Chapter 4

Initial Value Problems

We will study in detail the properties of initial value problems. This entails a precise definition of well-posed problems and some existence and uniqueness theorems for analytic systems. For linear equations we will generalize the Holmgren Theorem to obtain uniqueness also for non-analytic solutions. Finally, we will comment on the problem of stability.

4.1 Formally Well-Posed Problems

In Chapter 3 we obtained formal results on the number of arbitrary functions in the general solution of a differential equation. However, we could not prove that there actually exists a representation of the general solution to a certain non-negative solution of the diophantine system (3.23). In this chapter we will try to give a partial answer to this question using initial value problems. We will restrict ourselves almost always to analytic systems.

Before proving any existence and uniqueness theorem we must consider what kind of initial value problems make sense for involutive equations. The classical definition of a *well-posed* initial value problem is apparently due to Hadamard [82]. He required that the solution of the problem

(i) exists for arbitrary Cauchy data,

(ii) is determined uniquely by the Cauchy data,

(iii) depends continuously on the Cauchy data.

This is still fairly vague, as several key words are not precisely defined. We will give a formal and precise version of this definition omitting, however, the third point.

Definition 4.1 *An initial value problem is called* formally well-posed, *if it has a unique formal power series solution for arbitrary formal power series as Cauchy data.*

One could say that in such a problem exactly the right amount of Cauchy data is prescribed. If we posed more initial conditions, we would obtain compatibility conditions for the data; with less data the solution would no longer be unique. How the solution depends on the Cauchy data is not considered.

Corollary 4.2 *An initial value problem is formally well-posed, if and only if it is possible to construct a formal general power series solution such that there exists a bijective mapping between the parametric coefficients of the series and the Taylor coefficients of the Cauchy data.*

Obviously there arises now the question, how does one find formally well-posed initial value problems given a differential equation. In a fixed coordinate system Janet-Riquier Theory (cf. Appendix B) provides an easy method. Applying the Reid diagrams introduced in Section 2.3 one can in low-dimensional problems very simply read off correct initial prescriptions. For more complicated examples one can use an algorithm developed by Reid [124] which computes them automatically.

The idea behind this method is simple. So far we have always concentrated on the principal derivatives in our analysis. We have seen in Proposition 2.9, how one can obtain them all in a unique way by prolonging each equation with respect to its multiplicative variables only. Now we analyses in a similar way the parametric derivatives. Instead of a formal treatment we demonstrate the method with a few examples.

Before we should point out, however, that Reid's algorithm is designed for passive systems. We can nevertheless apply it to involutive equations provided we work in a δ-regular coordinate system and write the system in solved form with respect to a class-respecting ranking. This connection between the two concepts is established in Proposition B.3 in Appendix B.

Example 4.3 As first example we consider the Heat Equation $u_{xx} = u_t$. According to our ranking u_{xx} must be taken as principal derivative. Thus we take x as x^2. This leads to the diagram in Fig. 4.1. As usual the shaded area contains the principal derivatives.

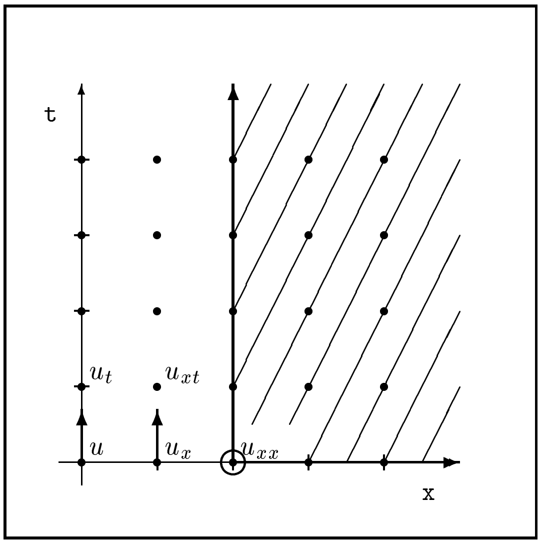

Figure 4.1: Reid diagram for the Heat Equation.

Obviously the parametric derivatives are of the form $p_{[n,0]}$ or $p_{[n,1]}$. Thus they lie all on two rays starting at u and u_x, respectively, and we prescribe for an arbitrary value x_0

$$u(t, x_0) = f(t)\,, \quad u_x(t, x_0) = g(t)\,. \tag{4.1}$$

It is easy to see that there is a bijection between the Taylor coefficients of f and the derivatives $p_{[n,0]}$, and the Taylor coefficients of g and the derivatives $p_{[n,1]}$, respectively. Thus this initial value problem is formally well-posed.

This is, however, not the initial value problem physicists usually consider for the Heat Equation but

$$u(x, t_0) = h(x) . \tag{4.2}$$

To obtain this one we must choose u_t as principal derivative, i.e. a derivative of lower order. There is a crucial difference between (4.1) and (4.2). Whereas for (4.1) Riquier's existence theorem (cf. Appendix B) ensures the convergence of the formal power series for arbitrary analytic functions f and g, we have no such result for the standard initial value problem. Indeed it is well-known that one must pose growth conditions on h to get a convergent series.

The second example considers a simple non-normal system. It shows an important fact that we will study later in more detail, namely that for such systems it is not possible to prescribe all initial data on a hypersurface of codimension 1. This might be obvious for over-determined systems, but here we consider an under-determined equation.

Example 2.29 (cont.) Take again the Maxwell Equation in two dimensions

$$\begin{aligned} u_{tt} - v_{xt} &= 0 , \\ u_{xt} - v_{xx} &= 0 . \end{aligned} \tag{4.3}$$

We have already seen in Section 2.6 that this equation is under-determined and that the coordinate system is δ-regular. We choose in each equation the u-derivative as principal derivative. Fig. 4.2 contains the Reid diagram for u. Since we have no v-derivatives as principal derivatives, v is unconstrained by the system and we can choose it arbitrarily.

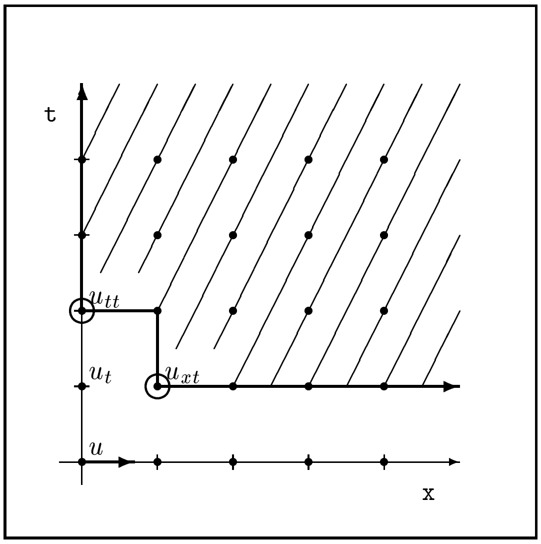

Figure 4.2: Reid diagram for the two-dimensional Maxwell Equation.

We see that in order to get all parametric derivatives, we need three initial conditions

$$u_t(x_0, t_0) = f , \quad u(x, t_0) = g(x) , \quad v(x, t) = h(x, t) . \tag{4.4}$$

Thus we prescribe some data on a zero-dimensional submanifold, the point (x_0, t_0), some data on a one-dimensional submanifold, the axis $t = t_0$, and some data on the whole manifold. This is a typical feature for non-normal systems. As soon as several pivots belong to the same dependent variable, the boundary of the area with the principal derivatives will have a kind of (in general higher-dimensional) echelon form. This entails the necessity to use submanifolds of different dimensions.

Finally, we want to illustrate with a simple example one of the reasons why it is important to use a δ-regular coordinate system. Besides the fact that the system might not be passive, e.g. because we gave a characteristic problem, the Janet-Riquier Theory sometimes yields "initial value problems", one would not necessarily call so.

Example 4.4 Consider the Wave Equation in characteristic (or light-cone) coordinates

$$u_{xy} = 0\,. \tag{4.5}$$

Obviously we are not dealing with a δ-regular coordinate system, but the system is nevertheless involutive and passive. Reid's algorithm leads to the "initial value problem"

$$u(x, y_0) = f(x)\,, \quad u_y(x_0, y) = g(y)\,. \tag{4.6}$$

(One could also take u in the second condition, but then one must require that $f(x_0) = g(y_0)$.) This represents, however, more a kind of boundary value problem than an initial value problem.

4.2 Intrinsic Description of Initial Value Problems

In the classical point of view of initial value problems Cauchy data are represented by an embedding of a submanifold of X with codimension 1 in an appropriate jet bundle [56]. This also corresponds to physical intuition: the data are prescribed on a space-like manifold and the differential equation determines its evolution. Furthermore the Cauchy-Kowalevsky Theorem (see next section) shows that in the case of normal systems this leads to unique solutions.

But the example of the two-dimensional Maxwell Equations in the last section has shown that this point of view can not be upheld as soon as we leave the category of normal equations. Physicists speak often of constrained systems: if Cauchy data is nevertheless prescribed on a submanifold of codimension 1, it cannot be chosen completely arbitrarily but must satisfy "constraints" (we will discuss this approach in Chapter 6 in more detail).

Janet-Riquier Theory shows how to avoid these constraints: initial data is prescribed on a flag of submanifolds instead of just one. Although this point of view is implicit in the Cartan-Kähler analysis of exterior systems, we know of only two articles [87, 88] (by the same author) where it has been used explicitly in a geometric approach to differential equations. There it was, however, always assumed that the problem is formally well-posed and no analysis of the conditions for this to be the case is given. We will give in the sequel an intrinsic description of these conditions using the k-tableaux. Simply said, we will show how to pick principal derivatives in a geometric way independent of a coordinate system.

The method to construct formally well-posed initial value problems hinted in the last section is very general on one side. By using the notion of a passive system one can treat on the same footing also characteristic problems, problems like (4.2) where one chooses derivatives of lower order as principal ones, and many more. The particularities of these problems are caught by the ranking used and different rankings will lead to different passive systems. On the other hand this method depends crucially on the used coordinate system. We prescribe the Cauchy data always on coordinate hyperplanes and rankings can be defined only with respect to a given coordinate system.

We will restrict our attention for most of the remainder of this chapter on first order equations. We will further assume that no lower order, i.e. algebraic, equations are present. They could always be solved for some of the dependent variables and thus used to eliminate them. These assumptions are mainly for notational simplicity. We know from Appendix C.2 that any equation can be transformed into an equivalent first order equation, thus there is no loss of generality.

Let $\chi^{(k)}, k = 1, \dots, n$ be a non-systatic basis of T^*X as defined in Section 2.7. We associate with it two flags of submanifolds. Let $X^k \subset X$ be a one-dimensional submanifold ("coordinate axis") such that $T^*X^k = <\chi^{(k)}>$, then we introduce $N_k = X^1 \times \cdots \times X^{k-1}$ ($N_1 = \{x_0\}$} for an arbitrary but fixed point $x_0 \in X$) and $N^k = X^k \times \cdots \times X^n$. The N^k have already been used in the definition of the k-tableaux. We call a flag such as N_k a *non-systatic flag of submanifolds.*

Since we consider only first order equations, the k-tableau is given by a mapping from $T^*N^k \otimes V\mathcal{E}$ to $\mathcal{E}'$. We search now for a flag of fibered submanifolds M^k of $\mathcal{E}$ such that $\emptyset \subseteq M^1 \subseteq \cdots \subseteq M^n \subseteq M^{n+1} = \mathcal{E}$ and

$$\operatorname{rank} \tau_{[\chi^{(k)},\dots,\chi^{(n)}]}\Big|_{T^*X^k \otimes VM^k} = \beta_q^{(k)} . \tag{4.7}$$

The subset condition reflects the fact that in a δ-regular coordinate system not the individual values $\beta_q^{(k)}$ become maximal but the sums $\sum_{i=k}^{n} \beta_q^{(i)}$.

Lemma 4.5 *Such a flag M^k exists for every involutive first-order equation $\mathcal{R}_1$. Furthermore one can choose M^k such that $\dim M^k = \beta_1^{(k)}$ (with $\beta_1^{(n+1)} = m$).*

PROOF. We choose a coordinate system associated with the $\chi^{(k)}$. Let $(x_0, u_0) \in \mathcal{E}$ be an arbitrary but fixed point. In a neighborhood of it we can describe the N^k by the equations $x^i = x_0^i$ for $i = 1, \dots, k-1$ and we claim that the fibered submanifolds given by the equations $u^\alpha = u_0^\alpha$ for $\alpha = \beta_1^{(k)} + 1, \dots, m$ can be taken as M^k.

Proposition 2.12 implies that if we choose the principal derivatives according to the inverse lexicographic ranking, then with p_k^α are also all p_i^α with $i > k$ principal derivatives. By the Implicit Function Theorem there exists a local description of $\mathcal{R}_1$ of the form

$$\Phi_k^\alpha(x, u, p) = p_k^\alpha - \varphi_k^\alpha(x^i, u^\beta, p_j^\gamma) = 0 , \qquad \begin{cases} \alpha = 1,\dots,\beta_1^{(k)} , \\ j = 1,\dots,k , \\ \gamma = \beta_1^{(j)}+1,\dots,m . \end{cases} \tag{4.8}$$

Thus there exist principal derivatives of class k only for $1 \leq \alpha \leq \beta_1^{(k)}$. But restricting the k-tableau to $T^*X^k \otimes VM^k$ means nothing else but taking only the corresponding equations into account for its construction. Hence we get for its rank $\beta_1^{(k)}$. The M^k define indeed a flag due to Corollary 2.13 and have obviously the correct dimensions. •

From (4.8) it is easy to read off a formally well-posed initial value problem:

$$\begin{aligned} u^\alpha(x_0) &= f^\alpha \ (= Const.) , \qquad \alpha = 1, \dots, \beta_1^{(1)} , \\ u^\alpha(x^1, \dots, x^k, x_0^{k+1}, \dots, x_0^n) &= f^\alpha(x^1, \dots, x^k) , \\ &\qquad \alpha = \beta_1^{(k)} + 1, \dots, \beta_1^{(k+1)} . \end{aligned} \tag{4.9}$$

In order to give an intrinsic description of these initial conditions, we introduce the fibered submanifolds M_k by $M^k = M_1 \times \cdots \times M_k$. Thus $\dim M_k = \beta_1^{(k)} - \beta_1^{(k-1)}$. Let us denote by μ_k the natural projection from $\mathcal{E}$ to M_k[1]. Instead of the functions f^α we use now fiber

[1] One should perhaps remark that in general M_k is *not* a subbundle of $\mathcal{E}$. This will, however, always be the case, if $\mathcal{E}$ is trivial. In principle it makes no sense at all to speak about global properties of the M_k, as they exist in general only locally. Furthermore it may happen that the values $\beta_q^{(k)}$ vary on $\mathcal{R}_q$; then one must restrict the definition of the M_k to a neighborhood with constant values.

preserving mappings $\rho_k : N_k \to M_k$. If the section $\sigma \in \Gamma_{loc}(\mathcal{E})$ is a local solution of $\mathcal{R}_1$, then (4.9) can be written as

$$\mu_k(\sigma)\Big|_{N_k} = \rho_k\,. \tag{4.10}$$

Proposition 4.6 *$\mathcal{R}_1$ together with (4.10) is a formally well-posed initial value problem.*

Example 2.29 (cont.) To make this fairly awkward construction more transparent, we detail the involved submanifolds for the two-dimensional Maxwell Equation. In order to rewrite it as a first-order system, we introduce the additional dependent variables $w^2 = u_t, w^1 = u_x$ and $z^2 = v_t, z^1 = v_x$

$$\begin{aligned}
&u_t - w^2 = 0\,, && u_x - w^1 = 0\,,\\
&v_t - z^2 = 0\,, && v_x - z^1 = 0\,,\\
&w_t^2 - z_t^1 = 0\,, && w_x^2 - z_x^1 = 0\,,\\
&w_t^1 - w_x^2 = 0\,, && \\
&z_t^1 - z_x^2 = 0\,. &&
\end{aligned} \tag{4.11}$$

This yields $\beta_1^{(2)} = 5$ (left column) and $\beta_1^{(1)} = 3$ (right column). We take in each equation the first term as principal derivative. A non-systatic basis of T^*X is given by $\chi^{(1)} = dx, \chi^{(2)} = dt$.

We can now choose the submanifolds $N_3 = X$, $N_2 = \{t = t_0\}$, $N_1 = \{(x_0, t_0)\}$ and $M^3 = \mathcal{E}$, $M^2 = \{z^2 = z_0^2\}$, $M^1 = \{z^2 = z_0^2, z^1 = z_0^1, w^1 = w_0^1\}$. This can be seen as follows: The system (4.11) has been set up in such a way that it is preferable to choose $x^2 = t, x^1 = x$. Hence the values of the N_k. M^2 comes from the fact that there are no principal derivatives involving z^2. For all other dependent variables, we have a principal derivative of class 2. To get M^1 we observe that only for u, v, w^2 principal derivatives of class 1 occur.

Summarizing we get the initial value problem

$$\begin{aligned}
&z^2(x,t) = f_1(x,t)\,, && z^1(x,t_0) = f_2(x)\,, && w^1(x,t_0) = f_3(x)\,,\\
&w^2(x_0,t_0) = f_4\,, && v(x_0,t_0) = f_5\,, && u(x_0,t_0) = f_6\,.
\end{aligned} \tag{4.12}$$

The connection with (4.4) is given by

$$f = f_4\,,\quad g(x) = \int_{x_0}^{x} f_3(\rho)d\rho + f_6\,,\quad h(x,y) = \int_{t_0}^{t} f_1(x,\tau)d\tau + \int_{x_0}^{x} f_2(\rho)d\rho + f_5\,. \tag{4.13}$$

The initial conditions (4.10) are sufficient to select a unique formal power series solution. The natural question now is under what conditions does this formal series converge, i.e. we should look for a strong existence theorem. For general equations this is only possible, if everything (equations, Cauchy data, solution) is analytic. This will be the topic of the next section. In Section 4.4 we will extend the uniqueness also to non-analytic solutions, however only for linear equations.

Example 4.7 Getting rid of the restriction to analytic equations is much harder due to the existence of smooth involutive equations without solutions. Lewy [96] found the first example of such a system. He showed that there exists a $\mathcal{C}^\infty$ function F such that the linear(!), normal system

$$\begin{aligned}
u_x - v_y + 2yu_z + 2xv_z &= F(x,y,z)\,,\\
v_x + u_y - 2xu_z + 2yv_z &= 0
\end{aligned} \tag{4.14}$$

has no $\mathcal{C}^\infty$ solution (actually not even a $\mathcal{C}^1$ solution).

4.3 Cartan-Kähler Theorem

In this section we will consider only analytic equations and we will look only for analytic solutions. This implies that the Cauchy data as well as the submanifolds on which they are prescribed are also analytic. The "mother of all existence theorems" in this category is the well-known Cauchy-Kowalevsky Theorem which covers normal equations.

Theorem 4.8 (Cauchy-Kowalevsky) *Let $\phi^\alpha(x^i, u^\beta, p_j^\beta)$ and $f^\alpha(x^1, \dots, x^{n-1})$ where $\alpha, \tau = 1, \dots, m$, $i = 1, \dots, n$, $j = 1, \dots, n-1$ be (real) analytic functions of all their arguments in a neighborhood of the origin. Then the initial value problem*

$$\begin{aligned} p_n^\alpha &= \phi^\alpha(x^i, u^\beta, p_j^\beta)\,, \\ u^\alpha(x^1, \dots, x^{n-1}, 0) &= f^\alpha(x^1, \dots, x^{n-1})\,, \end{aligned} \tag{4.15}$$

has a unique solution that is analytic in a neighborhood of the origin of X. Unique meaning that no other analytic solution exists.

Since the proof of this theorem can be found in almost any text book on partial differential equations (e.g. [82, 116]), we omit it here. Most books consider only quasi-linear equations. One constructs a power series solution along the lines described in Section 2.5 and proves then (this is the hard part) its convergence by the method of majorants. One can easily reduce the non-linear case to the quasi-linear one by differentiating (4.15) with respect to x^n and introducing the first order derivatives $u^\alpha_{x^i}$ as new dependent variables p_i^α [33]. This leads to the quasi-linear system

$$\begin{aligned} u^\alpha_{x^n} &= p_n^\alpha\,, \\ p^\alpha_{j,x^n} &= p^\alpha_{n,x^j}\,, \qquad\qquad j = 1, \dots, n-1\,, \\ p^\alpha_{n,x^n} &= \Phi^\tau_{x^n} + \sum_{\beta=1}^m \Phi^\tau_{u^\beta} p_n^\beta + \sum_{\beta=1}^m \sum_{j=1}^{n-1} \Phi^\tau_{p_j^\beta} p^\beta_{n,x^j}\,, \end{aligned} \tag{4.16}$$

which is obviously again in the form required by Theorem 4.8. If we augment it with the initial conditions

$$\begin{aligned} u^\alpha(x^1, \dots, x^{n-1}, 0) &= f^\alpha(x^1, \dots, x^{n-1})\,, \\ p_j^\alpha(x^1, \dots, x^{n-1}, 0) &= \partial_{x^j} f^\alpha(x^1, \dots, x^{n-1})\,, \\ p_n^\alpha(x^1, \dots, x^{n-1}, 0) &= \Phi^\tau(x^1, \dots, x^{n-1}, 0, f^\beta, \partial_{x^j} f^\beta)\,, \end{aligned} \tag{4.17}$$

then every solution of it yields a solution of (4.15) and vice versa.

We can compare the results of Theorem 4.8 and Theorem 3.9. Since for normal equations $\beta_1^{(n)} = m, \beta_1^{(n-1)} = \cdots = \beta_1^{(1)} = 0$, we obtain $f_{n-1} = m, f_n = f_{n-2} = \cdots = f_1 = 0$ as solution of the recursion relation (3.16) according to Corollary 3.10 in perfect agreement with (4.15) where m functions of $n-1$ variables are prescribed. Note, however, the crucial difference between the two theorems: Theorem 3.9 makes no statement about the existence or uniqueness of solutions.

The Cauchy-Kowalevsky Theorem can be extended from normal equations to arbitrary involutive ones. This generalization is usually known as Cartan-Kähler Theorem. As the name already indicates, it stems originally from the Cartan-Kähler Theory of exterior differential systems (cf. Section A). A first non-constructive proof within the framework of formal theory was given by Goldschmidt [58] using the Spencer cohomology.

We will give here an adaption of the original proof successively applying the Cauchy-Kowalevsky Theorem, as this demonstrates further interesting properties of involutive equations. It will also allow us to generalize other classical results like the Holmgren Theorem (cf. Section 4.4).

Theorem 4.9 (Cartan-Kähler) *Let the differential equation $\mathcal{R}_1$ be analytic and involutive. There exists one and only one analytic solution σ of $\mathcal{R}_1$ satisfying the initial conditions (4.10) for analytic mappings ρ_k and analytic flags M_k, N_k.*

Proof. We will use the local formulation of this initial value problem given by (4.8,4.9) and construct the solution step by step. In the first step, we treat only the equations of class 1:

$$p_1^\alpha = \varphi_1^\alpha(x^i, u^\beta, p_1^\gamma)\,, \qquad \begin{cases} \alpha = 1,\dots,\beta_1^{(1)}\,, \\ \gamma = \beta_1^{(1)}+1,\dots,m\,. \end{cases} \tag{4.18}$$

If we evaluate (4.18) at $x^j = x_0^j$ for $j > 1$, we can substitute u^β for $\beta > \beta_1^{(1)}$ by the f^β (restricted to N_1) and we obtain a normal system with dependent variables u^α for $\alpha = 1, \dots, \beta_1^{(1)}$, independent variable x^1 and initial conditions $u^\alpha(0) = f^\alpha$. The Cauchy-Kowalevsky Theorem yields thus a unique analytic solution $U^\alpha(x^1, x_0^2, \dots, x_0^n)$ satisfying $U^\alpha(x_0) = f^\alpha$.

The next step takes the equations of class 2 to extend this one-dimensional solution to a two-dimensional one $U^\alpha(x^1, x^2, x_0^3, \dots, x_0^n)$ applying again the Cauchy-Kowalevsky Theorem after the corresponding substitution. This inductive process continues till class n. We omit the details and go directly to step n. The other steps run similarly.

We suppose that we know a solution $U^\alpha(x^1, \dots, x^{n-1}, x_0^n)$ for all equations up to class $n-1$. The equations of class n are

$$p_n^\alpha = \varphi_n^\alpha(x^i, u^\beta, p_j^\gamma)\,, \qquad \begin{cases} \alpha = 1,\dots,\beta_1^{(n)} \\ \gamma = \beta_1^{(j)},\dots,m\,, \\ j = 1,\dots,n\,, \end{cases} \tag{4.19}$$

and they form a normal system, this time with dependent variables u^α for $\alpha = 1, \dots, \beta_1^{(n)}$ and independent variables $x^1, \dots, x^n$. As initial conditions we take

$$u^\alpha(x^1, \dots, x^{n-1}, x_0^n) = \begin{cases} U^\alpha(x^1, \dots, x^{n-1}, x_0^n) & \text{if } 1 \le \alpha \le \beta_1^{(n-1)}\,, \\ f^\alpha(x^1, \dots, x^{n-1}) & \text{if } \beta_1^{(n-1)} + 1 \le \alpha \le \beta_1^{(n)}\,. \end{cases} \tag{4.20}$$

Since these functions are all analytic, we can invoke the Cauchy-Kowalevsky Theorem. The solution of the initial value problem yields the wanted functions $U^\alpha(x)$. u^α for $\alpha > \beta_1^{(n)}$ is unconstrained by the system (4.8) and directly given by the initial conditions (4.9).

We must, however, still show that these functions $U^\alpha(x)$ also solve the full equations of lower class, for in the previous steps we have always treated some independent variables as constants. That is where involution proves to be crucial. If the functions U^α satisfy all equations of system (4.8), then the functions

$$\Delta_i^\alpha(x) = \Phi_i^\alpha(j_1(U^\beta)(x)) \tag{4.21}$$

must vanish in the considered neighborhood of x_0 (this is a formal way to express the substitutions $u^\alpha = U^\alpha(x)$, $p_i^\alpha = \partial U^\alpha(x)/\partial x^i$).

Since system (4.8) is involutive, the prolongation of an equation with respect to a non-multiplicative variable must lead to a linear combination of equations obtained by differentiating all equations with respect to their multiplicative variables only (Proposition 2.9, cf. also the discussion following Equation (2.16)). Thus we have for $i < n$ (among others) the following relations between the prolonged equations

$$D_n \Phi_i^\alpha = \sum_{\substack{1 \le k \le n, 1 \le l \le k \\ 1 \le \beta \le \beta_1^{(k)}}} A_{l\beta}^k(x,u,p) D_l \Phi_k^\beta + \sum_{k,\beta} B_\beta^k(x,u,p) \Phi_k^\beta \,, \tag{4.22}$$

as we differentiate on the left hand side with respect to a non-multiplicative variable. The ranges of the indices on the right hand side are so chosen, that only prolongations with respect to multiplicative variables are taken into account.

If we substitute the functions U^α and their derivatives for the jet variables $u^\alpha, p_i^\alpha, \ldots$, we can easily evaluate the formal derivatives

$$(D_j \Phi_i^\alpha)(j_2(U^\alpha)(x)) = \frac{\partial}{\partial x^j} \Phi_i^\alpha(j_1(U^\alpha)(x)) \,. \tag{4.23}$$

Now (4.22) becomes a normal system

$$\frac{\partial}{\partial x^n} \Delta_i^\alpha = \sum A_{l\beta}^k \frac{\partial}{\partial x^l} \Delta_k^\beta + \sum B_\beta^k \Delta_k^\beta \tag{4.24}$$

and according to our induction hypothesis that $U^\alpha(x^1, \ldots, x^{n-1}, x_0^n)$ satisfy the equations up to class $n-1$, if $x^n = x_0^n$, we get the initial conditions

$$\Delta_j^\beta(x^1, \ldots, x^{n-1}, x_0^n) = 0 \,. \tag{4.25}$$

Applying for the last time the Cauchy-Kowalevsky Theorem, we get that this system has one and only one analytic solution which is trivially $\Delta_j^\beta(x) \equiv 0$. •

One should note that formal integrability is not sufficient for this proof. The first part, i.e. the step by step construction of $U^\alpha(x^1, \ldots, x^n)$ could of course be done for any differential equation even for a not formally integrable one. It is, however, unclear for such a system, whether the so obtained U^α also solve the equations of lower class. Only involution provides the key to set up the final normal system (4.24). Nevertheless, we can conclude that every analytic formally integrable system has an analytic solution, as it will become involutive after a finite number of prolongation.

Again it is instructive to compare this theorem with the results of Chapter 3. The Cauchy data prescribed in (4.9) consist of $f_n = m - \beta_1^{(n)}$ functions of n variables and $f_k = \beta_1^{(k+1)} - \beta_1^{(k)}$ functions of k variables. These are exactly the values predicted by Corollary 3.10 for an algebraic representation of the general solution in the case of a first-order equation.

If we start with an equation $\mathcal{R}_q$ of order q, it can be reduced by one of the two methods outlined in Appendix C.2 to a first-order equation. Directly reducing $\mathcal{R}_q$ leads by (C.21) to the prescription of $\alpha_q^{(k)} - \alpha_q^{(k+1)}$ functions of k arguments. This corresponds exactly to the values computed in (3.27) for a representation with $(q-1)$-dimensional integrals. If we reduce $\mathcal{R}_{q+1}$, we prescribe according to (C.23) $\alpha_q^{(k)}$ functions of k arguments. Comparing with (3.26) we see that this corresponds to a representation with q-dimensional integrals. These integrations occur of course during the transformation of a solution of the reduced equation into one of the original equation $\mathcal{R}_q$.

Example 4.10 A simple example [17] may clarify this construction and the importance of involution. We consider a first-order system with two independent and three dependent variables of the form

$$\begin{aligned} u^1_x &= F(x,y,u^1,u^2,u^3)\,, \\ u^1_y &= G(x,y,u^1,u^2,u^3,u^1_x,u^2_x,u^3_x)\,, \\ u^2_y &= H(x,y,u^1,u^2,u^3,u^1_x,u^2_x,u^3_x) \end{aligned} \tag{4.26}$$

with the initial conditions

$$u^1(0,0) = f^1\,,\ u^2(x,0) = f^2(x),\ u^3(x,y) = f^3(x,y)\,. \tag{4.27}$$

We have two equations of class 2 and one of class 1. We begin with the latter one. This leads to the initial value problem

$$u^1_x = F\big(x,0,u^1,f^2(x),f^3(x,0)\big)\,, \quad u^1(0,0) = f^1 \tag{4.28}$$

which has the unique solution $U^1(x,0)$. In the next step, we consider the problem

$$\begin{aligned} u^1_y &= G\big(x,y,u^1,u^2,f^3(x,y),u^1_x,u^2_x,f^3_x(x,y)\big)\,, \\ u^2_y &= H\big(x,y,u^1,u^2,f^3(x,y),u^1_x,u^2_x,f^3_x(x,y)\big)\,, \\ u^1(x,0) &= U^1(x,0), \quad u^2(x,0) = f^2(x)\,. \end{aligned} \tag{4.29}$$

Its unique solution be the two functions $U^1(x,y), U^2(x,y)$. $U^3(x,y)$ is of course simply $f^3(x,y)$.

But now we must check whether U^1, U^2, U^3 also solve the first equation. We define

$$\Delta(x,y) = U^1_x(x,y) - F\big(x,y,U^1(x,y),U^2(x,y),U^3(x,y)\big)\,. \tag{4.30}$$

Because of the first step $\Delta(x,0) = 0$. Furthermore,

$$\begin{aligned} \Delta_y = \ & G_{u^1}\Delta + \\ & \big[G_x + \textstyle\sum\big(G_{u^\alpha}u^\alpha_x + G_{u^\alpha_x}u^\alpha_{xx}\big) - F_y - \sum F_{u^\alpha}u^\alpha_y\big]\big|_{u^\alpha=U^\alpha}\,. \end{aligned} \tag{4.31}$$

The square bracket contains the integrability condition obtained by cross-differentiating the first and the second equation of (4.26). If we assume that the system is involutive, it must vanish. Thus Δ satisfies

$$\Delta_y = G_{u^1}\Delta\,, \quad \Delta(x,0) = 0\,. \tag{4.32}$$

By the usual uniqueness theorem for ordinary differential equations, it follows that Δ vanishes everywhere.

4.4 Holmgren Theorem

In this section we will try to extend the uniqueness to non-analytic solutions. Using a technique apparently introduced by Holmgren [77] this is fairly easily possible for linear systems. Before we get to the proof of Holmgren's theorem, we have to introduce a few facts about linear differential operators. We will again restrict without loss of generality to first order equations. Let L be a linear first order differential operator

$$Lu = \sum_{i,\alpha} A^{i\tau}_\alpha(x)p^\alpha_i + \sum_\alpha B^\tau_\alpha(x)u^\alpha\,. \tag{4.33}$$

Its *adjoint operator* $\tilde{L}$ is given by

$$\tilde{L}v = -\sum_{\tau,i}\partial_{x^i}\Big[((A^i)^T)^\alpha_\tau(x)v^\tau\Big] + \sum_\tau (B^T)^\alpha_\tau(x)v^\tau\,. \tag{4.34}$$

(Note that we do not require that the matrices A^i and B are square.) Obviously $\tilde{L}$ is normal, if and only if L is normal.

The Lagrange-Green Identity [82] asserts that for vector valued $\mathcal{C}^1$ functions u and v

$$\int_\Omega v^T Lu\,dV = \int_\Omega (\tilde{L}v)^T u\,dV + \int_{\partial\Omega} v^T Au\,dS\,, \tag{4.35}$$

where $\partial\Omega$ denotes the boundary of a region $\Omega \subset X$ and $A = \sum A^i(x)\zeta_i$ with the exterior unit normal ζ of $\partial\Omega$. The notion of an adjoint operator and the Lagrange-Green identity can readily be extended to higher order operators. The surface integral in (4.35) becomes then more complicated and its form is no longer unique but depends on the order in which one performs the partial integrations needed to prove (4.35).

Theorem 4.11 (Holmgren) *Let L be a normal linear differential operator with analytic coefficients. Let $M \subset X$ be a non-characteristic submanifold of dimension $n-1$. Any $\mathcal{C}^1$ solution of the initial value problem*

$$\begin{aligned} Lu &= 0\,, \\ u\big|_M &= 0 \end{aligned} \tag{4.36}$$

vanishes in a neighborhood of M.

Proof. This is the classical theorem as given by Holmgren [77]. We will follow the proof outlined in [33]. The linearity of L is now of course also a question of the coordinate system. We will assume in the sequel always that all considered regions are so small that they can be covered by one δ-regular chart and that in it L is linear.

Let $U(x_0) \subset M$ be a neighborhood of a point $x_0 \in M$. We consider a region $\Omega \subset X$ bounded by $U(x_0)$ and another $(n-1)$-dimensional submanifold N. The initial value problem

$$\begin{aligned} \tilde{L}v &= p\,, \\ v\big|_N &= 0\,, \end{aligned} \tag{4.37}$$

where the components of p are polynomials in the chosen coordinate system, possesses by the Cauchy-Kowalevsky Theorem a solution defined in Ω, if Ω is chosen sufficiently small.

Applying the Lagrange-Green Identity (4.35) yields

$$\int_\Omega p^T u\,dV = 0\,. \tag{4.38}$$

Since this must hold for arbitrary p and the polynomials lie dense in $\mathcal{C}^1$, this implies that $u \equiv 0$ in Ω as claimed. •

This proof is at one point a bit weak. We assumed tacitly that we can find a finite region Ω where a solution of (4.37) exists for all p with polynomial components. One can remedy this and at the same time extend Theorem 4.11 to larger regions by introducing the notion of an *analytic field* S_λ, i.e. a family of submanifolds in co-dimension 1 which sweeps out a region $\Omega \subset X$ for $\lambda \in [a,b] \subset \mathbb{R}$ [82]. Such a region Ω is then called a *domain of determinacy* for M.

Theorem 4.12 *Let L be an involutive linear differential operator with analytic coefficients. Let $x_0 = N_0 \subset N_1 \subset \cdots \subset N_n = X$ be a non-systatic flag of submanifolds with $\dim N_i = i$. Then any $\mathcal{C}^1$ solution of the initial value problem*

$$\begin{aligned} &Lu = 0\,, \\ u^\alpha\big|_{N_k} = 0\,, \qquad &\alpha = \beta_1^{(k)}+1,\dots,\beta_1^{(k+1)} \end{aligned} \tag{4.39}$$

vanishes in a neighborhood of x_0.

PROOF. This extension to a more general situation is straightforward considering the proof we gave for the Cartan-Kähler Theorem. Let u be any solution of the considered initial value problem. Then $u|_{N_1}$ must satisfy (4.18) with vanishing Cauchy data. Thus according to Theorem 4.11 $u|_{N_1} \equiv 0$. Similarly we can conclude following step by step the construction in the proof of Theorem 4.9 that $u|_{N_k} \equiv 0$ for $k = 1, \ldots, n$. Each of these intermediate initial value problems will have vanishing Cauchy data as a consequence of the previous step and thus Holmgren's Theorem is throughout applicable. •

4.5 Stability

To conclude this chapter on initial value problems we make a few remarks on the problem of stability being omitted in our definition of a formally well-posed problem. It is especially of importance for numerical calculations, for they make sense only if the solutions of problems with "similar" Cauchy data are "similar". We have constructed in Section 4.2 a formally well-posed problem for any involutive differential equation. But this problem will be unstable for many equations.

Example 4.13 The classical example of an ill-posed problem is the Cauchy problem for the Laplace Equation

$$u_{xx} + u_{yy} = 0\,. \tag{4.40}$$

Hadamard [33] was the first to show that the solution of the following sequence of initial value problems does not depend continuously on the Cauchy data:

$$u(x,0) = 0\,, \quad u_y(x,0) = f_n(x) = \frac{1}{n^2}\sin(nx)\,. \tag{4.41}$$

Obviously each problem is formally well-posed; the formal power series solutions even converge, as the solution is given by

$$u_n(x,y) = \frac{1}{n^2}\sin(nx)\sinh(ny)\,. \tag{4.42}$$

$f_n(x) \to 0$ for $n \to \infty$, but the solutions $u_n(x,y)$ do not converge to the solution of the initial value problem with $u_y(x,0) = 0$ which is of course $u(x,y) = 0$. Thus the larger y the larger the difference between two solutions with arbitrarily close Cauchy data becomes.

The question is how to define rigorously when the Cauchy data of two different initial value problems are close or similar. One must distinguish here between two possibilities to deform Cauchy data. In numerical analysis one considers usually the situation found in Example 4.13: how is the solution affected by a slight variation in the functions prescribed. But in a geometric formulation like (4.10) one must also take deformations of the submanifolds on which the data is given into account.

Shih used in his thesis [141] the fact that it is possible to define topologies on the space of embeddings [59] to introduce a topology on the space of all Cauchy data for a given equation. He defined a problem to be well-defined, if it is an interior point of the subspace of problems having at least one solution. Thus "small" is defined to mean that there exist an open neighborhood in the used topology such that every problem in it is well-posed.

This approach tackles the second possibility of deformation while neglecting the first one. It considers the Laplace Equation as well-posed, for it only requires that the "neighboring" initial value problems have solutions. But it does not require any relation between their solutions. Shih's formulation is restricted to the standard problem where all data is prescribed on a submanifold of codimension 1. But it should be straightforward to extend it to our more general formulation.

Restricting ourselves to this notion of stability, we can easily prove that our formulation is stable under deformation of the flag N_k. As described in Section 4.2 these nested submanifolds are characterized by a non-systatic basis of T^*X. Thus to study a deformation of them it suffice to analyze deformations of such bases. This avoids the use of complicated topologies, since we can use the standard topology of $\mathbb{R}^n$.

Proposition 4.14 *A non-systatic basis of T^*X is stable under deformations.*

PROOF. In a given coordinate system the conditions for a basis to be non-systatic can be expressed by requiring that certain subdeterminants of the k-tableaux do not vanish. But this implies that systatic bases can be characterized by the vanishing of certain rational functions of the coefficients of the forms. Thus we can define a homeomorphic mapping f from the space of all bases of T^*X into an $\mathbb{R}^p$ such that $f^{-1}(0)$ gives all systatic bases. But the image of a closed set is again closed, thus the complement of $f^{-1}(0)$, the space of all non-systatic bases is open. Since transition functions are smooth, this statement holds independently of the particular coordinate system used. •