# Toward a Geometry of Differential Equations

**Robert Bryant**
*Duke University*
*Durham, NC 27708*

**Phillip Griffiths**
*Institute for Advanced Study*
*Princeton, NJ 08540*

**Lucas Hsu**
*Institute for Advanced Study*
*Princeton, NJ 08540*

In this expository paper we will discuss the geometry of differential equations. On the one hand this subject has very classical origins going back to the earliest systematic studies of differential equations. On the other hand, in much of this half-century, its main thrust has been more towards structural and foundational matters within the theory than towards applications to particular equations. However, recently this balance has been changing and there are indications that this type of geometric understanding of differential equations is of value in understanding geometric and analytic problems.

In this paper the emphasis will be on illustrating the general theory rather than explaining it. Thus we will attempt to illustrate how the geometry of differential equations may be used to gain insight into selected particular equations and classes of equations. The examples we have chosen will for the most part be elementary. The presentation is designed to show how one may look at familiar materials in PDE theory in a somewhat more geometric fashion.

What do we mean by a "*geometry*"? In the first approximation, this will mean a pair

$$(B_G \to M,\, \phi)$$

where $M$ is an $n$-manifold, $G \subset \mathrm{GL}(n, \mathbb{R})$ is a Lie subgroup and $B_G \to M$ is a $G$-structure on $M$, and $\phi$ is a connection on $B_G \to M$. (Actually, $B_G$ will sometimes turn out to be a sub-bundle of a higher order coframe bundle and $\phi$ will be a pseudo-connection, but these refinements will not be insisted upon here.) To

This research was supported in part by NSF Grant DMS 9205222 (Bryant), and the Institute for Advanced Study (Griffiths and Hsu).

**Geometry, Topology, and Physics for Raoul Bott**



Cambridge, MA 02138 USA

the geometry $(B_G \to M,\, \phi)$ we may invariantly attach other objects, such as the *curvature* and *geodesics* of the connection $\phi$.

The main general observation is that to a differential equation (ordinary or partial) one may canonically attach a geometry in the above sense. More precisely, given a differential equation $E$ *and* a group $\Gamma$ of admissable transformations (for example gauge, point, or contact transformations, as explained below), there is, under mild regularity assumptions, an intrinsically associated geometry. For general PDE systems, including many interesting special cases, the pseudo-connection $\phi$ will be unique. However, the construction

$$(E, \Gamma) \rightsquigarrow (B_G \to M,\, \phi)$$

is somewhat subtle, and in fact the general result that the construction is a finite process has (to our knowledge) never really been proved. In this paper we will usually think of this process as a "black box"; in the literature it is more formally known as the *equivalence problem.* The consequences of the construction will be discussed in the text for a number of examples. In the Appendix we have worked through the "black box" in these examples. Here we emphasize the fundamental point that *there is a natural geometry associated to a differential equation.* Thus to a differential equation one may invariantly attach its *curvatures* (classically known as differential invariants) and the notion of *completeness* (which pertains to existence of global solutions).

Although the association of a geometry to a differential equation may not be generally familiar, certain aspects of this construction—"partial geometries", so to speak—are, of course, well known. For example, many equations are written in "almost canonical" coordinates that do have physical significance, so that coordinate calculations have at least partially invariant meaning. Or again, the *symbol* (and sometimes *sub-principal symbol*) of a partial differential equation, which is usually discussed in a non-invariant way, turns out to have geometric meaning. Our point is that consideration of the full geometry associated to a differential equation can sometimes supplement these existing methods.

Before turning to a discussion of the contents of this paper, we would like to suggest what the objectives of a "geometry of differential equations" should be. First and foremost, the theory should provide new information and perspectives on solutions[1] to differential equations. In fact, the classical theory was originally developed in order to provide explicit solutions to interesting differential equations, and then later it sought to provide explicit methods (integral formulas, etc.) for constructing the solutions (see *[17]*). A second objective of the theory is to treat $(B_G \to M,\, \phi)$ as a geometric object of interest in its own right. Thus, as we shall illustrate below, certain interesting differential equations may be uniquely characterized in terms of their associated geometry.

Perhaps one may think here of the subject of algebraic geometry. Its original objective was to solve algebraic equations. Later the geometric objects these equations define—the algebraic varieties—became an object of interesting study in their own right. In algebraic geometry the above two objectives are now basically coincident. As to what will happen with the geometric theory of differential equations, it has been said that it is easier to invent the future than to predict it.

[1]In fact, as we shall illustrate below, the theory once again raises the question "What is meant by a *solution* to a differential equation?"



We now proceed to give an outline of the contents of this paper.

In Section 0 we give a brief discussion of differential equations and their equivalence under various natural groups of transformations. In addition, we give a sketch of how one may canonically associate an exterior differential system to a differential equation. We do not belabor the technicalities of this construction as these may be found in the literature (see for instance, *[2]*) and, in any case, they will be explained in the examples.

In Section 1 we study two elementary examples—the first being the geometry associated to the second order ODE

$$\frac{d^2y}{dx^2} = f\left(x, y, \frac{dy}{dx}\right) \tag{1}$$

with the group being induced by coordinate transformations in the $xy$-plane. The geometry associated to this simple example is extremely rich and, with hindsight, may be seen to portend the double fibration construction of twistor theory. The geodesics of the canonical connection associated to (1) turn out to give solution curves to the equation together with a distinguished choice of parameter defined up to a linear fractional transformation; this leads to an intrinsic notion of completeness for solutions of (1). An interesting observation is that the second order ODE above has a natural "dual equation". It turns out that the "curvature" of (1) has two principal components—the vanishing of one (or the other) being equivalent to the condition that the equation (or its dual) is the geodesic equation of a projective connection. É. Cartan observed that in the latter case, generically, equation (1) has a complete set of first integrals expressed rationally in terms of $f$ and its derivatives.

In the Appendix we have worked through the equivalence problem for (1). In fact, this simple example is somewhat subtle and illustrates many aspects of the procedure of associating a geometry to a differential equation.

In our second example, we study the geometry associated to the non-linear scalar PDE

$$u_t + g(x,t,u)u_x = h(x,t,u), \qquad g_u \neq 0. \tag{2}$$

It is observed that this geometry is equivalent to the geometry associated to (1). This leads to an intrinsic notion of completeness for solutions of (2), and of global integral surfaces of the exterior differential system associated to (2). We will show that the sign of the curvature associated to the geometry governs the development of singularities of classical smooth solutions $u(x,t)$ to (2).

Next, in Section 2, we turn to the geometry associated to the non-linear scalar conservation law

$$u_t + \partial_x(F(x,t,u)) = 0, \qquad F_{uu} \neq 0 \tag{3}$$

where $\partial_x$ denotes the total derivative with respect to the $x$ variable. This geometry is a refinement of that associated to (2) since two distinct classes of scalar conservation laws may be equivalent as PDEs. There is again a notion of completeness and of global integral surfaces of the exterior differential system associated to (3). For example, any translation-invariant conservation law

$$u_t + f(u)u_x = 0, \qquad f'(u) \neq 0 \tag{4}$$

where $f : \mathbb{R} \to \mathbb{R}$ is a smooth function defined on all of $\mathbb{R}$, is complete. Again we interpret classical results on the existence or non-existence of global smooth solutions $u(x,t)$ in terms of the sign of the curvature of (3). Finally, for equations such as (4) the classical results on global existence and uniqueness of "shock solutions" (see *[20]*) to the equation are discussed in terms of the geometry of the global, smooth integral surface of the associated exterior differential system.

In the Appendix we discuss the equivalence problem associated to the scalar conservation law (3).

In Section 3.1 we study the exterior differential system associated to a hyperbolic PDE system

$$\begin{aligned} u_y &+ au_x + bv_x + f = 0 \\ v_y &+ cu_x + ev_x + g = 0 \end{aligned} \tag{5}$$

where the coefficients functions are functions of $x, y, u$ and $v$. This exterior differential system is given by the data $(M; \Omega_1, \Omega_2)$ consisting of a 4-manifold $M$ together with a pair of transverse, decomposable 2-forms $\Omega_1, \Omega_2$. This relatively simple structure turns out to have a very rich geometry, the basic aspects of which we explain in the Appendix. For example, many interesting non-linear PDEs (5) are explicitly linearizable as exterior differential systems—i.e., they may be linearized by a suitable contact transformation. In the Appendix we give necessary and sufficient conditions on the torsion and curvature of the geometry associated to (5) that its associated exterior differential system be linearizable.

In Section 1 we discuss conditions on the curvature of (1) that imply that there be a complete set of first integrals of (1). For PDEs one analog of having a complete set of first intergals is that there should exist explicit formulae for the general solution of (5) of the form

$$\begin{aligned} u &= U\big(x, y, \alpha(x), \alpha'(x), \ldots, \alpha^{(k)}(x), \beta(y), \beta'(y), \ldots, \beta^{(k)}(y)\big) \\ v &= V\big(x, y, \alpha(x), \alpha'(x), \ldots, \alpha^{(k)}(x), \beta(y), \beta'(y), \ldots, \beta^{(k)}(y)\big) \end{aligned} \tag{6}$$

where $\alpha$ and $\beta$ are arbitrary functions of one variable. This is closely related to the concept of *Darboux integrability* of the exterior differential system associated to (5), and in Section 3.2 we will express the necessary and sufficient conditions for Darboux integrability at level one (roughly, this is equivalent to the existence of (6) with $k = 1$) in terms of the vanishing of suitable components of the torsion and curvature of (5). This result is then illustrated by a number of examples.



In Section 4.1 we begin a discussion of translation-invariant hyperbolic systems of the form

$$\begin{aligned} u_y &+ a(u,v)u_x + b(u,v)v_x = 0 \\ v_y &+ c(u,v)u_x + e(u,v)v_x = 0. \end{aligned} \tag{7}$$

The exterior differential systems associated to (7) are always linearizable. For a general hyperbolic system (5) we may intrinsically formulate the condition (which we call *characteristic completeness*) that the non-characteristic initial value problem of the associated exterior differential problem admit global, smooth integral surfaces. Then we show that a wide class of equations (7) satisfy the condition to be characteristically complete, and we illustrate this result in several examples.

In Section 4.2 we tie in the classical concept of genuine non-linearity of (7) with the discussion of the elementary equation (2). Then we derive, from our perspective, the well-known result that genuinely non-linear systems (7) do not admit global smooth solutions with compactly supported initial data. On the other hand we will show that in this case there exists a unique, global smooth integral surface for the initial value problem of the associated exterior differential system.

Finally, in Section 4.3 we discuss the case where (7) is given as a hyperbolic system of conservation laws

$$\begin{aligned} u_y &+ \partial_x\big(f(u,v)\big) = 0 \\ v_y &+ \partial_x\big(g(u,v)\big) = 0. \end{aligned} \tag{8}$$

As mentioned above, for smooth non-characteristic initial data there is a unique, global smooth integral surface $S$ for the associated exterior differential system. For $\mathbb{R}^2_+ = \{(x,y) : y \geq 0\}$ the mapping

$$\pi : S \to \mathbb{R}^2_+ \tag{9}$$

is shown to be proper. We may think of $S$ as the "geometric" solution of (8).

On the other hand, it is known *[20]* that under suitable conditions on the initial data and with the assumption that (8) is genuinely non-linear, there is a unique global shock solution. It is of interest to compare the geometric and shock solutions.

For a scalar conservation law, the shock solution arises by taking a suitable cross-section of (9). We may say that the geometric solution "captures" the shock solution.

For a system of conservation laws, the situation is more complex. It turns out that on the $uv$-plane there are two naturally defined pairs of families of curves. The first pair consists of the two foliations $\rho_\lambda$, $\rho_\mu$ given by the level sets of the Riemann invariants. The other is the two families $\delta_\lambda$, $\delta_\mu$ of jump curves obtained by imposing the Rankine-Hugonoit (or jump) conditions. At each point of the $uv$-plane the $\delta$-curves and $\rho$-curves osculate to second, but generally not to third, order.

In general, the two families of $\delta$-curves do *not* define a pair of (local) foliations of the $uv$-plane—the relation defined by a $\delta$-curve is symmetric and reflexive but not necessarily transitive. The condition that $\delta_\lambda$ and $\delta_\mu$ define a pair of foliations is seen to be given by a pair of second-order PDEs for the defining functions $f(u,v)$ and $g(u,v)$ of the conservation laws (8). (These equations may be explicitly solved.) We shall call such hyperbolic systems of conservation laws *special*. For special systems, the $\delta$-curves and $\rho$-curves coincide since two foliations which are everywhere tangent must coincide.

We formulate and study the *Riemann problem* for the exterior differential system associated to (8). For this case, it will be shown that the geometric solution "captures" the shock solution if, and only if, the system is special. Special systems form a remarkable class of hyperbolic conservation laws—for arbitrary smooth initial data, there appears to be a unique shock solution to each of these systems arising as a cross-section of (9); in particular, the shock solution should be piecewise smooth if the initial data is piecewise smooth. On the other hand, for non-special systems, it seems necessary to perform some sort of "PDE surgery" on $S$ in order to obtain an integral surface that captures the shock solution. We conclude the paper with a brief discussion of this point.

## 0 Preliminaries

Recall that a first order PDE system may be written in coordinates as

$$F^\mu(x^i,\, u^\alpha(x),\, \partial u^\alpha(x)/\partial x^i) = 0, \qquad \mu = 1,\dots,m \tag{1}$$

where $x = (x^1,\dots,x^n) \in X \subset \mathbb{R}^n$ and $u = (u^1,\dots,u^s) \in U \subset \mathbb{R}^s$ are the independent and dependent variables respectively. Introducing variables $p_i^\alpha$ to stand for the derivatives $\partial u^\alpha/\partial x^i$, we consider the locus in $xup$-space defined by

$$M = \{F^\mu(x^i,\, u^\alpha,\, p_i^\alpha) = 0, \quad \mu = 1,\dots,m\}.$$

We assume $M$ to be a smooth manifold. Clearly, a graph mapping $X \to M$ given by $x \mapsto (x, u(x), p(x))$ is a solution to (1) if, and only if, the differential forms

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

restrict to be zero on $M$. This suggests that we more generally consider immersions

$$f : N \to M$$

of a manifold $N$ into $M$ which satisfy

$$f^*\theta^\alpha = 0, \qquad \alpha = 1,\dots,s. \tag{2}$$

Since

$$f^* : \Omega^*(M) \to \Omega^*(N)$$

is a mapping of differential graded algebras, i.e., $f^*$ commutes with exterior derivative and wedge products, we should consider the *differential ideal* $\mathcal{I}$ in $\Omega^*(M)$ generated by the $\theta^\alpha$. Immersions satisfying

$$f^*\theta = 0, \qquad \theta \in \mathcal{I}$$



are the same as those satisfying (2) and are called *integral manifolds* of the differential ideal $\mathcal{I}$. The pair $(M, \mathcal{I})$ consisting of a manifold $M$ together with a differential ideal $\mathcal{I} \subset \Omega^*(M)$ is by definition an *exterior differential system.*

This construction of associating an exterior differential system to a PDE system given in coordinates may be extended to higher order systems, either directly using higher order jet coordinates and the associated contact forms (see, for instance, [2]) or by rewriting the higher order system as a (larger) first order system.

How invariant is the above coordinate construction of an exterior differential system associated with a PDE system? In an increasing order of generality there are four natural types of change of variables associated to (1):

**Classical transformations:** These are induced by a change $\bar{x} = \bar{x}(x)$, $\bar{u} = \bar{u}(u)$ of independent and dependent variables separately. Geometrically we are viewing (1) as imposing conditions on the differential of a mapping

$$\begin{array}{ccc} X & \longrightarrow & U \\ \cap & & \cap \\ \mathbb{R}^n & & \mathbb{R}^s. \end{array}$$

**Gauge transformations:** These are induced by a change of variables of the form $\bar{x} = \bar{x}(x)$, $\bar{u} = \bar{u}(x,u)$. Here we are viewing (1) as imposing differential conditions on cross sections of a fibration

$$\begin{array}{c} U \subset \mathbb{R}^{n+s} \\ \downarrow \\ X \subset \mathbb{R}^n. \end{array}$$

**Point transformations:** These are induced by the following change of variables $\bar{x} = \bar{x}(x,u)$, $\bar{u} = \bar{u}(x,u)$. Now we are viewing (1) as imposing differential conditions on an immersion

$$\begin{array}{ccc} X & \longrightarrow & \mathbb{R}^{n+s} \\ \cap & & \\ \mathbb{R}^n & & \end{array}$$

where we have happened to write the immersion in (1) locally in the form of a graph $(x^i) \mapsto (x^i,\, u^\alpha(x))$.

In all the above transformations, the change $\bar{p} = \bar{p}(x,u,p)$ is induced by differentiating the change among the $x, u$ variables. The last type of transformation we shall consider is of a different character.

**Contact transformations:** These are induced by a change of variables $\bar{x} = \bar{x}(x,\, u,\, p)$, $\bar{u} = \bar{u}(x,\, u,\, p)$, and $\bar{p} = \bar{p}(x,\, u,\, p)$ which map $M$ to itself and preserve the differential ideal generated by the $\theta^\alpha$.

More or less by construction, the association of an exterior differential system to a PDE system is contact invariant. For a given problem we may therefore choose the group $\Gamma$ to be generated by the classical, gauge, point or contact transformations of (1). Under point transformations there is no notion of "independent" and "dependent" variables. More generally, under contact transformations there is not even a notion of "base" variables. To illustrate the difference, we note that under contact transformations the following equations for a function $u(x,\, y)$:

$$u_{xx} - u_{yy} = 0 \qquad \text{and} \qquad u_{xx}u_{yy} - u_{xy}^2 = -1$$

are globally equivalent even though there is no change of variables in $xyu$-space which will convert one of these equations into the other. Special contact transformations (Legendre transformation, hodographic method) have of course been classically used to study differential equations (see, for instance, *[12]*). But the systematic utilization of the invariants of a differential equation under any of the above groups (as in Klein's Erlangen Program) has not been so much an aspect of the modern theory of differential equations. In the following we shall illustrate the use of each of the above groups.

# 1 Two Elementary Examples and their Geometries

In this section, with the help of two elementary examples, we shall show how the geometry associated with a differential equation might be used to deduce interesting properties of the solutions of the equation. For both examples, the geometry of the defining equation have been worked out in the Appendix.

EXAMPLE 1: *The second-order ODE.* Consider the differential equation

$$\frac{d^2y}{dx^2} = f\left(x,\, y,\, \frac{dy}{dx}\right). \tag{1}$$

Of course, such equations have been objects of study for over three centuries, with a primary goal being to describe the solutions in a reasonably explicit way. The classical theory generally concentrates on constructing "first integrals" of the equation, with the most common technique being that of assuming some sort of symmetry, so that some version of Noether's Theorem can be applied to yield a conserved quantity, i.e., a first integal.

We shall study (1) under the equivalence relation given by point transformations. Thus, we shall think of (1) as a second-order ODE for curves on a surface $S$, where $x$, $y$ are local coordinates on $S$ and (1) is the ODE for those curves $(x,\, y(x))$ which are graphs over the $x$-axis.

To study (1) geometrically we consider the exterior differential system $(\mathcal{I}, \underline{\alpha})$ in $xyp$-space, henceforth denoted by $M$, generated by the 1-forms

$$\underline{\theta} = dy - p\,dx \qquad \text{and} \qquad \underline{\omega} = dp - f(x,\, y,\, p)\,dx$$



with independence condition given by the 1-form $\underline{\alpha} = dx \neq 0$. Integral curves $c : \mathbb{R} \to M$ of the differential system

$$\underline{\theta} = \underline{\omega} = 0$$

satisfying the transversality condition $c^*(\underline{\alpha}) \neq 0$ are in one-to-one correspondence with solution curves of (1).

In intrinsic terms, $M = \mathbb{P}(TS)$ is the projectived tangent bundle of $S$ and $x$, $y$, $p$ are local coordinates in the open set of tangent directions $\xi$ for which $dx(\xi) \neq 0$. The equations $\underline{\theta} = \underline{\omega} = 0$ define a field of line elements on $M$ whose integral curves project to $S$ to give solutions to (1).

By construction, $M$ is a contact manifold with contact structure defined by the 1-form $\underline{\theta}$. Diffeomorphisms of $M$ which preserve the differential ideal generated by $\underline{\theta}$ are classically called *contact transformations.* Under such transformations, a century-old theorem of Sophus Lie states that every ODE of the form (1) is locally equivalent to the "flat" model

$$\frac{d^2y}{dx^2} = 0.$$

Although this is an interesting geometric fact, it is well known that finding a contact transformation taking (1) into the flat model involves solving a PDE for the generating function of the contact transformation. Thus Lie's result is generally of little help in explicitly solving (1).

In the following, we shall instead study the geometry of (1) with respect to the smaller pseudo-group $\Gamma$ of those contact transformations of $M$ which preserve the differential ideal $\mathcal{J}$ generated by the 1-forms

$$\underline{\alpha} = dx \qquad \text{and} \qquad \underline{\theta} = dy - p\,dx.$$

Clearly this is the same as the ideal generated by $dx$ and $dy$ and so such contact transformations are in fact induced by point transformations. Following Cartan, we will explain the consequences of a geometric construction that singles out a large class of second-order ODEs that may be (globally) explicitly solved, in spite of there being no assumption of symmetry. This will illustrate how a geometric study of a differential equation can provide new perspectives on solutions to the equation.

The starting point is to observe that the ideals $\mathcal{I}$ and $\mathcal{J}$ on $M$ determine a canonical double fibration

$$\begin{array}{ccc} & M & \\ {}^{\lambda}\swarrow & & \searrow^{\rho} \\ S & & \Sigma \end{array}$$

where $\ker \lambda_* = \{\underline{\alpha}, \underline{\theta}\}^\perp$ and $\ker \rho_* = \{\underline{\theta}, \underline{\omega}\}^\perp$ define a pair of line fields on $M$ with $S$ and $\Sigma$ the locally defined quotient surfaces. Geometrically, viewing (1) as defining a 2-parameter family $\Sigma$ of paths on a surface $S$—each point $s \in S$ and each tangent vector $\xi \in T_sS$ determining a unique path $c \in \Sigma$—the quotient surface $\Sigma$ may be viewed as the "space of solutions" of the second-order ODE (1) on $S$.

In local coordinates $x, y$ on $S$ and $\sigma, \tau$ on $\Sigma$, the solutions of (1) may be written as

$$F(x, y, \sigma, \tau) = 0 \tag{2}$$

where for each fixed $\sigma, \tau$ the above relation gives a solution curve to (1)—i.e., $\sigma, \tau$ are constants of integration. Explicitly, the equation $F = 0$ describes the image of the mapping $(\lambda, \rho) : M \to S \times \Sigma$.

Note that this geometric picture is symmetric with respect to $S$ and $\Sigma$. In other words, while each point of $\Sigma$ represents a curve in $S$, dually, each point $p \in S$ represents a curve $\rho(\lambda^{-1}(p))$ in $\Sigma$. It is easy to see that the 2-parameter family of curves in $\Sigma$ defined by (2) are the solutions to a certain second-order ODE in the $\sigma\tau$-plane (i.e., $\Sigma$)

$$\frac{d^2\sigma}{d\tau^2} = \varphi\left(\sigma, \tau, \frac{d\sigma}{d\tau}\right). \tag{3}$$

The relation between (1) and (3) is symmetric. In the classical literature, these two second order ODEs were said to be "dual" equations. By construction, functions on $\Sigma$ are first integrals of (1), and dually, functions on $S$ are first integrals of (3).

The goal of Cartan's approach was to find an explicit procedure for finding functions on $M$ which are the pull back of (non-constant) functions on $\Sigma$. Clearly, the level sets of such functions implicitly define the solution curves of (1). More exactly, Cartan wanted to describe classes of equations (1) for which such procedures existed.

Cartan's method is quite general, but here we want to describe one particular such class of equations. Recall that the geodesics of a Riemannian metric on $S$ are locally given by an ODE (1) where $f$ is a cubic polynomial in $dy/dx$. More generally, Cartan *[6]* has defined the concept of a *normal projective connection*, and shown that the geodesics of such a connection satisfy an equation of the form (1) where

$$\frac{\partial^4 f}{\partial p^4} = 0. \tag{4}$$

Conversely, every such equation can be realized as the geodesic equation of a unique normal projective connection on $S$. Curvature invariants of this connection then produce functions on $S$ which are generally non-constant.

Cartan *[7]* showed that the condition on (1) that its dual equation (3) describe the geodesics of a projective connection on $\Sigma$ is that the function $f$ satisfy the differential equation

$$\frac{d^2}{dx^2}(f_{pp}) - 4\frac{d}{dx}(f_{py}) + f_p\left(4f_{py} - \frac{d}{dx}(f_{pp})\right) - 3f_y f_{pp} + 6f_{yy} = 0 \tag{5}$$

where $d/dx$ denotes total derivative with respect to $x$. Moreover, he showed that, in this case, starting with only the knowledge of the ideals $\mathcal{I}$ and $\mathcal{J}$ on $M$, which, as we have seen, are completely determined by equation (1), there is an algorthmic procedure to compute the curvatures of the projective connection on $\Sigma$, thus yielding first integrals of the original equation (1). It follows that the class of equations (1) satisfying (5) should be *geometrically integrable* in the sense that a complete set of first integrals (or conservation laws) may be determined explicitly from $f$ and its derivatives.[2] Thus, Cartan *[7]* was able to state the following result:

> *Every second-order ODE (1) satisfying (5) may be integrated by a process involving only algebraic operations, differentiation, and quadratures.*

**Remark:** Strictly speaking, Cartan's statement is not quite correct. It turns out that, up to equivalence, there is a 1-parameter family of such equations whose integration by Cartan's method requires solving a Ricatti equation, which is well-known not to be solvable by quadrature. (This is caused by the fact that the projective curvatures on $\Sigma$ turn out to be constant.) This expectional family of equations is uniquely characterized by the condition that its symmetry group is $\mathrm{SL}(2, \mathbb{R})$ (see *[5]* for more details).

The actual computation of the first integrals of the above equations requires a detailed study of the $G$-structure associated with the above double fibration. (The $G$-structure construction for (1) is described in the Appendix.) Associated to the $G$-structure is an intrinsic connection $\phi$ whose curvature $\Phi$ has two principal components, which we denote by $K_1$ and $K_2$. These components have the property that $K_2$ vanishes if, and only if, the ODE (1) is the geodesic equation of a projective connection $\phi_S$ on $S$ and $K_1$ vanishes if, and only if, the ODE (3) is the geodesic equation of a projective connection $\phi_\Sigma$ on $\Sigma$. Furthermore, the curvatures of $\phi_\Sigma$ may be computed in terms of $\Phi$ and therefore in terms of the original equation (1). In fact, they may be expressed rationally in terms of $f$ and its derivatives. The functions on $M$ constructed from the curvatures of $\phi_\Sigma$ together with their covariant derivatives will then generate the desired set of first integrals of (1).

Cartan's approach to integrating differential equations is rather general. It has been extended to third order ODEs by Chern *[10]* and to fourth order equations by Bryant *[1]* in connection with his work on exceptional holonomies. In fact, Chern's paper directly anticipates several aspects of twistor theory while Bryant's makes explicit reference to it. To paraphrase a famous remark, one may see in the double fibration picture a "piece of twistor theory that fell into the early 20th century".

EXAMPLE 2: *Scalar first order PDE.* We shall now discuss the geometry associated to a first order quasi-linear PDE

$$u_t + g(x, t, u)u_x = h(x, t, u). \tag{6}$$

This equation is usually introduced at the beginning of introductory textbooks on PDE theory and one may well wonder what is new and interesting that the geometry

[2] In recent years there has been much interest in "integrable systems". The Cartan approach provides a complementary, perhaps more directly geometric approach to integrating a differential equation.

associated to (6) can tell us? We shall show that when (6) is *genuinely non-linear* (a condition defined in the PDE literature on this equation to mean that $g_u \neq 0$), the geometry associated to (6) under equivalence by gauge transformations is equivalent to the geometry associated to the ODE (1) above under equivalence by point transformations. Having already worked out the latter geometry we shall then be able to apply the results to draw conclusions for (6).

To explain this identification we return briefly to EXAMPLE 1. We may equivalently give the above data $(M; \mathcal{I}, \mathcal{J})$ of the 3-manifold $M$ with Pfaffian systems $\mathcal{I}$, $\mathcal{J}$ as the data $(M; \Phi, \Omega)$ consisting of the 3-manifold $M$ together with a pair of everywhere linearly independent 2-forms $\Phi$, $\Omega$, defined up to non-zero multiples and satisfying the non-degeneracy condition that the field of 2-planes $\Phi^\perp{\wedge}\Omega^\perp$ be a contact structure. To explain this we observe that $\Phi^\perp$ and $\Omega^\perp$ are linearly independent line fields on $M$, and therefore span a field $\Phi^\perp{\wedge}\Omega^\perp$ of 2-plane elements. Our non-degeneracy condition is that this rank 2 distribution be non-integrable. Given $\mathcal{I}$ and $\mathcal{J}$ as in EXAMPLE 1 where locally $\mathcal{I}$ is generated by 1-forms $\theta$, $\omega$ and $\mathcal{J}$ is generated by 1-forms $\alpha$, $\theta$, we may set

$$\Phi = \theta\wedge\omega \qquad \text{and} \qquad \Omega = \alpha\wedge\theta.$$

Then $\Phi^\perp{\wedge}\Omega^\perp$ is the field of 2-planes defined by $\theta = 0$ and the non-degeneracy condition is $\theta{\wedge}d\theta \neq 0$. The construction

$$(M; \mathcal{I}, \mathcal{J}) \to (M; \Phi, \Omega)$$

is easily seen to be well-defined and reversible.

We now proceed to show that the PDE (6) gives rise to a $(M; \Phi, \Omega)$ structure. Letting $M$ denote $xtu$-space, we consider the 2-forms on $M$ defined by

$$\Phi = (du - h\,dt)\wedge(dx - g\,dt) \qquad \text{and} \qquad \Omega = dx\wedge dt.$$

The motivation for introducing $\Phi$ is that on graph surfaces $(x, t) \mapsto (x, t, u(x, t))$ the exterior equation

$$\Phi = 0$$

is equivalent to the PDE (6). Thus the exterior differential system $\Phi = 0$ with independence condition $\Omega \neq 0$ models (6).

Clearly the 2-forms $\Phi$ and $\Omega$ are well-defined up to non-zero factors under gauge transformations. Furthermore, since $\Omega = (dx - g\,dt){\wedge}dt$, we see that $\Phi$ and $\Omega$ have the common linear factor

$$\theta = dx - g\,dt.$$

Our assumption of genuine non-linearity now implies that

$$\theta\wedge d\theta = g_u\, du\wedge dx\wedge dt \neq 0$$

and so $\theta$ defines a contact structure on $M$. Thus, to the gauge-equivalence class of non-linear PDEs (6) is associated a $(M; \Phi, \Omega)$ structure, and conversely.



Now, in view of the above identification of the two geometries associated with (1) and (6), the $G$-structure construction for (1) now associates to the gauge-equivalence class of (6) a principal $G$-bundle $(B_G \to M, \phi)$ with connection $\phi$. In the following we shall discuss how the geometry $(B_G \to M, \phi)$ determines the global behavior of solutions of (6).

Here, *global* has the following intrinsic meaning: Associated to the connection $\phi$ are its geodesics, which are a special class of curves in $B_G$ equipped with a distinguished parameter $\tau$ (see the Appendix for details). Two geodesics in $B_G$ that project to the same curve $\gamma$ in $M$ have their parameters related by a linear fractional transformation

$$\tilde{\tau} = \frac{a\tau + b}{c\tau + d}, \qquad ad - bc = 1.$$

We shall say that $\gamma$ is complete in case its parameter $\tau$ takes values in all of $\mathbb{R}$ viewed as a subset of the projective line $\mathbb{P}$. The geometry $(B_G \to M, \phi)$ is then defined to be *complete* in case all such curves $\gamma$ are complete.

In the case of the ODE (1) above, the curves $\gamma$ are the canonical lifts to $\mathbb{P}(TS)$ of solution curves to (1). In the case of the non-linear PDE (6), the curves $\gamma$ are the characteristic curves (to be explained momentarily). In both cases the completeness of the curves $\gamma$ gives an intrinsic meaning to the concept of a *global solution* to (1) or (6). In examples this concept turns out to agree with what one usually thinks of as a global solution to a differential equation.

We now formulate the initial value problem of the PDE (6). Recall that associated to (6) is the *characteristic vector field*

$$X = \partial_t + g\,\partial_x + h\,\partial_u$$

on $M$. Invariantly, $X$ is the vector field uniquely defined up to scaling by the relation $X \lrcorner \Phi = 0$. We note that $X$ depends only on the exterior differential system $\Phi = 0$ and not on the independence condition $\Omega$.

Non-characteristic initial data is given by an immersed curve $\Gamma$

$$s \mapsto (x(s), t(s), u(s))$$

which is nowhere tangent to $X$. Flowing $\Gamma$ along integral curves of $X$ then generates the general solution surface $N$ to the exterior differential system $\Phi = 0$. At points of $N$ where $\Omega \neq 0$ the mapping

$$N \to xt\text{-plane}$$

is locally one-to-one. Thus, near such points, $N$ is locally given as a graph

$$(x, t) \mapsto (x, t, u(x, t))$$

where $u(x, t)$ is a solution to (6). This is the method of characteristics, here illustrated in its simplest form.

Now what does the geometry $(B_G \to M, \phi)$ tells us about (6)? First, in the flat case, where the curvature of $\phi$ is zero, working through the identification of the geometries associated to (1) and (6) shows that the PDE (6) is gauge equivalent to the (inviscid) Burgers' equation

$$u_t + uu_x = 0. \tag{7}$$

If, in this case, we consider the corresponding exterior differential system

$$\Phi = du \wedge (dx - u\,dt) = 0 \tag{8}$$

on $M$, then there are global solutions to (8) as follows: Given an initial curve

$$s \mapsto \big(s,\, 0,\, u_0(s)\big), \qquad s \in \mathbb{R},$$

we may uniquely extend this to a mapping of $\mathbb{R}^2$ giving an integral surface of (8) by

$$(s,\, t) \mapsto \big(s + tu_0(s),\, t,\, u_0(s)\big).$$

This solution may be said to be global for the geometric reason that the characteristic curves $s =$ const are geodesics for the connection $\phi$ that are *complete* in the sense described above. The question of whether the projection $N \to \mathbb{R}^2$ is one-to-one or onto will be taken up in the next section when we discuss the "shock" behavior of (7).

What does this perfectly global solution to (8) say about classical solutions to (7)? We let $\mathbb{R}^2_+$ and $\mathbb{R}^3_+$ denote the half-space $t \geq 0$ in $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively, and we have the picture

$$\begin{array}{ccccc} \Gamma \subset & N & \subset \mathbb{R}^3_+ & \ni & (x,\, t,\, u) \\ {\scriptstyle \delta_0}\uparrow & \downarrow{\scriptstyle \pi} & & & \downarrow \\ \mathbb{R} \subset & \mathbb{R}^2_+ & & \ni & (x,\, t) \end{array}$$

where $\mathbb{R}$ is the $x$-axis $t = 0$ and the initial data is given by $\delta_0$, and we seek to extend $\delta_0$ to as large a neighborhood of the initial curve $t = 0$ as possible. To see that we may in general expect difficulties we note that up on $N$

$$\pi^*(dx \wedge dt) = \big(1 + tu_0'(s)\big)\, ds \wedge dt. \tag{9}$$

Thus, if there are points on the initial curve where $u_0'(s) < 0$ (which will certainly happen if $u_0(s)$ is compactly supported and not identically zero) then $\delta_0$ cannot be extended to a smooth map $\delta$ defined on all of $\mathbb{R}^2_+$. This is also evident from the explicit formula

$$\delta(x,\, t) = \big(x,\, t,\, u(x,\, t)\big)$$

where $u(x,\, t)$ is defined implicitly by the equation

$$u = u_0(x - tu). \tag{10}$$



Returning to the general discussion we ask what is the meaning of the vanishing of the curvatures $K_1$, $K_2$ of $\phi$? The condition $K_1 = 0$ is sufficient to guarantee that the characteristic equation associated to $\Phi = 0$ has a complete set of first integrals. Thus the characteristic curves of the PDE (6) may be found explicitly. In this case we may then write down an implicit equation for $u$ which solves the initial value problem, as was done in (10) for Burgers' equation.

Perhaps more interesting is the condition $K_2 = 0$, which is the case when the second order ODE (1) associated to (6) is the equation of geodesics of a projective connection on a surface. For example, suppose that (6) is

$$u_t + f(u)u_x = h(x,\, t,\, u).$$

Our non-linearity assumption is $f'(u) \neq 0$, and so we may write

$$h(x,\, t,\, u) = H(x,\, t,\, u)/f'(u)$$

for some function $H$. The condition $K_2 = 0$ then implies that

$$H(x,\, t,\, u) = A + Bf(u) + Cf(u)^2 + Df(u)^3.$$

Geometrically the ODE associated to (6) arises as follows: In the picture

$$\begin{array}{ccc} \mathbb{R}^3 \ni & (x,\, t,\, u) \\ \downarrow{\scriptstyle \pi} & \downarrow \\ \mathbb{R}^2 \ni & (x,\, t) \end{array}$$

the projection to $\mathbb{R}^2$ of the characteristic curves of this equation will, by our non-linearity assumption, generate a 2-parameter family of curves which are the solutions of an ODE of the form (1) in $\mathbb{R}^2$. For instance, in the case $K_2 = 0$ as above the ODE is

$$\frac{d^2x}{dt^2} = A + B\left(\frac{dx}{dt}\right) + C\left(\frac{dx}{dt}\right)^2 + D\left(\frac{dx}{dt}\right)^3. \tag{11}$$

The initial value problem for (6) then has the following meaning: Along an initial curve $\gamma$

$$s \mapsto \big(x(s),\, t(s)\big)$$

in the $xt$-plane we prescribe the initial values $(dt/dx)(s)$. Then we take the integral curves of (1) emanating from $\gamma$ with the given initial values. As long as these curves don't cross or focus, we will have a classical solution to (6). In fact, *it is exactly the focusing of geodesics that gives classical development of singularities*. This phenomenon is well-known in the study of wave fronts and the Hamilton-Jacobi equation.

Thus for example, suppose that (11) is the geodesic equation for a complete Riemannian metric in the $xt$-plane whose Gauss curvature $K \geq 0$. (For Burgers' equation (7) this is the situation, with $K = 0$.) Suppose also that $\gamma$ is a geodesic and that there is a sub-interval of $\gamma$ along which the angle of the initial data is decreasing. (This is the geometric form of the assumption $u_0' < 0$ in Burgers' equation.) Then the geodesics which represent the projections of the characteristic curves on the surface $N$ will focus in the $xt$-plane and thus singularities will develop. If, however, the Gauss curvature $K$ is strictly negative and the initial data is appropriately slowly varying, then the spreading of geodesics as given by the Jacobi equation will imply that there is no focussing and global solutions to the PDE will exist. Pictorially, moderate ocean waves will not break along a negatively curved beach.

## 2 The Scalar Conservation Law

In EXAMPLE 2 above we considered the geometry associated to the scalar first order PDE

$$u_t + g(x, t, u)u_x = h(x, t, u). \tag{1}$$

We assumed that $g_u \neq 0$ (i.e., genuine non-linearity) and saw that the geometry of (1) was equivalent to that of a second order ODE

$$\frac{d^2y}{dx^2} = f\left(x, y, \frac{dy}{dx}\right). \tag{2}$$

The initial value problem and development of singularities of solutions to (1) then had interpretations in terms of the "path geometry" defined by (2).

In this section we will look more deeply into singularity development and so-called "shock solutions" in terms of a geometry associated to a special subclass of equations (1)—the ones which can be written as a scalar conservation law

$$u_t + \partial_x\big(F(x,t,u)\big) = 0.$$

To motivate the introduction of this class of equations, we assume given a PDE (1) whose associated geometry is complete. Given a non-characteristic initial curve $\Gamma$, we sweep out a surface $S$ by flowing $\Gamma$ along the characteristic vector field $X$. This constructs $S$ as an integral surface of the exterior differential system $\Phi = 0$ associated to (1). If $\Gamma$ is given in coordinates by $s \mapsto \big(s,\, 0,\, u_0(s)\big)$ and we set $\mathbb{R}^2_+ = \{(x,\, t) : t \geq 0\}$, then suitable conditions on the initial data (see below) plus the completeness of $X$ will insure that

$$\pi : S \to \mathbb{R}^2_+ \tag{3}$$

is surjective. Singularities of a solution $u(x,\, t)$ to (1), defined say for $0 \leq t < t_0$, will then develop in relation to the projection in the $xt$-plane of the *fold locus* $\mathcal{F}$ where the differential $\pi_*$ drops rank. We may seek to understand this situation by applying singularity theory to the projection $\pi$.[3]

[3]The idea of multi-valued solutions to PDEs is of course classical. The issue is to show in examples that these exist and have interesting properties. In this regard we would like to call attention to the recent interesting paper *[9]* by Caflish et al which studies the branching behavior of multi-valued solutions to certain non-linear PDEs.



Typically $\mathcal{F}$ will be a union of curves, with the typical one looking something like a classical "cusp" (Figure 2.1). Over the lines $t = t_-$, $t = t_0$, $t = t_+$ the graphs of the projection will resemble Figure 2.2. More generally the fold locus will project to a curve that might contain several such cusps (Figure 2.3) with the following cross-sections of $S$ over the lines $t = t_i$ (Figure 2.4).

So far, the whole picture is gauge invariant. However, if, as is commonly done in the PDE literature, we seek to define a "generalized solution" to (1) as a cross-section (Figure 2.5) satisfying certain local conditions (which are typically *not* invariant under gauge transformations), then we need more information than that given by the gauge equivalence class of (1). To determine the cross-section $u$, we need to know the *breaking curve* along which $u$ jumps from one sheet of $S$ to another, and we also need to know to which sheet $u$ will jump as it crosses this curve. This data cannot be determined from the gauge equivalence class of (1), so we will consider a special subclass for which a more restrictive geometry can be defined.

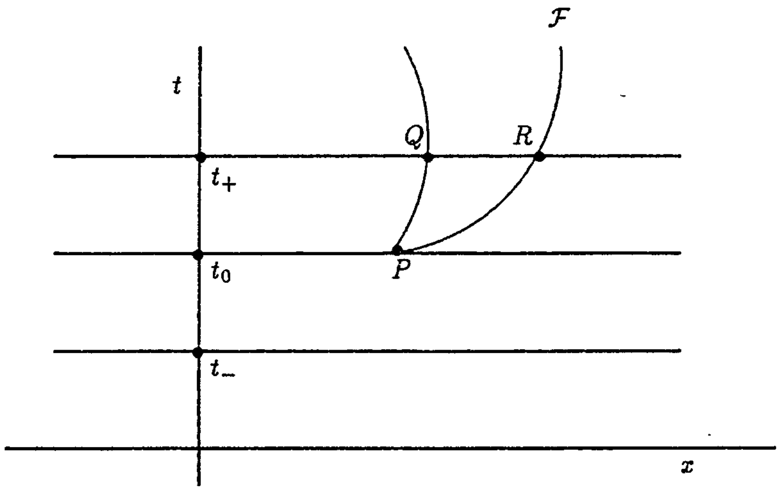


**Figure 2.1**

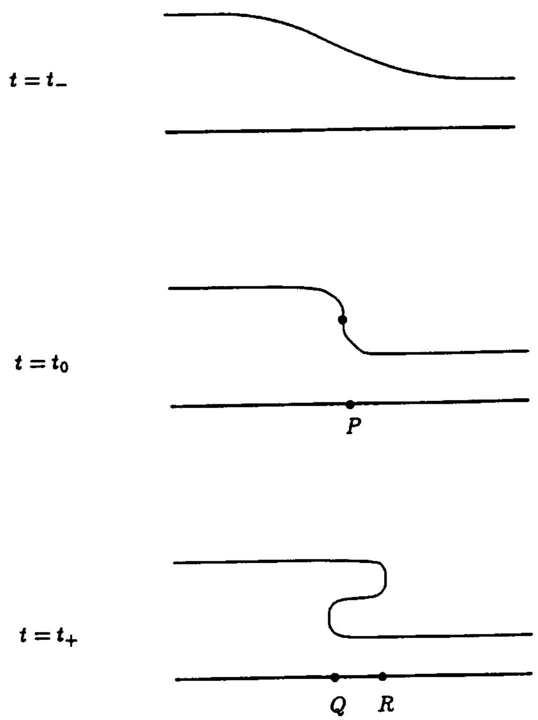


**Figure 2.2**

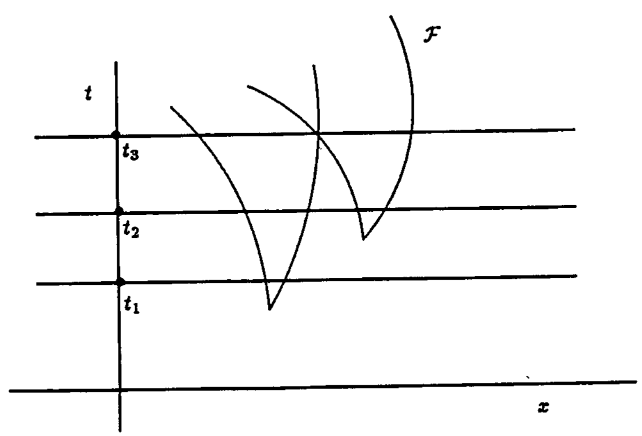


**Figure 2.3**



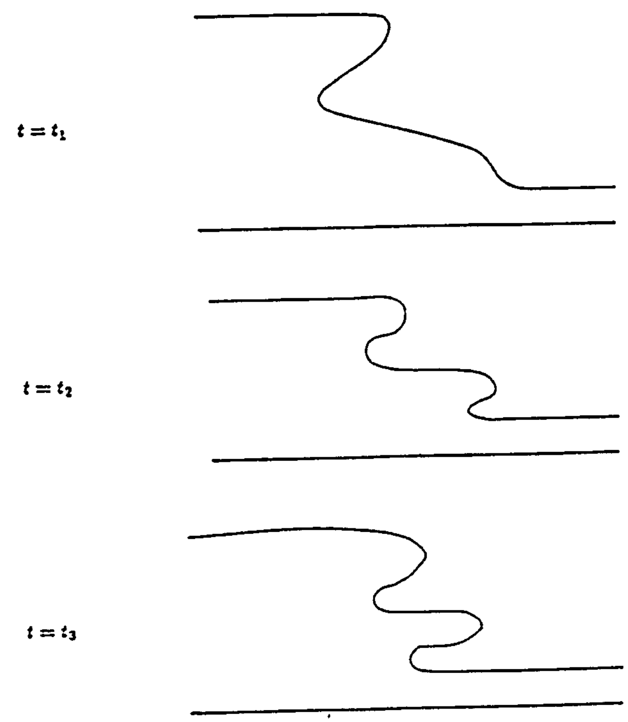


**Figure 2.4**

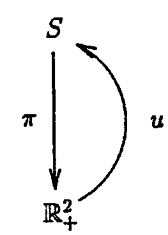


**Figure 2.5**

EXAMPLE 3: *The scalar conservation law.* We now discuss the geometry associated to a scalar conservation law

$$u_t + \partial_x(F(x, t, u)) = 0 \tag{4}$$

where $\partial_x$ denotes total derivative with respect to $x$. This is an equation of the form (1) where $g = F_u$ and $h = -F_x$. The associated exterior differential system is

$$0 = \Phi = (du + F_x\,dt) \wedge (dx - F_u\,dt) = d\big(u\,dx - F(x,t,u)\,dt\big).$$

Note that $\Phi$ is an *exact* 2-form.[4]

This suggests that we try to construct a geometry from the data $(M; \Phi, \Omega)$ consisting of a 3-manifold $M$ endowed with two linearly independent 2-forms $\Phi$ (which is closed) and $\Omega$ which is invariant under diffeomorphisms which preserve $\Phi$ and preserve $\Omega$ up to a non-vanishing factor. We shall always make the non-degeneracy assumption on our data that the common linear factor of $\Phi$ and $\Omega$ define a contact structure on $M$. For simplicity of exposition, we assume that $M$ is connected and simply connected.

In the Appendix we have worked out the corresponding $G$-structure problem for $(M; \Phi, \Omega)$. It turns out that there is a principal relative invariant $K$, and when $K = 0$ there is a secondary relative invariant $L$. The geometry of $(M; \Phi, \Omega)$ thus divides into three cases:

(i) $K \neq 0$; (ii) $K = 0, L \neq 0$; (iii) $K = L = 0$.

In all three cases we will see that there is a canonical *affine* connection on $M$ such that the characteristic curves of (4) are geodesics, with a canonical parameter $\tau$ defined up to an affine transformation $\tau \to a\tau + b$. We shall then assume that this geometry is *complete*. An integral surface $S$ of the differential system $\Phi = 0$ with initial curve $\Gamma$ as above may then be said to be *global* in the intrinsic sense that each characteristic curve in $S$ is defined for $\tau \in [0, \infty)$, where $\tau = 0$ gives a point on $\Gamma$.

**Definition:** Let $S$ be an integral surface for $\Phi = 0$. We will say that $S$ has a *singularity* at $p \in S$ in case $\Omega$ pulled back to $S$ vanishes at $p$.

The points of $S$ where $\Omega(p) = 0$ are precisely where the differential of $\pi$ in (3) drops rank. Thus, when the surface $S$ constructed from $\Gamma$ has singularities in this sense, a global classical solution to (4) with the given initial data will not exist. We will now establish the following result:

*Let $\Gamma$ be a non-characteristic initial data for $(M; \Phi, \Omega)$ and let $\lambda_0(s)$ be the real-valued function on $\Gamma$ constructed from $\Gamma$ and $(M; \Phi, \Omega)$ below. There is a curvature function $Q$ associated with the geometry of $(M; \Phi, \Omega)$ such that if*

$$Q \leq 0, \qquad \textit{and} \qquad \lambda_0(\underline{s}) > 0$$

*for some $\underline{s}$, then the global integral surface of the exterior system $\Phi = 0$ with initial data $\Gamma$ develops singularities. On the other hand, assuming that the curvatures of $(M; \Phi, \Omega)$ are constants, if the curvature function $Q > 0$ and the initial data $\lambda_0(s)$ is small relative to $Q$, then there exists a global solution to the PDE which $(M; \Phi, \Omega)$ models.*

[4] More generally we may consider equations $\partial_t\big(F(x, t, u)\big) - \partial_x\big(G(x, t, u)\big) = 0$, expressing the condition that the 1-form $\varphi = F\,dx + G\,dt$ be closed on solutions.



These results are similar to known statements in PDE theory (although there may not be an exact analog of the existence result); our main point is that they may be intrinsically formulated in terms of the geometry associated to the PDE.

Intuitively, the function $\lambda_0$ measures the tangent of the "angle" of the unique integral element at each point of $\Gamma$ relative to $\Omega$. Thus $\lambda_0(p) = \infty$ is equivalent to the condition $\Omega(p) = 0$.

A special case occurs when (4) is a translation invariant conservation law

$$u_t + f(u)u_x = 0, \qquad f'(u) > 0. \tag{5}$$

In this case, the curvature $Q$ vanishes identically and the condition on the initial data for blowup is the classical one

$$u_0'(\underline{s}) < 0.$$

We will now give a proof of the above result in the cases (i) and (ii) above. Using the results in the Appendix, the argument will apply to case (iii) as well.

Since $M$ is simply connected, in each of the cases (i)—(iii) there will be a coframing $\alpha$, $\theta$, $\omega$ of $M$ such that

$$\Phi = \theta \wedge \omega \qquad \text{and} \qquad \Omega = \alpha \wedge \theta.$$

The characteristic vector field $X = \partial_\alpha$ is then dual to $\alpha$, and we will assume it to be complete. Although the structure equations will be a little different in each of these cases, they all have in common the equation

$$d\theta \equiv \alpha \wedge \omega \mod \theta \tag{6}$$

reflecting the non-linearity of (4).

An *integral element* of the exterior differential system $\Phi = 0$ with the independence condition $\Omega \neq 0$ is by definition given by a point $p \in M$ and 2-plane $E \subset T_pM$ such that

$$\Phi|_E = 0 \qquad \text{and} \qquad \Omega|_E \neq 0.$$

Integral elements are thus defined by a linear equation in the tangent space

$$\omega - \lambda\,\theta = 0, \qquad \lambda \in \mathbb{R} \tag{7}$$

and so the set of all such integral elements forms the manifold

$$M^{(1)} = M \times \mathbb{R}$$

underlying the *first prolongation* $\mathcal{I}^{(1)}$ of the differential ideal $\mathcal{I}$ generated by $\Phi$. On $M^{(1)}$ we consider the exterior differential system $\mathcal{I}^{(1)}$ generated by the 1-form

$$\sigma = \omega - \lambda\,\theta;$$

thus by definition $d\sigma$ lies in $\mathcal{I}^{(1)}$ as does $\Phi = \theta{\wedge}\sigma$. Every integral surface $S \subset M$ of $\Phi = 0$ satisfying $\Omega \neq 0$ lifts to a unique integral surface $S^{(1)} \subset M^{(1)}$ of $\sigma = 0$ satisfying $\Omega \neq 0$ and moreover, every integral surface of $\sigma = 0$ satisfying $\Omega \neq 0$ is the lifting $S^{(1)} \subset M^{(1)}$ of an integral surface $S \subset M$ of $\Phi = 0$ satisfying $\Omega \neq 0$. The structure equations in both case (i) and case (ii) will give that

$$d\lambda \wedge \theta \equiv (\lambda^2 - Q)\,\alpha \wedge \theta \bmod \mathcal{I}^{(1)} \tag{8}$$

where $Q$ is a curvature of the system.

Now let $\Gamma \subset M$ be a non-characteristic initial curve with parameter $s$ along which $\theta \neq 0$. We let $S$ be the flow of $\Gamma$ under $X$ for $0 \leq t < \infty$. Then $S$ has coordinates $s,\ t$ where $t \geq 0$ and is a global integral surface of $\Phi = 0$.

The open subset $U \subset S$ where $\Omega(p) \neq 0$ will contain a neighborhood of $\Gamma$. In $U$ the condition $\sigma = 0$ defines a function $\lambda = \lambda(s,\ t)$. We write $d\lambda \equiv \dot\lambda\,\alpha \bmod \theta$ where $\dot\lambda = \mathcal{L}_X\lambda$ is the Lie derivative of $\lambda$ along the characteristic vector field $X$. By (8)

$$\dot\lambda = \lambda^2 - Q. \tag{9}$$

Thus, for each $s$ the function $\lambda(s,\ t)$ satisfies a Ricatti equation in $t$ with initial value $\lambda(s,\ 0)$. Under the conditions $Q \leq 0$ and $\lambda(\underline{s},\ 0) > 0$ for some $\underline{s}$, the usual Ricatti analysis then implies that we will have finite time blowup. This is the proof of the first part of the above result.

We now formulate and prove a converse statement in the special case where the structure equations have the form

$$d\alpha = 0, \qquad d\theta = \alpha \wedge \omega, \qquad d\omega = -Q\,\alpha \wedge \theta \tag{10}$$

where $Q$ is a constant.

Setting $\Phi = \theta{\wedge}\omega$ and $\Omega = \alpha{\wedge}\theta$, it is straightforward to verify that the system $(M;\ \Phi, \Omega)$ models the PDE

$$u_t + uu_x - Qx = 0. \tag{11}$$

In fact, in $M = \mathbb{R}^3$ with coordinates $x,\ t,\ u$ we may take

$$\alpha = dt, \qquad \theta = dx - u\,dt, \qquad \omega = du - Qx\,dt, \tag{12}$$

and then (10) is satisfied. Moreover, the vector fields $\partial_\theta = \partial_x,\ \partial_\omega = \partial_u$, and $\partial_\alpha = \partial_t + u\,\partial_x + Qx\,\partial_u$ are complete, and thus the 1-forms (12) are the Maurer-Cartan forms on the unique simply connected Lie group with the structure equations (10). We will show that

> *For $Q > 0$ with initial data $u_0(x)$ satisfying $|u_0'(x)| < \sqrt{Q}$, the solution to (11) exists for all time.*



To see why this result should be true (but not yet to give a proof), we write $Q = q^2$ and examine the ODE (9)

$$\dot\lambda = \lambda^2 - q^2$$

with initial condition $\lambda(0) = \lambda_0$. For $|\lambda_0| < q$ this equation has a solution for $0 \leq t < \infty$. In fact, $\lambda(t)$ is obtained by solving for $\lambda$ in the linear fractional equation

$$\frac{q-\lambda}{q+\lambda} = \left(\frac{q-\lambda_0}{q+\lambda_0}\right) e^{2qt}.$$

In case $(M;\ \Phi,\ \Omega)$ models the PDE (11) we have

$$\lambda_0(s) = u_0'(s).$$

**Proof:** For notational simplicity we specialize to $Q = 1$ so that the characteristic vector field becomes

$$X = \partial_t + u\,\partial_x + x\,\partial_u.$$

With a given initial curve

$$s \mapsto (s,\ 0,\ u_0(s))$$

in $xtu$-space the global integral surface of $\Phi = 0$ is

$$(s,t) \mapsto \left(\tfrac{1}{2}\big(s-u_0(s)\big)e^{t}+\tfrac{1}{2}\big(s-u_0(s)\big)e^{-t}, t, \tfrac{1}{2}\big(s+u_0(s)\big)e^{t}-\tfrac{1}{2}\big(s-u_0(s)\big)e^{-t}\right).$$

The projection $\pi : S \to \mathbb{R}^2_+$ is given explicitly by

$$2x = \big(s + u_0(s)\big)e^{t} + \big(s - u_0(s)\big)e^{-t} \qquad \text{and} \qquad t = t.$$

If $\pi(s,\ t) = \pi(\bar s,\ \bar t)$ then $t = \bar t$ and, assuming $s \geq \bar s$ and writing

$$u_0(s) - u_0(\bar s) = u_0'(\xi)(s - \bar s), \qquad \bar s \leq \xi \leq s$$

we have

$$(s - \bar s)\left[\big(1 + u_0'(\xi)\big)e^{t} + \big(1 - u_0'(\xi)\big)e^{-t}\right] = 0.$$

With the assumption $|u_0'| < 1$ this implies that $s = \bar s$ and $\pi : S \to \mathbb{R}^2_+$ is one-to-one. To show that $\pi$ is onto we note that

$$2\partial_s x = \big(1 + u_0'(s)\big)e^{t} + \big(1 - u_0'(s)\big)e^{-t} > 0$$

so that for each fixed $t$

$$x(-\infty,\ t) = -\infty \qquad \text{and} \qquad x(+\infty,\ t) = +\infty$$

and so the mapping $\pi : S \to \mathbb{R}^2_+$ is onto. □

We now return to our discussion of equation (5) where we make the genuine non-linearity assumption $f'(u) > 0$. We assume initial data given in the form

$$x \mapsto (x,\ 0,\ u_0(x)) \tag{13}$$

where $u_0(x)$ is constant outside of a compact set, say for $|x| > R$, but where $u_0'(x_0) \neq 0$ for some $x_0$. Then the situation is this:

(i) *the geometry of the equation is complete, so that there is a unique global, smooth integral surface $S$ of the associated exterior differential system*

$$\Phi = du \wedge \left(dx - f(u)\,dt\right) = 0$$

*with the initial curve $\Gamma$ corresponding to* (13);

(ii) *the mapping $\pi : S \to \mathbb{R}^2_+$ is surjective but develops a singularity.*

We have proved everything except for the surjectivity of $\pi$. In fact, $\pi$ is *proper*. To see this we think of $S$ as the half-plane $\{(s, t) : -\infty < s < \infty, t \geq 0\}$. The line $t = 0$ maps bijectively to $\Gamma$. For each fixed $s_0$ the half-line $s = s_0$, $t \geq 0$ maps bijectively to the half-line $\left(s_0 + tu(x_0), t\right), t \geq 0$. For $x_0 > R$ we therefore have a family of parallel lines, and similarly for $x_0 < -R$. The picture is something like Figure 2.6 and from this it follows that $\pi$ is proper.

Now in view of (ii) above a natural question is whether or not the scalar conservation (5) admits some sort of global "generalized solution" $u(x, t)$. Since (5) is a non-linear differential equation care must be taken in defining a generalized solution. The key observation is that, writing $f(u)u_x = \partial_x\left(F(u)\right)$ for a smooth function $F(u)$ with $F'(u) = f(u)$, a locally bounded, measurable function $u(x, t)$ may be defined to be a weak solution to (5) if the equation holds in the distribution sense, i.e., if

$$\int\!\!\int \left(\alpha_t u + \alpha_x F(u)\right) dx dt = 0 \tag{14}$$

holds for all smooth functions $\alpha$ with compact support in $\mathbb{R}^2_+ = \{(x, t) : t > 0\}$. We will be primarily interested in the situation where $u$ is locally bounded and of class $C^1$ outside a set of *break curves* across which $u$ has a jump discontinuity. These break curves will consist of piecewise $C^1$ arcs, and if $\gamma = \left(x(\sigma), t(\sigma)\right)$ is one such arc then (14) is easily seen to imply the *Rankine-Hugonoit* (or jump) condition

$$[F(u)] = s[u] \tag{15}$$

where $s = x'(\sigma)/t'(\sigma)$ is the propagation speed of the discontinuity and [ ] represents the magnitude of the jump across $\gamma$ (thus $[u] = u_- - u_+$ where $u_-$

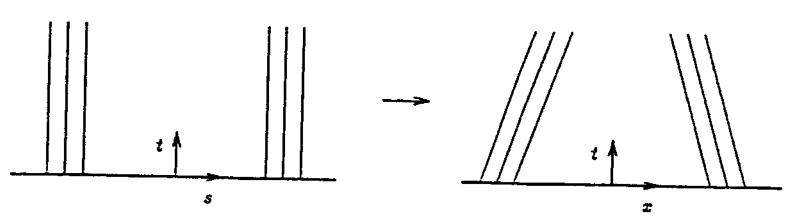


**Figure 2.6**



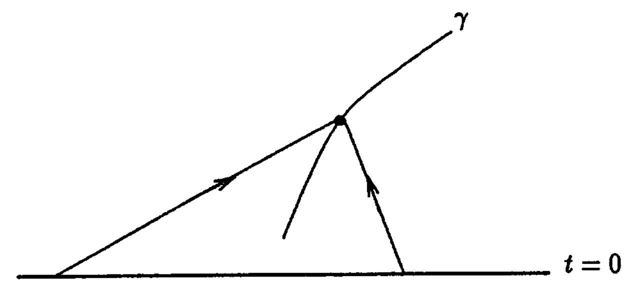


**Figure 2.7**

and $u_+$ are respectively the left and right hand limits of $u(x, t)$ at a point of $\gamma$). In general the jump condition (15) are not sufficient to uniquely specify a weak solution to (5), and so one traditionally adds the *entropy* condition

$$f'(u_+) < s < f'(u_-). \tag{16}$$

Geometrically this mean that at a point of the break curve $\gamma$ the two characteristics should impinge (Figure 2.7). With our assumption that $f' > 0$ the entropy condition means that $[u] > 0$, i.e., $u$ should "jump down" across $\gamma$.

If one now defines a *shock solution* to (5) to be a weak solution satisfying the entropy condition, then under rather general assumptions on the initial data (for instance that it be smooth with compact support) there is a unique, global shock solution to the initial value problem for (5) (for details, see *[20]*).

The question now arises: *Does the shock solution arise by taking a suitable cross-section of the proper mapping $\pi : S \to \mathbb{R}^2_+$?* In particular for the shock solution $u$, does $\left(x, t, u(x, t)\right)$ lie on the geometric solution surface $S$ for all $(x, t) \in \mathbb{R}^+_2$? We shall argue that this is indeed the case, at least when the singularities of $\pi$ are generic.

Referring to Figure 2.1 and Figure 2.2 above, there will be a break curve $\gamma$ corresponding to each component of the fold locus $\mathcal{F}$. Inside $\mathcal{F}$ the jump will be known since we must jump from the top sheet to the bottom sheet (jumping to the intermediate sheet will run us into a singularity on $\mathcal{F}$). Thus, inside the cusp defined by $\mathcal{F}$ we may view (15) as a differential equation for the break curve $\gamma$. More precisely, in this region, equation (15) defines a vector field $X$. The genuine non-linearity assumption $F'' > 0$ implies that, aside from the cusp points themselves, at each point along $\mathcal{F}$, the vector field $X$ will be transverse to $\mathcal{F}$. Thus there is a unique integral curve of $X$ which emanates from the cusp point of $\mathcal{F}$, represented in Figure 2.8 by the dashed line. This is the desired break curve.

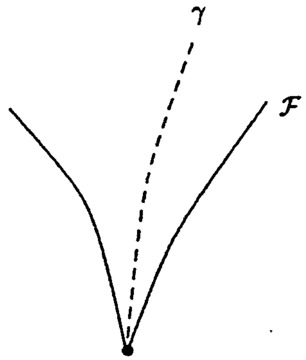


**Figure 2.8**

In summary, under the assumption that the singularities of the projection $\pi$ are generic, we have given heuristic reasoning to explain the well-known existence and uniqueness of shock solutions to (5). The point is that the exterior differential system perspective naturally leads to the "picture" of the shock solution.

In the following sections, we will see that the situation for hyperbolic systems of conservation laws (as opposed to a single conservation law) is both more complex and also more interesting.

## 3 Geometry of Hyperbolic Systems

**3.1** *Hyperbolic Exterior Differential Systems.* In the preceeding sections we have studied the geometry associated to a second-order ODE, a first-order non-linear scalar PDE and a non-linear scalar conservation law. In each case the geometry is given by a suitable $G$-structure $B_G \longrightarrow M$ with an intrinsic connection $\phi$ and the geodesics of $\phi$ gives rise to solution curves in the case of the ODE, and characteristic curves in the other two cases. Moreover, each geodesic $\gamma$ has a natural parameter $\tau$ defined up to an affine transformation $\tau \longrightarrow a\tau + b$ once we fix a point of $\gamma$, and we may then speak of what it means for $\gamma$ to be *complete.* With this concept in hand we could define what it means for an integral surface $S \subset M$ passing through a given initial curve to be *global.*

The curvatures of $\phi$ had several interpretations. One is that the vanishing of certain components implied that there were a complete set of first integrals for the solution curves of the ODE. Another is that the sign of suitable curvature components could be interpreted as a spreading or focusing of geodesics, thereby relating to the existence or non-existence of global classical solutions to the non-linear scalar PDE. Finally, the exterior differential system point of view led to a picture of the classical existence and uniqueness of shock solutions to the non-linear conservation law in terms of the geometry of the proper map $\pi : S \longrightarrow \mathbb{R}^2_+$.

When we seek to study the geometry of less elementary differential equations the situation is more interesting and correspondingly more complex. We shall now



discuss certain aspects of the geometry associated to a hyperbolic PDE system. For further details on this geometry, consult the Appendix and *[4].*

EXAMPLE 4: Beyond the scalar equation already discussed, the simplest non-linear PDE is a first order quasi-linear system for two unknowns over a domain $U \subset \mathbb{R}^2$. We assume that the system is hyperbolic. Then, after a suitable change of dependent and independent variables, the system will assume the form

$$\begin{aligned} u_y + a\,u_x + b\,v_x + f &= 0 \\ v_y + c\,u_x + e\,v_x + g &= 0 \end{aligned} \tag{1}$$

where the coefficients $a$, $b$, ... etc. are functions of $x$, $y$, $u$, $v$. Hyperbolicity of (1) amounts to the condition that the matrix

$$A = \begin{pmatrix} a & b \\ c & e \end{pmatrix}$$

have everywhere distinct real eigenvalues. Examples of such equations include hyperbolic systems of conservation laws for two unknowns in $(1+1)$-dimensional spacetime.

To write (1) as an exterior differential system, we set

$$\begin{aligned} \Phi &= -du{\wedge}dx + (a\,du + b\,dv){\wedge}dy + f\,dx{\wedge}dy \\ \Psi &= -dv{\wedge}dx + (c\,du + e\,dv){\wedge}dy + g\,dx{\wedge}dy. \end{aligned}$$

Integral surfaces of the exterior differential system $\Phi = \Psi = 0$ on which $dx{\wedge}dy \neq 0$ are locally in one-to-one correspondence with solutions to (1). A little computation shows that

$$(\Phi + \xi\Psi)^2 = 2\big(b + (e-a)\xi - c\xi^2\big)\,du \wedge dv \wedge dx \wedge dy.$$

The discriminant of the quadratic polynomial in $\xi$ is

$$(e-a)^2 + 4bc = (\operatorname{tr} A)^2 - 4\det A.$$

Thus, by hyperbolicity, the equation

$$(\Phi + \xi\Psi)^2 = 0$$

has two distinct real roots $\xi_1 < \xi_2$, giving rise to two linearly independent 2-forms $\Omega_1 = \Phi + \xi_1\Psi$ and $\Omega_2 = \Phi + \xi_2\Psi$ satisfying

$$\begin{cases} \Omega_1{\wedge}\Omega_1 = 0 = \Omega_2{\wedge}\Omega_2 \\ \Omega_1{\wedge}\Omega_2 \neq 0. \end{cases} \tag{2}$$

**Definition:** A *hyperbolic exterior differential system* (of class $s = 0$) on a 4-manifold $M$ is a differential system which can be generated locally by a pair $\Omega_1$ and $\Omega_2$ of 2-forms satisfying (2).

As we have seen, the quasi-linear system (1) above generates a hyperbolic exterior differential system. A number of well-known examples of system (1) will be given below.

Another interesting source of examples are the hyperbolic Monge-Ampere equations of the form

$$E\,(z_{xx}z_{yy} - z_{xy}^2) + A\,z_{xx} + 2B\,z_{xy} + C\,z_{yy} + D = 0 \tag{3}$$

where the coefficient functions $A$, $B$, $C$, $D$, and $E$ are functions of $x$, $y$, $p = z_x$ and $q = z_y$ alone, i.e., they have no explicit dependence on $z$. In this case, the exterior differential system in $xypq$-space generated by the pair of 2-forms

$$\begin{aligned} \Phi &= dp{\wedge}dx + dq{\wedge}dy \\ \Psi &= E\,dp{\wedge}dq + A\,dp{\wedge}dy + B\,(dq{\wedge}dy + dx{\wedge}dp) + C\,dx{\wedge}dq + D\,dx{\wedge}dy \end{aligned}$$

is a hyperbolic exterior differential system. Again, solution surfaces of $\Phi = \Psi = 0$ on which $dx{\wedge}dy \neq 0$ are locally in one-to-one correspondence with solutions to (3). The hyperbolicity of (3) is equivalent to the equation $(\Phi + \xi\Psi)^2 = 0$ having distinct real roots. Among the many well-known hyperbolic Monge-Ampere equation of this type, we cite only the classical hyperbolic Monge-Ampere equation

$$z_{xx}z_{yy} - z_{xy}^2 = -1$$

and the analog, for timelike surfaces in Minkowski 3-space, of the classical minimal surface equation

$$z_y^2\,z_{xx} - 2(1 + z_x z_y)z_{xy} + z_x^2\,z_{yy} = 0. \tag{4}$$

(This is the Euler-Lagrange equation for the area functional for surfaces of the form $(x, y, z(x, y))$ in 3-space with metric $dz^2 + 2\,dx \circ dy$.)

Before going further in the study of hyperbolic exterior differential systems, we want to impose some non-degeneracy conditions. Recall that a non-zero exterior 2-form $\Omega$ is *decomposable* in the sense that $\Omega$ can be locally written in the the form $\Omega = \alpha{\wedge}\beta$ for a pair of (linearly independent) 1-forms $\alpha$ and $\beta$ if, and only if, $\Omega{\wedge}\Omega = 0$. We say that such an $\Omega$ is *integrable* if it can be locally written in the form $\Omega = f\,dx{\wedge}dy$ for some local functions $x$, $y$, and $f$. By an elementary application of the Frobenius's theorem, the integrability of $\Omega$ is equivalent to the condition that there exist a 1-form $\omega$ so that $d\Omega = \omega{\wedge}\Omega$. In this case, the integral manifolds of $\Omega$ can locally be described by elementary methods. In particular, in a region where we can write $\Omega = f\,dx{\wedge}dy$, the integral manifolds are just given as hypersurfaces cut out by a non-trivial equation of the form $F(x, y) = 0$.

For a hyperbolic exterior differential system $(M; \Omega_1, \Omega_2)$, the case where either of $\Omega_i$ is integrable may thus be analyzed by essentially elementary methods and will not be further discussed here. Thus, we shall assume without further mention that the system is *non-degenerate* in the sense that neither $\Omega_1$ nor $\Omega_2$ is integrable. In this case one may seek to determine the geometry associated to



$(M; \Omega_1, \Omega_2)$ with the non-degeneracy assumption built in from the outset, in a manner similar to the use of the non-linearity assumption in EXAMPLE 2 and EXAMPLE 3 above. We shall discuss some of the specific implications of the equivalence problem in this case.

Before doing this, a remark on the group $\Gamma$ of allowable coordinate transformations is in order. In the case of EXAMPLE 1 we took $\Gamma$ to be the group of point transformations while in EXAMPLE 2 we used for $\Gamma$ the gauge group, since under the larger group of contact transformations any equations of either EXAMPLE 1 or EXAMPLE 2 are locally equivalent to the corresponding trivial model. In EXAMPLE 3 we used a subgroup of gauge transformations which preserve a closed 2-form, since it is only under this smaller group that the "jump conditions" for weak solutions are invariant. For hyperbolic systems, however, we shall use for $\Gamma$ the full group of contact transformations. Here the geometry is already quite rich.

For instance, suppose we define a hyperbolic system $(M; \Omega_1, \Omega_2)$ to be *linear* if it is locally contact equivalent to the exterior differential system arising from a linear hyperbolic PDE system (1). Linear hyperbolic systems include many arising from non-linear PDEs, such as the equation $z_{xx}z_{yy} - z_{xy}^2 = -1$. It turns out (see the Appendix and *[4]*) that linear hyperbolic exterior differential systems may be characterized by the vanishing of certain "curvature" components in the geometry associated to $(M; \Omega_1, \Omega_2)$ and that in turn this geometry induces on solution surfaces an intrinsic pseudo-Riemannian metric whose own geometry (constancy of Gauss curvature, etc.) has meaning for solutions to the original PDE.

In the Appendix we study the equivalence problem associated with the non-degenerate hyperbolic system $(M; \Omega_1, \Omega_2)$ and derive its geometry. A natural question to ask is: What can that geometry tell us about the original PDE? By analogy with EXAMPLE 1 and EXAMPLE 2 in §1 above we could ask

(i) for conditions on the curvatures that would allow us to explicitly "solve" the PDE;

(ii) for a notion of completeness that would guarantee that the exterior differential system has "global" integral surfaces which project onto the half-plane $\mathbb{R}^2_+ = \{(x, y) : y \geq 0\}$ to provide global multi-valued solutions to the PDE system; and

(iii) for curvature-type conditions that will imply that there are no global smooth classical solutions to (1).

In this section we will take up (i); in later sections there will be a discussion of (ii) and (iii).

**3.2** *Remarks on Integration by the Method of Darboux.* For a hyperbolic system given by (1) or (3) what does it mean to explicitly "solve" the PDE? One notion is that a general solution is given by an explicit formula

$$
\begin{aligned}
x &= X\big(s,t,\ \alpha(s),\alpha'(s),\ldots,\alpha^{(k)}(s);\ \beta(t),\beta'(t),\ldots,\beta^{(k)}(t)\big)\\
y &= Y\big(s,t,\ \alpha(s),\alpha'(s),\ldots,\alpha^{(k)}(s);\ \beta(t),\beta'(t),\ldots,\beta^{(k)}(t)\big)\\
u &= U\big(s,t,\ \alpha(s),\alpha'(s),\ldots,\alpha^{(k)}(s);\ \beta(t),\beta'(t),\ldots,\beta^{(k)}(t)\big)\\
v &= V\big(s,t,\ \alpha(s),\alpha'(s),\ldots,\alpha^{(k)}(s);\ \beta(t),\beta'(t),\ldots,\beta^{(k)}(t)\big)
\end{aligned}
\tag{5}
$$

where $\alpha$ and $\beta$ are arbitrary functions of one variable. Note that the "function count" is correct, since, in particular, initial data for a hyperbolic system of type (1) or (3) is given by two arbitrary functions of one variable. A more refined notion of "explicitly solving the PDE" is that there should be explicit expressions of the general form (5) in terms of the initial data.

The existence of formulae of the above form for the general integral surface of a hyperbolic exterior exterior differential system is part of what constitutes the phenomenon of *integrability by the method of Darboux.* For a more complete discussion of this method, see *[13, 17]* and *[4]*, but intuitively this corresponds to the equations being solvable in the above form.

If a hyperbolic system is Darboux-integrable, then the values of a solution at $(x, y)$ depend only on the initial data at the two points $p$ and $q$ where the characteristics enamating from $(x, y)$ meet the curve along which the initial data is posed. This is not true for the general equation (1) where, even in the linear case, the solution usually depends on the initial data along the whole segment between $p$ and $q$ on the initial curve. Whether this might be an analytic characterization of Darboux integrability is an interesting question.

We now provide some examples of Darboux-integrable equations:

EXAMPLE: Consider a surface endowed with a Lorentzian metric (i.e., a metric of signature $(1, 1)$) of constant curvature $-1$. A natural coordinate representation of the equation

$$\Box z = -k(1-k)z, \qquad k \in \mathbb{R}^*$$

(here $\Box$ denotes the Lorentzian Laplacian of the metric) is given by

$$z_{xy} = k(1-k)(x-y)^{-2}z$$

where the domain is taken to be the half-plane $x - y > 0$. This equation admits a general solution of the form (5) exactly when $k$ is a positive integer, and in that case the formula is

$$z = (x-y)^k \frac{\partial^{2k-2}}{\partial x^{k-1}\partial y^{k-1}}\left(\frac{f(x)-g(y)}{x-y}\right)$$

where $f$ and $g$ are arbitrary functions of one variable.

EXAMPLE: Equation (4) for a "maximal" time-like surface in Minkowski 3-space is Darboux-integrable, with a formula of the form (5) with $k = 2$. This formula for the general solution is the analog of the well-known Weierstrass representation for the minimal surface equation in Euclidean 3-space.



In fact every time-like surface $\big(x,\ y,\ z(x,\ y)\big)$ in $\mathbb{R}^3$ where $z$ satisfies (4) is seen to be a surface of translation, where the two translation curves are null curves of the quadratic form $dz^2 + 2dx \circ dy$. This allows us to write the general solution in the form

$$
\begin{aligned}
x(s,\ t) &= \alpha''(s) + \tfrac{1}{2}t^2\beta''(t) - t\beta'(t) + \beta(t)\\
y(s,\ t) &= \beta''(t) + \tfrac{1}{2}s^2\alpha''(s) - s\alpha'(s) + \alpha(s)
\end{aligned}
$$

and

$$z(x,\ t) = s\alpha''(s) - \alpha'(s) + t\beta''(t) - \beta'(t)$$

where $\alpha$ and $\beta$ are arbitrary functions of one variable. (Note that, in this case, we cannot eliminate the "characteristic" parameters $s$ and $t$ to get a formula for $z$ directly as an expression in $x$ and $y$ and two arbitrary functions of them.)

For hyperbolic systems our assumption of non-degeneracy rules out the possibility of a "level zero" formula

$$
\begin{aligned}
u &= U\big(\alpha(x),\ \beta(y)\big)\\
v &= V\big(\alpha(x),\ \beta(y)\big)
\end{aligned}
$$

(i.e., one for which no derivatives of the arbitrary functions are needed) such as one obtains for the $s = 0$ classical wave system

$$
\begin{aligned}
u_x &= 0\\
v_y &= 0
\end{aligned}
$$

this being the model system when both $\Omega_1$ and $\Omega_2$ are integrable. We therefore turn to the question of the existence of a "level one" formula, this being a representation (5) when $k = 1$. It may be shown that the necessary and sufficient conditions for such a formula are

$$p_3 = p_4 = q_1 = q_2 = 0; \qquad (k_{13}-1) = (k_{31}-1) = k_{14} = k_{32} = 0 \tag{6}$$

where the $p_i$ and $q_i$ are components of the torsion and the $k_{ij}$ are suitable components of the curvature derived from the geometry associated to the hyperbolic system (cf. the Appendix). This gives in practice an algorithm to test if a hyperbolic system has an explicit level one general solution of the form (5). Perhaps more interestingly, it leads to the result that there are exactly two contact equivalence classes of hyperbolic systems for which this is true.

EXAMPLE: If (6) holds together with

$$p_1 = q_3 = 0,$$

then (cf. again the Appendix) the hyperbolic system is linear. Moreover, the induced metric on solution surfaces has constant Gauss curvature $K = 1$ and invariant $F = 0$ (for this terminology, see the Appendix). From this it follows that it is locally equivalent as an exterior differential system to the system in $xyuv$-space generated by

$$\Omega_1 = \left(du - \frac{v\,dy}{\cos(x+y)}\right) \wedge dx$$
$$\Omega_2 = \left(dv - \frac{u\,dx}{\cos(x+y)}\right) \wedge dy$$

where $|x+y| < \pi/2$. This exterior differential system models the hyperbolic PDE system

$$u_y = \frac{v}{\cos(x+y)}$$
$$v_x = \frac{u}{\cos(x+y)}.$$

The general solution to this system is provided by the formulas

$$u(x,\,y) = g(y)\sec(x+y) \;+ f(x)\tan(x+y) \;+ f'(x)$$
$$v(x,\,y) = f(x)\sec(x+y) \;+ g(y)\tan(x+y) \;+ g'(y)$$

where $f$ and $g$ are arbitrary functions of one variable.

If $p_1q_3 \neq 0$ then it can be shown that the hyperbolic exterior differential system is equivalent to one modeling the $s = 0$ Liouville system

$$u_y = e^v$$
$$v_x = e^u.$$

The representation (5) of a general solution is

$$e^u = \frac{-a'(x)}{a(x)+b(y)}$$
$$e^v = \frac{-b'(y)}{a(x)+b(y)}.$$

# 4 Translation Invariant Hyperbolic Systems

**4.1** *Characteristic Completeness and Global Integral Surfaces.* We will now discuss the geometry of translation invariant, quasi-linear hyperbolic systems

$$\begin{aligned} u_y \;+ a(u,\,v)u_x \;+ b(u,\,v)v_x &= 0 \\ v_y \;+ c(u,\,v)u_x \;+ e(u,\,v)v_x &= 0. \end{aligned} \tag{1}$$

For such systems the natural group of equivalences to consider are classical transformations which change the independent and dependent variables separately. It will be seen that the exterior differential system $(M;\,\Omega_1,\,\Omega_2)$ associated to (1) is *linear* in the appropriate sense, and in this section we will use that observation to establish a global existence and uniqueness result for integral surfaces of the exterior differential system. In subsequent sections we will discuss the geometry of these surfaces as they pertain to classical solutions of (1).



Before doing this we want to comment on the initial value problem for integral surfaces of a hyperbolic system. Initial data is given by a mapping $\phi : [0,1] \to M$ whose image is an immersed, non-characteristic curve, where, by *non-characteristic*, we mean that $\phi'(t) \lrcorner \Omega_i \neq 0$ for $i = 1$ or 2. We consider the unit square $\Sigma = \{(s,\,t) \in \mathbb{R}^2 : 0 \leq s,\, t \leq 1\}$ and set $\Delta$ to be the diagonal in $\Sigma$. The *initial value problem* seeks to extend the initial data $\phi$ to a mapping

$$f_\phi : U_\phi \to M \tag{2}$$

of a neighborhood $U_\phi \subset \Sigma$ of the diagonal, which is an integral surface of the hyperbolic system $(M;\,\Omega_1,\,\Omega_2)$, with the additional property that the characteristic foliations pull-back to be the $s =$ const and $t =$ const lines. This last requirement removes the reparametrization ambiguity usually associated with integral surfaces of differential systems since any diffeomorphism of the unit square fixing the diagonal pointwise and the foliations $s =$ const and $t =$ const must be the identity.

Now, the standard existence and uniqueness theorems for hyperbolic PDEs imply that there is a neighborhood $U$ of the diagonal such that the desired extension (2) exists and is unique. We shall say that the solution is *characteristically complete* in case we may take $U$ to be the whole unit square $\Sigma$. That is, the solution is characteristically complete if we are able to extend the mapping $\phi$ to include all points lying on the characteristics emanating from the initial curve to the point where they meet. Pictorially, we have the following situation (see Figure 4.1). As stated above, the initial data, prescribed on the diagonal $\Delta$, may always be extended to give a solution surface in a neighborhood $U$ of $\Delta$. However, if an extension could not be further extended beyond the $U$ depicted, the two characteristic curves drawn could not be extended until they met in $U$, so such a solution would be incomplete. In fact, unless we can extend the mapping $f_\phi$ to the whole square $\Sigma$, this condition of characteristic completeness fails to be satisfied. The question of existence of complete solutions with arbitrary non-characteristic initial data is obviously a global one. We shall give a large class of hyperbolic systems for which such complete integral surfaces exist with arbitrary initial data. By illustration we shall also show that in some examples one may expect characteristic completeness for some but not all initial data.

We now return to the system (1), which we write in vector form as

$$\mathbf{u}_y + \mathbf{a}(\mathbf{u})\mathbf{u}_x = 0 \tag{3}$$

where $\mathbf{u} = {}^t(u,\,v)$ and $\mathbf{a}(\mathbf{u})$ is a 2-by-2 matrix which has distinct real eigenvalues $\lambda = \lambda(\mathbf{u})$ and $\mu = \mu(\mathbf{u})$ at all points under consideration. The exterior differential system associated to (3) is given by

$$\Phi = d\mathbf{u} \wedge dx - \mathbf{a}(\mathbf{u})\,d\mathbf{u} \wedge dy = 0 \tag{4}$$

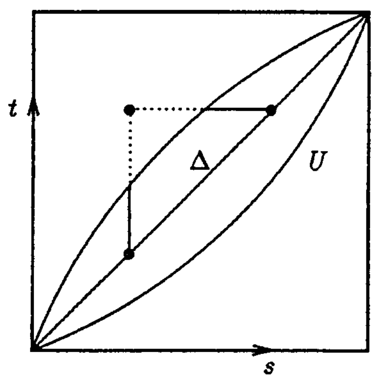


**Figure 4.1**

where $\Phi$ is an $\mathbb{R}^2$-valued 2-form. Let $\mathbf{c}(\mathbf{u})$ be a 2-by-2 matrix such that

$$\mathbf{c}(\mathbf{u})\mathbf{a}(\mathbf{u})\mathbf{c}(\mathbf{u})^{-1} = \Lambda(\mathbf{u}) = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

Setting $\Pi = {}^t(\pi_1, \pi_2) = \mathbf{c}(\mathbf{u})\, d\mathbf{u}$, the hyperbolic system (4) becomes

$$\Pi \wedge dx - \Lambda(\mathbf{u})\, \Pi \wedge dy = 0$$

or equivalently, in components

$$\begin{aligned} \pi_1 \wedge (dx - \lambda\, dy) &= 0 \\ \pi_2 \wedge (dx - \mu\, dy) &= 0. \end{aligned} \tag{5}$$

In the following, we shall define a notion of completeness for the PDE system (3). We shall show that for such systems the corresponding exterior differential system has a unique, complete smooth integral surface for arbitrary non-characteristic initial data.

We begin by noticing that for the PDE system (3), the 1-forms $\pi_1$ and $\pi_2$ defined above are integrable in the $uv$-plane and hence there are locally defined functions $p = p(u, v)$ and $q = q(u, v)$ such that $\pi_1$ is a multiple of $dp$ and $\pi_2$ is a multiple of $dq$. These functions $p(u, v)$ and $q(u, v)$ are the well-known *Riemann invariants* associated with hyperbolic PDE systems.

A hyperbolic system (3) is said to be *complete* on a domain $D$ in the $uv$-plane if we can choose the Riemann invariants $p$ and $q$ to be global coordinates on $D$ in such a way that the image under the mapping $(p, q) : D \to \mathbb{R}^2$ is a coordinate box in the $pq$-plane. If (3) is complete on $D = \mathbb{R}^2$, then we simply say that (3) is complete. We will say that non-characteristic initial data

$$s \mapsto \big(x(s), y(s), u(s), v(s)\big)$$



is complete if the image of the data in the $uv$-plane lies in a domain $D$ in which the equation (3) is complete.

We can now establish the following result:

*For any hyperbolic system (3), any complete non-characteristic initial curve extends to a unique, smooth characteristically complete integral surface of the associated exterior differential system. Furthermore, the integral surface is an immersion in a neighborhood of the initial curve.*

Before proceeding to the proof of this result, we consider an example:

EXAMPLE: Recall that the Fermi-Pasta-Ulam (FPU) equation is given by

$$z_{yy} - k^2(z_x) z_{xx} = 0, \tag{6}$$

where $k$ is a smooth positive function. This is a special case of the Monge-Ampere equation (3) introduced in the preceeding section. The exterior differential system corresponding to (6) is generated by

$$\begin{aligned} \Phi_1 &= du \wedge dx + dv \wedge dy \\ \Phi_2 &= dv \wedge dx + \big(k(u)\big)^2\, du \wedge dy. \end{aligned}$$

Integral surfaces of the system $\Phi_1 = \Phi_2 = 0$ on which $dx \wedge dy \neq 0$ are (up to an additive constant) locally in one-to-one correspondence with solutions to the FPU equation. Setting

$$\begin{aligned} \pi_1 &= dv + k(u)\, du, & \omega^1 &= dx + k(u)\, dy \\ \pi_2 &= dv - k(u)\, du, & \omega^2 &= dx - k(u)\, dy \end{aligned}$$

and

$$\Omega_1 = \pi_1 \wedge \omega^1, \qquad \Omega_2 = \pi_2 \wedge \omega^2$$

we note that span $\{\Phi_1, \Phi_2\}$ = span $\{\Omega_1, \Omega_2\}$. Thus we obtain an exterior differential system of the form (5) where $\pi_1$, $\pi_2$ are exact. In fact, letting $K(u)$ be an anti-derivative of the function $k(u)$ we have

$$\begin{aligned} dp &= dv + k(u)\, du = d\big(v + K(u)\big) \\ dq &= dv - k(u)\, du = d\big(v - K(u)\big). \end{aligned}$$

From this it follows that the condition for the FPU system to be complete is that the one-to-one mapping $K : \mathbb{R} \to \mathbb{R}$ should be surjective. More generally, if the image of $K$ is some proper subinterval $(a, b)$ of the line (where either $a = -\infty$ or $b = \infty$ is allowed, but not both), then initial data of the form

$$\phi(s) = \big(x_0(s),\, y_0(s),\, u_0(s),\, v_0(s)\big)$$

is complete if and only if, for all $s$ and $t$ in the domain of $\phi$, we have

$$a < \tfrac{1}{2}\big(K(u_0(s)) + K(u_0(t)) + v_0(s) - v_0(t)\big) < b$$

(note that this inequality clearly holds when $s$ and $t$ are sufficiently close, since the value in the middle will then be close to $K(u_0(s))$ ).

We now turn to the proof of the above claim. Let the initial data be given by a complete non-characteristic curve written in suitable $xypq$-coordinates in the form

$$\phi(s) = (x_0(s),\, y_0(s),\, p_0(s),\, q_0(s))$$

where $0 \le s \le 1$. We seek a mapping $f : \Sigma \to M = \mathbb{R}^2 \times P \times Q$, where $P$ and $Q$ are the image intervals in $\mathbb{R}$ of the functions $p_0$ and $q_0$ respectively, of the form

$$(s,\, t) \mapsto (x(s,\, t),\, y(s,\, t),\, p(s,\, t),\, q(s,\, t)) \tag{7}$$

and satisfying the initial conditions

$$x(s,\, s) = x_0(s), \quad y(s,\, s) = y_0(s), \quad p(s,\, s) = p_0(s), \quad q(s,\, s) = q_0(s) \tag{8}$$

as well as the exterior equations

$$\begin{aligned} dp(s,\, t)\wedge ds &= 0, & (dx(s,\, t) - \lambda(s,\, t)\, dy(s,\, t))\wedge ds &= 0 \\ dq(s,\, t)\wedge dt &= 0, & (dx(s,\, t) - \mu(s,\, t)\, dy(s,\, t))\wedge dt &= 0 \end{aligned} \tag{9}$$

where $\lambda(s,\, t) = \lambda\big(p(s,\, t),\, q(s,\, t)\big)$ and $\mu(s,\, t) = \mu\big(p(s,\, t),\, q(s,\, t)\big)$. In fact, these equations imply not only that $\Omega_1$ and $\Omega_2$ pull back to zero under the mapping (7), but also that the two linear factors of $\Omega_1$ will each pull back to define the characteristic foliation $s = \text{const}$, and similarly the linear factors of $\Omega_2$ will define the curves $t = \text{const}$.

The first set of equations in (9) above together with the initial conditions (8) imply that

$$p(s,\, t) = p_0(s), \qquad q(s,\, t) = q_0(t).$$

(Note that this does well-define $p$ and $q$ as functions of $s$ and $t$ since the allowable values of $(p, q)$ form a coordinate box in the $pq$-plane.) In particular, $\lambda(s,\, t)$ and $\mu(s,\, t)$ are determined from the initial data. The second set of equations in (9) are equivalent to

$$\begin{aligned} x_t - \lambda(s,t) y_t &= 0, \\ x_s - \mu(s,t) y_s &= 0. \end{aligned} \tag{10}$$

This is a *linear* hyperbolic PDE system posed in the unit square $\Sigma = [0,\, 1] \times [0,\, 1]$, with characteristics given by the line $s = \text{const}$ and $t = \text{const}$ and with initial data given on the diagonal. By standard existence and uniqueness theorems there is a unique solution $x(s,\, t)$, $y(s,\, t)$ to the system (10).

The condition that the initial data be non-characteristic may be easily seen to be equivalent to (7) being an immersion along the diagonal, and hence in a neighborhood of the diagonal. □

In the next section we will show that, under certain restrictions on the initial data, the mapping (7) is an immersion outside a compact set in the $st$-plane and that it is everywhere an immersion for "generic" initial data. It may be that (7) is always an immersion everywhere for all initial data but we have not been able to establish this in general.



We now consider equations which are not complete—many interesting equations, such as those arising from gas dynamics, belong to this class. For such equations, the above proof may still allow us to establish a global existence result for some (but not all) initial data. We illustrate this with the following

EXAMPLE: Consider the coupled Burgers' equations

$$\begin{aligned} u_y + v v_x &= 0 \\ v_y + u u_x &= 0. \end{aligned} \tag{11}$$

Note that this system is hyperbolic whenever $uv > 0$, and so in what follows we shall restrict to the region $\mathbb{R}^4_+ = \{(x,\, y,\, u,\, v) \in \mathbb{R}^4 : u > 0, v > 0\}$. The corresponding exterior differential system on $\mathbb{R}^4_+$ has the form

$$\begin{aligned} \Phi &= -du\wedge dx + v\, dv\wedge dy = 0 \\ \Psi &= -dv\wedge dx + u\, du\wedge dy = 0. \end{aligned}$$

From the relation

$$(\Phi + \xi\Psi)^2 = 2(v - \xi^2 u)(du \wedge dv \wedge dx \wedge dy)$$

we see that $\xi = \pm\sqrt{v/u}$ gives the decomposable linear combinations of $\Phi$ and $\Psi$. A little computation shows that (up to a multiple) these decomposables are

$$\begin{aligned} \Omega_1 &= \tfrac{3}{2}(\sqrt{u}\, du + \sqrt{v}\, dv)\wedge(dx - \sqrt{uv}\, dy), \\ \Omega_2 &= \tfrac{3}{2}(\sqrt{u}\, du - \sqrt{v}\, dv)\wedge(dx + \sqrt{uv}\, dy). \end{aligned}$$

In what follows, we shall find it convenient to introduce the following change of variables

$$p = \tfrac{1}{2}(u^{3/2} + v^{3/2}), \qquad q = \tfrac{1}{2}(u^{3/2} - v^{3/2}).$$

Now, the exterior differential system on $M = \{(x,\, y,\, p,\, q) \in \mathbb{R}^4 : p > |q|\}$ is given by

$$\begin{aligned} \Omega_1 &= dp\wedge\big(dx - (p^2 - q^2)^{1/3}\, dy\big) = 0, \\ \Omega_2 &= dq\wedge\big(dx + (p^2 - q^2)^{1/3}\, dy\big) = 0. \end{aligned}$$

Notice that this system is *not* complete on its domain of hyperbolicity.

We now seek solutions $f : \Sigma \to M$ of the above hyperbolic system. As in the proof above, we have

$$p(s,t) = p_0(s), \qquad q(s,t) = q_0(t). \tag{12}$$

In order to have $p(s,t) > |q(s,t)|$ for all $(s,t) \in \Sigma$, the initial data must clearly satisfy

$$\max_s |q_0(s)| < \min_s p_0(s)$$

and these are necessary and sufficient restrictions on the initial data to have a characteristically complete integral surface.

**4.2** *The Geometry of Global Integral Surfaces.* We continue our discussion of integral surfaces of the exterior differential system on $M \subset \mathbb{R}^4$ associated to the PDE system

$$\mathbf{u}_y + \mathbf{a}(\mathbf{u})\mathbf{u}_x = 0. \tag{1}$$

As was established in the previous section we may locally write this exterior differential system as

$$\begin{aligned}\Omega_1 &= dp\wedge(dx - \lambda\, dy) = 0\\ \Omega_2 &= dq\wedge(dx - \mu\, dy) = 0\end{aligned} \tag{2}$$

where $p$ and $q$ are functions of $u$ and $v$ for which $dp\wedge dq \neq 0$. Henceforth, we shall assume that the system (1) above is complete and so we may choose $p$, $q$ to be global coordinates on the $uv$-plane. In particular, $\lambda$ and $\mu$ are then well-defined functions of $p$ and $q$.

We now consider special solution surfaces of the above exterior differential system given by a mapping $f : U \to M$ where the domain $U$ is an open subset of the $st$-plane. These solution surfaces are characterized by the dimension of their image under the projection to the $uv$-plane.

TYPE 0: These are solution surfaces $S \subset M$ on which

$$\pi_1 = \pi_2 = 0 \tag{3}$$

or equivalently $p =$ const and $q =$ const. The image of such a solution surface is thus a point in the $uv$-plane.

In a region of the $st$-plane where (3) is valid, $x$ and $y$ satisfy a linear hyperbolic PDE system with constant coefficients (cf. equation (10) of §4.1).

TYPE 1: These solutions $S \subset M$ are defined by the condition that

$$\pi_1 \wedge \pi_2 = 0, \tag{4}$$

and in addition are not of TYPE 0. Thus the image of a TYPE 1 solution surface in the $uv$-plane is a curve. These solutions are classically known as *simple waves.*

Note that for simple waves, the hyperbolic system $(M;\, \Omega_1,\, \Omega_2)$ reduces to the system $(N_1;\, \Omega_1)$ or $(N_2;\, \Omega_2)$ where $N_1$ and $N_2$ are the hypersurfaces in $M$ defined respectively by $q = q_0$ and $p = p_0$. In what follows we shall have occasion to study classical solutions of (1) which then naturally requires us to introduce the independence 2-form $\Omega = dx\wedge dy$. Now recall that the condition for genuine non-linearity of $(N_A;\, \Omega_A,\, \Omega)$ is that the common linear factor $\theta_A$ of $\Omega_A$ and $\Omega$ should be a contact form on $N_A$.

From (2) above we have that

$$\begin{aligned}\theta_1 &= dx - \lambda(p, q_0)\, dy,\\ \theta_2 &= dx - \mu(p_0, q)\, dy,\end{aligned}$$



and hence

$$\begin{aligned}\theta_1\wedge d\theta_1 &= \lambda_p\, dp\wedge dx\wedge dy,\\ \theta_2\wedge d\theta_2 &= \mu_q\, dq\wedge dx\wedge dy.\end{aligned}$$

This leads to the classical definition of *genuine non-linearity* for the PDE system (1), namely, that it should satisfy the conditions

$$\lambda_p \neq 0 \qquad \text{and} \qquad \mu_q \neq 0. \tag{5}$$

This concept is intrinsic to the data $(M;\, \Omega_1,\, \Omega_2;\, \Omega)$ consisting of the hyperbolic exterior differential system $(M;\, \Omega_1,\, \Omega_2)$ together with the independence condition $\Omega$.

TYPE 2: These are solutions on which

$$\pi_1 \wedge \pi_2 \neq 0.$$

Clearly, the image of such a solution surface is an open subset of the $uv$-plane.

We now proceed to develop a "picture" of the solutions to the system (2). For this we assume given classical initial data

$$x \mapsto (x,\, 0,\, u_0(x),\, v_0(x)) \tag{6}$$

where $u_0(x)$ and $v_0(x)$ are constant outside a compact set, say for $|x| > R$. We then use coordinates

$$\xi = \tfrac{1}{2}(s + t) \qquad \text{and} \qquad \eta = \tfrac{1}{2}(s - t)$$

in the $st$-plane and assume, as we may, that the initial data (6) is given by

$$\xi \mapsto (\xi,\, 0,\, p_0(\xi),\, q_0(\xi))$$

where $p_0$ and $q_0$ are constant for $|\xi| > R$. We also recall that

$$\begin{aligned}p(s,\, t) &= p(s,\, s) = p_0(s)\\ q(s,\, t) &= q(t,\, t) = q_0(t)\end{aligned} \tag{7}$$

are determined by the initial curve. This suggests we consider the following picture in the $\xi\eta$-plane

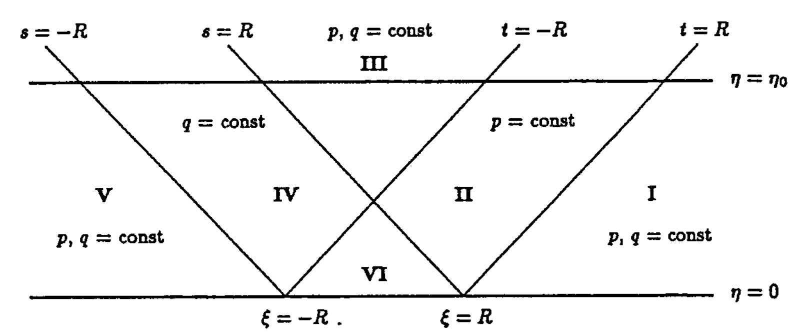


**Figure 4.2**

From (7) it follows that $p$ and $q$ are constant in regions **I**, **III**, **V**, so that the solution there is of TYPE 0. Moreover, $p$ is constant in region **II** and $q$ is constant in region **IV**, so that the solution there is of TYPE 1. For fixed $\eta = \eta_0 > R$ the solution along the $\xi$-axis

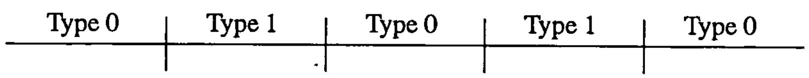


**Figure 4.3**

is of a wave that is alternately constant and simple with the simple wave components separating as $\eta$ increases. We also see from the above picture that the boundary of a region where we have a solution of TYPE 0 consists of characteristic curves and in the adjacent region the solution is of TYPE 1. Furthermore, outside of region **VI** and its mirror image in the half-plane $\eta < 0$, the mapping giving the complete integral surface with initial data (6) is an immersion.

We shall use the above picture to deduce the following well-known fact (see *[19]*):

*If the PDE system (1) is genuinely non-linear then any classical solution $\mathbf{u}(x, y)$ with non-constant compactly supported initial data will develop a singularity in finite time.*

**Proof:** The proof is by contradiction. Setting

$$\omega^1 = dx(\xi,\eta) - \lambda(\xi,\eta)\,dy(\xi,\eta),$$
$$\omega^2 = dx(\xi,\eta) - \mu(\xi,\eta)\,dy(\xi,\eta),$$

where $\lambda(\xi,\eta)$, $\mu(\xi,\eta)$ are given with the non-linearity property (5), we will show that $\Omega = \omega^1 \wedge \omega^2$ vanishes somewhere in the $\xi\eta$-plane.

Assuming this is not the case there will exist smooth functions $\rho(\xi,\eta)$ and $\sigma(\xi,\eta)$ such that

$$\pi^1 - \rho\,\omega^1 = 0 \qquad \text{and} \qquad \pi^2 - \sigma\,\omega^2 = 0.$$

The exterior derivative of the first of these equations gives

$$(\lambda - \mu)\,d\rho + \rho(\lambda_p\rho + \lambda_q\sigma)\,\omega^2 \equiv 0 \mod \omega^1.$$

Recall that for genuinely non-linear systems $\lambda_p \neq 0$. Now, because we want to illustrate the geometric approach rather than derive the most general result, we will assume that $\lambda_p$ is bounded away from zero, say $\lambda_p \geq \lambda_0 > 0$.

Note that for a fixed $s = s_0$ characteristic line we may assume that the parametrization is chosen so that $\omega^2 = (\mu - \lambda)\,dt$. Along this characteristic we then have

$$\rho_t = \rho(\lambda_p\rho + \lambda_q\sigma). \tag{8}$$



By assumption the initial data vanishes outside a compact set, say $|\xi| > R$. Thus we have

$$\left.\begin{aligned} p(\xi,0) &= 0 \\ q(\xi,0) &= 0 \end{aligned}\right\} \qquad \text{for } |\xi| > R.$$

Now notice that along the initial curve $\eta = 0$ we have

$$\rho(\xi,0) = \alpha(\xi)p_\xi(\xi,0) \qquad \text{and} \qquad \sigma(\xi,0) = \beta(\xi)q_\xi(\xi,0) \tag{9}$$

where $\alpha(\xi)$ and $\beta(\xi)$ are nowhere vanishing functions. Thus we have that on the set $|\xi| > R$

$$\rho(\xi,0) = \sigma(\xi,0) = 0.$$

Pictorially we have Figure 4.4. From equation (8), we have that $\rho = 0$ along $s =$ const characteristics that begin to the right of $\xi = R$. And by assumption there will be a point $P$ on the line $\eta = 0$ where $p_\xi(P, 0)$ vanishes, and changes sign as we cross from one side of $P$ to the other. Again by (8) $\rho$ will vanish along the characteristic represented by the dashed line, while we may assume that $\rho$ is positive at $Q$. Finally, along the characteristic emanating from $Q$ we will have $\sigma = 0$ and hence along this characteristic equation (8) gives

$$\rho_t = \lambda_p\rho^2 \geq \lambda_0\rho^2$$

with

$$\lambda_0 > 0 \qquad \text{and} \qquad \rho|_Q > 0$$

from which we infer that $\rho$ blows up in finite time. □

Recall that for a complete hyperbolic system (1), smooth initial data

$$s \mapsto (s, 0, u_0(s), v_0(s))$$

defined along the $x$-axis extends to a unique, smooth complete integral surface given by

$$f : \mathbb{R}^2_{(s,t)} \to \mathbb{R}^4$$

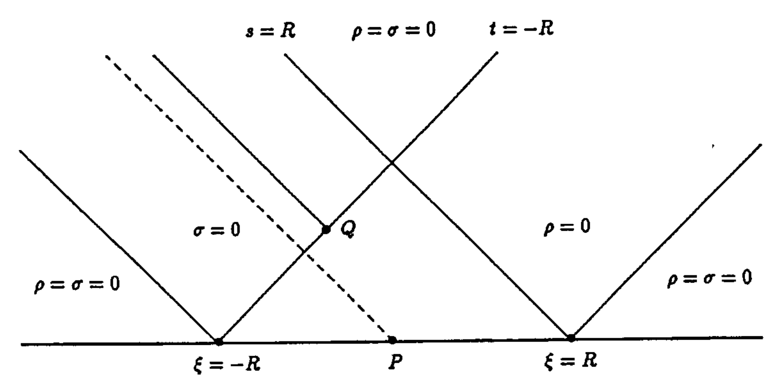


**Figure 4.4**

where $\mathbb{R}^2_{(s,t)}$ denotes the $st$-plane. Let us consider the mapping

$$\pi : \mathbb{R}^2_{(s,t)} \to \mathbb{R}^2_{(x,y)} \tag{10}$$

where $\pi$ is the composition of $f$ with the projection onto the $xy$-plane. By the above result we see that for compactly supported initial data the differential of $\pi$ must drop rank somewhere. In fact, we have the relation (cf. equation (10) of §4.1)

$$\det\begin{pmatrix} x_s & x_t \\ y_s & y_t \end{pmatrix} = (\mu - \lambda) y_s y_t$$

which shows that the curve where $\pi_*$ drops rank is given by

$$y_s y_t = 0.$$

Thus it has two "branches", one where $y_s = 0$ (and thus $x_s = 0$) and the other $y_t = 0$ (and thus $x_t = 0$). These branches intersect at a point where $x_s = x_t = y_s = y_t = 0$, which is a cusp for the mapping $\pi$. Along the branch $y_s = 0$, $y_t \neq 0$ the kernel of $\pi_*$ is spanned by $\partial_s$, i.e., it is a characteristic direction.

We now address the question of whether the mapping (10) is surjective; i.e., does our construction give a global albeit "multi-valued" solution to (1)? In the following we shall show that

*If the initial data is constant outside of a compact set, then the mapping (10) is proper. In particular, it is surjective.*

**Proof:** Explicitly, the mapping (10) is given by

$$\pi(s,t) = \big(x(s,t),\, y(s,t)\big)$$

where $x(s,t)$ and $y(s,t)$ satisfy the linear PDE (10) of the preceding section. To show that $\pi$ is proper we need to understand the behavior of the characteristic curves, which are the images of the straight lines parallel to the coordinate axes in the $st$-plane. Thus we consider Figure 4.5. The characteristic curves in the $xy$-plane are the integral curves of

$$\begin{aligned} dx - \lambda(p,q)\,dy &= 0 \\ dx - \mu(p,q)\,dy &= 0. \end{aligned}$$

Recall that $p = p(s,t) = p(s)$ and $q = q(s,t) = q(t)$ by the argument in the preceding section. It follows that

$p = \text{const}$ in regions **I**, **III**, **V** and **II**,
$q = \text{const}$ in regions **I**, **III**, **V** and **IV**.

Thus, the characteristic curves in the images of regions **I**, **III** and **V** are two families of parallel lines.



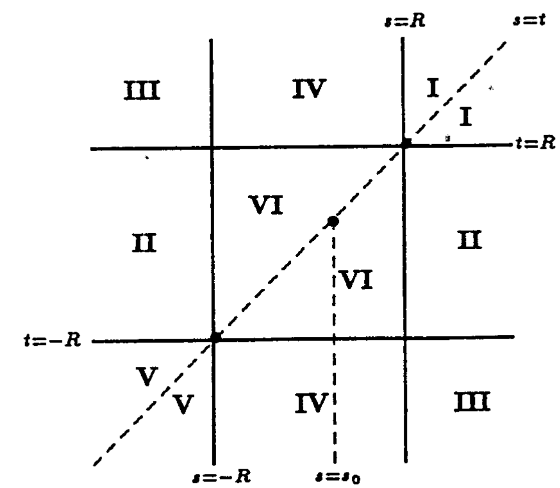


**Figure 4.5**

We will now examine what happens to the images of the lines $s = s_0$ as $s_0$ moves from $R$ to $-R$. In region **IV**, $q$ is a constant $q_0$ and along the dotted vertical line $p = p(s_0)$ is also constant. Thus the image of the $t \le -R$ part of the dotted line in region **IV** will be a line in the $xy$-plane since it satisfies

$$dx - \lambda\big(p(s_0), q_0\big) dy = 0.$$

As $s_0$ moves from right to left from $R$ to $-R$ these lines will not be parallel but will have slope $\lambda\big(p(s_0), q_0\big)$ that changes continuously for $-R \le s_0 \le R$. (Thus in the $xy$-plane they may focus, causing folds for $\pi$ in region **IV**). However, we see from this picture that the mapping will be proper in region **IV** as $s$ tends to $\infty$. Similar arguments apply to the other region **IV** and to the two regions **II**. From this we may conclude that the mapping $\pi : \mathbb{R}^2_{(s,t)} \to \mathbb{R}^2_{(x,y)}$ is proper. □

**4.3** *The Riemann Problem for Hyperbolic Exterior Differential Systems.* In §2 we have considered the geometry associated to the non-linear, scalar conservation law

$$u_y + f'(u)u_x = 0 \tag{1}$$

where the non-linearity is expressed by the condition $f''(u) \neq 0$. Setting $\Phi = du \wedge \big(dx - f'(u)\,dy\big) = d\varphi$ for some 1-form $\varphi = u\,dx - f(u)\,dy$, the associated exterior differential system is given by $\Phi = 0$ on the manifold $M = \mathbb{R}^3_+ = \{(x,y,u) : y \ge 0\}$. Initial data given by $\big(x, 0, u_0(x)\big)$ where $u_0(x)$ is smooth and constant for $|x| > R$ gives rise to an initial curve $\gamma \subset M$ which then extends to a unique, smooth global integral surface $S \subset M$ such that the projection

$$\pi : S \to \mathbb{R}^2_+ = \{(x,y) : y \ge 0\} \tag{2}$$

is proper. (Here global means that $S$ is characteristically complete in the sense that the characteristic curves are complete geodesics in the intrinsic geometry associated to (1).) Furthermore, it is well known that there is a unique, global shock solution $u(x, y)$ of (1) associated to the above initial data, and in §2 we gave heuristic reasoning to the effect that $u(x, y)$ arises by taking a suitable (possibly discontinuous) cross section

$$(x, y) \mapsto (x, y, u(x, y)) \in S$$

of the projection (2).

We now turn to a hyperbolic system of conservation laws

$$\begin{cases} u_y + \partial_x(f(u, v)) = 0 \\ v_y + \partial_x(g(u, v)) = 0 \end{cases} \tag{3}$$

which we also write as

$$\mathbf{u}_y + \partial_x(\mathbf{f}(\mathbf{u})) = 0$$

where $\mathbf{u} = {}^t(u, v)$ and $\mathbf{f} = {}^t(f, g)$. This gives rise to the hyperbolic exterior differential system $\Phi = \Psi = 0$ on $M = \mathbb{R}^4_+ = \{(x, y, u, v) : y \geq 0\}$ where $\Phi = d\phi$ and $\Psi = d\psi$ for 1-forms $\phi, \psi$ defined by

$$\begin{cases} \phi = u\,dx - f(u, v)\,dy \\ \psi = v\,dx - g(u, v)\,dy. \end{cases} \tag{4}$$

With an appropriate assumption on the completeness of (3) we shall show that smooth initial data

$$x \mapsto (x, 0, \mathbf{u}_0(x))$$

satisfying $\mathbf{u}_0(x) = \text{const}$ for $|x| > R$ gives a unique, smoothly parametrized global integral surface $S \subset M$ such that the projection

$$\pi : S \to \mathbb{R}^2_+ = \{(x, y) : y \geq 0\} \tag{5}$$

is proper. (Here, global means that $S$ should be characteristically complete as explained above.)

In PDE theory (see *[20]* for a general reference and *[14]* for recent developments) it is proved that if (3) is genuinely non-linear and the initial data has small mean oscillation, then there is a unique shock solution $\mathbf{u}(x, y)$ to (3) with the given initial data.

A natural question to ask is whether this shock solution arises by taking a suitable cross-section of (5)? The situation is much more subtle than in the scalar case, and as we shall see when we discuss the Riemann problem below the answer appears to be no in general. However, for a special class of hyperbolic systems which we shall describe in the following, this construction is indeed possible.

Briefly, the problem is that the Rankine-Hugonoit conditions for a weak solution—the analog of the jump condition (14) in §2—are now the pair of equations

$$[f(u, v)] = s[u] \qquad \text{and} \qquad [g(u, v)] = s[v]. \tag{6}$$



On a solution surface $S$, this is an over-determined set of equations for the shock speed $s$. We shall see that, at least for the Riemann problem, these conditions can be satisfied on $S$ only when the defining functions $f(u, v)$ and $g(u, v)$ of the hyperbolic system (3) satisfy a pair of second-order PDEs (to be derived below). We shall call such hyperbolic systems of conservation laws *special.*

We now proceed to study the analogue for exterior differential systems of the classical Riemann problem for (1) and (3). Recall that this is the PDE initial value problem for the initial data

$$u_0(x) = \begin{cases} u_- & \text{for } x < 0, \\ u_+ & \text{for } x > 0, \end{cases} \tag{7a}$$

in the case of equation (1) and

$$\mathbf{u}_0(x) = \begin{cases} \mathbf{u}_- & \text{for } x < 0, \\ \mathbf{u}_+ & \text{for } x > 0, \end{cases} \tag{7b}$$

in the case of equation (3).

Notice that the initial data (7) gives a disconnected curve $\gamma_0$ in $M$. In the following we shall show how $\gamma_0$ may be uniquely completed to a connected, continuous and piecewise smooth curve $\gamma$ such that there is a unique, global weak solution surface $S$ to the corresponding exterior differential system with the additional property that $S$ consists only of Type 0 or Type 1 solutions in the sense of the preceeding section. (The reason that we restrict $S$ to consists of Type 0 and Type 1 solutions is that the graph of the classical shock solutions have this property. Furthermore, this is a natural geometric condition to impose on a solution surface.) In the scalar case the shock solution to the Riemann problem arises as a suitable cross-section of

$$\pi : S \to \mathbb{R}^2_+,$$

but this is not in general true for (3). As stated above, this will be valid only when the system (3) is special.

We begin with a brief discussion of the non-linear scalar conservation law (1). In the following we shall see that there are two basic cases:

(i) $u_- \geq u_+$ — breaking waves; (ii) $u_- \leq u_+$ — rarefaction waves.

We will show that "upstairs" on the geometric solution surface $S$ these two cases are essentially the same, which is not the case "downstairs" for the classical shock solution $\mathbf{u}(x, y)$.

We prescribed the initial data by a mapping

$$s \mapsto \gamma(s) = (x(s), 0, u_0(s)), \qquad s \in \mathbb{R}$$

as follows: For $a \leq b$ with $b - a = |f'(u_+) - f'(u_-)|$ we set

$$\gamma(s) = \begin{cases} (s - a, 0, u_-) & \text{for } s \leq a \\ (s - b, 0, u_+) & \text{for } s \geq b \end{cases}$$

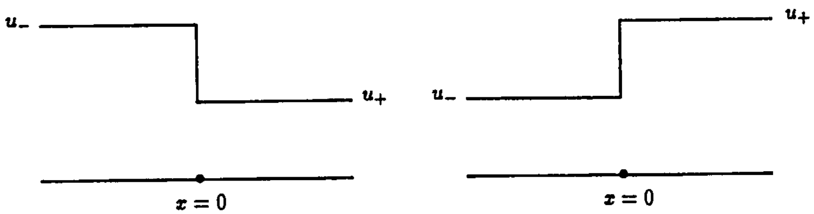


**Figure 4.6**

while for $a \le s \le b$, the data $\gamma(s)$ is defined by

$$x(s) = 0 \qquad \text{and} \qquad f'(u_0(s)) = \pm s$$

where $u_0(a) = u_-$ and $u_0(b) = u_+$. Here, $\pm$ is the sign of $u_+ - u_-$ (thus $+$ for rarefaction and $-$ for breaking waves). The picture is as follows (see Figure 4.6). From the relation

$$f''(u_0(s))\,u_0'(s) = \pm 1$$

together with the non-linearity assumption $f'' > 0$, we see that the sign is determined by whether $u_0(s)$ is increasing or decreasing on the vertical segment.

The characteristic vector field is

$$X = \partial_y + f'(u)\,\partial_x$$

and we define $S$ to be the image under the mapping $\Gamma$ of $\mathbb{R}^2_+ = \{(s,t) : t \ge 0\}$ obtained by flowing $\gamma$ along the integral curves of $X$. Explicitly

$$\Gamma(s,t) = \big(x(s) + tf'(u_0(s)),\, t,\, u_0(s)\big). \tag{8}$$

This is a continuous, piecewise smooth mapping with (finite) jumps in the differential $\Gamma_*$ arising from

$$f''(u_0(s))\,u_0'(s) = \pm 1, \qquad a \le s \le b,$$

which gives the jump at $s = a$

$$[u_0'] = \pm\frac{1}{f''(u_-)},$$

and similarly at $s = b$. We note that in $a \le s \le b$ we have

$$u_0(s) = (f')^{-1}(\pm s)$$

according to whether $u_- \le u_+$ or $u_- \ge u_+$. Thus, upstairs a shock wave looks like a "backwards" rarefaction wave.



The mapping (8) is a *weak solution* to the exterior differential system

$$d\varphi = 0 \tag{9}$$

in the following sense: First, $\Gamma^*(\varphi)$ is a continuous 1-form in $\mathbb{R}^2_+$. Secondly, for any function $\alpha$ with compact support in $\mathbb{R}^3_+ = \{(x, y, u) : y > 0\}$, $\Gamma^*(\alpha)$ is a compactly supported piecewise smooth function in $\mathbb{R}^2_+ = \{(s, t) : t > 0\}$ and thus the equation

$$\int d\Gamma^*(\alpha) \wedge \Gamma^*(\varphi) = 0 \tag{10}$$

makes sense, and is satisfied for all $\alpha$.

For $S_0 = S - \gamma$ the mapping $\pi : S_0 \to \tilde{\mathbb{R}}^2_+ = \{(x, y) : y > 0\}$ is one-to-one in the rarefaction case, while in the shock case the image in the $xy$-plane of the region $a \le s \le b$, $t > 0$ is covered three times by the mapping $\pi$. The entropy condition (see §2) tells us that along the line

$$\frac{x}{y} = \frac{f(u_-) - f(u_+)}{u_- - u_+}$$

we should jump from the top to the bottom sheet of $\pi^{-1}(x, y)$ to obtain the shock solution.

Turning to the Riemann problem for the exterior differential system arising from (3), it turns out that through each point of the $uv$-plane there are two pairs of curves, one arising from the level sets of Riemann invariants and the other arising from the solution curves to the jump conditions (6). In order to understand the Riemann problems "upstairs" and "downstairs", as well as their relationship, we shall need to understand the local geometry of these curves and how they relate to each other.

Recall that a *Riemann invariant* is a function $r(u, v)$ that is constant on one of the families of characteristics on solutions $\big(u(x, y), v(x, y)\big)$ of (3). The characteristics are integral curves of

$$\text{(i) } dx - \lambda\, dy = 0 \qquad \text{or} \qquad \text{(ii) } dx - \mu\, dy = 0.$$

We shall denote by $r_\lambda, r_\mu$ the Riemann invariants corresponding to (i) and (ii) respectively.

Next, we observe that the form of the equation (3) remains invariant under the affine linear transformation

$$\tilde{\mathbf{u}} = \mathbf{a}_0\mathbf{u} + \mathbf{v}_0 \tag{11}$$

where $\mathbf{a}_0$ is an invertible 2-by-2 matrix and $\mathbf{v}_0$ is a vector. Thus, we may intrinsically think of $(u, v)$ as coordinates in an affine vector space. For each point $\mathbf{u}$ in the affine space we denote by $\mathbf{e}_\lambda(\mathbf{u}), \mathbf{e}_\mu(\mathbf{u})$ eigenvectors of $\mathbf{f}'(\mathbf{u})$ corresponding to $\lambda(\mathbf{u}), \mu(\mathbf{u})$ respectively. We may identify the tangent space at each point $\mathbf{u}$ with the affine space itself and so in this way we may think of $\mathbf{e}_\lambda, \mathbf{e}_\mu$ as a pair of direction fields.

**Lemma 1:** *The level sets $r_\lambda$ = const and $r_\mu$ = const are the integral curves of $\mathbf{e}_\lambda$ and $\mathbf{e}_\mu$ respectively.*

**Proof:** We recall that a simple wave for (3) is a solution $\mathbf{u}(x, y)$ whose image in the $uv$-plane is a 1-dimensional curve. We shall show that the level sets of simple waves are also level sets of Riemann invariants as well as integral curves of the above direction fields.

For a simple wave the level sets $\mathbf{u}(x, y)$ = const are a family of curves giving a foliation $\xi(x, y)$ = const in an open set in the $xy$-plane. This implies that $\mathbf{u} = \mathbf{u}(\xi)$, a function of 1-variable. The PDE (3) gives

$$(\xi_y + \xi_x \mathbf{f}'(\mathbf{u}))\mathbf{u}' = 0$$

which means that $\lambda = -\xi_y/\xi_x$ is an eigenvalue of $\mathbf{f}'(\mathbf{u})$ with eigenvector $\mathbf{e}_\lambda = \mathbf{u}'$. Thus we have

$$\text{(i) } \xi_y + \lambda(\mathbf{u}(\xi))\xi_x = 0; \qquad \text{(ii) } \mathbf{u}'(\xi) = \mathbf{e}_\lambda(\mathbf{u}(\xi)).$$

Rewriting equation (i) as

$$(\partial_y + \lambda(\mathbf{u}(\xi))\partial_x)\xi = 0$$

and using the identity

$$\langle dx - \lambda\, dy,\ \partial_y + \lambda\, \partial_x \rangle = 0$$

we see that the curve $\xi \mapsto \mathbf{u}(\xi)$ in the $uv$-plane is a level set of the Riemann invariant $r_\lambda$.

Now, from equation (ii) we have

$$0 = dr_\lambda(\mathbf{u}(\xi)) = \langle dr_\lambda, \mathbf{e}_\lambda \rangle (\mathbf{u}(\xi))$$

and so $\langle dr_\lambda, \mathrm{e}_\lambda \rangle = 0$ as require. □

To study the geometry of the various curves in a neighborhood of a point $\mathbf{u}_0$, we make an affine transformation (11) so that $\mathbf{u}_0 = (\mathbf{0})$ is the origin and $\mathrm{e}_\lambda(0) = \partial_u, \mathrm{e}_\mu(0) = \partial_v$. Then we have the series expansions

$$\begin{aligned} f(u,v) &= \lambda_0 u + \tfrac{1}{2} f_{uu} u^2 + f_{uv} uv + \tfrac{1}{2} f_{vv} v^2 + \mathrm{O}(3) \\ g(u,v) &= \mu_0 v + \tfrac{1}{2} g_{uu} u^2 + g_{uv} uv + \tfrac{1}{2} g_{vv} v^2 + \mathrm{O}(3) \end{aligned} \tag{12}$$

where O($k$) denotes terms of order $k$.

**Lemma 2:** *In a neighborhood of the origin $\mathbf{u} = (\mathbf{0})$, we have*

$$\begin{aligned} \lambda(u, v) &= \lambda_0 + f_{uu} u + f_{uv} v + \mathrm{O}(2), \\ \mu(u, v) &= \mu_0 + g_{vv} v + g_{uv} u + \mathrm{O}(2). \end{aligned} \tag{i}$$

*Futhermore, we may choose the eigenvectors and Riemann invariants such that*

$$\begin{aligned} \mathbf{e}_\lambda(\mathbf{u}, \mathbf{v}) &= \partial_u + (\lambda_0 - \mu_0)^{-1}(g_{uu} u + g_{uv} v)\partial_v + \mathrm{O}(2), \\ \mathbf{e}_\mu(\mathbf{u}, \mathbf{v}) &= \partial_v + (\mu_0 - \lambda_0)^{-1}(f_{vv} v + f_{uv} u)\partial_u + \mathrm{O}(2), \end{aligned} \tag{ii}$$

*and*

$$\begin{aligned} r_\lambda(u, v) &= v - (\lambda_0 - \mu_0)^{-1}(\tfrac{1}{2} g_{uu} u^2 + g_{uv} uv + \tfrac{1}{2} g_{vv} v^2) + \mathrm{O}(3), \\ r_\mu(u, v) &= u - (\mu_0 - \lambda_0)^{-1}(\tfrac{1}{2} f_{uu} u^2 + f_{uv} uv + \tfrac{1}{2} f_{vv} v^2) + \mathrm{O}(3). \end{aligned} \tag{iii}$$

We remark that $\mathbf{e}_\lambda$, $\mathbf{e}_\mu$ and $r_\lambda$, $r_\mu$ are not unique—i.e., the eigenvectors are defined up to a non-zero factor while any non-constant function of the Riemann invariants will still be a Riemann invariant. The above formula should be interpreted as giving expressions for representatives of equivalence classes.

**Proof:** The linear terms in $\lambda$ and $\mu$ are determined from the equations

$$\begin{pmatrix} 1+\alpha & \beta \\ \gamma & 1+\delta \end{pmatrix}\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}\begin{pmatrix} 1-\alpha & -\beta \\ -\gamma & 1-\delta \end{pmatrix} = \begin{pmatrix} \lambda_0 + f_{uu} u + f_{uv} v & f_{uv} u + f_{vv} v \\ g_{uu} u + g_{uv} v & \mu_0 + g_{uv} u + g_{vv} v \end{pmatrix}$$

where $\alpha$, $\beta$, $\gamma$, $\delta$ are homogeneous linear forms in $u$, $v$. Expanding this out gives (i).

For (ii) we write $\mathrm{e}_\lambda = {}^t(1, \eta) = \partial_u + \eta\,\partial_v$ where $\eta$ is homogeneous and linear in $u$, $v$, and then we solve for $\eta$ in

$$\begin{pmatrix} \lambda_0 + f_{uu} u + f_{uv} v & f_{uv} u + f_{vv} v \\ g_{uu} u + g_{uv} v & \mu_0 + g_{uv} u + g_{vv} v \end{pmatrix}\begin{pmatrix} 1 \\ \eta \end{pmatrix} = (\lambda_0 + f_{uu} u + f_{uv} v)\begin{pmatrix} 1 \\ \eta \end{pmatrix}$$

to obtain

$$\eta = (\lambda_0 - \mu_0)^{-1}(g_{uu} u + g_{uv} v) + \mathrm{O}(2).$$

A similar argument gives the expression for $\mathbf{e}_\mu$.

Finally, (iii) follows from the previous lemma and

$$\begin{aligned} \langle dr_\lambda, \mathbf{e}_\lambda \rangle &\equiv 0 \mod \mathrm{O}(2) \\ \langle dr_\mu, \mathbf{e}_\mu \rangle &\equiv 0 \mod \mathrm{O}(2). \end{aligned}$$

□

From equation (iii) above we see that the level sets of the Riemann invariants, passing through the origin, are given respectively by

$$v = \tfrac{1}{2}(\lambda_0 - \mu_0)^{-1} g_{uu} u^2 + \mathrm{O}(3) \tag{14a}$$

and

$$u = \tfrac{1}{2}(\mu_0 - \lambda_0)^{-1} f_{vv} v^2 + \mathrm{O}(3). \tag{14b}$$

We now consider the solution curves of the jump equations (6) where we assume (without loss of generality) that the starting point is again $\mathbf{u} = (\mathbf{0})$. These equations have the form

$$\begin{aligned} (s - \lambda_0)u &= P(u, v) \\ (s - \mu_0)v &= Q(u, v) \end{aligned} \tag{15}$$

where

$$P(u, v) = \tfrac{1}{2} f_{uu} u^2 + f_{uv} uv + \tfrac{1}{2} f_{vv} v^2 + O(3)$$
$$Q(u, v) = \tfrac{1}{2} g_{uu} u^2 + g_{uv} uv + \tfrac{1}{2} g_{vv} v^2 + O(3).$$

Eliminating $s$ gives

$$(\lambda_0 - \mu_0)uv + vP - uQ = 0.$$

This equation has two smooth branches, with tangent vectors $\partial_u$ and $\partial_v$, given respectively by

$$v = \tfrac{1}{2}(\lambda_0 - \mu_0)^{-1} g_{uu} u^2 + O(3) \tag{16a}$$

and

$$u = \tfrac{1}{2}(\mu_0 - \lambda_0)^{-1} f_{vv} v^2 + O(3). \tag{16b}$$

Comparing with (14) we see that:

> *Through each point of the uv-plane, the solutions with $s \neq 0$ of the jump equations (6) form a pair of curves which osculate to second, but not in general to third, order to the pair of level curves of the Riemann invariants.*

Clearly the level sets of the Riemann invariants $r_\lambda$ and $r_\mu$ define a pair of local foliations, henceforth denoted by $\rho_\lambda$ and $\rho_\mu$ respectively, of the $uv$-plane. Somewhat more subtle is the fact that the two families of jump curves, now denoted by $\delta_\lambda$ and $\delta_\mu$, *need not* necessarily form a pair of local foliations.

The problem is that the relation defined by, say, the $\delta_\lambda$ curves is symmetric but not necessarily transitive. Consider two nearby points $P$, $Q$ in the $uv$-plane. We say that $Q$ is $\lambda$-related to $P$ in case $Q$ lies on the $\lambda$-jump curve of $P$, denoted by $\delta_\lambda(P)$. Since the jump equation (6) is symmetric, it is clear that $P$ is also $\lambda$-related to $Q$, i.e., $P$ lies on the $\lambda$-jump curve of $Q$, $\delta_\lambda(Q)$, but in general $\delta_\lambda(Q)$ will be different from $\delta_\lambda(P)$ (see Figure 4.7). For the point $R$ as depicted above we have that $Q$ and $R$ are $\lambda$-related but in general $P$ and $R$ will *not* be $\lambda$-related. This situation is clarified by the following

**Lemma 3:** *The two families of jump curves form a pair of local foliations if and only if they coincide with the two Riemann foliations, and furthermore are local foliations of the uv-plane by straight lines.*

An obvious case when the $\delta$ and $\rho$ foliations coincide is when the system (3) *uncouples*, in the sense that there is an affine linear change of variables such that (3) becomes

$$u_y + \partial_x(f(u)) = 0$$
$$v_y + \partial_x(g(u)) = 0.$$

In this case the $\delta$-curves are two families of *parallel* lines. In general, however, the conditions of the lemma may be satisfied by two families of non-parallel lines (although, clearly, if the foliations are *global* in the $uv$-plane then uncoupling will occur).



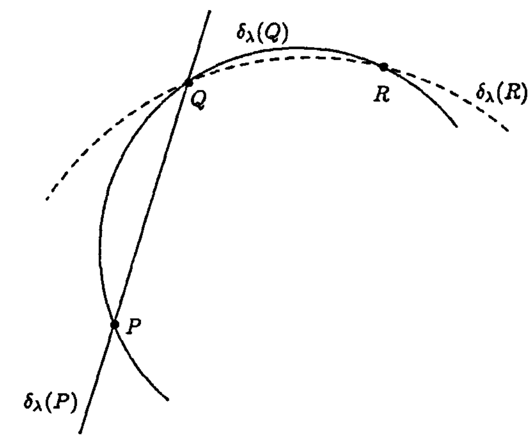


**Figure 4.7**

**Proof:** For $\mathbf{v}$ close to $\mathbf{u}$ we consider the locus $\Sigma$ defined by the equations

$$\mathbf{f}(\mathbf{v}) - \mathbf{f}(\mathbf{u}) = s(\mathbf{v} - \mathbf{u})$$

in $(\mathbf{u}, \mathbf{v}, s) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}$. From the previous lemma, we may infer that

$$\Sigma = \Sigma_\Delta \cup \Sigma_\lambda \cup \Sigma_\mu$$

where $\Sigma_\Delta$, $\Sigma_\lambda$, $\Sigma_\mu$ are smooth 3-manifolds with

$$\Sigma_\lambda = \{(\mathbf{u}, \mathbf{v}, s) \in \mathbb{R}^5 : \mathbf{v} \text{ is } \lambda\text{-related to } \mathbf{u}\},$$
$$\Sigma_\mu = \{(\mathbf{u}, \mathbf{v}, s) \in \mathbb{R}^5 : \mathbf{v} \text{ is } \mu\text{-related to } \mathbf{u}\},$$

and

$$\Sigma_\Delta = \{(\mathbf{u}, \mathbf{v}, s) \in \mathbb{R}^5 : \mathbf{u} = \mathbf{w} \text{ and } \mathbf{v} = \mathbf{w}\}.$$

Thus, $(\mathbf{u}, \mathbf{v}, s) \in \Sigma_\lambda$ means that $\mathbf{v} \in \delta_\lambda(\mathbf{u})$ and that if $\mathbf{v} \neq \mathbf{u}$ then $s$ is uniquely determined by the jump equations above. We note that, as $\mathbf{v}$ approaches $\mathbf{u}$ along $\delta_\lambda(\mathbf{u})$, $s$ approaches $\lambda(\mathbf{u})$. This suggests that we take $(\mathbf{u}, s)$ as local coordinates on $\Sigma_\lambda$ and use $s - \lambda(\mathbf{u})$ as parameter along $\delta_\lambda(\mathbf{u})$. We then write the Taylor's series for $\mathbf{v}$ along $\delta_\lambda(\mathbf{u})$ as

$$\mathbf{v}_\lambda(\mathbf{u}, s) = \mathbf{u} + (s - \lambda(\mathbf{u}))\mathbf{e}_\lambda(\mathbf{u}) + O((s - \lambda(\mathbf{u}))^2).$$

We now determine the conditions that the $\delta_\lambda$-curves define a transitive relation. Thus suppose that $\mathbf{v}_1 = \mathbf{v}(\mathbf{u}, r)$ and $\mathbf{v}_2 = \mathbf{v}(\mathbf{u}, s)$ are $\lambda$-related to $\mathbf{u}$, i.e.,

$$\mathbf{f}(\mathbf{v}_\lambda(\mathbf{u}, r)) - \mathbf{f}(\mathbf{u}) = r(\mathbf{v}_\lambda(\mathbf{u}, r) - \mathbf{u})$$
$$\mathbf{f}(\mathbf{v}_\lambda(\mathbf{u}, s)) - \mathbf{f}(\mathbf{u}) = s(\mathbf{v}_\lambda(\mathbf{u}, s) - \mathbf{u}).$$

Transitivity now implies that $\mathbf{v}_2$ is $\lambda$-related to $\mathbf{v}_1$ and so there a $t = t_\lambda(\mathbf{u}, s, r)$ such that

$$\mathbf{f}(\mathbf{v}_\lambda(\mathbf{u}, r)) - \mathbf{f}(\mathbf{v}_\lambda(\mathbf{u}, s)) = t_\lambda(\mathbf{u}, s, r)(\mathbf{v}_\lambda(\mathbf{u}, r) - \mathbf{v}_\lambda(\mathbf{u}, s)).$$

This equation uniquely determines $t_\lambda(\mathbf{u}, s, r)$ as a continuous function of $(\mathbf{u}, s, r)$ for $s$ and $r$ distinct and close to $\lambda(\mathbf{u})$. Moreover, setting $t_\lambda(\mathbf{u}, s, s) = \lambda(\mathbf{v}_\lambda(\mathbf{u}, s))$ defines $t_\lambda(\mathbf{u}, s, r)$ continuously for all $(\mathbf{u}, s, r)$ with $s$ and $r$ close to $\lambda(\mathbf{u})$.

The above equations give

$$r(\mathbf{v}_\lambda(\mathbf{u}, r) - \mathbf{u}) - s(\mathbf{v}_\lambda(\mathbf{u}, s) - \mathbf{u}) = t_\lambda(\mathbf{u}, s, r)(\mathbf{v}_\lambda(\mathbf{u}, r) - \mathbf{v}_\lambda(\mathbf{u}, s)).$$

Now, fix $\mathbf{u}$ and consider the following curve in the $uv$-plane

$$\gamma(s) = \mathbf{v}_\lambda(\mathbf{u}, s) - \mathbf{u}.$$

The previous equation now gives

$$0 = (r\gamma(r) - s\gamma(s)) \wedge (\gamma(r) - \gamma(s)) = (s - r)\,\gamma(r) \wedge \gamma(s).$$

Thus $\gamma(s)$ lies on a line through the origin, as was to be shown.

Conversely, if the $\delta$-curves are straight lines then transitivity is clear. $\square$.

**Definition:** We will say that (3) is a *special* hyperbolic system of conservation laws in case the level sets of the Riemann invariants are straight lines.

We note that this is an intrinsic condition on a hyperbolic system of conservation laws (since only affine linear transformations of the $uv$-plane will preserve the form of (3)). We now proceed to describe all such special systems.

For an eigen-direction $\mathbf{e}(\mathbf{u}, \mathbf{v})$ of $\mathbf{f}'(\mathbf{u})$ we define

$$\dot{\mathbf{e}} = \nabla_{\mathbf{e}}\mathbf{e}.$$

Then $\dot{\mathbf{e}}$ is well-defined modulo $\mathbf{e}$ and

$$\mathbf{e} \wedge \dot{\mathbf{e}} = 0$$

is the condition that the integral curves of $\mathbf{e}$ be straight lines. This equation is a second-order condition on $\mathbf{f}(\mathbf{u})$ at each point of the $uv$-plane—taking into account both eigen-directions, we see that special systems of hyperbolic conservation laws are given by (3) with $\mathbf{f}(\mathbf{u}) = {}^t(f(u, v), g(u, v))$ satisfying a pair of second-order PDEs. We may suspect that locally the solutions to this system depend on 4 arbitrary functions of 1 variable. We will see that this is indeed the case by the following geometric construction.



Each point $\mathbf{u}_0$ has a neighborhood consisting of a quadrilateral made up of lines in the two foliations. By an affine transformation we may assume that two edges of the quadrilateral are the $u$ and $v$ axes and that the $\lambda$ and $\mu$ foliations are described by lines of the form

$$u = \alpha(p)v + p, \qquad 0 \le p < p_0,\ \alpha(0) = 0,$$
$$v = \beta(q)u + q, \qquad 0 \le q < q_0,\ \beta(0) = 0,$$

respectively. We may solve for $u$ and $v$ to obtain

$$\begin{pmatrix} u \\ v \end{pmatrix} = (1 - \alpha(p)\beta(q))^{-1} \begin{pmatrix} p + q\alpha(p) \\ q + p\beta(q) \end{pmatrix}.$$

As long as $\alpha(p)\beta(q) < 1$ this gives a valid formula and furthermore, if $(1+\alpha'(p)v)(1+\beta'(q)u) \neq 0$ it will be smoothly solvable for $p$ and $q$ as functions of $u$ and $v$.

Now, we can construct another map to a $\tilde{u}\tilde{v}$-plane by considering the equations

$$\tilde{u} = \alpha(p)\tilde{v} + \varphi(p)$$
$$\tilde{v} = \beta(q)\tilde{u} + \psi(q)$$

which solves to give

$$\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = (1 - \alpha(p)\beta(q))^{-1} \begin{pmatrix} \varphi(p) + \alpha(p)\psi(q) \\ \psi(q) + \beta(q)\varphi(p) \end{pmatrix}.$$

It is not hard to see that, as long as $\varphi$ and $\psi$ satisfy an open condition, the induced mapping $\mathbf{f}(u, v) = (\tilde{u}, \tilde{v})$ will be hyperbolic and genuinely non-linear. Moreover, this mapping carries each $p$-line to a $p$-line of the same slope and each $q$-line to a $q$-line of the same slope. Thus, it has the desired properties, and consequently, the solutions to our problem depend on four functions of one variable.

Note that if $\alpha$ and $\beta$ are not constant, this system will not uncouple.

We now return to our general discussion where the $\delta$ and $\rho$ curves osculate to second order but need not coincide. Assuming that (3) is genuinely non-linear we shall show that each of the $\delta$ and $\rho$ curves has an *orientation*. These orientations will be constructed from the entropy conditions used in the construction of shock solutions to the Riemann problem.

Recall that the conditions for genuine non-linearity are

$$\langle d\lambda, \mathbf{e}_\lambda \rangle \neq 0 \qquad \text{and} \qquad \langle d\mu, \mathbf{e}_\mu \rangle \neq 0.$$

From Lemma 2 we have that $f_{uu}g_{vv} \neq 0$. For definiteness we will assume that

$$f_{uu} > 0 \qquad \text{and} \qquad g_{vv} > 0.$$

Now, a shock solution $\mathbf{u}(x, y)$ to (3) is given for $y > 0$ by (see *[20]*)

$$\mathbf{u} = \begin{cases} \mathbf{u}_- & \text{for } x/y < s, \\ \mathbf{u}_+ & \text{for } x/y > s. \end{cases}$$

In addition, one of the following *entropy conditions*

$$s < \lambda(\mathbf{u}_-) \qquad \text{and} \qquad \lambda(\mathbf{u}_+) < s < \mu(\mathbf{u}_+)$$

or

$$\mu(\mathbf{u}_+) < s \qquad \text{and} \qquad \lambda(\mathbf{u}_-) < s < \mu(\mathbf{u}_-)$$

must be satisfied. A shock solution satisfying the first set of entropy conditions above is called a 1-*shock* while a solution satisfying the second set of conditions is called a 2-*shock*.

Now, let us set $\mathbf{u}_- = (\mathbf{0})$ and assume that $\mathbf{u}_+$ is in a neighborhood of the origin. The jump conditions specify that $\mathbf{u}_+$ must lie on the $\delta_\lambda(\mathbf{0})$ or $\delta_\mu(\mathbf{0})$ curve. We shall show that the condition for a 1-shock solution is that

$$\mathbf{u}_+ \in \delta_\lambda^-(0) = \{(u, p(u)) \in \delta_\lambda(0) : u < 0\}.$$

**Proof:** Solving for $s$ in (15) along the arc $\delta_\lambda(\mathbf{0})$ gives

$$s = \lambda_0 + \tfrac{1}{2} f_{uu} u + O(2).$$

The conditions for a 1-shock are

$$\lambda_0 + \tfrac{1}{2} f_{uu} u + O(2) < \lambda_0$$

and

$$\lambda_0 + \tfrac{1}{2} f_{uu} u + O(2) < \lambda_0 + \tfrac{1}{2} f_{uu} u + O(2) < \mu_0 + \tfrac{1}{2} g_{uv} u + O(2)$$

and these are satisfied only for $u < 0$, $|u| \leq \epsilon$. Similar calculations show that the conditions for a 2-shock cannot be satisfied anywhere along $\delta_\lambda(\mathbf{0})$. □

The orientation on $\delta_\lambda$ and $\delta_\mu$ is now clear: The positive direction is from $\delta_\lambda^-$ to $\delta_\lambda^+$, and similarly for $\delta_\mu$. Since the $\rho_\lambda$, $\rho_\mu$ curves are everywhere tangent to the $\delta_\lambda$, $\delta_\mu$ curves, the $\rho_\lambda$, $\rho_\mu$ curves also inherit an orientation.

We are now ready to proceed with the construction of the initial curve $\gamma$ upstairs to show the existence of a unique "weak" integral surface $S$ emanating from $\gamma$, and to contrast $S$ with the graph of the classical shock solution to the Riemann problem.

Let us assume that the initial data $\mathbf{u}_-$ and $\mathbf{u}_+$ for the Riemann problem (7b) is sufficiently close so that $\mathbf{u}_-$ and $\mathbf{u}_+$ are related as in Figure 4.8. We prescribed the initial data by a mapping

$$s \mapsto \gamma(s) = (x(s), 0, \mathbf{u}_0(s)), \qquad s \in \mathbb{R}$$



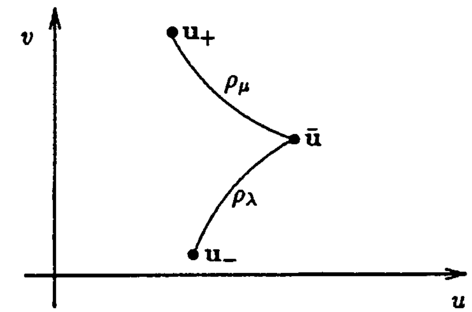


**Figure 4.8**

as follows: For an interval $[a, c] \subset \mathbb{R}$ to be determined below, we set

$$\gamma(s) = \begin{cases} (s - a, 0, \mathbf{u}_-) & \text{for } s \leq a \\ (s - c, 0, \mathbf{u}_+) & \text{for } s \geq c \end{cases}$$

and furthermore we set

$$x(s) = 0, \qquad \mathbf{u}_0(s) \in \rho_\lambda, \qquad \text{for } s \in [a, b],$$

and

$$x(s) = 0, \qquad \mathbf{u}_0(s) \in \rho_\mu, \qquad \text{for } s \in [b, c].$$

Pictorially we have Figure 4.9, where the $uv$-plane lies over $x = 0$. The speed at which $\mathbf{u}_0(s)$ transverses $\rho_\lambda$ and $\rho_\mu$ is specified in the following. Before doing so, we notice that the condition of genuine non-linearity implies that $\lambda$ and $\mu$ are strictly monotone functions on $\rho_\lambda$ and $\rho_\mu$ respectively. In what follows, we shall asume that $\lambda$ and $\mu$ are increasing functions along $\rho_\lambda$ and $\rho_\mu$ (a similar construction as below would work for the other cases as well).

For $\xi$ in the interval with endpoints $\lambda(\mathbf{u}_-)$, $\lambda(\bar{\mathbf{u}})$ we solve the ODE

$$\frac{d}{d\xi}\mathbf{u}_\lambda(\xi) = \mathbf{e}_\lambda(\mathbf{u}_\lambda(\xi)) \qquad \text{with} \qquad \mathbf{u}_\lambda(\lambda(\mathbf{u}_-)) = \mathbf{u}_-, \ \mathbf{u}_\lambda(\lambda(\bar{\mathbf{u}})) = \bar{\mathbf{u}}.$$

For $\eta$ in the interval with endpoints $\mu(\bar{\mathbf{u}})$, $\mu(\mathbf{u}_+)$ we solve the ODE

$$\frac{d}{d\eta}\mathbf{u}_\mu(\eta) = \mathbf{e}_\mu(\mathbf{u}_\mu(\eta)) \qquad \text{with} \qquad \mathbf{u}_\mu(\mu(\mathbf{u}_+)) = \mathbf{u}_+, \ \mathbf{u}_\mu(\mu(\bar{\mathbf{u}})) = \bar{\mathbf{u}}.$$

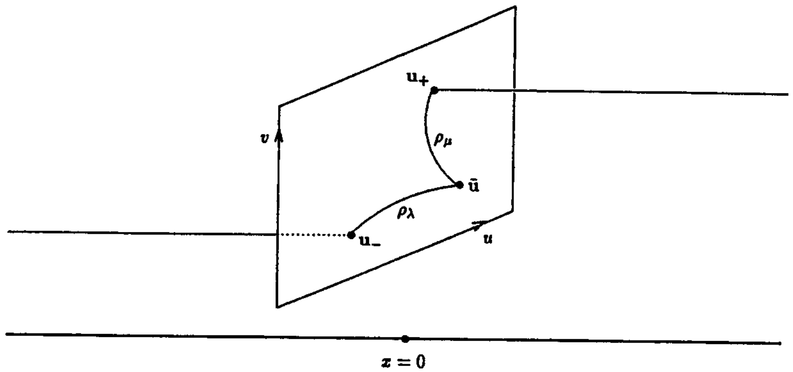


Figure 4.9

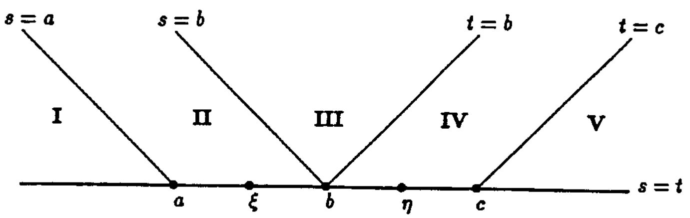


Figure 4.10

We now proceed to construct a mapping of $\mathbb{R}^2_+ = \{(s,\,t) : s-t>0\}$ into $M$ which will be the unique piecewise $C^1$ integral surface $S$ of the exterior differential system associated to (3) consisting only of TYPE 0 and TYPE 1 solutions. We begin by dividing $\mathbb{R}^2_+$ into 5 regions as in Figure 4.10. The mapping for regions **I** and **V** are given respectively by

$$(s,\,t)\mapsto\big(\lambda(\mathbf{u}_-)t+s-a,\,t,\,\mathbf{u}_-\big),$$



and

$$(s,\,t)\mapsto\big(\mu(\mathbf{u}_+)t+s-c,\,t,\,\mathbf{u}_+\big).$$

Thus the solution in each of these regions is of TYPE 0.

We now consider region **II**. As mentioned above, we are assuming that $\lambda(\mathbf{u})$ is increasing along $\rho_\lambda$ as we pass from $\mathbf{u}_-$ to $\bar{\mathbf{u}}$. Now we set $a=\lambda(\mathbf{u}_-)$ and $b=\lambda(\bar{\mathbf{u}})$. Along the lines $s=\xi$, the mapping is given by

$$(s,\,t)\mapsto\big(\lambda(\mathbf{u}_\lambda(s))t,\,t,\,\mathbf{u}_\lambda(s)\big)$$

i.e., the lines $s=\xi$ map to the $xy$-plane to the fan of lines with $x/y$-slopes increasing from $\lambda(\mathbf{u}_-)$ to $\lambda(\bar{\mathbf{u}})$. Symmetrically, region **IV** is mapped into $M$ using the lines $t=\eta$ in place of $s=\xi$, and where the fan in the $xy$-plane consists of lines with slopes running between $\mu(\bar{\mathbf{u}})$ and $\mu(\mathbf{u}_+)$. The solution in each of these region is of TYPE 1.

Finally, the boundary of region **III** is mapped into $M$ as indicated in Figure 4.12. Now, let $T$ be the unique linear mapping of the $st$-plane to the $xy$-plane taking $\partial_t$ to $\frac{1}{\sqrt{2}}(-\partial_x+\partial_y)$ and $\partial_s$ to $\frac{1}{\sqrt{2}}(\partial_x+\partial_y)$. In region **III**, we set

$$(s,\,t)\mapsto\big(T(s,\,t),\bar{\mathbf{u}}\big).$$

Clearly, the solution in this region is of TYPE 0.

It is clear that $S$ is the image of a continuous, piecewise $C^1$ mapping. It is also clear that it is a solution surface, in the usual sense, to the exterior differential system $\Phi=\Psi=0$ in each region. Somewhat less clear is that accross the

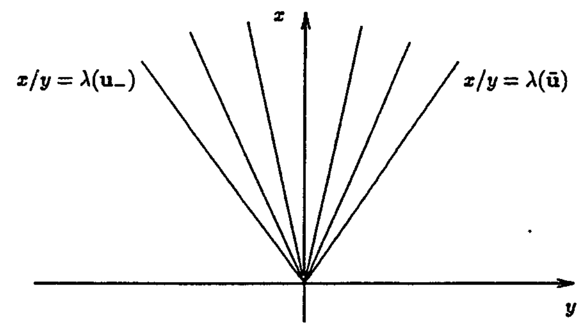


Figure 4.11

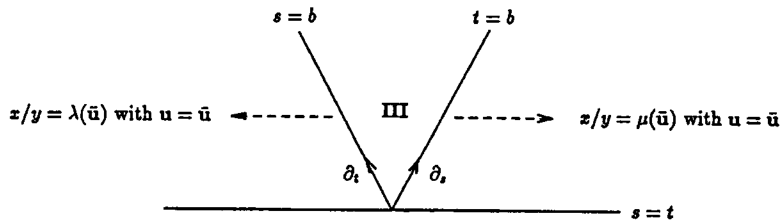


**Figure 4.12**

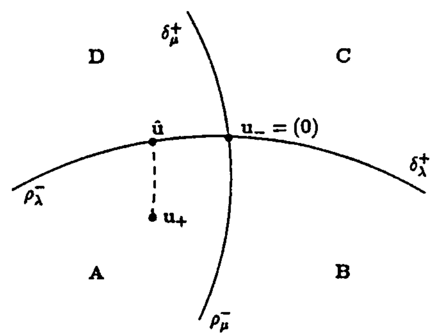


**Figure 4.13**

boundaries it is a weak solution to $d\varphi = d\psi = 0$. But by our construction the calculation here is essentially the same as in the scalar case.

It remains to compare $S$ with the graph of the classical shock solution $\mathbf{u}(x, y)$. Recall that in a neighborhood of the origin in the $uv$-plane we have Figure 4.13. These curves are of class $C^2$ but not generally of class $C^3$ at the origin. Without loss of generality we set $\mathbf{u}_- = (\mathbf{0})$, and for $\mathbf{u}_+$ lying in each of regions **A**, **B**, **C**, **D**, the solution $\mathbf{u}(x, y)$ consists respectively of a 1-shock followed by a 2-shock, a 1-rarefaction wave followed by a 2-shock, a 1-rarefaction wave followed by a 2-rarefaction wave, and a 1-shock followed by a 2-rarefaction wave. For example, if $\mathbf{u}_+$ lies in region **A**, then there is a unique $\hat{\mathbf{u}}$ as depicted above, and the picture



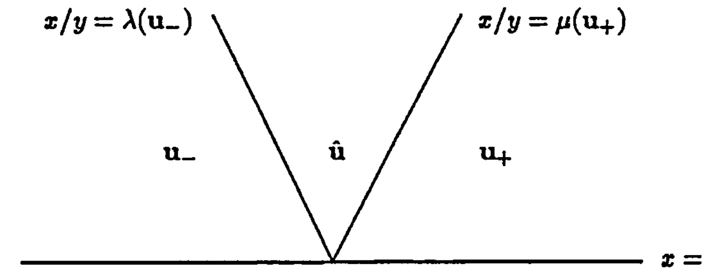


**Figure 4.14**

of the shock solution $\mathbf{u}(x, y)$ is shown in Figure 4.14.

In general $\rho_\lambda^-$ will not coincide with $\delta_\lambda^-$, in which case $\hat{\mathbf{u}} \neq \bar{\mathbf{u}}$ and so the graph of $\mathbf{u}(x, y)$ will in general not lie on the geometric surface $S$. However, as we have seen, an exception occurs in case (3) is a special hyperbolic system of conservation laws. For such systems we have the following result:

*For the Riemann problem associated to a special hyperbolic system of conservation laws, the shock solution arises by taking a cross-section of the geometric solution.*

For smooth initial data with suitable bounds on the oscillation there is the Glimm scheme *[16]* leading to a shock solution. Since this scheme is based on iteratively solving Riemann problems one may suspect that here also the shock solution $\mathbf{u}(x, y)$ arise by taking a cross-section of the geometric solution. In this case the singularities of $\mathbf{u}(x, y)$ would reflect the singularities of the map. In particular, $\mathbf{u}(x, y)$ would be piecewise smooth if the initial data is piecewise smooth.

In summary, for Riemann problems we have the following results:

(i) *For special hyperbolic systems of conservation laws, the geometric solution captures the shock solution;*

(ii) *For non-special systems, we must "cut" the geometric solution along suitable break curves and glue in other pieces of integral surfaces to obtain a surface upstairs that captures the shock solution.*

One may reasonably conjecture that both of these statements remains true for arbitrary smooth compactly supported initial data.

## Appendix: The Equivalence Problem

In the preceeding sections we have made extensive use of the geometry associated to a number of differential equations to deduce information about the behavior of their solutions. In this appendix we shall explain how to associate a geometry to each of these equations. The main tool is É. Cartan's theory of $G$-structures—commonly known in the literature as the Method of Equivalence. Aside from Cartan's own exposition *[8]*, there are several other sources for this material, notably Chern *[11]*, Gardner *[15]*, and Bryant-Griffiths *[3]*, which the reader may consult for more examples and information.

**A.1** *The equivalence problem for second-order ODE.* We seek to attach an intrinsic geometry to a second-order ordinary differential equation of the form

$$\frac{d^2y}{dx^2} = f\left(x,\, y,\, \frac{dy}{dx}\right). \tag{1}$$

Thinking of $x$ and $y$ as local coordinates on a surface $S$, the solutions of equation (1) form a 2-parameter family $\Sigma$ of curves on $S$ in the domain of the coordinates $x$ and $y$.

Conversely, suppose that we are given a surface $S$ together with a 2-parameter family $\Sigma$ of curves on $S$ with the property that, through each point $p$ of $S$, the curves in $\Sigma$ which pass through $p$ form a smooth curve $\Sigma_p \subset \Sigma$ smoothly parametrized by the assignment $C \mapsto T_pC \in \mathbb{P}(T_pS)$. Such a family $\Sigma$ is called a *path geometry* on $S$. Two path geometries, $\Sigma_1$ on $S_1$ and $\Sigma_2$ on $S_2$, are to be regarded as equivalent if there exists a diffeomorphism $\phi : S_1 \to S_2$ which carries the curves of $\Sigma_1$ onto the curves of $\Sigma_2$.

In any local coordinate system $(x, y)\colon U \to \mathbb{R}^2$ on an open set $U \subset S$, the curves of a path geometry $\Sigma$ are the solutions of a differential equation of the form (1). We say that two such equations are equivalent if they induce equivalent path geometries on their domains of definition.

Given a path geometry $\Sigma$ on $S$, we can define an incidence correspondence

$$M = \{(p, C) \in S \times \Sigma \,|\, p \in C\}.$$

By hypothesis, the assignment $(p, C) \mapsto T_pC$ defined for all $(p, C) \in M$ induces a smooth embedding of $M$ as a subset of $\mathbb{P}(TS)$, the projectivized tangent bundle of $S$. For all intents and purposes, we may therefore regard $M$ as an open subset of $\mathbb{P}(TS)$, which we do from now on.

As we remarked in §1, this picture is symmetric with respect to interchanging the roles of $S$ and $\Sigma$, since, for each fixed $p$, the set of curves in $\Sigma$ which contain $p$ is itself a curve $\Sigma_p$ in $\Sigma$. Thus, $S$ defines a path geometry on $\Sigma$ which is called the *dual* path geometry, with a corresponding local representation as an ordinary differential equation called the *dual equation* to (1).

Now, for a fixed $C \in \Sigma$, as $p$ varies on $C$, the assignment $p \mapsto T_C\Sigma_p$ then immerses $C$ *as a curve* into the 1-dimensional projective space $\mathbb{P}(T_C\Sigma)$. Thus, every curve in a path geometry inherits a canonical projective structure and hence



a parameter well-defined up to a linear fractional change of parameter. Already, one can see that there is quite a bit of local and global geometry attached to a second order ODE. One could say, for example, that an equation of the form (1) (or more generally, a path geometry) is *projectively complete* if the canonical projective structure on its curves is projectively complete. (Note that projective completeness of a projective structure on a curve does not imply that the curve is closed.)

As we discussed in §1, this leads to the double fibration

$$\begin{array}{ccc} & M & \\ {\scriptstyle\lambda}\swarrow & & \searrow{\scriptstyle\rho} \\ S & & \Sigma \end{array}$$

where $\lambda(p, C) = p$ and $\rho(p, C) = C$ and each of $\lambda$ and $\rho$ is a smooth submersion. It follows that $M$ can be covered by open sets $U$ on which there exist coframings $(\underline{\alpha},\, \underline{\theta},\, \underline{\omega})$ so that on $U$, we have

$$\ker \lambda_* = \{\underline{\alpha},\, \underline{\theta}\}^\perp \qquad \text{and} \qquad \ker \rho_* = \{\underline{\theta},\, \underline{\omega}\}^\perp. \tag{2}$$

We say that a coframing $(\alpha, \theta, \omega)$ is *0-adapted* to the double fibration if the relations in (2) are satisfied. Note that the 1-form $\underline{\theta}$ is well-defined up to a scalar multiple and that, as explained in §1, it is a contact form, i.e., $\underline{\theta}\wedge d\underline{\theta}$ is nowhere vanishing.

Now, any other 0-adapted coframing in $U$ is seen to be of the form

$$\begin{pmatrix} \alpha \\ \theta \\ \omega \end{pmatrix} = \begin{pmatrix} a_1 & b_1 & 0 \\ 0 & a_3 & 0 \\ 0 & b_2 & a_2 \end{pmatrix}^{-1} \begin{pmatrix} \underline{\alpha} \\ \underline{\theta} \\ \underline{\omega} \end{pmatrix} \tag{3}$$

where the "transition matrix" takes values in the obvious 5-dimensional subgroup $G_0$ of $\mathrm{GL}(3, \mathbb{R})$. Thus the local 0-adapted coframings of the double fibration are the local sections of a principal $G_0$-subbundle $B_{G_0} \to M$ of the bundle of general coframes of $M$. In other words, $B_{G_0}$ is a $G_0$-structure on $M$.

It is important to note that the induced $G_0$-structure on $M \subset \mathbb{P}(TS)$ can be computed *without* knowing the solutions of (1) in advance. For example, using coordinates $x$, $y$, and $p$ $(= dy/dx)$ on $\mathbb{R}^3 \subset \mathbb{P}(T\mathbb{R}^2)$, we can use

$$\begin{pmatrix} \underline{\alpha} \\ \underline{\theta} \\ \underline{\omega} \end{pmatrix} = \begin{pmatrix} dx \\ dy - p\,dx \\ dp - f(x, y, p)\,dx \end{pmatrix}$$

as a 0-adapted coframing on $M$, even though we may not know explicit functions $\sigma$ and $\tau$ on $M$ whose differentials have the same span as $\{\underline{\theta}, \underline{\omega}\}$, i.e., a pair of independent first integrals of (1). Indeed, one of the reasons for studying the $G_0$-structure $B_{G_0}$ in the first place is to find ways to compute such functions $\sigma$ and $\tau$.

Following Cartan, we now apply the equivalence method to this $G_0$-structure to compute its invariants. Differentiating (3), and collecting terms as much as possible, we see that there exist 1-forms $\varphi_1, \varphi_2, \varphi_3, \mu_1$, and $\mu_2$ and a function $t$ on $B_{G_0}$ so that

$$d\begin{pmatrix}\alpha\\ \theta\\ \omega\end{pmatrix} = -\begin{pmatrix}\varphi_1 & \mu_1 & 0\\ 0 & \varphi_3 & 0\\ 0 & \mu_2 & \varphi_2\end{pmatrix}\wedge\begin{pmatrix}\alpha\\ \theta\\ \omega\end{pmatrix} + \begin{pmatrix}0\\ t\,\alpha\wedge\omega\\ 0\end{pmatrix}$$

In the terminology of the equivalence method, the $\mathfrak{g}_0$-valued 1-form

$$\begin{pmatrix}\varphi_1 & \mu_1 & 0\\ 0 & \varphi_3 & 0\\ 0 & \mu_2 & \varphi_2\end{pmatrix}$$

is called the *pseudo-connection matrix* and the term $t\,\alpha\wedge\omega$ constitutes the *torsion* of the structure. The pseudo-connection 1-forms are not uniquely determined by the structure equation and we will return to this point below.

Now, the contact condition $\theta\wedge d\theta \neq 0$ implies that $t \neq 0$, which suggests that we may restrict to coframings which satisfy $t = 1$. Thus, we say that a 0-adapted local coframing $(\underline{\alpha}, \underline{\theta}, \underline{\omega})$ is *1-adapted* if, in addition to being 0-adapted, it satisfies $d\underline{\theta} \equiv \underline{\alpha}\wedge\underline{\omega} \mod \theta$. The local coframings which are 1-adapted define a $G_1$-structure on $M$ which is a sub-structure of the given $G_0$-structure. Its structure group is the 4-dimensional subgroup

$$G_1 = \left\{ \begin{pmatrix}a_1 & b_1 & 0\\ 0 & a_1 a_2 & 0\\ 0 & b_2 & a_2\end{pmatrix} \;\middle|\; a_1, a_2 \neq 0 \right\} \subset G_0 .$$

The structure equation of $B_{G_1}$ now assumes the form

$$d\begin{pmatrix}\alpha\\ \theta\\ \omega\end{pmatrix} = -\begin{pmatrix}\underline{\varphi}_1 & \underline{\mu}_1 & 0\\ 0 & \underline{\varphi}_1 + \underline{\varphi}_2 & 0\\ 0 & \underline{\mu}_2 & \underline{\varphi}_2\end{pmatrix}\wedge\begin{pmatrix}\alpha\\ \theta\\ \omega\end{pmatrix} + \begin{pmatrix}0\\ \alpha\wedge\omega\\ 0\end{pmatrix}$$

for some 1-forms $\underline{\varphi}_1, \underline{\varphi}_2, \underline{\mu}_1$, and $\underline{\mu}_2$ which are, however, not uniquely determined by this equation.

It is easy to see that if $\varphi_1, \varphi_2, \mu_1$, and $\mu_2$ are any 1-forms on $B_{G_1}$ which satisfy

$$d\begin{pmatrix}\alpha\\ \theta\\ \omega\end{pmatrix} = -\begin{pmatrix}\varphi_1 & \mu_1 & 0\\ 0 & \varphi_1 + \varphi_2 & 0\\ 0 & \mu_2 & \varphi_2\end{pmatrix}\wedge\begin{pmatrix}\alpha\\ \theta\\ \omega\end{pmatrix} + \begin{pmatrix}0\\ \alpha\wedge\omega\\ 0\end{pmatrix}, \qquad (4)$$

then there exist functions $p_1, p_2, q_1$, and $q_2$ so that

$$\begin{aligned}\varphi_1 &= \underline{\varphi}_1 + p_1\,\theta,\\ \varphi_2 &= \underline{\varphi}_2 + p_2\,\theta,\end{aligned} \qquad \text{and} \qquad \begin{aligned}\mu_1 &= \underline{\mu}_1 + q_1\,\theta + p_1\,\alpha,\\ \mu_2 &= \underline{\mu}_2 + q_2\,\theta + p_2\,\omega,\end{aligned}$$



and conversely that, for any such functions on $B_{G_1}$ the 1-forms defined by the above equations satisfy (4). For such a set of 1-forms satisfying (4), we will say that the collection of 1-forms $(\alpha, \theta, \omega; \varphi_1, \varphi_2, \mu_1, \mu_2)$ define a $(1,1)$-adapted coframing of the 7-manifold $B_{G_1}$. It is clear that the $(1,1)$-adapted coframings are the sections of a $G_1^{(1)}$-structure $B_{G_1^{(1)}}$ on $B_{G_1}$ where $G_1^{(1)} \subset \mathrm{GL}(7,\mathbb{R})$ is a certain 4-dimensional Lie group. Note that $B_{G_1^{(1)}}$ is a sub-bundle of the second-order coframe bundle of $M$ and defines what is often called a second-order $G$-structure on $M$.

We now want to apply the equivalence method to the $G_1^{(1)}$-structure $B_{G_1^{(1)}}$. The structure equation of $B_{G_1^{(1)}}$ consists of two parts—the first is equation (4). The second will consist of the formulae for $d\varphi_1, d\varphi_2, d\mu_1$, and $d\mu_2$. We can get information on these by differentiating (4), yielding

$$\begin{aligned}(d\varphi_1 + \mu_1\wedge\omega)\wedge\alpha + (d\mu_1 - \varphi_2\wedge\mu_1)\wedge\theta &= 0\\ (d\varphi_2 - \mu_2\wedge\alpha)\wedge\omega + (d\mu_2 - \varphi_1\wedge\mu_2)\wedge\theta &= 0\end{aligned}$$

and

$$d(\varphi_1 + \varphi_2) - \mu_1\wedge\omega + \mu_2\wedge\alpha \equiv 0 \mod \theta.$$

From these relations, it follows that there are 1-forms $\pi_1, \pi_2$ and $\psi_1, \psi_2$ and a function $t_0$ on $B_{G_1^{(1)}}$ such that

$$\begin{aligned}d\varphi_1 &= -\pi_1\wedge\theta - \mu_1\wedge\omega - 2\mu_2\wedge\alpha + t_0\,\alpha\wedge\omega\\ d\varphi_2 &= -\pi_2\wedge\theta + \mu_2\wedge\alpha + 2\mu_1\wedge\omega + t_0\,\omega\wedge\alpha\end{aligned}$$

and

$$\begin{aligned}d\mu_1 &= -\psi_1\wedge\theta - \pi_1\wedge\alpha + \varphi_2\wedge\mu_1\\ d\mu_2 &= -\psi_2\wedge\theta - \pi_2\wedge\omega + \varphi_1\wedge\mu_2.\end{aligned}$$

Note that $(\pi_1, \pi_2, \psi_1, \psi_2)$ are the pseudo-connection 1-forms of the $G_1^{(1)}$-structure.

We now compute how the function $t_0$ (which is the only non-constant component of the torsion) varies on the fibers of $B_{G_1^{(1)}} \to B_{G_1}$. Differentiating the above structure equations yields

$$dt_0 \equiv t_0\,(\varphi_1 + \varphi_2) + 2(\pi_1 - \pi_2) \mod \alpha, \theta, \omega.$$

Thus, the equation $t_0 = 0$ defines a submanifold $B_{G_2} \subset B_{G_1^{(1)}}$ which is a $G_2$-structure where $G_2$ is a 3-dimensional subgroup of $G_1^{(1)}$. Because of the formula for $dt_0$, we see that, on $B_{G_2}$, there must exist a 1-form $\sigma$ and functions $s_1$ and $s_2$ so that

$$\left.\begin{aligned}\pi_1 &\equiv -\sigma + s_1\,\alpha + s_2\,\omega\\ \pi_2 &\equiv -\sigma - s_1\,\alpha - s_2\,\omega\end{aligned}\right\} \mod \theta.$$

Thus, by modifying the forms $\psi_1$ and $\psi_2$, the structure equations can be reduced to the form

$$d\varphi_1 = \sigma\wedge\theta - \mu_1\wedge\omega - 2\mu_2\wedge\alpha - (s_1\,\alpha + s_2\,\omega)\wedge\theta$$
$$d\varphi_2 = \sigma\wedge\theta + \mu_2\wedge\alpha + 2\mu_1\wedge\omega + (s_1\,\alpha + s_2\,\omega)\wedge\theta$$

and

$$d\mu_1 = -\psi_1\wedge\theta + \sigma\wedge\alpha + \varphi_2\wedge\mu_1 + s_2\,\alpha\wedge\omega$$
$$d\mu_2 = -\psi_2\wedge\theta + \sigma\wedge\omega + \varphi_1\wedge\mu_2 + s_1\,\alpha\wedge\omega.$$

The functions $s_1$ and $s_2$ constitute the non-constant terms in the torsion of this $G_2$-structure.

To calculate how $s_1$ and $s_2$ vary on the fibers of $B_{G_2} \to B_{G_1}$, we differentiate the above structure equations, yielding

$$\left.\begin{aligned} ds_1 &\equiv s_1\,(2\varphi_1 + \varphi_2) - 3\psi_2 \\ ds_2 &\equiv s_2\,(2\varphi_2 + \varphi_1) - 3\psi_1 \end{aligned}\right\} \mod \alpha,\, \theta,\, \omega.$$

Thus, the equations $s_1 = s_2 = 0$ define a submanifold $B_{G_3} \subset B_{G_2}$ which is a $G_3$-structure on $B_{G_1}$ where the structure group $G_3$ is now 1-dimensional. On $B_{G_3}$, we can now differentiate the identity

$$\tfrac{1}{3}d(\varphi_2 - \varphi_1) = \mu_1\wedge\omega + \mu_2\wedge\alpha$$

and use the structure equations found so far to conclude that

$$\psi_1\wedge\omega + \psi_2\wedge\alpha \equiv 0 \mod \theta,$$

which implies that there exist functions $K_1$, $K_2$ and $S$ so that

$$\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} \equiv \begin{pmatrix}K_2 & S\\ S & K_1\end{pmatrix}\begin{pmatrix}\omega\\ \alpha\end{pmatrix} \mod \theta.$$

It follows that by adding appropriate multiple of $\theta$ to each of $\psi_1$, $\psi_2$, and $\sigma$, the structure equations reduce to the form

$$d\varphi_1 = \sigma\wedge\theta - \mu_1\wedge\omega - 2\mu_2\wedge\alpha$$
$$d\varphi_2 = \sigma\wedge\theta + \mu_2\wedge\alpha + 2\mu_1\wedge\omega$$
$$d\mu_1 = \sigma\wedge\alpha + \varphi_2\wedge\mu_1 + K_2\,\theta\wedge\omega$$
$$d\mu_2 = \sigma\wedge\omega + \varphi_1\wedge\mu_2 + K_1\,\theta\wedge\alpha.$$

Now, the remaining pseudo-connection 1-form $\sigma$ is uniquely defined by the above structure equations. Differentiating these equations yields

$$d\sigma = (\varphi_1 + \varphi_2)\wedge\sigma - \mu_1\wedge\mu_2 + \theta\wedge(L_1\,\alpha + L_2\,\omega)$$

for some functions $L_1$ and $L_2$ on $B_{G_3}$.



At this stage we have determined a canonical parallelism on $B_{G_3}$ with structure equations given by

$$d\begin{pmatrix}\alpha\\ \theta\\ \omega\end{pmatrix} = -\begin{pmatrix}\varphi_1 & \mu_1 & 0\\ 0 & \varphi_1+\varphi_2 & 0\\ 0 & \mu_2 & \varphi_2\end{pmatrix}\wedge\begin{pmatrix}\alpha\\ \theta\\ \omega\end{pmatrix} + \begin{pmatrix}0\\ \alpha\wedge\omega\\ 0\end{pmatrix}$$

where

$$d\varphi_1 = \sigma\wedge\theta - \mu_1\wedge\omega - 2\mu_2\wedge\alpha$$
$$d\varphi_2 = \sigma\wedge\theta + \mu_2\wedge\alpha + 2\mu_1\wedge\omega$$
$$d\mu_1 = \sigma\wedge\alpha + \varphi_2\wedge\mu_1 + K_2\,\theta\wedge\omega$$
$$d\mu_2 = \sigma\wedge\omega + \varphi_1\wedge\mu_2 + K_1\,\theta\wedge\alpha$$

and

$$d\sigma = (\varphi_1 + \varphi_2)\wedge\sigma - \mu_1\wedge\mu_2 + \theta\wedge(L_1\,\alpha + L_2\,\omega).$$

Moreover, as the method of equivalence shows, any self-equivalence of $S$ with itself preserving the paths geometry $\Sigma$ lifts to a unique diffeomorphism of $B_{G_3}$ with itself which preserves the eight 1-forms in this coframe. In particular, the space of automorphisms of the path geometry $\Sigma$ on $S$ is a Lie group of dimension at most 8. In fact, the dimension of this group can be 8 only if the functions $K_i$ and $L_i$ are constants.

These structure equations can be collected into a more coherent form as follows: Define an $\mathfrak{sl}(3,\mathbb{R})$-valued matrix $\phi$ by

$$\phi = \begin{pmatrix} -\tfrac{1}{3}(2\varphi_1+\varphi_2) & -\mu_2 & \sigma\\ \alpha & \tfrac{1}{3}(\varphi_1-\varphi_2) & \mu_1\\ \theta & \omega & \tfrac{1}{3}(\varphi_1+2\varphi_2)\end{pmatrix}.$$

Then the structure equations above take the form

$$d\phi + \phi\wedge\phi = \begin{pmatrix}0 & K_1\alpha\wedge\theta & \theta\wedge(L_1\alpha + L_2\omega)\\ 0 & 0 & K_2\theta\wedge\omega\\ 0 & 0 & 0\end{pmatrix}.$$

Setting $\Phi = d\phi + \phi\wedge\phi$, the Bianchi identity $d\Phi = \Phi\wedge\phi - \phi\wedge\Phi$ then gives rise to the equations

$$dK_1 \equiv (3\varphi_1 + \varphi_2)K_1 - L_1\,\omega \mod \theta, \alpha$$
$$dK_2 \equiv (3\varphi_2 + \varphi_1)K_2 - L_2\,\alpha \mod \theta, \omega$$
$$dL_1 \equiv (3\varphi_1 + 2\varphi_2)L_1 - K_1\mu_1 + J\omega \mod \theta, \alpha$$
$$dL_2 \equiv (3\varphi_2 + 2\varphi_1)L_2 + K_2\mu_2 + J\alpha \mod \theta, \omega$$

where $J$ is some function on $B_{G_3}$.

From these structure equations, it follows, for example, that the expressions

$$\mathcal{K}_1 = (K_1\,\alpha - L_1\theta) \otimes (\alpha\wedge\theta)$$
$$\mathcal{K}_2 = (K_2\,\omega + L_2\theta) \otimes (\theta\wedge\omega)$$

are well-defined sections of $T^* \otimes \Lambda^2 T^*$ on $M$. Thus, these define tensors which are the simplest invariants of a path geometry or second order ODE up to point transformations. In the classical terminology, the functions $K_1$ and $K_2$ are what is known as *relative invariants*. Furthermore, a little computation shows that $J$ is a relative invariant and that the cubic form $\mathcal{J} = J\,\theta^3$ is well-defined on $M$ and hence is another invariant of the path geometry.

Note however, that these invariants are defined on $M$, and not on either $S$ or $\Sigma$. As an example of an invariant function on $M$, the ratio $J^4/(K_1K_2)^3$ is seen to be constant on the fibers of $B_{G_3} \to M$ and hence is a well-defined function on $M$.

Let us consider the meaning of the vanishing of the first two invariants, i.e., the case where $K_1$ and $K_2$ vanish identically. (Note that, by the Bianchi identities, this is the only way that $K_1$ and $K_2$ can be constants anyway.) By the Bianchi identities, this forces $L_1$ and $L_2$ to vanish identically as well, so that $\Omega$ vanishes identically. Then, since $d\phi = -\phi\wedge\phi$, it follows that there exists a local diffeomorphism $B_{G_3} \to \mathrm{SL}(3,\mathbb{R})$ which identifies $\phi$ with the canonical left-invariant 1-form on $\mathrm{SL}(3,\mathbb{R})$. This identification induces an identification of $S$ (which can be thought of as the leaves in $B_{G_3}$ of the integrable distribution $\alpha = \theta = 0$) with the left cosets in $\mathrm{SL}(3,\mathbb{R})$ of the subgroup $P$ whose Lie algebra is of the form

$$\begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}.$$

In other words, $S$ is is locally identified with $\mathrm{SL}(3,\mathbb{R})/P = \mathbb{RP}^2$. Moreover, we also get an identification of $\Sigma$ (which can be thought of as the leaves in $B_{G_3}$ of the integrable distribution $\omega = \theta = 0$) with the left cosets in $\mathrm{SL}(3,\mathbb{R})$ of the subgroup $\Pi$ whose Lie algebra is of the form

$$\begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & * \end{pmatrix}.$$

Of course, the homogeneous space $\mathrm{SL}(3,\mathbb{R})/\Pi$ is also diffeomorphic to $\mathbb{RP}^2$, but its points can be regarded as the lines in the first $\mathbb{RP}^2$. Thus, the path geometry in this case is just the classical path geometry of straight lines in the plane. Moreover, the intrinsic geometric structure $B_{G_3}$ is identified with the group of symmetries of this path geometry, i.e., the projective transformations of $\mathbb{RP}^2$. Note that a consequence of our calculations is that this is the only path geometry (up to local equivalence) for which the group of symmetries is actually of dimension 8.

Now let us consider the meaning of the vanishing of each of $K_1$ or $K_2$ separately. In terms of the ODE (1), Cartan showed that $K_2$ vanishes if and only if the function $f$ satisfies $\partial^4 f/\partial p^4 = 0$, i.e., if and only if $f$ is a cubic polynomial in $p$. On the other hand $K_1$ vanishes if and only if $f$ satisfies equation (5) of §1.



Now, in the first case, when $K_2$ vanishes, it follows that the ODE (1) is the local equation for the geodesics of a projective connection on $S$ in the usual sense of differential geometry. Since a projective connection on a surface has a curvature tensor which is well-defined on that surface, it should not be surprising that when $K_2$ vanishes we can create expressions on $M$ which are well-defined on $S$. In fact, the vanishing of $K_2$ implies directly that the tensor $\mathcal{K}_1$ (which is clearly semi-basic for the map $\lambda : M \to S$) is actually well-defined on $S$. Computing higher derivatives of $K_1$ as functions on $B_{G_3}$ and taking the right combinations yields functions which are actually well-defined on $S$. Thus, they are first integrals of the dual ODE on $\Sigma$.

Perhaps more interesting is the other side. When $K_1$ vanishes, it follows that the dual equation to (1) (which we may not know explicitly) is actually the local equation of the geodesics of a projective connection. Again, the same arguments show that by taking the right combinations of the (iterated) derivatives of $K_2$ (which we can compute from $f$ *without having to solve any ODE*), we can arrive at functions on $B_{G_3}$ which are constant on the fibers of the map $B_{G_3} \to \Sigma$. Hence, these will be functions on $M$ which are first integrals of the equation (1). Thus, equations (1) satisfying the curvature condition $K_1 = 0$ can usually be integrated by differentiation alone! (The reason we have to say "usually" is that there are certain special cases where this procedure won't lead to a pair of independent first integrals. However, these cases are very rare and can only happens when the path geometry has a non-trivial symmetry group, so that other methods, due to Lie, can be applied.)

Although Cartan did not pursue this, there are higher order conditions like $K_1$ vanishing which will guarantee that the equations which satisfy them can be integrated by an algorithm (see *[5]*). Unfortunately, we do not have time to pursue this topic here.

**A.2** *The equivalence problem for scalar conservation laws.* In §2 we have shown that a scalar conservation law determines the data $(M;\,\Phi,\,\Omega)$ consisting of a 3-manifold $M$ endowed with two linearly independent 2-forms $\Phi$ (which is closed) and $\Omega$ (which is nowhere vanishing and well-defined up to a multiple). In addition the non-linearity of the conservation law is expressed by the condition that the common linear factor of $\Phi$ and $\Omega$ defines a contact structure on $M$. We shall study the geometry associated with the above $(M;\,\Phi,\,\Omega)$ structure under diffeomorphisms of $M$ which preserve $\Phi$ and preserve $\Omega$ up to a non-vanishing factor.

Let $(\underline{\alpha},\,\underline{\theta},\,\underline{\omega})$ be a local coframing of $M$ such that

$$\Phi = \underline{\theta}\wedge\underline{\omega} \qquad \text{and} \qquad \Omega = \underline{\alpha}\wedge\underline{\theta} \tag{1}$$

with $\underline{\theta}$ defining a contact structure on $M$, i.e.,

$$\underline{\theta}\wedge d\underline{\theta} \neq 0. \tag{2}$$

We shall say that a coframing $(\alpha, \theta, \omega)$ is 0-*adapted* to the $(M; \Phi, \Omega)$-structure if the relations in (1) and (2) are satisfied with $\Phi$ preserved and $\Omega$ preserved up to a non-vanishing multiple. It is easy to see that any two 0-adapted coframes are related as follows

$$\begin{pmatrix} \bar\alpha \\ \bar\theta \\ \bar\omega \end{pmatrix} = \begin{pmatrix} * & * & 0 \\ 0 & a^{-1} & 0 \\ 0 & * & a \end{pmatrix}^{-1} \begin{pmatrix} \alpha \\ \theta \\ \omega \end{pmatrix} \tag{3}$$

where the "transition matrices" take values in a 4-dimensional subgroup $G_0$ of $GL(3, \mathbb{R})$. Thus these local 0-adapted coframings define a $G_0$-structure $B_{G_0}$ on $M$.

We now apply the equivalence method to compute the invariants of the above $G_0$-structure. Differentiating (3) and using the condition $d\Phi = 0$, we see that on $B_{G_0}$ we have the following structure equation

$$d\begin{pmatrix} \alpha \\ \theta \\ \omega \end{pmatrix} = -\begin{pmatrix} \varphi_1 & \mu_1 & 0 \\ 0 & -\varphi_2 & 0 \\ 0 & \mu_2 & \varphi_2 \end{pmatrix} \wedge \begin{pmatrix} \alpha \\ \theta \\ \omega \end{pmatrix} + \begin{pmatrix} 0 \\ t\,\alpha\wedge\omega \\ 0 \end{pmatrix} \tag{4}$$

where $t$ is a nowhere vanishing function on $B_{G_0}$. The $G_0$-action on $t$ can be deduced by differentiating (4) which yields the relation

$$dt \equiv t\,(\varphi_1 + 2\varphi_2) \mod \alpha, \theta, \omega.$$

This suggests that we may restrict to local coframings on $M$ which satisfy $t = 1$. We call such coframings 1-*adapted*. They are the sections of a $G_1$-structure $B_{G_1}$ on which the structure equation has the form

$$d\begin{pmatrix} \alpha \\ \theta \\ \omega \end{pmatrix} = -\begin{pmatrix} 2\varphi & \mu_1 & 0 \\ 0 & \varphi & 0 \\ 0 & \mu_2 & -\varphi \end{pmatrix} \wedge \begin{pmatrix} \alpha \\ \theta \\ \omega \end{pmatrix} + \begin{pmatrix} t_0\,\alpha\wedge\omega \\ \alpha\wedge\omega \\ 0 \end{pmatrix}$$

for a smooth function $t_0$ on $B_{G_1}$. The structure group $G_1$ is now 3-dimensional.

Again, we deduce the $G_1$-action on the function $t_0$ by differentiating the above equation which yields

$$dt_0 \equiv -t_0\,\varphi - 3\mu_1 \mod \alpha, \theta, \omega.$$

This suggests that we further restrict to the submanifold $B_{G_2} \subset B_{G_1}$ defined by $t_0 = 0$. On $B_{G_2}$ we have

$$d\begin{pmatrix} \alpha \\ \theta \\ \omega \end{pmatrix} = -\begin{pmatrix} 2\varphi & 0 & 0 \\ 0 & \varphi & 0 \\ 0 & \mu & -\varphi \end{pmatrix} \wedge \begin{pmatrix} \alpha \\ \theta \\ \omega \end{pmatrix} + \begin{pmatrix} K\,\theta\wedge\omega \\ \alpha\wedge\omega \\ 0 \end{pmatrix}$$

where $K$ is a smooth function on $B_{G_2}$. Note that the structure group $G_2 \subset G_3$ is now 2-dimensional.



At this stage we notice that the pseudo-connection form $\varphi$ is uniquely determined by the above structure equation. Thus the 1-form $\alpha$ is now well-defined up to scale. Furthermore, from

$$\alpha \wedge d\alpha = K\,\alpha\wedge\theta\wedge\omega$$

we see that the invariant $K$ vanishes identically if and only if $\alpha$ is integrable.

Now, the $G_2$-action on the function $K$ is expressed by the relation

$$dK \equiv -2K\,\varphi \mod \alpha, \theta, \omega.$$

The equivalence problem now branches into two cases according to whether $K$ vanishes or not.

- $K = 0$—In this case the structure equation reduces to

$$d\begin{pmatrix} \alpha \\ \theta \\ \omega \end{pmatrix} = -\begin{pmatrix} 2\varphi & 0 & 0 \\ 0 & \varphi & 0 \\ 0 & \mu & -\varphi \end{pmatrix} \wedge \begin{pmatrix} \alpha \\ \theta \\ \omega \end{pmatrix} + \begin{pmatrix} 0 \\ \alpha\wedge\omega \\ 0 \end{pmatrix}.$$

Since $\varphi$ is now canonically defined, we can differentiate the above equation to get

$$d\varphi = -\mu\wedge\alpha$$

which in turn (together with the above structure equation) uniquely determines the pseudo-connection form $\mu$. Further differentiation yields

$$d\mu = -2\mu\wedge\varphi + L\,\alpha\wedge\theta$$

for some function $L$ on $B_{G_2}$.

At this stage that the 1-forms $(\alpha, \theta, \omega; \varphi, \mu)$ determines a canonical parallelism on $B_{G_2}$ with structure equations given as above. In the case $L = 0$, it is easy to show that $B_{G_2}$ is a Lie group with $(\alpha, \theta, \omega; \varphi, \mu)$ giving a basis for the left-invariant 1-forms. On the other hand, if $L \neq 0$, then a further differentiation of the structure equations above gives

$$dL \equiv 5L\,\varphi \mod \alpha, \theta.$$

Thus we can further restrict to the submanifold $B_{G_3} \subset B_{G_2}$ defined by $L = 1$. On $B_{G_3}$ we can set

$$\varphi = P_0\,\alpha + \tfrac{1}{2}P\,\theta$$

and hence from the above structure equations we obtain

$$dP_0 \equiv -\mu + \tfrac{1}{2}P\,\omega \mod \alpha, \theta.$$

Now restricting to the submanifold defined by $P_0 = 0$, we obtain an $e$-structure on $M$, i.e., a canonical parallelism on $M$ with structure equations

$$d\alpha = P\,\alpha\wedge\theta, \qquad d\theta = \alpha\wedge\omega, \qquad d\omega = (Q\,\alpha - P\,\omega)\wedge\theta.$$

Notice that $P$ and $Q$ are well-defined functions on $M$. In what follows, they will be refered to as the curvatures of the associated conservation law.

• $K \neq 0$—In this case, assuming, say, that $K > 0$, we may further restrict to the submanifold $B_{G_3} \subset B_{G_2}$ defined by $K = 1$. On $B_{G_3}$ the structure equation has the form

$$d\begin{pmatrix}\alpha\\ \theta\\ \omega\end{pmatrix} = -\begin{pmatrix}0&0&0\\0&0&0\\0&\mu&0\end{pmatrix}\wedge\begin{pmatrix}\alpha\\ \theta\\ \omega\end{pmatrix} + \begin{pmatrix}-2\varphi\wedge\alpha+\theta\wedge\omega\\ -\ \varphi\wedge\theta\ +\alpha\wedge\omega\\ \varphi\wedge\omega\end{pmatrix}$$

where

$$\varphi \equiv 0 \quad \mathrm{mod}\,\alpha,\,\theta,\,\omega.$$

An easy computation shows that we can further reduce to an $e$-structure on $M$, i.e., we have a canonical parallelism on $M$ with structure equations given by

$$d\alpha = -2\varphi\wedge\alpha+\theta\wedge\omega, \qquad d\theta = -\varphi\wedge\theta+\alpha\wedge\omega, \qquad d\omega = -\mu\wedge\theta$$

where

$$\varphi \equiv 0 \quad \mathrm{mod}\,\theta,\,\omega \qquad \text{and} \qquad \mu \equiv 0 \quad \mathrm{mod}\,\alpha,\,\theta,\,\omega.$$

Expanding out these congruences and substituting them into the structure equations above produce a collection of well-defined functions on $M$ which have interpretations as curvature functions of the associated conservation law.

**A.3** *The equivalence problem for hyperbolic systems.* In the following we will show how to associate a geometry to a (non-degenerate) hyperbolic exterior differential system $(M;\,\Omega_1,\,\Omega_2)$.

Recall that a hyperbolic system is given by a transverse pair of decomposable 2-forms $\Omega_1, \Omega_2$ on a 4-manifold $M$. This implies that there are local coframings $(\omega^1,\,\omega^2,\,\omega^3,\,\omega^4)$ on $M$ such that

$$\Omega_1 = \omega^1\wedge\omega^2, \qquad \Omega_2 = \omega^3\wedge\omega^4. \tag{1}$$

Clearly such coframings are determined up to the group $\mathrm{GL}(2,\mathbb{R})\times\mathrm{GL}(2,\mathbb{R})$. Thus the Pfaffian systems $\Xi_1$ and $\Xi_2$ generated respectively by $\omega^1$, $\omega^2$ and $\omega^3$, $\omega^4$ are well-defined on $M$. Furthermore on any solution surface $S \subset M$ to the hyperbolic system $\Omega_1 = \Omega_2 = 0$, each of $\Xi_1$ and $\Xi_2$ will have rank exactly equal to one and will therefore induce two foliations on $S$ by curves. These are the classical *characteristic curves* and for this reason we shall call $\Xi_1$ and $\Xi_2$ the *characteristic systems*. The interaction of the geometry with the characteristic systems and their prolongations is probably the deepest aspect of our subject. In the following we shall only consider hyperbolic systems for which the behavior of the characteristic systems—such as the ranks of their derived flags—is symmetric.

For a non-degenerate hyperbolic system the first derived system of each characteristic system has rank 1. This suggests that we restrict to local coframings satisfying (1) and the conditions

$$d\omega^1 \equiv 0 \quad \mathrm{mod}\,\omega^1,\,\omega^2, \qquad d\omega^3 \equiv 0 \quad \mathrm{mod}\,\omega^3,\,\omega^4. \tag{2}$$



Thus, $\omega^1$ and $\omega^3$ span the first derived systems of the two characteristic systems.

Now, for any coframing satisfying (1) and (2), there must exist functions $A$ and $C$ so that

$$d\omega^2 \equiv A\,\omega^3\wedge\omega^4 \quad \mathrm{mod}\,\omega^1,\,\omega^2 \qquad d\omega^4 \equiv C\,\omega^1\wedge\omega^2 \quad \mathrm{mod}\,\omega^3,\,\omega^4. \tag{3}$$

By assumption, $\Xi_1$ and $\Xi_2$ are non-integrable. It follows from this and (2) that neither $A$ nor $C$ can vanish and thus we can further restrict to coframings which satisfy $A = C = 1$. We shall say that a coframing is 1-adapted to the hyperbolic system if it satisfies the conditions (1), (2), and (3) with $A = C = 1$.

If $\omega = (\omega^1,\,\omega^2,\,\omega^3,\,\omega^4)$ is a 1-adapted coframing on a domain $U \subset M$, then any other coframing on $U$, say $\bar\omega = (\bar\omega^1,\,\bar\omega^2,\,\bar\omega^3,\,\bar\omega^4)$ is seen to be 1-adapted if and only if there exist functions $a^2_1$, $a^4_3$, $a^2_2 \neq 0$, and $a^4_4 \neq 0$ on $U$ so that

$$\begin{pmatrix}\bar\omega^1\\ \bar\omega^2\\ \bar\omega^3\\ \bar\omega^4\end{pmatrix} = \begin{pmatrix}a^4_4/a^2_2 & 0 & 0 & 0\\ a^2_1 & a^2_2 & 0 & 0\\ 0 & 0 & a^2_2/a^4_4 & 0\\ 0 & 0 & a^4_3 & a^4_4\end{pmatrix}\begin{pmatrix}\omega^1\\ \omega^2\\ \omega^3\\ \omega^4\end{pmatrix}.$$

The "transition matrices" take values in a certain 4-dimensional lower triangular subgroup of $\mathrm{GL}(4,\mathbb{R})$ which we shall henceforth denote $G$. Thus, the local coframings which are 1-adapted to the hyperbolic system are the local sections of a principal $G$-bundle $B_G \to M$ which is a subbundle of the bundle of all coframes of $M$. In other words, $B_G$ is a $G$-structure on $M$ in the usual sense. We will refer to $B_G$ as the $G$-structure *associated to* (or *determined by*) the non-degenerate hyperbolic exterior differential system $(M;\,\Omega_1,\,\Omega_2)$.

Now we shall apply the equivalence method to the $G$-structure $B_G$ in order to understand its invariants. Accordingly, we write the structure equations on $B_G$ in the form

$$d\begin{pmatrix}\omega^1\\ \omega^2\\ \omega^3\\ \omega^4\end{pmatrix} = -\begin{pmatrix}\phi_{44}-\phi_{22} & 0 & 0 & 0\\ \phi_{21} & \phi_{22} & 0 & 0\\ 0 & 0 & \phi_{22}-\phi_{44} & 0\\ 0 & 0 & \phi_{43} & \phi_{44}\end{pmatrix}\wedge\begin{pmatrix}\omega^1\\ \omega^2\\ \omega^3\\ \omega^4\end{pmatrix} + \begin{pmatrix}T^1\\ T^2\\ T^3\\ T^4\end{pmatrix} \tag{4}$$

where, in the terminology of the equivalence method, the $\phi_{ij}$ are the *pseudo-connection forms* and the $T^i$ are the *torsion* terms (which are semi-basic[5]). These forms are not uniquely determined by these equations, and, following the usual method of equivalence, we now want to understand how modifications of the pseudo-connection forms can be employed to simplify the torsion terms.

Now, by the defining properties of the $G$-structure $B_G$, we have

$$\begin{aligned} d\omega^1 &\equiv 0 && \mathrm{mod}\,\omega^1,\omega^2 \\ d\omega^2 &\equiv \omega^3\wedge\omega^4 && \mathrm{mod}\,\omega^1,\omega^2 \\ d\omega^3 &\equiv 0 && \mathrm{mod}\,\omega^3,\omega^4 \\ d\omega^4 &\equiv \omega^1\wedge\omega^2 && \mathrm{mod}\,\omega^3,\omega^4 \end{aligned} \qquad \text{so} \qquad \begin{aligned} T^1 &\equiv 0 && \mathrm{mod}\,\omega^1,\omega^2 \\ T^2 &\equiv \omega^3\wedge\omega^4 && \mathrm{mod}\,\omega^1,\omega^2 \\ T^3 &\equiv 0 && \mathrm{mod}\,\omega^3,\omega^4 \\ T^4 &\equiv \omega^1\wedge\omega^2 && \mathrm{mod}\,\omega^3,\omega^4. \end{aligned}$$

[5] I.e., these terms have the form $T^i = T^i_{jk}\,\omega^j\wedge\omega^k$ for some functions $T^i_{jk} = -T^i_{kj}$ on $B$.

It follows that there exist 1-forms $\chi_1, \chi_2, \chi_3$, and $\chi_4$ which are linear combinations of the $\omega^i$ so that

$$T^2 = \omega^3\wedge\omega^4 + \chi_1\wedge\omega^1 + \chi_2\wedge\omega^2,$$
$$T^4 = \omega^1\wedge\omega^2 + \chi_3\wedge\omega^3 + \chi_4\wedge\omega^4.$$

The equations for $d\omega^2$ and $d\omega^4$ can therefore be written in the form

$$\begin{aligned} d\omega^2 &= -(\phi_{21} - \chi_1)\wedge\omega^1 - (\phi_{22} - \chi_2)\wedge\omega^2 + \omega^3\wedge\omega^4 \\ d\omega^4 &= -(\phi_{43} - \chi_3)\wedge\omega^3 - (\phi_{44} - \chi_4)\wedge\omega^4 + \omega^1\wedge\omega^2. \end{aligned}$$

It follows that we may assume that the $\phi_{ij}$ have been chosen so that

$$T^2 = \omega^3\wedge\omega^4 \qquad \text{and} \qquad T^4 = \omega^1\wedge\omega^2,$$

so we assume this from now on. This condition still does not determine the $\phi_{ij}$ since making the replacements

$$\begin{pmatrix}\phi_{21}\\ \phi_{22}\\ \phi_{43}\\ \phi_{44}\end{pmatrix} \mapsto \begin{pmatrix}\phi_{21} + a_1\,\omega^1 + a_2\,\omega^2\\ \phi_{22} + a_2\,\omega^1 + a_3\,\omega^2\\ \phi_{43} + c_3\,\omega^3 + c_4\,\omega^4\\ \phi_{44} + c_4\,\omega^3 + c_1\,\omega^4\end{pmatrix}$$

in the above equations will clearly not affect $T^2$ or $T^4$. However, the above congruences on $T^1$ and $T^3$ imply that

$$\begin{aligned} T^1 &\equiv T^1_{13}\,\omega^1\wedge\omega^3 + T^1_{14}\,\omega^1\wedge\omega^4 \quad \mod \omega^2 \\ T^3 &\equiv T^3_{31}\,\omega^3\wedge\omega^1 + T^3_{32}\,\omega^3\wedge\omega^2 \quad \mod \omega^4 \end{aligned}$$

and the above replacements can be chosen so that $T^1_{13} = T^1_{14} = T^3_{31} = T^3_{32} = 0$. Note that the only replacements of the above form which preserve these latter conditions are ones with $a_2 = a_3 = c_4 = c_1 = 0$.

The upshot of this discussion is that, for the $G$-structure we have associated to a non-degenerate hyperbolic system, there is a choice of pseudo-connection so that the torsion takes the form

$$\begin{pmatrix}T^1\\ T^2\\ T^3\\ T^4\end{pmatrix} = \begin{pmatrix}\omega^2\wedge(p_1\,\omega^1 + p_3\,\omega^3 + p_4\,\omega^4)\\ \omega^3\wedge\omega^4\\ \omega^4\wedge(q_3\,\omega^3 + q_1\,\omega^1 + q_2\,\omega^2)\\ \omega^1\wedge\omega^2\end{pmatrix} \tag{5}$$

Moreover, with the structure equations in this form, the 1-forms $\phi_{22}$ and $\phi_{44}$ are unique, the form $\phi_{21}$ is determined up to the addition of a multiple of $\omega^1$, and the form $\phi_{43}$ is determined up to the addition of a multiple of $\omega^3$. At this stage, no further reduction of this $G$-structure can be made without making some non-vanishing assumptions on the invariants.

To complete the discussion of the structure equations, it will be necessary to compute their "Bianchi identities" by differentiating the equations in (5). We will not give the details of the calculations (which are straightforward, if tedious), but shall describe the results. First of all, differentiation of the equations (5) and reduction of the results modulo various combinations of the $\omega^i$ shows that there are relations of the form



$$\left.\begin{aligned} dp_1 &\equiv p_1\,\phi_{22} && -q_2\,\phi_{43}\\ dp_3 &\equiv p_3\,(3\phi_{22}-2\phi_{44}) && +p_4\,\phi_{43}\\ dp_4 &\equiv p_4\,(2\phi_{22}) &&\\ dq_3 &\equiv q_3\,\phi_{44} && -p_4\,\phi_{21}\\ dq_1 &\equiv q_1\,(3\phi_{44}-2\phi_{22}) && +q_2\,\phi_{21}\\ dq_2 &\equiv q_2\,(2\phi_{44}) && \end{aligned}\right\} \mod \omega^1,\omega^2,\omega^3,\omega^4. \tag{6}$$

We shall use the notation $\nabla p_4$ to mean $dp_4 - 2p_4\,\phi_{22}$, i.e., the semi-basic part of the exterior derivative of $p_4$, and similarly for the other quantities.

If we now introduce "curvature" 2-forms $\Phi_{22}$, $\Phi_{44}$, $\Phi_{21}$, and $\Phi_{43}$ by the equations

$$\begin{aligned} d\phi_{22} &= -\phi_{21}\wedge\left(p_3\,\omega^3 + p_4\,\omega^4\right) + q_3\,\omega^1\wedge\omega^2 + \tfrac{1}{2}p_3q_1\,\omega^2\wedge\omega^4 + \Phi_{22}\\ d\phi_{44} &= -\phi_{43}\wedge\left(q_1\,\omega^1 + q_2\,\omega^2\right) + p_1\,\omega^3\wedge\omega^4 + \tfrac{1}{2}q_1p_3\,\omega^4\wedge\omega^2 + \Phi_{44}\\ d\phi_{21} &= -\phi_{21}\wedge\left(\phi_{44} - 2\phi_{22} - p_1\,\omega^2\right) + \Phi_{21}\\ d\phi_{43} &= -\phi_{43}\wedge\left(\phi_{22} - 2\phi_{44} - q_3\,\omega^4\right) + \Phi_{43}, \end{aligned} \tag{7}$$

the exterior derivatives of the equations (5) become

$$\begin{aligned} 0 &= (\Phi_{22}-\Phi_{44})\wedge\omega^1 - \left(\nabla p_1\wedge\omega^1 + \nabla p_3\wedge\omega^3 + \nabla p_4\wedge\omega^4 - q_3p_3\,\omega^3\wedge\omega^4\right)\wedge\omega^2\\ 0 &= -\Phi_{21}\wedge\omega^1 - \Phi_{22}\wedge\omega^2 - \omega^1\wedge\omega^2\wedge\omega^3\\ 0 &= (\Phi_{44}-\Phi_{22})\wedge\omega^3 - \left(\nabla q_3\wedge\omega^3 + \nabla q_1\wedge\omega^1 + \nabla q_2\wedge\omega^2 - p_1q_1\,\omega^1\wedge\omega^2\right)\wedge\omega^4\\ 0 &= -\Phi_{43}\wedge\omega^3 - \Phi_{44}\wedge\omega^4 - \omega^3\wedge\omega^4\wedge\omega^1. \end{aligned} \tag{8}$$

(Note that because $\phi_{21}$ and $\phi_{43}$ are not canonical, the expression $\nabla p_3$ is actually only well-defined modulo $\omega^3$. However, since this term only occurs wedged with $\omega^3$, the resulting term is well defined. A similar comment applies to the other ambiguities caused by the ambiguity in the pseudo-connection.) The identities (8) now give relations among the coefficients of the derivatives of the primary invariants (i.e., the torsion coefficients) and the curvature coefficients. It is not useful to write these out here; the form (8) will suffice for our purposes.

A little exterior algebra shows that the relations (8) imply that $\Phi_{22}$ and $\Phi_{44}$ are semi-basic 2-forms, i.e., they are quadratic expressions in the $\omega^i$. In fact,

$$\Phi_{22} = -\kappa_1\wedge\omega^1 - \kappa_2\wedge\omega^2 \qquad \text{and} \qquad \Phi_{44} = -\kappa_3\wedge\omega^3 - \kappa_4\wedge\omega^4 \tag{9}$$

where

$$\kappa_i = k_{ij}\omega^j$$

for functions $k_{ij}$, suitably skew-symmetrized so as to be well-defined. Using this plus some more exterior algebra, it follows that there are 1-forms $\phi_{211}$ and $\phi_{433}$ so that

$$\begin{aligned}\Phi_{21} &= -(\kappa_1 - \omega^3)\wedge\omega^2 - \phi_{211}\wedge\omega^1 \\ \Phi_{43} &= -(\kappa_3 - \omega^1)\wedge\omega^4 - \phi_{433}\wedge\omega^3.\end{aligned} \tag{10}$$

We now want to interpret the vanishing of the torsion coefficients in (5) in terms of integrability of various bundles intrinsically associated to the original hyperbolic system and use this to derive (local) normal forms in various special cases. We will then use the structure equations to develop a test for 'linearizability' of non-degenerate hyperbolic systems.

We begin our first interpretation by noting that the rank 2 Pfaffian system $\Theta = \{\omega^1, \omega^3\}$ spanned by the first derived systems of $\Xi_1$ and $\Xi_2$ is well-defined. Indeed, from the structure equations, the 2-form

$$\Omega = \omega^1 \wedge \omega^3$$

itself is well-defined, since the scalings of $\omega^1$ and $\omega^3$ cancel. The integrability of $\Omega$ has the following interpretation:

*For any non-degenerate hyperbolic system, the system $\Theta = \{\omega^1, \omega^3\}$ is Frobenius if and only if the torsion coefficients satisfy $p_4 = q_2 = 0$. Moreover, the rank one systems $\{\omega^1\}$ and $\{\omega^3\}$ are integrable if and only if we have, in addition to the above, the relations $p_3 = q_1 = 0$.*

In these special cases we have the following normal forms result:

**Proposition:** *Let $(M; \Omega_1, \Omega_2)$ be a non-degenerate hyperbolic system with $p_4 = q_2 = 0$. If $p_3 q_1 \neq 0$, then the hyperbolic system is locally generated by the PDE system*

$$\begin{aligned} u_y + u u_x &= C, \qquad & C_v &\neq 0, \\ v_x + v v_y &= D, & D_u &\neq 0 \end{aligned}$$

*where $C$ and $D$ are functions of $x, y, u$ and $v$. On the other hand, if $p_3 = q_1 = 0$, then the local model is given by the following (wave equations)*

$$\begin{aligned} u_y &= C, \qquad & C_v &\neq 0, \\ v_x &= D, & D_u &\neq 0. \end{aligned}$$

**Remark:** Non-degeneracy of the hyperbolic system is expressed by the coupling condition $C_v D_u \neq 0$.

As another example of the use of the invariants of $B_G$ to understand normal forms, we shall give a characterization of linear systems of PDE in terms of these invariants.

**Proposition:** *A non-degenerate hyperbolic system $(M; \Omega_1, \Omega_2)$ satisfies $p_4 = q_2 = p_3 = q_1 = p_1 = q_3 = 0$ and $\Phi_{22} + \Phi_{44} = F\omega^1\wedge\omega^3$ for some function $F$ if and only if it is locally the hyperbolic system associated to a linear first order hyperbolic PDE system.*



**Remark:** In this case, a little algebra shows that

$$(\Phi_{22} - \Phi_{44}) = K\omega^1 \wedge \omega^3.$$

Furthermore, the quadratic differential form

$$ds^2 = \omega^1 \circ \omega^3$$

is well-defined and induces on solution surfaces an intrinsic pseudo-Riemannian metric whose null-geodesics are the characteristic curves. The Gauss curvature of this metric may then be verified to be the function $K$ above. In this way the geometry associated to a PDE induces, in an intrinsic manner, a geometry on solution surfaces.

## References


[1] R. Bryant, Two exotic holonomies in dimension four, path geometries and twistor theory, in: *Complex Geometry and Lie Theory, Proc. Sym. Pure Math.* **53** (1991), 33–88.

[2] R. Bryant, S.-S. Chern, R. Gardner, H. Goldschmidt, and P. Griffiths, Exterior differential systems, MSRI Publications #18, Springer-Verlag (1991).

[3] R. Bryant and P. Griffiths, Characteristic cohomology of differential systems (I): General theory, IAS preprint (1993); (II): Conservation Laws for a Class of Parabolic Equations, IAS preprint (1993).

[4] R. Bryant, P. Griffiths and L. Hsu, Hyperbolic exterior differential systems and their conservation laws, IAS preprint (1994).

[5] R. Bryant and L. Hsu, Geometric integrability and Cartan's class (C) equations, (in preparation).

[6] É. Cartan, Sur les variétés à connexion projective, Bull. Soc. Math. France **52** (1924), 205–241.

[7] É. Cartan, Les espaces généralisés et l'intégration de certaines classes d'équations différentielles, C. R. Acad. Sci. **206** (1938), 1689–1693.

[8] É. Cartan, Les problèmes d'équivalence, in Œuvres Complètes, Partie II, Éditions du CNRS, Paris (1984).

[9] R. E. Caflisch, N. Ercolani, T. Y. Hou and Y. Landis, Multi-valued solutions and branch point singularities for non-linear hyperbolic or elliptic systems, Comm. Pure and Applied Math. **46** (1993), 453–499.

[10] S.-S. Chern, The geometry of the differential equation $y''' = F(x, y, y', y'')$, Sci. Rep. Tsing Hua Univ. **4** (1940), 97–111.

[11] S.-S. Chern, Geometry of $G$-structures, Bulletin of the AMS **72** (1966), 167–219.

[12] R. Courant and D. Hilbert, Methods of Mathematical Physics, Volume I, Interscience, New York (1953), Volume II, Interscience, New York (1962).

[13] G. Darboux, Leçons sur la theorie general des surfaces, Tome II, Hermann, Paris (1915).

[14] Ph. Le Floch and Z. Xin, Uniqueness via the adoint problems for systems of conservation laws, Comm. Pure and Applied Math. **46** (1993), 1499–1533.

[15] R. B. Gardner, The Method of Equivalence and Its Applications, CBMS–NSF Regional Conference Series in Applied Math. **58**, SIAM, Philadelphia (1989).

[16] J. Glimm, Solutions in the large for non-linear hyperbolic systems of equations, Comm. Pure and Applied Math. **18** (1965), 697–715.

[17] E. Goursat, Leçons sur l'intégration des équations aux dérivées partielles du second ordre à deux variables indépendantes, Tome 1, Hermann, Paris (1896), Tome 2, Hermann, Paris (1898).

[18] J. Hurtubise and N. Kamran, Projective connections, double fibrations, and formal neighbourhoods of lines, Math. Ann. **292** (1992), 383–409.

[19] P. Lax, Development of Singularities of Solutions of Nonlinear Hyperbolic Partial Differential Equations, J. Math. Phys. **5** (1964), 611–613.

[20] J. Smoller, Shock Waves and Reaction-Diffusion Equations, Springer, New York (1983).

# Operators on Noncommutative Differential Forms and Cyclic Homology

**Joachim Cuntz**
*Mathematisches Institut*
*Universität Heidelberg*
*In Neuenheimer Feld 288*
*D-69120 Heidelberg, Germany*

**Daniel Quillen**
*Mathematical Institute*
*University of Oxford*
*24 St. Giles'*
*Oxford OX1 3LB, ENGLAND*

Let $A$ be an associative unital algebra over the complex numbers and let $\Omega A$ be the differential graded algebra of its noncommutative differential forms (*[3]*). The first topic of this paper is a family of canonical operators $b, d, \kappa, B, P, G$ on $\Omega A$ arising in the following way.

On $\Omega A$ we have two differentials

$$\Omega^0 A \underset{d}{\overset{b}{\leftrightarrows}} \Omega^1 A \underset{d}{\overset{b}{\leftrightarrows}} \Omega^2 A \underset{d}{\overset{b}{\leftrightarrows}}$$

where $d$ is the differential associated to the DG algebra structure, and where $b$, the Hochschild differential, is such that $(\Omega A, b)$ is the standard normalized complex calculating the Hochschild homology $HH_n(A)$.

Let us now view $\Omega A$ as analogous to the space of differential forms on a Riemannian manifold, with $b$ playing the role of the adjoint of $d$. The 'Laplacian' $bd + db$ then has the form $1 - \kappa$, where $\kappa$ is an operator introduced by Karoubi (*[6]*, 2.12). The Karoubi operator $\kappa$ satisfies a polynomial identity in each degree, hence $\Omega A$ decomposes into generalized eigenspaces for the Laplacian. Thus we have the 'harmonic decomposition'

$$\Omega A = P\Omega A \oplus P^{\perp}\Omega A \tag{1}$$

where $P$ is the spectral projection onto the generalized nullspace for the Laplacian and $P^{\perp} = 1 - P$. The 'Green's operator' $G$ is then defined to be zero on $P\Omega A$ and the inverse of the Laplacian on $P^{\perp}\Omega A$. Finally, the Connes boundary operator $B$

**Geometry, Topology, and Physics for Raoul Bott**

Cambridge, MA 02138 USA