1. Formal Geometry of Differential Equations

He who can digest a second or third fluxion, a second or third difference, need not, we think, be squeamish about any point of divinity.

George Berkeley

In this chapter we lay the geometric foundations of the so-called *formal theory of differential equations*. The name “formal theory” stems from the fact that, at least indirectly, much of it is based on the analysis of formal power series solutions, i. e. one ignores the question whether the series actually converge. Another interpretation of the name is that one tries to extract as much information as possible on the solution space by purely formal operations like algebraic manipulations of the equations or their differentiations without actually solving the given equation.

The basic tool for a geometric approach to differential equations is the jet bundle formalism. The first two sections of this chapter give an introduction to it; first in a more pedestrian way considering jets as a geometric approach to power series, then in a more abstract, intrinsic way stressing certain structural properties which will later be of importance.

The third section starts with the geometric definition of differential equations as fibred submanifolds of jet bundles and introduces the two basic operations of prolongation to higher order and projection to lower order. This leads almost immediately to the notion of formal integrability, the topic of the next section. From a more computational point of view this is related to the problem of constructing order by order formal power series solutions for a given differential equation. For a formally integrable equation such a construction is always possible, as no obstructions in the form of integrability conditions exist.

The fifth section introduces the geometric and the principal symbol of a differential equation and related notions like characteristics. At this point it does not yet become apparent why the symbol is so important, but in later chapters we will see that not only the concept of involution is based on it but that it is also of fundamental importance in computations.

In the final section we present some classical differential equations of mathematical physics. In the subsequent chapters, they will be repeatedly used as examples in order to demonstrate that the theory can be used in realistic applications and not only in toy problems.

1.1 A Pedestrian Approach to Jet Bundles

Let $\mathcal{X}$ and $\mathcal{U}$ be two (finite-dimensional) manifolds, for example two open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let furthermore $\phi : \mathcal{X} \to \mathcal{U}$ be a smooth map between these two manifolds. If we take some charts on $\mathcal{X}$ and $\mathcal{U}$ with local coordinates $\mathbf{x} = (x^1, \dots, x^n)$ and $\mathbf{u} = (u^1, \dots, u^m)$, respectively, we can locally write ϕ in the form $u^\alpha = \phi^\alpha(x^1, \dots, x^n)$ for $1 \leq \alpha \leq m$. In the context of differential equations, the coordinates $\mathbf{x}$ are often called *independent variables*, whereas the coordinates $\mathbf{u}$ represent the *dependent variables*.

According to Taylor's Theorem we can expand the components ϕ^α of the map ϕ in the neighbourhood of an arbitrary but fixed point $\mathbf{x}_0 \in \mathcal{X}$

$$\phi^\alpha(\mathbf{x}) = \sum_{0 \leq |\mu| \leq q} \frac{p^\alpha_\mu}{\mu!}(\mathbf{x} - \mathbf{x}_0)^\mu + R^\alpha_q(\mathbf{x}, \mathbf{x}_0) \; . \tag{1.1}$$

Here $q \geq 0$ is an arbitrary integer and we use the standard multi index notation of multivariate analysis (see Appendix A.1). It is well-known that the coefficients $p^\alpha_\mu \in \mathbb{R}$ of the expansion (1.1) are the derivatives of the functions ϕ^α evaluated at $\mathbf{x}_0$

$$p^\alpha_\mu = \frac{\partial^{|\mu|}\phi^\alpha}{\partial x^\mu}(\mathbf{x}_0) \tag{1.2}$$

and that in the Lagrange representation the remainder term takes the form

$$R^\alpha_q(\mathbf{x}, \mathbf{x}_0) = \sum_{|\mu| = q+1} \frac{\partial^{q+1}\phi^\alpha}{\partial x^\mu}\big(\mathbf{x}_0 + \theta(\mathbf{x} - \mathbf{x}_0)\big)\frac{(\mathbf{x} - \mathbf{x}_0)^\mu}{\mu!} \tag{1.3}$$

for some $\theta \in]0, 1[$.

We introduce an equivalence relation on the space of all maps $\mathcal{X} \to \mathcal{U}$. The maps ϕ and ψ are equivalent to order q at $\mathbf{x}_0 \in \mathcal{X}$, written $\phi \sim_q \psi$, if their expansions up to order q at $\mathbf{x}_0$ are identical. Geometrically, this means that the graphs of the two maps have at x_0 a contact of order q. Thus if ϕ is given by (1.1) and ψ by a similar expansion with coefficients $\bar{p}^\alpha_\mu$, then $\phi \sim_q \psi$ implies that $p^\alpha_\mu = \bar{p}^\alpha_\mu$ for $1 \leq \alpha \leq m$ and $0 \leq |\mu| \leq q$. A q-*jet* at x_0 is defined as an equivalence class

$$[\phi]^{(q)}_{x_0} = \left\{\psi \in C^\infty(\mathcal{X}, \mathcal{U}) \mid \psi \sim_q \phi \text{ at } x_0\right\} . \tag{1.4}$$

Loosely speaking, the jet bundle of order q is the space of all power series truncated at order q, i.e. a point in it is such an equivalence class $[\phi]^{(q)}_{x_0}$. Note that we always consider only a finite part of the series; thus it does not matter whether the series actually converge and we do not have to restrict to analytic functions.

Definition 1.1.1. *The* jet bundle of order q *over the manifolds* $\mathcal{X}$ *and* $\mathcal{U}$ *is the space* $J_q(\mathcal{X}, \mathcal{U})$ *of all* q*-jets* $[\phi]^{(q)}_x$ *with* $x \in \mathcal{X}$.

This definition is very similar to the introduction of the tangent bundle in Appendix C.1. A point in it was also defined as an equivalence class of functions (more precisely, of curves). As for the tangent bundle, it is not difficult to show that $J_q(\mathcal{X}, \mathcal{U})$ is a manifold, too. In fact, it is obvious what one may use as local coordinates: given the coordinates x on $\mathcal{X}$ and $\mathbf{u}$ on $\mathcal{U}$, we take as coordinates for the q-jet $[\phi]^{(q)}_{x_0}$ the tuple (x^i_0, p^α_μ) with $0 \leq |\mu| \leq q$. The jet bundle of order 0 is the product $\mathcal{X} \times \mathcal{U}$, as a 0-jet tells us only the result $\phi(x_0)$ of evaluating any map ϕ contained in it at the point x_0.

In concrete calculations one often ignores that points in $J_q(\mathcal{X}, \mathcal{U})$ are actually equivalence classes of functions. One simply thinks of $J_q(\mathcal{X}, \mathcal{U})$ as a space whose coordinates are besides the independent variables x and dependent variables $\mathbf{u}$ all derivatives p^α_μ with $|\mu| \leq q$.

$J_q(\mathcal{X}, \mathcal{U})$ is not just any manifold but possesses many internal structures. In particular, jet bundles form a natural hierarchy. If $r > q$, we have an obvious projection map $\pi^r_q : J_r(\mathcal{X}, \mathcal{U}) \to J_q(\mathcal{X}, \mathcal{U})$ defined by "forgetting" the higher-order terms of the power series. We can project even further by forgetting everything but the expansion point; this yields maps $\pi^q : J_q(\mathcal{X}, \mathcal{U}) \to \mathcal{X}$. All these projection maps are surjective submersions, so that $J_q(\mathcal{X}, \mathcal{U})$ may be endowed in many different ways with the structure of a fibred manifold (cf. Appendix C.2).

The dimension of a jet bundle is easily calculated by some elementary combinatorics. As $\dim \mathcal{X} = n$, $\dim \mathcal{U} = m$, we get $\dim J_0(\mathcal{X}, \mathcal{U}) = n + m$. If we go from $J_{q-1}(\mathcal{X}, \mathcal{U})$ to $J_q(\mathcal{X}, \mathcal{U})$, we add one new coordinate for each derivative of order q of m functions of n variables. Thus, by (A.4a),

$$\dim J_q(\mathcal{X}, \mathcal{U}) - \dim J_{q-1}(\mathcal{X}, \mathcal{U}) = m\binom{n+q-1}{q} , \tag{1.5}$$

which, by (A.4c), implies that

$$\dim J_q(\mathcal{X},\mathcal{U}) = n + m\binom{n+q}{q}\,. \tag{1.6}$$

In practice, one considers usually only the fibre dimension, i. e. the number of variables p^α_μ with $0 \leq |\mu| \leq q$, and omits the summand n stemming from the base space $\mathcal{X}$.

Example 1.1.2. Assume that $\mathcal{X}$ is two-dimensional with coordinates (x^1, x^2) and $\mathcal{U}$ one-dimensional with coordinate u^1. The jet bundle $J_0(\mathcal{X},\mathcal{U}) = \mathcal{X}\times\mathcal{U}$ is then three-dimensional with coordinates $(x^1, x^2; u^1)$. The first-order jet bundle $J_1(\mathcal{X},\mathcal{U})$ is five-dimensional with coordinates $(x^1, x^2; u^1; p^1_{[1,0]}, p^1_{[0,1]})$ and for the eight-dimensional second-order jet bundle $J_2(\mathcal{X},\mathcal{U})$ we use the coordinates $(x^1, x^2; u^1; p^1_{[1,0]}, p^1_{[0,1]}, p^1_{[2,0]}, p^1_{[1,1]}, p^1_{[0,2]})$ and so on. The fibre dimension is always two less.

Obviously, the multi index notation is rather cumbersome in such concrete examples. For this reason, we restrict its use to theoretical parts and use more natural variable names in examples. Thus we may take variables (x, y) for $\mathcal{X}$ and u for $\mathcal{U}$. The derivatives in $J_2(\mathcal{X},\mathcal{U})$ are then denoted by $(x, y; u; u_x, u_y, u_{xx}, u_{xy}, u_{yy})$ which brings us back to the standard notations in differential equations. $\lhd$

Any smooth map $\phi : \mathcal{X} \to \mathcal{U}$ can be *prolonged* to a smooth map

$$j_q\phi\,:\,\begin{cases}\mathcal{X} \longrightarrow J_q(\mathcal{X},\mathcal{U})\\ \mathbf{x} \longmapsto [\phi]^{(q)}_x = \left(x^i, \phi^\alpha(\mathbf{x}), \frac{\partial^{|\mu|}\phi^\alpha}{\partial x^\mu}(\mathbf{x})\right)\end{cases}\,. \tag{1.7}$$

Of course, not all maps $\mathcal{X} \to J_q(\mathcal{X},\mathcal{U})$ are of this special form, but these prolonged maps are particularly important. In local coordinates, a general map $\mathcal{X} \to J_q(\mathcal{X},\mathcal{U})$ has the form $\mathbf{x} \mapsto \left(x^i, \phi^\alpha(\mathbf{x}), \phi^\alpha_\mu(\mathbf{x})\right)$ without any relationship between the functions ϕ^α_μ and ϕ^α.

Those maps that come from a prolongation can be characterised with the help of the *contact structure* of the jet bundle $J_q(\mathcal{X},\mathcal{U})$. The name refers to the fact that the equivalence relation underlying our definition of the jet bundle identifies maps whose graphs have a contact of order q. The structure is often written in terms of differential one-forms on $J_q(\mathcal{X},\mathcal{U})$. Given a one-form $\omega \in \Omega^1(J_q(\mathcal{X},\mathcal{U}))$, we can compute for an arbitrary n-dimensional fibred submanifold $\mathcal{N} \subset J_q(\mathcal{X},\mathcal{U})$, i. e. satisfying $\pi^q\mathcal{N} = \mathcal{X}$, the pull-back $\iota^*\omega$ where $\iota : \mathcal{N} \to J_q(\mathcal{X},\mathcal{U})$ is the natural inclusion map. A one-form on $J_q(\mathcal{X},\mathcal{U})$ is called a *contact form*, if this pull-back vanishes, if and only if $\mathcal{N}$ is the image of a prolonged map $j_q\phi$. The contact forms span the *contact codistribution* $\mathcal{C}^0_q \subset T^*J_q\mathcal{E}$.

Locally, $\mathcal{C}_q^0$ is generated by the following one-forms[1]

$$\omega_\mu^\alpha = dp_\mu^\alpha - p_{\mu+1_i}^\alpha dx^i \tag{1.8}$$

with $1 \leq \alpha \leq m$ and $0 \leq |\mu| < q$. Thus $\dim \mathcal{C}_q^0 = \dim J_{q-1}\mathcal{E}$ (the fibre dimension!). It is easy to show that the forms ω_μ^α are indeed contact forms and generate $\mathcal{C}_q^0$ [86]. Because of the transversality condition $\pi^q\mathcal{N} = \mathcal{X}$, we may use the coordinates $\mathbf{x}$ as parameters for $\mathcal{N}$ and describe it in the form $(\mathbf{x}, \mathbf{p}) = \big(\mathbf{x}, \psi(\mathbf{x})\big)$ with some function ψ. Hence pulling-back the contact forms by the natural inclusion map $\iota : \mathcal{N} \hookrightarrow J_q(\mathcal{X}, \mathcal{U})$ yields

$$(\iota^*\omega_\mu^\alpha)(\mathbf{x}) = \left(\frac{\partial \psi_\mu^\alpha}{\partial x^i}(\mathbf{x}) - \psi_{\mu+1_i}^\alpha(\mathbf{x})\right) dx^i \ . \tag{1.9}$$

The pulled-back form vanishes, if and only if the components of ψ satisfy the relations $\psi_\mu^\alpha = \partial^{|\mu|}\psi_0^\alpha / \partial x^\mu$. But this is only the case, if $\psi = j_q\psi_0$.

An alternative description of the contact structure uses vector fields instead of differential forms: the *contact distribution* $\mathcal{C}_q \subset TJ_q\mathcal{E}$ consists of all vectors that are annihilated by the contact one-forms. It follows from (1.8) that $\dim \mathcal{C}_q = \dim J_q\mathcal{E} - \dim J_{q-1}\mathcal{E} + \dim \mathcal{X}$. $\mathcal{C}_q$ is locally generated by the following two sets of vector fields:

$$C_i = \partial_{x^i} + p_{\mu+1_i}^\alpha \partial_{p_\mu^\alpha} \ , \qquad 1 \leq i \leq n \tag{1.10a}$$

(summation over all $1 \leq \alpha \leq m$ and all multi indices μ with $0 \leq |\mu| < q$ in the second term) and

$$C_\alpha^\mu = \partial_{p_\mu^\alpha} \ , \qquad 1 \leq \alpha \leq m \ , \quad |\mu| = q \ . \tag{1.10b}$$

Now we may characterise those fibred submanifolds $\mathcal{N} \subset J_q\mathcal{E}$ that are the image of a prolonged map $j_q\phi$ by the fact that their tangent spaces are subspaces of $\mathcal{C}_q$.

For another kind of prolongation we need the *formal* or *total derivatives*. They take as argument a function $\Phi : J_q(\mathcal{X}, \mathcal{U}) \to \mathbb{R}$ and return a function $J_{q+1}(\mathcal{X}, \mathcal{U}) \to \mathbb{R}$. In local coordinates, the formal derivative D_i with respect to the independent variable x^i is defined as

$$D_i\Phi(\mathbf{x}, \mathbf{p}) = \frac{\partial \Phi}{\partial x^i}(\mathbf{x}, \mathbf{p}) + \frac{\partial \Phi}{\partial p_\mu^\alpha}(\mathbf{x}, \mathbf{p})\, p_{\mu+1_i}^\alpha \ . \tag{1.11}$$

[1] Here we use on the right hand side the *summation convention* that a summation is understood over repeated indices, i. e. in this case over i. The range of the summation should be clear from the context; here it is obviously $1 \leq i \leq n$. In the sequel we will mostly omit the sums in formulae.

Note that the function $D_i\Phi$ is always *quasi-linear*, i. e. linear in the derivatives of maximal order: as Φ depends only on the derivatives p^α_μ with $|\mu| \leq q$, derivatives of order $q+1$ appear in (1.11) only in the terms $p^\alpha_{\mu+1_i}$ with $|\mu| = q$.

Eq. (1.11) represents the well-known *chain rule* of calculus within the jet bundle formalism. Indeed, suppose that we consider the function Φ restricted to the graph of a map $j_q\psi : \mathcal{X} \to J_q(\mathcal{X}, \mathcal{U})$. Then we can substitute the variables p^α_μ by $\partial^{|\mu|}\psi^\alpha/\partial x^\mu$ and obtain a function $\tilde{\Phi} : \mathcal{X} \to \mathcal{U}$. The partial derivative $\partial\tilde{\Phi}/\partial x_i$ must be computed by the chain rule. Exactly the same result is reached, if we perform the substitutions in the function $D_i\Phi$.

Example 1.1.3. Let (x, t) be coordinates on $\mathcal{X}$ and u on $\mathcal{U}$. We consider the function $\Phi(x, t, u, u_x, u_t) = (u_t)^2 - xu_x$ defined on $J_1(\mathcal{X}, \mathcal{U})$. Its formal derivatives with respect to t and x, respectively, are given by the following two functions defined on $J_2(\mathcal{X}, \mathcal{U})$

$$D_x\Phi(x, t, u, u_x, u_t, u_{xx}, u_{xt}, u_{tt}) = 2u_tu_{xt} - xu_{xx} - u_x\ , \tag{1.12a}$$

$$D_t\Phi(x, t, u, u_x, u_t, u_{xx}, u_{xt}, u_{tt}) = 2u_tu_{tt} - xu_{xt}\ . \tag{1.12b}$$

These are indeed quasi-linear functions, i. e. linear in all second-order derivatives, although Φ is quadratic in u_t.

If we substitute u by some function $\psi(x, t)$ in Φ, we obtain a function $\tilde{\Phi}(x, t) = \Phi\big(x, t, \psi(x, t), \psi_x(x, t), \psi_t(x, t)\big)$. Differentiating this function with respect to x yields according to the chain rule

$$\frac{\partial\tilde{\Phi}}{\partial x}(x, t) = 2\psi_t(x, t)\psi_{xt}(x, t) - x\psi_{xx}(x, t) - \psi_x(x, t) \tag{1.13}$$

and correspondingly for $\partial\tilde{\Phi}/\partial t$. The back substitutions $\psi \to u$, $\psi_x \to u_x$, $\psi_{xx} \to u_{xx}$ etc. lead exactly to the above formal derivatives. $\lhd$

We have mentioned above that jet bundles form a natural hierarchy. The projections $\pi^q_{q-1} : J_q(\mathcal{X}, \mathcal{U}) \to J_{q-1}(\mathcal{X}, \mathcal{U})$ have an additional structural property that will become rather important for us in the sequel.

Proposition 1.1.4. *The jet bundle $J_q(\mathcal{X}, \mathcal{U})$ of order q is affine over the jet bundle $J_{q-1}(\mathcal{X}, \mathcal{U})$ of order $q - 1$.*

Before we prove this proposition, we discuss what it means. Recall that a function is affine, if it is of the form $x \mapsto ax + b$ for constants a, b. Proposition 1.1.4 claims that the highest-order derivatives p^α_μ with $|\mu| = q$ transform under a change of coordinates in $\mathcal{X}$ and $\mathcal{U}$ as affine functions with coefficients a, b depending on the remaining coordinates p^α_μ with $|\mu| < q$. A precise definition of an affine bundle is given in Appendix C.2.

Proof. We study the effect of changes of coordinate $\bar{\mathbf{x}} = \bar{\mathbf{x}}(\mathbf{x})$ and $\bar{\mathbf{u}} = \bar{\mathbf{u}}(\mathbf{u})$ in the manifolds $\mathcal{X}$ and $\mathcal{U}$, respectively. They induce a change of coordinates $\bar{\mathbf{p}} = \bar{\mathbf{p}}(\mathbf{x}, \mathbf{p})$ in the jet bundle $J_q(\mathcal{X}, \mathcal{U})$. It can be determined either by using the chain rule or, equivalently, by requiring that the contact forms (1.8) remain invariant. In either case the result is that in repeated index notation

$$\bar{p}^{\alpha}_{j_1\cdots j_q} = \left(\frac{\partial \bar{u}^{\alpha}}{\partial u^{\beta}}\frac{\partial x^{i_1}}{\partial \bar{x}^{j_1}}\cdots\frac{\partial x^{i_q}}{\partial \bar{x}^{j_q}}\right) p^{\beta}_{i_1\cdots i_q} + \cdots \tag{1.14}$$

where the dots represent a complicated expression which, however, does not depend on any derivative p^{α}_{μ} with $|\mu| = q$ and where $\partial\mathbf{x}/\partial\bar{\mathbf{x}}$ represents the inverse of the Jacobian $\partial\bar{\mathbf{x}}/\partial\mathbf{x}$. But this implies that (1.14) is indeed affine in the derivatives of order q as claimed. □

1.2 An Intrinsic Approach to Jet Bundles

We describe now a direct intrinsic construction of the jet bundles based on Proposition 1.1.4. This construction does not seem to be widely known in the literature, but it is very convenient for our purposes. Strictly speaking, we do not construct the jet bundles as introduced in the previous section but a bundle canonically diffeomorphic to it. Thus we identify it with the jet bundle. We also slightly generalise: instead of jets of maps $\mathcal{X} \to \mathcal{U}$ we consider jets of sections of a fibred manifold $\pi : \mathcal{E} \to \mathcal{X}$ (see Appendix C.2 for some basic material on fibred manifolds). Note that this affects only global properties; locally there is no real difference. The situation of the last section is recovered, if we take the trivial fibration $\mathrm{pr}_1 : \mathcal{E} = \mathcal{X} \times \mathcal{U} \to \mathcal{X}$ with the canonical projection on the first factor.

Definition 1.2.1. *Let $\pi : \mathcal{E} \to \mathcal{X}$ be a fibred manifold. The* first-order jet bundle *over $\mathcal{E}$ is the affine bundle $\pi^1_0 : J_1\mathcal{E} \to \mathcal{E}$ whose fibre at the point $\xi \in \mathcal{E}$ with $\pi(\xi) = x \in \mathcal{X}$ is given by the affine space*

$$(J_1\mathcal{E})_{\xi} = \mathcal{A}^{(1)}_{\xi} = \left\{\gamma \in T^*_x\mathcal{X} \otimes T_{\xi}\mathcal{E} \mid T_{\xi}\pi \circ \gamma = \mathrm{id}_{T_x\mathcal{X}}\right\} \tag{1.15}$$

modelled on the vector space

$$\mathcal{V}^{(1)}_{\xi} = T^*_x\mathcal{X} \otimes V_{\xi}\mathcal{E}\ . \tag{1.16}$$

If we take two arbitrary maps $\gamma_1, \gamma_2 \in \mathcal{A}^{(1)}_{\xi}$, then their difference satisfies $T_{\xi}\pi \circ (\gamma_1 - \gamma_2) = \mathrm{id}_{T_x\mathcal{X}} - \mathrm{id}_{T_x\mathcal{X}} = 0$. Hence $\gamma_1 - \gamma_2$ maps any vector in $T_x\mathcal{X}$ into a vertical vector and $\gamma_1 - \gamma_2 \in \mathcal{V}^{(1)}_{\xi}$. So $\mathcal{A}^{(1)}_{\xi}$ is indeed an affine space modelled on the vector space $\mathcal{V}^{(1)}_{\xi}$. Any element of $\mathcal{A}^{(1)}_{\xi}$ has the form $\gamma(v) = \big(v, B(v)\big)$ with some linear map B.

The affine bundle $J_1\mathcal{E}$ is modelled on the vector bundle $\pi^*(T^*\mathcal{X})\otimes V\mathcal{E} \to \mathcal{E}$. Here the first factor is the pull-back of $T^*\mathcal{X} \to \mathcal{X}$ over $\pi : \mathcal{E} \to \mathcal{X}$. As the pull-back notation is rather cumbersome, we will write this bundle as $T^*\mathcal{X} \underset{\mathcal{E}}{\otimes} V\mathcal{E}$. The $\mathcal{E}$ under the tensor product sign signals that both factors are to be considered as bundles over $\mathcal{E}$ using, if necessary, appropriate pull-backs which are usually obvious from the context.

At first sight, Definition 1.2.1 seems to be completely unrelated to our previous Definition 1.1.1. So we show next that points in $J_1\mathcal{E}$ may be interpreted as equivalence classes of sections $\sigma \in \Gamma_{loc}(\mathcal{X},\mathcal{E})$. For such a section we define its (first) *prolongation* $j_1\sigma \in \Gamma_{loc}(\mathcal{X}, J_1\mathcal{E})$ by

$$j_1\sigma : \begin{cases} \mathcal{X} & \longrightarrow & J_1\mathcal{E} \\ x & \longmapsto & (\sigma(x), T_x\sigma) \end{cases} . \tag{1.17}$$

The tangent map $T_x\sigma$ indeed belongs to the space $\mathcal{A}^{(1)}_\xi$ with $\xi = \sigma(x)$. As σ is a map from $\mathcal{X}$ to $\mathcal{E}$, its tangent map $T_x\sigma$ is a linear map $T_x\mathcal{X} \to T_\xi\mathcal{E}$ and thus an element of $T^*_x\mathcal{X} \otimes T_\xi\mathcal{E}$. Furthermore, by definition of a section, σ satisfies $\pi\circ\sigma = \mathrm{id}_\mathcal{X}$ and hence by the chain rule $T_\xi\pi\circ T_x\sigma = \mathrm{id}_{T_x\mathcal{X}}$. Conversely, to any $\gamma \in \mathcal{A}^{(1)}_\xi$ many sections $\sigma \in \Gamma_{loc}(\mathcal{X},\mathcal{E})$ exist such that $\sigma(x) = \xi$ and $T_x\sigma = \gamma$. The bijection

$$\gamma \in \mathcal{A}^{(1)}_\xi \longleftrightarrow [\sigma] = \bigl\{\sigma \in \Gamma_{loc}(\mathcal{X},\mathcal{E}) \mid \sigma(x) = \xi,\ T_x\sigma = \gamma\bigr\} \tag{1.18}$$

gives us thus a canonical identification between the jet bundles introduced in the last section and those of Definition 1.2.1. Hence we will not distinguish between them in the sequel.

In adapted coordinates $(\mathbf{x},\mathbf{u})$ on $\mathcal{E}$, we may locally write the section σ as $\sigma(\mathbf{x}) = \bigl(\mathbf{x}, s(\mathbf{x})\bigr)$. For the tangent map $T_x\sigma$ we obtain then the local expression $(\mathbf{x},\dot{\mathbf{x}}) \mapsto \bigl(\sigma(\mathbf{x}), d\sigma(\mathbf{x})\cdot\dot{\mathbf{x}}\bigr)$ with the matrix

$$d\sigma(\mathbf{x}) = \begin{pmatrix} \frac{\partial s^1}{\partial x^1}(\mathbf{x}) & \cdots & \frac{\partial s^1}{\partial x^n}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial s^m}{\partial x^1}(\mathbf{x}) & \cdots & \frac{\partial s^m}{\partial x^n}(\mathbf{x}) \\ 1 & 0\cdots 0 & 0 \\ & \ddots & \\ 0 & 0\cdots 0 & 1 \end{pmatrix} . \tag{1.19}$$

It consists of the Jacobian of the function s evaluated at the fixed point $x \in \mathcal{X}$ stacked over the $n \times n$ identity matrix. Conversely, the matrix of any linear map contained in $\mathcal{A}^{(1)}_\xi$ has the form (1.19), i. e. an arbitrary $m \times n$ matrix stacked over the $n \times n$ identity matrix.

Remark 1.2.2. A point $(\xi, \gamma_\xi) \in J_1\mathcal{E}$ induces a decomposition of the tangent space $T_\xi\mathcal{E}$. By definition, the fibre component γ_ξ represents an element of $A^{(1)}_\xi$, i.e. a linear map $T_x\mathcal{X} \to T_\xi\mathcal{E}$ (where again $x = \pi(\xi)$). Because of the condition $T_\xi\pi \circ \gamma_\xi = \mathrm{id}_{T_x\mathcal{X}}$, the intersection $\operatorname{im}\gamma_\xi \cap V_\xi\mathcal{E}$ contains only the zero vector and $T_\xi\mathcal{E} = \operatorname{im}\gamma_\xi \oplus V_\xi\mathcal{E}$. This implies that a (global) section γ of the bundle $\pi^1_0 : J_1\mathcal{E} \to \mathcal{E}$ with $\gamma(\xi) = (\xi, \gamma_\xi)$ may be interpreted as a *connection* (see Appendix C.2) on $\mathcal{E}$ where the horizontal space $H_\xi\mathcal{E}$ at the point $\xi \in \mathcal{E}$ is given by $\operatorname{im}\gamma_\xi$.

In fact, *any* connection on $\mathcal{E}$ can be generated that way. More precisely, there exists a bijection between the space of all connections on $\mathcal{E}$ and $\Gamma(\mathcal{E}, J_1\mathcal{E})$, the space of all global sections $\mathcal{E} \to J_1\mathcal{E}$. Given a section $\gamma \in \Gamma(\mathcal{E}, J_1\mathcal{E})$, we define a connection via its *horizontal lift* $A : \mathcal{E} \underset{\mathcal{X}}{\times} T\mathcal{X} \to T\mathcal{E}$ by $A(\xi, v_x) = \big(\xi, \gamma_\xi(v_x)\big)$. Conversely, given a horizontal lift A, we define a section $\gamma \in \Gamma(\mathcal{E}, J_1\mathcal{E})$ by setting $\gamma_\xi(v_x) = A(\xi, v_x)$. $\triangleleft$

The geometry of the jet bundle is to a large extent determined by its *contact structure*. In the previous section we introduced it in terms of the contact forms locally spanned by the one-forms ω^α_μ defined by (1.8). Their definition is obviously not intrinsic, although our characterisation of contact forms via pull-backs is. In this section we take another approach to the contact structure via the *contact map*. Using Definition 1.2.1, it becomes a simple evaluation map:

$$\Gamma_1 : \begin{cases} J_1\mathcal{E} \underset{\mathcal{X}}{\times} T\mathcal{X} \longrightarrow T\mathcal{E} \\ (\xi, \gamma_\xi, v_x) \longmapsto \big(\xi, \gamma_\xi(v_x)\big) \end{cases} . \tag{1.20}$$

Note the close relation of the contact map to the horizontal lift A introduced in the remark above: given a global section $\gamma : \mathcal{E} \to J_1\mathcal{E}$ we can write $A(\xi, v_x) = \Gamma_1\big(\gamma(\xi), v_x\big)$. The *complementary contact map* is defined by

$$\theta_1 : \begin{cases} J_1\mathcal{E} \underset{\mathcal{E}}{\times} T\mathcal{E} \longrightarrow V\mathcal{E} \\ (\xi, \gamma_\xi, w_\xi) \longmapsto w_\xi - \Gamma_1\big(\xi, \gamma_\xi, T_\xi\pi(w_\xi)\big) \end{cases} . \tag{1.21}$$

Like Γ_1 is related to the horizontal lift A, the complementary contact map θ_1 is related to the vertical projector P^V of the connection induced by a section $\gamma : \mathcal{E} \to J_1\mathcal{E}$, namely $P^V(\xi, w_\xi) = \theta_1\big(\gamma(\xi), w_\xi\big)$.

As the map Γ_1 is linear in its last argument, we may interpret it alternatively as a map $J_1\mathcal{E} \to T^*\mathcal{X} \underset{\mathcal{E}}{\otimes} T\mathcal{E}$. Then its local coordinate form is $\Gamma_1 : (x^i, u^\alpha, p^\alpha_i) \mapsto \big(x^i, u^\alpha; dx^i \otimes (\partial_{x^i} + p^\alpha_i \partial_{u^\alpha})\big)$. It is no accident that the vector fields in the parentheses look exactly like the contact fields (1.10) in the case $q = 1$; after an appropriate pull-back they are indeed identical. Similarly, we interpret θ_1 as a map $J_1\mathcal{E} \to T^*\mathcal{E} \underset{\mathcal{E}}{\otimes} V\mathcal{E}$ with the local coordinate form $\theta_1 : (x^i, u^\alpha, p_i^\alpha) \mapsto \big(x^i, u^\alpha; (du^\alpha - p_i^\alpha dx^i) \otimes \partial_{u^\alpha}\big)$. Now we recover (after some appropriate pull-back) in the parentheses the contact forms (1.8).

The jet bundles of higher order permit a similar direct construction, although it becomes somewhat less elegant, as the characterisation of the appearing affine spaces is more cumbersome. We proceed inductively and define $J_q\mathcal{E}$ as an affine bundle over $J_{q-1}\mathcal{E}$.

Definition 1.2.3. *Let $\pi : \mathcal{E} \to \mathcal{X}$ be a fibred manifold and $J_{q-1}\mathcal{E}$ the jet bundle of order $q-1$. The* qth order jet bundle *over $\mathcal{E}$ is the affine bundle $\pi^q_{q-1} : J_q\mathcal{E} \to J_{q-1}\mathcal{E}$ whose fibre at the point $\rho \in J_{q-1}\mathcal{E}$ with $\pi_0^{q-1}(\rho) = \xi \in \mathcal{E}$ and $\pi(\xi) = x \in \mathcal{X}$ is given by the affine space*

$$\begin{aligned}(J_q\mathcal{E})_\rho = \mathcal{A}_\xi^{(q)} = \Big\{\gamma \in (T_x^*\mathcal{X})^q \otimes T_\xi\mathcal{E} \mid \\ \gamma(v_1, \dots, v_q) = \big(v_q, B_q(v_1, \dots, v_q)\big) \\ \text{with } B_q \text{ totally symmetric}\Big\}\end{aligned} \tag{1.22}$$

modelled on the vector space

$$\mathcal{V}_\xi^{(q)} = S_q(T_x^*\mathcal{X}) \otimes V_\xi\mathcal{E} \tag{1.23}$$

where S_q denotes the q-fold symmetric product.

The elements of $\mathcal{A}_\xi^{(q)}$ may also be considered as maps $T_x\mathcal{X} \to \mathcal{A}_\xi^{(q-1)}$. It follows by the same argument as in the case of the first-order jet bundle that $\mathcal{A}_\xi^{(q)}$ is indeed an affine space. If we take the difference of two elements $\gamma_1, \gamma_2 \in \mathcal{A}_\xi^{(q)}$, we obtain a linear symmetric map whose value is a vertical vector, i. e. an element of $\mathcal{V}_\xi^{(q)}$. The affine bundle $\pi^q_{q-1} : J_q\mathcal{E} \to J_{q-1}\mathcal{E}$ is thus modelled on the vector bundle $S_q(T^*\mathcal{X}) \underset{J_{q-1}\mathcal{E}}{\otimes} V\mathcal{E}$ (writing out the involved pull-backs, this is the bundle $(\pi^{q-1})^*S_q(T^*\mathcal{X}) \otimes (\pi_0^{q-1})^*V\mathcal{E} \to J_{q-1}\mathcal{E}$). It follows from (1.5) (or (A.4a), respectively) that the fibre dimension of this vector bundle is

$$\dim \big(S_q(T^*\mathcal{X}) \underset{J_{q-1}\mathcal{E}}{\otimes} V\mathcal{E}\big) = m\binom{n+q-1}{q}\,. \tag{1.24}$$

Remark 1.2.4. As in the first-order case any global section $J_{q-1}\mathcal{E} \to J_q\mathcal{E}$ defines a connection on $J_{q-1}\mathcal{E}$. We do not give here any details how the splitting of the tangent bundle into a horizontal and a vertical subbundle is performed. This will become obvious below, when we study the relation to iterated jet bundles. ◁

We still have to show that the thus constructed bundle may indeed be identified with the jet bundle. In the case of the first-order jet bundle this was straightforward, as the elements of $\mathcal{A}^{(1)}_\xi$ are of the form $T_x\sigma$ for sections $\sigma : \mathcal{X} \to \mathcal{E}$. For the higher-order jet bundles it is similarly possible to identify the elements of $\mathcal{A}^{(q)}_\xi$ with higher-order tangent maps, i. e. tangent maps of tangent maps etc. As this becomes notationally very messy, we show it explicitly only for the second-order jet bundle.

Consider a section $\sigma : \mathcal{X} \to \mathcal{E}$ with $\sigma(x) = \xi = \big(x, s(x)\big)$. Then $T_x\sigma$ is a map $T_x\mathcal{X} \to T_\xi\mathcal{E}$. Let $w = T_x\sigma(v)$ for some vector $v \in T_x\mathcal{X}$; we have seen above that we may write $w = \big(v, (\partial s/\partial x)v\big)$. The second-order tangent map is a linear map $T_{(x,v)}(T_x\sigma) : T_{(x,v)}(T_x\mathcal{X}) \to T_{(\xi,w)}(T_\xi\mathcal{E})$. A vector $z \in T_{(x,v)}(T_x\mathcal{X})$ can be split into two parts $z = (z_1, z_2)$ with z_1 being "tangent in x-direction" and z_2 in "v-direction". Then

$$T_{(x,v)}(T_x\sigma)(z) = \Big(z_1, \frac{\partial s}{\partial x}v, z_2, \frac{\partial^2 s}{\partial x^2}(v, z_1) + \frac{\partial s}{\partial x}z_2\Big) \,. \tag{1.25}$$

Obviously, this map is completely determined by the Jacobian $\partial s/\partial x$ and the Hessian $\partial^2 s/\partial x^2$. The Jacobian is already encoded in the point $\rho \in J_1\mathcal{E}$ where we consider the fibre of $J_2\mathcal{E}$. The Hessian evaluated at a fixed point $x \in \mathcal{X}$ defines a linear symmetric map B_2 as in Definition 1.2.3. Conversely, to every linear symmetric map B_2 many sections σ exist such that their Hessian at the single point $x \in \mathcal{X}$ coincides with B_2. So we obtain again a bijection between a point in $J_q\mathcal{E}$ and an equivalence class of local sections of $\mathcal{E}$.

As each jet bundle $J_q\mathcal{E}$ is also a fibred manifold over the base space $\mathcal{X}$, we may construct iterated jet bundles $J_r(J_q\mathcal{E})$. Another way to view $J_{q+1}\mathcal{E}$ is then as a submanifold of $J_1(J_q\mathcal{E})$. The mapping $\iota_{q,1} : J_{q+1}\mathcal{E} \to J_1(J_q\mathcal{E})$ describing this identification is defined intrinsically as the prolongation of the identity map on $J_q\mathcal{E}$, i. e. as the unique mapping satisfying for any section $\sigma \in \Gamma_{\mathrm{loc}}(\mathcal{X}, \mathcal{E})$ the condition $\iota_{q,1} \circ j_{q+1}\sigma = j_1(j_q\sigma)$.

As by Definition 1.2.1 the fibre components of elements of $J_1(J_q\mathcal{E})$ may be considered as maps $\gamma : T\mathcal{X} \to TJ_q\mathcal{E}$ satisfying $T\pi^q \circ \gamma = \mathrm{id}_{T\mathcal{X}}$, the same may be done with the elements of $\mathcal{A}^{(q)}$. More precisely, $\mathcal{A}^{(q)}$ is isomorphic to an affine subspace of the space of all such maps γ which is characterised by a symmetry condition.

If we take on $J_q\mathcal{E}$ our usual local coordinates (x^i, p^α_μ), then local coordinates on $J_1(J_q\mathcal{E})$ are $(x^i, p^\alpha_\mu; p^\alpha_{\mu,i})$, i. e. all derivatives p^α_μ are considered as dependent variables and all their first-order derivatives $p^\alpha_{\mu,i}$ are introduced. The submanifold of $J_1(J_q\mathcal{E})$ with which we identify $J_{q+1}\mathcal{E}$ is described by the local equations $p^\alpha_{\mu,i} = p^\alpha_{\nu,j}$ whenever $\mu + 1_i = \nu + 1_j$. Obviously, these equations express that partial derivatives commute.

Similarly, we introduce mappings $\iota_{q,r} : J_{q+r}\mathcal{E} \to J_r(J_q\mathcal{E})$ by the requirement $\iota_{q,r} \circ j_{q+r}\sigma = j_r(j_q\sigma)$ for any section $\sigma \in \Gamma_{\mathrm{loc}}(\mathcal{X}, \mathcal{E})$. Local coordinates on $J_r(J_q\mathcal{E})$ are of the form $p^\alpha_{\mu,\nu}$ with $0 \leq |\mu| \leq q$ and $0 \leq |\nu| \leq r$. The image of $J_{q+r}\mathcal{E}$ in $J_r(J_q\mathcal{E})$ is then described by the equations $p^{\alpha}_{\mu,\nu} = p^{\alpha}_{\bar{\mu},\bar{\nu}}$ whenever $\mu + \nu = \bar{\mu} + \bar{\nu}$.

More generally, we define the prolongation of any map $\Phi : J_q\mathcal{E} \to \mathcal{E}'$ where $\mathcal{E}'$ is another fibred manifold over $\mathcal{X}$ and Φ is fibred over the identity. The r-fold prolongation $\rho_r\Phi$ is a map $J_{q+r}\mathcal{E} \to J_r\mathcal{E}'$ such that the diagram

$$\begin{tikzcd}
J_{q+r}\mathcal{E} \arrow[r, "\pi^{q+r}_r"] \arrow[rrr, bend left, "\rho_r\Phi"] & J_q\mathcal{E} \arrow[r, "\Phi"] & \mathcal{E}' & J_r\mathcal{E}' \arrow[l, "\pi^r_0"'] \\
& \mathcal{X} \arrow[ul, "j_{q+r}\sigma"] \arrow[u, "j_q\sigma"'] \arrow[r, equal] & \mathcal{X} \arrow[u, "\sigma'"] \arrow[ur, "j_r\sigma'"'] &
\end{tikzcd} \tag{1.26}$$

commutes for any local section $\sigma \in \Gamma_{loc}(\mathcal{X}, \mathcal{E})$, i.e. the prolonged map satisfies $j_r(\Phi \circ j_q\sigma) = (\rho_r\Phi) \circ (j_{q+r}\sigma)$.

If we use adapted coordinates $(\mathbf{x}, \mathbf{u})$ on $\mathcal{E}$ and $(\mathbf{x}, \mathbf{v})$ on $\mathcal{E}'$, so that locally the map Φ is of the form $v^{\tau} = \Phi^{\tau}(\mathbf{x}, \mathbf{p})$, then the prolongation $\rho_r\Phi$ has the coordinate form $q^{\tau}_{\sigma} = D_{\sigma}\Phi^{\tau}(\mathbf{x}, \mathbf{p})$ with $0 \leq |\sigma| \leq r$ (we denote the jet coordinates of $J_r\mathcal{E}'$ by q^{τ}_{σ}). Here $D_{\sigma} = D_1^{\sigma_1} \cdots D_n^{\sigma_n}$ with D_i the formal derivative defined by (1.11). Thus (1.26) is just an abstract representation of the chain rule.

The *contact structure* of the higher-order jet bundle may be introduced in a similar manner to the first-order case. The q-contact map is a map $\Gamma_q : J_q\mathcal{E} \times_{\mathcal{X}} T\mathcal{X} \to TJ_{q-1}\mathcal{E}$ such that the diagram

$$\begin{tikzcd}
J_q\mathcal{E} \times_{\mathcal{X}} T\mathcal{X} \arrow[rr, "\Gamma_q"] & & TJ_{q-1}\mathcal{E} \\
& T\mathcal{X} \arrow[ul, "((j_q\sigma)\circ\tau_{\mathcal{X}})\times \mathrm{id}_{T\mathcal{X}}"] \arrow[ur, "Tj_{q-1}\sigma"'] &
\end{tikzcd} \tag{1.27}$$

commutes for any section $\sigma \in \Gamma_{loc}(\mathcal{X}, \mathcal{E})$. This condition defines it uniquely. It is easy to see that Γ_q is a linear fibred morphism over π^q_{q-1} which can be expressed by the fact that the diagram

$$\begin{tikzcd}
J_q\mathcal{E} \times_{\mathcal{X}} T\mathcal{X} \arrow[r, "\Gamma_q"] \arrow[d, "\mathrm{pr}_1"'] & TJ_{q-1}\mathcal{E} \arrow[d, "\tau_{J_{q-1}\mathcal{E}}"] \\
J_q\mathcal{E} \arrow[r, "\pi^q_{q-1}"'] & J_{q-1}\mathcal{E}
\end{tikzcd} \tag{1.28}$$

commutes. Here pr_1 is the projection on the first factor and τ the tangent bundle projection. The linearity implies that we may consider Γ_q as a map

$J_q\mathcal{E} \to T^*\mathcal{X} \underset{J_{q-1}\mathcal{E}}{\otimes} TJ_{q-1}\mathcal{E}$ which takes in adapted coordinates the concise form $\Gamma_q : (x^i, u^\alpha, p^\alpha_\mu) \mapsto \bigl(x^i, u^\alpha, p^\alpha_\mu; dx^i \otimes (\partial_{x^i} + p^\alpha_{\mu+1_i}\partial_{p^\alpha_\mu})\bigr)$. Note that on the right hand side only multi indices with $0 \leq |\mu| < q$ appear whereas on the left hand side $|\mu| = q$ is admitted, too. As in the first-order case, it is not difficult to recognise the contact fields (1.10) in the parentheses.

The *complementary contact map* of the qth order jet bundle is the map $\theta_q : J_q\mathcal{E} \underset{J_{q-1}\mathcal{E}}{\times} TJ_{q-1}\mathcal{E} \to VJ_{q-1}\mathcal{E}$ defined by $\theta_q(\rho, w) = w - \Gamma_q\bigl(\rho, T_{\bar\rho}\pi^{q-1}(w)\bigr)$ where $\bar\rho = \pi^q_{q-1}\rho$. It takes in adapted coordinates on the jet bundle $J_q\mathcal{E}$ the form $\theta_q : (x^i, u^\alpha, p^\alpha_\mu) \mapsto \bigl(x^i, u^\alpha, p^\alpha_\mu; (dp^\alpha_\mu - p^\alpha_{\mu+1_i}dx^i) \otimes \partial_{p^\alpha_\mu}\bigr)$ where now the contact forms (1.8) reappear. Given a global section $J_{q-1}\mathcal{E} \to J_q\mathcal{E}$ defining a connection on $J_{q-1}\mathcal{E}$ we get the same relation between Γ_q and the horizontal lift on one hand and θ_q and the vertical projector on the other hand as in the first-order case.

1.3 Differential Equations

With the concept of jet bundles at our hands we give a geometric definition of (systems of) differential equations. It does not require coordinates or local equations but defines a differential equation as an intrinsic geometric object, namely a fibred submanifold.

Definition 1.3.1. *A (non-linear)* differential equation *of order* q *is a fibred submanifold* $\mathcal{R}_q$ *of the jet bundle* $J_q\mathcal{E}$*.*

Locally, i. e. within some open subset $\mathcal{V}_q \subseteq J_q\mathcal{E}$, we may always assume that a fibred map

$$\Phi : \begin{cases} \mathcal{V}_q \subseteq J_q\mathcal{E} \longrightarrow \quad \mathcal{E}' \\ \quad (\mathbf{x}, \mathbf{p}) \quad \longmapsto \Phi(\mathbf{x}, \mathbf{p}) \end{cases} \tag{1.29}$$

exists taking values in a vector bundle $\pi' : \mathcal{E}' \to \mathcal{X}$ such that $\mathcal{R}_q \cap \mathcal{V}_q$ is the zero set of Φ. Thus locally the differential equation $\mathcal{R}_q$ can be described by a system of equations $\Phi^\tau(\mathbf{x}, \mathbf{p}) = 0$ with $1 \leq \tau \leq t = \dim \mathcal{E}'$ which brings us back to the traditional picture.

In the sequel we will mostly assume that there exists a *global* morphism $\Phi : J_q\mathcal{E} \to \mathcal{E}'$ such that $\mathcal{R}_q$ is its zero set, i. e. that our differential equation is defined by the *differential operator* $\Delta[\Phi] : \Gamma_{loc}(\mathcal{E}) \to \Gamma_{loc}(\mathcal{E}')$ with $\Delta[\Phi](\sigma) = \Phi \circ j_q\sigma$. As we have mainly local applications in mind, this assumption is not a serious restriction but simplifies the notation at some places. If $\pi : \mathcal{E} \to \mathcal{X}$ is also a vector bundle, we define *linear differential equations* by the condition that Φ is a vector bundle morphism. Note that the linearity refers only to the derivatives; the dependency on the independent variables may be arbitrary.

Definition 1.3.1 does *not* distinguish between scalar equations and systems. In fact, we will use in the sequel the following convention. When we speak of a *differential equation*, we always think of it as such a geometric object $\mathcal{R}_q$. When we speak of a *system*, we think of a local representation of $\mathcal{R}_q$ given by a map of the form (1.29). Of course, such a system consists again of equations $\Phi(x, p) = 0$, so that it is somewhat difficult to strictly maintain this distinction ...

Definition 1.3.1 requires $\mathcal{R}_q$ to be a *fibred* submanifold. This condition implies that $\pi^q(\mathcal{R}_q) = \mathcal{X}$. If it was violated, the independent variables would not be really independent, as the differential equation would impose relations between them. More precisely, some local representations of $\mathcal{R}_q$ would contain equations of the form $\Phi(\mathbf{x}) = 0$ which obviously should not happen.

Remark 1.3.2. First-order *ordinary* differential equations $\mathcal{R}_1 \subset J_1\mathcal{E}$ possess a special geometric structure. If $\mathcal{R}_1$ is the image of a section $\gamma : \mathcal{E} \to J_1\mathcal{E}$, we may interpret it according to Remark 1.2.2 as a connection on $\mathcal{E}$. Locally, such equations correspond to systems of the form $\dot{\mathbf{u}} = \boldsymbol{\phi}(t, \mathbf{u})$, i. e. we are dealing with a pure differential equation without algebraic constraints. If constraints are present, they determine a fibred submanifold $\mathcal{C} \subset \mathcal{E}$ and we may consider equations defined by sections $\gamma : \mathcal{C} \to J_1\mathcal{E}$. They correspond to connections on $\mathcal{C}$. Obviously, this can straightforwardly be extended to an identification of qth-order ordinary differential equations with connections on $J_{q-1}\mathcal{E}$.

In the case of *partial* differential equations, the images of global sections $\gamma : \mathcal{E} \to J_1\mathcal{E}$ yield *equations of finite type*, also called *maximally overdetermined* equations (see Remark 1.5.2). If such an equation is of first order, it can locally be represented by a system of the form $\mathbf{p} = \boldsymbol{\phi}(\mathbf{x}, \mathbf{u})$ where the vector $\mathbf{p}$ contains all first-order derivatives. Again the inclusion of constraints, i. e. of additional algebraic equations defining a fibred submanifold $\mathcal{C} \subset \mathcal{E}$, is trivial. $\triangleleft$

Two natural geometric operations with differential equations exist: *prolongation* and *projection*. The former one “lifts” an equation to higher order; the latter one lowers the order. Both are of central importance for all what we will do with differential equations.

Prolonging a differential equation is easy to perform effectively, once a local representation is given. Its intrinsic geometric description is a bit more involved. We need again the embeddings $\iota_{q,r} : J_{q+r}\mathcal{E} \hookrightarrow J_r(J_q\mathcal{E})$. As $\mathcal{R}_q$ is a fibred manifold, too, we can construct jet bundles over it, for example $J_r\mathcal{R}_q$. It may be consider in a natural way as a submanifold of $J_rJ_q\mathcal{E}$. We denote the corresponding embedding by $\bar{\iota}_{q,r}$. Then we define the prolongation

$$\mathcal{R}_{q+r} = \iota_{q,r}^{-1}\bigl(\bar{\iota}_{q,r}(J_r\mathcal{R}_q) \cap \iota_{q,r}(J_{q+r}\mathcal{E})\bigr) \subseteq J_{q+r}\mathcal{E} \ . \tag{1.30}$$

In words, this somewhat awkward construction has a fairly simple explanation. We consider both $J_r\mathcal{R}_q$ and $J_{q+r}\mathcal{E}$ as submanifolds of $J_r(J_q\mathcal{E})$ using the corresponding embeddings. Then we take their intersection and consider the result as a submanifold of $J_{q+r}\mathcal{E}$. As already discussed in the last section, the intersection enforces the commutativity of partial differentiations.

If our differential equation $\mathcal{R}_q$ comes from a differential operator, its prolongation can be expressed via the prolongation of the corresponding global map Φ. We get the following commuting diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{R}_{q+r} \arrow[r] \arrow[d, "\pi^{q+r}_q"'] & J_{q+r}\mathcal{E} \arrow[r, "\rho_r\Phi"] \arrow[d, "\pi^{q+r}_q"'] & J_r\mathcal{E}' \arrow[d, "(\pi')^r_0"] \\
0 \arrow[r] & \mathcal{R}_q \arrow[r] & J_q\mathcal{E} \arrow[r, "\Phi"] & \mathcal{E}'
\end{tikzcd} \tag{1.31}$$

and the prolonged equation $\mathcal{R}_{q+r}$ is the zero set of the prolonged map $\rho_r\Phi$.

In local coordinates, the prolongation is computed with the help of the formal derivative (1.11). If $\Phi(\mathbf{x}, \mathbf{p}) = 0$ is a local representation of $\mathcal{R}_q$, then $\mathcal{R}_{q+r}$ is described locally by the system

$$\mathcal{R}_{q+r} : \begin{cases} \Phi^\tau(\mathbf{x}, \mathbf{p}) &= 0 \ , \qquad 1 \le \tau \le t \ , \\ D_j\Phi^\tau(\mathbf{x}, \mathbf{p}) &= 0 \ , \qquad 1 \le j \le n \ , \\ \quad \vdots & \\ D_\nu\Phi^\tau(\mathbf{x}, \mathbf{p}) &= 0 \ , \qquad |\nu| = r \ . \end{cases} \tag{1.32}$$

Indeed, these are just the coordinate expressions of $\rho_r\Phi$.

In contrast, projections are easily described geometrically but difficult to perform effectively in local coordinates. Geometrically we simply apply the natural projections $\pi^q_{q-r} : J_q\mathcal{E} \to J_{q-r}\mathcal{E}$ and define for $r \le q$

$$\mathcal{R}^{(r)}_{q-r} = \pi^q_{q-r}(\mathcal{R}_q) \ . \tag{1.33}$$

To obtain a local representation of the differential equation $\mathcal{R}^{(r)}_{q-r} \subseteq J_{q-r}\mathcal{E}$, one takes a local representation of $\mathcal{R}_q$, say $\Phi^\tau(x, p) = 0$, and tries to eliminate by purely algebraic operations the derivatives of order greater than $q-r$ in as many equations as possible. If we cannot construct any equation depending only on lower-order derivatives, then $\mathcal{R}^{(r)}_{q-r} = J_{q-r}\mathcal{E}$. Otherwise it is a proper submanifold locally described by all such equations.

We will see later many examples of such computations. Fortunately, we mostly project differential equations generated by prolongations. In this case, the elimination of the higher-order derivatives requires only linear algebra. A slight complication arises from the fact that the prolongation or the projection of a given differential equation $\mathcal{R}_q$ is not necessarily a submanifold anymore. For projections this is rather obvious and the following example shows a counter example for a prolongation.

Example 1.3.3. We consider the *non-linear* second-order equation

$$\mathcal{R}_2 : \begin{cases} u_{yy} - \frac{1}{2}(u_{xx})^2 = 0 , \\ u_{xy} - u_{xx} \qquad = 0 . \end{cases} \tag{1.34}$$

Adding the x- and the y-derivative of the second equation and subtracting the x-derivative of the first one yields $(u_{xx} - 1)u_{xxx} = 0$. Thus $\mathcal{R}_3$ lies in the union of the two hyperplanes $u_{xx} = 1$ and $u_{xxx} = 0$. Such a union is *not* a submanifold, as in the neighbourhood of the intersection it is not homeomorphic to an open subset of some $\mathbb{R}^k$. We must distinguish two cases in which the differential equation might show quite different behaviour; in particular, the solution spaces will generally differ. $\lhd$

For simplicity, we will always assume that all performed projections and prolongations yield again submanifolds. This means that we consider only *regular* equations. Regularity can always be achieved by restricting to some submanifold of $\mathcal{R}_q$ but is difficult to check effectively.

One could be tempted to think that prolongation and projection are some kind of "inverse" operations: if one first prolongs a differential equation $\mathcal{R}_q$ to $\mathcal{R}_{q+r}$ for some $r > 0$ and subsequently projects back to $J_q\mathcal{E}$ with π^{q+r}_q, one might naively expect that the obtained equation $\mathcal{R}^{(r)}_q$ is identical with the original equation $\mathcal{R}_q$. However, this is in general not correct, as *integrability conditions* may arise: we only get that always $\mathcal{R}^{(r)}_q \subseteq \mathcal{R}_q$.

Example 1.3.4. A trivial example may explain this phenomenon and also demonstrate explicitly the operations of prolongation and projection. We consider the first-order equation $\mathcal{R}_1$ defined by the linear system

$$\mathcal{R}_1 : \begin{cases} u_z + yu_x = 0 , \\ u_y \qquad\quad = 0 . \end{cases} \tag{1.35}$$

In order to obtain the system for the prolonged equation $\mathcal{R}_2$ we must formally differentiate all equations in (1.35) with respect to all independent variables x, y, and z. The arising equations are added to the original system and we get

$$\mathcal{R}_2 : \begin{cases} u_z + yu_x = 0 , & u_y = 0 , \\ u_{xz} + yu_{xx} = 0 , & u_{xy} = 0 , \\ u_{yz} + yu_{xy} + u_x = 0 , & u_{yy} = 0 , \\ u_{zz} + yu_{xz} = 0 , & u_{yz} = 0 . \end{cases} \tag{1.36}$$

In order to compute the projection to $J_1\mathcal{E}$ of the prolongation $\mathcal{R}_2$ we must try to obtain by algebraic manipulations as many first-order equations as possible. Since we are dealing with a linear system, this is not very difficult. The first two equations are obviously first order, as they stem from the original system. In the first equation of the third row we can eliminate both second-order derivatives using equations of the second column. No other eliminations are possible. So we find the integrability condition $u_x = 0$ and a local representation of the projected equation $\mathcal{R}_1^{(1)}$ is given by

$$\mathcal{R}_1^{(1)} : \begin{cases} u_z = 0 , \\ u_y = 0 , \\ u_x = 0 . \end{cases} \tag{1.37}$$

It is easy to see that further prolongations and subsequent projections do not yield any additional conditions and thus $\mathcal{R}_1^{(1)} = \mathcal{R}_1^{(2)} = \mathcal{R}_1^{(3)} = \cdots$. ◁

It is important to note that in Example 1.3.4 it is not possible to construct the integrability condition by purely algebraic operations without differentiations. This property may serve as an informal definition of integrability conditions: they are equations algebraically independent of the original system that can be generated using formally differentiated equations. Obviously, any sufficiently smooth solution of the system *automatically* satisfies the integrability conditions; thus they do not represent additional restrictions to the solution space but they are equations that have been "hidden" in the structure of the system.

In Example 1.3.4 we see the classical form of an integrability condition. It arises as a generalised cross-derivative: a linear combination of differentiated equations that yields a lower-order equation. Obviously, cross-derivatives make sense only for partial differential equations. However, a second mechanism for the generation of integrability conditions exists which also works for ordinary differential equations: if the original system contains equations of differing orders, the differentiation of a lower-order equation may generate an integrability condition.

Example 1.3.5. We consider the trivial second-order equation $\mathcal{R}_2$ described in local coordinates by the linear system

$$\mathcal{R}_2 : \begin{cases} u_{yy} = 0 , \\ u_x = 0 . \end{cases} \tag{1.38}$$

Prolongation to third order yields a system that contains the second-order equations $u_{xy} = u_{yy} = 0$. Thus the projection to $J_2\mathcal{E}$ is not surjective and we must add these two integrability conditions to the system (1.38) in order to obtain a local representation of $\mathcal{R}_2^{(1)}$. $\lhd$

Geometrically, it makes no sense to speak of individual integrability conditions or to distinguish between cross-derivatives and the prolongation of lower-order equations. Instead both kinds of integrability conditions have the same effect, namely that for some $r > 0$ the differential equation $\mathcal{R}_q^{(r)}$ is a proper subset of $\mathcal{R}_q$. More precisely, it is a submanifold of lower dimension. In Example 1.3.4 $\mathcal{R}_1$ is two-dimensional (the fibre dimension of $J_1\mathcal{E}$ is four and we have two independent equations), whereas $\dim \mathcal{R}_1^{(1)} = 1$.

Finally, we need of course the notion of a solution. Classically, solutions are functions that entered into a given system let it vanish identically. In our geometric framework, functions are substituted by sections.

Definition 1.3.6. *Let* $\mathcal{R}_q \subseteq J_q\mathcal{E}$ *be a differential equation of order* q. *A* (local) solution *is a smooth section* $\sigma \in \Gamma_{\mathrm{loc}}(\mathcal{E})$ *such that* $\operatorname{im} j_q\sigma \subseteq \mathcal{R}_q$.

In local coordinates, such a section $\sigma : \mathcal{V} \subseteq \mathcal{X} \to \mathcal{E}$ can be written as a mapping $x \mapsto \big(x, \mathbf{s}(x)\big)$. Its prolongation $j_q\sigma : \mathcal{V} \to J_q\mathcal{E}$ has locally the form $x \mapsto \big(x, \mathbf{s}(x), \partial^{|\mu|}\mathbf{s}(x)/\partial x^\mu\big)$. If the differential equation $\mathcal{R}_q$ is locally represented by the system $\Phi(x, \mathbf{p}) = 0$, the condition $\operatorname{im} j_q\sigma \subseteq \mathcal{R}_q$ requires that $\Phi(x, \partial^{|\mu|}\mathbf{s}(x)/\partial x^\mu) = 0$ for all $x \in \mathcal{V}$, i. e. the (smooth) function $\mathbf{s}(x)$ defined on $\mathcal{V}$ is a solution in the conventional sense.

Finally, we compare our definition of a differential equation with the situation in algebraic geometry. There a variety is the zero set of some polynomials, i. e. every point on the variety represents a solution of the algebraic equations defined by the polynomials. Let the differential equation $\mathcal{R}_q \subseteq J_q\mathcal{E}$ be locally represented by the system $\Phi(x, \mathbf{p}) = 0$. It may be understood as a system of *algebraic* equations on $J_q\mathcal{E}$ and the points on $\mathcal{R}_q$ are the solutions of these equations. As in algebraic geometry, the functions Φ do not really matter but only the ideal in $\mathcal{F}(J_q\mathcal{E})$ generated by them.

Recall that a point of the jet bundle $J_q\mathcal{E}$ was defined as an equivalence class of sections. Let the class $[\sigma]_x^{(q)}$ be a point in $J_q\mathcal{E}$ and $\mathcal{V} \subseteq \mathcal{X}$ a neighbourhood of x such that σ is defined on $\mathcal{V}$. If the section σ is a solution of the differential equation $\mathcal{R}_q$ in the sense of Definition 1.3.6, the points $[\sigma]_x^{(q)}$ (for all $x \in \mathcal{V}$) must lie on the submanifold $\mathcal{R}_q$. However, the converse does not necessarily hold and the equivalence classes $[\sigma]_x^{(q)}$ in general contain many sections that are *not* solutions of $\mathcal{R}_q$. This is easy to understand, as the equivalence relation $\sim_q$ takes into account only the values of the derivatives of σ up to order q at a *single* point x.

1.4 Formal Integrability

We have seen that in general integrability conditions arise, if we prolong and project a differential equation. We now show that such conditions are obstructions for the construction of power series solutions. Obviously, such a computation can only be performed in local coordinates. Thus we assume that we are given a differential equation $\mathcal{R}_q$ described in some neighbourhood of the point $\mathbf{x}_0 \in \mathcal{X}$ by the system $\Phi(\mathbf{x}, \mathbf{p}) = 0$ and make the usual ansatz for a power series solution

$$u^\alpha(x) = \sum_{|\mu|=0}^{\infty} \frac{a^\alpha_\mu}{\mu!}(\mathbf{x} - \mathbf{x}_0)^\mu \tag{1.39}$$

with real coefficients a^α_μ. Note that we are only dealing with formal series, as we do not discuss their convergence (but see Chapter 5).

Of course, without initial or boundary conditions the equation possesses in general many solutions and we cannot determine uniquely all coefficients in (1.39). But we can at least exhibit some relations between them. We enter the ansatz into the system describing $\mathcal{R}_q$ and evaluate the arising expressions at the point $\mathbf{x} = \mathbf{x}_0$. One easily checks that this yields the following *algebraic* equations for the coefficients a^α_μ with $|\mu| \leq q$:

$$\Phi^\tau(x^i_0, a^\alpha_\mu) = 0\,, \qquad 1 \leq \tau \leq t\,, \tag{1.40}$$

i. e. in the functions Φ^τ we only substitute the derivatives p^α_μ by the coefficients a^α_μ of the ansatz (1.39). In general, (1.40) represent a non-linear system of equations. Hence its solution space might have a very complicated structure. For example, it could consist of disjoint parts of different dimensions. Each of these must be treated separately.

Now we turn our attention to the prolonged equation $\mathcal{R}_{q+1}$. According to (1.32) it is described in a neighbourhood of $\mathbf{x}_0$ by the original equations $\Phi(\mathbf{x}, \mathbf{p}) = 0$ and in addition by the equations $D_i\Phi(\mathbf{x}, \mathbf{p}) = 0$ for $1 \leq i \leq n$. Again we enter our ansatz into this system and evaluate at $\mathbf{x} = \mathbf{x}_0$. The first part of the arising algebraic system for the coefficients a^α_μ with $|\mu| \leq q + 1$ has already been handled in the previous step. The only new equations are

$$D_j\Phi^\tau(x^i_0, a^\alpha_\mu) = 0\,, \qquad \begin{cases} 1 \leq \tau \leq t\,, \\ 1 \leq j \leq n\,. \end{cases} \tag{1.41}$$

If we assume that all relations between the coefficients up to order q are already taken into account with (1.40), then (1.41) is an inhomogeneous *linear* system for the coefficients of order $q + 1$. Its matrix and right hand side depends on the coefficients of lower order. The linearity of the system stems from the quasi-linearity of the formal derivative (1.11).

In this manner we proceed order by order. At order $q + r$ we get from the local representation of $\mathcal{R}_{q+r}$ as additional equations for the coefficients a^α_μ with $|\mu| \leq q + r$

$$D_\nu \Phi^\tau(x^i_0, a^\alpha_\mu) = 0 \ , \qquad \begin{cases} 1 \leq \tau \leq t \ , \\ |\nu| = r \ . \end{cases} \tag{1.42}$$

Again we may consider it as an inhomogeneous linear system for the coefficients of order $q + r$.

In the end we obtain a system of infinitely many equations for the infinitely many coefficients of the ansatz (1.39). Each solution of this system defines a formal power series solution. Except (1.40) all systems are linear and their solution spaces have a very simple structure. But if we choose different solutions of (1.40), the solution spaces of the linear systems (1.42) may change their dimension, as their matrices depend on the coefficients of lower order. This corresponds to the existence of so-called *singular solutions* briefly discussed in Section 4.1.

Each of these infinitely many equations can be solved for a different coefficient a^α_μ of our ansatz. These coefficients are called *principal*, the remaining ones are the *parametric coefficients*. Of course, this distinction is not uniquely defined; in general, there are many ways to solve the equations for some coefficients. The parametric coefficients may be chosen arbitrarily; they parametrise the formal solution space. The principal coefficients are determined by the differential equation $\mathcal{R}_q$ and its prolongations.

In a concrete computation we can perform only a finite number of prolongations. Thus we must stop the above construction at some order $s \geq q$. Of course, we expect that with the help of the so far found equations for the coefficients a^α_μ with $0 \leq |\mu| \leq s$ we obtain at least a correct truncation of power series solutions. However, this is only true, if no integrability conditions of order less than or equal to s are hidden in the system describing $\mathcal{R}_s$, as otherwise some further conditions on the coefficients a^α_μ with $0 \leq |\mu| \leq s$ exist which we have not taken into account.

Example 1.4.1. In Example 1.3.4 $\mathcal{R}^{(1)}_1 \neq \mathcal{R}_1$. Entering a power series ansatz into (1.35) yields only that we must have $a_{[0,0,1]} + y_0 a_{[1,0,0]} = 0$ which can be satisfied in many ways. Using (1.37) we get that in fact $a_{[0,0,1]} = a_{[1,0,0]} = 0$ for any power series solution of $\mathcal{R}_1$. ◁

As this example demonstrates, one should always remember that integrability conditions are not additional restrictions of the solution space but that any solution of the original system *automatically* satisfies them.

Definition 1.4.2. *The differential equation* $\mathcal{R}_q$ *is called* formally integrable*, if for all* $r \geq 0$ *the equality* $\mathcal{R}^{(1)}_{q+r} = \mathcal{R}_{q+r}$ *holds.*

Thus an equation is formally integrable, if at no order of prolongation integrability conditions arise. For such equations the above method for the construction of formal power series solutions works without problems which explains the name "formal integrability". Furthermore, we can be sure that we do not overlook any hidden conditions on the lower order coefficients, if we build only a truncated series. In Example 1.3.4, the equation $\mathcal{R}_1$ is obviously not formally integrable, but one easily shows that $\mathcal{R}^{(1)}_1$ is.

Example 1.4.3. In Remark 1.3.2 we introduced differential equations of finite type and saw that we may interpret them as connections. Now we exploit this observation for the integration of such equations. For simplicity of notation, we restrict to first-order equations $\mathcal{R}_1 \subset J_1\mathcal{E}$ which are the image of a section $\gamma : \mathcal{E} \to J_1\mathcal{E}$. Locally, we choose a representation of the form

$$\mathcal{R}_1 \,:\, \{\; p^\alpha_i = \phi^\alpha_i(x, \mathbf{u})\;, \qquad 1 \leq i \leq n\;. \tag{1.43}$$

A natural question is when such an equation is formally integrable. As (1.43) contains no equations of lower order, the only possibility for integrability conditions are cross-derivatives

$$0 = p^\alpha_{ij} - p^\alpha_{ji} = D_i\phi^\alpha_j - D_j\phi^\alpha_i = \Delta^\alpha_{ij}\;. \tag{1.44}$$

Evaluation of these on the submanifold $\mathcal{R}_1$ yields the integrability conditions

$$\Delta^\alpha_{ij}\Big|_{\mathcal{R}_1} = \frac{\partial \phi^\alpha_j}{\partial x^i} + \frac{\partial \phi^\alpha_j}{\partial u^\beta}\phi^\beta_i - \frac{\partial \phi^\alpha_i}{\partial x^j} + \frac{\partial \phi^\alpha_i}{\partial u^\beta}\phi^\beta_j = 0\;. \tag{1.45}$$

Obviously, the satisfaction of these equations is a necessary condition for the formal integrability of $\mathcal{R}_1$. In our particular case of a maximally overdetermined equation, it is not difficult to see that this is also a sufficient condition: any cross-derivative in a prolonged equation $\mathcal{R}_{1+r}$ is a prolongation of one of the $\Delta^\alpha_{ij}|_{\mathcal{R}_1}$. We do not show this here in detail, as we will prove a much more general result in Section 3.2.

A more geometric way to view these integrability conditions starts with the introduction of the vector fields $X_i \in \mathfrak{X}(\mathcal{E})$ locally defined by

$$X_i\Big|_{(x,\mathbf{u})} = \partial_{x^i} + \phi^\alpha_i(x, \mathbf{u})\partial_{u^\alpha} \qquad 1 \leq i \leq n\;. \tag{1.46}$$

They span the horizontal distribution of the connection on $\mathcal{E}$ defined by the section γ. Indeed, if we interpret as in Section 1.2 the fibre component of each point $(x, \mathbf{u}, \mathbf{p}) \in \mathcal{R}_1$ as a linear map $T_x\mathcal{X} \to T_{(x,\mathbf{u})}\mathcal{E}$, then in our local coordinates its matrix is given by stacking the $m \times n$ matrix defined by ϕ_i^α over the $n \times n$ identity matrix and the horizontal lift of the basis vector ∂_{x^i} yields the vector X_i.

The integrability conditions (1.45) are equivalent to the vanishing of all Lie brackets $[X_i, X_j]$. In other words, $\mathcal{R}_1$ is formally integrable, if and only if the section γ defines a *flat* connection with an involutive horizontal distribution. By the Frobenius Theorem (see Appendix C.3), any involutive distribution is integrable and the n-dimensional horizontal distribution $\mathcal{H} = \langle X_1, \dots, X_n \rangle \subset T\mathcal{E}$ possesses n-dimensional integral manifolds $\mathcal{N} \subset \mathcal{E}$.

We claim that these integral manifolds are images of solutions of $\mathcal{R}_1$. Indeed, let $\sigma : \mathcal{U} \subseteq \mathcal{X} \to \mathcal{E}$ be a (local) solution defined in some open set $\mathcal{U}$. In our case this implies that $\gamma(\sigma(x)) = j_1\sigma(x)$ for all $x \in \mathcal{U}$. According to our intrinsic approach to jet bundles, we can write $j_1\sigma(x) = \big(\sigma(x), T_x\sigma\big)$ and obtain that $\operatorname{im} T_x\sigma = \mathcal{H}_{\sigma(x)}$. But this equality entails that $\operatorname{im}\sigma$ is an integral manifold of the horizontal distribution $\mathcal{H}$.

Conversely, let $\mathcal{N} \subset \mathcal{E}$ be an n-dimensional integral manifold of $\mathcal{H}$. Obviously, it is transversal to the fibration $\pi : \mathcal{E} \to \mathcal{X}$ and hence the image of a (local) section $\sigma : \mathcal{U} \subseteq \mathcal{X} \to \mathcal{E}$. By definition of an integral manifold, this implies that $\operatorname{im} T_x\sigma = \mathcal{H}_{\sigma(x)}$ or, equivalently, $\gamma(\sigma(x)) = j_1\sigma(x)$ for all $x \in \mathcal{U}$. Thus $\operatorname{im} j_1\sigma \subset \mathcal{R}_1$ and σ is a solution.

This result has also practical implications: the concrete determination of integral manifolds of an involutive distribution requires only the solution of *ordinary* differential equations, although we are dealing with a partial differential equation. This was already shown by Lie (see [107]) interpreting the vector fields X_i as generators of Lie point symmetries [129] of $\mathcal{R}_1$. Indeed, let an n-dimensional integral manifold $\mathcal{N}$ of $\mathcal{H}$ be locally described by the m equations $\psi^\alpha(\mathbf{x}, \mathbf{u}) = 0$. As the vector fields X_i span the tangent space of $\mathcal{N}$, we must have $d\psi^\alpha(X_i) = X_i\psi^\alpha = 0$ (cf. Appendix C.3). Thus it suffices to determine all solutions of the overdetermined system $X_i v = 0$ for one unknown $v(\mathbf{x}, \mathbf{u})$. Such systems are studied in Example 3.2.5 where it is shown that their solution requires only the integration of ordinary differential equations. $\triangleleft$

1.5 (Geometric) Symbol and Principal Symbol

In our introduction of jet bundles we have stressed several times the fact that $J_q\mathcal{E}$ is an affine bundle over $J_{q-1}\mathcal{E}$ modelled on the vector space $S_q(T^*\mathcal{X}) \otimes V\mathcal{E}$; in fact our intrinsic approach to jet bundles was built on it. The reason for this emphasis becomes apparent now. It allows us to introduce a very useful geometric object, namely the symbol of a differential equation. Many properties of a differential equation can be read off its symbol and it effectively reduces some problems to questions in linear algebra.

Definition 1.5.1. *Let* $\mathcal{R}_q \subseteq J_q\mathcal{E}$ *be a differential equation. The* (geometric) symbol $\mathcal{M}_q$ *of* $\mathcal{R}_q$ *is the family of vector spaces over* $\mathcal{R}_q$ *with the vector space at the point* $\rho \in \mathcal{R}_q$ *given by*

$$(\mathcal{M}_q)_\rho = V^{(q)}_\rho \mathcal{R}_q \subseteq V^{(q)}_\rho J_q\mathcal{E} \; . \tag{1.47}$$

Here $V^{(q)}_\rho J_q\mathcal{E} = \ker T_\rho \pi^q_{q-1}$ *is the vertical bundle with respect to the fibration* $\pi^q_{q-1} : J_q\mathcal{E} \to J_{q-1}\mathcal{E}$.

Thus the geometric symbol is the vertical part of the tangent space of $\mathcal{R}_q$ with respect to the fibration π^q_{q-1} (here we identify $T\mathcal{R}_q$ with a subspace of $TJ_q\mathcal{E}$). But as mentioned above, for this fibration the vertical part is isomorphic to $S_q(T^*\mathcal{X}) \underset{J_{q-1}\mathcal{E}}{\otimes} V\mathcal{E}$, so that we may consider $\mathcal{M}_q$ as a subspace of the latter vector bundle. If the differential equation $\mathcal{R}_q$ is globally described by the map $\Phi : J_q\mathcal{E} \to \mathcal{E}'$ with a vector bundle $\pi' : \mathcal{E}' \to \mathcal{X}$, then we introduce the *symbol map* $\sigma : S_q(T^*\mathcal{X}) \underset{J_{q-1}\mathcal{E}}{\otimes} V\mathcal{E} \to T\mathcal{E}'$ given by $\sigma = T\Phi\big|_{V^{(q)}J_q\mathcal{E}}$ and define $\mathcal{M}_q = \ker \sigma$.

Locally, this leads to the following picture. Let $(\mathbf{x}, \mathbf{p})$ be local coordinates on $J_q\mathcal{E}$ in a neighbourhood of ρ. We must first determine $T_\rho\mathcal{R}_q$ as a subspace of $T_\rho J_q\mathcal{E}$. Let $(\mathbf{x}, \mathbf{p}; \dot{\mathbf{x}}, \dot{\mathbf{p}})$ be the induced coordinates on $T_\rho J_q\mathcal{E}$; then every vector $X \in T_\rho J_q\mathcal{E}$ can be written

$$X = \dot{x}^i \partial_{x^i} + \dot{p}^\alpha_\mu \partial_{p^\alpha_\mu} \; . \tag{1.48}$$

Assuming that a local representation of the submanifold $\mathcal{R}_q$ is given by the equations $\Phi^\tau(\mathbf{x}, \mathbf{p}) = 0$ with $\tau = 1, \dots, p$, its tangent space $T_\rho\mathcal{R}_q$ consists of all vectors $X \in T_\rho J_q\mathcal{E}$ such that $d\Phi^\tau(X) = 0$ (see Remark C.3.5) which yields the linear system

$$\sum_{i=1}^{n} \frac{\partial \Phi^\tau}{\partial x^i}(\mathbf{x}, \mathbf{p})\, \dot{x}^i + \sum_{\substack{1\leq\alpha\leq m \\ 0\leq|\mu|\leq q}} \frac{\partial \Phi^\tau}{\partial p^\alpha_\mu}(\mathbf{x}, \mathbf{p})\, \dot{p}^\alpha_\mu = 0 \; . \tag{1.49}$$

The symbol is by definition the vertical part of this tangent space. This means that we are only interested in those solutions of (1.49) where $\dot{\mathbf{x}} = 0$ and $\dot{p}^\alpha_\mu = 0$ whenever $|\mu| < q$. Thus $\mathcal{M}_q$ can be locally described as the solution space of the following system of linear equations

$$(\mathcal{M}_q)_\rho : \left\{ \sum_{\substack{1\leq\alpha\leq m \\ |\mu|=q}} \frac{\partial \Phi^\tau}{\partial p^\alpha_\mu}(\mathbf{x}, \mathbf{p})\, \dot{p}^\alpha_\mu = 0 \; , \quad \tau = 1, \dots, p \; . \right. \tag{1.50}$$

This is a system with real coefficients, as the derivatives $\partial\Phi^\tau/\partial p^\alpha_\mu$ are evaluated at the point $\rho \in \mathcal{R}_q$. If there are n independent and m dependent variables, the coefficient matrix has p rows (labelled by τ) and $m\binom{n+q-1}{n-1}$ columns (labelled by α and μ). We call it *symbol matrix* denoted by $M_q(\rho)$. It is also the matrix of the symbol map σ in local coordinates.

Obviously, $\dim(\mathcal{M}_q)_\rho$ (or equivalently $\operatorname{rank} M_q(\rho)$) might vary with ρ. For this reason, we speak in Definition 1.5.1 only of a family of vector spaces and not of a vector bundle. Only if the dimension remains constant over $\mathcal{R}_q$, the symbol $\mathcal{M}_q$ is a vector subbundle of $S_q(T^*\mathcal{X}) \underset{J_{q-1}\mathcal{E}}{\otimes} V\mathcal{E}$. In the sequel we will always assume that this is the case if not explicitly stated otherwise. In particular, $\mathcal{M}_q$ is a vector bundle, if $\mathcal{R}_q$ is globally defined as zero set of a fibred map $\Phi : J_q\mathcal{E} \to \mathcal{E}'$.

The symbol is most easily understood for linear differential equations. Loosely speaking, the geometric symbol is then simply the highest order or principal part of the system. For non-linear systems we perform a brute force linearisation at the point ρ in order to obtain $(\mathcal{M}_q)_\rho$.

Remark 1.5.2. A particular class of differential equations are those where the dimension of the symbol is zero. These equations are called *maximally overdetermined* or of *finite type*. The latter name stems from the fact that, as we will see later, the formal solution space of such an equation is always finite-dimensional. $\triangleleft$

If we are interested in the symbol $\mathcal{M}_{q+r} \subseteq S_{q+r}(T^*\mathcal{X}) \underset{J_{q+r-1}\mathcal{E}}{\otimes} V\mathcal{E}$ of the prolonged equation $\mathcal{R}_{q+r} \subseteq J_{q+r}\mathcal{E}$, there is no need to explicitly compute a local representation of $\mathcal{R}_{q+r}$ in order to obtain local equations for its symbol. We can directly derive them from a local representation of $\mathcal{R}_q$. This follows easily from the coordinate expression (1.11) of the formal derivative.

By (1.50) and (1.32), we need only the partial derivatives $\partial D_i\Phi^\tau/\partial p^\alpha_\nu$ with $1 \leq i \leq n$ and $|\nu| = q+1$, i.e. the highest-order part of the formal derivative $D_i\Phi^\tau$, for determining the symbol $\mathcal{M}_{q+1}$. According to (1.11) it is given by $\partial D_i\Phi^\tau/\partial p^\alpha_\nu = \partial\Phi^\tau/\partial p^\alpha_{\nu-1_i}$ (if $\nu_i = 0$, the derivative vanishes). Thus a local representation of the first prolongation $\mathcal{M}_{q+1}$ is

$$(\mathcal{M}_{q+1})_\rho : \left\{ \sum_{\substack{1\leq\alpha\leq m\\ 0\leq|\mu|\leq q}} \frac{\partial\Phi^\tau}{\partial p^\alpha_\mu}(\mathbf{x},\mathbf{p})\,\dot{p}^\alpha_{\mu+1_i} = 0\,, \quad \begin{array}{l}\tau = 1,\dots,p\,,\\ i = 1,\dots,n\,.\end{array} \right. \tag{1.51}$$

At first sight one might be tempted to think that the linear systems (1.50) and (1.51) have the same coefficient matrices. This is of course not true. The matrix M_{q+1} of (1.51) has np rows and $m\binom{n+q}{n-1}$ columns and is thus much larger. Because of the particular way in which the index i appears in (1.51), one may say that the matrix M_{q+1} emerges from M_q by a kind of combinatorial blowing up. For each row in M_q we obtain n rows in M_{q+1} and similarly for the columns. These n rows contain exactly the same entries but distributed in different patterns.

Example 1.5.3. We consider a scalar second-order equation in two independent variables x, y with the local representation

$$\mathcal{R}_2 : \{ \Phi(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0 \; . \tag{1.52}$$

If we define the three functions $a(\rho) = \partial\Phi/\partial u_{xx}(\rho)$, $b(\rho) = \partial\Phi/\partial u_{xy}(\rho)$ and $c(\rho) = \partial\Phi/\partial u_{yy}(\rho)$, each $\mathcal{R}_2 \to \mathbb{R}$, then the matrix of the symbol $\mathcal{M}_2$ at a point $\rho \in \mathcal{R}_2$ contains only one row

$$M_2(\rho) = \begin{pmatrix} a(\rho) & b(\rho) & c(\rho) \end{pmatrix} \; . \tag{1.53}$$

If no point $\rho \in \mathcal{R}_2$ exists where all three functions a, b, c vanish simultaneously, $\mathcal{M}_2$ is a two-dimensional vector bundle.

For the prolonged symbol $\mathcal{M}_3$ at a point $\rho' \in \mathcal{R}_3$ such that $\pi^3_2(\rho') = \rho$ we obtain the matrix

$$M_3(\rho') = \begin{pmatrix} a(\rho) & b(\rho) & c(\rho) & 0 \\ 0 & a(\rho) & b(\rho) & c(\rho) \end{pmatrix} \; . \tag{1.54}$$

Here the columns correspond to the unknowns $\dot{u}_{xxx}$, $\dot{u}_{xxy}$, $\dot{u}_{xyy}$ and $\dot{u}_{yyy}$. If $\mathcal{M}_2$ is a vector bundle, then $\mathcal{M}_3$ is too. $\triangleleft$

The considerations above are straightforwardly extended by iteration to higher prolongations and we obtain as local representation

$$(\mathcal{M}_{q+r})_\rho : \begin{cases} \displaystyle\sum_{\substack{1\leq\alpha\leq m \\ 0\leq|\mu|\leq q}} \frac{\partial\Phi^\tau}{\partial p^\alpha_\mu}(\mathbf{x}, \mathbf{p})\, \dot{p}^\alpha_{\mu+\nu} = 0 \; , & \tau = 1, \dots, p \; , \\ & |\nu| = r \; . \end{cases} \tag{1.55}$$

The size of the corresponding symbol matrices M_{q+r} is growing rapidly. Because of (1.24), one easily calculates that M_{q+r} has $p\binom{n+r-1}{n-1}$ rows and $m\binom{n+q+r-1}{n-1}$ columns.

The matrix of the symbol and in particular of its prolongations indirectly already appeared in Section 1.4 during the construction of a formal power series solution (1.39). The rank of the symbol matrix M_q tells us how many coefficients of order q are determined by (1.40) (nothing is said about the coefficients of lower order!). The matrices M_{q+r} of the prolonged symbols appear, if we interpret (1.42) as an inhomogeneous linear system for the coefficients of order $q + r$: its matrix is precisely M_{q+r}.

In Section 1.4 we saw the this simple algorithm for the construction of formal power series solutions works only, if we are dealing with a formally integrable equation. The symbol aids us in deciding whether or not a given equation is formally integrable. It indicates whether during the prolongation of a differential equation integrability conditions may have occurred and it allows us to construct them using only linear algebra. The proof of the next theorem shows this explicitly.

Theorem 1.5.4. *If the symbol $\mathcal{M}_{q+1}$ is a vector bundle, then*

$$\dim \mathcal{R}_q^{(1)} = \dim \mathcal{R}_{q+1} - \dim \mathcal{M}_{q+1}\,. \tag{1.56}$$

Proof. The dimension of a submanifold is determined by counting how many independent equations are needed to describe it locally, i. e. by the rank of a Jacobian. We assume that $\mathcal{R}_q$ has the local representation $\Phi^\tau(\mathbf{x}, \mathbf{p}) = 0$. The part of the Jacobian of $\mathcal{R}_{q+1}$ relevant for the determination of the fibre dimension can be divided into four blocks:

$$\left(\begin{array}{|c|c|}\hline \dfrac{\partial D_i\Phi^\tau}{\partial p^\alpha_\mu}\,,\ |\mu| = q+1 & \dfrac{\partial D_i\Phi^\tau}{\partial p^\alpha_\mu}\,,\ 0 \le |\mu| \le q \\ \hline 0 & \dfrac{\partial \Phi^\tau}{\partial p^\alpha_\mu}\,,\ 0 \le |\mu| \le q \\ \hline \end{array}\right)\,. \tag{1.57}$$

The lower part stems from the local representation of $\mathcal{R}_q$; its rank is $\operatorname{codim}\mathcal{R}_q$. The upper part comes from the prolonged equation; the left block is the symbol matrix M_{q+1}. We must distinguish two cases. If M_{q+1} has maximal rank, then $\dim\mathcal{R}_{q+1} = \dim\mathcal{R}_q + \dim\mathcal{M}_{q+1}$ and thus $\mathcal{R}_q^{(1)} = \mathcal{R}_q$.

If M_{q+1} has not maximal rank, one can construct by linear operations a row with only zeros in the left block. The question is how this row looks like in the right block. If it is independent of the rows in the lower part, we have found an integrability condition. It is constructed by performing exactly the same linear operations on the full equations $D_i\Phi^\tau = 0$. Otherwise, we have only obtained a redundant description of $\mathcal{R}_{q+1}$. This leads to *identities* or *compatibility conditions*, if we add a right hand side to the differential equation. Comparing ranks yields the claim for one point, but since $\mathcal{M}_{q+1}$ is a vector bundle, it holds everywhere. □

In the classical theory of partial differential equations a different notion of symbol is used which should not be confused with the geometric symbol introduced above: the classical symbol is *not* an intrinsic object. Our notion of symbol is closely related to what is traditionally called the principal symbol which is an intrinsically defined object.

Assume we are given a one-form $\chi \in T^*\mathcal{X}$. It induces for every $q > 0$ a map $\iota_{\chi,q} : V\mathcal{E} \to S_q(T^*\mathcal{X}) \otimes V\mathcal{E}$ defined by $\iota_{\chi,q}(\nu) = \chi^q \otimes \nu$ where χ^q denotes the q-fold symmetric product of χ. In local coordinates, we write $\chi = \chi_i dx^i$ and obtain $\iota_{\chi,q} : \nu^\alpha \partial_{u^\alpha} \mapsto \chi_\mu \nu^\alpha dx^\mu \otimes \partial_{u^\alpha}$ where μ runs over all multi indices of length q and $\chi_\mu = \chi_1^{\mu_1} \cdots \chi_n^{\mu_n}$. Now we may define the principal symbol in terms of the symbol map σ.

Definition 1.5.5. *Let $\chi \in T^*\mathcal{X}$ be a one-form over $\mathcal{X}$ and $\mathcal{R}_q \subseteq J_q\mathcal{E}$ a differential equation globally described by the map $\Phi : J_q\mathcal{E} \to \mathcal{E}'$. The* principal symbol *of $\mathcal{R}_q$ is the linear map $\tau_\chi : V\mathcal{E} \to T\mathcal{E}'$ defined by $\tau_\chi = \sigma \circ \iota_{\chi,q}$.*

The one-form χ plays the role of a parameter in the principal symbol. For many purposes, it is of great importance how the properties of the linear map τ_χ change with χ. Locally, we can associate a matrix $T[\chi]$ with τ_χ:

$$T^\tau_\alpha[\chi] = \sum_{|\mu|=q} \frac{\partial \Phi^\tau}{\partial p^\alpha_\mu} \chi^\mu \ . \tag{1.58}$$

If $\dim \mathcal{E} = m$ and $\dim \mathcal{E}' = p$, it has p rows and m columns. Its entries are homogeneous polynomials of degree q in the coefficients of χ. We may think of $T[\chi]$ as a kind of contraction of the symbol matrix M_q of $\mathcal{R}_q$. Both matrices have the same number of rows. The column with index α of $T[\chi]$ is a linear combination of all columns in M_q corresponding to a variable $\dot{p}^\alpha_\mu$ with the coefficients given by χ_μ.

Example 1.5.6. Consider again the equation $\mathcal{R}_2$ described by (1.52). Its symbol matrix is the 1×3 matrix shown in (1.53). The matrix of its principal symbol is 1×1 with the single entry $a(\rho)\chi_x^2 + b(\rho)\chi_x\chi_y + c(\rho)\chi_y^2$. $\lhd$

Definition 1.5.7. *Let $\mathcal{R}_q$ be an involutive differential equation. We call a non-zero one-form $\chi \in T^*\mathcal{X}$* characteristic *at a point $\rho \in \mathcal{R}_q$, if the principal symbol τ_χ is not injective at ρ. The set of all characteristic one-forms at ρ is the* characteristic variety *C_ρ.*

C_ρ is indeed a variety in the strict algebraic meaning of the word. The condition that τ_χ is not injective at a point ρ can be expressed as the vanishing of certain subdeterminants of the matrix $T[\chi]$, i. e. as the vanishing of (homogeneous) polynomials in the coefficients χ_i. For *semi-linear* equations, where the symbol space $(\mathcal{M}_q)_\rho$ depends only on $x = \pi^q(\rho) \in \mathcal{X}$, one can define characteristic one-forms over points $x \in \mathcal{X}$.

Example 1.5.8. In introductory textbooks on differential equations one usually does not consider characteristic one-forms but *characteristic surfaces*. We show now for linear first-order systems the relation between the two approaches. Consider

$$\mathcal{R}_1 : \left\{ \sum_{\alpha=1}^{m} \sum_{i=1}^{n} A_\alpha^{\tau i}(\mathbf{x}) p_i^\alpha + \sum_{\alpha=1}^{m} B_\alpha^\tau(\mathbf{x}) + C^\tau(x) = 0 \;, \qquad 1 \le \tau \le p \right. \tag{1.59}$$

and the $(n-1)$-dimensional hypersurface $\Sigma \subset \mathcal{X}$ described by the equation

$$\Sigma = \left\{ x \in \mathcal{X} \mid x^n = \phi(x^1, \dots, x^{n-1}) \right\} \,. \tag{1.60}$$

Such a surface is called characteristic, if prescribing the value of $\mathbf{u}$ on Σ does not suffice to compute all first-order derivatives on Σ. Obviously, this condition concerns only the derivatives $(\partial \mathbf{u}/\partial x^n)\big|_\Sigma$; all other ones follow easily from our "initial values".

Thus assume that $\mathbf{u}\big|_\Sigma = f(x^1, \dots, x^{n-1})$. Using the chain rule we obtain

$$\frac{\partial \mathbf{u}}{\partial x^i}\bigg|_\Sigma + \frac{\partial \phi}{\partial x^i} \frac{\partial \mathbf{u}}{\partial x^n}\bigg|_\Sigma = \frac{\partial f}{\partial x^i} \,, \qquad 1 \le i < n \,. \tag{1.61}$$

Substituting these relations into (1.59) and restricting to Σ yields a linear system for the derivatives $(\partial \mathbf{u}/\partial x^n)\big|_\Sigma$ whose matrix T is

$$T_\alpha^\tau = A_\alpha^{\tau n} - \sum_{i=1}^{n-1} A_\alpha^{\tau i} \frac{\partial \phi}{\partial x^i} \,. \tag{1.62}$$

If $\operatorname{rank} T < m$, i. e. the associated linear map is not injective, then it is *not* possible to determine all derivatives on the hypersurface Σ and it is characteristic. The matrix of the principal symbol of (1.59) is $T_\alpha^\tau[\chi] = A_\alpha^{\tau i} \chi_i$ which coincides with (1.62) for the one-form $\chi = dx^n - (\partial \phi/\partial x^i) dx^i$. Hence the hypersurface Σ is characteristic, if and only if this one-form is characteristic. More abstractly, we may say that the conormal space $N^*\Sigma$ (see Appendix C.3) is spanned by the one-form χ. ◁

1.6 Some Examples

In this section we briefly discuss some classical differential equations which will be used as examples at many places later on. They stem from physics and are of great importance in applications. Most of them are already formally integrable, however, we have not yet the necessary tools to prove this.

Example 1.6.1. Maxwell's equations form the foundations of electrodynamics. They describe the evolution of the electric and magnetic fields and fluxes. Their most general (differential) form is

$$\mathbf{D}_t = \nabla \times \mathbf{H} - \mathbf{J}_e\,, \qquad \mathbf{B}_t = -\nabla \times \mathbf{E} - \mathbf{J}_m\,, \tag{1.63a}$$
$$0 = \nabla \cdot \mathbf{D} - \rho_e\,, \qquad 0 = \nabla \cdot \mathbf{B} - \rho_m\,. \tag{1.63b}$$

Here ∇ denotes the gradient with respect to the spatial variables only; $\cdot$ and $\times$ represent the scalar and the vector product, respectively. $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic field, respectively, $\mathbf{D}$ and $\mathbf{B}$ the corresponding fluxes, $\mathbf{J}_e$ and $\mathbf{J}_m$ are currents and finally, ρ_e and ρ_m charge densities. In physics, $\mathbf{J}_m$ and ρ_m are usually set to zero, as nobody has ever observed experimentally magnetic charges or currents. But there is no mathematical reason for not adding such terms (at the level of (1.64); in gauge theory the situation changes).

(1.63) is yet underdetermined. In a simple medium, we have the constitutive relations $\mathbf{D} = \epsilon\mathbf{E}$, $\mathbf{B} = \mu\mathbf{H}$ and $\mathbf{J}_e = \sigma\mathbf{E}$ where ϵ, μ and σ are the permittivity, the permeability and the conductivity, respectively. If the medium is isotropic and homogeneous, these quantities are constant scalars. In general, they are tensor valued functions of space-time. We mainly study Maxwell's equations in vacuum. There $\sigma = 0$ and $\epsilon\mu = 1/c$ where c is the speed of light. In natural units $c = 1$ and we obtain the following system:

$$\mathbf{E}_t = \nabla \times \mathbf{B}\,, \qquad \mathbf{B}_t = -\nabla \times \mathbf{E}\,, \tag{1.64a}$$
$$0 = \nabla \cdot \mathbf{E}\,, \qquad 0 = \nabla \cdot \mathbf{B}\,. \tag{1.64b}$$

The equations in (1.64a) describe the time evolution of the fields. The equations in (1.64b) are often called the *Gauß* or *divergence laws*; they render (1.64) overdetermined, as we have got eight equations for six field components. Because of the well-known identity $\nabla \cdot (\nabla \times \mathbf{X}) = 0$ for any vector field $\mathbf{X}$, we obtain no integrability conditions, if we first prolong and then project back (1.64). Note that the same is true for the general form (1.63), if we ignore the currents and charges.

This identity also implies another interesting property. Let us take only the evolutionary part (1.64a); it forms a hyperbolic (in t-direction) system in Cauchy–Kovalevskaya form. So we may consider the initial value problem for it. Assume that we prescribe for $t = 0$ fields $\mathbf{E}_0(\mathbf{x})$ and $\mathbf{B}_0(\mathbf{x})$ satisfying the divergence laws (1.64b). Then the solution $\mathbf{E}(\mathbf{x},t)$, $\mathbf{B}(\mathbf{x},t)$ of (1.64a) for these initial data satisfy (1.64b) for all times t, i.e. it is a solution of the full Maxwell equations (1.64). In Section 5.4 we will see that this is no accident but in fact a characteristic and very important property of involutive equations.

If we include the currents $\mathbf{J}_e$ and $\mathbf{J}_m$ and the charge densities ρ_e and ρ_m, we see that they must satisfy the *continuity equations*

$$(\rho_e)_t + \boldsymbol{\nabla} \cdot \mathbf{J}_e = 0 \,, \qquad (\rho_m)_t + \boldsymbol{\nabla} \cdot \mathbf{J}_m = 0 \,. \tag{1.65}$$

Otherwise the equations become inconsistent and do not admit any solutions. These are examples of *compatibility conditions.* ◁

Example 1.6.2. The fundamental equations governing the flow of fluids and gas are the *Navier–Stokes equations*. We are particularly interested in *incompressible flows.* The flow lives in some domain $\Omega \subseteq \mathbb{R}^3$; for simplicity we assume it to be open. Our base manifold $\mathcal{X}$ is then given by the product $\Omega \times \mathbb{R}$ where the factor $\mathbb{R}$ represents time. The dependent variables are the velocity of the flow and the pressure. Thus we may use as total space $\mathcal{E} = T\Omega \times \mathbb{R}^2$ with local coordinates $(\mathbf{x}, t, \mathbf{u}, p)$: $\mathbf{x}$ represent a point in Ω, t the time, $\mathbf{u}$ are coordinates of a tangent vector in $T_{\mathbf{x}}\Omega$ and p is the pressure.

The incompressible Navier–Stokes equations are

$$\mathbf{u}_t + (\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{u} = \nu \triangle \mathbf{u} - \boldsymbol{\nabla} p \,, \tag{1.66a}$$

$$\boldsymbol{\nabla} \cdot \mathbf{u} = 0 \,. \tag{1.66b}$$

Here we use again notations from vector analysis: $\boldsymbol{\nabla}$ is the gradient with respect to the spatial variables $\mathbf{x}$ and the scalar product $\boldsymbol{\nabla} \cdot \mathbf{u}$ corresponds to the divergence of $\mathbf{u}$. The left hand side of (1.66a) represents the convective derivative of $\mathbf{u}$ expressing that our coordinate system moves with the flow; in components it is given by $\partial u^i/\partial t + u^j \partial u^i/\partial x^j$.

The coefficient ν is the *viscosity* of the flow. If it is zero, we obtain *Euler's equations* of inviscid flow. They can be described in a much more geometric way as a flow on the group of volume preserving diffeomorphisms [9]. With respect to the formal analysis, the Euler and the Navier–Stokes equations behave very similar.

(1.66) describes a submanifold $\mathcal{R}_2 \subset J_2\mathcal{E}$. As the incompressibility constraint (1.66b) is of first order, $\mathcal{R}_2$ is *not* formally integrable; trivial integrability conditions are given by the prolongations of (1.66b) to second order. Thus $\mathcal{R}_2^{(1)}$ is described by (1.66) plus the equations

$$\boldsymbol{\nabla} \cdot \mathbf{u}_t = 0 \,, \qquad \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{u}) = 0 \,. \tag{1.67}$$

$\mathcal{R}_2^{(1)}$ is still not formally integrable, as we find in the next prolongation a non-trivial integrability condition: if we take the divergence of (1.66a), we can eliminate the third-order term $\triangle(\boldsymbol{\nabla} \cdot \mathbf{u})$ by the divergence of the second equation in (1.67). There remains the second-order equation

$$\triangle p = -\boldsymbol{\nabla} \cdot \big((\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{u}\big) \,, \tag{1.68}$$

which may be interpreted as a Poisson equation for the pressure. It plays an important role in the numerical integration of the Navier–Stokes equations. One can check that the equation $\mathcal{R}_2^{(2)}$ defined by (1.66), (1.67), and (1.68) is formally integrable. ◁

The following two examples are actually of a very geometric nature. Nevertheless, we describe them only in local coordinates. There are two reasons for this. First of all, we do not want to introduce all the necessary differential geometry. Secondly, we will take at many places a more computationally point of view. In order to be able to demonstrate many of our computational techniques with the help of these examples, we must write them down in local coordinates.

Example 1.6.3. In modern elementary particle physics three of the four fundamental interactions are described by *Yang–Mills theories*. The proper framework for them are principal fibre bundles, but we take here a simplified approach. Underlying the theory is the space-time, an n-dimensional manifold equipped with a Lorentzian metric. Again for simplicity we restrict here to the case of a flat space-time with the Minkowski metric η^{ij} defined by $\eta^{ii} = 1$ for $1 \leq i < n$, $\eta^{nn} = -1$ and $\eta^{ij} = 0$ in all other cases.

Assume we are given a d-dimensional Lie group $\mathcal{G}$. The basic field of the corresponding Yang–Mills theory, the *vector potential*, is a one-form A over space-time taking values in the Lie algebra $\mathfrak{g}$ of $\mathcal{G}$. If a basis of $\mathfrak{g}$ is given by the elements T_a with $1 \leq a \leq d$, such a one-form has the local form $A = A_i^a T_a dx^i$. Thus we are dealing with a fibre bundle over space-time whose fibre dimension is dn.

Let the structure constants of the Lie algebra $\mathfrak{g}$ be $C^c_{\ ab}$, i. e. we have $[T_a, T_b] = C^c_{\ ab} T_c$. Then the exterior derivative of A yields the *field strength* two-form $F = dA = F^a_{ij} T_a dx^i \wedge dx^j$ where the coefficients are given by

$$F^a_{ij} = \partial_{x^i} A^a_j - \partial_{x^j} A^a_i + C^a_{bc} A^b_i A^c_j \ . \tag{1.69}$$

They allow us to write the field equations in a concise form:

$$\Phi^a_\ell \left(A^b_i, \partial_{x^j} A^b_i, \partial_{x^j x^k} A^b_i\right) = \eta^{ij} \left(\partial_{x^i} F^a_{j\ell} + C^a_{bc} A^b_i F^c_{j\ell}\right) = 0 \ . \tag{1.70}$$

This is a semi-linear system of dn second-order equations. The non-linearity is introduced by the third term in the definition of the field strengths which is quadratic in the vector potentials. It disappears, if the Lie algebra $\mathfrak{g}$ is Abelian, as in this case all structure constants are zero.

In a rather lengthy computation one can show that the partial differential equation $\mathcal{R}_2$ defined locally by (1.70) is formally integrable. The key is the *Noether identity*

$$\eta^{ij}\left(D_i\Phi^a_j + C^a_{bc}A^b_i\Phi^c_j\right) \equiv 0 \tag{1.71}$$

stemming from the properties of the structure constants C^a_{bc} due to the Jacobi identity. It implies that $\mathcal{R}^{(1)}_2 = \mathcal{R}_2$ and, as we will see later, this suffices to prove formal integrability.

The Yang–Mills equations for the Lie group $U(1)$ are closely related to Maxwell's equations. As we are dealing with a one-dimensional group, we skip the upper index on the coefficients of the potential A and the field strength F. The coefficients of latter form thus in four dimensions a 4×4 antisymmetric matrix related to the electric and magnetic field:

$$F = \begin{pmatrix} 0 & B^{(z)} & -B^{(y)} & -E^{(x)} \\ -B^{(z)} & 0 & B^{(x)} & -E^{(y)} \\ B^{(y)} & -B^{(x)} & 0 & -E^{(z)} \\ E^{(x)} & E^{(y)} & E^{(z)} & 0 \end{pmatrix} . \tag{1.72}$$

In other words, if we build the three-dimensional vector $\mathbf{A}$ consisting of the first three coefficients of A, then $\mathbf{B} = \nabla \times \mathbf{A}$ and $\mathbf{E} = -\partial_t\mathbf{A} - \nabla \cdot A_4$. Note that these relations automatically imply half of the equations in (1.64), namely $\nabla \cdot \mathbf{B} = 0$ and $\mathbf{B}_t = -\nabla \times \mathbf{E}$. The remaining four equations are equivalent to the four Yang–Mills equations (1.70). $\triangleleft$

Example 1.6.4. General relativity describes the fourth fundamental interaction, gravity; a cornerstone of the theory are *Einstein's equations.* We consider only the vacuum case without any matter fields but in arbitrary dimensions.[2] Space-time is modelled as an n-dimensional Lorentzian manifold, i. e. as a manifold $\mathcal{M}$ equipped with a metric g_{ij} that has the same signature as the Minkowski metric η_{ij}. This manifold $\mathcal{M}$ is our base space $\mathcal{X}$. The dependent variables are the components g_{ij} of the metric. Thus $\mathcal{E}$ is the vector bundle of symmetric bilinear forms; its fibre dimension is $m = n(n+1)/2$.

The vacuum Einstein equations express the condition that the Ricci tensor R_{ij} of the Levi–Civita connection for the metric g_{ij} vanishes:

$$R_{ij}[g_{k\ell}] = 0 . \tag{1.73}$$

As the Ricci tensor is symmetric, this is a system of $n(n+1)/2$ equations so that we have as many equations as unknown functions. (1.73) defines a submanifold $\mathcal{R}_2 \subset J_2\mathcal{E}$, since the Ricci tensor depends on the second-order derivatives of the metric.

Writing out the components R_{ij} of the Ricci tensor in terms of the metric and its derivatives leads to rather complicated non-linear expressions which

[2] Strictly speaking, we assume that the dimension is at least four; for topological reasons two and three dimensions are special cases where different equations have to be used.

can be found in any textbook on either Riemannian geometry or general relativity. We omit them here and use instead a little trick for simplifying the coordinate form of (1.73): we take *locally geodesic coordinates*. Such coordinates exist for every point $x_0 \in \mathcal{M}$ and are characterised by two properties: (i) the metric becomes the Minkowski metric at the point x_0, $g_{ij}(x_0) = \eta_{ij}$, and (ii) the Christoffel symbols vanish at x_0. At x_0, (1.73) takes in these coordinates the following simple linear form

$$\begin{aligned} R_{ij}(x_0) = \frac{1}{2}\eta^{k\ell}\Big[\partial_{x^i x^k} g_{j\ell}(x_0) + \partial_{x^j x^\ell} g_{ik}(x_0) - {}& \\ \partial_{x^i x^j} g_{k\ell}(x_0) - \partial_{x^k x^\ell} g_{ij}(x_0)\Big] = 0\,. \end{aligned} \tag{1.74}$$

Note that the defining properties of a locally geodesic coordinate system hold only *at the single point* x_0 and generally not even in some neighbourhood of it. Thus (1.74) should not be considered as a differential equation but only as a relation between the fibre coordinates of $J_2\mathcal{E}$ in the fibre over the point $x_0 \in \mathcal{M}$. In particular, we cannot expect that local representations of prolongations $\mathcal{R}_{2+r} \subset J_{2+r}\mathcal{E}$ are obtained by formally differentiating (1.74). However, this holds for the first prolongation of $\mathcal{R}_2$, as $\partial_{x^k} g_{ij}(x_0) = 0$ because of the metricity of the Levi–Civita connection. In further prolongations, lower-order terms appear. Fortunately, it turns out that $\mathcal{R}_3$ suffices for the formal analysis of the Einstein equations.

Because of the *contracted Bianchi identity* (a formal analogue to the Noether identity (1.71) of the Yang–Mills equations)

$$g^{ki}\partial_{x^k} R_{ij} - \partial_{x^j}(g^{ki}R_{ki}) = \Gamma^k_{ij} g^{i\ell} R_{k\ell} - \Gamma^\ell_{\ell k} g^{ik} R_{ij} \tag{1.75}$$

(here g^{ki} is the inverse to g_{ij}, i.e. $g^{ki}g_{ij} = \delta^k_j$), we find that $\mathcal{R}_2^{(1)} = \mathcal{R}_2$. We will see later that this suffices to prove the formal integrability of Einstein's equations. $\triangleleft$

1.7 Notes

Implicitly, jet bundles have been used for a long time. They were formally introduced by Ehresmann in a series of papers [44, 45] (see also [43]). They have become a standard tool for the geometric analysis of differential equations. By now, many approaches to jet bundles exist with various degrees of abstraction and sophistication. A standard reference is the book by Saunders [155]; useful introductions are also contained in [61, 129].

The pedestrian approach based on power series we presented in Section 1.1 is completely sufficient for all computational purposes. It can also be described in a coordinate free manner. This requires only an intrinsic formulation of the equivalence relation $\sim_q$. We say $\phi \sim_1 \psi$ at $x_0 \in \mathcal{X}$, if the tangent maps satisfy $T_{x_0}\phi = T_{x_0}\psi$. Then we extend recursively: $\phi \sim_q \psi$ at $x_0 \in \mathcal{X}$, if $T_{x_0}\phi \sim_{q-1} T_{x_0}\psi$ at all points $v \in T_{x_0}\mathcal{X}$. This formulation is detailed in [61].

Another approach is discussed in [132, Ch. IV]. It requires that $\mathcal{U}$ is a real vector space, which is the case in most applications. We define for any point $x_0 \in \mathcal{X}$ the ideal $\mathcal{I}_{x_0}(\mathcal{X})$ of all smooth functions $\phi : \mathcal{X} \to \mathbb{R}$ with $\phi(x_0) = 0$. Then we introduce the spaces

$$\mathcal{Z}^{(q)}_{x_0}(\mathcal{X}, \mathcal{U}) = [\mathcal{I}_{x_0}(\mathcal{X})]^q \cdot C^\infty(\mathcal{X}, \mathcal{U}) . \tag{1.76}$$

All functions in $\mathcal{Z}^{(q)}_{x_0}(\mathcal{X}, \mathcal{U})$ possess at x_0 an at least q-fold zero and thus all their derivatives up to order $q-1$ vanish at this point. Finally, we identify q-jets at x_0 with the elements of the quotient space $C^\infty(\mathcal{X}, \mathcal{U})/\mathcal{Z}^{(q+1)}_{x_0}(\mathcal{X}, \mathcal{U})$.

The intrinsic approach to (first-order) jet bundle presented in Section 1.2 does not seem to be very popular. It is rarely found in the literature, although it has a number of advantages for studying the geometry of jet bundles. In recent times, it has gained some importance in the context of multisymplectic geometry, as the affine dual of $J_1\mathcal{E}$ carries a natural multisymplectic structure [62]. Our treatment of the contact structure is taken from Modugno [125].

We have seen that there are some analogies between our geometric view of differential equations and the situation in algebraic geometry. However, there are also some apparent differences. In particular, not every section contained in an equivalence class $[\sigma]^{(q)}_x \in \mathcal{R}_q$ is a solution of the differential equation. If one wants such a one-to-one relation between points and solutions, one must take *all* prolongations of $\mathcal{R}_q$ into account. This leads to the introduction of the jet bundle $J_\infty\mathcal{E}$ of infinite order a point of which defines a unique formal power series.

The proper definition of $J_\infty\mathcal{E}$ requires some mathematical care, as it is an infinite-dimensional space and it is not obvious how to equip it with a topological and a differentiable structure. This is achieved using the notions of projective and inductive limits in directed systems of sets [17]. As the jet bundles of different orders are related by the natural projections π^r_q, we may define

$$J_\infty\mathcal{E} = \operatorname*{proj\,lim}_{q\to\infty} J_q\mathcal{E} . \tag{1.77}$$

In order to introduce locally analytic functions on $J_\infty\mathcal{E}$ we exploit the natural inclusions $C^\omega(J_q\mathcal{E}) \subset C^\omega(J_r\mathcal{E})$ for $r > q$ and set

$$C^\omega(J_\infty\mathcal{E}) = \operatorname*{ind\,lim}_{q\to\infty} C^\omega(J_q\mathcal{E}) . \tag{1.78}$$

Note that the definition of an inductive limit implies that any function in $C^\omega(J_\infty\mathcal{E})$ is an element of at least one space $C^\omega(J_q\mathcal{E})$. Thus although these functions are defined on an infinite-dimensional manifold, they always depend only on a finite number of variables. In a similar, though slightly more complicated manner, one constructs the tangent and cotangent bundle of $J_\infty\mathcal{E}$ [17].

Then we introduce the infinite prolongation $\mathcal{R}_\infty \subseteq J_\infty\mathcal{E}$ of a differential equation. As each point on $\mathcal{R}_\infty$ represents a formal power series solution of the equation, the submanifold $\mathcal{R}_\infty$ may be identified with the formal solution space. Thus at the formal level we are in the same situation as in algebraic geometry. This point of view is much used by the school of Vinogradov (see e. g. [5, 38, 96, 181]) who calls such a submanifold $\mathcal{R}_\infty \subseteq J_\infty\mathcal{E}$ a *differential variety* or short *diffiety*. For our purposes we do not need this concept.

An interesting aspect of $J_\infty\mathcal{E}$ is that the contact codistribution $\mathcal{C}^0_\infty$ is involutive. In a jet bundle $J_q\mathcal{E}$ of finite order q, this can never be the case, as one can easily see from the basic contact forms ω^α_μ defined in (1.8). Indeed, if we take the exterior derivative of a form ω^α_μ, we obtain $d\omega^\alpha_\mu = -dp^\alpha_{\mu+1_i} \wedge dx^i$. If $|\mu| < q-1$, we may rewrite this as $d\omega^\alpha_\mu = -\omega^\alpha_{\mu+1_i} \wedge dx^i$ and the derivative vanishes modulo $\mathcal{C}^0_q$. But if $|\mu|$ takes the maximal value $q-1$, this is not possible and the exterior derivative does not vanish. Hence $\mathcal{C}^0_q$ is not involutive for any finite order q. In $\mathcal{C}^0_\infty$ no maximal value for $|\mu|$ exists and thus it is involutive.

Similarly, the contact distribution $\mathcal{C}_\infty$ is involutive. It consists of the n vector fields C_i given by (1.10a). Note that they are now defined by a series. Their application to a function yields its formal derivatives.