Overdetermined PDEs[1]

Maciej Dunajski

Department of Applied Mathematics and Theoretical Physics
University of Cambridge
Wilberforce Road, Cambridge CB3 0WA, UK

June 3, 2008

[1]Please send an email to `m.dunajski@damtp.cam.ac.uk` if you find typos/errors in these notes.

Introduction

This course treats geometric approaches to differential equations (DEs), both ordinary (ODEs) and partial (PDEs). Geometry in this context means that certain results do not depend on coordinate choices made to write down a DE, and also that structures like connection and curvature are associated to DEs.

The subject can get very technical but we shall take a low technology approach. This means that some times, for a sake of explicitness, a coordinate calculation will be performed instead of presenting an abstract coordinate–free argument. We shall also skip some proofs, and replace them by examples illustrating the assumptions and applications. The proofs can be found in [3]. See also [11] and [7]. I shall assume that you have a basic knowledge of differential forms, vector fields and manifolds.

Given a system of DEs it is natural to ask the following questions

- Are there any solutions?
- If yes, how many?
- What data is sufficient to determine a unique solution?
- How to construct solutions?

These are all local questions, i.e. we are only interested in a solution in a small neighbourhood of a point in a domain of definition of dependent variables. We shall mostly work in smooth category, except when a specific reference to Cauchy–Kowalewska theorem is made. This theorem holds only in the real analytic category.

Problem 1. Consider an ODE

$$\frac{du}{dx} = F(x, u) \tag{1}$$

where F and $\partial_u F$ are continuous[1] in some open rectangle

$$U = \{(x, u) \in \mathbb{R}^2, a < x < b, c < u < d\}.$$

The Picard theorem states that for all $(x_0, u_0) \in U$ there exists an interval $I \subset \mathbb{R}$ containing x_0 such that there is a unique function $u : I \to \mathbb{R}$ which satisfies (1) and such that $u(x_0) = u_0$. We say that the general solution to this first order ODE depends on one constant. The unique solution in Picard's theorem arises as a limit

$$u(x) = \lim_{n\to\infty} u_n(x)$$

[1] In fact it is sufficient if F is Lipschitz.

of uniformly convergent sequence of functions $\{u_n(x)\}$ defined iteratively by

$$u_{n+1}(x) = u_0 + \int_{x_0}^{x} F(t, u_n(t))dt.$$

One can treat a system of n first order ODEs with n unknowns in the same way: The unique solution depends on n constants of integration.

The conditions in Picard's theorem always need to be checked and should not be taken for granted. For example the ODE

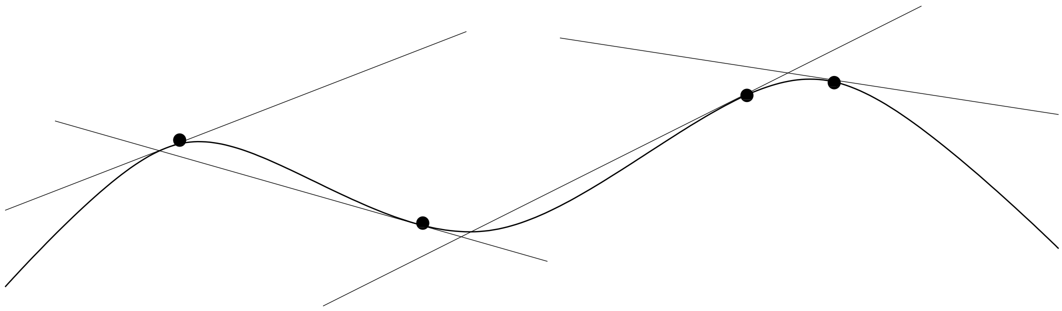
$$\frac{du}{dx} = u^{1/2}, \qquad u(0) = 0$$

has two solutions: $u(x) = 0$ and $u(x) = x^2/4$.

More geometrically, the solutions to (1) are curves tangent to a vector field

$$X = \frac{\partial}{\partial x} + F(x, u)\frac{\partial}{\partial u}.$$

The Picard theorem states that the tangent directions always fit together to form a curve.



One can rephrase this in a language of differential forms. The one–form annihilated by X is (a multiple of) $\theta = du - F dx$ and a parametrised curve $x \to (x, u(x))$ is an integral curve of (1) if θ (or any of its multiples) vanishes on this curve. In general, if θ is a k–form on a manifold M the submanifold $S \subset M$ is an integral of θ if $f^*(\theta) = 0$, where $f : S \to M$ is an immersion.

One aim of this course is to reformulate systems of DEs as vanishing of a set of differential forms (in general of various degree). This gives a coordinate invariant formulation of DEs as Exterior Differential Systems, and allows a discussion of dimension of integral manifold.

Problem 2. Consider a system of PDEs

$$u_x = A(x, y, u), \qquad u_y = B(x, y, u) \tag{2}$$

where $u_x = \partial_x u$ etc. Both derivatives of u are determined at each point $(x, y, u) \in \mathbb{R}^3$ where A, B, A_u, B_u are continuous. This gives rise to a two–dimensional plane spanned by two vectors

$$X_1 = \frac{\partial}{\partial x} + A\frac{\partial}{\partial u}, \qquad X_2 = \frac{\partial}{\partial y} + B\frac{\partial}{\partial u}.$$

Do these planes fit together to form a solution surface in a neighbourhood of (say) $(0, 0, u_0) \in \mathbb{R}^3$? Let us try two successive applications of Picard's theorem.

- Set $y = 0, u(0,0) = u_0$. The Picard theorem guarantees the existence of the unique $\tilde{u}(x)$ such that
$$\frac{d\tilde{u}}{dx} = A(x, 0, \tilde{u}), \qquad \tilde{u}(0) = u_0.$$
- Consider $\tilde{u}(x)$ and hold x fixed, regarding it as a parameter. Picard's theorem gives the unique $u(x,y)$ such that
$$\frac{du}{dy} = B(x, y, u), \qquad u(x, 0) = \tilde{u}(x).$$

We have therefore constructed a function $u(x,y)$ but it may not satisfy the original PDE (2) which is overdetermined and requires that the compatibility condition

$$(u_x)_y = (u_y)_x$$

holds. Expanding the mixed partial derivatives yields

$$A_y - B_x + A_u B - B_u A = 0. \tag{3}$$

Do we need more compatibility conditions arising from differentiating (3) and using (2) to get rid of u_x, u_y ? The answer is no. It follows from the Frobenius theorem which we are going to prove in §1 (the left hand side of (3) is the obstruction to the vanishing of the commutator $[X_1, X_2]$). If (3) holds then solving the pair of ODEs gives the solution surface depending on one constant.

What happens if (3) does not hold?

- If u does not appear in (3) then (3) is a curve in $\mathbb{R}^2$ and there is no solution in an open set containing $(0, 0, u_0)$.
- If (3) gives an implicit algebraic relation between (x, y, u), then solve this relation to get a surface $(x, y) \to (x, y, u(x, y))$. This may or may not be a solution to the original pair of PDEs (2). In particular the initial condition may not be satisfied.

This simple example raises a number of questions. How to deal with more complicated compatibility conditions? When can we stop cross-differentiating? Theorems 2.1 and 2.2 proved in §2 and more generally the Cartan Test 5.5 discussed in §5 give some of the answers.

Problem 3. Consider a system of linear PDEs

$$u_x = \alpha u + \beta v, \qquad u_y + v_x = \gamma u + \delta v, \qquad v_y = \epsilon u + \phi v \tag{4}$$

where $\alpha, \beta, \ldots, \phi$ are some functions of (x, y) defined on an open set $U \subset \mathbb{R}^2$. This is an overdetermined system as there are three equations for two unknowns, but (unlike the system (2)) it is not overdetermined enough, as the partial derivatives are not specified at each point. Therefore we cannot start the process of building the solution surface as we can not specify the tangent planes. One needs to use the process of *prolongation* and introduce new variables for unknown derivatives hoping to express derivatives of these variables using the (differential consequences of) the original system. In our case it is enough to define

$$w = u_y - v_x$$

(there are other choices, for example $w = u_y$, but the solutions surface will not depend on the choices made). Now

$$u_y = \frac{1}{2}(\gamma u + \delta v + w), \quad v_x = \frac{1}{2}(\gamma u + \delta v - w)$$

and we can impose the compatibility conditions

$$(u_y)_x = (u_x)_y, \quad (v_y)_x = (v_x)_y.$$

These conditions will lead to expressions

$$w_x = \ldots, \quad w_y = \ldots$$

where $(\ldots)$ denote terms linear in (u, v, w). The system is now closed as first derivatives of (u, v, w) are determined at each point thus specifying a family of two–dimensional planes in $\mathbb{R}^5$. Do these two planes fit in to form a solution surface

$$(x, y) \longrightarrow (x, y, u(x, y), v(x, y), w(x, y))$$

in $\mathbb{R}^5$? Not necessarily, as there are more compatibility conditions to be imposed (for example $(w_x)_y = (w_y)_x$). These additional conditions will put restrictions of the functions $(\alpha, \beta, \ldots, \phi)$. In the course we shall see how to deal with the prolongation procedure systematically.

This simple example of prolongation arises naturally in geometry of surfaces. Assume you are given a metric (a first fundamental form) on a surface

$$g = Edx^2 + 2Fdxdy + Gdy^2.$$

Does there exist a one form $K = udx + vdy$ such that the Killing equations

$$\nabla_{(i}K_{j)} = 0, \qquad x^i = (x, y)$$

are satisfied, where ∇ is the Levi–Civita connection of g? Expanding the Killing equation in terms of the Christoffel symbols leads to a system (4) where the six functions $(\alpha, \beta, \ldots, \phi)$ are given in terms of E, F, G and their derivatives. The consistency conditions for the prolonged system to admit non–zero solutions give differential constraints on E, F, G. These constrains can be expressed in tensor form as differential invariants of the metric g. In §3.1 we shall discuss an approach to constructing such invariants.

Chapter 1

Exterior Differential System and Frobenius Theorem

Definition 1.1 *An exterior differential system (EDS) is a pair* $(M, \mathcal{I})$ *where* M *is a smooth manifold and* $\mathcal{I} \subset \Omega^*(M)$ *is a graded differential ideal in a ring of differential forms that is closed under exterior differentiation:*

$$d\theta \in \mathcal{I} \qquad \textit{if} \qquad \theta \in \mathcal{I}.$$

For example the set of forms

$$\{dy - pdx,\ dp \wedge dx,\ dx\}$$

forms EDS where $M = \mathbb{R}^3$. We shall use the following notation: $\mathcal{I}^k = \mathcal{I} \cap \Omega^k(M)$ is a set of all forms of degree k in $\mathcal{I}$. The evaluation of a form θ at $x \in M$ will be denoted θ_x and $\mathcal{I}_x$ will denote the evaluation of all forms in $\mathcal{I}$ at x.

One way to present EDS is by specifying the set of differential generators

$$< \theta^1, \ldots, \theta^n >_{\text{diff}} := \{\gamma_1 \wedge \theta^1, \ldots, \gamma_n \wedge \theta^n, \beta_1 \wedge d\theta^1, \ldots, \beta_n \wedge d\theta^n\}$$

where γ, β are arbitrary differential forms. We shall assume that none of the generators are 0–forms (i.e. functions). Otherwise we shall restrict the EDS to submanifolds on which these functions vanish. An EDS whose generators are 1–forms is called a Pfaffian system.

We shall also use the notation

$$< \theta^1, \ldots, \theta^n >_{\text{alg}} := \{\gamma_1 \wedge \theta^1, \ldots, \gamma_n \wedge \theta^n >$$

to denote the set of forms generated algebraically by exterior multiplication.

Definition 1.2 *An integral manifold of* $\mathcal{I}$ *is a submanifold* $f : S \to M$ *such that* $f^*(\theta) = 0$ *for all* $\theta \in \mathcal{I}$*.*

In particular S is an integral submanifold of $\mathcal{I} =< \theta^1, \ldots, \theta^n >_{\text{diff}}$ iff $f^*(\theta^i) = 0$.

- **Example.** A system of N first order ODEs

$$\frac{du^\alpha}{dx} = F^\alpha(x, u^1, \ldots, u^N), \qquad \alpha = 1, \ldots, N$$

is modelled by the EDS $\mathcal{I}$ generated by N 1–forms $< du^\alpha - F^\alpha dx >_{\text{diff}}$ on an open set in $\mathbb{R}^{N+1}$. The integral manifolds of this EDS are integral curves of a vector field

$$X = \frac{\partial}{\partial x} + \sum_{\alpha=1}^{N} F^\alpha \frac{\partial}{\partial u^\alpha}$$

which annihilates all forms in $\mathcal{I}$.

- **Example.** The pair of PDEs $u_x = A(x, y, u), u_y = B(x, y, u)$ is modelled by an ideal generated by one 1-form

  $$I =< du - Adx - Bdy > .$$

  The vectors $\partial_x + A\partial_u, \partial_y + B\partial_u$ annihilating this one form are tangent to the integral surface if one exists. There are no integral surfaces if the compatibility (3) does not hold.

Two EDS $(M, \mathcal{I})$ and $(\hat{M}, \hat{\mathcal{I}})$ are equivalent if there exist a diffeomorphism such that

$$f : M \longrightarrow \hat{M}, \qquad f^*(\hat{\mathcal{I}}) = \mathcal{I}.$$

This notion can be applied to determine whether two systems of DEs are equivalent and, in particular, to linearise some DEs. In the following two examples we shall use the following notation: if $\theta = \theta_j(x^1, \dots, x^i)\, dx^j$ then $\hat{\theta} = \theta_j(\hat{x}^1, \dots, \hat{x}^i)\, d\hat{x}^j$ where x^j and $\hat{x}^j$ are local coordinates on M and $\hat{M}$ respectively.

- **Example.** Consider the Monge–Ampere equation

  $$u_{xx}u_{yy} - u_{xy}^2 = 1,$$

  where $u = u(x, y)$. This nonlinear equation is modelled by the EDS

  $$< \theta^1 = du - pdx - qdy, \theta^2 = dp \wedge dq - dx \wedge dy >_{\text{diff}} \tag{1.1}$$

  on $\mathbb{R}^5$. In particular, it is not a Pfaffian system. Consider $f : \mathbb{R}^5 \to \mathbb{R}^5$ given by

  $$f(x, y, u, p, q) = (\hat{x}, \hat{y}, \hat{u}, \hat{p}, \hat{q}) := (x, q, u - qy, p, -y).$$

  We verify that

  $$f^*(\hat{\theta}^1) = d\hat{u} - \hat{p}d\hat{x} - \hat{q}d\hat{y} = du - pdx - qdy, \qquad f^*(\hat{\theta}^2) = d\hat{p} \wedge d\hat{q} - d\hat{x} \wedge d\hat{y} = dy \wedge dp + dq \wedge dx.$$

  The integral manifolds of the pulled back ideal are

  $$du - pdx - qdy = 0, \qquad dy \wedge dp + dq \wedge dx = 0.$$

  Vanishing of the one form gives $p = u_x, q = u_y$, and vanishing of the two–form gives the linear Laplace equation

  $$u_{xx} + u_{yy} = 0.$$

  Some care needs to be taken with this example: We have established a one-to-one correspondence between integral surfaces of the Laplace equation and the Monge–Ampere equation, but not between solutions as some integral surfaces may have $dx \wedge dy = 0$.

- **Example.** The similar procedure can be used to reduce the general four–dimensional Ricci–flat Kähler metric with a tri–holomorphic Killing vector to the Gibbons–Hawking form where the non–linear Ricci–flat condition reduces to the Laplace equation on $\mathbb{R}^3$. Consider a Kähler metric in an open ball in $\mathbb{C}^2$ with local holomorphic coordinates (w, z) given in terms of the (non–holomorphic) Kähler potential $\Omega : \mathbb{C}^2 \longrightarrow \mathbb{R}$

$$g = \Omega_{w\bar{w}} dw\, d\bar{w} + \Omega_{w\bar{z}} dw\, d\bar{z} + \Omega_{z\bar{w}} dz\, d\bar{w} + \Omega_{z\bar{z}} dz\, d\bar{z}. \tag{1.2}$$

The Ricci–flat condition on g gives a non–linear Monge–Ampere equation on Ω

$$\Omega_{w\bar{w}}\Omega_{z\bar{z}} - \Omega_{w\bar{z}}\Omega_{z\bar{w}} = 1. \tag{1.3}$$

Assume that this metric admits a Killing vector[1] $K = i(\partial_w - \partial_{\bar{w}})$. The Killing equations yield $K(\Omega) = 0$ and the Monge–Ampere equation reduces to

$$\Omega_{vv}\Omega_{z\bar{z}} - \Omega_{vz}\Omega_{v\bar{z}} = 1,$$

where $\Omega = \Omega(z, \bar{z}, v)$ and $v = i(\bar{w} - w) \in \mathbb{R}$. This nonlinear PDE is modelled by the EDS generated by

$$< \theta^1 = d\Omega - pdv - qdz - \bar{q}d\bar{z},\, \theta^2 = dq \wedge dp \wedge dz - dz \wedge d\bar{z} \wedge dv >_{\text{diff}}$$

together with the independence condition $dz \wedge d\bar{z} \wedge dv \neq 0$ on an open set in $\mathbb{R}^7$. Consider

$$f(\Omega, z, v, p, q) = (\hat{\Omega}, \hat{z}, \hat{v}, \hat{p}, \hat{q}) = (\Omega - pv, z, p, -v, q).$$

Vanishing of the forms

$$f^*(\hat{\theta}^1) = d\Omega - pdv - qdz - \bar{q}d\bar{z}, \quad f^*(\hat{\theta}^2) = -dq \wedge dv \wedge dz - dz \wedge d\bar{z} \wedge dp$$

gives the Laplace equation on $\mathbb{R}^3$

$$\Omega_{vv} + \Omega_{z\bar{z}} = 0.$$

In this derivation we assumed non-vanishing of $d\hat{z} \wedge d\bar{\hat{z}} \wedge d\hat{p}$. If this three–form vanishes then $\hat{\Omega}$ is linear in $\hat{v}$ and the Monge–Ampere equations (which hats over all variables) implies that the resulting metric $\hat{g}$ is flat.

Exercise. Implement the change of coordinates at the level of $\hat{g}$ given by (hatted version of) (1.2) to show that it is equivalent to the Gibbons–Hawking form

$$g = Vd\mathbf{x}^2 + V^{-1}(d\tau + A)^2$$

where $\hat{z} = x + iy, \hat{w} = (\tau + iv)/2$, the coordinates (x, y, v, τ) are real, $\mathbf{x} = (x, y, p)$ and (A, V) are a one–form and a harmonic function which satisfy

$$dA = *dV$$

as a consequence of the Laplace equation (here $*$ is the Hodge operator on $\mathbb{R}^3$ with its flat Euclidean metric).

[1] In fact using the freedom $\Omega \to \Omega + \kappa + \bar{\kappa}$ where $\kappa = \kappa(w, z)$ is holomorphic and redefining the holomorphic coordinates $(w, z) \to (\hat{w}(w, z), \hat{z}(w, z))$ one can show that this is the most general form of a Killing vector which Lie–derives the Kähler form and the holomorphic two–form $dw \wedge dz$.

We shall now prove the existence theorem of integral manifolds which applies to ideals generated by one–forms.

Theorem 1.3 (Frobenius - Version 1) *Let $\mathcal{I}$ be a differential ideal generated algebraically by one–forms $\theta^1, \dots, \theta^{n-r}$ on some n–dimensional manifold M such that*

$$d\theta^i = \sum_{j=1}^{n-r} \gamma^i_j \wedge \theta^j \tag{1.4}$$

for some one forms γ^i_j (so that $\mathcal{I}$ is closed). In sufficiently small neighbourhood of a point where θ^i are linearly independent there exists a coordinate system $(y^1, \dots, y^n)$ such that $\mathcal{I}$ is generated by $dy^{r+1}, \dots, dy^n$ and the maximal, r–dimensional integral manifolds are

$$y^{r+1} = const,\ y^{r+2} = const, \quad \dots, \quad y^n = const.$$

Proof. Let W_x =span$(\theta^i|_x) \subset T^*_x M$ and let $W_x{}^\perp \subset T_x M$ be an r–dimensional subspace[2] of vectors annihilating $(\theta^i)_x$.

We shall follow the proof given in [3] and proceed by induction with respect to r. If $r = 1$ then $W^\perp$ is spanned by one vector field X. The Picard existence theorem for ODEs implies the existence of a local coordinate system[3] $y^1, \dots, y^n$ such that $X = \partial/\partial y^1$. Therefore W_x =span$\,(dy^2, \dots, dy^n)$ and we are done. Note that no integrability condition is needed for existence of integral curves so we did not have to use (1.4) which in fact holds identically if $r = 1$.

Now assume that $r > 1$ and suppose that the Theorem holds for $r - 1$ (which is to say that it holds for $(n - r + 1)$ one forms). Let x^i be local coordinates such that the set of one–forms $\mathcal{I}' := \{\theta^1, \dots, \theta^{n-r}, dx^r\}$ is linearly independent. The forms $\theta^1, \dots, \theta^{n-r}$ satisfy the closure condition (1.4) and so this condition is also satisfied by the generators of $\mathcal{I}'$. Therefore, by the inductive hypothesis, there exist coordinates $y^1, \dots, y^n$ such that $dy^r, \dots, dy^n$ span $\mathcal{I}'$ and so $x^r = x^r(y^r, \dots, y^n)$. Assume, without loss of generality, that $\partial x^r/\partial y^r \neq 0$ (no summation!) and solve the relation

$$dx^r = \frac{\partial x^r}{\partial y^r} dy^r + \sum_{i=1}^{n-r} \frac{\partial x^r}{\partial y^{r+i}} dy^{r+i}$$

for dy^r. The one–forms θ^i are in the span of $dy^r, \dots, dy^n$. Therefore, substituting for dy^r, we get

$$\theta^i = b^i dx^r + \sum_{j=1}^{n-r} a^i_j dy^{j+r}, \quad i = 1, \dots, n - r.$$

The forms θ^i and dx^r are linearly independent so the matrix $(a^i{}_j)$ is non–singular, or otherwise $\sum_i V_i(\theta^i - b^i dx^r) = 0$ for some $V \in \ker(a)$. Thus $a^{-1}\theta$ gives a new set of generators

$$\tilde{\theta}^i = dy^{r+i} + p^i dx^r, \quad i = 1, \dots, n - r.$$

[2] The integral manifolds in the Frobenius theorem are leaves of r–dimensional foliation of M by a distribution $W^\perp := \cup_x W_x{}^\perp \subset TM$.

[3] To see it set $X = \partial/\partial y^1$ at $x = (0, 0, \dots, 0)$. Then, there is a unique integral curve through each point $(0, a^2, \dots, a^n)$. If a point x lies on the integral curve through this point we can use $(y^2, \dots, y^n)$ as the last $(n - 1)$ coordinates of x and the time interval it takes the curve to get to x as the first coordinate.

The closure condition (1.4) gives

$$d\tilde{\theta}^i = dp^i \wedge dx^r = \sum_{k=1}^{r-1} \frac{\partial p^i}{\partial y^k} dy^k \wedge dx^r = 0 \quad \mathrm{mod} \quad \tilde{\theta}^i$$

(recall that dy^r is a combination of dx^r and dy^{r+i} so it does not appear in the summation). Therefore

$$p^i = p^i(y^r, y^{r+1}, \dots, y^n)$$

and the $(n-r)$ forms $\theta^1, \dots, \theta^{n-r}$ satisfy the Frobenius condition (1.4) in $(n-r+1)$ coordinates. This case corresponds to $r = 1$ and was dealt with at the beginning of the proof.

□

We shall now give two more formulations of the Frobenius theorem. One in terms of vector fields, and one in terms of overdetermined PDEs.

Assume that the Frobenius condition holds and extend the ideal $\mathcal{I}$ to a basis

$$\theta^1, \dots, \theta^{n-r}, \theta^{n-r+1}, \dots, \theta^n$$

of T_x^*M so that

$$d\theta^i = \frac{1}{2} \sum_{j,k=1}^{n} C^i_{jk} \theta^j \wedge \theta^k, \qquad i = 1, \dots, n$$

for some C^i_{jk}. The closure condition (1.4) is equivalent to

$$C^m_{pq} = 0, \qquad m = 1, \dots, n-r, \quad p, q = (n-r+1), \dots, n.$$

Define the dual basis X_i of T_xM by

$$df = \sum_{i=1}^{n} X_i(f)\theta^i,$$

where f is any function on M. Differentiating this relation gives

$$0 = d^2 f = \sum_{i,j} X_j(X_i(f))\theta^j \wedge \theta^i + \frac{1}{2} \sum_{i,j,k} X_i(f) C^i_{jk} \theta^j \wedge \theta^k,$$

and finally

$$[X_p, X_q] = -C^s_{pq} X_s, \qquad p, q, s = (n-r+1), \dots, n$$

where the vectors $\{X_{n-r+1}, \dots, X_n\}$ span the distribution $W^\perp$. However the same distribution is spanned by $\{\partial/\partial y^1, \dots, \partial/\partial y^r\}$ which gives

Theorem 1.4 (Frobenius - Version 2) *Let $\{X_{n-r+1}, \dots, X_n\}$ be an r–dimensional distribution on M such that*

$$[X_p, X_q] = -C^s_{pq} X_s, \qquad p, q, s = (n-r+1), \dots, n. \tag{1.5}$$

In sufficiently small neighbourhood of a point where X_i are linearly independent there exists a coordinate system $(y^1, \dots, y^n)$ such that

$$\mathit{span}\{X_{n-r+1}, \dots, X_n\} = \mathit{span}\{\partial/\partial y^1, \dots, \partial/\partial y^r\}.$$

For the last formulation of the Frobenius theorem consider a system of PDEs

$$\frac{\partial u^\rho}{\partial x^i} = \psi_i^\rho(x, u), \qquad i = 1, \dots, n, \quad \rho = 1, \dots, N, \tag{1.6}$$

where $u : \mathbb{R}^n \longrightarrow \mathbb{R}^N$. We want to construct a solution through each point

$$(x^1, \dots, x^n, u^1, \dots u^N) \in \mathbb{R}^{n+N}.$$

This is the same as constructing a foliation of $\mathbb{R}^{n+N}$ by n–dimensional integral surfaces of the ideal generated by

$$< \theta^\rho = du^\rho - \psi_i^\rho dx^i >_{\text{diff}}, \qquad \rho = 1, \dots, N.$$

The annihilator $W^\perp$ of this ideal is spanned by the vector fields

$$X_i = \frac{\partial}{\partial x^i} + \sum_\rho \psi_i^\rho \frac{\partial}{\partial u^\rho}, \qquad i = 1, \dots, n.$$

The Frobenius integrability condition

$$[X_i, X_j] = 0$$

gives the necessary and sufficient condition for the existence of the integral manifolds. Note that in this case the commutators must vanish exactly as there is no way of generating $\partial/\partial x^i$ on the RHS of the commutator. Expanding the commutators yields

Theorem 1.5 (Frobenius - Version 3) *The necessary and sufficient conditions for the unique solution $u^\alpha = u^\alpha(x)$ to the system* (1.6) *such that $u(x_0) = u_0$ to exist for any initial data $(u_0, x_0) \in \mathbb{R}^{n+N}$ is that the relations*

$$\frac{\partial \psi_i^\alpha}{\partial x^j} - \frac{\partial \psi_j^\alpha}{\partial x^i} + \sum_\beta \Big(\frac{\partial \psi_i^\alpha}{\partial u^\beta} \psi_j^\beta - \frac{\partial \psi_j^\alpha}{\partial u^\beta} \psi_i^\beta \Big) = 0, \qquad i, j = 1, \dots, n, \quad \alpha, \beta = 1, \dots, N \tag{1.7}$$

hold.

- **Example.** The one–form

  $$\theta = du - A(x, y, u)\, dx - B(x, y, u)\, dy$$

  in $\mathbb{R}^3$ satisfies (1.4) iff

  $$d\theta = \gamma \wedge \theta$$

  for some one–form γ, or, equivalently, iff

  $$\theta \wedge d\theta = 0.$$

  This condition holds iff the compatibility condition (3) for a pair of overdetermined PDEs $u_x = A, u_y = B$ are satisfied. The Frobenius theorem implies that in this case $\theta = \mu\, df$ where μ, f are some functions of (x, y, u) and that $f = \text{const}$ is the solution surface in $\mathbb{R}^3$.

- **Example.** Another simple application of the Frobenius theorem is used in General Relativity. Any metric g with a Killing vector K on an n dimensional manifold can locally be written as

$$g = Vh + V^{-1}(d\tau + A)^2$$

where $(\tau, x^1, \ldots, x^{n-1})$ is a local coordinate system such that $K = \partial/\partial\tau$ and

$$V = V(x), \quad A = A_i(x)\, dx^i, \quad h = h_{ij}\,(x) dx^i\, dx^j.$$

Moreover in the twist–free case $K \wedge dK{=}0$ one can redefine the coordinates, the function V and the metric h to set $A = 0$ (we follow the usual abuse of notation and denote the vector K and the one–form $g(K, \ldots)$ by the same symbol).

Chapter 2

Involutivity

Any system of DEs can be rewritten as a system of algebraic equations on a manifold where higher derivatives are regarded as independent variables. This idea is formalised by the apparatus of jet spaces. Let $u : \mathbb{R}^n \longrightarrow \mathbb{R}^N$, so that we can write $u = u^\alpha(x^i)$. The space of k-jets $J^k(\mathbb{R}^n, \mathbb{R}^N)$ is the space of Taylor polynomials of u of degree k. It is a smooth manifold of dimension

$$n + N\binom{n+k}{k}$$

with local coordinates

$$\{x^i, u^\alpha, p_i^\alpha, p_{ij}^\alpha, \ldots, p_{i_1 i_2 \ldots i_k}^\alpha\}, \quad \alpha = 1, \ldots, N, \ i = 1, \ldots, n.$$

Any map $u : \mathbb{R}^n \longrightarrow \mathbb{R}^N$ can be lifted to a k–graph of u (a section of the jet bundle $J^k(\mathbb{R}^n, \mathbb{R}^N) \to \mathbb{R}^n$) by

$$u^\alpha = u^\alpha(x), \quad p_i^\alpha = \frac{\partial u^\alpha}{\partial x^i}(x), \quad \ldots, \quad p_{i_1 i_2 \ldots i_k}^\alpha = \frac{\partial^k u^\alpha}{\partial x^{i_1} \partial x^{i_2} \ldots \partial x^{i_k}}(x).$$

The system of r k-th order PDEs

$$F^\rho\Big(x^i, u^\alpha, \frac{\partial u^\alpha}{\partial x^i}, \ldots, \frac{\partial^k u^\alpha}{\partial x^{i_1} \partial x^{i_2} \ldots \partial x^{i_k}}\Big) = 0, \qquad \rho = 1, \ldots, r \tag{2.1}$$

gives a submanifold $M^{(k)}$ of co–dimension r in $J^k(\mathbb{R}^n, \mathbb{R}^N)$ and a k–graph of the solution to (2.1) is an n–dimensional integral submanifold $S \subset M^{(k)}$ of the ideal associated to (2.1) such that $dx^1 \wedge dx^2 \ldots \wedge dx^n \neq 0$ on S. The $(k+1)$st graph of the solution lies in a manifold $M^{(k+1)} \subset J^{k+1}(\mathbb{R}^n, \mathbb{R}^N)$ called a prolongation of $M^{(k)}$. The manifold $M^{(k+1)}$ is defined as a zero locus

$$F^\rho = 0, \quad \frac{dF^\rho}{dx^i} = 0, \qquad \rho = 1, \ldots, r, \quad i = 1, \ldots, n$$

in $J^{k+1}(\mathbb{R}^n, \mathbb{R}^N)$.

For any integer $l \geq 0$ define the family of projections

$$\pi_l : J^{l+1}(\mathbb{R}^n, \mathbb{R}^N) \longrightarrow J^l(\mathbb{R}^n, \mathbb{R}^N)$$

by

$$\pi_l(x^i, u^\alpha, p_i^\alpha, \ldots, p_{i_1 i_2 \ldots i_l}^\alpha, p_{i_1 i_2 \ldots i_l i_{l+1}}^\alpha) = (x^i, u^\alpha, p_i^\alpha, \ldots, p_{i_1 i_2 \ldots i_l}^\alpha).$$

Therefore $\mathrm{Im}(M^{(k+1)}) \subset M^{(k)}$ (this is obvious as $F^\rho = 0$ holds on $M^{(k+1)}$) but π_k does not have to be surjective: differentiating the PDEs (2.1), mixing partial derivatives, and using (2.1) gives rise to new PDEs of order lower than k. So the image of $M^{(k+1)}$ under π_k will in general be a submanifold of $M^{(k)}$ of some non–zero co–dimension. Therefore the k–jets of a solutions do not have to extend to $(k+1)$ jets. We keep differentiating and adding lower order conditions restricting $M^{(k)}$. When can we stop this process? The combined system of equations and lower order conditions must be involutive. In general one needs Cartan test which will be discussed in §5. Theorems 2.1 and 2.2 which we will prove in this section answer this question for systems of 1st order PDEs (1.6)

$$\frac{\partial u^\rho}{\partial x^i} = \psi^\rho_i(x, u), \qquad i = 1, \dots, n, \quad \rho = 1, \dots, N.$$

If the Frobenius integrability conditions (1.7) hold, the general solution of (1.6) depends on N arbitrary constants. Otherwise (1.7) give a set of algebraic equations

$$F_1(u, x) = 0$$

which must be satisfied by any solution to (1.6). Differentiating these equations and eliminating the derivatives of u using (1.6) leads to a new set of equations

$$F_2(u, x) = 0.$$

Proceeding in this way we get a sequence of sets of equations

$$F_1(u, x) = 0, \quad F_2(u, x) = 0, \quad F_3(u, x) = 0, \dots \ .$$

If the system (1.6) admits a solution there must be an integer K such that the equations in the set $F_{K+1} = 0$ are satisfied as a consequence of the equations in the first K sets. Otherwise we would obtain more than N independent conditions on $(u^1, \dots, u^N)$ which would imply a relation between the independent variables. In particular we must have $K \leq N$. This proves the 'only if' statement in the following

Theorem 2.1 *The system* (1.6) *admits solution if and only if there exists a positive integer $K \leq N$ such that the set of algebraic equations*

$$F_1 = F_2 = \dots = F_K = 0$$

is compatible for all $x \in U \subset \mathbb{R}^n$ and that the set $F_{K+1} = 0$ is satisfied identically. If p is the number of independent equations in the first K sets, then the general solution depends on $(N - p)$ arbitrary constants.

Proof. It remains to prove the 'if' part. We follow the classical treatment given for example in [12]. Assume that the first K independent sets impose $p < N$ independent conditions

$$G_\nu(u, x) = 0, \qquad \nu = 1, \dots, p. \tag{2.2}$$

Therefore

$$\mathrm{rank}\Big(\frac{\partial G_\nu}{\partial u^\alpha}\Big) = p$$

and, by implicit function theorem, the relations (2.2) can be solved for (say) the first p functions $u^1, \dots, u^p$

$$u^\lambda = \phi^\lambda(u^{p+1}, \dots, u^N, x), \qquad \lambda = 1, \dots, p.$$

Differentiate this and use (1.6) to eliminate the derivatives

$$\psi_i^\lambda - \sum_{\nu=p+1}^{N} \frac{\partial \phi^\lambda}{\partial u^\nu} \psi_i^\nu - \frac{\partial \phi^\lambda}{\partial x^i} = 0.$$

These equations belong to the set $F_{K+1} = 0$ so they hold by assumption. We rewrite the above equations substituting $\psi_i^\lambda = \partial u^\lambda / \partial x^i$ and subtracting:

$$\frac{\partial u^\lambda}{\partial x^i} - \psi_i^\lambda - \sum_{\nu=p+1}^{N} \frac{\partial \phi^\lambda}{\partial u^\nu} \Big(\frac{\partial u^\nu}{\partial x^i} - \psi_i^\nu \Big) = 0$$

so

$$\frac{\partial u^\nu}{\partial x^i} = \overline{\psi}_i^\nu (u^{p+1}, \dots, u^N, x) \tag{2.3}$$

where $\nu = p+1, \dots, N$ and

$$\overline{\psi}_i^\nu = \psi_i^\nu|_{u^\lambda = \phi^\lambda(u^{p+1}, \dots, u^N, x)}.$$

The system (2.3) is Frobenius integrable as the consistency belongs to the set

$$F_1 = \dots = F_K = 0$$

so, by the Frobenius Theorem 1.5, there is a solution which involves $(N-p)$ constants.

□

In many applications the functions ψ_i^α in (1.6) are linear and homogeneous in u^ρ. This allows the following geometric interpretation of the last theorem. Let us write the system of linear homogeneous PDEs

$$\frac{\partial u^\rho}{\partial x^i} = \psi_{\gamma i}^\rho(x) u^\gamma$$

as

$$d\mathbf{u} + \mathbf{\Omega u} = 0 \tag{2.4}$$

where $\mathbf{u} = (u^1, \dots, u^N)^T$ and $\mathbf{\Omega} = -\psi_{\gamma i}^\rho dx^i$ is a matrix valued one–form on an open set $U \subset \mathbb{R}^n$. Therefore solutions to (2.4) correspond to parallel sections $\mathbf{u} : U \to \mathbb{E}$ of a rank N vector bundle $\mathbb{E} \to U$ with connection $D = d + \mathbf{\Omega}$. Locally the total space of this bundle is an open set in $\mathbb{R}^{n+N}$. To simplify notation let us assume that $n = 2$ and (x^1, x^2) are local coordinates in $U \subset \mathbb{R}^2$.

Differentiating (2.4) and eliminating $d\mathbf{u}$ yields $\mathbf{Fu} = 0$, where

$$\begin{aligned} \mathbf{F} &= d\mathbf{\Omega} + \mathbf{\Omega} \wedge \mathbf{\Omega} = (\partial_1 \mathbf{\Omega}_2 - \partial_2 \mathbf{\Omega}_1 + [\mathbf{\Omega}_1, \mathbf{\Omega}_2]) dx^1 \wedge dx^2 \\ &= F dx^1 \wedge dx^2 \end{aligned}$$

is the curvature of D. Thus we need

$$F\mathbf{u} = 0, \tag{2.5}$$

where $F = F(x^1, x^2)$ is an N by N matrix. This is the first set of conditions $F_1 = 0$ in Theorem 2.1. In this case these conditions are just linear homogeneous equations. If $F = 0$ and the connection is flat, there exist N–independent parallel sections. In this case the Frobenius integrability conditions (1.7) hold. On the other hand if $\det(F) \neq 0$ then no non–zero parallel sections exists.

In general we want to determine the dimension of the space of parallel sections. To achieve it, differentiate the condition (2.5) and use (2.4) to obtain

$$0 = dF\mathbf{u} - F\mathbf{\Omega}\mathbf{u} = \Big((\partial_i F - F\mathbf{\Omega}_i)\mathbf{u}\Big)dx^i$$

Using $F\mathbf{u} = 0$ we rewrite this as

$$(D_i F)\mathbf{u} = 0,$$

where $D_i F = \partial_i F + [\mathbf{\Omega}_i, F]$.

We continue differentiating to produce algebraic matrix equations

$$F\mathbf{u} = 0, \quad (D_i F)\mathbf{u} = 0, \quad (D_i D_j F)\mathbf{u} = 0, \quad (D_i D_j D_k F)\mathbf{u}, \quad \dots .$$

These are the conditions $F_1 = 0, F_2 = 0, F_3 = 0, \dots$ in Theorem 2.1. After K differentiations this leads to $r(K)$ linear equations which we write as

$$\mathcal{F}_K \mathbf{u} = 0,$$

where $\mathcal{F}_K$ is a $r(K)$ by N matrix. We also set $\mathcal{F}_0 = F$. Theorem 2.1 adapted to (2.5) and (2.4) tells us when we can stop the process

Theorem 2.2 *Assume that the ranks of the matrices $\mathcal{F}_K, K = 0, 1, 2, \dots$ are maximal and constant*[1]*. Let K_0 be the smallest natural number such that*

$$rank\,(\mathcal{F}_{K_0}) = rank\,(\mathcal{F}_{K_0+1}). \tag{2.6}$$

If K_0 exists then $rank(\mathcal{F}_{K_0}) = rank(\mathcal{F}_{K_0+k})$ for $k \in \mathbb{N}$ and the space of parallel sections (2.4) of $d + \mathbf{\Omega}$ has dimension $(N - rank(\mathcal{F}_{K_0}))$.

Thus if the curvature of $(\mathbb{E}, D)$ does not vanish, the non–zero solutions to the system of linear PDEs can exist if the holonomy D on the lies in some proper subgroup of $GL(N, \mathbb{R})$.

[1] This can always be achieved by restricting to a sufficiently small neighbourhood of some point $x \in U$.

Chapter 3

Prolongation

The theorems presented in the last section apply to systems of 1st order PDEs. Given an arbitrary system of PDEs we could aim to represent it as a first order system on a jet space of higher dimension by introducing new variables for second, and higher derivatives. This process will however lead to systems where not all first derivatives are determined (compare the system (4) in the Introduction) and Theorems 1.5 and 2.1 can not be applied to construct the solution surfaces. The idea of prolongation is to introduce more new variables for unknown derivatives aiming to express derivatives of these variables using the (differential consequences of) the original system. Apriori it is not clear that this process will work (i.e. the process of adding new variables may never terminate). The relevant theorems which state under what circumstances the prolongation works were, in case of linear PDEs, given independently by Spencer, Kuranishi and Goldschmidt. See Chapter 5 of [3] for complete exposition of these ideas.

Let $P : E_1 \longrightarrow E_2$ be a linear kth order differential operator between two smooth vector bundles over a manifold M. In local coordinates

$$P(v) = a^{i_1 i_2 \dots i_k} \frac{\partial^k v}{\partial x^{i_1} \partial x^{i_2} \dots \partial x^{i_k}} + \dots$$

where $(\dots)$ denote the lower order terms. The leading term $a^{i_1 i_2 \dots i_k}$ transforms as a tensor under the change of coordinates and gives rise to a bundle map called a symbol of P

$$\sigma(P) : \odot^k \Lambda^1(M) \otimes E_1 \longrightarrow E_2.$$

Thus the symbol is a matrix whose components are polynomials homogeneous of degree k:

$$\sigma(P) = (a^{i_1 i_2 \dots i_k} \xi_{i_1} \xi_{i_2} \dots \xi_{i_k})^\beta_\alpha, \qquad \alpha = 1, \dots, \text{rank}\,(E_1),\ \ \beta = 1, \dots, \text{rank}\,(E_2).$$

For any integer $s \geq k$ define the vector spaces

$$V_s := (\odot^k \Lambda^1(M) \otimes E_1) \cap (\odot^{(s-k)} \Lambda^1(M) \otimes \text{Ker}(\sigma(P))).$$

The system (P, E_1, E_2) is said to be of *finite type* if $V_s = 0$ for s sufficiently large. The seminal result of Spencer [10] is that for systems of finite type the equation

$$P(v) = 0$$

is equivalent to a closed system of PDEs of the form (1.6), where all partial derivatives of the dependent variables are determined. The criterion for a given system to be of finite type is given in [10], but in practice it can be difficult to implement, as the vector spaces V_s can not be easily constructed. For systems not of finite type the process of adding new variables and cross differentiating the equations will never end.

In the last Section we explained how to regard this closed system as a vector bundle $\mathbb{E}$ with a connection D. In the work of Spencer the bundle $\mathbb{E}$ arises as a direct sum $\oplus_s V_s$. Theorem 2.2 can be adapted to the systems of finite type:

Theorem 3.1 *For systems of finite type there exists a vector bundle $\mathbb{E} \to M$ with a connection D and a bijection*

$$\{v \in \Gamma(E_1) \text{ such that } P(v) = 0\} \to \{\mathbf{u} \in \Gamma(\mathbb{E}),\, D\mathbf{u} = 0\}.$$

The dimension of the kernel of P is bounded by the rank of $\mathbb{E}$.

The determined system of equations for $D\mathbf{u} = 0$ is the prolongation of the system $P(v) = 0$. Theorem 2.2 can now be applied to give an algorithm for calculating the dimension of the kernel of D. In many geometric applications, where P is build out of covariant derivatives of some connection on TM, the bundle with connection $(\mathbb{E}, D)$ is called the tractor bundle.

- **Example.** Let (M, g) be an n–dimensional (pseudo) Riemannian manifold and let ∇ be a Levi–Civita connection of g. The Killing equations

$$\nabla_{(i}v_{j)} = 0 \tag{3.1}$$

can be put into the framework described in this section with

$$E_1 = \Lambda^1(M), \quad E_2 = \Lambda^1(M) \odot \Lambda^1(M).$$

The system (3.1) is equivalent to the first order system

$$\begin{aligned} \nabla_i v_j &= \mu_{ij} \\ \nabla_i \mu_{jk} &= R_{jki}{}^m v_m \end{aligned}$$

where μ_{ij} is anti–symmetric, $R_{jki}{}^m$ is the Riemann curvature of g and we arrived at the second equation by using $\nabla_{[i}\mu_{jk]} = 0$ and commuting the covariant derivatives on v. We combine (v_i, μ_{ij}) into a section

$$\mathbf{u} = \left(\begin{array}{c} v_i \\ \mu_{ij} \end{array}\right)$$

of a vector bundle $\mathbb{E} = \Lambda^1(M) \oplus \Lambda^2(M)$ with connection D

$$\left(\begin{array}{c} v_i \\ \mu_{ij} \end{array}\right) \stackrel{D_i}{\longmapsto} \left(\begin{array}{c} \nabla_i v_j - \mu_{ij} \\ \nabla_i \mu_{jk} - R_{jki}{}^m v_m \end{array}\right). \tag{3.2}$$

The solutions of the Killing equation (3.1) are in one–to–one correspondence with parallel sections of D. The number of these parallel sections does not exceed

$$\text{rank}(\mathbb{E}) = \text{rank}(\Lambda^1) + \text{rank}(\Lambda^2) = \frac{n(n+1)}{2}.$$

This upper bound is also the dimension of the Lie algebra of the orthogonal group. It is achieved for spaces of constant curvature.

- **Example.** The Cauchy–Riemann equations

$$f_x = g_y, \qquad f_y = -g_x$$

where f and g are functions of (x, y) are not of finite type (the reader is invited to try first few iterations of the prolongation procedure). In fact no uniqueness result analogous to Theorem 2.2 is expected to hold. The general solution to the CR equations depends on one holomorphic function of $(x + iy)$ rather than on a finite number of constants.

3.1 Differential invariants

The prolongation procedure together with Theorem 2.2 gives a straightforward algorithm of constructing invariants which obstruct existence of certain geometric structures: We shall look at two examples: a relatively simple (but sufficiently nontrivial!) example of Killing equations in Riemannian geometry and more involved problem of existence of metric connections in a given projective class [5]. Our treatment of the subject is based on restricting a holonomy of a connection of some vector bundle. The more common principal bundle approach (due to Cartan) is used in [3].

- **Question**. Let g be a (pseudo) Riemannian metric on an open set U in $\mathbb{R}^2$. When is g a metric on a surface of revolution?

Any metric on surface of revolution takes a form

$$g = dx^2 + f(x)dy^2$$

in some coordinates where $f = f(x)$ is a non–vanishing function of one variable. This metric admits a Killing vector $v = \partial/\partial y$. Conversely, the existence of a non-trivial solution to the Killing equations (3.1) guarantees the existence of this coordinate system. Therefore an equivalent form of the question is: When does a metric on a surface admits a solution to (3.1)? The answer must have been known to the classical differential geometers in 19th century: Darboux states it in his book [6]. We shall give the answer as a vanishing of two weighted scalar invariants constructed out of g: one invariant of order 4 and one invariant of order 5.

The metrics of constant curvature admit three Killing vectors (which is the maximal number). The following theorem (also known to Matveev and proved in [8] using different methods) applies to metrics with non-constant curvature.

Theorem 3.2 *A (pseudo) Riemannian metric g with non–constant scalar curvature R admits a Killing vector in a neighbourhood of a point $x \in U$ such that $dR \neq 0$ at x iff*

$$I_1 := dR \wedge d(|R|^2) = 0, \qquad I_2 := dR \wedge d(\triangle(R)) = 0, \tag{3.3}$$

where

$$|R|^2 = g^{ij}\nabla_i R \nabla_j R, \qquad \triangle(R) = g^{ij}\nabla_i\nabla_j R.$$

Proof. Solutions to the Killing equation (3.1) are in one to one correspondence with parallel sections of the connection (3.2). We want to find necessary and sufficient conditions for the existence of one such section. The prolongation procedure simplifies in two dimensions. Firstly any two–form is a multiple of a (chosen) volume form, thus we can write

$$\mu_{ij} = |g|^{1/2} \varepsilon_{ij}\, \mu$$

where $|g| = |\det g|$ for some section of the canonical bundle μ. Moreover the Riemann tensor is determined by the scalar curvature R

$$R_{ijkl} = \frac{R}{2}(g_{ik}g_{jl} - g_{jk}g_{il}).$$

With these simplifications the connection (3.2) reduces to a connection D on a rank three vector bundle $\mathbb{E} \to U$

$$\begin{pmatrix} v_j \\ \mu \end{pmatrix} \xmapsto{D_i} \begin{pmatrix} \nabla_i v_j - |g|^{1/2}\varepsilon_{ij}\, \mu \\ \nabla_i \mu - \frac{1}{2}|g|^{-3/2}\varepsilon_i^{\ j} R\, v_j \end{pmatrix},$$

Using $\nabla_{[i}\nabla_{j]}\mu = 0$ and elliminating the first derivatives of $\mathbf{u} = (v_i, \mu)^T$ gives

$$v^i \nabla_i R = 0, \tag{3.4}$$

where $v^i = g^{ij} v_j$. This is the condition (2.5) leading to Theorem 2.2 where the curvature of the tractor connection is given by a 3 by 3 matrix of rank one

$$F = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \nabla^1 R & \nabla^2 R & 0 \end{pmatrix}.$$

Differentiating (3.4), or equivalently differentiating the tractor curvature F covariantly with respect to D gives two more conditions

$$|g|^{3/2}(\nabla^i R)\mu + \varepsilon^{ji}(\nabla_j \nabla_k R) v^k = 0. \tag{3.5}$$

Therefore the determinant of a 3 by 3 matrix

$$\begin{pmatrix} \nabla_1 R & \nabla_2 R & 0 \\ -\nabla_2\nabla_1 R & -\nabla_2\nabla_2 R & |g|^{3/2}\nabla^1 R \\ \nabla_1\nabla_1 R & \nabla_1\nabla_2 R & |g|^{3/2}\nabla^2 R \end{pmatrix} \tag{3.6}$$

should vanish for non–zero parallel sections of $(\mathbb{E}, D)$ to exist. Calculating this determinant yields the first obstruction I_1 in (3.3). This is the necessary condition for the existence of a Killing vector. Assume that this condition holds. The rank of the matrix (3.6) has to be smaller than three. It is equal to zero if the scalar curvature R is constant. In this case tractor connection is flat. Otherwise, in a neighbourhood of a point where $\nabla_i R \neq 0$, the rank is equal to two and constant. Theorem 2.2 implies that the sufficient conditions are obtained by demanding that the rank of the 6 by 3 matrix obtained from the matrix (3.6) and the second derivatives of (3.4) does not go up and is equal to two. This could a priori lead to three additional obstructions. However only one of them is a new condition and the other two follow as differential consequences of (3.4). To see it write the first algebraic obstruction (3.4) as

$$\mathbf{V} \cdot \mathbf{u} = 0,$$

where $\mathbf{V} = (\nabla_1 R, \nabla_2 R, 0)$. Let $\mathbf{V}_{ij...k}$ denote the vector in $\mathbb{R}^3$ orthogonal to $\mathbf{u}$ which is obtained by eliminating the derivatives of $\mathbf{u}$ from $\partial_i \partial_j ... \partial_k(\mathbf{V} \cdot \mathbf{u}) = 0$. Vanishing of the first obstruction (3.4) implies the linear dependence condition

$$c\mathbf{V} + c_1 \mathbf{V}_1 + c_2 \mathbf{V}_2 = 0 \tag{3.7}$$

for some functions c, c_1, c_2 on U. Assume that we add one more condition

$$e\mathbf{V} + e_1 \mathbf{V}_1 + e_2 \mathbf{V}_2 + e_{12} \mathbf{V}_{12} = 0$$

for some functions $(e, \dots, e_{12})$ on U. This gives an obstruction $I_2 := \det(\mathbf{V}, \mathbf{V}_i, \mathbf{V}_{12}) = 0$ where i equals to 1 or 2 (there is only one obstruction because of the earlier linear dependence condition). Now differentiating (3.7) with respect to x^i and using $\mathbf{V}_{12} = \mathbf{V}_{21}$ which holds modulo lower order terms, implies that $\mathbf{V}_{11}$ and $\mathbf{V}_{22}$ are in the span of $\mathbf{V}, \mathbf{V}_1, \mathbf{V}_2$ and no additional conditions need to be added. To write the second obstruction I_2 we could differentiate (3.5) and take a determinant of one of resulting 3 by 3 matrices. Alternatively we can take Laplacian of (3.4) and eliminate the first derivatives of $\mathbf{u}$. This leads to linear dependence of dR and $d(\triangle(R))$ which is equivalent to vanishing of I_2 in (3.3). Both methods lead to obstructions of differential order 5 in components of the metric g. The argument presented above shows that the resulting sets of obstructions are equivalent. This completes the proof.

$\square$

We shall give one more example of using the prolongation procedure and Theorem 2.2 to produce differential invariants. This time two iterations of the prolongation procedure will be needed to close the system. The aim is to answer the following

- **Question.** Cover a two–dimensional plane with curves, one curve through each point in each direction. How can you tell whether these curves are the geodesics of some metric?

This is an old problem which goes back at least to the work of Roger Liouville [9] in 1887. The soultion was given in [5]. The following discussion summarises the main results. Assume that the curves are presented to us as integral curves of a second order ODE

$$\frac{d^2y}{dx^2} = \Lambda\Big(x, y, \frac{dy}{dx}\Big).$$

Thus we want to find conditions on the ODE so that its integral curves are unparametrised geodesics of some metric connection. First of all they need to be geodesics of some symmetric connection with Christoffel symbols Γ_{ij}^k. Eliminating the parameter t between the geodesic equations

$$\ddot{x}^i + \Gamma^i_{jk} \dot{x}^j \dot{x}^k \sim \dot{x}^i, \qquad x^i = x^i(t)$$

with $(x^1, x^2) = (x, y)$ yields a second order ODE of the form

$$\frac{d^2y}{dx^2} = A_0(x,y) + A_1(x,y)\frac{dy}{dx} + A_2(x,y)\Big(\frac{dy}{dx}\Big)^2 + A_3(x,y)\Big(\frac{dy}{dx}\Big)^3, \tag{3.8}$$

where

$$A_0 = -\Gamma^2_{11}, \quad A_1 = \Gamma^1_{11} - 2\Gamma^2_{12}, \quad A_2 = 2\Gamma^1_{12} - \Gamma^2_{22}, \quad A_3 = \Gamma^1_{22}\,.$$

Conversely, any ODE of the form (3.8) defines an equivalence class of connections which share the same unparametrised geodesics. Thus

$$\frac{\partial^4\Lambda}{\partial(y')^4} = 0$$

is the first necessary condition for metricity of paths. One can check that this condition is invariant under the coordinate transformations $(x, y) \to (\hat{x}(x, y), \hat{y}(x, y))$.

Now assume that there exists a (pseudo) Riemannian metric

$$g = Edx^2 + 2Fdxdy + Gdy^2$$

such that Γ^k_{ij} is the Levi–Civita connection of g. Following R. Liouville [9] introduce a 2 by 2 matrix

$$\sigma^{ij} = \left(\begin{array}{cc} \psi_1 & \psi_2 \\ \psi_2 & \psi_3 \end{array}\right),$$

where

$$E = \psi_1/\Delta, \quad F = \psi_2/\Delta, \quad G = \psi_3/\Delta, \quad \Delta = (\psi_1\psi_3 - {\psi_2}^2)^2.$$

Calculating the Levi–Civita connection in terms of the ψs shows that the integral curves of the ODE (3.8) are metrisable on a neighbourhood of a point $x \in U$ iff there exists σ^{ij} such that $\det(\sigma) \neq 0$ does not vanish at x and following set of equations holds[1]

$$\begin{aligned}
\frac{\partial\psi_1}{\partial x} &= \frac{2}{3}A_1\psi_1 - 2A_0\psi_2, \\
\frac{\partial\psi_3}{\partial y} &= 2A_3\psi_2 - \frac{2}{3}A_2\psi_3, \\
\frac{\partial\psi_1}{\partial y} + 2\frac{\partial\psi_2}{\partial x} &= \frac{4}{3}A_2\psi_1 - \frac{2}{3}A_1\psi_2 - 2A_0\psi_3, \\
\frac{\partial\psi_3}{\partial x} + 2\frac{\partial\psi_2}{\partial y} &= 2A_3\psi_1 - \frac{4}{3}A_1\psi_3 + \frac{2}{3}A_2\psi_2.
\end{aligned} \tag{3.9}$$

We need to prolong this system and look for integrability conditions, but let us first rewrite the system in more invariant form. Recall that a projective structure on an open set $U \subset \mathbb{R}^2$ is an equivalence class of torsion free connections $[\Gamma]$. Two connections Γ and $\hat{\Gamma}$ are projectively equivalent if they share the same unparametrised geodesics. The analytic expression for this equivalence class is

$$\hat{\Gamma}^k_{ij} = \Gamma^k_{ij} + \delta^k_i\omega_j + \delta^k_j\omega_i, \qquad i, j, k = 1, 2 \tag{3.10}$$

[1]Calculating the expressions $A_0, \ldots, A_3$ directly in terms of (E, F, G) and their first derivatives without introducing ψs would lead to non–linear relations.

for some one–form $\omega = \omega_i dx^i$.

Thus, in the language of projective differential geometry, we are looking for local conditions on a connection Γ^k_{ij} for the existence of a one form ω_i and a symmetric non-degenerate tensor g_{ij} such that the projectively equivalent connection is the Levi-Civita connection for g_{ij}, i. e.

$$\Gamma^k_{ij} + \delta^k_i\omega_j + \delta^k_j\omega_i = \frac{1}{2}g^{kl}\Big(\frac{\partial g_{il}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l}\Big).$$

This is an overdetermined system: there are six components in Γ^k_{ij} and five components in the pair (g_{ij}, ω_i).

Let $\Gamma \in [\Gamma]$ be a connection in the projective class. Its curvature is defined by

$$[\nabla_i, \nabla_j]X^k = R_{ijl}{}^k X^l$$

and can be uniquely decomposed as

$$R_{ijl}{}^k = \delta^k_i \mathrm{P}_{jl} - \delta^k_j \mathrm{P}_{il} + \beta_{ij}\delta^k_l \tag{3.11}$$

where β_{ij} is skew. In dimensions higher than 2 there would be another term (the projective Weyl tensor) in this curvature but in two dimensions it vanishes identically.

If we change the connection in the projective class using (3.10) then

$$\hat{\mathrm{P}}_{ij} = \mathrm{P}_{ij} - \nabla_i\omega_j + \omega_i\omega_j, \quad \hat{\beta}_{ij} = \beta_{ij} + 2\nabla_{[i}\omega_{j]}.$$

If the de Rham cohomology class $[\beta] \in H^2(U, \mathbb{R})$ vanishes then we can set β_{ij} to 0 by a choice of ω_i in (3.10). We are looking for a local metrisability condition on U so we shall assume that this global cohomological obstruction vanishes. The residual freedom in changing the representative of the equivalence class (3.10) is given by gradients $\omega_i = \nabla_i f$, where f is a function on U.

Now $\mathrm{P}_{ij} = \mathrm{P}_{ji}$ and the Ricci tensor of Γ is symmetric. The Bianchi identity implies that Γ is flat on a bundle of volume forms on U. Thus the equivalent way to normalise ∇_i is to require the existence of a volume form ε^{ij} such that

$$\nabla_i\varepsilon^{jk} = 0.$$

We shall use the volume forms to raise and lower indices according to $z_i = \varepsilon_{ij}z^j, z^i = z_j\varepsilon^{ji}$. Locally, such a volume form is unique up to scale: let us fix one.

With these preliminaries there exists a representative Γ in a projective class such that the linear system (3.9) becomes

$$\nabla_{(i}\sigma_{jk)} = 0,$$

where $\sigma_{ij} = \varepsilon_{il}\varepsilon_{jk}\sigma^{kl}$. Its prolongation gives rise to a connection on a rank 6 vector bundle $\mathbb{E}$ over U. Specifically, sections of this bundle comprise triples of contravariant tensors $\mathbf{u} = (\sigma^{ij}, \mu^i, \rho)$ with σ^{ij} being symmetric. The connection is given by

$$\left(\begin{array}{c} \sigma^{jk} \\ \mu^j \\ \rho \end{array}\right) \stackrel{D_i}{\longmapsto} \left(\begin{array}{c} \nabla_i\sigma^{jk} - \delta^j_i\mu^k - \delta^k_i\mu^j \\ \nabla_i\mu^j - \delta^j_i\rho + \mathrm{P}_{ik}\sigma^{jk} \\ \nabla_i\rho + 2\mathrm{P}_{ij}\mu^j - 2Y_{ijk}\sigma^{jk} \end{array}\right), \tag{3.12}$$

where $Y_{ijk} = \frac{1}{2}(\nabla_i \mathrm{P}_{jk} - \nabla_j \mathrm{P}_{ik})$ the Cotton tensor. The curvature of the connection D obtained from $\nabla_{[i}\nabla_{j]} = 0$. It is a 6×6 matrix of rank one. The first condition analogous to (2.5) is

$$5Y_i\mu^i + (\nabla_i Y_j)\sigma^{ij} = 0, \qquad \text{where} \quad Y_k = \varepsilon^{ij} Y_{ijk}.$$

Differentiating this equation twice and eliminating the first derivatives shows that the 6×6 matrix

$$\mathcal{M} = \left(\begin{pmatrix} 0 \\ 5Y_k \\ \nabla_{(j}Y_{k)} \end{pmatrix}, D_i \begin{pmatrix} 0 \\ 5Y_k \\ \nabla_{(j}Y_{k)} \end{pmatrix}, D_{(i}D_{j)} \begin{pmatrix} 0 \\ 5Y_k \\ \nabla_{(k}Y_{l)} \end{pmatrix} \right) \tag{3.13}$$

must be singular. Its determinant gives the first obstruction to metrisability of a projective structures. More detailed calculation shows that the expression for $\det(\mathcal{M})$ involves raising an index 14 times using the volume form ε and gives rise to a projectively invariant section of a 14th power of the canonical bundle

$$\det(\mathcal{M})\,(dx \wedge dy)^{\otimes 14}$$

which gives a projective invariant.

Analysis of the necessary conditions using Theorem 2.2 leads to higher order obstructions. If $\det(\mathcal{M}) = 0$ and $\mathrm{rank}(\mathcal{M}) = 5$ there will be two additional obstructions of order 6 in the components of a connection. If $2 < \mathrm{rank}(\mathcal{M}) < 5$ then there is one obstruction of order 7 in the rank 4 case and of order 8 in the rank 3 case. If $\mathrm{rank}(\mathcal{M}) = 2$ there always exists a four dimensional space of metrics compatible with the projective structure. Finally if $\mathrm{rank}(\mathcal{M}) < 2$ then Γ is projectively flat. See [5] for details and proofs.

Chapter 4

Method of Characteristics

If a differential ideal on M generated by one 1–form θ is closed then the Frobenius Theorem 1.3 provides a simple local normal form: There exist functions μ, y on M such that $\theta = \mu\, dy$. The next theorem gives a stronger result and can be applied to the case when the Frobenius conditions do not hold

Theorem 4.1 (Pfaff) *Let $(M, \mathcal{I})$ be an EDS such that $\mathcal{I} =< \theta >_{\text{diff}}$ for some non–vanishing one–form θ and let $r \geq 0$ be the smallest integer such that*

$$\theta \wedge d\theta^{r+1} = 0.$$

Set $\dim(M) = N$. *For each $x \in M$ such that $\theta \wedge d\theta^r \neq 0$ at x there exists a coordinate system*

$$(v, y^1, \ldots, y^r, q_1, \ldots, q_r, z^{2r+2}, \ldots, z^N)$$

in the neighbourhood of x such that $\mathcal{I} =< dv >$ if $r = 0$ and, if $r > 0$,

$$\mathcal{I} =< dv - q_1 dy^1 - \ldots - q_r dy^r >_{\text{diff}}$$

and moreover

- *There exists a maximal $(N - r - 1)$–dimensional integral manifold of $\mathcal{I}$*

$$v = q_1 = q_2 = \ldots = q_r = 0.$$

- *Any integral manifold near this one depends on one arbitrary function of r variables, $f(y^1, \ldots, y^r)$ and is given by*

$$v = f(y^1, \ldots, y^r), \quad q_k = \frac{\partial f}{\partial y^k}(y^1, \ldots, y^r), \quad k = 1, \ldots, r.$$

This theorem is proved in [3]. We shall not reproduce this proof, but instead concentrate on one important application: The method of characteristics.

Consider a single first order PDE

$$F\Big(x^1, \ldots, x^n, u, \frac{\partial u}{\partial x^1}, \ldots, \frac{\partial u}{\partial x^n}\Big) = 0. \tag{4.1}$$

This PDE defines a co–dimension one manifold $M \subset J^1(\mathbb{R}^n, \mathbb{R})$ of the $(2n+1)$ dimensional first jet space $J^1(\mathbb{R}^n, \mathbb{R})$ with coordinates $(x^i, u, p_i := \partial u/\partial x^i)$. If we assume that F is smooth and not all partial derivatives $\partial F/\partial p_i$ vanish at any single point then the implicit function theorem implies that the surface M given by

$$F(x^1, \dots, x^n, u, p_1, \dots, p_n) = 0$$

is a smooth manifold. The PDE (4.1) is modeled by an EDS $\mathcal{I}$ generated on M by one–form

$$\theta = du - p_i dx^i.$$

On M the one–forms $\{dx^i, dp_i, du\}$ are linearly dependent as

$$0 = dF = \frac{\partial F}{\partial x^i} dx^i + \frac{\partial F}{\partial p_i} dp_i + \frac{\partial F}{\partial u} du.$$

Moreover $\theta \wedge (d\theta)^n = 0, \theta \wedge (d\theta)^{n-1} \neq 0$. Therefore the Pfaff theorem 4.1 implies the existence of a coordinate system

$$(v, y^1, \dots, y^{n-1}, q_1, \dots, q_{n-1}, z)$$

such that

$$\theta = \mu(dv - q_1 dy^1 - \dots - q_{n-1} dy^{n-1})$$

for some non–vanishing function μ. The vector field

$$\frac{\partial}{\partial z}$$

is a characteristic vector field as it satisfies[1]

$$\frac{\partial}{\partial z} \lrcorner \, \theta = 0, \quad \frac{\partial}{\partial z} \lrcorner \, d\theta = 0.$$

Using the original coordinate system we verify that the vector field

$$Z = \frac{\partial F}{\partial p_i} \frac{\partial}{\partial x^i} + p_i \frac{\partial F}{\partial p_i} \frac{\partial}{\partial u} - \Big(\frac{\partial F}{\partial x^i} + p_i \frac{\partial F}{\partial u} \Big) \frac{\partial}{\partial p_i} \tag{4.2}$$

on $J^1(\mathbb{R}^n, \mathbb{R})$ is tangent to the level set $M = F^{-1}(0)$ and it satisfies

$$Z \lrcorner \, \theta = 0, \quad Z \lrcorner \, d\theta = 0 \text{ mod } \theta.$$

Thus, $Z = \nu \, \partial/\partial z$ for some non–vanishing function ν. The initial value problem for the PDE (4.1) can now be solved in the following steps

- The initial data for (4.1) is an $(n-1)$ dimensional submanifold Σ of $\mathbb{R}^{n+1}$ given in parametric form by

  $$(s_1, \dots, s_{n-1}) \longrightarrow (x^i(s), u(s)) \subset \mathbb{R}^{n+1}.$$

  The natural lift of this submanifold to a graph in $J^1(\mathbb{R}^n, \mathbb{R})$ gives an $(n-1)$ dimensional integral manifold $\Sigma \subset M$ of $\mathcal{I}$ that is transverse to Z.

[1] In general Z is a Cauchy characteristic vector field if $Z \lrcorner \, \theta \in \mathcal{I}$ for all $\theta \in \mathcal{I}$

- Construct an n–dimensional integral manifold by solving a system of ODEs to find integral curves of Z (called the characteristic curves) and taking the union of these curves through Σ. If a characteristic curve has a point in common with a graph of a solution, it lies entirely on the graph.

- In the coordinates of Pfaff theorem 4.1 the n–dimensional integral manifold is give by

$$v = f(y^1, \dots, y^{n-1}), \qquad q_i = \frac{\partial f}{\partial y^i}(y^1, \dots, y^{n-1})$$

  for some function f which should be determined from the initial data.

Consider the special case of quasilinear PDE (4.1) where

$$F\Big(x, u, \frac{\partial u}{\partial x^i}\Big) = R^i(x, u)\frac{\partial u}{\partial x^i} + S(x, u)$$

and the Cauchy characteristic vector field (4.2) is

$$Z = R^i \frac{\partial}{\partial x^i} + p_i R^i \frac{\partial}{\partial u} - \Big(p_i \frac{\partial R^i}{\partial x^j} + \frac{\partial S}{\partial x^j} + p_j\Big(\frac{\partial R_i}{\partial u}p_i + \frac{\partial S}{\partial u}\Big)\Big)\frac{\partial}{\partial p_j}.$$

The classical treatment of this quasilinear problem does not use the jet space formalism. Evaluating Z at $F = 0$ shows that the integral curves of Z project to curves on the solution surface $x \to (x, u = u(x))$ which are integral curves of

$$\widetilde{Z} = R^i \frac{\partial}{\partial x^i} - S\frac{\partial}{\partial u}.$$

The PDE $F = 0$ can be rewritten as

$$\widetilde{\mathbf{Z}} \cdot \mathbf{n} = 0,$$

where the vector

$$\mathbf{n} = (\partial_1 u, \dots, \partial_n u, -1)$$

is normal to a solution surface $u = u(x)$ in $\mathbb{R}^{n+1}$. Therefore $\widetilde{\mathbf{Z}}$ is tangent to this surface. The characteristic curves which foliate the solution surface are solutions to the system of ODEs

$$\dot{x}^i = R^i(x, u), \quad \dot{u} = -S(x, u), \qquad i = 1, \dots, n$$

(where $\dot{=}\partial/\partial z$) with the initial conditions given by the initial data for (4.1)

$$x^i(0) = x^i(s_1, \dots, s_{n-1}), \quad u(0) = u_0(s_1, \dots, s_{n-1}).$$

The method breaks down if the initial data is not transverse of $\widetilde{\mathbf{Z}}$. A surface tangent to $\widetilde{Z}$ is called characteristic. Thus initial data specified along the characteristic surface does not determine the solution uniquely.

- **Example.** Consider the initial value problem for the dispersionless KdV equation

$$u_t + uu_x = 0, \qquad u(x, 0) = f(x).$$

The characteristic equations are

$$\frac{dx}{dz} = u, \quad \frac{dt}{dz} = 1, \quad \frac{du}{dz} = 0.$$

The solution surface must contain the curve $\Sigma \subset \mathbb{R}^3$ which we parametrise as

$$s \longrightarrow (x(s), t(s), u(s)) = (s, 0, f(s)).$$

Using this as the initial condition for the characteristic ODEs yields

$$x(s, z) = f(s)z + s, \quad t(s, z) = z, \quad u(s, z) = f(s).$$

Eliminating (s, z) between these formulae yields the general solution in the implicit form

$$u = f(x - ut).$$

This will be valid in the domain where the coordinates (s, z) are well defined and can be used instead of (x, t). In general one needs to analyse the Jacobian of the transformation to specify the domain of solution. In our case we can proceed as follows: The characteristic curves project to the straight lines $x(s) = f(s)t(s) + s$ in the domain of (x, t) in $\mathbb{R}^2$. These lines have different slopes for different values of s (say s_1 and s_2) and thus they can intersect. The intersection will take place at a point $(x, t) \in \mathbb{R}^2$ where

$$t = \frac{s_2 - s_1}{f(s_1) - f(s_2)}.$$

At this point the solution becomes multivalued, taking values $f(s_1)$ and $f(s_2)$. To understand it better, differentiate the implicit solution to find

$$u_x = \frac{f(s)'}{1 + tf(s)'}.$$

Hence if $f(s)' < 0$ the derivative u_x becomes infinite at the finite positive time $t = -(f(s)')^{-1}$. At this time the solution experiences the gradient catastrophe.

Chapter 5

Cartan–Kähler Theorem

The Frobenius Theorem 1.3 gives the criterion for the existence of integral manifolds for EDS generated algebraically by one–forms. The Cartan–Kähler theorem (proved by Cartan for Pfaffian systems, and extended to the general case by Kähler) deals with arbitrary EDSes. Our brief presentation of the subject in this section follows [4].

The proof of the Frobenius Theorem 1.3 was based on Picard's existence theorem for ODEs. Thus the Frobenius theorem works in the smooth category. The proof of the Cartan–Kähler theorem involves the Cauchy–Kowalewska existence theorem for PDEs. The Cauchy–Kowalewska theorem which we shall state below is valid in the real–analytic category. Let $u : \mathbb{R}^{n+1} \to \mathbb{R}^N$. Thus the collection of functions $u^\alpha, \alpha = 1, \ldots, N$ depends on $(n+1)$ independent variables $(x^i, t), i = 1, \ldots, n$. The system of PDEs in the Cauchy form is

$$\begin{aligned} \frac{\partial u^\alpha}{\partial t} &= F^\alpha(x^i, t, u^\alpha, \frac{\partial u^\alpha}{\partial x^i}). \\ u^\alpha|_{t=t_0} &= g^\alpha(x^i) \end{aligned} \tag{5.1}$$

Theorem 5.1 (Cauchy–Kowalewska) *If the equation* (5.1) *and the initial data are real–analytic then there exists a unique solution in the form of a power series*

$$u^\alpha(t, x) = g^\alpha(x^i) + g_1^\alpha(x^i)(t - t_0) + \frac{1}{2} g_2^\alpha(x^i)(t - t_0)^2 + \ldots$$

which converges on some domain containing $t = t_0$.

This theorem can be refined if the first derivatives of u with respect to t are specified only for the first r components of u, i.e. if $\alpha = 1, \ldots, r < N$ in (5.1). In this case the system is under-determined as there are fewer equations than unknowns. The general analytic solution to (5.1) depends on $(N - r)$ arbitrary functions. This is quite obvious, a choice of $(N - r)$ functions is needed to put the equation in the 'determined form' (5.1) with $\alpha = 1, \ldots, N$.

Definition 5.2 *A k–dimensional subspace $E \subset T_xM$ is an integral element of $\mathcal{I}$ if*

$$\theta(e_1, \ldots, e_k) = 0$$

for all $\theta \in \mathcal{I}^k$ and $e_i \in E, i = 1, \ldots, k$.

The set of all k–dimensional integral elements is denoted $V_k(\mathcal{I})$. It is clear that tangent space to any k–dimensional integral manifold is an integral element. We aim to answer the following

- **Question**. When is an integral element tangent to an integral manifold?

Certainly not always, as obstructions can arise from Frobenius theorem

- **Example.** The EDS
$$\mathcal{I} =< dx \wedge dz, dy \wedge (dz - ydx) >_{\text{diff}}$$
has a two–dimensional integral element
$$\{\partial_x + y\partial_z, \partial_y\}$$
at each point, but no two–dimensional integral manifolds as the vectors spanning E do not satisfy the Frobenius condition (1.5).

If $E \subset V_k(\mathcal{I})$ and $G \subset E$ is a p–dimensional subspace of E then $G \subset V_p(\mathcal{I})$. Thus restrictions of integral elements are integral elements. But the converse is not true, and not every extension of integral element may be an integral element.

Definition 5.3 *Let $E \subset V_k(\mathcal{I})$ be spanned by $\{e_1, \ldots, e_k\}$. A polar space of E is*
$$H(E) = \{v \in T_xM, \theta(v, e_1, \ldots, e_k) = 0, \;\forall\theta \in \mathcal{I}^{k+1}\} \subset T_xM.$$

The polar space is a vector space containing E, but it does not have to be an integral element. However if $v \in H(E)$ and v is not an element of E then a direct sum $E \oplus \text{span}\{v\}$ is a $(k+1)$ dimensional integral element. Thus $H(E)$ is the space of possible one–dimensional extensions of a given integral element. Constructing H from a given E comes down to solving a set of linear homogeneous equations for components of v. In practice to compute a polar space of a k–dimensional integral element E, contract all $(k+1)$ forms in the ideal with all vectors in E. The resulting one–forms should be annihilated by all vectors in $H(E)$. An integral element E is called regular if the dimension of the polar space is constant in a neighbourhood of E in $V_k(\mathcal{I})$. Moreover E is called ordinary if the intersection of $V_k(\mathcal{I})$ with an open neighbourhood of E is a smooth submanifold of the Grassmanian $\text{Gr}_k(TM)$ of all k planes in TM.

For a given $E \in V_k(\mathcal{I})$ define
$$r(E) = \dim\left(H(E)\right) - k - 1$$
to be a dimension of the set of $(k+1)$ integral elements that contain E with $r(E) = -1$ if there are no such elements.

- **Example [4].** Let
$$\mathcal{I} =< dx \wedge dz, dy \wedge (dz - ydx) >_{\text{diff}}$$
be an EDS on $\mathbb{R}^3$. One–dimensional integral element E is spanned by $e_1 = a\partial_x + b\partial_y + c\partial_z$. The vector $v = f\partial_x + g\partial_y + h\partial_z$ is in $H(E)$ if two linear equations
$$cf - ah = 0, \qquad -ybf - (c - ya)g + bh = 0$$
for (f, g, h) hold. If $c - ya \neq 0$ we get $H(E) = E$ and thus $r(E) = -1$. If $c - ya = 0$ then $\dim\left(H(E)\right) = 2$ and $r(E) = 0$. In particular E is not a regular integral element.

An integral manifold $S \subset M$ is called ordinary/regular iff all of its tangent spaces are ordinary/regular elements. For regular integral manifolds we define $r(S) = r(T_x S)$.

Theorem 5.4 (Cartan–Kähler) *Let $(M, \mathcal{I})$ be a real analytic EDS and let $\Sigma \subset M$ be an n–dimensional analytic submanifold whose tangent spaces are regular integral elements such that $\dim(H(T_x\Sigma)) = n + 1$. Then there exists an open neighbourhood of $x \in \Sigma$ and a unique analytic $(n + 1)$–dimensional integral manifold $S \subset U$ containing $\Sigma \cap U$.*

The Cartan–Kähler theorem states when an n–dimensional integral manifold can be thickened to an $(n+1)$ dimensional integral manifold. This theorem needs to be modified by introducing so called restraining manifold if the dimension of the polar space of $T_x\Sigma$ is greater that $(n+1)$. This is needed for uniqueness. The restraining R manifold is an analytic submanifold of M of co–dimension $r(\Sigma)$ such that $\Sigma \subset R$ and $T_xR \cap H(T_x\Sigma)$ has dimension $(n+1)$ for all $x \in \Sigma$. Then there exists a unique connected $(n+1)$ analytic integral manifold S which satisfies $\Sigma \subset S \subset R$. The reader is referred to [3] where this is discussed.

The proof of the Cartan–Kähler theorem is obtained by adopting local coordinates and reducing the problem to a solution of the system of PDEs of the form (5.1). This uses the Cauchy–Kowalewska theorem and so one needs to require real–analyticity. Again, consult [3] or [7] for details.

The integral manifolds can in principle be constructed successively using the Cartan–Kähler theorem. At each step the integral manifold is determined by a choice of restraining manifolds and the arbitrary functions in the maximal integral manifold parametrise these choices. We shall now discuss the Cartan test which gives a handle on how to calculate this freedom in the 'general solution'. Applying the Cartan–Kähler theorem succesivelly, starting from one–dimensional integral manifolds gives a sufficient condition for the existence of an integral manifold tangent to a given integral element: If $E \subset V_n(\mathcal{I})$ contains a flag of subspaces

$$\{0\} = E_0 \subset E_1 \subset \ldots \subset E_n = E \subset T_xM,$$

where the integral elements $E_k \subset V_k(\mathcal{I})$ are regular, then there exists a real analytic n–dimensional integral manifold $\Sigma \subset M$ passing through x and satisfying $T_xP = E$.

This corollary from theorem 5.4 is not of great practical significance, as the regularity assumption needs to be checked at each step. Also, it gives a sufficient condition which is not necessary as not all integral manifolds have tangent spaces which are final objects in a flag of regular integral elements.

To get around this, consider the integral flag $\mathcal{F} = (E_0, \ldots, E_n)$, not necessarily regular, and set

$$c(E_k) := \dim(T_xM) - \dim H(E_k), \qquad k = 1, 2, \ldots, n$$

and let $c(E_{-1}) = 0$. The Cartan characters of the flag $\mathcal{F}$ are the non–negative numbers defined by

$$s_k(\mathcal{F}) := c(E_k) - c(E_{k-1}).$$

Theorem 5.5 (Cartan Test) *Let $(M, \mathcal{I})$ be an EDS and let $\mathcal{F} = (E_0, \ldots, E_n)$ be an integral flag of $\mathcal{I}$. Then $V_n(\mathcal{I})$ has co–dimension at least*

$$c(\mathcal{F}) := c(E_0) + c(E_1) + \ldots + c(E_{n-1})$$

in the Grassmannian[1] $Gr_n(TM)$ *at* E_n*. Moreover* $V_n(\mathcal{I})$ *is a smooth submanifold of* $Gr_n(TM)$ *of co–dimension* $c(\mathcal{F})$ *iff the flag* $\mathcal{F}$ *is regular.*

Performing Cartan's test on a given flag is just a matter of linear algebra. If the flag passes the test and therefore is regular, the Cartan–Kähler theorem implies the existence of at least one real–analytic n–dimensional integral manifold $S \subset M$ such that $T_xS = E_n$ (there will be exactly one such manifold if the $r(E_{n-1}) = 0$. Otherwise the restraining manifold has to be chosen). Of course for a given integral element E_n there may be more than one flag which terminates at E_n. It practice it makes sense to choose the first element in a flag such that the first Cartan character s_k is as large as possible, then choose the second element such that the next character is is as large as possible etc. The sum $s_1 + s_2 + \ldots + s_n$ is fixed regardless of these choices. In what follows we shall drop the reference to the flag and write s_k instead of $s_k(\mathcal{F})$. The highest k such that $s_k \neq 0$ is called Cartan character. Moreover let $c(E_n) = s$. Using the definitions of Cartan characters and $c_n = \dim M - n$ we can write

$$s_0 + s_1 + \ldots + s_k = c_k$$

and can rewrite the inequality in Cartan's test as

$$\dim (V_n(\mathcal{I})) - \dim (M) \leq s_1 + 2s_2 + \ldots + ns_n, \tag{5.2}$$

where the LHS is the fiber dimension of $V_n(\mathcal{I})$.

Given a flag $\mathcal{F}$ which passes the test it is possible to chose a coordinate system

$$(x^1, \ldots, x^n, u^1, \ldots, u^s)$$

centered at $x \in U \subset \mathbb{R}^{n+s}$ such that E_k is spanned by

$$\{\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^k}\}, \qquad 0 \leq k < n$$

and elements $H(E_k)$ are annihilated by the one forms

$$\{du^1, \ldots, du^{c_k}\}.$$

Let $\mathcal{S}$ be the collection of the real analytic integral manifolds near S. This means that $\hat{S} \in \mathcal{S}$ if it can be represented by

$$u^\alpha = F^\alpha(x^1, \ldots, x^n),$$

where the analytic functions F^α are defined in the neighbourhood of $\mathbf{x} = 0$. Then the collection $\mathcal{S}$ depends on s_0 constants, s_1 functions of one variable, …, s_n functions of n variables. Thus the integers $(s_0, s_1, \ldots, s_n)$ measure the arbitrariness of the general integral manifold.

[1] Recall that the Grassmannian $\mathrm{Gr}_k(E)$ is a set of k–dimensional subspaces of a vector space E. It is a smooth manifold of dimension $k(\dim E - k)$. The set of all k–dimensional subspaces in T_xM as x varies over M is denoted $\mathrm{Gr}_k(TM)$. It is a manifold of dimension $\dim (M) + k(\dim M - k)$. Given a k–plane E in T_xM on which $dx^1 \wedge \ldots \wedge dx^k \neq 0$ we can choose coordinates $(x^1, \ldots, x^k, u^1, \ldots u^s)$ on M such that E is spanned by vectors

$$\frac{\partial}{\partial x^i} + \sum_{\alpha=1}^{s} p_i^\alpha(E) \frac{\partial}{\partial u^\alpha}, \qquad i = 1, \ldots, k,$$

where $p_i^\alpha = p_i^\alpha(E)$ are coordinates on the fibres of $\mathrm{Gr}_k(TM) \to M$.

- **Example.** A Lagrangian submanifold of $\mathbb{R}^{2n}$ is an integral manifold of an ideal generated by a symplectic structure

$$\theta = dx^1 \wedge du^1 + dx^2 \wedge du^2 + \ldots + dx^n \wedge du^n.$$

Choose a flag

$$\{0\} \subset \{\frac{\partial}{\partial x^1}\} \subset \{\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}\} \subset \ldots \subset \{\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \ldots, \frac{\partial}{\partial x^n}\}.$$

For this flag $H(E_0)$ is the whole tangent space, so $c_0 = 0$. Then $H(E_1)$ consists of all vectors which annihilate du^1 and more generally

$$H(E_k) = \{du^1, \ldots, du^k\}^{\perp}.$$

Thus

$$c_0 = 0, c_1 = 1, c_2 = 2, \ldots, c_n = n$$

which implies

$$s_0 = 0, \quad s_1 = s_2 = \ldots = s_n = 1.$$

To calculate the fiber dimension of $V_n(\mathcal{I})$ note that the vectors spanning E_n annihilate the one forms $du^k, k = 1, \ldots, n$. The nearby integral planes are given by

$$du^k = \sum_{j=1}^{n} p^{jk} dx^j, \quad k = 1, \ldots, n,$$

and the total number of symmetric coefficients $p^{jk} = p^{kj}$ is the fibre dimension of $V_n(\mathcal{I})$ which appears on the LHS of the inequality (5.2). This number is

$$\binom{n+1}{2}$$

which is also equal to the RHS of (5.2). Thus we have an equality and the flag is regular. The general integral manifold depends on one function of n variables and functions of lower number of variables. Explicitly

$$u^k = \frac{\partial f}{\partial x^k}, \qquad k = 1, \ldots, n,$$

where $f = f(x^1, x^2, \ldots, x^n)$.

- **Example.** Consider the EDS (1.1) in $\mathbb{R}^5$

$$< \theta^1 = du - pdx - qdy, \, \theta^2 = dp \wedge dq - dx \wedge dy >_{\text{diff}}$$

for the Monge–Ampere equation. The four–dimensional space $V_1(\mathcal{I}) = \{\theta^1\}^{\perp}$ of one–dimensional integral elements is spanned by

$$\{\partial_x + p\partial_u, \partial_y + q\partial_u, \partial_p, \partial_q\}.$$

Pick $E_1 = \{\partial_p\}$ to be the first element in the flag. The two–forms in the ideal are

$$\theta^2, \quad d\theta^1 = dx \wedge dp + dy \wedge dq, \quad \theta^1 \wedge \gamma$$

where γ is any one form. Therefore the polar space $H(E_1)$ will consist of all vectors annihilating $\partial_p \lrcorner\, \theta_2, \partial_p \lrcorner\, d\theta^1$ and θ^1. Thus $H(E_1) = \{dq, dx, \theta^1\}^{\perp}$. This space is two dimensional and there is a unique extension of E_1 to an integral element

$$E_2 = \{\partial_p, \partial_y + q\partial_u\}$$

(we would have got the same integral element E_2 if we had picked $E_1 = \{\partial_y + q\partial_u\}$). Contracting the two vectors in E_2 with all three forms in the ideal shows that this can not be further extended i.e. $H(E_2) = E_2$. Therefore we pick a flag

$$\{0\} \subset \{\partial_p\} \subset \{\partial_p, \partial_y + q\partial_u\} = E \subset T_x\mathbb{R}^5.$$

This flag has $c_0 = 5-4 = 1, c_1 = 5-2 = 3, c_2 = 5-2 = 3$ and so $s_0 = 1, s_1 = 2$. To perform the Cartan test we need to compute the co–dimension of $V_2(\mathcal{I})$ in the Grassmannian of two planes. The two–planes close to E_2 are spanned by

$$v_1 = \partial_p + \alpha(\partial_x + p\partial_u) + \beta\partial_q + \gamma\partial_u, \;\; v_2 = \partial_y + q\partial_u + \delta(\partial_x + p\partial_u) + \epsilon\partial_q + \phi\partial_u$$

for some $(\alpha, \beta, \ldots, \phi)$. The conditions

$$\theta^1(v_1) = 0, \quad \theta^1(v_2) = 0, \quad d\theta^1(v_1, v_2) = 0, \quad \theta^2(v_1, v_2) = 0$$

give four linear equations

$$\gamma = 0, \quad \phi = 0, \quad \epsilon - \alpha = 0, \quad \beta + \delta = 0.$$

Thus the fibre co–dimension of $V_2(\mathcal{I})$ is 4 which is equal to $c_0 + c_1 + c_2$. The Cartan test holds and the general solution to the Monge–Ampere equation depends on two functions of one variable.

In theory one could always reduce a problem to analysis of a Pfaffian system (i.e. one where $\mathcal{I}$ is generated by one–forms) as any EDS can be prolonged to such system. If a Pfaffian system is generated by

$$< \theta^1, \ldots, \theta^N >$$

then the vectors $\{e_1, e_2, \ldots, e_k\}$ spanning E_k in a flag

$$\{0\} \subset E_1 \subset E_2 \subset \ldots \subset E_n \subset T_xM$$

are found by solving a system

$$e_i \lrcorner\, \theta^\alpha = 0, \quad e_i \lrcorner\, (e_j \lrcorner\, d\theta^\alpha) = 0, \qquad i, j = 1, \ldots, k, \quad \alpha = 1, \ldots, N.$$

This however comes at a price of introducing more variables and working in spaces of high dimension: For a system of r PDEs of order k for N functions of n unknowns

$$F^\rho\Big(x^i, u^\alpha, \frac{\partial u^\alpha}{\partial x^i}, \ldots, \frac{\partial^k u^\alpha}{\partial x^{i_1}\partial x^{i_2}\ldots\partial x^{i_k}}\Big) = 0, \qquad \rho = 1, \ldots, r, \quad \alpha = 1, \ldots N, \quad i = 1, \ldots, n$$

the Pfaffian system is generated by one forms

$$du^\alpha - p^\alpha_i dx^i, \quad dp^\alpha_i - p^\alpha_{ij} dx^j, \quad \ldots, \quad dp^\alpha_{i_1 i_2 \ldots i_{k-1}} - p^\alpha_{i_1 i_2 \ldots i_k} dx^{i_k}$$

on a manifold M given by the zero locus

$$F^\rho\Big(x^i, u^\alpha, p^\alpha_i, \ldots, p^\alpha_{i_1 i_2 \ldots i_k}\Big) = 0$$

in the kth jet space $J^k(\mathbb{R}^n, \mathbb{R}^N)$. There are however more economical tricks to reduce a problem to a Pfaffian system. The following example, modified from [1], shows one such trick.

- **Example.** Consider the Ricci–flat Kähler equation (1.3) in four dimensions. We are interested in the real analytic solutions, so we can complexify the dependent and independent variables and regard $(w, z, \bar{w}, \bar{z})$ as independent holomorphic coordinates on an open ball in $\mathbb{C}^4$. The equation (1.3) is modelled by the ideal generated by one 1–form and one 4–form on $\mathbb{C}^9$

  $$< d\Omega - pdw - \bar{p}d\bar{w} - qdz - \bar{q}d\bar{z},\ dp \wedge dq \wedge dw \wedge dz - d\bar{w} \wedge d\bar{z} \wedge dw \wedge dz >_{\text{diff}}$$

  together with the independence condition $dw \wedge dz \wedge d\bar{w} \wedge d\bar{z} \neq 0$. To reformulate the problem as a Pfaffian system rewrite the vanishing of the 4-form as

  $$d(pdq - \bar{w}d\bar{z}) \wedge dw \wedge dz = 0.$$

  The independence condition implies $dw \wedge dz \neq 0$. Thus locally there exist functions a, b, Σ such that

  $$pdq - \bar{w}d\bar{z} = d\Sigma - adz - bdw$$

  on integral manifolds. Conversely equation (1.3) can be modelled as a Pfaffian EDS

  $$\mathcal{I} =< \theta^1 = d\Omega - pdw - \bar{p}d\bar{w} - qdz - \bar{q}d\bar{z},\ \theta^2 = d\Sigma - adz - bdw - pdq + \bar{w}d\bar{z} >_{\text{diff}}$$

  in $\mathbb{C}^{12}$ with coordinates $(w, z, \bar{w}, \bar{z}, p, q, \bar{p}, \bar{q}, a, b, \Omega, \Sigma)$.

  The space of one–dimensional integral elements $\{\theta^1, \theta^2\}^\perp$ is 10 dimensional, thus $c_0 = 2$ and $s_0 = 2$. Let $E_1 = \{e_1\}$. The polar space of E_1 is the 8 dimensional vector space

  $$H(E_1) = \{\theta^1, \theta^2\ e_1 \lrcorner\, d\theta^1, e_1 \lrcorner\, d\theta^2\}^\perp.$$

  Thus $c_1 = 4$ and $s_1 = c_1 - c_0 = 2$. Let $E_2 = \{e_1, e_2\}$. Then

  $$H(E_2) = \{\theta^1, \theta^2, e_1 \lrcorner\, d\theta^1, e_1 \lrcorner\, d\theta^2, e_2 \lrcorner\, d\theta^1, e_2 \lrcorner\, d\theta^2\}^\perp,$$

  so $c_2 = 6$ and $s_2 = 2$. We continue looking for polar spaces and extending the integral elements. Let $E_3 = \{e_1, e_2, e_3\}$. This gives[2] $c_3 = 8, s_3 = 2$. Pick some $e_4 \in H(E_3)$ and set $E_4 = \{e_1, e_2, e_3, e_4\}$. Now

  $$H(E_4) = \{\theta^1, \theta^2, e_i \lrcorner\, d\theta^1, e_i \lrcorner\, d\theta^2\}^\perp, \qquad i = 1, \ldots, 4$$

[2]The flag must be chosen carefully for this to be true. The choice (5.3) will do.

and $\dim H(E_4) \leq 4$. However $E_4 \subset H(E_4)$ and we must have $H(E_4) = E_4$ and the integral element E_4 is not extendable. We have $c_4 = 12 - 4 = 8$ and $s_4 = 0$. Thus the maximal integral manifolds may be at most four dimensional if we can pick a regular flag. We can verify the computations of Cartan characters by choosing the flag with

$$e_1 = \partial_w + \partial_{\bar{w}} + (p + \bar{p})\partial_\Omega + b\partial_\Sigma, \quad e_2 = \partial_p - \partial_{\bar{p}}, \quad e_3 = \partial_{\bar{q}} + \partial_a, \quad e_4 = \partial_{\bar{q}}. \tag{5.3}$$

Then

$$\begin{aligned}
H(E_1) &= \{\theta^1, \theta^2, dp + d\bar{p}, db + d\bar{z}\}^\perp \\
H(E_2) &= \{\theta^1, \theta^2, dp + d\bar{p}, db + d\bar{z}, dw - d\bar{w}, dq\}^\perp \\
H(E_3) &= \{\theta^1, \theta^2, dp + d\bar{p}, db + d\bar{z}, dw - d\bar{w}, dq, d\bar{z}, dz\}^\perp \\
H(E_4) &= E_4.
\end{aligned}$$

The codimension of $V_4(\mathcal{I})$ around $E = E_4$ can be now computed as in the last example. The Cartan test holds and thus the general real–analytic Ricci–flat Kähler metric in four dimensions depends on two arbitrary functions of three variables.

Bibliography

[1] Boyer, C. P. & Plebański, J. F. (1977) Heavens and their integral manifolds. J. Math. Phys. **18**, 1022–1031.

[2] Branson, T. P., Čap, A., Eastwood, M. G., and Gover, A. R., *Prolongations of geometric overdetermined systems*, Internat. J. Math. **17** (2006) 641–664.

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

[4] Bryant, R. L. (2002) Lectures on Exterior Differential Systems.

[5] Bryant, R. L, Dunajski, M. & Eastwood M.(2008) Metrisability of two-dimensional projective structures `arXiv:0801.0300` .

[6] Darboux G. (1896) Lecons sur la thorie gnrale des surfaces, Vol. III, Chelsea Publishing.

[7] Ivey, T. A. & Landsberg, J. M. (2003) Cartan for Beginners: Differential Geometry via Moving Frames and Exterior Differential Systems, AMS.

[8] Kruglikov, B. (2007) Invariant characterisation of Liouville metrics and polynomial integrals. `arXiv:0709.0423`

[9] Liouville, R. (1887) Sur une classe d'équations différentielles, parmi lesquelles, en particulier, toutes celles des lignes géodésiques se trouvent comprises, Comptes rendus hebdomadaires des seances de l'Academie des sciences **105**, 1062-1064.

[10] Spencer, D. C. (1969) Overdetermined systems of linear partial differential equations, Bull. Amer. Math. Soc. **75**, 179-239.

[11] Spivak, M. A comprehensive Introduction to Differential Geometry, vol 5.

[12] Veblen, O, & Thomas, J. M. (1926) Projective invariants of affine geometry of paths. Ann. of Math. **27**, 279–296.