Chapter 2

Formal Theory

We give a general introduction into the ideas and the machinery of the formal theory of partial differential equations. We will define involutive equations and show how to construct power series solutions for them. We will study the properties of involutive symbols and show how to complete a given equation to an involutive one. Introducing tableaux we will be able to define characteristics, over- and underdeterminacy and tackle δ-regularity.

2.1 Geometric Theory

The formal theory is based on a differential geometric approach to differential equations using the jet bundle formalism. An exact introduction to this framework can be found in [117, 130]; here we will only give an informal description mainly in local coordinates. It should, however, be noted that all constructions can be performed in an intrinsic, coordinate-free manner.

The basic idea of jet bundles is to consider the derivatives of the dependent variables with respect to the independent ones as additional, *algebraically* independent variables. This gives naturally rise of a bundle structure. The starting point is a bundle $\pi : \mathcal{E} \to X$ with the independent variables $x^i,\ i = 1, \dots, n$, as coordinates of the base space and the dependent variables $u^\alpha,\ \alpha = 1, \dots, m$, as fiber coordinates.

Over this bundle we construct further bundles: The first order jet bundle $J_1\mathcal{E}$ has X as base and as fiber coordinates the dependent variables plus the first order derivatives $p_i^\alpha = \partial u^\alpha / \partial x^i$. Correspondingly, as fiber coordinates of the q-th order jet bundle $J_q\mathcal{E}$ all derivatives up to order q are occurring. We can also use a lower order jet bundle as base space instead of X; the canonical projection is then denoted by $\pi_q^{q+r} : J_{q+r}\mathcal{E} \to J_q\mathcal{E}$. A simple combinatorial calculation yields the (fiber) dimension of $J_q\mathcal{E}$:

$$\dim J_q\mathcal{E} = m\binom{n+q}{q}. \tag{2.1}$$

We will mainly use the standard multi-index notation of multivariate calculus. A multi-index is an n-tuple $\mu = [\mu_1, \dots, \mu_n]$ and its length is defined by $|\mu| = \mu_1 + \cdots + \mu_n$. Then the jet variable p_μ^α represents the derivative $\partial^{|\mu|} u^\alpha / \partial(x^1)^{\mu_1} \cdots \partial(x^n)^{\mu_n}$. Multi-indices can be added and subtracted like vectors provided all entries remain non-negative. As shorthand we write $\mu \pm a_i$ for the multi-index $[\mu_1, \dots, \mu_i \pm a, \dots, \mu_n]$. In concrete examples, we will often use the notation $p^\alpha_{ij\dots}$ for a derivative with respect to $x^i, x^j, \dots$ This should not lead to confusion, as multi-indices will always be denoted by greek letters and the indices of the independent variables by Latin letters.

Sections of the bundle $\mathcal{E}$ can be considered as natural generalization of functions (or better: of their graphs). They are defined as mappings $\sigma : X \to \mathcal{E}$ satisfying $\pi \circ \sigma = \mathrm{id}_X$; i.e. in local coordinates σ has the form $\sigma : x \mapsto (x, f(x))$. The sheaf of all local sections of $\mathcal{E}$ will be denoted by $\Gamma_{loc}(\mathcal{E})$; local meaning that a section σ needs to be defined only in a subset $U_\sigma \subset X$. We can *prolong* a section to a section of the q-th order jet bundle. Locally this is simply done by adding the derivatives of f: $j_q(\sigma) : x \mapsto \left(x, f(x), \partial^{|\mu|} f(x)/\partial^\mu x\right)$. This explains also a different point of view of jets. There a point $(x^i, u^\alpha, p^\alpha_\mu)$ in $J_q\mathcal{E}$ is considered as an equivalence class of functions with the same power series expansion up to order q around the point x.

Definition 2.1 *A (non-linear)* differential equation *of order q is a fibered submanifold $\mathcal{R}_q$ of $J_q\mathcal{E}$. A* solution *is a section $\sigma \in \Gamma_{loc}(\mathcal{E})$ such that its prolongation $j_q(\sigma)$ lies entirely in $\mathcal{R}_q$.*

Fibered means that $\mathcal{R}_q$ can be projected onto the base space X. In other words, the differential equation poses no restrictions on the independent variables.This definition, which does not distinguish between a single equation and a system of equations, is motivated by two main ideas: The differential equation is "reduced" to an algebraic equation in $J_q\mathcal{E}$ and we are now treating a geometric object — namely a submanifold — which is independent of a particular set of equations or a particular coordinate system. If $\mathcal{E}$ is a vector bundle, then linear equations possess in addition the structure of a vector subbundle.

Usually we will be given a map with local representation

$$\Phi : \left\{ \begin{array}{ccc} J_q\mathcal{E} & \longrightarrow & \mathcal{E}' \\ (x^i, u^\alpha, p^\alpha_\mu) & \longmapsto & \Phi^\tau(x^i, u^\alpha, p^\alpha_\mu) \end{array} \right. \tag{2.2}$$

where $\mathcal{E}'$ is another bundle over X such that $\mathcal{R}_q$ is the kernel of this map. For topological reason not every differential equation in the sense of Definition 2.1 has such a local representation [117]. Since, however, all equations treated in the sequel are of this form, we will always assume the existence of such a mapping Φ and furthermore that this description is non-degenerate, i.e. the functions Φ^τ are functionally independent. Typically $\mathcal{E}$ will be of the form $\mathbb{R}^{m+n}$ and $\mathcal{E}'$ of the form $\mathbb{R}^p$ and one can use the functions Φ^τ globally.

There are two basic geometric operations that can be performed with a differential equation: We can *prolong* it into a higher-order jet bundle or we can *project* it into a lower-order one. Prolongation corresponds of course to differentiation of the equations. In the jet bundle formalism the *formal derivative* serves this purpose:

$$D_i\Phi^\tau(x^i, u^\alpha, p^\alpha_\mu) = \frac{\partial \Phi^\tau}{\partial x^i} + \sum_\alpha \frac{\partial \Phi^\tau}{\partial u^\alpha} p^\alpha_i + \sum_{\alpha,\mu|} \frac{\partial \Phi^\tau}{\partial p^\alpha_\mu} p^\alpha_{\mu+1_i} \,. \tag{2.3}$$

This definition reflects the interpretation of the jet variables p^α_μ as derivatives of the dependent variables u^α and respects the chain rule.

The prolongation $\mathcal{R}_{q+1} \subset J_{q+1}\mathcal{E}$ of the differential equation $\mathcal{R}_q$ is defined by the combined set of equations $\Phi^\tau = 0,\ D_i\Phi^\tau = 0$ for $i = 1, \dots, n$. An intrinsic definition is given by $\mathcal{R}_{q+r} = J_r\mathcal{R}_q \cap J_{q+r}\mathcal{E}$. The intersection means essentially that we identify derivatives of the derivatives in $J_r\mathcal{R}_q$ with higher-order derivatives of the original dependent variables u^α. This is necessary, as they are otherwise treated as additional dependent variables.

In the sequel, we will assume throughout that $\mathcal{R}_q$ is a *regular equation*, i.e. all prolongations define again fibered submanifolds. This does not hold automatically, as the following example will show, but can always be achieved by using an appropriate restriction of $\mathcal{R}_q$.

Example 2.2

$$\mathcal{R}_2 : \begin{cases} u_{yy} - \frac{1}{2}(u_{xx})^2 &= 0\,, \\ u_{xy} - u_{xx} &= 0\,. \end{cases} \tag{2.4}$$

Adding the x- and y-derivative of the second equation and subtracting the x-derivative of the first one yields the equation $(u_{xx} - 1)u_{xxx} = 0$. Thus $\mathcal{R}_3$ is not a fibered submanifold, as its dimension changes wherever $u_{xx} = 1$. We must consider two distinct cases, in which the differential equation can show quite different behavior.

At first sight, it may surprise that the projection $\mathcal{R}_q^{(1)} = \pi_q^{q+1}(\mathcal{R}_{q+1}) \subset J_q\mathcal{E}$ of the prolongation of a differential equation does not always recover the original equation $\mathcal{R}_q$ but sometimes only a subset. This signals the existence of integrability conditions. In terms of the local representation $\Phi^\tau = 0$ they arise mostly as linear combinations of prolonged equations (in the simplest case as cross-derivatives) and represent additional equations which are automatically satisfied by any solution of the original differential equation but which are nevertheless algebraically independent of it. More generally, we introduce $\mathcal{R}_{q+r}^{(s)} = \pi_{q+r}^{q+r+s}(\mathcal{R}_{q+r+s}) \subset J_{q+r}\mathcal{E}$.

Example 2.3 A trivial example may explain the phenomenon. Consider the first-order system

$$\mathcal{R}_1 : \begin{cases} u_z + yu_x &= 0\,, \\ u_y &= 0\,. \end{cases} \tag{2.5}$$

Its prolongation $\mathcal{R}_2$ contains the equation $u_{yz} + yu_{xy} + u_x = 0$ (from the first equation). It follows, however, from the second equation that both second-order derivatives vanish. Hence, we get the integrability condition $u_x = 0$, which must be added to $\mathcal{R}_1$ to get $\mathcal{R}_1^{(1)}$. It is important to note that it is not possible to construct this equation without differentiation.

Whereas this corresponds exactly to what is usually called an integrability condition, in the geometric theory there exists a second reason why the projection may not be surjective, namely if the local representation of $\mathcal{R}_q$ contains equations of different order. The prolongation of these yields equations of order less than or equal to q, thus they have to be taken into account when determining $\mathcal{R}_q^{(1)}$.

Example 2.4 Consider the simple second order equation

$$\mathcal{R}_2 : \begin{cases} u_{xx} &= 0\,, \\ u_y &= 0\,. \end{cases} \tag{2.6}$$

$\mathcal{R}_3$ contains also the equations $u_{xy} = u_{yy} = 0$. Thus the projection is not surjective and we have to add these two equations to $\mathcal{R}_2$ to obtain $\mathcal{R}_2^{(1)}$.

The occurrence of integrability conditions thoroughly disturbs the construction of power series solutions order by order. To build such a solution for an equation $\mathcal{R}_q$ one starts with the coefficients up to order q. Some of them can be freely chosen; the other ones can be expressed by these using the differential equation. Analogously one proceeds with the coefficients of the next order; this time applying the prolonged equation $\mathcal{R}_{q+1}$. (We will detail this construction in Section 2.5). If there occur, however, some integrability conditions during the prolongations, we suddenly get additional equations for the coefficients of *lower order*. Thus the construction of a power series solution order by order can be performed only for an equation containing all its integrability conditions.

Definition 2.5 *A differential equation $\mathcal{R}_q$ is* formally integrable, *if $\mathcal{R}^{(1)}_{q+r} = \mathcal{R}_{q+r}$ for all $r \geq 0$, i.e. if it contains all its integrability conditions.*

2.2 Involutive Equations

Formally integrable equations are often called involutive. This is, however, not correct. Involution comprises formal integrability but requires additional properties which are of a more algebraic nature. To give an exact definition we have to introduce the symbol of a differential equation which is closely connected with the construction of power series solutions.

Definition 2.6 *Let the differential equation $\mathcal{R}_q$ be locally described by $\Phi^\tau(x^i, u^\alpha, p^\alpha_\mu) = 0$ for $\tau = 1, \dots, p$. The solution space $\mathcal{M}_q$ of the following system of linear equations in the unknowns v^α_μ with $\alpha = 1, \dots, m,\ |\mu| = q$*

$$\mathcal{M}_q\ :\ \left\{ \sum_{\alpha, |\mu|=q} \left(\frac{\partial \Phi^\tau}{\partial p^\alpha_\mu} \right) v^\alpha_\mu = 0\,, \qquad \tau = 1, \dots, p \right. \tag{2.7}$$

is called the symbol *of $\mathcal{R}_q$. (By abuse of language, we will often also call the linear system symbol instead of its solution space).*

Differential geometrically seen the symbol is a family of vector spaces over $\mathcal{R}_q$. It can be defined intrinsically as $\mathcal{M}_q = V\mathcal{R}_q \cap (S_q T^*X \otimes V\mathcal{E})$ where S_q denotes the q-fold symmetric product, T^*X the cotangent bundle of X and V the verticle bundle[1]. Alternatively we can use the local representation (2.2) of $\mathcal{R}_q$ and introduce the symbol map

$$\sigma\ :\ \left\{ \begin{array}{ccc} S_q T^*X \otimes V\mathcal{E} & \longrightarrow & \mathcal{E}' \\ v^\alpha_\mu & \longmapsto & \displaystyle\sum_{\alpha, |\mu|=q} \left(\frac{\partial \Phi^\tau}{\partial p^\alpha_\mu} \right) v^\alpha_\mu \end{array} \right. . \tag{2.8}$$

It is easy to see that this map is intrinsically defined, i.e. independent of the choice of the local representation Φ, and that its kernel is just $\mathcal{M}_q$.

(2.7) represents at each point of $\mathcal{R}_q$ a linear system in a vector space of the dimension

$$\dim S_q T^*X \otimes V\mathcal{E} = m \binom{n+q-1}{n-1}\,. \tag{2.9}$$

This is of course nothing else than the number of derivatives p^α_μ of order q. We will make the assumption that $\mathcal{M}_q$ is a vector bundle over $\mathcal{R}_q$, i.e. the dimensions of the vector spaces are constant. Again this can always be achieved by taking an appropriate restriction.

The formal derivative applied to a function over $J_q\mathcal{E}$ yields a function on $J_{q+1}\mathcal{E}$ which is *linear* in the derivatives of order $q+1$. This allows to compute directly without further differentiations the equations of the prolonged symbols $\mathcal{M}_{q+r}$. Applying Definition 2.6 to $\mathcal{R}_{q+1}$ yields the equations

$$\mathcal{M}_{q+1}\ :\ \left\{ \sum_{\alpha, |\nu|=q+1} \left(\frac{\partial D_i \Phi^\tau}{\partial p^\alpha_\nu} \right) v^\alpha_\nu = 0\,, \qquad \begin{array}{rcl} \tau & = & 1, \dots, p\,, \\ i & = & 1, \dots, n\,. \end{array} \right. \tag{2.10}$$

[1]Here we use the fact that $J_q\mathcal{E}$ is an affine bundle over $J_{q-1}\mathcal{E}$ modeled on the vector bundle $S_q T^*X \otimes V\mathcal{E}$. This follows easily from the transformation and symmetry properties of q-th order derivatives [117]. One can thus consider $S_q T^*X \otimes V\mathcal{E}$ as the "vector space of all q-th order derivatives".

But by the definition (2.3) of the formal derivative, $\partial D_i\Phi^\tau/\partial p^\alpha_\nu$ can be different from zero only, if there exist multi-indices μ of order q with $\nu = \mu + 1_i$ for some i such that $\partial\Phi^\tau/\partial p^\alpha_\mu \neq 0$. The same argument holds of course for higher prolongations, too. Thus we arrive at the simple formula

$$\mathcal{M}_{q+r} \,:\, \left\{ \sum_{\alpha,|\mu|=q} \left(\frac{\partial\Phi^\tau}{\partial p^\alpha_\mu}\right) v^\alpha_{\mu+\nu} = 0\,, \qquad \begin{array}{rl} \tau &= 1,\dots,p\,, \\ |\nu| &= r\,. \end{array} \right. \tag{2.11}$$

The notation sometimes gives the impression that the matrices of the linear systems $\mathcal{M}_q$ and $\mathcal{M}_{q+r}$ are identical. This is of course not true; the latter one is much larger. It contains more rows and columns, as it is defined by more equations in a larger vector space. But we can see that all rows with the same index τ contain the same entries, because the coefficients are independent of ν, and these are basically the same entries as in the row τ of the first symbol $\mathcal{M}_q$. The only thing that happens is that they appear several times and that they are distributed in a different way.

The symbol indicates whether or not during a prolongation integrability conditions may have occurred and allows us to construct them using only linear algebra. To show this explicitly, we outline a proof of the next theorem.

Theorem 2.7 *If $\mathcal{M}_{q+1}$ is a vector bundle, then*

$$\dim \mathcal{R}^{(1)}_q = \dim \mathcal{R}_{q+1} - \dim \mathcal{M}_{q+1}\,. \tag{2.12}$$

Proof. We can compute the dimension of a submanifold by counting how many independent functions are needed to describe it locally, i.e. by the rank of a Jacobi matrix. The spaces occurring in (2.12) are defined by the following equations:

$$\mathcal{R}_q \,:\, \left\{ \; \Phi^\tau(x^i, u^\alpha, p^\alpha_\mu) = 0\,, \right. \tag{2.13}$$

$$\mathcal{R}_{q+1} \,:\, \left\{ \begin{array}{rl} \Phi^\tau(x^i, u^\alpha, p^\alpha_\mu) &= 0\,, \\ D_i\Phi^\tau(x^i, u^\alpha, p^\alpha_\mu) &= 0\,, \end{array} \right. \tag{2.14}$$

$$\mathcal{M}_{q+1} \,:\, \left\{ \sum_{\alpha,|\mu|=q} \left(\frac{\partial\Phi^\tau}{\partial p^\alpha_\mu}\right) v^\alpha_{\mu+1_i} = 0\,. \right. \tag{2.15}$$

The Jacobi matrix of $\mathcal{R}_{q+1}$ can be divided into blocks

$$\left(\begin{array}{|c|c|c|} \hline \dfrac{\partial D_i\Phi^\tau}{\partial p^\alpha_\mu}\,,\ |\mu| = q+1 & \dfrac{\partial D_i\Phi^\tau}{\partial p^\alpha_\mu}\,,\ |\mu| \leq q & \dfrac{\partial D_i\Phi^\tau}{\partial u^\alpha} \\ \hline 0 & \dfrac{\partial \Phi^\tau}{\partial p^\alpha_\mu}\,,\ |\mu| \leq q & \dfrac{\partial \Phi^\tau}{\partial u^\alpha} \\ \hline \end{array}\right)\,. \tag{2.16}$$

The lower part stems from the original equations. Hence its rank is $\operatorname{codim}\mathcal{R}_q$. The upper part comes from the prolonged equations. We consider first the left block. Comparison with (2.10) shows that it is just the symbol $\mathcal{M}_{q+1}$. We must now distinguish two cases. If this symbol has maximal rank, then $\dim\mathcal{R}_{q+1} = \dim\mathcal{R}_q + \dim\mathcal{M}_{q+1}$ and thus $\mathcal{R}^{(1)}_q = \mathcal{R}_q$.

If, however, $\mathcal{M}_{q+1}$ has not maximal rank, then it is possible to construct by linear operations a row with only zeros in the first block. The question is now, how does this row look like in the two blocks to the right. If it is independent of the rows in the lower part, then 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 will lead either 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. •

Example 2.3 (cont.) We can demonstrate this construction for our previous example. The first prolongation of this equation is given by

$$\mathcal{R}_2 : \left\{ \begin{array}{ll} u_z + yu_x = 0\,, & u_y = 0\,, \\ u_{xz} + yu_{xx} = 0\,, & u_{xy} = 0\,, \\ u_{yz} + yu_{xy} + u_x = 0\,, & u_{yy} = 0\,, \\ u_{zz} + yu_{xz} = 0\,, & u_{yz} = 0\,. \end{array} \right. \tag{2.17}$$

Obviously these equations are independent. With $\dim J_2\mathcal{E} = 10$ we get $\dim \mathcal{R}_2 = 10 - 8 = 2$. The equations for the symbol are

$$\mathcal{M}_2 : \left\{ \begin{array}{ll} v_{xz} + yv_{xx} = 0\,, & v_{xy} = 0\,, \\ v_{yz} + yv_{xy} = 0\,, & v_{yy} = 0\,, \\ v_{zz} + yv_{xz} = 0\,, & v_{yz} = 0\,. \end{array} \right. \tag{2.18}$$

The third equation is linearly dependent of the second and sixth. With $\dim S_2T^*X \otimes V\mathcal{E} = 6$ we find $\dim \mathcal{M}_2 = 6 - 5 = 1$. Applying Theorem 2.7 yields $\dim \mathcal{R}_1^{(1)} = 2 - 1 = 1$. The dimension of $\mathcal{R}_1$ is however 2. Hence an integrability condition has occurred and we can construct it by subtracting the eighth and y times the fourth equation from the fifth in (2.17) getting $u_x = 0$.

To proceed further, we must introduce an apparently coordinate dependent notion: the *class* of a multi-index. We will study in Section 2.7 under what conditions the results are nevertheless intrinsic. The multi-index $\mu = [\mu_1, \mu_2, \ldots, \mu_n]$ is said to be of class k, if the first non-vanishing entry is μ_k. We can order the columns of the symbol by the class of the multi-index of the corresponding v_μ^α (the ordering within a class is of no importance) and then compute a row echelon form. In this form, the class of each row is given by the class of its first non-vanishing entry, the pivot. We associate with a row of class k the *multiplicative variables* $x^1, x^2, \ldots, x^k$. The symbol is now in *solved form*. We denote by $\beta_q^{(k)}$ the number of rows of class k [2].

The jet variables corresponding to the pivots are called *principal derivatives*, all others are *parametric derivatives*. The meaning of this distinction will become clearer in Section 2.5 where we construct power series solutions. Principal derivatives are obtained using the differential equations; the parametric ones must be prescribed by initial or boundary conditions. The idea is now to find a unique way to compute all derivatives of the principal derivatives. This leads to the notion of an *involutive* symbol.

An intrinsic definition of an involutive symbol can be obtained from the Spencer cohomology [58, 117]. Since this is fairly abstract, we will mention here only a simple criterion for an involutive symbol using the above introduced multiplicative variables. It can also be easily applied in concrete computations.

[2] Note that this number could take different values at different points of $\mathcal{R}_q$. But like we always assume that $\mathcal{M}_q$ is a vector bundle, we will always assume that the values of $\beta_q^{(k)}$ are constant on $\mathcal{R}_q$.

Definition 2.8 *The symbol $\mathcal{M}_q$ is called* involutive, *if*

$$\operatorname{rank} \mathcal{M}_{q+1} = \sum_{k=1}^{n} k\beta_q^{(k)} . \tag{2.19}$$

The differential equation $\mathcal{R}_q$ is called involutive *or* in involution, *if it is formally integrable and if its symbol $\mathcal{M}_q$ is involutive.*

2.3 Involutive Symbols and Multiplicative Variables

The definition given at the end of the last section looks at first sight rather strange. It has, however, a straightforward explanation, which is so important that we will give it in form of a proposition for easier reference.

Proposition 2.9 *If the symbol $\mathcal{M}_q$ of the differential equation $\mathcal{R}_q$ is involutive, then we obtain (in a given local representation) all independent equations of order $q+1$ of the prolongation $\mathcal{R}_{q+1}$ by differentiating each equation of $\mathcal{R}_q$ with respect to its multiplicative variables only.*

Proof. Obviously we need only linear operations to transform a symbol into the above described solved form. Thus we can assume without loss of generality, that our local representation of $\mathcal{R}_q$ is such that we obtain the symbol directly in solved form by applying (2.7). Now it is easy to see that all equations obtained by prolonging each equation with respect to its multiplicative variables only are independent, as they all have distinct pivots. Since there are $\beta_q^{(k)}$ equations of class k in $\mathcal{R}_q$, this yields $\sum k\beta_q^{(k)}$ independent equations of order $q+1$ in the prolongation $\mathcal{R}_{q+1}$. Thus we know a lower bound for the number of such equations. By assumption $\operatorname{rank} \mathcal{M}_{q+1}$ equals this number. Hence there cannot be more independent equations of order $q+1$. Note, however, that there may very well exist further independent equations in $\mathcal{R}_{q+1}$, but they must be of lower order, i.e. they yield integrability conditions (cf. the proof of Theorem 2.7). •

This approach to involution via analyzing pivots is inspired by the Janet-Riquier Theory (cf. Appendix B). In Chapter 3 we will see that involution of a symbol means that we can calculate the ranks of its prolongations $\mathcal{M}_{q+r}$ directly from the numbers $\beta_q^{(k)}$ without having to compute explicitly their equations. This will allow statements about the arbitrariness of the general solution.

Example 2.10 A simple example of a differential equation with a non-involutive symbol is given by

$$\mathcal{R}_2 : \begin{cases} u_{xx} = 0 , \\ u_{yy} = 0 . \end{cases} \tag{2.20}$$

It does not matter, whether we choose x or y as x^2, in each case we have one equation of class 2 and one of class 1, thus 3 multiplicative variables. However, the rank of $\mathcal{M}_3$ is 4, as the prolongation contains 4 independent equations. Thus $\mathcal{M}_2$ is not involutive but $\mathcal{M}_3$. It has 5 multiplicative variables and the rank of $\mathcal{M}_4$ is indeed 5. As (2.10) generates obviously no integrability conditions, we have here an example of a formally integrable equation which is not involutive. More generally, one can show (cf. Section 5.4) that the equation

$$\mathcal{R}_q : \begin{cases} u_{[0,q]} = 0 , \\ u_{[q,0]} = 0 \end{cases} \tag{2.21}$$

is formally integrable for any q but becomes involutive only after $q-1$ prolongations.

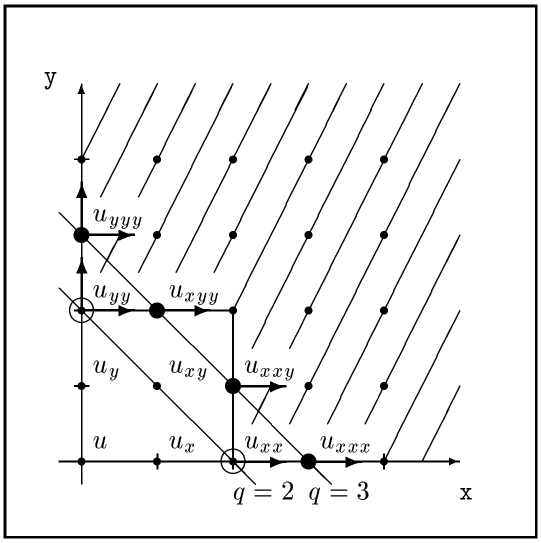

Figure 2.1: Reid diagram for (2.20).

Reid [124] has developed a nice graphical representation of principal and parametric derivatives which can also be used to explain the meaning of involution. Fig. 2.1 shows the Reid diagram for (2.20).

The axes represent the independent variables, the dots the different derivatives of u. The circles around u_{xx} and u_{yy} indicate that they are pivots in the symbol. All derivatives in the dashed area can be obtained from them by differentiating the equations, i.e. they are principal derivatives. All other derivatives are parametric ones. The full circles for the third-order derivatives depict the pivots of the prolonged symbol $\mathcal{M}_3$. The arrows indicate the multiplicative variables taking $x^1 = x$, $x^2 = y$.

We see that it is not possible to get all derivatives in the shaded area, if we prolong $\mathcal{M}_2$ only with respect to its multiplicative variables. We miss u_{xxy} . $\mathcal{M}_3$ is involutive and indeed every principal derivative of order 4 or higher can be obtained by differentiating one of its equations with respect to its multiplicative variables.

In order to study further properties of involutive symbols, we need a systematic way to order derivatives. In the definition of the solved form of a symbol we used the class, but this provides only a partial ordering. We introduce therefore the concept of a ranking.

Definition 2.11 *A* ranking *of derivatives is a total ordering defined on the space of all derivatives p^α_μ for fixed values of m, n which satisfies two conditions:*

(i) If $|\mu| < |\nu|$, then $p^\alpha_\mu < p^\alpha_\nu$.

(ii) If $p^\alpha_{\mu_1} < p^\alpha_{\mu_2}$, then $p^\alpha_{\mu_1+\nu} < p^\alpha_{\mu_2+\nu}$ for any multi-index ν.

If the first property holds not only for derivatives of the same dependent variable u^α but for all, then the ranking is called a *total-degree* or *orderly* ranking. In such a ranking derivatives are always ordered first by their order and then by other criteria. A total-degree ranking *respects classes*, if $p^\alpha_\mu < p^\beta_\nu$ for $\operatorname{class} p^\alpha_\mu < \operatorname{class} p^\beta_\nu$.

A ranking of derivatives induces of course also a ranking of the unknowns v^α_μ used as place holders in the symbol. From now on, we will always assume that the columns of the symbol are ordered using a ranking respecting the classes, i.e. if one derivative has a higher class than another one, then it is also higher in the ranking.

An important example for this provides the *inverse lexicographic ranking*. It is a total-degree ranking. If two derivatives have the same order, then we look at the first differing entry in the multi-index. The one with the higher entry is taken as higher in the ranking (hence inverse). Thus it respects classes. Of two derivatives with the same multi-index we take the one as higher which belongs to the dependent variable with the higher index.

Using rankings one can make useful statements about the pivots in the solved form of involutive symbols [117]. They allow one to construct to given values of m, n, q all possible sets of pivots with a given number of equations. Furthermore we will derive several very important properties of involutive symbols (and equations) as corollaries of the following rather technical

Proposition 2.12 *Let $\mathcal{M}_q$ be an involutive symbol in solved form, where the columns have been ordered according to a class-respecting ranking.*

(i) If v^α_μ is a pivot of class i, then $v^\alpha_{\mu-1_i+1_j}$ is also a pivot for all $j > i$.

(ii) If there are pivots v^α_μ of class i, then the entries μ_i of their multi-indices take all values between 1 and a maximal value.

Proof. We start with Part (i). Since v^α_μ is a pivot of class i and $j > i$, x^j is a non-multiplicative variable for the corresponding equation. We prolong it with respect to x^j. Because of the second property of a ranking, $v^\alpha_{\mu+1_j}$ is the pivot of the prolonged equation. For an involutive symbol one obtains, however, all independent equations of the prolonged symbol by prolonging with respect to the multiplicative variables only (Proposition 2.9). Since $v^\alpha_{\mu+1_j}$ is still of class i, it must be possible to get it by prolonging a pivot v^α_ν of higher or equal class with respect to x^i. Thus $\nu + 1_i = \mu + 1_j$ or $\nu = \mu - 1_i + 1_j$.

Part (ii) follows easily from Part (i). Let v^α_μ be the pivot of class i with minimal value of μ_i. But according to (i) $v^\alpha_{\mu-1_i+1_j}$ is also a pivot for $j > i$ and for $\mu_i > 1$ it is also of class i. This contradicts the assumption that μ_i was minimal. Thus $\mu_i = 1$ is the minimal value and one sees at once that with $\mu_i = a$ also $\mu_i = a - 1$ occurs. •

Corollary 2.13 *Let $\mathcal{M}_1$ be the involutive symbol of a first-order equation. Then*

$$0 \le \beta_1^{(1)} \le \beta_1^{(2)} \le \cdots \le \beta_1^{(n)} \le m\,. \tag{2.22}$$

Proof. This is a direct consequence of Part (i) of Proposition 2.12. Let v^α_μ be a pivot of class k, i.e. $\mu_i = \delta_{ik}$. Then v^α_ν with $\nu = \mu - 1_k + 1_{k+1}$, i.e. $\nu_i = \delta_{i,k+1}$, is a pivot of class $k+1$ occurring in the symbol $\mathcal{M}_1$. Thus $\beta_1^{(k+1)} \ge \beta_1^{(k)}$. $\beta_q^{(n)}$ can never be greater than m for any value of q, as there exist only m different derivatives of class n. •

Theorem 2.14 *Let $\mathcal{M}_q$ be the involutive symbol of the differential equation $\mathcal{R}_q$. Then*

(i) $\mathcal{M}_{q+1}$ is involutive, too.

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

Proof. We start with Part (i); Part (ii) will then be a corollary. Without loss of generality we can assume that $\mathcal{M}_q$ is in solved form with the columns ordered according to a class-respecting ranking. Let v^α_μ be the pivots of $\mathcal{M}_q$. Then by Proposition 2.9 the pivots of $\mathcal{M}_{q+1}$ are $v^\alpha_{\mu+1_\ell}$ where $\ell = 1, \dots, \operatorname{class}\mu$. The class of such a pivot is ℓ.

To prove Part (i), we must show that prolonging an equation in $\mathcal{M}_{q+1}$ with respect to a non-multiplicative variable leads to a linearly dependent equation. Thus let $k = \operatorname{class}\mu$ and let $\ell \le k$. Prolonging the equations of $\mathcal{M}_{q+1}$ with respect to all independent variables leads to pivots of the form $v^\alpha_{\mu+1_\ell+1_j}$ with $j = 1, \dots, n$. Three cases must be distinguished:

$j \leq \ell$: These are the prolongations with respect to multiplicative variables. Obviously all these equations are linearly independent.

$\ell < j \leq k$: Since $j \leq k$ x^j is a multiplicative variable for the equation in $\mathcal{M}_q$ with the pivot v^α_μ. Thus we can first prolong this equation with respect to x^j and then the result with respect to x^ℓ. Thus each prolongation is only with respect to multiplicative variables and the resulting equation is already contained in the equations of the previous case.

$j > k$: Let us denote the equation by $\phi = 0$. As in the previous case we interchange the prolongations, thus prolong first with respect to x^j and then with respect to x^ℓ to obtain ϕ. Since $j > k$, $v^\alpha_{\mu-1_k+1_j}$ is by Proposition 2.12 also a pivot in $\mathcal{M}_q$. Its class is greater than or equal k; thus x^k is a multiplicative variable for the corresponding equation. By prolonging it with respect to x^k we see that $v^\alpha_{\mu+1_j}$ is among the pivots of $\mathcal{M}_{q+1}$. Since x^ℓ is multiplicative for this equation, we have found an equation contained in the first case with pivot $v^\alpha_{\mu+1_\ell+1_j}$.

If we compare the two equations with pivot $v^\alpha_{\mu+1_\ell+1_j}$, then we see that they are both prolongations with respect to x^ℓ of equations with the pivot $v^\alpha_{\mu+1_j}$. One of these was obtained by prolonging an equation of $\mathcal{M}_q$ with respect to a non-multiplicative variable. Since $\mathcal{M}_q$ is involutive, it is not contained in the generating set of independent equations we use for $\mathcal{M}_{q+1}$ but is a linear combination of those. Thus ϕ can be written as x^ℓ-prolongation of this linear combination. If x^ℓ is a non-multiplicative variable for some of the equations involved, then we can apply the same analysis again. Since the considered pivots will become smaller and smaller, this process will eventually terminate and then we will have expressed ϕ as a linear combination of equations contained in the first case.

Thus we have shown that the only independent equations in $\mathcal{M}_{q+2}$ can be obtained by prolonging the equations of $\mathcal{M}_{q+1}$ with respect to their multiplicative variables only. But this means nothing else than that $\mathcal{M}_{q+1}$ is involutive.

Part (ii) may look trivial at first sight, but it is not. At the left hand side only the integrability conditions which arise during the projection from order $q+1$ to order q are taken into account; then one prolongs again to order $q+1$. On the right hand side, however, we project from order $q+2$ to order $q+1$. Hence we have to show that all integrability conditions which occur during this projection are prolongations of integrability conditions of the projection of order $q+1$ to q.

In order to achieve this we make basically the same considerations as in the proof of Part (i) but now with the full equations instead of only with their highest-order part. Let Φ^τ be an equation of class k in a local description of $\mathcal{R}_q$. We have to analyses all equations of the form $D_\ell D_j \Phi^\tau$. We distinguish two cases assuming always $\ell \leq j$:

$j \leq k$: In this case $D_\ell D_j \Phi^\tau$ contains only prolongations with respect to multiplicative variables. As we have seen in Proposition 2.9, these are the only independent equations of order $q+2$ in our local description of $\mathcal{R}_{q+2}$.

$k < j$: This implies that $D_j \Phi^\tau$ was a prolongation with respect to a non-multiplicative variable. By Proposition 2.9 we can express this equation as a linear combination of prolongations of equations of $\mathcal{R}_q$ with respect to multiplicative variables and perhaps some integrability conditions. Following again the argument of the third case above, we find that $D_\ell D_j \Phi^\tau$ is a linear combination of the equations contained in the first case and of prolonged integrability conditions.

Note that it is not necessary here to distinguish between $\ell \leq k$ and $\ell > k$, as we are allowed to use all prolongations to construct $(\mathcal{R}_q^{(1)})_{+1}$. •

Finally we prove a generalization of an old result by Finzi [53, 111] for systems with the same number of equations and dependent variables. It emphasizes the distinguished position of normal equations (cf. Section 2.6).

Proposition 2.15 *A differential equation $\mathcal{R}_q$ has either identities or integrability conditions, if and only if $\beta_q^{(n-1)} > 0$.*

PROOF. If $\beta_q^{(n-1)} > 0$, there exist equations with non-multiplicative variables, as the only equations without are those of class n. Thus according to Proposition 2.9 the prolongation of these equations with respect to x^n will lead either to identities or to integrability conditions. For the converse, we note that if $\beta_q^{(n-1)} = 0$ all $\beta_q^{(k)}$ with $k < n-1$ vanish, too. This can be shown similarly to Corollary 2.13. Thus $\mathcal{R}_q$ contains only equations of class n and all prolonged equations are independent. •

2.4 Cartan-Kuranishi Theorem

Theorem 2.14 has a very important application. It yields an easy criterion for an involutive equation. Recall that Definition 2.5 of a formally integrable equation requires to check infinitely many conditions, namely that at no order of prolongation integrability conditions arise. If we have, however, an involutive symbol, this can be reduced to just one prolongation.

Theorem 2.16 (Criterion of Involution) *$\mathcal{R}_q$ is an involutive equation, if and only if its symbol $\mathcal{M}_q$ is involutive and $\mathcal{R}_q^{(1)} = \mathcal{R}_q$.*

PROOF. Since $\mathcal{M}_q$ is involutive, $\mathcal{R}_{q+1}^{(1)} = (\mathcal{R}_q^{(1)})_{+1} = \mathcal{R}_{q+1}$ according to Theorem 2.14 and our assumptions. But all prolonged symbols $\mathcal{M}_{q+r}$ are also involutive and we can apply this argument at any order. Thus $\mathcal{R}_{q+r}^{(1)} = \mathcal{R}_{q+r}$ for all $r \geq 1$ and $\mathcal{R}_q$ is formally integrable. •

Based on this theorem we will now show that any differential equation $\mathcal{R}_q$ can be completed to an equivalent equation, i.e. one with the same solution space, which is involutive. Thus in the sequel we can always assume that we deal with involutive equations. We will later see that this has many advantages, as involution entails many special properties.

Theorem 2.17 (Cartan-Kuranishi) [3] *For every differential equation $\mathcal{R}_q$ there exist two integers r, s such that $\mathcal{R}_{q+r}^{(s)}$ is an involutive equation with the same solution space.*

PROOF. We refer for a detailed proof of the theorem to [117]. We here outline a proof by giving a concrete algorithm for the completion and showing its correctness and termination. It consists of two nested loops. The inner one prolongs the equation, until its symbol is involutive. The outer one performs the projections; i.e. it computes the arising integrability conditions. In a more formal language, we get the algorithm shown in Fig. 2.2.

To prove the correctness and termination of this algorithm, we need the following key theorem of the formal theory of partial differential equations which we cite without proof.

[3] The name *Cartan-Kuranishi Theorem* stems from a similar theorem in Cartan-Kähler Theory [92].

Algorithm `Completion` $(\mathcal{R}_q)$

Input: Equation $\mathcal{R}_q$
Output: Equation $\mathcal{R}^{(s)}_{q+r}$ in involution

[1] r ← 0; s ← 0
[2] `until` $\mathcal{R}^{(s)}_{q+r}$ involutive `repeat`
[2.1] compute $\mathcal{M}^{(s)}_{q+r}$ {*extract symbol*}
[2.2] `while not(`$\mathcal{M}^{(s)}_{q+r}$ involutive`) repeat`
[2.2.1] r ← r+1
[2.2.2] compute $\mathcal{M}^{(s)}_{q+r}$ {*prolong symbol*}
[2.3] compute $\mathcal{R}^{(s)}_{q+r}$ {*prolong equation*}
[2.4] `if` $\mathcal{R}^{(s+1)}_{q+r} \subset \mathcal{R}^{(s)}_{q+r}$ `then`
[2.4.1] s ← s+1
[2.4.2] compute $\mathcal{R}^{(s)}_{q+r}$ {*project equation*}
[3] `return` $\mathcal{R}^{(s)}_{q+r}$

Figure 2.2: Algorithm for the Cartan-Kuranishi Theorem (sketchy version).

Theorem 2.18 (Prolongation Theorem) *Let $\mathcal{M}_q$ be the symbol of an equation $\mathcal{R}_q$. Then there exists an integer $\hat{q}(n,m,q) \geq q$ such that $\mathcal{M}_{\hat{q}}$ is involutive.*

The formal proof of this theorem is highly nontrivial and was first given by Sweeny [150]. It makes decisive use of the Spencer cohomology and also yields a bound for $\hat{q}(n,m,q)$. In our algorithm, it guarantees the termination of the inner loop in step [2.2].

An important point for the correctness of the algorithm is the question, whether it really results in an equation of the form $\mathcal{R}^{(s)}_{q+r}$. Due to the structure of the loops, it will generate equations of the form $(((((\mathcal{R}_{q+r_1})^{(1)})_{+r_2})^{(1)})\cdots)_{+r_s}$. Since the proposed algorithm first prolongs to an involutive symbol and then projects, we can always apply Theorem 2.14 and get indeed an equation of the predicted form.

The check whether a symbol or an equation is involutive makes no difficulties with our criterion (2.8) and Theorem 2.16. To decide whether or not $\mathcal{R}_q$ and $\mathcal{R}^{(1)}_q$ are equal, it suffices to compare their dimensions, as the second is a submanifold of the first. The dimension of $\mathcal{R}^{(1)}_q$ can be computed indirectly using Theorem 2.7. The proof of that theorem indicated clearly the construction of the integrability conditions.

Finally, we show by a Noetherian argument that the algorithm will terminate. To simplify the notation, we set $m = 1$. The proof runs along the same lines for higher m. Since we know already from Theorem 2.18 that the inner loop terminates, it suffices to consider the equations $\mathcal{R}^{(1)}_{q+r_1}, \mathcal{R}^{(2)}_{q+r_2}, \ldots$ obtained after each repetition of the outer loop and the corresponding symbols $\mathcal{M}^{(1)}_{q+r_1}, \mathcal{M}^{(2)}_{q+r_2}, \ldots$ We will associate with each an ideal in the polynomial ring $\mathbb{R}[\chi_1, \ldots, \chi_n]$.

Let the symbol $\mathcal{M}_q$ of the original equation be described by the system $A^{\tau\mu}(P)v_\mu = 0$ for an arbitrary point $P \in \mathcal{R}_q$. This defines a set of polynomials through the substitution

$$v_\mu = v_{[\mu_1 \cdots \mu_n]} \mapsto \chi_\mu = \chi_1^{\mu_1} \cdots \chi_n^{\mu_n} . \tag{2.23}$$

Let I_0 denote the ideal generated by these polynomials.

Let the basis of the ideal I_i be given by the basis of I_{i-1} plus all polynomials obtained by applying (2.23) to the equations of the symbol $\mathcal{M}^{(i)}_{q+r_i}$. This means essentially that we add one new polynomial for each new integrability condition, since differentiation with respect to x^i corresponds now to multiplication with χ_i. Thus we obtain a new independent generator in the ideal basis only, if an integrability condition has occurred.

This procedure yields an ascending chain of ideals $I_0 \subseteq I_1 \subseteq \cdots \subseteq I_s \subseteq \cdots$. It is well-known from algebra that such a chain must become stationary. Consequently there exists an s after which no more integrability conditions have to be added. But then the equation is involutive and the algorithm terminates.

To finish the proof of the Cartan-Kuranishi Theorem, we note that every solution of $\mathcal{R}_q$ is also a solution of $\mathcal{R}^{(s)}_{q+r}$, as by the definition of an integrability condition every solution of the original equation also satisfies all integrability conditions. Prolongations have of course no effect on the solution space. (Recall that we consider only smooth solutions!) •

With the help of these considerations, we can design a more detailed algorithm using essentially only linear algebra and partial derivatives. It is shown in Fig. 2.3. There remains essentially only one *algebraic* problem, namely determining the dimensions of the submanifolds $\mathcal{R}^{(s)}_{q+r}$.

Algorithm `Completion` $(\mathcal{R}_q)$

Input: Equation $\mathcal{R}_q$
Output: Equation $\mathcal{R}^{(s)}_{q+r}$ in involution

[1]	r $\leftarrow$ 0; s $\leftarrow$ 0	
[2]	compute $\mathcal{R}_{q+1}$	{*prolong*}
[3]	compute $\mathcal{M}_q, \mathcal{M}_{q+1}$	{*extract symbols*}
[4]	`until` $\mathcal{R}^{(s)}_{q+r}$ involutive `repeat`	
[4.1]	`while` $\#multVar(\mathcal{M}^{(s)}_{q+r}) \neq \operatorname{rank} \mathcal{M}^{(s)}_{q+r+1}$ `repeat`	
[4.1.1]	r $\leftarrow$ r+1	{*counter for prolongations*}
[4.1.2]	compute $\mathcal{R}^{(s)}_{q+r+1}$	{*prolong*}
[4.1.3]	compute $\mathcal{M}^{(s)}_{q+r+1}$	{*extract symbol*}
[4.2]	`if` $\dim \mathcal{R}^{(s)}_{q+r+1} - \dim \mathcal{M}^{(s)}_{q+r+1} < \dim \mathcal{R}^{(s)}_{q+r}$ `then`	
[4.2.1]	s $\leftarrow$ s+1	{*counter for projections*}
[4.2.2]	compute $\mathcal{R}^{(s)}_{q+r}$	{*add integrability conditions*}
[4.2.3]	compute $\mathcal{R}^{(s)}_{q+r+1}$	{*prolong*}
[4.2.4]	compute $\mathcal{M}^{(s)}_{q+r}, \mathcal{M}^{(s)}_{q+r+1}$	{*extract symbols*}
[5]	`return` $\mathcal{R}^{(s)}_{q+r}$	

Figure 2.3: Algorithm for the Cartan-Kuranishi Theorem (detailed version).

The Cartan-Kuranishi Theorem also shows that it is no restriction to assume from now on that we deal exclusively with involutive equations, as we possess an algorithm to complete every differential equation in a finite number of steps to an involutive one.

Example 2.19 We demonstrate this algorithm at a standard example in formal theory [120] originally given by Janet. It comprises two linear equations of second order in one dependent and three independent variables:

$$\mathcal{R}_2 : \begin{cases} u_{zz} + y\,u_{xx} = 0\,, \\ u_{yy} = 0\,. \end{cases} \tag{2.24}$$

We take the ordering $x^3 = z$, $x^2 = y$, $x^1 = x$ for the definition of the classes. The algorithm in Fig. 2.3 tells us to prolong first until the symbol becomes involutive. $\mathcal{M}_2$ comprises one equation of class 3 and one of class 2; this means there are 5 multiplicative variables but $\operatorname{rank}\mathcal{M}_3 = 6$. Hence according to Definition 2.8 $\mathcal{M}_2$ is not involutive. Since $\mathcal{M}_3$ has 11 multiplicative variables and $\operatorname{rank}\mathcal{M}_4 = 11$, the inner loop terminates after one iteration with the involutive symbol $\mathcal{M}_3$.

We must check whether there are integrability conditions: $\dim\mathcal{R}_4 = 15$ and $\dim\mathcal{M}_4 = 4$, thus with Theorem 2.12 $\dim\mathcal{R}_3^{(1)} = 11 < \dim\mathcal{R}_3 = 12$. This implies the existence of one integrability condition. The construction outlined in Section 2.2 leads to the equation

$$\partial_{yy}(u_{zz} + y\,u_{xx}) - \partial_{zz}u_{yy} - y\partial_{xx}u_{yy} = u_{xxy} = 0\,. \tag{2.25}$$

After the first iteration of the outer loop we have arrived at the equation $\mathcal{R}_3^{(1)}$. We start the same procedure again. Since $\mathcal{M}_3^{(1)}$ is not involutive, we must prolong once to get the involutive symbol $\mathcal{M}_4^{(1)}$. But again an integrability condition occurs, namely

$$\partial_{xxy}(u_{zz} + y\,u_{xx}) - y\partial_{xx}u_{xxy} - \partial_{zz}u_{xxy} = u_{xxxx} = 0\,. \tag{2.26}$$

This completes the second iteration of the outer loop and we must start anew with $\mathcal{R}_4^{(2)}$.

But we detect again a non-involutive symbol. Its prolongation $\mathcal{M}_5^{(2)}$ vanishes ($\dim\mathcal{M}_5^{(2)} = 0$) and is therefore trivially involutive. Since $\dim\mathcal{R}_5^{(3)} = 12 = \dim\mathcal{R}_5^{(2)}$, no integrability conditions occur and we have finished. Because of the vanishing symbol, $\mathcal{R}_5^{(2)}$ is a finite type equation (see Section 3) and has a solution space of dimension $12 = \dim\mathcal{R}_5^{(2)}$. $\mathcal{R}_4^{(2)}$ provides a further example of an equation which is formally integrable, but not involutive.

Note that in this example a direct integration would have been easily possible. It yields a simple polynomial solution depending on 12 arbitrary constants. If we perform, however, a complicated non-linear change of coordinates in (2.19), then we will probably have problems with finding the explicit solution. In contrast the here outlined computation will hardly change and still be readily performed.

As mentioned above, the proof of Theorem 2.18 as given by Sweeny yields an upper bound for the number of prolongations needed to render a symbol $\mathcal{M}_q$ involutive. This bound depends only on the order q, the number of independent variables n and the number of dependent variables m. It can be expressed in form of a recursion relation [117]:

$$\begin{aligned} \hat{q}(n,m,q) &= \hat{q}(n, m\binom{q+n-1}{n}, 1)\,,\\ \hat{q}(n,m,1) &= m\binom{\hat{q}(n-1,m,1)+n}{n-1} + \hat{q}(n-1,m,1) + 1\,,\\ \hat{q}(0,m,1) &= 0\,. \end{aligned} \tag{2.27}$$

Table 2.4 contains $\hat{q}(n,m,1)$ for different values of m and n. One sees that the values explode with growing n and m. The situation is still worse for higher order equations. It is not known, whether this bound is sharp, i.e. whether or not there are symbols which need so many prolongations to become involutive. But fortunately, in applications $\hat{q}(n,m,q)$ yields usually a coarse over-estimate of the actually needed number of prolongations.

The "worst" symbol we will encounter occurred in Example 2.10. The symbol of (2.21) becomes involutive after $q-1$ prolongations. Evaluation of (2.27) yields however

$$\hat{q}(2,1,q) = \frac{1}{4}q^4 + \frac{1}{2}q^3 + \frac{9}{4}q^2 + 2q + 2\,. \tag{2.28}$$

This yields e.g. $\hat{q}(2,1,1) = 7$ and $\hat{q}(2,1,2) = 23$.

$n\backslash m$	1	2	3	4
1	2	3	4	5
2	7	14	23	34
3	53	287	999	2.699
4	29.314	8.129.858	503.006.503	13.151.182.504

Figure 2.4: $\hat{q}(n, m, 1)$ for different values of m and n.

2.5 Power Series Solutions

Definition 2.1 reminds of algebraic geometry, where one associates an algebraic variety with an ideal of polynomials, namely the set of common zeros of all elements of the ideal. There is, however, one big difference between this variety and the submanifolds $\mathcal{R}_q$ considered in the geometric theory of partial differential equations: points in $\mathcal{R}_q$ are no solutions! We will see in this section, how it is possible to obtain something similar for differential equations by studying power series solutions. We will restrict throughout this section our attention to analytic equations.

In order to construct a formal power series solution we need all prolongations of the given differential equation. Thus we need a jet bundle of infinite order. Such a bundle can be introduced using the notions of projective and inductive limits in directed systems of sets [94]. These limits allow also to "transport" the topology and differential structure of the finite dimensional jet bundles. We define

$$J_\infty\mathcal{E} = \operatorname*{proj\,lim}_{q\to\infty} J_q\mathcal{E}\,, \tag{2.29}$$

for the jet bundles of different order are related by the natural projections π_q^{q+r}. Local analytic functions on $J_\infty\mathcal{E}$ can be introduced by

$$\mathcal{F}^\omega = \operatorname*{ind\,lim}_{q\to\infty} \mathcal{C}^\omega(J_q\mathcal{E})\,, \tag{2.30}$$

as the spaces $\mathcal{C}^\omega(J_q\mathcal{E})$ are naturally related by inclusion maps. Note that the definition of an inductive limit implies that every element of $\mathcal{F}^\omega$ is also element of at least one $\mathcal{C}^\omega(J_q\mathcal{E})$, i.e. although these functions are defined on an infinite-dimensional manifold, each of them depends only on a finite number of arguments. In a similar manner, although slightly more complicated, one can construct the tangent and co-tangent bundle of $J_\infty\mathcal{E}$ [11], but this is not necessary for our purposes.

Let us introduce the following ideals of local analytic functions vanishing on the differential equation $\mathcal{R}_q$ and its prolongations:

$$I_{q+r} = \left\{ f \in \mathcal{C}^\omega(J_{q+r}\mathcal{E}) \mid f\big|_{\mathcal{R}_{q+r}} \equiv 0 \right\}. \tag{2.31}$$

Then we can define

$$I_\infty \subset \mathcal{F}^\omega = \operatorname*{ind\,lim}_{r\to\infty} I_{q+r}\,. \tag{2.32}$$

and consequently the infinite prolongation $\mathcal{R}_\infty = V(I_\infty) \subset J_\infty\mathcal{E}$ as the set of common zeros of all functions in I_∞. For an involutive differential equation we could alternative use that since $\mathcal{R}^{(s)}_{q+r} = \mathcal{R}_{q+r+s}$ for all r, s, we have natural projections $\pi^k_l : \mathcal{R}_k \to \mathcal{R}_l$ for $k > l$. Thus we can directly write

$$\mathcal{R}_\infty = \operatorname*{proj\,lim}_{q\to\infty} \mathcal{R}_q \,. \tag{2.33}$$

$\mathcal{R}_\infty$ is also called the space of formal solutions of $\mathcal{R}_q$ or a *diffiety* [89] as short form for differential variety. It can be considered as differential analogue to an algebraic variety. The study of its properties is the subject of the so-called *formal geometry* [89, 157]. We will not pursue it further; we only need I_∞ to define general formal power series solutions.

The generic form of a formal power series around an arbitrary but fixed point $x_0 \in X$ is given by

$$u^\alpha(x) = \sum_{|\mu|=0}^{\infty} \frac{a^\alpha_\mu}{\mu!}(x - x_0)^\mu \,, \tag{2.34}$$

with the standard notations: $\mu! = \mu_1!\,\mu_2!\cdots\mu_n!$ and $(x-x_0)^\mu = (x^1-x^1_0)^{\mu_1}\cdots(x^n-x^n_0)^{\mu_n}$. To simplify the notations, we will regard from now on the dependent variables u^α as derivatives of order zero, i.e. we will write them also in the form p^α_μ with $|\mu| = 0$.

Plugging this ansatz into a local representation of $\mathcal{R}_q$ yields algebraic equations for the coefficients a^α_μ. These equations contain still the independent variables x^i. If (2.34) is supposed to be a solution, they must be satisfied for all values of x in a neighborhood of x_0. Instead of expanding we do the same with all prolongations $\mathcal{R}_{q+r}$ and evaluate each system at $x = x_0$. Then we can identify the jet coordinates p^α_μ with the Taylor coefficients a^α_μ and each prolongation leads to one set of equations:

$$\begin{array}{lll} \mathcal{R}_q: & \Phi^\tau(x_0, a^\alpha_\mu) = 0\,, & 0 \le |\mu| \le q\,, \\ \mathcal{R}_{q+1}: & (D_i\Phi^\tau)\,(x_0, a^\alpha_\mu) = 0\,, & 0 \le |\mu| \le q+1\,, \\ \mathcal{R}_{q+r}: & (D_\nu\Phi^\tau)\,(x_0, a^\alpha_\mu) = 0\,, & 0 \le |\mu| \le q+r\,, |\nu| = r\,. \end{array} \tag{2.35}$$

The union of the left hand sides of all these equations yields a basis of the ideal $I_\infty(x_0) \subset I_\infty$ of functions defined in a neighborhood of x_0. Any solution a^α_μ of these infinitely many algebraic equations yields a formal power series solution of the differential equation $\mathcal{R}_q$.

Definition 2.20 *A* general formal power series solution *of the differential equation $\mathcal{R}_q$ in a neighborhood of $x_0 \in X$ consists of a sequence of manifolds $W_0 \subset W_1 \subset W_2 \subset \cdots$ (not necessarily distinct) and mappings*

$$a^\alpha_\mu : W_q \to \mathbb{R}\,, \qquad \begin{cases} \alpha = 1,\dots,m\,, \\ |\mu| = q\,, \end{cases} \tag{2.36}$$

such that

$$f \in I_\infty(x_0) \iff f\Big(x_0, a^\alpha_\mu(W_q)\Big) \equiv 0 \quad \text{(where defined)}. \tag{2.37}$$

The manifolds W_q in this definition represent the arbitrary parameters of the general solution. Note that it is decisive in (2.37) that there is an equivalence between the two statements. One direction ensures that (2.34) yields a formal solution for any value of the parameters. The other one ensures that we have indeed a general solution, i.e. no other differential equation than those contained in $I_\infty(x_0)$ are solved by (2.34). (More on defining general solutions can be found in Section 3.1).

Obviously there are many ways to write a general formal power series solution. The most important ones for us are those, where we take a basis of $I_\infty(x_0)$ in which all functions are of the form

$$p^\alpha_\mu = \varphi^\alpha_\mu(x^i, u^\beta, p^\gamma_\nu) \tag{2.38}$$

with different left hand sides. Such a basis entails a natural distinction into principal and parametric derivatives: principal are all derivatives appearing on the left hand side; all others are parametric. The mappings a^α_μ take now a very simple form. The parametric coefficients are identified with the arbitrary parameters; the principal ones are obtained by using the functions φ^α_μ.

Example 2.21 As a simple example we consider the Heat Equation $u_{xx} = u_t$. Taking $x^2 = x$ and $x^1 = t$ we can also write it as $p_{[0,2]} = p_{[1,0]}$. If we choose $p_{[1,0]}$ as the principal derivative, then the parametric ones are all derivatives of the form $p_{[0,q]}$ (pure x-derivatives). We obtain the following general formal power series solution: We set $W_q = \mathbb{R}^{2q+1}$ and denote the coordinates in these spaces with w_i, $i = 1, \dots$ Then we take the mappings $a_{[r,s]} = w_{2r+s+1}$. It is straightforward to check that this is a general formal power series solution of the Heat Equation, since $a_{[r,s]} = a[0, 2r+s]$. Note that this provides indeed only a formal solution, as the resulting series will not converge for arbitrary values of the parameters w_i.

Our main interest will lie in formal power series solution which are even more specialized. We say that a power series can be constructed *order by order*, if in the basis (2.38) all functions φ^α_μ depend only on derivatives p^γ_ν with $|\nu| \leq |\mu|$, i.e. the principal derivatives of order q depend only on parametric derivatives of the same or lower order. The power series solution constructed for the Heat Equation was not of this type. We expressed the principal coefficient $a_{[1,0]}$ through the parametric one $a_{[0,2]}$, i.e. one of higher order.

We will describe now a general procedure to construct general formal power series solutions for involutive equations order by order. We know from Section 2.2 that the formal derivative is linear in the highest-order derivatives. Thus we can rewrite in (2.35) the equations for $r > 0$ as inhomogeneous linear systems, where the right hand sides depend only on coefficients of lower order and the left hand sides are the symbols $\mathcal{M}_{q+r}$. For the coefficients of order $q + r$ we get from (2.11) the following system

$$\sum_{\alpha,|\mu|=q} \left(\frac{\partial \Phi^\tau}{\partial p^\alpha_\mu}\right)(x_0, a^\beta_\lambda)\, a^\alpha_{\mu+\nu} = B^\tau_\nu(x_0, a^\gamma_\rho)\,, \qquad \begin{cases} |\lambda|\leq q\,, \\ |\nu|=r\,, \\ |\rho|<q+r\,. \end{cases} \tag{2.39}$$

The B^τ_ν contain the remaining terms of the formal derivative.

In general, system (2.39) will be under-determined and there will be different possibilities to choose the coefficients $a^\alpha_{\mu+\nu}$ for which the system will be solved. These are the principal coefficients. All other coefficients of order $q + r$ are parametric. It is important to note that no choice is distinguished and that the choices at the different orders can be made completely independently.

We are thus lead to the "algorithm" in Fig. 2.5. It is valid for any method of choosing the principal derivatives in step [2.2]. The remaining problem is essentially to solve the non-linear equations for the coefficients of order q and lower in step [1], i.e. an algebraic problem. Note that in general the solution space of such an algebraic system may contain several components of different dimensions. According to Definition 2.1, $\mathcal{R}_q$ is, however, a submanifold (this can for instance be achieved by selecting one smooth component). Thus we can solve by the Implicit Function Theorem for $\operatorname{codim} \mathcal{R}_q$ principal coefficients of order q or less.

All formal power series solutions generated with this "algorithm" have one thing in common: At order $q+r$ there are exactly $\dim \mathcal{M}_{q+r}$ parametric derivatives. These numbers will be fairly important in Chapter 3 for measuring the arbitrariness. We will give them therefore a special name.

Algorithm `PowerSeriesSolution` $(\mathcal{R}_q, maxOrd)$

Input: Involutive equation $\mathcal{R}_q$; maximal order $maxOrd$
Output: Formal power series solution

[1] Compute the general solution of $\Phi^\tau(x_0, a^\alpha_\mu) = 0$ {*non-linear equations!*}
[2] **for** r **in** $1..maxOrd$ **repeat**
[2.1] Set up system (2.39)
[2.2] Choose principal coefficients of order $q + r$
[2.3] Solve for principal coefficients of order $q + r$ {*linear equations!*}
[3] **return** coefficients

Figure 2.5: "Algorithm" to compute power series solution of involutive $\mathcal{R}_q$.

Definition 2.22 *The* Hilbert polynomial *of an involutive symbol $\mathcal{M}_q$ is defined by*

$$H_q(r) = \dim \mathcal{M}_{q+r} \,. \tag{2.40}$$

If $\mathcal{R}_q$ is an involutive differential equation with the symbol $\mathcal{M}_q$, then it denotes the number of arbitrary Taylor coefficients in a general formal power series solution of $\mathcal{R}_q$ which can be constructed order by order.

The name Hilbert polynomial expresses a similarity between this analysis of power series solutions and the analysis of the coordinate ring of a monomial ideal I in algebraic geometry as introduced by Hilbert [73]. There the affine Hilbert function $HF_I(s)$ is defined as the number of monomials of total degree less than or equal to s not contained in I. A well-known theorem [34] asserts that for s large enough the Hilbert function becomes a polynomial $HP_I(s)$. The smallest integer s_0 for which $HF_I(s) = HP_I(s)$ is called the index of regularity. (The restriction to monomial ideals is not severe, as an arbitrary ideal has the same Hilbert function as the ideal spanned by all its leading monomials).

These ideas can be directly related to partial differential equations with the mapping (2.23) used in the proof of the Cartan-Kuranishi Theorem. If we set again for simplicity $m = 1$, then it associated with a derivative p_μ the monomial $\chi_1^{\mu_1} \cdots \chi_n^{\mu_n}$. The total degree of this monomial is obviously the order of the derivative. If we consider the ideal spanned by the monomials obtained from the principal derivatives, then the parametric derivatives correspond to the monomials not contained in the ideal. At this stage it is not at all clear that $H_q(r)$ is indeed a polynomial. But in Section 3.2 we will give an explicit expression for it, from which it will become evident. Thus we see that the order at which a symbol becomes involutive provides an upper bound for the index of regularity of the associated monomial ideal. More on this relationship can be found in [25].

The here used notation for the Hilbert polynomial leads sometimes to confusion. The index q does not distinguish different polynomials but refers to the order used as basis for the argument r. Thus the value $H_q(3)$ gives the number of arbitrary coefficients at order $q + 3$ and not at order 3. This is especially important, if one wants to compare the Hilbert polynomials of two different differential equations. Let us assume that their respective orders are q_1 and q_2 with $q_2 > q_1$. In order to see which equations allows for more arbitrary coefficients at a given order, one must compare $H_{q_1}(r + q_2 - q_1)$ and $H_{q_2}(r)$.

2.6 Characteristics and the Tableau

Recalling the definition (2.8) of the symbol map σ we define a further linear map parameterized by a one-form over the base manifold X. This new way to look at the symbol will turn out to be quite fruitful. It will bring us not only characteristics but will also allow us to define rigorously under- and over-determinacy, respectively, of a differential equation.

Definition 2.23 *Let $\chi = \chi_i dx^i \in T^*X$ be an arbitrary one-form over X and let $\mathcal{R}_q$ be a differential equation locally defined by the map $\Phi : J_q\mathcal{E} \to \mathcal{E}'$. The* tableau *$\tau_\chi$ of $\mathcal{R}_q$ is the linear mapping*[4]

$$\tau_\chi \,:\, \left\{ \begin{array}{ccccc} V\mathcal{E} & \xrightarrow{\iota_\chi} & S_qT^*X \otimes V\mathcal{E} & \xrightarrow{\sigma} & \mathcal{E}' \\ v^\alpha & \longmapsto & \chi_\mu v^\alpha & \longmapsto & \displaystyle\sum_{\alpha, |\mu|=q} \left(\frac{\partial \Phi^\tau}{\partial p^\alpha_\mu}\right) \chi_\mu v^\alpha \end{array} \right. \tag{2.41}$$

where $\chi_\mu = (\chi_1)^{\mu_1} \cdots (\chi_n)^{\mu_n}$.

The tableau is sometimes also called *principal symbol* [123]. Its matrix

$$\mathcal{T}^\tau_\alpha[\chi] = \sum_{|\mu|=q} \left(\frac{\partial \Phi^\tau}{\partial p^\alpha_\mu}\right) \chi_\mu \tag{2.42}$$

has p rows and m columns, where p is the dimension of $\mathcal{E}'$. The entries of this matrix are homogeneous polynomials of degree q in the coefficients of χ.

Definition 2.24 *We call a non-zero one-form $\chi \in T^*X$* characteristic *in $P \in \mathcal{R}_q$, if the tableau τ_χ is not injective at the point P. The set of all characteristic one-forms at P is the* characteristic variety *C_P.*

C_P is indeed a variety in the strict algebraic meaning of the word. The condition that τ_χ is not injective can be expressed as the vanishing of certain subdeterminants of the matrix $\mathcal{T}[\chi]$, i.e. as the vanishing of polynomials in χ_i. Characteristics are most important for semi-linear equations, i.e. equations where the symbol depends only on the independent variables. Then we can define them at points $x_0 \in X$.

Example 2.25 To see the connection between this definition and the usual one let us consider a linear first-order system

$$\mathcal{R}_1 \,:\, \left\{ \sum_{i,\alpha} A^{\tau i}_\alpha(x) p^\alpha_i + \sum_\alpha B^\tau_\alpha(x) u^\alpha + C^\tau(x) = 0\,, \quad \tau = 1, \dots, p \right. \tag{2.43}$$

and an $(n-1)$-dimensional hypersurface Σ in solved form

$$\Sigma \,:\, x^n = \phi(x^1, \dots, x^{n-1})\,. \tag{2.44}$$

Such a surface is usually called characteristic, if prescribing the values of u^α on Σ does not suffice to compute all derivatives on the surface.

We assume that $u^\alpha\big|_\Sigma = f^\alpha(x^1, \dots, x^{n-1})$. This yields using the chain rule

$$\left.\frac{\partial u^\alpha}{\partial x^i}\right|_\Sigma + \frac{\partial \phi}{\partial x^i} \left.\frac{\partial u^\alpha}{\partial x^n}\right|_\Sigma = \frac{\partial f^\alpha}{\partial x^i}\,, \quad i = 1, \dots, n-1\,. \tag{2.45}$$

[4] This is a kind of "dual" definition to [17] where τ_χ is called symbol and its solution space tableau.

Substituting these relations into (2.43) yields a linear system for the derivatives $\partial u^\alpha/\partial x^n|_\Sigma$ with matrix

$$A_\alpha^{\tau n} - \sum_{i=1}^{n-1} A_\alpha^{\tau i}\frac{\partial\phi}{\partial x^i}\,. \tag{2.46}$$

If the rank of this matrix is less than m, i.e. the associated linear mapping is not injective, then it is not possible to determine all derivatives on Σ. But this means nothing else than that the one-form $\chi = dx^n - \sum_{i=1}^{n-1}(\partial\phi/\partial x^i)dx^i$ is characteristic according to Definition 2.24. Hence $\chi = d\Phi$ where $\Phi = x^n - \phi(x^1,\dots,x^{n-1}) = 0$ describes Σ. If we introduce the co-normal variety $N_x^*\Sigma$ as the set of all one-forms annihilating the tangent space of the surface[5], i.e. $N_x^*\Sigma = \{\omega \in T_x^*X \mid \omega(T_x\Sigma) = 0\}$, then we see that $N^*\Sigma$ is generated by χ.

The tableau τ_χ allows us to give a precise definition of under- and over-determined equations. Intuitively we would consider a differential equation as under-determined, if its general solution contains one (or more) arbitrary function of all independent variables $x^1,\dots,x^n$, as this implies that the equation poses no restrictions at all for at least one of the dependent variables u^α.

Definition 2.26 *The involutive equation $\mathcal{R}_q$ is said to be* under-determined *at $P \in \mathcal{R}_q$, if there exists no one-form χ such that τ_χ is injective. If there exists a one-form χ such that τ_χ is bijective, then the equation is* well-determined *or* normal. *Otherwise the equation is called* over-determined.

This definition of an under-determined equation is more precise than the one given by Pommaret [117]. It is a straightforward generalization of the one given by Olver [111] for systems with the same number of equations and dependent variables. We will see later in Chapter 3, that it is indeed equivalent to requiring that the general solution contains arbitrary functions of all independent variables. Olver's definition of over-determined equations is not satisfactory, as he does not require the equation to be in involution. Thus it can happen that an over-determined system becomes under-determined by *adding* an integrability condition!

Theorem 2.27 *Let $\mathcal{R}_q$ be a differential equation with tableau τ_χ. In different coordinate systems we might get different values for $\beta_q^{(n)}$. Let $\tilde\beta_q^{(n)}$ be the maximum of these values. Then*

$$\max_\chi\ \operatorname{rank}\tau_\chi = \tilde\beta_q^{(n)}\,. \tag{2.47}$$

Proof. By definition of $\tilde\beta_q^{(n)}$ we can find a local description of $\mathcal{R}_q$ where the first $\tilde\beta_q^{(n)}$ equations $\Phi^\tau(x^i,u^\alpha,p_\mu^\alpha) = 0$ satisfy

$$\frac{\partial\Phi^\tau}{\partial p^\alpha_{[0,\dots,0,q]}} = \delta_\alpha^\tau\,. \tag{2.48}$$

The diagonal elements of the top left $\tilde\beta_q^{(n)} \times \tilde\beta_q^{(n)}$ submatrix of $\mathcal{T}[\chi]$ are now of the form $(\chi_n)^q + \dots$ and the monomial $(\chi_n)^q$ occurs in no other place in the matrix. If we choose $\chi = dx^n$ all entries will vanish with the exception of the above mentioned diagonal elements which will become one. Hence the rank of τ_χ is at least $\tilde\beta_q^{(n)}$.

Assume now the rank be greater than $\tilde\beta_q^{(n)}$. We choose in a neighborhood of a point $x_0 \in X$ a new coordinate system $\bar x^i$ given by $\bar x^i = \psi^i(x^j)$ where the function ψ^n satisfies

$$\frac{\partial\psi^n}{\partial x^i}(x_0) = \chi_i\,. \tag{2.49}$$

[5] In the language of Appendix A $N^*\Sigma$ provides a Pfaffian representation of Σ.

Obviously this is always possible. For the derivatives of class n we obtain

$$\frac{\partial \bar{\Phi}^\tau}{\partial \bar{p}^\alpha_{[0,\dots,0,q]}} = \frac{\partial \Phi^\tau}{\partial p^\alpha_\nu} \prod_{i=1}^{n} \left(\frac{\partial \psi^n}{\partial x^i} \right)^{\nu_i} . \tag{2.50}$$

At the point x_0 we can evaluate the derivatives of ψ and the right hand side becomes the matrix of the tableau τ_χ. By assumption its rank is greater than $\tilde{\beta}^{(n)}_q$. But this implies that we have found a coordinate system where more than $\tilde{\beta}^{(n)}_q$ equations are of class n which clearly contradicts our definition of $\tilde{\beta}^{(n)}_q$. •

Corollary 2.28 *An involutive equation $\mathcal{R}_q$ is under-determined, if and only if $\tilde{\beta}^{(n)}_q < m$. $\mathcal{R}_q$ is normal, if and only if $\tilde{\beta}^{(n)}_q = m$ and all other $\tilde{\beta}^{(k)}_q = 0$.*

PROOF. By definition $\mathcal{R}_q$ is under-determined, if and only if there exists no χ such that $\operatorname{rank} \tau_\chi = m$. But by the preceding theorem this happens if and only if $\tilde{\beta}^{(n)}_q < m$. Similarly we get the value of $\tilde{\beta}^{(n)}_q$ for a normal equation. It follows from the bijectivity of the tableau that the other $\tilde{\beta}^{(k)}_q$ must vanish. •

One-forms χ for which the rank of the tableau is less than $\tilde{\beta}^{(n)}_q$ are called *systatic* [117]. For equations which are not under-determined systatic is equivalent to characteristic. It is important to note that under-determinacy depends only on $\tilde{\beta}^{(n)}_q$ and not on the total number p of equations used to describe $\mathcal{R}_q$, as only the former value is intrinsically defined (cf. next section). Of course a system which can be locally described with less equations than unknowns will always be under-determined, but the converse is not true.

Example 2.29 Consider the Maxwell Equation in two dimensions

$$\begin{aligned} u_{tt} - v_{xt} &= 0 \,, \\ u_{xt} - v_{xx} &= 0 \,. \end{aligned} \tag{2.51}$$

The tableau of this system is given by the matrix

$$\mathcal{T}[\chi] = \begin{pmatrix} (\chi_1)^2 & -\chi_1\chi_2 \\ \chi_1\chi_2 & -(\chi_2)^2 \end{pmatrix} . \tag{2.52}$$

If we multiply the first row by χ_2 and the second by χ_1, we obtain the same. Thus an upper bound for the rank of the tableau is 1 and it takes indeed this value for every non-vanishing one-form χ. This yields $\tilde{\beta}^{(2)}_2 = 1$ and it is not possible to find a coordinate system in which there occur two derivatives of class 2. The tableau is never injective and every one-form is characteristic. Hence the system is indeed under-determined, although there are as many equations as dependent variables. We will see later that this is typical for gauge theories and related to a symmetry. Since $\mathcal{T}[\chi]$ has the same rank for every non-vanishing one-form χ there are no systatic one-forms and every coordinate system is δ-regular (see below).

Corollary 2.30 *If there exists a one-form χ such that the tableau τ_χ of the differential equation $\mathcal{R}_q$ is surjective, then $\mathcal{R}_q$ is involutive.*

PROOF. τ_χ can be only surjective, if the number p of equations in $\mathcal{R}_q$ is less than or equal to the number of dependent variables. The maximal rank of the tableau is then p. It follows from Theorem 2.27 that all equations are of class n. But this implies immediately that $\mathcal{M}_q$ is involutive and that no integrability conditions arise; thus $\mathcal{R}_q$ is involutive. •

An alternative way of dealing with non-linear equations is to define characteristics with respect to a given solution. For this purpose we introduce the *linearization of an equation* $\mathcal{R}_q$ *at a solution* τ. If $\mathcal{R}_q$ is locally described by $\Phi^\tau(x^i, u^\alpha, p^\alpha_\mu) = 0$ for $\tau = 1, \ldots, p$ and the section σ by $x \mapsto (x, f^\alpha(x))$, we obtain the linearization by expanding around $u^\alpha = f^\alpha, p^\alpha_\mu = \partial_\mu f^\alpha$ and taking only the first order terms. This yields the linear equation

$$\mathcal{P}_q(\sigma) \,:\, \left\{ \sum_\alpha \frac{\partial \Phi^\tau}{\partial u^\alpha}\bigg|_{u=f} \bar{u}^\alpha + \sum_{\alpha,\mu} \frac{\partial \Phi^\tau}{\partial p^\alpha_\mu}\bigg|_{u=f} \bar{p}^\alpha_\mu = 0 \,, \right. \tag{2.53}$$

where $u = f$ abbreviates the substitution of $f^\alpha(x)$ and its derivatives for u^α, p^α_μ. The one-form χ is now said to be characteristic for the solution σ, if χ is characteristic for $\mathcal{P}_q(\sigma)$ [123].

2.7 δ-Regularity and the k-Tableaux

δ-Regularity concerns our criterion (2.19) for involutive symbols. We had to introduce the class of a multi-index to define the $\beta_q^{(k)}$, an obviously coordinate dependent concept. One can show, however, that almost every coordinate system will lead to the same values for the $\beta_q^{(k)}$. In this section we will discuss this problem in greater detail.

As the class is defined for multi-indices, we need to consider only changes of the independent coordinates x^i. We will see later that it even suffices to take into account only linear changes. Such a transformation is characterized by a matrix A with constant coefficients a_{ij}. If we perform a generic change of coordinates, i.e. without specifying the matrix A, the coefficients a_{ij} will appear in the symbol in a complicated, non-linear way.

If we order the columns by class and compute a row echelon form of the symbol in this generic coordinate system, we will obtain the maximal value for $\beta_q^{(n)}$ admitted by the considered differential equation which we called in the last section $\tilde{\beta}_q^{(n)}$. Furthermore this yields the maximal admitted values for the sums $\beta_q^{(n)} + \beta_q^{(n-1)}$, $\beta_q^{(n)} + \beta_q^{(n-1)} + \beta_q^{(n-2)}$, $\ldots$, $\beta_q^{(n)} + \cdots + \beta_q^{(1)}$. These are the intrinsically defined values for the $\beta_q^{(k)}$. Let us denote them for a moment by $\tilde{\beta}_q^{(k)}$. Later we will omit the tilde and always assume that the $\beta_q^{(k)}$ take these values. A more direct definition of them can be given using the Spencer cohomology [117]. Definition 2.8 and all the results in Section 2.3 apply only to the $\tilde{\beta}_q^{(k)}$.

Definition 2.31 *Let $\mathcal{M}_q$ be the symbol of differential equation; let $\beta_q^{(k)}$ be the number of equations of class k in a given coordinate system. Then the coordinate system is called* δ-regular *for $\mathcal{M}_q$, if $\beta_q^{(k)} = \tilde{\beta}_q^{(k)}$.*

Note that this definition depends crucially on the symbol under consideration. Especially after a projection, when new equations are added to the symbol, the coordinate system must be checked again, even if it was δ-regular before. Prolongations are of course harmless. We further see, why it suffices to consider linear changes of the independent variables. For the analysis of the symbol only one aspect is important: which derivatives of order q do occur after the transformation. But this is determined by the Jacobi matrix. Thus the linearized transformation contains already all relevant information.

For special choices of the coefficients a_{ij} a pivot in the row echelon form may vanish and a row may change to a lower class. Then the corresponding matrix A leads to a δ-singular coordinate system. As the coefficients a_{ij} enter the symbol algebraically, the set of δ-singular systems can be described by an algebraic variety in the space of $n \times n$-matrices. If we make a random choice of A, the probability to pick a singular one is zero, as this variety is a set of measure zero. Nevertheless it is of course possible to get one! But this offers an alternative to performing the expensive generic calculation: one generates a random coordinate system and completes the differential equation in this one to an involutive system.

Example 2.32 δ-regularity is sometimes considered as a purely technical nuisance in geometric approaches to differential equations. This is, however, not true, as it is connected with characteristics. As a simple example, we consider the Wave Equation

$$u_{tt} - u_{xx} = 0\,. \tag{2.54}$$

The coordinate system (x, t) is obviously δ-regular, no matter how we order x and t. In a general coordinate system $\eta = \eta(x, t), \zeta = \zeta(x, t)$, Equation (2.54) becomes

$$\begin{gathered}[\eta_t^2 - \eta_x^2]u_{\eta\eta} + 2[\eta_t\zeta_t - \eta_x\zeta_x]u_{\eta\zeta} + [\zeta_t^2 - \zeta_x^2]u_{\zeta\zeta} + \\ [\eta_{tt} - \eta_{xx}]u_\eta + [\zeta_{tt} - \zeta_{xx}]u_\zeta = 0\,.\end{gathered} \tag{2.55}$$

No matter how we order the new coordinates, we will loose the derivative of class 2, if $\zeta_t = \pm\zeta_x$ and $\eta_t = \mp\eta_x$. But these are exactly the equations for the two characteristics of the Wave Equation. It is not harmful if one of the coordinates is characteristic, as we can always choose the other one as $\bar{x}^2$. This comes from the fact that of course only the coordinates x^k with $\beta_q^{(k)} > 0$ are important for δ-regularity.

One could be tempted to think that since almost every coordinate system is δ-regular, δ-regularity is not important in applications. Unfortunately, this is not true. It turns out that one encounters fairly often differential equations which are not given in such a coordinate system (e.g. "light-cone coordinates"). Thus one needs a practical tool to cope with this problem. In principle, the above said solves it: one must simply compute in a generic coordinate system. In practice this is, however, not feasible, as the computations simply blow up.

There are two reasons for this effect. First we introduce n^2 indeterminates which enter the computations in a complicated non-linear way. Thus even if the original system has constant coefficients, they become polynomials in these n^2 variables. Second, and this is usually more severe, we make the system dense. A typical differential equation of order q does not depend on all jet variables up to order q. Introducing the generic coordinate system generates many new jet variables and all the matrices become dense. One can see this effect very clearly at the symbol. Typically it is a fairly sparse matrix. Now it becomes a dense matrix and the entries are much more complicated, as they contain more jet variables and the new indeterminates.

But one could argue that making a complete change of coordinates is an "over-kill". As we have seen in Example 2.32, only those coordinates x^k where $\beta_q^{(k)} > 0$ are important. Thus one might wonder, whether it is not possible to look only at those. Furthermore one could argue whether one really wants to change the coordinate system. There might be good reasons to use a special one. What we really need, is an intrinsic method to compute the numbers $\beta_q^{(k)}$.

We have already found a way to calculate $\beta_q^{(n)}$ intrinsically through Theorem 2.27. The proof of this theorem showed that in a δ-regular coordinate system the tableau τ_χ has maximal rank, if we choose $\chi = dx^n$. We will now introduce generalized tableaux depending on up to n one-forms $\chi^{(k)}$ such that in a δ-regular coordinate system these tableaux have maximal rank, if we choose $\chi^{(k)} = dx^k$, and such that conversely, if $\chi^{(k)}$ are a set of one-forms which lead to maximal rank of all these tableaux, then the coordinate system defined by $d\bar{x}^k = \chi^{(k)}$ is δ-regular. Furthermore the maximal ranks of the generalized tableaux will directly determine the $\beta_q^{(k)}$.

The transformation of derivatives under changes of coordinates look rather messy in multi-index notation. We must hence introduce some notations. If μ is a multi-index of length q, we will denote by $\mathcal{S}_\mu$ the set of all possible realizations of it with repeated indices. An element $s \in \mathcal{S}_\mu$ is thus a tuple of q integers containing μ_i times the value i; e.g. with $n = 2$ one gets $\mathcal{S}_{[1,2]} = \{(1,2,2); (2,1,2); (2,2,1)\}$. We will denote the k-th element of s by s_k. Furthermore we will abbreviate by $r(\mu)$ the realization of μ as a repeated index in which the entries are sorted (smallest first).

Definition 2.33 *Let $\chi^{(k)}, \dots, \chi^{(n)}$ be $n-k+1$ linearly independent one-forms in T^*X for $1 \leq k \leq n$ and let $N^k \subset X$ be a submanifold such that the one-forms represent a basis of T^*N^k. Let the differential equation $\mathcal{R}_q$ be locally described by the map $\Phi : J_q\mathcal{E} \to \mathcal{E}'$. The* k-tableau $\tau_{[\chi^{(k)},\dots,\chi^{(n)}]}$ *is the linear mapping*

$$\tau_{[\chi^{(k)},\dots,\chi^{(n)}]} : \begin{cases} S_qT^*N^k \otimes V\mathcal{E} & \xrightarrow{\iota_{[\chi]}} & S_qT^*X \otimes V\mathcal{E} & \xrightarrow{\sigma} & \mathcal{E}' \\ v_\nu^\alpha & \longmapsto & C_\mu^\nu v_\nu^\alpha & \longmapsto & \displaystyle\sum_{\alpha,|\mu|=q} \left(\frac{\partial\Phi^\tau}{\partial p_\mu^\alpha}\right) C_\mu^\nu v_\nu^\alpha \end{cases} \tag{2.56}$$

where ν is a multi-index with n entries of length q and $\operatorname{class}\nu \geq k$ *and the factor C_μ^ν is defined by*

$$C_\mu^\nu = \sum_{s\in\mathcal{S}_\nu} \prod_{i=1}^{q} \chi_{m_i}^{(s_i)}\,, \qquad m = r(\mu)\,. \tag{2.57}$$

We see that the only difference to Definition 2.23 lies in the first map of (2.56) which is generalized to one parameterized by several one-forms. The multi-index ν provides a convenient way to introduce local coordinates in $S_qT^*N^k \otimes V\mathcal{E}$; one could also say that we choose local coordinates such that N^k is given by $x^1 = const, \dots, x^{k-1} = const$. The matrix of the k-tableau is then given by

$$\mathcal{T}_\alpha^{\tau\nu}[\chi^{(k)}, \dots, \chi^{(n)}] = \sum_{|\mu|=q} \left(\frac{\partial\Phi^\tau}{\partial p_\mu^\alpha}\right) C_\mu^\nu\,, \qquad \text{with } \operatorname{class}\nu \geq k\,,\ |\nu| = q\,. \tag{2.58}$$

It has p rows indexed by τ and mr_k columns indexed by α and ν where p is again the dimension of $\mathcal{E}'$ and $r_k = \dim S_qT^*N^k \otimes V\mathcal{E} = \binom{n-k+q}{n-k}$. The n-tableau is the usual one as defined in the last section. Since $N^1 = X$, the 1-tableau as the largest one has the same size as the symbol. Indeed we will later see that we can identify it with the symbol in a generic coordinate system.

Lemma 2.34 *The matrix of the k-tableau contains the matrices of all ℓ-tableaux with $\ell > k$ as submatrices.*

Proof. We can view the matrix $\mathcal{T}[\chi^{(k)}, \dots, \chi^{(n)}]$ as a horizontal concatenation of r_k matrices with p rows and m columns each generated similar to $\mathcal{T}[\chi]$ in Section 2.6 but with different mappings ι_χ. The lemma stems from the fact that in (2.58) all multi-indices with $\operatorname{class}\nu \geq k$ are taken into account. This includes of course all multi-indices with $\operatorname{class}\nu \geq \ell$ for $\ell > k$. (2.57) shows that for such multi-indices the entries of the k-tableau do not depend on $\chi^{(k)}, \dots, \chi^{(\ell-1)}$. Thus we obtain all columns of the ℓ-tableau. •

Theorem 2.35 *Let $\mathcal{R}_q$ be a differential equation with characters $\beta_q^{(n)}$. Then for $1 \le k \le n$*

$$\max_{\chi^{(k)},\dots,\chi^{(n)}} \operatorname{rank} \tau_{[\chi^{(k)},\dots,\chi^{(n)}]} = \sum_{i=k}^{n} \beta_q^{(i)} . \tag{2.59}$$

PROOF. The proof is a straightforward generalization of the one of Theorem 2.27. In a δ-regular coordinate system there exists a local description of $\mathcal{R}_q$ such that the first $\beta_q^{(n)}$ equations are solved for derivatives of class n, the next $\beta_q^{(n-1)}$ equations are solved for derivatives of class $n-1$ and so on. We order the columns of the k-tableau in such a way that the first m columns represent the n-tableau, the first mr_{n-1} columns the $n-1$ tableau and so on. This is always possible according to Lemma 2.34.

Choosing $\chi^{(k)} = dx^k, \dots, \chi^{(n)} = dx^n$, we can easily evaluate the factors C_μ^ν. They vanish whenever $\mu \neq \nu$ and are one otherwise. Thus we have eliminated all contributions to the k-tableau by derivatives whose class is less than k. If the principal derivative of equation τ is p_μ^α, then in row τ the column indexed by α and μ contains the first non-vanishing entry. Since our local representation was chosen in such a way that the symbol is in solved form, all principal derivatives are different and we have at least $\sum_{i=k}^{n} \beta_q^{(i)}$ linearly independent rows.

On the other hand assume that there are one-forms $\chi^{(k)}, \dots, \chi^{(n)}$ such that the rank of the k-tableau is larger than the claimed value. Then we apply a coordinate transformation $\bar{x}^j = \psi^j(x)$ where the functions ψ^j satisfy in a point $x_0 \in X$

$$\frac{\partial \psi^j}{\partial x^i}(x_0) = \chi_i^{(j)} , \quad j = k, \dots, n . \tag{2.60}$$

Due to the linear independence of the one-forms $\chi^{(i)}$ this is always possible. The effect of this transformation on the symbol is determined by the highest order part of the transformation law for the derivatives of order q. It is given by

$$\frac{\partial p_\mu^\alpha}{\partial \bar{p}_\nu^\alpha} = \sum_{s \in \mathcal{S}_\nu} \prod_{i=1}^{g} \left(\frac{\partial \psi^{s_i}}{\partial x^{m_i}} \right) , \quad |\mu| = |\nu| = q,\ m = r(\mu) . \tag{2.61}$$

Since we are interested in the values of $\beta_q^{(k)}, \dots, \beta_q^{(n)}$, we consider only the following derivatives

$$\frac{\partial \bar{\Phi}^\tau}{\partial \bar{p}_\nu^\alpha} = \sum_{|\mu|=q} \left(\frac{\partial \Phi^\tau}{\partial p_\mu^\alpha} \right) \frac{\partial p_\mu^\alpha}{\partial \bar{p}_\nu^\alpha} , \quad \operatorname{class} \nu \ge k . \tag{2.62}$$

But at x_0 we can evaluate the derivatives of ψ^j occurring at the right hand side when we plug in (2.61). Comparing (2.62) and (2.58) we immediately recognize the k-tableau at the right hand side. Thus we have found a coordinate system in which we have more equations of class greater than or equal to k than given by $\sum_{i=k}^{n} \beta_q^{(i)}$. But this contradicts the definition of the $\beta_q^{(k)}$. •

Definition 2.36 *Let $\chi^{(i)} \in T^*X,\ i = 1, \dots, n$ be n linearly independent one-forms. They define a* non-systatic basis *of T^*X, if*

$$\operatorname{rank} \tau_{[\chi^{(k)},\dots,\chi^{(n)}]} = \sum_{i=k}^{n} \beta_q^{(i)} , \qquad k = 1, \dots, n . \tag{2.63}$$

Coordinates x^i such that $dx^i = \chi^{(i)}$ are called an associated coordinate system.

Different associated coordinate system differ only by constants, namely the choice of an origin. Note that in this definition only the $\chi^{(k)}$ where $\beta_q^{(k)} > 0$ really matter. For the remaining one-forms one can choose any which complete to a basis of T^*X. Theorem 2.35 implies now immediately the following

Corollary 2.37 *A coordinate system on X is δ-regular, if and only if it is associated to a non-systatic basis of T^*X.*

From (2.61) and (2.62) we also see that if we write down the symbol in a generic linear coordinate system, i.e. in coordinates $\bar{x}^j = \sum a_i^j x^i$, then we get the 1-tableau after identifying a_i^j with $\chi_i^{(j)}$. Thus the introduction of the k-tableau allows us to introduce these generic coordinates step by step. Since we often must change only some of the coordinates, this can save considerable computing time in typical applications, as one rarely has to go until the 1-tableau. Note furthermore that Lemma 2.34 allows us to organize the computation in an efficient way. If we calculate the rank of the k-tableau as column rank, we can keep the results from the $(k+1)$-tableau, as we only add further columns to the matrix.

Example 2.38 We demonstrate the use of the k-tableaux in an example stemming from a concrete application [83]. The analysis of moving pseudo-spherical surfaces in $\mathbb{R}^3$ leads to the so-called Bianchi equation. This is a third-order equation of the form:

$$\begin{aligned} & u_{xyt} - u_{yt}u_x \cot u + u_{xt}u_y \tan u = 0\,, \\ & \left(\frac{u_{xt}}{\cos u}\right)_x - K(K\sin u)_t - \frac{u_{yt}u_y}{\sin u} = 0\,, \\ & \left(\frac{u_{yt}}{\sin u}\right)_y + K(K\cos u)_t - \frac{u_{xt}u_x}{\cos u} = 0\,, \end{aligned} \tag{2.64}$$

where K is some given function of t. Obviously the used coordinate system is not δ-regular no matter how we order the coordinates, for there is no derivative of class 3. We will use the ordering $x^1 = t$, $x^2 = x$ and $x^3 = y$. Since each equation in (2.64) contains only one third order derivative, the symbol takes after some trivial manipulations the simple form

$$\begin{aligned} v_{xyt} &= 0\,, \\ v_{xxt} &= 0\,, \\ v_{yyt} &= 0\,. \end{aligned} \tag{2.65}$$

Now the matrix of the 3-tableau is readily computed.

$$\mathcal{T}[\chi^{(3)}] = \begin{pmatrix} \chi_1^{(3)}\chi_2^{(3)}\chi_3^{(3)} \\ \chi_1^{(3)}(\chi_2^{(3)})^2 \\ \chi_1^{(3)}(\chi_3^{(3)})^2 \end{pmatrix}. \tag{2.66}$$

We find two families of characteristic one-forms

$$\begin{aligned} \omega_1 &= \alpha\, dx^1\,, \\ \omega_2 &= \beta\, dx^2 + \gamma\, dx^3\,. \end{aligned} \tag{2.67}$$

This implies especially that the three coordinate forms dx^1, dx^2, dx^3 are all characteristic. Thus it is not surprising that the coordinate system is not δ-regular. The maximal rank of the 3-tableau is 1 and obtained e.g. with the choice $\chi^{(3)} = dx^1 + dx^3$.

Next we compute the 2-tableau. There exist four third order derivatives of class 2, namely [0,0,3], [0,1,2], [0,2,1] and [0,3,0]. Hence its matrix has four columns:

$$\mathcal{T}[\chi^{(2)},\chi^{(3)}] = \left(\begin{array}{ll} \chi_1^{(3)}\chi_2^{(3)}\chi_3^{(3)} & \chi_1^{(2)}\chi_2^{(3)}\chi_3^{(3)} + \chi_1^{(3)}\chi_2^{(2)}\chi_3^{(3)} + \chi_1^{(3)}\chi_2^{(3)}\chi_3^{(2)} \\ \chi_1^{(3)}(\chi_2^{(3)})^2 & \chi_1^{(2)}(\chi_2^{(3)})^2 + 2\chi_1^{(3)}\chi_2^{(3)}\chi_2^{(2)} \\ \chi_1^{(3)}(\chi_3^{(3)})^2 & \chi_1^{(2)}(\chi_3^{(3)})^2 + 2\chi_1^{(3)}\chi_3^{(3)}\chi_3^{(2)} \end{array}\right.$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \tag{2.68}$$

$$\left.\begin{array}{cc} \chi_1^{(2)}\chi_2^{(2)}\chi_3^{(3)} + \chi_1^{(2)}\chi_2^{(3)}\chi_3^{(2)} + \chi_1^{(3)}\chi_2^{(2)}\chi_3^{(2)} & \chi_1^{(2)}\chi_2^{(2)}\chi_3^{(2)} \\ \chi_1^{(3)}(\chi_2^{(2)})^2 + 2\chi_1^{(2)}\chi_2^{(2)}\chi_2^{(3)} & \chi_1^{(2)}(\chi_2^{(2)})^2 \\ \chi_1^{(3)}(\chi_3^{(2)})^2 + 2\chi_1^{(2)}\chi_3^{(2)}\chi_3^{(3)} & \chi_1^{(2)}(\chi_3^{(2)})^2 \end{array}\right) .$$

One can now compute the maximal rank of this matrix using Gaussian elimination. It is, however, much easier to substitute the above mentioned non-characteristic one-form $\chi^{(3)} = dx^1 + dx^3$. This yields the following, much simpler matrix

$$\mathcal{T}[\chi^{(2)}, dx^1 + dx^3] = \begin{pmatrix} 0 & \chi_2^{(2)} & \chi_1^{(2)}\chi_2^{(2)} + \chi_2^{(2)}\chi_3^{(2)} & \chi_1^{(2)}\chi_2^{(2)}\chi_3^{(2)} \\ 0 & 0 & (\chi_2^{(2)})^2 & \chi_1^{(2)}(\chi_2^{(2)})^2 \\ 1 & \chi_1^{(2)} + 2\chi_2^{(2)} & (\chi_3^{(2)})^2 + 2\chi_1^{(2)}\chi_3^{(2)} & \chi_1^{(2)}(\chi_3^{(2)})^2 \end{pmatrix} . \tag{2.69}$$

One sees at once that the rank of this matrix is 3 whenever $\chi_2^{(2)} \neq 0$. Thus we get $\beta_3^{(3)} = 1$ and $\beta_3^{(2)} = 2$ and we have also found a non-systatic basis of T^*X

$$\begin{aligned} \chi^{(3)} &= dx^1 + dx^3 , \\ \chi^{(2)} &= dx^2 , \\ \chi^{(1)} &= dx^1 , \end{aligned} \tag{2.70}$$

An associated δ-regular coordinate system is given by $\bar{x}^3 = y + t$, $\bar{x}^2 = x$ and $\bar{x}^1 = t$. This yields the following form for the symbol

$$\begin{aligned} \bar{v}_{333} + \bar{v}_{331} &= 0 , \\ \bar{v}_{332} + \bar{v}_{321} &= 0 , \\ \bar{v}_{322} + \bar{v}_{221} &= 0 , \end{aligned} \tag{2.71}$$

where one can easily read off the $\beta_3^{(k)}$ directly.