}} - \partial_{u_{xx}} D_x^2 L_{u_{xx}} + \partial_{u_x} D_x L_{u_{xx}} = \\
&= L_{u_x u_x} - 2L_{u u_{xx}} - \partial_{u_{xx}}(-L_u + D_x L_{u_x}) + (D_x \partial_{u_x} + \partial_u) L_{u_{xx}} = \\
&= L_{u_x u_x} - \partial_{u_{xx}} D_x L_{u_x} + D_x \partial_{u_x} L_{u_{xx}} = \\
&= L_{u_x u_x} - (D_x \partial_{u_{xx}} + \partial_{u_x}) L_{u_x} + D_x \partial_{u_x} L_{u_{xx}} = 0.
\end{aligned}$$

Therefore, if we choose A and A^x as we did in (2.88) and (2.89), then ω is closed when $\mathcal{E}_u L = 0$.

Now we basically understood how we can construct Λ and one finds out that the general formula is

$$\Lambda(x,u,u_x,...) = \int_\gamma \left\langle \begin{pmatrix} L - u_x(L_{u_x} - D_x L_{u_{xx}} \pm ...) - u_{xx}(L_{u_{xx}} - D_x L_{u_{xxx}} \pm ...) - ... \\ L_{u_x} - D_x L_{u_{xx}} + D^2 L_{u_{xxx}} \pm ... \\ L_{u_{xx}} - D_x L_{u_{xxx}} \pm ... \\ \vdots \end{pmatrix}, \begin{pmatrix} dx \\ du \\ du_x \\ \vdots \end{pmatrix} \right\rangle . \tag{2.90}$$

A similar formula holds for PDEs in any dimension, for example, see the slightly different formula 5.109 in (Olv86, p.363).

There are now at least two ways how to prove Lemma 2.7.1. The first way is to prove that we can always determine the coefficients A,A^x,...,$A^{(k-1)}$ in

$$\omega = Ldx + A\Theta + A^x\Theta_x + ... + A^{(k-1)}\Theta_{(k-1)},$$

such that $d\omega = 0$ on J^kE and then we use the standard Poincaré lemma to integrate ω and we get a function Λ. To prove that there exist coefficients A,A^x,...,$A^{(k-1)}$ such that $d\omega = 0$ is not obvious, since dx,Θ,Θ_x,...,$\Theta_{(k-1)}$ do not form a basis of differential 1-forms on J^kE. The previous calculation has also shown that it can get complicated to determine these coefficients. We know that dx,Θ,Θ_x,...,$\Theta_{(k-1)}$,$du_{(k)}$ form a basis on J^kE and the freedom of choosing the coefficient in front of $du_{(k)}$ is missing. Then it would be trivial to find coefficients A,A^x,...,$A^{(k)}$ such that $d\omega = 0$. But actually we do not need to determine the coefficient $A^{(k)}$ in front of $du_{(k)}$, since we have the additional condition $\mathcal{E}_\alpha L = 0$, which can be used.

That we can choose A,A^x,...,$A^{(k-1)}$, such that $d\omega = 0$, can probably easier be seen when understanding the Euler-Lagrange operator $\mathcal{E}_\alpha$ in the variational sequence. There, $\mathcal{E}_\alpha L = 0$ is equivalent to $[d(Ldx)] = 0$ and this is equivalent to $[d(Ldx + A\Theta + A^x\Theta_x + ... + A^{(k-1)}\Theta_{(k-1)})] = 0$ (in the equivalence class of the variational sequence). Then with the help of $[d(Ldx)] = 0$ it can be shown that this (locally) leads to

$$Ldx = d\eta + \text{contact form}, \tag{2.91}$$

where η is some function on J^kE. The forms $\{\Theta,\Theta_x,...,\Theta_{(k-1)}\}$ are a basis of contact forms on J^kE. Therefore, the contact form in (2.91) can be written as $-A\Theta - A^x\Theta_x - ... - A^{(k-1)}\Theta_{(k-1)}$, where we choose suitable coefficients A,A^x,...,$A^{(k-1)}$ and we get

$$Ldx + A\Theta + A^x\Theta_x + ... + A^{(k-1)}\Theta_{(k-1)} = d\eta.$$

It is also interesting to note here that the non-triviality of this problem - because we can only choose A,A^x,...,$A^{(k-1)}$ - cannot be seen directly when using the infinite jet space $J^\infty E$. Since there dx,Θ,Θ_x,...,$\Theta_{(k)}$,... is a basis of 1-forms. In this case, the Euler-Lagrange operator is also defined differently and not as a quotient mapping. Since we have not developed both of these theories very well so far, we also do a straight forward calculation below.

A second, more straight forward, proof is to integrate the formula from above by hand and then show that $L = D_x\Lambda$ is satisfied. In any case, one has to do a more or less notation heavy calculation at some point and the question is where we want to do that.

<u>Proof of Lemma 2.7.1:</u> We only prove it for first order Lagrangians $L = L(x,u,u_x)$. Actually, we have already proven it above. But now we want to find a different proof, without using A and B directly. It is helpful to rewrite (2.90) as

$$\Lambda(x,u,u_x) := \int_0^1 \left\langle \begin{pmatrix} L(t) - u_x t L_{u_x}(t) \\ L_{u_x}(t) \end{pmatrix}, \begin{pmatrix} x \\ u \end{pmatrix} \right\rangle dt,$$

where

$$\gamma = \begin{pmatrix} tx \\ tu \end{pmatrix}, t \in [0,1], \quad L(t) := L(tx,tu,tu_x), \quad L_{u_x}(t) := L_{u_x}(tx,tu,tu_x).$$

More generally, we define $g(t) := g(tx,tu,tu_x,...)$ for every function g. We assume that there are local coordinates such that $(x,u) = (0,0)$ is in this coordinate system and that the coordinate system is a starlike set (see (GM10, p.45)). We want to prove the following formula

$$\frac{d}{dt}[tL(t)] + \frac{d}{dt}[(1-t)u_x t L_{u_x}(t)] = D_x \left\langle \begin{pmatrix} L(t) - u_x t L_{u_x}(t) \\ L_{u_x}(t) \end{pmatrix}, \begin{pmatrix} x \\ u \end{pmatrix} \right\rangle, \tag{2.92}$$

then

$$\begin{aligned} D_x \int_0^1 \left\langle \begin{pmatrix} L(t) - u_x t L_{u_x}(t) \\ L_{u_x}(t) \end{pmatrix}, \begin{pmatrix} x \\ u \end{pmatrix} \right\rangle dt &= \int_0^1 \frac{d}{dt}[tL(t) + (1-t)u_x t L_{u_x}(t)]dt = \\ &= L(x,u,u_x) \end{aligned}$$

and we are done.

<u>Proof of (2.92):</u>

$$\begin{aligned} &D_x \left\langle \begin{pmatrix} L(t) - u_x t L_{u_x}(t) \\ L_{u_x}(t) \end{pmatrix}, \begin{pmatrix} x \\ u \end{pmatrix} \right\rangle = \\ &= \underbrace{x D_x[L(t) - u_x t L_{u_x}(t)]}_{=:\mathrm{I}} + \underbrace{u D_x[L_{u_x}(t)]}_{=:\mathrm{II}} + \underbrace{[L(t) - u_x t L_{u_x}(t)] + u_x L_{u_x}(t)}_{=:\mathrm{III}}. \end{aligned} \tag{2.93}$$

We start with discussing II:

$$\begin{aligned}\mathrm{II}:\quad uD_x[L_{u_x}(t)] &= u[tL_{xu_x}(t) + tu_xL_{uu_x}(t) + tu_{xx}L_{u_xu_x}(t)] = \\ &= u[(t-1)L_{xu_x}(t) + (D_xL_{u_x})(t)] = \\ &= u[(t-1)L_{xu_x}(t) + L_u(t)] = \\ &= (1-t)u[-L_{xu_x}(t) + L_u(t)] + tuL_u(t).\end{aligned} \tag{2.94}$$

Then we can use the second last line in (2.94) to discuss I:

$$\begin{aligned}\mathrm{I}:\quad & xD_x[L(t) - u_xtL_{u_x}(t)] = \\ =&x\{D_x[L(t)] - u_{xx}tL_{u_x}(t) - u_xt[(t-1)L_{xu_x}(t) + L_u(t)]\} = \\ =&x[tL_x(t) - u_xt(t-1)L_{xu_x}(t)]\end{aligned}$$

and III can be written as

$$\mathrm{III}:\quad L(t) - u_xtL_{u_x}(t) + u_xL_{u_x}(t) = L(t) + (1-t)u_xL_{u_x}(t).$$

We continue with (2.93) and we order all terms with respect to t and $(1-t)$-terms and we use $L_{u_xu_x} = 0$, which is the highest order coefficient in the equation $\mathcal{E}_uL = 0$. Then we get

$$\begin{aligned}& D_x\left\langle \begin{pmatrix} L(t) - u_xtL_{u_x}(t) \\ L_{u_x}(t)\end{pmatrix}, \begin{pmatrix} x \\ u\end{pmatrix}\right\rangle = \\ =&x[tL_x(t) - u_xt(t-1)L_{xu_x}(t)] + (1-t)u[-L_{xu_x}(t) + L_u(t)] + tuL_u(t) + \\ &\qquad\qquad + L(t) + (1-t)u_xL_{u_x}(t) = \\ =&t[xL_x(t) + uL_u(t) + u_xL_{u_x}(t)] + L(t) \\ &+ (1-t)[xu_xtL_{xu_x}(t) - uL_{xu_x}(t) + uL_u(t) + (1-2t)u_xL_{u_x}(t)] = \\ =&\frac{d}{dt}[tL(t)] + (1-t)[xu_xtL_{xu_x}(t) - u\underbrace{L_{xu_x}(t)}_{\substack{=(D_xL_{u_x})(t) \\ -tu_xL_{uu_x}(t)}} + uL_u(t)] + [\frac{d}{dt}(t-t^2)]u_xL_{u_x}(t) = \\ =&\frac{d}{dt}[tL(t)] + (1-t)[xu_xtL_{xu_x}(t) + tuu_xL_{uu_x}(t)] + [\frac{d}{dt}(t-t^2)]u_xL_{u_x}(t) = \\ =&\frac{d}{dt}[tL(t)] + (1-t)tu_x\frac{d}{dt}L_{u_x}(t) + [\frac{d}{dt}(t-t^2)]u_xL_{u_x}(t) = \\ =&\frac{d}{dt}[tL(t)] + \frac{d}{dt}[(1-t)tu_xL_{u_x}(t)],\end{aligned}$$

which proves Lemma 2.7.1. □

This proof is a good example, where we can see that the computation is getting much more complicated in a slightly more general situation, compared to the calculation at the beginning of this subsection. There we used the special form $L = A + u_xB$ and found out pretty easily that $(A,B) = (\phi_x,\phi_u)$ and then we computed Λ. However, to be fair, there we referred to the standard Poincaré lemma to get this result and here we have basically proven a version of the Poincaré lemma. Surprisingly, proving the homotopy formula in the next subsection will be simpler.

2.7.2. The Inverse Helmholtz Mapping

Similar to the previous subsection, we want to prove the following lemma:

Lemma 2.7.3. *If f_α satisfies the Helmholtz conditions, then locally there exists a Lagrangian L, such that $f_\alpha = \mathcal{E}_\alpha L$. Locally, a Lagrangian is given as*

$$L = \int_0^1 f_\alpha(x,tu^\beta,tu^\beta_x,tu^\beta_{xx},...,tu^\beta_{(k)})u^\alpha dt. \tag{2.95}$$

Again, before we start with the proof, let us explain the idea. Variational equations correspond to vector fields, which can be written as gradients, or in the language of differential forms, they correspond to 1-forms which are (locally) exact.

Finite dimensional analog in $\mathbb{R}^n$: Let us assume that the vector field w on $\mathbb{R}^n$ can be written as a gradient field $w = \nabla\phi$ for some function ϕ. To construct the potential ϕ, we have to integrate along a 1-parameter curve γ_t, where $\gamma_{t=0} = x_0$, $\gamma_{t=1} = x$ and

$$\phi(x) = \int_0^1 < w(\gamma_t),\dot{\gamma}_t >_{\mathbb{R}^n} dt, \tag{2.96}$$

where $\dot{\gamma}_t := \frac{d}{dt}\gamma_t$. In (2.96) we used the Euclidean scalar product.

The ∞-dimensional analog in the calculus of variations: We do the same in the calculus of variations, where we have an ∞-dimensional space S with scalar product given as an integral $\int dx$ over some set $U^0 \subset M$. Let γ_t be a 1-parameter family of sections on E such that $\gamma_{t=1} = \sigma \in \Gamma(E)$ and $\gamma_{t=0} = \sigma_0 \in \Gamma(E)$. Then, the analog of (2.96) is

$$I(\sigma) = \int_{U^0} \mathrm{pr}^k\sigma^*(Ldx) = \int_0^1\int_{U^0} f(\mathrm{pr}^k\gamma_t)\cdot\dot{\gamma}_t dxdt = \int_0^1 < f(\mathrm{pr}^k\gamma_t),\dot{\gamma}_t >_{L^2(U^0)} dt,$$

where $\dot{\gamma}_t := \frac{d}{dt}\gamma_t$. Note that $\dot{\gamma}_t = V_t$ is a vertical vector field on E and $\mathrm{supp}V_t \subset \pi^{-1}(U^0)$. Interchanging the integrals $\int_0^1 dt$ and $\int_{U^0} dx$ shows that

$$L = \int_0^1 f_\alpha(\mathrm{pr}^k\gamma_t)\dot{\gamma}^\alpha_t dt$$

should be the Lagrangian, where

$$\begin{aligned}\mathrm{pr}^k\gamma_t &= (x,u^1_t,...,u^m_t,u^1_{t,x},...,u^m_{t,x},...,u^1_{t,(k)},...,u^m_{t,(k)}),\\ \mathrm{pr}^k\dot\gamma_t &= (\underbrace{0,V^{u,1}_t,...,V^{u,m}_t}_{=(\dot\gamma^\alpha_t)},D_xV^{u,1}_t,...,D_xV^{u,m}_t,...,D^k_xV^{u,1}_t,...,D^k_xV^{u,m}_t),\end{aligned}$$

since we consider vertical perturbations. In a more special case, where

$$\begin{aligned}\mathrm{pr}^k\gamma_t &= (x,tu^1,tu^2,...,tu^m,tu^1_x,...,tu^m_x,...),\\ \mathrm{pr}^k\dot\gamma_{t=0} &= (\underbrace{0,u^1,u^2,...,u^m}_{=(\dot\gamma^\alpha)},u^1_x,...,u^m_x,...),\end{aligned}$$

we get

$$L = \int_0^1 f_\alpha(x,tu^\beta,tu^\beta_x,tu^\beta_{xx},...,tu^\beta_{(k)})u^\alpha dt \tag{2.97}$$

and this is called the Vainberg-Tonti Lagrangian (for example, see (KM10)). Note that there is no t in front of x in (2.97), since we consider vertical perturbations.

Example: As we already discussed above, the equation $f = u_{xx} = 0$ is variational and usually we use $L = -\frac{1}{2}u_x^2$ as Lagrangian. Formula (2.95) tells us

$$L = \int_0^1 tu_{xx}udt = \frac{1}{2}uu_{xx} = D_x(\frac{1}{2}uu_x) - \frac{1}{2}u_x^2 = -\frac{1}{2}u_x^2 + D_x\Lambda,$$

where $D_x\Lambda = D_x(\frac{1}{2}uu_x)$ is a trivial Lagrangian. Formula (2.95) has the disadvantage that it does not deliver the lowest order possible Lagrangian, instead a Lagrangian of the same order as the differential expression f_α. Finding lowest order possible Lagrangians (also invariant Lagrangians) is a topic of its own and we will not discuss it here.

Proof of Lemma 2.7.3: We prove it only for second order Lagrangians. Let us consider

$$\begin{aligned}&\mathcal{E}_\alpha \int_0^1 f_\beta(x,tu^\gamma,tu^\gamma_x,tu^\gamma_{xx})u^\beta dt =\\ =&(\partial_{u^\alpha} - D_x\partial_{u^\alpha_x} + D_x^2\partial_{u^\alpha_{xx}})\int_0^1 f_\beta(x,tu^\gamma,tu^\gamma_x,tu^\gamma_{xx})u^\beta dt =\\ =&\int_0^1 [tf_{\beta,u^\alpha}(x,tu^\gamma,tu^\gamma_x,tu^\gamma_{xx})u^\beta + f_\alpha(x,tu^\gamma,tu^\gamma_x,tu^\gamma_{xx})]dt-\\ &-\int_0^1 D_x[tf_{\beta,u^\alpha_x}(x,tu^\gamma,tu^\gamma_x,tu^\gamma_{xx})u^\beta]dt + \int_0^1 D_x^2[tf_{\beta,u^\alpha_{xx}}(x,tu^\gamma,tu^\gamma_x,tu^\gamma_{xx})u^\beta]dt.\end{aligned} \tag{2.98}$$

To compute the integrals in (2.98), we need the following short notation

$$g(t) := g(x,tu^\alpha,tu^\alpha_x,tu^\alpha_{xx},...),$$

for every function g and we need

$$\begin{aligned} D_x[g(t)] &= D_x[g(x,tu^\alpha,tu^\alpha_x,tu^\alpha_{xx},...)] = \\ &= g_x(t) + tu^\alpha_x g_{u^\alpha}(t) + tu^\alpha_{xx} g_{u^\alpha_x}(t) + tu^\alpha_{xxx} g_{u^\alpha_{xx}}(t) + ... = \\ &= (D_x g)(t). \end{aligned}$$

Then, (2.98) can be written as

$$\begin{aligned} &\int_0^1 [tf_{\beta,u^\alpha}(t)u^\beta + f_\alpha(t)]dt - \int_0^1 t[(D_x f_{\beta,u^\alpha_x})(t)u^\beta + f_{\beta,u^\alpha_x}(t)u^\beta_x]dt + \\ &+ \int_0^1 t[(D_x^2 f_{\beta,u^\alpha_{xx}})(t)u^\beta + 2(D_x f_{\beta,u^\alpha_{xx}})(t)u^\beta_x + f_{\beta,u^\alpha_{xx}}(t)u^\beta_{xx}]dt = \end{aligned}$$

$$\begin{aligned} = &\int_0^1 [tf_{\beta,u^\alpha}(t)u^\beta + f_\alpha(t)]dt + u^\beta \int_0^1 t \underbrace{[-(D_x f_{\beta,u^\alpha_x})(t) + (D_x^2 f_{\beta,u^\alpha_{xx}})(t)]}_{=f_{\alpha,u^\beta}(t) - f_{\beta,u^\alpha}(t)} dt + \\ &+ u^\beta_x \int_0^1 t \underbrace{[-f_{\beta,u^\alpha_x}(t) + 2(D_x f_{\beta,u^\alpha_{xx}})(t)]}_{=f_{\alpha,u^\beta_x}(t)} dt + u^\beta_{xx} \int_0^1 t \underbrace{f_{\beta,u^\alpha_{xx}}(t)}_{=f_{\alpha,u^\beta_{xx}}(t)} dt, \end{aligned}$$

where we used the Helmholtz expressions in (2.68) and the corresponding Helmholtz conditions. Therefore, (2.98) can be written as

$$\begin{aligned} &\int_0^1 [tf_{\beta,u^\alpha}(t)u^\beta + f_\alpha(t)]dt + u^\beta \int_0^1 t[f_{\alpha,u^\beta}(t) - f_{\beta,u^\alpha}(t)]dt + u^\beta_x \int_0^1 tf_{\alpha,u^\beta_x}(t)dt + \\ &\qquad + u^\beta_{xx} \int_0^1 tf_{\alpha,u^\beta_{xx}}(t)dt = \\ &= \int_0^1 \frac{d}{dt}[tf_\alpha(t)]dt = f_\alpha(x,u^\beta,u^\beta_x,u^\beta_{xx}) \end{aligned}$$

and we have proven Lemma 2.7.3. □

Together with the previous section, we have proven the local exactness of the sequence

$$0 \to \mathbb{R} \to C^\infty(J^k E) \xrightarrow{D_x} \{L\} \xrightarrow{\mathcal{E}_\alpha} \{f_\alpha\} \xrightarrow{\mathcal{H}^\gamma_{\alpha\beta}} \{H_{\alpha\beta},H^x_{\alpha\beta},H^{xx}_{\alpha\beta}\} \longrightarrow ....$$

2.8. Conservation Laws

So far, we explained a lot about variational equations and we discussed the most important concepts and objects. Now we want to talk about something different, namely conservation laws.

Definition 2.8.1. *Let f_α be a differential expression with corresponding source form $\Delta = f_\alpha du^\alpha \wedge dx$. A* ***conservation law for*** *f_α is a horizontal differential 1-form $Fdx \in \Omega^1(J^kE)$ (we think of a function F in local coordinates), which satisfies the following two properties*

$$\begin{cases} F(x,u^\beta,u^\beta_x,...,u^\beta_{(k)}) = D_xC, & \text{for some function } C \in C^\infty(J^kE), \\ F(x,u^\beta,u^\beta_x,...,u^\beta_{(k)}) = 0, & \text{whenever } f_\alpha = 0. \end{cases} \tag{2.99}$$

Note that F is not a function on J^kE because when we change coordinates then

$$F = D_xC = (D_yC)\frac{\partial y}{\partial x} \neq D_y\tilde{C}$$

in general. Therefore, we need an object which is invariant under coordinate transformation and the differential form Fdx does exactly what we want.

For example, let us consider the coordinate system (x,u) and (y,v) and the coordinate transformation $x = e^y$ for $x \in \mathbb{R}_+$ and $u = v$ on $\mathbb{R}$. Then $\frac{\partial y}{\partial x} = \frac{\partial \ln x}{\partial x} = \frac{1}{x} = e^{-y}$. Let $C = u = v$, then $D_xC = u_x = v_ye^{-y}$ is not a total derivative in the (y,v)-coordinate system, but in the (x,u)-coordinate system (see Proposition 2.4.3 for coordinate transformations).

In the above definition, we could also say that F is a trivial Lagrangian, which vanishes on solutions of $f_\alpha = 0$. This is also the definition for PDEs, where conservation laws are divergence expressions and these are also the trivial Lagrangians.

The more **classical definition of conservation law** is a divergence expression D_xC, which vanishes on solutions of $f_\alpha = 0$, that is, $\mathrm{pr}^k\sigma^*(D_xC) = 0$ whenever $\mathrm{pr}^k\sigma^* f_\alpha = 0$. For further details, see (Olv86), especially on page 265 and the following ones. The classical definition is slightly differently compared to Definition 2.8.1, since a differential equation may not always have a solution, but Definition 2.8.1 may still hold. For example, solutions of the differential equation

$$(f_\alpha) = \begin{pmatrix} u - u_x \\ u - 1 \end{pmatrix} = 0 \tag{2.100}$$

must satisfy $u = ce^x, c \in \mathbb{R}$, because of the first entry in (f_α) and $u \equiv 1$ because of the second entry. Both conditions will never be satisfied (except for one very special value x). But

$$u_x(u-1) = D_x(\frac{1}{2}u^2 - u)$$

is a conservation law in the sense of Definition 2.8.1 (the points in J^1E, which satisfy (2.100), are $u = 1, u_x = 1$ and all x). Even more cases have to be investigated. For example, let us consider

$$f = 1 + u_x^2.$$

Then $f = 0$ has no (real) solution, either in J^1E or as differential equation $\mathrm{pr}^k\sigma^* f = 0$. But there is a formal conservation law of the form

$$Qf = u_{xx}(1 + u_x^2) = D_x(u_x + \frac{1}{3}u_x^3),$$

where $Q := u_{xx}$ and Definition 2.8.1 still holds (this example also works with $f = 1 + u^2$ and $Q = u_x$).

We say **we are on the equation** $f_\alpha = 0$ if we mean the set of points in J^kE given by a section σ, or let us say by a solution, such $\mathrm{pr}^k\sigma^* f_\alpha = 0$. If we have a conservation law for a differential equation $f_\alpha = 0$, we get $D_xC = 0$ on the equation and this can be integrated very easily and forces $C = const.$ on the equation. We also say that $C = const.$ is a **first integral** for the differential equation.

If we have $F = Q^\alpha f_\alpha = D_xC$ for some generalized vector field $Q^\alpha \partial_{u^\alpha}$ on J^kE, then both conditions in (2.99) are satisfied. We call Q^α the **characteristics of the conservation law**, see (Olv86, p.270) (recall the transformation property of f_α with corresponding source form $\Delta = f_\alpha du^\alpha \wedge dx$). For example, let $f = u + u_{xx}$, then we can choose $Q = u_x$ and we get

$$Qf = u_x(u + u_{xx}) = D_x(\frac{1}{2}u^2 + \frac{1}{2}u_x^2) \quad \text{for all points in } J^kE.$$

It turns out that the special form $F = Q^\alpha f_\alpha = D_xC$ is what we will need in the following. Therefore, we formulate a second definition of conservation law as follows:

Definition 2.8.2. *Let f_α be a differential expression with corresponding source form $\Delta = f_\alpha du^\alpha \wedge dx$. A* ***conservation law for f_α in characteristic form*** *is a horizontal differential 1-form $F dx \in \Omega^1(J^kE)$ (we think of a function F in local coordinates), which satisfies the following two properties*

$$\begin{cases} F(x,u^\beta,u_x^\beta,\ldots,u_{(k)}^\beta) = D_xC, & \textit{for some function } C \in C^\infty(J^kE), \\ F(x,u^\beta,u_x^\beta,\ldots,u_{(k)}^\beta) = Q^\alpha f_\alpha, & \textit{for some generalised vector field } Q^\alpha\partial_{u^\alpha} \textit{ on } J^kE. \end{cases}$$

See (Olv86, p.270). In the following, we will always use this definition instead of Definition (2.8.1) and we simply call it conservation law for f_α. Let us explain why this restriction is reasonable:

- The best answer is probably that, later, we will have even further restrictions and we will only allow characteristics of the form $Q^\alpha = V^\alpha - u_x^\alpha V^x$, where

$V = V^x\partial_x + V^\alpha\partial_{u^\alpha}$ is a projectable vector field on E. The reason is that Noether's first theorem does not allow any more freedom (in the classical formulation with projectable vector fields) and it therefore only makes sense to consider such characteristics in Takens' problem (in the classical formulation with projectable vector fields). We will explain this later in more detail, when we have discussed Noether's theorems.

- One can find an explanation in Peter Olver's book (Olv86, p.270) (especially Exercise 2.33. therein). The statement is the following: We can prove the special form $F = Q^\alpha f_\alpha$ if we assume that f_α is totally non-degenerate (in this case, we do not necessarily require the special form $Q^\alpha = V^\alpha - u_x^\alpha V^x$). For the definition of totally non-degenerate see Definition 2.83 in (Olv86, p.171). We will explain in *Appendix D* how to prove the special form $F = Q^\alpha f_\alpha$ in this case.

Now we want to find a way how to define a conservation law in an intrinsic way. That is, we rewrite Definition 2.8.2. Let $Q = Q^\alpha\partial_{u^\alpha}$ be a generalized π-vertical vector field. We consider

$$\begin{aligned}\mathrm{pr}^k\sigma^*(dC) = \mathrm{pr}^k\sigma^*(C_x dx + C_u du + C_{u_x} du_x + ...) = \mathrm{pr}^k\sigma^*[(D_xC)dx] =\\ = \mathrm{pr}^k\sigma^*[\iota_{\mathrm{pr}Q}\Delta] = \mathrm{pr}^k\sigma^*[Q^\alpha f_\alpha dx] \quad \text{for all } \sigma \in \Gamma(E)\end{aligned}$$

and this equation tells us how to define it:

Definition 2.8.3. *A source form Δ on J^kE admits a conservation law, if there exists a generalized π-vertical vector field $Q = Q^\alpha\partial_{u^\alpha}$ and a function $C \in C^\infty(J^kE)$, such that*

$$pr^k\sigma^*(\iota_{prQ}\Delta - dC) = 0$$

for all sections σ on E (we can also say that $\iota_{prQ}\Delta - dC$ is a contact form).

Note that conservation laws have the structure of a vector space. This will be important later, since they are connected to symmetry vector fields which also have the structure of a vector space.

Also note that there are different kinds of trivial conservation laws. For example, $c \in \mathbb{R}$ is always a trivial conservation law for any differential equation, since $D_xc = 0$ is always satisfied, whether we are on the equation or not. Another kind of triviality is that $D_x^l f_\alpha$, where $l \geq 1$, is always a conservation law in the classical definition of conservation law. Further information can be found in (Olv86) on pages 268-270.

2.9. Symmetries and the ECS

Finally, we are able to formulate Takens' problem in the notation we have developed so far. We only need one more definition, to clarify what we mean with symmetry of a differential equation or source form.

Intuitively, a symmetry of a differential equation $f = 0$ is a transformation T, such that Tu is again a solution whenever u is a solution and with solution we actually mean a section $(x,u^\alpha(x))$. This definition alone would immediately cause some problems, since do we mean solutions for some initial or boundary problems or do we not restrict to such additional constraints? Therefore and for other reasons, we define a symmetry of a differential equation or source form slightly differently, actually as it is usually defined. Formally, it is defined as follows: We take any section or formal object u, not necessarily a solution, and if

$$f(u) = f(Tu) \quad \text{for all } u, \tag{2.101}$$

then we call the transformation T a symmetry of f. For many problems, this definition is not sufficient and we need a slightly weaker condition, namely

$$Tf(u) = f(Tu) \quad \text{for all } u. \tag{2.102}$$

That is, the transformation T also induces a transformation on f in a certain way (which can differ from the transformation on u). Functions f which satisfy (2.102) are also called **equivariant** with respect to the transformation T. Here, f, u and T are formal objects. Actually, we think of a Lie group, where g is an element in the Lie group and $T = g$. Then $Tu = g \cdot u$ is a group action and f is a function or differential form. In the case of differential equations, usually, we have to prolong the group action. Note that if (2.101) holds for all u then it of course also holds for solutions. Sometimes it is too complicated to consider symmetries only for solutions, because this would mean we need to have knowledge about the solutions and it can be very hard to solve differential equations or find even any properties of the solutions. From a physical point of few, the definition for all u in (2.101) or (2.102) also makes sense, since differential equations describe physical laws and the physical laws itself should have symmetries and not only the solutions. In fundamental physical laws and differential equations, we have actually almost always the situation in (2.102) (it depends on how the laws are formulated). Especially, when we are considering source forms and symmetries of source forms then f_α does not satisfy (2.101) rather a version of the condition (2.102).

Now let us define what we mean with symmetry precisely and one has to convince oneself that it is equivalent to the formal definition in (2.101) or (2.102).

Definition 2.9.1. *A projectable vector field V on E, such that $\mathcal{L}_{prV}\Delta = 0$ (for all points in J^kE), is called a **symmetry of** Δ.*

We also have to keep in mind that $\mathcal{L}_{\mathrm{pr}V}\Delta = 0$ is not equivalent to $\mathcal{L}_{\mathrm{pr}V} f_\alpha = 0$ and that the source form induces a transformation for f_α. For example, $f = u + u_{xx}$ satisfies the symmetry $x\partial_x$, but the corresponding source form $\Delta = (u + u_{xx})du \wedge dx$ does not satisfy this symmetry.

Usually, symmetries are of course described by Lie groups and Lie algebras. However, Takens' problem can be solved without having the structure of a Lie algebra and this is important to note here (see Theorem 1.0.2). But let us say a few more words about Lie algebras. It turns out that the set of all symmetries of Δ is a Lie algebra. To prove this, we basically have to show that when V,W are projectable vector fields on E, then $[V,W]$ is also a projectable vector field on E and that $[\mathrm{pr}V,\mathrm{pr}W] = \mathrm{pr}[V,W]$. Also see Lemma 4 in (Kru15, p.172). Since we will not need this structure to prove Theorem 1.0.2 and 1.0.3 this exercise is left to the reader.

Let us briefly say a few words about the special form of $Q^\alpha = V^\alpha - u_x^\alpha V^x$, why the conservation laws are connected to the symmetries and if Theorem 1.0.2 makes sense from a physical point of few. Since we know from Noether's theorem (see next section) that variational equations $f_\alpha = \mathcal{E}_\alpha L$ have conservation laws of such a form, where the characteristics are $Q^\alpha = V^\alpha - u_x^\alpha V^x$ (when using projectable vector fields), it is absolutely necessary, from a pure mathematical point of few, to consider this special form of characteristics. More precisely, in Noether's theorem the invariance of Ldx leads to

$$\begin{aligned} 0 = \mathcal{L}_{\mathrm{pr}V}(Ldx) &= [Q^\alpha L_{u^\alpha} + (D_x Q^\alpha) L_{u_x^\alpha} + D_x(V^x L)]dx = \\ &= [Q^\alpha \mathcal{E}_\alpha L + D_x(Q^\alpha L_{u_x^\alpha} + V^x L)]dx = \\ &= (Q^\alpha f_\alpha + D_x C)dx, \end{aligned}$$

where

$$C := Q^\alpha L_{u_x^\alpha} + V^x L = V^\alpha L_{u_x^\alpha} + V^x(L - u_x^\alpha L_{u_x^\alpha})$$

is the conserved quantity. For example, in classical mechanics $V^x(L - u_x^\alpha L_{u_x^\alpha})$ describes the kinetic energy (Legendre transformation of L) and $V^\alpha L_{u_x^\alpha}$, describes momentum- and angular momentum. However, at this point it is not clear if the relation $Q^\alpha = V^\alpha - u_x^\alpha V^x$ also makes sense in general and what kind of meaningful conservation laws we can assume in physics. We will discuss this later in Section 4.3.

To be able to prove Theorem 1.0.2, we need the following lemma, which we only derive for vertical vector fields on E. Actually, we need the identity (2.103) in the Lemma 2.9.2 for projectable vector fields on E and the more general version can be found in (AP94, p.202) in Theorem 2.6, or see the remark below.

Lemma 2.9.2. *Let $\Delta = f_\alpha du^\alpha \wedge dx$ be a second order source form defined on J^kE and $V = V^\beta \partial_{u^\beta}$ be a vertical vector field on E. Then*

$$\mathcal{L}_{prV}\Delta = [\mathcal{E}_\alpha(Q^\beta f_\beta) + H_{\alpha\beta}Q^\beta + H_{\alpha\beta}^x(D_x Q^\beta) + H_{\alpha\beta}^{xx}(D_x^2 Q^\beta)]du^\alpha \wedge dx, \qquad (2.103)$$

where $\mathcal{E}_\beta$ is the Euler-Lagrange operator, $H_{\alpha\beta}$,$H^x_{\alpha\beta}$,$H^{xx}_{\alpha\beta}$ are the Helmholtz expressions and $Q^\beta = V^\beta$ is the corresponding characteristic of V.

<u>Proof:</u> Since we are only considering vertical vector fields, we have $V^\alpha = Q^\alpha$ and

$$\begin{aligned}\mathrm{pr}^2 V &= V^\alpha \partial_{u^\alpha} + (D_x V^\alpha)\partial_{u^\alpha_x} + (D_x^2 V^\alpha)\partial_{u^\alpha_{xx}} = \\ &= Q^\alpha \partial_{u^\alpha} + (D_x Q^\alpha)\partial_{u^\alpha_x} + (D_x^2 Q^\alpha)\partial_{u^\alpha_{xx}}.\end{aligned}$$

We get

$$\begin{aligned}\mathcal{L}_{\mathrm{pr}V}\Delta =&\iota_{\mathrm{pr}V} d\Delta + d(\iota_{\mathrm{pr}V}\Delta) = \\ =&\iota_{\mathrm{pr}V}\Big(f_{\alpha,u^\beta} du^\beta \wedge du^\alpha \wedge dx + f_{\alpha,u^\beta_x} du^\beta_x \wedge du^\alpha \wedge dx + \\ &\qquad + f_{\alpha,u^\beta_{xx}} du^\beta_{xx} \wedge du^\alpha \wedge dx\Big) + d(Q^\beta f_\beta dx) = \\ =&\Big(f_{\alpha,u^\beta} Q^\beta + f_{\alpha,u^\beta_x} D_x Q^\beta + f_{\alpha,u^\beta_{xx}} D_x^2 Q^\beta\Big) du^\alpha \wedge dx - \\ &- Q^\alpha \Big(f_{\alpha,u^\beta} du^\beta \wedge dx + f_{\alpha,u^\beta_x} du^\beta_x \wedge dx + f_{\alpha,u^\beta_{xx}} du^\beta_{xx} \wedge dx\Big) + \\ &+ 0 \quad (\text{since } \iota_{\mathrm{pr}V} dx = 0 \text{ for vertical vector fields } V) \\ &+ \partial_{u^\alpha}(Q^\beta f_\beta) du^\alpha \wedge dx + Q^\beta f_{\beta,u^\alpha_x} du^\alpha_x \wedge dx + Q^\beta f_{\beta,u^\alpha_{xx}} du^\alpha_{xx} \wedge dx. \end{aligned} \tag{2.104}$$

All terms with a du^β_x or du^β_{xx} basis element are canceling out (This is surely the case, since if Δ is a source form then $\mathcal{L}_{\mathrm{pr}V}\Delta$ is also of source form type). We can also change some indices α,β in (2.104) and then it becomes

$$\begin{aligned}\mathcal{L}_{\mathrm{pr}V}\Delta = &\left(f_{\alpha,u^\beta} - f_{\beta,u^\alpha}\right) Q^\beta du^\alpha \wedge dx + \partial_{u^\alpha}(Q^\beta f_\beta) du^\alpha \wedge dx + \\ &+ \Big(f_{\alpha,u^\beta_x} D_x Q^\beta + f_{\alpha,u^\beta_{xx}} D_x^2 Q^\beta\Big) du^\alpha \wedge dx.\end{aligned} \tag{2.105}$$

Now we will rewrite the second line in (2.105). Before we continue with rewriting (2.105), we have to derive two identities, with a sort of partial integration technique. Let a,b functions on $J^k E$. Then, the first identity is formally (without summation) of the form

$$aD_x b = aD_x b + aD_x b - D_x(ab) + bD_x a$$

and in our case, with summation over β,

$$\begin{aligned}f_{\alpha,u^\beta_x} D_x Q^\beta &= f_{\alpha,u^\beta_x} D_x Q^\beta + f_{\beta,u^\alpha_x} D_x Q^\beta - f_{\beta,u^\alpha_x} D_x Q^\beta = \\ &= f_{\alpha,u^\beta_x} D_x Q^\beta + f_{\beta,u^\alpha_x} D_x Q^\beta - D_x\left(f_{\beta,u^\alpha_x} Q^\beta\right) + Q^\beta D_x f_{\beta,u^\alpha_x}.\end{aligned}$$

The second identity is formally of the form

$$\begin{aligned}aD_x^2 b &= aD_x^2 b - aD_x^2 b + aD_x^2 b = \\ &= (a-a)D_x^2 b + D_x^2(ab) - 2(D_x a)(D_x b) - bD_x^2 a\end{aligned}$$

and in our case, with summation over β,

$$\begin{aligned}
&f_{\alpha,u^\beta_{xx}} D_x^2 Q^\beta = \\
=&f_{\alpha,u^\beta_{xx}} D_x^2 Q^\beta - f_{\beta,u^\alpha_{xx}} D_x^2 Q^\beta + f_{\beta,u^\alpha_{xx}} D_x^2 Q^\beta = \\
=&\left(f_{\alpha,u^\beta_{xx}} - f_{\beta,u^\alpha_{xx}}\right) D_x^2 Q^\beta + D_x^2\left(f_{\beta,u^\alpha_{xx}} Q^\beta\right) - 2\left(D_x f_{\beta,u^\alpha_{xx}}\right) D_x Q^\beta - \left(D_x^2 f_{\beta,u^\alpha_{xx}}\right) Q^\beta.
\end{aligned}$$

Now we use these identities and (2.105) becomes

$$\begin{aligned}
\mathcal{L}_{\mathrm{pr}V}\Delta = &\left(f_{\alpha,u^\beta} - f_{\beta,u^\alpha}\right) Q^\beta du^\alpha \wedge dx + \partial_{u^\alpha}(Q^\beta f_\beta) du^\alpha \wedge dx + \\
&+ \left[f_{\alpha,u^\beta_x} D_x Q^\beta + f_{\beta,u^\alpha_x} D_x Q^\beta - D_x\left(f_{\beta,u^\alpha_x} Q^\beta\right) + Q^\beta D_x f_{\beta,u^\alpha_x}\right] du^\alpha \wedge dx + \\
&+ \Big[\left(f_{\alpha,u^\beta_{xx}} - f_{\beta,u^\alpha_{xx}}\right) D_x^2 Q^\beta + D_x^2\left(f_{\beta,u^\alpha_{xx}} Q^\beta\right) - 2\left(D_x f_{\beta,u^\alpha_{xx}}\right) D_x Q^\beta - \\
&\qquad - \left(D_x^2 f_{\beta,u^\alpha_{xx}}\right) Q^\beta\Big] du^\alpha \wedge dx.
\end{aligned} \tag{2.106}$$

Since $Q^\alpha = Q^\alpha(x,u^\beta)$ for vertical V, and therefore

$$\begin{aligned}
f_{\beta,u^\alpha_x} Q^\beta &= \partial_{u^\alpha_x}(f_\beta Q^\beta) \quad \text{and} \\
f_{\beta,u^\alpha_{xx}} Q^\beta &= \partial_{u^\alpha_{xx}}(f_\beta Q^\beta),
\end{aligned} \tag{2.107}$$

we can write (2.106) as

$$\begin{aligned}
\mathcal{L}_{\mathrm{pr}V}\Delta = &\left[\left(\partial_{u^\alpha} - D_x \partial_{u^\alpha_x} + D_x^2 \partial_{u^\alpha_{xx}}\right)(Q^\beta f_\beta)\right] du^\alpha \wedge dx + \\
&+ \left[f_{\alpha,u^\beta} - f_{\beta,u^\alpha} + \left(D_x f_{\beta,u^\alpha_x}\right) - \left(D_x^2 f_{\beta,u^\alpha_{xx}}\right)\right] Q^\beta du^\alpha \wedge dx + \\
&+ \left[f_{\alpha,u^\beta_x} + f_{\beta,u^\alpha_x} - 2D_x\left(f_{\beta,u^\alpha_{xx}}\right)\right] D_x Q^\beta du^\alpha \wedge dx + \\
&+ \left(f_{\alpha,u^\beta_{xx}} - f_{\beta,u^\alpha_{xx}}\right) D_x^2 Q^\beta du^\alpha \wedge dx = \\
\\
=&\mathcal{E}_\alpha(Q^\beta f_\beta) du^\alpha \wedge dx + \\
&+ H_{\alpha\beta} Q^\beta du^\alpha \wedge dx + H^x_{\alpha\beta}(D_x Q^\beta) du^\alpha \wedge dx + H^{xx}_{\alpha\beta}(D_x^2 Q^\beta) du^\alpha \wedge dx,
\end{aligned}$$

where $\mathcal{E}_\alpha$ is the Euler-Lagrange operator for second order ODEs and $H_{\alpha\beta}$,$H^x_{\alpha\beta}$,$H^{xx}_{\alpha\beta}$ are the Helmholtz expressions defined in (2.68). $\square$

Remark: Note that for projectable vector fields $V \in \mathfrak{X}(E)$ we can also write

$$\mathrm{pr}^2 V = V^x D_x + Q^\alpha \partial_{u^\alpha} + (D_x Q^\alpha)\partial_{u^\alpha_x} + (D_x^2 Q^\alpha)\partial_{u^\alpha_{xx}}$$

and

$$\begin{aligned}
\mathcal{L}_{\mathrm{pr}V}\Delta &= (\mathcal{L}_{\mathrm{pr}V} f_\alpha) du^\alpha \wedge dx + f_\alpha \mathcal{L}_{\mathrm{pr}V}(du^\alpha \wedge dx) = \\
&= (\mathcal{L}_{\mathrm{pr}V} f_\alpha) du^\alpha \wedge dx + f_\alpha d[\iota_{\mathrm{pr}V}(du^\alpha \wedge dx)] = \\
&= (\mathcal{L}_{\mathrm{pr}V} f_\alpha) du^\alpha \wedge dx + f_\alpha d(V^\alpha dx - V^x du^\alpha) = \\
&= (\mathcal{L}_{\mathrm{pr}V} f_\alpha) du^\alpha \wedge dx + f_\alpha (V^\alpha_{u^\beta} du^\beta \wedge dx + V^x_x du^\alpha \wedge dx) = \\
&= [V^x D_x f_\alpha + Q^\beta f_{\alpha,u^\beta} + (D_x Q^\beta) f_{\alpha,u^\beta_x} + (D^2_x Q^\beta) f_{\alpha,u^\beta_{xx}}] du^\alpha \wedge dx + \\
&\qquad\qquad + f_\alpha (V^\alpha_{u^\beta} du^\beta \wedge dx + V^x_x du^\alpha \wedge dx) = \\
&= [D_x(V^x f_\alpha) + Q^\beta f_{\alpha,u^\beta} + (D_x Q^\beta) f_{\alpha,u^\beta_x} + (D^2_x Q^\beta) f_{\alpha,u^\beta_{xx}}] du^\alpha \wedge dx + \\
&\qquad\qquad + [\partial_{u^\beta}(f_\alpha V^\alpha) - f_{\alpha,u^\beta} V^\alpha] du^\beta \wedge dx = \\
&= [Q^\beta (f_{\alpha,u^\beta} - f_{\beta,u^\alpha}) + (D_x Q^\beta) f_{\alpha,u^\beta_x} + (D^2_x Q^\beta) f_{\alpha,u^\beta_{xx}}] du^\alpha \wedge dx + \\
&\qquad\qquad + D_x(V^x f_\alpha) du^\alpha \wedge dx + \underbrace{[\partial_{u^\beta}(f_\alpha V^\alpha) - f_{\alpha,u^\beta} u^\alpha_x V^x]}_{=\partial_{u^\beta}(f_\alpha Q^\alpha)} du^\beta \wedge dx
\end{aligned}$$

Then we do similar reformulations as we have done above, when we applied a kind of partial integration. The difference is that the identities in (2.107) will be slightly differently and we will get, for example,

$$\begin{aligned}
(D_x Q^\beta) f_{\beta,u^\alpha_x} &= D_x(Q^\beta f_{\beta,u^\alpha_x}) - Q^\beta D_x f_{\beta,u^\alpha_x} = \\
&= D_x \partial_{u^\alpha_x}(Q^\beta f_\beta) + D_x(V^x f_\alpha) - Q^\beta D_x f_{\beta,u^\alpha_x}
\end{aligned}$$

and the term in the middle in the last line will cancel with other terms.

We need another proposition, to be able to prove Theorem 3.1.1.

Proposition 2.9.3. *Let V be a projectable vector field on E, such that $\mathcal{L}_{prV}\Delta = 0$. Furthermore, we have a corresponding conservation laws, i.e. $\mathcal{E}_\alpha(Q^\beta f_\beta) = \mathcal{E}_\alpha(D_x C) = 0$. Then*

$$Q^\beta H_{\alpha\beta} + (D_x Q^\beta) H^x_{\alpha\beta} + (D^2_x Q^\beta) H^{xx}_{\alpha\beta} = 0. \tag{2.108}$$

The proof easily follows from the identity in (2.103), the fact that the Euler-Lagrange operator $\mathcal{E}_\alpha$ annihilates total derivatives, and by the assumption that we assume $Q^\alpha f_\alpha$ is a total derivative, i.e. a corresponding conservation law.

For a set of symmetries $\{V\}$, equations (2.108) are the key to solve Takens' problem and we call them the **equations of conservation laws and symmetries (ECS)**. Already Takens derived these equations and they can also be found in a paper of Ian M. Anderson and Juha Pohjanpelto in a more general version (AP94). For arbitrary n,m, the ECS is

$$Q^\beta H_{\alpha\beta} + (D_i Q^\beta) H^i_{\alpha\beta} + (D_{ij} Q^\beta) H^{ij}_{\alpha\beta} = 0,$$

where we also have summation over i,j. This equation can slightly be simplified and it can be written as

$$Q^{\beta}_{;I} H^{I}_{\alpha\beta} = 0,$$

where I is a multi index of length $|I| \leq 2$ and $;I$ denotes total derivative $Q^{\beta}_{;I} = D_I Q^{\beta}$.

2.10. Noether's First and Second Theorem

This section is not relevant for the proof of Theorem 1.0.2 and 1.0.3, it rather provides further information and connections to other fields in mathematics and physics. Solving Takens' problem can be considered as inverting Noether's theorem and we want to explain this in more detail in this section. In (AP94, p.192) we can find the following theorem (we changed it slightly and the theorem is only formally true):

Generalized Noether's Theorem (AP94, p.192): Any two of the following three statements imply the third:
(A) Δ is invariant, i.e. $\mathcal{L}_{prV}\Delta = 0$
(B) The symmetry vector fields V generate corresponding conservation laws.
(C) Δ is locally variational.

Takens' problem is to show (A),(B) $\Rightarrow$ (C) and Noether's theorem are basically all the other possibilities. In the original paper of Noether (Noe18), (B),(C) $\Rightarrow$ (A) is called *reversal* (in German: Umkehrung) and from physical point of view, (A),(C) $\Rightarrow$ (B) is probably the most interesting direction, at least it is what is usually taught when studying physics. It is probably also the simplest direction. Note that (A) is equivalent to the invariance of the first variation δI and in the original work of Noether, she only considered the invariance of I, which is slightly weaker. The invariance of δI implies the invariance of I, but not in reversal direction.

In each cases in Noether's work, we have the so-called first and second Noether's theorem. Roughly speaking, the first is about finite dimensional Lie groups and the second is about infinite dimensional Lie groups. We will now formulate (A),(C) $\Rightarrow$ (B) for finite- and infinite dimensional Lie groups. In both cases we will formulate the theorems only formally.

Noether's first Theorem (finite dimensional): If Ldx (or δI) is invariant under projectable symmetry vector fields (and therefore f_α is variational), then there exist conservation laws and they can explicitly be described.

Proof: We only prove the case where I, or equivalently Ldx, satisfies these symmetries. For simplicity, we assume first order Lagrangians $L = L(x,u,u_x)$. Then the proof is to apply the Lie derivative and do one time partial integration, i.e.

$$\begin{aligned} 0 = \mathcal{L}_{\mathrm{pr}V}(Ldx) &= [Q^\alpha L_{u^\alpha} + (D_x Q^\alpha) L_{u_x^\alpha} + D_x(V^x L)]dx = \\ &= [Q^\alpha \mathcal{E}_\alpha L + D_x(Q^\alpha L_{u_x^\alpha} + V^x L)]dx = \\ &= (Q^\alpha f_\alpha + D_x C)dx, \end{aligned} \tag{2.109}$$

where we define the conserved quantity C as

$$C = Q^\alpha L_{u_x^\alpha} + V^x L = V^\alpha L_{u_x^\alpha} + V^x(L - u_x^\alpha L_{u_x^\alpha}). \tag{2.110}$$

This proves the formal version of Noether's first theorem. □

Note that equation (2.109) is pretty similar to the ECS. Noether's second theorem needs the definition of so-called *differential identities*. Let $r \in \mathbb{N}_0$ be some fixed integer. Then any equation of the form

$$\sum_{l=0}^{r} A_l^\alpha D_x^l f_\alpha = 0 \quad \text{for all points in } J^k E,$$

where A_l^α are certain functions on $J^k E$, is called **differential identity for** f_α. The definition is relative to some local coordinates and we do not further show how to get a coordinate independent formulation (this must be done in more concrete examples as, for example, for Maxwell's equations, see below). Let us briefly consider PDEs, to show an interesting example. Let $n = m$, then $i = 1,2,...,n$ and $\alpha = 1,2,...,n$ and we write f^i instead of f_α. Then $D_i f^i = \mathrm{div}(f^i) = 0$ is a differential identity for (f^i) and it is called: (f^i) is *divergence-free.*

For example, Maxwell's equations in vacuum are

$$D_i F^{ij} = 0, \quad D_i \tilde{F}^{ij} = 0,$$

where $\tilde{F}^{ij}$ is the dual field strength tensor of F^{ij} and we get the differential identities $D_j D_i F^{ij} = 0$ and $D_j D_i \tilde{F}^{ij} = 0$, since D_{ij} is symmetric in i,j and F^{ij},$\tilde{F}^{ij}$ are skew-symmetric in i,j. Actually, the definition of D_i is slightly differently in case of Maxwell's equations and should be replaced by $(D_i) = (D_t,D_x,D_y,D_z)$ and $(D^i) = (D_t, -D_x, -D_y, -D_z)$, i.e. we have to use the relativistic 4-gradient in Minkowski space (this will be needed to define F^{ij}). Using differential forms, we can also write Maxwell's equations in vacuum as

$$dF = 0, \quad d * F = 0,$$

where $*$ is the Hodge-star-operator and then it is immediately clear that $ddF = 0$ and $d(d * F) = 0$. For the definition of F and further details see (AF01). In the case of Maxwell's equations, these differential identities describe *charge conservation.*

Noether's second Theorem (infinite dimensional): If Ldx (or δI) is invariant under projectable symmetry vector fields $V = V^x\partial_x + V^\alpha\partial_{u^\alpha}$ (and therefore f_α is variational) and V^x,V^α depend linearly on arbitrary functions $p = p(x)$ and their derivatives, i.e

$$V^x = V^x(x,p(x),p_x(x),...),$$
$$V^\alpha = V^\alpha(x,u^\beta,p(x),p_x(x),...),$$

then there are differential identities for f_α and they can explicitly be described.

Note that if the projectable symmetry vector field $V = V^\alpha\partial_{u^\alpha} + V^x\partial_x$ depends linearly on p,p_x,... then $Q^\alpha = V^\alpha - u^\alpha_x V^x$ also depends linearly on p,p_x,....

Proof: For simplicity, we assume first order Lagrangian $L = L(x,u^\beta,u^\beta_x)$ and V^x,V^α depend linearly on p,p_x. In the following, we will just write p,p_x, but we have to keep in mind that we actually mean $p(x)$,$p_x(x)$. We get

$$\begin{aligned} Q^\alpha(x,u^\beta,u^\beta_x,p,p_x) &= V^\alpha(x,u^\beta,p,p_x) - u^\alpha_x V^x(x,p,p_x) = \\ &= (pa^\alpha + p_x b^\alpha) - u^\alpha_x(pc + p_x d) = \\ &= (a^\alpha - u^\alpha_x c)p + (b^\alpha - u^\alpha_x d)p_x, \end{aligned}$$

where a^α,b^α,c,d are functions on E. Then we derive the same identity as in (2.109) from Noether's first theorem (without writing dx) and we can write

$$\begin{aligned} 0 &= Q^\alpha f_\alpha + D_x C = \\ &= [(a^\alpha - u^\alpha_x c)p + (b^\alpha - u^\alpha_x d)p_x]f_\alpha + D_x C = \\ &= (a^\alpha - u^\alpha_x c)pf_\alpha + D_x[(b^\alpha - u^\alpha_x d)pf_\alpha] - pD_x[(b^\alpha - u^\alpha_x d)f_\alpha] + D_x C = \\ &= p\{(a^\alpha - u^\alpha_x c)f_\alpha - D_x[(b^\alpha - u^\alpha_x d)f_\alpha]\} + D_x[(b^\alpha - u^\alpha_x d)pf_\alpha + C]. \end{aligned} \tag{2.111}$$

In (2.110) we can observe that C depends linearly on p and their derivatives and the second term in (2.111) is a total derivative. Since p is arbitrary, we can choose p with compact support in U^0, and then we get

$$0 = \int_{U^0} \mathrm{pr}^k\sigma^* \left(p\{(a^\alpha - u^\alpha_x c)f_\alpha - D_x[(b^\alpha - u^\alpha_x d)f_\alpha]\}dx\right)$$

for all p with compact support in U^0 and all sections $\sigma \in \Gamma(E)$. Using Du Bois-Reymond's lemma shows that

$$0 = (a^\alpha - u^\alpha_x c)f_\alpha - D_x[(b^\alpha - u^\alpha_x d)f_\alpha] \tag{2.112}$$

for all points in J^kE and this is the differential identity for f_α. This is the idea of the proof of Noether's second theorem and a more general proof can be done in a similar way. □

Note that the second term in (2.111) must now also vanish and we get further differential identities. In the case of PDEs, where $n \geq 2$, we get a divergence equation at this point, which might be hard to solve in general. However, for $n = 1$, this equation is very simple and we can write it as

$$\begin{aligned} c_1 &= (b^\alpha - u_x^\alpha d) p f_\alpha + C \overset{(2.110)}{=} \\ &= p(b^\alpha - u_x^\alpha d) f_\alpha + Q^\alpha L_{u_x^\alpha} + V^x L = \\ &= p(b^\alpha - u_x^\alpha d) f_\alpha + [p(a^\alpha - u_x^\alpha c) + p_x(b^\alpha - u_x^\alpha d)] L_{u_x^\alpha} + (pc + p_x d) L = \\ &= p[(b^\alpha - u_x^\alpha d) f_\alpha + (a^\alpha - u_x^\alpha c) L_{u_x^\alpha} + cL] + p_x[(b^\alpha - u_x^\alpha d) L_{u_x^\alpha} + Ld], \end{aligned} \tag{2.113}$$

where $c_1 \in \mathbb{R}$. Since we can choose p arbitrary, the two summands in (2.113) must vanish separately and c_1 must also vanish. Then we get the two additional differential identities

$$0 = (b^\alpha - u_x^\alpha d) f_\alpha + (a^\alpha - u_x^\alpha c) L_{u_x^\alpha} + cL, \tag{2.114}$$

$$0 = (b^\alpha - u_x^\alpha d) L_{u_x^\alpha} + Ld. \tag{2.115}$$

Equation (2.115) is a very simple differential equation for L, with solution

$$L = \frac{1}{d}(b^\alpha - u_x^\alpha d) c_\alpha, \quad c_\alpha = c_\alpha(x, u^\beta)$$

in the case when $d \neq 0$ and $L = L(x, u^\beta)$ when $d = 0$. This shows that we get very strong restrictions for L and f_α. Now we could try to solve equation (2.114) and (2.112) to get more information, but we will stop the discussion here. It would also be interesting if (2.112),(2.114) and (2.115) are somehow dependent. It seems that this has probably not been discussed extensively in the literature and one is usually only interested in the identity (2.112). At least in the original paper of Emmy Noether (Noe18, p.243) there is only one remark, namely: *"... Aus (15) und (16) folgt noch $Div(B - \Gamma) = 0$..."*, which means in our notation the vanishing of the second term in (2.111).

There are two modifications of Noether's theorem. The first modification is to assume a weaker symmetry condition and the second is an implicit version of Noether's theorem.

<u>Weaker symmetries (because of equivalent Lagrangians):</u> Noether's first theorem basically says that we can compute the conserved quantity with the help of the Lagrangian L. However, we already know that any two Lagrangians L and $L + D_x\Lambda$ are equivalent. Therefore, the question is if we can also compute the conserved quantity with the help of the Lagrangian $L + D_x\Lambda$. To understand this modification, let us consider a trivial Lagrangian $L = D_x\Lambda$. Such a Lagrangian leads to the trivial

Euler-Lagrange expression $f_\alpha \equiv 0$ for all points in J^kE. Then, going back to the symmetry equation (2.109), we get

$$\mathcal{L}_{\mathrm{pr}V}(Ldx) = \mathcal{L}_{\mathrm{pr}V}(D_x\Lambda dx) = (D_xC)dx \overset{(2.110)}{=} D_x(Q^\alpha L_{u^\alpha_x} + V^x L)dx.$$

This equation shows that trivial Lagrangians always lead to a total derivative expression $(D_xC)dx$, when we apply the Lie derivative to Ldx. Therefore, we can use a **weaker symmetry** assumption in Noether's first theorem, namely

$$(D_xA)dx = \mathcal{L}_{\mathrm{pr}V}(Ldx) = (Q^\alpha f_\alpha + D_xC)dx,$$

for some function $A \in C^\infty(J^kE)$, where we now also allow $D_xA \neq 0$. This leads to the **modified conservation law**

$$Q^\alpha f_\alpha + D_x(C - A) = 0,$$

with modified conserved quantity $\tilde{C} = C - A$.

Lemma 2.10.1. *Let $(D_x\Lambda)dx$ be a trivial Lagrange form on J^kE and V a projectable vector field on E. Then $\mathcal{L}_{prV}(D_x\Lambda dx) = (D_xA)dx$ for some function A on J^kE.*

The interesting observation is here that we do not get a statement modulo contact forms. That is, we do not get a trivial Lagrange form plus a contact form, when applying the Lie derivative to a trivial Lagrange form, we rather get exactly a trivial Lagrange form.

Proof: We have to do the calculation in (2.109) for higher order Lagrangians. First, we need the identity

$$\begin{aligned}(D_x^kQ^\alpha)L_{u^\alpha_{(k)}} &= D_x[(D_x^{k-1}Q^\alpha)L_{u^\alpha_{(k)}}] - (D_x^{k-1}Q^\alpha)D_xL_{u^\alpha_{(k)}} = \\ &= D_x[(D_x^{k-1}Q^\alpha)L_{u^\alpha_{(k)}} - (D_x^{k-2}Q^\alpha)D_xL_{u^\alpha_{(k)}}] + (D_x^{k-2}Q^\alpha)D_x^2L_{u^\alpha_{(k)}} = \\ &= ... = \\ &= D_x[\sum_{l=0}^{k-1}(-1)^l(D_x^{k-1-l}Q^\alpha)D_x^lL_{u^\alpha_{(k)}}] + (-1)^kQ^\alpha D_x^kL_{u^\alpha_{(k)}} = \\ &= D_x[...] + (-1)^kQ^\alpha D_x^kL_{u^\alpha_{(k)}},\end{aligned}$$

for all $k \geq 1$, to compute

$$\begin{aligned}\mathcal{L}_{\mathrm{pr}V}(Ldx) &= [D_x(V^xL) + Q^\alpha L_{u^\alpha} + (D_xQ^\alpha)L_{u^\alpha_x} + ... + (D_x^kQ^\alpha)L_{u^\alpha_{(k)}}]dx = \\ &= [Q^\alpha\mathcal{E}_\alpha L + D_x(V^xL + Q^\alpha L_{u^\alpha_x} + ...)]dx = \\ &= (Q^\alpha f_\alpha + D_xA)dx,\end{aligned}$$

where

$$
\begin{aligned}
A &:= \\
&= V^x L + Q^\alpha L_{u^\alpha_x} + [(D_x Q^\alpha) L_{u^\alpha_{xx}} - Q^\alpha D_x L_{u^\alpha_{xx}}] + \ldots \sum_{l=0}^{k-1} (-1)^l (D_x^{k-1-l} Q^\alpha) D_x^l L_{u^\alpha_{(k)}} = \\
&= V^x L + \sum_{r=1}^{k-1} \sum_{l=0}^{r-1} (-1)^l (D_x^{r-1-l} Q^\alpha) D_x^l L_{u^\alpha_{(r)}}, \quad k \geq 1.
\end{aligned}
$$

If f_α vanishes for all points in $J^k E$, i.e. if $\mathcal{E}_\alpha L = 0$, or in other words, when $L = D_x \Lambda$, then $\mathcal{L}_{\mathrm{pr}V}(Ldx) = (D_x A)dx$ is a trivial Lagrange form. □

Implicit version of Noether's theorem: To understand this version of Noether's theorem, we consider the following lemma:

Lemma 2.10.2. *If Ldx (or I) is invariant under some projectable symmetry vector fields then the corresponding variational source form Δ (or δI) is also invariant under the same symmetry vector fields. The opposite direction is not true in general.*

Proof: For simplicity, we consider first order Lagrangians $L = L(x,u^\beta,u^\beta_x)$. Let $V = V^x \partial_x + V^\alpha \partial_{u^\alpha}$ be a projectable vector field on E such that $\mathcal{L}_{\mathrm{pr}V}(Ldx) = 0$. Then we can write

$$
\begin{aligned}
0 = \mathcal{L}_{\mathrm{pr}V}(Ldx) &= (\mathcal{L}_{\mathrm{pr}V} L)dx + L V^x_x dx = \\
&= [D_x(LV^x) + Q^\alpha L_{u^\alpha} + (D_x Q^\alpha) L_{u^\alpha_x}]dx = \\
&= [Q^\alpha \mathcal{E}_\alpha L + D_x(LV^x + Q^\alpha L_{u^\alpha_x})]dx = \\
&= [Q^\alpha f_\alpha + D_x C]dx, \qquad (2.116)
\end{aligned}
$$

where $C := LV^x + Q^\alpha L_{u^\alpha_x}$. Furthermore,

$$
\begin{aligned}
\mathcal{L}_{\mathrm{pr}V}\Delta &= [\mathcal{E}_\alpha(Q^\beta f_\beta) + Q^\beta H_{\alpha\beta} + (D_x Q^\beta) H^x_{\alpha\beta} + (D_x^2 Q^\beta) H^{xx}_{\alpha\beta}] du^\alpha \wedge dx = \\
&= [\mathcal{E}_\alpha(Q^\beta f_\beta)] du^\alpha \wedge dx, \qquad (2.117)
\end{aligned}
$$

what we have already derived in Lemma 2.9.2 (only for vertical V) and the Helmholtz expressions vanish, since Δ is variational. Since Ldx in (2.116) satisfies the symmetry condition, we get $Q^\alpha f_\alpha = -D_x C$ and if we plug this into (2.117) then we get

$$
\begin{aligned}
\mathcal{L}_{\mathrm{pr}V}\Delta &= [\mathcal{E}_\alpha(Q^\beta f_\beta)] du^\alpha \wedge dx = \qquad (2.118) \\
&= [\mathcal{E}_\alpha(-D_x C)] du^\alpha \wedge dx = 0, \qquad (2.119)
\end{aligned}
$$

since the Euler-Lagrange operator $\mathcal{E}_\alpha$ annihilates total derivatives $D_x(-C)$. Therefore, Δ satisfies the symmetry condition if Ldx satisfies the symmetry condition.

To show that the opposite direction is not true in general, we consider the following simple example: Let $Ldx = (D_x\Lambda)dx = xdx$ be a trivial Lagrangian with corresponding $f_\alpha \equiv 0$ for all points in J^kE. Then Ldx is not invariant with respect to ∂_x-symmetry, since

$$\begin{aligned}\mathcal{L}_{\mathrm{pr}V}(Ldx) &= [Q^\alpha \mathcal{E}_\alpha L + D_x(LV^x + Q^\alpha L_{u_x^\alpha})]dx = \\ &= [D_x(x \cdot 1)]dx = dx \neq 0,\end{aligned}$$

but the corresponding source form $\Delta \equiv 0$ satisfies every symmetry condition, since it is identically zero for all points on J^kE. □

Now equation (2.117) describes the implicit version of Noether's first theorem. If a variational source form Δ is invariant under some symmetry vector fields, then $\mathcal{E}_\alpha(Q^\beta f_\beta) = 0$ and this gives an implicit version of Noether's theorem. Implicit because we do not get the conserved quantity C directly, but we know there must be a C such that $Q^\beta f_\beta = D_x C$ (locally exact sequence). We could invert the operator D_x and compute C, but it needs slightly more work or we could compute the Lagrangian L, which also needs slightly more work.

When we allow the weaker symmetry condition in (2.116), i.e.

$$(D_x A)dx = \mathcal{L}_{\mathrm{pr}V}(Ldx) = [Q^\alpha f_\alpha + D_x C]dx,$$

then we get

$$\begin{aligned}\mathcal{L}_{\mathrm{pr}V}\Delta &= [\mathcal{E}_\alpha(Q^\beta f_\beta)]du^\alpha \wedge dx = \\ &= [\mathcal{E}_\alpha D_x(A - C)]du^\alpha \wedge dx = 0\end{aligned}$$

in (2.117). Therefore, Ldx (or I) satisfies a weak symmetry if and only if Δ (or δI) satisfies this symmetry (in the usual sense).

We finish this section with a short summary:

invariance of Ldx or I	$\Leftrightarrow\ 0 = Q^\alpha f_\alpha + D_x C$ (explicitly)
weak invariance of Ldx or I	$\Leftrightarrow\ 0 = Q^\alpha f_\alpha + D_x(C - A)$ (explicitly)
invariance of Δ or δI	$\Leftrightarrow\ 0 = \mathcal{E}_\alpha(Q^\beta f_\beta)\ \Leftrightarrow\ Q^\alpha f_\alpha = D_x\Lambda$ (implicitly)

The symmetry equation in Noether's first theorem and the ECS are pretty similar and they are given by the two equations

$$\mathcal{L}_{\mathrm{pr}V}(Ldx) = [D_x(LV^x + Q^\alpha L_{u_x^\alpha}) + Q^\alpha \mathcal{E}_\alpha L]dx, \tag{2.120}$$

$$\mathcal{L}_{\mathrm{pr}V}\Delta = [\mathcal{E}_\alpha(Q^\beta f_\beta) + Q^\beta H_{\alpha\beta} + (D_x Q^\beta)H^x_{\alpha\beta} + (D_x^2 Q^\beta)H^{xx}_{\alpha\beta}]du^\alpha \wedge dx, \tag{2.121}$$

where D_x,$\mathcal{E}_\alpha$ are the operators used in (2.120) and $\mathcal{E}_\alpha$,$\mathcal{H}^\gamma_{\alpha\beta}$ are the operators used in (2.121). We could say that the Lie derivative $\mathcal{L}_{\mathrm{pr}V}$ can be decomposed into these two operators. Similar formulas also hold for higher order and more general differential forms in the variational sequence and the decomposition is also reflected by Cartan's formula

$$\mathcal{L}_{\mathrm{pr}V}\omega = (d\iota_{\mathrm{pr}V} + \iota_{\mathrm{pr}V}d)\omega$$

for general differential forms ω. Further information can be found in (Kru15, p.174), especially Theorem 2 therein.

2.11. The Order of Jet Coordinates, Part II

We already introduced the notation of order in Section 2.1 and Section 2.5. Here we introduce a second notation which is slightly shorter and which will help us to solve Takens' problem. A function φ is of order k, in short $\varphi = O(k)$, if $\varphi \in C^\infty(J^kE)$. Total derivatives increase the order by one when applied to functions, i.e. $D_x : C^\infty(J^kE) \to C^\infty(J^{k+1}E)$. More precisely, the order is increased affine linear in the highest order jet coordinates, i.e.

$$\begin{aligned} D_x\varphi &= \varphi_x + u^\alpha_x \varphi_{u^\alpha} + u^\alpha_{xx}\varphi_{u^\alpha_x} + \ldots + u^\alpha_{(k)}\varphi_{u^\alpha_{(k-1)}} + u^\alpha_{(k+1)}\varphi_{u^\alpha_{(k)}} = \\ &= O(k) + u^\alpha_{(k+1)}\varphi_{u^\alpha_{(k)}} \quad \in C^\infty(J^{k+1}E) \end{aligned} \tag{2.122}$$

for every function $\varphi \in C^\infty(J^kE)$. This is a very simple but important observation and will be crucial later. We also want to introduce the notation $\varphi = O_1(k)$ if φ is affine linear in the k-th order jet coordinates (see (2.122), where the expression is affine linear in $u^\alpha_{(k+1)}$). Sometimes we will also write $O_1(k) = O_{lin}(k)$. In general, $\varphi = O_P(k)$ if φ is a polynomial of degree P in k-th order jet coordinates. However, note that later $O_{\mathscr{A}}(1)$ does not indicate a polynomial of degree $\mathscr{A}$, it rather labels the different kinds of symmetries. Sometimes we will write a few indices on the expression $O(k)$, for example $O^\beta_{ij}(k)$, and always when we use the indices $\mathscr{A}$,α,β,i,j, then we do not describe polynomial degree (this will also be clear from the context). Also note that the definition of objects $O_P(k)$ is invariant under local coordinate transformations, in other words, it is well-defined. See Proposition 2.4.3, where the transformation of the highest order jet coordinates can be found and highest order does never change in the sense of $O(k)$-notation. This is not true for lower order. The notation of $O_P(k)$ satisfies some nice properties, like

$$\begin{aligned} O_{P_1}(k)O_{P_2}(k) &= O_{P_1+P_2}(k), \quad \text{for all } k \geq 1, \\ O_{P_1}(k)O_{P_2}(l) &= O_{P_1}(k), \quad \text{for all } k > l \geq 0. \end{aligned}$$

We can further develop similar properties when needed.

2.12. The Expressions for Arbitrary PDEs

Since we mostly discussed ODEs so far and we also want to solve Takens' problem for PDEs, we briefly have to explain the more general notation. Most of the generalizations are straight forward, except the definition of conservation law and conserved quantity will be different in some sense and we have to say a few words about that.

The Expressions for arbitrary PDEs can be found in (AP94) and references therein. However, sometimes we use slightly different notation here. In the case where $n = 3$, we write $(x^i) = (x,y,z)$, and similar for $m = 3$, we will write $(u^\alpha) = (u,v,w)$. We also want to refer to (Kru15) for the PDE case.

Now let us discuss some of these expressions. The Euler-Lagrange expression in PDE case is

$$\mathcal{E}_\alpha L = L_{u^\alpha} - D_i L_{u_i^\alpha} + c_{ij} D_{ij} L_{u_{ij}^\alpha} \pm ...$$

and the Helmholtz expressions for second order f_α are

$$\begin{aligned} H_{\alpha\beta} &= f_{\alpha,u^\beta} - f_{\beta,u^\alpha} + D_i f_{\beta,u_i^\alpha} - c_{ij} D_{ij} f_{\beta,u_{ij}^\alpha}, \\ H_{\alpha\beta}^i &= f_{\alpha,u_i^\beta} + f_{\beta,u_i^\alpha} - 2c_{ij} D_j f_{\beta,u_{ij}^\alpha}, \\ H_{\alpha\beta}^{ij} &= c_{ij}(f_{\alpha,u_{ij}^\beta} - f_{\beta,u_{ij}^\alpha}), \end{aligned}$$

where c_{ij} is defined as

$$c_{ij} := \begin{cases} 1, & i = j, \\ \frac{1}{2}, & i \neq j. \end{cases}$$

The constants c_{ij} are needed to avoid *double counting* of partial derivatives, like $\partial_{u_{xy}^\alpha}, \partial_{u_{yx}^\alpha}$. We already see here that the factors c_{ij} are a bit annoying, but we need them when we want to use Einstein summation in the Euler-Lagrange and Helmholtz expressions. Note that Einstein summation in $c_{ij} D_{ij} \partial_{u_{ij}^\alpha}$ means that we have Einstein summation in $D_{ij} \partial_{u_{ij}^\alpha}$ and c_{ij} are just numerical factors, that is, there is no second Einstein summation with $c_{ij} \partial_{u_{ij}^\alpha}$ and therefore we could also hide the factors c_{ij} in the expression

$$\partial_\alpha^{ij} := c_{ij} \partial_{u_{ij}^\alpha}.$$

This notation is also used in (Poh95) and later in Section 3.7 we will also use it. The so-called double counting also occurs in the following situations:

- Let us consider the case where $n = 2$ and $m = 1$. Then, for example, in the summation $u_{ij} A^{ij} = u_{xx} A^{xx} + u_{xy} A^{xy} + u_{yx} A^{yx} + u_{yy} A^{yy}$ we have the double counting of u_{xy}, u_{yx}, which are (by definition) the same coordinates in the jet bundle (and we can define $A^{xy} = A^{yx}$).

- Again, let us consider $n = 2$. Differential forms like $A_{ij}dx^i \wedge dx^j = A_{xy}dx \wedge dy + A_{yx}dy \wedge dx = (A_{xy} - A_{yx})dx \wedge dy$ also have this double counting in a different way.

In short, the PDE case is getting more complicated and one reason is the **double counting**, discussed above. Alternatively, we could use **ordered sums**, instead of Einstein summation over all indices, but this can also get complicated (from notation point of view). Most of the time we will use Einstein summation without any ordering. The factors c_{ij} are also important in the total derivative

$$D_i = \partial_{x^i} + u_i^\alpha \partial_{u^\alpha} + u_{ik}^\alpha \partial_{u_k^\alpha} + c_{kl} u_{ikl}^\alpha \partial_{u_{kl}^\alpha} + ...$$

and it seems that there is no way out of using these factors. Similar, as we defined c_{ij}, we can also define c_{ijk} and so on to avoid double counting of higher order jet coordinates. Most of the time it will be sufficient to use the coefficients c_{ij} and therefore we do not further stress this notation. Also see (Poh95, p.344) (in general $c_I := \frac{I!}{|I|!}$). In the PDE case, we consider jet coordinates

$$(x^i, u^\alpha, u_{j_1}^\alpha, u_{j_1 j_2}^\alpha, ..., u_I^\alpha)$$

on $J^k E$, where $I = j_1 j_2 ... j_k$ is a multi index of length $|I| = k$, and where $1 \leq j_r \leq n$. Let $0 \leq l \leq k$, then the indices $j_1 j_2 ... j_l$ are **unordered** and

$$u_{j_1 j_2 ... j_l}^\alpha = u_{\pi(j_1)\pi(j_2)...\pi(j_l)}^\alpha \tag{2.123}$$

defines the same jet coordinate for every permutation π. For example, as we already mentioned, $u_{xy}^\alpha = u_{yx}^\alpha$, as well as $u_{xxy}^\alpha = u_{xyx}^\alpha = u_{yxx}^\alpha$. Even if we allow such unordered indices, local coordinates on $J^k E$ are also given as $(x^i, u^\alpha, u_{j_1}^\alpha, u_{j_1 j_2}^\alpha, ..., u_I^\alpha)$, where

$$u_{j_1 j_2 ... j_l}^\alpha, \quad 1 \leq j_1 \leq j_2 \leq ... \leq j_l \leq n, \quad l = 0,1,2,...,k. \tag{2.124}$$

Also see (And89, p.3). But it is reasonable to work with the unordered indices, where we get the equivalent expressions in (2.123).

Now let us continue with the definition of Lagrange and source forms. The Lagrange form is

$$\lambda = L dx^1 \wedge dx^2 \wedge ... \wedge dx^n = L dx$$

and the source form is a $(n+1)$-form

$$\Delta = f_\alpha du^\alpha \wedge dx^1 \wedge dx^2 ... \wedge dx^n = f_\alpha du^\alpha \wedge dx.$$

Vector fields on E are written as

$$V = V^i \partial_{x^i} + V^\alpha \partial_{u^\alpha}$$

and when using concrete labels, like $i = 1,\alpha = 2$, we will write

$$V = V^{x,1}\partial_{x^1} + V^{u,2}\partial_{u^2},$$

since V^1 does not indicate if it belongs to ∂_{x^1} or ∂_{u^1}. The prolongation of vector fields is

$$\mathrm{pr}^k V = V^i D_i + Q^\alpha \partial_{u^\alpha} + (D_i Q^\alpha)\partial_{u_i^\alpha} + c_{ij}(D_{ij}Q^\alpha)\partial_{u_{ij}^\alpha} + ...$$

and higher order coefficients can be found in (AP94, p.197). To define conservation laws, we need to consider horizontal $(n-1)$-forms

$$C = \sum_{j=1}^{n} (-1)^{j+1} C^j dx^1 \wedge ... \wedge dx^{j-1} \wedge dx^{j+1} \wedge ... \wedge dx^n$$

(the conserved quantities in some sense). We took this definition from (AP94, p.199), except we wrote C,C^j instead of ω,V^j. Then a conservation law is a trivial Lagrangian, i.e. a horizontal n-form, which is a divergence expression (for f_α) of the form

$$\begin{aligned} F dx &= \sum_{j=1}^{n} (-1)^{j+1} (D_j C^j) dx^j \wedge dx^1 \wedge ... \wedge dx^{j-1} \wedge dx^{j+1} \wedge ... \wedge dx^n = \\ &= \sum_{j=1}^{n} (D_j C^j) dx^1 \wedge ... \wedge dx^{j-1} \wedge dx^j \wedge dx^{j+1} \wedge ... \wedge dx^n = \\ &= (D_j C^j) dx. \end{aligned}$$

Note that such trivial Lagrangians can also be written as $F dx = d_H C$, where d_H is the horizontal exterior derivative and this operator is similar to the usual exterior derivative d, except it treats the coordinates $u^\alpha,u_i^\alpha,...$ as functions depending on $x^1,...,x^n$. Usually, d_H is defined on the infinite jet bundle $J^\infty E$ (see (AP94)). The operator d_H is also similar to the operator D_j, which also treats the coordinates $u^\alpha,u_i^\alpha,...$ as functions depending on $x^1,...,x^n$ and then D_j is just imitating the standard partial derivative ∂_{x^i}. From that point of view, the coordinate invariance of divergence expressions should be clear. However, we will also show it below by direct computation in the case where $n = 2$. The form $F dx$ can also be written as dC plus a contact form. Now let us do the calculation for $n = 2$, where the horizontal $(n-1)$-form C transforms like

$$\begin{aligned} C = C^x dy - C^y dx &= C^x \left(\frac{\partial y}{\partial \tilde{x}} d\tilde{x} + \frac{\partial y}{\partial \tilde{y}} d\tilde{y} \right) - C^y \left(\frac{\partial x}{\partial \tilde{x}} d\tilde{x} + \frac{\partial x}{\partial \tilde{y}} d\tilde{y} \right) = \\ &= \left(C^x \frac{\partial y}{\partial \tilde{x}} - C^y \frac{\partial x}{\partial \tilde{x}} \right) d\tilde{x} + \left(C^x \frac{\partial y}{\partial \tilde{y}} - C^y \frac{\partial x}{\partial \tilde{y}} \right) d\tilde{y} = \\ &= -\tilde{C}^y d\tilde{x} + \tilde{C}^x d\tilde{y}, \end{aligned}$$

where

$$\tilde{C}^y := -\left(C^x \frac{\partial y}{\partial \tilde{x}} - C^y \frac{\partial x}{\partial \tilde{x}}\right), \quad \tilde{C}^x := C^x \frac{\partial y}{\partial \tilde{y}} - C^y \frac{\partial x}{\partial \tilde{y}}.$$

Furthermore,

$$\begin{aligned} d\tilde{x} \wedge d\tilde{y} &= \left(\frac{\partial \tilde{x}}{\partial x} dx + \frac{\partial \tilde{x}}{\partial y} dy\right) \wedge \left(\frac{\partial \tilde{y}}{\partial x} dx + \frac{\partial \tilde{y}}{\partial y} dy\right) = \\ &= \left(\frac{\partial \tilde{x}}{\partial x}\frac{\partial \tilde{y}}{\partial y} - \frac{\partial \tilde{x}}{\partial y}\frac{\partial \tilde{y}}{\partial x}\right) dx \wedge dy. \end{aligned}$$

Then, using the transformation of D_x,D_y, see Proposition 2.4.3, we get

$$\begin{aligned} &(D_{\tilde{x}} C^{\tilde{x}} + D_{\tilde{y}} C^{\tilde{y}}) d\tilde{x} \wedge d\tilde{y} = \\ =&\left[\left(\frac{\partial x}{\partial \tilde{x}} D_x + \frac{\partial y}{\partial \tilde{x}} D_y\right) C^{\tilde{x}} + \left(\frac{\partial x}{\partial \tilde{y}} D_x + \frac{\partial y}{\partial \tilde{y}} D_y\right) C^{\tilde{y}}\right] d\tilde{x} \wedge d\tilde{y} = \\ =&\left[\left(\frac{\partial x}{\partial \tilde{x}} D_x + \frac{\partial y}{\partial \tilde{x}} D_y\right)\left(C^x \frac{\partial y}{\partial \tilde{y}} - C^y \frac{\partial x}{\partial \tilde{y}}\right) - \right. \\ &\left. - \left(\frac{\partial x}{\partial \tilde{y}} D_x + \frac{\partial y}{\partial \tilde{y}} D_y\right)\left(C^x \frac{\partial y}{\partial \tilde{x}} - C^y \frac{\partial x}{\partial \tilde{x}}\right)\right] d\tilde{x} \wedge d\tilde{y} = \end{aligned} \tag{2.125}$$

Let us continue here with (2.125)

$$\begin{aligned} =&\left[\frac{\partial x}{\partial \tilde{x}}\frac{\partial y}{\partial \tilde{y}} D_x C^x + \frac{\partial y}{\partial \tilde{x}}\frac{\partial y}{\partial \tilde{y}} D_y C^x + C^x \frac{\partial^2 y}{\partial \tilde{y} \partial \tilde{x}} - \right. \\ &- \left(\frac{\partial x}{\partial \tilde{x}}\frac{\partial x}{\partial \tilde{y}} D_x C^y + \frac{\partial y}{\partial \tilde{x}}\frac{\partial x}{\partial \tilde{y}} D_y C^y + \frac{\partial^2 x}{\partial \tilde{y} \partial \tilde{x}} C^y\right) - \\ &- \left(\frac{\partial x}{\partial \tilde{y}}\frac{\partial y}{\partial \tilde{x}} D_x C^x + \frac{\partial y}{\partial \tilde{y}}\frac{\partial y}{\partial \tilde{x}} D_y C^x + C^x \frac{\partial^2 y}{\partial \tilde{x} \partial \tilde{y}}\right) + \\ &\left. + \frac{\partial x}{\partial \tilde{y}}\frac{\partial x}{\partial \tilde{x}} D_x C^y + \frac{\partial y}{\partial \tilde{y}}\frac{\partial x}{\partial \tilde{x}} D_y C^y + \frac{\partial^2 x}{\partial \tilde{x} \partial \tilde{y}} C^y\right] d\tilde{x} \wedge d\tilde{y} = \\ =&\left(\frac{\partial x}{\partial \tilde{x}}\frac{\partial y}{\partial \tilde{y}} - \frac{\partial x}{\partial \tilde{y}}\frac{\partial y}{\partial \tilde{x}}\right)(D_x C^x + D_y C^y)\left(\frac{\partial \tilde{x}}{\partial x}\frac{\partial \tilde{y}}{\partial y} - \frac{\partial \tilde{x}}{\partial y}\frac{\partial \tilde{y}}{\partial x}\right) dx \wedge dy = \\ =&(D_x C^x + D_y C^y) dx \wedge dy, \end{aligned}$$

where we used $\det(A^{-1}) = (\det A)^{-1}$ for every invertible matrix A in the last line. We also used

$$\begin{aligned} 1 &= \partial_{\tilde{x}} \tilde{x} = ..., \\ 0 &= \partial_{\tilde{y}} \tilde{x} = \left(\frac{\partial x}{\partial \tilde{y}} \partial_x + \frac{\partial y}{\partial \tilde{y}} \partial_y\right) \tilde{x} = \frac{\partial x}{\partial \tilde{y}}\frac{\partial \tilde{x}}{\partial x} + \frac{\partial y}{\partial \tilde{y}}\frac{\partial \tilde{x}}{\partial y}, \\ 0 &= \partial_{\tilde{x}} \tilde{y} = ..., \\ 1 &= \partial_{\tilde{y}} \tilde{y} = ... \end{aligned}$$

to show that

$$AA^{-1} = \begin{pmatrix} \frac{\partial \tilde{x}}{\partial x} & \frac{\partial \tilde{x}}{\partial y} \\ \frac{\partial \tilde{y}}{\partial x} & \frac{\partial \tilde{y}}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial \tilde{x}} & \frac{\partial x}{\partial \tilde{y}} \\ \frac{\partial y}{\partial \tilde{x}} & \frac{\partial y}{\partial \tilde{y}} \end{pmatrix} = \begin{pmatrix} ... & \frac{\partial \tilde{x}}{\partial x}\frac{\partial x}{\partial \tilde{y}} + \frac{\partial \tilde{x}}{\partial y}\frac{\partial y}{\partial \tilde{y}} \\ ... & ... \end{pmatrix} = \begin{pmatrix} ... & 0 \\ ... & ... \end{pmatrix}.$$

A similar calculation holds for arbitrary n. Note that considering the n-tupel

$$(C^1,C^2,...,C^n)$$

as a sort of conserved quantity seems to be confusing at first sight. However, it is well-known that the 4-tupel

$$(C^1,C^2,C^3,C^4) = (\rho,j^1,j^2,j^3),$$

with corresponding divergence

$$D_iC^i = D_t\rho + D_{x^1}j^1 + D_{x^2}j^2 + D_{x^3}j^3,$$

describes the conservation of electromagnetic charge, where ρ is the charge density and (j^1,j^2,j^3) the current density in Maxwell's equations. In the following, let $dx := dx^1 \wedge dx^2 \wedge dx^3$, and we integrate over a compact set $\tilde{U}^0$, which corresponds to the coordinates (x^1,x^2,x^3). Then, by Gauss's theorem, we get

$$0 = \int_{\tilde{U}^0} \mathrm{pr}^k\sigma^*(D_iC^i dx) = \partial_t \int_{\tilde{U}^0} \mathrm{pr}^k\sigma^*(\rho dx) + \oint_{\partial\tilde{U}^0} \mathrm{pr}^k\sigma^*[\sum_{i=1}^{3}(D_ij^i)dS], \quad (2.126)$$

for all solutions $\sigma \in \Gamma(E)$ of the corresponding differential equation (Maxwell's equations). Therefore, the total charge

$$\int_{\tilde{U}^0} \mathrm{pr}^k\sigma^*(\rho dx)$$

in some compact set $\tilde{U}^0$ can only change in time when we have a flow (j^1,j^2,j^3) of current into or out of this set. Thus, it makes still sense to speak about conservation laws, even when we do not mean constant functions here, as we have in the case where $n = 1$. Now we consider again, as usually, $dx := dx^1 \wedge dx^2 \wedge ... \wedge dx^n$. A conservation law for f_α is a divergence expression Fdx and a generalized vertical vector field $Q^\alpha\partial_{u^\alpha}$ on E, such that

$$\iota_{\mathrm{pr}Q}\Delta = (D_iC^i)dx = Fdx$$

for all points in J^kE. This is equivalent to saying that $\iota_{\mathrm{pr}Q}\Delta - dC$ is a contact form or $\iota_{\mathrm{pr}Q}\Delta = d_HC$, where d_H is the horizontal exterior derivative in the variational bicomplex (And89).

3. The Main Result and What We Need to Prove it

In the first part of this chapter, we only consider second order source forms, and in Section 3.8, we start with the discussion of fourth order source forms. The ECS (2.108) for second order PDE source forms is

$$Q^\beta H_{\alpha\beta} + (D_i Q^\beta) H^i_{\alpha\beta} + (D_{ij} Q^\beta) H^{ij}_{\alpha\beta} = 0. \tag{3.1}$$

This equation can also be found in (AP94). Note that $Q^\beta = V^\beta - u^\beta_i V^i$, where $V = V^i \partial_{x^i} + V^\alpha \partial_{u^\alpha}$ is a the symmetry vector field $V \in \mathcal{V}$. Roughly speaking, we want to solve this equation for the unknown Helmholtz expressions $H_{\alpha\beta}$,$H^i_{\alpha\beta}$,$H^{ij}_{\alpha\beta}$. More precisely, we want to derive that the Helmholtz conditions are satisfied, i.e. that $H_{\alpha\beta} = 0$,$H^i_{\alpha\beta} = 0$,$H^{ij}_{\alpha\beta} = 0$. The ECS (3.1) is a linear equation for the Helmholtz expressions. Actually, it is a partial differential equation, but when we consider $H_{\alpha\beta}$,$H^i_{\alpha\beta}$,$H^{ij}_{\alpha\beta}$ as arbitrary expressions, then it can be considered as a linear system in terms of linear algebra. Solving the pure linear algebra system does in general not solve Takens' problem, since then it has a non-trivial kernel. Therefore, we also have to take in account that $H_{\alpha\beta}$,$H^i_{\alpha\beta}$,$H^{ij}_{\alpha\beta}$ are not arbitrary expression, but Helmholtz expressions and they must satisfy certain conditions, also called integrability conditions. There are two ways how to formulate that $H_{\alpha\beta}$,$H^i_{\alpha\beta}$,$H^{ij}_{\alpha\beta}$ are Helmholtz expressions:

- We simply write

$$\begin{aligned} H_{\alpha\beta} &= f_{\alpha,u^\beta} - f_{\beta,u^\alpha} + D_i f_{\beta,u^\alpha_i} - c_{ij} D_{ij} f_{\beta,u^\alpha_{ij}}, \\ H^i_{\alpha\beta} &= f_{\alpha,u^\beta_i} + f_{\beta,u^\alpha_i} - 2c_{ij} D_j f_{\beta,u^\alpha_{ij}}, \\ H^{ij}_{\alpha\beta} &= c_{ij}(f_{\alpha,u^\beta_{ij}} - f_{\beta,u^\alpha_{ij}}), \end{aligned} \tag{3.2}$$

  and therefore the ECS is a partial differential equation for f_α.

- We extend the locally exact sequence

$$0 \to \mathbb{R} \to C^\infty(J^k E) \to ... \xrightarrow{D_i} \{L\} \xrightarrow{\mathcal{E}_\alpha} \{f_\alpha\} \xrightarrow{\mathcal{H}^\gamma_{\alpha\beta}} \{H_{\alpha\beta}, H^i_{\alpha\beta}, H^{ij}_{\alpha\beta}\} \xrightarrow{\mathcal{F}} ...,$$

where we introduced an operator $\mathcal{F}$ and if (formally)

$$\mathcal{F}\{H_{\alpha\beta},H^i_{\alpha\beta},H^{ij}_{\alpha\beta}\} = 0, \tag{3.3}$$

then $H_{\alpha\beta},H^i_{\alpha\beta},H^{ij}_{\alpha\beta}$ can locally be written as (3.2).

The conditions (3.3) are then called integrability conditions for the expressions $H_{\alpha\beta},H^i_{\alpha\beta},H^{ij}_{\alpha\beta}$. These conditions are relatively complicated when written down in local coordinates. Let us only mention the following condition

$$H^{xx}_{\alpha\beta,u^\gamma_{xx}}|_{[\alpha,\beta,\gamma]} = 0, \tag{3.4}$$

where $[\alpha,\beta,\gamma]$ means skew-symmetrization in α,β,γ. A direct computation shows

$$\begin{aligned}6H^{xx}_{\alpha\beta,u^\gamma_{xx}}|_{[\alpha,\beta,\gamma]} =&\partial_{u^\gamma_{xx}}(f_{\alpha,u^\beta_{xx}} - f_{\beta,u^\alpha_{xx}}) - \partial_{u^\gamma_{xx}}(f_{\beta,u^\alpha_{xx}} - f_{\alpha,u^\beta_{xx}})+\\ &+ \partial_{u^\alpha_{xx}}(f_{\beta,u^\gamma_{xx}} - f_{\gamma,u^\beta_{xx}}) - \partial_{u^\alpha_{xx}}(f_{\gamma,u^\beta_{xx}} - f_{\beta,u^\gamma_{xx}})-\\ &- \partial_{u^\beta_{xx}}(f_{\alpha,u^\gamma_{xx}} - f_{\gamma,u^\alpha_{xx}}) + \partial_{u^\beta_{xx}}(f_{\gamma,u^\alpha_{xx}} - f_{\alpha,u^\gamma_{xx}}) = 0.\end{aligned}$$

Also see (KM10) and (Mal09) for further details. Most of the time it seems to be easier to use the expressions in (3.2) directly. In either case, solving Takens' problem means to solve a system of partial differential equations

$$\begin{cases}\text{ECS:} \quad Q^\beta H_{\alpha\beta} + (D_iQ^\beta)H^i_{\alpha\beta} + (D_{ij}Q^\beta)H^{ij}_{\alpha\beta} = 0,\\ \text{integrability conditions for } H_{\alpha\beta},H^i_{\alpha\beta},H^{ij}_{\alpha\beta},\end{cases}$$

and (3.2), (3.3) are locally equivalent. There is one additional information we have to understand, when we want to solve Takens' problem. Surprisingly, the Helmholtz expressions and conditions are not independent and we have the condition

$$\begin{aligned}0 =&H_{\alpha\beta} + H_{\beta\alpha} - D_iH^i_{\alpha\beta} + D_{ij}H^{ij}_{\alpha\beta},\\ 0 =&H^i_{\alpha\beta} - H^i_{\beta\alpha} - 2D_jH^{ij}_{\alpha\beta},\\ 0 =&H^{ij}_{\alpha\beta} + H^{ij}_{\beta\alpha}.\end{aligned} \tag{3.5}$$

We discovered such dependencies already earlier in (2.78). Also see Proposition 3 in (AP12) and Proposition 3.1 in (AP96), where such dependencies are formulated. These identities can be proven by a direct computation, when using the expressions in (3.2). Note that the (formal) operator $\mathcal{F}$ always differentiates the expressions $H_{\alpha\beta},H^i_{\alpha\beta},H^{ij}_{\alpha\beta}$. For example, the operator $\mathcal{F}$ leads to the condition (3.4). Since there are no partial derivatives applied on the first two summands $H_{\alpha\beta} + H_{\beta\alpha}$ in (3.5), the first equation cannot be obtained from the operator $\mathcal{F}$ directly. But of course indirectly, since when (3.3) is satisfied then we get (3.2) and then (3.5). A similar observation holds for the expressions $H^i_{\alpha\beta} - H^i_{\beta\alpha}$ and $H^{ij}_{\alpha\beta} + H^{ij}_{\beta\alpha}$ in (3.5), where also no partial derivatives are applied. We call (3.5) the **Helmholtz dependencies** and we want to distinguish them from the integrability conditions. The

Helmholtz dependencies are really key to solve Takens' problem, also see Section 3.8, where we discuss them in more detail.

Therefore, solving Takens' problem means to solve the system

$$\begin{cases} 0 = Q^\beta H_{\alpha\beta} + (D_i Q^\beta) H^i_{\alpha\beta} + (D_{ij} Q^\beta) H^{ij}_{\alpha\beta}, & \text{(ECS)}, \\ \text{integrability conditions for } H_{\alpha\beta}, H^i_{\alpha\beta}, H^{ij}_{\alpha\beta}, & \\ 0 = H_{\alpha\beta} + H_{\beta\alpha} - D_i H^i_{\alpha\beta} + D_{ij} H^{ij}_{\alpha\beta}, & \text{(Helmholtz dependencies)}, \\ 0 = H^i_{\alpha\beta} - H^i_{\beta\alpha} - 2 D_j H^{ij}_{\alpha\beta}, & \text{(Helmholtz dependencies)}, \\ 0 = H^{ij}_{\alpha\beta} + H^{ij}_{\beta\alpha}, & \text{(Helmholtz dependencies)}, \end{cases} \tag{3.6}$$

which still seems to be a pretty complicated system of PDEs (even if we have added additional equations). This system is also complicated, because Q^β is a function and if we try to solve this system of PDEs explicitly, then we have to solve it for every admissible function Q^β (or characteristic). In general, there is not a high chance of solving such a system explicitly in this form. However, when we consider a set of symmetry vector fields with corresponding set of characteristics $\{Q^\beta\}$ then we get a set of ECS. In some situations this allows us to eliminate some of the unknowns $H_{\alpha\beta}, H^i_{\alpha\beta}, H^{ij}_{\alpha\beta}$ and we get a simplified ECS, which could be solvable. Also discussing the order of the equations in (3.6) step by step will simplify the problem. In the next sections we discuss how to solve (3.6) in several simple cases. After this discussion we prove Theorem (1.0.2) and (1.0.3). However, before we continue with this discussion, we need to formulate a technical theorem to be able to prove Theorem (1.0.2) and (1.0.3).

3.1. Local Properties

Takens' problem is to prove Theorem 1.0.2, 1.0.3 or similar ones. This mainly means that we have to solve the ECS somehow. Let us focus on Theorem 1.0.2 and we will rewrite the statement slightly differently (actually we can prove it with slightly weaker assumptions). For the calculations later, it is reasonable to introduce the label $\mathscr{A}$ and to work only with a finite set of symmetry vector fields. Theorem 1.0.2 follows from the following theorem:

Theorem 3.1.1. *Let $n,m \in \mathbb{N}$ be arbitrary and let $U \subset E$ be open. Furthermore, let $\Delta = f_\alpha du^\alpha \wedge dx^1 \wedge ... \wedge dx^n$ be a second order source form defined on $(\pi^{2,0})^{-1} U \subset J^2 E$. Assume:*
i) There are projectable vector fields $V_\mathscr{A}$, $\mathscr{A} = 1,2,...,n+m$ on U which are symmetries of Δ, and they satisfy span$\{V_{\mathscr{A},p}$, $\mathscr{A} = 1,2,...,n+m\} = T_p E$ for all $p \in U$.
ii) Each $V_\mathscr{A}$, $\mathscr{A} = 1,2,...,n+m$ generates a conservation law of the from $Q^\alpha_\mathscr{A} f_\alpha =$

$D_i C^i_{\mathscr{A}}$ on $(\pi^{2,0})^{-1}U \subset J^2E$, where $Q^\alpha_{\mathscr{A}} = V^\alpha_{\mathscr{A}} - u^\alpha_i V^i_{\mathscr{A}}$ are the characteristics. Then f_α satisfies the Helmholtz conditions in $(\pi^{2,0})^{-1}(U) \subset J^2E$.

Note that the statement: f_α satisfies the Helmholtz conditions in $(\pi^{2,0})^{-1}(U) \subset J^2E$, and Δ is locally variational in $(\pi^{2,0})^{-1}(U) \subset J^2E$ are equivalent (see Proposition 2.4.2 and Lemma 2.7.3). Also note that if f_α satisfies the Helmholtz conditions in $(\pi^{2,0})^{-1}(U)$ then there does not necessarily exist a Lagrange form on the whole space $(\pi^{2,0})^{-1}(U)$, but on certain subsets of $(\pi^{2,0})^{-1}(U)$. We can formulate a similar theorem to be able to prove Theorem 1.0.3. This technical formulation simplifies the proof of Theorem 1.0.2 and 1.0.3 and it is easier to explain and understand how many equations and unknowns we will have later. The whole fiber bundle E is the union of such sets $U \subset E$ and when we show that the Helmholtz conditions are satisfied in all of the subsets $(\pi^{2,0})^{-1}U \subset J^2E$ then f_α satisfies the Helmholtz conditions on J^kE, which means that Δ is locally variational on J^kE.

Proposition 3.1.2. *Let $\{V\}$ be a set of projectable vector fields on E such that span$\{V_p\} = T_pE$ for all $p \in E$. Then for every $p_0 \in E$ there exists a small neighbourhood $U_{p_0} \subset E$ of p_0 such that we can choose $n+m$ vector fields $\{V_1,V_2,...,V_{n+m}\} \subset \{V\}$ such that span$\{V_{1,p},V_{2,p},...,V_{n+m,p}\} = T_pE$ for all $p \in U_{p_0}$.*

Proof: Let $p_0 \in E$. Then, by assumption, there exist $\{V_1,V_2,...,V_{n+m}\} \subset \{V\}$ such that

$$\mathrm{span}_{\mathbb{R}}\{V_{1,p_0},V_{2,p_0},...,V_{n+m,p_0}\} = T_{p_0}E. \tag{3.7}$$

Let us write $V_l \in \{V_1,V_2,...,V_{n+m}\}$, $1 \le l \le n+m$, in local coordinates as

$$V_l = V_l^{i,x}\partial_{x^i} + V_l^{\alpha,u}\partial_{u^\alpha}.$$

The condition (3.7) forces that the matrix

$$\begin{pmatrix} V_1^{x,1} & V_1^{x,2} & ... & V_1^{x,n} & V_1^{u,1} & V_1^{u,2} & ... & V_1^{u,m} \\ V_2^{x,1} & V_2^{x,2} & ... & V_2^{x,n} & V_2^{u,1} & V_2^{u,2} & ... & V_2^{u,m} \\ \vdots & & & & & & & \vdots \\ V_{n+m}^{x,1} & V_{n+m}^{x,2} & ... & V_{n+m}^{x,n} & V_{n+m}^{u,1} & V_{n+m}^{u,2} & ... & V_{n+m}^{u,m} \end{pmatrix}$$

must be invertible at $p_0 \in E$. If a matrix is invertible at a point $p_0 \in E$ and the coefficients in this matrix are smooth function on E, then the matrix must also be invertible in a small neighbourhood $U_{p_0} \subset E$ of p_0. This is due to the fact that the determinant of this matrix is non-zero at $p_0 \in E$ and then the determinant must also be non-zero in a small neighborhood of p_0, which proves the proposition. $\square$

3.2. Counting the Unknowns and Equations

Labeling the symmetry vector fields V by $\mathscr{A}$, and using the notation of order, the ECS can be written as

$$(V^{\beta}_{\mathscr{A}} - u^{\beta}_{i} V^{i}_{\mathscr{A}}) H_{\alpha\beta} + (O^{\beta}_{\mathscr{A}}(1) - u^{\beta}_{ik} V^{i}_{\mathscr{A}}) H^{k}_{\alpha\beta} + (O^{\beta}_{\mathscr{A}}(2) - u^{\beta}_{ikl} V^{i}_{\mathscr{A}}) H^{kl}_{\alpha\beta} = 0,$$

where $\mathscr{A} = 1,...,(n+m)$, $i = 1,2,...,n$ and $\alpha = 1,2,...,m$. Therefore, we have

$$(n+m)m \tag{3.8}$$

equations. Note that the vector fields $\{V_{\mathscr{A}} : \mathscr{A} = 1,2,...,n+m\}$ span T_pE at each $p \in U \subset E$. Therefore, we get $n+m$ equations at each $p \in E$ for the different kinds of symmetries. Furthermore, since $\alpha = 1,2,...,m$ in the ECS, we get (3.8). The number of unknowns $H_{\alpha\beta}$,$H^{i}_{\alpha\beta}$,$H^{ij}_{\alpha\beta}$ is

$$\begin{aligned} \#H_{\alpha\beta} &= m^2, \\ \#H^{i}_{\alpha\beta} &= nm^2, \\ \#H^{ij}_{\alpha\beta} &= \frac{n(n+1)}{2}\frac{m(m-1)}{2}. \end{aligned}$$

This simply follows by counting the indices in the unknowns. Whether we say, as above, that $H^{ij}_{\alpha\beta}$ are $\frac{n(n+1)}{2}\frac{m(m-1)}{2}$ unknowns, or whether we start with n^2m^2 unknowns and then show that there are actually less unknowns, because of the Helmholtz dependencies (3.5), is a matter of taste. We presuppose from the beginning that we know the symmetry and skew-symmetry conditions for $H^{ij}_{\alpha\beta}$, since they can easily be observed in the Helmholz dependencies. For dependencies which involve $H_{\alpha\beta}$,$H^{i}_{\alpha\beta}$, the counting of indices more complicated and we do not further investigate $H_{\alpha\beta}$ and $H^{i}_{\alpha\beta}$. Together, the number of unknowns is

$$m^2 + nm^2 + \frac{n(n+1)}{2}\frac{m(m-1)}{2}.$$

Therefore, for large n,m, we have a highly under-determined ECS system and only for $m = 1$, we get that the system only allows the trivial solution in the case where we can span T_pE at each $p \in E$, see (1.2). In the case where we can only span T_qM for all $q \in M$, i.e. condition (1.3), the problem is more complicated and we have to investigate the non-trivial solutions of the ECS.

The idea how to prove Theorem 1.0.2 and 1.0.3 is very simple. The ECS is used to show that the Helmholtz conditions are satisfied. In general, since the ECS does not provide enough information, we also have to use the integrability conditions for the Helmholtz expressions and the Helmholtz dependencies. Discussing the order of these equations step by step will also help to solve the problem.

3.3. The Proof for $m = 1$, Arbitrary n and 2nd Order Source Forms

In the case where $m = 1$, the ECS (3.1) has a very special form, namely

$$
\begin{aligned}
0 &= Q^\beta H_{\alpha\beta} + (D_i Q^\beta) H^i_{\alpha\beta} = \\
&= QH + (D_i Q) H^i,
\end{aligned}
\tag{3.9}
$$

where $H^{ij}_{\alpha\beta} = 0$, $H^i_{\alpha\beta} = H^i = 2(f_{u_i} - c_{ik} D_k f_{u_{ik}})$ and we write $u^\alpha = u, Q^\beta = Q$. Furthermore, we have the Helmholtz dependency $2H = D_i H^i$ and we write $H_{\alpha\beta} = H$. This very special form allows us to formulate a simpler proof, compared to the more general situation when $m > 1$. Therefore, it is probably no surprise that also Takens considered this case in one of his proofs in the original paper (Tak77).

3.3.1. Full Rank, a Proof for $m = 1$, Arbitrary n and 2nd Order

The first and easiest method which we want to present how to solve Takens' problem is the following: If we can span T_pE at each point $p \in E$ with the symmetry vector fields, i.e. when (1.2) is satisfied, then this means that we must have at least $n+1$ symmetry vector fields (with certain properties). Therefore, we get at least $n+1$ equations and the ECS can be written in matrix vector form as

$$
\begin{pmatrix}
Q_1 & D_1Q_1 & D_2Q_1 & \dots & D_nQ_1 \\
Q_2 & D_1Q_2 & D_2Q_2 & \dots & D_nQ_2 \\
\vdots & & & & \vdots \\
Q_n & D_1Q_n & D_2Q_n & \dots & D_nQ_n \\
Q_{n+1} & D_1Q_{n+1} & D_2Q_{n+1} & \dots & D_nQ_{n+1}
\end{pmatrix}
\begin{pmatrix}
H \\ H^1 \\ \vdots \\ H^n
\end{pmatrix}
= 0.
$$

If we can show that the determinant

$$
\begin{vmatrix}
Q_1 & D_1Q_1 & D_2Q_1 & \dots & D_nQ_1 \\
Q_2 & D_1Q_2 & D_2Q_2 & \dots & D_nQ_2 \\
\vdots & & & & \vdots \\
Q_n & D_1Q_n & D_2Q_n & \dots & D_nQ_n \\
Q_{n+1} & D_1Q_{n+1} & D_2Q_{n+1} & \dots & D_nQ_{n+1}
\end{vmatrix}
\neq 0, \quad \text{(almost everywhere)}
\tag{3.10}
$$

then we know that the Helmholtz conditions must be satisfied. It is rather simple to show that the determinant is non-zero almost everywhere in J^kE (or subsets $(\pi^{k,0})^{-1}U$ of J^kE), in the case where we can span T_pE at each $p \in E$, i.e. when (1.2) is satisfied. By a continuity argument we get that the Helmholtz conditions must be satisfied everywhere in J^kE. Also see Case I in (AP94, p.207), where we can find a similar discussion and some examples.

Since this proof is rather simple, let us investigate if we can weaken the condition of spanning T_pE at each $p \in E$. That is, instead of the condition (1.2), we only want to assume the condition (1.3). This assumption leads us directly to the next subsection, where we solve the ECS explicitly for $n,m = 1$ and a single symmetry vector field. A generalization can be found in (AP94).

3.3.2. Solving the ECS Explicitly in the Simplest Case

A second method how to solve Takens' problem is to solve the ECS (3.9) explicitly for a single symmetry vector field (or characteristic Q). Because of the Helmholtz dependency $2H = D_iH^i$, we can write the ECS as

$$0 = \frac{1}{2}QD_iH^i + (D_iQ)H^i.$$

Let us first consider the simplest case, where $n = 1$ (and $m = 1$). Then the ECS is

$$0 = \frac{1}{2}QD_xH^x + (D_xQ)H^x. \tag{3.11}$$

This is probably one of the rare cases, where we are able to solve the ECS explicitly (for a single symmetry). Note that we want to solve this equation for H^x and for all admissible functions Q. Let us write $V = V^x\partial_x + V^u\partial_u$ for the symmetry vector field in $\mathcal{V}$ and $Q = V^u - u_xV^x$. The general solution is

$$H^x = \frac{c}{Q^2} = \frac{c}{[V^u(x,u) - u_xV^x(x)]^2}, \quad \text{where } c \in \mathbb{R}. \tag{3.12}$$

If $V^x(x) \neq 0$ (for at least one x), then $c = 0$, since H^x must be a non-singular function in the coordinates (x,u,u_x). The condition $V^x(x) \neq 0$ (for all x) can also be written as $\pi_*V \neq 0$. We could also investigate singular function H^x and compute the corresponding f, but we will not discuss this case in more detail in this subsection. If just $V^u(x,u) \neq 0$ for at least one (x,u) and $V^x \equiv 0$, then it does not follow that $c = 0$. Here we can observe once again, where we need the span-T_pE-condition (1.2) in the assumptions of Theorem 1.0.2 and 1.0.3 to be able to prove these theorems. Only for $m = 1$, the condition (1.3) is sufficient to show that the Helmholtz conditions for second order source forms must be satisfied and we do not need the stronger assumption (1.2). However, when we only assume (1.3) then we need the additional non-singular assumption of the Helmholtz expressions (or of f) to be able to solve the problem. For the more general case, where $m = 1$ and arbitrary n, see (AP94).

Note that we can pull-back the equation in (3.11) by a prolonged section $\sigma \in \Gamma(E)$, in local coordinates by a section $u(x)$, and then we get an ODE for $u(x)$, which has the general solution (3.12). Otherwise (3.11) has to be consider as a kind of PDE on J^2E and it might not be clear what we mean with general solution.

We finish this short section with two remarks: First, in special cases, like for $n,m = 1$, it is relatively easy to find conditions under which we can solve Takens' problem. For example, when we assume $\pi_* V \neq 0$. One should start with such simple cases to get a good understanding for the problem. Second, for arbitrary n,m, arbitrary order differential equations, and quite general symmetry assumptions, it is very hard to solve the problem. One part of solving Takens' problem is to find conditions, especially for the symmetries, under which we can solve it. The condition $\mathrm{span}\{V_p : V \in \mathcal{V}\} = T_pE$ for all $p \in E$ seems to be quite obvious, but actually we do not exactly know how to generalize this condition for arbitrary k-th order differential equations. For example, for k-th order source forms maybe the condition $\mathrm{span}\{V_p : V \in \mathcal{V}\} = T_pJ^lE$ for all $p \in J^lE$, where $l = k-2$, solves the problem? Also finding counter examples is a very important part of solving Takens' problem. Therefore, there are many open problems and it is fair to say that Takens' problem is only solved in some special cases and there is no general understanding and solution to this problem.

In the next sections, we will solve Takens' problem in more general situations, where the actual results of this dissertation are formulated.

3.4. The Proof for $n = 1$, Arbitrary m and 2nd Order Source Forms

In this section, we solve Takens' problem in the special case $n = 1$, arbitrary m and second order source forms. In this case, the Helmholtz expressions are given as

$$H_{\alpha\beta} = f_{\alpha,u^\beta} - f_{\beta,u^\alpha} + D_x f_{\beta,u^\alpha_x} - D_x^2 f_{\beta,u^\alpha_{xx}}, \tag{3.13}$$

$$H^x_{\alpha\beta} = f_{\alpha,u^\beta_x} + f_{\beta,u^\alpha_x} - 2D_x f_{\beta,u^\alpha_{xx}}, \tag{3.14}$$

$$H^{xx}_{\alpha\beta} = f_{\alpha,u^\beta_{xx}} - f_{\beta,u^\alpha_{xx}}. \tag{3.15}$$

Symmetries and conservation laws lead to the ECS (see (2.108))

$$Q^\beta_{\mathscr{A}} H_{\alpha\beta} + (D_x Q^\beta_{\mathscr{A}}) H^x_{\alpha\beta} + (D_x^2 Q^\beta_{\mathscr{A}}) H^{xx}_{\alpha\beta} = 0, \tag{3.16}$$

where $\mathscr{A} = 1,2,...,(1+m)$, since we can span T_pE at each $p \in U \subset E$ and $\dim E = 1+m$. Also see Section 3.1. Before we start with the proof of Theorem 1.0.2 in this special case, let us formulate a lemma. The following lemma also holds for PDEs and we will formulate it in that way, since we also need it later.

Lemma 3.4.1 (Local Simplification Lemma)**.** *Let*

$$V_{\mathscr{A}} = V^i_{\mathscr{A}}(x)\partial_{x^i} + V^\alpha_{\mathscr{A}}(x,u)\partial_{u^\alpha}, \quad \mathscr{A} = 1,2,...,n+m$$

be projectable vector fields on $U \subset E$. If $span_{\mathbb{R}}\{V_{\mathscr{A},p}\ , \mathscr{A} = 1,2,...,n+m\} = T_pE$ at each $p \in U$, then there exists a $(n+m) \times (n+m)$-matrix $C = C(x,u)$ on U, such that the $(n+m) \times (n+m)$-matrix $(V^i_{\mathscr{A}},V^\alpha_{\mathscr{A}})_{\mathscr{A};i,\alpha}$ satisfy the following condition

$$C \cdot (V^i_{\mathscr{A}},V^\alpha_{\mathscr{A}})_{\mathscr{A};i,\alpha} = id.$$

That is, the matrix $(V^i_{\mathscr{A}},V^\alpha_{\mathscr{A}})_{\mathscr{A};i,\alpha}$ is invertible for all $p \in U$ and C is the inverse matrix. Moreover, there exists a row $c^{\mathscr{A}} = c^{\mathscr{A}}(x,u)$ of the matrix C such that either

$$c^{\mathscr{A}} V^i_{\mathscr{A}} = \delta^{ij}, \quad c^{\mathscr{A}} V^\alpha_{\mathscr{A}} = 0, \tag{3.17}$$

or

$$c^{\mathscr{A}} V^i_{\mathscr{A}} = 0, \quad c^{\mathscr{A}} V^\alpha_{\mathscr{A}} = \delta^{\alpha\beta}, \tag{3.18}$$

where $\delta^{ij},\delta^{\alpha\beta}$ are Kronecker deltas.

The proof follows directly by definition of $span_{\mathbb{R}}\{V_{p,\mathscr{A}}\ , \mathscr{A} = 1,2,...,n+m\} = T_pE$ at each $p \in U \subset E$. □

For example, let $n,m = 1$ and $V_{\mathscr{A}} = V^x_{\mathscr{A}}\partial_x + V^u_{\mathscr{A}}\partial_u \in \mathcal{V}$, where $\mathscr{A} = 1{,}2$ such that $\operatorname{span}\{V_{p,\mathscr{A}}\ ,\ \mathscr{A} = 1{,}2\} = T_pE$ at each $p \in E$ is satisfied. Then

$$(V^i_{\mathscr{A}},V^\alpha_{\mathscr{A}})_{\mathscr{A};i,\alpha} = \begin{pmatrix} V^x_1 & V^u_1 \\ V^x_2 & V^u_2 \end{pmatrix} \quad \Rightarrow \quad C = \frac{1}{V^x_1V^u_2 - V^u_1V^x_2}\begin{pmatrix} V^u_2 & -V^u_1 \\ -V^x_2 & V^x_1 \end{pmatrix}$$

and furthermore

$$\begin{aligned} c^{\mathscr{A}}_1 &= \frac{1}{V^x_1V^u_2 - V^u_1V^x_2}(V^u_2, -V^u_1), \quad c^{\mathscr{A}}_2 = \frac{1}{V^x_1V^u_2 - V^u_1V^x_2}(-V^x_2,V^x_1), \\ \Rightarrow \quad c^{\mathscr{A}}_1V^x_{\mathscr{A}} &= \frac{1}{V^x_1V^u_2 - V^u_1V^x_2}(V^u_2, -V^u_1)\begin{pmatrix} V^x_1 \\ V^x_2 \end{pmatrix} = 1, \\ \Rightarrow \quad c^{\mathscr{A}}_2V^x_{\mathscr{A}} &= \frac{1}{V^x_1V^u_2 - V^u_1V^x_2}(-V^x_2,V^x_1)\begin{pmatrix} V^x_1 \\ V^x_2 \end{pmatrix} = 0. \end{aligned}$$

A similar calculation holds for $c^{\mathscr{A}}_2$. Also see Appendix E, where we explicitly compute the matrix C and the rows $c^{\mathscr{A}}$ in another example.

The reason why we formulate this lemma is the following: We will need a simple, or more precisely, the most simplest local coordinate formulation of the condition $\operatorname{span}\{V_{\mathscr{A},p}\ ,\ \mathscr{A} = 1{,}2{,}...{,}n+m\} = T_pE$ at each $p \in U \subset E$ and this is described by the conditions (3.17) and (3.18).

Although this lemma seems to be an obvious statement, let us notice the following: If $V_{\mathscr{A}} \in \mathfrak{X}(E)$ is a projectable symmetry of Δ, then $c^{\mathscr{A}}V_{\mathscr{A}}$ is not a symmetry of Δ in general, because $c^{\mathscr{A}}\mathrm{pr}V_{\mathscr{A}} \neq \mathrm{pr}(c^{\mathscr{A}}V_{\mathscr{A}})$ and only prolonged vector fields are allowed to describe symmetries. And, $\mathcal{L}_{c^{\mathscr{A}}\mathrm{pr}V_{\mathscr{A}}}\Delta \neq c^{\mathscr{A}}\mathcal{L}_{\mathrm{pr}V_{\mathscr{A}}}\Delta$ for source forms, or more generally, for any differential forms. Only for functions $\mathcal{L}_{c^{\mathscr{A}}\mathrm{pr}V_{\mathscr{A}}}f = c^{\mathscr{A}}\mathcal{L}_{\mathrm{pr}V_{\mathscr{A}}}f$ (the $\mathbb{R}$-span is to distinguish from the C^∞-span). In short: $c^{\mathscr{A}}V_{\mathscr{A}}$ is not a symmetry of Δ, but we want to use $c^{\mathscr{A}}V_{\mathscr{A}}$ to get simple local expressions.

Now let us prove Theorem 1.0.2 in the case $n = 1$ and arbitrary m. We divide the proof into different steps and the main results in every step will be written in a box. The results in the boxes will be needed in the next steps. Things which are not written in boxes are basically the proofs of what is written in the boxes.

Proof of Theorem 1.0.2 in the case $n = 1$ and arbitrary m:

Step 1 (transform the ECS): We want to investigate equation (3.16) and we want to reduce the number of unknowns $H_{\alpha\beta},H^x_{\alpha\beta},H^{xx}_{\alpha\beta}$. Since (3.16) is a linear equation for the unknowns, we can apply a sort of Gauss algorithm and eliminate some of the unknowns. Note that

$$\begin{aligned} D_xQ^\beta &= D_x(V^\beta - u^\beta_xV^x) = O^\beta(1) - u^\beta_{xx}V^x, \\ D^2_xQ^\beta &= D^2_x(V^\beta - u^\beta_xV^x) = D_x(O^\beta(1) - u^\beta_{xx}V^x) = O^\beta(2) - u^\beta_{xxx}V^x. \end{aligned}$$

Therefore, we can write the ECS as

$$(V^\beta_{\mathscr{A}} - u^\beta_x V^x_{\mathscr{A}})H_{\alpha\beta} + (O^\beta_{\mathscr{A}}(1) - u^\beta_{xx} V^x_{\mathscr{A}})H^x_{\alpha\beta} + (O^\beta_{\mathscr{A}}(2) - u^\beta_{xxx} V^x_{\mathscr{A}})H^{xx}_{\alpha\beta} = 0,$$

where $\mathscr{A} = 1,2,...,(m+1)$. Using the Local Simplification Lemma 3.4.1, we can take linear combinations

$$c^{\mathscr{A}}(V^\beta_{\mathscr{A}} - u^\beta_x V^x_{\mathscr{A}})H_{\alpha\beta} + c^{\mathscr{A}}(O^\beta_{\mathscr{A}}(1) - u^\beta_{xx} V^x_{\mathscr{A}})H^x_{\alpha\beta} + c^{\mathscr{A}}(O^\beta_{\mathscr{A}}(2) - u^\beta_{xxx} V^x_{\mathscr{A}})H^{xx}_{\alpha\beta} = 0,$$

such that we get the system

$$\begin{aligned} i) \quad & 0 = H_{\alpha\gamma} + O^\beta_\gamma(1)H^x_{\alpha\beta} + O^\beta_\gamma(2)H^{xx}_{\alpha\beta}, \\ ii) \quad & 0 = -u^\beta_x H_{\alpha\beta} + (O^\beta(1) - u^\beta_{xx})H^x_{\alpha\beta} + (O^\beta(2) - u^\beta_{xxx})H^{xx}_{\alpha\beta}. \end{aligned} \tag{3.19}$$

For equation $i)$ we used

$$c^{\mathscr{A}} V^x_{\mathscr{A}} = 0 \quad \text{and} \quad c^{\mathscr{A}} V^\beta_{\mathscr{A}} = \delta^{\beta\gamma}$$

and for equation $ii)$ we used

$$c^{\mathscr{A}} V^x_{\mathscr{A}} = 1 \quad \text{and} \quad c^{\mathscr{A}} V^\beta_{\mathscr{A}} = 0.$$

Then we add $i)$ and $ii)$ such that we eliminate $H_{\alpha\beta}$, i.e.

$$0 = ii) + \sum_{\gamma=1}^{m} u^\gamma_x i) = (\tilde{O}^\beta(1) - u^\beta_{xx})H^x_{\alpha\beta} + (\tilde{O}^\beta(2) - u^\beta_{xxx})H^{xx}_{\alpha\beta},$$

where $\tilde{O}^\beta(1) = u^\gamma_x O^\beta_\gamma(1) + O^\beta(1)$ and $\tilde{O}^\beta(2) = u^\gamma_x O^\beta_\gamma(2) + O^\beta(2)$. Therefore, we derived the **transformed system**

$$\boxed{\begin{aligned} i) \quad & 0 = H_{\alpha\gamma} + O^\beta_\gamma(1)H^x_{\alpha\beta} + O^\beta_\gamma(2)H^{xx}_{\alpha\beta}, \\ ii) \quad & 0 = (O^\beta(1) - u^\beta_{xx})H^x_{\alpha\beta} + (O^\beta(2) - u^\beta_{xxx})H^{xx}_{\alpha\beta} \end{aligned}} \tag{3.20}$$

and we will use it instead of the ECS (3.16). We should remember equations $i)$ and $ii)$, since we will need them several times and from now on we simply call them $i)$ and $ii)$. Note that $i)$ can be considered as the contribution of the vertical part of the symmetry vector fields and $ii)$ the horizontal part.

Step 2 (discuss $u^\gamma_{(4)}$): The Helmholtz expressions in (3.13)-(3.15) show that the order of $i)$ is

$$0 = \underbrace{H_{\alpha\gamma}}_{O_1(4)} + \underbrace{O^\beta_\gamma(1)H^x_{\alpha\beta} + O^\beta_\gamma(2)H^{xx}_{\alpha\beta}}_{O(3)}$$

and since the left hand side is zero, and therefore does not include fourth order coordinates, fourth order coordinates on the right hand side must vanish, as well. If f_α is of second order, then $H_{\alpha\beta}$ in (3.13) is of fourth order. More precisely,

$$H_{\alpha\beta} = O(3) - D_x^2 f_{\beta,u^\alpha_{xx}} = O(3) - u^\gamma_{xxxx} f_{\beta,u^\alpha_{xx}u^\gamma_{xx}}.$$

Therefore, we get $f_{\beta,u^\alpha_{xx}u^\gamma_{xx}} = 0$ for all β,α,γ. We can integrate this equation, which is of the form

$$G_{z^\alpha z^\gamma} = 0, \quad \text{where } G = f_\beta, \quad z^\alpha = u^\alpha_{xx}, \quad z^\gamma = u^\gamma_{xx}.$$

The solution is

$$G = \begin{cases} a(z^\alpha) + b(z^\gamma), & \text{for } z^\alpha \neq z^\gamma, \\ c + z^\alpha d \quad \text{or} \quad c + z^\gamma d, & \text{for } z^\alpha = z^\gamma, \end{cases}$$

where a is a functions which does not depend on z^γ and b is a function which does not depend on z^α. Furthermore, c and d are functions which do not depend on $z^\alpha = z^\gamma$. We solve this equation for all combinations α,$\gamma = 1$,2,...,m and together we get $G = c + z^\delta d_\delta$, where $\delta = 1$,2,...,m and c,d_δ do not depend on $(z^1,z^2,...,z^m)$. For f_β we get

$$\boxed{f_\beta = A_\beta + B_{\beta\gamma} u^\gamma_{xx},} \tag{3.21}$$

where A_β,$B_{\beta\gamma} = O(1)$.

Step 3 (restrictions for the Helmholtz expressions)

Lemma 3.4.2. *If $f_\alpha = A_\alpha + u^\beta_{xx} B_{\alpha\beta}$ is affine linear in the second order jet coordinates u^β_{xx}, where A_β,$B_{\alpha\beta} = O(1)$, then the Helmholtz expressions are of the form*

$$\boxed{\begin{aligned} H_{\alpha\beta} &= O(2) + u^\gamma_{xxx}(B_{\beta\gamma,u^\alpha_x} - B_{\beta\alpha,u^\gamma_x}), \\ H^x_{\alpha\beta} &= O(1) + u^\gamma_{xx}(B_{\alpha\gamma,u^\beta_x} + B_{\beta\gamma,u^\alpha_x} - 2B_{\beta\alpha,u^\gamma_x}), \\ H^{xx}_{\alpha\beta} &= B_{\alpha\beta} - B_{\beta\alpha} = O(1). \end{aligned}} \tag{3.22}$$

Proof of Lemma 3.4.2: Again, for $n = 1$, the Helmholtz expressions can be found in (3.13)-(3.15).
Derivation of $H_{\alpha\beta}$:

$$\begin{aligned} &f_{\alpha,u^\beta} - f_{\beta,u^\alpha} + D_x f_{\beta,u^\alpha_x} - D_x^2 f_{\beta,u^\alpha_{xx}} = \\ &= (A_{\alpha,u^\beta} + u^\gamma_{xx} B_{\alpha\gamma,u^\beta}) - (A_{\beta,u^\alpha} + u^\gamma_{xx} B_{\beta\gamma,u^\alpha}) + D_x(A_{\beta,u^\alpha_x} + u^\gamma_{xx} B_{\beta\gamma,u^\alpha_x}) - D_x^2 B_{\beta\alpha} = \\ &= O(2) + u^\gamma_{xxx} B_{\beta\gamma,u^\alpha_x} - D_x^2 B_{\beta\alpha} = \\ &= O(2) + u^\gamma_{xxx} B_{\beta\gamma,u^\alpha_x} - D_x(O(1) + u^\gamma_{xx} B_{\beta\alpha,u^\gamma_x}) = \\ &= O(2) + u^\gamma_{xxx}(B_{\beta\gamma,u^\alpha_x} - B_{\beta\alpha,u^\gamma_x}). \end{aligned}$$

Derivation of $H^x_{\alpha\beta}$:

$$\begin{aligned} f_{\alpha,u^\beta_x} + f_{\beta,u^\alpha_x} - 2D_x f_{\beta,u^\alpha_{xx}} &= (A_{\alpha,u^\beta_x} + u^\gamma_{xx} B_{\alpha\gamma,u^\beta_x}) + (A_{\beta,u^\alpha_x} + u^\gamma_{xx} B_{\beta\gamma,u^\alpha_x}) - 2D_x B_{\beta\alpha} = \\ &= O(1) + u^\gamma_{xx}(B_{\alpha\gamma,u^\beta_x} + B_{\beta\gamma,u^\alpha_x} - 2B_{\beta\alpha,u^\gamma_x}). \end{aligned}$$

Derivation of $H^{xx}_{\alpha\beta}$: The expression for $H^{xx}_{\alpha\beta}$ can be easily seen and therefore we have proven Lemma 3.4.2. □

Step 4 (discuss $u^\gamma_{(3)}$): Using Lemma 3.4.2, the transformed system (3.20) now has the following order

$$i) \quad 0 = \underbrace{H_{\alpha\gamma}}_{O_1(3)} + O^\beta_\gamma(1) \underbrace{H^x_{\alpha\beta}}_{O_1(2)} + O^\beta_\gamma(2) \underbrace{H^{xx}_{\alpha\beta}}_{O(1)}, \tag{3.23}$$

$$ii) \quad 0 = \underbrace{(O^\beta(1) - u^\beta_{xx})}_{O_1(2)} \underbrace{H^x_{\alpha\beta}}_{O_1(2)} + \underbrace{(O^\beta(2) - u^\beta_{xxx})}_{O_1(3)} \underbrace{H^{xx}_{\alpha\beta}}_{O(1)}. \tag{3.24}$$

Since $i)$ does not have order three on the left hand side, it must also vanish on the right hand side and this means

$$\boxed{0 = B_{\beta\gamma,u^\alpha_x} - B_{\beta\alpha,u^\gamma_x}.} \tag{3.25}$$

And, since $ii)$ does not have order three on the left hand side, it must also vanish on the right hand side and this means

$$\boxed{0 = H^{xx}_{\alpha\beta} = B_{\alpha\beta} - B_{\beta\alpha}.} \tag{3.26}$$

Step 5 (restrictions for the Helmholtz expressions): Condition (3.25) and (3.26) forces that $B_{\beta\gamma,u^\alpha_x}$ must be symmetric in all of the three indices α,β,γ and this leads to

$$0 = B_{\alpha\gamma,u^\beta_x} + B_{\beta\gamma,u^\alpha_x} - 2B_{\beta\alpha,u^\gamma_x}.$$

Therefore,

$$\boxed{\begin{aligned} H_{\alpha\beta} &= O(2), \\ H^x_{\alpha\beta} &= O(1), \\ H^{xx}_{\alpha\beta} &= 0. \end{aligned}}$$

Step 6 (discuss u^γ_{xx}): Now (3.24) becomes

$$ii) \quad 0 = \underbrace{(O^\beta(1) - u^\beta_{xx})}_{O_1(2)} \underbrace{H^x_{\alpha\beta}}_{O(1)},$$

and therefore $H^x_{\alpha\beta} = 0$. Then from (3.23) we get

$$i) \quad 0 = H_{\alpha\gamma}$$

and all Helmholtz conditions are satisfied. □

3.4.1. The Algorithm for $n = 1$, Arbitrary m and 2nd Order

Here, we formulate again the main steps in the proof:

Step 1:	transform the ECS, (simplify the ECS)
Step 2:	discuss $u^\gamma_{(4)}$, (fourth order)
Step 2:	restrictions for $H_{\alpha\beta}$,$H^x_{\alpha\beta}$,$H^{xx}_{\alpha\beta}$,
Step 4:	discuss $u^\gamma_{(3)}$, (third order)
Step 5:	restrictions for $H_{\alpha\beta}$,$H^x_{\alpha\beta}$,$H^{xx}_{\alpha\beta}$,
Step 6:	discuss u^γ_{xx}, (all Helmholtz conditions are satisfied).

Note that in a more general situation, for example when we consider PDEs, it is probably not clear whether we should start with the discussion of the different orders in i) or ii) first. In the proof above there was no big difference and we did not notice it explicitly in the algorithm. More precisely, in Step 4 there is no difference whether we first discuss i) and then ii) or first discuss ii) and then i). In the next section, we have to find a generalization of this algorithm and we also explain why we chose this algorithm.

3.5. The Problem of the Proof for $n = 2$, Arbitrary m and 2nd Order

Since the proof from the previous section for $n = 1$ and arbitrary m was quite nice, it is obvious to try to formulate a similar proof also for $n = 2$ and arbitrary m. However, we will find out that this is not so easy.

The following discussion shows that it is not a good idea to generalize the previous proof (the previous algorithm), especially when $n > 2$. It is probably even impossible to formulate such a proof for $n > 2$, because the notation is getting to complicated. However, the following (idea of a) proof has the advantage, that it is very systematically, because it is a sorting of polynomial degree and order of jet coordinates and it makes a lot of sense to do so, at least theoretically.

There are two reasons why we would like to present the following idea of a proof. The first reason is that it shows that there is no simple algorithm, which can be applied to solve Takens' problem. That is, a possible notation heavy, but systematically simple algorithm does not exist and this is an important observation (at least we did not find a simple algorithm). The second reason is that it gives an overview of techniques and expressions which will occur in Takens' problem. These expressions are especially the so called **Hyperjacobians**. Roughly speaking, Hyperjacobians are a generalization of affine linear from the previous section, where we found out that $f_\beta = A_\beta + u^\gamma_{xx} B_{\beta\gamma}$ in Step 2. When one wants to solve Takens' problem, the first thing one would try to do is probably to use this systematic approach. But then one will find out, that it is not the best method and one has to improve the notation and formalism. In Section 3.6, we present a general proof of Theorem 1.0.2 which is in some sense much shorter than the (idea of a) proof in this section. Actually, we will not present a complete proof of Theorem 1.0.2 in this section. We will just start with the proof and stop it when the notation is getting too complicated. Note that the following idea of a proof can be completed for $n = 2$ and arbitrary m, but we do not present it in this dissertation.

The main features of the following (idea of a) proof are:

- Generalize the nice proof from the previous section (it is obvious to try to do so).
- Discuss systematically the order and degree of jet coordinates (which makes sense in principle).
- Find out if a tensor is symmetric or skew-symmetric in some of the indices or both and combine these conditions.

The setting here is the same as in the ODE case. We want to prove Theorem 1.0.2 for $n = 2$, arbitrary m. Symmetries and conservation laws force the ECS

$$Q^\beta H_{\alpha\beta} + (D_i Q^\beta) H^i_{\alpha\beta} + (D_{ij} Q^\beta) H^{ij}_{\alpha\beta} = 0, \tag{3.27}$$

where

$$H_{\alpha\beta} = f_{\alpha,u^\beta} - f_{\beta,u^\alpha} + D_i f_{\beta,u^\alpha_i} - c_{ij} D_i D_j f_{\beta,u^\alpha_{ij}}, \tag{3.28}$$

$$H^i_{\alpha\beta} = f_{\alpha,u^\beta_i} + f_{\beta,u^\alpha_i} - 2c_{ij} D_j f_{\beta,u^\alpha_{ij}}, \tag{3.29}$$

$$H^{ij}_{\alpha\beta} = c_{ij}(f_{\alpha,u^\beta_{ij}} - f_{\beta,u^\alpha_{ij}}) \tag{3.30}$$

are the Helmholtz expressions (also see (AP94)) and

$$c_{ij} = \begin{cases} 1, & i = j, \\ \frac{1}{2}, & i \neq j. \end{cases}$$

The proof is again divided into different steps. We will discuss successively

$$O(4),$$
$$O(3),$$
$$O(2).$$

Again, the main results in every step are written in a box and the results in the boxes will be needed in the next steps. Things which are not written in a box are basically the proofs of what is written in the boxes.

Proof of Theorem 1.0.2 for $n = 2$, arbitrary m: The characteristics Q^β are of the form

$$Q^\beta = V^\beta(x^j,u^\alpha) - u^\beta_i V^i(x^j).$$

Therefore, equation (3.27) can be written as

$$(V^\beta_{\mathscr{A}} - u^\beta_i V^i_{\mathscr{A}}) H_{\alpha\beta} + (O^\beta_{\mathscr{A}}(1) - u^\beta_{ik} V^i_{\mathscr{A}}) H^k_{\alpha\beta} + (O^\beta_{\mathscr{A}}(2) - u^\beta_{ikl} V^i_{\mathscr{A}}) H^{kl}_{\alpha\beta} = 0, \tag{3.31}$$

where $\mathscr{A} = 1,...,(2+m)$ and $\alpha = 1,2,...,m$.

Step 1 (transform the ECS): We want to apply a sort of Gauss algorithm to the system (3.31) and thereby eliminate some of the unknowns $H_{\alpha\beta}, H^i_{\alpha\beta}, H^{ij}_{\alpha\beta}$. Counting the unknowns shows that it is in general only possible to eliminate $H_{\alpha\beta}$. Using the Local Simplification Lemma 3.4.1, we can take linear combinations of (3.31)

$$c^{\mathscr{A}}(V^\beta_{\mathscr{A}} - u^\beta_i V^i_{\mathscr{A}}) H_{\alpha\beta} + c^{\mathscr{A}}(O^\beta_{\mathscr{A}}(1) - u^\beta_{ik} V^i_{\mathscr{A}}) H^k_{\alpha\beta} + c^{\mathscr{A}}(O^\beta_{\mathscr{A}}(2) - u^\beta_{ikl} V^i_{\mathscr{A}}) H^{kl}_{\alpha\beta} = 0$$

such that we get

$$\begin{aligned} i) \quad & 0 = H_{\alpha\gamma} + O^{\beta}_{\gamma,k}(1)H^{k}_{\alpha\beta} + O^{\beta}_{\gamma,kl}(2)H^{kl}_{\alpha\beta}, \\ ii) \quad & 0 = -u^{\beta}_{j}H_{\alpha\beta} + (O^{\beta}_{jk}(1) - u^{\beta}_{jk})H^{k}_{\alpha\beta} + (O^{\beta}_{jkl}(2) - u^{\beta}_{jkl})H^{kl}_{\alpha\beta}. \end{aligned}$$

For $i)$ we used

$$c^{\mathscr{A}}V^{i}_{\mathscr{A}} = 0 \quad \text{and} \quad c^{\mathscr{A}}V^{\beta}_{\mathscr{A}} = \delta^{\beta\gamma}$$

and for $ii)$ we used

$$c^{\mathscr{A}}V^{i}_{\mathscr{A}} = \delta^{ij} \quad \text{and} \quad c^{\mathscr{A}}V^{\beta}_{\mathscr{A}} = 0.$$

We want to eliminate $H_{\alpha\beta}$ and therefore we consider

$$0 = ii) + \sum_{\gamma=1}^{m} u^{\gamma}_{j} i) = (\tilde{O}^{\beta}_{jk}(1) - u^{\beta}_{jk})H^{k}_{\alpha\beta} + (\tilde{O}^{\beta}_{jkl}(2) - u^{\beta}_{jkl})H^{kl}_{\alpha\beta}.$$

Then we get the **transformed system**

$$\boxed{\begin{aligned} i) \quad & 0 = H_{\alpha\gamma} + O^{\beta}_{\gamma}(1)H^{k}_{\alpha\beta} + O^{\beta}_{\gamma}(2)H^{kl}_{\alpha\beta}, \\ ii) \quad & 0 = (O^{\beta}_{jk}(1) - u^{\beta}_{jk})H^{k}_{\alpha\beta} + (O^{\beta}_{jkl}(2) - u^{\beta}_{jkl})H^{kl}_{\alpha\beta} \end{aligned}} \qquad \begin{aligned} (3.32) \\ (3.33) \end{aligned}$$

and we simply call these two equations $i)$ and $ii)$. We should remember the equations $i)$ and $ii)$, since we will need them several times.

Step 2 (discuss $u^{\gamma}_{(4)}$): Since the left hand side of $i)$ is zero, and therefore does not include fourth order coordinates, fourth order must also vanish on the right hand side and only $H_{\alpha\beta}$ can include fourth order (see (3.28)).[1] Fourth order in $H_{\alpha\beta}$ is given as

$$H_{\alpha\beta} = O(3) - c_{ij}D_iD_jf_{\beta,u^{\alpha}_{ij}} = O(3) - c_{ij}c_{kl}u^{\gamma}_{ijkl}f_{\beta,u^{\alpha}_{ij}u^{\gamma}_{kl}}. \tag{3.34}$$

It is important not to forget the factors c_{kl} in (3.34), otherwise the following discussion leads to a different result. We want to analyze the fourth order of (3.34) in detail:

$$u^{\gamma}_{xxxx}: \quad 0 = f_{\beta,u^{\alpha}_{xx}u^{\gamma}_{xx}}, \tag{3.35}$$
$$u^{\gamma}_{xxxy}: \quad 0 = f_{\beta,u^{\alpha}_{xx}u^{\gamma}_{xy}} + f_{\beta,u^{\alpha}_{xy}u^{\gamma}_{xx}}, \tag{3.36}$$
$$u^{\gamma}_{xxyy}: \quad 0 = f_{\beta,u^{\alpha}_{xx}u^{\gamma}_{yy}} + f_{\beta,u^{\alpha}_{xy}u^{\gamma}_{xy}} + f_{\beta,u^{\alpha}_{yy}u^{\gamma}_{xx}}, \tag{3.37}$$
$$u^{\gamma}_{xyyy}: \quad 0 = f_{\beta,u^{\alpha}_{xy}u^{\gamma}_{yy}} + f_{\beta,u^{\alpha}_{yy}u^{\gamma}_{xy}}, \tag{3.38}$$
$$u^{\gamma}_{yyyy}: \quad 0 = f_{\beta,u^{\alpha}_{yy}u^{\gamma}_{yy}}. \tag{3.39}$$

[1] It is also possible to consider the conservation law condition $\mathcal{E}_{\alpha}(f_{\beta}Q^{\beta}) = 0$, but discussing successively the order of the Helmholtz expressions seems to be more systematically here. Note that $E(\iota_{\mathrm{pr}Q}\Delta) = 0$ means that $\iota_{\mathrm{pr}Q}\Delta$ is a trivial Lagrangian and this is discussed in (AD80), especially Theorem 4.3 therein.

We can integrate these equations and find the general solution. However, it needs some work and notation. Surprisingly, the general solution can be described by a finite dimensional vector space, which is not obvious for PDEs. We start with (3.35). This equation is of the form

$$G_{z^\alpha z^\gamma} = 0, \quad \text{where } G = f_\beta, \quad z^\alpha = u^\alpha_{xx}, \quad z^\gamma = u^\gamma_{xx}$$

and the solution is

$$\begin{cases} G = a(z^\alpha) + b(z^\gamma), & \text{for } z^\alpha \neq z^\gamma, \\ G = c + z^\alpha d \quad \text{or} \quad G = c + z^\gamma d, & \text{for } z^\alpha = z^\gamma. \end{cases}$$

Together

$$G = c + z^\delta d_\delta,$$

where c,d_δ do not depend on z^δ, or in our case, $f_\beta = A + u^\gamma_{xx} B_\gamma$, where A,B_γ do not depend on u^α_{xx}. A similar calculation holds for (3.39) and together we get

$$\begin{aligned} f_\beta &= C + u^\gamma_{yy} D_\gamma + u^\gamma_{xx}(E_\gamma + u^\delta_{yy} F_{\gamma\delta}) = \\ &= C + u^\gamma_{yy} D_\gamma + u^\gamma_{xx} E_\gamma + u^\gamma_{xx} u^\delta_{yy} F_{\gamma\delta}, \end{aligned}$$

where C,D_γ,E_γ,$F_{\gamma\delta}$ do not depend on u^γ_{xx},u^γ_{yy}, but on u^γ_{xy} and first order coordinates. Note that we do not write the index β on C,D_γ,E_γ,$F_{\gamma\delta}$, to keep the notation as simple as possible. Then the condition (3.37) leads to

$$(F_{\alpha\gamma} + F_{\gamma\alpha}) + \frac{\partial^2}{\partial u^\alpha_{xy} \partial u^\gamma_{xy}} \left(C + u^\delta_{yy} D_\delta + u^\delta_{xx} E_\delta + u^\delta_{xx} u^\epsilon_{yy} F_{\delta\epsilon} \right) = 0$$

and this forces

$$E_{\delta, u^\alpha_{xy} u^\gamma_{xy}} = 0, \tag{3.40}$$

$$D_{\delta, u^\alpha_{xy} u^\gamma_{xy}} = 0, \tag{3.41}$$

$$F_{\delta\epsilon, u^\alpha_{xy} u^\gamma_{xy}} = 0, \tag{3.42}$$

$$F_{\alpha\gamma} + F_{\gamma\alpha} + \frac{\partial^2}{\partial u^\alpha_{xy} \partial u^\gamma_{xy}} C = 0. \tag{3.43}$$

We can integrate the equations (3.40), (3.41), (3.42), which are again of the form $G_{z^\alpha z^\gamma} = 0$, where $z^\alpha = u^\alpha_{xy}$, $z^\gamma = u^\gamma_{xy}$ and we get

$$\begin{aligned} E_\delta &= e_\delta + u^\eta_{xy} E_{\delta\eta}, \\ D_\delta &= d_\delta + u^\eta_{xy} D_{\delta\eta}, \\ F_{\delta\epsilon} &= f_{\delta\epsilon} + u^\eta_{xy} F_{\delta\epsilon\eta}, \end{aligned}$$

where e_δ,$E_{\delta\eta}$,d_δ,$D_{\delta\eta}$,$f_{\delta\epsilon}$,$F_{\delta\epsilon\eta}$ do not depend on u^γ_{xx},u^γ_{xy},u^γ_{yy}, i.e. they are of first order. Now f_β can be written as

$$\begin{aligned} f_\beta =& C + u^\alpha_{xx}(e_\alpha + u^\gamma_{xy}E_{\alpha\gamma}) + u^\alpha_{yy}(d_\alpha + u^\gamma_{xy}D_{\alpha\gamma}) + u^\alpha_{xx}u^\gamma_{yy}(f_{\alpha\gamma} + u^\delta_{xy}F_{\alpha\gamma\delta}) = \\ =& C + u^\alpha_{xx}e_\alpha + u^\alpha_{yy}d_\alpha + u^\alpha_{xx}u^\gamma_{xy}E_{\alpha\gamma} + u^\alpha_{yy}u^\gamma_{xy}D_{\alpha\gamma} + u^\alpha_{xx}u^\gamma_{yy}f_{\alpha\gamma} + u^\alpha_{xx}u^\gamma_{yy}u^\delta_{xy}F_{\alpha\gamma\delta}. \end{aligned}$$

Then the condition (3.36) leads to

$$f_{\beta,u^\alpha_{xx}u^\gamma_{xy}} + f_{\beta,u^\alpha_{xy}u^\gamma_{xx}} = E_{\alpha\gamma} + E_{\gamma\alpha} + u^\beta_{yy}(F_{\alpha\beta\gamma} + F_{\gamma\beta\alpha}) = 0$$

and this forces

$$\begin{aligned} 0 &= E_{\alpha\gamma} + E_{\gamma\alpha}, \\ 0 &= F_{\alpha\beta\gamma} + F_{\gamma\beta\alpha}. \quad (\beta \text{ is fixed in the middle position}) \end{aligned} \tag{3.44}$$

Then the condition (3.38) leads to

$$0 = f_{\beta,u^\alpha_{xy}u^\gamma_{yy}} + f_{\beta,u^\alpha_{yy}u^\gamma_{xy}} = D_{\gamma\alpha} + D_{\alpha\gamma} + u^\beta_{xx}(F_{\beta\gamma\alpha} + F_{\beta\alpha\gamma})$$

and this forces

$$\begin{aligned} 0 &= D_{\gamma\alpha} + D_{\alpha\gamma}, \\ 0 &= F_{\beta\gamma\alpha} + F_{\beta\alpha\gamma}. \quad (\beta \text{ is fixed in the first position}) \end{aligned} \tag{3.45}$$

From (3.44) and (3.45) we get

$$\begin{aligned} & 0 = F_{\alpha\beta\gamma} + F_{\gamma\beta\alpha}, \quad (\text{fix middle index } \beta) \\ & 0 = F_{\alpha\beta\gamma} + F_{\alpha\gamma\beta}, \quad (\text{fix first index } \alpha) \\ \Rightarrow \quad & 0 = F_{\alpha\beta\gamma} + F_{\beta\alpha\gamma}, \quad (\text{fix last index } \gamma) \end{aligned}$$

because of

$$0 = F_{\alpha\beta\gamma} + \underbrace{F_{\gamma\beta\alpha}}_{\text{fix first}} = F_{\alpha\beta\gamma} - \underbrace{F_{\gamma\alpha\beta}}_{\text{fix middle}} = F_{\alpha\beta\gamma} + F_{\beta\alpha\gamma}.$$

Therefore, $F_{\alpha\beta\gamma}$ has to be skew-symmetric whenever we change two indices.[2] Now the condition (3.43) can be written as

$$f_{\alpha\gamma} + f_{\gamma\alpha} + u^\eta_{xy}\underbrace{(F_{\alpha\gamma\eta} + F_{\gamma\alpha\eta})}_{=0} + \frac{\partial^2}{\partial u^\alpha_{xy}\partial u^\gamma_{xy}}C = f_{\alpha\gamma} + f_{\gamma\alpha} + \frac{\partial^2}{\partial u^\alpha_{xy}\partial u^\gamma_{xy}}C = 0. \tag{3.46}$$

[2]For $m = 2$, we know more than $F_{\alpha\gamma\eta} + F_{\gamma\alpha\eta} = 0$. Then we even know that $F_{\alpha\beta\gamma} = 0$. Because, for $m = 2$, two of the three indices α,β,γ have to be the same, since only $F_{111} = 0$, $F_{112} = 0$, $F_{221} = 0$ and $F_{222} = 0$ and permutations of that are possible.

We can integrate (3.46) and we get

$$C = -u^\alpha_{xy} u^\gamma_{xy} f_{\alpha\gamma} + u^\alpha_{xy} c_\alpha + h,$$

where $f_{\alpha\gamma}$,c_α,h must be of first order. Now f_β can be written as

$$\begin{aligned} f_\beta =& h + u^\alpha_{xy} c_\alpha + u^\alpha_{xx} e_\alpha + u^\alpha_{yy} d_\alpha - \\ & - u^\alpha_{xy} u^\gamma_{xy} f_{\alpha\gamma} + u^\alpha_{xx} u^\gamma_{xy} E_{\alpha\gamma} + u^\alpha_{yy} u^\gamma_{xy} D_{\alpha\gamma} + u^\alpha_{xx} u^\gamma_{yy} f_{\alpha\gamma} + u^\gamma_{xx} u^\delta_{xy} u^\epsilon_{yy} F_{\gamma\delta\epsilon} = \\ \\ =& h + u^\alpha_{xy} c_\alpha + u^\alpha_{xx} e_\alpha + u^\alpha_{yy} d_\alpha + \\ & + (u^\alpha_{xx} u^\gamma_{yy} - u^\alpha_{xy} u^\gamma_{xy}) f_{\alpha\gamma} + u^\alpha_{xx} u^\gamma_{xy} E_{\alpha\gamma} + u^\alpha_{yy} u^\gamma_{xy} D_{\alpha\gamma} + u^\gamma_{xx} u^\delta_{xy} u^\epsilon_{yy} F_{\gamma\delta\epsilon}, \end{aligned}$$

where $E_{\alpha\gamma}$ and $D_{\alpha\gamma}$ have to be skew-symmetric and there are no restrictions for $f_{\alpha\gamma}$. Let us write this differently as

$$\begin{aligned} f_\beta =& h + u^\alpha_{xy} c_\alpha + u^\alpha_{xx} e_\alpha + u^\alpha_{yy} d_\alpha + (u^\alpha_{xx} u^\gamma_{yy} - u^\alpha_{xy} u^\gamma_{xy}) f_{\alpha\gamma} + \\ & + \sum_{\alpha<\gamma} (u^\alpha_{xx} u^\gamma_{xy} - u^\gamma_{xx} u^\alpha_{xy}) E_{\alpha\gamma} + \sum_{\alpha<\gamma} (u^\alpha_{yy} u^\gamma_{xy} - u^\gamma_{yy} u^\alpha_{xy}) D_{\alpha\gamma} + u^\gamma_{xx} u^\delta_{xy} u^\epsilon_{yy} F_{\gamma\delta\epsilon} = \\ \\ =& h + u^\alpha_{xy} c_\alpha + u^\alpha_{xx} e_\alpha + u^\alpha_{yy} d_\alpha + (u^\alpha_x, u^\gamma_y) f_{\alpha\gamma} + \\ & + \sum_{\alpha<\gamma} (u^\alpha_x, u^\gamma_x) E_{\alpha\gamma} + \sum_{\alpha<\gamma} (u^\alpha_y, u^\gamma_y) D_{\alpha\gamma} + u^\gamma_{xx} u^\delta_{xy} u^\epsilon_{yy} F_{\gamma\delta\epsilon}, \end{aligned}$$

where we define the expressions

$$(u^\alpha_i, u^\beta_j) := \det(\nabla u^\alpha_i, \nabla u^\beta_j)) = \frac{D(u^\alpha_i, u^\beta_j)}{D(x,y)} = u^\alpha_{i1} u^\beta_{j2} - u^\alpha_{i2} u^\beta_{j1} \tag{3.47}$$

and they are so-called **Monge-Ampere expressions** or **Hyperjacobians** (note that $u_{xx} u_{yy} - u^2_{xy} = 0$ is called Monge-Ampere equation). For the more general definition of Hyperjacobians see (Olv83). We also define the determinant

$$(u^\gamma, u^\delta, u^\epsilon) := \begin{vmatrix} u^\gamma_{xx} & u^\delta_{xx} & u^\epsilon_{xx} \\ u^\gamma_{xy} & u^\delta_{xy} & u^\epsilon_{xy} \\ u^\gamma_{yy} & u^\delta_{yy} & u^\epsilon_{yy} \end{vmatrix},$$

which is not a Hyperjacobian. We finish Step 2 with some change in the notation and we will write

$$\boxed{f_\alpha = A_\alpha + u^\gamma_{ij} B^{ij}_{\alpha|\gamma} + (u^\gamma_i, u^\delta_j) C^{ij}_{\alpha|\gamma\delta} + (u^\gamma, u^\delta, u^\epsilon) D_{\alpha|\gamma\delta\epsilon},} \tag{3.48}$$

where $A_\alpha, B^{ij}_{\alpha|\gamma}, C^{ij}_{\alpha|\gamma\delta}, D_{\alpha|\gamma\delta\epsilon} = O(1)$, and furthermore $D_{\alpha|\gamma\delta\epsilon}$ is skew-symmetric in γ,δ,ϵ. Note that for $m = 2$, two of the three indices γ,δ,ϵ in $D_{\alpha|\gamma\delta\epsilon}$ have to be the same. For example, $D_{\alpha|112}$ or $D_{\alpha|221}$ and therefore $D_{\alpha|\gamma\delta\epsilon} = 0$ when $m = 2$.

Step 2 b) (explaining some problems in the case $n,m = 2$): For simplicity, we only consider $n,m = 2$ in Step 2 b). Compared to Step 2 in the ODE case in Section 3.4, we saw that we needed a lot more notation and work to get the result in Step 2 in this section. The more complicated notation causes problems in the following discussion of the proof. Let us explain this in more detail. However, note that this step is only partially relevant for the proof and one can also continue with Step 3 (that is why we also called it Step 2 b)). This step only explains some notational problems.

After discussing $O_1(4)$ in the ODE case in Section 3.4, we derived

$$f_\alpha = A_\alpha + u^\gamma_{xx} B_{\alpha\gamma} \quad (f_\alpha \text{ is affine linear in } u^\gamma_{xx})$$

and in the PDE case here, where $n = 2$, we derived

$$f_\alpha = A_\alpha + u^\gamma_{ij} B^{ij}_{\alpha|\gamma} + (u^\gamma_i,u^\delta_j) C^{ij}_{\alpha|\gamma\delta} + (u^\gamma,u^\delta,u^\epsilon) D_{\alpha|\gamma\delta\epsilon},$$

where the last expression $D_{\alpha|\gamma\delta\epsilon}$ vanishes, when $m = 2$. As we already mentioned, this is because $D_{\alpha|\gamma\delta\epsilon}$ is skew-symmetric in γ,δ,ϵ and at least two of the three indices γ,δ,ϵ have to be the same when $m = 2$ (this only holds for $m = 2$, otherwise we have to discuss $u^\alpha_{(3)} u^\beta_{(3)}$ as well). In the PDE case, we now have to handle some problems and the main problem is the so-called double counting, which was already mentioned in Section 2.12. The first observation is that

$$\begin{aligned} u^\gamma_{ij} B^{ij}_{\alpha|\gamma} &= u^\gamma_{xx} B^{xx}_{\alpha|\gamma} + u^\gamma_{xy} B^{xy}_{\alpha|\gamma} + u^\gamma_{yx} B^{yx}_{\alpha|\gamma} + u^\gamma_{yy} B^{yy}_{\alpha|\gamma} = \\ &= u^\gamma_{xx} B^{xx}_{\alpha|\gamma} + u^\gamma_{xy} (B^{xy}_{\alpha|\gamma} + B^{yx}_{\alpha|\gamma}) + u^\gamma_{yy} B^{yy}_{\alpha|\gamma} = \\ &= u^\gamma_{xx} \tilde{B}^{xx}_{\alpha|\gamma} + u^\gamma_{xy} \tilde{B}^{xy}_{\alpha|\gamma} + u^\gamma_{yy} \tilde{B}^{yy}_{\alpha|\gamma} \end{aligned}$$

and since $u^\gamma_{xy} = u^\gamma_{yx}$ (jet coordinates are $(x,u,u_x,u_y,u_{xx},u_{xy},u_{yy})$, see (2.123) and (2.124)) we define

$$\begin{aligned} \tilde{B}^{ij}_{\alpha|\gamma} &:= B^{ij}_{\alpha|\gamma} + B^{ji}_{\alpha|\gamma}, && \text{if } i < j, \\ \tilde{B}^{ii}_{\alpha|\gamma} &:= B^{ii}_{\alpha|\gamma}, && \text{if } i = j. \end{aligned}$$

Then we get the ordered sum

$$u^\gamma_{ij} B^{ij}_{\alpha|\gamma} = \sum_\gamma \sum_{i,j} u^\gamma_{ij} B^{ij}_{\alpha|\gamma} = \sum_\gamma \sum_{i \le j} u^\gamma_{ij} \tilde{B}^{ij}_{\alpha|\gamma}.$$

The ordered sum has the advantage that when we have, for example, an equation

$$\sum_\gamma \sum_{i \le j} u^\gamma_{ij} \tilde{B}^{ij}_{\alpha|\gamma} = 0,$$

then we get immediately that all coefficients $\tilde{B}^{ij}_{\alpha|\gamma} = 0$, whereas

$$\sum_{\gamma} \sum_{i,j} u^{\gamma}_{ij} B^{ij}_{\alpha|\gamma} = 0$$

does not imply $B^{ij}_{\alpha|\gamma} = 0$.

To order the Hyperjacobians needs more effort. There are sixteen terms and they can be written as

$$\begin{aligned}
(u^{\gamma}_i,u^{\delta}_j)C^{ij}_{\alpha|\gamma\delta} &= (u^{\gamma}_x,u^{\delta}_x)C^{xx}_{\alpha|\gamma\delta} + (u^{\gamma}_x,u^{\delta}_y)C^{xy}_{\alpha|\gamma\delta} + (u^{\gamma}_y,u^{\delta}_x)C^{yx}_{\alpha|\gamma\delta} + (u^{\gamma}_y,u^{\delta}_y)C^{yy}_{\alpha|\gamma\delta} = \\
&= (u^1_x,u^1_x)C^{xx}_{\alpha|11} + (u^1_x,u^2_x)C^{xx}_{\alpha|12} + (u^2_x,u^1_x)C^{xx}_{\alpha|21} + (u^2_x,u^2_x)C^{xx}_{\alpha|22} + \\
&+ (u^1_x,u^1_y)C^{xy}_{\alpha|11} + (u^1_x,u^2_y)C^{xy}_{\alpha|12} + (u^2_x,u^1_y)C^{xy}_{\alpha|21} + (u^2_x,u^2_y)C^{xy}_{\alpha|22} + \\
&+ (u^1_y,u^1_x)C^{yx}_{\alpha|11} + (u^1_y,u^2_x)C^{yx}_{\alpha|12} + (u^2_y,u^1_x)C^{yx}_{\alpha|21} + (u^2_y,u^2_x)C^{yx}_{\alpha|22} + \\
&+ (u^1_y,u^1_y)C^{yy}_{\alpha|11} + (u^1_y,u^2_y)C^{yy}_{\alpha|12} + (u^2_y,u^1_y)C^{yy}_{\alpha|21} + (u^2_y,u^2_y)C^{yy}_{\alpha|22} = \\
&= (u^1_x,u^2_x)(C^{xx}_{\alpha|12} - C^{xx}_{\alpha|21}) + \\
&+ (u^1_x,u^1_y)(C^{xy}_{\alpha|11} - C^{yx}_{\alpha|11}) + (u^1_x,u^2_y)(C^{xy}_{\alpha|12} - C^{yx}_{\alpha|21}) + \\
&+ (u^2_x,u^1_y)(C^{xy}_{\alpha|21} - C^{yx}_{\alpha|12}) + (u^2_x,u^2_y)(C^{xy}_{\alpha|22} - C^{yx}_{\alpha|22}) + \\
&+ (u^1_y,u^2_y)(C^{yy}_{\alpha|12} - C^{yy}_{\alpha|21}).
\end{aligned}$$

By definition of the Hyperjacobians we have

$$(u^1_x,u^1_x) = (u^2_x,u^2_x) = (u^1_y,u^1_y) = (u^2_y,u^2_y) = 0, \tag{3.49}$$

and, more generally, we have the identity

$$(u^{\alpha}_i,u^{\beta}_j) = u^{\alpha}_{ix}u^{\beta}_{jy} - u^{\alpha}_{iy}u^{\beta}_{jx} = -(u^{\beta}_j,u^{\alpha}_i). \tag{3.50}$$

Now we want to order the Hyperjacobians. Let us define

$$\begin{cases} \tilde{C}^{ii}_{\alpha\gamma\delta} = \tilde{C}^{ii}_{\alpha12} := C^{ii}_{\alpha12} - C^{ii}_{\alpha21}, & \text{if } i = j \text{ and } \gamma < \delta \\ \tilde{C}^{ij}_{\alpha\gamma\delta} = \tilde{C}^{xy}_{\alpha\gamma\delta} := C^{xy}_{\alpha\gamma\delta} - C^{yx}_{\alpha\delta\gamma}, & \text{if } i < j \text{ and for all } \gamma,\delta. \end{cases}$$

Then we can write

$$(u^{\gamma}_i,u^{\delta}_j)C^{ij}_{\alpha\gamma\delta} = \sum_{\substack{\gamma,\delta \\ \gamma<\delta \text{ if } i=j}} \sum_{i\le j} (u^{\gamma}_i,u^{\delta}_j)\tilde{C}^{ij}_{\alpha\gamma\delta}. \tag{3.51}$$

We have the ordered expressions (we write $u^1 = u$ and $u^2 = v$)

$$\begin{aligned}
z_1 &:= (u_x, v_x), \\
z_2 &:= (u_x, u_y), \\
z_3 &:= (v_x, u_y), \\
z_4 &:= (u_x, v_y), \\
z_5 &:= (u_y, v_y), \\
z_6 &:= (v_x, v_y),
\end{aligned} \tag{3.52}$$

and they are also characterized by the property that they are linearly independent (this is easy to check for $n = 2$ and one can do it as an exercise). The expressions in (3.52) are not completely independent, since there are non-linear relations like

$$z_1 z_5 + z_4 z_3 = z_2 z_6. \tag{3.53}$$

This is called a *Plücker relation*, see (KPS08). Note that this relation is a special case of Syzygy. When we consider $z_1, z_2, \ldots, z_6$ as new coordinates, they only form a five-dimensional manifold and not a six-dimensional one, as we would might expect, since they are generated by the six coordinates

$$u^1_{xx}, u^1_{xy}, u^1_{yy}, u^2_{xx}, u^2_{xy}, u^2_{yy}.$$

There are further relations like (we write $u^1_{xx} = u_{xx}$, $u^2_{xx} = v_{xx}$ and so on)

$$\begin{aligned}
v_{xx} z_2 - u_{xy} z_1 &= v_{xx}(u_{xx} u_{yy} - u^2_{xy}) - u_{xy}(u_{xx} v_{xy} - u_{xy} v_{xx}) = \\
&= v_{xx} u_{xx} u_{yy} - u_{xy} u_{xx} v_{xy} = u_{xx} z_3,
\end{aligned}$$

and therefore

$$v_{xx} z_1 = u_{xy} z_2 + u_{xx} z_3. \tag{3.54}$$

Note that in the case $n = 4$, there are also linear relations among the Hyperjacobians. For example,

$$\frac{D(u_1, u_2)}{D(x^3, x^4)} - \frac{D(u_1, u_3)}{D(x^2, x^4)} + \frac{D(u_1, u_4)}{D(x^2, x^3)} = 0. \tag{3.55}$$

We took (3.55) from (BCO81, p.155), where it can be found in formula (4.6) (there, one can also find the more general definition of Hyperjacobian). In (BCO81, p.156) is also written that it seems to be difficult to find the number of linearly independent Hyperjacobians for arbitrary n.

Step 3 (restrictions for the Helmholtz expressions):
Similar to Step 3 in the proof in the ODE case, see Lemma 3.4.2, now we have to compute the Helmholtz expressions explicitely, under the assumption that f_α can be written as (3.48). It turns out that these expressions are very complicated, and therefore we will stop the proof here. We will anyway formulate a better proof for arbitrary n in Section 3.6. Let us finish this section with the algorithm how the proof would work in principle.

3.5.1. The Algorithm for $n = 2$, Arbitrary m and 2nd Order

In this subsection, we explain the structure of the previous idea of a proof and we give some remarks on why we chose this algorithm, also why we chose a similar algorithm for $n = 1$ in Section 3.4. We want to understand the following steps:

Step 1:	transform the ECS,	
Step 2:	discuss $u^\gamma_{(4)}$,	(degree 1)
Step 3:	restrictions for $H_{\alpha\beta}$,$H^i_{\alpha\beta}$,$H^{ij}_{\alpha\beta}$,	
Step 4:	discuss $u^\gamma_{(3)}u^\delta_{(3)}$,	(degree 2)
Step 5:	restrictions for $H_{\alpha\beta}$,$H^i_{\alpha\beta}$,$H^{ij}_{\alpha\beta}$,	
Step 6:	discuss $u^\gamma_{(3)}u^\delta_{(2)}$,	(degree 2)
Step 7:	restrictions for $H_{\alpha\beta}$,$H^i_{\alpha\beta}$,$H^{ij}_{\alpha\beta}$,	
Step 8:	discuss $u^\gamma_{(3)}$,	(degree 1)
Step 9:	restrictions for $H_{\alpha\beta}$,$H^i_{\alpha\beta}$,$H^{ij}_{\alpha\beta}$,	
Step 10:	discuss $u^\gamma_{(2)}u^\delta_{(2)}$,	(degree 2)
Step 11:	restrictions for $H_{\alpha\beta}$,$H^i_{\alpha\beta}$,$H^{ij}_{\alpha\beta}$,	
Step 12:	discuss $u^\gamma_{(2)}$,	(degree 1)

which is the algorithm in Section 3.5. Theoretically, this algorithm works pretty good, also for arbitrary n,m. Practically, using this algorithm, the previous proof cannot be generalized to arbitrary n,m, since the expressions are getting too complicated and we cannot find a simple structure in the expressions. For $n = 1$ in Section 3.4, we had only to discuss polynomial degree one. Here, in Section 3.5, we have also to discuss polynomial degree two. For arbitrary n, we would have to discuss polynomial degree n.

Now we want to explain why we chose this algorithm, especially why we compute the restrictions for the Helmholtz expressions between these steps. For example, it is not a good idea to use the following algorithm

Step 1:	transform the ECS,
Step 2:	discuss $O(4)$,
Step 3:	discuss $O(3)$,
Step 4:	discuss $O(2)$,
Step 5:	consider all restrictions from the $O(4)$,$O(3)$,$O(2)$-discussion together,

which would also be possible. In the following, we give an explanation why this is not meaningful. After discussing $u^\gamma_{(4)}$ in Step 2, i.e. the equation

$$c_{ij}c_{kl}u^\gamma_{ijkl}f_{\beta,u^\alpha_{ij}u^\gamma_{kl}} = 0,$$

see (3.34), we could immediately discuss the $u^{\gamma}_{(3)} u^{\delta}_{(3)}$-terms in $H_{\alpha\beta}$ and in equation (3.32), without computing the restrictions for the Helmholtz expressions. That is, we could solve the equation

$$u^{\delta}_{ikl} u^{\gamma}_{jpq} c_{ij} c_{kl} c_{pq} f_{\beta, u^{\alpha}_{ij} u^{\delta}_{kl} u^{\gamma}_{pq}} = 0. \tag{3.56}$$

We observe that it is difficult to handle the conditions (3.56), but it is possible. After a longer calculation we get the following equations (let us write f instead of f_{β})

$$\begin{aligned}
&6x{,}0y \quad u^{\beta}_{xxx} u^{\gamma}_{xxx} : \quad && f_{u^{\alpha}_{xx} u^{\beta}_{xx} u^{\gamma}_{xx}} = 0, && (3.57)\\
&5x{,}1y \quad u^{\beta}_{xxx} u^{\gamma}_{xxy} : \quad && 2 f_{u^{\alpha}_{xx} u^{\beta}_{xx} u^{\gamma}_{xy}} + f_{u^{\alpha}_{xy} u^{\beta}_{xx} u^{\gamma}_{xx}} = 0, && (3.58)\\
&4x{,}2y \quad u^{\beta}_{xxx} u^{\gamma}_{xyy} : \quad && \ldots && (3.59)\\
&4x{,}2y \quad u^{\beta}_{xxy} u^{\gamma}_{xxy} : \quad && \ldots && (3.60)\\
&3x{,}3y \quad u^{\beta}_{xxx} u^{\gamma}_{yyy} : \quad && f_{u^{\alpha}_{xy} u^{\beta}_{xx} u^{\gamma}_{yy}} = 0 && (3.61)\\
&3x{,}3y \quad u^{\beta}_{xxy} u^{\gamma}_{xyy} : \quad && \ldots && (3.62)\\
&2x{,}4y \quad u^{\beta}_{xxy} u^{\gamma}_{yyy} : \quad && \ldots && (3.63)\\
&2x{,}4y \quad u^{\beta}_{xyy} u^{\gamma}_{xyy} : \quad && \ldots && (3.64)\\
&1x{,}5y \quad u^{\beta}_{xyy} u^{\gamma}_{yyy} : \quad && \ldots && (3.65)\\
&0x{,}6y \quad u^{\beta}_{yyy} u^{\gamma}_{yyy} : \quad && \ldots && (3.66)
\end{aligned}$$

We observe that a lot of these conditions are redundant. For example, the conditions (3.57) and (3.58) follow from (3.35). Not to have these redundant conditions, we constructed the following algorithm

- Discuss order,
- Restrictions for the Helmholtz expressions,
- Discuss order,
- Restrictions for the Helmholtz expressions,
- and so on...

In this algorithm we do not have (many) redundant conditions, since in every step we already restricted to such f which satisfy all the conditions from the previous steps. For example, from the conditions (3.57)-(3.66) we only need one condition, e.g. (3.61). Since (3.61) leads to $D_{\beta|\gamma\delta\epsilon} = 0$ and then $(u^{\gamma}, u^{\delta}, u^{\epsilon}) D_{\beta|\gamma\delta\epsilon} = 0$. That is, when this term vanishes, then we do not have to investigate all the remaining $u^{\gamma}_{(3)} u^{\delta}_{(3)}$-terms in $H_{\alpha\beta}$, since only $(u^{\gamma}, u^{\delta}, u^{\epsilon}) D_{\beta|\gamma\delta\epsilon}$ in f_{β} can generate $u^{\gamma}_{(3)} u^{\delta}_{(3)}$-terms in the Helmholtz expressions. Therefore, only with the condition (3.61) we can derive that

$$f_\alpha = A_\alpha + u^\gamma_{ij} B^{ij}_{\alpha|\gamma} + (u^\gamma_i, u^\delta_j) C^{ij}_{\alpha|\gamma\delta},$$

i.e. we get the Hyperjacobian structure in the second order coordinates of f_α and all other conditions (3.57)-(3.60) and (3.62)-(3.66) are redundant, when also using the conditions (3.35)-(3.39). Now, the idea is that these redundancies do not occur when we use the algorithm from the beginning of Subsection 3.5.1. However, it is hard to say if there never occur redundancies in this algorithm. But in any case there are less of them.

Let us finish this section with three remarks. First, the above algorithm is clever in some sense, but it is not clever enough to handle the problem for arbitrary n. Second, the systematic discussion of degree and order of jet coordinates makes sense in principle and gives an interesting view of a very rich mathematical structure. For example, the vanishing of fourth order does not force Hyperjacobian structure in f_α, since we also have terms of the form $(u^\alpha, u^\beta, u^\gamma)$ which are not Hyperjacobians. But we can also find a structure in these expressions and it would be nice to find out more about this structure in general. It seems that a lot of these expressions can be written as different kinds of determinants. It also seems that there is a sub-structure of the Hyperjacobians and not all Hyperjacobians are variational expressions. A possible question could be: Can we find an invariant definition of all of the expression which occur in every step? That is,

$$\underbrace{\{\text{all } f_\alpha\}}_{\text{Step 1}} \supset \underbrace{\{(u^\alpha, u^\beta, u^\gamma), \text{Hyperjacobians}\}}_{\text{Step 2}} \supset \ldots \supset \underbrace{\{\text{Hyperjacobians}\}}_{\text{Step 4}} \supset \ldots$$
$$\ldots \supset \underbrace{\{\text{variational expressions}\}}_{\text{Step 12}} \supset \ldots$$

The third remark is the following: Although the algorithm seems to be very systematically, it is actually not completely systematically. One reason is that computing the (restrictions for the) Helmholtz expressions is non-trivial and there are different ways how to write down these expressions. Let us explain this in the case $n = 1$ and arbitrary m. In Step 5 in the proof for $n = 1$ and arbitrary m in Section 3.4, we combined the conditions

$$0 = B_{\alpha\beta} - B_{\beta\alpha},$$
$$0 = B_{\beta\gamma, u^\alpha_x} - B_{\beta\alpha, u^\gamma_x}$$

and deduced that this forces

$$0 = B_{\alpha\gamma, u^\beta_x} + B_{\beta\gamma, u^\alpha_x} - 2B_{\beta\alpha, u^\gamma_x}.$$

Therefore, the last equation actually follows by Helmholtz dependency, i.e. relations among $H_{\alpha\beta}, H^x_{\alpha\beta}, H^{xx}_{\alpha\beta}$. A systematic discussion of polynomial degree and order of jet coordinates alone does not solve the problem in general and the algorithm does not tell us where and how we should use the Helmholtz dependencies to compute the restricted expression $H_{\alpha\beta}$,$H^{x}_{\alpha\beta}$,$H^{xx}_{\alpha\beta}$.

3.6. The Proof for $n \geq 2$, Arbitrary m and 2nd Order Source Forms

The proof in this section is probably shorter as expected, when we recall the algorithm and expressions of the proof in Section 3.5. The difficulty is rather to find the following steps of the proof. The proof is much shorter as expected, since we do not need to use the full properties of Hyperjacobians, other available information, and we can refer to the work of I.M. Anderson, J. Pohjanpelto and others, where some of the problems are already solved. In some sense we could even say that not using all the available information actually allows us to formulate a proof for arbitrary n,m. The previous algorithm in Subsection 3.5.1 seems not to be generalizable, since the expressions in the steps there are getting too complicated. As far as we know, and as it is written in (BCO81, p.156), it can be very complicated to determine the dimension of the set of Hyperjacobians in general. For $n = 2$, we presented relations in the previous section, see (3.53), (3.54) and (3.55). When we would try to prove Theorem 1.0.2 for arbitrary n,m with the algorithm from Subsection 3.5.1, then we would definitely have to know how many (linearly) independent Hyperjacobians there are and a method how to order them. Even when we would know all that, the proof would still be very long and complicated. There occur a lot of tensors, indices, symmetry/skew-symmetry conditions and it is hard to combine them systematically, if not even impossible. In the following proof, we also have to combine such symmetry/skew-symmetry conditions, but these combinations are relatively simple.

Now we will give an overview of the different steps in the proof, which is also a part of the proof. Details are then proven separately, if they need more space. Counter examples at the end of the proofs also show that we are indeed proving non-trivial statements and that they are no longer true under slightly weaker assumptions.

Proof of Theorem (1.0.2) for arbitrary n,m:

Step 1 (transform the ECS): Using the Local Simplification Lemma 3.4.1, we get the transformed ECS[3]

$$\begin{aligned} i)\quad & 0 = H_{\alpha\gamma} + O^{\beta}_{\gamma,k}(1)H^{k}_{\alpha\beta} + O^{\beta}_{\gamma,kl}(2)H^{kl}_{\alpha\beta}, \\ ii)\quad & 0 = (O^{\beta}_{jk}(1) - u^{\beta}_{jk})H^{k}_{\alpha\beta} + (O^{\beta}_{jkl}(2) - u^{\beta}_{jkl})H^{kl}_{\alpha\beta}, \end{aligned}$$

[3]This step was developed during my visit at Utah State University when working together with Ian M. Anderson.

where

$$\begin{aligned}
H_{\alpha\beta} &= f_{\alpha,u^\beta} - f_{\beta,u^\alpha} + D_k f_{\beta,u^\alpha_k} - c_{kl} D_{kl} f_{\beta,u^\alpha_{kl}},\\
H^i_{\alpha\beta} &= f_{\alpha,u^\beta_i} + f_{\beta,u^\alpha_i} - 2c_{ik} D_k f_{\beta,u^\alpha_{ik}},\\
H^{ij}_{\alpha\beta} &= c_{ij}(f_{\alpha,u^\beta_{ij}} - f_{\beta,u^\alpha_{ij}})
\end{aligned}$$

are the Helmholtz expressions ($c_{ij} = 1$ if $i = j$ and $c_{ij} = \frac{1}{2}$ if $i \neq j$). Note that i) and ii) are together $m^2 + nm$ equations and we have $m^2 + nm^2 + \frac{n(n+1)}{2}\frac{m(m-1)}{2}$ unknowns $H^{ij}_{\alpha\beta}$,$H^i_{\alpha\beta}$,$H_{\alpha\beta}$. That is, for large n,m the system i) and ii) is a highly under-determined system and only for $m = 1$ we can immediately solve it.

Step 2 (polynomial structure, we discuss $u^\alpha_{(4)}$- and $u^\alpha_{(3)}u^\beta_{(3)}$-terms, which are generated by the D_{ij}-derivatives): In this step, we solve the equations

$$u^\gamma_{ijkl} c_{kl} c_{ij} f_{\alpha,u^\beta_{ij} u^\gamma_{kl}} = 0 \quad \Leftrightarrow \quad \partial^{(kl}_\gamma \partial^{ij)}_\beta f_\alpha = 0 \tag{3.67}$$

and

$$u^\gamma_{kli} u^\delta_{jpq} c_{kl} c_{pq} c_{ij} f_{\alpha,u^\beta_{ij} u^\gamma_{kl} u^\delta_{pq}} = 0 \quad \Leftrightarrow \quad \partial^{(kl}_\gamma \partial^{i)(j}_\beta \partial^{pq)}_\gamma f_\alpha = 0 \tag{3.68}$$

and show the equivalence in (3.67) and (3.68). These equations result from the leading terms in $H_{\alpha\beta}$ in equation i). Note that we could also discuss the leading orders in $\mathcal{E}_\alpha(Q^\beta f_\beta) = 0$, which leads to the same result. Equations (3.67) and (3.68) force that f_α must be a polynomial of degree $\leq n$ in the second order jet coordinates.

We could also show the more precise result that f_α must be a sum of Hyperjacobians in the second order jet coordinates. However, we will actually only need the weaker statement that f_α must be a polynomial of degree $\leq n$ in the second order jet coordinates. For this statement we need results from a paper of I.M. Anderson and T. Duchamp (AD80, pp.786), which are pretty complicated to prove in general. However, for second order source forms it is not that complicated and we will prove it in Subsection 3.7.3.

Step 3 (we consider equation ii) and discuss $u^{\alpha_1}_{(2)}...u^{\alpha_n}_{(2)}$ and $u^\alpha_{(3)}$-terms): With the help of Step 2 we show that

$$\boxed{\begin{aligned} H^i_{\alpha\beta} &= O(2),\\ H^{ij}_{\alpha\beta} &= 0. \end{aligned}} \tag{3.69}$$

It needs a lot of work to deduce these equations and therefore we will prove it separately in Subsection 3.6.1.

This step can probably also be proven with the help of the so called d-fold operator used in the proof of Theorem 1.1 in (AP96, p.379) or in the proof of Theorem

1 in (MPV08, p.12).

Note that after deriving the conditions in (3.69) we have still $m^2 + nm^2 = m(m+nm)$ unknowns, but $i)$ and $ii)$ are only $m(n+m)$ equations, and therefore the problem is still non-trivial.

Step 4 (we consider equation $i)$ and discuss $u^\alpha_{(3)}$-terms): This step is quite simple and we can prove it here directly. From Step 3 we know that

$$i) \quad H_{\alpha\gamma} + \underbrace{O^\beta_{\gamma,k}(1)H^k_{\alpha\beta}}_{=O(2)} + \underbrace{O^\beta_{\gamma,kl}(2)H^{kl}_{\alpha\beta}}_{=0} = 0,$$

and therefore

$$\boxed{H_{\alpha\beta} = O(2).}$$

Step 5 (we consider the Helmholtz dependencies): This step is again very simple. We consider the Helmholtz dependency

$$\underbrace{H_{\alpha\beta} + H_{\beta\alpha}}_{=O(2)} = D_k H^k_{\alpha\beta} - D_{kl}\underbrace{H^{kl}_{\alpha\beta}}_{=0},$$

and therefore

$$c_{ij} u^\gamma_{kij} H^k_{\alpha\beta, u^\gamma_{ij}} = 0. \tag{3.70}$$

From (3.70) we get that $\partial_{u^\gamma_{(ij}} H^{k)}_{\alpha\beta} = 0$, where (ijk) means symmetrization in ijk. When we set $i = j = k$, then symmetrization leads to a single term and we get

$$\boxed{H^i_{\alpha\beta, u^\gamma_{ii}} = 0.} \tag{3.71}$$

Note that there is no summation over i in (3.71). As we mentioned at the beginning of this section, sometimes we do not need the full available information, i.e. in this case we do not need the full information of $\partial_{u^\gamma_{(ij}} H^{k)}_{\alpha\beta} = 0$.

Note that the equation $D_k H^k_{\alpha\beta} = O(2)$ is almost a trivial divergence $D_k H^k_{\alpha\beta} = 0$ and such equations are discussed in Subsection 3.7.1. The more information we can get is that $H^i_{\alpha\beta}$ must be a polynomial degree $\leq n-1$ in second order jet coordinates (see Subsection 3.7.1). However, it is reasonable only to use the minimal amount of needed information here. Especially when we try to find generalizations of this proof with weaker assumptions, where we do not have the stronger condition $D_k H^k_{\alpha\beta} = O(2)$, but the condition (3.71).

Step 6 (we consider equation *ii*) and discuss second order): We consider equation *ii*), condition (3.71) and show that

$$\boxed{H^i_{\alpha\beta} = 0.}$$

This step is quite difficult and we prove it separately in Subsection 3.6.2. Note that this step can probably also be proven with a different method, when using the so-called d-fold operator in (AP96, p.379) or (MPV08, p.12).

Step 7 (we consider equation *i*) and discuss second order): Then from equation *i*) we can easily conclude that $H_{\alpha\beta} = 0$, and therefore all Helmholtz conditions are satisfied. □

It remains to prove Step 3 and Step 6 in the following subsections. We also prove Step 2 in Subsection 3.7.3, although we could refer to the work of I.M. Anderson and T. Duchamp (AD80).

Let us reformulate the above algorithm as the following one, where the steps differ from the steps above, but the similarities to the algorithms in Subsection 3.4.1 and Subsection 3.5.1 can be seen more clearly:

$$\begin{aligned}
\text{Step 1}: &\quad \text{transform the ECS},\\
\text{Step 2}: &\quad \text{discuss } u^\alpha_{(4)} \text{ and } u^\alpha_{(3)}u^\beta_{(3)} \text{ simulatanously},\\
\text{Step 3}: &\quad \text{restrictions for } H_{\alpha\beta}, H^i_{\alpha\beta}, H^{ij}_{\alpha\beta},\ f_\alpha \text{ is a polynomial of degree} \leq n,\\
\text{Step 4}: &\quad \text{consider equation } ii) \text{ and discuss } u^\alpha_{(3)},\\
\text{Step 5}: &\quad \text{restrictions for } H_{\alpha\beta}, H^i_{\alpha\beta}, H^{ij}_{\alpha\beta},\\
\text{Step 6}: &\quad \text{consider equation } i) \text{ and discuss } u^\alpha_{(3)},\\
\text{Step 7 a)}: &\quad \text{restrictions for } H_{\alpha\beta}, H^i_{\alpha\beta}, H^{ij}_{\alpha\beta},\\
\text{Step 7 b)}: &\quad \text{consider the Helmholtz dependencies and}\\
&\quad \text{derive further restrictions for } H_{\alpha\beta}, H^i_{\alpha\beta}, H^{ij}_{\alpha\beta},\\
\text{Step 8}: &\quad \text{consider equation } ii) \text{ and discuss } O(2),\\
\text{Step 9}: &\quad \text{restrictions for } H_{\alpha\beta}, H^i_{\alpha\beta}, H^{ij}_{\alpha\beta},\\
\text{Step 10}: &\quad \text{consider equation } i) \text{ and discuss } O(2),\\
\text{Step 11}: &\quad \text{restrictions for } H_{\alpha\beta}, H^i_{\alpha\beta}, H^{ij}_{\alpha\beta}, \text{ all Helmholtz conditions satisfied.}
\end{aligned}$$

This is just one possibility how to formulate an algorithm similar to the algorithms in Subsection 3.4.1 and Subsection 3.5.1. Different formulations can be meaningful, especially discussing certain orders of jet coordinates and computing the restricted Helmholtz expressions can be formulated as one step. Polynomial degree of second order jet coordinates seems not to be necessarily split into different steps in this algorithm and this is why it works fairly well for arbitrary n (compare the algorithm in Subsection 3.5.1).

3.6.1. Proof of Step 3

Let us briefly explain the notation in this subsection and then we start with the proof. From Step 2 in Section 3.6 we know that f_α must be a polynomial of degree $\leq n$ in second order jet coordinates (or more precisely, a sum of Hyperjacobians). We will write

$$f_\alpha = A_0 + A_1 u_{(2)} + ... + A_n \underbrace{u_{(2)}...u_{(2)}}_{n \text{ times}},$$

where $u_{(2)}$ is a short notation for second order jet coordinates u^γ_{ij}, and $A_k = O(1)$ for all $k = 0,1,...,n$. That is, we use a symbolic notation, where we suppress some of the indices. The exact notation would be

$$f_\alpha = A_\alpha + A^{ij}_{\alpha|\gamma} u^\gamma_{ij} + ... + A^{i_1 j_1 ... i_n j_n}_{\alpha|\gamma_1...\gamma_n} u^{\gamma_1}_{i_1 j_1} ... u^{\gamma_n}_{i_n j_n},$$

and the coefficients $A^{...}_{...} = O(1)$ satisfy some symmetry and skew-symmetry conditions, which are relatively complicated and not important here, just Hyperjacobians. In the following, we only need that f_α is a polynomial of degree $\leq n$ in the second order jet coordinates. Then for $H^{ij}_{\alpha\beta}$ we get

$$\begin{aligned} H^{ij}_{\alpha\beta} &= c_{ij}(f_{\alpha,u^\beta_{ij}} - f_{\beta,u^\alpha_{ij}}) = \\ &\overset{\text{symbolic}}{=} c_{ij}(\partial_{u^\beta_{ij}} - \partial_{u^\alpha_{ij}})[A_0 + A_1 u_{(2)} + ... + A_n \underbrace{u_{(2)}...u_{(2)}}_{n \text{ times}}] = \\ &= \tilde{A}_1 + \tilde{A}_2 u_{(2)} + ... + \tilde{A}_n \underbrace{u_{(2)}...u_{(2)}}_{(n-1) \text{ times}}, \end{aligned}$$

i.e. $H^{ij}_{\alpha\beta}$ is of degree $\leq n-1$ in the second order jet coordinates, where $\tilde{A}_k = O(1)$. We wrote symbolic, since f_α and f_β have different coefficients A_k. In exact notation we get

$$\boxed{H^{ij}_{\alpha\beta,u^{\gamma_1}_{i_1 j_1}...u^{\gamma_n}_{i_n j_n}} = 0.} \tag{3.72}$$

For $H^i_{\alpha\beta}$ we can do a similar (symbolic) calculation and we get

$$\begin{aligned} H^i_{\alpha\beta} &= f_{\alpha,u^\beta_i} + f_{\beta,u^\alpha_i} - 2c_{ik}D_k f_{\beta,u^\alpha_{ik}} = \\ &= O_n(2) - 2c_{ik}D_k\partial_{u^\alpha_{ik}}[A_0 + A_1u_{(2)} + ... + A_n \underbrace{u_{(2)}...u_{(2)}}_{n \text{ times}}] = \\ &= O_n(2) - 2c_{ik}D_k[A_1 + A_2u_{(2)} + ... + A_n \underbrace{u_{(2)}...u_{(2)}}_{(n-1) \text{ times}}] = \\ &= O_n(2) - 2c_{ik}[\tilde{A}_2u_{(2)} + ... + \tilde{A}_n \underbrace{u_{(2)}...u_{(2)}}_{(n-2) \text{ times}} u_{(3)}] \end{aligned} \tag{3.73}$$

and $u_{(3)}$ is a short notation for third order jet coordinates u^γ_{ijk}. Note that for $n = 1$, $H^i_{\alpha\beta}$ does not depend on third order, and for $n = 2$, we only have combinations of the form $O(1)u^\gamma_{jtk}$ as third order terms. In general, the third order terms are quite specific and we will use this property below. In exact notation we get

$$\boxed{H^i_{\alpha\beta,u^{\gamma_1}_{i_1j_1}...u^{\gamma_{n-1}}_{i_{n-1}j_{n-1}}} = O_{lin}(2) = O(2).} \tag{3.74}$$

Note that only third order terms vanish here when applying the partial differential operator $\partial_{u^{\gamma_1}_{i_1j_1}}...\partial_{u^{\gamma_{n-1}}_{i_{n-1}j_{n-1}}}$ on $H^i_{\alpha\beta}$, since they are also of degree $\leq n-2$ in second order coordinates and, in general, second order coordinates in $H^i_{\alpha\beta}$ can occur up to degree $\leq n$.

Then we consider equation *ii*) and we apply the following differential operators

$$\begin{aligned} &a) \quad (\partial_{u^\gamma_{JJ}})^{n-1} = \underbrace{\partial_{u^{\gamma_1}_{JJ}}...\partial_{u^{\gamma_{n-1}}_{JJ}}}_{(n-1) \text{ times}} && (\text{to determine } H^{kl}_{\alpha\beta}), \\ &b) \quad \partial_{u^\delta_{jJ}}(\partial_{u^\gamma_{JJ}})^{n-2} = \partial_{u^\delta_{jJ}} \underbrace{\partial_{u^{\gamma_1}_{JJ}}...\partial_{u^{\gamma_{n-2}}_{JJ}}}_{(n-2) \text{ times}} && (\text{to determine } H^k_{\alpha\beta}), \\ &c) \quad \partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-2} = \partial_{u^\delta_{jj}} \underbrace{\partial_{u^{\gamma_1}_{JJ}}...\partial_{u^{\gamma_{n-2}}_{JJ}}}_{(n-2) \text{ times}} && (\text{to determine } H^k_{\alpha\beta}), \end{aligned}$$

where all the J can take all possible values and combinations such that $J \in \{1,2,...,n\}$ and $J \neq j$. It is important to understand that the expressions in *a*), *b*), *c*) can take all possible combinations of such J, i.e. J is not one fixed number in these expressions it is a combination of numbers. For example, let $j = 1$ and $n = 4$, then we get

$$(\partial_{u^\gamma_{JJ}})^2 = \partial_{u^{\gamma_1}_{22}}\partial_{u^{\gamma_1}_{22}} \quad \text{and} \quad (\partial_{u^\gamma_{JJ}})^2 = \partial_{u^{\gamma_1}_{23}}\partial_{u^{\gamma_2}_{24}} \tag{3.75}$$

and even more expressions. That is, we actually apply a set of operators

$$(\partial_{u^\gamma_{JJ}})^2 = \{\partial_{u^{\gamma_1}_{22}}\partial_{u^{\gamma_1}_{22}}, \partial_{u^{\gamma_1}_{23}}\partial_{u^{\gamma_2}_{24}}, ...\},$$

which will lead to a set of equations below. We also use the notation $(\gamma)^n = \gamma_1 ... \gamma_n$, i.e. multi index notation in $(\gamma)^n$ is assumed. Without using this simplified notation for J and $(\gamma)^n$ it would be very hard to write down all the expressions in a), b), c), or in (3.75). For example, for a) we would have to write

$$(\partial_{u^{\gamma}_{JJ}})^{n-1} = \left\{ \partial_{u^{\gamma_1}_{J_1 J_2}} \partial_{u^{\gamma_2}_{J_3 J_4}} ... \partial_{u^{\gamma_{n-1}}_{J_{2n-3} J_{2n-2}}}, \quad \text{for all } J_l \in \{1,2,...,n\} \text{ and } J_l \neq j \right\},$$

where $1 \leq l \leq 2n-2$, which is an equivalent notation to what we have written above and a similar notation holds for b) and c). Note that below we will also have further restrictions than $J \in \{1,2,...,n\}$, $J \neq j$ and then we simply write $J \in \{2,3,...,n\}$, $J \in \{3,4,...,n\}$ and so on. We do not always write $J \neq j$, since this is always be assumed.

Here the actual proof of Step 3 in Section 3.6 starts. The proof is based on a sort of induction. We will write Step 3.k for the different steps in the induction.

Proof of Step 3 in Section 3.6:

Step 3.0: We briefly repeat what we know from the discussion above. We know that

$$\boxed{\begin{aligned} 0 &= H^{kl}_{\alpha\beta, u^{\gamma_1}_{i_1 j_1} ... u^{\gamma_n}_{i_n j_n}}, \qquad (3.76) \\ O(2) &= H^{k}_{\alpha\beta, u^{\gamma_1}_{i_1, j_1} ... u^{\gamma_{n-1}}_{i_{n-1} j_{n-1}}}. \qquad (3.77) \end{aligned}}$$

Actually, we only need the weaker conditions written in the following boxes

$$a) \quad \text{apply } (\partial_{u^{\gamma}_{JJ}})^{n-0} = (\partial_{u^{\gamma}_{JJ}})^{n} \quad \Rightarrow \quad \boxed{(\partial_{u^{\gamma}_{JJ}})^{n} H^{kl}_{\alpha\beta} = 0, \quad J \in \{1,2...n\}}, \qquad (3.78)$$

$$b) \quad \text{apply } \partial_{u^{\delta}_{jJ}} (\partial_{u^{\gamma}_{JJ}})^{n-1} \quad \Rightarrow \quad \boxed{(\partial_{u^{\gamma}_{JJ}})^{n-1} H^{J}_{\alpha\beta} = O(2), \; J \in \{2,3...n\}}, \qquad (3.79)$$

$$c) \quad \text{apply } \partial_{u^{\delta}_{jj}} (\partial_{u^{\gamma}_{JJ}})^{n-1} \quad \Rightarrow \quad \boxed{(\partial_{u^{\gamma}_{JJ}})^{n-1} H^{j}_{\alpha\beta} = O(2), \; J \in \{2,3...n\}}. \qquad (3.80)$$

Without loss of generality, we always assume $j = 1$ and then J are all the other possibilities, or, when explicitly writing $J \in \{1,2,...,n\}$, then it is also clear which values J can take, for example, in (3.78). The zero in $n-0$ should coincide with Step 3.0 in the induction and this will be clear soon after we showed how the indiction works. We will also explain where and why we only need the weaker conditions in (3.78)-(3.80) and why we sometimes write *only if* or *also for*.

Step 3.1: Applying the operator $(\partial_{u^\gamma_{JJ}})^{n-1} = \partial_{u^\delta_{JJ}}(\partial_{u^\gamma_{JJ}})^{n-2}$ to equation $ii)$, we get

$$
\begin{aligned}
a) \quad 0 &= (\partial_{u^\gamma_{JJ}})^{n-1}[(O(1) - u^\beta_{jk})H^k_{\alpha\beta} + (O(2) - u^\beta_{jkl})H^{kl}_{\alpha\beta}] = \\
&= (O(1) - u^\beta_{jk}) \underbrace{H^k_{\alpha\beta,(u^\gamma_{JJ})^{n-1}}}_{\substack{=O(2) \text{ also for} \\ J\in\{2,3...n\} \\ \text{Step 3.0 (3.77)}}} + \underbrace{(\partial_{u^\gamma_{JJ}})^{n-1}[O(2)H^{kl}_{\alpha\beta}]}_{=O(2) \text{ always}} - u^\beta_{jkl} \underbrace{H^{kl}_{\alpha\beta,(u^\gamma_{JJ})^{n-1}}}_{\substack{\text{here we want to} \\ \text{get information}}}, \qquad (3.81)
\end{aligned}
$$

and since $H^{kl}_{\alpha\beta}$ is symmetric in k,l, we get

$$
\boxed{\text{Step 3.1 } a) \quad 0 = H^{kl}_{\alpha\beta,(u^\gamma_{JJ})^{n-1}}, \quad \begin{array}{l}\text{where } J \text{ is in a set of } n-1 \text{ numbers,} \\ \text{e.g. } j=1 \text{ then } J\in\{2,3,...,n\}.\end{array}} \qquad (3.82)
$$

We restricted to $J \in \{2,3,...,n\}$, because otherwise we cannot commute $(\partial_{u^\gamma_{JJ}})^{n-1}$ with u^β_{jk}. This restriction is crucial in the whole proof.[4] For example, for $n = 2$ we get $H^{kl}_{\alpha\beta,u^\gamma_{xx}} = 0$ and $H^{kl}_{\alpha\beta,u^\gamma_{yy}} = 0$, but we do not get $H^{kl}_{\alpha\beta,u^\gamma_{xy}} = 0$, since we only choose from a set of $n-1 = 1$ numbers.[5] Now we apply the operator $\partial_{u^\delta_{jJ}}(\partial_{u^\gamma_{JJ}})^{n-2}$, $J \in \{3,4,...,n\}$ to equation $ii)$ and we get

$$
\begin{aligned}
b) \quad 0 =& \partial_{u^\delta_{jJ}}(\partial_{u^\gamma_{JJ}})^{n-2}[(O(1) - u^\beta_{jk})H^k_{\alpha\beta} + (O(2) - u^\beta_{jkl})H^{kl}_{\alpha\beta}] = \\
=& O(1) \underbrace{\partial_{u^\delta_{jJ}}(\partial_{u^\gamma_{JJ}})^{n-2}H^k_{\alpha\beta}}_{\substack{=O(2) \text{ also for } J\in\{3,4...n\} \\ \text{Step 3.0 (3.77)}}} - \partial_{u^\delta_{jJ}}(\partial_{u^\gamma_{JJ}})^{n-2}[u^\beta_{jk}H^k_{\alpha\beta}] + \\
&+ \underbrace{\partial_{u^\delta_{jJ}}(\partial_{u^\gamma_{JJ}})^{n-2}[O(2)H^{kl}_{\alpha\beta}]}_{=O(2) \text{ always}} - u^\beta_{jkl} \underbrace{\partial_{u^\delta_{jJ}}(\partial_{u^\gamma_{JJ}})^{n-2}H^{kl}_{\alpha\beta}}_{\substack{=0 \text{ only if } J=\{3,4...n\} \\ \text{Step 3.1 a) (3.82)}}} = \\
=& O(2) - \partial_{u^\delta_{jJ}}(\partial_{u^\gamma_{JJ}})^{n-2}[u^\beta_{jk}H^k_{\alpha\beta}] = \\
=& O(2) - \partial_{u^\delta_{jJ}}[u^\beta_{jk}(\partial_{u^\gamma_{JJ}})^{n-2}H^k_{\alpha\beta}] = \\
=& O(2) - (\partial_{u^\delta_{JJ}})^{n-2}H^J_{\alpha\delta} - u^\beta_{jk} \underbrace{\partial_{u^\delta_{jJ}}(\partial_{u^\gamma_{JJ}})^{n-2}H^k_{\alpha\beta}}_{\substack{=O(2) \text{ also for } J\in\{3,4...n\} \\ \text{Step 3.0 (3.77)}}} = \\
=& O(2) - (\partial_{u^\gamma_{JJ}})^{n-2}H^J_{\alpha\delta}, \quad J \in \{3,4,...,n\}. \qquad (3.83)
\end{aligned}
$$

This computation needs some explanation. In the third line in (3.83) we had to choose J in a set of $n-2$ numbers (formally we wrote $J \in \{3,4,...,n\}$), such that the

[4] Step 3.0, i.e. (3.77) can also be applied when we would choose $J \in \{1,2,...,n\}$ and therefore we wrote *also for*, but then we would not be able to commute the differential operators with u^β_{jk}. Later, i.e. in Step 3.2, we cannot write *also for* any longer, and we will write *only if*.

[5] For $n = 3$, $j \hat{=} z$ and $J \hat{=} \{x,y\}$, we would get $H^{kl}_{\alpha\beta,u^\gamma_{xx}u^\delta_{xx}} = H^{kl}_{\alpha\beta,u^\gamma_{xx}u^\delta_{xy}} = H^{kl}_{\alpha\beta,u^\gamma_{xx}u^\delta_{yy}} = H^{kl}_{\alpha\beta,u^\gamma_{xy}u^\delta_{xy}} = 0$, but we would not get $H^{kl}_{\alpha\beta,u^\gamma_{xx}u^\delta_{yz}} = 0$.

set $\{j,3,4,...,n\}$ is a set of $n-1$ numbers and only then we can apply (3.82) from Step 3.1 a).[6] Therefore, we get

$$\boxed{\text{Step 3.1 b)}\quad O(2) = (\partial_{u^\gamma_{JJ}})^{n-2} H^J_{\alpha\delta}, \quad \begin{array}{l} J \text{ in a set of } n-2 \text{ numbers,} \\ \text{e.g. } J \in \{3,4,...,n\}. \end{array}} \tag{3.84}$$

Applying $\partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-2}$, $J \in \{3,4,...,n\}$, leads to

$$\begin{aligned}
c)\quad 0 =&\partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-2}[(O(1) - u^\beta_{jk})H^k_{\alpha\beta} + (O(2) - u^\beta_{jkl})H^{kl}_{\alpha\beta}] = \\
=&O(1) \underbrace{\partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-2} H^k_{\alpha\beta}}_{\substack{=O(2) \text{ also for } J \in \{3,4...n\} \\ \text{Step 3.0 (3.77)}}} - \partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-2}[u^\beta_{jk} H^k_{\alpha\beta}] + \\
&+ \underbrace{\partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-2}[O(2)H^{kl}_{\alpha\beta}]}_{=O(2) \text{ always}} - u^\beta_{jkl} \underbrace{\partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-2} H^{kl}_{\alpha\beta}}_{\substack{=0 \text{ only if } J=\{3,4...n\} \\ \text{Step 3.1 a) (3.82)}}} = \\
=&O(2) - \partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-2}[u^\beta_{jk} H^k_{\alpha\beta}] = \\
=&O(2) - \partial_{u^\delta_{jj}}[u^\beta_{jk}(\partial_{u^\gamma_{JJ}})^{n-2} H^k_{\alpha\beta}] = \\
=&O(2) - (\partial_{u^\delta_{JJ}})^{n-2} H^j_{\alpha\delta} - u^\beta_{jk} \underbrace{\partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-2} H^k_{\alpha\beta}}_{\substack{=O(2) \text{ also for } J \in \{3,4...n\} \\ \text{Step 3.0 (3.77)}}} = \\
=&O(2) - (\partial_{u^\gamma_{JJ}})^{n-2} H^j_{\alpha\delta}, \quad J \in \{3,4,...,n\}
\end{aligned}$$

and we conclude

$$\boxed{\text{Step 3.1 c)}\quad O(2) = (\partial_{u^\gamma_{JJ}})^{n-2} H^j_{\alpha\delta}, \quad \begin{array}{l} J \text{ in a set of } n-2 \text{ numbers,} \\ \text{e.g. } J = \{3,4,...,n\} \end{array}} \tag{3.85}$$

Together with a), b), c), i.e. with (3.82),(3.84) and (3.85), we get

$$0 = H^{kl}_{\alpha\beta,(u^\gamma_{JJ})^{n-1}}, \quad \text{where } J \text{ is in a set of } n-1 \text{ numbers,} \tag{3.86}$$

$$O(2) = H^k_{\alpha\delta,(u^\gamma_{JJ})^{n-2}}, \quad \text{where } J \text{ is in a set of } n-2 \text{ numbers.} \tag{3.87}$$

Note that there is always one derivative more on $H^{kl}_{\alpha\beta}$ in comparison with the derivatives on $H^k_{\alpha\beta}$ and the equations hold for all $k,l = 1,2...n$. We can also observe that we have to prove a) first, since it is needed to prove b) and c).

<u>Step 3.2:</u> Now we are almost ready to do this inductively, i.e. to repeat the same

[6] Step 3.0, that is (3.77), would also be applicable when we would choose from the set $J \in \{2,3,...,n\}$, and therefore we wrote *also for* in the second last line in (3.83).

argument. To get the induction working, we write down Step 3.2 once again, since there is a small difference in comparison with Step 3.1. In Step 3.2 a), we apply the operator $(\partial u^{\gamma}_{JJ})^{n-2}$, $J \in \{3,4...n\}$, and we get

$$
\begin{aligned}
a) \quad 0 &= (\partial_{u^{\gamma}_{JJ}})^{n-2}[(O(1) - u^{\beta}_{jk})H^{k}_{\alpha\beta} + (O(2) - u^{\beta}_{jkl})H^{kl}_{\alpha\beta}] = \\
&= (O(1) - u^{\beta}_{jk}) \underbrace{H^{k}_{\alpha\beta,(u^{\gamma}_{JJ})^{n-2}}}_{\substack{=O(2) \text{ only if} \\ J \in \{3,4...n\} \\ \text{Step 3.1 (3.87)}}} + \underbrace{(\partial_{u^{\gamma}_{JJ}})^{n-2}[O(2)H^{kl}_{\alpha\beta}]}_{=O(2) \text{ always}} - u^{\beta}_{jkl} \underbrace{H^{kl}_{\alpha\beta,(u^{\gamma}_{JJ})^{n-2}}}_{\substack{\text{here we want to} \\ \text{get information}}},
\end{aligned}
$$

that is, we get

$$
\boxed{\text{Step 3.2 } a) \quad 0 = H^{kl}_{\alpha\beta,(u^{\gamma}_{JJ})^{n-2}}, \quad \text{where } J \text{ is in a set of } n-2 \text{ numbers.}} \tag{3.88}
$$

Here we can observe that in Step 3.2 a), we wrote *only if* instead of *also for* and this is in contrast to Step 3.1 a). For commuting $(\partial_{u^{\gamma}_{JJ}})^{n-2}$ with u^{β}_{jk}, we could also have chosen $J \in \{2,3...n\}$, but then we would not be able to use (3.87) and therefore the set of possible J is again decreased by one. In Step 3.2 b) we apply $\partial u^{\delta}_{jJ}(\partial u^{\gamma}_{JJ})^{n-3}$, $J = \{4,5,...n\}$, which leads to

$$
\begin{aligned}
b) \quad 0 =& \partial_{u^{\delta}_{jJ}}(\partial_{u^{\gamma}_{JJ}})^{n-3}[(O(1) - u^{\beta}_{jk})H^{k}_{\alpha\beta} + (O(2) - u^{\beta}_{jkl})H^{kl}_{\alpha\beta}] = \\
=& O(1) \underbrace{\partial_{u^{\delta}_{jJ}}(\partial_{u^{\gamma}_{JJ}})^{n-3}H^{k}_{\alpha\beta}}_{\substack{=O(2) \text{ only if } J \in \{4,5,...n\} \\ \text{Step 3.1 (3.87)}}} - \partial_{u^{\delta}_{jJ}}(\partial_{u^{\gamma}_{JJ}})^{n-3}[u^{\beta}_{jk}H^{k}_{\alpha\beta}] + \\
&+ \underbrace{\partial_{u^{\delta}_{jJ}}(\partial_{u^{\gamma}_{JJ}})^{n-3}[O(2)H^{kl}_{\alpha\beta}]}_{=O(2) \text{ always}} - u^{\beta}_{jkl} \underbrace{\partial_{u^{\delta}_{jJ}}(\partial_{u^{\gamma}_{JJ}})^{n-3}H^{kl}_{\alpha\beta}}_{\substack{=0 \text{ only if } J = \{4,5,...n\} \\ \text{Step 3.2 a) (3.88)}}} = \\
=& O(2) - \partial_{u^{\delta}_{jJ}}(\partial_{u^{\gamma}_{JJ}})^{n-3}[u^{\beta}_{jk}H^{k}_{\alpha\beta}] = \\
=& O(2) - \partial_{u^{\delta}_{jJ}}[u^{\beta}_{jk}(\partial_{u^{\gamma}_{JJ}})^{n-3}H^{k}_{\alpha\beta}] = \\
=& O(2) - (\partial_{u^{\delta}_{JJ}})^{n-3}H^{J}_{\alpha\delta} - u^{\beta}_{jk} \underbrace{\partial_{u^{\delta}_{jJ}}(\partial_{u^{\gamma}_{JJ}})^{n-3}H^{k}_{\alpha\beta}}_{\substack{=O(2) \text{ only if } J \in \{4,5...n\} \\ \text{Step 3.1 (3.87)}}} = \\
=& O(2) - (\partial_{u^{\gamma}_{JJ}})^{n-3}H^{J}_{\alpha\delta}, \quad J \in \{4,5,...,n\},
\end{aligned}
$$

and therefore we get

$$
\boxed{\text{Step 3.2 } b) \quad O(2) = H^{J}_{\alpha\delta,(u^{\gamma}_{JJ})^{n-3}}, \quad \text{where } J \text{ in a set of } n-3 \text{ numbers.}}
$$

Note that if $J \in \{4,5...n\}$ is in a set of $n-3$ numbers, then $\{j,4,5...n\}$ is in a set of $n-2$ numbers and only in this case we can apply (3.87) and (3.88). In Step 3.2 c), we apply $\partial u^\delta_{jj}(\partial u^\gamma_{JJ})^{n-3}$, $J\in\{4,5...n\}$, and this leads to

$$
\begin{aligned}
c)\quad 0 =&\partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-3}[(O(1)-u^\beta_{jk})H^k_{\alpha\beta}+(O(2)-u^\beta_{jkl})H^{kl}_{\alpha\beta}] =\\
=&O(1)\underbrace{\partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-3}H^k_{\alpha\beta}}_{\substack{=O(2)\text{ only if } J\in\{4,5,...n\}\\ \text{Step 3.1 (3.87)}}} -\partial_{u^\delta_{jJ}}(\partial_{u^\gamma_{JJ}})^{n-3}[u^\beta_{jk}H^k_{\alpha\beta}]+\\
&+\underbrace{\partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-3}[O(2)H^{kl}_{\alpha\beta}]}_{=O(2)\text{ always}} - u^\beta_{jkl}\underbrace{\partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-3}H^{kl}_{\alpha\beta}}_{\substack{=0\text{ only if } J=\{4,5,...n\}\\ \text{Step 3.2 a) (3.88)}}} =\\
=&O(2)-\partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-3}[u^\beta_{jk}H^k_{\alpha\beta}] =\\
=&O(2)-\partial_{u^\delta_{jj}}[u^\beta_{jk}(\partial_{u^\gamma_{JJ}})^{n-3}H^k_{\alpha\beta}] =\\
=&O(2)-(\partial_{u^\delta_{JJ}})^{n-3}H^j_{\alpha\delta}-u^\beta_{jk}\underbrace{\partial_{u^\delta_{jJ}}(\partial_{u^\gamma_{JJ}})^{n-3}H^k_{\alpha\beta}}_{\substack{=O(2)\text{ only if } J\in\{4,5...n\}\\ \text{Step 3.1 (3.87)}}} =\\
=&O(2)-(\partial_{u^\gamma_{JJ}})^{n-3}H^j_{\alpha\delta},\quad J\in\{4,5,...,n\},
\end{aligned}
$$

and therefore we get

$$\boxed{\text{Step 3.2 } c)\quad O(2) =(\partial_{u^\gamma_{JJ}})^{n-3}H^j_{\alpha\delta},\quad \text{where } J \text{ in a set of } n-3 \text{ numbers.}}$$

Together with a), b), c) we get

$$\boxed{\begin{aligned} 0 =&H^{kl}_{\alpha\beta,(u^\gamma_{JJ})^{n-2}},\quad \text{where } J \text{ in a set of } n-2 \text{ numbers,}\\ O(2) =&H^k_{\alpha\delta,(u^\gamma_{JJ})^{n-3}},\quad \text{where } J \text{ in a set of } n-3 \text{ numbers.}\end{aligned}}$$

Now we are able to do this inductively, since we can repeat exactly the same argument from Step 3.l in Step 3.(l+1) (and we will always need a *only if*).

$$\vdots$$

Step 3.(n-2): In this step, we get

$$\boxed{\begin{aligned} 0 =&H^{kl}_{\alpha\beta,(u^\gamma_{JJ})^{n-(n-2)}} = H^{kl}_{\alpha\beta,u^{\gamma_1}_{JJ}u^{\gamma_2}_{JJ}},\quad J \text{ in a set of } n-(n-2)=2 \text{ numbers,}\\ O(2) =&H^k_{\alpha\delta,(u^\gamma_{JJ})^{n-(n-1)}} = H^k_{\alpha\delta,u^\gamma_{JJ}},\quad J \text{ in a set of } n-(n-1)=1 \text{ numbers.}\end{aligned}}$$

(3.89)

For example, for $n=3$, we get $H^k_{\alpha\delta,u^\gamma_{xx}}, H^k_{\alpha\delta,u^\gamma_{yy}}, H^k_{\alpha\delta,u^\gamma_{zz}}=0$. But we do not get one of the mixed derivatives $H^k_{\alpha\delta,u^\gamma_{xy}}=0$, $H^k_{\alpha\delta,u^\gamma_{xz}}=0$ or $H^k_{\alpha\delta,u^\gamma_{yz}}=0$. This will be crucial in the next step, since applying $\partial_{u^\delta_{jJ}}$ does not work any longer.

Step 3.(n-1): The second last step is different in comparison with all the previous steps, since applying the $\partial_{u^\gamma_{jJ}}$-operator does not work any longer, as we will see below. Formally, in this step we would get

$$\boxed{\begin{aligned} 0 =& H^{kl}_{\alpha\beta,(u^\gamma_{JJ})^{n-(n-1)}} = H^{kl}_{\alpha\beta,u^\gamma_{JJ}}, \quad J \text{ in a set of } n-(n-1)=1 \text{ numbers,} \\ O(2) =& H^{k}_{\alpha\delta,(u^\gamma_{JJ})^{n-n}} = H^k_{\alpha\delta} \quad J \text{ in a set of } n-n \text{ numbers, i.e. } J \in \emptyset \end{aligned}} \tag{3.90}$$

and since we cannot choose $J \in \emptyset$, there must be something wrong. However, part a) still works, where we get

$$\begin{aligned} a) \quad 0 =& (\partial_{u^\gamma_{JJ}})^{n-(n-1)}[(O(1)-u^\beta_{jk})H^k_{\alpha\beta} + (O(2)-u^\beta_{jkl})H^{kl}_{\alpha\beta}] = \\ =& (O(1)-u^\beta_{jk}) \underbrace{H^k_{\alpha\beta,u^\gamma_{JJ}}}_{\substack{=O(2) \text{ only if } J\in\{n\} \\ \text{Step 3.(n-2) (3.89)}}} + \underbrace{\partial_{u^\gamma_{JJ}}[O(2)H^{kl}_{\alpha\beta}]}_{=O(2) \text{ always}} - u^\beta_{jkl} \underbrace{H^{kl}_{\alpha\beta,u^\gamma_{JJ}}}_{\substack{\text{here we want to} \\ \text{get information}}}, \\ & J \in \{n\}, \end{aligned}$$

where J is in a set of one number. Otherwise (3.89) would not be applicable, i.e. both of the JJ must be the same. Therefore, we get

$$\boxed{\text{Step 3.(n-1) a)} \quad 0 = H^{kl}_{\alpha\beta,u^\gamma_{JJ}}, \quad \text{where } J \text{ is in a set of 1 number.}} \tag{3.91}$$

Therefore, and because of Step 3.(n-2) b), the Step 3.(n-1) does not work any longer, where we would formally get

$$\begin{aligned} b) \quad 0 =& \partial_{u^\delta_{jJ}}(\partial_{u^\gamma_{JJ}})^{n-1-(n-1)}[(O(1)-u^\beta_{jk})H^k_{\alpha\beta} + (O(2)-u^\beta_{jkl})H^{kl}_{\alpha\beta}] = \\ =& O(1) \underbrace{\partial_{u^\delta_{jJ}} H^k_{\alpha\beta}}_{\substack{=O(2) \text{ only if } J\in\emptyset \\ \text{Step 3.(n-2) (3.89)}}} - \partial_{u^\delta_{jJ}}[u^\beta_{jk}H^k_{\alpha\beta}] + \underbrace{\partial_{u^\delta_{jJ}}[O(2)H^{kl}_{\alpha\beta}]}_{=O(2) \text{ always}} - \\ & - u^\beta_{jkl} \underbrace{\partial_{u^\delta_{jJ}} H^{kl}_{\alpha\beta}}_{\substack{=0 \text{ only if } J\in\emptyset \\ \text{Step 3.(n-1) a) (3.91)}}} = \ldots \end{aligned}$$

Surprisingly, part c) still works and we get

$$
\begin{aligned}
c)\quad 0 =&\partial_{u^\delta_{jj}}(\partial_{u^\gamma_{JJ}})^{n-1-(n-1)}[(O(1)-u^\beta_{jk})H^k_{\alpha\beta}+(O(2)-u^\beta_{jkl})H^{kl}_{\alpha\beta}] = \\
=&O(1)\underbrace{\partial_{u^\delta_{jj}}H^k_{\alpha\beta}}_{\substack{=O(2)\text{ no }J\text{ necessary}\\ \text{Step 3.(n-2) (3.89)}}}-\partial_{u^\delta_{jj}}[u^\beta_{jk}H^k_{\alpha\beta}]+\underbrace{\partial_{u^\delta_{jj}}[O(2)H^{kl}_{\alpha\beta}]}_{=O(2)\text{ always}}- \\
&-u^\beta_{jkl}\underbrace{\partial_{u^\delta_{jj}}H^{kl}_{\alpha\beta}}_{\substack{=0\text{ no }J\text{ necessary}\\ \text{Step 3.(n-1) a) (3.91)}}} = \\
=&O(2)-\partial_{u^\delta_{jj}}[u^\beta_{jk}H^k_{\alpha\beta}] = \\
=&O(2)-H^j_{\alpha\delta}-u^\beta_{jk}\underbrace{\partial_{u^\delta_{jj}}H^k_{\alpha\beta}}_{\substack{=O(2)\text{ no }J\text{ necessary}\\ \text{Step 3.(n-2) (3.89)}}} = \\
=&O(2)-H^j_{\alpha\delta},\quad \text{no } J \text{ necessary.}
\end{aligned}
$$

Therefore, we get

$$\boxed{\text{Step 3.(n-1) c)}\quad O(2)=H^j_{\alpha\delta}\quad \text{for all } \alpha,\delta=1,2,...,m \text{ and for all } j=1,2,...,n.}$$

From a), b) and c) together, we get

$$\boxed{\begin{aligned} 0 =&H^{kl}_{\alpha\beta,(u^\gamma_{JJ})^{n-(n-1)}}=H^{kl}_{\alpha\beta,u^\gamma_{JJ}}, \quad J \text{ in a set of } n-(n-1)=1 \text{ number,}\\ O(2) =&H^k_{\alpha\delta,(u^\gamma_{JJ})^{n-n}}=H^k_{\alpha\delta}, \qquad \text{(formally for } J\in\emptyset).\end{aligned}} \tag{3.92}$$

Now we also want to show that $H^{kl}_{\alpha\beta}=0$.

Step 3.n: Formally, we would apply $(\partial_{u^\gamma_{JJ}})^{n-n}=1$ and deduce

$$
\begin{aligned}
a)\quad 0 =& (\partial_{u^\gamma_{JJ}})^{n-n}[(O(1)-u^\beta_{jk})H^k_{\alpha\beta}+(O(2)-u^\beta_{jkl})H^{kl}_{\alpha\beta}] = \\
=& (O(1)-u^\beta_{jk})\underbrace{H^k_{\alpha\beta}}_{\substack{=O(2)\\ \text{Step 3.(n-1) (3.92)}}}+\underbrace{O(2)H^{kl}_{\alpha\beta}}_{=O(2)\text{ always}}-u^\beta_{jkl}\underbrace{H^{kl}_{\alpha\beta}}_{\substack{\text{here we want to}\\ \text{get information}}},
\end{aligned}
$$

and therefore (since $H^{kl}_{\alpha\beta}$ is symmetric in k,l and third order must vanish) we get

$$\boxed{0 =H^{kl}_{\alpha\beta}.}$$

Equation b) and c) do not provide new information in Step 3.n. All together, we have proven what we wanted to prove and, again, the result is

$$\boxed{\begin{aligned} 0 =&H^{kl}_{\alpha\beta},\\ O(2) =&H^k_{\alpha\delta}.\end{aligned}}$$

For $n = 1$, we should probably do the proof separately, or we can use the slightly different proof in Section 3.4. $\square$

Remarks: The proof of Step 3 does not work any longer under slightly weaker assumptions, as we want to show here. Before we do so, we should be aware of what the assumptions have been in what we have proven above. We did not use the full structure of the Helmholtz expressions in the proof of Step 3. For example, we did not us the skew-symmetry of $H^{kl}_{\alpha\beta}$ in α,β.

Under weaker assumptions, here is a counter example for $n = 1$ and $m = 2$: If we assume f_α is a polynomial of degree 2, then the expressions for $H^k_{\alpha\beta}$ and $H^{kl}_{\alpha\beta}$ are of the form

$$\begin{aligned} H^x_{\alpha\beta} &= A_1 + A_{2\gamma} u^\gamma_{xx} + A_{3\gamma\delta} u^\gamma_{xx} u^\delta_{xx} + A_{4\gamma} u^\gamma_{xxx}, \\ H^{xx}_{\alpha\beta} &= B_1 + B_{2\gamma} u^\gamma_{xx}, \end{aligned}$$

where A_1,$A_{2\gamma}$,$A_{3\gamma\delta}$,$A_{4\gamma}$,B_1,$B_{2\gamma} = O(1)$. For example, we could choose

$$\begin{aligned} H^{xx}_{11} &= 0, & H^x_{11} &= -u^2_{xxx}, \\ H^{xx}_{12} &= u^1_{xx}, & H^x_{12} &= 0, \\ H^{xx}_{21} &= 0, & H^x_{21} &= 0, \\ H^{xx}_{21} &= 0, & H^x_{22} &= 0 \end{aligned}$$

and equation $ii)$

$$\begin{aligned} 0 &= u^\beta_{xx} H^x_{1\beta} + u^\beta_{xxx} H^{xx}_{1\beta} = -u^1_{xx} u^2_{xxx} + u^2_{xxx} u^1_{xx}, \\ 0 &= u^\beta_{xx} H^x_{2\beta} + u^\beta_{xxx} H^{xx}_{2\beta} \end{aligned}$$

would be satisfied, but $H^{xx}_{\alpha\beta} \neq 0$. Of course, a more interesting counter example would be if we define a function f_α which is a polynomial of degree 2 and then we compute $H^x_{\alpha\beta}$ and $H^{xx}_{\alpha\beta}$, such that $H^{xx}_{\alpha\beta} \neq 0$, but the equations $ii)$ are satisfied. This is probably not possible, since then $H^{xx}_{\alpha\beta}$ has definitely to be skew-symmetric in α,β and we would have additional restrictions. We do not discuss this in more detail.

3.6.2. Proof of Step 6

Let us directly start with the proof. Again, the proof will be a kind of induction and we will write Step 6.k for the different steps in the induction.

Step 6.0: From Step 5 in Section 3.6 we know that

$$\boxed{H^j_{\alpha\beta,u^\gamma_{jj}} = 0.} \tag{3.93}$$

We also know that $H^i_{\alpha\beta}$ is a polynomial of degree $\leq n$ in the second order jet coordinates (terms with third order jet coordinates even have degree $\leq n-2$ in the second order jet coordinates, but they already vanished). Therefore, we can assume that

$$\partial_{u^{\gamma_1}_{11}}\partial_{u^{\gamma_2}_{22}}...\partial_{u^{\gamma_n}_{nn}}H^j_{\alpha\beta} \neq 0,$$

in general. But this is actually not the case, since in the derivatives $\partial_{u^{\gamma_1}_{11}}\partial_{u^{\gamma_2}_{22}}...\partial_{u^{\gamma_n}_{nn}}$ we can definitely find a $\partial_{u^{\gamma}_{jj}}$-derivative and $H^j_{\alpha\beta,u^{\gamma}_{jj}} = 0$. Therefore, we actually get

$$\boxed{\partial_{u^{\gamma_1}_{11}}\partial_{u^{\gamma_2}_{22}}...\partial_{u^{\gamma_n}_{nn}}H^j_{\alpha\beta} = 0.}$$

Now we will use this property as the starting point in the induction and we will also use (3.93).

<u>Step 6.1 (apply n derivatives, then compute expression with $(n-1)$ derivatives):</u>
We consider equation $ii)$ and apply certain differential operators. Let $\partial_{u^{\gamma_1}_{11}}\partial_{u^{\gamma_2}_{22}}...\partial_{u^{\gamma_n}_{nn}}$ be a differential operator. Then we define

$$\partial_{u^{\gamma_1}_{11}}\partial_{u^{\gamma_2}_{22}}... \wedge_{jj} ...\partial_{u^{\gamma_n}_{nn}} := \partial_{u^{\gamma_1}_{11}}\partial_{u^{\gamma_2}_{22}}...\partial_{u^{\gamma_{j-1}}_{(j-1)(j-1)}}\partial_{u^{\gamma_{j+1}}_{(j+1)(j+1)}}...\partial_{u^{\gamma_n}_{nn}},$$

i.e. the notation $\wedge_{jj}$ means that the $\partial_{u^{\gamma_j}_{jj}}$-derivative is omitted (and all other derivatives are included). Now we apply the operator $\partial_{u^{\gamma_1}_{11}}\partial_{u^{\gamma_2}_{22}}...\partial_{u^{\gamma_n}_{nn}}$ to $ii)$, i.e.

$$ii) \quad 0 = (O^{\beta}_{jk}(1) - u^{\beta}_{jk})H^k_{\alpha\beta} \quad | \quad \partial_{u^{\gamma_1}_{11}}\partial_{u^{\gamma_2}_{22}}...\partial_{u^{\gamma_n}_{nn}}$$

and we get (there is no summation over j)

$$0 = O^{\beta}_{jk}(1) \underbrace{H^k_{\alpha\beta,u^{\gamma_1}_{11}...u^{\gamma_n}_{nn}}}_{\substack{=0,\text{ see Step 6.0,}\\ \text{definitely } \partial_{u^{\gamma}_{kk}}\text{-der.}\\ \text{included } \forall k}} - H^j_{\alpha\gamma_j,u^{\gamma_1}_{11}...\wedge_{jj}...u^{\gamma_n}_{nn}} - u^{\beta}_{jk} \underbrace{H^k_{\alpha\beta,u^{\gamma_1}_{11}...u^{\gamma_n}_{nn}}}_{\substack{=0,\text{ see Step 6.0,}\\ \text{definitely } \partial_{u^{\gamma}_{kk}}\text{-der.}\\ \text{included } \forall k}} .$$

This leads to

$$H^j_{\alpha\gamma_j,u^{\gamma_1}_{11}...\wedge_{jj}...u^{\gamma_n}_{nn}} = 0,$$

where we have $(n-1)$ derivatives on $H^j_{\alpha\gamma_j}$, all derivatives are different, and the $\partial_{u^{\gamma_j}_{jj}}$-derivative is omitted. For example, for $n = 2$ we get $H^x_{\alpha\beta,u^{\gamma}_{yy}} = 0$ and $H^y_{\alpha\beta,u^{\gamma}_{xx}} = 0$, but we do not (yet) get $H^x_{\alpha\beta,u^{\gamma}_{xx}} = 0$. But because of Step 6.0, see (3.93), we also know that

$$H^j_{\alpha\gamma_j,u^{\gamma_1}_{11}...\wedge_{kk}...u^{\gamma_n}_{nn}} = 0,$$

$(n-1)$ derivatives, all derivatives different,

the $\partial_{u^{\gamma_j}_{jj}}$-derivative is included, i.e. $k \neq j$.

Together we get

$$\boxed{\begin{gathered} H^j_{\alpha\gamma_j, u^{\gamma_1}_{11}\ldots\wedge_{kk}\ldots u^{\gamma_n}_{nn}} = 0 \\ (n-1) \text{ derivatives, all derivatives are different, for all } j,k = 1,2,\ldots n. \end{gathered}} \tag{3.94}$$

Note that the standard rule is: In Step 6.1 there is one derivative omitted, in Step 6.2 there will be two derivatives omitted and so on.

Step 6.2 (apply $(n-1)$ derivatives, compute expression with $(n-2)$ derivatives):
We apply the following differential operator to equation ii)

$$0 = (O(1) - u^\beta_{jk}) H^k_{\alpha\beta} \quad |\partial_{u^{\gamma_1}_{11}} \ldots \wedge_{ll} \ldots \partial_{u^{\gamma_n}_{nn}}.$$

If $j = l$, then we can commute all derivatives with u^β_{jk} and we do not get any new information. Therefore, let $j \neq l$, i.e. the $\partial_{u^{\gamma_j}_{jj}}$-derivative is included. Then we get

$$0 = O(1) \underbrace{H^k_{\alpha\beta, u^{\gamma_1}_{11}\ldots\wedge_{ll}\ldots u^{\gamma_n}_{nn}}}_{\substack{=0, \\ \text{Step 6.1, see (3.94)}}} - H^j_{\alpha\gamma_j, u^{\gamma_1}_{11}\ldots\wedge_{jj}\ldots\wedge_{ll}\ldots u^{\gamma_n}_{nn}} - u^\beta_{jk} \underbrace{H^k_{\alpha\beta, u^{\gamma_1}_{11}\ldots\wedge_{ll}\ldots u^{\gamma_n}_{nn}}}_{\substack{=0, \\ \text{Step 6.1, see (3.94)}}} \tag{3.95}$$

and we get

$$\begin{gathered} H^j_{\alpha\gamma_j, u^{\gamma_1}_{11}\ldots\wedge_{jj}\ldots\wedge_{ll}\ldots u^{\gamma_n}_{nn}} = 0, \\ (n-2) \text{ derivatives, where all derivatives different, } j \neq l. \end{gathered}$$

Note that when we would write $j = l$, then we would have to say which derivatives are included, since from the different derivatives $\partial_{u^{\gamma_1}_{11}}\ldots\partial_{u^{\gamma_n}_{nn}}$, we can only omit different derivatives. Also note that in the sum over k in (3.95) there exists a k such that $k = l$, and therefore it is not possible to apply Step 6.0 and we definitely need the result from Step 6.1 in (3.94). Because of Step 6.0, we can also write

$$\begin{gathered} H^j_{\alpha\gamma_j, u^{\gamma_1}_{11}\ldots\wedge_{rr}\ldots\wedge_{ll}\ldots u^{\gamma_n}_{nn}} = 0, \\ (n-2) \text{ derivatives, all derivatives different, } r \neq l \text{ and } r,l \neq j, \end{gathered}$$

i.e. the $\partial_{u^{\gamma_j}_{jj}}$-derivative in $\partial_{u^{\gamma_1}_{11}} \ldots \wedge_{rr} \ldots \wedge_{ll} \ldots \partial_{u^{\gamma_n}_{nn}}$ is included. Then, together we get

$$\boxed{\begin{gathered} H^j_{\alpha\gamma_j, u^{\gamma_1}_{11}\ldots\wedge_{rr}\ldots\wedge_{ll}\ldots u^{\gamma_n}_{nn}} = 0, \\ (n-2) \text{ derivatives, all derivatives are different, } r \neq l. \end{gathered}} \tag{3.96}$$

Step 6.3 (apply $(n-2)$ derivatives, compute expression with $(n-3)$ derivatives):

Now we are able to do this inductively, i.e. we can repeat exactly the same argument. Once again, we will briefly write down the next step. We apply the following differential operator to equation ii)

$$0 = (O(1) - u_{jk})H^k_{\alpha\beta} \quad |\partial_{u^{\gamma_1}_{11}} \ldots \wedge_{rr} \ldots \wedge_{ll} \ldots \partial_{u^{\gamma_n}_{nn}}.$$

Of course, $r \neq l$, since from different derivatives we can only omit different derivatives. If $j = l$ or $j = r$, i.e. the $\partial_{u^{\gamma_j}_{jj}}$-derivative is not included, then all derivatives commute with u^{β}_{jk} and we do not get any new information. Now let $r,l \neq j$ (and $r \neq l$), i.e. the $\partial_{u^{\gamma_j}_{jj}}$-derivative is included. Then we get

$$0 = O(1)\underbrace{H^k_{\alpha\beta,u^{\gamma_1}_{11}\ldots\wedge_{rr}\ldots\wedge_{ll}\ldots u^{\gamma_n}_{nn}}}_{\substack{=0,\\ \text{Step 6.2, see (3.96)}}} - H^j_{\alpha\gamma_j,u^{\gamma_1}_{11}\ldots\wedge_{jj}\ldots\wedge_{rr}\ldots\wedge_{ll}\ldots u^{\gamma_n}_{nn}} - u^{\beta}_{jk}\underbrace{H^k_{\alpha\beta,u^{\gamma_1}_{11}\ldots\wedge_{rr}\ldots\wedge_{ll}\ldots u^{\gamma_n}_{nn}}}_{\substack{=0,\\ \text{Step 6.2, see (3.96)}}},$$

and therefore we get

$$H^j_{\alpha\gamma_j,u^{\gamma_1}_{11}\ldots\wedge_{jj}\ldots\wedge_{rr}\ldots\wedge_{ll}\ldots u^{\gamma_n}_{nn}} = 0,$$
$$(n-3) \text{ derivatives, all derivatives are different and}$$
$$j,r,l \text{ are different (i.e. from different derivatives}$$
$$\text{we exclude three different derivatives)}$$

Note that because of symmetry, the index j in $H^j_{\ldots}$ can also be written as r or l. This is again true when the $\partial_{u^{\gamma_j}_{jj}}$-derivative is included, since then we can apply Step 6.0 and together we get

$$\boxed{\begin{array}{c} H^k_{\alpha\gamma_j,u^{\gamma_1}_{11}\ldots\wedge_{jj}\ldots\wedge_{rr}\ldots\wedge_{ll}\ldots u^{\gamma_n}_{nn}} = 0, \\ (n-3) \text{ derivatives, where all derivatives are different} \\ j,r,l \text{ are different, for all } k = 1,2,\ldots,n. \end{array}}$$

We repeat exactly the same argument from Step 6.l in Step 6.(l+1)

$$\vdots$$

until we get Step 6.(n-1).

Step 6.(n-1) (apply $n - (n-2) = 2$ derivatives, then compute expression with 1 derivative): We do the same calculation as before until we get

$$\boxed{\begin{array}{c} H^k_{\alpha\gamma_j,u^{\gamma_1}_{ll}} = 0, \\ (n-(n-1)) = 1 \text{ derivative, for all } k,l = 1,2,\ldots,n \end{array}}$$

and this is a generalization of Step 6.0, since the conditions now also holds for $k \neq l$.

Step 6.n (apply $n-(n-1)=1$ derivatives, compute expression with 0 derivatives):
In the last step we apply the differential operator $\partial_{u_{jj}^{\gamma_j}}$

$$0 = (O(1) - u_{jk}^{\beta}) H_{\alpha\beta}^{k} \quad |\partial_{u_{jj}^{\gamma_j}},$$

i.e. the $\partial_{u_{jj}^{\gamma_j}}$-derivative is included, as it always was the case above, since otherwise we do not get any new information, and we get

$$0 = O(1) \underbrace{H^{k}_{\alpha\beta,u_{jj}^{\gamma_j}}}_{\substack{=0,\\ \text{Step 6.(n-1)}}} - H^{j}_{\alpha\gamma_j} - u_{jk}^{\beta} \underbrace{H^{k}_{\alpha\beta,u_{jj}^{\gamma_j}}}_{\substack{=0,\\ \text{Step 6.(n-1)}}}.$$

Therefore,

$$\boxed{H^{j}_{\alpha\gamma_j} = 0}$$

and this proves the initial statement. □

Remarks: In the above proof, we did not even use that $H^k_{\alpha\beta}$ is a polynomial in the second order jet coordinates of degree $\leq n$. We only used the condition $H^{j}_{\alpha\beta,u_{jj}^{\gamma}} = 0$ at the beginning of the proof. This could be a difference to applying the d-fold operator, used in similar proofs, as in (AP96, p.379) or (MPV08, p.12).

If we would not be able to use the initial condition $H^{j}_{\alpha\beta,u_{jj}^{\gamma}} = 0$ (Helmholtz dependency) from Step 6.0 in Subsection 3.6.2, then we would not be able to do this proof. Since, for example, for $n=1$,$m=2$, the expressions $H^x_{\alpha 1} = -u^2_{xx}$ and $H^x_{\alpha 2} = u^1_{xx}$ satisfy the equation

$$0 = u^{\beta}_{xx} H^{x}_{\alpha\beta},$$

but it does not follow that $H^x_{\alpha\beta} = 0$.

More generally and what the proof of Step 6 is about: A system of the form

$$u^{\beta}_{ik} H^{k}_{\beta} = 0 \quad \text{for all } i = 1,2,...,n \tag{3.97}$$

has solutions of degree $\geq n$ and there are no non-trivial solutions of degree $< n$. To understand this, let us write (3.97) as

$$\begin{pmatrix} u^1_{xx} & u^1_{xy} & u^2_{xx} & u^2_{xy} \\ u^1_{yx} & u^1_{yy} & u^2_{yx} & u^2_{yy} \end{pmatrix} \begin{pmatrix} H^x_1 \\ H^y_1 \\ H^x_2 \\ H^y_2 \end{pmatrix} = 0, \tag{3.98}$$

where we consider $n,m = 2$ for simplicity and the more general case is analogous. To solve (3.98), we do the following trick: Let us write (3.98) as two terms of the form

$$\underbrace{\begin{pmatrix} u^1_{xx} & u^1_{xy} \\ u^1_{yx} & u^1_{yy} \end{pmatrix}}_{=:A} \begin{pmatrix} H^x_1 \\ H^y_1 \end{pmatrix} + \begin{pmatrix} u^2_{xx} & u^2_{xy} \\ u^2_{yx} & u^2_{yy} \end{pmatrix} \begin{pmatrix} H^x_2 \\ H^y_2 \end{pmatrix} = 0. \tag{3.99}$$

The matrix A can be inverted almost everywhere and multiplying the equation (3.99) by this almost everywhere invertible matrix, we get

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} H^x_1 \\ H^y_1 \end{pmatrix} + \frac{1}{u^1_{xx}u^1_{yy} - u^1_{xy}u^1_{xy}} \begin{pmatrix} u^1_{yy} & -u^1_{xy} \\ -u^1_{xy} & u^1_{xx} \end{pmatrix} \begin{pmatrix} u^2_{xx} & u^2_{xy} \\ u^2_{yx} & u^2_{yy} \end{pmatrix} \begin{pmatrix} H^x_2 \\ H^y_2 \end{pmatrix} = 0.$$

Therefore, the almost everywhere solution is

$$\begin{aligned} \begin{pmatrix} H^x_1 \\ H^y_1 \end{pmatrix} &= \frac{-1}{u^1_{xx}u^1_{yy} - u^1_{xy}u^1_{xy}} \begin{pmatrix} (u^1_{yy}u^2_{xx} - u^1_{xy}u^2_{yx}) & (u^1_{yy}u^2_{xy} - u^1_{xy}u^2_{yy}) \\ (u^1_{xx}u^2_{yx} - u^1_{xy}u^2_{xx}) & (u^1_{xx}u^2_{yy} - u^1_{xy}u^2_{xy}) \end{pmatrix} \begin{pmatrix} H^x_2 \\ H^y_2 \end{pmatrix} = \\ &= \frac{-1}{(u^1_x,u^1_y)} \begin{pmatrix} (u^2_x,u^1_y) & (u^2_y,u^1_y) \\ (u^1_x,u^2_x) & (u^1_x,u^2_y) \end{pmatrix} \begin{pmatrix} H^x_2 \\ H^y_2 \end{pmatrix}, \end{aligned}$$

where we used the notation of Hyperjacobians, see (3.47). Then we get

$$\begin{pmatrix} H^x_1 \\ H^y_1 \\ H^x_2 \\ H^y_2 \end{pmatrix} = \begin{pmatrix} \frac{(u^1_y,u^2_x)}{(u^1_x,u^1_y)} H^x_2 + \frac{(u^1_y,u^2_y)}{(u^1_x,u^1_y)} H^y_2 \\ \frac{(u^2_x,u^1_x)}{(u^1_x,u^1_y)} H^x_2 + \frac{(u^2_y,u^1_x)}{(u^1_x,u^1_y)} H^y_2 \\ H^x_2 \\ H^y_2 \end{pmatrix} = \begin{pmatrix} \frac{(u^1_y,u^2_x)}{(u^1_x,u^1_y)} \\ \frac{(u^2_x,u^1_x)}{(u^1_x,u^1_y)} \\ 1 \\ 0 \end{pmatrix} H^x_2 + \begin{pmatrix} \frac{(u^1_y,u^2_y)}{(u^1_x,u^1_y)} \\ \frac{(u^2_y,u^1_x)}{(u^1_x,u^1_y)} \\ 0 \\ 1 \end{pmatrix} H^y_2. \tag{3.100}$$

In the case where $H^x_2,H^y_2 = O(1)$, it is relatively easy to find out that there is only the trivial solution $H^k_\beta = 0$, or singular non-polynomial solutions. This follows, since there are no linear relations between the Hyperjacobinans over $O(1)$-coefficients for $n = 2$. This is no longer true for $n \geq 4$, see (3.55). In the case where $H^x_2,H^y_2 = O_{lin}(2)$ this kind of discussion is already more complicated and we need to take in account the relations in (3.54) and similar ones. Although it is more complicated, we can still investigate all the relations between the Hyperjacobians and we can show that there is only the trivial solution, or singular non-polynomial solutions. Now let us write $H^x_2 = (u^1_x,u^1_y)\lambda_1$, $H^y_2 = (u^1_x,u^1_y)\lambda_2$, where $\lambda_1,\lambda_2 = O(2)$ are smooth functions (polynomials in second order coordinates). Then we get

$$\begin{pmatrix} H^x_1 \\ H^y_1 \\ H^x_2 \\ H^y_2 \end{pmatrix} = \begin{pmatrix} (u^1_y,u^2_x) \\ (u^2_x,u^1_x) \\ (u^1_x,u^1_y) \\ 0 \end{pmatrix} \lambda_1 + \begin{pmatrix} (u^1_y,u^2_y) \\ (u^2_y,u^1_x) \\ 0 \\ (u^1_x,u^1_y) \end{pmatrix} \lambda_2.$$

and this shows that there are definitely polynomial solutions of degree ≥ 2 in second order coordinates. A similar calculation holds in any dimension and the equation $u^\beta_{ik} H^k_{\alpha\beta} = 0$ can always be written in the form

$$u^1_{ik} H^k_1 + \sum_{\beta=2}^{m} u^\beta_{ik} H^k_\beta = 0$$

and then we can multiply with the almost everywhere invertible matrix $(u^1_{ik})_{i,k}$, where the determinant of this matrix has degree n. This is the idea why we always get non-trivial polynomial solutions of degree $\geq n$ and why there are no non-trivial (non-singular and polynomial) solutions of degree $< n$. Of course, that the entries in the solution vector H^k_β, i.e. $(u^1_y, u^2_x) H^x_2 + (u^1_y, u^2_y) H^y_2$ and $(u^2_x, u^1_x) H^x_2 + (u^2_y, u^1_x) H^y_2$ in (3.100), do not factor through a certain Hyperjacobian simultaneously when H^k_β is of degree $< n$ has to be proven in more detail and we will not further discuss it here. This is probably a complicated discussion for general n, when one has to understand all the relations between the Hyperjacobians. Hilbert's syzygy theorem could also be of importance here, although we never investigated this problem in detail. There is definitely a rich algebraic structure behind this and we should also keep in mind the Plücker relations, which have been discussed earlier, see (3.53), (3.54) and (3.55).

Note, the fact that $u^\beta_{ik} H^k_{\alpha\beta} = 0$ has no non-trivial polynomial solutions of degree $< n$ can be proven with a similar proof as we have done above or see Lemma 2.3. in (AP95, p.629) or (MPV08, p.13).

Together, the above example shows that an equation of the form $u^\beta_{ik} H^k_{\alpha\beta} = 0$ has a non-trivial solution of degree $\geq n$ and has no non-trivial solution of degree $< n$. Moreover, $H^k_{\alpha\beta}$ has degree $\leq n$, see Step 2 in Section 3.6. This means that it is possible to get a non-trivial solution of degree n. However, the conditions $H^i_{\alpha\beta, u^\gamma_{ii}} = 0$ is sufficient to prove that there are only the trivial solutions and this is what Step 6 is about.

3.7. Methods from Multi-Linear Algebra

The goal in this section is to use some simple ideas from multi-linear algebra to analyze the large systems of differential equations which arise in problems in the calculus of variations. For example, this ideas can be used to solve equations (3.35)-(3.39) or (3.57)-(3.66) in some sense. We could say that there is a formalism which tells us how to solve such equations or to get properties of the solutions. We will start this section with some motivation, or more precisely, we will start with special algebraic equations and later we will apply the formalism to differential equations. The following discussion is about multi-linear forms with additional properties, like symmetry or skew-symmetry conditions. The results in this section are not new and they can be found in (And89) and (AD80).

Let $T : \mathbb{R}^n \times ... \times \mathbb{R}^n \to \mathbb{R}$ be a k-tensor. This map can be written as

$$T(A^1,...,A^k) = T^{i_1...i_k} A^1_{i_1}...A^k_{i_k},$$

where $A^1,...,A^k \in \mathbb{R}^n$ and $i_l \in \{1,2,...,n\}$, $1 \leq l \leq k$, label the components of the vectors $A^1,A^2,...,A^k \in \mathbb{R}^n$. The tensor T vanishes identically, by definition, if $T^{i_1...i_k} = 0$ for all indices $i_1,...,i_k$. Now there is an equivalent way how to show the vanishing of this tensor. To explain this, let us consider a few very simple examples to understand the idea, before we will do a more complicated calculation in the case where we will need similar techniques.

Linear forms: If we have a linear form $T : \mathbb{R}^n \to \mathbb{R}$ such that

$$0 = T(A) = T^i A_i \quad \text{for all } A \in \mathbb{R}^n,$$

then the form T vanishes identically.

Bilinear forms: If we have a bilinear form $T : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ such that

$$0 = T(A,B) = T^{ij} A_i B_j \quad \text{for all } A,B \in \mathbb{R}^n, \tag{3.101}$$

then, again, T vanishes identically. If we have a bilinear form T such that

$$0 = T(A,A) = T^{ij} A_i A_j \quad \text{for all } A \in \mathbb{R}^n,$$

then the form T must be skew-symmetric, i.e. $T(A,B) = -T(B,A)$ for all $A,B \in \mathbb{R}^n$. If however $n = 1$, then skew-symmetry implies $T \equiv 0$. How can we prove this? Well for $n = 1$, skew- symmetry is equivalent to (3.101), since for every A,B (except for $A = 0$) we can find a $\lambda \in \mathbb{R}$, such that $B = \lambda A$ and then

$$T(A,B) = \lambda T(A,A) = 0, \quad \text{for all } A,B \in \mathbb{R} \text{ with } A \neq 0,$$

and therefore for almost every A,B. When $A = 0$, it is clear that $T(A,B) = 0$, by linearity of T. Therefore, $T(A,B)$ vanishes for all A,B and this means $T \equiv 0$. Now the idea is the following, for small n (here we had $n = 1$) the following three statements are equivalent

$$T(A,A) = 0 \;\forall A \in \mathbb{R}^n \quad \Leftrightarrow \quad T(A,B) = 0 \;\forall A,B \in \mathbb{R}^n \quad \Leftrightarrow \quad T \equiv 0.$$

For large n (here we had for $n \geq 2$), this is not true, but we will **extend** the linear form T to a linear form $\tilde{T}(A^1,...,A^l)$, where $l > n$ and then it is true for the extended form $\tilde{T}$.

Let us do one more example.

<u>Trilinear forms:</u> If we have a trilinear form T such that

$$0 = T(A,B,C) = T^{ijk} A_i B_j C_k \quad \text{for all } A,B,C \in \mathbb{R}^n, \tag{3.102}$$

then the form T vanishes identically. If we have a form T, such that

$$\begin{aligned} 0 &= T(A,A,B) = T^{ijk} A_i A_j B_k \quad \text{for all } A,B \in \mathbb{R}^n, \\ 0 &= T(B,A,A) = T^{ijk} A_i A_j B_k \quad \text{for all } A,B \in \mathbb{R}^n, \end{aligned} \tag{3.103}$$

then in general $T \neq 0$. For example, the Levi-Civita tensor $T^{ijk} = \epsilon^{ijk}$ is not identical zero in general, but satisfies (3.103). In the case where we consider the Levi-Civita tensor and $n = 2$, (3.103) is equivalent to (3.102), since for almost every $A,B,C \in \mathbb{R}^2$ we can find $\lambda_1,\lambda_2 \in \mathbb{R}$ such that

$$C = \lambda_1 A + \lambda_2 B$$

and then we get

$$\begin{aligned} T(A,B,C) &\overset{\text{linearity}}{=} \lambda_1 T(A,B,A) + \lambda_2 T(A,B,B) = \\ &\overset{\text{skew-sym.}}{=} -\lambda_1 T(A,A,B) + \lambda_2 T(A,B,B) \overset{(3.103)}{=} 0. \end{aligned} \tag{3.104}$$

By a density argument this holds for every $A,B,C \in \mathbb{R}^2$, and therefore $T \equiv 0$. Here, density argument means that if the equation holds for almost every $A,B,C \in \mathbb{R}^2$ then it must also hold for every $A,B,C \in \mathbb{R}^2$. In other words, for $n = 2$, at least two of the three indices i,j,k in the Levi-Civita tensor ϵ^{ijk} must be the same, and therefore $\epsilon^{ijk} \equiv 0$.

In (3.104), we used (3.103) and also skew-symmetry of the Levi-Civita tensor. However, this kind of argumentation also works when we assume the conditions

$$\begin{aligned} 0 = &T(A,A,B) = T^{ijk} A_i A_j B_k \quad \text{for all } A,B \in \mathbb{R}^n, \\ &T(A,B,C) \quad \text{is symmetric in } A,B,C, \end{aligned}$$

since this implies (3.103) and we can do the calculation in (3.104), except with a plus instead of a minus in the second line (i.e. using symmetry instead of skew-symmetry).

Now we want to find a generalization of that and apply it to our problems, especially to Takens' problem. In the following, we have the additional property that all of our forms are symmetric in $A^1, A^2, ..., A^l$, i.e.

$$T(A^1, A^2, ..., A^l) = T(A^{\pi(1)}, A^{\pi(2)}, ..., A^{\pi(l)})$$

for all permutation π. The idea is formulated in the following formal lemma:

Formal Symmetric Tensor Lemma: Let T be a symmetric, multi-linear form, defined for $A^k \in \mathbb{R}^n$, where $k = 1,2,...,l$, i.e. a map of the form

$$T : \underbrace{\mathbb{R}^n \times \mathbb{R}^n \times ... \times \mathbb{R}^n}_{l-times} \to \mathbb{R}.$$

If $l \geq n + 1$ and

$$0 = T(A^1, A^1, A^2, ..., A^{l-1}) \quad \textit{for all } A^1, A^2, ..., A^{l-1} \in \mathbb{R}^n \tag{3.105}$$

then $T \equiv 0$.

The proof is simple and follows by the fact that in the case where $l \geq n + 1$, we get that $A^1, A^2, ..., A^l$ must always be linearly dependent and we can almost always write $A^l = \lambda_1 A^1 + ... + \lambda_{l-1} A^{l-1}$. Then by symmetric permutations we can use the condition (3.105). By a density argument we get that the form must vanish everywhere. Later, we will also need slightly modifications of that formal lemma.

3.7.1. Classification of Trivial Divergences of Second Order

In this subsection, we will discuss the first example, where we can use the techniques which we explained above. The goal is to characterizing trivial divergences.

Let $f^i = f^i(x^k, u^\alpha, u^\alpha_k, u^\alpha_{kl})$ be a second order tensor with vanishing divergence, i.e.

$$0 = D_i f^i = \underbrace{f^i_{x^i} + u^\alpha_i f^i_{u^\alpha} + u^\alpha_{ik} f^i_{u^\alpha_k}}_{=O(2)} + c_{kl} u^\alpha_{ikl} f^i_{u^\alpha_{kl}},$$

for all points in J^3E. A necessary condition is that third order terms must vanish, i.e.

$$c_{kl} u^\alpha_{ikl} f^i_{u^\alpha_{kl}} = 0 \tag{3.106}$$

and this condition classifies the leading order of f^i. For simplicity, let us consider the case where $n = 2$, the more general case is analogous. For $n = 2$, we write $(f^i) = (f^x, f^y)$.

When $n = 2$, from equation (3.106) we get four conditions

$$\begin{aligned} a) \quad & u^\alpha_{xxx}: \quad 0 = f^x_{u^\alpha_{xx}}, && (3.107)\\ b) \quad & u^\alpha_{xxy}: \quad 0 = f^y_{u^\alpha_{xx}} + f^x_{u^\alpha_{xy}}, && (3.108)\\ c) \quad & u^\alpha_{xyy}: \quad 0 = f^x_{u^\alpha_{yy}} + f^y_{u^\alpha_{xy}}, && (3.109)\\ d) \quad & u^\alpha_{yyy}: \quad 0 = f^y_{u^\alpha_{yy}}. && (3.110) \end{aligned}$$

We differentiate $b)$ and $c)$ with respect to $\partial_{u^\beta_{xx}}$ and $\partial_{u^\beta_{yy}}$ and we use $a)$ and $d)$ to deduce

$$\begin{aligned} b)\,i) \quad & f^y_{u^\beta_{xx}u^\alpha_{xx}} = 0,\\ b)\,ii) \quad & f^x_{u^\beta_{yy}u^\alpha_{xy}} = 0,\\ c)\,iii) \quad & f^y_{u^\beta_{xx}u^\alpha_{xy}} = 0,\\ c)\,iv) \quad & f^x_{u^\beta_{yy}u^\alpha_{yy}} = 0. \end{aligned}$$

Then we differentiate $b)$ and $c)$ with respect to $\partial_{u^\beta_{xy}}$ and we use $c)\,iii)$ and $b)\,ii)$ to deduce

$$\begin{aligned} b)\,v) \quad & f^x_{u^\beta_{xy}u^\alpha_{xy}} = 0, && (3.111)\\ c)\,vi) \quad & f^y_{u^\beta_{xy}u^\alpha_{xy}} = 0. && (3.112) \end{aligned}$$

Then, from equations (3.107)-(3.112) we can deduce that

$$\begin{aligned} & f^x_{u^\alpha_{xx}}, f^x_{u^\beta_{yy}u^\alpha_{xy}}, f^x_{u^\beta_{yy}u^\alpha_{yy}}, f^x_{u^\beta_{xy}u^\alpha_{xy}} = 0,\\ & f^y_{u^\alpha_{yy}}, f^y_{u^\beta_{xx}u^\alpha_{xx}}, f^y_{u^\beta_{xx}u^\alpha_{xy}}, f^y_{u^\beta_{xy}u^\alpha_{xy}} = 0 \qquad (3.113) \end{aligned}$$

and this means that all 2-combinations of second order derivatives applied to f^x or f^y must vanish. Therefore, f^i must be affine linear in u^α_{xx}, u^α_{xy}, u^α_{yy}, i.e.

$$f^i = a^i + b^{ikl}_\alpha u^\alpha_{kl},$$

where $a^i, b^{ikl}_\alpha = O(1)$. We could also solve equations $a)$, $b)$, $c)$ and $d)$ exactly, but we do not need the more precise result here.

For example, let us consider $n = 2$. Then

$$\begin{pmatrix} f^x \\ f^y \end{pmatrix} = \begin{pmatrix} u_{yy} \\ -u_{xy} \end{pmatrix}$$

is a trivial divergence of second order, which is a polynomial of degree $n-1=1$ in second order coordinates. Note that if $D_x f^x + D_y f^y = 0$, then for $n=2$, we could also have shown that locally

$$\begin{pmatrix} f^x \\ f^y \end{pmatrix} = \begin{pmatrix} D_y\phi \\ -D_x\phi \end{pmatrix} = \begin{pmatrix} O(1) + u_{xy}\phi_{u_x} + u_{yy}\phi_{u_y} \\ O(1) - u_{xx}\phi_{u_x} - u_{xy}\phi_{u_y} \end{pmatrix}$$

for some function $\phi = O(1)$, by standard exactness arguments and order discussion. However, for $n>2$, this is more complicated and, more precisely, the leading order is not obvious.

For $n=1$, only constants are trivial conservation laws, which are polynomials of degree $n-1=0$.

In a more general situation, we have to find a better method how to solve equations of the form (3.106) or similar ones. Let us now explain how this works. First, we define

$$\partial_\alpha^{kl} := c_{kl}\partial_{u^\alpha_{kl}}, \tag{3.114}$$

since this simplifies the expressions slightly and symmetrization can be seen explicitly (note that $c_{kl}=1$ if $k=l$ and $c_{kl}=\frac{1}{2}$ if $k\neq l$). The notation in (3.114) can be found in (Poh95, p.344). Then (3.106) can be written as

$$0 = u^\alpha_{ikl}\partial_\alpha^{kl} f^i \quad \Leftrightarrow \quad 0 = \partial_\alpha^{(kl} f^{i)},$$

where brackets (kli) mean symmetrization in k,l,i. Symmetrization occurs because u^α_{ikl} is symmetric in i,k,l. Next, we set

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \quad \in \mathbb{R}^2$$

and we consider the equation

$$\begin{aligned} 0 = A_iA_kA_l\partial_\alpha^{kl}f^i =& A_1A_1A_1\partial_\alpha^{11}f^1 + \\ &+A_1A_1A_2\partial_\alpha^{12}f^1 + A_1A_2A_1\partial_\alpha^{21}f^1 + A_2A_1A_1\partial_\alpha^{11}f^2 + \\ &+A_1A_2A_2\partial_\alpha^{22}f^1 + A_2A_1A_2\partial_\alpha^{12}f^2 + A_2A_2A_1\partial_\alpha^{21}f^2 + \\ &+A_2A_2A_2\partial_\alpha^{22}f^2 = \\ \\ =& A_1A_1A_1\partial_\alpha^{11}f^1 + \\ &+A_1A_1A_2(2\partial_\alpha^{12}f^1 + \partial_\alpha^{11}f^2) + \\ &+A_1A_2A_2(\partial_\alpha^{22}f^1 + 2\partial_\alpha^{12}f^2) + \\ &+A_2A_2A_2\partial_\alpha^{22}f^2. \end{aligned}$$

If we assume that this equations must be satisfied for every $A \in \mathbb{R}^2$ then we get

$$\begin{aligned}
0 &= \partial_\alpha^{11} f^1, \\
0 &= 2\partial_\alpha^{12} f^1 + \partial_\alpha^{11} f^2, \\
0 &= \partial_\alpha^{22} f^1 + 2\partial_\alpha^{12} f^2, \\
0 &= \partial_\alpha^{22} f^2
\end{aligned}$$

and this is exactly the system in (3.107)-(3.110). Therefore,

$$\begin{aligned}
& 0 = c_{kl} u^\alpha_{ikl} f^i_{,u^\alpha_{kl}}, \qquad (\text{for all } u^\alpha_{ikl}) \\
\Leftrightarrow \quad & 0 = A_i A_k A_l \partial_\alpha^{kl} f^i, \quad \text{for all } A \in \mathbb{R}^2 \text{ and all } \alpha
\end{aligned}$$

are equivalent. To assign (or replace) the polynomial $A_i A_k A_l$ to the coordinate u^α_{ikl} is a very special case of the **Gelfand Dikii transformation**, which can be found in (GD75) and (AP95, p.634). Note that the assignment of the polynomial $A_i A_k A_l$ to the coordinates u^α_{ikl} can be understood when we evaluate on a section

$$u^\alpha(x^j) = u_0^\alpha + u^\alpha_{0i} x^i + \frac{1}{2} u^\alpha_{0ik} x^i x^k + \frac{1}{6} c^\alpha A_i A_k A_l x^i x^k x^l.$$

This transformation has similarities to the Fourier transformation, where we also replace the expression $u'''(x)$ by the transformed expression $\xi^3 \hat{u}(\xi)$. (Roughly speaking, the third derivative of $u(x)$ can be identified with $\xi^3 = \xi\xi\xi$ in some sense. This is just the rough idea and why there are similarities, of course we cannot forget the factor $\hat{u}(\xi)$.) Beside some similarities, the Gelfand Dikii transformation is also different, since we can transform non-linear terms, like $u'(x)u''(x)$ and differential forms, as well, however the non-linearity has to be polynomial. We could also say that the Gelfand Dikii transformation is an appliance of a polarization operator (AP95, p.634). Let us consider the following notation

$$0 = A_k A_l A_i \partial_\alpha^{kl} f^i = A_k A_l A_i T_\alpha^{kli} = T_\alpha(A{,}A|A),$$

where we define $\partial_\alpha^{kl} f^i =: T_\alpha^{kli}$. We want to define T_α as a multi-linear form. Therefore, we define the map

$$T_\alpha(A{,}B|C) := A_k B_l C_i T_\alpha^{kli}, \quad \text{for all } A{,}B{,}C \in \mathbb{R}^2 \tag{3.115}$$

and we know that

$$T_\alpha(A{,}A|A) = 0 \quad \text{for all } A \in \mathbb{R}^2. \tag{3.116}$$

Assigning an object $T_\alpha(A{,}B|C)$ to the object $T_\alpha(A{,}A|A)$ can be considered as a kind of **polarization technique**, which is a well known theory for homogeneously polynomials. Roughly speaking, polarization extends an object, which is only defined for A, to an object, which is defined for $(A{,}B{,}C)$. Actually, we will only need the form $T_\alpha(A,A|B) = A_k A_l B_i T_\alpha^{kli}$ for all $A,B \in \mathbb{R}^2$, but it is reasonable to start with the more general definition in (3.115) when introducing this technique (otherwise the expressions are non-linear in A).

Note that if we would know that $T_\alpha(A,A|B) = 0$ for all $A,B \in \mathbb{R}^2$, then we would know that all second order derivatives of f^i must vanish, but we do not know that and it is not true in general.

Now we extend the form T_α by applying differential operators and we will get the **extended form** $T_{\beta\alpha}$. Note that we derived equations (3.113) by applying differential operators to a), b), c) and d) and now we want to do the same manipulation with algebraic methods. We define

$$\begin{aligned}T_{\beta\alpha}(A,A|B,B|C) :=& A_k A_l B_i B_j C_p T_{\beta\alpha}^{klijp} = \\ =& A_k A_l B_i B_j C_p \partial_\beta^{kl} \partial_\alpha^{ij} f^p, \quad \text{for all } A,B,C \in \mathbb{R}^2.\end{aligned}$$

Since partial derivatives commute (by Schwartz's theorem), we get the symmetry condition

$$T_{\beta\alpha}(A,A|B,B|C) = T_{\alpha\beta}(B,B|A,A|C), \quad \text{for all } A,B,C \in \mathbb{R}^2.$$

From equation (3.116) we deduce

$$\begin{aligned}T_{\beta\alpha}(A,A|B,B|B) &= A_k A_l \partial_\beta^{kl} T_\alpha(B,B|B) = 0, \quad \text{for all } A,B \in \mathbb{R}^2,\\ T_{\beta\alpha}(A,A|B,B|A) &= T_{\alpha\beta}(B,B|A,A|A) = 0, \quad \text{for all } A,B \in \mathbb{R}^2,\end{aligned}$$

but we do not know if $T_{\beta\alpha}(A,A|B,B|C) = 0$ for all A,B,C and this would mean that all 2-combinations of second order derivatives vanish.

Almost all randomly chosen vectors $A,B \in \mathbb{R}^2$ are linearly independent, since the set of linearly dependent vectors

$$R := \{(A,B) \in \mathbb{R}^2 \times \mathbb{R}^2 : \; A = \lambda B, \, \lambda \in \mathbb{R}\}$$

has measure zero in the set of all pairs (A,B), which is defined as

$$G := \{(A,B) \in \mathbb{R}^2 \times \mathbb{R}^2\}.$$

In other words, $G \setminus R$ is dense in G. Also note that the set G has dimension four, since we have four parameters (A_1,A_2,B_1,B_2) and the set R has dimension three, since it can be described by three parameters (A_1,A_2,λ). Now let $(A,B) \in G \setminus R$, then every $C \in \mathbb{R}^2$ can be written as linear combination of A,B for some $\lambda_1,\lambda_2 \in \mathbb{R}$ as

$$C = \lambda_1 A + \lambda_2 B. \tag{3.117}$$

Then by linearity of $T_{\beta\alpha}$, we get

$$T_{\beta\alpha}(A,A|B,B|C) = \lambda_1 \underbrace{T_{\beta\alpha}(A,A|B,B|A)}_{=0} + \lambda_2 \underbrace{T_{\beta\alpha}(A,A|B,B|B)}_{=0} = 0, \tag{3.118}$$

i.e. $T_{\beta\alpha} \equiv 0$ in $G \setminus R$. By continuity of linear forms and density of $G \setminus R$ in G, we get $T_{\beta\alpha} \equiv 0$ in G. This means all 2-combinations of second order derivatives applied to f^i must vanish, and therefore f^i must be a polynomial of degree $n-1=1$ in the second order jet coordinates.

For example, for $n=1$ and $B \neq 0$ we can always find a $\lambda \in \mathbb{R}$ such that $A = \lambda B$ for every A and $T_\alpha(A,A|B) = \lambda T_\alpha(A,A|A) = 0$, where $A,B \in \mathbb{R}$. Therefore, for $n=1$ the function f must be a polynomial of degree zero, i.e. a constant.

Solving the system of differential equations (3.107)-(3.110) directly or solving the system with the help of $T_{\beta\alpha}$ are two essentially different methods. Solving equations (3.107)-(3.110) directly does not require any density arguments and this is an essential difference. Since three vectors $A,B,C \in \mathbb{R}^2$ are always linearly dependent, we get the following observation: In the case where we write C as linear combination of A and B, we can immediately derive (3.118), and therefore the density argument is not needed there. But if we write A as linear combination of B and C, then we get

$$\begin{aligned} &T_{\beta\alpha}(\lambda_1 B + \lambda_2 C,\lambda_1 B + \lambda_2 C|B,B|C) = \\ =&\lambda_1 T_{\beta\alpha}(B,\lambda_1 B + \lambda_2 C|B,B|C) + \lambda_2 T_{\beta\alpha}(C,\lambda_1 B + \lambda_2 C|B,B|C) = \\ =&\lambda_1\lambda_1 T_{\beta\alpha}(B,B|B,B|C) + \lambda_1\lambda_2 T_{\beta\alpha}(B,C|B,B|C) + ..., \end{aligned}$$

i.e. expressions of the form $T_{\beta\alpha}(B,B|B,B|C), T_{\beta\alpha}(B,C|B,B|C)$, where we do not immediately know if they vanish or not. However, using the density argument, we do not have to consider such expressions explicitly and this is the advantage of this method.

For $n=2$, solving the system of differential equations (3.107)-(3.110) directly or solving the problem with the help of $T_{\beta\alpha}$ seems to be equivalently difficult, but for large n, solving the problem with the help of $T_{\beta\alpha}$ is much easier and it is almost impossible to solve the system systematically for large n in this or similar situations. However, note that we do not completely solve the equations, we only derive that the solutions must be polynomial of a certain degree.

Let us briefly present how we can solve the equation

$$0 = c_{kl} u^\alpha_{ikl} f^i_{u^\alpha_{kl}} \tag{3.119}$$

for arbitrary n. We define the extended form

$$T_{\alpha_1\ldots\alpha_n}(A^1,A^1|...|A^n,A^n|B) := A^1_{i_1} A^1_{j_1} ... A^n_{i_n} A^n_{j_n} B_i \partial^{i_1 j_1}_{\alpha_1} ... \partial^{i_n j_n}_{\alpha_n} f^i.$$

From (3.119) it follows that

$$T_{\alpha_1\alpha_2\ldots\alpha_n}(A^1,A^1|A^2,A^2|...|A^n,A^n|B) = 0,$$

whenever $A^k = B$ for at least one $k = 1,2,...,n$. Since $A^1,...,A^n,B \in \mathbb{R}^n$, these vectors must be linearly dependent and almost always we can write

$$B = \sum_{k=1}^{n} \lambda_k A^k.$$

Then by density, $T_{\alpha_1\alpha_2\ldots\alpha_n}$ vanishes for all A^k,$B \in \mathbb{R}^n$. Therefore, all n-combinations of second order derivatives applied to f^i must vanish and this means that f^i must be a polynomial of degree $n-1$.

For example, for $n=3$,

$$(f^i) = \begin{pmatrix} f^x \\ f^y \\ f^z \end{pmatrix} = \begin{pmatrix} \frac{D(u_y,u_z)}{D(y,z)} \\ -\frac{D(u_y,u_z)}{D(x,z)} \\ \frac{D(u_y,u_z)}{D(x,y)} \end{pmatrix} = \begin{pmatrix} u_{yy}u_{zz} - u_{yz}u_{yz} \\ -(u_{xy}u_{zz} - u_{yz}u_{xz}) \\ u_{xy}u_{yz} - u_{yy}u_{xz} \end{pmatrix}$$

is a trivial divergence since by direct calculation

$$\begin{aligned} D_i f^i =& u_{xyy}u_{zz} + u_{yy}u_{xzz} - 2u_{xyz}u_{yz} - \\ &- (u_{xyy}u_{zz} + u_{xy}u_{yzz}) + (u_{yyz}u_{xz} + u_{yz}u_{xyz}) + \\ &u_{xyz}u_{yz} + u_{xy}u_{yzz} - (u_{yyz}u_{xz} + u_{yy}u_{xzz}) = 0. \end{aligned}$$

We also observe that f^i is a polynomial of degree $n-1=2$.

The idea and proof in this subsection can be found in (AP96, p.377). Generalizations can be found in (AD80). In the next subsection we want to classify trivial Lagrangians with similar methods.

3.7.2. Classification of Trivial Lagrangians of Second Order, $n=2$

Here we want to present a second example, where we can use the techniques from multi-linear algebra which have been explained above. Let L be a second order Lagrangian (we also think of $L = Q^\alpha f_\alpha$) and, for simplicity, we first consider the case $n=2$. Later, we will generalize it. Solving

$$\mathcal{E}_\alpha L = 0 \quad \text{or} \quad \mathcal{H}^\gamma_{\alpha\beta} f_\gamma = 0 \tag{3.120}$$

forces to highest order the equations (3.35)-(3.39), i.e. $u^\alpha_{(4)}$-terms must vanish. From the $u^\alpha_{(3)}u^\beta_{(3)}$-terms, we also get equation (3.61). These equations force that L or f_α must be polynomials of degree two in second order jet coordinates, as we found out earlier. It took quite some effort to derive this, especially solving the equations (3.35)-(3.39). We now want to find a faster method, and later also a method for arbitrary n, how to prove this.

We apply $\partial_{u^\delta_{xx}}$,$\partial_{u^\delta_{xy}}$,$\partial_{u^\delta_{yy}}$ to (3.35) and (3.39) and we get (we write L instead of

f_β)

$$\begin{aligned}
0 &= L_{u^\alpha_{xx} u^\gamma_{xx} u^\delta_{xx}},\\
0 &= L_{u^\alpha_{xx} u^\gamma_{xx} u^\delta_{yy}},\\
0 &= L_{u^\alpha_{xx} u^\gamma_{xx} u^\delta_{xy}},\\
0 &= L_{u^\alpha_{yy} u^\gamma_{yy} u^\delta_{xy}},\\
0 &= L_{u^\alpha_{yy} u^\gamma_{yy} u^\delta_{xx}},\\
0 &= L_{u^\alpha_{yy} u^\gamma_{yy} u^\delta_{yy}}.
\end{aligned} \tag{3.121}$$

Then we apply $\partial_{u^\delta_{xx}}$,$\partial_{u^\delta_{yy}}$ to (3.37), we use the conditions in (3.121), and we deduce

$$\begin{aligned}
0 &= L_{u^\alpha_{xy} u^\gamma_{xy} u^\delta_{xx}},\\
0 &= L_{u^\alpha_{xy} u^\gamma_{xy} u^\delta_{yy}}.
\end{aligned} \tag{3.122}$$

Next, we apply $\partial_{u^\delta_{xy}}$ to equation (3.37) and we deduce

$$0 = L_{u^\alpha_{xx} u^\gamma_{yy} u^\delta_{xy}} + L_{u^\alpha_{yy} u^\gamma_{xx} u^\delta_{xy}} + L_{u^\alpha_{xy} u^\gamma_{xy} u^\delta_{xy}}. \tag{3.123}$$

Then we change $\delta \leftrightarrow \alpha$ and $\delta \leftrightarrow \gamma$ in (3.123), we add the changed equations, and we deduce

$$\begin{aligned}
0 =& L_{u^\delta_{xx} u^\gamma_{yy} u^\alpha_{xy}} + L_{u^\delta_{yy} u^\gamma_{xx} u^\alpha_{xy}} + L_{u^\alpha_{xy} u^\gamma_{xy} u^\delta_{xy}} + \\
&+ L_{u^\alpha_{xx} u^\delta_{yy} u^\gamma_{xy}} + L_{u^\alpha_{yy} u^\delta_{xx} u^\gamma_{xy}} + L_{u^\alpha_{xy} u^\delta_{xy} u^\gamma_{xy}} = \\
=& \partial_{u^\delta_{xx}} \underbrace{\left(L_{u^\gamma_{yy} u^\alpha_{xy}} + L_{u^\alpha_{yy} u^\gamma_{xy}}\right)}_{=0,\ (3.38)} + \partial_{u^\delta_{yy}} \underbrace{\left(L_{u^\gamma_{xx} u^\alpha_{xy}} + L_{u^\alpha_{xx} u^\gamma_{xy}}\right)}_{=0,\ (3.36)} + 2L_{u^\alpha_{xy} u^\gamma_{xy} u^\delta_{xy}} = \\
=& 2L_{u^\alpha_{xy} u^\gamma_{xy} u^\delta_{xy}}.
\end{aligned} \tag{3.124}$$

From the conditions in (3.121), (3.122) and (3.124), we get that all 3-combinations of partial derivatives

$$L_{u^\alpha_{ij} u^\gamma_{kl} u^\delta_{pq}} = 0$$

must vanish, except

$$L_{u^\alpha_{xx}, u^\gamma_{yy} u^\delta_{xy}}.$$

It is clear that at least one of these combinations cannot vanish, since (u^α,u^β,u^γ)-terms are still allowed after $u^\alpha_{(4)}$-discussion, see (3.48). After using the condition

(3.61), we get that L must be a polynomial of degree 2 in second order coordinates. It is an interesting observation that we do not need to discuss the $u^\alpha_{(3)}$-terms in (3.120) to derive this. The $u^\alpha_{(3)}$-terms are a result of D_i- and D_{ij}-derivatives, whereas the discussion above results from pure D_{ij}-derivatives.

Now let us discuss another method how to derive this result and how the discussion can be generalized for arbitrary n in the next subsection. From the results above it is clear that we need to consider $u^\alpha_{(4)}$- and $u^\alpha_{(3)}u^\beta_{(3)}$-terms when we want to derive that L is a polynomial of degree 2 (or of degree n in the next subsection). First, the $u^\alpha_{(4)}$-terms in (3.120) lead to

$$0 = c_{ij}c_{kl}u^\beta_{ijkl}L_{u^\alpha_{ij}u^\beta_{kl}} \quad \Leftrightarrow \quad 0 = \partial^{(ij}_\alpha \partial^{kl)}_\beta L. \tag{3.125}$$

We define

$$T_{\alpha\beta}(A,A|B,B) := A_iA_jB_kB_l\partial^{ij}_\alpha\partial^{kl}_\beta L = A_iA_jB_kB_lT^{ijkl}_{\alpha\beta}, \quad A,B \in \mathbb{R}^2.$$

Equation (3.125) is equivalent to

$$0 = T_{\alpha\beta}(A,A|A,A) = A_iA_jA_kA_l\partial^{ij}_\alpha\partial^{kl}_\beta L, \quad \text{for all } A \in \mathbb{R}^2.$$

In general, it is not true that

$$0 = T_{\alpha\beta}(A,A|B,B), \quad \text{for all } A,B \in \mathbb{R}^2,$$

since this would mean that all 2-combinations of second order partial derivatives must vanish, which is not the case. Second, the $u^\alpha_{(3)}u^\beta_{(3)}$-terms in (3.120) lead to

$$0 = c_{ij}c_{kl}c_{pq}u^\beta_{kli}u^\gamma_{jpq}L_{u^\beta_{kl}u^\alpha_{ij}u^\gamma_{pq}} \quad \Leftrightarrow \quad 0 = \partial^{(kl}_\beta\partial^{i)(j}_\alpha\partial^{pq)}_\gamma L. \tag{3.126}$$

The equivalence in (3.126) will be proven in Subsection 3.7.3 (we have summation over β,γ and we have to check that we indeed get the equivalence in (3.126)). We define the extended form

$$T_{\beta\alpha\gamma}(A,A|A,B|B,B) := A_kA_lA_iB_jB_pB_q\partial^{kl}_\beta\partial^{ij}_\alpha\partial^{pq}_\gamma L,$$

where (A,A,A) and (B,B,B) can be considered as the transformations of u^β_{kli} and u^γ_{jpq} (see Gelfand Dikii transformation (GD75, OS78, AP95)). If we would know that

$$\begin{aligned} 0 = T_{\beta\alpha\gamma}(A,A|C,C|B,B) &:= \\ &:= A_kA_lC_iC_jB_pB_q\partial^{kl}_\beta\partial^{ij}_\alpha\partial^{pq}_\gamma L \quad \text{for all } A,B,C \in \mathbb{R}^2, \end{aligned}$$

then all 3-combinations of second order partial derivatives would vanish. However, we only know

$$0 = T_{\beta\alpha\gamma}(A,A|A,B|B,B) \quad \text{for all } A,B \in \mathbb{R}^2,$$

since this is equivalent to (3.126). We have the symmetry

$$T_{\beta\alpha\gamma}(A{,}A|A{,}B|B{,}B) = A_k A_l A_i B_j B_p B_q \partial_\beta^{kl}\partial_\alpha^{ij}\partial_\gamma^{pq} L = \\ = A_k A_l B_j A_i B_p B_q \partial_\beta^{kl}\partial_\alpha^{ij}\partial_\gamma^{pq} L = T_{\beta\alpha\gamma}(A{,}A|B{,}A|B{,}B),$$

since ∂_α^{ij} is symmetric in $i{,}j$. For almost all randomly chosen $A{,}B{,}C \in \mathbb{R}^2$ we get that $A{,}B$ are linearly independent and that we can write

$$C = \lambda_1 A + \lambda_2 B$$

for some $\lambda_1{,}\lambda_2 \in \mathbb{R}$. Then we can write

$$\begin{aligned} T_{\beta\alpha\gamma}(A{,}A|C{,}C|B{,}B) &= T_{\beta\alpha\gamma}(A{,}A|\lambda_1 A + \lambda_2 B{,}\lambda_1 A + \lambda_2 B|B{,}B) = \\ &= \lambda_1^2 T_{\beta\alpha\gamma}(A{,}A|A{,}A|B{,}B) + \lambda_1\lambda_2 \underbrace{T_{\beta\alpha\gamma}(A{,}A|A{,}B|B{,}B)}_{=0} + \\ &+ \lambda_1\lambda_2 \underbrace{T_{\beta\alpha\gamma}(A{,}A|B{,}A|B{,}B)}_{=0} + \lambda_2^2 T_{\beta\alpha\gamma}(A{,}A|B{,}B|B{,}B) = \\ &= \lambda_1^2 T_{\beta\alpha\gamma}(A{,}A|A{,}A|B{,}B) + \lambda_2^2 T_{\beta\alpha\gamma}(A{,}A|B{,}B|B{,}B). \end{aligned} \tag{3.127}$$

Since we can extend the form $T_{\beta\alpha}$ and $T_{\alpha\gamma}$, for example,

$$\begin{aligned} B_p B_q \partial_\gamma^{pq} \underbrace{T_{\beta\alpha}(A{,}A|A{,}A)}_{=0} &= B_p B_q A_i A_j A_k A_l \partial_\gamma^{pq}\partial_\beta^{ij}\partial_\alpha^{kl} L = \\ &= T_{\beta\alpha\gamma}(A{,}A|A{,}A|B{,}B) = 0, \end{aligned}$$

we know that (3.127) must vanish for almost all $A{,}B{,}C \in \mathbb{R}^2$. By a density argument we get

$$T_{\alpha\beta\gamma}(A{,}A|C{,}C|B{,}B) = 0 \quad \text{for all } A{,}B{,}C \in \mathbb{R}^2.$$

This means that all 3-combinations of second order partial derivatives must vanish, and therefore L must be a polynomial of degree 2 in second order coordinates.

In the next subsection, we will do this calculation for second order trivial Lagrangians and arbitrary $n{,}m$, which is also the proof of Step 2 from Section 3.6. We will also show the equivalence in (3.126).[7]

3.7.3. Proof of Step 2

In this subsection, we want to solve the equations (3.35)-(3.39) and (3.57)-(3.66) for arbitrary n, i.e. we want to solve the equations (3.120) for arbitrary n and to

[7]It would also be interesting to find out what kind of expressions we would get, when solving only the equations from the $u^\alpha_{(3)} u^\beta_{(3)}$-terms in (3.120).

a certain order. The ideas how to solve this problem have been discussed at the beginning of this section. The equations (3.35)-(3.39) and (3.57)-(3.66) are a result of order discussion in the equations $H_{\alpha\beta} = 0$ or $\mathcal{E}_\alpha(Q^\beta f_\beta) = 0$ (or $H_{\alpha\beta} = O(2)$), when applying operators $D_{ij}\partial_{u^\alpha_{ij}}$ on the functions f_γ or L. Let us consider the equation $H_{\alpha\beta} = 0$, where $H_{\alpha\beta}$ depends on f_γ. However, in the following, we will write L instead of f_γ, since then we do not have to write the index γ and the relation to the equation $\mathcal{E}_\alpha(Q^\beta f_\beta) = \mathcal{E}_\alpha L = 0$ is also given. We consider the $u^\beta_{(4)}$- and $u^\beta_{(3)}u^\gamma_{(3)}$-terms in these equations, i.e.

$$0 = c_{ij}c_{kl}u^\beta_{klij}L_{u^\alpha_{ij}u^\beta_{kl}} \quad \Leftrightarrow \quad 0 = \partial^{(ij}_\alpha \partial^{kl)}_\beta L \tag{3.128}$$

and

$$0 = c_{ij}c_{kl}c_{pq}u^\beta_{kli}u^\gamma_{jpq}L_{u^\beta_{kl}u^\alpha_{ij}u^\gamma_{pq}} \quad \Leftrightarrow \quad 0 = \partial^{(kl}_\beta \partial^{i)(j}_\alpha \partial^{pq)}_\gamma L. \tag{3.129}$$

The equivalence in (3.128) is clear, but we have not yet shown the equivalence in (3.129) (roughly speaking, summation over β,γ does not affect the expressions). We have to distinguish the two cases:

First Case (only one term): Let $\beta = \gamma$ and without loss of generality $\beta = \gamma = 1$. Then the following equation for L has to be satisfied separately:

$$0 = c_{ij}c_{kl}c_{pq}u^1_{kli}u^1_{jpq}L_{u^1_{kl}u^\alpha_{ij}u^1_{pq}} \quad \Leftrightarrow \quad 0 = \partial^{(kl}_1 \partial^{i)(j}_\alpha \partial^{pq)}_1 L$$

and this is equivalent to (3.129) for $\beta = \gamma = 1$.

Second Case (two terms): Let $\beta \neq \gamma$ and without loss of generality $\beta = 1$, $\gamma = 2$ and $\beta = 2$, $\gamma = 1$. Then, the following equation has to be satisfied separately (it is possible that this equation splits up once more, this is the question here):

$$0 = c_{ij}c_{kl}c_{pq}u^1_{kli}u^2_{jpq}L_{u^1_{kl}u^\alpha_{ij}u^2_{pq}} + c_{ij}c_{kl}c_{pq}u^2_{kli}u^1_{jpq}L_{u^2_{kl}u^\alpha_{ij}u^1_{pq}}. \tag{3.130}$$

In the second term (3.130) we can change the following indices

$$i \leftrightarrow j, \quad p \leftrightarrow l, \quad k \leftrightarrow q$$

and we get

$$\begin{aligned} 0 &= c_{ij}c_{kl}c_{pq}u^1_{kli}u^2_{jpq}L_{u^1_{kl}u^\alpha_{ij}u^2_{pq}} + c_{ij}c_{kl}c_{pq}u^2_{qpj}u^1_{ilk}L_{u^2_{qp}u^\alpha_{ji}u^1_{lk}} = \\ &= c_{ij}c_{kl}c_{pq}u^1_{kli}u^2_{jpq}(L_{u^1_{kl}u^\alpha_{ij}u^2_{pq}} + L_{u^2_{qp}u^\alpha_{ji}u^1_{lk}}) = \\ &= 2c_{ij}c_{kl}c_{pq}u^1_{kli}u^2_{jpq}L_{u^1_{kl}u^\alpha_{ij}u^2_{pq}} \end{aligned} \tag{3.131}$$

$$\Leftrightarrow \quad 0 = \partial^{(kl}_1 \partial^{i)(j}_\alpha \partial^{pq)}_2 L,$$

i.e. this equation does not split into two conditions. That is, we get

$$0 = c_{ij}c_{kl}c_{pq}u^{\beta}_{kli}u^{\gamma}_{jpq}L_{u^{\beta}_{kl}u^{\alpha}_{ij}u^{\gamma}_{pq}} \quad \Leftrightarrow \quad 0 = \begin{cases} \partial^{(kl}_{\beta}\partial^{i)(j}_{\alpha}\partial^{pq)}_{\gamma}L, & \beta = \gamma, \\ 2\partial^{(kl}_{\beta}\partial^{i)(j}_{\alpha}\partial^{pq)}_{\gamma}L, & \beta \neq \gamma. \end{cases}$$

We define the form

$$T_{\alpha\beta}(A^1,A^1|A^2,A^2) := A^1_i A^1_j A^2_k A^2_l \partial^{ij}_{\alpha}\partial^{kl}_{\beta}L$$

and we know that

$$0 = T_{\alpha\beta}(A^1,A^1|A^1,A^1) \quad \text{for all } A^1 \in \mathbb{R}^n, \tag{3.132}$$

since this is equivalent to (3.128). Furthermore, we define the form

$$T_{\beta\alpha\gamma}(A^1,A^1|A^2,A^2|A^3,A^3) := A^1_k A^1_l A^2_i A^2_j A^3_p A^3_q \partial^{kl}_{\beta}\partial^{ij}_{\alpha}\partial^{pq}_{\gamma}L \quad \text{for all } A^1,A^2,A^3 \in \mathbb{R}^n$$

and we know that

$$0 = T_{\beta\alpha\gamma}(A^1,A^1|A^1,A^2|A^2,A^2) \quad \text{for all } A^1,A^2 \in \mathbb{R}^n. \tag{3.133}$$

Equations (3.132) and (3.133) are equivalent to (3.128) and (3.129). The extended form is defined as

$$\begin{aligned} &T_{\alpha_{n+1}\alpha_n\ldots\alpha_1}(A^{n+1},A^{n+1}|\ldots|A^3,A^3|A^2,A^2|A^1,A^1) := \\ =&A^{n+1}_{i_{n+1}}A^{n+1}_{j_{n+1}}\partial^{i_{n+1}j_{n+1}}_{\alpha_{n+1}}\ldots A^1_{i_1}A^1_{j_1}\partial^{i_1j_1}_{\alpha_1}L, \quad \text{for all } A^{n+1},A^n,\ldots,A^1 \in \mathbb{R}^n. \end{aligned} \tag{3.134}$$

Note that we also define the form in (3.134) when the pairs of vectors in $\ldots|A^k,A^k|\ldots$ are not equal, i.e. when we have entries of the form $\ldots|A^k,B^k|\ldots$ with $A^k,B^k \in \mathbb{R}^n$ and $A^k \neq B^k$. The forms $T_{\alpha\beta}$, $T_{\beta\alpha\gamma}$ and $T_{\alpha_{n+1}\alpha_n\ldots\alpha_1}$ satisfy the symmetry condition

$$T_{\alpha_{n+1}\ldots\alpha_k\ldots\alpha_l\ldots\alpha_1}(\ldots|A^k,A^k|\ldots|A^l,A^l|\ldots) = T_{\alpha_{n+1}\ldots\alpha_l\ldots\alpha_k\ldots\alpha_1}(\ldots|A^l,A^l|\ldots|A^k,A^k|\ldots), \tag{3.135}$$

since partial derivatives commute. Moreover, we have the symmetry

$$T_{\alpha_{n+1}\ldots\alpha_k\ldots\alpha_1}(\ldots|A^k,B^k|\ldots) = T_{\alpha_{n+1}\ldots\alpha_k\ldots\alpha_1}(\ldots|B^k,A^k|\ldots). \tag{3.136}$$

Note that the condition (3.135) is actually sufficient for the discussion below and we do not necessarily need the condition (3.136), because with (3.135) we can always bring the pairs A^k,A^k on the side where we want them (this will be needed in (3.138) below). The extended form $T_{\alpha_{n+1}\alpha_n\ldots\alpha_3\alpha_2\alpha_1}$ is generated by $T_{\beta\alpha\gamma}$ and $T_{\alpha\beta}$, i.e.

$$\begin{aligned} &T_{\alpha_{n+1}\alpha_n\ldots\alpha_3\alpha_2\alpha_1}(A^{n+1},A^{n+1}|\ldots|A^3,A^3|A^2,A^2|A^1,A^1) = \\ =&A^{n+1}_{i_{n+1}}A^{n+1}_{j_{n+1}}\partial^{i_{n+1}j_{n+1}}_{\alpha_{n+1}}\ldots A^4_{i_4}A^4_{j_4}\partial^{i_4j_4}_{\alpha_4}T_{\alpha_3\alpha_2\alpha_1}(A^3,A^3|A^2,A^2|A^1,A^1) = \\ =&A^{n+1}_{i_{n+1}}A^{n+1}_{j_{n+1}}\partial^{i_{n+1}j_{n+1}}_{\alpha_{n+1}}\ldots A^3_{i_3}A^3_{j_3}\partial^{i_3j_3}_{\alpha_3}T_{\alpha_2\alpha_1}(A^2,A^2|A^1,A^1). \end{aligned}$$

Since $T_{\alpha_{n+1}\alpha_n\ldots\alpha_3\alpha_2\alpha_1}$ is generated in this way, we know that

$$0 = T_{\alpha_{n+1}\alpha_n\ldots\alpha_1}(A^{n+1},A^{n+1}|\ldots|A^3,A^3|A^3,A^2|A^2,A^2)$$

for all $A^{n+1},A^n,\ldots,A^2 \in \mathbb{R}^n$, because of (3.133). Moreover, we know that

$$0 = T_{\alpha_{n+1}\alpha_n\ldots\alpha_1}(A^{n+1},A^{n+1}|\ldots|A^3,A^3|A^2,A^2|A^1,A^1),$$

whenever two of the $A^{n+1},A^n,\ldots,A^1 \in \mathbb{R}^n$ are the same, because of (3.132) and by the symmetry of commuting partial derivatives (3.135). If we would know that

$$0 = T_{\alpha_{n+1}\alpha_n\ldots\alpha_3\alpha_2\alpha_1}(A^{n+1},A^{n+1}|A^n,A^n|\ldots|A^3,A^3|A^2,A^2|A^1,A^1) \tag{3.137}$$

for all $A^{n+1},A^n,\ldots,A^1 \in \mathbb{R}^n$, then all $(n+1)$-combinations of second order derivatives applied to L must vanish, and therefore L must be polynomial of degree n in second order jet coordinates. We will show this now. The vectors

$$A^{n+1},A^n,\ldots,A^2,A^1 \in \mathbb{R}^n$$

are always linearly dependent and we can almost always write

$$A^2 = \sum_{\substack{i=1\\ i\neq 2}}^{n+1} \lambda_i A^i.$$

Then by multi-linearity, we get

$$\begin{aligned} &T_{\alpha_{n+1}\alpha_n\ldots\alpha_3\alpha_2\alpha_1}(A^{n+1},A^{n+1}|\ldots|A^3,A^3|A^2,A^2|A^1,A^1) = \\ =&\sum_{\substack{i=1\\ i\neq 2}}^{n+1}\sum_{\substack{j=1\\ j\neq 2}}^{n+1} \lambda_i\lambda_j T_{\alpha_{n+1}\alpha_n\ldots\alpha_3\alpha_2\alpha_1}(A^{n+1},A^{n+1}|\ldots|A^3,A^3|A^i,A^j|A^1,A^1) = 0, \end{aligned} \tag{3.138}$$

since $T_{\alpha_{n+1}\alpha_n\ldots\alpha_3\alpha_2\alpha_1}$ is generated by $T_{\alpha\beta}$ and $T_{\beta\alpha\gamma}$ and either (3.132) or (3.133) is satisfied. By a density argument we get that the equation (3.137) must always be satisfied for all $A^{n+1},A^n,\ldots,A^1 \in \mathbb{R}^n$, even when we cannot write A^2 as such a linear combination. Therefore, we get that L must be a polynomial of degree $\leq n$ in second order jet coordinates. $\square$

3.8. Helmholtz Dependencies and 4th Order Source Forms

In this section, we prove Theorem 1.0.3, which is again formulated as:

Theorem 3.8.1. *Let $\pi : E \to M$ be a fiber bundle of fiber dimension one and base dimension one. Furthermore, let $\Delta = f du \wedge dx$ be a 4-th order source form defined on J^4E. Assume:*
i) The set $\mathcal{V}$ of symmetries of Δ satisfies (1.2).
ii) Each $V \in \mathcal{V}$ generates a conservation law of the from $Q_V f = D_x C_V$, where $Q_V = V^u - u_x V^x$ are the characteristics.
Then Δ must be locally variational.

Note that Theorem 3.8.1 is no longer true when Δ is only defined on open subsets $\mathcal{R}^4 \subset J^4E$ (see the counter examples in Section 4.2). We conjecture that Theorem 3.8.1 also holds with the same assumption, but arbitrary n. The theorem can be proven by investigation of the Helmholtz dependencies, which are also interesting for other reasons and we start with the discussion of these relations. The investigation of the Helmholtz dependencies for $m = 1$ can also be found in (And89, p.76). However, it seems that this problem has not been investigated extensively in the literature, especially for $m > 1$.

For second order source forms and arbitrary n,m, we have the Helmholtz dependencies

$$\begin{aligned} &H_{\alpha\beta} + H_{\beta\alpha} = D_k H^k_{\alpha\beta} - D_{kl} H^{kl}_{\alpha\beta}, \\ &H^i_{\alpha\beta} - H^i_{\beta\alpha} = 2 D_k H^{ik}_{\alpha\beta}, \\ &H^{ij}_{\alpha\beta} + H^{ij}_{\beta\alpha} = 0. \end{aligned} \tag{3.139}$$

For arbitrary order source forms, $n = 1$ and arbitrary m, the Helmholtz conditions are

$$\partial_{u^\beta_{(j)}} f_\alpha - (-1)^j \partial_{u^\alpha_{(j)}} f_\beta - \sum_{i=j+1} \binom{i}{j} (-1)^i D_x^{i-j} \partial_{u^\alpha_{(i)}} f_\beta = 0, \quad \forall j = 0,1,2,\ldots \forall \alpha,\beta, \tag{3.140}$$

where the sum goes formally to infinity, but is actually finite, since every f_α is assumed to have finite order (we use this notation in this section, since it simplifies some of the calculations). These conditions can be found in (Kru97b, p.56) (also compare the conditions in (2.68) and Lemma 2.6.7). In the following, we consider only the case where n,$m = 1$, but arbitrary order source forms. Then the Helmholtz conditions (3.140) are

$$\sum_{i=1}(-1)^{i+1}\binom{i}{0}D_x^i f_{u_{(i)}} = 0, \quad j=0 \tag{3.141}$$

$$2f_{u_x}+\sum_{i=2}(-1)^{i+1}\binom{i}{1}D_x^{i-1} f_{u_{(i)}} = 0, \quad j=1 \tag{3.142}$$

$$\sum_{i=3}(-1)^{i+1}\binom{i}{2}D_x^{i-2} f_{u_{(i)}} = 0, \quad j=2 \tag{3.143}$$

$$2f_{u_{(3)}}+\sum_{i=4}(-1)^{i+1}\binom{i}{3}D_x^{i-3} f_{u_{(i)}} = 0, \quad j=3 \tag{3.144}$$

$$\vdots$$

Let us call (3.141), (3.143) **odd** and (3.142), (3.144) **even** Helmholtz conditions. For example, for fourth order source forms we get

$$\begin{aligned} H = H^0 =& D_x(f_{u_x} - D_x f_{u_{xx}} + D_x^2 f_{u_{(3)}} - D_x^3 f_{u_{(4)}}) = 0, & (3.145)\\ H^x = H^1 =& 2f_{u_x} - 2D_x f_{u_{xx}} + 3D_x^2 f_{u_{(3)}} - 4D_x^3 f_{u_{(4)}} = 0, & (3.146)\\ H^{xx} = H^2 =& D_x(3f_{u_{(3)}} - 6D_x f_{u_{(4)}}) = 0, & (3.147)\\ H^{(3)} = H^3 =& 2f_{u_{(3)}} - 4D_x f_{u_{(4)}} = 0. & (3.148) \end{aligned}$$

We can easily see that the Helmholtz condition (3.147) is unnecessary, since if (3.148) is satisfied, then (3.147) is automatically satisfied. Let us multiply (3.145) by 2, then we get

$$\begin{aligned} &2D_x(f_{u_x} - D_x f_{u_{xx}} + D_x^2 f_{u_{(3)}} - D_x^3 f_{u_{(4)}}) = \\ =& D_x(2f_{u_x} - 2D_x f_{u_{xx}} + 3D_x^2 f_{u_{(3)}} - 4D_x^3 f_{u_{(4)}}) + D_x^3(-f_{u_{(3)}} + 2D_x f_{u_{(4)}}), \end{aligned}$$

and therefore this conditions is also unnecessary, since if (3.146) and (3.148) are satisfied, then (3.145) is automatically satisfied. Therefore, in this case it is reasonable to define the integrability conditions (Helmholtz conditions) as

$$\begin{aligned} h^0 :=& (f_{u_x} - D_x f_{u_{xx}} + D_x^2 f_{u_{(3)}} - D_x^3 f_{u_{(4)}}) = 0,\\ h^1 :=& (f_{u_{(3)}} - 2D_x f_{u_{(4)}}) = 0, \end{aligned}$$

instead of using the partially redundant conditions (3.145)-(3.148). It turns out that we can write

$$\begin{aligned} H^0 &= D_x h^0,\\ H^1 &= 2h^0 + D_x^2 h^1,\\ H^2 &= 3D_x h^1,\\ H^3 &= 2h^1. \end{aligned} \tag{3.149}$$

More generally, we can write

$$\begin{aligned}
H^0 &= D_x h^0, \\
H^1 &= 2h^0 + D_x^2 h^1, \\
H^2 &= 3D_x h^1 + D_x^3 h^2, \\
H^3 &= 2h^1 + 4D_x^2 h^2 + D_x^4 h^3, \\
H^4 &= 5D_x h^2 + 5D_x^3 h^3 + D_x^5 h^4, \\
H^5 &= 2h^2 + 9D_x^2 h^3 + 6D_x^4 h^4 + D_x^6 h^5,
\end{aligned} \tag{3.150}$$

where h^j is defined in (3.152) below. Finding the dependencies (3.139) is important, however, finding non-dependent conditions is even more interesting. For arbitrary m (and $n = 1$) we also get such dependencies for all $\alpha = \beta$, and therefore the following investigation of such dependencies is also important in more generality. That we can always find such simpler conditions for $n,m = 1$ and arbitrary order is formulated in the following lemma:

Lemma 3.8.2. *Let $n,m = 1$. The dependent Helmholtz conditions*

$$H^j = \left(\partial_{u_{(j)}} - (-1)^j \partial_{u_{(j)}} - \sum_{i=j+1} \binom{i}{j} (-1)^i D_x^{i-j} \partial_{u_{(i)}} \right) f = 0, \quad \forall j = 0,1,2,... \tag{3.151}$$

also follow by the reduced Helmholotz conditions

$$h^j = \left(\sum_{i=j} \binom{i}{j} (-1)^{j+i} D_x^{i-j} \partial_{u_{(i+j+1)}} \right) f = 0, \quad \forall j = 0,1,2,.... \tag{3.152}$$

The proof can be shown by using the following identities (where $i \geq 0$ and $n \geq 0$)

$$\binom{2i+2n+1}{2i} = \sum_{p=0}^{i} \left[\binom{i+1+p}{i-p} + \binom{i+p}{i-(1+p)} \right] \binom{i+2n-p}{i+p}, \tag{3.153}$$

$$\binom{2i+2n+2}{2i} = \sum_{p=0}^{i} \left[\binom{i+1+p}{i-p} + \binom{i+p}{i-(1+p)} \right] \binom{i+2n-p+1}{i+p}, \tag{3.154}$$

$$\binom{2i+2n}{2i-1} = \sum_{p=0}^{i} \left[\binom{i+1+p}{i-p} + \binom{i+p}{i-(1+p)} \right] \binom{i+2n-1-p}{i-1+p}, \quad n \geq 1,$$

$$\binom{2i+2n+1}{2i-1} = \sum_{p=0}^{i} \left[\binom{i+1+p}{i-p} + \binom{i+p}{i-(1+p)} \right] \binom{i+2n-p}{i-1+p}, \tag{3.155}$$

which can be proven inductively. For the induction we will need all four identities at the same time, which makes it rather complicated (numerically, the proof can easily be done for very high numbers, which should be sufficient for practical reasons). Also see (And89, p.77), where such binomial coefficients occur. Note that by definition ($a,b \in \mathbb{Z}_0$)

$$\begin{aligned}
\binom{a}{b} &= 0, \quad \text{if } b < 0, \\
\binom{a}{b} &= 0, \quad \text{if } a < b \text{ and } b > 0, \\
\binom{0}{0} &= 1.
\end{aligned} \tag{3.156}$$

It seems that the Helmholtz conditions and Helmholtz form is not completely understood and we should probably find conditions and a form where we do not have such dependencies (i.e. we should find a certain representative in the equivalence class of the variational sequence). Then Takens' problem could probably be solved much easier. Note that the Helmholtz form (VU13, p.13)

$$H = \frac{1}{2}(H_{\alpha\beta} du^\beta \wedge du^\alpha + H^x_{\alpha\beta} du^\beta_x \wedge du^\alpha + H^{xx}_{\alpha\beta} du^\beta_{xx} \wedge du^\alpha + ...) \wedge dx$$

may have a coefficient $H_{\alpha\alpha}$ which does not vanish, but the term $H_{\alpha\beta} du^\beta \wedge du^\alpha \wedge dx$ vanishes for $\alpha = \beta$, since $du^\alpha \wedge du^\alpha = 0$. Moreover, in the case where $n,m = 1$, the term $H_{\alpha\beta} du^\beta \wedge du^\alpha \wedge dx$ vanishes completely. However, the term $H^{xx} du_{xx} \wedge du \wedge dx$ does not vanish and as we saw above $H^{xx} = H^2 = 3D_x h_1 + D^3_x h^2$. Therefore, we definitely get a redundant condition from this term. More precisely, we get an expression which can be written as an exact form plus a 3-contact form. We will not further discuss the Helmholtz form in this section and we only consider the redundant conditions in 3.151.

<u>Proof of Lemma 3.8.2:</u> Let us rewrite the equations (3.151) and (3.152) as

$$\left(\partial_{u_{(j)}} - (-1)^j \partial_{u_{(j)}} - \sum_{p=0} \binom{j+1+p}{j} (-1)^{j+1+p} D_x^{1+p} \partial_{u_{(j+1+p)}}\right) f = 0, \quad \forall j = 0,1,2,...$$

and

$$h^j = \left(\sum_{p=0} \binom{j+p}{j} (-1)^{2j+p} D_x^p \partial_{u_{(2j+1+p)}}\right) f = 0, \quad \forall j = 0,1,2,.... \tag{3.157}$$

Let us first consider the case where j is even, i.e. $j = 2i$ for some $i \in \mathbb{N}_0$. Then we get

$$\begin{aligned}
&\partial_{u_{(j)}} - (-1)^j \partial_{u_{(j)}} - \sum_{p=0} \binom{j+1+p}{j} (-1)^{j+1+p} D_x^{1+p} \partial_{u_{(j+1+p)}} = \\
=&D_x \sum_{p=0} \binom{2i+1+p}{2i} (-1)^p D_x^p \partial_{u_{(2i+1+p)}}.
\end{aligned} \tag{3.158}$$

The partial derivatives $D_x^p \partial_{u_{(2i+1+p)}}$ in (3.157) and (3.158) have the same structure, but the binomial coefficients are quite different. Let $n \in \mathbb{N}_0$. If $p = 2n+1$ is odd, then we use (3.154), and if $p = 2n$ is even, then we use (3.153) to write

$$
\begin{aligned}
&D_x \sum_{p=0} \binom{2i+1+p}{2i} (-1)^p D_x^p \partial_{u_{(2i+1+p)}} = \\
=&D_x \sum_{n=0} \binom{2i+1+2n}{2i} D_x^{2n} \partial_{u_{(2i+1+2n)}} - D_x \sum_{n=0} \binom{2i+2+2n}{2i} D_x^{2n+1} \partial_{u_{(2i+2n+2)}} = \\
=&D_x \sum_{n=0} \sum_{p=0}^{i} \left[\binom{i+1+p}{i-p} + \binom{i+p}{i-(1+p)} \right] \binom{i+2n-p}{i+p} D_x^{2n} \partial_{u_{(2i+1+2n)}} - \\
&- D_x \sum_{n=0} \sum_{p=0}^{i} \left[\binom{i+1+p}{i-p} + \binom{i+p}{i-(1+p)} \right] \binom{i+2n-p+1}{i+p} D_x^{2n+1} \partial_{u_{(2i+2n+2)}} = \\
=&D_x \sum_{p=0}^{i} \left[\binom{i+1+p}{i-p} + \binom{i+p}{i-(1+p)} \right] \times \\
&\times \sum_{n=0} \left[\binom{i+2n-p}{i+p} D_x^{2n} \partial_{u_{(2i+1+2n)}} - \binom{i+2n-p+1}{i+p} D_x^{2n+1} \partial_{u_{(2i+2n+2)}} \right] = \\
=&D_x \sum_{p=0}^{i} \left[\binom{i+1+p}{i-p} + \binom{i+p}{i-(1+p)} \right] \sum_{k=0} \binom{i+k-p}{i+p} (-1)^k D_x^k \partial_{u_{(2i+k+1)}}.
\end{aligned}
\tag{3.159}
$$

In the second sum over k in the last line in (3.159) we now use the condition

$$
i+k-p \geq i+p \quad \Leftrightarrow \quad k \geq 2p,
$$

since otherwise the binomial coefficient of this sum vanishes, see (3.156). Therefore, we can write (3.159) as ($k = 2p + l$, where $l = 0,1,2,...$)

$$
\begin{aligned}
&= D_x \sum_{p=0}^{i} \left[\binom{i+1+p}{i-p} + \binom{i+p}{i-(1+p)} \right] \sum_{k=0} \binom{i+k-p}{i+p} (-1)^k D_x^k \partial_{u_{(2i+k+1)}} = \\
&= D_x \sum_{p=0}^{i} \left[\binom{i+1+p}{i-p} + \binom{i+p}{i-(1+p)} \right] \underbrace{\sum_{l=0} \binom{i+p+l}{i+p} (-1)^{2p+l} D_x^{2p+l} \partial_{u_{(2i+2p+l+1)}}}_{=D_x^{2p} h^{i+p}} = \\
&= \sum_{p=0}^{i} \left[\binom{i+1+p}{i-p} + \binom{i+p}{i-(1+p)} \right] D_x^{2p+1} h^{i+p} = \qquad \text{note: } \binom{n}{k} = \binom{n}{n-k} \\
&= \sum_{p=0}^{i} \left[\binom{i+1+p}{1+2p} + \binom{i+p}{1+2p} \right] D_x^{2p+1} h^{i+p}.
\end{aligned}
\tag{3.160}
$$

For $j = 2i+1$ odd we can do a similar calculation, when we use the identity (3.155) and the identity above it. □

Let us consider three very simple examples, where we can use the binomial coefficients in the third line in (3.160), when $j = 2i$: For $i = 0$, we get

$$H^0 = \left[\binom{1}{0} + \binom{0}{-1}\right] D_x h^0 = D_x h^0.$$

For $i = 1$, we get

$$H^2 = D_x \left[\binom{2}{1} + \binom{1}{0}\right] h^{1+0} + D_x^3 \left[\binom{3}{0} + \binom{2}{-1}\right] h^{1+1} = 3D_x h^1 + D_x^3 h^2.$$

and for $i = 2$, we get

$$\begin{aligned} H^4 =& D_x \left[\binom{3}{2} + \binom{2}{1}\right] h^{2+0} + D_x^3 \left[\binom{4}{1} + \binom{3}{0}\right] h^{2+1} + D_x^5 \left[\binom{5}{0} + \binom{4}{-1}\right] h^{2+2} = \\ =& 5D_x h^2 + 5D_x^2 h^3 + D_x^5 h^4. \end{aligned}$$

Therefore, for $n = 1$ and arbitrary m, it is reasonable to define the Helmholtz conditions as

$$\boxed{\begin{aligned} H^j_{\alpha\beta} &= \partial_{u^\beta_{(j)}} f_\alpha - (-1)^j \partial_{u^\alpha_{(j)}} f_\beta - \sum_{i=j+1} \binom{i}{j} (-1)^i D_x^{i-j} \partial_{u^\alpha_{(i)}} f_\beta = 0;\ \forall j = 0,\dots;\ \alpha \neq \beta, \\ h^j_\alpha &= \left(\sum_{i=j} \binom{i}{j} (-1)^{i+j} D_x^{i-j} \partial_{u^\alpha_{(i+j+1)}} \right) f_\alpha = 0; \quad \forall j = 0,1,2,\dots; \quad \alpha = \beta, \end{aligned}}$$

but maybe there are still dependencies among these conditions. Now they do not have such a nice form as the usual Helmholtz conditions, but (some of) the dependencies are eliminated, and for $m = 1$, they reduce to the much simpler conditions (3.157). Now we can use the simpler conditions to solve Takens' problem.

3.8.1. The Proof for 4th Order Source Forms and $n,m = 1$

The reduced Helmholtz conditions allow us now to solve Takens' problem more systematically and we do not have to use the Helmholtz dependencies anymore at some stage in the proof. The proof is now more likely solving differential equations than discussing highest order jet coordinates. Solving differential equations in the whole range of definition also means that we are investigating a kind of global version of Takens' problem. Global in the sense that we are investigating the whole fibers of $J^k E$ over E. However, we do not investigate the global structure of E, and therefore the problem is still local in this sense.

Some of the solutions of the differential equations below will have singularities. We will discuss and define at the end of this subsection what we precisely mean with singularity. To define and explain it here would interrupt the flow of our discussion and intuitively it will be clear what we mean with singularity.

Let us consider $n,m = 1$ and 4-th order source forms. The ECS in the standard form is

$$0 = Q^\beta H_{\alpha\beta} + (D_x Q^\beta) H^x_{\alpha\beta} + (D_x^2 Q^\beta) H^{xx}_{\alpha\beta} + (D_x^3 Q^\beta) H^{(3)}_{\alpha\beta} + (D_x^4 Q^\beta) H^{(4)}_{\alpha\beta}. \tag{3.161}$$

Using the notation from Section 3.8, we can write the ECS as (see (3.149) and (3.150))

$$\begin{aligned} 0 =& QH^0 + (D_x Q)H^1 + (D_x^2 Q)H^2 + (D_x^3 Q)H^3 + (D_x^4 Q)H^4 = \\ =& QD_x h^0 + (D_x Q)(2h^0 + D_x^2 h^1) + (D_x^2 Q)(3D_x h^1 + D_x^3 h^2) + \\ &+ (D_x^3 Q)(2h^1 + 4D_x^2 h^2 + D_x^4 h^3) + (D^4 Q)(5D_x h^2 + 5D_x^3 h^3 + D_x^5 h^4). \end{aligned} \tag{3.162}$$

For 4-th order source forms we get immediately that $h^2,h^3,h^4 = 0$, see (3.157), which reduces the number of unknowns essentially. Then the ECS (3.162) can be written as

$$0 = [QD_x h^0 + 2(D_x Q)h^0] + [(D_x Q)D_x^2 h^1 + 3(D_x^2 Q)D_x h^1 + 2(D_x^3 Q)h^1] \tag{3.163}$$

and instead of considering the four unknowns H^0,H^1,H^2,H^3 (H^4 was anyway zero), we only have to consider the two unknowns h^0,h^1, but we have derivatives on these expressions, whereas (3.161) has no derivatives on the unknowns. This allows us now to solve the differential equation for h^0,h^1 and if the general solution is, for example, singular, or has some other properties which are not allowed in our setting, then the Helmholtz conditions must be satisfied. Note that the equation is linear and by multiplying h^0,h^1 with some constant we can always get the non-singular trivial solution.

Now we will discuss a few very interesting aspects of equation (3.163) and after that we prove Theorem 1.0.3 for 4-th order source forms.

Compared to (3.161), equation (3.162) looks very complicated. However, it can be written in a much simpler form. In the following, we will also consider h^2 in the higher order version of (3.162), see (3.150), to understand the general structure of this type of equation. Let us start with the following identities

$$QD_x h^0 + 2(D_x Q)h^0 = \frac{1}{Q} D_x[Q^2 h^0], \tag{3.164}$$

$$(D_x Q)D_x^2 h^1 + 3(D_x^2 Q)D_x h^1 + 2(D_x^3 Q)h^1 = D_x \left(\frac{D_x[(D_x Q)^2 h^1]}{D_x Q} \right), \tag{3.165}$$

$$(D_x^2 Q)D_x^3 h^2 + 4(D_x^3 Q)D_x^2 h^2 + 5(D_x^4 Q)D_x h^2 + 2(D_x^5 Q)h^2 = D_x^2 \left(\frac{D_x[(D_x^2 Q)^2 h^2]}{D_x^2 Q} \right). \tag{3.166}$$

It seems that (this is our conjecture), in general, we can write the ECS as

$$\begin{aligned} 0 &= \sum_{k=0} D_x^k \left(\frac{D_x[(D_x^k Q)^2 h^k]}{D_x^k Q} \right) = \\ &= \frac{1}{Q} D_x[Q^2 h^0] + D_x \left(\frac{D_x[(D_x Q)^2 h^1]}{D_x Q} \right) + D_x^2 \left(\frac{D_x[(D_x^2 Q)^2 h^2]}{D_x^2 Q} \right) + \dots . \end{aligned} \tag{3.167}$$

Note that this is a **formal notation** and there are no singularities in this equation, even if it seems that we dividing by $D_x^k Q$. The counter in the fraction has terms

$$\begin{aligned} D_x[(D_x^k Q)^2 h^k] &= 2(D_x^k Q)(D_x^{k+1} Q)h^k + (D_x^k Q)^2 D_x h^k = \\ &= (D_x^k Q)[2(D_x^{k+1} Q)h^k + (D_x^k Q)D_x h^k], \end{aligned}$$

and therefore it can always be divided by $(D_x^k Q)$, without producing singularities. It is very helpful to use this formal notation, since it explains the structure, whereas (3.163) has no structure. To prove the conjecture for arbitrary order, we have to understand the coefficients in (3.160). Hopefully we can solve this problem in the future and since we are now mostly investigating 4th order source forms, we are not confronted with that problem.

Investigating (3.167) immediately provides a few interesting results. The first, second, third,... summands in the sum (3.167) have separately solutions of the following form:

$$\begin{aligned} 0 &= \frac{1}{Q} D_x[Q^2 h^0] && \Rightarrow h^0 = \frac{c_0}{Q^2}, \quad c_0 \in \mathbb{R}, \\ 0 &= D_x \left(\frac{D_x[(D_x Q)^2 h^1]}{D_x Q} \right) && \Rightarrow h^1 = \frac{c_1 + Q d_1}{(D_x Q)^2}, \quad c_1, d_1 \in \mathbb{R}, \\ 0 &= D_x^2 \left(\frac{D_x[(D_x^2 Q)^2 h^2]}{D_x^2 Q} \right) && \Rightarrow h^2 = \frac{c_2 + (D_x Q)d_2 + [x(D_x Q) - Q]e_2}{(D_x^2 Q)^2}, \; c_2, d_2, e_2 \in \mathbb{R} \end{aligned} \tag{3.168}$$

and all of these solutions are singular, if $V^x \neq 0$, i.e. if we have a symmetry $V \in \mathcal{V}$ and corresponding conservation law such that $\pi_* V \neq 0$. The coupled system (3.167) for h^0,h^1,h^2,... has also non-singular solutions, for example,

$$h^0 = D_x^2 Q, \quad h^1 = -\frac{1}{2}Q, \quad h^2,h^3,... = 0.$$

This solution corresponds to translation symmetry ∂_x, corresponding characteristic $Q = -u_x$, source form

$$\Delta = f du \wedge dx = (\frac{1}{2}u_x u_{(3)} + u_{xx}^2) du \wedge dx$$

and conservation law

$$Qf = -\frac{1}{2}u_x^2 u_{(3)} - u_x u_{xx}^2 = -D_x(\frac{1}{2}u_x^2 u_{xx}).$$

The source form is not variational. This source form has also ∂_u-symmetry, i.e. translations in u-direction, but not corresponding conservation law. Therefore, Takens' question cannot be answered affirmatively for n,$m = 1$ and third order source forms if we only assume one symmetry $\pi_* V \neq 0$. However, maybe for two symmetries which span T_pE at every $p \in E$. Indeed, we can prove the existence of a variational formulation in this case and we will do that below. Therefore, we want to investigate the solutions of the coupled system in more detail.

The summands in (3.167) are very special: When we multiply them by Q, then they can (again) be written as total derivatives. For $k = 0$ this is clear. For $k = 1$ we apply the partial integration technique and we get

$$\begin{aligned} QD_x\left(\frac{D_x[(D_xQ)^2 h^1]}{D_xQ}\right) &= D_x\left(Q\frac{D_x[(D_xQ)^2 h^1]}{D_xQ}\right) - (D_xQ)\frac{D_x[(D_xQ)^2 h^1]}{D_xQ} = \\ &= D_x \underbrace{\left[Q\frac{D_x[(D_xQ)^2 h^1]}{D_xQ} - (D_xQ)^2 h^1\right]}_{=:g^1} = D_x g^1 \qquad (3.169) \end{aligned}$$

For general k we can also apply a sort of partial integration technique and we get inductively (we use the short notation $D_xQ = Q_{;x}$, $D_x^2Q = Q_{;xx}$ and $D_x^kQ = Q_{(k)}$)

$$\begin{aligned}
&QD_x^k\left(\frac{D_x[Q_{(k)}^2 h^k]}{Q_{(k)}}\right) = \\
=&D_x\left[QD_x^{k-1}\left(\frac{D_x[Q_{(k)}^2 h^k]}{Q_{(k)}}\right)\right] - Q_{;x}D_x^{k-1}\left(\frac{D_x[Q_{(k)}^2 h^k]}{Q_{(k)}}\right) = \\
=&D_x[...] - D_x\left[Q_{;x}D_x^{k-2}\left(\frac{D_x[Q_{(k)}^2 h^k]}{Q_{(k)}}\right)\right] + Q_{;xx}D_x^{k-2}\left(\frac{D_x[Q_{(k)}^2 h^k]}{Q_{(k)}}\right) = \\
=&... = \\
=&D_x[...] + (-1)^l Q_{(l)}D_x^{k-l}\left(\frac{D_x[Q_{(k)}^2 h^k]}{Q_{(k)}}\right) = , \quad \text{(we continue until } l = k) \\
=&D_x[...] + (-1)^k Q_{(k)}\left(\frac{D_x[Q_{(k)}^2 h^k]}{Q_{(k)}}\right) = \\
=&D_x \underbrace{[... + (-1)^k Q_{(k)}^2 h^k]}_{=:g^k} = D_x g^k.
\end{aligned}$$

This means that the ECS multiplied by Q is a total derivative and we can use this property to reduce the order of the differential equation. In other words, Q is an integrating factor for this differential equation.

Remark: Let us explain why the more general ECS expression multiplied by Q^α, i.e.

$$Q^\alpha Q^\beta H_{\alpha\beta} + Q^\alpha(D_iQ^\beta)H^i_{\alpha\beta} + Q^\alpha(D_{ij}Q^\beta)H^{ij}_{\alpha\beta} + ...,$$

is also a total derivative (or divergence) for arbitrary n,m. At least for first order source forms (we use the short notation $D_iQ^\alpha = Q^\alpha_{;i}$)

$$\begin{aligned}
Q^\alpha Q^\beta H_{\alpha\beta} + Q^\alpha Q^\beta_{;i}H^i_{\alpha\beta} &= \underbrace{Q^\alpha Q^\beta}_{\text{sym.}}(\underbrace{f_{\alpha,u^\beta} - f_{\beta,u^\alpha}}_{\text{skew-sym. in } \alpha,\beta} + D_i f_{\beta,u_i^\alpha}) + Q^\alpha Q^\beta_{;i}(f_{\alpha,u_i^\beta} + f_{\beta,u_i^\alpha}) = \\
&= Q^\alpha Q^\beta D_i f_{\beta,u_i^\alpha} + Q^\alpha Q^\beta_{;i}(f_{\alpha,u_i^\beta} + f_{\beta,u_i^\alpha}) = \\
&= D_i(Q^\alpha Q^\beta f_{\beta,u_i^\alpha}) \qquad (3.170)
\end{aligned}$$

we can immediately observe that $Q^\alpha Q^\beta H_{\alpha\beta} + Q^\alpha Q^\beta_{;i}H^i_{\alpha\beta}$ must be a divergence expression. When this expression vanishes, then it must be a trivial divergence with certain properties, see Subsection 3.7.1. This immediately forces strong restrictions for possible differential equations f_β. It could be an interesting open problem to investigate this structure in more detail. To understand why Q^α is an integrating factor for higher order ECS's, see the calculation in (2.67), which shows that for $V = W$ the expression $\mathcal{L}_{\mathrm{pr}V}(\iota_{\mathrm{pr}W}\Delta) - \iota_{\mathrm{pr}V}\mathcal{L}_{\mathrm{pr}W}\Delta$ must always vanish (also for non-variational source forms). The question is if the method of integrating factor can only be used for $n,m = 1$ or if it is also useful for $m,n > 1$. Note that for $m > 1$, the (system) ECS is transformed to a single equation, when multiplied by Q^α and summing over α. Therefore, the meaning of integrating factor is slightly different there and this method is possibly only useful for $m = 1$, but arbitrary n. Now let us continue with the case $n,m = 1$.

For simplicity, let us discuss equation (3.167) for $k = 0,1,2$, i.e.

$$0 = \frac{1}{Q}D_x[Q^2h^0] + D_x\left(\frac{D_x[(D_xQ)^2h^1]}{D_xQ}\right) + D_x^2\left(\frac{D_x[(D_x^2Q)^2h^2]}{D_x^2Q}\right). \tag{3.171}$$

We multiply (3.171) by Q and we get

$$0 = D_x(Q^2h^0) + D_x(g^1 + g^2).$$

This leads to

$$c_1 = Q^2h^0 + g^1 + g^2, \quad c_1 \in \mathbb{R}. \tag{3.172}$$

More generally, and as we mentioned above, it seems that we get an equation of the form $c_1 = Q^2h^0 + g^1 + g^2 + ... + g^k$, if the ECS in (3.167) can be written in such a way.

The next step is to solve equation (3.172). Therefore, we do the following trick: We multiply (3.172) by $\frac{D_xQ}{Q^2}$ and we can reduce the order of all terms, except the h^0-term, once again. We conjecture that this also works for any order and the more general expression $c_1 = Q^2h^0 + g^1 + g^2 + g^3 + ... + g^k$. Note that multiplying by $Q^2\frac{D_xQ}{Q^2} = D_xQ$ is in some sense equivalent to what we will consider later. But for technical reasons it is better to multiply by $\frac{D_xQ}{Q^2}$, which slightly simplifies the calculations below. Also note that multiplying by $\frac{D_xQ}{Q^2}$ seems to make sense only when $D_xQ \neq 0$, otherwise we get a trivial equation. For the moment, we ignore this aspect. For example, when we consider translation symmetries ∂_x,∂_u, where the characteristic Q for the ∂_u-symmetry is a constant, we get $D_xQ = 0$. Before we consider all terms in equation (3.172) together, let us separately compute

$$\begin{aligned}\frac{D_xQ}{Q^2}g^1 &= \frac{D_xQ}{Q^2}\left[Q\frac{D_x[(D_xQ)^2h^1]}{D_xQ} - (D_xQ)^2h^1\right] = \\ &= \frac{D_x[(D_xQ)^2h^1]}{Q} - \frac{(D_xQ)}{Q^2}(D_xQ)^2h^1 = \\ &= D_x\left(\frac{1}{Q}(D_xQ)^2h^1\right)\end{aligned}$$

and

$$\begin{aligned}
\frac{D_xQ}{Q^2}g^2 &= \frac{D_xQ}{Q^2}\left[QD_x\left(\frac{D_x[(D_x^2Q)^2h^2]}{D_x^2Q}\right)-(D_xQ)\frac{D_x[(D_x^2Q)^2h^2]}{D_x^2Q}+(D_x^2Q)^2h^2\right]=\\
&=\frac{D_xQ}{Q}D_x\left(\frac{D_x[(D_x^2Q)^2h^2]}{D_x^2Q}\right)-\frac{(D_xQ)^2}{Q^2}\frac{D_x[(D_x^2Q)^2h^2]}{D_x^2Q}+\frac{D_xQ}{Q^2}(D_x^2Q)^2h^2=\\
&=D_x\left(\frac{D_xQ}{Q}\frac{D_x[(D_x^2Q)^2h^2]}{D_x^2Q}\right)-\left(\frac{D_x^2Q}{Q}-\frac{(D_xQ)^2}{Q^2}\right)\frac{D_x[(D_x^2Q)^2h^2]}{D_x^2Q}-\\
&\quad-\frac{(D_xQ)^2}{Q^2}\frac{D_x[(D_x^2Q)^2h^2]}{D_x^2Q}+\frac{D_xQ}{Q^2}(D_x^2Q)^2h^2=\\
&=D_x\left(\frac{D_xQ}{Q}\frac{D_x[(D_x^2Q)^2h^2]}{D_x^2Q}\right)-\frac{D_x[(D_x^2Q)^2h^2]}{Q}+\frac{D_xQ}{Q^2}(D_x^2Q)^2h^2=\\
&=D_x\left(\frac{D_xQ}{Q}\frac{D_x[(D_x^2Q)^2h^2]}{D_x^2Q}-\frac{(D_x^2Q)^2h^2}{Q}\right).
\end{aligned}$$

That is, both of these expressions are total derivatives. Therefore, using these identities, and multiplying (3.172) by $\frac{D_xQ}{Q^2}$, we get

$$\frac{D_xQ}{Q^2}c_1=(D_xQ)h^0+D_x\left[\frac{(D_xQ)^2h^1}{Q}+\left(\frac{D_xQ}{Q}\frac{D_x[(D_x^2Q)^2h^2]}{D_x^2Q}-\frac{(D_x^2Q)^2h^2}{Q}\right)\right]. \tag{3.173}$$

The left hand side of (3.173) is also a total derivative, since

$$\frac{D_xQ}{Q^2}c_1=D_x\frac{-c_1}{Q}.$$

Then we bring this expression to the right hand side of (3.173) and we get

$$0=(D_xQ)h^0+D_x\left[\left(\frac{c_1}{Q}+\frac{(D_xQ)^2h^1}{Q}\right)+\left(\frac{D_xQ}{Q}\frac{D_x[(D_x^2Q)^2h^2]}{D_x^2Q}-\frac{(D_x^2Q)^2h^2}{Q}\right)\right]. \tag{3.174}$$

Now we assume 4th order source forms, and therefore $h^2=0$. Above, we just wanted to show that our reformulations work more generally. As we already mentioned, instead of multiplying by $\frac{D_xQ}{Q^2}$, we can also multiply (3.172) by D_xQ, and then we get

$$0=Q^2\left[(D_xQ)h^0+D_x\left(\frac{c_1}{Q}+\frac{(D_xQ)^2h^1}{Q}\right)\right]. \tag{3.175}$$

For technical reasons it is better to write (3.175) instead of (3.174), since otherwise we get singularities of the form $\frac{1}{Q^2}$.

Since one projectable symmetry $V = V^x\partial_x + V^u\partial_u \in \mathfrak{X}(E)$, where $V^x \neq 0$ is not sufficient to solve Takens' problem for 3rd or 4th order source forms, as we saw above, we now assume that we have two projectable symmetries $V_1,V_2 \in \mathfrak{X}(E)$ such that $\mathrm{span}\{V_{1p},V_{2p}\} = T_pE$ at each $p \in E$. For every symmetry vector field V_1,V_2 we get an equation of the form (3.175), i.e.

$$0 = Q_1^2\left[(D_xQ_1)h^0 + D_x\left(\frac{c_1}{Q_1} + \frac{(D_xQ_1)^2h^1}{Q_1}\right)\right], \tag{3.176}$$

$$0 = Q_2^2\left[(D_xQ_2)h^0 + D_x\left(\frac{c_2}{Q_2} + \frac{(D_xQ_2)^2h^1}{Q_2}\right)\right], \tag{3.177}$$

where $Q_1 = V_1^u - u_xV_1^x$ and $Q_2 = V_2^u - u_xV_2^x$. It is easy to see that we can eliminate h^0 with the help of (3.176) and (3.177) and then we can solve the remaining equation for h^1. However, as we mentioned above, in the case where $D_xQ_1 = 0$ or $D_xQ_2 = 0$, one of these equations is trivial (or even both). Therefore, in the case where $D_xQ_1 = 0$ or $D_xQ_2 = 0$, we rather use the original equation (3.171), i.e. the two equations

$$0 = \frac{1}{Q_1}D_x[Q_1^2h^0] + D_x\left(\frac{D_x[(D_xQ_1)^2h^1]}{D_xQ_1}\right), \tag{3.178}$$

$$0 = \frac{1}{Q_2}D_x[Q_2^2h^0] + D_x\left(\frac{D_x[(D_xQ_2)^2h^1]}{D_xQ_2}\right). \tag{3.179}$$

<u>Short conclusion of what we have found out so far:</u> Multiplying (3.178) and (3.179) by the integrating factor Q and by the factor D_xQ to derive (3.176) and (3.177) only makes sense if $D_xQ \neq 0$, otherwise we get a trivial equation (which is definitely the case for translation symmetries ∂_x,∂_u). Equations (3.178) and (3.179) are one order higher compared to (3.176) and (3.177), where the two constants c_1,c_2 already occur, because we integrated the equations one time. The above derivation of (3.176) and (3.177) shows how to reformulate the equations (3.178) and (3.179) to be able to solve them later. Furthermore, the ECS for $n,m = 1$ seems to have a very nice structure and we can assume that, more generally, multiplying by Q, then by D_xQ, and similar factors can help to find the general solution of the ECS. However, we do not want to consider the two different kinds of systems (3.176), (3.177) and (3.178), (3.179) simultaneously. These different kinds of systems occur, because we have to distinguish the two cases $D_xQ = 0$ and $D_xQ \neq 0$ and similar problems may also occur for higher order (maybe when $D_x^2Q = 0$ and so on). Deriving the equation (3.176) (or (3.177)) is important when we want to solve the ECS for a single symmetry vector field (or characteristic Q). But in the following, we want to solve the ECS for two or more symmetries such that (1.2) is satisfied and we can immediately use the condition (1.2) to simplify the problem.

Therefore, now we will start with (3.178) and (3.179) once again, but derive an equivalent system somehow differently, such that we do not have a problem when $D_xQ_1 = 0$ or $D_xQ_2 = 0$, and we can immediately use the assumption (1.2), since we want to prove Theorem 1.0.3 below. We can do this relatively fast, since we already explained the general structure of the ECS (3.167).

Let us multiply (3.178) and (3.179) by Q_1 and Q_2. It could be possible that Q_1 or Q_2 vanish somewhere, but this will not effect our calculations below (only $D_xQ = 0$ would be a problem later). Then we use the identity (3.169) and the short notation $D_xQ = Q_{;x}$ to derive the two equations

$$0 = D_x\left[Q_1^2 h^0 + \left(Q_1 \frac{D_x(Q_{1;x}^2 h^1)}{Q_{1;x}} - Q_{1;x}^2 h^1\right)\right], \tag{3.180}$$

$$0 = D_x\left[Q_2^2 h^0 + \left(Q_2 \frac{D_x(Q_{2;x}^2 h^1)}{Q_{2;x}} - Q_{2;x}^2 h^1\right)\right]. \tag{3.181}$$

Equations (3.180) and (3.181) are written in a formal notation and there are no singularities in these equations, even when we formally divide by $Q_{;x}$. Let us rewrite (3.180) and (3.181) as

$$0 = D_x\left[Q_1^2 h^0 + \frac{Q_1^2}{Q_{1;x}} D_x\left(\frac{Q_{1;x}^2 h^1}{Q_1}\right)\right], \tag{3.182}$$

$$0 = D_x\left[Q_2^2 h^0 + \frac{Q_2^2}{Q_{2;x}} D_x\left(\frac{Q_{2;x}^2 h^1}{Q_2}\right)\right], \tag{3.183}$$

which is again a formal notation and there do not occur singularities. From (3.182) and (3.183) we get

$$c_1 = Q_1^2 h^0 + \frac{Q_1^2}{Q_{1;x}} D_x\left(\frac{Q_{1;x}^2 h^1}{Q_1}\right), \quad c_1 \in \mathbb{R}, \tag{3.184}$$

$$c_2 = Q_2^2 h^0 + \frac{Q_2^2}{Q_{2;x}} D_x\left(\frac{Q_{2;x}^2 h^1}{Q_2}\right), \quad c_2 \in \mathbb{R}. \tag{3.185}$$

Now we are not multiplying by $\frac{D_xQ}{Q^2}$, as we did above, instead we do the following: We multiply (3.184) by Q_2^2, (3.185) by Q_1^2, and subtract both. This eliminates h^0 and we get (again, $Q = 0$ is not a problem here since it cannot vanish everywhere)

$$\begin{aligned} c_1 Q_2^2 - c_2 Q_1^2 =& \frac{Q_2^2 Q_1^2}{Q_{1;x}} D_x\left(\frac{Q_{1;x}^2 h^1}{Q_1}\right) - \frac{Q_1^2 Q_2^2}{Q_{2;x}} D_x\left(\frac{Q_{2;x}^2 h^1}{Q_2}\right) = \\ =& D_x\left(\frac{Q_2^2 Q_1^2}{Q_{1;x}} \frac{Q_{1;x}^2 h^1}{Q_1} - \frac{Q_1^2 Q_2^2}{Q_{2;x}} \frac{Q_{2;x}^2 h^1}{Q_2}\right) - \\ & - \frac{Q_{1;x}^2 h^1}{Q_1} D_x\left(\frac{Q_2^2 Q_1^2}{Q_{1;x}}\right) + \frac{Q_{2;x}^2 h^1}{Q_2} D_x\left(\frac{Q_1^2 Q_2^2}{Q_{2;x}}\right). \end{aligned} \tag{3.186}$$

Let us compute the last two terms in (3.186) separately (we can factor out h^1), i.e.

$$\begin{aligned}
&-\frac{Q_{1;x}^2}{Q_1}D_x\left(\frac{Q_2^2Q_1^2}{Q_{1;x}}\right)+\frac{Q_{2;x}^2}{Q_2}D_x\left(\frac{Q_1^2Q_2^2}{Q_{2;x}}\right)=\\
=&-\frac{Q_{1;x}^2}{Q_1}\left(2\frac{Q_2Q_{2;x}Q_1^2+Q_2^2Q_1Q_{1;x}}{Q_{1;x}}-\frac{Q_2^2Q_1^2Q_{1;xx}}{Q_{1;x}^2}\right)+\\
&+\frac{Q_{2;x}^2}{Q_2}\left(2\frac{Q_1Q_{1;x}Q_2^2+Q_1^2Q_2Q_{2;x}}{Q_{2;x}}-\frac{Q_1^2Q_2^2Q_{2;xx}}{Q_{2;x}^2}\right)=\\
=&(-2Q_2^2Q_{1;x}^2+Q_2^2Q_1Q_{1;xx})+(2Q_1^2Q_{2;x}^2-Q_1^2Q_2Q_{2;xx})=\\
=&2(Q_1^2Q_{2;x}^2-Q_2^2Q_{1;x}^2)+Q_1Q_2(Q_2Q_{1;xx}-Q_1Q_{2;xx})=\\
=&2(Q_1^2Q_{2;x}^2-Q_2^2Q_{1;x}^2)+Q_1Q_2D_x(Q_2Q_{1;x}-Q_1Q_{2;x})=\\
=&2(Q_1Q_{2;x}-Q_2Q_{1;x})(Q_1Q_{2;x}+Q_2Q_{1;x})+Q_1Q_2D_x(Q_2Q_{1;x}-Q_1Q_{2;x})=\\
=&2(Q_1Q_{2;x}-Q_2Q_{1;x})D_x(Q_1Q_2)+Q_1Q_2D_x(Q_2Q_{1;x}-Q_1Q_{2;x})=\\
=&-2AD_xB+BD_xA,
\end{aligned}$$

where we define

$$A:=Q_2Q_{1;x}-Q_1Q_{2;x}\quad\text{and}\quad B:=Q_1Q_2.$$

Using this identity, (3.186) becomes

$$\begin{aligned}
c_1Q_2^2-c_2Q_1^2=&D_x[(Q_2^2Q_1Q_{1;x}-Q_1^2Q_2Q_{2;x})h^1]+(BD_xA-2AD_xB)h^1=\\
=&D_x[Q_1Q_2(Q_2Q_{1;x}-Q_1Q_{2;x})h^1]+(BD_xA-2AD_xB)h^1=\\
=&D_x(BAh^1)+(BD_xA-2AD_xB)h^1.
\end{aligned}\tag{3.187}$$

Then we multiply (3.187) by A (note that $Q_{1;x}$ or $Q_{2;x}$ could vanish, but this does not cause problems, since A cannot vanish everywhere) and we get

$$\begin{aligned}
A(c_1Q_2^2-c_2Q_1^2)=&AD_x(BAh^1)+(BAD_xA-2A^2D_xB)h^1=\\
=&D_x(BA^2h^1)-(D_xA)BAh^1+(BAD_xA-2A^2D_xB)h^1=\\
=&D_x(BA^2h^1)-2(D_xB)A^2h^1.
\end{aligned}\tag{3.188}$$

Now we define the new unknown $F:=A^2h^1$ and (3.188) becomes

$$A(c_1Q_2^2-c_2Q_1^2)=D_x(BF)-2(D_xB)F.\tag{3.189}$$

The homogeneous solution is $F_{\text{hom}}=\lambda_1B$, where $\lambda_1\in\mathbb{R}$. Therefore,

$$h^1_{\text{hom}}=\frac{F_{\text{hom}}}{A^2}=\lambda_1\frac{B}{A^2}=\frac{\lambda_1Q_1Q_2}{(Q_2Q_{1;x}-Q_1Q_{2;x})^2},\quad\lambda_1\in\mathbb{R}.$$

We get the inhomogeneous solution with the help of variation of constants, i.e.

$$A(c_1Q_2^2-c_2Q_1^2)=D_x(\lambda_1B^2)-2(D_xB)\lambda_1B\overset{!}{=}(D_x\lambda_1)B^2$$

and this leads to

$$
\begin{aligned}
D_x\lambda_1 &\overset{!}{=} \frac{A(c_1Q_2^2 - c_2Q_1^2)}{B^2} = \frac{(Q_2Q_{1;x} - Q_1Q_{2;x})(c_1Q_2^2 - c_2Q_1^2)}{(Q_1Q_2)^2} = \\
&= (Q_2Q_{1;x} - Q_1Q_{2;x})(\frac{c_1}{Q_1^2} - \frac{c_2}{Q_2^2}) = \\
&= \frac{c_1Q_2}{Q_1}\left(\frac{Q_{1;x}}{Q_1} - \frac{Q_{2;x}}{Q_2}\right) - \frac{c_2Q_1}{Q_2}\left(\frac{Q_{1;x}}{Q_1} - \frac{Q_{2;x}}{Q_2}\right) = \\
&= \left(\frac{c_1Q_2}{Q_1} - \frac{c_2Q_1}{Q_2}\right) D_x(\ln Q_1 - \ln Q_2) = \\
&= \left(\frac{c_1Q_2}{Q_1} - \frac{c_2Q_1}{Q_2}\right) D_x \ln\frac{Q_1}{Q_2}.
\end{aligned}
\tag{3.190}
$$

The equation in (3.190) is of the form

$$
D_x\lambda_1 = \left(\frac{c_1}{g} - c_2g\right) D_x \ln g = \left(\frac{c_1}{g} - c_2g\right)\frac{D_xg}{g} = \tag{3.191}
$$

$$
= D_x\left(\frac{-c_1}{g} - c_2g\right), \quad \text{where} \quad g := \frac{Q_1}{Q_2}. \tag{3.192}
$$

Therefore, we get

$$
\lambda_1 = \frac{-c_1}{g} - c_2g + c_3, \quad c_1,c_2,c_3 \in \mathbb{R}.
$$

Then we can write

$$
\begin{aligned}
A^2h^1_{\text{inhom}} &= F_{\text{inhom}} = \lambda_1 B = (\frac{-c_1}{g} - c_2g + c_3)B = \left(-c_1\frac{Q_2}{Q_1} - c_2\frac{Q_1}{Q_2} + c_3\right)B = \\
&= -c_1Q_2^2 - c_2Q_1^2 + c_3Q_1Q_2
\end{aligned}
$$

and a inhomogeneous solution h^1_{inhom} is

$$
h^1_{\text{inhom}} = \frac{-c_1Q_2^2 - c_2Q_1^2 + c_3Q_1Q_2}{(Q_2Q_{1;x} - Q_1Q_{2;x})^2}.
$$

The general solution h^1_{gen} is

$$
\begin{aligned}
h^1_{\text{gen}} &= \frac{-c_1Q_2^2 - c_2Q_1^2 + c_3Q_1Q_2}{(Q_2Q_{1;x} - Q_1Q_{2;x})^2} + \frac{\lambda_1Q_1Q_2}{(Q_2Q_{1;x} - Q_1Q_{2;x})^2} = \\
&= \frac{C_1Q_2^2 + C_2Q_1^2 + C_3Q_1Q_2}{(Q_2Q_{1;x} - Q_1Q_{2;x})^2}, \\
&\quad C_1 := -c_1,\, C_2 := -c_2,\, C_3 := c_3 + \lambda_1 \in \mathbb{R}.
\end{aligned}
\tag{3.193}
$$

Since

$$\begin{aligned} Q_{2;x}Q_1 - Q_{1;x}Q_2 &= (O(1) - u_{xx}V_2^x)(V_1^u - u_xV_1^x) - (O(1) - u_{xx}V_1^x)(V_2^u - u_xV_2^x) = \\ &= O(1) - u_{xx}V_2^x(V_1^u - u_xV_1^x) + u_{xx}V_1^x(V_2^u - u_xV_2^x) = \\ &= O(1) + u_{xx}(-V_2^xV_1^u + V_1^xV_2^u) + u_{xx}u_x(V_2^xV_1^x - V_1^xV_2^x) = \\ &= O(1) + u_{xx}\underbrace{(-V_2^xV_1^u + V_1^xV_2^u)}_{\neq 0 \text{ for all } p \in U \subset E}, \end{aligned}$$

there will be points (x,u,u_x,u_{xx}) in J^kE, where this expression vanishes. Therefore, h^1_{gen} is singular, except $C_1,C_2,C_3 = 0$. The coordinates $(u_x,u_{xx},...)$ are always assumed to take all possible values in the whole domain of $\mathbb{R}$, which is a simple consequence of the definition of the jet space and the use of associated charts. Let us pull-back h^1_{gen} in (3.193) by a prolonged section $\sigma \in \Gamma(E)$. Then, in the one-dimensional case where $n = 1$, Picard-Lindelöf's theorem says: If we found a (singular) solution (3.193) for almost every x in the considered domain of definition, which tends to $\pm\infty$ for some values x, then there cannot exist a smooth continuation for all x in the considered domain.

If $h^1 \equiv 0$ then (3.182) and (3.183) tell us h^0 is singular or $h^0 \equiv 0$ and we have proven Theorem 1.0.3. The singularity in h^0 is caused by the term

$$h^0 = \frac{C_4}{Q^2}, \quad C_4 \in \mathbb{R},$$

where either $Q_1 = V_1^u - u_xV_1^x$ or $Q_2 = V_2^u - u_xV_2^x$ satisfy $V_1^x \neq 0$ or $V_2^x \neq 0$ for some x. It is also possible to argue that, if $h^1 \equiv 0$, then the two equations (3.184) and (3.185)

$$\begin{gathered} h^0 = \frac{C_4}{Q_1^2}, \quad h^0 = \frac{C_5}{Q_2^2}, \quad C_4,C_5 \in \mathbb{R} \\ (\Rightarrow \quad Q_2^2C_4 = Q_1^2C_5) \end{gathered}$$

cannot be satisfied at the same time. With the singular solution h^1_{gen} we could try to solve the equation for h^0 and then construct H^0,H^1.

We conjecture that we get a similar result for fourth order source forms, $m = 1$, but arbitrary n. Solving the differential equation for h^1 in this case could be quite difficult. Maybe we should only try to investigate the singularities and do a kind of perturbation theory. In any case, it seems to be a very interesting open problem. Because of lack of time we cannot solve it anymore here in this dissertation. The above calculations already prove Theorem 1.0.3, but we want to formulate the proof briefly again.

Proof of Theorem 1.0.3: Proving the reduced Helmholtz conditions in Lemma 3.8.2

is the first step. Then, the ECS can be written as

$$0 = [QD_xh^0 + 2(D_xQ)h^0] + [(D_xQ)D_x^2h^1 + 3(D_x^2Q)D_xh^1 + 2(D_x^3Q)h^1],$$

what we derived in (3.163). It turns out that this equation can also be written as

$$0 = \frac{1}{Q}D_x[Q^2h^0] + D_x\left(\frac{D_x[(D_xQ)^2h^1]}{D_xQ}\right).$$

what we derived in(3.171). Multiplying this equation by the integration factor Q leads to

$$c = Q^2h^0 + \frac{Q^2}{D_xQ}D_x\left(\frac{(D_xQ)^2h^1}{Q}\right), \quad c \in \mathbb{R}, \tag{3.194}$$

what we derived in (3.184) and (3.185). Since in the case where $\operatorname{span}\{V_p : V \in \mathcal{V}\} = T_pE$ for all $p \in U \subset E$, we have at least two symmetry vector fields V_1,V_2 with corresponding characteristics Q_1,Q_2. Therefore, the expression h^0 in (3.194) can be eliminated and we get the equation

$$A(c_1Q_2^2 - c_2Q_1^2) = D_x(BA^2h^1) - 2(D_xB)A^2h^1, \tag{3.195}$$

what we derived in (3.188), where

$$A := Q_2Q_{1;x} - Q_1Q_{2;x} \quad \text{and} \quad B := Q_1Q_2.$$

Equation (3.195) can be solved and the general solution is (3.193). The solution is singular, except $h^1 \equiv 0$. Then h^0 is singular or $h^0 \equiv 0$. Also see the remark below. □

We did not yet precisely define what we mean with *singular*, and therefore we need the following definition:

Definition 3.8.3. *We call a source form Δ **singular** if it is defined on some open subset $\mathcal{R}^k \subset J^kE$ and if there is no smooth continuation of Δ from $\mathcal{R}^k$ to J^kE. Otherwise we call a source form **non-singular**.*

A similar definition holds for general differential forms and functions on J^kE.[8] Note that every source form defined on $\mathcal{R}^k \subset J^kE$ is a smooth differential form on $\mathcal{R}^k$, whether it is singular or non-singular according to Definition 3.8.3.

Let us discuss further aspects of singular and non-singular expressions. For example, intuitively the expression

$$f = \frac{1}{u_x} \tag{3.196}$$

[8]Subsets $\mathcal{R} \subset J^kE$ are also used in (AT92, p.17ff), but there not necessarily open subsets, rather some kind of embedded manifolds.

should be a singular expression, defined on some open subset $\mathcal{R}^1 \subset J^1E$, where we have local coordinates (x,u,u_x). However, we have to clarify for which subset $\mathcal{R}^1$ we define the expression (3.196), how all possible continuations of (3.196) form $\mathcal{R}^1$ to J^1E can be described and where the local coordinates (x,u,u_x) are defined. Without specifying these things it does not make sense to talk about a singular or non-singular expression in (3.196). We do the discussion exemplary with the expression in (3.196) and generalizations are straight forward.

Let $U \subset E$ be open and $\varphi : U \to \Omega$ a local chart for E, where $\Omega \subset \mathbb{R}^{n+m}$. Furthermore, let $\varphi^0 : U^0 \to \Omega^0$, $U^0 = \pi(U)$, $\Omega^0 \subset \mathbb{R}^n$ be the corresponding local chart for M, such that the diagram

$$\begin{tikzcd} U \subset E \arrow[r, "\varphi"] \arrow[d, "\pi"'] & \Omega \subset \mathbb{R}^n \times \mathbb{R}^m \arrow[d, "\tilde{\pi}"] \\ U^0 \subset M \arrow[r, "\varphi^0"] & \Omega^0 \subset \mathbb{R}^n \end{tikzcd}$$

commutes ($\tilde{\pi}$ is the canonical projection). The chart φ defines a so called **associated chart on** J^kE, where local coordinates $(x^i,u^\alpha,u^\alpha_i,u^\alpha_{ij},...,u^\alpha_I)$, $|I| = k$, on J^kE are given as

$$u^\alpha_{i_1...i_l}(\mathrm{pr}^k\sigma(q)) = [D_{i_1}...D_{i_l}(u^\alpha \circ \sigma \circ (\varphi^0)^{-1})](x), \quad 1 \leq l \leq k, \tag{3.197}$$

and $\varphi^0(q) = (x^i) = (x^1,...,x^n)$. Associated charts on J^kE are written as φ^k. We took the definition for associated charts from (Kru97b, p.30).

Every locally defined smooth functions $u^\alpha(x)$, $\alpha = 1,2,...,m$ define a section $\sigma(q) := \varphi^{-1}(x(q),u^\alpha(x(q)))$ on E, where $x = \varphi^0(q)$. Prolonging this local section $\sigma \in \Gamma(E)$ at a point $q \in M$ defines corresponding coordinates $(x^i,u^\alpha,u^\alpha_i)$ on J^1E (see (3.197)). These local coordinates are defined in $\Omega \times \mathbb{R}^{nm}$, i.e.

$$(x^i,u^\alpha,u^\alpha_i) \in \Omega \times \mathbb{R}^{nm}.$$

Now we consider the situation the other way around, i.e. for every $(x^i,u^\alpha,u^\alpha_i) \in \Omega \times \mathbb{R}^{nm}$ there exists a local section $\gamma \in \Gamma(E)$ such that $\mathrm{pr}^1\gamma$ has these local coordinates at a point $x = \gamma(q)$, i.e.

$$\varphi^1(\mathrm{pr}^1\gamma(q)) = (x^i,u^\alpha,u^\alpha_i).$$

More importantly, $\varphi^1 : (\pi^{1,0})^{-1}(U) \to \Omega \times \mathbb{R}^{nm}$ is an (associated) chart for J^1E, i.e. a bijective map between these open sets.

We define the function (3.196) on $\mathcal{R}^1$, where

$$\mathcal{R}^1 := (\varphi^1)^{-1}(\Omega \times (\mathbb{R} \setminus \{0\})). \tag{3.198}$$

It can easily be seen that the function (3.196) is singular in the sense of Definition 3.8.3, when written in local coordinates and defined on $\Omega \times (\mathbb{R} \setminus \{0\})$. That is, there is no smooth continuation of $f(u_x) = \frac{1}{u_x}$ from $\Omega \times (\mathbb{R} \setminus \{0\})$ to $\Omega \times \mathbb{R}$. The occurrence of such singularities is independent of the choice of local coordinates. Therefore, when f is considered to be a function defined on $\mathcal{R}^1$, i.e.

$$f(u_x) = \tilde{f}(\mathrm{pr}^1\sigma) = \frac{1}{u_x(\mathrm{pr}^1\sigma)},$$

then there is no smooth continuation form $\mathcal{R}^1$ to J^1E and this means that f is singular. More general singular functions, differential forms and so on can be investigated in a similar way and usually we do not discuss this in such a detail.

Remark on the proof of Theorem 1.0.3: Let us consider the expression (3.193), let us explain why this expression is singular according to Definition 3.8.3, and we want to say a few more words about the proof of Theorem 1.0.3. First, we define an associated chart on J^kE and a certain open subset $\mathcal{R}^k \subset J^kE$, where (3.193) is defined. Then we do all the calculations in the proof of Theorem 1.0.3 and then we show that there is no smooth continuation of h^1 from $\mathcal{R}^k$ to J^kE, except when $h^1 \equiv 0$. Then we do the same for h^0 which completes the proof.

Note that open subsets $\mathcal{R}^k \subset J^kE$ do in general not have the structure of a jet bundle over E. When we do not have the structure of a jet bundle then it is not obvious what we mean with sections and prolonged vector fields on $\mathcal{R}^k$. However, projectable vector fields $V \in \mathfrak{X}(E)$ can be prolonged on J^kE and then we can restrict pr^kV to $\mathcal{R}^k$ and apply the restricted vector field on functions or the corresponding Lie derivative on differential forms.

Let us assume that we would derive in the proof of Theorem 1.0.3 that h^1 in (3.193) is of the form

$$h^1 = \frac{c}{u_x}, \quad c \in \mathbb{R}.$$

According to what we have discussed above means that h^1 can only be a non-singular expression on J^kE when $c = 0$. This holds in a similar way when

$$h^1 = \frac{c}{u_{xx}}, \quad h^1 = \frac{c}{u_{(3)}}, \quad h^1 = \frac{c}{u_{(4)}}, \ldots, \quad c \in \mathbb{R}. \tag{3.199}$$

But if we would derive that, for example,

$$h^1 = \frac{c}{u}, \quad c \in \mathbb{R} \tag{3.200}$$

then it does not necessarily follow that $c = 0$, since (3.200) is in general not a singular expression defined on some open subset $\mathcal{R}^k \subset J^kE$. For example, when the fiber bundle E is $\pi : \mathbb{R} \times (\mathbb{R} \setminus 0) \to \mathbb{R}$, where the local coordinates (x,u) are identified with points on E through the identity map, $\pi(x,u) = x$ and where (3.200) is defined (on open subsets) and non-singular. Under which conditions (3.200) is singular or non-singular on open subsets depends on the definition of E and the local charts there. The same happens when we consider $f = \frac{c}{x}$, $c \in \mathbb{R}$. The main difference between the expressions in (3.199) and (3.200) is that we always describe, in all calculations, the coordinates $(u_x,u_{xx},...,u_{(k)})$ by associated fiber bundle charts, but there are no such special charts for the expression in (3.200). This is why (3.200) can be defined on E (or open subsets of E) as a non-singular expression. In other words, for E we do not assume a certain topology and we do not have canonical charts on E. However, we always have canonical charts for the coordinates $(u_i^\alpha,u_{ij}^\alpha,...,u_I^\alpha)$ on J^kE, namely associated charts. The investigation of singularities can be considered as a question of topology and global analysis in what we have discussed above.

In this context, it is helpful to know that E can have non-trivial cohomology in the sense of De-Rham. The cohomology of J^kE (in the sense of the variational sequence) is the following $H^*(J^kE) \backsimeq H^*(E)$, see (KM10, Tak79) or (Kru15, p.xi).

Also note that local coordinate transformations of u_x are of the form

$$v_y = \frac{\partial x}{\partial y}\left(\frac{\partial v}{\partial x} + u_x\frac{\partial v}{\partial u}\right). \tag{3.201}$$

This shows that J^1E is not a vector bundle over E, rather an affine linear bundle. However, this is no longer true for higher order jet spaces. For example, J^2E is not an affine linear bundle over E. This can be seen in Proposition 2.4.3 and the local coordinate transformation for v_{yy}. Note that when u_x takes all possible values in $\mathbb{R}$ then v_y in (3.201) also takes all possible values in $\mathbb{R}$ (since $\frac{\partial x}{\partial y}\frac{\partial v}{\partial u}$ is non-vanishing everywhere). This is also true for higher order coordinates u_{xx}, v_{yy} and so on. Practically, we always use associated fiber bundle charts. Then the expressions (3.199) cannot be defined for all $u_x \in \mathbb{R}$, $u_{xx} \in \mathbb{R}$ and so on, which immediately shows they will be singular expressions when deriving them in the proof of Theorem 1.0.3. However, when we would derive the expression (3.200) in the proof of Theorem 1.0.3 then it does not necessarily follow that $c = 0$, since the x and u-coordinate do not have to be defined for all values in $\mathbb{R}$. This is the rough idea how to understand this problem, but of course the precise formulation is to consider the topology on E and the local charts there.

Although, we do not allow singular expression like (3.193), by assumptions of Theorem 1.0.3, it would be interesting to know how the corresponding f or source form Δ would look like. In other words, we would like to find the homotopy operator of h^0,h^1. We will partially investigate this problem in the next subsection, where we again consider $n,m = 1$. In the next subsection, we will consider a very special case, where we only investigate translation symmetries ∂_x,∂_u and third order source forms. We also do not introduce the homotopy operator in detail, we rather invert the operator h^0,h^1 by hand and we want to get a relatively fast result as a kind of verification or example of what we have proven above. In the next subsection, we do not use Lemma 3.8.2 or other techniques from this subsection. Therefore, it will be a second approach how to solve Takens' problem in a very special case and, indeed, a kind of verification of the result from above.

3.8.2. Complete Classification for Translation Symmetries

Let us consider the trivial fiber bundle $\pi : \mathbb{R}^2 \to \mathbb{R}$, with coordinates (x,u) and projection $\pi(x,u) = x$. Furthermore, we consider translation symmetries ∂_x,∂_u and source forms of third order. We want to classify all source forms which are invariant under the prolonged symmetries $\partial_x,\partial_u \in \mathcal{V}$ and which satisfy the corresponding conservation law conditions

$$\begin{aligned} \partial_u : &\quad Qf = 1 \cdot f = D_x g, \\ \partial_x : &\quad Qf = u_x f = D_x h, \end{aligned} \tag{3.202}$$

for some functions g,h. The statement is the following: All source forms which satisfy the above conditions are of the form

$$f = a - e - \frac{(-u_x e + s)u_{(3)}}{u_{xx}^2} + u_{xx} K_{u_x}(u_x)$$

where $a,s,e \in \mathbb{R}$ and K_{u_x} is an arbitrary function depending on u_x.

<u>Proof:</u> Since the source form is of third order, the conservation law conditions (3.202) lead to $g,h = O(2)$. Because Δ satisfies the ∂_x,∂_u symmetries, f cannot depend on (x,u). Furthermore, ∂_x,∂_u commute with D_x and u_x. Therefore, we get

$$\begin{aligned} \partial_x : &\quad 0 = \partial_x f = \partial_x D_x g, &\Rightarrow &\quad 0 = D_x g_x \\ \partial_x : &\quad 0 = \partial_x (u_x f) = \partial_x D_x h, &\Rightarrow &\quad 0 = D_x h_x \\ \partial_u : &\quad 0 = \partial_u f = \partial_u D_x g, &\Rightarrow &\quad 0 = D_x g_u \\ \partial_u : &\quad 0 = \partial_u (u_x f) = \partial_u D_x h, &\Rightarrow &\quad 0 = D_x h_u. \end{aligned}$$

These equations show that g_x,h_x,g_u,h_u must be constant. Integrating these conditions leads to

$$\begin{aligned} g &= xa + ub + r(u_x,u_{xx}), \\ h &= xc + ud + q(u_x,u_{xx}), \end{aligned}$$

where $a,b,c,d \in \mathbb{R}$ and r,q are arbitrary functions depending on (u_x,u_{xx}). We can write

$$u_x f = D_x h = u_x D_x g = D_x(u_x g) - u_{xx} g \quad \Rightarrow \quad u_{xx} g = D_x(u_x g - h).$$

That is, $u_{xx}g$ must be a total derivative and the Euler-Lagrange operator applied to this expression must vanish, i.e.

$$\begin{aligned}
0 &= (\partial_u - D_x\partial_{u_x} + D_x^2\partial_{u_{xx}})(u_{xx}g) = \\
&= u_{xx}b - D_x(u_{xx}r_{u_x}) + D_x^2(g + u_{xx}g_{u_{xx}}) = \\
&= 2u_{xx}b - D_x(u_{xx}r_{u_x}) + D_x^2(r + u_{xx}r_{u_{xx}}) = \\
&= 2u_{xx}b - D_x(u_{xx}r_{u_x}) + D_x(u_{xx}r_{u_x} + u_{(3)}r_{u_{xx}} + u_{(3)}r_{u_{xx}} + u_{xx}D_x r_{u_{xx}}) = \\
&= 2u_{xx}b + D_x(2u_{(3)}r_{u_{xx}} + u_{xx}D_x r_{u_{xx}}) = \\
&= D_x[2u_x b + (2u_{(3)}r_{u_{xx}} + u_{xx}D_x r_{u_{xx}})] = \\
&= D_x[2u_x b + 2u_{(3)}r_{u_{xx}} + u_{xx}(u_{xx}r_{u_x u_{xx}} + u_{(3)}r_{u_{xx}u_{xx}})] = \\
&= D_x[\underbrace{2u_x b + u_{xx}^2 r_{u_x u_{xx}}}_{=:\mathrm{I})} + u_{(3)}\underbrace{(2r_{u_{xx}} + u_{xx}r_{u_{xx}u_{xx}})}_{=:\mathrm{II})}].
\end{aligned} \tag{3.203}$$

The term I) in (3.203) must be constant and the term II) must vanish. To solve the equation for II), let us multiply II) by u_{xx}. Then the general solution of

$$\begin{aligned}
u_{xx}(2r_{u_{xx}} + u_{xx}r_{u_{xx}u_{xx}}) = \partial_{u_{xx}}(u_{xx}^2 r_{u_{xx}}) \quad &\Rightarrow \quad r_{u_{xx}} = -\frac{R(u_x)}{u_{xx}^2} \\
&\Rightarrow \quad r = \frac{R(u_x)}{u_{xx}} + K(u_x)
\end{aligned}$$

allows only singular solutions or solutions, where r only depends on u_x. In case of singular solutions, we can further determine r with the help of I), i.e.

$$2u_x b - R_{u_x}(u_x) = e,$$

where $e \in \mathbb{R}$. Therefore, the solution for R is

$$R(u_x) = -u_x^2 b - u_x e + s,$$

where $s \in \mathbb{R}$. Therefore, R has no singularities. Then we get

$$\begin{aligned}
f = D_x g = D_x(ax + ub + \frac{R(u_x)}{u_{xx}} + K(u_x)) = \\
= a + bu_x - \frac{u_{(3)}R(u_x)}{u_{xx}^2} + R_{u_x}(u_x) + u_{xx}K_{u_x}(u_x)
\end{aligned} \tag{3.204}$$

and this describes the singularity of third order f very precisely. However, we can determine R even more precisely, by using the fact that $u_x f$ must be a total derivative, i.e. $\mathcal{E}(u_x f) = 0$. This leads to the equation

$$0 = \mathcal{E}[u_x a + bu_x^2 - \frac{u_{(3)}u_x R(u_x)}{u_{xx}^2} + u_x R_{u_x}(u_x) + u_{xx}u_x K_{u_x}(u_x)].$$

The term au_x is a total derivative, and therefore vanishes for all a. The same happens with

$$u_{xx}u_xK_{u_x}(u_x) = D_x \int u_xK_{u_x}(u_x)du_x.$$

It remains to show that

$$\begin{aligned}
0 &= (\partial_u - D_x\partial_{u_x} + D_x^2\partial_{u_{xx}} - D_x^3\partial_{u_{(3)}})\left[bu_x^2 - \frac{u_{(3)}u_xR(u_x)}{u_{xx}^2} + u_xR_{u_x}(u_x)\right] = \\
&= D_x(-\partial_{u_x} + D_x\partial_{u_{xx}} - D_x^2\partial_{u_{(3)}})\left[-\frac{u_{(3)}u_xR}{u_{xx}^2} - bu_x^2\right] = \\
&= D_x\left[\frac{u_{(3)}(R+u_xR_{u_x})}{u_{xx}^2} - 2bu_x + 2D_x\frac{u_{(3)}u_xR}{u_{xx}^3} + D_x^2\frac{u_xR}{u_{xx}^2}\right] = \\
&= D_x\left[\frac{u_{(3)}(R+u_xR_{u_x})}{u_{xx}^2} - 2bu_x + 2D_x\frac{u_{(3)}u_xR}{u_{xx}^3} + \right. \\
&\qquad\qquad \left. + D_x\left(-2\frac{u_{(3)}u_xR}{u_{xx}^3} + \frac{D_x(u_xR)}{u_{xx}^2}\right)\right] = \\
&= D_x\left[\frac{u_{(3)}(R+u_xR_{u_x})}{u_{xx}^2} - 2bu_x + D_x\frac{u_{xx}R + u_{xx}u_xR_{u_x}}{u_{xx}^2}\right] = \\
&= D_x\left[-2bu_x + \frac{D_x(R+u_xR_{u_x})}{u_{xx}}\right] = \\
&= D_x\left[-2bu_x + \frac{u_{xx}R_{u_x} + u_{xx}R_{u_x} + u_{xx}u_xR_{u_xu_x}}{u_{xx}}\right] = \\
&= D_x\left[-2bu_x + 2R_{u_x} + u_xR_{u_xu_x}\right] = \\
&= D_x\left[-2bu_x + 2(-2bu_x - e) + u_x(-2b)\right] = \\
&= D_x\left[-8bu_x - 2e\right]
\end{aligned}$$

and this means $b = 0$. Therefore,

$$f = a - \frac{(-u_xe + s)u_{(3)}}{u_{xx}^2} - e + u_{xx}K_{u_x}(u_x) \tag{3.205}$$

and this exactly describes the singularity of f, and even more, all f which satisfy ∂_x,∂_u symmetry and corresponding conservation laws are of this form. □

We could also consider the equation $\mathcal{E}f = 0$ and investigate if we get further restrictions and we could try to compute the function q, but we will stop the discussion here.

A similar calculation holds for $n = 1$, arbitrary m and third order source forms, but the calculations are getting more complicated. We could generalize Theorem 1.0.3 in the following sense: Classify all source forms, which satisfy certain symmetries and corresponding conservation laws (similar to the calculation above).

Maybe there remains one open question: How does the expression in (3.205) transform under local coordinate transformations (and how can we describe the singularity invariantly)? We will not investigate this problem in more detail. Let us only mention that when vector fields $V,W \in \mathcal{V}$ commute and when they span T_pE at every $p \in E$, then we can always find local coordinates such that $V = \partial_x$ and $W = \partial_u$. Therefore, at least for such vector fields we get the same result as in (3.205) (modulo coordinate transformations). The effect of choosing certain local coordinates will be investigated in more detail in Section 4.4 and it is of great interest in general. Then we can indeed classify a wide variety of source forms, namely all those where the symmetry vector fields span T_pE at every $p \in E$ and $[V,W] = 0$.

4. Information Beyond the Proofs

In this chapter, we want to provide further information about Takens' problem. In Section 4.1, we show interesting counter examples, which show that Theorem 1.0.2 is sharp in some sense, and in Section 4.2, we show that Theorem 1.0.3 is sharp in some sense, as well. Then in Section 4.3, we want to investigate the question of applications. Finally, we investigate a kind of technical question in Section 4.4, namely if the special choice of local coordinates can simplify the proof of Theorem (1.0.2), (1.0.3) and similar ones.

4.1. Counter Examples, Part I

The following counter examples show that Theorem 1.0.2 is sharp in basically any sense.

The trivial source form $\Delta \equiv 0$ satisfies all symmetry conditions and corresponding conservation laws (and it is also variational). If Theorem 1.0.2 would imply that all second order source forms, which satisfy symmetries $V \in \mathcal{V}$ such that $\operatorname{span}\{V_p : V \in \mathcal{V}\} = T_pE$ at each $p \in E$ and the corresponding conservation laws hold, are trivial source forms, then we would have formulated the theorem in that way. Indeed, Theorem 1.0.2 allows a lot of non-trivial source forms. Let us discuss just two of them. Note that when not otherwise stated we will always assume the trivial fiber bundle $\pi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ with projection $\pi(x^i,u^\alpha) = (x^i)$. The source form

$$\Delta = u_{xx} du \wedge dx$$

satisfies ∂_x,∂_u symmetry and corresponding conservation laws. Laplace's equation in general satisfies these conditions. Let us consider $n = 2$ and

$$\Delta = (u_{xx} + u_{yy}) du \wedge dx \wedge dy,$$

which satisfies $\partial_x, \partial_y, \partial_u$ symmetry and corresponding conservation laws of the form

$$\begin{aligned} 1 \cdot \operatorname{div} \operatorname{grad} u &= \operatorname{div}(\operatorname{grad} u) = \operatorname{div} \begin{pmatrix} u_x \\ u_y \end{pmatrix}, \\ u_x(u_{xx} + u_{yy}) &= \operatorname{div} \begin{pmatrix} \frac{1}{2}(u_x^2 - u_y^2) \\ u_x u_y \end{pmatrix}, \\ u_y(u_{xx} + u_{yy}) &= \operatorname{div} \begin{pmatrix} u_x u_y \\ \frac{1}{2}(u_y^2 - u_x^2) \end{pmatrix}. \end{aligned}$$

The first conservation laws has probably a physical meaning in the sense of *Fick's law*. The second and third maybe not. We will come back to this in Section 4.3, where we briefly discuss applications and the physical meaning of conservation laws.

To formulate the counter examples, it is reasonable to consider the following definition:

Definition 4.1.1. *We call a k-th order source form* $\Delta = f_\alpha du^\alpha \wedge dx^1 \wedge ... \wedge dx^n$ ***non-degenerate*** *if* $f_{\alpha, u_I^\beta} \neq 0$ *for at least one* $\alpha, \beta = 1,2,...,m$ *and one multi-index* I *of length* k*. Otherwise we call* Δ ***degenerate****.*

1st Counter example (corresponding conservation laws are necessary): Let us consider the simplest case, where $n,m = 1$ and where we have the trivial fiber bundle $\pi : \mathbb{R}^2 \to \mathbb{R}$ and $\pi(x,u) = x$. In this case, we have global coordinates (x,u) and they will be identified with points on E through the identity map (similar for (x,u,u_x) on J^1E and for higher order jet coordinates). The second order (degenerate) source form

$$\Delta = f du \wedge dx = u_x du \wedge dx$$

satisfies translation symmetry in x and u directions, described by the vector fields $V_1 = \partial_x$ and $V_2 = \partial_u$ on E. The prolongations of V_1 and V_2 are

$$\begin{aligned} \operatorname{pr}^2 V_1 &= \partial_x + 0 \cdot \partial_u + 0 \cdot \partial_{u_x} + 0 \cdot \partial_{u_{xx}} = \partial_x, \\ \operatorname{pr}^2 V_2 &= \partial_u + 0 \cdot \partial_{u_x} + 0 \cdot \partial_{u_{xx}} = \partial_u. \end{aligned}$$

The characteristics are

$$Q_1 = -u_x \quad \text{and} \quad Q_2 = 1.$$

Only Q_2 generates a conservation law of the form

$$Q_2 f = 1 \cdot u_x = D_x u. \tag{4.1}$$

The source form is not variational (as we saw in Subsection 2.6.2). The symmetry vector fields V_1, V_2 span the tangent space T_pE at each $p \in E$. But we do not have corresponding conservation laws, since $V_1 = \partial_x$ does not generate a conservation law. Therefore, the assumption of having corresponding conservation law in Theorem 1.0.2 is necessary in general. Theorem 1.0.2 is sharp in this sense. This counter example can also be found in (AP94, p.215) in Example 3.12.

2nd Counter example (span$\{V_{p,\mathscr{A}}\} = T_pE$ is necessary, vertical symmetry): We consider the same situation as in the 1st Counter example, except we only consider the symmetry vector field $V_2 = \partial_u$ on E and corresponding conservation law (4.1). Then all assumption in Theorem 1.0.2 are satisfied, except we cannot span T_pE with all of the symmetry vector fields which satisfy corresponding conservation laws. Therefore, the assumption span$\{V_{p,\mathscr{A}}\} = T_pE$ is necessary in general and Theorem 1.0.2 is sharp in this sense.

In this counter example, ∂_u is a vertical symmetry and we also want to find a symmetry where $\pi_* V \neq 0$. It is clear (see Subsection 3.3.2) that we have to consider systems if we want to find non-singular second order counter examples. A singular counter example would be $f = \frac{1}{u_x}$ with ∂_x symmetry. It is also possible to consider third and higher order source forms.

3rd Counter example (span$\{V_{p,\mathscr{A}}\} = T_pE$ is necessary, non-vertical symmetry, 3rd order): Let us consider the 3rd order source form

$$\Delta = (\frac{1}{2}u_x u_{(3)} + u_{xx}^2)du \wedge dx,$$

which satisfies ∂_x symmetry and corresponding conservation law

$$-Qf = u_x(\frac{1}{2}u_x u_{(3)} + u_{xx}^2) = D_x(\frac{1}{2}u_x^2 u_{xx}),$$

but is not variational.

4th Counter example (span$\{V_{p,\mathscr{A}}\} = T_pE$ is necessary, non-vertical symmetry, 2nd order): We consider ∂_x symmetry and the trivial fiber bundle $\pi : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ with coordinates (x,u,v) and projection $\pi(x,u,v) = x$. Let us define

$$\Delta := Av_x v_x du \wedge dx - Au_x v_x dv \wedge dx$$

where $A = A(u,v)$. This source form is ∂_x invariant and has a trivial conservation law of the form

$$u_x^\alpha f_\alpha = u_x A v_x v_x - v_x A u_x v_x = 0 = D_x \,\text{const}.$$

The source form is not variational when, for example, $A \equiv 1$, since then

$$\begin{aligned} H_{uv} = f_{u,v} - f_{v,u} + D_x f_{v,u_x} &= v_x v_x A_v + u_x v_x A_u - D_x(v_x A) = \\ &= v_x v_x A_v + u_x v_x A_u - v_{xx}A - v_x(u_x A_u + v_x A_v) = -v_{xx}A \neq 0. \end{aligned}$$

This counter example is similar to the construction of counter examples in (Poh95, p.354).

5th Counter example (third order source forms): In (MPV08, p.14) we can find a counter example, but it is relatively complicated (it is also for $n \geq 3$). Therefore, we construct a simpler counter example. Similar to the proof in Subsection 3.8.2, we can classify source forms, which satisfy translation symmetries ∂_x,∂_u,∂_v and corresponding conservation laws. Let us define

$$\Delta := D_x(u_{xx}v_{xx})du \wedge dx + D_x(-u_{xx}^2)dv \wedge dx.$$

It is clear that this is a third order irreducible source form, which satisfies ∂_u,∂_v symmetries and corresponding conservation laws, since the differential equation itself is a total derivative. Furthermore, the source form is ∂_x invariant and has corresponding conservation law of the form

$$\begin{aligned} u_x f_1 + v_x f_2 = u_x D_x(u_{xx}v_{xx}) - v_x D_x(u_{xx}^2) = \\ = D_x(u_x u_{xx} v_{xx} - u_{xx}^2 v_x). \end{aligned}$$

The source form is not variational.

6th Counter example (globally variational source forms): We consider the fiber bundle $\pi : (\mathbb{R} \times (\mathbb{R}^2 \setminus 0)) \to \mathbb{R}$ with coordinates (x,u,v) and projection $\pi(x,u,v) = x$. Furthermore, let us define

$$\Delta := \left(\frac{v}{u^2+v^2} du - \frac{u}{u^2+v^2} dv \right) \wedge dx. \tag{4.2}$$

Then Δ satisfies the Helmholtz conditions

$$\begin{aligned} H_{\alpha\beta} = f_{\alpha,u^\beta} - f_{\beta,u^\alpha} = \partial_v \frac{v}{u^2+v^2} + \partial_u \frac{u}{u^2+v^2} = , \quad \alpha = 1,\beta = 2 \\ = \frac{1}{u^2+v^2} - \frac{v2v}{(u^2+v^2)^2} + \frac{1}{u^2+v^2} - \frac{u2u}{(u^2+v^2)^2} = 0 \end{aligned}$$

and all other Helmholtz conditions are also satisfied. A local Lagrangian is

$$L = \arctan \frac{u}{v},$$

since

$$\begin{aligned} L_u &= \partial_u \arctan \frac{u}{v} = \frac{\frac{1}{v}}{1 + (\frac{u}{v})^2} = \frac{\frac{1}{v}v^2}{v^2+u^2}, \\ L_v &= \partial_v \arctan \frac{u}{v} = \frac{\frac{-u}{v^2}}{1 + (\frac{u}{v})^2} = \frac{1}{v^2}v^2 \frac{-u}{v^2+u^2}. \end{aligned}$$

The Lagrangian is not globally defined, since it is defined on $\mathbb{R} \times (\mathbb{R}^2 \setminus \{v = 0\})$ and for all $u \neq 0$ and there is no smooth continuation from that set to $\mathbb{R} \times (\mathbb{R}^2 \setminus 0)$. There does not exist a global Lagrangian, since this would mean

$$dL = L_u du + L_v dv = \frac{v}{u^2 + v^2} du - \frac{u}{u^2 + v^2} dv = \Delta$$

and then

$$\oint_\gamma dL = 0$$

for every closed integral in E. But for $\gamma = (x,u(t),v(t)) = (x, \cos t, \sin t)$, $t \in [0,2\pi]$, which is a closed curve in E, we get

$$\oint_\gamma \left(\frac{v}{u^2 + v^2} du - \frac{u}{u^2 + v^2} dv \right) = \int_0^{2\pi} \left\langle \begin{pmatrix} \frac{v(t)}{u^2(t)+v^2(t)} \\ \frac{-u(t)}{u^2(t)+v^2(t)} \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \right\rangle dt = -2\pi \neq 0.$$

Furthermore, the source form satisfies translation symmetry ∂_x and corresponding conservation law. The idea is that

$$u_x^\alpha f_\alpha = u_x \frac{v}{u^2 + v^2} - v_x \frac{u}{u^2 + v^2} = D_x \arctan \frac{u}{v}$$

is a total derivative. But we have to verify that on the whole space $J^k E$, i.e. also where $v = 0$ and $u \neq 0$. Actually, we only need to verify $\mathcal{E}_\beta(u_x^\alpha f_\alpha) = 0$, since this is the local conservation law condition assumed in Theorem 1.0.2. We get

$$\begin{aligned}
\mathcal{E}_u(u_x^\alpha f_\alpha) &= (\partial_u - D_x \partial_{u_x})(\frac{u_x v}{u^2 + v^2} - \frac{v_x u}{u^2 + v^2}) = \\
&= -\frac{u_x v 2u}{(u^2 + v^2)^2} - \frac{v_x}{u^2 + v^2} + \frac{v_x u 2u}{(u^2 + v^2)^2} - D_x \frac{v}{u^2 + v^2} = \\
&= \frac{-2u_x vu - v_x(u^2 + v^2) + 2v_x u^2}{(u^2 + v^2)^2} - \frac{v_x}{u^2 + v^2} + \frac{v2(uu_x + vv_x)}{(u^2 + v^2)^2} = \\
&= \frac{-2u_x vu - 2v_x(u^2 + v^2) + 2v_x u^2 + 2v(uu_x + vv_x)}{(u^2 + v^2)^2} = 0.
\end{aligned}$$

A similar calculation holds for $\mathcal{E}_v(u_x^\alpha f_\alpha)$ (by $u \leftrightarrow v$ symmetry of Δ). Furthermore, the source form satisfies scale and rotation invariance

$$\begin{pmatrix} x \\ u \\ v \end{pmatrix} \to \begin{pmatrix} x \\ e^t u \\ e^t v \end{pmatrix} \quad \Leftrightarrow \quad V := u\partial_u + v\partial_v, \quad \text{(scale invariance)}$$

$$\begin{pmatrix} x \\ u \\ v \end{pmatrix} \to \begin{pmatrix} x \\ u \cos t - v \sin t \\ u \sin t + v \cos t \end{pmatrix} \quad \Leftrightarrow \quad W := -v\partial_u + u\partial_v \quad \text{(rotation invariance)},$$

since

$$\begin{aligned}\mathcal{L}_{\mathrm{pr}V}\Delta =&\mathcal{L}_{\mathrm{pr}V}\left(\frac{v}{u^2+v^2}du-\frac{u}{u^2+v^2}dv\right)\wedge dx=\\ =&\left(\mathcal{L}_{\mathrm{pr}V}\frac{v}{u^2+v^2}\right)du\wedge dx+\frac{v}{u^2+v^2}\mathcal{L}_{\mathrm{pr}V}(du\wedge dx)-\\ &-\left(\mathcal{L}_{\mathrm{pr}V}\frac{u}{u^2+v^2}\right)dv\wedge dx-\frac{u}{u^2+v^2}\mathcal{L}_{\mathrm{pr}V}(dv\wedge dx)=\\ =&\left(\frac{v}{u^2+v^2}-\frac{v2(u^2+v^2)}{(u^2+v^2)^2}+\frac{v}{u^2+v^2}\right)du\wedge dx-\\ &-\left(\frac{u}{u^2+v^2}-\frac{u2(u^2+v^2)}{(u^2+v^2)^2}+\frac{u}{u^2+v^2}\right)dv\wedge dx=0,\end{aligned}\tag{4.3}$$

where we use

$$\mathcal{L}_{\mathrm{pr}V}\frac{v}{u^2+v^2}=\frac{v}{u^2+v^2}-\frac{v(2u^2+2v^2)}{(u^2+v^2)^2}=-\frac{v(u^2+v^2)}{(u^2+v^2)^2}$$

and

$$\mathcal{L}_{\mathrm{pr}V}(du\wedge dx)=(\mathcal{L}_{\mathrm{pr}V}du)\wedge dx=du\wedge dx.$$

In a similar way we can compute it for W, where we use

$$\mathcal{L}_{\mathrm{pr}W}\frac{1}{u^2+v^2}=-\frac{\mathcal{L}_{\mathrm{pr}W}(u^2+v^2)}{(u^2+v^2)^2}=\frac{-v2u+u2v}{(u^2+v^2)^2}=0\tag{4.4}$$

and the same manipulations in the second and third line in (4.3), we get

$$\begin{aligned}\mathcal{L}_{\mathrm{pr}W}\Delta =&\frac{u}{u^2+v^2}du\wedge dx+\frac{v}{u^2+v^2}\underbrace{\mathcal{L}_{\mathrm{pr}W}(du\wedge dx)}_{=-dv\wedge dx}-\\ &-\frac{-v}{u^2+v^2}dv\wedge dx-\frac{u}{u^2+v^2}\underbrace{\mathcal{L}_{\mathrm{pr}W}(dv\wedge dx)}_{=du\wedge dx}=0.\end{aligned}$$

The corresponding (trivial) conservation laws are

$$\begin{aligned}Q^\alpha f_\alpha &= u\frac{v}{u^2+v^2}+v\frac{-u}{u^2+v^2}=0=D_xc,\quad c\in\mathbb{R},\\ Q^\alpha f_\alpha &= -v\frac{v}{u^2+v^2}+u\frac{-u}{u^2+v^2}=-1=D_x(-x).\end{aligned}\tag{4.5}$$

In (4.5) we can observe that the differential equation has no solution, but this does not contradict any assumptions in Theorem 1.0.2.

Let us check if we can span T_pE for every $p\in E$. The three symmetries are

$$\begin{pmatrix}\partial_x\\ u\partial_u+v\partial_v\\ u\partial_v-v\partial_u\end{pmatrix}=\begin{pmatrix}1&0&0\\0&u&v\\0&-v&u\end{pmatrix}\begin{pmatrix}\partial_x\\ \partial_u\\ \partial_v\end{pmatrix}$$

and since the determinant

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & u & v \\ 0 & -v & u \end{vmatrix} = u^2 + v^2 \neq 0 \quad \text{on } E,$$

we can span T_pE for every $p \in E$. When we add a non-degenerate second order variational source form to (4.2), which satisfies ∂_x,V,W symmetries and corresponding conservation laws, then we get a non-degenerate counter example of second order.

4.2. Counter Examples, Part II

In this section, we want to show that Theorem 1.0.3 is sharp. First, we want to investigate if there are any non-trivial source forms of 4th order which satisfy symmetries $V \in \mathcal{V}$ and corresponding conservation laws such that $\text{span}\{V_p : V \in \mathcal{V}\} = T_pE$ for all $p \in E$. Therefore, let us consider the source form

$$\Delta = (2u_{xx}u_{(3)} + u_x u_{(4)})du \wedge dx,$$

which satisfies ∂_x,∂_u symmetry and corresponding conservation laws

$$\begin{aligned} \partial_u : \quad & Qf = 1 \cdot f = D_x(\frac{1}{2}u_{xx}^2 + u_x u_{(3)}), \\ \partial_x : \quad & Qf = u_x f = D_x(u_x^2 u_{(3)}), \end{aligned}$$

and this source form is also variational with Lagrangian $L = u(u_{xx}u_{(3)} + \frac{1}{2}u_x u_{(4)})$.

<u>1st Counter example (the assumption non-singular is necessary)</u>: In this context, non-singular is equivalent to say that Δ is defined on the whole space J^4E. Let us consider the trivial fiber bundle $\pi : \mathbb{R}^2 \to \mathbb{R}$ with coordinates (x,u) and projection $\pi(x,u) = x$. Furthermore, we consider the singular source form

$$\Delta = -\frac{u_{(3)}}{u_{xx}^2} du \wedge dx = f du \wedge dx, \tag{4.6}$$

which satisfies ∂_x,∂_u symmetry and corresponding conservation laws of the form

$$\begin{aligned} \partial_u : \quad & Qf = 1 \cdot f = -\frac{u_{(3)}}{u_{xx}^2} = D_x \frac{1}{u_{xx}}, \\ \partial_x : \quad & Qf = u_x f = -\frac{u_x u_{(3)}}{u_{xx}^2} = D_x(\frac{u_x}{u_{xx}} - x). \end{aligned}$$

The source form is not variational since the Helmlholtz condition $f_{u_{(3)}} = 0$ is not satisfied. Therefore, if we allow singular source forms, then Theorem 1.0.3 is no longer true for third order source forms and the theorem is sharp in this sense.

2nd Counter example (the assumption $\operatorname{span}\{V_p : V \in \mathcal{V}\} = T_pE$ is necessary):
The following counter example was already discussed in Subsection 3.8.1. Again, let us consider the trivial fiber bundle from 1st Counter example above and we define the source form

$$\Delta := f du \wedge dx = (\frac{1}{2}u_x u_{(3)} + u_{xx}^2) du \wedge dx,$$

The source form satisfies translation symmetry ∂_x and corresponding conservation law

$$-u_x f = -\frac{1}{2}u_x^2 u_{(3)} - u_x u_{xx}^2 = -D_x(\frac{1}{2}u_x^2 u_{xx}).$$

The source form is not variational. The source form also satisfies ∂_u-symmetry, but not corresponding conservation law. Therefore, the assumption of having corresponding conservation laws is also necessary.

4.3. Applications and Corollaries

In this section, we want to investigate Theorem 1.0.2 in the light of applications, especially in physics. It is relatively easy to accept that fundamental physical theories should be invariant in some sense, as we motivated in Section 1.2. Moreover, it is also relatively easy to define what we mean with invariance, i.e. we apply the Lie derivative to certain objects and investigate when the expressions vanish or do not vanish. It is much more complicated to find reasonable conservation laws, connected to the symmetries or not. For example, if the physical theory should satisfy some kind of energy conservation, then the question is of course: What is the energy (function) in general and how do we define it? For example, what is the energy function for fourth order differential equations?, or does it only make sense to talk about energy for second order differential equations? To investigate and formulate it more precisely would be a topic of its own and we are not able to do that here. It is also possible, and even most likely, that there does not exist such a definition of energy for very general classes of differential equations. In the following, our focus is rather to discuss some examples, explain the relation $Q^\alpha = V^\alpha - u_i^\alpha V^i$ between symmetries and conservation laws, and investigate the conservation laws, or better differential identities, in Noether's second theorem. Our standard notation is that the vector field $V = V^i\partial_{x^i} + V^\alpha\partial_{u^\alpha} \in \mathcal{V}$ describes a symmetry of Δ.

The idea of this section is to replace the relation $Q^\alpha = V^\alpha - u_i^\alpha V^i$ by *divergence-free* or related conditions and to apply Noether's second theorem. In many cases in Noether's second theorem, we do not have the components V^α,V^i in the conservations laws or differential identities. More precisely, the arbitrary functions p in Noether's second theorem do not occur explicitly in the conservation laws, as we explained in Section 2.10. Only the components a^α,b^α,c,d do occur and in many cases these are constants. We also know that symmetries are described by vector fields and a vector space structure, where such constants do not have an explicit meaning in applications (only how they relate different components has a meaning). In this sense we do not directly have such a strong relation between symmetries and corresponding conservation laws, as we have in Noether's first theorem, where we have the condition $(V^\alpha - u_i^\alpha V^i)f_\alpha = D_iC^i$. On the other hand, in Noether's second theorem we need an infinite dimensional symmetry group and this seems to be a stronger condition as in Noether's first theorem. However, infinite dimensional symmetry groups can be explained much better in some applications as the relation in Noether's first theorem. For example, let γ be a curve in $\mathbb{R}^n$. Then we can define the curvature κ of such curves γ. Furthermore, let us think of a physical differential equation which depends on the curvature κ and derivatives of κ (derivatives in the sense of derivation with respect to arc length). The curvature is invariant under reparametrization of γ and this is an infinite dimensional symmetry group (diffeomorphism group). Since the curvature could be of interest in applications, it is obvious that diffeomorphism symmetries could be of interest in applications and it does not seem to be a strong restriction from that point. For example, Einstein's field equation has diffemorphism invariance. We will now explain this in more detail and discuss applications.

Before we do so, it is fair to say that finding applications is not easy and we would have to spend more effort on that. One reason is that applications can get quite complicated and we would have to introduce more notation. Hopefully we can solve these problems in future and when we systematically write down physical meaningful symmetries and conservation laws or differential identities.

The conservation law condition $Q^\alpha f_\alpha = D_i C^i$ is problematic from an application point of view for at least two reasons, where the first reason is more drastic as the second:

- From an application point of view, we cannot explain the very special relation $Q^\alpha = V^\alpha - u_i^\alpha V^i$. We can only show that $Q^\alpha f_\alpha = D_i C^i$ if f_α is non-degenerate. But in general, Q^α does not have the special form $Q^\alpha = V^\alpha - u_i^\alpha V^i$, even if f_α is non-degenerate. (For the definition of non-degenerate see (Olv86, p.171).)

- Non-degenerate is a nice mathematical condition, but how can we explain it from an application point of view.

In some situations, the condition *divergence-free* or *vanishing of covariant derivative* solves both of these problems. Note that divergence-free and vanishing of covariant derivative have sometimes the same meaning. However, we want to distinguish them in the following as we will explain below. **Divergence-free** is nothing else than $\operatorname{div} f = 0$. This only works when $n = m$ and it can also be written as $\operatorname{div} f = D_i f^i = 0$, where we write f^i with upper indices $i = 1,2,...,n$ instead of f_α. **Vanishing covariant derivative** is nothing else than $\nabla_{i;} f^i = 0$, where $\nabla_{i;}$ is the covariant derivative in some metric field theory. Note that in (AP12, p.4) divergence-free means vanishing covariant derivative in our notation here. We want to distinguish metric fields from vector fields, diffeomorphism invariance of metrics and gauge transformations for vector fields. The reason is that gauge transformations are vertical fields, see (AP96, p.370) and (MPV08, p.3), but the transformations for vanishing covariant derivative (AP12, p.3) are non-vertical, and therefore they can be quite different (especially when we want to investigate if they span T_pE or T_pJ^1E and so on).

Let us start the discussion from the perspective of physics. Maxwell's- and Einstein's-field equation satisfy the divergence-free or vanishing covariant derivative condition, and in this case, it is also called charge-,energy- and mass-conservation. Note that in the following, we will always write f^i for the differential equation instead of

f_α. Formally, let $f^i = 0$ be Maxwell's- or Einstein's equation in vacuum, where i has to be chosen such that it labels all components of the corresponding differential equation. Note that Einstein's equation is usually written as $G_{\mu\nu} = 0$ and (one of) Maxwell's equation is written as $\partial_\mu F^{\mu\nu} = 0$. The equation where we have charge- or mass sources are formally written as $f^i = J^i$ where (J^i) is the current or energy-momentum-tensor. The equation $f^i = J^i$ is called equation in matter. In experiments we observe that we have charge- and mass-conservation and this is formulated as $D_i J^i = 0$ or $\nabla_{i;} J^i = 0$. Now it is reasonable to assume that this property is transferred to f^i, i.e. $\operatorname{div} f = D_i f^i = 0$ or $\nabla_{i;} f^i = 0$, where these identities hold for all values in $J^k E$, not only for solutions of the differential equation. That physical equations in matter can be written as $f = J$ can probably be found out in experiments. For example, in the case of Einstein's field equation, $f = J$, or as it is usually written $G_{\mu\nu} = T_{\mu\nu}$, roughly says that the mass is the source which curves the space and the curvature is somehow proportional to the presence of mass or energy (this is just the rough idea). In the case of Maxwell's equation, $f = J$, or as it is usually written $\partial_\mu F^{\mu\nu} = J^\nu$, roughly says that the charge and current is the source which generates the electric and magnetic fields, and the electromagnetic field is somehow proportional to the presence of charge and current (this is just the rough idea and we did not write the homogeneous Maxwell's equation). We could also say that this is in some sense the simplest coupling method of different kinds of fields, like metric fields to energy and electromagnetic fields. This of course has to be explained in more detail. But for us this should be sufficient motivation here. Now we want to find the link between $\operatorname{div} f = 0$ or $\nabla_{i;} f^i = 0$ and $Q^\alpha = V^\alpha - u_i^\alpha V^i$, i.e. we want to find the correspondence of symmetries and conservation laws.

Let us first explain the idea before we show a concrete example. Vanishing covariant derivative will be explained below, it is a bit more complicated. If (f^i) is divergence-free, i.e. if $D_i f^i = 0$, then also $p D_i f^i = 0$ for every function $p = p(x)$ (we could also allow $p = p(x,u,u_x,...)$). Using partial integration, we can show that $\operatorname{div} f = 0$ provides a conservation law of the form

$$0 = p D_i f^i = D_i(p f^i) - (D_i p) f^i$$

for every function $p = p(x)$ and for every point in $J^k E$. This equation can we rewritten as

$$(D_i p) f^i = D_i(p f^i) = D_i C^i, \quad \text{where } C^i := p f^i$$

and $Q_i = D_i p$ is the characteristic in this case. Usually, we wrote $Q^\alpha = V^\alpha - u_i^\alpha V^i$ for the characteristics, but the notation here is $Q_i = V_i - u_{i,j} V^j = D_i p$, where we write u_i instead of u^α. Higher order jet coordinates are written as $u_{i,j} = D_j u_i$ and more generally $u_{i,I} = D_I u_i$, where I is a multi-index.

Since $p = p(x)$, and therefore we do not have a $u_{i,j}$-coordinate in the characteristic, the only possibility for the corresponding projectable symmetry vector field is

$V_i = D_i p$. Considering the symmetry vector field $V = V_i \partial_{u_i}$ and the corresponding transformation, we observe that

$$u_i \to u_i + t D_i p, \quad t \in \mathbb{R}$$

and this is called **gauge transformation**. Roughly speaking, we can add a gradient field to the field u_i and this does not change the considered differential equation or source form. Therefore, we found the link between symmetries and conservation laws and it is

$$\begin{aligned} D_i f^i &= 0, \quad \text{conservation law, } \textbf{divergence-free}, \\ V &= (D_i p)\partial_{u_i}, \quad \text{for all } p = p(x), \ \textbf{gauge transformation}. \end{aligned} \tag{4.7}$$

Since we can choose any function p it is an infinity dimensional symmetry group, and therefore related to Noether's second theorem. Sometimes, divergence-free is also called **charge-conservation**. Also see (2.126), where we discuss the physical meaning of conservation laws.

The idea is that we would like to prove the following corollary of Theorem 1.0.2:

<u>Formal Corollary of Theorem 1.0.2:</u> Let $n = m$. If a second order source form $\Delta = f^i du_i \wedge dx^1 \wedge \ldots \wedge dx^n$ is gauge invariant and f^i is divergence-free, then Δ must be locally variational.

But it turns out that the corollary is not true as it is formulated and we need additional conditions. The problem is that the corresponding symmetries of gauge transformations do not span T_pE for every $p \in E$ (gauge transformations are vertical vector fields). Therefore, we need additional symmetries and corresponding conservation laws to be able to span T_pE. Otherwise the corollary is not true in general (also see the counter examples in Section 4.1).

The next question which arises is if $\operatorname{div} f = 0$ can also be used for symmetries, where $\pi_* V \neq 0$ or if we need additional conservation laws or differential identities. To answer this question, let us consider the following: What are physically reasonable symmetries? Let us assume **translation symmetries** of the x coordinates, i.e. $V = \partial_{x^i}$. From a physical point of view, on the base manifold M, we have Poincaré transformations and translations of the x-coordinates are a very special case of that. The characteristic for the vector field ∂_{x^i} is $Q_k = V_k - u_{k,i}V^i = -u_{k,i}$ and a simple calculation

$$Q_k f^k = -u_{k,i} f^k = -(D_i u_k) f^k = -D_i(u_k f^k) + u_k \underbrace{D_i f^k}_{\neq \operatorname{div} f}$$

shows that we cannot use $\operatorname{div} f = 0$ as corresponding conservation law for the ∂_{x^i}-symmetries (at least not so easily). Therefore, we need to assume further that

$-u_{k,i}f^k$ is a divergence expression. This is problematic when we cannot explain the physical meaning of this conservation law. As far as we know it has no physical meaning, or the meaning is not yet known. Although the ∂_{x^i}-symmetry and corresponding conservation law seems to be some kind of energy conservation, the exact interpretation is missing.

Note that in (AP96, p.370) and (MPV08, p.3) $-u_{k,i}f^k$ is required to be a divergence expression and it cannot be derived from the divergence-free condition $D_i f^i = 0$. But as it is shown in (AP96, p.371), gauge symmetry, ∂_{x^i}-symmetry and the assumption that $-u_{k,i}f^k$ is a divergence expression implies divergence-free. However, we would like to have the opposite direction. Anyway, we will not further discuss this problem here. Instead, we want to consider metric field theories below. Before we discuss this problem, we have to say a few words about vector- and metric field theories.

Let A_i describe the vector potential in Maxwell's equation and g_{ij} the metric in a metric field theory. We think of Einstein's field equation, however, we consider the Riemannian case for simplicity. The vector potential A_i transforms according to the 1-form $A := A_i dx^i$ and the metric g_{ij} according to $g = g_{ij}dx^i \otimes dx^j$. Therefore, in the discussion above we ignored some non-trivial transformations for A_i and g_{ij}. Let $\xi^i\partial_{x^i}$ be a vector field on M. This vector field induces a transformation for A and g in the following form

$$\begin{aligned} A = A_i dx^i \qquad &\Rightarrow \qquad \xi^i\partial_{x^i} - \xi^k_{,i}A_k\partial_{A_i} \in \mathfrak{X}(E), \\ g = g_{ij}dx^i \otimes dx^j \qquad &\Rightarrow \qquad \xi^i\partial_{x^i} - 2\xi^i_{,(k}g_{l)i}\partial_{g_{kl}} \in \mathfrak{X}(E). \end{aligned} \tag{4.8}$$

The transformation for g can be found in (AP12, p.3). When doing a more detailed discussion, this has to be taken in account of course and can change the situation which we discussed above. This depends on which kinds of symmetries we are investigating. Note that in the case where we assume the gauge transformation (4.7) this does not change what we have discussed above.

Metric field theories: In metric field theories we have the coordinates

$$(x^i,g_{kl},g_{kl,j},...,g_{kl,I})$$

on J^kE, where I is a multi-index of length k and all indices have values in 1,2,...,n. Furthermore, g_{kl} is symmetric in k,l and $\det(g_{kl}) \neq 0$. The source form is given as

$$\Delta = f^{ij}dg_{ij} \wedge dx^1 \wedge ... \wedge dx^n,$$

where we now consider f^{ij} instead of f_α, see (AP12, p.6). We assume diffeomorphism invariance of g, when pull-backed to M. This transformation is described by the vector field in (4.8). The corresponding characteristics $Q^\alpha = V^\alpha - u_i^\alpha V^i$ are in this case of the form

$$Q_{kl} = -2\xi^i_{,(k} g_{l)i} - \xi^i g_{kl,i}. \tag{4.9}$$

The expression (4.9) can also be found in (AP12, p.6). The conservation laws are of the form

$$Q_{kl} f^{kl} = -(2\xi^i_{,(k} g_{l)i} + \xi^i g_{kl,i}) f^{kl} = ...$$

To complete this equation, we have to investigate the Christoffel symbols Γ^b_{cd}, which are defined as

$$\Gamma^b_{cd} = \frac{1}{2} g^{bk} (g_{dk,c} + g_{ck,d} - g_{cd,k}).$$

We also need the identity $g_{bc} g^{bk} = \delta^k_c$, i.e. g^{bk} is the inverse matrix of g_{bc}. Then we get

$$\begin{aligned} g_{be}\Gamma^b_{cd} &= \frac{1}{2} g_{be} g^{bk} (g_{dk,c} + g_{ck,d} - g_{cd,k}) = \\ &= \frac{1}{2}\delta^k_e (g_{dk,c} + g_{ck,d} - g_{cd,k}) = \\ &= \frac{1}{2}(g_{de,c} + g_{ce,d} - g_{cd,e}). \end{aligned}$$

Let us investigate the following equation:

$$\begin{aligned} 0 &= \xi^e g_{be}(D_a f^{ab} + \Gamma^b_{cd} f^{cd}) = \\ &= D_a(\xi^e g_{be} f^{ab}) - f^{ab} D_a(\xi^e g_{be}) + \xi^e g_{be} \Gamma^b_{cd} f^{cd} = \\ &= D_a(\xi^e g_{be} f^{ab}) - f^{ab} D_a(\xi^e g_{be}) + \frac{1}{2}\xi^e \underbrace{(g_{de,c} + g_{ce,d} - g_{cd,e}) f^{cd}}_{c\to a,\ d\to b} = \\ &= D_a(\xi^e g_{be} f^{ab}) - f^{ab}(\xi^e_{,a} g_{be} + \xi^e g_{be,a}) + \frac{1}{2}\xi^e (g_{be,a} + g_{ae,b} - g_{ab,e}) f^{ab} = \\ &= D_a(\xi^e g_{be} f^{ab}) + [-(\xi^e_{,a} g_{be} + \xi^e g_{be,a}) + \frac{1}{2}\xi^e (g_{be,a} + g_{ae,b} - g_{ab,e})] f^{ab} = \\ &= D_a(\xi^e g_{be} f^{ab}) + [-\xi^e_{,a} g_{be} + \frac{1}{2}\xi^e (-g_{be,a} + g_{ae,b} - g_{ab,e})] f^{ab}. \end{aligned} \tag{4.10}$$

In (4.10) we assume that f^{ab} is symmetric in a,b. Then we can write

$$\begin{aligned} 0 &= \xi^e g_{be}(D_a f^{ab} + \Gamma^b_{cd} f^{cd}) = \\ &= D_a(\xi^e g_{be} f^{ab}) + [-\xi^e_{,(a} g_{b)e} + \frac{1}{2}\xi^e(-g_{ab,e})] f^{ab} = \\ &= D_a(\xi^e g_{be} f^{ab}) - \frac{1}{2}(2\xi^e_{,(a} g_{b)e} + \xi^e g_{ab,e}) f^{ab}. \end{aligned}$$

This means that the characteristic in (4.9) generates a conservation law (we have to multiply the expression by 2) if the covariant derivative

$$0 = D_a f^{ab} + \Gamma^b_{cd} f^{cd} \tag{4.11}$$

vanishes. This is also well-know by Noether's second theorem. When we can show that the vector fields in (4.8) span T_pE at every $p \in E$, then we get the following formal corollary:

Formal Corollary of Theorem 1.0.2: If a second order symmetric source form $\Delta = f^{ij} dg_{ij} \wedge dx^1 \wedge \ldots \wedge dx^n$ is diffeomorphism invariant on M, i.e. under the vector fields (4.8) *and the covariant derivative* (4.11) *vanishes, then Δ must be locally variational.*

Such a theorem is also proven in (AP12) for third order source forms and it is a highly complicated calculation. Note that it is a non-trivial calculation to show (or disprove) that the vector fields in (4.8) span T_pE at every $p \in E$ and we will not investigate this problem in more detail here.

As far as we understand the problem, the setting with the metric fields seems to be one of the rare direct applications of Theorem 1.0.2, where vanishing of covariant derivative and diffeomorphism invariance can be assumed to be more or less necessary assumptions in applications in physics. Historically, Einstein had different versions of his metric field equations and some of them did not satisfy the necessary conservation laws. Later, it turned out that the field equation, which can be derived from the Einstein Hilbert functional, describes the right physical laws. Therefore, it could be possible that this setting is also historically mostly relevant, when searching for applications of Theorem 1.0.2 and Takens' problem. Hopefully we are able to find more applications in the future.

4.4. Do Coordinate Transformations Simplify the Problem

In this short section, we want to investigate the following question: Are there local coordinates, such that a projectable vector field

$$V = V^i(x^j)\partial_{x^i} + V^\alpha(x^j,u^\beta)\partial_{u^\alpha} \quad \in \mathfrak{X}(E)$$

has constant coefficients? More generally, let $\{V_{\mathscr{A}}\ , \mathscr{A} = 1,2,...,n+m\}$ be a set of projectable vector fields on E. Can we take linear combinations over $\mathbb{R}$ and find local coordinates, such that $\{V_{\mathscr{A}}\ , \mathscr{A} = 1,2,...,n+m\} = \{\partial_{x^i},\partial_{u^\alpha}\ , i = 1,...,n,\ \alpha = 1,...,m\}$, i.e. such that these vector fields describe translation symmetries in a certain coordinate system on E. If the symmetry vector fields in Theorem 1.0.2 or 1.0.3 are $\{\partial_{x^i},\partial_{u^\alpha}\ , i = 1,...,n,\ \alpha = 1,...,m\}$ then we can immediately derive from the ECS that $H_{\alpha\beta} = 0$ and this simplifies the proof in some sense. In the following, we only consider the case $n,m = 1$ for simplicity.

According to the above question, we cannot always find such local coordinates. For example, if V is vertical, i.e. $\pi_* V = 0$ then we cannot choose coordinates such that $V = \partial_x$, because then $\pi_* V \neq 0$. Second example, in the case where we have two vector fields V,W on E, we can also use the vector space structure and try to find simpler vector fields, by taking linear combinations over $\mathbb{R}$ and doing local coordinate transformations, as well. Let us consider $n,m = 1$ and we assume $\operatorname{span}\{V_p, W_p\} = T_pE$ for every $p \in E$. We want to find local coordinates and $\lambda_1,\lambda_2,\lambda_3,\lambda_4 \in \mathbb{R}$ such that

$$\begin{aligned} \lambda_1 V + \lambda_2 W &= \partial_x, \\ \lambda_3 V + \lambda_4 W &= \partial_u. \end{aligned} \tag{4.12}$$

The vector fields ∂_x,∂_u in (4.12) satisfy $[\partial_x,\partial_u] = 0$. Therefore, we get

$$\begin{aligned} 0 = [\partial_x,\partial_u] &= [\lambda_1 V + \lambda_2 W,\lambda_3 V + \lambda_4 W] = \\ &= \lambda_1\lambda_4[V,W] + \lambda_2\lambda_3[W,V] = (\lambda_1\lambda_4 - \lambda_2\lambda_3)[V,W]. \end{aligned} \tag{4.13}$$

This means that if the linear combination is invertible, i.e. $(\lambda_1\lambda_4 - \lambda_2\lambda_3) \neq 0$, then this implies $[V,W] = 0$. Moreover, the linear combination in (4.12) must be invertible, since ∂_x,∂_u are linearly independent. More precisely, let us consider

$$\begin{aligned} a\partial_x + b\partial_u &= a(\lambda_1 V + \lambda_2 W) + b(\lambda_3 V + \lambda_4 W) = \\ &= (a\lambda_1 + b\lambda_3)V + (a\lambda_2 + b\lambda_4)W, \quad a,b \in \mathbb{R}. \end{aligned} \tag{4.14}$$

If the linear combination in (4.12) would not be invertible, then

$$\begin{pmatrix} \lambda_1 & \lambda_3 \\ \lambda_2 & \lambda_4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0$$

in (4.14) would have a non-trivial solution $(a,b) \neq 0$ which is a contradiction. In general, the commutator $[V,W]$ does not always vanish of course. For example, for the vector fields $V = \partial_x$ and $W = e^x \partial_u$ we get $[V,W] \neq 0$. With the following lemma and proposition we want to investigate this question more systematically. More generally, we would like to prove them for arbitrary n,m, but for simplicity, we only consider the case where $n,m = 1$.

Lemma 4.4.1. *Let $n,m = 1$ and V,W be projectable vector fields on E such that span$\{V_p,W_p\} = T_p$ for every $p \in E$. Then, the commutator $[V,W]$ vanishes if and only if we can take linear combinations over $\mathbb{R}$ and find local coordinates such that $\tilde{V} = \partial_x, \tilde{W} = \partial_u$, where $\tilde{V} = V\lambda_1 + W\lambda_2$ and $\tilde{W} = \lambda_3 V + \lambda_4 W$ and $\lambda_1,\lambda_2,\lambda_3\lambda_4 \in \mathbb{R}$.*

Proposition 4.4.2. *Let $n,m = 1$ and V,W be projectable vector fields on E such that span$\{V_p,W_p\} = T_pE$ for every $p \in E$ and $[V,W] = 0$. Then we can take linear combinations over $\mathbb{R}$ and find local coordinates such that*

$$\begin{aligned} &\lambda_1 V + \lambda_2 W = \partial_v, \\ &\lambda_3 V + \lambda_4 W = \partial_y + B(y)\partial_v, \quad \lambda_1,\lambda_2,\lambda_3,\lambda_3 \in \mathbb{R}. \end{aligned} \tag{4.15}$$

Furthermore, in any local coordinates there exists a function $C = C(x,u)$ and a function $F = F(x,u) \neq 0$ for all $p \in E$ such that V,W have the following form:

$$\begin{aligned} V &= c_1 A(x)\partial_x + c_1(C(x,u) + F(x,u))\partial_u, \qquad (\pi_* V \neq 0), \\ W &= c_2 A(x)\partial_x + c_2(C(x,u) - F(x,u))\partial_u, \qquad (\pi_* V \neq 0), \end{aligned} \tag{4.16}$$

where $c_1,c_2 \in \mathbb{R} \setminus 0$ and the function A vanishes nowhere, or

$$\begin{aligned} V &= c_1 A(x)\partial_x + c_1 C(x,u)\partial_u, \qquad (\pi_* V \neq 0), \\ W &= c_2 F(x,u)\partial_u, \qquad (\pi_* W = 0), \end{aligned} \tag{4.17}$$

or

$$\begin{aligned} V &= c_1 F(x,u)\partial_u, \qquad (\pi_* V = 0), \\ W &= c_2 A(x)\partial_x + c_2 C(x,u)\partial_u, \qquad (\pi_* W \neq 0). \end{aligned} \tag{4.18}$$

Roughly speaking, Proposition 4.4.2 says that we can find simple local coordinate expressions, like (4.15), if and only if the vector fields are already written in a very simple form, namely of the form (4.16), (4.17) or (4.18). More precisely, in any local coordinates it is relatively easy to see that we can take linear combinations over $\mathbb{R}$ such that one of the vector fields is vertical. This is obvious for (4.17) and (4.18), but it also holds for (4.16). Moreover, with the help of Proposition 4.4.2 we can prove Lemma 4.4.1.

Proof of Proposition 4.4.2: By assumptions we have $[V,W] = 0$, i.e.

$$\begin{aligned}0 &= [V,W] = [V^x\partial_x + V^u\partial_u, W^x\partial_x + W^u\partial_u] = \\ &= [V^x\partial_x, W^x\partial_x] + [V^x\partial_x, W^u\partial_u] + [V^u\partial_u, W^x\partial_x] + [V^u\partial_u, W^u\partial_u] = \\ &= (V^xW^x_x - W^xV^x_x)\partial_x + V^xW^u_x\partial_u - W^xV^u_x\partial_u + (V^uW^u_u - W^uV^u_u)\partial_u = \\ &= (V^xW^x_x - W^xV^x_x)\partial_x + (V^xW^u_x + V^uW^u_u - W^xV^u_x - W^uV^u_u)\partial_u. \end{aligned} \tag{4.19}$$

The coefficient in front of ∂_x tells us

$$0 = V^xW^x_x - W^xV^x_x = \left\langle \begin{pmatrix} V^x \\ W^x \end{pmatrix}, \begin{pmatrix} W^x_x \\ -V^x_x \end{pmatrix} \right\rangle.$$

This means that $(W^x_x, -V^x_x)$ must be orthogonal to (V^x, W^x). Since we can span T_pE for every $p \in E$, the coefficient V^x and W^x cannot vanish at the same point, and therefore

$$\begin{pmatrix} W^x_x \\ -V^x_x \end{pmatrix} = \kappa \begin{pmatrix} W^x \\ -V^x \end{pmatrix}$$

describes the one-dimensional solution space, where $\kappa = \kappa(x)$. The solution of these two differential equations are

$$W^x = c_2 \exp(\int \kappa dx), \quad V^x = c_1 \exp(\int \kappa dx), \tag{4.20}$$

where $c_1, c_2 \in \mathbb{R}$ and $A := \exp(\int \kappa dx)$. By definition of A, it is clear that A vanishes nowhere (this can also be seen by the condition $\operatorname{span}\{V_p, W_p\} = T_pE$). Now let us again consider (4.19), which can now be written as

$$\begin{aligned}0 &= V^xW^u_x + V^uW^u_u - W^xV^u_x - W^uV^u_u = \\ &= (c_1W^u_x - c_2V^u_x)A + (V^uW^u_u - W^uV^u_u).\end{aligned} \tag{4.21}$$

The coefficients V^x, W^x in (4.20) show that c_1, c_2 cannot both vanish at the same time, since we can span T_pE at every $p \in E$. We have to distinguish the two cases:

- $c_1 \neq 0$ and $c_2 \neq 0$,
- without loss of generality $c_1 \neq 0$ and $c_2 = 0$.

In the following, we will also need local coordinate transformations of vector fields, which can be found in Proposition 2.4.3 and they are of the form

$$\begin{aligned}V &= V^x(x)\partial_x + V^u(x,u)\partial_u = V^x(\frac{\partial y}{\partial x}\partial_y + \frac{\partial v}{\partial x}\partial_v) + V^u\frac{\partial v}{\partial u}\partial_v = \\ &= V^x\frac{\partial y}{\partial x}\partial_y + (V^x\frac{\partial v}{\partial x} + V^u\frac{\partial v}{\partial u})\partial_v\end{aligned}$$

and

$$\begin{aligned}W &= W^x(x)\partial_x + W^u(x,u)\partial_u = W^x(\frac{\partial y}{\partial x}\partial_y + \frac{\partial v}{\partial x}\partial_v) + W^u\frac{\partial v}{\partial u}\partial_v = \\ &= W^x\frac{\partial y}{\partial x}\partial_y + (W^x\frac{\partial v}{\partial x} + W^u\frac{\partial v}{\partial u})\partial_v.\end{aligned}$$

First Case ($c_1 \neq 0$ and $c_2 \neq 0$): In this case, we take linear combinations over $\mathbb{R}$, such that

$$\begin{aligned}&c_2V - c_1W = \\ &= (c_2V^x\frac{\partial y}{\partial x} - c_1W^x\frac{\partial y}{\partial x})\partial_y + \left[c_2(V^x\frac{\partial v}{\partial x} + V^u\frac{\partial v}{\partial u}) - c_1(W^x\frac{\partial v}{\partial x} + W^u\frac{\partial v}{\partial u})\right]\partial_v = \\ &= (c_2V^x - c_1W^x)\frac{\partial y}{\partial x}\partial_y + \left[(c_2V^x - c_1W^x)\frac{\partial v}{\partial x} + (c_2V^u - c_1W^u)\frac{\partial v}{\partial u}\right]\partial_v = \\ &= (c_2V^u - c_1W^u)\frac{\partial v}{\partial u}\partial_v\end{aligned}$$

and

$$\begin{aligned}&c_2V + c_1W = \\ &= (c_2V^x\frac{\partial y}{\partial x} + c_1W^x\frac{\partial y}{\partial x})\partial_y + \left[c_2(V^x\frac{\partial v}{\partial x} + V^u\frac{\partial v}{\partial u}) + c_1(W^x\frac{\partial v}{\partial x} + W^u\frac{\partial v}{\partial u})\right]\partial_v = \\ &= (c_2V^x + c_1W^x)\frac{\partial y}{\partial x}\partial_y + \left[(c_2V^x + c_1W^x)\frac{\partial v}{\partial x} + (c_2V^u + c_1W^u)\frac{\partial v}{\partial u}\right]\partial_v = \\ &= 2c_1c_2A\frac{\partial y}{\partial x}\partial_y + \left[2c_1c_2A\frac{\partial v}{\partial x} + (c_2V^u + c_1W^u)\frac{\partial v}{\partial u}\right]\partial_v,\end{aligned}$$

where the determinant of the transformation does not vanish, i.e.

$$\begin{pmatrix} c_2 & -c_1 \\ c_2 & c_1 \end{pmatrix}\begin{pmatrix} V \\ W \end{pmatrix}, \quad \begin{vmatrix} c_2 & -c_1 \\ c_2 & c_1 \end{vmatrix} = 2c_1c_2 \neq 0.$$

Since A vanishes nowhere, as well as $\frac{\partial y}{\partial x}$, we can do local coordinate transformations such that

$$2c_1c_2A\frac{\partial y}{\partial x} = 1.$$

Furthermore, since V,W span T_pE at every $p \in E$, we get $c_2V^u - c_1W^u \neq 0$ everywhere, otherwise

$$\begin{aligned}c_2V &= c_2(V^x\partial_x + V^u\partial_u) = c_2(c_1A\partial_x + V^u\partial_u), \\ c_1W &= c_1(W^x\partial_x + W^u\partial_u) = c_1(c_2A\partial_x + W^u\partial_u)\end{aligned}$$

would be linearly dependent over $\mathbb{R}$. Then we can do the local coordinate transformations such that we also get

$$(c_2V^u - c_1W^u)\frac{\partial v}{\partial u} = 1. \tag{4.22}$$

Then we derive the following identity

$$\begin{aligned}
&(c_2V^u - c_1W^u)(c_2V^u_u + c_1W^u_u) - (c_2V^u_u - c_1W^u_u)(c_2V^u + c_1W^u) = \\
=&[c_2^2V^uV^u_u + c_1c_2V^uW^u_u - c_1c_2W^uV^u_u - c_1^2W^uW^u_u] - \\
&- [c_2^2V^u_uV^u + c_1c_2V^u_uW^u - c_1c_2W^u_uV^u - c_1^2W^u_uW^u] = \\
=&2c_1c_2(V^uW^u_u - W^uV^u_u).
\end{aligned}$$

We use this identity, (4.22) and the identity

$$\begin{aligned}
0 = \partial_u \cdot 1 = \partial_u[(c_2V^u - c_1W^u)\frac{\partial v}{\partial u}] = \\
= (c_2V^u_u - c_1W^u_u)\frac{\partial v}{\partial u} + (c_2V^u - c_1W^u)\frac{\partial^2 v}{\partial u^2}
\end{aligned}$$

to derive

$$\begin{aligned}
&2c_1c_2(V^uW^u_u - W^uV^u_u) = \\
&= (c_2V^u - c_1W^u)(c_2V^u_u + c_1W^u_u) - (c_2V^u_u - c_1W^u_u)(c_2V^u + c_1W^u) = \\
&= \frac{1}{\frac{\partial v}{\partial u}}\underbrace{\frac{\partial v}{\partial u}(c_2V^u - c_1W^u)}_{=1}(c_2V^u_u + c_1W^u_u) - \frac{1}{\frac{\partial v}{\partial u}}\underbrace{\frac{\partial v}{\partial u}(c_2V^u_u - c_1W^u_u)}_{-(c_2V^u - c_1W^u)\frac{\partial^2 v}{\partial u^2}}(c_2V^u + c_1W^u) = \\
&= \frac{1}{\frac{\partial v}{\partial u}}(c_2V^u_u + c_1W^u_u) + \frac{1}{\frac{\partial v}{\partial u}}(c_2V^u - c_1W^u)\frac{\partial^2 v}{\partial u^2}(c_2V^u + c_1W^u) = \\
&= \frac{1}{\frac{\partial v}{\partial u}}(c_2V^u_u + c_1W^u_u) + \frac{1}{\left(\frac{\partial v}{\partial u}\right)^2}\underbrace{\frac{\partial v}{\partial u}(c_2V^u - c_1W^u)}_{=1}\frac{\partial^2 v}{\partial u^2}(c_2V^u + c_1W^u) = \\
&= \frac{1}{\left(\frac{\partial v}{\partial u}\right)^2}\left[\frac{\partial v}{\partial u}(c_2V^u_u + c_1W^u_u) + \frac{\partial^2 v}{\partial u^2}(c_2V^u + c_1W^u)\right] = \\
&= \frac{1}{\left(\frac{\partial v}{\partial u}\right)^2}\partial_u[\frac{\partial v}{\partial u}(c_2V^u + c_1W^u)] = 2c_1c_2(c_2V^u_x - c_1W^u_x)A,
\end{aligned} \tag{4.23}$$

where we used (4.21) in the last step. Now we apply the partial derivative ∂_u on the second coefficient in the vector field $c_2V + c_1W$, i.e. on

$$2c_1c_2A\frac{\partial v}{\partial x} + (c_2V^u + c_1W^u)\frac{\partial v}{\partial u} \tag{4.24}$$

and we use (4.23) to derive

$$
\begin{aligned}
&2c_1c_2A\frac{\partial^2 v}{\partial x\partial_u} + \partial_u[(c_2V^u + c_1W^u)\frac{\partial v}{\partial u}] = \\
=&2c_1c_2A\partial_x\frac{\partial v}{\partial u} + \partial_u[(c_2V^u + c_1W^u)\frac{\partial v}{\partial u}] = \\
=&2c_1c_2A\partial_x\frac{\partial v}{\partial u} + \left(\frac{\partial v}{\partial u}\right)^2 2c_1c_2A\partial_x(c_2V^u - c_1W^u) = \\
=&2c_1c_2A\left[\partial_x\frac{\partial v}{\partial u} + \left(\frac{\partial v}{\partial u}\right)^2\partial_x\frac{1}{\frac{\partial v}{\partial u}}\right] = 0.
\end{aligned}
$$

This means that the coefficient (4.24) does not depend on u, and therefore does not depend on v, but it can depend on x or on y and we can write

$$
2c_1c_2A\frac{\partial v}{\partial x} + (c_2V^u + c_1W^u)\frac{\partial v}{\partial u} = B(y)
$$

for some function $B = B(y)$. Therefore, we derived

$$
\begin{aligned}
c_2V - c_1W &= \partial_v, \\
c_2V + c_1W &= \partial_y + B(y)\partial_v.
\end{aligned}
$$

Now let us consider

$$
\begin{aligned}
2c_2V &= (c_2V - c_1W) + (c_2V + c_1W) = \partial_v + \partial_y + B(y)\partial_v = \partial_y + (1 + B(y))\partial_v, \\
2c_1W &= (c_2V + c_1W) - (c_2V - c_1W) = \partial_y + B(y)\partial_v - \partial_v = \partial_y + (B(y) - 1)\partial_v.
\end{aligned}
$$

Furthermore, let us consider arbitrary fiber preserving local coordinate transformations of the vector fields

$$
\begin{aligned}
&\partial_y + (1 + B(y))\partial_v, \\
&\partial_y + (B(y) - 1)\partial_v,
\end{aligned}
$$

which describe the vector fields V,W up to the constants $\frac{1}{2c_2}$,$\frac{1}{2c_1}$. The first transforms like

$$
\begin{aligned}
\partial_y + (1 + B(y))\partial_v &= (\frac{\partial x}{\partial y}\partial_x + \frac{\partial u}{\partial x}\partial_u) + (1 + B(y))\frac{\partial u}{\partial v}\partial_u = \\
&= \frac{\partial x}{\partial y}\partial_x + \left[\frac{\partial u}{\partial x} + (B(y) + 1)\frac{\partial u}{\partial v}\right]\partial_u
\end{aligned}
$$

and the second transforms like

$$
\begin{aligned}
\partial_y + (B(y) - 1)\partial_v &= (\frac{\partial x}{\partial y}\partial_x + \frac{\partial u}{\partial x}\partial_u) + (B(y) - 1)\frac{\partial u}{\partial v}\partial_u = \\
&= \frac{\partial x}{\partial y}\partial_x + \left[\frac{\partial u}{\partial x} + (B(y) - 1)\frac{\partial u}{\partial v}\right]\partial_u.
\end{aligned}
$$

We define

$$C := \frac{\partial u}{\partial x} + B\frac{\partial u}{\partial v},$$
$$F := \frac{\partial u}{\partial v} \neq 0$$

and then we can write

$$V = \frac{1}{2c_2}\frac{\partial x}{\partial y}\partial_x + \frac{1}{2c_2}(C(x,u) + F(x,u))\partial_u,$$
$$W = \frac{1}{2c_1}\frac{\partial x}{\partial y}\partial_x + \frac{1}{2c_1}(C(x,u) - F(x,u))\partial_u,$$

where $F \neq 0$ for all $p \in E$. We define $\tilde{c}_1 := \frac{1}{2c_2}$, $\tilde{c}_2 := \frac{1}{2c_1}$ and we have proven what we wanted to prove in the first case.

Second Case ($c_1 \neq 0$ and $c_2 = 0$): The calculation is similar to the first case and we can do it faster. In this case, the vector fields $V = V^x\partial_x + V^u\partial_u$ and $W = W^u\partial_u$ can be written as

$$V = V^x\frac{\partial y}{\partial x}\partial_y + (V^x\frac{\partial v}{\partial x} + V^u\frac{\partial v}{\partial u})\partial_v = c_1A\frac{\partial y}{\partial x}\partial_y + (c_1A\frac{\partial v}{\partial x} + V^u\frac{\partial v}{\partial u})\partial_v$$

and

$$W = W^u\frac{\partial v}{\partial u}\partial_v$$

after local coordinate transformations. Since we can span T_pE for every $p \in E$, we get $W^u \neq 0$ for every $p \in E$. Then we can choose local coordinates such that

$$c_1A\frac{\partial y}{\partial x} = 1,$$
$$W^u\frac{\partial v}{\partial u} = 1. \tag{4.25}$$

Then the second coefficient in V, i.e.

$$c_1A\frac{\partial v}{\partial x} + V^u\frac{\partial v}{\partial u}$$

cannot depend on u or v. To show this, let us go back to (4.21), which can be written as

$$0 = c_1W^u_xA + (V^uW^u_u - W^uV^u_u) = c_1W^u_xA - (W^u)^2\partial_u(\frac{V^u}{W^u}) =$$
$$= (W^u)^2\left(\frac{c_1AW^u_x}{(W^u)^2} - \partial_u(\frac{V^u}{W^u})\right). \tag{4.26}$$

We apply ∂_x on (4.25) and we get

$$0 = \partial_x \cdot 1 = W^u_x \frac{\partial v}{\partial u} + W^u \frac{\partial^2 v}{\partial_x \partial u} = W^u_x \frac{\partial v}{\partial u} + \frac{1}{\frac{\partial v}{\partial u}} \frac{\partial^2 v}{\partial_x \partial u}.$$

This equation can be solved with respect to W^u_x and we plug W^u_x into (4.26) to derive

$$\begin{aligned}
0 &= (W^u)^2 \left(\frac{c_1 A W^u_x}{(W^u)^2} - \partial_u(\frac{V^u}{W^u}) \right) = (W^u)^2 \left(\frac{-c_1 A \frac{\partial^2 v}{\partial x \partial u}}{(W^u)^2 (\frac{\partial v}{\partial u})^2} - \partial_u(\frac{V^u}{W^u}) \right) = \\
&= (W^u)^2 \left(\frac{-c_1 A \frac{\partial^2 v}{\partial x \partial u}}{1} - \partial_u(\frac{V^u}{W^u}) \right) = (W^u)^2 \partial_u \left(-c_1 A \frac{\partial v}{\partial x} - \frac{V^u}{W^u} \right) = \\
&= -(W^u)^2 \partial_u \left(c_1 A \frac{\partial v}{\partial x} + V^u \frac{\partial v}{\partial u} \right) = 0.
\end{aligned}$$

Therefore, we can write

$$\begin{aligned}
V &= c_1 A \frac{\partial y}{\partial x} \partial_y + B(y) \partial_v, \\
W &= W^u \frac{\partial v}{\partial u} \partial_v,
\end{aligned} \tag{4.27}$$

where $B = B(y)$ is a function depending only on y. (Note that this case can also be proven when considering V and $V + W$ and then we can use the proof from the First Case.) Local coordinate transformations lead to

$$\begin{aligned}
V &= c_1 A \frac{\partial y}{\partial x} \partial_y + B(y) \partial_v = c_1 A \frac{\partial y}{\partial x} \left(\frac{\partial x}{\partial y} \partial_x + \frac{\partial u}{\partial y} \partial_u \right) + B(y) \frac{\partial u}{\partial v} \partial_u = \\
&= c_1 A \partial_x + \left(c_1 A \frac{\partial y}{\partial x} \frac{\partial u}{\partial y} + B(y) \frac{\partial u}{\partial v} \right) \partial_u
\end{aligned}$$

and

$$W = W^u \partial_u.$$

Therefore, we have proven Proposition 4.4.2. □

<u>Proof of Lemma 4.4.1:</u> One direction is simple, since if $\tilde{V} = \partial_x$ and $\tilde{W} = \partial_u$ then we get $\mathrm{span}\{\tilde{V}_p, \tilde{W}_p\} = T_p E$ for every $p \in E$ and by the calculation in (4.13) we get $[V, W] = 0$.

For the other direction, we use Proposition 4.4.2 and the expressions in (4.15). Then we do a second local coordinate transformation to derive

$$\begin{aligned}
\tilde{V} &= \partial_u = \frac{\partial v}{\partial u} \partial_v, \\
\tilde{W} &= \partial_x + B(x) \partial_u = \frac{\partial y}{\partial x} \partial_y + \left(\frac{\partial v}{\partial x} + B(x) \frac{\partial v}{\partial u} \right) \partial_v,
\end{aligned}$$

where we consider $B(x)$ as a function depending on x (the new transformation has noting to do with the old transformation in Proposition 4.4.2). We want to find a local coordinate transformations such that

$$\frac{\partial v}{\partial u} \stackrel{!}{=} 1,$$
$$\frac{\partial v}{\partial x} + B(x)\frac{\partial v}{\partial u} \stackrel{!}{=} 0$$

and this can be satisfied by the transformation

$$y = x,$$
$$v(x,u) = u - \int B(x)dx.$$

Always when $\frac{\partial y}{\partial x} \neq 0$ and $\frac{\partial v}{\partial u} \neq 0$, the transformation must be a diffeomorphism and corresponds to an allowed coordinate transformation. Therefore, we have proven Lemma 4.4.1. $\square$

For example, the vector fields

$$V = \partial_u,$$
$$W = \partial_x + e^x \partial_u.$$

are already in the form, described in Proposition 4.4.2. The vector fields

$$V = \partial_x,$$
$$W = \partial_x + e^u \partial_u$$

are not in such a form, but we can bring them in such a form, since they span T_pE for every $p \in E$ and the commutator vanishes. We can do the same with

$$V = e^x \partial_x + \partial_u,$$
$$W = \partial_u,$$

but it does not work for

$$V = \partial_x + \partial_u,$$
$$W = e^x \partial_u,$$

since there the commutator does not vanish.

Also see (Spi99, p.158) and (Tal11), where a theorem, similar to Lemma 4.4.1, can be found.

5. Open Problems and Conclusion

In this chapter, we discuss open problems, possible further research and we provide a critical discussion of Theorem 1.0.2 and 1.0.3, especially concerning applications. We also discuss a new question similar to Takens' problem.

Open problems regarding Takens' question in general:

- Find conditions, like $\mathrm{span}\{V_p : V \in \mathcal{V}\} = T_pE$ for all $p \in E$, for the symmetries and conservation laws, under which Takens' problem can be solved for higher order, or even arbitrary order source forms. The condition $\mathrm{span}\{(\mathrm{pr}^1 V)_p : V \in \mathcal{V}\} = T_pJ^1E$ for all $p \in J^1E$ could probably solve the problem for third order source forms and similar conditions might solve it for arbitrary order.

- Find the general solution of the ECS for arbitrary order and arbitrary n,m. For example, when we have only one symmetry vector field $V \neq 0$, or when $\mathrm{span}\{\pi_* V_p : V \in \mathcal{V}\} = T_{\pi(p)}M$ for all $p \in E$. Similar conditions are thinkable as well.

- Assume Δ satisfies symmetries and corresponding conservation laws, but is not variational. Classify such source forms under the given symmetry and conservation law assumption. We gave an example for such classifications in Subsection 3.8.2 and further discussion can also be found in (Poh95).

- Formulate the proofs with the help of differential invariants, which are used in, for example, (KO03).

- Investigate generalized symmetries in Takens' problem, where the characteristics $Q^\alpha = Q^\alpha(x^i,u^\beta,...,u^\beta_I)$ can depend on higher order jet coordinates. Due to the proof in Section 3.6, which relies on order discussion of polynomial expressions, we can expect that then Theorem 1.0.2 is no longer true in general.

According to the counter examples in Section 4.1, and according to the assumptions in Theorem 1.0.2, it seems that we have found the most general theorem for second order source forms and arbitrary systems of PDEs in some sense. The similarities in the proofs in (AP96, MPV08) and Section 3.6 also show that we cannot expect many generalizations in theorems for second order source forms and arbitrary systems of PDEs. The only substantial generalization would probably be to investigate generalized symmetries.

Open problems, specifically concerning this dissertation:

- Further develop the induction method used in the proofs of Step 3 and Step 6 in Section 3.6. For example, as a starting point, let $H^k_{\alpha\beta}$ be a polynomial of degree $\leq n-1$ in second order coordinates u^γ_{ij}. Show that $u^\beta_{ik}H^k_{\alpha\beta} = 0$ has only the trivial solution with the induction method and with the d-fold operator used in (AP96, p.12) and compare the differences. More generally, investigate if Lemma 2.3 in (AP95, p.629) can be proven with the induction method.

- Prove Theorem 1.0.3 for $m = 1$ and arbitrary n or find a counter example.

- Find the general solution of the ECS for $m = 1$, arbitrary n and 4th order source forms, similar to the calculation in Subsection 3.8.1. In case that finding the general solution is too complicated, find and characterize the singularities with a perturbation theory or with other methods. The methods which we present in Section 3.8 can help to solve this problem. Also see (AP94) for a singular problem.

- Investigate the Helmholtz dependencies for arbitrary order and arbitrary n,m. How many unknowns in Takens' problem can thereby be eliminated and how do the reduced Helmholtz expressions simplify all these proofs? Find a new Helmholtz form with non-dependent Helmholtz coefficients. For $m = 1$, these dependencies can be found in (And89, p.76).

- Check if, and in which sense, the theorems in (AP96, MPV08, AP12) (for second order source forms) could be corollaries of Theorem 1.0.2. Formulate two or three theorems, such that all other theorems, which have been proven so far, are corollaries of them. More generally, it would be nice to have only one or two theorems for Takens' problem, which includes all other theorems.

Open problems concerning applications:

- Find physically meaningful symmetries and especially conservation laws, such that Theorem 1.0.2 and 1.0.3 can be applied. This is of great interest and the short discussion we give in Section 4.3 does not fairly discuss this aspect.

- Find an explanation why, in physical theories, symmetries should be connected to conservation laws. Moreover, why should they be connected in the very special form $Q^\alpha_V = V^\alpha - u^\alpha_i V^i$ and $Q^\alpha_V f_\alpha = D_i C^i_V$?

- Find an explanation why a physical theory should be described by a source form and not by some other formulation, like an equation $f_\alpha = 0$ with certain properties, which cannot be described by a source form. For example, if the number of equations is greater than the number of dependent coordinates (see Maxwell's equations, formulated with the electromagnetic fields $\boldsymbol{E}$,$\boldsymbol{B}$).

The question of applications is important, since most of the motivation in the introduction comes from applications in physics. It is fair to say that it is hard to find many meaningful applications. Of course, the question, formulated by Takens, makes a lot of sense in principle and should definitely be investigated. Even if there are open problems, these theorems could just be the starting point for theorems with many more applications. However, let us now discuss some of these problems in hope that they will be solved in the future. We also give some remarks below how these problems can possibly be solved.

Beside the problems we already mentioned, there are at least three main issues. All three problems come from the requirement that we have to assign a source form to a differential equation $f_\alpha = 0$. Even if it seems that we only make assumptions on symmetries and corresponding conservation laws in Theorem 1.0.2 and 1.0.3, which makes sense from a physical perspective somehow, there are actually more hidden assumptions. Let us discuss them in more detail:

- Are physical differential equations really given by source forms and, more importantly, do we require the symmetry of source forms, or can we just postulate symmetry of f_α in a certain way? The symmetries $\mathcal{L}_{\mathrm{pr}V} f_\alpha = 0$ for $\alpha = 1,2,...,m$ and $\mathcal{L}_{\mathrm{pr}V}\Delta = 0$ are not equivalent in general. More precisely, as we showed in Section 2.3 in formula (2.21), f_α transforms like $f_\alpha \frac{\partial u^\alpha}{\partial v^\beta}\frac{\partial x}{\partial y}$. Why should a physical theory have such a transformation property in general?

- As we already mentioned, why should we restrict to differential equations, where the number of equations and number of dependent coordinates are the same, i.e. f_α and u^α have the same indices $\alpha = 1,2,...,m$? This does not seem to be a necessary requirement in physics.

- Let $f = (f_1,f_2)$ be given functions which describe a differential equation $f = 0$. The source form $\Delta = (f_1 du^1 + f_2 du^2) \wedge dx$ might be variational, but the source form $\Delta = (f_2 du^1 + f_1 du^2) \wedge dx$ might not be variational. For $m > 1$, there is no natural way how to assign a source form to a given differential equation. In this sense, we could even say that it is an unnatural question to ask if a given differential equation is variational or not. Moreover, for Maxwell's equations, it is not even clear if we should formulate the equations with the electromagnetic fields $\boldsymbol{E}$,$\boldsymbol{B}$, or rather formulate them with the vector potential A_μ. We would have to include all equivalent transformations of differential equations to get a more natural question.

  Therefore, a better question would probably be: If a source form satisfies certain symmetries and corresponding conservation laws, then does there exist a variational multiplier (AT92, p.6) such that the differential equation is variational? However, this does not include all equivalence transformations either and the variational multiplier also changes the symmetries of the corresponding source forms in general. Therefore, the definition of symmetry is more complicated then. For reformulations of differential equations also see (Ton84).

A lot of physical differential equations are linear or quasi linear and they are polynomial in their fields and derivatives. We can expect that generalizations of the nowadays known physical theories, for example, the standard model, will probably also be polynomial in their fields and derivatives. Or at least, they are not arbitrarily complicated, since otherwise we would not be able to discover them and verify in experiments. For polynomial equations and source forms, Theorem 2.1 in (AP95) basically solves the question of Takens in full generality. Furthermore, the order of physical differential equations does usually not exceed order two or three. Theorems 1.0.2, 1.0.3 and the theorems in (AP12, AP96, MPV08, AP94) give fairly strong answers to Takens' question with regard to this aspect.

In our opinion, the priority should now be to find more applications of these theorems, which can also mean to give the question of Takens a new formulation. Beside the open problems we discussed above, we would suggest the following continuation: A good starting point could be to prove Theorem 1.0.3 for 4th order, $m = 1$ and arbitrary n, or to find a counter example. Also investigating the Helmholtz dependencies in more generality and for $m > 1$ could be of independent interest. But then, to avoid the above problems concerning the interpretation of source forms in applications, we would suggest the following: Investigate Takens' problem together with the variational multiplier problem (AT92, p.6), or any equivalence transformations of differential equations. This could contribute a new value to Takens' question and connect it to other problems. More precisely, we would like to extend the question of Takens and ask if a differential equation, which satisfies certain symmetries and conservation laws, always allows a variational multiplier.

A. Appendix, Proof of Theorem 5,2.

Takens (Tak77, p.599) formulates the proof differently, but we will use our notation here and do a straight forward discussion of orders in the ECS. The proof can be done to any order and also for PDEs, but we only show it for second order and ODEs. The idea is the same and the PDE case just needs more notation.

Proof: We consider the ECS

$$\begin{aligned} 0 &= Q^\beta H_{\alpha\beta} + (D_x Q^\beta) H^x_{\alpha\beta} + (D_x^2 Q^\beta) H^{xx}_{\alpha\beta} = \\ &= (V^\beta - u^\beta_x V^x) H_{\alpha\beta}(x) + (O^\beta(1) - u^\beta_{xx} V^x) H^x_{\alpha\beta}(x) + (O^\beta(2) - u^\beta_{xxx} V^x) H^{xx}_{\alpha\beta}(x). \end{aligned} \tag{A.1}$$

If $f_\alpha = a_{\alpha\beta}(x) u^\beta + b_{\alpha\beta}(x) u^\beta_x + c_{\alpha\beta}(x) u^\beta_{xx}$ is linear in $(u^\beta, u^\beta_x, u^\beta_{xx})$, then $H_{\alpha\beta}, H^x_{\alpha\beta}, H^{xx}_{\alpha\beta}$ do not depend on the coordinates $(u^\beta, u^\beta_x, u^\beta_{xx})$, they only depend on x. This can easily be seen by the definition of $H_{\alpha,\beta}, H^x_{\alpha\beta}, H^{xx}_{\alpha\beta}$. Therefore, if $V^x \neq 0$ for all x, then there is a $u^\beta_{xxx}, u^\beta_{xx}, u^\beta_x$-cascade of independent terms in (A.1) which shows that all Helmholtz-conditions must be satisfied. That is, we start with the highest order coordinate u^β_{xxx} which shows that $H^{xx}_{\alpha\beta} = 0$. Then the next coordinate is u^β_{xx} which shows $H^x_{\alpha\beta} = 0$, and after that we get $H_{\alpha\beta} = 0$. $\square$

B. Appendix, Prolongation of Vector Fields, Part II

We want to find the local coordinate expression for prolonged vector fields for arbitrary n,m. A general projectable vector field V on E can be written as

$$V = V^i(x)\partial_{x^i} + V^\alpha(x,u)\partial_{u^\alpha} \quad \in \mathfrak{X}(E). \tag{B.1}$$

The general prolongation formula is (Kru97b, p.32)

$$\mathrm{pr}^k V(\mathrm{pr}^k\sigma(q)) = \frac{d}{dt}\{\mathrm{pr}^k[\phi_t \circ \sigma \circ \phi^0_{-t}](\phi^0_t(q))\}|_{t=0}, \tag{B.2}$$

where ϕ_t is the flow of V and ϕ^0_t is the flow of $\pi_* V$. In the following, we simply compute the expression (B.2) in local coordinates. We will use the same ϕ_t and ϕ^0_t for the corresponding transformations of local coordinates (x^i) and (x^i,u^α). Let us write $x = (x^i) = (x^1,x^2,...,x^n)$, $u = (u^\alpha) = (u^1,u^2,...,u^m)$ and $\sigma(x) = (x^i,u^\alpha(x))$ for the local coordinate expression of the section σ. Again, we will use the same σ for local coordinates as for sections on E. The local coordinate x on M transforms according to ϕ^0_t as $\phi^0_t(x) = y_t(x)$ and we write

$$x = (x^1,x^2,...,x^n) \longrightarrow y_t = (y^1_t(x),y^2_t(x),...,y^n_t(x)).$$

We know that

$$y_t(x) = \begin{pmatrix} y^1_t(x) \\ y^2_t(x) \\ \vdots \\ y^n_t(x) \end{pmatrix} = \begin{pmatrix} x^1 + tV^1(x) + O(t^2) \\ x^2 + tV^2(x) + O(t^2) \\ \vdots \\ x^n + tV^n(x) + O(t^2) \end{pmatrix}, \tag{B.3}$$

since $y_{t=0} = x$ and $\frac{d}{dt}y_t|_{t=0} = V$, simply by the fact that $\phi^0_t(y) = y_t$ must satisfy the property of flow. We also know the inverse map

$$y_{-t}(x) = \begin{pmatrix} y^1_{-t}(x) \\ y^2_{-t}(x) \\ \vdots \\ y^n_{-t}(x) \end{pmatrix} = \begin{pmatrix} x^1 - tV^1(x) + O(t^2) \\ x^2 - tV^2(x) + O(t^2) \\ \vdots \\ x^n - tV^n(x) + O(t^2) \end{pmatrix}, \tag{B.4}$$

since, again, then property of flow $\phi^0_{-t} \circ \phi^0_t(x) = x$ must be satisfied. One can easily check that y_{-t} must be of the form (B.4), since

$$x = y_{-t}(y_t(x)) = \begin{pmatrix} y_t^1(x) - tV^1(y_t(x)) + O(t^2) \\ y_t^2(x) - tV^2(y_t(x)) + O(t^2) \\ \vdots \\ y_t^n(x) - tV^n(y_t(x)) + O(t^2) \end{pmatrix} \tag{B.5}$$

must be equal to x up to first order in t, when plugging in (B.3). The transformation ϕ_t acts on the coordinates (x,u) as $\phi_t(x,u) = (y_t(x),v_t(x,u))$, because of fiber preserving transformations. Also see Proposition 2.4.3 and the notation there, where (x^i,u^α) are the untransformed coordinates and $\phi_t(x^i,u^\alpha) = (y_t^i,v_t^\alpha)$ are the transformed coordinates. Similar to (B.3), the general transformation for sections $\phi_t(x,u(x)) = (y_t(x),v_t(x,u(x)))$ can be written as

$$\phi_t \circ \sigma(x) = \phi_t(x,u(x)) = \begin{pmatrix} y_t(x) \\ v_t(x,u(x)) \end{pmatrix} = \begin{pmatrix} y_t^1(x) \\ y_t^2(x) \\ \vdots \\ y_t^n(x) \\ v_t^1(x,u(x)) \\ v_t^2(x,u(x)) \\ \vdots \\ v_t^m(x,u(x)) \end{pmatrix}. \tag{B.6}$$

Note that the transformation ϕ_t only acts on the coordinates (x^i,u^α), but not on the x-coordinates of $u^\alpha(x)$. Roughly speaking, the x-coordinate in $u^\alpha(x)$ is invisible for the transformation ϕ_t, since ϕ_t considers $(x^i,u^\alpha(x))$ as a point (x^i,u^α) and transforms this point according to the rules of flow and corresponding vector field. In the following, we only write y_{-t} for simplicity, instead of $y_{-t}(x)$. We get

$$\phi_t \circ \sigma \circ \phi^0_{-t}(x) = \phi_t(y_{-t}(x),u(y_{-t}(x))) = \begin{pmatrix} y_t(y_{-t}(x)) \\ v_t(y_{-t},u(y_{-t})) \end{pmatrix} = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \\ v_t^1(y_{-t},u(y_{-t})) \\ v_t^2(y_{-t},u(y_{-t})) \\ \vdots \\ v_t^m(y_{-t},u(y_{-t})) \end{pmatrix}$$

and we can prolong this section at the point x. This means that we can take partial derivatives with respect to x, i.e.

$$\mathrm{pr}^k(\phi_t \circ \sigma \circ \phi^0_{-t})(x) = \begin{pmatrix} x \\ v_t(y_{-t},u(y_{-t})) \\ [\partial_{x^i} v^\alpha_t(y_{-t},u(y_{-t}))]_{\alpha,i} \\ [\partial_{x^j}\partial_{x^i} v^\alpha_t(y_{-t},u(y_{-t}))]_{\alpha,i,j} \\ \vdots \end{pmatrix} =$$

$$= \begin{pmatrix} x \\ v_t(y_{-t},u(y_{-t})) \\ [v^\alpha_{t,k}(y_{-t},u(y_{-t}))\partial_{x^i} y^k_{-t}]_{\alpha,i} \\ \{\partial_{x^j}[v^\alpha_{t,k}(y_{-t},u(y_{-t}))\partial_{x^i} y^k_{-t}]\}_{\alpha,i,j} \\ \vdots \end{pmatrix}.$$

Above, $v^\alpha_{t,k}$ means partial derivative with respect to all y^k_{-t}-coordinates, i.e.

$$v^\alpha_{t,k}(y_{-t},u(y_{-t})) = \partial_{y^k_t} v^\alpha_t(y_{-t},u(y_{-t})).$$

Therefore, we can write $v^\alpha_{t,k} = D_k v^\alpha_t$ (similar for $v^\alpha_{t,kl} = D_{kl} v^\alpha_t$ and higher order). The next step is to compute $\partial_{x^i} y^k_{-t}(x)$, $\partial_{x^j}\partial_{x^i} y^k_{-t}(x)$ and so on. Therefore,

$$\partial_{x^i} y^k_{-t}(x) = \partial_{x^i}(x^k - tV^k(x) + O(t^2)) = \delta^k_i - tV^k_{x^i}(x) + O(t^2),$$

where we used (B.4). For $\partial_{x^j}\partial_{x^i} y^k_{-t}(x)$ we get

$$\partial_{x^j}\partial_{x^i} y^k_{-t}(x) = \partial_{x^j}(\delta^k_i - tV^k_{x^i}(x) + O(t^2)) = -tV^k_{x^j x^i}(x) + O(t^2)$$

and higher order derivatives can be computed in a similar way. Then we consider

$$\mathrm{pr}^k[\phi_t \circ \sigma \circ \phi^0_{-t}](\phi^0_t(x)) =$$
$$= \begin{pmatrix} y_t(x) \\ v_t(x,u(x)) \\ \{v^\alpha_{t,k}(x,u(x))[\delta^k_i - tV^k_{x^i}(y_t(x)) + O(t^2)]\}_{\alpha,i} \\ \{v^\alpha_{t,kl}(x,u(x))[\partial_{x^j} y^l_{-t}(...)]\partial_{x^i} y^k_{-t}(...) + v^\alpha_{t,k}(x,u(x))[-tV^k_{x^j x^i}(y_t(x)) + O(t^2)]\}_{\alpha,i,j} \\ \vdots \end{pmatrix}. \tag{B.7}$$

A short auxiliary calculation shows that

$$V^k_{x^i}(y_t(x)) = V^k_{x^i}(x + tV(x) + O(t^2)) = V^k_{x^i}(x) + O(t),$$

where we used (B.3). A similar calculation holds for $V^k_{x^j x^i}(y_t(x))$. Furthermore,

$$[\partial_{x^j} y^l_{-t}(...)]\partial_{x^i} y^k_{-t}(...) = [\delta^l_j - tV^l_{x^j}(y_t(x)) + O(t^2)][\delta^k_i - tV^k_{x^i}(y_t(x)) + O(t^2)] =$$
$$= \delta^l_j \delta^k_i - t[\delta^k_i V^l_{x^j}(x) + \delta^l_j V^k_{x^i}(x)] + O(t^2).$$

Then, (B.7) can be written as

$$
\begin{aligned}
&\mathrm{pr}^k[\phi_t \circ \sigma \circ \phi^0_{-t}](\phi^0_t(x)) = \\
&= \begin{pmatrix} y_t(x) \\ v_t(x,u(x)) \\ \{v^\alpha_{t,k}(x,u(x))[\delta^k_i - tV^k_{x^i}(x) + O(t^2)]\}_{\alpha,i} \\ \{v^\alpha_{t,kl}(...)[\delta^l_j\delta^k_i - t(\delta^k_i V^l_{x^j}(x) + \delta^l_j V^k_{x^i}(x))] + v^\alpha_{t,k}(...)[-tV^k_{x^jx^i}(x)] + O(t^2)\}_{\alpha,i,j} \\ \vdots \end{pmatrix}.
\end{aligned}
$$

Now we want to apply $\frac{d}{dt}$ and evaluate at $t = 0$. When doing so, all terms $O(t^2)$ will vanish and the remaining terms are

$$
\begin{aligned}
&\frac{d}{dt}\{\mathrm{pr}^k[\phi_t \circ \sigma \circ \phi^0_{-t}](\phi^0_t(x))\}|_{t=0} = \\
&= \begin{pmatrix} [V^i(x)]_i \\ [V^\alpha(x,u(x))]_\alpha \\ \{\partial_{x^i}V^\alpha(x,u(x)) - u^\alpha_k(x)V^k_{x^i}(x)\}_{\alpha,i} \\ \{\partial_{x^i}\partial_{x^j}V^\alpha(x,u(x)) - [u^\alpha_{il}(x)V^l_{x^j}(x) + u^\alpha_{kj}(x)V^k_{x^i}(x)] - u^\alpha_k(x)V^k_{x^jx^i}(x)\}_{\alpha,i,j} \\ \vdots \end{pmatrix} = \\
&= \begin{pmatrix} (V^i)_i \\ (V^\alpha)_\alpha \\ \{D_iV^\alpha - u^\alpha_k V^k_{x^i}\}_{\alpha,i} \\ \{D_{ij}V^\alpha - (u^\alpha_{il}V^l_{x^j} + u^\alpha_{kj}V^k_{x^i}) - u^\alpha_k V^k_{x^jx^i}\}_{\alpha,i,j} \\ \vdots \end{pmatrix}.
\end{aligned}
$$

When computing $\frac{d}{dt}v^\alpha_{t,k}(x) = \partial_{x^k}\frac{d}{dt}v^\alpha_t(x)$, we interchange the derivatives. When we want to prove it for higher order, then it is reasonable do to find an inductive procedure for computing these expressions. For our purposes, doing the calculation up to second order is sufficient. Note that the first order coefficient can be written as

$$
D_iV^\alpha - u^\alpha_k V^k_{x^i} = D_i(V^\alpha - u^\alpha_k V^k) + u^\alpha_{ik}V^k
$$

and the second order coefficient as

$$
\begin{aligned}
&D_{ij}V^\alpha - (u^\alpha_{il}V^l_{x^j} + u^\alpha_{jk}V^k_{x^i}) - u^\alpha_k V^k_{x^jx^i} = \\
=&D_{ij}V^\alpha - (u^\alpha_{il}V^l_{x^j} + u^\alpha_{jk}V^k_{x^i}) - D_{ij}(u^\alpha_k V^k) + u^\alpha_{ijk}V^k + u^\alpha_{ik}V^k_{x^j} + u^\alpha_{jk}V^k_{x^i} = \\
=&D_{ij}(V^\alpha - u^\alpha_k V^k) + u^\alpha_{ijk}V^k,
\end{aligned}
$$

where we find the characteristics $Q^\alpha = V^\alpha - u^\alpha_k V^k$.

C. Appendix, the Helmholtz Form

The Helmholtz form H can be found in (VU13, p.13) or (VK15, p.3). Also see Section 2.6.2. Let us consider the expression $\iota_{\mathrm{pr}W}\iota_{\mathrm{pr}V}H$, which can be written as

$$\begin{aligned}
&\iota_{\mathrm{pr}W}\iota_{\mathrm{pr}V}H = \\
=&\iota_{\mathrm{pr}W}\iota_{\mathrm{pr}V}\frac{1}{2}(H_{\alpha\beta}du^{\beta}\wedge du^{\alpha} + H^{x}_{\alpha\beta}du^{\beta}_{x}\wedge du^{\alpha} + H^{xx}_{\alpha\beta}du^{\beta}_{xx}\wedge du^{\alpha})\wedge dx = \\
=&\frac{1}{2}[H_{\alpha\beta}(V^{\beta}W^{\alpha}-W^{\beta}V^{\alpha}) + H^{x}_{\alpha\beta}(V^{\beta}_{;x}W^{\alpha}-W^{\beta}_{;x}V^{\alpha}) + H^{xx}_{\alpha\beta}(V^{\beta}_{;xx}W^{\alpha}-W^{\beta}_{;xx}V^{\alpha})]dx,
\end{aligned} \tag{C.1}$$

where V,$W \in \mathfrak{X}(E)$ are vertical vector fields. We want to bring all D_x-derivatives to V^{β}. Therefore, we apply the following partial integration method

$$\begin{aligned}
H^{x}_{\alpha\beta}V^{\alpha}D_{x}W^{\beta} =&D_x(H^{x}_{\alpha\beta}V^{\alpha}W^{\beta}) - W^{\beta}D_x(H^{x}_{\alpha\beta}V^{\alpha}) = \\
=&D_x(H^{x}_{\alpha\beta}V^{\alpha}W^{\beta}) - W^{\beta}V^{\alpha}D_xH^{x}_{\alpha\beta} - W^{\beta}H^{x}_{\alpha\beta}V^{\alpha}_{;x}
\end{aligned}$$

and

$$\begin{aligned}
H^{xx}_{\alpha\beta}V^{\alpha}D^{2}_{x}W^{\beta} =&D_x(H^{xx}_{\alpha\beta}V^{\alpha}W^{\beta}_{;x}) - W^{\beta}_{;x}D_x(H^{xx}_{\alpha\beta}V^{\alpha}) = \\
=&D_x[H^{xx}_{\alpha\beta}V^{\alpha}W^{\beta}_{;x} - W^{\beta}D_x(H^{xx}_{\alpha\beta}V^{\alpha})] + W^{\beta}D^{2}_{x}(V^{\alpha}H^{xx}_{\alpha\beta}) = \\
=&D_x[H^{xx}_{\alpha\beta}V^{\alpha}W^{\beta}_{;x} - W^{\beta}D_x(H^{xx}_{\alpha\beta}V^{\alpha})] + W^{\beta}D_x(V^{\alpha}_{;x}H^{xx}_{\alpha\beta} + V^{\alpha}D_xH^{xx}_{\alpha\beta}) = \\
=&D_x[H^{xx}_{\alpha\beta}V^{\alpha}W^{\beta}_{;x} - W^{\beta}D_x(H^{xx}_{\alpha\beta}V^{\alpha})] + W^{\beta}(V^{\alpha}_{;xx}H^{xx}_{\alpha\beta} + 2V^{\alpha}_{;x}D_xH^{xx}_{\alpha\beta}+ \\
&+ V^{\alpha}D^{2}_{x}H^{xx}_{\alpha\beta}),
\end{aligned}$$

where we use the short notation $D_xV^{\alpha} = V^{\alpha}_{;x}$, $D^{2}_{x}V^{\alpha} = V^{\alpha}_{;xx}$ and similar for W^{α}. With the help of these two identities we can rewrite (C.1) as

$$\begin{aligned}
&\iota_{\mathrm{pr}W}\iota_{\mathrm{pr}V}H = \\
=&\frac{1}{2}[H_{\alpha\beta}(V^{\beta}W^{\alpha}-W^{\beta}V^{\alpha}) + H^{x}_{\alpha\beta}V^{\beta}_{;x}W^{\alpha} + H^{xx}_{\alpha\beta}V^{\beta}_{;xx}W^{\alpha}]dx+ \\
&+\frac{1}{2}[-D_x(H^{x}_{\alpha\beta}V^{\alpha}W^{\beta}) + W^{\beta}V^{\alpha}D_xH^{x}_{\alpha\beta} + W^{\beta}H^{x}_{\alpha\beta}V^{\alpha}_{;x}]dx+ \\
&+\frac{1}{2}\{-D_x[H^{xx}_{\alpha\beta}V^{\alpha}W^{\beta}_{;x} - W^{\beta}D_x(H^{xx}_{\alpha\beta}V^{\alpha})] - W^{\beta}(V^{\alpha}_{;xx}H^{xx}_{\alpha\beta} + 2V^{\alpha}_{;x}D_xH^{xx}_{\alpha\beta}+ \\
&+ V^{\alpha}D^{2}_{x}H^{xx}_{\alpha\beta})\}dx =
\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{2}[H_{\alpha\beta}-H_{\beta\alpha}+D_xH^x_{\beta\alpha}-D_x^2H^{xx}_{\beta\alpha}]V^\beta W^\alpha dx+\\
&+\frac{1}{2}[H^x_{\alpha\beta}+H^x_{\beta\alpha}-2D_xH^{xx}_{\beta\alpha}]V^\beta_{;x}W^\alpha dx+\\
&+\frac{1}{2}[H^{xx}_{\alpha\beta}-H^{xx}_{\beta\alpha}]V^\beta_{;xx}W^\alpha dx-\\
&-\frac{1}{2}D_x[H^x_{\alpha\beta}V^\alpha W^\beta+H^{xx}_{\alpha\beta}V^\alpha W^\beta_{;x}-W^\beta D_x(H^{xx}_{\alpha\beta}V^\alpha)]dx. \qquad \text{(C.2)}
\end{aligned}
$$

In (C.2) we can show that

$$
\begin{aligned}
&H_{\alpha\beta}-H_{\beta\alpha}+D_xH^x_{\beta\alpha}-D_x^2H^{xx}_{\beta\alpha}=2H_{\alpha\beta}\\
&H^x_{\alpha\beta}+H^x_{\beta\alpha}-2D_xH^{xx}_{\beta\alpha}=2H^x_{\alpha\beta}\\
&H^{xx}_{\alpha\beta}-H^{xx}_{\beta\alpha}=2H^{xx}_{\alpha\beta}
\end{aligned}
$$

and these relations are the Helmholtz dependencies. Note that it seems that considering $\iota_{\mathrm{pr}W}\iota_{\mathrm{pr}V}H$ and rewriting the coefficients in this form is one possibility to get the Helmholtz dependencies. All together, (C.2) can be written as

$$
\begin{aligned}
\iota_{\mathrm{pr}W}\iota_{\mathrm{pr}V}H =&(H_{\alpha\beta}V^\beta W^\alpha+H^x_{\alpha\beta}V^\beta_{;x}W^\alpha+H^{xx}_{\alpha\beta}V^\beta_{;xx}W^\alpha)dx+\\
&+\frac{1}{2}D_x[W^\beta D_x(H^{xx}_{\alpha\beta}V^\alpha)-H^x_{\alpha\beta}V^\alpha W^\beta-H^{xx}_{\alpha\beta}V^\alpha W^\beta_{;x}]dx. \qquad \text{(C.3)}
\end{aligned}
$$

Note that the total-derivative expression in (C.3), i.e.

$$
\begin{aligned}
&W^\beta D_x(H^{xx}_{\alpha\beta}V^\alpha)-H^x_{\alpha\beta}V^\alpha W^\beta-H^{xx}_{\alpha\beta}V^\alpha W^\beta_{;x}=\\
=&W^\beta V^\alpha D_xH^{xx}_{\alpha\beta}+W^\beta V^\alpha_{;x}H^{xx}_{\alpha\beta}-H^x_{\alpha\beta}V^\alpha W^\beta-H^{xx}_{\alpha\beta}V^\alpha W^\beta_{;x}=\\
=&W^\beta V^\alpha(D_xH^{xx}_{\alpha\beta}-H^x_{\alpha\beta})+W^\beta V^\alpha_{;x}H^{xx}_{\alpha\beta}-H^{xx}_{\alpha\beta}V^\alpha W^\beta_{;x}=\\
=&W^\beta V^\alpha[D_x(f_{\alpha,u^\beta_{xx}}-f_{\beta,u^\alpha_{xx}})-(f_{\alpha,u^\beta_x}+f_{\beta,u^\alpha_x}-2D_xf_{\beta,u^\alpha_{xx}})]+W^\beta V^\alpha_{;x}H^{xx}_{\alpha\beta}-\\
&\qquad\qquad-H^{xx}_{\alpha\beta}V^\alpha W^\beta_{;x}=\\
=&W^\beta V^\alpha[D_x(f_{\alpha,u^\beta_{xx}}+f_{\beta,u^\alpha_{xx}})-(f_{\alpha,u^\beta_x}+f_{\beta,u^\alpha_x})]+(W^\beta V^\alpha_{;x}-V^\alpha W^\beta_{;x})H^{xx}_{\alpha\beta} \qquad \text{(C.4)}
\end{aligned}
$$

seems to be different compared to the expression in (2.69), i.e.

$$
\begin{aligned}
&V^\beta W^\alpha f_{\beta,u^\alpha_x}+f_{\beta,u^\alpha_{xx}}V^\beta(D_xW^\alpha)-W^\alpha D_x(f_{\beta,u^\alpha_{xx}}V^\beta)=\\
=&V^\beta W^\alpha(f_{\beta,u^\alpha_x}-D_xf_{\beta,u^\alpha_{xx}})+f_{\beta,u^\alpha_{xx}}V^\beta W^\alpha_{;x}-W^\alpha V^\beta_{;x}f_{\beta,u^\alpha_{xx}}=\\
=&V^\beta W^\alpha(f_{\beta,u^\alpha_x}-D_xf_{\beta,u^\alpha_{xx}})+(V^\beta W^\alpha_{;x}-W^\alpha V^\beta_{;x})f_{\beta,u^\alpha_{xx}}. \qquad \text{(C.5)}
\end{aligned}
$$

For example, when $f_{\alpha,u^\beta_{xx}}=0$ for all $\alpha,\beta=1,2,...,m$ then there is no symmetric summation over f_{β,u^α_x} in (C.5), but there is symmetric summation over $f_{\alpha,u^\beta_x}+f_{\beta,u^\alpha_x}$ in (C.4).

D. Appendix, Proof of Proposition 2.10.

We would like to prove the special form

$$F(x,u^\beta,u^\beta_x,...,u^\beta_{(k)}) = Q^\alpha(x,u^\beta,u^\beta_x,...,u^\beta_{(k)})f_\alpha(x,u^\beta,u^\beta_x,...,u^\beta_{(k)}) \tag{D.1}$$

if f_α is totally-non-degenerate (see Definition 2.83. in (Olv86, p.171)) and if F is a conservation law for f_α according to the classical definition of conservation law (see (Olv86, p.265)). However, instead of proving this more complicated case, we only prove Proposition 2.10. in Peter Olver's book (Olv86, p.84). This proposition provides the main idea how to prove (D.1).

Proposition 2.10. (see (Olv86, p.84)): Let $f : J \to \mathbb{R}^m$ be of maximal rank on the submanifold $S = \{x \in J : f(x) = 0\}$. Then a real-valued function $F : J \to \mathbb{R}$ vanishes on S if and only if there exist smooth functions $Q^1(x),...,Q^m(x)$ such that

$$F(x) = Q^1(x)f_1(x) + ... + Q^m(x)f_m(x)$$

for all $x \in J$.

We changed the notation slightly and we think of $J \hat{=} J^kE$, $x \hat{=} (x^i,u^\alpha,u^\alpha_i,...,u^\alpha_I)$, Q are characteristics and f represents a differential equation.

Proof (1st Case, $\dim J = k > m$): The implicit function theorem says that we ca solve $f(x) = 0$ locally near S. For all

$$y = (y^1,...,y^k) \in J \quad \text{which satisfy} \quad f(y) = 0,$$

there exist coordinates $t = (t^1,...,t^{k-m})$ on S and smooth functions g^l such that

$$y = (t^1,...,t^{k-m},g^{k-m+1}(t),...,g^k(t)).$$

That is, S is a $(k - m)$-dimensional sub-manifold of J. Furthermore, there exist coordinates $z = (z^1,...,z^k)$ on J such that

$$f(x(z)) = \tilde{f}(z) = 0$$

if and only if

$$z = (z^1,...,z^{k-m},0,...,0).$$

See Lemma 1.12. in (Olv86, p.11). By an improved Taylor theorem, i.e. in one dimension: if $f(x^0) = 0$ then $f(x) = f(x^0) + \int_{x^0}^{x} f'(y)dy = f'(\xi(x))(x - x^0) = a(x)(x - x^0)$ by mean value theorem, we can write

$$\tilde{f}(z) = \begin{pmatrix} \tilde{f}_1(z) \\ \vdots \\ \tilde{f}_m(z) \end{pmatrix} = \begin{pmatrix} a_{11}(z)z^{k-m+1} + a_{12}(z)z^{k-m+2} + ... + a_{1m}(z)z^k \\ \vdots \\ a_{m1}(z)z^{k-m+1} + a_{m2}(z)z^{k-m+2} + ... + a_{mm}(z)z^k \end{pmatrix},$$

where $a_{pq}(z)$ are smooth functions. Due to the property of local coordinate transformations, the Jacobi-matrix $J_x f(x)$ and $J_z \tilde{f}(z)$ must have the same rank on S. Therefore,

$$\begin{aligned} & J_z\tilde{f}(z^1,...,z^{k-m},0,...,0) = \\ &= \begin{pmatrix} \frac{\partial \tilde{f}_1}{\partial z^1} & \frac{\partial \tilde{f}_1}{\partial z^2} & ... & \frac{\partial \tilde{f}_1}{\partial z^{k-m}} & \frac{\partial \tilde{f}_1}{\partial z^{k-m+1}} & ... & \frac{\partial \tilde{f}_1}{\partial z^k} \\ \vdots & & & & & & \vdots \\ \frac{\partial \tilde{f}_m}{\partial z^1} & \frac{\partial \tilde{f}_m}{\partial z^2} & ... & \frac{\partial \tilde{f}_m}{\partial z^{k-m}} & \frac{\partial \tilde{f}_m}{\partial z^{k-m+1}} & ... & \frac{\partial \tilde{f}_m}{\partial z^k} \end{pmatrix} |_{(z^1,...,z^{k-m},0,...,0)} = \\ &= \begin{pmatrix} 0 & 0 & ... & 0 & a_{11} & a_{12} & ... & a_{1m} \\ \vdots & & & & & & & \vdots \\ 0 & 0 & ... & 0 & a_{m1} & a_{m2} & ... & a_{mm} \end{pmatrix} |_{(z^1,...,z^{k-m},0,...,0)}. \end{aligned}$$

Since $J_z\tilde{f}$ must have rank m on S, the matrix

$$\Lambda := \begin{pmatrix} a_{11} & a_{12} & ... & a_{1m} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & ... & a_{mm} \end{pmatrix}$$

must be invertible on S. But then, Λ is also invertible in a small neighbourhood of S. In this small neighbourhood of S we can write

$$\begin{aligned} & \Lambda^{-1}(z)\tilde{f}(z) = \Lambda^{-1}(z) \begin{pmatrix} \tilde{f}_1(z) \\ \vdots \\ \tilde{f}_m(z) \end{pmatrix} = \\ &= \Lambda^{-1}(z) \begin{pmatrix} a_{11}(z)z^{k-m+1} + a_{12}(z)z^{k-m+2} + ... + a_{1m}(z)z^k \\ \vdots \\ a_{m1}(z)z^{k-m+1} + a_{m2}(z)z^{k-m+2} + ... + a_{mm}(z)z^k. \end{pmatrix} = \begin{pmatrix} z^{k-m+1} \\ \vdots \\ z^k \end{pmatrix}. \end{aligned}$$

Since F also vanishes on S, we get by the improved Taylor expansion

$$\begin{aligned} F(x(z)) = \tilde{F}(z) &= A_1(z)z^{k-m+1} + A_2(z)z^{k-m+2} + ... + A_m(z)z^k = \\ &=< A(z),\Lambda^{-1}(z)\tilde{f}(z) >= \\ &=< A(z(x)),\Lambda^{-1}(z(x))f(x) > . \end{aligned}$$

Therefore, F must depend linearly on $f(x)$ and we get

$$F(x) = Q^1(x)f_1(x) + Q^2(x)f_2(x) + ...Q^m(x)f_m(x) \tag{D.2}$$

near S. At points where f does not vanish, it is clear that we can also write $F(x)$ as we did in (D.2). $\square$

Proof (2nd Case, dim$J = k \leq m$): In this case, the set S are discrete points. This case does not occur in what we want to discuss, since $f_\alpha = f_\alpha(x^i,u^\beta,u_i^\beta,...)$, and therefore we have always more variables $(x^i,u^\beta,u_i^\beta,...)$ than equations $f_\alpha = 0$.

E. Appendix, Example for the Local Simplification Lemma

Here we want to present an example, where we demonstrate the notation of $O(1)$ and apply the Local Simplification Lemma 3.4.1. Let us consider the fiber bundle $\pi : E = \mathbb{R} \times (\mathbb{R}^2 \setminus 0) \to \mathbb{R}$ with coordinates (x,u,v) and projection $\pi(x,u,v) = x$. Furthermore, we consider translation, scale and rotation symmetry ∂_x, $u\partial_u + v\partial_v$ and $u\partial_v - v\partial_u$. The matrix $(V^i_{\mathscr{A}}, V^\alpha_{\mathscr{A}})_{\mathscr{A};i,\alpha}$ is given by these symmetries as

$$\begin{pmatrix} \partial_x \\ u\partial_u + v\partial_v \\ u\partial_v - v\partial_u \end{pmatrix} \quad \leftrightarrow \quad (V^i_{\mathscr{A}}, V^\alpha_{\mathscr{A}})_{\mathscr{A};i,\alpha} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & u & v \\ 0 & -v & u \end{pmatrix}.$$

Since the determinant

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & u & v \\ 0 & -v & u \end{vmatrix} = u^2 + v^2 \neq 0 \quad \text{on } E,$$

we can span T_pE at every $p \in E$. The inverse matrix of $(V^i_{\mathscr{A}}, V^\alpha_{\mathscr{A}})_{\mathscr{A};i,\alpha}$ is the matrix C

$$C = \frac{1}{u^2+v^2} \begin{pmatrix} (u^2+v^2) & 0 & 0 \\ 0 & u & -v \\ 0 & v & u \end{pmatrix} = \begin{pmatrix} c_1^{\mathscr{A}} \\ c_2^{\mathscr{A}} \\ c_3^{\mathscr{A}} \end{pmatrix}.$$

The ECS is

$$Q^\beta_{\mathscr{A}} H_{\alpha\beta} + (D_x Q^\beta_{\mathscr{A}}) H^x_{\alpha\beta} + (D_x^2 Q^\beta_{\mathscr{A}}) H^{xx}_{\alpha\beta} = 0$$

and for $\mathscr{A} = 1,2,3$, we get the characteristics

$$\begin{aligned} Q_1^\beta &= V_1^\beta - u_x^\beta V_1^x = -u_x^\beta, \\ Q_2^\beta &= V_2^\beta - u_x^\beta V_2^x = u^\beta, \\ Q_3^\beta &= V_3^\beta - u_x^\beta V_3^x = u\delta^{\beta v} - v\delta^{\beta u}, \end{aligned}$$

where $u = u^1$ corresponds to the index 1 and $v = u^2$ corresponds to the index 2. Therefore, it is convenient also to use u,v as indices. However, in the following, u^2 and v^2 are the square of u and v. Then we get

$$\begin{aligned}
c_1^{\mathscr{A}} Q_{\mathscr{A}}^{\beta} &= 1 \cdot Q_1^{\beta} + 0 \cdot Q_2^{\beta} + 0 \cdot Q_3^{\beta} = -u_x^{\beta}, \\
c_2^{\mathscr{A}} Q_{\mathscr{A}}^{\beta} &= 0 \cdot Q_1^{\beta} + \frac{u}{u^2+v^2} Q_2^{\beta} + \frac{-v}{u^2+v^2} Q_3^{\beta} = \frac{uu^{\beta} - v(u\delta^{\beta v} - v\delta^{\beta u})}{u^2+v^2} = \delta^{\beta u}, \\
c_3^{\mathscr{A}} Q_{\mathscr{A}}^{\beta} &= 0 \cdot Q_1^{\beta} + \frac{v}{u^2+v^2} Q_2^{\beta} + \frac{u}{u^2+v^2} Q_3^{\beta} = \frac{vu^{\beta} + u(u\delta^{\beta v} - v\delta^{\beta u})}{u^2+v^2} = \delta^{\beta v},
\end{aligned}$$

which actually was clear by definition of $c^{\mathscr{A}}$. Furthermore, we get

$$\begin{aligned}
c_1^{\mathscr{A}} D_x Q_{\mathscr{A}}^{\beta} &= 1 \cdot D_x Q_1^{\beta} + 0 \cdot D_x Q_2^{\beta} + 0 \cdot D_x Q_3^{\beta} = -u_{xx}^{\beta}, \\
c_2^{\mathscr{A}} D_x Q_{\mathscr{A}}^{\beta} &= 0 \cdot D_x Q_1^{\beta} + \frac{u}{u^2+v^2} D_x Q_2^{\beta} + \frac{-v}{u^2+v^2} D_x Q_3^{\beta} = \frac{uu_x^{\beta} - v(u_x\delta^{\beta v} - v_x\delta^{\beta u})}{u^2+v^2} = \\
&= \frac{uu_x + vv_x}{u^2+v^2}\delta^{\beta u} + \frac{uv_x - vu_x}{u^2+v^2}\delta^{\beta v} = O(1), \\
c_3^{\mathscr{A}} D_x Q_{\mathscr{A}}^{\beta} &= 0 \cdot D_x Q_1^{\beta} + \frac{v}{u^2+v^2} D_x Q_2^{\beta} + \frac{u}{u^2+v^2} D_x Q_3^{\beta} = \frac{vu_x^{\beta} + u(u_x\delta^{\beta v} - v_x\delta^{\beta u})}{u^2+v^2} = \\
&= \frac{vu_x - uv_x}{u^2+v^2}\delta^{\beta u} + \frac{vv_x + uu_x}{u^2+v^2}\delta^{\beta v} = O(1).
\end{aligned}$$

We observe that $O(1)$ does not vanish and that $O(1)$ also cannot be written as $\delta^{\beta u}$ or $\delta^{\beta v}$. In general, lower order expressions can get quite complicated.

Bibliography

[AD80] I.M. Anderson and T. Duchamp. On the Existence of Global Variational Principles. Amer. J. Math., 102(5):781–868, 1980.

[AF01] I. Agricola and T. Friedrich. Globale Analysis, Differentialformen in Analysis, Geometrie und Physik. Friedr. Vieweg Sohn Verl. mbH, Braunschweig, Wiesbaden, 2001.

[And89] I.M. Anderson. The Variational Bicomplex. Utah State University Technical Report, 1989.

[AP94] I.M. Anderson and J. Pohjanpelto. Variational principles for differential equations with symmetries and conservation laws I. Second order scalar equations. Math. Ann., 299:191–222, 1994.

[AP95] I.M. Anderson and J. Pohjanpelto. Variational principles for differential equations with symmetries and conservation laws II Polynomial differential equations. Math. Ann., 301:627–653, 1995.

[AP96] I.M. Anderson and J. Pohjanpelto. Symmetries, conservation laws and variational principles for vector field theories. Math. Proc. Camb. Philos. Soc., 120:369–384, 1996.

[AP12] I.M. Anderson and J. Pohjanpelt. Variational Principles for Natural Divergence-free Tensors in Metric Field Theories. J. Geom. Phys., 62(12):2376–2388, 2012.

[AT92] I.M. Anderson and G. Thompson. The Inverse Problem of the Calculus of Variations for Ordinary Differential Equations. Amer. Math. Soc., 98(473):1–110, 1992.

[BCO81] J.M. Ball, J.C. Currie, and P.J. Olver. Null Lagrangians, Weak Continuity, and Variational Problems of Arbitrary Order. Jorn. Funct. Analy., 41:135–174, 1981.

[DEF+99] P. Deligne, P. Etingof, D.S. Freed, L.C. Jeffry, D. Kazhdan, J.W. Morgan, D.R. Morrison, and E. Witten. Quantum Fields and Strings: A Course for Mathematicians, Volume 2. Amer. Math. Soc. Inst. Adv. Stud., 1999.

[Dou41] J. Douglas. Solution of the Inverse Problem of the Calculus of Variations. Trans. Amer. Math. Soc., 50(1):71–128, 1941.

[GD75] I.M. Gel'fand and L.A. Dikki. Asymptotic behaviour of Sturm-Liouville equations and the algebra of the Korteweg-de Vries equations. Russ. Math. Surveys, 30(5):77–113, 1975.

[GM10] A. Galbis and M. Maestre. Vector Analysis Versus Vector Calculus. Springer New York Dordrecht London, 2010.

[Hel87] H. Helmholtz. Über die physikalische Bedeutung des Prinzips der kleinsten Wirkung. Jorn. rein. u. angew. Math., 100:137–166, 1887.

[KM10] O. Krupková and R. Malíková. Helmholtz conditions and their generalizations. Balkan J. Geom. and Applications, 15(1):80–89, 2010.

[KO03] I.A. Kogan and P. Olver. Invariant Euler-Lagrange Equations and the Invariant Variational Bicomplex. Acta Appl. Math., 76:137–193, 2003.

[KPS08] A. Kasman, A. Pedings, K.and Reiszl, and T. Shiota. Universality of Rank 6 Plücker Relations and Grassmann Cone Preserving Maps. Proc. Amer. Math. Soc., 136(1):77–87, 2008.

[Kru97a] D. Krupka. Variational sequences in mechanics. Springer-Verlag, Calc. Var., 5:557–583, 1997.

[Kru97b] O. Krupková. The Geometry of Ordinary Variational Equations. Lecture notes in mathematics, 1678, Springer-Verlag Berlin Heidelberg New York, 1997.

[Kru15] D. Krupka. Intorduction to Global Variational Geomety. Atlantis Press, 2015.

[KS08] D. Krupka and D. Saunders. Jet manifolds and natural bundles. Elsevier B.V., Handbook of Global Analysis, pages 1035–1068, 2008.

[Ku04] D. Krupka and J. Šeděnková. Variational sequences and Lepage forms. Preprint Series in Global Analysis and Applications, pages 1–11, 2004.

[Lee13] J.M. Lee. Introduction to Smooth Manifolds, Second Edition. Graduate Texts in Mathematics 218, Springer New York Heidelberg, 2013.

[LL74] D.L. Lee and A.P. Lightman. Conservation laws and variational principles in metric theories of gravity. Phys. Rev. D, 10(6):1685–1700, 1974.

[Mal09] R. Malíková. On a generalization of Helmholtz conditions. Acta Math. Univ. Ostrav., 17(1):11–21, 2009.

[Mat98] P.C. Matthews. Vector Calculus. Springer undergr. math. ser., 1998.

[MPV08] G. Manno, J. Pohjanpelto, and R. Vitolo. Gauge invariance, charge conservation, and variational principles. J. Geom. Phys., 58:996–1006, 2008.

[Noe18] E. Noether. Invariante Variationsprobleme. Nachr. Ges. Wiss. Gött., pages 235–257, 1918.

[Olv83] P.J. Olver. Hyperjacobians, determinantal ideals and weak solutions to variational problems. Proc. Royal Soc. Edinb., 95A:317–340, 1983.

[Olv86] P.J. Olver. Applications of Lie Groups to Differential Equations. Springer-Verlag New York Berlin Heidelberg, Graduate Texts in Mathematics 107, 1986.

[OS78] P.J. Olver and C. Shakiban. A Resolution of the Euler Operator. I. Proc. Amer. Math. Soc., 69(2):223–229, 1978.

[Per02] G. Perelman. The entropy formula for the Ricci flow and its geometric applications. arXiv:math/0211159 [math.DG], pages 269–279, 2002.

[Poh95] J. Pohjanpelto. Takens' problem for systems of first order differential equations. Ark. Math., 33:343–356, 1995.

[PS95] M.E. Peskin and D.V. Schroeder. An Intorduction to Quantum Field Theory. Pers. Books Pub. L.L.C., 1995.

[Spi99] M. Spivak. A Comprehensive Introduction to Differential Geometry. Vol. 1, 3rd Edition, Publish or Perish, 1999.

[StbM12] N.Ja. Sonin and R.Ya. transl. by Matsyuk. About determining maximal and minimal properties. Lep. Inst. Arch., (1):1–42, 2012.

[Tak77] F. Takens. Symmetries, conservation laws and variational principles. Springer-Verlag, New York, Lecture Notes in Mathematics, 597:581–603, 1977.

[Tak79] F. Takens. A Global Version of the Inverse Problem of the Calculus of Variations. Jour. Diff. Geom., 14:543–562, 1979.

[Tal11] I.A. Tall. Flow-Box Theorem and Beyond. Afr. Diasp. Jour. Math., 11(1):75–102, 2011.

[Ton84] E. Tonti. Variational Formulation for Every Nonlinear Problem. Int. J. Engng Sci., 22(11/12):1343–1371, 1984.

[VK15] N. Voicu and D. Krupka. Canonical variational completion of differential equations. Jour. Math. Phys., 56(043507):1–11, 2015.

[VU13] J. Volná and Z. Urban. The interior Euler-Lagrange operator in field theory. Preprint Ser. Glob. Var. Geom., Lepage Research Institut, pages 1–14, 2013.

[Wan94] D. Wands. Extended gravity theories and the Einstein-Hilbert action. Class. Quant. Grav., 11(1):269–279, 1994.

Selbstständigkeitserklärung

Hiermit versichere ich, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Markus Dafinger

Oldenburg, den 12. Februar 2018

Lebenslauf

Persönliche Angaben

Name	Markus Dafinger
Geburtsdatum	05. Januar 1983
Geburtsort	Erding

Bildungsgang

2004–2006	Berufsoberschule in Landshut Abschluss: fachgebunde Hochschulreife
2006–2012	Diplomstudium allgemeine Physik an der Technischen Universität München Abschluss: Diplom
2012–2018	Promotionsstudium Mathematik an der Carl von Ossietzky Universität Oldenburg Abschluss: Dr. rer. nat.